Vinesh Raja

Kiran J. Fernandes

Editors

Reverse Engineering

An Industrial Perspective

Springer

Springer Series in Advanced Manufacturing

Series Editor

Professor D.T. Pham
Manufacturing Engineering Centre
Cardiff University
Queen's Building
Newport Road
Cardiff CF24 3AA
UK

Other titles in this series

Vinesh Raja and Kiran J. Fernandes (Eds.)

Reverse Engineering

An Industrial Perspective

Springer

Vinesh Raja, PhD, C.Eng
Warwick Manufacturing Group
University of Warwick
Coventry
UK

Kiran J. Fernandes, PhD, C.Eng
The York Management School
University of York
Heslington
York
UK

About the Editors

Prof. Vinesh Raja is a Professorial Fellow in Informatics at the University of Warwick. He is incharge of the Informatics Group, which encompasses the Virtual Reality Center (VRC) and the Collaborative Product Commerce Center (CPC) at the Warwick Manufacturing Group. He focuses on augmenting and extending everyday learning and work activities with interactive technologies that move 'beyond the desktop'. This involves designing enhanced user experiences through appropriating and assembling a diversity of technologies including haptics, handheld and pervasive computing. The main focus of his research is not the technology per se but the design and integration of the digital representations that are presented *via* them to support social and cognitive activities in ways that extend current capabilities.

Dr. Kiran Jude Fernandes is the 40th Anniversary Research Lecturer in Management at the University of York. He has been a pioneer in the systematic study of information modeling techniques and tools and has studied their evolution using techniques from the biological sciences. His research and teaching interests include strategic uses of information systems, information management, and the impact of information technology on the risks and benefits of outsourcing and strategic alliances. Before joining the University of York, Kiran worked at the University of Warwick and the NASA John C. Stennis Space Center.

Preface

Congratulations and thank you for reading this book! You hold in your hand perhaps the first book solely written on mechanical reverse engineering from an industry perspective. The motivation for this book originates from the needs of today's global industry.

We recall an incident during one of our industrial trips to a local manufacturing company. The office secretary was photocopying documents for this meeting, when the manufacturing manager remarked, "Wouldn't it be nice if I could do the same with mechanical parts, it would save me and my team a lot of time and money." "Have you not heard of reverse engineering?" we asked him. "Reverse engineering, isn't that something to do with programming computers?" "No," we replied. "Reverse engineering (RE) refers to creating a computer-aided design (CAD) model from an existing physical object, which can be used as a design tool for producing a copy of an object, extracting the design concept of an existing model, or reengineering an existing part." His eyes lit up. Such situations are not uncommon in today's manufacturing arena.

With globalization and trade liberalization, manufacturing companies face increasing competition from goods and services produced in lower wage economies. Countries in the West cannot compete against low wages and must therefore depend on raising innovation and best practices to create better products. In an attempt to compete in such a volatile environment, companies are looking to lean and agile strategies to compete and survive. Lean or world class manufacturing is principally aimed at reducing waste and controlling things that can be measured and controlled. On the other hand, agility deals with things that cannot be controlled.

To be agile and lean, companies cannot apply traditional approaches that often result in problems with inventories, overhead, and inefficiencies. Companies need to create small quantities of highly customized, designed-to-order parts that meet the needs of the global customer. The swift trend toward a multiplicity of finished products with short development and production lead times has led many companies into problems with inventories, overhead, and inefficiencies. They are trying to apply the traditional mass-production approach

without realizing that the whole environment has changed. Mass production does not apply to products where the customers require small quantities of highly custom, designed-to-order products, and where additional services and value-added benefits such as product upgrades and future reconfigurations are as important as the product itself. Approaches such as rapid prototyping (RP) and reverse engineering (RE) are helping to solve some of these problems.

Rapid Prototyping – Rapid prototyping is a relatively new class of technology used for building physical models and prototype parts from 3-D computer-aided design (CAD) data. Unlike milling machines (which are subtractive), RP systems join liquid, powder, and sheet materials together to form complex parts. Layer-by-layer, RP machines fabricate plastic, wood, ceramic, and metal objects based on thin horizontal cross sections taken from computer models.

Reverse Engineering – Reverse engineering encompasses a variety of approaches to reproduce a physical object with the aid of drawings, documentation, or computer model data. In the broadest sense, reverse engineering is whatever it takes–manual or under computer control–to reproduce something.

This is a book for people interested in RE from an industrial perspective. Several journal papers have discussed issues related to RE, but from our conversations with industrialists and our personal experience, there is a huge gap in practical literature in this field, especially in manufacturing. Hence, this book is written for the benefit of the industrialist, who might not have the time to "scramble" through libraries and other sources to read journal papers. Although this book is written for such "novice" engineers, we expect the reader to be familiar with basic computer-aided design and manufacturing principles. The main theme of the book is to get you started using RE as quickly as possible. We have provided examples from the aerospace, automotive, and medical equipment industries to familiarize you with the principles and techniques of reverse engineering.

Probably the most unusual thing about this book is that we start with practical examples from industry. This approach has worked well in the classroom when we teach our course to industrial students. We expect the same results with this book. Many of the examples used in the later chapters depend on the principles and techniques of RE, so, it is essential that you read the first few chapters.

Having taught reverse engineering to industrial managers and engineers for a number of years, we are beginning to see that RE is finally receiving the respect it deserves.

Acknowledgments

Vinesh Haridas Raja
Kiran Jude Fernandes

Contents

List of Contributors

Mongi Abidi
University of Tennessee
Department of Electrical and
Computer Engineering
Imaging, Robotics, and Intelligent
Systems Lab
330 Ferris Hall, Knoxville TN 37996
USA

Brian C. Beherens
Department of Decision Sciences
and MIS Saint Louis University
Cook School of Business
3674 Lindell Blvd.
Saint Louis, MO 63108
USA

Frits Feenstra
TNO Science and Industry,
P.O. Box 6235
5600 HE, Eindhoven
Netherlands
e-mail: f.feenstra@ind.tno.nl

Kiran Jude Fernandes
Department of Management
Studies
University of York
Heslington, York YO10 5DD
UK
e-mail: kf501@york.ac.uk

Ping Fu
Geomagic
3200 East Hwy 54
Cape Fear Building, Suite 300
Research Triangle Park
NC, 27709
USA
e-mail: pfu@geomagic.com

L.C. Hieu
Manufacturing Engineering Centre
Cardiff University
Queen's Building, Newport Road
Cardiff CF24 3AA
UK

John Keast
Warwick Manufacturing Group
International Manufacturing Centre
University of Warwick
Coventry, CV4 7AL
UK
e-mail: J.E.Keast@warwick.ac.uk

Andreas Koschan
University of Tennessee
Department of Electrical and
Computer Engineering
Imaging, Robotics, and Intelligent
Systems Lab
330 Ferris Hall, Knoxville TN 37996
USA
e-mail: akoschan@utk.edu

Reuven R. Levary
Department of Decision Sciences
and MIS Saint Louis University
Cook School of Business
3674 Lindell Blvd.
Saint Louis, MO 63108
USA
e-mail: levarypr@slu.edu

H.T. Loh
Department of Mechanical
Engineering
National University of Singapore
21 Lower Kent Ridge Road
Singapore 119077

Eef Moeskops
TNO Science and Industry
P.O. Box 6235
5600 HE, Eindhoven
Netherlands

David Page
University of Tennessee
Department of Electrical and
Computer Engineering
Imaging, Robotics, and Intelligent
Systems Lab
330 Ferris Hall, Knoxville TN 37996
USA

D.T. Pham
Manufacturing Engineering Centre
Cardiff University
Queen's Building, Newport Road
Cardiff CF24 3AA
UK
e-mail: phamdt@cardiff.ac.uk

Vinesh Raja
Warwick Manufacturing Group
International Manufacturing Centre
University of Warwick
Coventry, CV4 7AL
UK
e-mail: vinesh.raja@warwick.ac.uk

Y.S. Wong
Department of Mechanical
Engineering
National University of Singapore
21 Lower Kent Ridge Road
Singapore 119077

Y.F. Zhang
Department of Mechanical
Engineering
National University of Singapore
21 Lower Kent Ridge Road
Singapore 119077
e-mail: mpezyf@nus.edu.sg

Chapter 1

Introduction to Reverse Engineering

Vinesh Raja, University of Warwick, UK

Abstract

This chapter introduces readers to the term reverse engineering (RE), and to the associated techniques that can be used for scanning physical parts. In addition, the chapter presents the process of reverse engineering and the strategy for scanning and converting the scanned data into a 3-D surface or solid model.

1.1 Introduction

In today's intensely competitive global market, product enterprises are constantly seeking new ways to shorten lead times for new product developments that meet all customer expectations. In general, product enterprise has invested
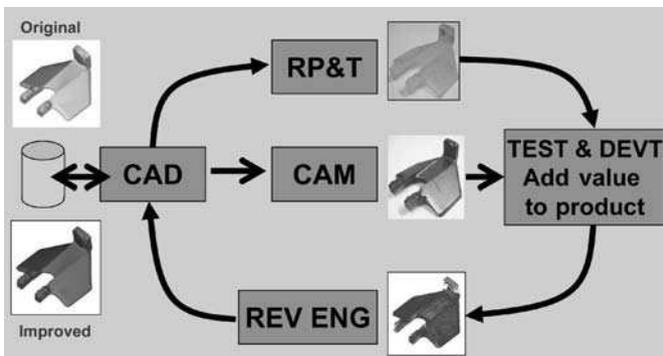
Figure 1.1. Product development cycle

in CADCAM, rapid prototyping, and a range of new technologies that provide business benefits. Reverse engineering (RE) is now considered one of the technologies that provide business benefits in shortening the product development cycle. Figure 1.1 below depicts how RE allows the possibilities of closing the loop between what is "as designed" and what is "actually manufactured".

1.2 What Is Reverse Engineering?

Engineering is the process of designing, manufacturing, assembling, and maintaining products and systems. There are two types of engineering, forward engineering and reverse engineering. Forward engineering is the traditional process of moving from high-level abstractions and logical designs to the physical implementation of a system. In some situations, there may be a physical part/ product without any technical details, such as drawings, bills-of-material, or without engineering data. The process of duplicating an existing part, subassembly, or product, without drawings, documentation, or a computer model is known as reverse engineering. Reverse engineering is also defined as the process of obtaining a geometric CAD model from 3-D points acquired by scanning/ digitizing existing parts/products. The process of digitally capturing the physical entities of a component, referred to as reverse engineering (RE), is often defined by researchers with respect to their specific task (Motavalli & Shamsaasef 1996). Abella *et al.* (1994) described RE as, "the basic concept of producing a part based on an original or physical model without the use of an engineering drawing". Yau *et al.*(1993) define RE, as the "process of retrieving new geometry from a manufactured part by digitizing and modifying an existing CAD model".

Reverse engineering is now widely used in numerous applications, such as manufacturing, industrial design, and jewelry design and reproduction For example, when a new car is launched on the market, competing manufacturers may buy one and disassemble it to learn how it was built and how it works. In software engineering, good source code is often a variation of other good source code. In some situations, such as automotive styling, designers give shape to their ideas by using clay, plaster, wood, or foam rubber, but a CAD model is needed to manufacture the part. As products become more organic in shape, designing in CAD becomes more challenging and there is no guarantee that the CAD representation will replicate the sculpted model exactly.

Reverse engineering provides a solution to this problem because the physical model is the source of information for the CAD model. This is also referred to as the physical-to-digital process depicted in Figure 1.2. Another reason for reverse engineering is to compress product development cycle times. In the intensely competitive global market, manufacturers are constantly seeking new ways to shorten lead times to market a new product. Rapid product development (RPD) refers to recently developed technologies and techniques that assist manufacturers and designers in meeting the demands of shortened product development time. For example, injection-molding companies need to shorten tool and die
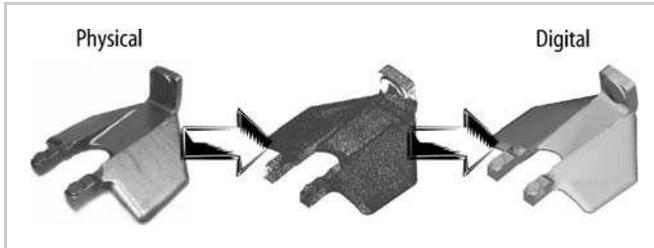
Figure 1.2. Physical-to-digital process

development time drastically. By using reverse engineering, a three-dimensional physical product or clay mock-up can be quickly captured in the digital form, remodeled, and exported for rapid prototyping/tooling or rapid manufacturing using multi-axis CNC machining techniques.

1.3 Why Use Reverse Engineering?

Following are some of the reasons for using reverse engineering:

- The original manufacturer no longer exists, but a customer needs the product, *e.g.*, aircraft spares required typically after an aircraft has been in service for several years.
- The original manufacturer of a product no longer produces the product, *e.g.*, the original product has become obsolete.
- The original product design documentation has been lost or never existed.
- Creating data to refurbish or manufacture a part for which there are no CAD data, or for which the data have become obsolete or lost.
- Inspection and/or Quality Control–Comparing a fabricated part to its CAD description or to a standard item.
- Some bad features of a product need to be eliminated *e.g.*, excessive wear might indicate where a product should be improved.
- Strengthening the good features of a product based on long-term usage.
- Analyzing the good and bad features of competitors' products.
- Exploring new avenues to improve product performance and features.
- Creating 3-D data from a model or sculpture for animation in games and movies.
- Creating 3-D data from an individual, model or sculpture to create, scale, or reproduce artwork.
- Architectural and construction documentation and measurement.
- Fitting clothing or footwear to individuals and determining the anthropometry of a population.

- Generating data to create dental or surgical prosthetics, tissue engineered body parts, or for surgical planning.
- Documentation and reproduction of crime scenes.

The above list is not exhaustive and there are many more reasons for using reverse engineering, than documented above.

1.4 Reverse Engineering–The Generic Process

The generic process of reverse engineering is a three-phase process as depicted in Figure 1.3. The three phases are scanning, point processing, and application-specific geometric model development. Reverse engineering strategy must consider the following:
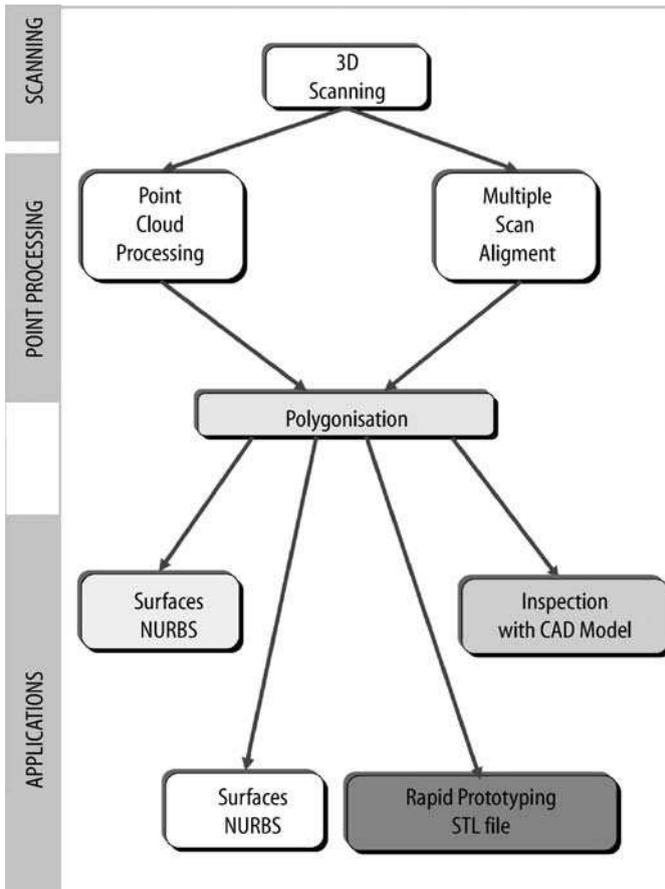
Figure 1.3. Reverse engineering – the generic process

- Reason for reverse engineering a part
- Number of parts to be scanned–single or multiple
- Part size–large or small
- Part complexity–simple or complex
- Part material–hard or soft
- Part finish–shiny or dull
- Part geometry–organic or prismatic and internal or external
- Accuracy required–linear or volumetric

1.5 Phase 1–Scanning

This phase is involved with the scanning strategy–selecting the correct scanning technique, preparing the part to be scanned, and performing the actual scanning to capture information that describes all geometric features of the part such as steps, slots, pockets, and holes. Three-dimensional scanners are employed to scan the part geometry, producing clouds of points, which define the surface geometry. These scanning devices are available as dedicated tools or as add-ons to the existing computer numerically controlled (CNC) machine tools. There are two distinct types of scanners, contact and noncontact.

1.5.1 Contact Scanners

These devices employ contact probes that automatically follow the contours of a physical surface (Figure 1.4). In Ithe current marketplace, contact probe
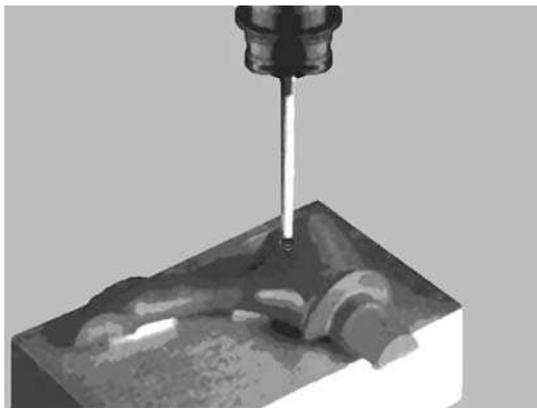
Figure 1.4. Contact scanning touch probe. Originally published in *Rapid Prototyping Casebook*, McDonald, J.A., Ryal, C.J. and Wimpenny, D.I., 2001, © John Wiley and Sons Limited. Reproduced with permission.

scanning devices are based on CMM technologies, with a tolerance range of +0.01 to 0.02 mm. However, depending on the size of the part scanned, contact methods can be slow because each point is generated sequentially at the tip of the probe. Tactile device probes must deflect to register a point; hence, a degree of contact pressure is maintained during the scanning process. This contact pressure limits the use of contact devices because soft, tactile materials such as rubber cannot be easily or accurately scanned.

1.5.2 Noncontact Scanners

A variety of noncontact scanning technologies available on the market capture data with no physical part contact. Noncontact devices use lasers, optics, and charge-coupled device (CCD) sensors to capture point data, as shown in Figure 1.5. Although these devices capture large amounts of data in a relatively short space of time, there are a number of issues related to this scanning technology.

- The typical tolerance of noncontact scanning is within ±0.025 to 0.2 mm.
- Some noncontact systems have problems generating data describing surfaces, which are parallel to the axis of the laser (Figure 1.6).
- Noncontact devices employ light within the data capture process. This creates problems when the light impinges on shiny surfaces, and hence some surfaces must be prepared with a temporary coating of fine powder before scanning.

Figure 1.5. Optical scanning device. Originally published in *Rapid Prototyping Casebook*, McDonald, J.A., Ryal, C.J. and Wimpenny, D.I., 2001, © John Wiley and Sons Limited. Reproduced with permission.
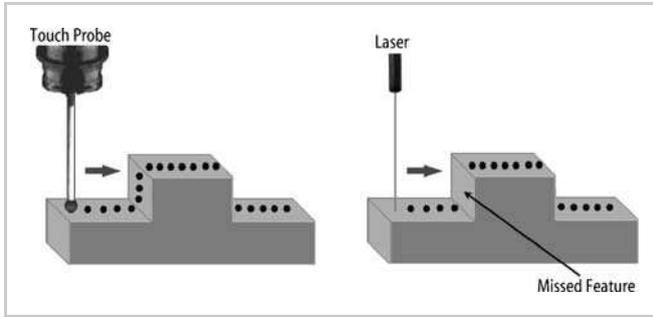
Figure 1.6. Vertical faces–touch probe versus a laser. Originally published in *Rapid Prototyping Casebook*, McDonald, J.A., Ryal, C.J. and Wimpenny, D.I., 2001, © John Wiley and Sons Limited. Reproduced with permission.

These issues restrict the use of remote sensing devices to areas in engineering, where the accuracy of the information generated is secondary to the speed of data capture. However, as research and laser development in optical technology continue, the accuracy of the commercially available noncontact scanning device is beginning to improve.

The output of the scanning phase is point cloud data sets in the most convenient format. Typically, the RE software provides a variety of output formats such as raw (*X, Y, Z* values separated by space or commas).

1.6 Phase 2–Point Processing

This phase involves importing the point cloud data, reducing the noise in the data collected, and reducing the number of points. These tasks are performed using a range of predefined filters. It is extremely important that the users have very good understanding of the filter algorithms so that they know which filter is the most appropriate for each task. This phase also allows us to merge multiple scan data sets. Sometimes, it is necessary to take multiple scans of the part to ensure that all required features have been scanned. This involves rotating the part; hence each scan datum becomes very crucial. Multiple scan planning has direct impact on the point processing phase. Good datum planning for multiple scanning will reduce the effort required in the point processing phase and also avoid introduction of errors from merging multiple scan data. A wide range of commercial software is available for point processing.

The output of the point processing phase is a clean , merged, point cloud data set in the most convenient format. This phase also supports most of the proprietary formats mentioned above in the scanning phase.

1.7 Phase 3–Application Geometric Model Development

In the same way that developments in rapid prototyping and tooling technologies are helping to shorten dramatically the time taken to generate physical representations from CAD models, current RE technologies are helping to reduce the time to create electronic CAD models from existing physical representations. The need to generate CAD information from physical components will arise frequently throughout any product introduction process.

The generation of CAD models from point data is probably the most complex activity within RE because potent surface fitting algorithms are required to generate surfaces that accurately represent the three-dimensional information described within the point cloud data sets. Most CAD systems are not designed to display and process large amounts of point data; as a result new RE modules or discrete software packages are generally needed for point processing. Generating surface data from point cloud data sets is still a very subjective process, although feature-based algorithms are beginning to emerge that will enable engineers to interact with the point cloud data to produce complete solid models for current CAD environments.

The applications of RE for generating CAD data are equally as important as the technology which supports it. A manager's decision to employ RE technologies should be based on specific business needs.

This phase depends very much on the real purpose for reverse engineering. For example, if we scanned a broken injection molding tool to produce a new tool, we would be interested in the geometric model and also in the ISO G code data that can be used to produce a replacement tool in the shortest possible time using a multi-axis CNC machine. One can also use reverse engineering to analyze "as designed" to "as manufactured". This involves importing the as designed CAD model and superimposing the scanned point cloud data set of the manufactured part. The RE software allows the user to compare the two data sets (as designed to as manufactured). This process is also used for inspecting manufactured parts. Reverse engineering can also be used to scan existing hip joints and to design new artificial hips joint around patient- specific pelvic data. This creates the opportunity for customized artificial joints for each patient.

The output of this phase is geometric model in one of the proprietary formats such as IGES, VDA, STL, DXF, OBJ, VRML, ISO G Code, *etc.*

This chapter defined the term "reverse engineering" followed by reasons for using reverse engineering. It also introduced the reverse engineering strategy, the three phases of the reverse engineering generic process, contact and noncontact scanning, point processing, and application geometric model development.

Chapter 2 builds on Chapter 1 by providing an in-depth depiction of methodologies and techniques for reverse engineering.

Chapter 3 presents information on reverse engineering hardware and software and also provides excellent information on commercially available reverse engineering hardware and software.

Chapter 4 provides a structured approach for selecting the appropriate reverse engineering system for a specific business need.

Chapter 5 introduces the fundamentals of rapid prototyping with an in-depth description of various commercially available rapid prototyping technologies.

Chapter 6 provides excellent information on the relationship between reverse engineering and rapid prototyping.

Chapter 7 provides cases studies of successful applications of reverse engineering in the automotive industry.

Chapter 8 provides cases studies of successful applications of reverse engineering in the aerospace industry.

Chapter 9 provides cases studies of successful applications of reverse engineering in the medical equipment industry.

Chapter 10 describes the legal aspects and implications of reverse engineering.

Chapter 11 discusses the barriers to adopting reverse engineering.

Chapter 2

Methodologies and Techniques for Reverse Engineering–The Potential for Automation with 3-D Laser Scanners

David Page, Andreas Koschan, and Mongi Abidi, University of Tennessee, USA

Abstract

In this chapter, we present methodologies and technologies for automating reverse engineering (RE) through digital imaging and computer vision. We begin this chapter with a definition of RE in terms of generating computer-aided design (CAD) models from existing objects and components. We use the term computer-aided reverse engineering (CARE) to describe this process. With this definition, we present a brief overview of the traditional approach to RE using coordinate measuring machines (CMMs). Then, we begin the main focus of the chapter where we explore digital imaging and computer vision as an alternative to CMMs. This exploration begins with data acquisition, where we present laser-based range scanners as a promising approach. For these scanners, we explain and highlight differences in data resolution and scanning rates and contrast those to CMM performance. Next, we present a processing pipeline for creating CAD models using these scanners. This processing includes tasks such as view registration, surface integration, patch reconstruction, model fitting, noise removal, and other functions. This chapter explains these topics to help the reader understand their importance in designing an RE imaging system and the impact that various parameters have on modeling accuracy.

2.1 Computer-aided Reverse Engineering

Reverse engineering (RE) has many meanings to many different people. As we begin this chapter, we first focus our attention on the specific meaning that we intend to address and then extend our attention to the automation of RE, which

we call computer-aided reverse engineering (CARE). To this end, this chapter discusses CARE in general and then explores 3-D laser scanning in depth as an emerging CARE technology from the computer vision community, as introduced in Varady *et al.* (1997), Bernardini *et al.* (1999), and Page *et al.* (2003). The chapter details the methodologies and techniques associated with computer vision scanning and notes specific challenges for the CARE problem.

2.1.1 What Is Not Reverse Engineering?

Each discipline of engineering has a different definition for RE. Computer engineers and computer scientists, for example, refer to RE when they speak of determining the algorithmic functionality of a software package when they have no prior knowledge of the original software design. Engineers and programmers attempt to develop functional block diagrams of the software through interaction with the interface and to develop high-level code descriptions from raw machine code. This software definition is not the scope of our RE discussion. Another example of RE that might be familiar–but also outside the scope of this chapter–concerns revealing the inner workings of a machine to figure out what makes it tick. This form of RE is also a systems level approach where an engineer disassembles the item of interest to develop an understanding of the functional relationship of components or to gain insight into the types of materials used to fabricate the components. As with software RE, the goal is to develop a high-level description of a system without *a priori* knowledge. These two examples are common applications that use the term RE, but we wish to emphasize that our definition of RE is not related to these examples, but is instead related to the area of computer-aided engineering (CAE).

2.1.2 What is Computer-aided (Forward) Engineering?

In the late 1970s and into the 1980s, computer-aided design (CAD)–a component of CAE–began to revolutionize engineering disciplines. The peak of this revolution occurred in 1990 with the design of the Boeing 777; the entire aircraft was designed and preassembled through a virtual CAD) simulation. According to Boeing, the first 777 to roll out of the production hangar in 1994 was just hundredths of an inch–about the thickness of a playing card–within alignment. This precision contrasts with the half-inch alignments common with most aircraft parts before that time–an improvement of several orders of magnitude. This revolution in technology has continued into the 1990s and today with the emergence and growth of computer-aided manufacturing (CAM). CAM is the automation of the manufacturing process itself–beyond just the design process–where machines such as computerized numerically controlled (CNC) mills allow precise fabrication of objects directly from CAD descriptions. With CAM, a designer can rapidly move from a conceptual CAD description to a real-world

tangible object. We might use the term *forward* engineering—in a tongue-in-cheek manner—to describe this type of design, and the term CAE to describe the automation of forward engineering through CAD and CAM technologies.

2.1.3 What Is Computer-aided Reverse Engineering?

CAE through CAD and CAM technologies is the automation of engineering and fabrication, where a design formalizes ideas through computer modeling and then fabricates those models into real-world objects. CARE flows in the opposite direction. CARE creates a computer model of an object through measurements of the object, as it exists in the real world. In this context, we define CARE as the *reversal* of CAE or the ability to generate a CAD model from a real-world tangible object. We illustrate this flow in Figure 2.1. The disc brake appears on the left side of this figure and its CAD model appears on the right. We acquired this particular brake from a local junkyard and cleaned the surfaces of rust and dirt. Then, we used a laser-based range scanner to create the CAD model. This model is metrically accurate to within a few millimeters of the original junkyard brake. By analogy, one might think of this capability as a 3-D fax. Just as a facsimile (fax) machine converts a hand-written document into digital form, we have converted a tangible object (the disc brake) into a computer model. This figure illustrates our definition of CARE. To be more concrete, we define CARE in terms of the geometry and shape of an object and not in terms of its functionality as with the previous RE counterexamples in Section 2.1.1.
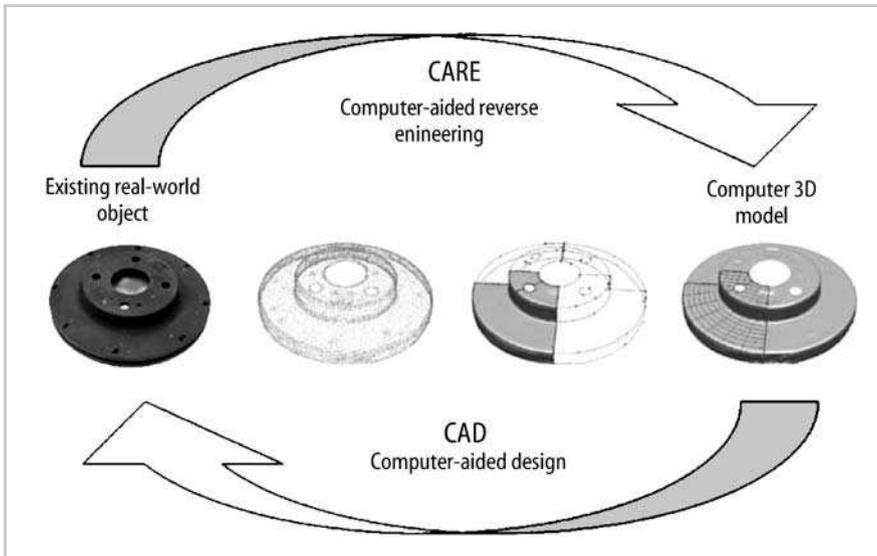
Figure 2.1. Computer-aided reverse engineering (CARE) process

To understand the CARE steps, consider the stages shown in Figure 2.1 from left to right. The first step in the CARE process is to make measurements at points along the surface of the brake. Each point has an x, y, and z coordinate locating the point in 3-D space. For a given object, a CARE system will measure hundreds or even thousands of such points depending on the nature of the object and the type of CARE system. The collection of these points is known as a *point cloud*; an example appears in the second picture from the left in Figure 2.1. In most applications, the point cloud is a sufficient description of the object. However, higher levels are possible as the remaining two pictures on the right show. The third picture from the left is a feature description of the object, in which the CARE system has detected surface edges and creases from the point cloud data. The final picture on the right is a full and complete CAD description of the object. For this description, the CARE system uses the point cloud and the detected features to fit surfaces for modeling the entire geometry of the object.

For both industrial and military applications, CARE offers many advantages to the engineering design, manufacturing fabrication, and field support of a part or component. For example, CARE allows rapid inspection and validation in real time at the production line based on the original CAD designs. A production technician can quickly evaluate tolerances relative to the CAD models. Such feedback enables tighter control over the manufacturing process or may aid future redesigns of the component. As another example, because CARE creates digital models, the technology is ideal for electronic dissemination of information. A manufacturing center in Asia can transmit as-built models of components to a design center in North America or *vice versa*. Although many thousands of miles separate these two centers, CARE enables more efficient collaboration. The North American team can transmit their CAD design to Asia *via* electronic mail, and the Asian team can return a CARE model of a fabricated prototype. Previously, such electronic sharing was a one-way endeavour where the physical prototype would require expensive courier-based delivery back to North America. Finally, archival problems of field modifications and out-of-production components are another area for which CARE systems offer promise.

Engineers often encounter situations where nontechnical personnel develop novel and important modifications of their designs but only after they are deployed. Such field-operations personnel are inadequately equipped to communicate their modifications to the engineers. These modifications are typically *ad hoc* with little or no documentation, such as CAD drawings. Thus, engineers have little hope if they wish to capture these modifications and incorporate them into future designs. However, the potential exists for even nontechnical personnel to send back a 3-D fax of the modification to the engineer. Additionally, engineers themselves face a similar problem when asked to modify an out-of-production component for which CAD documentation does not exist or has been lost. The engineer has the tedious task of generating a CAD model without a CARE system. This ability is the promise of CARE to allow both technical and nontechnical individuals to generate engineering quality CAD models of existing objects quickly and automatically.

If a CARE system automatically captures the geometric structure of an object and stores the subsequent shape and topology information as a CAD model, the next question might be how to achieve this goal. How can a CARE system automatically measure the geometry of an object? In industry, the most common answer is through a coordinate measuring machine (CMM). Although CMMs are not the computer vision solution that this chapter explores, they are the starting point as an industry standard for a discussion of CARE. So, in the next section, we quickly review the capabilities of a CMM and then compare those capabilities to the computer vision approach that laser scanners offer.

2.2 Computer Vision and Reverse Engineering

Computer vision bridges diverse fields from electrical engineering to computer science to cognitive psychology. Computer vision systems seek to develop computer models of the real world through processing of image data from sensors such as video cameras or–as in our case–3-D range scanners. Because computer vision is relatively new to RE, we begin this section by first investigating traditional (noncomputer vision) approaches to RE, and then use these methods as a backdrop for laser range scanners.

2.2.1 Coordinate Measuring Machines

This photograph shows a measurement of the disc brake using calipers. Calipers are a common *ad hoc* approach to RE. These measurement devices allow engineers and machinists to determine accurate diameters, lengths, and other dimensions of objects. This approach to RE is a manual process that requires significant effort for complicated objects and surfaces. CMM technology is the first effort to automate the RE process. Before CMM and probably still popular for most simple tasks, engineers and machinists have used measurement calipers. For the disc brake, we could use calipers to measure the diameters of the various holes and cylinders that comprise the basic shape of the brake, as in Figure 2.2. Then, from these measurements, we could manually lay out a computer model of the brake using CAD primitives. For a simple object, this manual process of RE is straightforward, but as the complexity of the object shape increases, the basic CAD primitives such as planar and quadric surfaces are no longer suitable. A free-form surface, for example, that is nonplanar and nonquadratic (Campbell and Flynn 2000) does not lend itself readily to characterization with just calipers. Free-form surfaces require special consideration and attention. Calipers are not practical for capturing their subtleties.

As an alternative, CMMs first appeared in the early 1960s and are a more practical means for characterizing and inspecting free-form surfaces. A CMM consists of a probe supported on three mutually perpendicular (x, y, and z) axes; each axis has a built-in reference standard. Figure 2.3 provides a conceptual view of

a CMM. The probe allows accurate measurements along each axis relative to the standard. Thus, a CMM generates 3-D coordinate points as the probe moves across a surface. Operators may run a CMM in a manual mode where they maneuver the probe around an object and collect coordinate measurements, or they
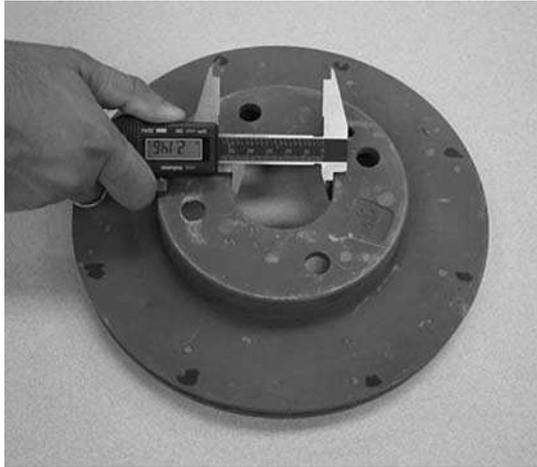
Figure 2.2. Measuring the disk brake using a caliper
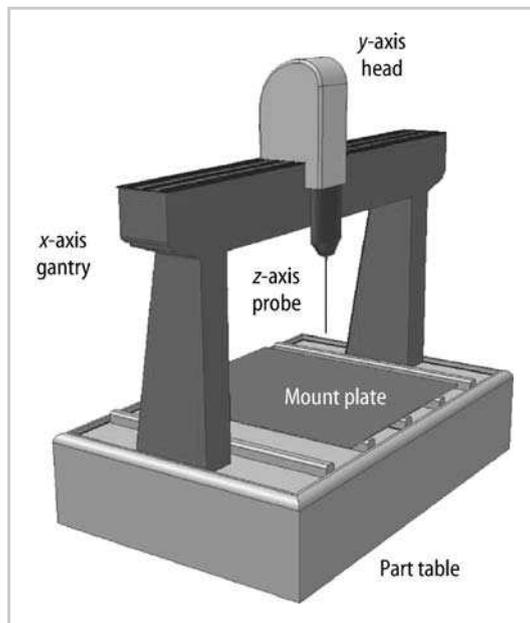
Figure 2.3. Conceptual view of a CMM that illustrates the major components of most systems

may program the probe to maneuver automatically on its own. This latter method is more relevant to the CARE definition under discussion. Different CMM manufacturers offer different schemes to help operators plan the path that the probe will follow. The more advanced CMM systems allow operators to upload a CAD model of the object and then the CMM uses this model for the path planning strategy. The CMM will analyze the CAD model to identify critical points and regions such as discontinuity creases or tight spaces. Tight spaces are a point of emphasis because the probe must come in contact with or be close to the point on the surface where a measurement occurs. If the probe cannot reach the point due to extension constraints (*i.e.,* the point is too far) or limited space (*i.e.,* the probe is too large), then the system cannot measure at that location. The probe must touch or be near the measurement location. This constraint is the motivation for computer vision approaches to CARE because computer vision offers a more standoff measurement.

2.2.2 Active Illumination 3-D Stereo

Computer vision attempts to develop solutions through sensors such as video cameras or other imaging-based systems. For an introduction, see Faugeras (1993), Suk and Bhandarkar (1992), or Klette *et al.* (1998). To acquire 3-D information, the most well-known methodology that probably comes to mind is stereovision, in which two cameras operate in a manner similar to human eyes. Unfortunately, although stereo-based imaging has become useful in such applications as navigation planning and obstacle avoidance for Mars Rover missions, passive stereo does not offer the accuracy that industrial RE applications demand. The major drawback of stereovision is the need to establish correspondence between a point in one camera's image and the same point in the second camera's image. If we can establish this correspondence, we can construct an accurate stereo depth map, or range image, but correspondence is an elusive problem. The term *passive* stereo is important because it indicates that the cameras in the system do not rely on active forms of illumination, except ambient light. One way to overcome the correspondence problem is to control illumination in an intelligent manner and thus simplify the correspondence search. Such methods are known as *active* illumination stereo (or more simply active stereo) where we replace one of the stereo cameras with a well-defined light source. A laser is typically the light source of choice.

Active stereo using laser-based illumination allows more accurate depth measurements than traditional passive illumination stereo. (We use the term *depth measurement* as an alternative to a 3-D point measurement from CMMs. At this point, the subtle nuance between these two terms is not important, but they are interchangeable because a depth measurement is a 3-D point measurement.) Although passive stereo is not viable for a CARE system, active stereo does offer potential. To understand active stereo, we should consider three popular approaches to illumination control. These major categories are (a) continuous wave

modulation, (b) time-of-flight estimation, and (c) structured-light triangulation. Continuous wave systems measure depth by using a coherent light source, *i.e.*, a laser, and measure phase differences of the light wave as it travels to and from the object. Figure 2.4 shows an example. The distance of an object from the camera is proportional to the phase difference between the transmitted wave and the returning wave.

The accuracy of these systems correlates with the accuracy in measuring the phase differences and the wavelength of the laser. Time-of-flight methods operate similarly to radar where the system measures the time required for a laser pulse to travel to and return from an object. Figure 2.5 illustrates such a system. Highly accurate depth measurements require precision electronics to measure speed-of-light time increments. As a third option for active stereo, structured-light techniques compute depth through geometric triangulation. The camera, the object, and the light source form the triangulation geometry. See Figure 2.6. This configuration is similar to passive stereo geometry, except that we have replaced one of the cameras with a light source. The accuracy of these methods is a function of the camera resolution, geometric dimensions, and illumination precision, but the primary parameter for controlling and increasing accuracy is the camera resolution. System geometry and illumination are not as critical. Thus, structured-light systems offer a solution more practical than passive stereo in achieving the accuracy necessary for an RE system. In a later section, we will explore structured-light techniques more in depth, but first we seek to compare and contrast the benefits of active stereo to CMM.
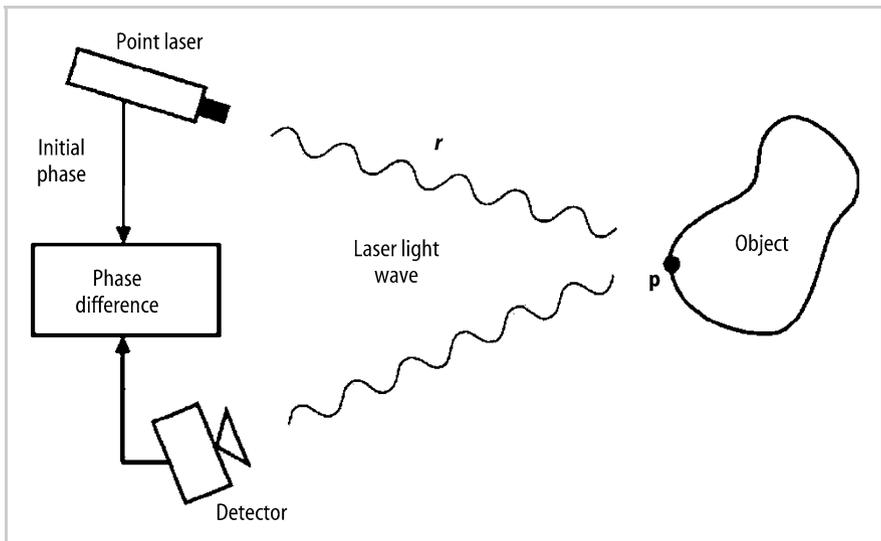
Figure 2.4. Active stereo example of a continuous wave system. The laser travels as a light wave from the laser source to the object and back to the detector. As a wave, the laser light undergoes a phase change as it travels. The difference in phase reveals the object's distance.
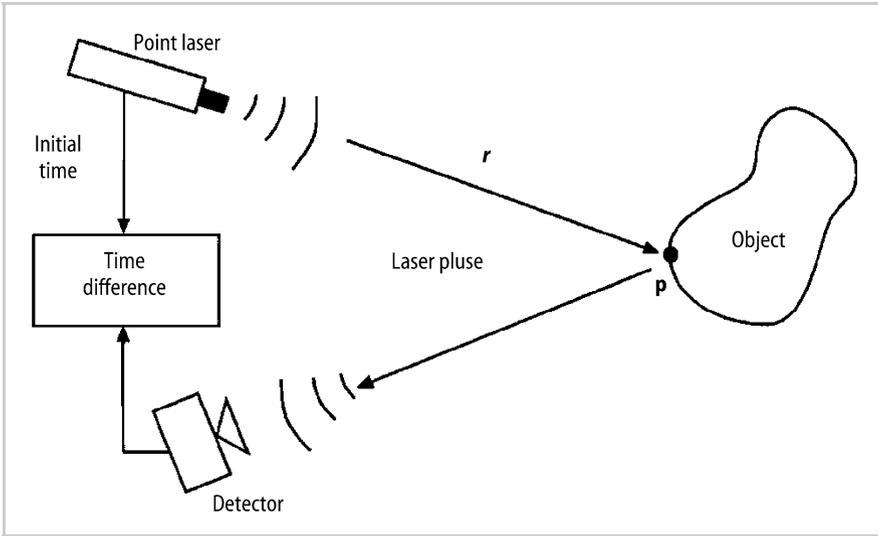
Figure 2.5. Active stereo example of a time-of-flight system. A point laser emits a pulse that is reflected from the object of interest. The difference between the initial time when the pulse was transmitted and the time that it returns to the detector correlates with the object's distance.
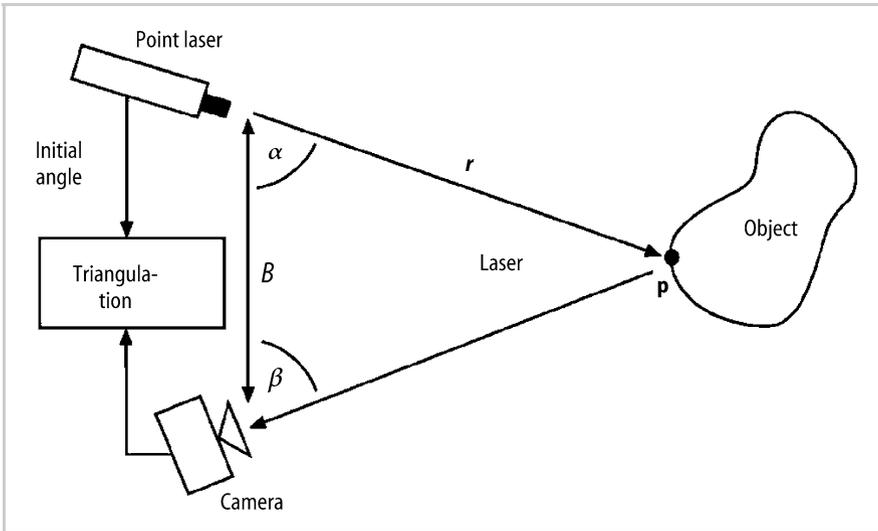
Figure 2.6. Active stereo example of a structured-light system. This example shows a triangulation-based approach where B, β, and α are typically known through system calibration.

2.2.3 Benefits and Drawbacks

In the previous sections, we discussed CMMs and active stereo as methods for CARE. Now we explore the benefits and drawbacks of these systems in cost, accuracy, speed, and reliability. Because we are not discussing specific manufacturers and because active stereo is an emerging technology for CARE, we limit ourselves to a qualitative comparison at a more theoretical level, rather than a quantitative comparison at a marketing level.

To begin, we note that CMMs are generally mechanical systems. Although they have computer control and use state-of-the-art electronics for measurements, they are bound by physical principles where a measurement probe must physically maneuver around an object close to its surfaces. Cameras and range scanners, on the other hand, are solid-state sensors that take measurements from a distance. Thus they require less mechanical precision in movement. The accuracy of a CMM is based directly on the mechanical control of the probe, but the accuracy of an active stereo system depends more on its solid-state sensors. An insightful analogy would be to consider the difference between analogue watches with gears and springs and digital watches with transistors and quartz crystals. With solid-state systems, this analogy brings to mind other points of interest such as the benefits of fewer moving parts, the advantages of speed, and the ease of software reconfigurability.

As noted previously, CMMs must contact the object of interest to measure its geometry. By contrast, computer vision systems are noncontact in that a camera makes measurements at some standoff distance. (We should clarify that some commercial CMM systems claim to be noncontact devices, but they do not necessarily allow standoff measurements. They still require a measurement probe to be near the point of measurement, just not touching it.) The standoff nature of computer vision systems offers two advantages. The first is that they avoid damage to either the object or the measurement probe. As a probe scans and traverses a surface, it may scratch that surface if the user fails to handle the probe properly or if the surface itself is too brittle or highly malleable. Additionally, if the temperature of the surface is too hot or too cold, the heat transfer could damage the measurement probe. These scenarios are not a major concern with computer vision systems. A typical laser-based range scanner can measure points on an object at a distance of 10 to 20 centimeters, if not more.

The second advantage to the standoff capability is the simplification of path planning. With a touch probe, a CMM (or a user) must carefully select a measurement path that properly covers the surfaces of an object but that avoids wedging the probe into tight spaces. The CMM must use a path that covers the object and yet obeys the physical constraints imposed by interaction between the object and the probe. Typically, a CAD model of the object is necessary to aid path planning by serving as a predictor of the object shape. Path planning for a computer vision system is less restrictive. A camera or laser range scanner can hover around the object and collect measurements just as a person taking pictures of the object might do. Some planning is necessary to ensure that we

measure the entire object, but the system does not have to check physical constraints as with a CMM probe. Again, an analogy is illustrative. Consider two men one of whom is wearing a blindfold. Both men are asked to describe an object. The man with no blindfold will simply pick up the object, turn it around a few times to obtain different views, and then report his description. Because he cannot see, the second man must be a little more careful as he uses his fingers to probe around the object and explore its general structure by touch. His path planning is obviously more complicated than that of the first man. This analogy lacks sufficient sophistication, as with most generalizing analogies. It unintentionally implies that CMMs and touch probes are less accurate than computer vision methods, but the analogy does convey the essence of path planning with each approach.

The previous discussion highlights the main advantages of computer vision systems for CARE, but in practice, their disadvantages hinder their practical use over CMMs. As a newcomer to RE, computer vision solutions are a young technology that has not reached the commercial maturity that CMMs have attained. Currently, CMM technology has tremendous market penetration. The existing capital investment in CMM tilts industry toward upgrading rather than replacing CMM technology. The old adage–if it ain't broke, don't fix it–gives credence to this resistance to adopting computer vision technology. Also, commercial computer vision systems for CARE are just beginning to emerge from the research community and many research questions are still unresolved. Among these questions are how to improve accuracy. Beyond market penetration, the most significant disadvantage of computer vision systems is accuracy. Ultra-high-accuracy CMMs work in the 1–2 micron range, and more moderate (in terms of cost) CMMs in the 10–20 micron range. Computer vision methods cannot compete—as of yet—with these levels where most systems operate in the submillimeter range of accuracy. The trend, however, indicates that CMMs have plateaued. Only a few microns difference in accuracy can result in more than a \$100 K increase in cost. On the other hand, computer vision research indicates that greater accuracy is yet to come with more precise lasers and higher resolution imaging sensors. With this hope, the next section will delve deeper into the most common computer vision approach a structured-light system, and later we investigate many of the research problems in this approach.

2.3 Structured-light Range Imaging

We begin this section by first outlining three categories of structured-light solutions to 3-D range imaging. Then, we select one of these categories, the sheet-of-light method, and discuss it in more detail. This section seeks to explain the manner in which a computer vision system collects 3-D point information. A broader presentation is given in Klette *et al.* (1998). Integrated Vision Products (2000) gives specifics for a commercially available system.

2.3.1 Source Illumination Categories

Although we have mentioned several approaches to range imaging in a previous section, the most promising for CARE are the structured-light approaches. Recall that these are triangulation methods, in which the geometry of the laser, camera, and object are exploited to recover the distance from the camera to the object. This distance, or range value, is a function of three parameters: the baseline distance between the light source and the camera, the viewing angle of the camera, and the illumination angle of the light source. Recall the geometry in Figure 2.6. We have the following equation for the range value, r:

$$r = F(\alpha, B, i, j, \dots), \tag{2.1}$$

where F is a function, possibly nonlinear, α is the illumination angle, and B is the baseline distance between the camera and the light source. The variables i and j represent the horizontal and vertical positions, respectively, of a pixel in the camera image. (A pixel is the fundamental picture element from a digital camera. We commonly represent images as a matrix where a pixel is an element of the matrix. We classify the quality of a digital camera in terms of the number of pixels. For example, a 1-megapixel camera usually has 960 rows $\times$ 1280 columns for its pixel matrix.) This pixel is the brightest spot where the light source falls on the imaging plane of the camera. The point on the image directly correlates with the range, *i.e.*, distance, of the object from the laser. The sequence in Figure 2.7 demonstrates how the pixel location changes with increasing range. Due to the view disparity between the light source, as the object moves further away from the light source, in this case a point laser, the projection of the laser onto the image plane from where it strikes the object moves as well. When the object is closest to the laser, Figure 2.7a, the projection is near the left edge of the image plane, but when the object is farthest, Figure 2.7c, the projection is near the right edge.

In Figure 2.7, we have used a point laser for illumination, but other types of light sources are possible. The three most common sources of illumination are single-point lasers, sheet-of-light lasers, and coded-light patterns. As already discussed, Figure 2.7 shows the point laser. If we do not move the object, we can scan the laser in a raster fashion across the scene to obtain a range measurement for every pixel. If the image has $M \times N$ pixels (*i.e.*, $0 \le i < M - 1$, $0 \le j < N - 1$), then we require $M \times N$ measurements. The second type of illumination is the sheet-of-light laser shown in Figure 2.8. It projects a plane of light onto the object so that this plane intersects the object and forms a line. That line becomes a projection onto the image plane of the camera. This method is a direct extension of the point laser approach, except that the line represents a collection of simultaneous point sources. Thus, one can readily see that a sheet-of-light system allows one to increase scanning speed over a point laser with almost no increase in effort. The sheet-of-light method requires only M measurements for an $M \times N$ image. The third method is the coded-light source. This method illustrated in Figure 2.9 is
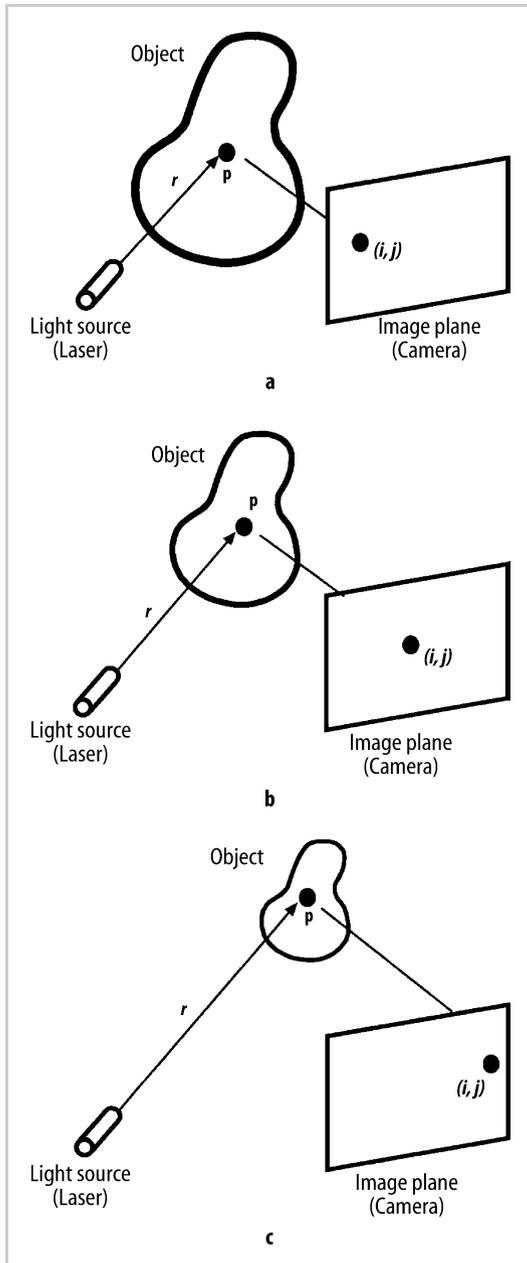
Figure 2.7. Demonstration of how a structured-light system determines distance. (**a**) The object is close to the laser source. (**b**) The object has been moved a slight distance further away. (**c**) The object is at an even greater distance. Note the change where the laser projects onto the image plane of the camera as the object moves.
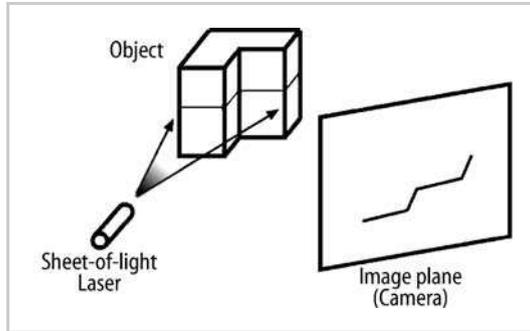
Figure 2.8. Sheet-of-light range scanner. The laser source fans out to create a plane that intersects the object of interest. The resulting line of light on the object projects as a line onto the imaging plane of the camera.
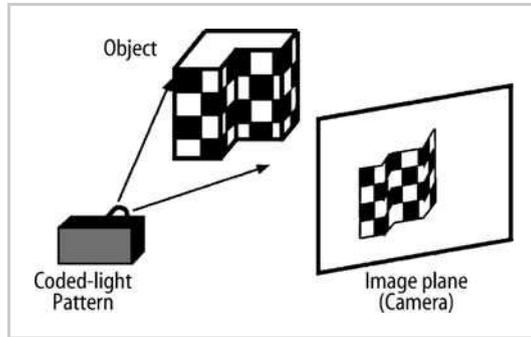
Figure 2.9. Coded-light range scanner. The light source is usually a projector with a screen pattern placed over the lens. The pattern projects onto the object and subsequently projects onto the image plane of the camera.

much different from either of the prior two methods. Usually, this method does not use a laser source but rather uses a slide projector with patterns placed in the slide carousel. The system projects these patterns as coded light onto the object, as the figure illustrates. The patterns take different forms but one technique uses a checkerboard. The range resolution of this method is a function of the number of patterns used. Thus, an $M \times N$ image would require R measurements where $R = \log(M)$ if we seek a resolution similar to a sheet-of-light scanner (Integrated Vision Products 2000).

The primary difference between these three illumination methods is the speed with which one can collect data. Accuracy varies little among them. Initially, one might think that the coded-light method offers the fastest scanning speeds, but in practice, the mechanisms necessary to exchange and display patterns onto an object are a bottleneck. The sheet-of-light method lends itself to the fastest implementation. Commercial sheet-of-light scanners are available that can generate 10,000 data points per second, and this number is growing as technology advances. Such speeds are beyond the capabilities of a mechanical CMM probe

that would need to bounce around the object at a very high rate. These data rates are the unique advantage that computer vision laser scanners–specifically sheet-of-light scanners–bring to CARE. In the next section, we take a more in-depth look at these scanners.

2.3.2 Sheet-of-light Range Imaging

Sheet-of-light scanners offer the greatest speed advantage in collecting 3-D data compared to other laser scanners and thus are the most suitable for CARE applications. Their basic operation is such that the laser projects a line onto the object of interest, and triangulation among this line, the laser, and a camera yield 3-D measurements. The photograph in Figure 2.10a is an example of a laser line projecting onto a set of objects, and Figure 2.10b is an example of a system, the Ranger Scanner developed by Integrated Vision Products (IVP).

In the figure, the objects are resting on a conveyor belt. This belt is one method for obtaining a full scan of an object. With the IVP Ranger, as with other sheet-of-light scanners, a single measurement results in a single line–a profile–of data. This profile is the 3-D data for points where the laser line falls on the object. A set of such profiles across the object is necessary to form a complete 3-D description of the object. Thus the conveyor belt moves the object under the laser line to scan the profiles over the entire object. If we stack the resulting profiles together, we have a 3-D model of the object. The conveyor generates profile slices of the object. Another approach is to not move the object but
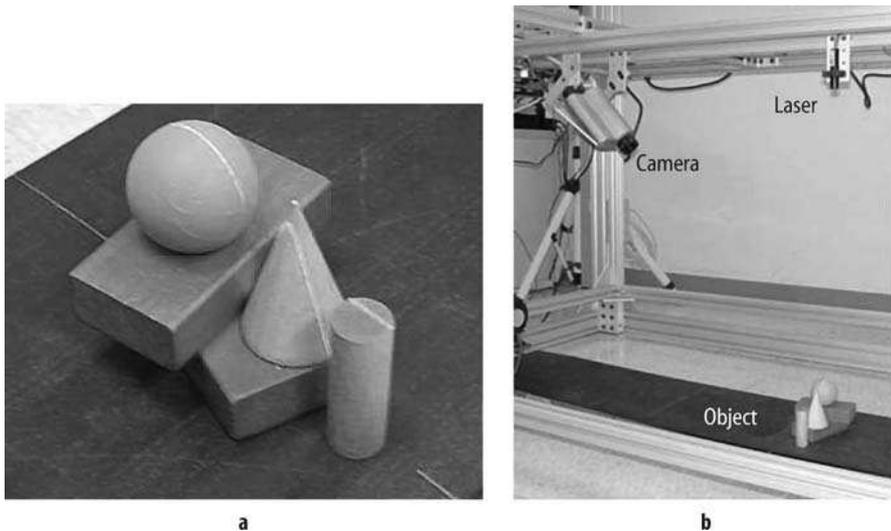
Figure 2.10. Examples of a sheet-of-light system. (**a**) The laser sheet casts a line on the objects under measurement. (**b**) This system is the IVP Ranger Scanner. A color reproduction of this figure can be seen in the Color Section (pages 219–230).

rather to move the scanner. Some sheet-of-light systems take this approach where usually the camera and laser are mounted on a wand. The user swipes this wand around the object of interest. With the conveyor, the spacing between the profiles is simply a function of the traveling speed of the conveyor and the acquisition rate of the scanner. With the wand, a more complicated tracking system is necessary to determine the interprofile relations. Commercial solutions range from magnetic position trackers to CMM probes.

As for accuracy, the geometry defined in Figure 2.6 and the resolution of the camera govern the system. The most common arrangement is for the view angle α to be in the 30–60° range, and camera resolutions are of the order of 512×512 pixels or greater. These pixel resolutions correlate with more than nine bits of range resolution. As the view angle α increases toward 90° or decreases toward zero, the range accuracy decreases. At these extreme viewing angles, the camera has difficulty detecting the positional changes in the laser. First, let us consider a zero view angle. If the object surfaces are planar and perpendicular to the laser plane, the camera will not see the laser because of the oblique angle, but this angle yields the best projection for depth changes. Secondly, consider the 90° angle. With this view, the offset camera is almost looking straight down the laser plane. The camera detects the laser at the same position regardless of depth changes on the object's surface. This view offers the worst projection to the camera, but it gives the camera the best views of the laser reflection. Thus, the 30–60° range balances these two issues of nonoblique viewing and projection resolution. This trade-off is also a function of the baseline distance B between the camera and the laser. As in traditional passive stereo, this distance should be of the order of the same magnitude as the range increments of interest.

The final problem we consider is occlusion, which is a common problem with computer vision systems. More specifically, the problem is known as self-occlu-
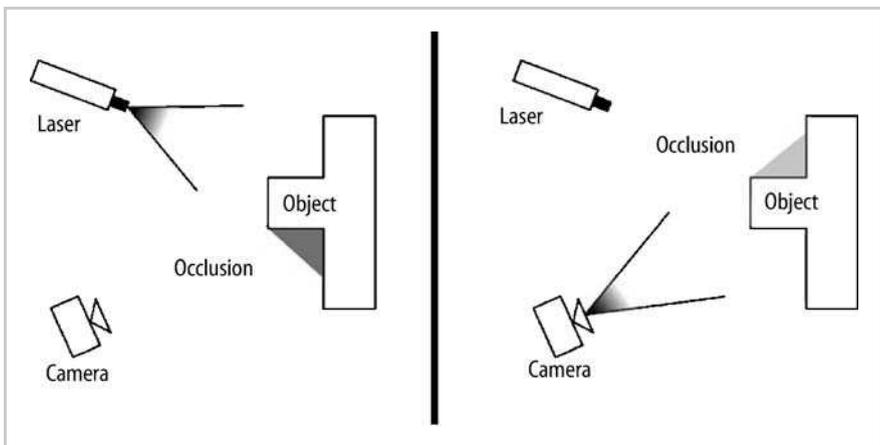
Figure 2.11. The occlusion problem with a sheet-of-light scanner. The object may occlude both the laser and the camera from viewing different surfaces on the object.

sion where a part of an object occludes the viewing of another part of the same
object. A sphere is a simple but not necessarily obvious example. When a cam-
era views a sphere, the camera is able to image the front side, but it cannot im-
age the reverse side. The front occludes the back. Shapes that are more compli-
cated than a sphere have more intricate occlusions. Because the camera and the
laser must view the object, a sheet-of-light range scanner has two possibilities
for occlusions. As Figure 2.11 illustrates, the camera and laser cannot view the
same surfaces of the object. The camera may image a surface that self-occlusion
does not allow the laser to illuminate. Conversely, the laser may illuminate
a surface that self-occlusion hides from the camera. A variety of strategies are
possible to minimize occlusions, but we cannot eliminate the problem, as the
simple sphere example demonstrates. In the next section, we discuss a pipeline
for building CAD models using laser range scanners. We specifically discuss
multiple views as a necessary solution to occlusion.

2.4 Scanner Pipeline

The goal of this section is to outline the steps that are necessary for a laser range
scanner to capture 3-D data of a real object and to translate the data into a CAD
description. We formulate these steps as a pipeline in Figure 2.12. Because many
of these steps are current areas of active research in computer vision, we present
an overview of the pipeline itself and identify the research challenges that face the
practical deployment of a computer vision CARE scanner. Várady *et al.* (1997),
Thompson *et al.* (1999), Guo (1997), and Campbell and Flynn (2001) also define
pipelines similar to Figure 2.12, and we refer the reader to these papers for more
specific details of implementation.

2.4.1 Data Collection

The first step for a CARE scanner is data collection. We have already covered
this issue fairly well in the previous section on structured-light techniques.
Sheet-of-light systems are probably best suited for the CARE application. How-
ever, we should discuss a few other important topics, and we note that Suk and
Bhandarkar (1992) provide additional points of discussion.

The first topic we consider is calibration. Figure 2.13 demonstrates how cali-
bration allows transforming a range image into a 3-D point cloud. We have been

Existing real-world object → Laser-based range scanner → Mesh reconstruc-tion → Surface fitting → CAD Model of object
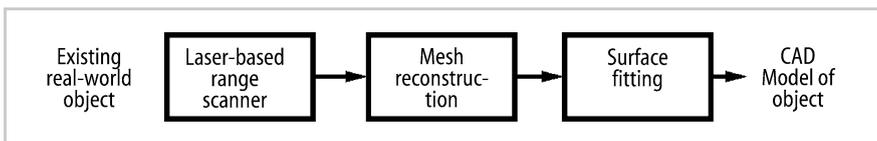
Figure 2.12. Block diagram of a CARE system based on a laser range scanner
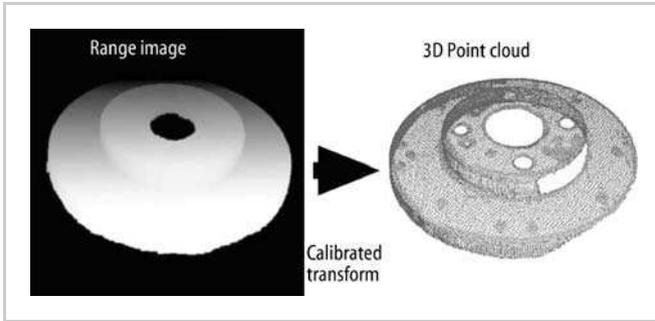
Figure 2.13. Example of a range image for the brake. The gray level of a pixel in the left image represents the distance between the brake and the sheet-of-light scanner. The darker pixels are farther away. If we calibrate the scanner, we can transform this range image into a 3-D point cloud, as shown on the right. The points (dots) exist in 3-D and each have x, y, and z coordinates.

using the term range image somewhat loosely without specific definition. A range image is similar to a conventional digital image that one might find with a digital camera. Consider a pixel (i, j), as in Figure 2.7. The distance to the object in the scene is the range value r where the units for r are relative to the configuration of the system, which may not be directly known. Thus, the raw data from a range scanner is an image with pixel data (r, i, j), but we do not necessarily know the units for r nor the relation of pixels to those units. Calibration determines the relationship between (r, i, j) and some world coordinate system (x, y, z) in units such as meters or inches. With calibration, we can define a transform T that converts a range image into a 3-D point cloud,

$$\mathbf{p} = (x, y, z) = T(r, i, j), \tag{2.2}$$

where $\mathbf{p}$ is a point in the cloud. The image on the left of Figure 2.13 is a range image; the one on the right is a screen shot of a 3-D point cloud. The point cloud is difficult to visualize on a 2-D sheet of paper, but using a 3-D viewer, we would be able to navigate in and around the points that the screen shot shows statically.

After calibration, our next important topic is view registration. In the previous section, we mentioned that occlusions are problems that lead to missing data. We noted that one way to overcome this problem is to move the range scanner around the object to obtain different viewing angles. Figure 2.14 illustrates how at least two views are necessary to collect measurements of the entire brake. If the scanner is on one side of the brake, small portions of the other side are occluded. The challenge with multiple views, however, is that we must now register the subsequent coordinate systems. Consider that a point $\mathbf{p}$ from one view has coordinates (x, y, z) relative to one world coordinate system, but that same point in the other view has coordinates $\mathbf{p}' = (x', y', z')$ relative to a second coordinate system. Although $\mathbf{p}$ and $\mathbf{p}'$ represent physically the same point on a surface, the different view measurements lead to different coordinate systems.
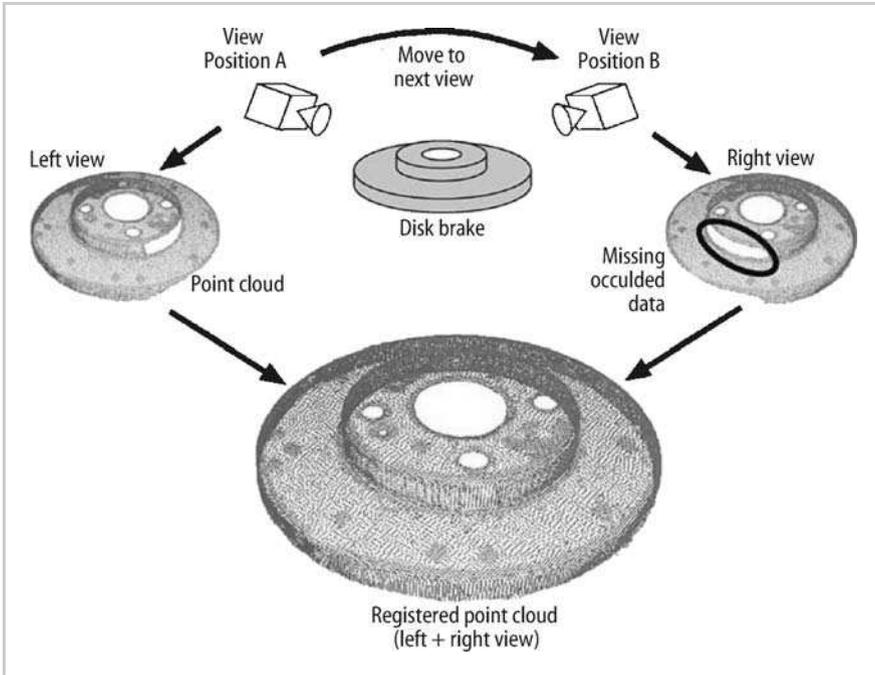
Figure 2.14. Multiple view scanning and registration overcome occlusions. Part of the brake is excluded in each of the views above. When these views are registered together, one view fills the occluded regions of the other view. A color reproduction of this figure can be seen in the Color Section (pages 219–230).

Registration is the process of recovering the rigid transformation that relates these two coordinate systems. The equation for this transformation is

$$\mathbf{p} = \mathbf{R}\mathbf{p}' + \mathbf{t}, \tag{2.3}$$

where the matrix $\mathbf{R}$ is a rotation and the vector $\mathbf{t}$ is a translation. The problem is that $\mathbf{R}$ and $\mathbf{t}$ are unknown. Registration algorithms recover these variables using the raw data from different views of the object. Horn *et al.* (1989) present a closed-form solution for the recovery itself. One criterion is the amount of overlap among the views. If no overlap exists, registration is practically impossible unless we impose higher order constraints. Some partial overlap is necessary to resolve the registration. A CARE scanner using laser ranging must define the amount of overlap between views, as the scanner collects data around the object.

2.4.2 Mesh Reconstruction

Once we have collected various views of an object and registered those views to create a point cloud, the next step for the CARE scanner is to reconstruct a mesh

that models the surface of the object–to connect the dots of the point cloud, so to speak. Without much thought, most people initially consider mesh reconstruction trivial, and it probably is trivial for the human mind, but automatic reconstruction by a computer is not so easy. Hoppe *et al.* (1992) first addressed this topic for the general case of an unorganized point cloud. Since then, many researchers have presented their solutions, and a survey appeared in Campbell and Flynn (2001). Figure 2.15 illustrates mesh reconstruction. This figure is a zoom view of the disc brake to show the triangle mesh recovered from the point cloud. Triangles are a default standard as a surface primitive for a variety of reasons, but mainly they simplify computer visualizations because they have guaranteed convexity and thus are useful as a first-order approximation of an object. (Later we will discuss surface fitting, which leads to higher order approximations.)

For reconstruction, sampling defines the number of points in the point cloud and their relative distance to each other (*i.e.*, their density) on the object. Sampling is a trade-off between data set size and object coverage. A computer can more readily store and manipulate a smaller data set, but a larger set more accurately captures the fidelity of the object. Undersampling leads to incomplete and inaccurate models, but oversampling tends to overwhelm computing resources. Mesh reconstruction algorithms must define their sampling requirements, and users must tailor their CARE applications accordingly. Additionally, inherent in sampling is measurement error. Although this error is a data acquisition problem, most mesh reconstruction algorithms take this problem into account to minimize the effects of such error. If we have a 3-D point **p** on the surface of an object, measuring that point with a scanner inevitably introduces an error such that the observed point **q** has the following relation:

$$\mathbf{q} = \mathbf{p} + \mathbf{e}, \tag{2.4}$$

where the vector **e** is the error. Mesh reconstruction algorithms attempt to minimize the impact of **e**. In addition to measurement error, a reconstruction algorithm must consider three factors. One is the topology of the object. Topology concerns the global structure of the object. For example, a ball and a doughnut
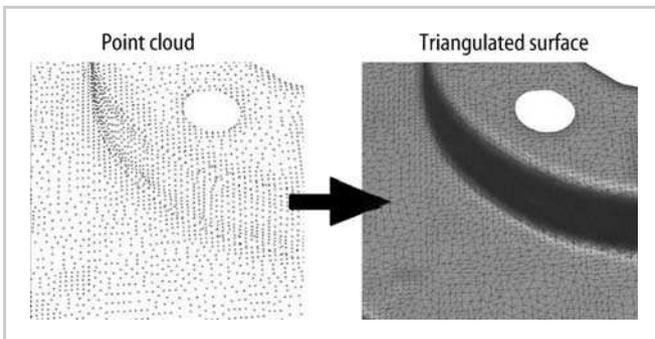
Figure 2.15. Reconstruction recovers the triangulated surface from the 3-D point cloud data

have different topologies. However, a block of wood and the ball have the same topology. The topological hole in the middle of the doughnut is the difference. A discussion of topology is beyond the scope of this chapter, but we want to emphasize that it is an important consideration for reconstruction. The second factor is the amount of overlap between different views of the object. Previously, we have noted in Figure 2.14 the need to collect multiple views to overcome self-occlusion. Typically, these multiple views overlap in their coverage primarily to aid registration for Equation (2.3), but overlap is also useful for minimizing measurement error. A reconstruction algorithm requires a strategy to handle multiple views and their overlap. The third factor is missing data, which lead to holes in the mesh. These holes are not topological holes as with the doughnut but gaps in the mesh like cutouts on a sheet of paper. Generally, researchers have developed two categories of mesh reconstruction algorithms to deal with these factors of topology, overlap, and missing data. These categories are surface-based such as presented by Bernardini *et al.* (1999) and volumetric-based methods such as developed by Curless and Levoy (1996).

2.4.3 Surface Fitting

Once we have a first-order approximation of the object from mesh reconstruction, the final stage of the pipeline is to generate higher order descriptions that are more appropriate for CAD applications. In some cases, a triangle mesh itself is sufficient as the final product of a CARE scanner. In other cases, parametric representations such as nonuniform rational B-splines (NURBS) are necessary. We broadly label this final stage as surface fitting, but in practice, it takes many forms and is highly dependent on the CARE application. Eck and Hoppe (1996) present a method for the recovery of splines from range scans, and their paper demonstrates the challenges that we outline. A parametric representation of a surface takes the following form:

$$S(u,v) = \big(f(u,v), g(u,v), h(u,v) \big) \tag{2.5}$$

where u and v parameterize the surface and the functions f, g, and h return the x, y, and z coordinates, respectively. For most complex objects, a single parameterized surface is not realizable and not practical for the entire object. Rather, we must divide the object into surface patches and then model each subsequent patch with its own parameterized surface. Thus, our first task is to segment the object into appropriate patches and then parameterize those patches individually. The major challenge is identifying appropriate patches, particularly, because parameterized surfaces assume a rectangular structure. Rectangles do not fit the arbitrary boundaries that often occur with patch selection. As a result, an additional step must trim the surface to these boundaries. With a patch identified, the fitting for recovering the parameters for f, g, and h is straightforward. Both direct and iterative optimization solutions are common. Figure 2.16, is a simple illustration of surface fitting for the now familiar brake. In this figure,
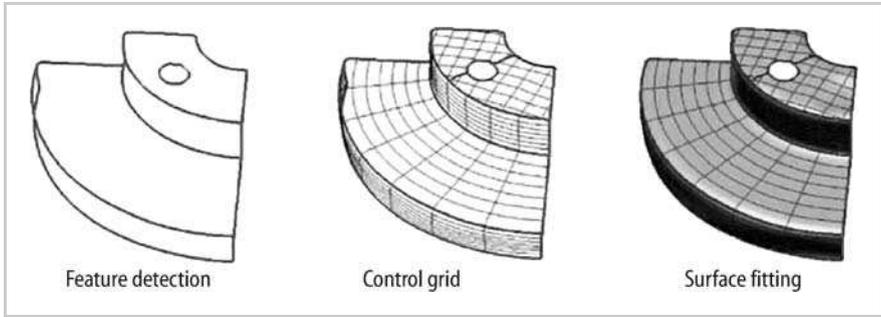
Figure 2.16. Typical surface fitting sequence. The leftmost image shows the feature detection to establish potential surface discontinuities. The next image to the right shows the control grid to support the surface fitting process. Finally, the rightmost image shows the result of fitting a higher order surface to the measurement data.

we use feature detection of surface discontinuities to serve as the patch boundaries and then fit NURBS to the resulting patches. This figure concludes this chapter and finalizes the CARE procedure, originally shown in Figure 2.1.

2.5 Conclusions

In this chapter, we have presented a definition for CARE with a brief background on the use of CMM technology. Based on this definition, we then outlined computer vision solutions as an alternative to CMMs. The focus of this discussion was laser range scanners. Of these various scanners, we identified the sheet-of-light scanner as the most suitable for CARE applications. Finally, we concluded this chapter with an overview of the processing pipeline that is necessary to capture 3-D data for an object and to translate that data into a CAD description. In closing, we reiterate the central theme of this chapter. Computer vision offers tremendous potential for advancing CARE technology to a point where an untrained user can generate CAD models of existing world objects. The caveat, however, is that improvements in measurement accuracy and commercial maturity are necessary to gain an advantage over existing CMM systems.

Acknowledgments

Chapter 3

Reverse Engineering–Hardware and Software

D.T. Pham and L.C. Hieu, Cardiff University, UK

3.1 Introduction

Reverse engineering (RE) is generally defined as a process of analyzing an object or existing system (hardware and software) to identify its components and their interrelationships and to investigate how it works to redesign or produce a copy without access to the design from which it was originally produced (Wikipedia, 2005). In areas related to 3-D graphics and modeling, RE technology is used for reconstructing 3-D models of an object in different geometric formats.

RE hardware is used for RE data acquisition, which for 3-D modeling, is the collection of geometric data that represent a physical object. There are three main technologies for RE data acquisition: contact, noncontact and destructive. Outputs of the RE data acquisition process are 2-D cross-sectional images and point clouds that define the geometry of an object.

RE software is employed to transform the RE data produced by RE hardware into 3-D geometric models. The final outputs of the RE data processing chain can be one of two types of 3-D data: (i) polygons or (ii) NURBS (non-uniform rational B-splines). Polygon models, which are normally in the STL, VRML, or DXF format, are commonly used for rapid prototyping, laser milling, 3-D graphics, simulation, and animations. NURBS surfaces or solids are frequently used in computer-aided design, manufacturing, and engineering (CAD-CAM-CAE) applications.

In this chapter, hardware and software for RE are presented. Commercially available RE hardware based on different 3-D data collection techniques is briefly introduced. The advantages and disadvantages of various data acquisition methods are outlined to help in selecting the right RE hardware for specific applications. In the RE software section, end-use RE applications are classified and typical commercial RE packages are reviewed. The four RE phases used in

a RE data processing chain are highlighted, and the fundamental RE operations that are necessary for completing the RE data processing chain are presented and discussed in detail.

3.2 Reverse Engineering Hardware

3.2.1 Contact Methods

Contact methods use sensing devices with mechanical arms, coordinate measurement machines (CMM), and computer numerical control (CNC) machines, to digitize a surface. There are two types of data collection techniques employed in contact methods:

(i)   point-to-point sensing with touch-trigger probes and

(ii) analogue sensing with scanning probes.

In the point-to-point sensing technique, a touch-trigger probe is used that is installed on a CMM or on an articulated mechanical arm to gather the coordinate points of a surface. A manually operated, articulated mechanical arm with a touch-trigger probe allows multiple degrees of freedom (DOF) of movement to collect the measurement points (Figure 3.2). A CMM with a touch-trigger probe can be programmed to follow planned paths along a surface. A CMM provides more accurate measurement data compared to the articulated arm. However, the limitation of using CMM is the lack of number of DOF so that a CMM cannot be used to digitize complex surfaces in the same way as an articulated arm.

In analogue sensing, a scanning probe is used that is installed on a CMM or CNC machine (Figure 3.3). The scanning probe provides a continuous deflection output that can be combined with the machine position to derive the location of the surface. When scanning, the probe stylus tip contacts the feature and then
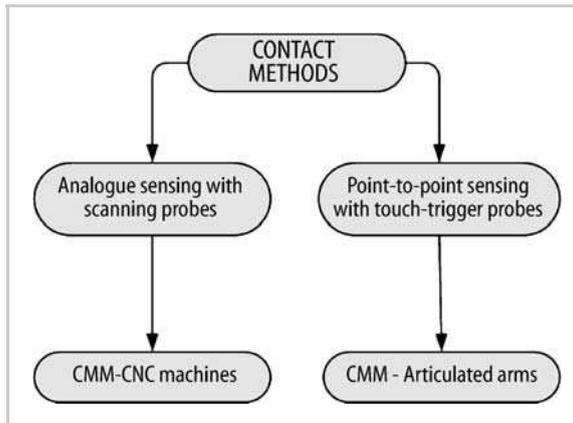
Figure 3.1. RE hardware classification–contact methods
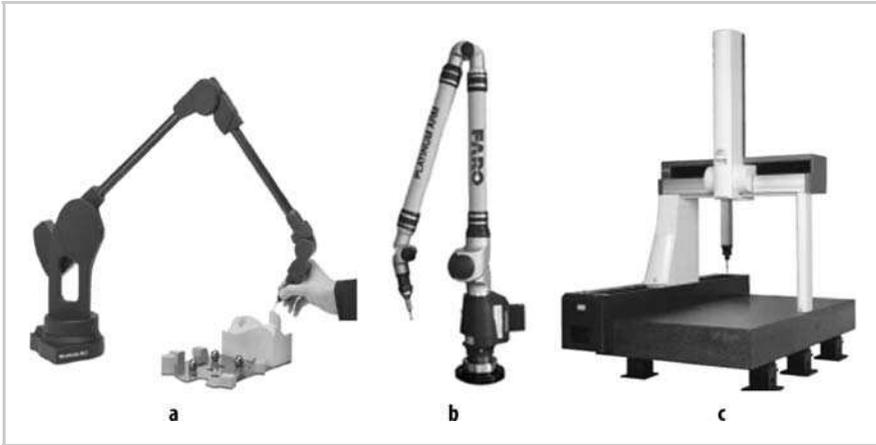
Figure 3.2. (**a**) MicroScribe MX Articulated Arm from Immersion Corporation (Immersion, 2005). (**b**) Faro Arm–Platinum articulated arm from FARO Technologies (FARO, 2005). (**c**) Mitutoyo CMM machine–CRA Apex C model (Mitutoya, 2005).
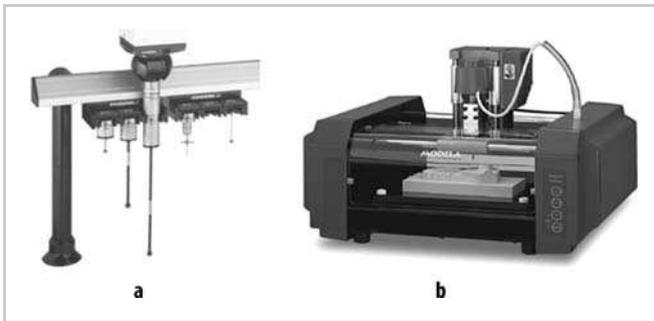
Figure 3.3. (**a**) SP25M scanning probes from Renishaw Inc (2005). (**b**) Roland DGA Corp. MDX-15/20 scanning and milling machine, using the Roland Active Piezo Sensor for 3-D scanning (Roland, 2005).

moves continuously along the surface, gathering data as it moves. Therefore, throughout the measurement, it is necessary to keep the deflection of the probe stylus within the measurement range of the probe. The scanning speed in analogue sensing is up to three times faster than in point-to-point sensing.

Table 3.1 gives examples of typical commercial RE hardware that employs contact methods for data acquisition. The advantages and disadvantages of contact methods compared to noncontact methods are as follows.

Advantages:
(i)   high accuracy,
(ii) low costs,
(iii) ability to measure deep slots and pockets, and
(iv) insensitivity to color or transparency.

Table 3.1. Contact methods–typical commercial RE hardware

Technology	Company	Model	Volume (mm)	Accuracy, resolution, and speed	Operation
Point-to-point sensing with a touch-trigger probe, mechanical arms	Faro Technologies	FaroArm Advantage	1200–3700	Accuracy: ±0.090 to ±0.431 mm	Manual
		FaroArm Platinum	1200–3700	Accuracy: ±0.018 to ±0.086 mm	
	Immersion Corp.	MicroScribe MX	1270	Accuracy: 0.1016 mm	Manual
		MicroScribe MLX	1670	Accuracy: 0.1270 mm	
Analogue sensing with a scanning probe, CNC machines	Roland DGA Corp.	Picza PIX-30	$305 \times 203 \times 60$	Scan pitch in Y,Y,Z axis: + (X, Y): 0.05–5.0 mm in steps of 0.05 mm. + Z: 0.025 mm	Programmed
		MDX-15	$150 \times 100 \times 60$		
		MDX-20	$200 \times 150 \times 60$		
Point-to-point sensing with a touch- trigger probe, CMM	Mitutoyo	Euro-C-121210	$1205 \times 1205 \times 1005$	Accuracy: 0.001 mm	Programmed
Analogue sensing with a scanning probe, CMM and CNC machines	Renishaw Inc.	Renscan 200	Based on the CMM and CNC machine volume	+ Speed: 508–1016 mm/min + Max data rate: 70 points/s	Programmed

Disadvantages:

(i) slow data collection and

(ii) distortion of soft objects by the probe.

3.2.2 Noncontact Methods

In noncontact methods, 2-D cross-sectional images and point clouds that represent the geometry of an object are captured by projecting energy sources (light, sound, or magnetic fields) onto an object; then either the transmitted or the reflected energy is observed. The geometric data for an object are finally calculated by using triangulation, time-of-flight, wave-interference information, and image processing algorithms. There is no contact between the RE hardware and an object during data acquisition.

There are different ways to classify RE hardware that uses noncontact RE methods for data acquisition. These classifications are based on the sensor technologies (Tamas *et al.* 2005) or data acquisition techniques (Alain 1999; Rocchini *et al.* 2001) employed. Figure 3.4 presents a classification of noncontact RE hardware based on data acquisition techniques.

The advantages and disadvantages of noncontact methods compared to contact methods are as follows.

Advantages:

(i)   no physical contact;

(ii) fast digitizing of substantial volumes;

(iii) good accuracy and resolution for common applications;
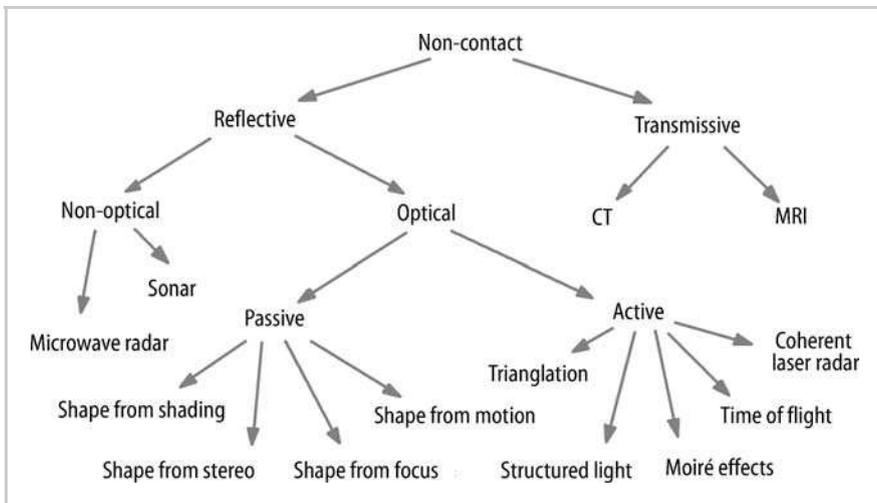
Figure 3.4. RE hardware classification–noncontact methods

(iv) ability to detect colors; and

(v) ability to scan highly detailed objects, where mechanical touch probes may be too large to accomplish the task.

Disadvantages:

(i) possible limitations for colored, transparent, or reflective surfaces and

(ii) lower accuracy.

Table 3.2 presents some typical commercial RE hardware using noncontact methods for data acquisition. The following sections introduce the most commonly available noncontact RE data acquisition techniques.

3.2.2.1 Optical Techniques

3.2.2.1.1 Triangulation

Most laser scanners use straightforward geometric triangulation to determine the surface coordinates of an object. Triangulation is a method that employs locations and angles between light sources and photosensitive devices (CCD–charge-coupled device camera) to calculate coordinates.

Figure 3.5. shows two variants of triangulation schemes using CCD cameras: single and double CCD camera. In a single camera system, a device transmits a light spot (or line) on the object at a defined angle. A CCD camera detects the
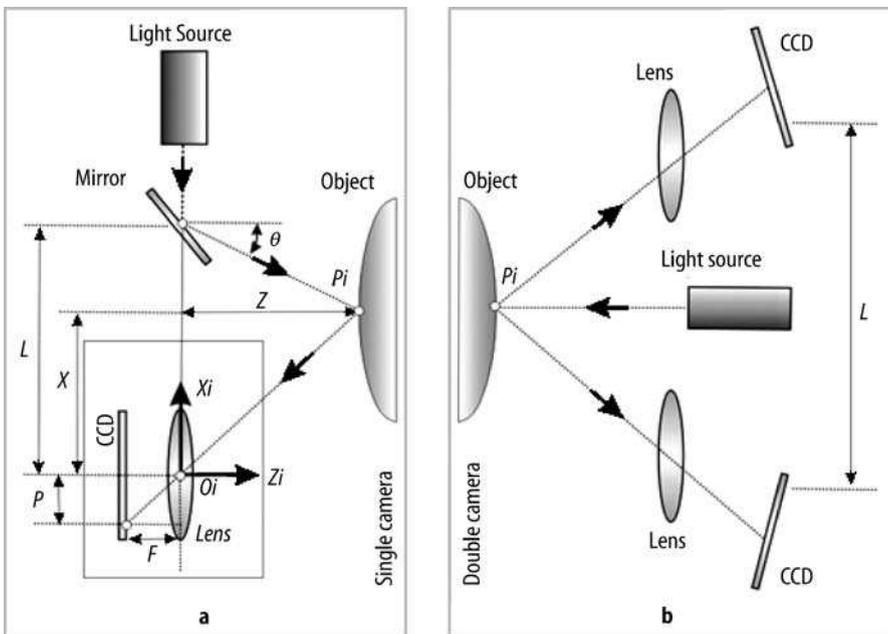
Figure 3.5. Triangulation methods: (**a**) single and (**b**) double camera arrangement

position of the reflected point (or line) on the surface. In a double camera system, two CCD cameras are used. The light projector is not involved in any measuring functions and may consist of a moving light spot or line, moving stripe patterns, or a static arbitrary pattern (Böhler *et al.* 2001).

The principle of the triangulation method is shown in Figure 3.5a. A high-energy light source is focused and projected at a prespecified angle (θ) onto the surface of an object. A photosensitive device senses the reflection from the illuminated point on the surface. Because the fixed baseline length (L) between the light source and the camera is known from calibration, using geometric triangulation from the known angle (θ), the focal length of the camera (F), the image coordinate of the illuminated point (P), and fixed baseline length (L), the position of the illuminated point (Pi) with respect to the camera coordinate system can be calculated as follows (Park and DeSouza 2005):

$$Z = \frac{FL}{P + F\tan\theta}$$

$$X = L - Z\tan\theta$$

The measurement errors in P and θ can be determined from the following equation:

$$\Delta Z = \frac{Z^2}{FL}\Delta P + \frac{Z^2\sec^2\theta}{L}\Delta\theta$$

The error in the Z measurement is directly proportional to Z^2 but inversely proportional to the focal length and the baseline length. Therefore, increasing the baseline length can produce higher accuracy in the measurement. For practical reasons, the baseline length cannot be increased at will, and it is limited by the hardware structure of the scanners. Therefore, triangulation scanners are commonly used for scanning small objects over short distances.

If the single-point or sheet-of-light pattern is used as the light source, the triangulation scanner is mounted on the travel platform so that it can produce multiple surface scans. Triangulation scanners are supplied both as complete systems and as self-contained scanning heads for mounting on standard touch-probe arms or CMMs.

3.2.2.1.2 Structured Light

In structured-light techniques (Park *et al.* 2001; Pagès *et al.* 2003; Caspi *et al.* 1998; Page *et al.* 2003; Szymon *et al.* 2002; Chen and Kak 1987; Joaquim *et al.* 2004; Salvi *et al.* 1998; Kiyasu *et al.* 1995; Grin *et al.* 1992; Morano *et al.* 1998), a light pattern is projected at a known angle onto the surface of interest and an image of the resulting pattern, reflected by the surface, is captured. The image is then analyzed to calculate the coordinates of the data point on the surface.

A light pattern can be (i) a single point; (ii) a sheet of light (line); and (iii) a strip, grid, or more complex coded light (Caspi *et al.* 1998) (Figure 3.6).

The most commonly used pattern is a sheet of light that is generated by fanning out a light beam. When a sheet of light intersects an object, a line of light is formed along the contour of the object. This line is detected and the X,Y,Z coordinates of hundreds of points along the line are simultaneously calculated by triangulation. The sheet of light sweeps the object as the linear slide carrying the scanning system moves it in the X direction while a sequence of images is taken by the camera in discrete steps. An index number k is assigned to each of the images in the order they are taken. Therefore, each k corresponds to the X position of the sheet of light. For each image k, a set of image coordinates (i, j) of the pixels in the illuminated stripe is obtained. The triples (i, j, k)'s are the range image coordinates; they are transformed to (x, y, z) world coordinates using a calibration matrix.

To improve the capturing process, a light pattern containing multiple strips is projected onto the surface of an object. To distinguish between different strips, they must be coded approximately so that the correspondence problem is solved without ambiguity (Park *et al.* 2001; Pagès *et al.* 2003; Caspi *et al.* 1998; Page *et al.* 2003; Szymon *et al.* 2002; Chen and Kak 1987; Joaquim *et al.* 2004; Salvi *et al.* 1998; Kiyasu *et al.* 1995; Grin *et al.* 1992; Morano *et al.* 1998).

Structured-light systems have the following strong advantages compared to laser systems, and these features have resulted in favoring structured-light systems for digitizing images of human beings:

(i)   the data acquisition is very fast (up to millions of points per second)

(ii) color texture information is available

(iii) structured-light systems do not use a laser.
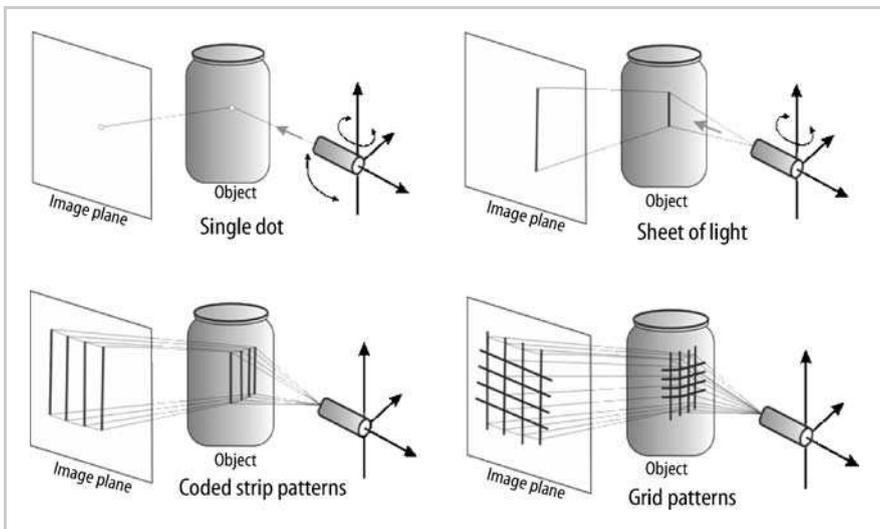
Figure 3.6. Different light patterns used in structured-light techniques
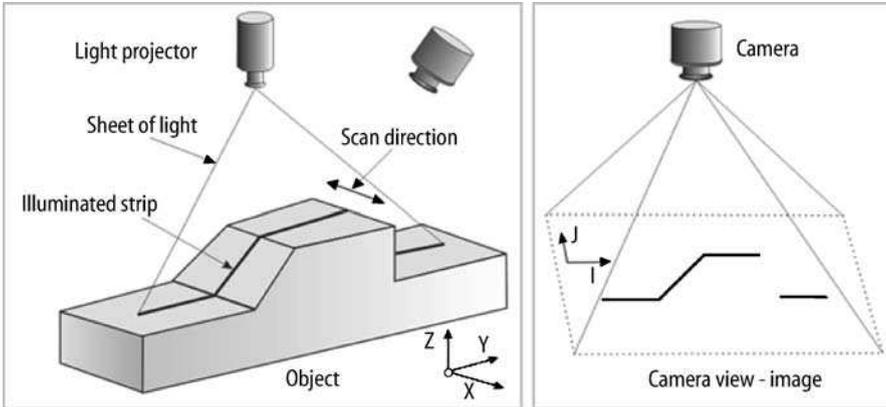
Figure 3.7. 2-D image acquisition in the structured-light technique using a sheet-of-light pattern

3.2.2.1.3 Interferometry (Moiré Effects)

The interferometry technique is well known in dimensional inspection as well as in flatness and deformation measurements (Reid *et al.* 1988; Pallek *et al.* 1999; Suzuki and Kanaya 1988), in which structured-light patterns are projected onto a surface to produce shadow moiré effects (Creath and Wyant 1992; Kafri and Glatt 1978). The light contours produced by moiré effects are captured in an image and analyzed to determine distances between the lines. This distance is proportional to the height of the surface at the point of interest, and so the surface coordinates can be calculated.

The moiré technique gives accurate results for 3-D reconstruction and measurement of small objects and surfaces. However, it has limitations for larger objects because precision is sacrificed for range.

Figure 3.8. shows the formation of moiré fringes by superimposing a line pattern with concentric circles and two other line patterns that vary in line spacing and rotation.
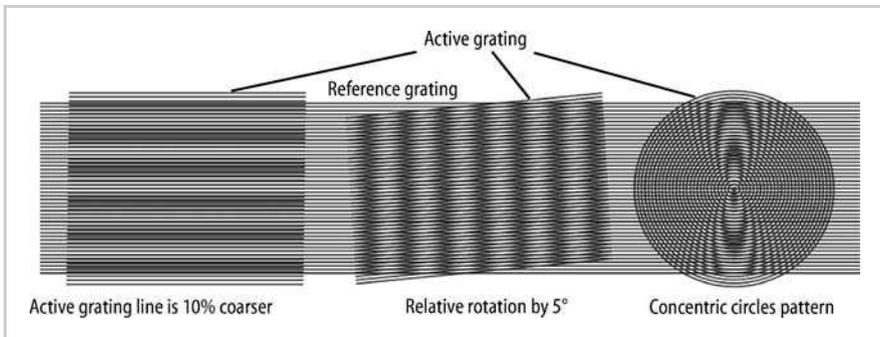
Figure 3.8. Formation of moiré fringes

3.2.2.1.4 Time of Flight

The principle behind all time-of-flight (TOF) (Bellian *et al.* 2005; Sekimoto *et al.* 2004; Bruno and Nick 2004; Lichti and Harvey 2002; Szymon *et al.* 2002) implementations is to measure the amount of time (t) that a light pulse (*i.e.*, laser electromagnetic radiation) takes to travel to the object and return. Because the speed of light (C) is known, it is possible to determine the distance traveled. The distance (D) of the object from the laser would then be equal to approximately one half of the distance the laser pulse traveled: $D = C \times t/2$.

Figure 3.9 illustrates in block diagram form how a time-of-flight laser scanner works. For all practical purposes, the angle θ is very small and thus has no effect on the accuracy of the TOF distance measurement. The high velocity of light allows TOF scanners to make hundreds, or even thousands of measurements per second. The advantage of TOF techniques is that they can digitize large, distant objects such as buildings and bridges. The accuracy of RE hardware based on TOF is reasonable and approximately between a few millimeters and two or three centimeters for long-range scanners. The accuracy depends on the pulse width of the laser, the speed of the detector, and the timing resolution; the shorter the pulse and the faster the detector, the higher the accuracy of the measurement.

The main disadvantage is that TOF scanners are large and do not capture an object's texture, only its geometry. They are not practical for fast digitization of small and medium-sized objects. Moreover, it takes time to complete the digitization process because the object (or environment) has to be swept during scanning.

A variation on the TOF method is the phase shift method for determining distance measurements. Distance is computed by comparing the phase shift between an emitted wavelength and the received light. The Surphaser Model 25 developed by Surphaser Inc. (2005) is a typical commercial system. The accuracy of a phase-shift system is higher than that of traditional TOF machines. The range accuracy of the Surphaser Model 25 is 25 µm, and the angular accuracy is
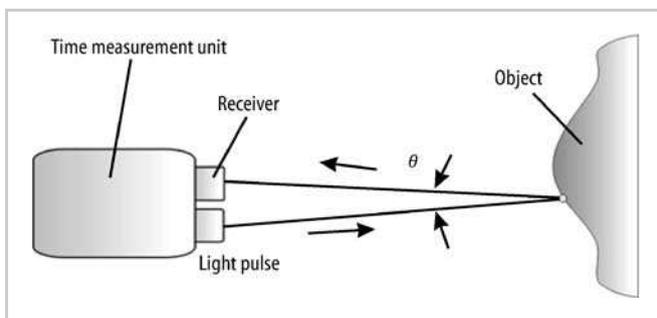
Figure 3.9. Principle of TOF scanners

0.003°. An important consequence of using phase-shift detection is that the system uses a single line of sight for its work. This means that the laser light travels the same path from the scanner to the surface and back again which enables scanning the inside of holes, cavities, and concave surfaces.

3.2.2.1.5 Passive Methods

Passive methods reconstruct a 3-D model of an object by analyzing the images to determine coordinate data. It is similar to (active) structured-light methods in its use of imaging frames for 3-D reconstruction; however, in passive methods, there is no projection of light sources onto the object for data acquisition.

There are many different passive methods, including shape from shading, shape from stereo, shape from motion, shape from focus/defocus, shape from silhouette, and volumetric reconstruction. The typical passive methods are shape from shading (Horn and Brooks 1989; Horn 1990; Reinhard *et al.* 1999; Kimmel *et al.* 1995; Bichsel and Pentland 1992; Dupuis and Oliensis 1992; Oliensis and Dupuis 1993; Lee and Kuo 1993) and shape from stereo (Hoff and Ahuja 1989; Chang *et al.* 2000; Sun 2002).

Shapes from shading (SFS) methods are used to reconstruct a 3-D representation of an object from a single image (2-D input) based on shading information. The first SFS technique was developed by Horn in the early 1970s (Horn and Brooks 1989; Horn 1990). These are the main disadvantages of this method (Park and DeSouza 2005):

(i) the shadow areas of an object cannot be recovered reliably because they do not provide enough intensity information;
(ii) the method cannot be applied to general objects because it assumes that the entire surface of an object has the same reflectance;
(iii) the method is very sensitive to noise because the computation of surface gradients is involved.

Shape from stereo or stereovision refers to the extension of SFS to a class of methods that use two or more images from different viewpoints for shading-based 3-D shape recovery. Normally, two cameras are coordinated to generate 3-D information about an object by automatically finding corresponding features in each of the two images; then triangulation is used to measure the distance to objects containing these features by intersecting the lines of sight from each camera to the object. Compared to SFS methods, there is improved accuracy. However, finding correspondence between images is extremely difficult and can produce erroneous results from mismatches.

To solve the problem of finding correspondence, stereovision techniques can be combined with color structured-light techniques for 3-D range data acquisition (Chen *et al.* 1997).

Although they require very simple hardware, passive methods do not produce accurate 3-D data. Active optical methods can overcome many of the problems in passive methods and thus result in more accurate solutions.

3.2.2.1.6 Coherent Laser Radar

Recently, the advent of a new type of laser radar frequency-modulated coherent laser radar (FMCLR), created a new generation of FMCLR instruments. They can measure large-scale geometry precisely. A typical commercial RE machine in this category is a MetricVision system (MV224 and MV260 models) from Metris (2005). The accuracy (2σ) of the MetricVision system is 16 μm at 1 m, 100 μm at 10 m, and 240 μm at 24 m. The MetricVision system operates by using a sensor to direct a focused invisible infrared laser beam to a point and coherently processes the reflected light. As the laser light travels to and from the target, it also travels through a reference path of calibrated optical fiber in an environmentally controlled module. The two paths are combined to determine the absolute range to the point. A very wide laser-modulation bandwidth (100 GHz) makes precise measurement possible on a millisecond timescale. The distance measurement is then combined with positional information from two precision encoders to determine a point on a surface in space.

3.2.2.2 Nonoptical Techniques

Nonoptical approaches include acoustic (active sonar) (Sea-Image 2005; Johnson and Herbert 1990, 1996; Zerr and Stage 1996) and microwave radar (radio detecting and ranging) (Valle *et al.* 2000). The principle of these techniques for 3-D reconstruction is measuring distances from the sensing device to objects by measuring the time delay between the transmitted and returned signals.

Sonar techniques are normally used in 3-D underwater mapping; they provide mariners with a major advancement in obstacle avoidance and navigation. Sonar range sensors are inexpensive, but their accuracy is not high and they do not have high acquisition speeds. Acoustic interference or noise is often a problem, as is determining the focused point location.

Radar is typically intended for use with long-range remote sensing, especially in airline applications. Commercial airliners are equipped with radar devices that warn of obstacles in or approaching their path and give accurate altitude readings. In 3-D reconstruction applications, radar is used to measure distances and map geographic areas and to navigate and fix positions at sea.

3.2.2.3 Transitive Techniques

Computerized tomography (CT) is a powerful transmissive approach for 3-D reconstruction. CT has revolutionized the medical diagnostic field since the 1970s (Kak and Malcolm 2001; ASTM 1992; Stewart and Bu 2000; Brooks and Di Chiro 1976). It has also been called computerized axial tomography (CAT), computerized transaxial tomography (CTAT), and digital axial tomography (DAT). CT is a nondestructive method that allows three-dimensional visualization of the internals of an object. It provides a large series of 2-D X-ray cross-sectional images taken around a single rotational axis.

Figure 3.10 presents the CT working principle of generating 2-D cross-sectional images. By projecting a thin X-ray or Y-ray beam through one plane of an object from many different angles and measuring the amount of radiation that passes through the object along various lines of sight, a map of attenuation coefficients (a density map or cross-sectional image) for the scanned surface is reconstructed.

CT is widely used for medical applications; however, it has been extended and adapted to a wide variety of industrial and 3-D modeling tasks (Iovea *et al.* 1994; Beall *et al.* 1996; Johns *et al.* 1993; Ketcham and Carlson 2001, Michael *et al.* 1985). Today, industrial CT (Iovea *et al.* 1994; Paulusa *et al.* 2000; Toshiba 2005; Akira 2001) and related technologies (digital computed laminography) (Gondrom and Schropfer 1999; Rooks *et al.* 1995; Gondrom *et al.* 1999) are commercially available and specialized for industrial applications. High-resolution X-ray CT and micro CT scanners can resolve details as small as a few tens of microns, even when imaging objects are made of high-density materials. It is applicable to a wide range of materials, including rock, bone, ceramic, metal, and soft tissue.

Magnetic resonance imaging (MRI) (Donald *et al.* 2004; Mark and Richard 2003) is a state-of-the-art imaging technology that uses magnetic fields and radio waves to create high-quality, cross-sectional images of the body without using radiation. When hydrogen protons in the human body are placed in a strong magnetic field, by sending in (and stopping) electromagnetic radio-frequency pulses, these protons emit signals. These signals are collected and processed to construct cross-sectional images. Compared to CT, MRI gives superior quality images of soft tissues such as organs, muscle, cartilage, ligaments, and tendons in many parts of the body.

CT and MRI are powerful techniques for medical imaging and reverse engineering applications; however, they are the most expensive in terms of both hardware and software for data processing.
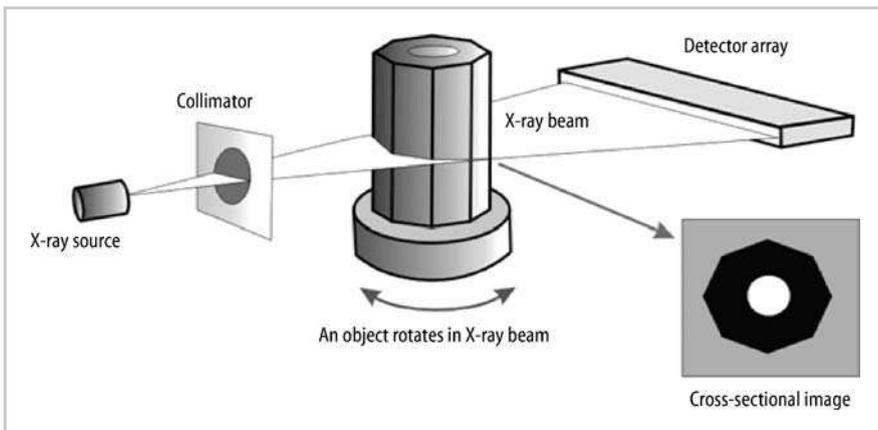
Figure 3.10. Working principle of the CT scanner

3.2.3 Destructive Method

The RE destructive method is useful for reverse engineering small and complex objects in which both internal and external features are scanned. A CNC milling machine exposes 2-D cross-sectional (slice) images, which are then gathered by a CCD camera. The scanning software automatically converts the digital bitmap image to edge detected points, as the part is scanned. The company, CGI Inc., produces a destructive system and calls this technology cross-sectional scanning (CSS) (CGI 2005).

In RP processes, the part is built layer-by-layer based on 2-D slice data (Pham and Dimov 2001). The destructive RE process is the reverse of this. To remodel the part, 2-D slice images of the part are gathered by destroying the part layer-by-layer.

The data acquisition procedure of a destructive system is presented in Figure 3.11. The disadvantage of this method is the destruction of the object. However, the technique is fast. The accuracy is acceptable; the repeatability is ±0.0127 mm (CGI systems). The layer thickness is from 0.0127–0.254 mm. The method allows capturing internal structures. In addition, a destructive system can work with any machinable object, including parts made from aluminum alloys, plastics, steel, cast iron, stainless steel, copper, and wood.
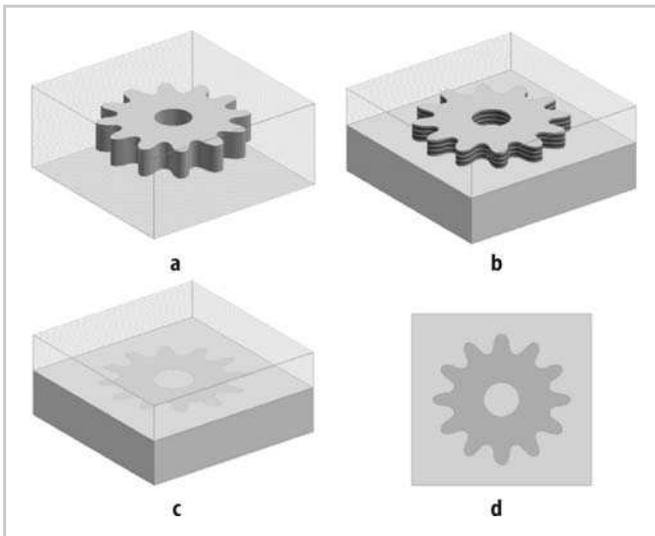
Figure 3.11. A procedure for data acquisition in a destructive RE system. (**a**) The part and matrix combination is embedded in a contrasting colored plastic matrix. (**b**, **c**) The part is machined layer-by-layer to expose the cross-sectional images. (**d**) The newly exposed surface of the part. A color reproduction of this figure can be seen in the color section (pages 219–230).

Table 3.2. Noncontact methods – typical commercial RE hardware

Technology	Company	Model	Volume, field of view, envelope (mm)	Accuracy, resolution, pitch (mm)	Color	Speed (points/s)	Main applications
Laser triangulation	Cyberware	3030RGB ScanHead	FOV: $260 \times 340 \times 300$ FOV = Y (at minimum Z) $\times Y$ (at maximum Z) $\times Z$.	Sampling pitch X: 0.250–1.000 Y: 0.700 Z: 0.100 (minimum)	Yes	14,580	Head and face 3-D color scanning. medical, anthropometry, human interface, and portrait sculpture.
Laser triangulation	Cyberware	3030RGB MS Model Shop 3-D Scanner	FOV: $1000 \times 260 \times 300$ FOV = Y (at Minimum Z) $\times$ Y (at Maximum Z) $\times$ Z.	Sampling Pitch: X: 0.250–1.000 Y: 0.700 Z: 0.100 (minimum)	Yes	14,580	Scanning large sculptures, maquettes, and industrial models.
Laser triangulation	Cyberware	WB4 Whole Body Colour 3-D Scanner	Cylindrical FOV: Depth: 120 Height: 200	Sampling pitch: Horizontal (X): 5.0 Vertical (Y): 2.0 Depth (Z): 0.5	Yes	60,000	Animation, fashion design; anthropology, animation
Laser triangulation	Konica Minolta	VIVID 910	Scanning Volume: Tele: $111 \times 84 \times 40$ Mid: $196 \times 153 \times 70$ Wide: $355 \times 266 \times 92$ (lens at 0.6 m)	X: ± 0.220, Y: ± 0.160, Z: ± 0.100 to the Z reference plane. (Condition: Tele or fine mode)	Yes	122,800 and 256,000 (fine and fast mode)	Common RE and quality inspection: 3-D graphics and animation, sculpturing, product development and design, tool and die modification, medical modeling, *etc*.

Table 3.2. (continued)

Technology	Company	Model	Volume, field of view, envelope (mm)	Accuracy, resolu-tion, pitch (mm)	Color	Speed (points/s)	Main applications
Laser triangulation	Konica Minolta	VIVID 700	Scanned area $(X \times Y)$: 70×70 to 1100×1100	Resolution (X, Y, Z): $200 \times 200 \times 256$ points	Yes	66,666	Common RE and inspec-tion
Interferometry	Optimet	Mini-Conoscan 3000	Scanning volume: $120 \times 120 \times 0.2$ to $120 \times 120 \times 70$	Accuracy: 0.5 to 25 μm	Yes	850	Common RE and quality inspection
Laser triangulation	Perceptron	ScanWorks Lite	Stand-off: 107 Volume: based on CMM	Accuracy: ± 0.050 at 3σ.	No	23,040	Common RE and quality inspection.
Transmission: X-ray computed tomography	Aracor	ICT 2500	Max. object diameter: 2500	Pixel size: 1.220 Resolution: 3.000 Accuracy: 0.250	No	$1024 \times 1024 \times 1024$ scan in 16 h	Internal viewing, common RE, and quality inspection
Coherent laser radar	Metris	MV224 and MV260	Range: MV224: 1–24 m and MV260: 1– 60 m. FOV: $\pm 45 \times 360°$	Accuracy (2σ): 16 μm at 1 m. 100 μm at 10 m. 240 μm at 24 m	No	Max: 1000	Static and dynamic inspections of aircraft, automotive, and heavy equipment tooling assemblies.
Laser triangulation	Steinbichler Optotech-nik	COMET T-Scan	Measuring depth: 75 Scan width: 90 Mean distance: 83	Accuracy: ± 0.03. Length error: ± 0.1 at 3 m distance to a tracking system.	No	Max: 6,666	Common RE and inspec-tion

Table 3.2. (continued)

Technology	Company	Model	Volume, field of view, envelope (mm)	Accuracy, resolution, pitch (mm)	Color	Speed (points/s)	Main applications
Structured light: Triangulation with white light fringe projection	Steinbichler Optotechnik	COMET C100 VZ	Measuring area in high resolution and standard mode: 100×80	Accuracy: ±0.020	No	Max: 6,666	Common RE and inspection
Structured light: Triangulation with white light fringe projection	Steinbichler Optotechnik	C400 VZ	Measuring area in high resolution and standard mode: 420×340	Accuracy: ±0.070	No	Max: 6,666	Common RE and inspection
Time of flight: Long-range 3-D laser scanner	Riegl	LMS Z210i	Range: 4–400 m. FOV: 80°×360° (vertical×horizontal angle)	Accuracy: ±25.000 (single shot) ±15.000 (averaged)	Yes	8,000–12,000	Process automation and robotics, topography, city modeling and urban planning.
Time of flight: Long-range 3-D laser scanner	Riegl	LMS Z420i	Range: 2–1000 m. FOV: 80°×360° (vertical×horizontal angle)	Accuracy: ±10.000 (single shot) ±5.000 (averaged)	Yes	8,000–12,000	Topography, architectural measurement, as-built surveying, archaeology and cultural heritage documentation, city modeling.

Table 3.2. (continued)

Technology	Company	Model	Volume, field of view, envelope (mm)	Accuracy, resolution, pitch (mm)	Color	Speed (points/s)	Main applications
Time of flight: Long-range 3-D laser scanner	Callidus Precision Systems GmbH	CP 3200	Max range: 40 or 80 m FOV = 40°×360° to 180°×360° (vertical×horizontal angle)	Accuracy: 5.000	Yes	1,750	Process automation and robotics, architectural and facade measurement, civil engineering, city modeling, mining, geology.
Structured light: Triangulation with a white-light pattern projection	Genex	Rainbow 250	FOV: 250×200×200	Accuracy: 0.250	Yes	442,368	Common RE and inspection, face recognition, medicine and dentistry.
Structured light: Triangulation with a white-light pattern projection	Genex	Rainbow 25	FOV: 32×25×20	Accuracy: 0.025	Yes	442,368	Common RE and inspection, face recognition, medicine and dentistry.
Laser triangulation	Metris	LC15	Width of view: 15 Depth of view: 15 Standoff: 70	Accuracy: 8 μm (1σ sphere fit)	No	19200	Inspection of small and detailed objects; RE with specific needs for highest accuracy.

Table 3.2. (continued)

Technology	Company	Model	Volume, field of view, envelope (mm)	Accuracy, resolution, pitch (mm)	Color	Speed (points/s)	Main applications
Time of flight-phase Shift: Short-range 3-D laser scanner	Surphaser	Surphaser Model 25	Max Range: 400–5,000 FOV: 150–3,000 View angle: 22°, 30°, 40°	Range accuracy: 0.025 Angular accuracy: 0.0030	Yes	200,000	Common RE and inspection, scanning the inside of holes, cavities,. and concave surfaces
Transmission: X-ray computed tomography	Aracor	Konoscope 200	Max object diameter: 200	Pixel size: 0.200 Resolution: 0.400 Accuracy: 0.040	No	$1024 \times 1024 \times 1024$ scan in 16 h	Internal viewing, common RE, and quality inspection
Transmission: X-ray computed tomography	Aracor	Konoscope 160	Max object diameter: 160	Pixel size: 0.100 Resolution: 0.250 Accuracy: 0.020	No	$1024 \times 1024 \times 1024$ scan in 16 h	Internal viewing, common RE, and quality inspection
Interferometry	Inspeck	3-D Mega Capturor II	FOV: 435×350 Standoff distance: 900 Reference distance ($z = 0$): 1,100	Resolution: $X = 0.300$ $Y = 0.300$ $Z = 0.400$	Yes	1.3 million points in 0.7 s.	Common RE and quality inspection
Interferometry	Inspeck	3-D Capturor II	FOV: 360×270 Standoff distance: 900 Reference distance ($z = 0$): 1,100	Resolution: $X = 0.600$ $Y = 0.600$ $Z = 0.400$	Yes	0.3 million points in 0.4 s.	Common RE and quality inspection

Table 3.2. (continued)

Technology	Company	Model	Volume, field of view, envelope (mm)	Accuracy, resolution, pitch (mm)	Color	Speed (points/s)	Main applications
Laser triangulation	Callidus Precision Systems GmbH	CT 900	- Max. object size: 1600×1400 Measuring range: + Horizontal: 180 + Depth: 700	Accuracy: 0.100 Resolution: + Near: 0.010 + Far: 0.070	No	4,000	Common RE and quality inspection
Laser triangulation	Callidus Precision Systems GmbH	CT180	Max. object size: 350×375. Measuring range: + Horizontal: 180 + Depth: 700	Accuracy: 0.1 Resolution: + Near: 0.025 + Far: 0.070	No	4,000	Common RE and quality inspection
Transmission: X-ray computed tomography	Micro Photonics Inc.	Skyscan 1172	Max. object diameter: 68	Pixel size: 0.9–1.6 µm Resolution: 5–8 µm Smallest detail detectability: 1 µm	No	6.8 s per cross section 1024×1024 pixels	Biomedical, biological specimens, medical imaging, bone analysis, nondestructive inspection, diamonds, electronic components and packaging.
Laser triangulation	Metris	XC50 cross scanner	Width of view: 3×15 Depth of view: 3×15 Standoff: 70	Accuracy: 15 µm (1σ sphere fit)	No	3×64000	A multistrip laser sensor is used: more efficient scanning of features such as holes, slots, and gaps that are typically inspected in automotive applications.

3.3 Reverse Engineering Software

3.3.1 Reverse Engineering Software Classification

There is no single RE software that can completely satisfy the requirements of RE data processing and geometric modeling. The selection of RE software depends on the specific requirements of RE projects.

Table 3.3. Reverse engineering software classification based on application

No.	Application	Main functions	Software
1	Hardware control	Control RE hardware for data acquisition. Normally, basic data processing operations and data conversion are also provided.	Mitutoyo Cosmos, Hymarc, Metris Scan, Cyberware CyDir and GSI Crystal Studio.
2	CAD entity manipulation	Manipulate CAD entities that are extracted from point clouds and polygon meshes. CAD entities include points, contour lines, and CAD primitives such as circles, rectangles, cylinders, and boxes.	ICEM surf, Imageware and other common CAD packages such as UG, Pro Engineer and Solidworks.
3	Polygon manipulation	3-D polygon data editing, modification, and optimization.	Magics RP, DeskArtes, Catia Shape Sculptor and Viscam RP.
4	Polygon and NURBS surface construction	Provide a complete set of RE data processing tools from working with point clouds and polygons to constructing NURBS surfaces as well as 3-D inspection.	GSI Studio, CopyCAD, Rapidform, Geomagics, Polyworks (Modeler) and Paraform.
5	2-D Scan Image Processing and 3-D modeling	Used for processing 2-D scan images (CT/MRI) and 3-D reconstruction.	Mimics, Rapidform, BioBuild, Velocity2, Amira, Scan IP, Analyze and 3-D Doctors.
6	3-D Inspection	Used for 3-D inspection, error map creation and analysis, inspection report and documentation.	COMETinspect, Metris Focus Inspection, Power INSPECT, PolyWorks Inspector and Geomagic Qualify.
7	NURBS surface and solid modeling	Provide NURBS modeling and editing tools based on basic CAD entities and primitives.	Pro Engineers, UG, Solidworks, Catia and Rhino.

Based on applications, RE software can be classified into the following groups: Hardware control, CAD entity manipulation, polygon manipulation, polygon and NURBS surface construction, 2-D scan image processing and 3-D modeling, 3-D inspection, and NURBS surface and solid modeling. Table 3.3 presents these RE software groups with representative commercial packages.

3.3.2 Reverse Engineering Phases

For an overall view of RE software operation, the different RE data processing phases will first be described. The required RE operations are then considered. The complete RE data processing chain, from scan data to the final NURBS

POINT AND IMAGES PHASE

- **Registration:** Manual and Automatic Alignment
- **Data Optimisation:** Noise and point redundancy reduction, Sampling points and Identifying primitives
- **Basic Operations:** Rotate and Move, Datum Control and Separating and Merging Point Clouds
- **Image processing:** Region Growing and Thresholding

POLYGON PHASE

- **Polygon Optimization:** Noise Reduction and Cleaning, Abnormal Faces Cleaning, Polygon Mesh Refinement and Polygon Mesh Decimation
- **Polygon Edit and Control:** Filling Holes, Defeaturing, Edge Detection and Sharpening Control, Primitives Fitting, Polygon Editing and Remeshing, and Boundary Control and Editing
- **Basic Operations:** Rotate and Move, Datum Control, Boolean, Offset, Shell, Thicken, Cut and Mirror

CURVE PHASE

- **Primitive Fitting:** Circle, Cylinder and Plane
- **Curves Construction:** Cross Section and Curve Fitting from Points
- **Curve Modification & Editing:** Curve Re-parameterization, Curve Degree Conversion, Curve Smooth and Clean, Control Point Edit, Transition and Extension, Point Generation and Curve Re-direction

NURBS SURFACE PHASE

- **Surfaces from Curves:** Loft, UV-Network & Extrude
- **Patch Creation & Control:** Curvature Detect, Patch Editing and Patch Template Re-use
- **NURBS Surface Creation and controls:** Grids, NURBS Patches Merging, NURBS surface Smoothing and Editing
- **Evaluation:** Point to CAD, Polygon to CAD, CAD to CAD

Figure 3.12. Four phases of the RE data processing chain with fundamental RE operations

model, can be divided into four main phases: points and images, polygon, curves, and NURBS surfaces.

Figure 3.12 presents the four phases of the RE data processing chain with the fundamental RE operations. These RE operations are necessary and are available with the most commonly used commercial RE software such as GSI Studio, Geomagics, CopyCAD, Rapidform, Polyworks, Paraform, ICEM surf, and Magics RP (Delcam 2005; Geomagic 2005; INUS 2005; Paraform 2005; Magics 2005; GSI 2005; Polyworks 2005).

3.3.2.1 Points and Images Phase

In the points and images phase, scan data are registered, prepared, and optimized for constructing 3-D polygon models. Figure 3.13 shows a flowchart for transforming RE scan data into 3-D polygon models.

Outputs from the RE data acquisition process are 2-D cross-sectional images or point clouds. RE systems that use transitive techniques such as CT and MRI provide a large series of 2-D cross-sectional images of an object. Systems that use the remaining RE techniques such as laser triangulation, TOF, and structured-light provide point cloud data.

In medical applications, CT/MRI images in the DICOM format (2005) are used as the input for image processing and 3-D reconstruction of anatomical
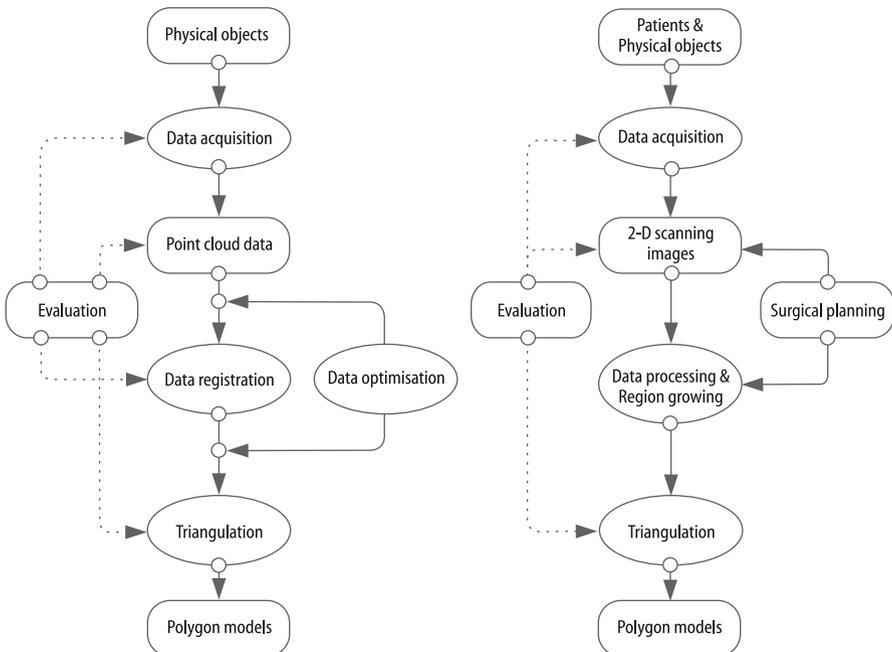
Figure 3.13. A flowchart for transforming RE scan data into 3-D triangle mesh models. (**a**) Point cloud data as the input. (**b**) 2-D scan images as the input.

structures. When working with 2-D cross-sectional images, segmentation by thresholding techniques can be used to define a region of interest that presents the object for 3-D reconstruction; segmentation is based on the gray-level value of image pixels. The object can be defined by using one or two thresholds. In the former case, the segmented object will contain all pixels in the images with a gray-level value higher (or lower) than or equal to the threshold. In the latter case, the pixel gray-level value must be between both thresholds to be part of the segmented object.

A region growing technique provides the capacity to split the segmentation into separate objects; for example, it allows creating separate models or 3-D views of a crane and the mandible, and the acetabulum and the femur. Figure 3.14 presents a segmented image and models of anatomical structures of a head; these 3-D models were reconstructed from CT scan images by using region growing and thresholding techniques.

Surgical planning based on patient scan data is normally required, especially in complex operations (Hieu *et al.* 2002, 2003, 2005). Clinical and technical constraints resulting from surgical planning are used for the data processing. Finally, the optimal point cloud data and region of interest are transformed into 3-D polygon models in the polygon phase (Figures 3.12 and 3.13).
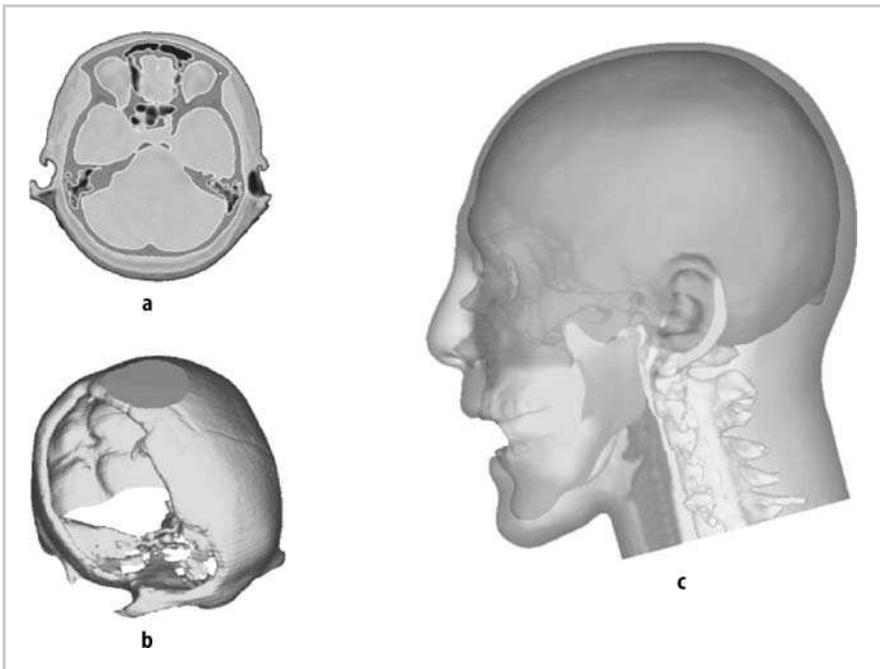
Figure 3.14. (**a**) A skull defined by the thresholding technique. (**b**, **c**) 3-D models of the cranial defect skull (**b**), and other anatomical structures of the head (**c**). A color reproduction of this figure can be seen in the Color Section (pages 219–230).

3.3.2.2 Polygon Phase

In this phase, polygon models are constructed. They are then manipulated and controlled to meet the requirements of the applications. The resulting 3-D polygon models are directly employed for applications such as rapid prototyping, 3-D graphics and animations, or are used as reference data for creating CAD entities (points, curves, and primitives) and constructing NURBS surfaces for CAD-CAM-CAE applications.

3.3.2.3 Curve Phase

In many RE projects, especially RE of mechanical parts, CAD entities are mainly used as the reference data for geometric modeling in CAD packages. The CAD entities are constructed directly from point clouds or indirectly from polygon models by manual editing, fitting, and sectioning operations (Figure 3.12). They are finally imported into CAD packages to complete the geometric modeling.

For simple geometries, a limited number of reference points may be sufficient to model the part using a CAD package (Figure 3.15a), and contact methods with mechanical touch probes are normally employed to collect the data. Curves are necessary for creating complex parts with free-form surfaces (Figure 3.15b); in this case, more point data are needed, and therefore, noncontact methods are likely to be used for data acquisition.
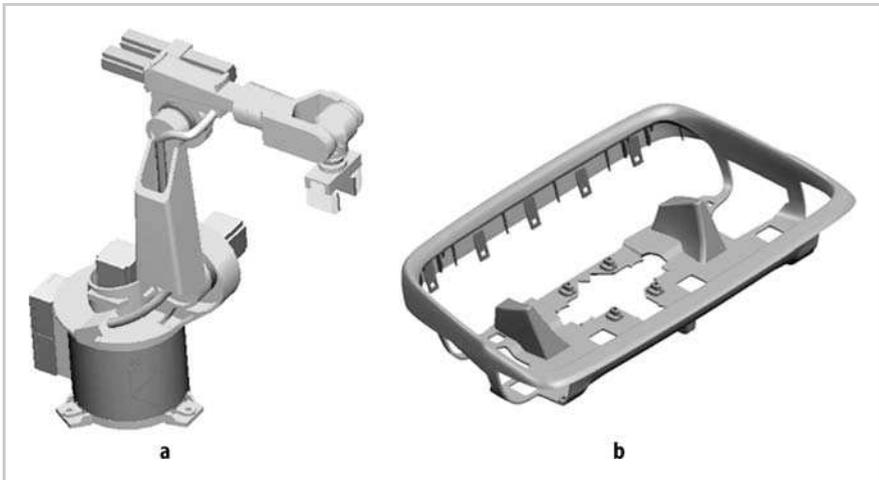
Figure 3.15. (**a**) An RE model of a mechanical part with simple geometries (KUKA robot). (**b**) An RE model of a complex part with free-form surfaces.

3.3.2.4 NURBS Surface Phase

NURBS surfaces are sometimes the ultimate RE output for CAD-CAM-CAE applications. NURBS surfaces can be constructed based on the CAD entities extracted from the curve phase or by using polygon meshes for surface fitting.

NURBS are an accurate way to define free-form curves and surfaces. NURBS are useful for the following reasons: (i) they offer one common mathematical form for both standard analytical shapes and free-form shapes; (ii) they provide the flexibility to design a large variety of shapes; (iii) they reduce the memory consumption when storing shapes (compared to simpler methods); (iv) they can be evaluated reasonably fast by numerically stable and accurate algorithms; (v) they are invariant under affine as well as perspective transformations; and (vi) they are generalizations of nonrational B-splines and nonrational and rational Bézier curves and surfaces (David 2001). Therefore, NURBS are commonly used in CAD-CAM-CAE systems.

A flowchart summarizing the steps in NURBS construction in a RE data processing chain is presented in Figure 3.16. There are three approaches for creating NURBS surfaces: (i) manual creation of NURBS from basic CAD entities; (ii) manual creation of NURBS from patches; and (iii) automatic creation of NURBS from polygon models. Finally, constructed NURBS surfaces can be imported into CAD/CAM systems to build NURB CAD solid models.
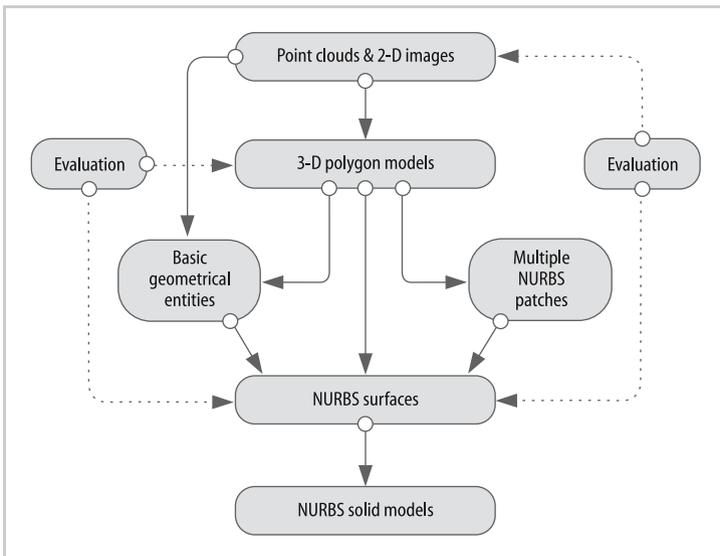
Figure 3.16. Different RE approaches for NURBS construction

3.3.2.4.1 Manual Creation of NURBS from Basic CAD Entities

This RE approach is the same as CAD modeling, in which CAD models are built from basic CAD entities: points, primitives, and curves. The fundamental difference is that, in RE projects, CAD entities are constructed based on RE scan data in the points and images phase as well as in the curve phase; they are then imported into CAD packages as references for modeling the object. In traditional CAD modeling approaches, CAD entities are directly constructed based on technical drawings or specifically assigned dimensions. Figure 3.15 presents NURBS solid models of an industrial robot and a free-form surface part using a manual creation of NURBS from basic CAD entities.

3.3.2.4.2 Manual Creation of NURBS from Patches

In this approach, the model is fitted with a patch structure of quadrangular shapes, from which NURBS patches are constructed. A patch is defined by four polylines traced on the polygonal surface. Figure 3.17 shows (a) the model of an infant's head with a NURBS patch layout, (b) NURBS patches with UV grids of control points, and (c) a NURBS solid model.

The quality of the NURBS surfaces depends on the layout of the patches and the number of control points in the U and V directions for defining a NURBS patch surface.

This is the most efficient NURBS surface fitting approach, especially when working with very complex shapes such as anatomical structures (Figure 3.14c). The optimal patch layout helps generating good grids of control points for constructing NURBS patches. A smaller number of control points will produce an inaccurate NURBS surface for the patch; on the other hand, one with many control points will create a CAD model with a large file size. Depending on the required accuracy of the final NURBS model, the number of control points for the
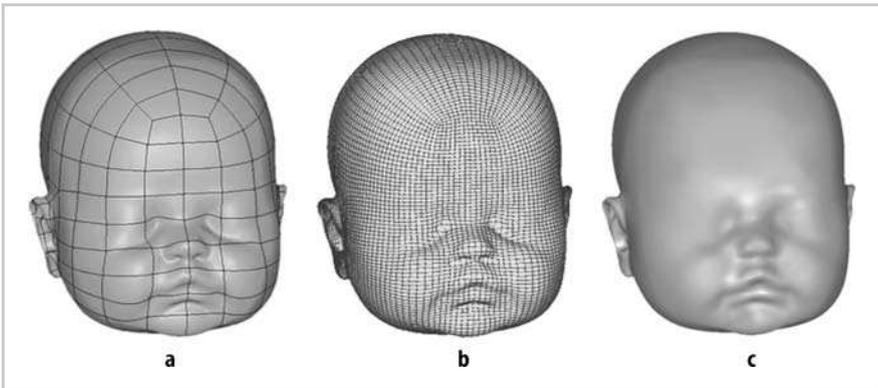
Figure 3.17. (**a**) An infant's head model with a NURBS patch layout. (**b**) NURBS patches with UV grids of control points. (**c**) A NURBS solid model of the head.
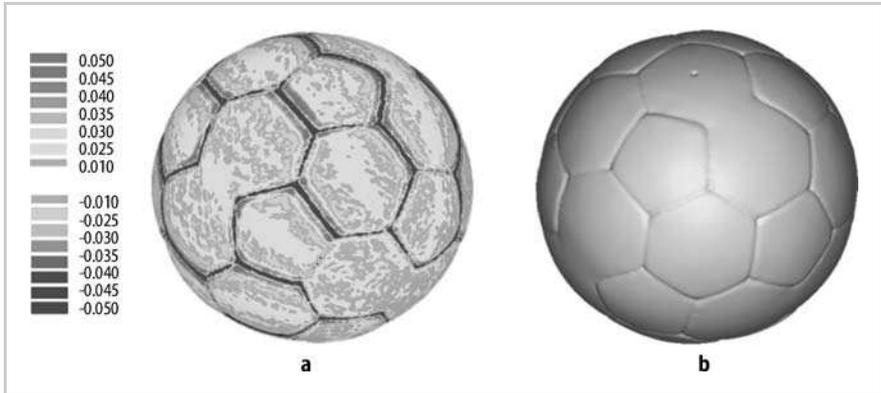
Figure 3.18. An error map (**a**) between an original polygon football model and (**b**) a resulting NURBS model. A color reproduction of this figure can be seen in the Color Section (pages 219–230).

NURB patch is flexibly adjusted. The error map (Figure 3.18) between a resulting NURBS model with an original scan point cloud (or a polygon mesh) is normally used to evaluate and verify the final result.

3.3.2.4.3 Automatic Creation of NURBS from Polygon Models

Most of the common RE software packages provide a one-step process for creating NURBS surfaces from a polygon model. The process combines the most frequently used operations in the NURBS patch approach. This automated method gives quick results. However, it is suitable only for simple geometries or when a draft NURBS surface model is needed for a specific application.

3.3.3 Fundamental Reverse Engineering Operations

3.3.3.1 Points and Images Phase

3.3.3.1.1 Data Registration

Although most scanners allow scanning an object from different angles with certain provided degrees of freedom, multiple scans of the object are required to capture the entire geometry of an object or to avoid any occlusions (undercuts). When using different scan setups, the point cloud from one series of scans is not accurately oriented with respect to the point cloud from another series. Data registration is needed to combine, align, or merge these point clouds so that all point clouds in the series are arranged in their proper orientation relative to one another in a common coordinate system.

Registration is very important for downstream RE data processing steps. Therefore, it is necessary to verify the output data carefully. If the number of point clouds is not enough to cover the entire geometry of an object or the
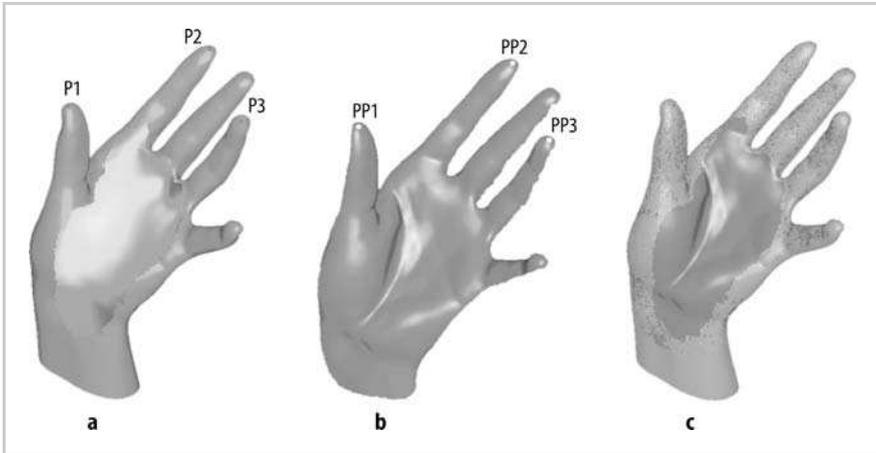
Figure 3.19. (**a**) The first point cloud with three landmark points (P1, P2, and P3). (**b**) The second point cloud with three landmark points (PP1, PP2, and PP3) that correspond to those of the first point cloud. (**c**) The resulting point cloud after registration. A color reproduction of this figure can be seen in the Color Section (pages 219–230).

expected tolerances are not achieved, more data acquisition as well as scanning schemes are required (Figure 3.13a). Data optimization may be applied to prepare good point clouds or triangulation meshes for registration.

There are two data registration approaches: manual and automatic alignment. In manual alignment, landmark points are manually assigned for fixed and floating point clouds, and they are used as references for alignment. Floating point clouds are translated and rotated to align with the fixed cloud based on these references. Figure 3.19 illustrates a manual registration process using three landmark points. The manual registration results in multiple point clouds that need to be fine-tuned to obtain an optimal cloud. This is done in the automatic alignment operation, in which the tolerance between the fixed and floating point clouds is used as the constraint for the alignment process.

3.3.3.1.2 Data Optimization

(i) Noise and Point Redundancy Reduction
A certain amount of error is always introduced into scan data, and points may be placed in undesirable regions or overlap because of points that have been scanned more than once when scanning complex shapes. Moreover, when point cloud registration is applied, the aligned scan data normally contains overlapping points.

Noise reduction tools are used for both manually and automatically removing the noise in scanned data. With automatic approaches, the noise removal operation determines where the points should lie, then moves them to these locations based on statistical information about the point data. If the point set represents a free-form or organic shape, the operation reduces the noise with respect to
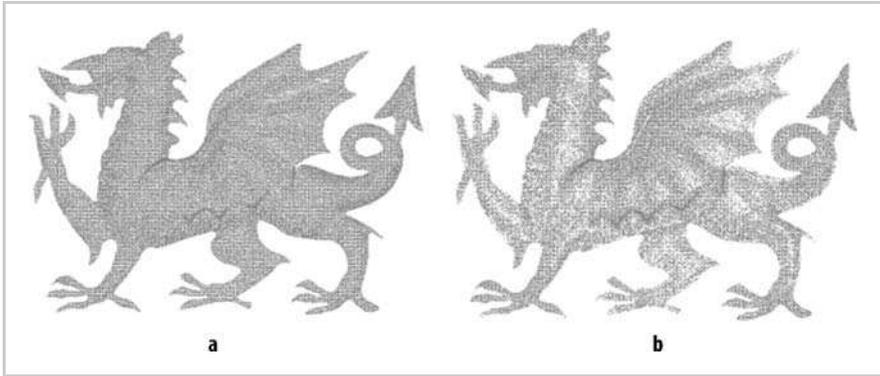
Figure 3.20. Images before (**a**) and after (**b**) applying curvature sampling to a point set

surface curvature. When working with a mechanical or prismatic shape, the operation helps keep features such as edges and sharp corners.

The redundancy reduction tool is used to reduce the number of points in the point cloud when points are too close to one another or overlap.

(ii) Sampling Points

The sampling function is used to minimize the number of points in the point cloud data and to make the data well-structured so that it is easier to handle.

There are three sampling methods: curvature, random, and uniform; they are based on a curvature, random, and proportional basis. In curvature sampling (Figure 3.20), fewer points are deleted in a region of high curvature than in a low curvature region to maintain the accuracy of the curvature. Random sampling is a random sampling of points within a specified region or over the entire model, based on the percentage of total points that need to be reduced. Uniform sampling uniformly reduces the number of points in a point set; it subdivides the model space into equally sized cubic cells and deletes all but one point from each cell.

(iii) Identifying Primitives

Identifying primitives such as planes, cylinders, and spheres is very important in the RE process, especially when working with mechanical shapes. This helps to correct imperfections in the scan data, such as noise captured during the scanning process. The primitive creation operation inserts mathematically perfect primitives within a model.

3.3.3.2 Polygon Phase

3.3.3.2.1 Optimizing Polygon Models

The main objective of these operations is to improve the quality of the polygon models. The most common and important operations for optimizing polygon models are noise reduction and cleaning, abnormal face cleaning, polygon mesh refinement, and polygon mesh decimation.

(i) Noise Reduction and Cleaning

Noise reduction in the polygon phase is the same concept as that in the point phase; noise introduced into polygon models is filtered and removed. The operation reduces noise with respect to the surface curvature or helps keep features such as edges and sharp corners. Normally, different levels of noise reduction can be selected so that the mesh quality is controlled. The operation uses statistics such as the maximum distance, average distance, and standard deviation of the points from their original positions.

Cleaning makes the polygonal surface conform to the shape designated by the point set. Depending on the situation, the operation can remove dents, smooth cylindrical sections, or sharpen edges, as well as produce a very good mesh and perform a degree of relaxation or smoothing.

(ii) Abnormal Face Cleaning

This operation cleans up abnormal faces in polygon models. Typical abnormal faces in the polygon phase include the following:

- Nonmanifold faces: In manifold topology, two faces can define an edge. When faces share three or more edges, the topology is nonmanifold.
- Redundant faces: This occurs when a number of faces and edges share the same vertex.
- Crossing faces: This occurs when edges that share a vertex intersect with one another.

(iii) Polygon Mesh Refinement and Decimation

Polygon mesh refinement improves the surface of a polygon model by adding new vertices and adjusting the coordinates of existing vertices, resulting in a greater number of triangles in the selected region and a smoother surface. The
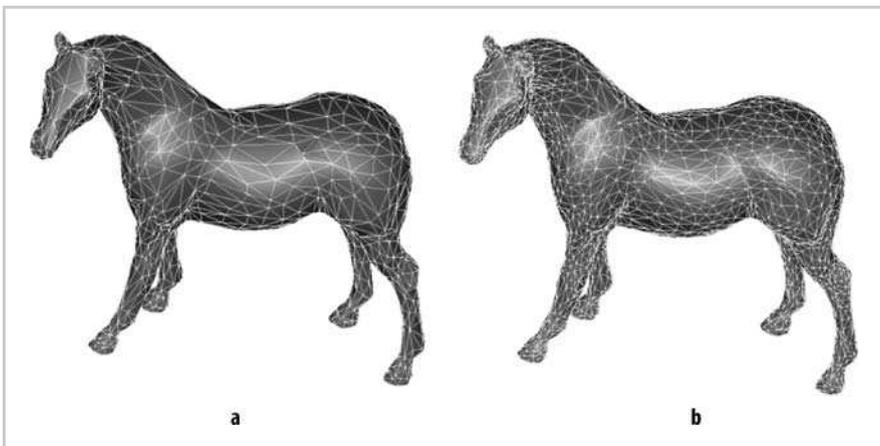
Figure 3.21. (**a**) A horse model with 2500 triangles created by applying polygon decimation to (**b**) the same model with 6000 triangles

operation subdivides selected triangles, producing either three or four triangles for every original triangle. On the other hand, polygon mesh decimation reduces the number of triangles without compromising surface integrity or detail (Figure 3.21).

3.3.3.2.2 Editing and Controlling Polygon Models

The main objective of these operations is to modify the triangles and add new features to the polygon mesh. The following operations are necessary for editing and controlling polygon models: filling holes, defeaturing, edge detection and control, primitives fitting, polygon editing and remeshing, boundary control, and basic operations (Boolean, cut, mirror, offset, rotate, and move).

(i) Filling Holes

This operation fills gaps in a model that may have been introduced during scanning or because of errors in converting point clouds into polygon models (Figure 3.22).

A polygonal structure is used to fill the hole, and both the hole and the surrounding region are remeshed so that the polygonal layout is organized and continuous. A curvature-based filling option ensures that the polygonal structure used to fill holes in high-curvature regions is curved to match the surrounding area.

(ii) Defeaturing

This operation allows refitting selected regions with a new triangulated polygon surface using a curvature-based method. Features in the selected region are deleted. The operation is useful for refining and smoothing the selected region.

a   b
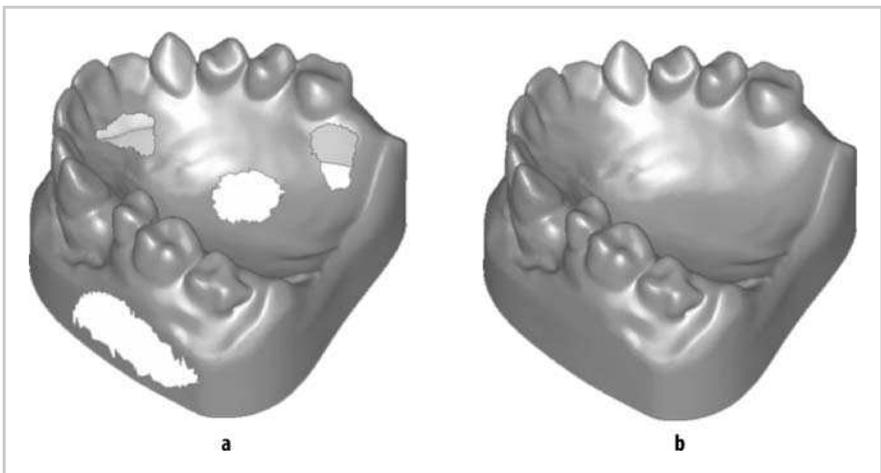
Figure 3.22. Images before (**a**) and after (**b**) the hole filling operation

(iii) Edge Detection and Sharpening Control
Scanning devices are often unable adequately to capture sharp features, such as creases or mechanical edges. When transforming to a polygon mesh, these features are usually filleted or rounded. This operation reproduces an edge by redefining it mathematically and then extending the polygon model to that newly defined sharp edge.

(iv) Primitives Fitting
The primitives fitting operation fits selected regions to primitives such as planes, cylinders, and spheres. It is especially useful for mechanical parts.

(v) Polygon Editing and Remeshing
Polygon editing and remeshing provides precise control over the polygonal mesh structure down to the triangle level and includes the five following options: splitting edge, flipping edge, contracting edge, adding triangle, and remeshing. The splitting edge option divides an edge into two parts, adding two additional triangles on either side of the selected edge. The flipping edge option moves an edge to another location in the quadrangular patch, altering the structure of the patch. However, there is only one possible location in each patch to which the edge can be moved. The contracting edge option removes individual triangle edges by contracting between two vertices. The adding triangle option inserts another triangle into the polygonal mesh based on three insertion points.

The quality of a polygon mesh depends on the form of the triangles that constitute it. A perfect mesh would consist of triangles that are all equilateral. The remeshing option allows optimizing the triangle mesh. Figure 3.23 shows the mesh before and after remeshing.

(vi) Boundary Control and Editing
Boundary control and editing are necessary when working with open polygon surfaces. This operation is used to repair the boundary edges of a model. It
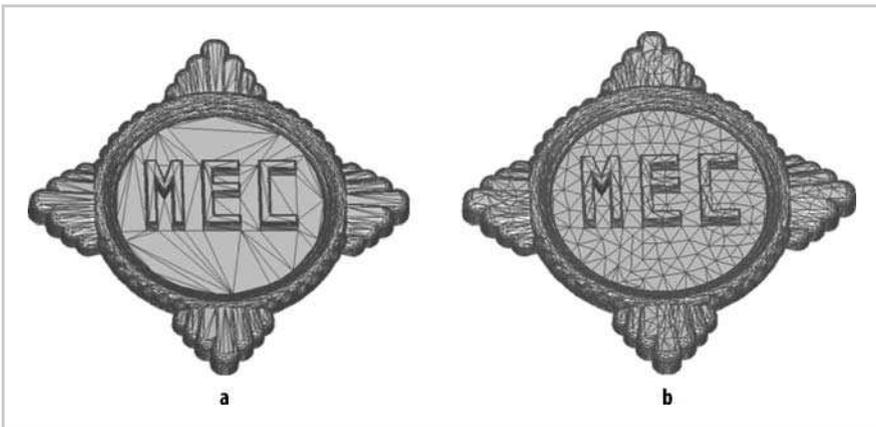
Figure 3.23. A polygon model (**a**) before and (**b**) after applying the remeshing option
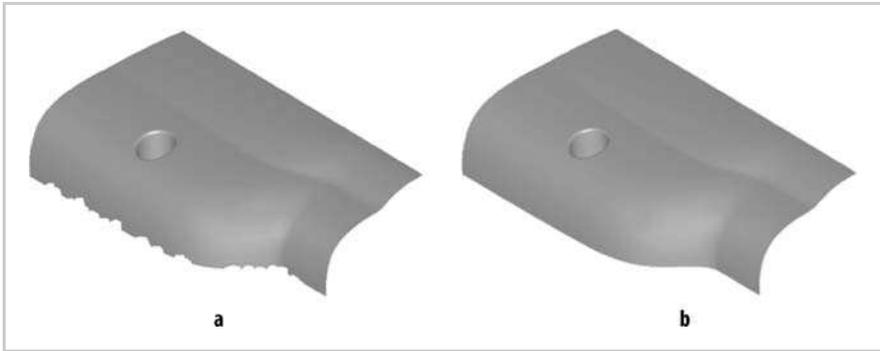
Figure 3.24. (a) The part with a bad boundary and (**b**) the good one created by boundary control and editing

provides the ability to mark edges of triangles to create boundaries, which are a series of triangle edges. Normally, the newly created boundary can be edited and modified by changing the number of control points, moving the control points, or adjusting the curvature of the boundary.

The editing and modifying options for a boundary include straightening, smoothing, fitting it to primitives and paths, and projecting it onto a user-defined plane (Figure 3.24). These operations move the boundary away from triangle edges or marked edges, and the surrounding polygonal structure is readjusted to accommodate the new boundary.

(vii) Basic Polygon Operations
The basic polygon operations are very familiar in CAD modeling packages. They include Boolean, trim, shell, offset, thicken, cut, mirror, rotate and move, and datum control.

Boolean operations allow union, subtraction, and intersection of open and closed polygon models. Trim cuts off the selected region defined by given reference curves. Offset creates a new surface offset from the original that maintains the original width. Shell increases the width of a model along a trajectory. Thicken converts an open shell into a volume shell with a constant thickness. Cut allows separation of selected polygons. Mirror creates mirrored polygons from selected ones. Finally, rotate, move, and datum control allow repositioning models in the global or local coordinate system, as well as controlling the datum for flexible data manipulation.

3.3.3.3 Curve Phase

The NURBS surface is defined by a network of curves. Therefore, when applying the "Manual Creation of NURBS from Basic CAD Entities" approach (Figure 3.16), point clouds and polygon models are normally used for creating curves, especially for parts with free-form surfaces.

Basic CAD entities such as circles, rectangles, and lines can be easily created by manual fitting methods based on reference points. The following operations are the most important in the curve phase: cross section, 3-D curve fitting from points, and curve modification.

Cross-sectioning creates cross-sectional points or curves based on a plane that intersects the model (Figure 3.25). The 3-D curve fitting from points operation allows creating a curve from a set of points. Curve modification allows controlling a curve to meet modeling requirements. The most useful curve modification options are as follows:

(i) **Curve Reparameterization**: Changing the number of control points or redistributing control points along the curve.
(ii) **Curve Degree Conversion**: Changing the degree of a curve with a specified tolerance.
(iii) **Curve Smoothing and Cleaning**: Smoothing a curve, cleaning, and removing unnecessary control points.
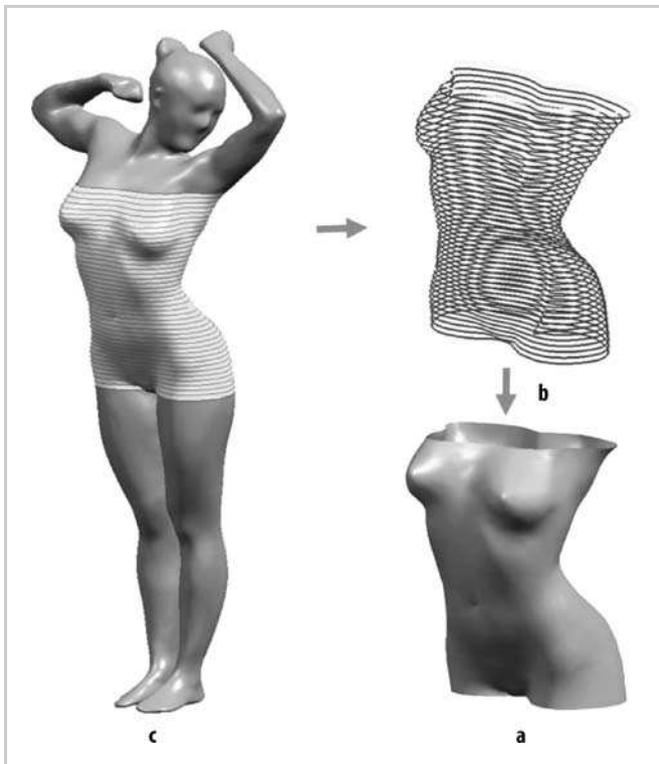
Figure 3.25. (**a**) A NURBS surface model of a woman's body (**b**) constructed by lofting curves (**c**) that were created by cross-sectioning a polygon model

(iv) **Control Point Editing**: Modifying the control points manually. The control points are moved to the specified positions to change the shape of a curve as desired.

(v) **Point Generation**: Creating a specified number of points from a curve with random or uniform distribution.

(vi) **Curve Redirection, Transition, and Extension**: Changing the direction of the curves, stitching two curves together to make a new one, and extending a curve to a point or distance with tangent or curvature continuity.

3.3.3.4 NURBS Surface Phase

The main objective of the NURBS surface phase in the RE data processing chain is to prepare a patch structure of quadrangular shapes for NURBS surface construction. The patches (or loops) can be drawn on the polygon model manually, semiautomatically, or automatically based on a target patch count and the curvature of a model. Although most RE software packages provide an automatic method to generate patch structures quickly, manual organization of the patch structure is always required to obtain the optimal patch layout because, in many cases, features such as edges and primitives cannot be recognized and the patches do not adequately define the surface curvature.

The most useful RE operations in the NURBS surface phase are curvature detection, patch editing, patch template reuse, and NURBS patch merging.

Curvature detection identifies components of the model based on surface curvature. The boundary (or feature) curves that define the components of the model are automatically detected. They are then manually edited to produce optimal curves. The boundary curves eventually define panels of rectangular patches. Preparing a good boundary curve layout ensures that the future patch structure maintains all necessary topological detail. Patch editing helps the optimal organization of a patch structure. Figure 3.26 shows examples of boundary curves, the patches designated by boundary curves, and the UV meshes of control points that define NURBS patches.

Figure 3.26. (**a**) Boundary curves define the components of the model. (**b**) Boundary curves designate panels of rectangular patches. (**c**) The UV meshes of control points define the NURBS patches of the component.

Patch template reuse saves the patch structure from one model and applies it to another model. This helps speed up the surfacing process for a model whose geometry is similar to an already surfaced model. It is especially useful for creating the patch structure for a model that contains symmetrical features (Figures 3.17 and 3.26). In this case, the patch layout for one half is mirrored to create the patch layout for the other half. Finally, the complete patch layout is saved as the template, which is finally applied to the model to construct a complete patch structure.

NURBS patch merging combines two or more NURBS patches into a single patch. This helps reduce the number of patches on the model and makes the data structure simpler for the final NURBS surface model.

3.4 Conclusion

In this chapter, fundamental information on RE hardware and software was presented, and commercial technologies for RE data acquisition and processing were introduced.

Today, there is a variety of RE hardware for reverse engineering both the internal and external structures of an object. However, selecting the right RE hardware and software for an RE project is not easy. Of the many selection criteria, the nature of the application is the most important.

Although the chapter presented limited application examples, RE applications have been introduced in many areas. Typical application include (i) inspection: Quality control and inspection for mechanical parts and machine components; (ii) product development and manufacture: CAD-CAM-CAE, rapid prototyping, and rapid tooling; (iii) biomechanics and medical applications: Biomechanics research on bone structures, design and manufacturing of medical models, surgical aid tools, implants, and medical devices; (iv) 3-D graphics and animations: Virtual reality, games, and films; (v) fashion design and arts: 3-D art modeling, portrait sculpturing, and prototyping; and (vi) topography, architectural and facade measurement, as-built surveying, archaeology and cultural heritage documentation, and city modeling.

There are three main technical specifications when selecting RE hardware: volume (field of view), accuracy, and speed. There is no single software package that can completely satisfy all requirements in RE data manipulation and modeling work. For inspection and polygon based applications such as rapid prototyping, 3-D graphics and animation, one RE package may be enough to complete the data processing chain. However, for most RE projects, especially in product development (CAD-CAM-CAE), at least two software packages are normally required: one for RE data manipulation and the other for the geometric modeling work, which is based on CAD entities or NURBS surfaces constructed from the RE data processing chain.

With advances in computer graphics and geometric modeling technologies, many new powerful RE operations and modules are introduced regularly in RE and CAD software. These help to reduce the RE data processing time and close the gap for data exchange and manipulation between traditional RE and CAD packages. Finally, in selecting RE software and CAD models for an RE project; the information provided in this chapter should help evaluate computing solutions.

Chapter 4

Selecting a Reverse Engineering System

John Keast, University of Warwick, UK

Abstract

In recent years, the market for systems to generate geometric representations of physical objects has exploded in size and complexity, as has the number of potential applications. Applications range from scanning the bodies and tracking the motion of actors for computer animated films through surveying buildings, plants, and landscape, medical applications from brain and prenatal scans to rapid prototyping of prosthetic devices and models of the patient for rehearsal of surgery, to the more conventional engineering applications of inspection, reverse engineering of obsolete parts, measurement of deformation on crash testing, and so on. Each application has unique requirements.

The common thread in all these applications is that the starting point is the gathering of accurate x-y-z data over the surface of the object, and then from this, the construction of a geometric model in a computer system that can then be used in a variety of ways. Because of the wide variety of systems and applications, finding the right system for a particular application becomes difficult. This chapter offers a structured process for system selection which should lead to a sound system purchase. The different approaches to each stage of the reverse engineering process are classified and discussed, as are the pros and cons of each. Because the author has no links to any particular vendor, the chapter does not attempt to endorse or recommend any one approach or any one system. Each has a realm of application for which, given the constraints of cost, accuracy, and speed, it may be the best solution. The trick is to match the application requirements with the unique set of capabilities of a system!

4.1 Introduction

Reverse engineering, the practice of studying an existing object to create a replica or a model of the object, is no longer a new concept. What has changed in recent years is that bewildering arrays of technologies have been made available at affordable prices. Because many software packages are available for reverse engineering software, that is, decompiling code, this chapter is not concerned with this form of reverse engineering. The chapter focuses on the tools and techniques for capturing and re-creating either an electronic model or a physical replica of one or more physical objects. The starting point for all of these tools and techniques is the acquisition of a set of *X-Y-Z* coordinates in space, called "points", that are then manipulated in software to provide an appropriate output for the re-creation phase of the process.

But this is where the problems start, for now, a plethora of measuring devices using a variety of measuring methods are available to measure a set of points, usually called a point cloud, and to store them in a computer. Techniques are available for just about anything from teeth to landscapes, model cars to full-size aircraft, buildings, humans, and so on. The range of applications of the reverse engineering tools and techniques is equally broad. In the realm of mechanical engineering alone reverse engineering approaches can be used for

- Inspection of parts and comparison of the actual geometry with computer-aided design (CAD) models
- Capturing the shape of an existing part, possibly one that is obsolete, so that a replacement part can be manufactured, perhaps through a rapid prototyping approach, perhaps through numerically controlled (NC) machining of surfaces interpolated through the captured points
- Capturing the deformed shape of an object, perhaps after some sort of impact test, to compare the actual deformed shape with that predicted by finite element analysis
- Capturing the geometry of an existing factory or process plant, to allow using a CAD system for piping; heating, ventilation, and air conditioning (HVAC) design, and so on
- Wear measurement of tooling

But if the list of mechanical engineering applications alone seems endless, there are applications in medicine, such as prosthetic design and manufacture, archaeology, animation, civil engineering, and many other fields. Thus, the attempt to create a generic selection approach that is equally applicable to all of the above, without adaptation, is almost inevitably doomed to failure. Consequently, this chapter will focus primarily on the mechanical engineering domain, but will make comments about other domains where necessary.
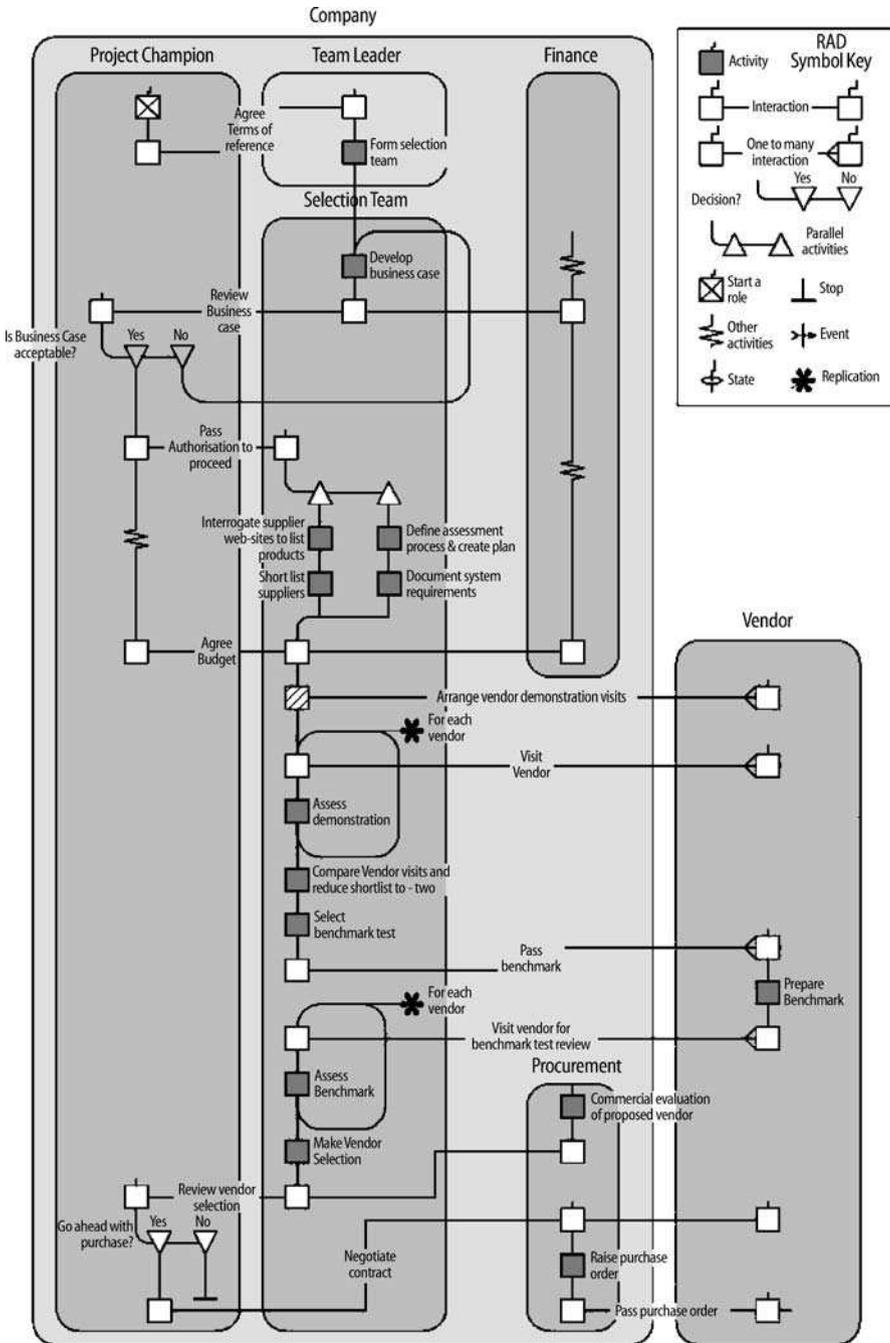
Figure 4.1. A role activity diagram (RAD) of a typical selection process

Further, even a cursory study of available reverse engineering hardware and software reveals everything from "build your own" instructions from a few cheap, readily available parts, through to highly accurate systems costing hundreds of thousands of dollars. The effort put into a system selection process has to be proportionate to the cost of the proposed system and the risks of making a less than optimal choice. Consequently, selection teams can vary from a single person working part-time to large multiperson teams with at least a core team working full-time. This chapter will assume that a team approach will be used but that the process developed will be readily adaptable to a single selection person. It will further assume that the purchase price is sufficient to justify a thorough selection process. Again, shortcuts for less expensive systems will be readily apparent.

The aim of any selection process is to reach a purchasing decision that satisfies as many of the stakeholder requirements as possible, as quickly as possible. It will usually not be possible to satisfy all stakeholders completely, and therefore, any decision will involve an element of compromise between stakeholders. Similarly, no system is ever going to be a perfect fit to requirements, and vendors in rapidly changing technological fields will always have a new hardware model or a new revision of the software available shortly, which they believe will take them ahead of the competition. This leads to the danger, endemic in many selection processes observed by the author, of delaying the decision until a new model or version is available and then finding that much of the information obtained about other systems is no longer relevant and needs to be updated. System selection can become a life's work for the unwary!

Similarly, anyone who has bought any consumer product from a showroom with a good salesman will be aware of the phenomenon of "requirements creep", and its consequent inflationary effect on the cost of the purchase. This is a temptation that needs to be strictly controlled. The aim should be to obtain best value for the money, which is not necessarily the same as the lowest price for a given requirement. Functionality in excess of that in the requirements document should be justified on the basis of its potential return in terms of benefit over time, not just on a feeling that it would be "nice to have".

Therefore it is, necessary to use a well-structured, systematic selection process. This chapter will attempt to provide such a process and to describe some of the techniques that can be used along the way. A graphical representation of a system selection process is provided in Figure 4.1 using the role activity diagramming notation described by Ould (Ould, 1995). The notation shows clearly the points where interactions between roles must take place and can be easily translated into a project plan to allow allocating the duration of activities and to monitor progress. The next section describes the process shown.

4.2 The Selection Process

The system selection process for any major system should comprise two components, a technical evaluation for choosing the most appropriate hardware or software from a technical perspective and then a commercial evaluation to ensure that the vendor chosen has the right attributes in financial stability, reputation, local support outlets, and so on.

4.2.1 Team Formation

The technical evaluation begins when a project champion, usually a member of senior management, initiates the formation of a team tasked with selecting the system. It is, of course, possible to use consultants. This has the benefit of potentially shortcutting the entire selection process because a good consultant should be able to match requirements rapidly with appropriate solutions. This approach is appropriate where there are no internal people available to participate in the selection process. However, many companies will opt for an internal selection team because of the costs of a consultant and the possibility of consultant bias toward a particular system. This has the benefits of allowing potential future users of the system to become conversant with the company aims and objectives for the technology, and increases the sense of "ownership". This in turn may help to reduce any internal resistance to implementation.

Typically, a team will consist of an "inner team" of one or two people who are involved in the selection project on a full-time basis and an "outer team" of people who will typically be representatives from the various stakeholder groups. The "outer team" will assist with requirements definition, attend vendor visits and presentations, and contribute to the final decision. The major requirement for the "outer team" is that its members remain the same throughout the selection process and that the whole team attends all vendor events to ensure consistency in forming opinions. If the company has existing reverse engineering systems, with vocal advocates of particular vendors, it will be necessary to ensure an even balance of prior opinion.

4.2.2 Identify the Business Opportunity and Technical Requirements

Once the team has understood its terms of reference and read this book, its initial task is to develop a business case document that outlines the business opportunity to be addressed, the major technical requirements of the system, and the likely revenues or cost savings from the system over time. This should lead to an initial determination of a budget and a view of whether an outright purchase or a lease is likely to be appropriate. The business case should also consider whether the volume of work is sufficient to justify an in-house facility, with

all the costs of facilities provision, operating costs, including power, lighting, staff, and training, or whether it is cheaper to use a service provider. The business case should be formally reviewed by the project champion and the senior management team and authorization to proceed obtained.

It is now necessary to develop a detailed understanding of the system requirements. This will include the requirements themselves and also how satisfaction of a particular requirement will be evaluated. For example, will the accuracy of the scanning system be assessed qualitatively based on a team opinion, or quantitatively based on the vendor's published data, a benchmark test, or some other approach? It will also be necessary to attempt to quantify the relative importance of various attributes. For example, is scanning speed more important than accuracy or *vice versa*? This will lead to the development of a rank-weighting assessment matrix for use later on in the process.

4.2.3 Vendor and System Information Gathering

In parallel with the creation of a requirements specification, it is appropriate at this stage for the team to start collecting information about the technical capabilities of a potential vendor's systems. This will provide the basis for the initial triage of possible systems from inappropriate systems and inform the development of the requirements document. Though the initial data gathering can be done by contacting vendors and requesting information, by far the easiest way to get a broad view of the system marketplace is by searching on the Internet. An exact phrase search on "Reverse Engineering" alone tends to return systems for software de-compilation. However, addition of keywords such as "CAD/CAM", "scanner", "digitizer", and so on will soon produce a lengthy list of vendors. To shortcut some of this process, a list of vendors is available at the end of this chapter. However, the list is not complete and should be used only as a starting point.

4.2.4 Vendor Short-listing

Short-listing potential suppliers requires a process based on fairly rapid evaluation of each vendor. This should be based around a decision tree of key questions which can lead to rapid exclusion of some vendors. Most reverse engineering systems comprise some form of scanner or digitizer and some software to reduce the number of points in a point cloud and to create polygons between the points. Additionally, software to fit curves or surfaces to the points to produce a CAD model of the object, or an export facility to some rapid prototyping technology may be required. It is important that the scanning hardware should interoperate with the point cloud software and that curve and surface fitting software should produce models that can be passed to the target CAD system, and so on. One way to ensure that this is possible would be to buy a complete solution from a single vendor. This obviates the

need to worry about the interface standards between hardware and software but may be at the expense of buying system components which are not quite "best in class". Also such vendors tend to be primarily CAD/CAM vendors. Thus, if you already use a particular CAD/CAM system, the choice of integrated solution suppliers may not be large. However, if you strongly prefer an integrated solution, then component suppliers can be rapidly eliminated.

Similarly, scanner/digitizers broadly divide into two categories: contact measurement systems, based usually on coordinate measuring machines or robot held mechanical probes, and noncontact systems based on laser or white light, magnetic resonance imaging, acoustic methods, and so on. Again, key questions will help to narrow the field rapidly. For example, if the objects to be scanned are large, easily deformed, difficult to clamp and support, or animate and therefore moving, then a noncontact method may be preferable. Conversely, if the objects to be scanned are relatively small, rigid, easily clamped and supported and if great accuracy is required, then, a contact method may be appropriate. If there is a clear preference for contact scanners then noncontact vendors can be eliminated. As always, a key parameter is cost, and some highly desirable systems may be simply too expensive!

Of course this process can safely be taken only so far because there are areas where the attributes of different approaches overlap. For example, there are overlaps between the measurement accuracy of some contact and noncontact methods. Although at first sight, noncontact methods acquire points faster, by the time software has been run to reduce the number of points acquired for planar surfaces, to eliminate noise, or to interpolate gaps in the data, and to deal with shading and screening issues, it may prove quicker to collect a smaller more structured, point cloud using a slower, contact measuring system.

4.2.5 Visit the Short-listed Vendors

Once this preliminary triage is complete, a short list of 5–10 suppliers should have emerged. It is now necessary to open communications with the potential suppliers. This can be done in a number of ways. A version of the requirements document can be sent to each supplier with a request that they confirm their ability to meet each requirement. This assumes that your requirements are written in an unambiguous way that the vendor can understand. However, vendor responses may lead to further reduction of the list. But at this stage, it is usual to visit the vendor for a demonstration or to ask a salesperson to visit the customer. The benefit of taking the team to the vendor is that the vendor may be able to involve a broader range of people, who are experts in some aspect of the system and is likely to be set up to provide a much fuller demonstration. Set against this, of course, is the cost of transporting the team to the vendor. While at the vendor, it is appropriate to discuss your requirements and to seek the advice of vendors on any aspect that may have been omitted.

4.2.6 Detailed Vendor Assessment

To compare the systems demonstrated meaningfully, it will be necessary first to have prepared a matrix of desired system attributes and their overall importance. The importance of a particular attribute is usually ranked on a scale from 1–5 where, for example, 1 may mean "not at all important" and 5 may mean "essential". Then as the team watches the demonstrations and questions the vendor about its product, each member should be trying form an opinion of the capability of the vendor against each of the attributes.

As soon as possible after the vendor visit and certainly within days rather than weeks, the team should meet and agree on a score for the vendor against each attribute. This can be done as a mean and standard deviation of each team member's individual scores or by discussion until a consensus is reached. Finally, the scores against each attribute are multiplied by the importance of each attribute and the resultant weighted scores are then totaled for each vendor. This gives the relative ranking of each vendor and may be used to reduce the list of potential vendors to two or three. Figure 4.2 shows a sample rank weighting grid. In this example, vendor 1 and vendor 2 both have comparable scores and should be considered; vendor 3 is poorer and should be rejected. The nature of the process described here requires that members of the team make fairly subjective decisions about a particular vendor's capability against an attribute and about the importance of an attribute. Clearly, therefore the outcome of this process should not be regarded as absolute truth. It is sometimes useful to perform a sensitivity analysis on the results to see how the outcome varies with small changes in the assumptions. For example, the value of each importance factor could be randomly varied by plus or minus an amount, and the totals reevaluated to see whether the outcome changes. Similarly, if team mean and standard deviations were calculated, it is possible to vary a vendor's score randomly where a standard deviation above an agreed value indicates a lack of agreement among the team members. For this process to be useful, it is essential that the same people see all the vendors and contribute to the scoring process. Otherwise the outcome may be more indicative of who attended each visit rather than the relative merits of the systems under consideration.

Attributes	Importance of feature	Vendor 1 scores		Vendor 2 scores		Vendor 3 scores	
		Score	Weighted score	Score	Weighted score	Score	Weighted score
Attribute 1	3	4	12	3	9	2	6
Attribute 2	2	2	4	2	4	3	6
Attribute 3	4	3	12	5	20	1	4
Attribute 4	3	3	9	3	9	3	9
Attribute 5	4	5	20	4	16	3	12
Vendor Totals		17	57	17	58	12	37

Figure 4.2. Example of a rank weighting grid

4.2.7 Benchmarking

Based on the outcomes of this work, a measure of judgement, and perhaps some preliminary views on the commercial stability of the proposed vendor, the successful two or three vendors should be asked to prepare a benchmark test. Because this will involve providing the vendor with one or more items that are representative of the typical work that the system will have to perform, first it may be necessary to ask the vendor to sign a nondisclosure agreement. The vendor is then asked to scan the component and create the outputs that you require from the process, perhaps a CAD model for your CAD system, perhaps an input file for your preferred rapid prototyping approach, perhaps a point cloud that can be used as a benchmark to evaluate software vendors. To assess the outcome of this stage, it will be necessary for the team to visit the vendors again to see the benchmark exercise being performed, for the team to study the deliverables from the benchmark in detail, and then to go back to the matrix of vendor capability attributes and decide whether any scores need to change. This should lead to a firm technical decision as to the desired system and a runner-up in case of unforeseen difficulties in the next phases.

4.2.8 Perform a Commercial Evaluation of the Vendor Chosen

The final stage is to perform a commercial evaluation of the supplier selected, to review with management and the project champion the outcome of the process, and then, if approved, to enter into negotiations with the supplier. The commercial evaluation will try to ensure that the proposed vendor is sufficiently financially stable and well established to meet its contractual commitments in terms of delivery, warranty, after-sales service and support, possible midlife upgrades to hardware and software, and so on. Negotiations are necessary to ensure that the contract is satisfactory to both parties and to determine whether or not there can be any reduction from the list price. As with any negotiations, it may be possible to use the runner-up as a lever to obtain better terms, providing you are prepared to buy from the runner-up if negotiations with the preferred vendor are unsatisfactory. Make sure that the total cost of ownership is understood, not just the purchase price. There may be fairly significant ongoing costs for hardware and software maintenance, extended warranty costs, consultancy charges if you want the vendor to help to install or customize the system, and so on. The end point of the selection process is the dispatch of a purchase order to the winning vendor.

4.3 Some Additional Complexities

The process just described and illustrated by the role activity diagram (Figure 4.1) has described a process where one vendor supplies the whole system. As

indicated earlier, the entire reverse engineering system finally implemented may comprise components from a number of vendors. For example, just as there are many pieces of hardware and software for obtaining and manipulating a point cloud from a scanner, so there are many pieces of software that can take a point cloud in a variety of formats and create a CAD model. Some of these software products will work or interface better with some scanning products than others. As a further complication, some vendors may have more than one offering in a particular space. Thus, it becomes necessary to consider "the vendor" as a particular combination of vendors and particular products from each vendor. This requires a modular approach to each product. In other words, the assessment matrix for a particular product may be reused in combination with other products, provided it can be established that there is a reliable and efficient interface between the proposed components. The quality of the interface between different vendor's components must also form part of the assessment process. To prevent the number of permutations from becoming unmanageable, it is strongly recommended that the selection process begin with the hardware for obtaining the point cloud. When the possibilities for this system have been reduced to a handful, it should be apparent which software may be compatible with each, and the selection process can be repeated for the software components. From time to time, it may be necessary to reconsider the hardware assumptions and revise or refine them in the light of the emerging preferences for software.

4.4 Point Capture Devices

As already suggested, the market divides broadly into two parts, contact and noncontact point capture devices. Some argue that there is a third category, a hybrid category, where, for example, noncontact scanning devices can be attached to the head of a coordinate measuring machine, a device more normally associated with some form of contact probe. Perhaps it is better to think in terms of the probing device, the work-holding device and the probe-positioning device. To compare probes, the major attributes to be considered are accuracy, speed of data gathering, and repeatability of the measurements. Figure 4.3 offers a taxonomy of measuring systems.

4.4.1 Contact Devices–Hard or Manual Probe

The oldest and slowest form of contact probe is usually known as a "hard probe". It requires that the user manually position the probe on the surface of the object to be measured, allow the machine to settle, and then manually tell the coordinate measuring machine (CMM) to record a measurement. The coordinates of the point are automatically calculated by the CMM's software, making
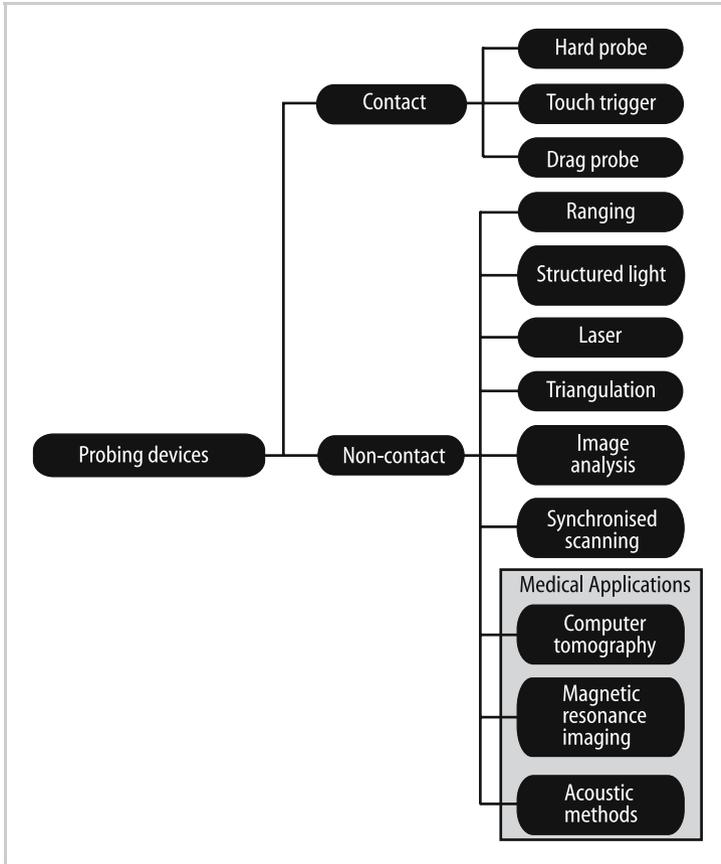
Figure 4.3. A taxonomy of measuring systems

due allowance for the diameter of the stylus tip. This form of probe is the slowest, and its accuracy and repeatability depend on the operator's "feel". In the hands of a skilled operator, the accuracy and repeatability can be good, but where multiple operators are used, there is a risk that each will "feel" the probe touching with slightly differing amounts of force and thus get very slightly different readings. However, systems based on this sort of probe are at the cheaper end of the spectrum.

4.4.2 Touch-trigger Probe

The touch-trigger probe overcomes some of the limitations of the hard probe in that the device itself triggers the taking of a measurement when the probe contacts the surface of the object. Typically, an LED flashes or a beep is heard to

indicate that a point has been recorded. This sort of probe is typically mounted on a manual or computer-controlled coordinate measuring machine and is the staple of much engineering metrology. The advantage over the hard probe is that it eliminates the element of operator feel, and when used in conjunction with a computer-controlled measuring machine, can be a reasonably rapid and very repeatable way to obtain data. Further enhancements to look out for include piezoelectric sensors that allow the software to compensate for deflection of the stylus caused by variable forces when the probes touches the workpiece, and strain gauge technology which ensures that the force applied by the probe is constant when a measurement is taken. Though they add cost, these features also improve accuracy.

4.4.3 Continuous Analogue Scanning Probe

The final class of probe is the drag probe or continuous analogue scanning probe. This type of probe is literally "dragged" across the surface of the object to be measured and continuously records the location of the probe. This dramatically increases the speed of data acquisition, but speed can suffer from variability in the surface finish and from vibration, as the probe is dragged across the surface. Some systems provide software that can filter out these effects. "Closed-loop" systems adjust the position of the stylus relative to the surface based on trends in the measurements being taken. This allows measuring a workpiece of unknown topology easily. However, where the shape is already known, from a CAD model, for example, the stylus can be driven over the surface by a prewritten "part program". This allows comparing the actual part with the CAD model, and reports of variance can be easily generated.

4.4.4 Other Facets of Probe Selection

Two additional facets of selecting a probing system are deciding whether or not to buy a motorized probing head and which styli to buy. When taking measurements, it may be necessary to adjust the angle of the stylus to maintain optimal contact with the surface or to access difficult to reach areas. Manual heads require operator intervention; motorized heads can be changed directly by the CMM controller. Styli are available in a bewildering array. The stylus is the part that actually contacts the surface of the workpiece. Typically, styli have spherical tips that are available in a range of diameters and materials. Larger diameters are less affected by surface finish but make access to small features problematic and *vice versa*. Styli are also available with cylindrical, hemispherical, star, and disc configurations. Styli also come in a variety of lengths. The shorter the stylus, the less it will deflect. However, access to some feature may require long styli. Some vendors provide optical systems to measure the deflection of the

stylus and hence compensate for stylus deflection. Stylus tips are available in a variety of materials; each material has accuracy and wear implications. To increase stylus configurational flexibility further, stylus extension bars are also available, but each additional joint provides the possibility of additional joint deflection during measuring. Again, the stem of the stylus can be made from a variety of materials, each has differing deflection properties and coefficients of thermal expansion. Using the right stylus for a particular feature is essential to maximize the accuracy of measurements. For some jobs, more than one stylus may be required. Some vendors offer automated stylus changing without the need to requalify the new stylus.

4.4.5 Noncontact Devices

Noncontact devices are typically based on some form of light measuring system, although for medical applications, sound, X-rays and magnetic resonance are all used regularly. Noncontact devices are now available which can be mounted on the head of CMMs and thus provide further flexibility for measuring objects which might deflect if a contact probing approach were used. However noncontact devices can also be free-standing. There are so many variations of device in this category that it is quite difficult to produce a convincing classification. However, there are some common themes.

4.5 Triangulation Approaches

Triangulation is a common approach. In this method, a spot or line of light from a high-intensity light source, often a laser, is deflected by a rotating mirror onto the surface of the object. The reflected light is focused by a lens onto one or more position-sensing devices (PSD) in older types of hardware or charge-coupled devices (CCD) in more modern hardware. The position of the surface of the object can be calculated by simple trigonometry from a determination of the position of the reflected light point. The position-sensing device uses a photodiode to determine the location of the mean of the light quantity distribution, which is not necessarily the point of maximum light intensity. This means that PSDs are seriously affected by changes in the reflectivity or color of the surface being measured, by the perpendicularity of the light to the surface, by ambient light changes, and by stray reflections. The newer charge-coupled devices use a matrix of photosensitive elements, each of which can determine the intensity of the light being received. Thus, it becomes relatively easy to determine the location of the pixel receiving the maximum light intensity. Algorithms that look at the intensity of light being received by adjacent pixels can provide subpixel resolution of the position of the point of maximum intensity, further enhancing the accuracy of this sort of device. This type of detector overcomes many of the

problems of the PSD. It is not so sensitive to surface finish, color, or ambient light changes, and black or shiny surfaces are not a problem. It can also use class I or II lasers which eliminates/reduces the safety issues of more powerful class III lasers. The accuracy of laser triangulation systems is very good, provided the laser is close to the object. The accuracy decreases linearly as the distance increases due to measurement problems with small angles.

Systems are available that project multiple lines of light simultaneously, thus increasing the speed of data gathering without sacrificing accuracy. Sophisticated software algorithms that allow changing the direction of the laser beam within the frame rate of the camera and then interpreting the multiple images on each frame, can further increase the speed of data capture Such systems can be used on animate objects where some movement is inevitable. Additionally, some systems record the pixel light intensity readings separately to allow constructing photorealistic images at the same time a point cloud is collected. Other systems combine video cameras to capture surface texture and color and hence produce full color photorealistic images. Such systems are widely used in the entertainment and film industries.

4.6 "Time-of-flight" or Ranging Systems

Alternatives to the triangulation approach include "time-of-flight" devices. These can be pulse-based or phase-based. In a pulse-based system, a pulse of laser light is emitted from the transmitter, and the time taken for the light to be reflected back to the sensor is calculated. The distance is then half the time multiplied by the speed of light. Another approach uses a continuous beam of laser energy, whose amplitude is modulated sinusoidally. The phase shift between the transmitted and received beams allows calculating the distance. Both approaches provide good accuracy. Pulse-based systems are relatively slow but have the advantage of allowing a large "standoff" distance from the object being scanned. This can make them suitable for civil and land-mapping applications where a relatively sparse set of data is required from distances up to 500 m. Phase-based systems, by contrast, can gather data much faster, but typically have shorter standoff distances.

4.7 Structured-light and Stereoscopic Imaging Systems

There are numerous systems using nonlaser light, which, of course, have the advantage of being inherently eye-safe. Structured-light systems project a predetermined pattern of light onto the object; a moiré fringe pattern is a popular choice. CCD cameras capture the distorted pattern as it is reflected from the object and the comparison of the two allows calculating distances by triangulation. These systems provide extremely fast data gathering, which coupled with

their eye safety, tends to make them the first choice for scanning the whole or part of the human body for medical or cinematic applications. Image-based systems use stereoscopic cameras to take digital pictures of the object from two or more locations. Image processing algorithms then attempt to register common points in the resultant images and hence to calculate distances. Though theoretically simple and cheap, difficulties in repeatable registration point recognition result in a variety of system enhancements that add complexity and cost. This sort of system is also very slow. In general, nonlaser light systems are faster than laser systems but tend to be less accurate and are limited to smaller scanning volumes.

4.8 Issues with Light-based Approaches

Light-based devices all suffer from the related problems of "shading" and "screening". Screening occurs when a point is illuminated by the transmitter but the reflected light cannot reach the receiver because another feature of the workpiece is in the way. Shading occurs when the receiver can "see" a workpiece location, but the transmitted light cannot reach the point. Light-based approaches also have difficulty when the object is fully or partially transparent.

The wavelength and power rating of a laser light source can also be an issue. Here, there is a compromise between eye-safety and noise minimization. Wavelengths in the visible spectrum are typically chosen to obtain a Class II rating but the trade-off is with increased "noise" from ambient light and objects whose color is complementary to the light source, *e.g.*, red objects with green light sources. Infrared low-power lasers are generally Class III as are higher power lasers using visible light. However, the classification system is complicated. The main point is that Class I lasers require no precautions, Class II devices require a few precautions Although "eye-safe", particular geometries on the workpiece could accidentally focus the beam to make it unsafe. Class III lasers require safety training and fairly strict safety procedures, but provided these are followed, they are quite safe.

The wavelength of the light transmitted from a laser system will vary with temperature, and this can have a significant effect on the accuracy of the measurements. Most systems have built-in temperature compensation techniques but if wide variations in temperature are a part of the likely usage conditions of the system selected, it may be wise to ask for evidence of adequate temperature compensation technology and/or make a temperature variation test part of the benchmark. Similarly, condensing humidity may be an issue. Make sure that the equipment chosen can work in the range of humidity that may be encountered.

4.9 Tracking Systems

A further variation on a theme is the use of noncontact measurement approaches to "track" locations on a moving object. Again, there are many varieties of technology. Laser trackers require fastening reflective markers to the moving object. This approach is often used for accurate alignment of plant and machinery over large distances. Light-emitting diodes (LEDs) can be attached to the moving object and their position tracked by cameras using triangulation approaches. High-speed modulation of the light from the LEDs provides each LED with a unique identity signal and thus allows tracking many locations simultaneously. A variation of the technique uses small orthogonal coils attached to the moving object to broadcast an electromagnetic field from each of the three orthogonal coils. This approach offers the advantage of tracking positions that are not in the "line of sight" from a light source to a receiver. However, the result may be affected by the proximity of metal objects. A variant of this approach uses magnetic tracking to determine the position of a handheld laser scanner which can then be swept over the workpiece in a natural, intuitive way to gather a point cloud for the object, the system software combines the positional measurement of the scanning head from a datum with the measurement from the scanner to the workpiece.

4.10 Internal Measurement Systems

Internal measurement systems are divided into destructive and nondestructive systems. Systems for medical applications include magnetic resonance imaging (MRI) systems, acoustic methods, such as ultrasound scanning in prenatal checks, and X-ray tomography. These will not be discussed in detail. However, now there are X-ray tomography systems intended for engineering applications such as void detection in castings and welds and measuring the internal dimensions of castings.

4.10.1 X-ray Tomography

X-ray tomography is based on a device that passes a highly collimated X-ray beam through an object and a detector, which measures how much of the original X-ray has been absorbed by the object or body being scanned. The result is given as a Hounsfield number, named after Dr G.N. Hounsfield, the developer of the first practical system. Higher density objects have higher Hounsfield numbers. By rotating the beam and detector around the object, measurements can be calculated for each X-Y location in a "slice" of the object. Then, by translating the beam and detector by a Z-increment, another slice can be obtained. By measuring a series of slices, a 3-D image of the object can be constructed. Each

2-D slice can be represented as an image where shades of gray correspond to different ranges of Hounsfield numbers. (The eye cannot distinguish among the numbers of shades of gray that would be required to use one shade for each Hounsfield number.) Algorithms that look for changes in Hounsfield number between adjacent pixels in the image can draw edge profile curves. Where the change in density is very distinct, this is relatively easy to do. However, in medical applications, the differences among the densities of tissues and fluids, for example, can make edge determination quite difficult. More recent systems use variations on the original theme. Some use two-dimensional or three-dimensional arrays of detectors, some use a "fan" of X-ray beams, and some use a spiral technique where the translation and rotation of the beam are continuous. These differences are highly significant in medical applications because they significantly affect the speed of measurement and the X-ray dose the patient receives. Measuring times in recent systems can be as low as 50 ms per slice, which is fast enough even for moving organs of the body, such as a beating heart. However, dose and speed issues are less significant for static inanimate objects. Here accuracy is more important, and these can be as low as a few microns. The cost of such systems, however, is high!

4.11 Destructive Systems

Destructive systems embed and fill the object to be measured with colored resin and then use, for example, CNC milling technology to remove a thin slice of material. The shapes of the boundaries of each slice are determined by using one of the imaging or scanning technologies already discussed. A further slice is then removed and the process repeated until data are available for layers through the object. The downside of this process is that it is slow and results in the complete destruction of the original object. It is not suitable for Ming vases then! It also assumes that the material of the object can be machined.

4.12 Some Comments on Accuracy

One of the fundamental questions in assessing the merits of a measuring system is "How accurate is it?". The response of most vendors is a single number. However, a complete accuracy assessment of a given device is a complex business, and reducing the results to a single number can be misleading. First, accuracy of all measurement systems will vary depending on where within the working volume the measurement is taken. As already mentioned, the accuracy of triangulation systems decreases with distance from the light source. So for this type of system, the standoff distance at which a measurement was taken is an important parameter. Similarly, a robotic arm carrying a measuring probe will deflect more at the full extent of its reach than at closer points. So, the accuracy is really a distribution of numbers and depends on a whole raft of other parameters. In

reducing this complexity to a single number, some vendors will quote the mean value; others will quote the root mean square value which at least gives some idea of the mean and standard distribution of the measurement accuracy. The mean value will be a lower number than the RMS value.

Therefore, before the accuracy of different systems can be compared, it is necessary to know which number has been quoted. There are standard protocols for determining the accuracy of measurement. Again, ideally, the vendor should declare which protocol they have used so that comparisons can be made. Then, it is necessary to know the resolution of the system in terms of the number of points per unit length that the system is capturing. (In triangulation systems based on angles, the maximum resolution is determined by the smallest angle though which the mirror can be rotated. The resolution in points per unit length will decrease as the standoff distance increases. so, the resolution should be specified at a particular standoff distance.) Although accuracy and resolution are not directly related, the resolution of the system can make a difference in the accuracy, for example, of finding an edge. Suppose that one measurement just misses the edge. The next measurement is at a resolution of, say, 1 point per mm. Thus, the next measurement will be nearly 1 mm inside the true position of the edge. So, even if the accuracy of measuring the point is high, the measurement of the edge position will be inaccurate by as much as 1 mm! Similarly, if the workpiece is larger than the working volume of the measuring system, one approach is to gather a point cloud for different sections of the object and then join the point clouds together. If the error in the registration of one cloud to another is, say, 0.1 mm, then 10 sections along there is a possibility of a tolerance stack up error of 1 mm. There are a variety of approaches to solving this and related problems. Ask the vendor how he deals with it.

The final point concerning accuracy is that for reasons of repeatability, vendors are likely to have performed their accuracy and calibration tests in very tightly controlled laboratory conditions, which may not be typical of the environment in which the end user will use the system. Temperature and humidity variations can have significant effects on the accuracy of the system as can the way the user takes measurements. Thus, ideally, the benchmark part of the selection process should attempt to verify the accuracy of the system in an environment as close as possible to the working environment. The results of such a test may be quite different from the numbers listed in the vendors' specifications! This raises the question "How accurate does the system need to be?". The answer to this depends upon the reason for performing the measurements in the first place. In addition to the data gathering accuracy issues just discussed, the data gathered from the scanning phase can be postprocessed in a number of ways depending on the required outcome. All of these processes carry with them the risk of introducing additional errors, for example, the errors in fitting a smooth, fair curve to a set of data points, the accuracy errors in a rapid prototyping machine, and so on. So, the real question is whether or not the final output of the process is "fit for its intended purpose". Again, a benchmark test that follows through the whole of the intended process and allows comparing the
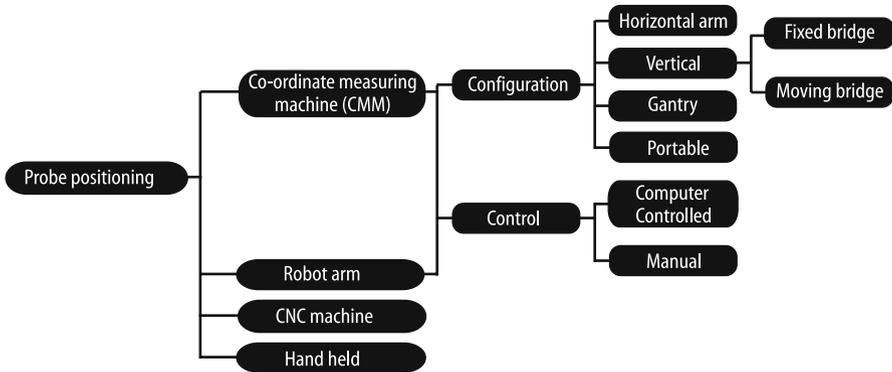
Figure 4.4. A taxonomy of positioning devices

inputs against the outputs may provide at least a partial answer to this question. For example, if a mechanical part is measured by laser scanning, a surface model created in a CAD system, and a replacement part NC machined or rapid proto-typed, it would be possible to measure the original and the replacement part using perhaps a CMM, and form a quantitative, statistical view as to whether the new and the original part are sufficiently similar. Though no amount of clever postprocessing can "add" accuracy to a poor original data set, the acid test really is whether or not the final output is satisfactory. This must be borne in mind when comparing vendors' accuracy numbers.

4.13 Positioning the Probe

There are a wide variety of ways to position the measuring device relative to the object, or the object relative to the measuring device. Coordinate measuring machines are a well tried and tested approach. They are broadly configured as horizontal, vertical, or gantry systems. Horizontal and vertical relate to the di-rection of the primary axis of the machine; gantry machines tend to be aimed at measuring large objects. CMMs can be manually driven or driven by computer-controlled motorized drives. If the latter is chosen, then there may also be a need for off-line programming software, to create measuring programs from CAD models. However, once programmed, the same measuring routine can be exe-cuted any number of times for inspection of production parts. As discussed before, most CMM vendors will offer both contact and non-contact probes, and single or multi-axis, motorized or manual heads.

An alternative method of positioning the measuring device is by using a multi-axis robot arm that can be driven manually or by a computer-controlled motorized drive. They range in size from desktop devices for very small objects, to arms that can measure whole cars. Another alternative is to fit the measuring device to an existing CNC machine tool to allow in-process measurement.

Finally, some systems can simply be handheld and swept over the surface of the object to be measured.

The other approach is to keep the measuring equipment stationary. With some body measuring systems, multiple measuring systems are positioned at different angles around the subject so that $360°$ measurements can be made very rapidly. The other alternative is to have a single measuring position and rotate the object. Rotary tables are commonly used. Make sure that your point post-processing software is capable of rotary table registration of the various scans. Figure 4.4 offers a taxonomy of measuring systems.

4.14 Postprocessing the Captured Data

There are a number of routes than can be followed once a set of data has been acquired. Some vendors call these routes workflows and offer workflow templates to guide the user through the process steps to achieve a desired output. Figure 4.5 shows a generalized workflow, which is intended to help clarify some
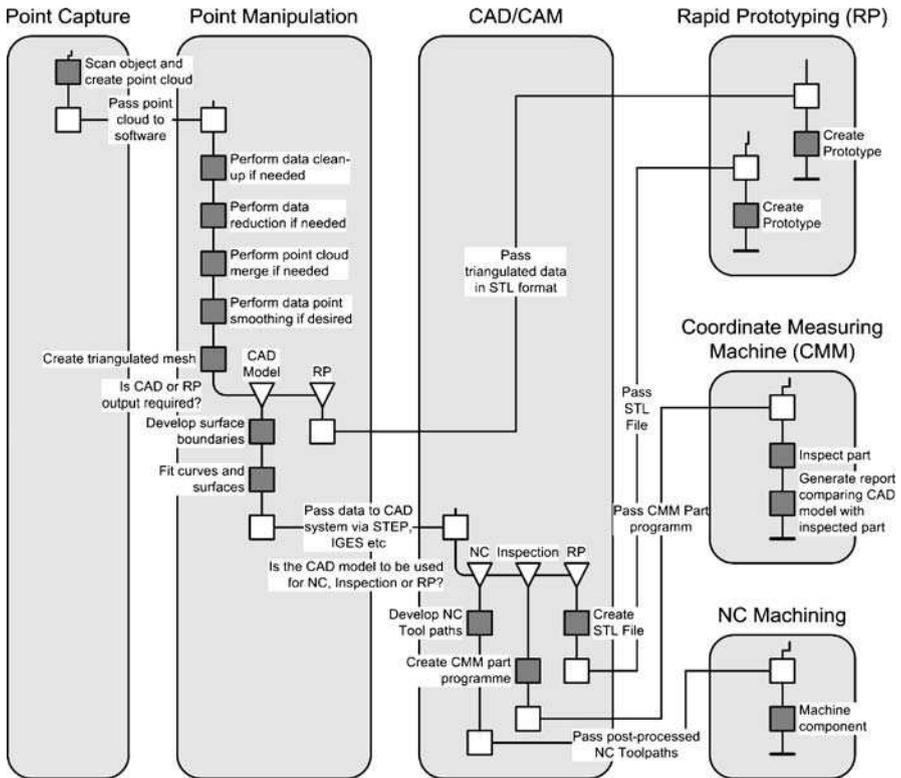
Figure 4.5. Generalized mechanical engineering reverse engineering workflow

of the following discussion. To keep the role activity diagram reasonably simple, some detail has been omitted.

The primary output from any of the scanning or measuring techniques discussed before is a very large set of *X-Y-Z* coordinates. File sizes can run to hundreds of megabytes in some systems. But little can be done with the data without some sophisticated software to postprocess the points. There are a number of classes of software for dealing with the point cloud. The first class in terms of the workflow of reverse engineering is those that manipulate the points in some way, usually leading to a polygon/triangulated mesh. This may be sufficient for some purposes. For example, rapid prototyping or finite element analysis may be possible directly from the polygon mesh. However, for other types of work, it may be necessary to create a CAD model. This leads to the second class of software where curves are smoothed or fitted through the data points and then surfaces can be interpolated between the curves. Some vendors include all the foregoing functionality in the software provided with their data acquisition system. Other vendors provide point cloud manipulation in their data acquisition system, but rely on niche market vendors to provide curve and surface manipulation tools. Finally, some of the major CAD vendors have modules dedicated to the special needs of reverse engineering. The trade-off is between stand-alone software modules, which in general are the most capable but require interfaces between modules and slightly less capable but seamlessly integrated software suites. Whichever approach is adopted, make sure that the interfaces needed for your circumstances work well.

4.15 Handling Data Points

Though any CAD system can read in a set of points that can then be used for surfacing, not all CAD systems can cope with large point clouds. First the data, particularly data from scanning systems, may require some cleaning up. Because scanning systems can produce millions of points in a few seconds, the cleanup process has to be largely automated. The exact sequence of operations, referred to by some vendors as a workflow, may vary from system to system and user to user. Typically, some data from scanning systems will be rogue data, caused perhaps by stray reflections or other "noise". Often, these points will be well separated from the main body of the point cloud and are relatively easily removed by automatic algorithms. Usually, to scan an object completely, a number of scans will have been performed from different locations. Each scan will create point coordinates relative to a local coordinate system. To merge them together successfully, it may be necessary to transpose the coordinates to a global coordinate system, perhaps based on reference markers or some vendor provided system for relating the scan positions to a common coordinate system. Some vendors offer semi- or fully automated algorithms for merging different scanned points together, based on relating the implied curvatures of the surfaces of the scanned

Point Number	X	Y	Z	Resultant planar polygon
P1	2.3564	4.5673	7.3428	
P2	3.5674	7.6784	7.3428	
P3	7.4536	5.9876	7.3428	
Edge	Start vertex	End vertex		
E1	P3	P2		
E2	P2	P1		
E3	P1	P3		

Figure 4.6. Topology of a triangle

object. Similarly, vendors may offer algorithms for reducing the number of points in areas of large radius of curvature, while retaining a higher point density in areas of rapidly changing or small radius of curvature. Some vendors offer this as a postprocessing operation, and some offer it as a real-time calculation while the data are being scanned, to minimize the number of points written to the storage device of the scanning system (often, a powerful laptop or desktop PC).

The next stage may be to perform some smoothing of the data points. Suppose that a surface was truly flat. Then, all data points should sit on a mathematically flat plane. However, due to measurement errors, as discussed before, it may be that some points are just above the plane, and others are just below. If curves are fitted to the points (*i.e.,* created so as to pass through all points), the result would be a "wavy" curve, when it should be a straight line. Similarly, it may be assumed that the "design intent" of a curving surface was a continuous variation in the radius of curvature of the surface. A point smoothing algorithm will provide some preconditioning of the data points, before curves are fitted, to remove oscillations around the assumed "design intent" curve. (Alternatively a smoothing algorithm may be used when curves are created. This will be discussed later.) Once the raw data have been manipulated to the user's satisfaction, the next task will be to create connectivity between the points. This is done by associating polygon topology with the coordinates of the raw data, as shown in Figure 4.6, by using some form of Delaunay algorithm.

Good algorithms will create well-conditioned polygons without small gaps, incorrect topology, spikes, inconsistent normals, and so on. Once the polygon mesh is complete, it should be possible to create shaded images of the point cloud. This will also help to highlight visually any problems areas where edges have not been correctly determined, where holes have been incorrectly filled, or where holes that should not be there remain. In some cases, it may be possible to fill such holes based on tangency or curvature continuity algorithms. In other cases, it may be necessary to rescan an area and merge the new points with the old. There may also be algorithms for reducing the number of triangles based on some user-defined error bound between the "true" shape of the surface and the plane of the polygon, for registering and merging polygon meshes from different scans or different models, and so on. There may also be a range of

tools for slicing polygon meshes, adding thickness, creating closed solids, creating parting lines, morphing meshes, creating offset meshes, trimming meshes with projected curves, mirroring, translating, zooming, scrolling, measuring, and so on. Make sure that there is enough functionality for your intended application, but don't get too carried away by the need for extra commands.

Given that the data sets can be very large, speed is a key parameter. Because this will depend in part on the specification of the computer that is running the software but will also be a function of the total number of points. This is an area where a benchmark test that uses a cloud of the "typical" size to be encountered, run on a computer of the type of specification intended, may be needed. Check the speed of decimation algorithms, but also check the speed of routine commands such as scrolling and zooming. It will also be useful to check the decimation ratios (*i.e.*, the ratio of the number of points before running the point or polygon reduction algorithm divided by the number of points remaining) and whether or not the decimation algorithm has resulted in some loss of accuracy. In addition, the usual software assessment criteria such as the quality of the user interface, the quality of the documentation, and so on should be checked.

Exporting the triangulated file and if necessary writing scripts to convert from one format to another is not usually very difficult, but it is worth checking to ensure that formats such as. STL (stereolithography format), ASCII, IGES, and STEP are available to interface with the applications to be used next in the intended workflow. If it is intended to allow collaborators to view the tessellated object over the Internet, make sure that the point cloud can be decimated sufficiently to provide "lightweight" viewing in formats such as virtual reality modeling language (VRML).

4.16 Curve and Surface Creation

To create a surface model of the object with mathematically defined curves and surfaces, it is first necessary to create a network of curves that fit or smooth the underlying data points. Most vendors will offer non-uniform rational B-splines (NURBS) as their core curve technology. These consist of segments of a given polynomial degree that are joined together with tangency and curvature continuity at the boundaries between segments. The result is a single curve defined by a set of "control points". Different vendors will offer differing maximum degrees for their N-splines. Why is this important? It relates to the argument between those who insist that the curve must fit all data points exactly and those who want to smooth or "fair" the curve. In general, n data points can be fitted exactly by a single segment curve of degree $n-1$. However, fitting even a good set of data with uniform curvatures with a high-degree curve will produce a wavy result, even though the curve passes through all points. Most vendors limit the maximum degree of their curves to the range of 3 to around 20, which means that to "fit" the data points, the number of segments has to be increased. Sometimes this can be to the point where there is one segment per pair of points. Again, this

will probably result in wavy curves and because of the large number of control points near the curve, it will be very difficult to manipulate manually.

Alternatively, data can be "smoothed". This involves creating a curve that passes close to but not necessarily through all of the data points used to initiate the smoothing algorithm. The user has to decide how close is close enough. In general, lower degree, lower number of segment curves will produce "fairer" curves, but at the expense of slightly increased tolerances on data points. To determine the fairness of a curve, there should be a range of tools available for displaying the radius and direction of the curvatures, the rate of change of curvature, and so on. Creating good quality curves is a skilled, labor-intensive activity, but good curves form the basis for good surfaces. Packages that are tailored to the needs of reverse engineering will offer a range of functions for selecting points from the polygon mesh that will form the starting point for either a curve fitting or curve smoothing algorithm. They will include chains of points created by cutting planar slices through the object that would lead naturally to a "lofted" surface, selecting point chains based on a predetermined template of curves projected onto the mesh by rate of change of curvature analysis or by algorithms that recognize features such as edges of holes, edges of the object, and so on. There should also be the possibility of picking points manually. Tedious but sometimes necessary!

Once a set of curves has been created, it may be necessary to create "contiguity" between the ends of curves. (*i.e.,* adjacent ends of two curves that share a single common end location) This is needed by some surfacing algorithms. It may also be necessary to manipulate the end conditions of adjacent curves to ensure tangency or curvature continuity. Although this can be done manually by manipulating the control points of the curves, software commands to reduce the effort required will be extremely useful.

Once a suitable network of curves has been created, surfaces can be fitted or smoothed. This is similar to curve fitting and raises the same issues of fitting versus smoothing, the number of segments, and so on. Typically, a surface patch will be created by selecting four bounding curves and a network of internal data points to be smoothed or fitted. The surface will pass exactly through the edge curves, but if smoothed, may not pass through all internal data points exactly. Alternatively, a surface can be created by lofting through a series of planar cross-sectional curves. This is typically the approach used for pipes, ducting, and so on. Some vendors offer tools to fit a surface to an area of a polygon mesh without the need for initially creating curves or to create a surface from sets of curves running in two directions that are not necessarily contiguous at the corners. Again, it may be possible for the system to recognize features such as cylindrical bosses or holes and create an appropriate mathematically precise surface. It should also be possible to sweep curves along lines and arbitrary drive curves and rotate around an axis.

It should be possible to extend surfaces, trim surfaces, offset surfaces, and possibly to "sew" surfaces together to make solid models.

Again, tools will be required to manipulate the contiguity, tangency, and curvature of adjacent surfaces. There will also need to be tools to analyze the quality

of surfaces by cutting cross sections, tangency and curvature analysis, light line analysis, and so on. Finally, it may be necessary to export the curves and surfaces to a CAD package. This is not trivial and will require IGES, VDA-FS, or preferably a STEP AP203 or AP214 translator or a dedicated translator. Finally, and perhaps seemingly perversely, it should be possible to create a polygon mesh from the mathematical surface, so that quickly shaded pictures and export of the modified object to rapid prototyping or finite element analysis are possible.

Given all this capability, it is possible with time and patience to alleviate many of the ills of a poor initial set of point data, although no amount of postprocessing can improve on the basic accuracy of measurement. However a good data set will significantly ease the curve and surface fitting process. This is where the trade-off between lots of points and a smaller set of points lies. The former may be captured very fast but take a long time to postprocess, whereas the latter, gathered more slowly, may be much faster to post-process.

4.17 Inspection Applications

Sometimes, it is necessary to compare an actual physical component and the original CAD model from which the physical object was created for first-article inspection reporting, wear measurements, permanent deformation measurement, and so on. The starting point is a point cloud that may need to be assembled from a number of scans, as discussed before. The point cloud will then need to be aligned with the CAD geometry. This is done either by relating a minimum of three reference points or by a "best fitting" algorithm that manipulates the position and orientation of the point cloud to minimize the error between the cloud and the model. Once this has been achieved, typically, the model can be measured and compared to dimensions on the CAD model. Check how automated this is. Geometric tolerances can be inspected, for example, flatness, circularity, cylindricity, and concentricity, and reports can be generated. Dimensional tolerance can also be inspected and go no-go reports generated. Cross sections of the point cloud and the CAD model can be created and compared, and in some systems, a colored contour map of the deviations of the measured object from the CAD model can be created. Some vendors provide tools to modify the original CAD model automatically to represent the measured physical model more accurately. This involves modifying the sizes of the CAD model parametrically and does not usually modify the topology of the model if features have been added or removed. However, this can be extremely useful where press tooling has been modified in the shop to make it work properly and a record of the modifications is needed so that repeat tools can be made "right the first time".

Where the aim is to measure the permanent deformation of an object caused by load or impact, it will usually be necessary to place markers or a predefined pattern on the object. Some systems provide coded markers that the system can automatically recognize, thus allowing a comparison of the original marker positions before deformation with the positions after deformation.

Reports can usually be provided in a variety of formats including Microsoft Word and Excel, PDF, html, and so on. Capability of this sort can dramatically increase the throughput of an inspection facility.

The software chosen must be able to import from the particular CAD system or systems that are of interest. This can be done *via* a neutral format exchange approach using STEP, IGES, or may be a direct import routine for a particular system. Some vendors offer both approaches and there is little to choose between them, provided the exchange software is robust.

4.18 Manufacturing Approaches

An in-depth discussion of the manufacturing methods that are available is beyond the scope of this chapter. However, the rapid prototyping (RP) route is one possibility. Most RP machines require an STL file format or a similar triangulated format, so provided this can be generated from the point cloud postprocessing software, there should not be a problem. Similarly, most CAD modelers now offer STL file generation from the CAD model, so there is another route to rapid prototyping. Some vendors offer software that will generate CNC tool path data directly from a point cloud or polygon mesh. Again, there is an alternative route from the CAD model.

4.19 Conclusion

Choosing a system is not an easy task. It is the author's belief that this chapter presents a workable, practical approach to the problem. However, it is not possible to provide a totally generic approach that will fit all situations exactly. Thus, the principles discussed before in the selection process section may need to be adapted to the particular local situation and to the time available to choose the system. As in so many areas of life, the Internet makes gathering the basic data much easier than it used to be, but there is still no substitute for visiting the vendors, seeing the system in action, perhaps even visiting vendor reference sites. In attempting to provide a basic categorization of the various options for measuring and data postprocessing, it is hoped that the author has gone some way to providing a road map through the maze of options. The appendix provides a list of some vendors and the product categories that they offer. It could never be complete but should provide at least a starting point for the data gathering stage. The author apologizes in advance to vendors who have been inadvertently omitted. The systems discussed here have opened up whole new worlds of opportunity. There is almost certainly a system out there that will suit your needs. I hope this chapter helps you find it quickly and efficiently.

4.20 Appendix

4.20.1 Data Capture Vendors

Company	Relevant products	Web address
Browne and Sharpe	CMMs, contact and noncontact probes, CMM software	www.brownandsharpe.com
Renishaw	CMMs, contact and noncontact probes	www.renishaw.com
Romer/ CimCore	Portable CMMs based on arm technology	www.romer.com or www.cimcore.com
Steinbichler	White light and laser 3-D digitizing systems	www.steinbichler.de
Polhemus	Laser measuring and tracking systems	www.polhemus.com
GOM	ATOS and TriTOP noncontact measuring system with integrated software suite	www.gom.com
FARO Technologies Inc	Laser scanner, laser tracker, arms, inspection software, control station	www.faro.com
Immersion Corporation	Portable, low cost desktop measuring arm with contact and noncontact measurement systems	www.immersion.com
Kreon Technologies	CMMS, noncontact measuring probes, software	www.kreon3d.com
Garda	Portable CMMs based on arm technology noncontact measurement probe, software	www.garda-misure.com
Leica GeoSystems	Portable, 3-D laser scanning system that captures, visualizes and models complex building/plant structures and sites	hds.leica-geosystems.com
Metris	Laser scanners and reverse engineering and point cloud manipulation software	www.metris.com
Ziess	Industrial measurement systems and software	www.zeiss.com
Perceptron	Noncontact laser measurement, arm-based system with scan control software	www.perceptron.com
KonicaMinolta	3-D scanners based on laser measurement methods	kmpi.konicaminolta.us

4.20.2 Postprocessing Vendors

Company	Relevant products	Web address
Parametric Technology Corporation	ProEngineer Interactive surface design, ProEngineer reverse engineering, CAD/CAM products	www.ptc.com
Dassault Systems/IBM	CATIA reverse engineering 2 (RE2), CATIA digitized shape editor 2 (DSE), Solidworks, CAD/CAM products	www.3ds.com or www.ibm.com
UGS	NX Imageware, eM-probe CAD, CAD/CAM products	www.ugs.com
Delcam	CopyCAD, powerINSPECT, CAD/CAM products software	www.delcam.com
INUS Technology Inc	Rapidform XO, scan workbench, polygon workbench, curve workbench, surface workbench, inspection workbench, feature workbench, inspect workbench, etc.	www.rapidform.com
HiRes Inc	Reverse engineering software modules to integrate with major CAD vendors' suites	www.reverse-it.com
Geomagic Inc	Inspection, reverse engineering and surfacing software	www.geomagic.com
Revware Inc	Reverse engineering software for use with Solidworks	www.revworks.com
Innovmetric	Integrated point cloud manipulation, surfacing, inspection and reverse engineering suite	www.innovmetric.com
Robert McNeel & Associates	Rhinoceros NURBS modeling software with reverse engineering and rapid prototyping interfaces	www.rhino3d.com

Chapter 5

Introduction to Rapid Prototyping

Eef Moeskops and Frits Feenstra, TNO Science and Industry, Netherlands

Abstract

The term rapid prototyping (RP) refers to a class of technologies that are used to produce physical objects layer-by-layer directly from computer-aided design (CAD) data. These techniques allow designers to produce tangible prototypes of their designs quickly, rather than just two-dimensional pictures. Besides visual aids for communicating ideas with co-workers or customers, these prototypes can be used to test various aspects of their design, such as wind tunnel tests and dimensional checks. In addition to the production of prototypes, rapid prototyping techniques can also be used to produce molds or mold inserts (rapid tooling) and even fully functional end-use parts (rapid manufacturing). Because these are nonprototyping applications, rapid prototyping is often referred to as

Figure 5.1. Prototype of a talking barcode scanner

solid free-form fabrication or layered manufacturing. For small series and complex parts, these techniques are often the best manufacturing processes available. They are not a solution to every part fabrication problem. After all, CNC technology and injection molding are economical, widely understood, available, and offer wide material selection.

In rapid prototyping, the term "rapid" is relative; it aims at the automated step from CAD data to machine, rather than at the speed of the techniques. Depending on the dimensions of the object, production times can be as long as a few days, especially with complex parts or when long cooling times are required. This may seem slow, but it is still much faster than the time required by traditional production techniques, such as machining. This relatively fast production allows analyzing parts in a very early stage of designing, which decreases the resulting design cost. The costs can also be reduced because rapid prototyping processes are fully automated. and therefore, need the skill of individual craftsmen for no more than finishing the part.

General advantages:

- Freedom of design: The production of complex parts is reduced to the accumulation of layers.
- Well automated: No supervision is needed during the build process.
- Relative easy to employ: Only little preparation and postprocessing are required.
- Avoiding the high cost of prototype tooling, allowing (more) design iterations.
- Physical models are easy to check for errors.

General disadvantages:

- Accuracy generally >0.1 mm.
- Material properties: products can be very fragile, and some need postprocessing before they can be handled (as with 3-DP).
- Staircasing effect. Because an inclined surface is constructed using several layers, staircasing will occur.

5.1 The Basic Process

Rapid prototyping techniques share the following process steps (see Figure 5.2):

1. Creating a CAD model either by designing a new or scanning an existing object.
2. Converting the CAD data to STL format. Because the various CAD packages apply a number of different algorithms to represent solid objects, the STL format (Standard Triangulation Language) has been adopted as the standard of the rapid prototyping industry to establish consistency. This STL file is a concrete visualization of the product geometry, built up from triangles. Using triangles to describe a surface, curved surfaces can only be approached. Increasing the number of triangles (*i.e.,* increasing the resolution) yields a better approach. However, it also enlarges the STL file. So, one has to find the optimum balance between file size and part accuracy.
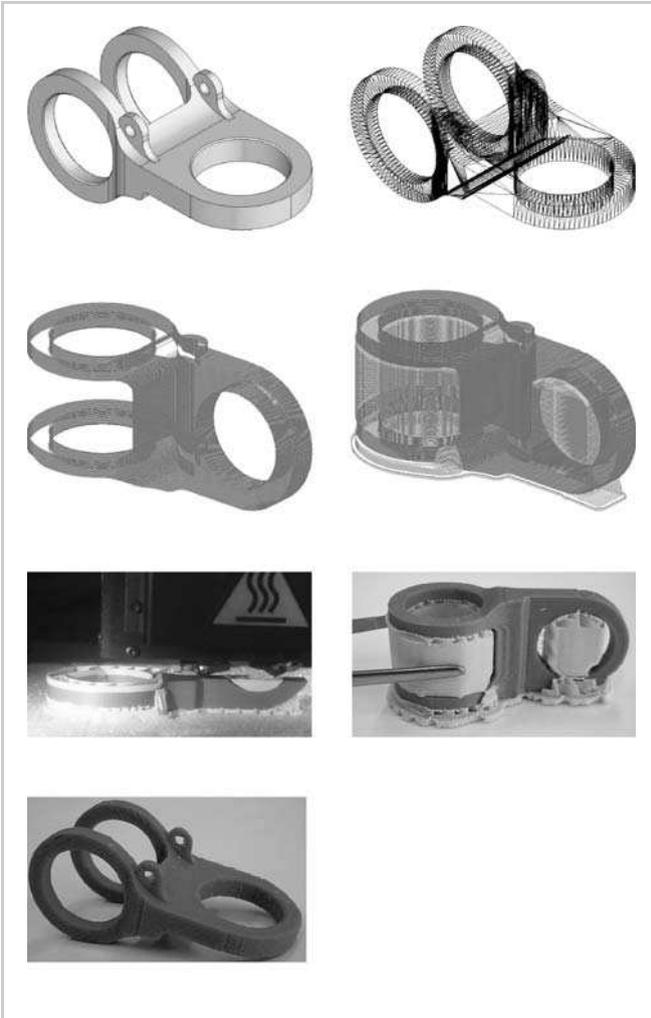
Figure 5.2. The basic RP process (FDM). A color reproduction of this figure can be seen in the Color Section (pages 219–230).

3. Slicing the STL file into thin cross-sectional layers. After the STL file has been sized and oriented, it is sliced in layers with a predefined thickness.

4. Generation of a support structure. This additional step is not required for all techniques. Because the model is built up in layers, there may be areas that could float away or have overhanging features, which could distort the resulting model. A base and support structures have to be added, which can be easily removed after the building step.

5. Producing the model layer-by-layer. The generated slices are reconstructed in the machine by building one layer at a time. This can be fully automatic.
6. Postprocessing. This step enhances cleaning and finishing the model and (if a base or support structure was built) removing the support structure. Some materials need to be postcured or infiltrated to achieve optimal properties.

5.2 Current Techniques and Materials

A wide range of techniques and materials can be used for rapid prototyping. There are more than ten commercial rapid prototyping processes and more than five concept modeling processes; all have unique properties. Due to worldwide research, this range is growing quickly. Commercial techniques are available to produce objects from numerous plastics, ceramics, metals, and wood-like paper. Among these techniques are

- Stereolithography
- Selective laser sintering
- Fused deposition modeling
- Three-dimensional printing
- Laminated object manufacturing
- Multijet modeling
- Laser-engineered net shaping

5.2.1 Stereolithography

Stereolithography (SLA-stereolithography apparatus), launched by 3D Systems Inc. in 1987, is the first and most commercially used rapid prototyping method. A platform is placed in a bath of photosensitive UV-curable resin at a level that leaves a small layer of resin between the top of the platform and the surface of the bath. A laser (often He-Cd or argon ion to produce UV radiation of about 320–370 nm wavelength) then strikes the desired areas, thereby curing the resin selectively.

As the layer is completed, the platform descends allowing liquid resin to flow over the previously cured area. A wiper blade clears the excess fluid from the top of the surface. This sweep is essential to achieve consistent layer thickness and prevent air entrapment. As the new layer is cured, it sticks to the preceding layer. This process continues until the object is completed. On completion, the object raises above the fluid, so that resin can drain out. The object is carefully removed and washed in a solvent to remove uncured resin. The cleaned object has to be placed in a UV oven to ensure that all resin is cured. During the process, features that lean over have to be supported. This support structure can easily be generated by software and consists of a series of slender sacrificial columns or lattices.
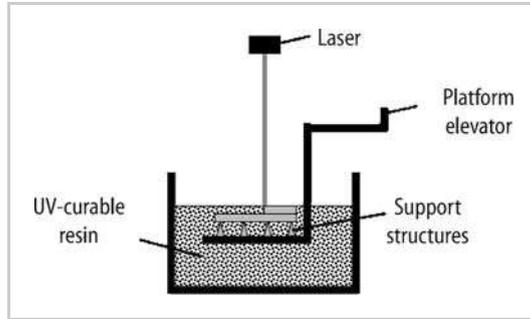
Figure 5.3. Schematic of the SLA process
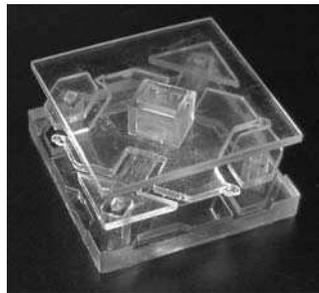
Figure 5.4. A detailed DMD mirror reproduction using microSLA

A lattice structure is also created as a base to prevent the model from sticking to the building platform. Thus, additional hand-finishing will be needed to remove these supporting structures and to sand any small stubs from the surface.

A large variety of photosensitive polymers is commercially available, including clear, water resistant, and flexible resins that simulate the properties of, for example PA, ABS, PP, and rubber-like materials. Process times, tolerances, and surface finish depend on layer thickness, which is controlled by the amount the platform is lowered into the resin. Generally, layer thicknesses vary from 0.05–0.5 mm. Thinner layers can be applied with digital light processing using a technique called perfactory, which is based on the standard SLA process. Instead of describing a cross section with a laser, a normal beamer covers the entire cross section at once. Due to the high resolution of the beamer (pixel size: 39 μm) and the accurate positioning system of the platform (layer thickness: 25 μm), the parts produced can contain highly detailed features.

Characteristics:

- Long-term curing can lead to overcuring which leads to warpage.
- Parts can be quite brittle.
- Support structures are required.
- Uncured material can be toxic.

5.2.2 Selective Laser Sintering

Selective laser sintering (SLS) is a process that was patented in 1989 by Carl Deck-ard, University of Texas. A layer of powder (particle size approximately $50\,\mu$m) is spread over a platform and heated to a temperature just below the melting tem-perature. A carbon dioxide laser needs to raise the temperature only slightly and selectively to melt the powder particles. As the layer is finished, the platform moves down by the thickness of one layer (approximately 0.10–0.15 mm), and new powder is spread. When the laser exposes the new layer, it melts and bonds to the previous layer. The process repeats until the part is complete.

On completion, the built volume has to cool down to room temperature after which the processed objects can be removed from the powder bed by brushing away excess powder. Sandblasting the objects removes all unsintered particles. Surrounding powder particles act as supporting material for the objects, so no
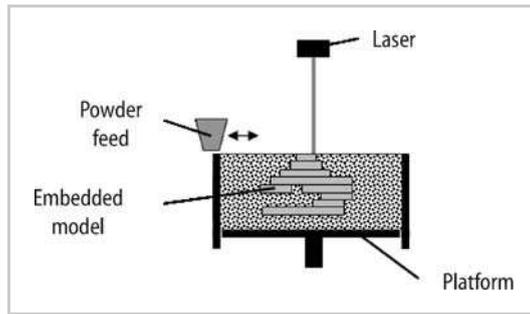
Figure 5.5. Schematic of the SLS process
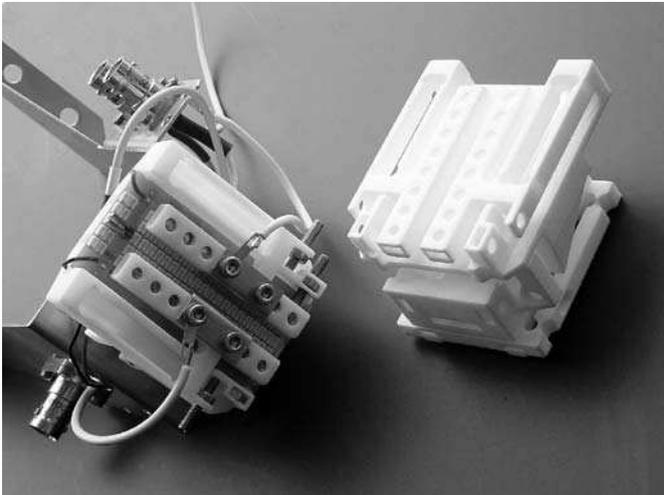
Figure 5.6. Accurate positioning elements with internal hinges produced by SLS

additional structures are needed. Furthermore, more objects can be built at the same time because they can be meshed above/in each other. Excess powder can be reused. However, it needs to be mixed with virgin powder to guarantee good part quality. Commonly used materials for SLS are nylon (polyamide-12), glass-filled nylon, and polystyrene. The method has also been extended to direct fabrication of metal and ceramic objects and tooling inserts.

Characteristics:

- Key advantage of making functional parts in essentially final materials.
- Good mechanical properties, though depends on building orientation.
- Powdery surface
- Many variables to control
- No support required

5.2.3 Fused Deposition Modeling

Fused deposition modeling (FDM), developed by Stratasys, is the second most widely used rapid prototyping process. A filament thread of plastic is unwound from a coil and fed into an extrusion head, where it is heated and extruded through a small nozzle. Because the extrusion head is mounted on a mechanical stage, the required geometry can be described, one layer at a time. The molten plastic solidifies immediately after being deposited and bonds to the layer below. Support material is laid down similarly through another extrusion head. The platform on which the object is built steps down by the thickness of a single layer.

The entire system is contained within a heated oven chamber which is held at a moderate temperature above the glass transition temperature of the polymer. This provides much better control of the process because stresses can relax.

As in the SLA process, overhanging features need to be supported. This support material needs to be removed in secondary operations. Commercially available
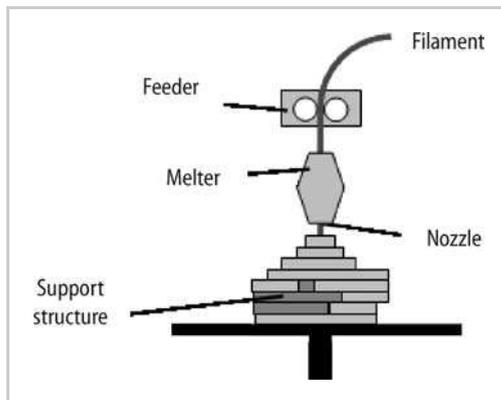
Figure 5.7. Schematic of the FDM process

Figure 5.8. Scanned archery handle and an FDM reproduction

water-soluble support materials facilitate this final step. ABS, polycarbonate, and poly(phenyl)sulfone are commonly used materials in the FDM process.
    Characteristics:

- Office-friendly and quiet.
- FDM is fairly fast for small parts.
- Good mechanical properties, so suitable for producing functional parts.
- Wide range of materials.

5.2.4 Three-dimensional Printing

In some textbooks, the term "three-dimensional printing" (3-DP) is used for all rapid prototyping processes. The process developed at MIT is referred to here. In this process, a layer of powder is spread over a platform. The particles are bonded together selectively by a liquid adhesive (binder solution). This liquid is deposited in a two-dimensional pattern by a multichannel jetting head. As the current layer is completed, the platform moves down by the thickness of a layer, so that a new layer can be spread. This process is repeated until the entire object is formed within the powder bed. On completion, the object is elevated and the
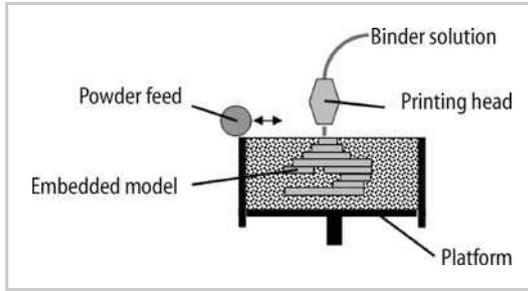
Figure 5.9. Schematic of the 3-DP process
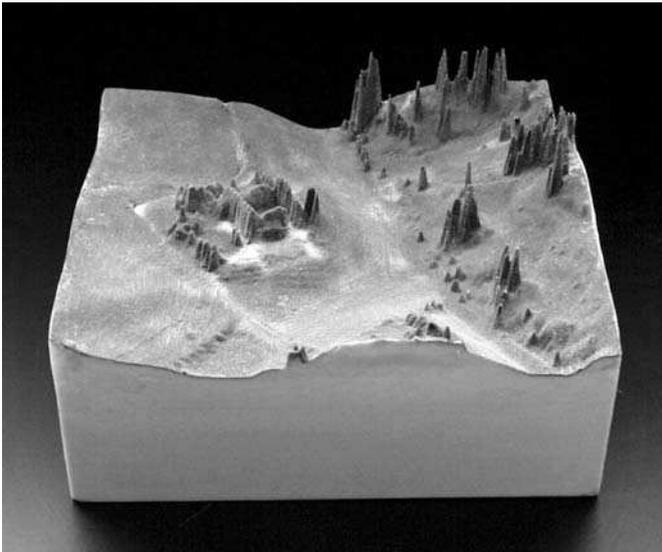
Figure 5.10. 3-D printed landscape

extra powder is brushed away, leaving a fragile "green" object. It is necessary to infiltrate the part with another material to improve mechanical characteristics.

No support structures are required because the surrounding powder particles support overhanging features. By adding color to the binder solution, objects can be produced in every desired color. Starch, plaster, medicines (for producing controlled-dosage pharmaceuticals), ceramics, and metals are commonly used materials (powders) for 3-DP.

Characteristics:

- Limitations on resolution and surface finish.
- Fragile objects need to be infiltrated.

5.2.5 Laminated Object Manufacturing

In laminated object manufacturing (LOM), a sheet of paper (unwound from a feed roll) with a polyethylene coating on the reverse side is placed on a platform. This coating is melted by a heated roller, making the paper adhere to the platform. Then, a carbon dioxide laser cuts out the cross section of the object and a border. The laser also creates hatch marks, or cubes that surround the pattern within the border. These cubes behave as a support structure for the model. When the laser has finished the layer, a new paper sheet is applied.

Upon completion, the model is captured within a block of paper. When all of the surrounding cubes have been removed, the unfinished part is sanded down. The humidity and temperature dependency of the paper material can be reduced by coating the model. The finish and accuracy are not as good as with some other methods; however, objects have the look and feel of wood and can be worked and finished like wood.
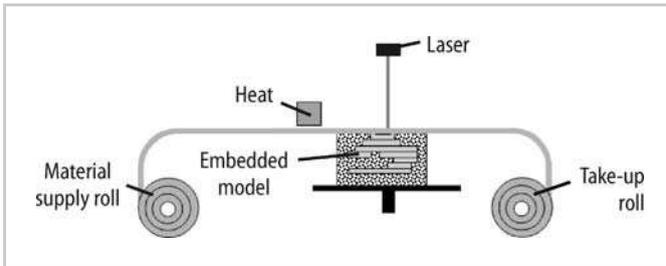
Figure 5.11. Schematic of the LOM process
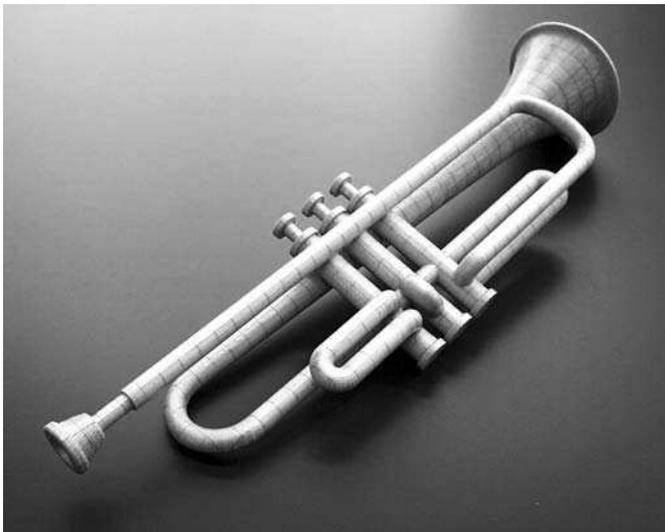
Figure 5.12. Trumpet prototype using LOM

5.2.6 Multijet Modeling

Multijet modeling (MJM) uses multiple print heads to deposit droplets of material in successive, thin layers. Two major MJM techniques can be distinguished (see http://www.3dsystems.com/ for more information): ThermoJet™. A 96-element print head deposits droplets of wax. Because of its relatively fast production, this technique is marketed to the engineering or design office for quick form studies (concept modeling). However, wax models can also be used as master patterns for investment casting, as will be explained later.

InVision™. A print head jets two separate materials, an acrylic UV-curable photopolymer-based model material and a wax-like material to produce support structures for the model. Due to the relative good quality of the models,
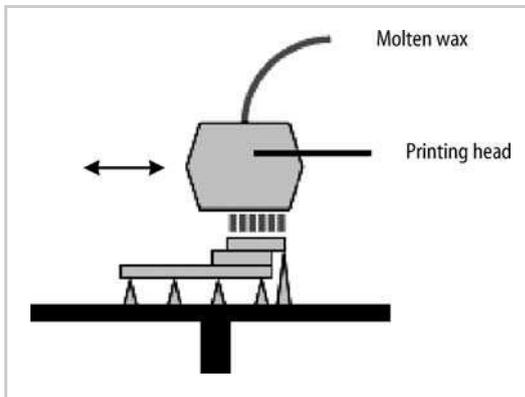
Figure 5.13. Schematic of the ThermoJetTM process
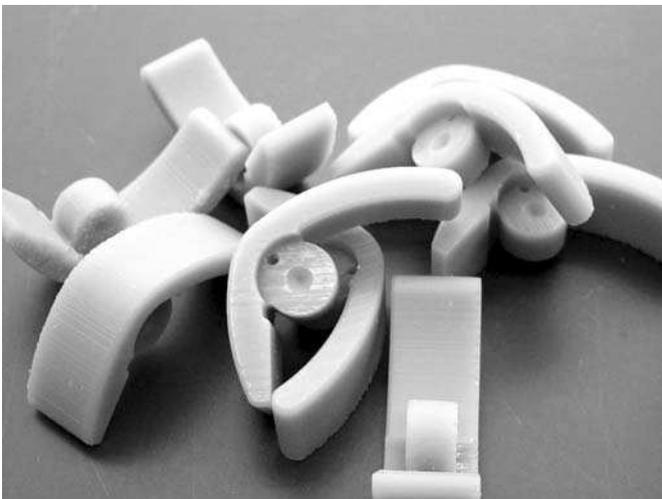
Figure 5.14. Wax models produced by MJM
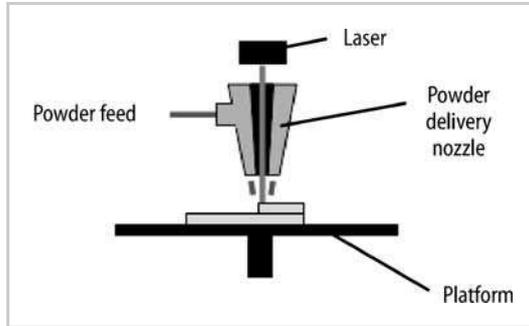
Figure 5.15. Schematic of the LENS process

production speed, and surface finish, applications range from preliminary pro-
totypes to mock-ups for concept proposals or marketing models.

5.2.7 Laser-engineered Net Shaping

In laser-engineered net shaping (LENS), a laser beam focuses onto a metal sub-
strate to melt the upper surface. A deposition head then applies metal (powder or
fine wire) into the molten puddle to increase the material volume. By moving the
platform in raster fashion, each layer of the object is fabricated. An inert gas is used
to shield the melt puddle from atmospheric oxygen for better control of properties
and to promote layer-to-layer adhesion by providing better surface wetting.

Fully dense metal parts (made of stainless steel, aluminum, copper, Inconel,
titanium, *etc.*) can be produced by LENS. It is even possible to change the mate-
rial composition dynamically, which lead to objects with properties that might
be mutually exclusive using traditional fabrication methods. Although produced
parts are near net shape, they generally require postprocessing. Applications of
LENS are injection molding tools and aerospace parts.

5.3 Applications

As stated before, RP models were primarily used as visual aids (for engineering,
tool-making, quote requests, and proposals). Due to the use of other more dura-
ble materials, the field of application has enlarged substantially. Companies are
increasingly applying additive RP processes to produce tools for various indus-
trial production processes (rapid tooling). The quality of RP models (durability,
accuracy, reproducibility) makes them suitable for end uses (rapid manufactur-
ing). The following graph shows the use of RP models. Of all RP models, 33.9% is
used for fit and function applications. More than one-fourth of all RP models are
being used as patterns for prototype tooling and metal casting, as well as for tool-
ing inserts. The category "direct manufacturing" represents 6.6%.

5.3.1 Rapid Prototyping

As stated before, RP models improve communication to ensure that all parties in a design process have a complete understanding of the design. So RP models provide the ability to manage, control, and detect changes and required modifications. However, RP techniques can also be applied to the needs of disciplines outside of design engineering. Scanned human parts (bones, tissue) can easily be reproduced using additive RP processes. In 1998, the head of a 16-year-old

Quoting	1.3%
Ergonomic studies	3.0%
Proposals	3.9%
Other	4.6%
Visual aids for toolmakers	5.0%
Direct manufacturing	6.6%
Tooling components	7.2%
Patterns for cast metal	8.3%
Patterns for prototype tooling	10.6%
Fit/assembly	12.5%
Visual aids for engineering	15.6%
Functional models	21.4%
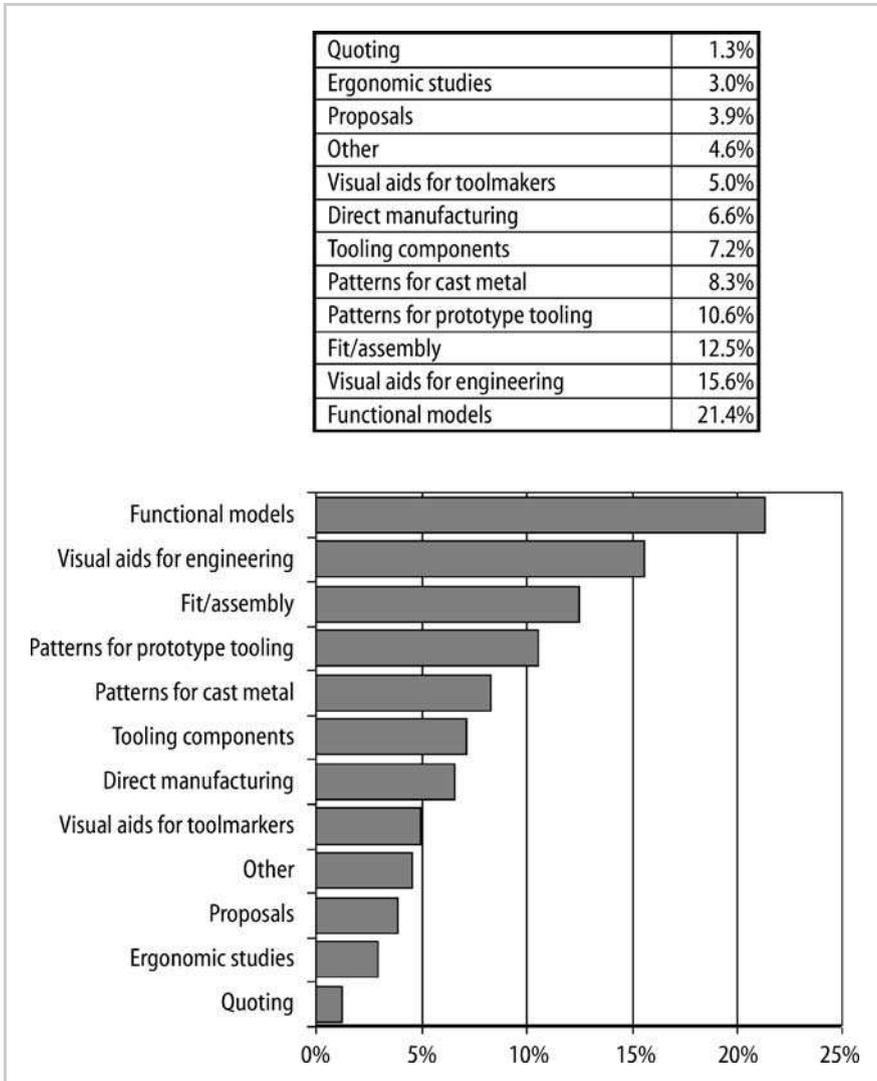
Figure 5.16. Use of RP models

Egyptian girl named Sensaos, who died in 109 AD, was reconstructed after researchers scanned Sensaos' mummy with a tomographic scanner. The computerized model was used to reproduce the skull using RP at TNO Science and Industry (The Netherlands). Realistic facial features could than be added to the skull to reconstruct the girl's head.

5.3.2 Rapid Tooling

The automated production of hard and wear-resistant parts with low surface roughness using rapid prototyping techniques is referred to as (direct) rapid tooling. Metal molds and dies can be produced by layered manufacturing, which in
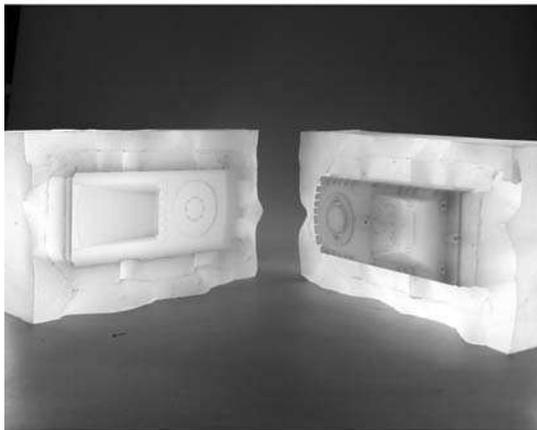
Figure 5.17. Vacuum casting of a talking barcode scanner. A color reproduction of this figure can be seen in the Color Section (pages 219–230).

turn can be used in limited mass production. The use of conformal-cooling channels is a method of rapid tooling that offers the potential benefit of improved thermal properties which cannot be realized with conventional machine tools. Because the mold (or mold insert) is produced in layers, cooling channels that closely follow the contours of the part can be incorporated into the mold. So, conformal cooling removes heat from the mold or die faster than the straight-line channels in machined tools, thereby reducing molding cycle time and tool distortion and improving tool life. Rapid prototyping models can be indirectly used for reproduction. Vacuum casting is the simplest and oldest indirect rapid tooling technique. Using this technique, a model is submerged in a bath of liquid silicone. When the silicone hardens, it is cut open and the model is removed. Reassembling the silicone parts leaves a gap that has the exact shape of the original model. By filling this gap with a two-component resin under vacuum, a reproduction can be created. Silicone molds can last for as many as 20 reproductions.

Investment casting is another indirect rapid tooling process. A wax model is placed in a bed of ceramic powder. When heated, the ceramic forms a hard shell as the model melts. The remaining gap can be filled with liquid metal. After cooling, the shell has to be removed.

5.3.3 Rapid Manufacturing

In general, there are two types of models in free-form fabrication:

- Shape (or concept) model, when using the model to evaluate the shape and dimensions of a product.
- Functional model. If the mechanical behavior of the model suits the designer's wishes for its application, the model is called a functional model.

Combining both types, as in producing fully functional end-use products, is termed rapid manufacturing (RM). The time to market and product expense can be further decreased when using RP to produce finished manufactured parts. RM has the ability to create end-use products with almost unlimited complexity. Furthermore, the techniques are suitable for mass customization.

Examples are hearing aids and nearly invisible teeth aligners made by Align Technology (Santa Clara, California). Especially, when only a small number of products is needed, RM can replace conventional production techniques. Because there is almost no limit to the complexity of products in layered manufacturing, RM can be a very helpful tool in topology optimization (*i.e.,* optimization for weight or material). In topology optimization, an engineer defines the building envelope and the requirements of a part. Commercially available software optimizes the shape according to those requirements. In most cases, the optimal structure becomes very complex, maybe even hollow, which makes it hard (or impossible) to produce this kind of shape using conventional techniques. So the engineer needs to reconstruct the optimized shape. Because of compromises to

Figure 5.18. A Typical RM application. A color reproduction of this figure can be seen in the Color Section (pages 219–230).

be made, this final shape will not be the best shape, but just an approximation of it. Using layered manufacturing, the last reconstruction operation becomes superfluous because the technique can create this complex shape, probably without any problem.

5.4 Future

As stated before, RP is changing the way companies design and build products. Rapid prototyping has developed into three primary technological and application areas: concept modeling, rapid prototyping for fit and function applications, and rapid manufacturing. It is believed that these three forms of free-form fabrication will grow exponentially over time, as indicated in Figure 5.19.

	Concept modeling	Rapid prototypig	Rapid manufacturing		Concept modeling	Rapid prototypig	Rapid manufacturing
0	3	5	1.2956	26	15.2	6.04	5009.913399
1	4	5.04	1.780107772	27	15.4	6.08	6883.440704
2	5	5.08	2.445804012	28	15.6	6.12	9457.599794
3	6	5.12	3.360446689	29	15.8	6.16	12994.40174
4	7	5.16	4.617132809	30	16	6.2	17853.84032
5	8	5.2	6.34377431	31	16.2	6.24	24530.53405
6	9	5.24	8.71611759	32	16.4	6.28	33704.07094
7	10	5.28	11.97563188	33	16.6	6.32	46308.18046
8	11	5.32	16.4540872	34	16.8	6.36	63625.77333
9	11.5	5.36	22.60732363	35	17	6.4	87419.52268
10	12	5.4	31.06164904	36	17.2	6.44	120111.2779
11	12.2	5.44	42.67758787	37	17.4	6.48	165028.5731
12	12.4	5.48	58.63746978	38	17.6	6.52	226743.3201
13	12.6	5.52	80.56577313	39	17.8	6.56	311537.1614
14	12.8	5.56	110.6944727	40	18	6.6	428040.8476
15	13	5.6	152.0902217	41	18.2	6.64	588112.7195
16	13.2	5.64	208.9664909	42	18.4	6.68	808045.7105
17	13.4	5.68	287.1124378	43	18.6	6.72	1110225.725
18	13.6	5.72	394.4821564	44	18.8	6.76	1525410.189
19	13.8	5.76	542.0042855	45	19	6.8	2095858.701
20	14	5.8	744.6943817	46	19.2	6.84	2879634.425
21	14.2	5.84	1023.183279	47	19.4	6.88	3956514.063
22	14.4	5.88	1405.817001	48	19.6	6.92	5436107.928
23	14.6	5.92	1931.541965	49	19.8	6.96	7469016.648
24	14.8	5.96	2653.869144	50	20	7	10262160.07
25	15	6	3646.320691				

□ Concept modeling ■ Rapid prototypig ■ Rapid manufacturing
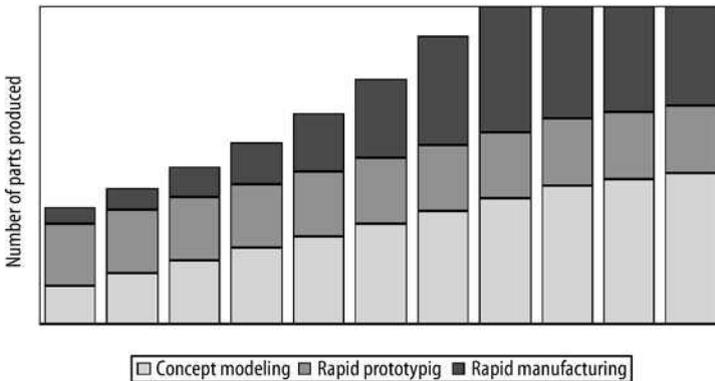
Figure 5.19. Growth of free-form fabrication over time

To achieve this growth, several technological developments are needed to revolutionize manufacturing:

Materials. Although some techniques can produce parts with fair mechanical properties, more rugged materials are necessary for parts that are subjected to actual service conditions. The number of commercially available metal, composite, and ceramic materials is growing rapidly due to worldwide research in labs and universities.

Size. On the one hand, part size is limited by the capacity of RP machines. For instance, the biggest build volume in SLS is $700 \times 380 \times 580$ mm (EOSINT P700). Larger parts must be built in sections and joined by hand. On the other hand, part size is limited by the accuracy and layer thickness of the process. The smallest detail of a model can be approximately 0.08 mm. Improvements in materials, optics, and machine controls should increase this accuracy.

Figure 5.20. A rapid fabrication dream

Speed. Some RP techniques are still slow by some standards. Faster computers, more complex control systems, and improved materials help to reduce build times. This reduction will make RP economical for a wider variety of products.

Costs. The reduction in production time will also reduce the cost per part.

Worldwide, a lot of research is carried out to optimize the above mentioned aspects. At TNO Science and Industry (The Netherlands), a rapid fabrication process for dental elements was investigated. In this process, a dental laboratory scans the tooth stump. The scan data are transported to a knowledge system that creates the digital product geometry and steers a production machine. The production time for a complex crown is two weeks by traditional techniques. Using the new process, the production time can be reduced to a maximum of four days. Figure 5.20 illustrates a rapid fabrication dream: fast production of customized dental elements in a dentist's workroom.

Production Volume	Injection moulding	SLA	SLS	FDM
0		5.2	2.2	4.6
50	580.7964567	5.2	2.2	4.6
100	291.7671282	5.2	2.2	4.6
150	195.0472485	5.2	2.2	4.6
200	146.5712404	5.2	2.2	4.6
250	117.4346494	5.2	2.2	4.6
300	97.98333803	5.2	2.2	4.6
350	84.07360208	5.2	2.2	4.6
400	73.63107919	5.2	2.2	4.6
450	65.50217152	5.2	2.2	4.6
500	58.99411062	5.2	2.2	4.6
550	53.66570141	5.2	2.2	4.6
600	49.22260943	5.2	2.2	4.6
650	45.46093643	5.2	2.2	4.6
700	42.2349571	5.2	2.2	4.6
750	39.43774963	5.2	2.2	4.6
800	36.98908331	5.2	2.2	4.6
850	34.8275772	5.2	2.2	4.6
900	32.90547016	5.2	2.2	4.6
950	31.18504082	5.2	2.2	4.6
1000	29.63610063	5.2	2.2	4.6
1050	28.23420247	5.2	2.2	4.6
1100	26.95933731	5.2	2.2	4.6
1150	25.79497059	5.2	2.2	4.6
1200	24.72731923	5.2	2.2	4.6
1250	23.74480203	5.2	2.2	4.6
1300	22.83761672	5.2	2.2	4.6
1350	21.99741093	5.2	2.2	4.6
1400	21.21702362	5.2	2.2	4.6
1450	20.49027995	5.2	2.2	4.6
1500	19.81182705	5.2	2.2	4.6
1550	19.17700167	5.2	2.2	4.6
1600	18.58172254	5.2	2.2	4.6
1650	18.0224023	5.2	2.2	4.6
1700	17.49587496	5.2	2.2	4.6

Production Volume	Injection moulding	SLA	SLS	FDM
1750	16.99933578	5.2	2.2	4.6
1800	16.53029115	5.2	2.2	4.6
1850	16.08651663	5.2	2.2	4.6
1900	15.66602154	5.2	2.2	4.6
1950	15.26701904	5.2	2.2	4.6
2000	14.88790069	5.2	2.2	4.6
2050	14.52721473	5.2	2.2	4.6
2100	14.18364743	5.2	2.2	4.6
2150	13.85600706	5.2	2.2	4.6
2200	13.54321008	5.2	2.2	4.6
2250	13.24426907	5.2	2.2	4.6
2300	12.95828238	5.2	2.2	4.6
2350	12.68442493	5.2	2.2	4.6
2400	12.42194032	5.2	2.2	4.6
2450	12.17013376	5.2	2.2	4.6
2500	11.92836599	5.2	2.2	4.6
2550	11.69604777	5.2	2.2	4.6
2600	11.47263515	5.2	2.2	4.6
2650	11.25762513	5.2	2.2	4.6
2700	11.05055194	5.2	2.2	4.6
2750	10.85098361	5.2	2.2	4.6
2800	10.65851897	5.2	2.2	4.6
2850	10.47278497	5.2	2.2	4.6
2900	10.29343424	5.2	2.2	4.6
2950	10.12014293	5.2	2.2	4.6
3000	9.952608724	5.2	2.2	4.6
3050	9.790549117	5.2	2.2	4.6
3100	9.633699792	5.2	2.2	4.6
3150	9.481813189	5.2	2.2	4.6
3200	9.334657193	5.2	2.2	4.6
3250	9.192013951	5.2	2.2	4.6
3300	9.053678791	5.2	2.2	4.6
3350	8.91945924	5.2	2.2	4.6
3400	8.789174129	5.2	2.2	4.6
3450	8.662652774	5.2	2.2	4.6
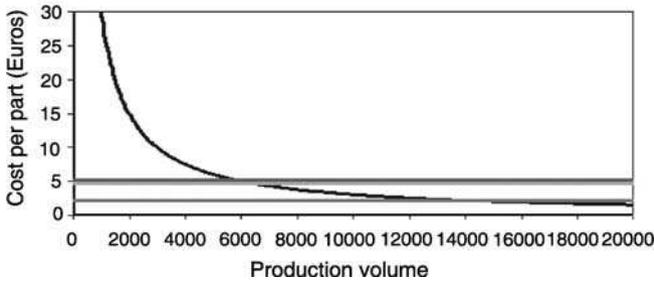
Figure 5.21. Relationship between cost per part and production volume

Chapter 6

Relationship Between Reverse Engineering and Rapid Prototyping

Y.F. Zhang, Y.S. Wong, and H.T. Loh,
National University of Singapore, Singapore

Abstract

In recent years, advanced reverse engineering (RE) techniques have been developed to convert point cloud data, obtained through digitization (by contact or noncontact scanning), into CAD models either in NURBS or STL (triangular mesh) format. These CAD models can be subsequently used for fabrication by using either conventional material removal methods (*e.g.*, milling) or material incremental methods, *i.e.*, rapid prototyping (RP). On the other hand, as a nonconventional fabrication method, RP has been gaining greater popularity in industry recently due to its capability of creating three-dimensional (3-D) parts with complex geometries. Typically, the CAD model in the STL format is first sliced into 2-D layers before RP fabrication can be carried out. The fabrication efficiency and accuracy are directly affected by the thickness of the layers.

Recently, there has been much research on developing adaptive slicing techniques to determine the minimum number of layers required under an allowable shape error. Unrtil now, the modeling algorithms developed in RE and adaptive slicing algorithms in RP are quite capable of handling problems in their own domains. When using RP to fabricate a part based on its point cloud data, the RE modeling and RP slicing algorithms are used sequentially. This, however, can result in serious problems in error control because in either of the two processes, shape error is controlled within the given tolerance band independently. Therefore, the shape error between the finally fabricated part and the original point cloud data cannot be controlled effectively. To solve this problem, the RE modeling process and the RP slicing process must be integrated by following a model-for-RP approach.

This chapter focuses on issues in RE for direct RP by describing an innovative modeling method that slices the point cloud directly along the part building direction and generates a layer-based model (RP model) that can be used directly for fabrication by RP techniques. With this method, the integration of RE and RP in shape-error control can be

effectively achieved. Furthermore, the slicing algorithm is also adaptive; thus the RP model generated has a minimum number of layers. Issues discussed include cloud data filtering, 2-D curve generation from point cloud data, and optimal layer thickness. Finally, the efficacy of the algorithms is demonstrated by several case studies.

6.1 Introduction

6.1.1 Modeling Cloud Data in Reverse Engineering

Reverse engineering (RE) is the process of creating a CAD model from a physical part or prototype (Varady *et al.*, 1997). In the conventional product development cycle, the initial conceptual or aesthetic design of sculptured surfaces is often described by stylists who formalize their ideas by making clay or wooden models. Recently, RE is increasingly employed in medical applications. For example, physical models are molded from a part of the body, such as the stump of an amputated limb, and an appropriate model is then captured and developed from the mold and used to custom-fit an artificial limb. Necessary CAD models can be generated using RE technology based on the corresponding geometric database, which is usually created by capturing the shape information from the original physical model by mechanical contact or optical noncontact measuring techniques. The mechanical contact measurement equipment, such as a coordinate measurement machine (CMM) equipped with a touch probe, is an important digitizing tool for data acquisition because its measurement accuracy can achieve up to 10 μm or less. However, compared with the optical noncontact measurement equipment, such as a laser scanner, the contact measurement speed is very slow. Furthermore, the measurement tends to be time-consuming and even difficult for free-form surfaces. A laser scanner is an effective digitizing device for 3-D geometric shapes, from which sufficient view scans of an object can be executed rapidly and a cloud data set obtained. Figure 6.1a shows a cloud data set obtained by digitizing a face mask with a laser scanner. The data set is combined with four range data patches that contain 104,175 points. In the RE process, the faster noncontact optical measuring techniques have significantly reduced the time for data acquisition. However, the captured cloud data set is usually highly dense, randomly distributed, and partially overlapping.

Conventional approaches to transforming a cloud data set to a CAD representation can be classified in two categories: the triangular polyhedral mesh method and the segment-and-fit method. In the first method, an initial triangular mesh is constructed to capture the unknown topological structure of the scattered data. The mesh is then optimized to reduce redundant vertices, and thereafter a curvature-continuous surface is reconstructed. Many triangulation techniques have been reported. The 2-D Delaunay triangulation method was extended to 3-D by Fang and Piegl (1995). Another Delaunay triangulation technique, based on a divide-and-conquer paradigm, was proposed by Cignoni *et al.* (1998).
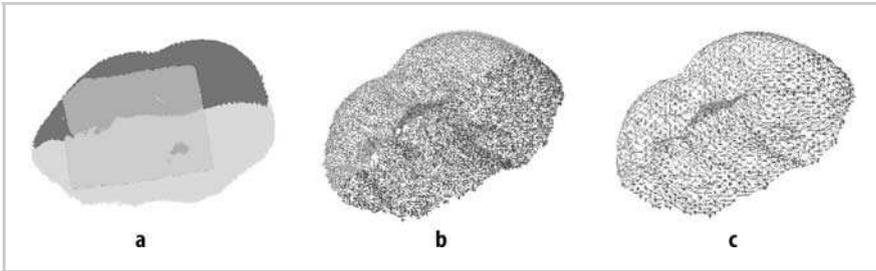
Figure 6.1. The point cloud data of a face mask and its triangular meshes. (**a**) Cloud data set. (**b**) Triangular mesh model with 5410 vertices. (**c**) Triangular mesh model with 1441 vertices. Reprinted from *Computer-Aided Design*, Vol 33, Sun W, Bradley C, Zhang Y, Loh H, Cloud data modelling employing a unified, non-redundant triangular mesh, pp. 188–189, Copyright (2001) with permission from Elsevier. A color reproduction of this figure can be seen in the Color Section (pages 219–230).

Lawson (1977) used geometric reasoning to construct a triangular facet mesh, and subsequently Choi *et al.* (1988) extended the same method using a vector angle order instead of Euclidean distance to determine the linkage of data points. Recently, Hoppe *et al.* (1992) developed a signed distance function by estimating the local tangential plane and using a marching cube method to extract a triangular polyhedral mesh.

In part digitization, fine scanning resolution is generally applied to capture all fine features on the surface of an object. This produces a cloud data set of very high density. Therefore, the point cloud data need to be thinned before a triangular mesh is generated. However, arbitrary removal of cloud data points, based on simple data thinning methods (Weir *et al.* 1996), may not ensure that the necessary accuracy between the triangulation and original data is achieved. Therefore, cloud data reduction methods are required that retain sufficient cloud data to satisfy the error specification between the triangulation and the original data. Sun *et al.* (2001) and Shi (2002) extended the voxel binning method (Weir *et al.* 1996) by directly relating the bin size to the specified shape-error bound using a local surface fitting method. The triangular mesh model generated using this method retains a close to minimum number of points that meet the given shape-error bound. By applying their algorithm to the cloud data set in Figure 6.1a, the triangular mesh models shown in Figure 6.1b (with 5,410 vertices) and Figure 6.1c (with 1,441 vertices) were developed when the shape-error bounds were 0.4 and 1.3 mm, respectively.

In the segment-and-fit method, as discussed by Hoffman and Jain (1987), the cloud data are segmented into bounded patches, each representing a discrete surface region on the physical object. Cloud data segmentation, accomplished either manually or automatically, defines the patch boundary curves. Data modeling methods, such as those employing parametric (Varady *et al.* 1997), quadric functions (Chivate and Jablokow, 1993; Weir *et al.*, 1996; Yang and Lee 1999), and nonuniform rational B-spline (NURBS) curves and surfaces (Chen and Liu, 1997) are then applied to fit appropriate surfaces to the data patches. Because

there is no prior knowledge of the patches that form the entire surface represented only by the point cloud, segmentation of the entire set of measured data points into regions according to shape-change detection has been a challenging and long-standing research issue. This will subsequently present enormous difficulty in controlling the shape error between the fitted surface patches and the point cloud.

6.1.2 Data Processing for Rapid Prototyping

Rapid prototyping (RP) is a nonconventional fabrication process that builds parts layer-by-layer. The use of RP can significantly reduce prototyping time and also the impact of the geometric complexity of the part on the fabrication process. It has been widely used in industry recently for prototype making, mechanism interference checking, conceptual design, and rapid tooling. Although many RP systems, such as stereolithography apparatus (SLA), fused deposition modeling (FDM), and selective laser sintering (SLS), are available on the market, the basic difference among these technologies is the methods by which the materials are processed. The underlying principle for the data processing of most RP systems is the same: slicing the 3-D part CAD model into 2-D layers.

The slicing process generally begins with a given 3-D CAD model in the STL format (CAD models in other formats, *e.g.*, NURBS, need to be converted into the STL format first). It cuts the 3-D model into layers to obtain a layer-based model, each layer represented by a 2-D contour and its thickness. The simplest method is to use a constant thickness for all layers, but the shape errors (staircase error) vary due to changes in the surface geometry. To meet the given shape-error tolerance, the layer with the minimum thickness must be used for all layers. This will make RP fabrication inefficient. Therefore, adaptive slicing has become a keen research topic Because the mid-1990s (Dolenc and Mäkelä 1994; Kulkarni and Dutta 1995; Sabourin *et al.* 1997; Tyberg and Bohn 1998; Lin 2001). The adaptive slicing technique determines the minimum number of layers required under the given shape-error bound.

6.1.3 Integration of RE and RP for Layer-based Model Generation

The cloud data modeling algorithms developed in RE and the adaptive slicing algorithms in RP have effective shape-error control mechanisms toward their own final output, respectively, However, problems arise when converting the point cloud data to a layer-based model for RP by applying the RE modeling algorithm and the RP adaptive slicing algorithm sequentially. First, shape-error control becomes very difficult, if not impossible. If a segment-and-and fit algorithm is applied, a surface model (*e.g.*, NURBS) will be generated in the RE process. Subsequently, the surface model needs to be converted to an STL model, and

finally the STL model is adaptively sliced to generate a layer-based model (RP model).

The shape error of the final RP model (between the RP model and the cloud data) comes from three sources: (1) shape error between the cloud data and the surface model, (2) shape error between the surface model and the STL model, and (3) shape error between the STL model and the layer-based RP model. Because the three processing stages are carried out sequentially, these errors are not directly related and thus are very difficult to control. When a triangulation algorithm is applied, an STL model is generated. However, shape errors from the last two processing stages will contribute to the shape error of the RP model, which makes it still difficult to control the overall shape error, although it is slightly better than that in the first case. On the other hand, when the data processing in RE and RP is carried out separately, it also results in rather expensive remodeling computations and large file transmissions. To resolve the error control and low efficiency problems when converting point cloud data to an RP model, data processing in RE and RP must be integrated.

To combine the data processing stages in RE and RP, one promising approach is to slice the point cloud into layers of points and for each layer, to construct a 2-D contour with the same thickness to represent the layer of points. For each layer, a single shape error between the layer of contour and the layer of points can be identified and effectively controlled by varying the layer thickness. It is also an adaptive slicing method, but the object is the point cloud, not the STL model.

Using this approach, a team at the National University of Singapore has developed several algorithms. Liu *et al.* (2002) developed an automated segmentation approach for generating a layer-based model directly from point cloud data. This is accomplished in three steps. First, the cloud data are adaptively subdivided into a set of regions according to a given subdivision error (the maximum distance between cloud data points and their respective projected plane), and the data in each region are compressed by keeping the feature points (FPs) within the user-defined shape tolerance using a digital image reduction method. Second, based on the FPs of each region, an intermediate point-based curve model is constructed, and RP layer contours are then extracted directly from the models. Finally, the RP layer contours are faired and subsequently closed to generate the final layer-based RP model.

Applying this algorithm to the point cloud data shown in Figure 6.1a, two layer-based models were developed in two orthogonal slicing directions, as shown in Figure 6.2a and b, respectively. It was demonstrated that the system developed can generate a layer-based model from point cloud data. However, the subdivision error, which is used to control the layer thickness, does not have an explicit relationship with the shape error, thus making the actual shape error still difficult to control. Later, Wu *et al.* (2004) developed an intuitive method of point cloud data segmentation by using the shape error to control the layer thickness directly so that each layer will yield the same shape error. This advanced modeling
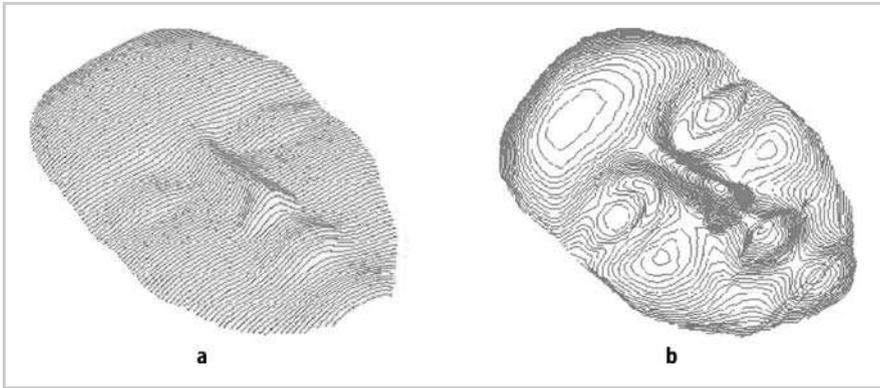
Figure 6.2. Two RP models and one RP fabricated part using SLA. (**a**) The point cloud data. (**b**) The layer-based model. Reprinted from *Computer-Aided Design*, Vol 35, Liu G, Wong Y, Zhang Y, Loh H, Error-based segmentation of cloud data for direct rapid prototyping, p. 641, Copyright (2002) with permission from Elsevier.

method will be described in this chapter. To implement this method for RP manufacturing, the RP machine must be able to change the layer thickness during material deposition.

6.2 The Adaptive Slicing Approach for Cloud Data Modeling

There are two basic steps in the proposed adaptive slicing approach for cloud data modeling: (1) slicing the point cloud data into a number of layers of points and (2) constructing a layer represented by a 2-D contour and a certain thickness. In the first step, the cloud data are segmented into a number of layers by slicing the point cloud in a user-specified direction (building direction). The data points in each layer are projected onto a plane that is perpendicular to the building direction. An example of this processing step is illustrated in Figure 6.3. In the second step, for each layer of points, these projected data points are used to reconstruct a 2-D polygon approximating the contour. Finally, the 2-D polygon is extruded in the building direction, with the same thickness as the layer of points in hand, to approximate the layer of points. In this processing step, some points are retained to form the vertices of the 2-D polygon while the rest are removed; thus data thinning is effectively carried out. The shape error between the original points in the layer and the 2-D polygon-based layer can be obtained by calculating the maximum deviation between the original points and the 2-D polygon on the projecting plane. The maximum shape error among all layers represents the shape error between the fabricated RP part and the point cloud data. Given a shape-error tolerance, the shape error of each layer must be kept below the tolerance value. Preferably, the shape error for each layer is very close
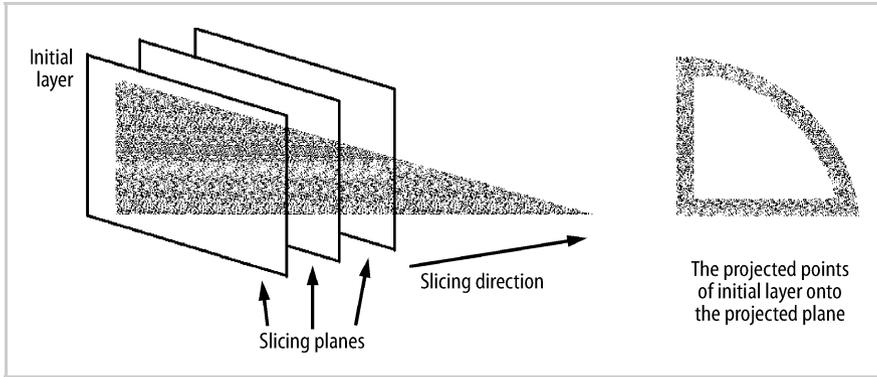
Figure 6.3. Point cloud slicing and projecting. Reprinted from *Computer-Aided Design*, Vol 36, Wu Y, Wong Y, Loh H, Zhang Y, Modelling cloud data using an adaptive slicing approach, pp. 232–239, Copyright (2004) with permission from Elsevier.

to the shape-error tolerance, thus maximizing the thickness of each layer. To achieve this, an adaptive slicing method is used.

To implement the adaptive slicing method, determination of the maximum layer thickness must be iterative because the actual shape error for each layer can be calculated only after the corresponding 2-D polygon is constructed. This iterative process can be illustrated using Figure 6.3. Given a slicing direction (also the RP building direction), the extreme points from both ends can be identified for the cloud data. The initial layer is obtained from one end with a sufficiently small thickness. The mid-plane in the initial layer is used as the projection plane, and the points within the layer are projected onto this plane. The 2-D points on the plane are then used to construct a closed polygonal curve. The distances between the points and the polygon are then calculated as the actual shape errors (σ_{actual}). If σ_{actual} is smaller than the given shape tolerance (σ_{given}), the layer thickness is adaptively increased until that σ_{actual} is very close to σ_{given}. This final thickness is the maximum allowable thickness for the first layer. The first layer is thus constructed by extruding the polygon in the slicing direction with the determined maximum allowable thickness. The subsequent layers are constructed, one by one, in the same manner. The details of the adaptive slicing algorithm, including approximating 2-D points in a contour using a 2-D polygon and searching for the maximum thickness of a layer, are described in the following sections.

6.3 Planar Polygon Curve Construction for a Layer

Once a layer of points is obtained, the points within the layer are projected (in the slicing direction) onto the projection plane. The next step is to construct one or several closed polygon curves to represent accurately the shape defined by

these points. Because each polygon curve is closed, these polygon curves are constructed one by one, and different curves (in the case of multiple loops) are split naturally. Here, we discuss only single- loop curve construction.

Curve reconstruction is to approximate the unorganized point set by a curve. Because the projected points in our application have local linearity, we can use line segments to represent the local shape of points and thus a 2-D polygon is formed. To approximate the point set accurately, the polygon must keep the feature points (or inflection points) of the original shape defined by the point set. Liu *et al.* (2002) present an algorithm to construct a feature-based planar curve from the unorganized data points. In their algorithm, the data points are first sorted based on the estimated oriented vector to generate the initial curve. It begins with a fixed point and then this point and its neighborhood point center determine the oriented vector of a point, from which the next point can be retrieved. The feature points are determined repeatedly. The data are compressed by removing redundant points other than the feature points. Finally, the curve is obtained by linking all of the feature points using straight line segments. However, when determining the oriented vector of a point in their algorithm, they used a fixed radius of neighborhood points. This could result in losing some feature points if the radius chosen is too large. If the radius chosen is too small, the algorithm lacks efficiency. Lee (2000) used the concept of correlation in probability theory (Pitman, 1992) to compute a regression curve. Here, this correlation concept is adopted to determine the radius of neighborhood adaptively in the process of curve construction.

6.3.1 Correlation Coefficient

Correlation refers to the degree of association between two or more quantities. In a 2-D plot, the correlation coefficient is used to measure the strength of the linear relationship between two variables on the two axes. Let X and Y be two variables. Then, the correlation coefficient of X and Y can be defined as

$$\rho(X,Y) = \frac{Cov(X,Y)}{S(X)S(Y)}, \tag{6.1}$$

where $Cov(X,Y) = E\big[(X-E(X))(Y-E(Y)\big] = E(XY) - E(X)E(Y)$ and $E(\zeta)$ denotes an expectation of a random variable ζ. $S(\zeta)$ represents a standard deviation of a random variable ζ. Let (X, Y) stand for a set of N data points $\{P_i = (x_i, y_i), |i = 1, \ldots, N\}$; then Equation (6.1) can be rewritten as

$$\rho(X,Y) = \left| \frac{\sum_{i=1}^{N}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sqrt{\sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2}\sqrt{\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2}} \right|, \tag{6.2}$$
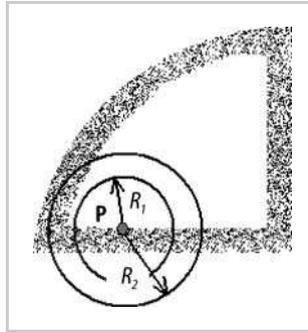
Figure 6.4. Correlation coefficients of neighborhood points of point **P**

where $\bar{x}$ and $\bar{y}$ are the average values of $\{x_i\}$ and $\{y_i\}$, respectively. $\rho\,(X,\,Y)$ has a value between 0 and 1 representing the degree of linear dependence between X and Y. In our application, we use this idea to check the linearity of the points with a neighborhood. Figure 6.4 shows a point P with two neighborhood radii, R_1 and R_2. The correlation coefficients for R_1 and R_2 are 0.928 and 0.602, respectively. Points within R_1 have better linearity. This is obvious because the neighborhood of **P** within R_2 includes some inflection points.

In the problem of 2-D polygon construction, we need to find the maximal neighborhood for each segment, in which a line segment can accurately fit the points. Using this idea of correlation coefficient, we can determine the neighborhood radius adaptively.

6.3.2 Initial Point Determination

The initial point (IP) is a reference point to start the construction of the first segment of the 2-D polygon from planar data points. Because the points are unorganized and error-filled, IP selection is very important. Liu *et al.* (2002) proposed a random search method in which a point (starting point) is randomly selected from the data points and the points within a certain neighborhood of this point are identified. The center of these points is then calculated and the point closest to the center is selected as the IP. The problem with this method is that if the randomly selected point is very close to an inflection point, the IP, subsequently identified by using a fixed neighborhood radius, may deviate too much from the original shape. An example is shown in Figure 6.5. If point **Q** is selected as the starting point, the center point **O** of the neighborhood will be far away from the original shape, and the closest point to **O** will also be the worst point to be the IP.

To resolve this problem, it is necessary to make sure that the points within the first neighborhood have good linearity. In our approach, first we randomly select a starting point, and then use a fixed radius to find its neighborhood points. The correlation coefficient (ρ) of this neighborhood is then calculated.
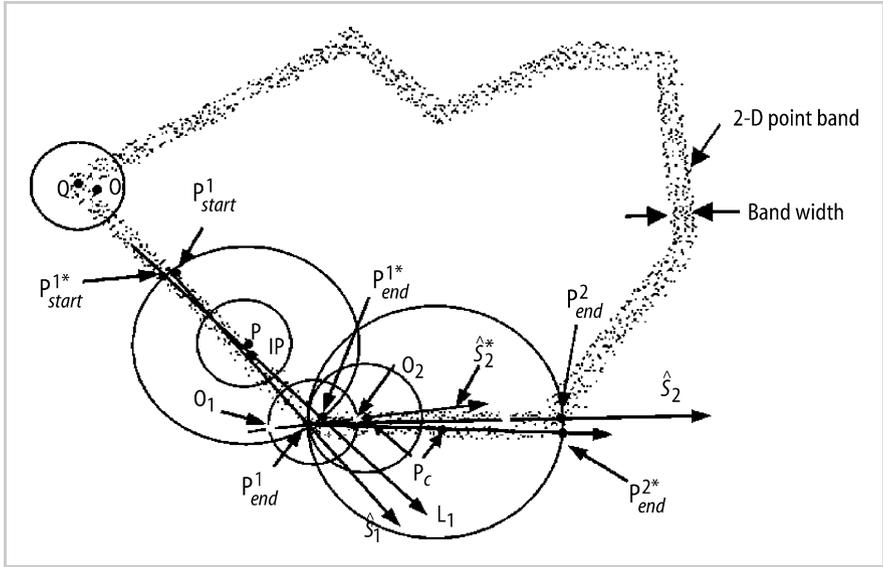
Figure 6.5. IP determination and first and second segment construction. Reprinted from *Computer-Aided Design*, Vol 36, Wu Y, Wong Y, Loh H, Zhang Y, Modelling cloud data using an adaptive slicing approach, pp. 232–239, Copyright (2004) with permission from Elsevier.

If ρ is larger than a preset bound, this neighborhood is used to find the IP, *i.e.*, the point that is nearest to the center of this neighborhood. Otherwise, we will reselect a point and repeat this checking process. For the case in Figure 6.5, point **Q** will be dropped due to its poor linearity, whereas point **P** can be used as the starting point to find the IP. The IP can then be used as a reference point for the first segment construction.

6.3.3 Constructing the First Line Segment (S^1)

After the IP is identified, its neighborhood circle (*N*-circle) for the first line segment, S^1, is obtained so that ρ satisfies the user requirement. At the same time, it is necessary to make the radius *R* as large as possible, so that the resultant polygon has the minimum number of line segments. Hence, *R* needs to be determined adaptively. In our approach, we start with a conservatively small value of *R* and search for the close to optimal *N*-circle radius based on the correlation coefficient. A small ρ means poor linearity, and thus we need to reduce the *N*-circle radius. A large ρ means good linearity, and we can increase the *N*-circle radius. This iterative process is as follows:

Algorithm find_neighborhood_S^1

Given a planar data set **C**, the IP, the initial radius of N-circle R, the increment of radius ΔR, and the predefined lower bound of correlation coefficient ρ_{low} and higher bound $\rho_{high,}$

(a) Select all the points Pi from C, such that $\|\mathbf{P}_i - \mathbf{IP}\| \leq R, \mathbf{P}_i \in C$, to form a data set $\mathbf{C}_1$.

(b) Compute the correlation coefficient ρ of data set $\mathbf{C}_1$ using Equation (6.2).

(c) If $\rho > \rho_{high,}$

    $R = R + 2\Delta R$, go to step (a).

Else, if $\rho < \rho_{low,}$

    $R = R - \Delta R$, go to step (a).

Else

    Return R and points from $\mathbf{C}_1$, stop.

END IF

The points within the N-circle have good linearity, so, we can construct a straight line segment that fits these points locally. Here, we use a least-squares method to compute a regression line, which passes the $\mathbf{IP}(x_{IP}, y_{IP})$ and best fits the points within the N-circle. Let $\mathbf{C}_1 = \{\mathbf{P}_i (x_i, y_i) | i = 1, \ldots, N\}$ be the points within the N-circle. Then, a straight line, $\mathbf{L}_1$: $y = a(x - x_{IP}) + y_{IP}$, can be computed by minimizing a quadratic function:

$$\varepsilon = \sum_{i=1}^{N} \left(a\,(x_i - x_{IP}) + y_{IP} - y_i \right)^2 \tag{6.3}$$

As shown in Figure 6.5, line $\mathbf{L}_1$ has two intersection points, $\mathbf{P}_{start}^{1*}$ and $\mathbf{P}_{end}^{1*}$, with the N-circle (centered at the IP with radius R). In theory, $\mathbf{P}_{start}^{1*}$ and $\mathbf{P}_{end}^{1*}$ can be considered the starting and ending points of the first segment. However, they may not be among the points within the N-circle. Thus, we select two points, which are the closest to $\mathbf{P}_{start}^{1*}$ and $\mathbf{P}_{end}^{1*}$, respectively, within the N-circle, as the starting and ending points, *i.e.*, the closest to $\mathbf{P}_{start}^{1*}$ as $\mathbf{P}_{start}^{1}$, and the closest to $\mathbf{P}_{end}^{1*}$ as $\mathbf{P}_{end}^{1}$. Therefore, $\mathbf{S}^1$ is obtained. $\mathbf{P}_{start}^{1}$ $\mathbf{P}_{end}^{1}$ also define the unit oriented vector of this neighborhood ($\hat{s}_1 = \mathbf{P}_{end}^{1} - \mathbf{P}_{start}^{1}$)/ $\|\mathbf{P}_{end}^{1} - \mathbf{P}_{start}^{1}\|$). Using $\mathbf{P}_{start}^{1}$ $\mathbf{P}_{end}^{1}$ as the diameter, a new N-circle is obtained. Then, we delete all other points within this circle. The remaining planar data set **C** is also updated.

In the aforementioned procedure for constructing the first line segment, the initial R plays an important role. This can be illustrated by the example shown in Figure 6.6 in which the cloud data represents two linear segments. Starting from the IP, if the initial R is too small, *e.g.*, R_1, only a few points are included for the first round, which gives a poor correlation coefficient. This leads to the reduction
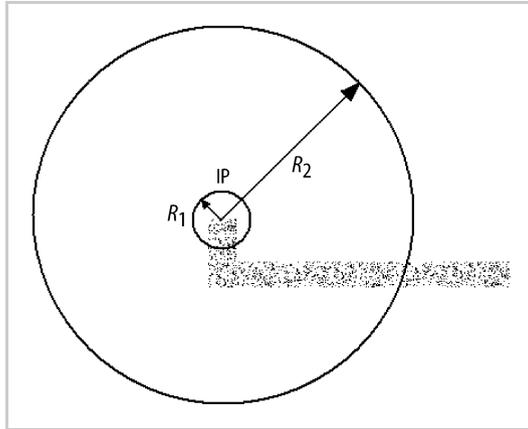
Figure 6.6. Possible problems with the selection of the initial R. Reprinted from *Computer-Aided Design*, Vol 36, Wu Y, Wong Y, Loh H, Zhang Y, Modelling cloud data using an adaptive slicing approach, pp. 232–239, Copyright (2004) with permission from Elsevier.

of R, and the iteration ends with an even smaller R (with two to three points inside). This is certainly not desired. On the other hand, if the initial R is too large, *e.g.*, R_2, we may have a satisfactory correlation coefficient at the first try, but this may lead to the loss of the fine corner feature. In our algorithm, we select the initial R so that there are 30–50 data points in the region selected. This generally produces satisfactory results. However, this number also depends on the scanning resolution. In our application, we use a laser scanner with a resolution of 0.001 mm. Moreover, if there are fine features on the scanned part, it is assumed that a fine resolution is used, so that there are sufficient data points representing these fine features.

6.3.4 Constructing the Remaining Line Segments (S^i)

The method for constructing the remaining segments is slightly different from that of the first segment. We begin with P_{end}^1 as the start point for the second segment, *i.e.*, P_{start}^2. Then, we adaptively determine the N-circle for S^2. Because the same algorithm is used for constructing the remaining segments, we denote the starting point as P_{start}^i ($i \geq 2$). The algorithm to find the radius of the N-circle of S^i is as follows:

Algorithm find_neighborhood_S^i
Given a planar data set **C**, P_{start}^i, initial radius of N-circle R, increment of radius Δ, and the predefined lower bound ρ_{low} and higher bound ρ_{high},

(a) Construct an N-circle (radius R) that is centered at $\mathbf{P}^i_{start}$. Select all points $\mathbf{P}_k$ from $\mathbf{C}$, so that $\left\| \mathbf{P}_k - \mathbf{P}^i_{start} \right\| \leq R$, to form a data set $\mathbf{C}_i$ ($k = 1, 2, \ldots, n$). Compute ρ of data set $\mathbf{C}_i$.

> If $\rho < \rho_{low,}$
>> $R = R - \Delta R$, go to step (a).
>
> Else

Use the least-squares method (Section 6.3.3) to compute a regression line that passes through $\mathbf{P}^i_{start}$. This line has two intersection points, $\mathbf{O}_1$ and $\mathbf{O}_2$, with the N-circle.

Let $\mathbf{P}_{ave} = \sum_{k=1}^{n} \mathbf{P}_k / n$.

If $\left\| \mathbf{P}_{ave} - \mathbf{O}_1 \right\| > \left\| \mathbf{P}_{ave} - \mathbf{O}_2 \right\|$, $\hat{s}^*_i = (\mathbf{O}_2 - \mathbf{O}_1) / \left\| \mathbf{O}_2 - \mathbf{O}_1 \right\|$;

otherwise; $\hat{s}^*_i = (\mathbf{O}_1 - \mathbf{O}_2) / \left\| \mathbf{O}_2 - \mathbf{O}_1 \right\|$.

> End IF

(b) Construct an N-circle that is centered at $\mathbf{P}_c$ ($\mathbf{P}_c = \mathbf{P}^i_{start} + R\,\hat{s}^*_i$) and has a radius R. Select all the points $\mathbf{P}_k$ from $\mathbf{C}$, such that $\left\| \mathbf{P}_k - \mathbf{P}_c \right\| \leq R$, to form a data set $\mathbf{C}_i$. Compute ρ of data set $\mathbf{C}_i$.

(c) If $\rho > \rho_{high,}$
> $R = R + 2\Delta R$, go to step (d).
>
> Else if $\rho < \rho_{low,}$
> $R = R - \Delta R$, go to step (d).
>
> Else
>> Return $\mathbf{P}^i_{start}$ and $\mathbf{P}^{i*}_{end}$, and all the points from $\mathbf{C}_i$, stop.
>
> End if

(d) Use the least-squares method (Section 6.3.3) to compute a regression line that passes through $\mathbf{P}^i_{start}$. This line has two intersection points, $\mathbf{P}^i_{start}$ and $\mathbf{P}^{i*}_{end}$, with the N-circle. Set $\hat{s}^*_i = (\mathbf{P}^{i*}_{end} - \mathbf{P}^i_{start}) / \left\| \mathbf{P}^{i*}_{end} - \mathbf{P}^i_{start} \right\|$. Go to step (b).

Because we do not have any prior knowledge about the neighborhood of $\mathbf{S}^i$, i.e., the unit oriented vector $\hat{s}_i$, we need to find a reasonable estimate to start the iterative process. This is achieved in step (a) of Algorithm find neighborhood $\mathbf{S}^i$. We start by choosing a small R to create an N-circle (centered at $\mathbf{P}^i_{start}$, as shown in Figure 6.5), so that the points within this circle have good linearity. Then, we compute a regression line that passes through $\mathbf{P}^i_{start}$, which helps determine a good estimate of $\hat{s}^*_i$. From step (b), we start with an N-circle (centered at $\mathbf{P}_c = \mathbf{P}^i_{start} + R\,\hat{s}^*_i$) and adaptively find the maximal allowable neighborhood

radius. The example shown in Figure 6.5 illustrates this process for the construction of S^2. $\mathbf{P}_{end}^{2*}$ is obtained from the final N-circle of S^2. Then, the closest point to $\mathbf{P}_{end}^{2*}$, within this N-circle, is found as $\mathbf{P}_{end}^2$. The other point worth mentioning in the above procedure is that in each round, a regression procedure is executed. This may cause long computation time. A trade-off solution is to use the $\hat{s}_i^*$ obtained from step (a) throughout the remaining iterative process, so that the computation becomes more efficient.

The outputs from the above procedure are $\mathbf{P}_{start}^i$ and $\mathbf{P}_{end}^i$, and all the points from $\mathbf{C}_i$. Using $\mathbf{P}_{start}^i$ $\mathbf{P}_{end}^i$ as the diameter, a new N-circle is obtained. Then, we then delete all other points within this circle. The remaining planar data set $\mathbf{C}$ is also updated. The above algorithm is then applied to construct $\mathbf{S}^{i+1}$, until the remaining planar data set $\mathbf{C}$ is null.

6.4 Determination of Adaptive Layer Thickness

Upon completing the 2-D polygon construction for the initial layer, the thickness will be adjusted by using the given shape-error tolerance (ε) as the control parameter. The shape error (σ) of the initial layer is obtained by calculating the distances from all points in the projection plane to the polygon and selecting the maximum distance. If $\sigma < \varepsilon$, the thickness of the initial layer is increased; otherwise, the thickness of the initial layer is reduced. The points within the updated initial layer are then projected to the projection plane. A new polygon is obtained through the curve construction process described in Section 6.3.3. The shape error is then recalculated and compared with ε. Subsequently, a decision is made whether to increase or reduce the thickness of the initial layer. This iterative process is continued until the shape error of the initial layer is slightly less than ε (*e.g.*, between 0.9ε and ε). The construction of the first layer is then completed.

The construction of the subsequent layers is similar to that of the first layer, *i.e.*, (1) creating an initial layer with a preset thickness; (2) projecting the data points within the initial layer to a 2-D plane; (3) constructing a polygon from the data points in the plane; (4) calculating the shape error of the initial layer; and (5) adaptively increasing or reducing the thickness of the initial layer until the shape error is just within ε, *e.g.*, between 0.9ε and ε. In this way, an RP model is generated layer-by-layer adaptively. To implement the aforementioned iterative procedure, a binary search algorithm is developed to find the thickness of a given layer. This algorithm is as follows:

Algorithm find_thickness_layer
Given an initial layer thickness h (a relatively large value) and shape-error tolerance ε,

(a) Set $h_{new} = h$

while (h_{new} < the total height of the data cloud)

{

If $\sigma(h_{new}) < \varepsilon$,

$$h_{new} = h_{new} + h;$$

Else

Return h_{new}, go to (b);

End if

}

(b) $H_{low} = 0$, $H_{high} = h_{new}$

while   ($H_{low} < H_{high}$)

{

$H_{mid} = (H_{low} + H_{high})/2$;

If bandwidth (H_{mid}) > 2ε

$$H_{high} = H_{mid};$$

Else if $\sigma(H_{mid}) > \varepsilon$

$$H_{high} = H_{mid};$$

Else if $0.9\varepsilon < \sigma(H_{mid}) < \varepsilon$

Return H_{mid};

Else

$$H_{low} = H_{mid};$$

End if

}

There are two steps in the above algorithm. In step (a), the search range of the layer thickness, h_{new}, is determined. In step (b), a binary search approach is employed. It can be seen that before $\sigma(H_{mid})$ is checked, bandwidth (H_{mid}) is checked first to decide whether to halve the search range. Because the calculation of σ (H_{mid}) involves a 2-D polygon construction, the process is computationally heavy. This checking step is incorporated to avoid the unnecessary calculation of σ (H_{mid}), and thus to improve efficiency. The bandwidth (H) is defined as follows: Within the projection plane, the projected points of layer H resemble a band, and at any position of the band, there is a bandwidth (see Figure 6.5). The polygon generated to represent the band will be within the boundary of the band. Therefore, if the maximum bandwidth of the band is larger than 2ε, the shape error of the generated polygon will, most likely, be larger than ε. In this case, the search range for the thickness of the layer is halved straightaway.

A simple method is developed to estimate the bandwith. First, we place the 2-D points of layer h into S_h. Then, we obtain another layer between h and $h + \Delta h$ and project its data points onto the same plane. The 2-D points within Δh are placed
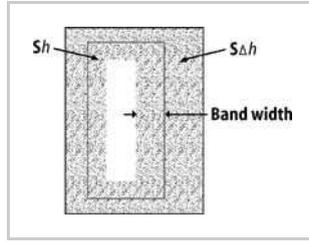
Figure 6.7. Estimation of the bandwidth of the 2-D data points

into $S_{\Delta h}$. An example of these two sets of data points is shown in Figure 6.7. For every point in S_h, P_i ($i = 1, 2, \ldots, n$), we find the corresponding point (Q_i) in $S_{\Delta h}$, so that among all points in $S_{\Delta h}$, P_iQ_i gives the minimum distance (L_i). Then, among all L_i ($i = 1, 2, \ldots, n$), the maximum (L_{max}) is taken as the bandwidth of S_h. Using this estimation method, we find that the points in S_h that give the maximum distances are along the interior boundary of S_h. Errors may occur in the estimation of the bandwidth for some points along the interior boundary (*e.g.*, at corners in the example in Figure 6.7). However, because we are interested only in L_{max}, we find that L_{max} will never overestimate the actual bandwidth. Because we set up 2 ε as the lower band of the bandwidth to determine whether to proceed with the polygon construction, this estimation suits our algorithm very well.

Using this method, our algorithm can handle cases where a very thin layer is needed with relatively good efficiency. For example, when the surface of the part is nearly parallel to the slicing plane, the initial relatively large h will result in diffuse points in the 2-D plane, or points that have a large bandwidth. With this bandwidth checking, the algorithm find_thickness_layer can reduce the search range quickly and proceed to find a satisfactory layer thickness quite efficiently. It is, however, worth mentioning that when the surface of the part is parallel to the slicing plane, this slicing approach will not work properly. In this case, a different slicing direction should be chosen manually.

6.5 Some Application Examples

The algorithm described above was implemented with C/C++ in the OpenGL environment. Three case studies are presented here to illustrate the efficacy of the algorithm for constructing a direct RP model. The first two case studies are based on simulated data sets in which the original cloud data are generated by mathematical equations, so that the theoretical shape errors can be obtained accurately and compared directly. The third is an actual case, and its cloud data were obtained with a laser scanner. The results after processing are input to an SLS machine for fabrication.

In the first case study, a sphere is selected by taking the advantage of its known geometry, so that the shape error of the actual slicing can be compared accurately with the theoretical. Although its geometry is simple, slicing it using a constant layer thickness will produce varying staircase errors. Therefore, it is a good case to test the adaptive slicing algorithm. The equation of a sphere with a radius of 2 is given as follows (random error is incorporated in the equations to simulate noise in the point cloud):

$$\begin{cases} x = 2\cos\beta\cos\alpha + \tau, \\ y = 2\cos\beta\sin\alpha + \tau, \\ z = 2\sin\beta, \end{cases}$$

τ is randomly distributed in $[-0.01, +0.01]$.
$\beta = [-\pi/2, +\pi/2]$.
$\alpha = [0, 2\pi]$.

We use a sampling increment of 0.01 and 0.02 to sample β and α, respectively, to obtain the cloud data for the sphere. There are totally 98,721 points generated. The original cloud data are shown in Figure 6.8a.

For data processing, the initial layer thickness is set at 0.04, and the initial N-circle radius is 0.1. ρ_{low} and ρ_{high} are set at 0.85 and 0.9, respectively. Employing a shape-error tolerance of 0.08, the direct RP model of the sphere shown in Figure 6.8b is obtained. This model contains 11,812 vertices distributed in 74 layers. Figure 6.9a shows the maximum shape error of each layer in the RP model. It can be seen that the maximum shape errors of all the layers are very close to 0.08. The sphere with a radius of 2 is then sliced into 74 layers according to the layer thickness in the generated model. The theoretical shape errors of all corresponding layers are shown in Figure 6.9b. It can be seen that the theoretical errors are close to 0.08 too (except for those close to the two pole areas).

The second case study uses an object composed of four spherical patches (see Figure 6.10a). Compared with the first case, this case poses a more difficult problem for adaptive slicing, which is caused by the combination of the four spherical patches of different diameters and at different positions in space. The parameter of the larger sphere is the same as that in the first case study, and the equations of the three smaller half-spheres are based on the following basic form:

$$\begin{cases} x = \cos\beta\cos\alpha + \tau, \\ y = \cos\beta\sin\alpha + \tau, \\ z = \sin\beta, \end{cases}$$

τ is randomly distributed in $[-0.001, +0.001]$.
$\beta = [0, +\pi/2]$.
$\alpha = [0, 2\pi]$.

The three half-spheres are then formed by transforming the basic form as follows:

(1) Translate the basic form along the z-axis by 1.732 to obtain the first half-sphere.
(2) Rotate the basic form clockwise by 60° around the y-axis, and translate it along the z-axis by 1.732 to form the second half-sphere.
(3) Rotate the basic form counterclockwise by 120° around the y-axis, and translate it along the z-axis by 1.732 to form the third half-sphere.
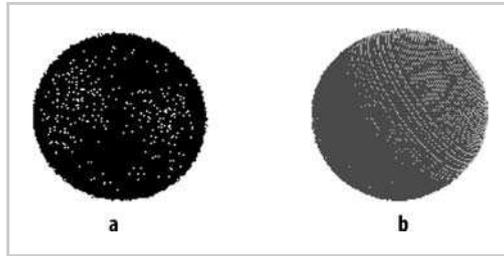
Figure 6.8. The original cloud data and the RP model in case study one. (**a**) The original cloud data. (**b**) The RP model. Reprinted from *Computer-Aided Design*, Vol 36, Wu Y, Wong Y, Loh H, Zhang Y., Modelling cloud data using an adaptive slicing approach, pp. 232–239, Copyright (2004) with permission from Elsevier. A color reproduction of this figure can be seen in the Color Section (pages 219–230).
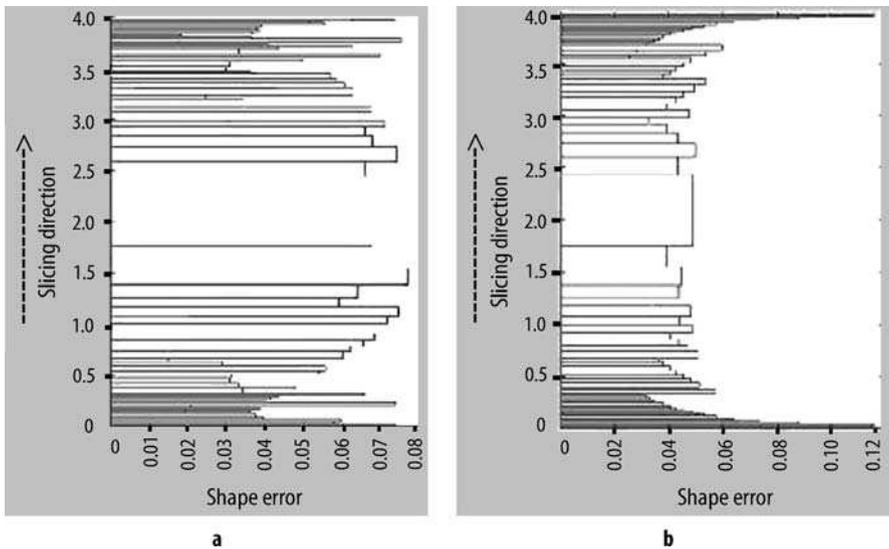
Figure 6.9. Shape-error comparison in case study one. (**a**) Maximum shape error in each layer. (**b**) Theoretical maximum shape error. Reprinted from *Computer-Aided Design*, Vol 36, Wu Y, Wong Y, Loh H, Zhang Y., Modelling cloud data using an adaptive slicing approach, pp. 232–239, Copyright (2004) with permission from Elsevier.

We use sampling increments of 0.02 and 0.05 to sample β and α, respectively, to obtain the cloud data for the object. There are totally 46,057 points, which are shown in Figure 6.10a.

The initial layer thickness is set at 0.04, and the initial N-circle radius is 0.1. ρ_{low} and ρ_{high} are set at 0.85 and 0.9, respectively. Employing a shape-error tolerance of 0.06, the direct RP model of the sphere shown in Figure 6.8b is obtained. This model contains 21,306 vertices distributed in 88 layers. Figure 6.11a shows the maximum shape errors in each layer. It can be seen that the shape errors of all layers are very close to 0.06. Figure 6.11b shows the theoretical maximum
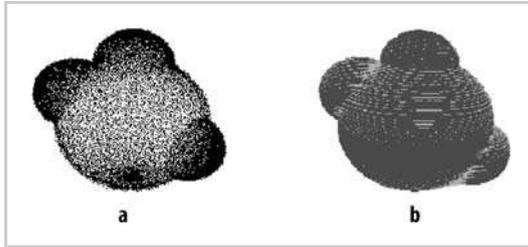
Figure 6.10. The original cloud data and the RP model in case study two. (**a**) The original cloud data. (**b**) The RP model. Reprinted from *Computer-Aided Design*, Vol 36, Wu Y, Wong Y, Loh H, Zhang Y., Modelling cloud data using an adaptive slicing approach, pp. 232–239, Copyright (2004) with permission from Elsevier.

shape error of each layer when directly slicing the object surface using the same corresponding layer thickness as in the RP model. Most of the shape errors in each layer (theoretical) are close to 0.06, although it is not as uniform as in the one in the case study. This may be due to the complexity of the object. On the other hand, in both case studies one and two, the theoretical shape errors in the two pole areas are much larger than the given shape-error tolerance because of the zero radius at the poles. Similarly, pole problems exist for the three smaller half-spheres, which can be seen from Figure 6.11b.
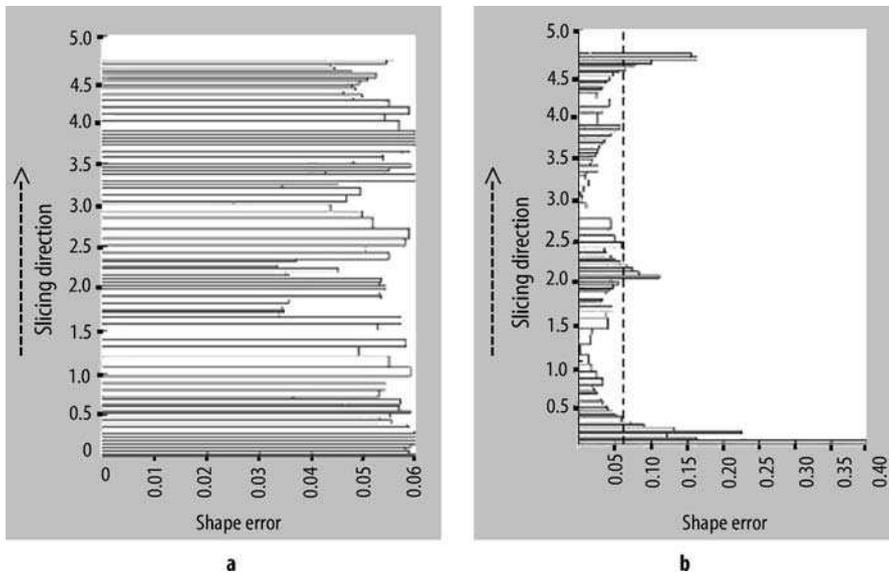
Figure 6.11. Shape-error comparison in case study two. (**a**) Maximum shape error in each layer. (**b**) Theoretical maximum shape error. Reprinted from *Computer-Aided Design*, Vol 36, Wu Y, Wong Y, Loh H, Zhang Y., Modelling cloud data using an adaptive slicing approach, pp. 232–239, Copyright (2004) with permission from Elsevier.
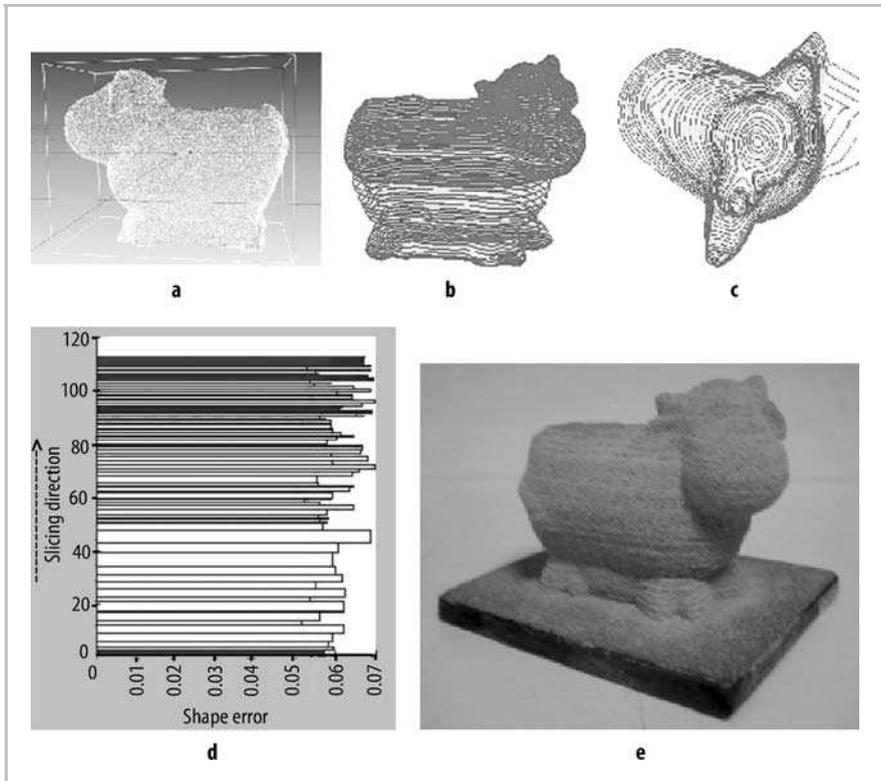
Figure 6.12. Case study three: (**a**) The original cloud data; (**b**) the generated RP model; (**c**) a zoom-in view at the head area on the RP model; (**d**) shape errors at each layer in the RP model; (**e**) the toy cow fabricated by using a SLS machine. Reprinted from Computer-Aided Design, Vol 36, Wu Y, Wong Y, Loh H, Zhang Y., Modelling cloud data using an adaptive slicing approach, pp. 232–239, Copyright (2004) with permission from Elsevier.

The object used in the third case study is a toy cow shown in Figure 6.12a. It has a very complex geometry, in particular, around the head. The original object occupies a volume of $150 \times 120 \times 90$ mm and was digitized with a laser scanner, Minolta VIVID-900. The data sets were obtained from different view angles and then merged to produce a cloud data set of 1,098,753 points. The adaptive slicing algorithm was then applied to the cloud data employing an error tolerance of 0.7 mm, initial layer thickness of 0.2 mm, and an initial N-circle radius of 0.2 mm. ρ_{low} and ρ_{high} were set at 0.85 and 0.9, respectively. This resulted in the direct RP model shown in Figure 6.12b with 115 layers and 59,686 points. In the model construction process, the head area (ears and horns) has a very complex shape and poses a multiple-loop problem. This can be seen clearly in Figure 6.12c, in which the multiple loops are separated successfully and the corresponding layers are generated. The shape error of each layer in the generated

model is shown in Figure 6.12d,which clearly shows that the shape errors are within 0.7 mm.

It took about 30 minutes for the adaptive slicing algorithm to generate the RP model using a PC of 1.5 GHz (Pentium III) CPU. The direct RP model was then converted to a layer-based RP slice-data file in CLI format and fed to a SLS machine. For this RP machine, a uniform thickness of layers is required, and hence the direct RP model of 115 layers was further sliced into 535 layers, with a layer thickness of 0.2 mm (the thinnest layer in the model). It took about 6 hours to complete the fabrication. Figure 6.12e shows the workpiece fabricated by the RP machine based on the direct RP model. However, if a RP machine that can deposit material with changing layer thickness could be used, the time saving could be as much as 78.5%.

6.6 Conclusions

In this chapter, the relationship between data processing in reverse engineering and rapid prototyping is discussed. To fabricate a part using RP processes directly from a representation in the form of point cloud data, the necessity of integrating data processing in RE and RP is comprehensively established. Following that, a method for the direct generation of RP models from arbitrarily scattered cloud data is presented. The modeling process consists of several steps: (1) The cloud data set is segmented into several layers in the RP building direction; (2) points within each layer are treated as planar data, and a polygon is constructed to best-fit the points; (3) the thickness of each layer is determined adaptively so that the surface error is kept just within a given error bound. The algorithm has been implemented. Test results on both simulated and real cases show that the algorithm is effective. The characteristics of the developed algorithm can be summarized in two aspects:

(1) The polygonal curve construction algorithm is adaptive. It can automatically find a feasible starting point and identify the maximally allowable neighborhood for each segment. It can also deal effectively with segments that have multiple-loop profiles.
(2) The thickness of each layer is determined adaptively, based on a given surface tolerance. This provides an intuitive control parameter to users, and the resultant model needs close to minimum RP building time.

Acknowledgments

Chapter 7

Reverse Engineering in the Automotive Industry

Ping Fu, Geomagic, USA

7.1 Introduction

Ford Motor Company celebrated its 100-year anniversary in 2003. Looking back over the decades, automakers have seen five influential trends in automotive design:

- the invention of the horseless carriage itself, a reengineering of a proven design;
- the speed and productivity of Henry Ford's mass production system, which was reengineered from the Chicago meatpacking industry;
- an attempt by Alfred Sloan, president of General Motors, to appeal to individual customers' desires by offering a greater variety of product models within the limits of mass production;
- the use of a wide variety of quality assurance methods to produce standard and exchangeable parts to reduce manufacturing costs; and
- attempting to fulfill the customer's desire to have it all–the economy of mass production, design options to fit an individual's taste and needs, and the highest possible quality.

It turns out that reverse engineering methods are as much a key to this fifth phase as the assembly line was to mass production. The modern automotive industry uses reverse engineering for these key reasons:

1. to create free-form shapes that are difficult to model in CAD software;
2. to overcome obstacles in data exchange and data integrity;
3. to create complex geometries that might not have a CAD model;
4. to resolve and correct problems arising from discrepancies between the CAD master model and the actual tooling or as-built part;

5. to speed up innovation in such areas as ergonomic design, retro-inspired design, combustion design, and aerodynamics, where CAD modeling often adds an unnecessary step; and

6. to ensure quality and performance through computer-aided inspection and engineering analysis.

In this chapter, we will look at several examples of reverse engineering in the automotive industry.

7.2 Reverse Engineering–Workflow for Automotive Body Design

A new car model typically starts with six full-scale designs. Sophisticated methods are used to extract information from consumer surveys and focus groups and to analyze the design model with mathematical simulations and production constraints. Based on customer feedback and manufacturing constraints, two models will be redesigned. In the end, one makes its way through production to showrooms.

Using traditional computer-aided design, engineering, and manufacturing processes, a typical full-scale car design takes more than 3 months to complete. This is not acceptable to Japanese auto companies, which are working to shorten the process to 3 days. The only way this can be done is through a drastic change in the design flow, such as the reverse engineering process illustrated below.

The 3-day process starts with the designer creating a one-quarter-scale clay model of the car body. The small-scale model is measured using a 3-D optical
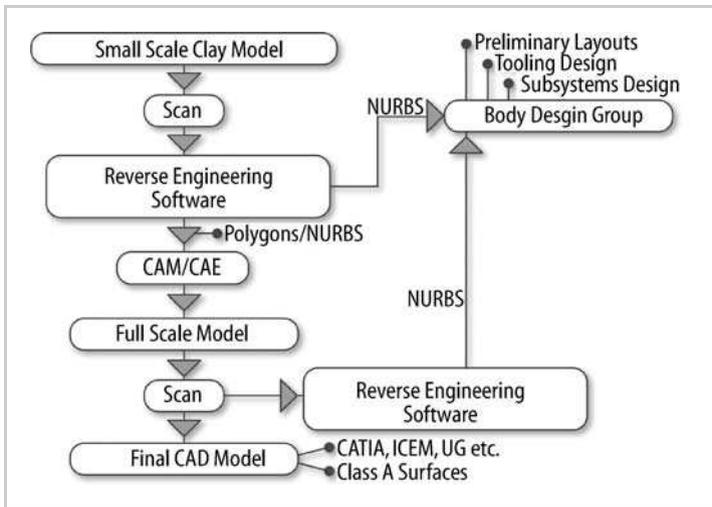
Figure 7.1. Reverse engineering workflow used by Japanese automotive companies

noncontact scanner, which outputs multiple pieces of dense 3-D point cloud data. The point cloud data are aligned and merged into a single point cloud that is dense enough to represent the shape. Reverse engineering software is used to construct the surface in a polygonal mesh for output to engineering analysis software or a rapid prototyping machine, and as a nonuniform rational B-spline (NURBS) surface for output to computer-aided design and manufacturing software.

The resulting digital model is sent to the surface design group for refinement and to the manufacturing group for initial calculation of manufacturing constraints. The digital model that goes through manufacturing is scaled up to full size and milled out in clay or another material. An optical noncontact scanner then measures the full-sized physical model. The dense point cloud from the scanner is used to verify the tolerance between the as-built physical model and the digital model used to mill the part. If it passes, designers proceed to reconstruct the full-size CAD surface using reverse engineering software for surface design and traditional CAD software for functional design. If it fails, the clay model is modified, and the scan, remodel, and compare steps are carried out again until the physical model is within tolerance and captures the design intent.

The final CAD model created by the engineering group and the final surface model created by the body design group are merged as the "golden" model–the digital design master for the remaining production.

The new reverse engineering process creates a closed digital loop between physical and digital worlds and among design, engineering, and manufacturing departments. Working with the same digital model and being able to compare it to an as-built product enables dramatically faster design iterations. This allows automakers to offer almost endless product variety and options at the same cost or to use design innovation to differentiate their products at a premium price point.

7.3 Inside GM's Virtual NASCAR Engine Block

They call it stock car racing, but it's anything but stock. By the time a NASCAR racing car hits the track, only the outline of the car's body conforms to the corresponding model on the showroom floor.

Some of the biggest changes are reserved for the engine block. For teams such as Richard Childress Racing–home of Nextel Cup drivers, Kevin Harvick, Jeff Burton, and Clint Bowyer–the challenge is to take the basic SB2 General Motors engine and squeeze out the most horsepower and torque possible within NASCAR rules.

The original SB2 design dates back 50 years, and until recently, there was no 3-D digital CAD model of the engine block. GM wanted an accurate digital model of the engine block for finite element analysis (FEA) and computational fluid dynamics (CFD) studies and to create a complete digital engine assembly in the future.
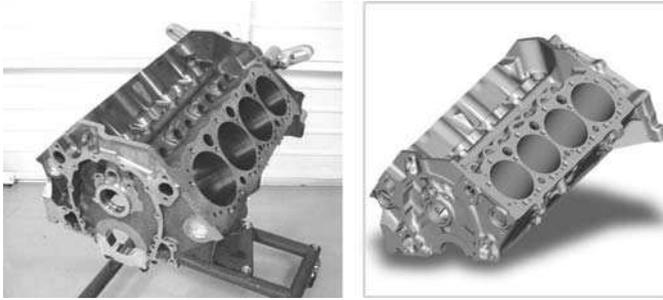
Figure 7.2. The General Motors SB2 racing car engine block and its digital duplication. Copyright © Richard Childress Racing Enterprises, Inc., Welcome, NC, USA, www.rcrracing.com. Reproduced with permission.

Richard Childress Racing (RCR), well known for its innovative use of reverse engineering technology, won the bid to create the first 3-D digital model of the SB2 engine block. The team has extensive experience using scanners and reverse engineering software to fine-tune cylinder head ports for its cars. The process enables RCR to generate highly accurate ports in about one-sixth the time it took to do the work by hand.

The GM project embodied five of the six main reasons that automotive engineers use reverse engineering:

1. to create free-form shapes that are difficult to model in CAD software;
2. to overcome obstacles in data exchange and data integrity;
3. to create complex geometries that might not have a CAD model;
4. to resolve and correct problems arising from discrepancies between the CAD master model and the actual tooling or as-built part; and
5. to ensure quality and performance through computer-aided inspection and engineering analysis.

The typical starting point for a reverse-engineering project without 3-D CAD data would be the 2-D drawings. In this case, however, the 2-D prints did not contain the design GM wanted to re-create. When 3-D molds and cores were constructed in the 1950s and later years, changes were made and they were not detailed in the 2-D prints. RCR wanted to capture these details as they appear in the as-built part. Complicating matters further, the cast-feature shapes are difficult to reproduce using direct CAD geometry, even after precise measurement. Radius rounds from an actual casting, for example, might be spline-shaped, not spherical with a single radius as depicted in CAD models. "We wanted a model of what we have, rather than an idealization of what we think we have", said Clifton Kiziah, the RCR engineer who managed the project. "The 3-D model needed to match the physical block as accurately as possible."

Accuracy for GM requires casting tolerances of ± 0.030 of an inch. Anything less might compromise the credibility of FEA and CFD test results. Meeting the

project's goals required a reverse engineering system that would capture data from an actual SB2 engine block and turn it into a highly accurate 3-D CAD model. The software needed to bridge the gap between the physical and digital worlds.

The project was particularly challenging for the following reasons:

1. The cylinder block of a racing car engine contains very complex topology.
2. Interior cavities cannot be captured by a range-image scanner (also called a line-of-sight scanner).
3. GM wanted a true CAD model in which parts of the assembly could move during simulation.

RCR was ready for the challenge. To capture data in cavities, urethane was poured into the actual sand core molds used in the block-casting process. A structured-light scanner was used to capture the cores and engine block as eight separate point cloud files. The engine block was a 20-MB file and the cores were about 15 MB each.

The point cloud files were brought into reverse engineering software and processed to reduce noise and delete outliers. The point clouds were already in the same coordinate space, so no alignment was needed. The software then automatically created a triangulated surface from the point cloud, and holes were filled to create a watertight mesh surface. Because of the need to produce CAD models with moving parts, the mesh surfaces were parameterized to create a nonuniform rational B-spline (NURBS) surface, a mathematical representation used by most CAD software. The reverse engineering software computed the data for aligning the water jacket surface with cores to create cavities, making the model ready to be used by CAD software.

The eight surface files–complete with datum planes, axes, and references– were imported individually into CAD software, and the cores were assembled as they would be to cast an actual water jacket in an engine block. A cut-out feature was used to subtract cores from the block.

The engine was reverse engineered to achieve three times the accuracy required by casting, tightening the tolerance to 0.010 of an inch, and in some cases

Figure 7.3. Reverse-engineered model of urethane molds of the internal cavities from the SB2 engine. Copyright © Richard Childress Racing Enterprises, Inc., Welcome, NC, USA, www.rcrracing.com. Reproduced with permission.

Figure 7.4. The completed digital model of the SB2 engine block with rotating assembly. Copyright © Richard Childress Racing Enterprises, Inc., Welcome, NC, USA, www.rcrracing.com. Reproduced with permission. A color reproduction of this figure can be seen in the Color Section (pages 219–230).

to within 0.005. "It would probably have taken several months if we had to model it directly, and I'm not sure we would have been able to capture the complexity of the actual cast surfaces", Kiziah said. "The finished model looks and measures the same as the real block."

GM and its racing teams now have an accurate digital model of the SB2 engine for CFD tests to optimize cooling, and secondary machining simulations to check for clearances and fit of new parts. FEA simulations will be used to determine where material can be removed in secondary machining without affecting the strength of the block.

Reverse engineering of the SB2 engine block could be just the start in RCR's quest to digitize parts that do not have 3-D CAD models. The team is working on a digital model of a total car assembly, complete with surface models of the chassis and the entire engine.

7.4 Ferrari Speed Not Confined to Race Track

When the Ferrari 550 Maranello crossed the finish line in front of all other GTS-class cars at the 24 Hours of Le Mans in June 2003, it was not only the car that showed unprecedented speed.

The Prodrive-run team won Le Mans in only its second attempt and had just joined forces with Veloqx 5 months before the race. In the course of its preparation, the Veloqx Prodrive team embraced new technologies that enabled an alternative rear-wing design in just 6 weeks. The task of studying and modifying the rear wing was assigned to Advantage CFD, a company that specializes in services to motor sport customers. Because airflow is severely affected by the rest of the car body upstream, the first challenge was to obtain accurate CAD data for the entire car body and the job had to be done by companies on separate continents. Prodrive assembled the required car panels at its Banbury, U.K., headquarters. 3-D Scanners, based in Coventry, England, scanned one side of the car using a handheld, noncontact scanner with an articulated arm. Everything was scanned

Figure 7.5. Advanced CFD Model X with a Faro arm is used to capture half of the Ferrari 550 Maranello body. Copyright © Advantage CFD, Brackley, Northamptonshire, UK, www.advantage-cfd.co.uk. Reproduced with permission.

except for the front bumper, the only part for which Prodrive had a CAD model. The point cloud of the half car body contained 40 million points.

The eight scan files were read into reverse engineering software, where they were merged and triangulated into a watertight polygonal mesh. A mirroring function created the other half of the car body that was not scanned, and different
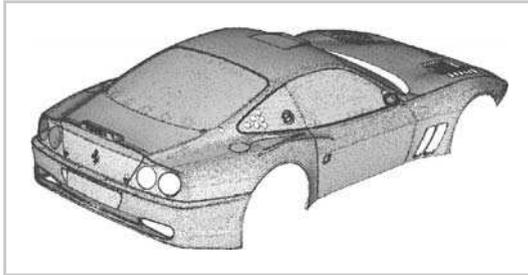
Figure 7.6. Aligned and merged point clouds of Ferrari 550
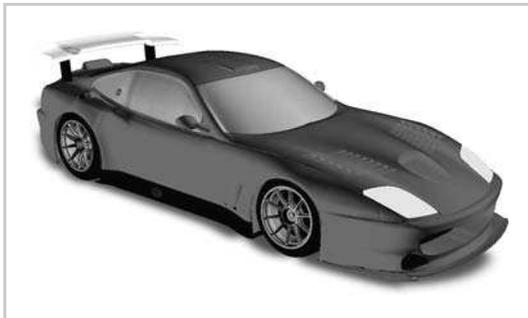
Figure 7.7. Ferrari 550 Maranello body reverse engineered in Geomagic Studio software and assembled and visualized in Pro/E software

sections of the body were aligned and merged together. The software then automatically parameterized the polygonal mesh, computed the nonuniform rational B-spline (NURBS) surface, and output the surface in IGES file format. The IGES files were imported into a CAD assembly along with the front bumper data.

After assembling a complete CAD model of the car's external geometry, engineers defined a surrounding fluid volume and created a hexahedral mesh extruded from the majority of the body surfaces. A structured hexahedral mesh was also used to resolve the critical rear-wing flow, the flow under the flat floor, and the wake behind the car. The remaining volume was filled with tetrahedral cells. The 3-D model was tested initially without a wing to assess flow curvatures around the rear of the car and through the volume where the rear wing is positioned. Based on these findings, a 2-D analysis was done on a wide range of wing angles. The analysis showed that the center of pressure for the new wing section should be moved further back.
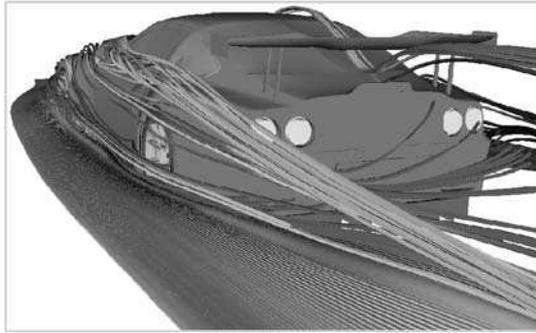
Figure 7.8. Visualized results from CFD analysis generated by CEI's EnSight software. Copyright © Advantage CFD, Brackley, Northamptonshire, UK, www.advantage-cfd.co.uk. Reproduced with permission. A color reproduction of this figure can be seen in the Color Section (pages 219–230).
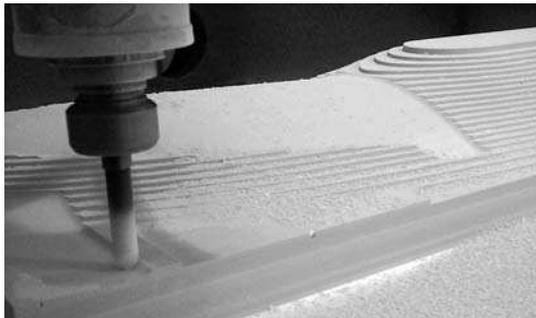
Figure 7.9. Manufacturing the rear wing. Copyright © Advantage CFD, Brackley, Northamptonshire, UK, www.advantage-cfd.co.uk. Reproduced with permission.

Several new wing designs were evaluated on the 3-D model, simulating performance at the same speed as used on track tests. Visualizations of the first few design iterations showed regions of flow separation caused by the original wing mounts. The wing mounting system was redesigned to address the issue, and further CFD tests were run.

A comparison of oil flow on the rear wings clearly indicated flow separation on the mounts of the original wing and significant improvement in this area with the new wing design. The new design had a small separation at the trailing edge, but that was removed when a 6-mm gurney was attached to the wing. The final CAD model was loaded into CNC machines to produce tooling blocks needed to manufacture the new wing.

A new rear wing was made in time for track tests 6 weeks after the project had begun. The track tests and further running at the Le Mans prequalifying event confirmed that the new wing reduced drag by 2.5% for the same level of downforce.

Lap times on Le Mans test day confirmed what the race itself would later prove: The Veloqx Prodrive Ferraris were easily the quickest cars in the GTS class–with both the new and old wings. Because of time constraints, the new wing could not be tested for durability, a key factor in the 24-hour race. Not wanting to risk going into the race without durability testing and knowing that its cars were still the fastest, Veloqx Prodrive decided to stick with the old wing design. The wisdom of the decision was borne out by the victory of the young racing team.

Buoyed by the success of the Le Mans project and continually looking for a competitive edge, Prodrive developed an additional rear wing for the Ferrari 550 Maranello. The goal this time was to provide greater downforce for the American Le Mans series. A similar design approach was used for the new wing, and once again, it went straight from computational flow analysis to the track for testing. The wing exceeded predictions from CFD analysis when tested on the car and was used for the rest of the season, as Prodrive teams won the GTS class in all three of the American Le Mans series races in which they competed.

7.5 Reverse Engineering for Better Quality

For a leading European automaker, reverse engineering is a major tool in continual engine improvement. The company is using reverse engineering to implement a new computer-aided inspection (CAI) process for a plant that builds nearly 600,000 engines a year.

More than 900 specialists oversee quality assurance on the plant's production lines and operating units. Core engine block components, such as the crankcase, crankshaft, connecting rods, and the cylinder head, are verified by spot checks in the company's precision measurement laboratory, where engine components must conform to tolerances of hundredths of a millimeter.

Faced with these rigorous quality demands, the company wanted to move beyond traditional inspection processes using coordinate measurement machines (CMMs). CMMs collect a sample of discrete points on a part, one at a time. The process is slow and does not adequately address surface-to-surface inspection required to verify the accuracy of sheet metal or free-form surfaces. Results are recorded in a 2-D geometric dimensioning and tolerance (GD&T) report that does not directly correlate with the 3-D CAD model of the part.

The CAI process uses noncontact scanners to collect millions of points in seconds. Software based on reverse engineering principles then processes the information automatically to compare an as-built part to its CAD representation. The process creates an interactive loop among the design, manufacturing, and quality control divisions.

The first step in the CAI process is to capture accurate geometry and dimensions by placing target points, which are used to align multiple scans, on the surface of the existing engine component. The engine part is then scanned with a noncontact white-light scanner that generates a polygonal model.

Jagged edges, holes, and surfaces on the polygonal model are smoothed out, and the model is cleaned to remove extraneous points or noise that might undermine the data. The completed files are then merged, aligned, and saved in STL format. The STL model is imported into CAI software, which automatically allows engineers to align and compare the STL model with the original CAD data
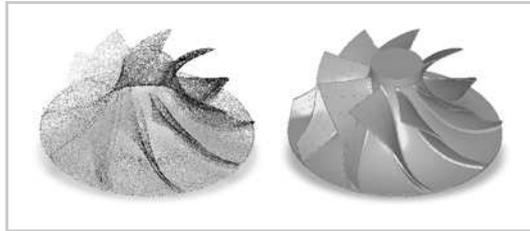
Figure 7.10. A point cloud of an automotive intake turbine scanned from the physical part (left) and the original CAD reference model (right)
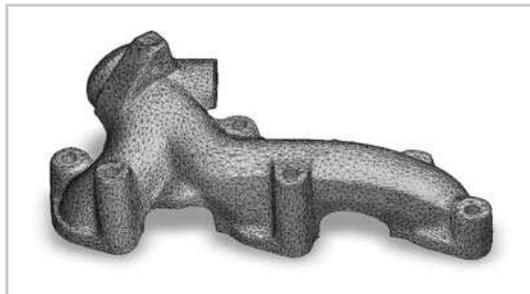
Figure 7.11. A polygonal model of an automotive exhaust manifold reverse engineered from the actual part

to determine exactly where variations in the geometry occur and to analyze how deviations might impact the part's functionality.
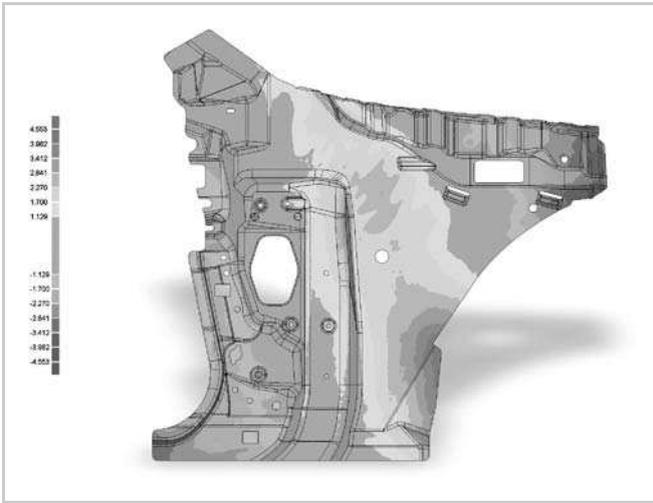
Figure 7.12. Color maps visualize the deviation between the scanned sheet metal part and its corresponding reference CAD model. A color reproduction of this figure can be seen in the Color Section (pages 219–230).
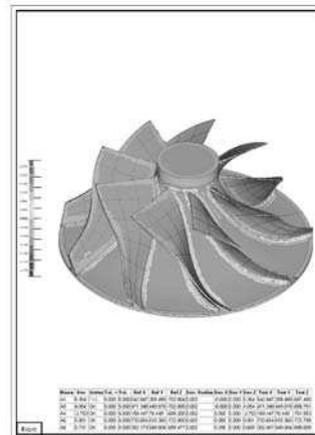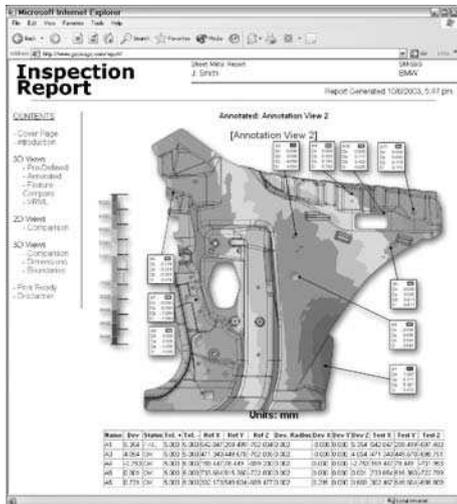
Figure 7.13. An HTML report of a sheet metal part and a PDF report of an intake turbine. A color reproduction of this figure can be seen in the Color Section (pages 219–230).

Even miniscule differences between the physical part and the CAD data can result in performance flaws and inaccurate engineering analysis, so the auto maker imposes a tolerance level of 0.02 to 0.03 millimeters–less than the width of a human hair. If parts differ from the original design by more than that, they are sent back to be reworked. Part discrepancies are communicated through reports with visual images and numerical results that are generated automatically by the CAI software. Reports can be output in many standard formats, including HTML, PDF, Microsoft Word, and Excel, or as customizable graphics. The reports are also leveraged by suppliers, quality engineers, and others in the company's supply chain.

The European automaker is joining its Japanese counterparts in finding that the CAI process based on reverse engineering saves time and increases quality by ensuring dimensional accuracy and engineering performance. According to the manufacturer, design engineers have better information about the quality of parts early in the product life cycle, which saves the company money and produces a better end product for consumers.

7.6   A Look Ahead–Convergence of Digital and Physical Worlds

The term *engineering* generally has a positive meaning. It is used to describe the act of creating something beneficial. But, that is not the case with *reverse engineering* (RE), which has suffered from a negative connotation. Reverse engineering has been associated with the often illegal act of copying an original design for competitive purposes.

In the manufacturing world today, however, the concept of reverse engineering is being legally applied for producing new products or variations of old products. The term *reverse* comes from the concept of bi-directional data exchange between the digital and physical worlds.

The primary thrust in the early development of computer-aided design (CAD), engineering (CAE) and manufacturing (CAM) was to create a product in a computer and bring the results out to the real world. This process became known as *forward engineering.* As envisioned by CAD/CAM companies, a designer sat down with the known specifications of the design object based on its form, fit, and function, and crafted the item entirely within a computer. CAD was supposed to be able to define a simple part or a complex assembly entirely from its dimensional characteristics. CAE components, such as structural or thermal analysis software, would take this digital representation and analyze it. The CAM software would take this same electronic definition and create the paths to cut the tools for part manufacture.

Where this process falls apart is when analytical features cannot define the form of the part. This occurs usually when the forces of nature are involved directly in shaping the design. The dimensions are hard to apply where the shape is controlled by one or more of the following requirements:

- ergonomics–seats, dashboard, helmets
- aesthetics–free-form, car styling
- aerodynamics–car bodies
- flow dynamics–engine, transmission
- time-worn–working parts, stamping dies, aging parts

The auto industry, the largest manufacturing sector, has long been a leader in applying computer methods to streamline design and manufacturing. It also represents one of the most challenging arenas for computer-aided modeling because the geometric requirements are so diverse. Whereas exterior design is driven largely by human emotion and aesthetic concerns, other requirements are driven by ergonomics, functional fit, structural integrity, aerodynamics, combustion processes, and manufacturability. Geometry has many masters in the automotive industry.

Challenges also arise from multiple large-scale suppliers that are dispersed across the globe. Only automotive companies with efficient processes and reliable supply chains can survive in this environment. For more than 30 years, CAD/CAM methods have been a partial solution to these challenges. The elusive goal has been to create a shared computer model for use by all engineering functions and all suppliers, from concept through manufacturing and service. Billions of dollars have been invested toward this goal.

Traditional CAD has had many successes in the auto industry, but there have also been failures. Automotive stylists still prefer the tactile feedback of sculpting in clay. Inspectors struggle to compare golden master parts against solid models. Differences among CAD systems frequently break the continuity of design and manufacturing processes. The list goes on.

What automotive designers are now discovering, much as physicists learned a century ago, is that a single geometric representation lacks the power of dual representations. For physicists, the discovery was that light and energy were best represented as both discrete particles (photons) and as waves (electromagnetic spectrum). Problems that were difficult to solve in one domain proved much easier in another. This need for dual representations has long been known with automotive NVH (noise, vibration, harshness) analysis. Problems that prove difficult to solve in the discrete time domain are often best solved in the frequency domain–and *vice versa*.

Traditional CAD is based on the mathematics of continuous curves and surfaces. This is a powerful technique, as evidenced by today's leading CAD systems, but it has many limitations. The complement to continuous mathematics is *discrete* mathematics–handling geometry as sets of discrete points. Until recently, this was a cumbersome process, inhibited by labor-intensive coordinate measuring and hands-on processing by early reverse-engineering software. The first generation of reverse-engineering algorithms focused on fitting curves over cross sections of points and then falling back to the curve-creation and lofting methods commonly used in CAD and modeling software. Early software

failed to make the process easier and did not break away from the limitation of continuous mathematics.

In the last 10 years, new computational algorithms have been developed to triangulate and parameterize the entire point cloud. The approach is global and automatic and depends far less on user input. Coupled with increased computational power, more computer memory, and high-speed noncontact scanning devices, discrete geometry has gained increasing importance in automotive design, manufacturing, and quality assurance.

The real power in all this is not competition between continuous and discrete mathematics, but convergence of the two. The key is using the two representations where they fit best, and then seamlessly translating back and forth, independent of whether the starting point is a physical part or a digital model.

The world has witnessed three digital convergences during the past three decades. Each time new technologies break down the barrier between physical and digital forms, new products and new markets are created. The 1970s ushered in digitized sound using signal processing (one-dimensional), which made analogy and digital conversion (A to D, D to A conversion) part of a common language in the telecom industry. The 1980s brought digitized fonts and pictures using image processing (two-dimensional). The convenience of switching between electronic and paper documents changed the publishing industry and the way we store and share information.

The third convergence, beginning in the 1990s, centers on digitizing the physical world using geometry processing (three-dimensional). The convergence of physical and digital worlds enabled by reverse and forward engineering technologies should fundamentally change the way products are designed, manufactured, and marketed. If we can create a digital duplicate of our world as easily as taking a digital picture of what we see, the biggest breakthrough of the twenty-first century will be in manufacturing. In the automotive world, that will mean achieving the ultimate goal of fulfilling the customer's desire to have it all.

Acknowledgments

The vendors of the technologies used in these case studies include

- 3-D Scanners–noncontact scanner (Figure 7.5)
- Advanced CFD (Figures 7.8 and 7.9)
- Analytical Methods–2-D analysis software
- Beta CAE Systems–meshing software
- CATIA–CAD software
- CEI – 3-D visualization software
- ESI Group–meshing software
- Faro–portable CMM

- Fluent–CFD software
- GOM–structured-light scanner
- Parametric Technology Corp.–CAD software
- Pointwise–meshing software
- Geomagic–reverse engineering and computer-aided inspection software (Figures 7.1, 7.6, 7.7, and 7.10–7.13)
- Richard Childress Racing (Figures 7.2–7.4)

Chapter 8

Reverse Engineering in the Aerospace Industry

Ping Fu, Geomagic, USA

8.1 Introduction

December 17, 2003 marked the 100th anniversary of the Wright brothers' first successful airplane flight–the first time a heavier-than-air vehicle with an engine successfully propelled itself into flight. Aviation has come a long way since 1903. We've broken the sound barrier; we can fly more than 500 people on a single jumbo jet; and we've landed on the moon. All in less than a century!

The aerospace industry comprises companies producing aircraft, guided missiles, space vehicles, aircraft engines, propulsion units, and related parts. Aircraft repair and modification and aerospace research and development also are included. The combination of advanced production processes, a highly trained workforce, and significant research and development has allowed the aerospace industry to attain major milestones on the technology front.

The biggest technology achievement in the last decade came from the Boeing Commercial Airplane Group. The company's 777 division was recognized in 1995 for the first airplane to be 100% digitally designed and pre-assembled on computer.

In creating the 777, Boeing used fundamentally new approaches to designing and building an airplane. The 777 program established design/build teams to develop each element of the airplane's airframe or system. Under this approach, all of the different specialties involved in airplane development–designers, manufacturing representatives, tooling, engineers, finance, suppliers, customers, and others–worked jointly to create the airplane's parts and systems. Based at the same location, team members worked concurrently, sharing their knowledge rather than applying their skills sequentially.

Communication among the program's 238 design/build teams was enhanced by sophisticated computers linked by the largest mainframe installation of its kind in the world. Using 3-D digital software, designers could see parts as solid images and then simulate the assembly of those parts on the screen, easily correcting misalignments and other fit or interference problems.

As a result of these digital processes, the 777 program exceeded its goal of reducing change, errors and rework by 50%. Parts and systems have fit together better than anticipated and at the highest level of quality. The first 777 was just 0.023 of an inch–about the width of a playing card–away from perfect alignment, compared to a half inch for most airplanes.

"Digital preassembly helped us significantly improve our engineering and manufacturing processes and the overall quality of our product", said Charlie Kyle, Boeing chief project engineer, Airplane Integration. "We've been able to use this technology to lower our costs and decrease the time it traditionally would have taken to introduce this airplane into the marketplace. The 777 is a marvellous achievement of teamwork and technology."

Though standard forms of data representation and computer-aided design tools such as those used in the 777 projects have led to remarkable progress during the last 50 years, the aerospace industry still faces big challenges in using digital tools to manufacture airplanes. Maintaining hard tooling for spare parts costs Boeing $4 billion or more a year. And, if Boeing has this problem, the U.S. Air Force has it ten times over. Even Boeing's digital process for the 777 faces data incompatibility issues. Much of the engineering data is in polyhedra (also called faceted solids), which have limited use in today's CAD environment.

Reverse engineering is a process that can be used by Boeing and other aerospace companies to create digital inventories of spare parts or to convert legacy data into today's CAD environments. It turns out that reverse engineering methods are as much a key to the future of aerospace manufacturing as CAD has been for the all-digital 777. The modern aerospace industry uses reverse engineering for these key reasons:

1. to create legacy parts and hard tooling that do not have CAD models;
2. to overcome obstacles in data exchange and data integrity;
3. to resolve and correct problems arising from discrepancies between the CAD master model and the actual tooling or as-built part; and
4. to ensure quality and performance through computer-aided inspection and engineering analysis.

In this chapter, we will look at several examples of the way reverse engineering is transforming design, engineering, inspection, and manufacturing processes in the aerospace industry.

8.2 RE in Aerospace–A Work in Progress

Reverse engineering faces a bit of a chicken-and-egg dilemma in the aerospace industry. Companies know the value of digitizing legacy parts to reduce physical inventories, maintain digital data on parts that are no longer manufactured in the country of origin, and modify products for which they don't have blueprints. But instituting change in a massive aerospace organization with hundreds of suppliers is not a simple matter.

"A company is lucky if it has digital models of 1% of inventory", said Jeff Brehm, president of Computer-Aided Measurement Surfaces (CAMS), a company that provides data capture services for the aerospace industry. "Companies need a real good excuse to create a digital model of a part–it might be a unique part or a part that will never be made again in this country."

The automotive industry has had some decided advantages over aerospace in implementing reverse engineering: Automotive products are not nearly as large and their assemblies not as complex as those in the aerospace industry. But although these are obstacles, they also serve as compelling reasons why reverse engineering could be more beneficial to aerospace companies.

The other big reason is legacy parts. GM can leave obtaining a transmission for a 1968 Pontiac GTO to hobbyists and collectors, but the 1968 models of aircraft, with designs barely changed, are likely to be more prevalent in the air than models from the twenty-first century. This creates a huge inventory of legacy parts to be stored and maintained.

Most airplanes designed 30 years ago are still flying today, and they were designed without any 3-D CAD models. Keeping aging airplanes flying is one of the most cost-efficient and technically challenging areas for aerospace companies, and one that deserves more investment.
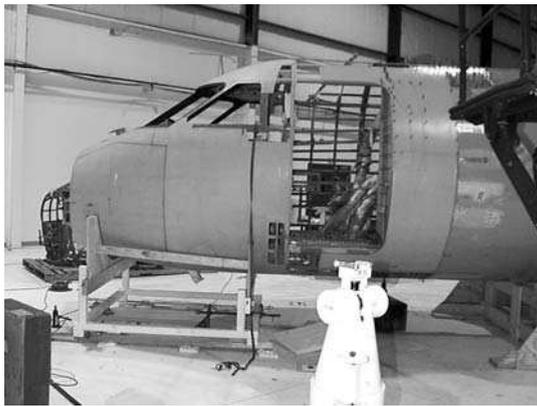
Figure 8.1. A Navy C20G (Gulfstream IV) that suffered nose, wing, and tail damage from a tornado

Figure 8.2. Digital surface of landing gear door generated from data captured by a laser tracker

Most aerospace companies have been doing reverse engineering for about 5 or 6 years. The processes have become faster and better with improvements in scanners, surfacing software, and computers, but they are still performed by specialists, those with what Brehm calls "tribal knowledge".

On the surface, reverse engineering appears to be a simple process: measure, reverse engineer with software, and verify. But, those doing the work need to know what tools to use for what jobs–collecting too much data bogs down the computing process; not enough data requires going back to rescan the physical part.

Different data capture hardware is typically needed according to the job. The more complex the shape geometry, the more points are needed. That's where the tribal knowledge comes in. Those with experience in reverse engineering can look at a part or assembly and know the level of detail that has to be captured. It's a fine balance between capturing enough for accuracy, but not so much to bring processing to a standstill.

"It would be a waste to capture a wing with a structured-light scanner that might generate a model with 40 million points", said Brehm. "On the other hand, you might need 40–100 million points to capture the radii and complex shapes of an actuator."

For master parts that require capturing a large shape, the cuts for the wings and fuselage of a 747, for example, a laser tracker is often used. Laser trackers can capture large surfaces fairly quickly, as long as high detail is not required. Complex parts with tight contours that require lots of data to generate an accurate model are captured by a structured-light or laser scanner. For small, intricate parts, a scanning arm with an integrated scanner head is likely to be used.

The size and type of data also determine what kind of software is used to process it. Traditional surfacing and CAD software cannot handle large point clouds. The new generation of reverse engineering software can handle very large data sets, although for interactive editing and viewing, it is difficult to get real-time graphic response if the model has more than 5 million points or

Figure 8.3. Boeing 747 landing gear door
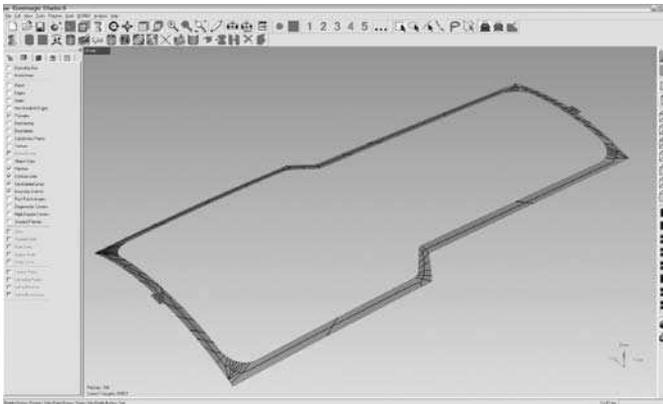
Figure 8.4. Digital surface of a landing gear door generated from data captured by a laser tracker

10 million triangles. For computation only, new reverse engineering software packages such as Geomagic Studio, Polyworks, or Rapidform can handle 50–100 million points.

As might be expected, accuracy requirements vary widely. Hole patterns require accuracy within ± 0.005 of an inch, whereas a mold line requires much less accuracy. For a large 1968 part, all that might be required is to verify that the cuts are good.

During the coming years, aerospace should see more automation from reverse engineering software. Moreover, the combination of reverse engineering and surfacing capability will be able to create a surface that is not evident in the point cloud data but was part of the original design intent.

The obvious justification for the aerospace industry to invest in reverse engineering is the cost of storage space and keeping parts maintained. In the next case study, we will take a look at cost reductions of hard tooling.

8.3 Reducing Costs of Hard Tooling

At the turn of the century, one of Boeing's top goals was to reduce the costs of hard tooling. Many airplanes built 30 years ago are still flying, and hard tooling for spare parts is critical for repairs. Tooling is the standard for manufacturing parts because drawings and digital models either do not exist or are not accurate enough.

Boeing is not alone in dealing with the hard-tooling problem; it is widespread in the aerospace and defense industries. Keeping aging aircraft flying is as critical to aerospace as keeping computers running is to information technology. From an economic point of view, less than 10% of the investment in a new airplane can keep 80% of aging airplanes flying. Reducing hard tooling costs is a sure way to boost profit margins.

Boeing decided it needed a reverse engineering process to digitize physical models rapidly. The principal goals were to reduce costs and ultimately eliminate hard tooling. Manufacturing Research & Development (MR&D) within Boeing's Commercial Aircraft Group was charged with developing a reverse engineering strategy and a standard process to support this initiative.

In 2000, a Boeing 747 wing tip was used to examine the reverse engineering process. The project required digitally duplicating the wing tip, creating a CAD model, and manufacturing the part within one working day. The completed part needed to meet tolerances of 0.003 of an inch.

The wing tip was mounted on a platform for scanning by Boeing's custom-made system, which comprised a line-laser scan head and a mechanical motion controller. Several types of reverse engineering software were used to create the CAD model, and the part was manufactured on site with a process that is consistent with Boeing's new product design. The goal was that once the CAD model was obtained, the manufacturing process would be the same whether it was for a new airplane or for the spare parts of an aging airplane. If this process was relatively easy and cost-effective, there would be a huge potential to reduce the costs of maintaining hard tooling.
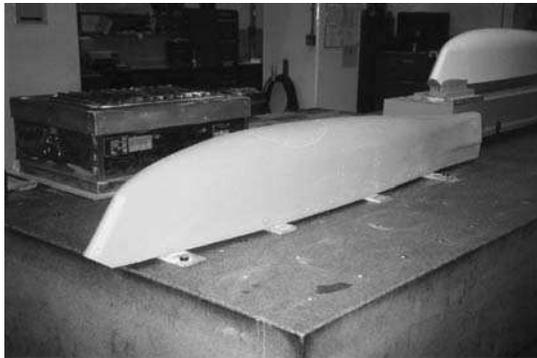
Figure 8.5. Boeing 747 wing tip
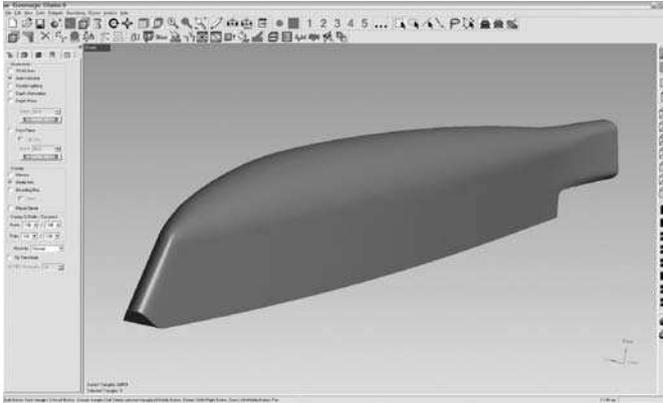
Figure 8.6. Digital duplicate of the wing tip
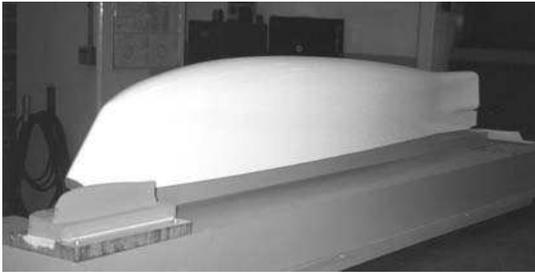
Figure 8.7. Newly manufactured wing tip from the spare part

Using the reverse engineering process, a database of 3-D CAD files for hard tools can be stored in the computer. If an airplane is on ground for repairs, parts can be made using an on-demand manufacturing process.

At the time of Boeing's wing-tip project, software vendors were offering a fraction of the functionality available today. The Boeing team was using prerelease betas of first-generation reverse engineering software. Although the old wing was replicated within tolerance, the process took more than a day. Simply put, the technology was not mature enough for prime time.

Today, reverse engineering technology can reproduce a part like this within one working day. But so far, the new generation of reverse-engineering software has not been used widely to reduce the costs of hard-tooling maintenance. One of the reasons is that the process has not been tested in a real production environment. Despite the huge cost-saving potential, companies are more reluctant to spend money on spare parts then they are to spend on developing new products. Nevertheless, there is real value in creating a digital inventory of hard tooling, especially in an economic climate where profit matters. Digital hard-tooling inventory is one area in which return on investment can be measured and both short- and long-term benefits can be realized.

8.4 Digitizing a NASA Space Vehicle

In the early 1990s, NASA started to develop the Autonom Crew Return Vehicle X-38, an aircraft designed to bring astronauts back to earth safely in an emergency.

3-D digital models were needed to confirm and refine the design for simulation. An accurate match between the actual model and the digital model was important in verifying the accuracy of computational fluid dynamics (CFD) simulations.

Reverse engineering for the X-38 model started with digitizing, using a high-resolution 3-D scanner. CAD models were created from the scanner's point cloud data and used in the simulations. NASA contracted with a Southern California company called Capture3-D to do this work.

The X-38 test vehicle is a working prototype that is approximately 33 feet long, 12 feet tall, and 10 feet wide. Although the prototype is 20% smaller than the actual vehicle, it was too big to be captured by a coordinate measurement machine (CMM). Previous efforts to capture the X-38 prototype with a laser tracker met accuracy requirements, but not the data density needed to recreate surface details for reverse engineering. To overcome these obstacles, NASA used a combination of a white-light scanner and a photogrammetry system.

The X-38 model had to be scanned in a working vehicle bay. No spraying of the model was allowed to reduce shiny surfaces. During the scanning, bay doors were opening and closing, ambient lighting changed, and objects were moved. Furthermore, there was only three feet of clearance between the underside of the vehicle and the floor.

These conditions required that the scanning system perform normally under external influences while keeping operators informed if the conditions might influence data integrity. The scanning system had to be able to work in both large and small areas with restricted space under the object.
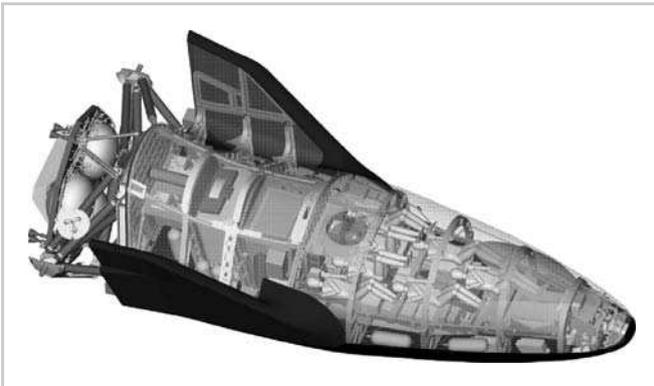
Figure 8.8. X-38 model. A color reproduction of this figure can be seen in the Color Section (pages 219–230).
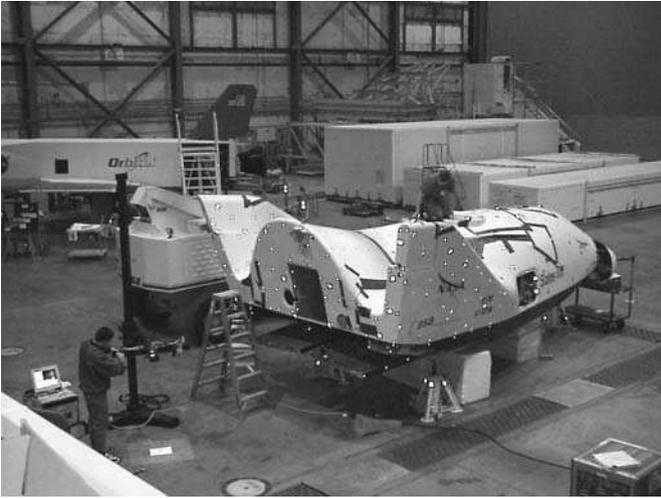
Figure 8.9. Setting up measurement

Figure 8.10. Application of the markers

To start the measurement, coded and noncoded markers and two scale bars were applied to the model. Flat white targets were also placed on the model to define the coordinate system of a tooling hole. The digital images were then captured, allowing the photogrammetry system to define the exact coordinates of all markers. From this measurement, the reference file for the detailed digitizing was derived.

Figure 8.11. Photogrammetry measurement
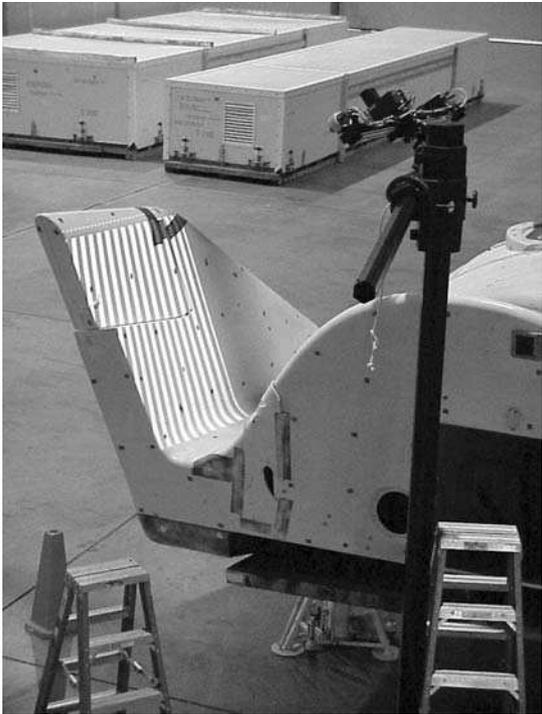
Figure 8.12. Scanning the inner side of the wing

Figure 8.13. Scanning the outer side

Measuring volume was set to $800 \times 640 \times 640$ mm to digitize the upper and side surfaces of the model. When scanning the underside, the measuring volume was adjusted to $360 \times 280 \times 280$ mm with a shorter standoff distance due to the limited space under the object. The adjustment, including the calibration of the system, was completed in 15 minutes.

The captured data were automatically combined into one point cloud, and a polygonal model with the required data density was computed. From the polygonal model, additional computation was done to create cross sections, feature lines, and other attributes. A CAD model was then generated using reverse-engineering software.

Figure 8.14. Individual measurement at one-quarter of the maximum data density
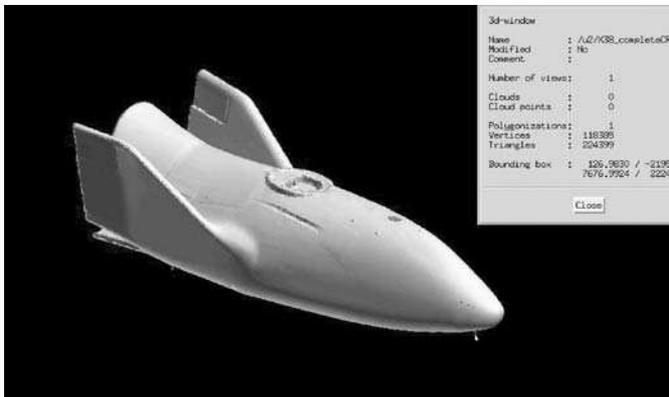
Figure 8.15. Dense polygonal data with 6 million triangles

Measurement, data calculation, and data postprocessing took just 4 days. For the CFD analysis, a CAD model was derived in eight 8 hours based on decimated polygonal data. For detailed analysis, a CAD model based on the dense polygonal data was built in 5 days.

Color visualization showed very little deviation between the actual model and the CAD data. Only the symmetry of the wings was slightly out of tolerance. In addition, the CFD analysis showed that a simulated hard landing–a condition where a vehicle hits the ground harder than expected–had not caused damage to the model. Based on its findings, NASA was able to make some corrections in the prototype to improve performance.
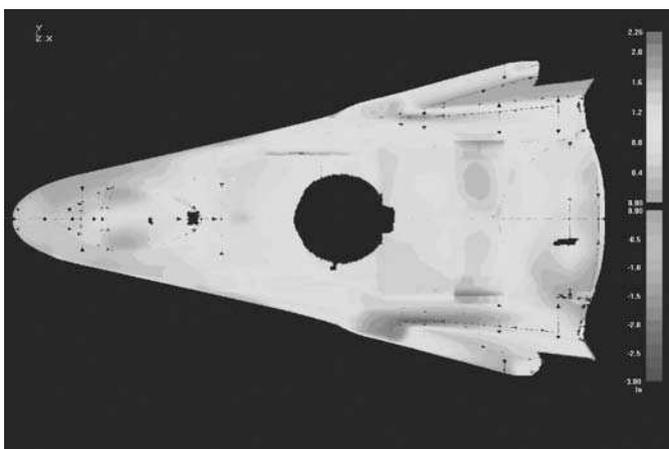
Figure 8.16. Top view of as-built compared to as-designed. Copyright © Aerospace Manufacturing Technologies, Inc., Arlington, WA, USA, http://www.amtnw.com/. Reproduced with permission. A color reproduction of this figure can be seen in the Color Section (pages 219–230).
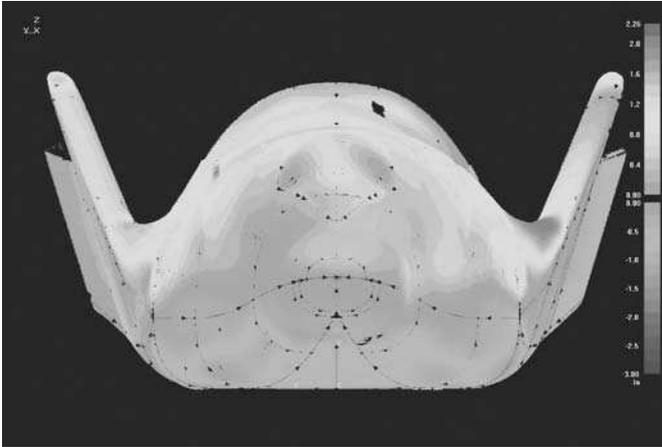
Figure 8.17. Rear view of as-built compared to as-designed. Copyright © Aerospace Manufacturing Technologies, Inc., Arlington, WA, USA, http://www.amtnw.com/. Reproduced with permission. A color reproduction of this figure can be seen in the Color Section (pages 219–230).

The success of this project has led NASA to adopt reverse engineering as part of its processes, providing the agency with the equipment to digitize an entire full-sized X-38 CRV in-house.

8.5 Inspection in Half the Time

Delays in the aerospace supply chain have the same effect as lane closings on highways. Nowhere is this more apparent than in quality inspection. "Better processes in design, analysis, machine tools, and cutter technology have improved manufacturing reliability and speed tremendously, but quality inspection still lags behind", said Matt Reinhard, CEO of Aerospace Manufacturing Technologies (AMT). "Manufacturing has a fire-hose capacity, but quality inspection clamps it down to a drizzle."

AMT manufactures 3-, 4-, and 5-axis milled parts, stretched-formed parts, and sheet metal parts of all sizes for the aerospace industry to speed its part verification. The company recently turned to noncontact laser scanning and computer-aided inspection (CAI) based on reverse engineering. Traditional coordinate measurement machines (CMMs) are used for final quality inspections.

Quality has always been a big focus for the aerospace industry but it also has been a huge bottleneck, perhaps the biggest in the design–engineering–manufacturing cycle. Pressures to eliminate the inspection bottleneck have led AMT and other aerospace contractors to examine new methods of speeding the process.
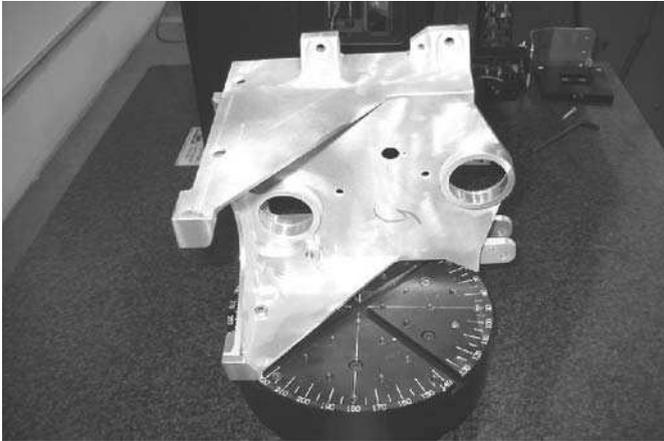
Figure 8.18. Gearbox placed on rotating table ready for scanning. Copyright © Aerospace Manufacturing Technologies, Inc., Arlington, WA, USA, http://www.amtnw.com/. Reproduced with permission

The new CAI process starts with a laser scanner that captures a manufactured part as a highly detailed 3-D point cloud. The engineer places the part ready for inspection on a rotating table and scans it at a rate up to 5,000 points per second. If the part is too large for the cube, the scanner can be moved and configured in a larger area to capture the necessary points.

The large numbers of points captured during scanning enables close inspection of complex contours, forms, shapes, and holes quickly and accurately. This is a major advantage over CMMs, which can scan only 30–40 points in the time it takes the laser system to scan thousands.

Another major advantage is automation: The laser scanner's automatic control and calibration doesn't take highly skilled engineers to perform the inspection. For AMT, it also means not tying up valuable time on two CMMs that cost nearly $275,000 each. Before implementing the laser scanner, AMT's first-article quality inspection could take half a day to set up and hours to run. By changing to noncontact part verification and automating much of the work, inspections can be set up in as little as half an hour. The actual inspection might take only a few minutes. The reduced time saves manpower, cuts costs, and adds flexibility for design changes.

"Faster response times are critical", Reinhard said. "While CMMs are accurate, the high cost, slower speeds and difficulty correlating to 3-D CAD models make them less conducive to adjustments during the manufacturing process. Customers like Boeing are pressured by OEMs as they try to cut inventory, and we need to decrease the time involved in the inspection phase of manufacturing."

The second part of the new inspection process involves processing the 3-D point cloud data for the manufactured part and comparing it with the corresponding 3-D CAD model. That requires the latest generation of CAI software.
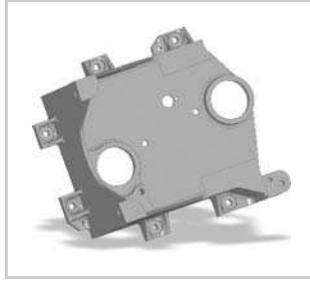
Figure 8.19. CAD model of a Boeing gearbox. Copyright © Aerospace Manufacturing Technologies, Inc., Arlington, WA, USA, http://www.amtnw.com/. Reproduced with permission.

The CAD model supplied by the customer and the point cloud captured by the 3-D laser scanner are loaded into the CAI software. The CAD model and the point cloud are automatically aligned, and color maps are generated to depict deviations. As a user runs graphical comparisons for dimensions, data, features, and other quality factors, the software automatically records the session for reporting purposes.

Reports from the CAI software are used to help designers decide what alterations are necessary, to make changes to milling machines, or to certify that the part is suitable for further production. Reports can be formatted in HTML, PDF, Microsoft Word, and Excel, allowing them to be used in print or online for collaborative assessments.

Behind the scenes, a template of information gathered by the scanner and the CAI software during the first inspection of a particular part allows subsequent inspections of the same part without a new setup. AMT is using this capability to develop a library of 200 different part inspection profiles that will further automate its processes.

Once the library is implemented, AMT will be able to put a part on the scanner, hit a barcode, and automatically scan the part. The CAI software will bring in data from the inspection library and then develop a finished inspection report without any user intervention.

Eventually, CAI systems could extend beyond first-article inspection and dimensional verification to encompass the entire quality inspection process. But for now, industry standards dictate using CMMs for quality inspection before shipping parts to customers. The obstacle is getting customers familiar with the first-article inspection process. An industry standard for noncontact part verification will improve the flow of the supply chain and relieve pressure on aerospace suppliers' customers as they move to just-in-time inventory and other cost-saving measures.
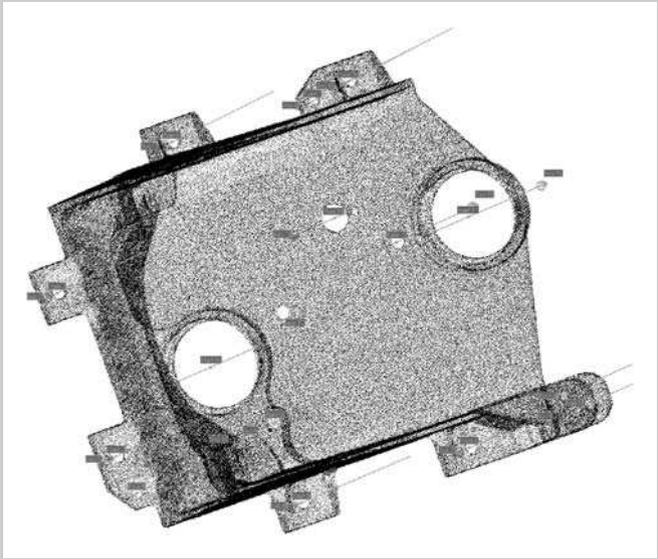
Figure 8.20. Point cloud from scanned gearbox. Copyright © Aerospace Manufacturing Technologies, Inc., Arlington, WA, USA, http://www.amtnw.com/. Reproduced with permission.
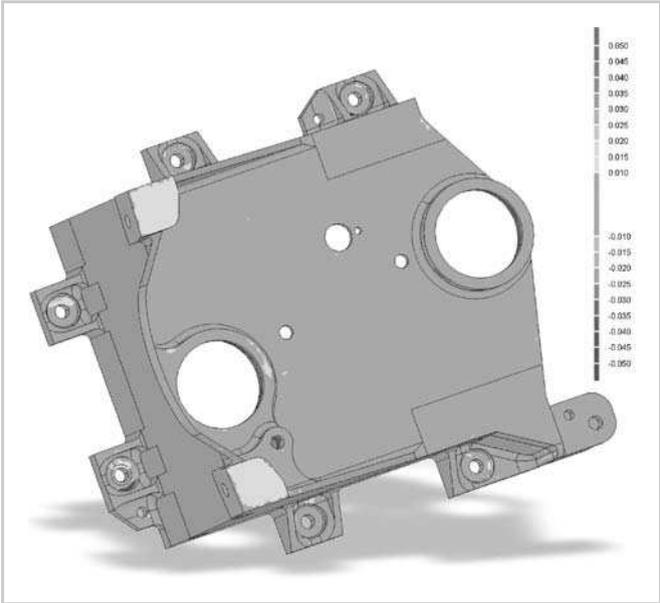
Figure 8.21. Comparison model detailing deviations between the CAD model and the scanned part. A color reproduction of this figure can be seen in the Color Section (pages 219–230).
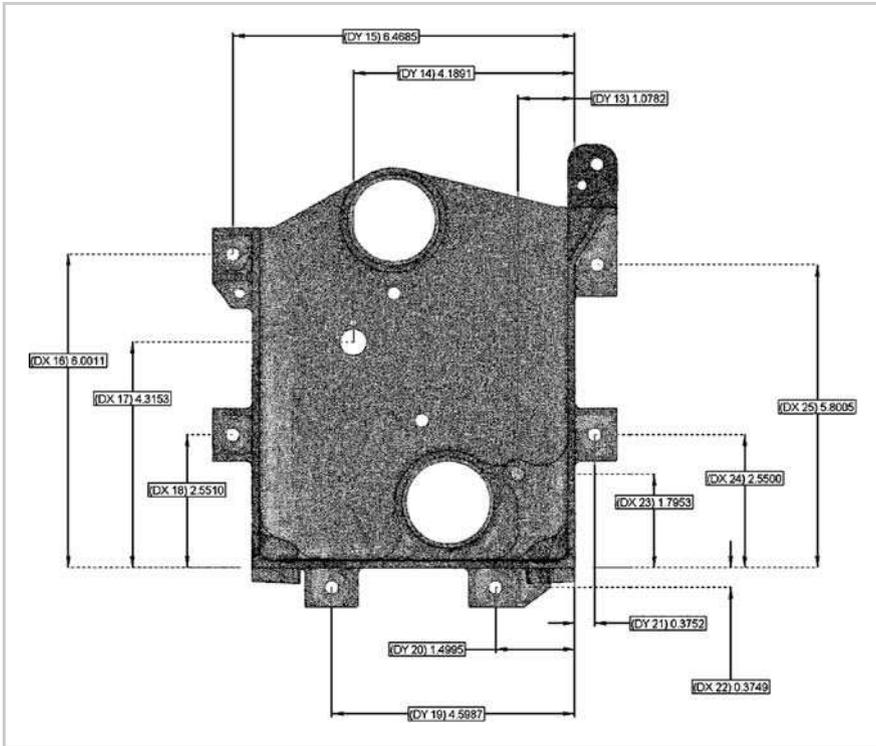
Figure 8.22. Dimensions rendering automatically generated by CAI software

8.6 Making the Next Great Leap

Boeing invested at least $1 billion (insiders say much more) in CAD infrastructure for the design of the Boeing 777. The company based its process on a commercial CAD package and in-house digital pre-assembly applications. Much of the same technology was used on the B-2 program.

The marketing hype has certainly distorted the picture, but the reality , nonetheless, is that Boeing reaped huge benefits from design automation. In comparison to earlier aircraft designs such as those for the 757 and 767, Boeing achieved the following:

- eliminated more than 3,000 assembly interfaces, without any physical prototyping;
- reduced engineering change requests by 90%;
- reduced cycle time for engineering change requests by 50%;
- reduced material rework by 90%; and
- gained a 50X improvement in assembly tolerances for the fuselage.

Though the Boeing 777 results are impressive and worthy of deep respect, the process shows the glaring limitations of existing CAD tools. Hard tooling still requires physical storage because there is no easy way to update the CAD model when design changes occur during tooling and fabrication and there is no fast and easy way to inspect as-built parts against their corresponding CAD models.

Although there is finite-element modeling of static stresses and loads, all dynamic modeling of actual flight, including aerodynamics and structures, is done by conventional CFD and flight simulation with no connection to 3-D solid modeling. Each of the individual modeling efforts benefits from the separate CAD tools available in their specialized domains, but this is far from the highly integrated environment that is envisioned for the future and far from the advanced practices in many other industries.

The next major step for Boeing and other aerospace companies is a much higher level of integration for the design process and the parts used for new and aging airplanes. Reverse engineering holds the potential to provide a closed loop among CAD, CAM, CAE, hard tooling and inspection, making it much easier to exchange data. Even the seemingly simple process of measuring, modeling, and verifying a physical part requires companies to look beyond traditional solid modeling systems to new reverse engineering technology that can capture the physical world and represent it with accurate, manufacturable digital models. Aerospace engineers view this step as an enormous challenge–the sheer scale of tracking millions of components can be conceptually and computationally overwhelming–but one that must be faced.

It might help to take a deep breath and look back at the 100 years of aviation. It's an amazingly compressed success story, from the early days of scepticism about the Wrights' initial flight to the development of the Boeing 777. Aviation has progressed in leaps and bounds since 1903, pushing flight beyond the skies and into space. During the next decade, sights will be set on the frontier traversed by reverse engineering: conquering the divide between the physical and digital worlds and creating full integration among all facets of the aerospace manufacturing process.

Acknowledgments

The vendors of the technologies used in these case studies include

- GOM–structured light scanner
- Leica–laser tracker
- Romer–portable CMM
- Faro–portable CMM
- UGS–CAD software
- Geomagic–reverse engineering and computer-aided inspection software (Figure 8.5)

- Verisurf–computer-aided inspection software
- Metron–laser scanner

Chapter 9

Reverse Engineering in the Medical Device Industry

Ping Fu, Geomagic, USA

9.1 Introduction

Since the advent of computer-aided design (CAD), researchers have tried to create products in digital form. But, despite major advances in mechanical CAD, 99% of the things we use or treasure in our daily lives do not have manufacturable digital representations. Nowhere has this problem been more of an obstacle to progress than in the medical device market.

Medical devices are an important class of products that help us overcome disability and improve our quality of life. But for the most part, the medical device industry has not benefited widely from the digital revolution. Most technology developed in the last 30 years has been aimed at mass production, where the first step is to develop a product from a concept. In the medical device field, the process often starts by capturing a 3-D shape from a person's unique form to ensure that the device is designed for a perfect fit.

In recent years, technology advances from automotive and aerospace industries have made their way into the medical device market. 3-D scanners and industrial desktop CT have advanced to where measuring firm surfaces, such as dental and ear impressions, is accessible and cost-effective. Custom software is now simple enough to be used by nontechnical operators, and automated fabrication systems with new medical-grade materials are production-friendly and efficient.

New medical device production systems usually consist of four major components:

1. Measurement system: measurement hardware that captures human anatomy and outputs measurement data in spatial points.

2. Software design system: a software package that reads spatial points, automates the design process, and outputs manufacturable digital data that combines the human form and product function and fit.

3. Fabrication system: a hardware system that can produce large quantities of devices from one-of-a-kind digital design data.

4. New materials: materials developed for new fabrication systems that are equal to or better than materials used in traditional manual processes.

Turnkey systems that combine the four technological advances have made a significant impact in orthodontics, general dental, hearing instruments, and other medical fields. They are fueling some of the most important developments in medical device history. In this chapter, we examine the new generation of medical device manufacturing processes, as well as research based on reverse engineering that could benefit millions of lives.

9.2 Orthodontics Without Wires and Brackets

Most of us agree that a trip to the orthodontist isn't one of our favorite things. The procedures are expensive and painful with less than optimal results. On top of that, the wires and braces required for treatment leave patients feeling unattractive. What if we could make orthodontic devices that are invisible, removable, and comfortable? This question led to the start of Align Technology and its quest for a digital manufacturing process that sets the milestone for mass customizaion of medical devices at the beginning of the twenty-first century.

We are all born with uniquely shaped teeth that cannot be captured with primitive geometry. The only way to model teeth accurately is first to scan them. For Align Technology, reverse engineering is not just a way to improve an existing process–it is the key technology that enables the company to introduce a disruptive new product with distinct customer benefits.

Align Technology designs, manufactures, and markets the Invisalign System, a proprietary method for treating malocclusion, or the misalignment of teeth. The Invisalign system corrects malocclusion using a series of clear, nearly invisible, removable aligners that gently move teeth to a desired final position. Because it

Figure 9.1. At left, traditional braces with wires and brackets. At right, invisible aligners from Align Technology.

Figure 9.2. A sequence of invisible aligners

does not rely on the use of brackets and wires, the Invisalign system significantly reduces the discomfort, embarrassment, and other limitations of braces.

In Align's treatment regime, anywhere from 12–50 invisible aligners are supplied within the course of a typical treatment. Aligners are manufactured to an accuracy of 0.005 of an inch for each stage of an individual's treatment.

Three major technologies drive the company's highly customized design, engineering, and manufacturing processes:

1. *high-resolution scanners* that digitally measure the dental impression of an individual's upper and lower arches;

2. *computer graphics technology* to compute and visualize the movement of teeth during each treatment stage; and

3. *advanced digital manufacturing systems* such as stereolithography machines to make individualized molds for shaping aligners.

Jon Fjeld, vice president of engineering at Align, says that Invisalign isn't a new idea. The concept first appeared in orthodontic literature in the 1940s. The Invisalign concept wasn't realistic, however, before digital processes became available and practical to use. According to Fjeld, "It's only possible because we have the ability to model in software, treat in software, and then manufacture the appliances".

Align has more than 40 U.S. patents filed for technologies related to its processes. The patents encompass everything from mold acquisition and bite registration, to tooth-path modeling, orthodontic visualization, computerized machining, and data compression for transmitting images over the Internet.

The success did not come easily. Align went through many iterations of trial and error because the technologies it was using were not yet mature. This is what distinguishes a visionary company from followers. Other manufacturers had access to the same technologies. Align had a vision and was willing to take the risk to create something new and better.

9.3 Improving the Scanning Process

An example of the obstacles Align had to overcome can be seen in the process of scanning dental impressions. In 1998, when Align started its work, most imaging-based scanners could not capture data in between teeth and obtain the details of the gingival line at the speed and accuracy required by this application.

For its first attempt, Align chose a CGI scanner, a destructive method that requires messy casting and one additional plaster model to be made with special material. The process started with two plaster models made from a dental impression. One of the plaster models, white in color, was encased in black material and hardened for about 17 hours into a block. The block was shaved in thin layers, with a digital image captured for each layer. An image processing technique was used to obtain contour lines from each layer. The images were combined and processed in software to create 3-D digital models of dental arches. Approximately 14 arches were processed in each batch.

This process worked adequately for research, but was not fast enough for production purposes. It was also not environmentally friendly because the process produced a lot of waste materials. After careful evaluation, Align had no choice but to build its own automatic scanners for production.

The second process used a 5-axis precision machine, a structured-light scanner head, and home-grown software. The project took longer and cost more than Align anticipated, but the result was a production-ready turnkey system that could scan and digitally duplicate dental arches in minutes with 0.1 mm accuracy. This process required only one plaster model and was clean and fast, but it scanned only one model at a time.

Today, Align uses a desktop industrial CT to scan the dental impression (native mold), saves the step of making a plaster model (positive) and enables the company to scan multiple impressions at the same time.

In the future, Align wants the scanning to be done at the orthodontist's office. This requires an intra-oral scanner that can capture dimensions inside the mouth. Once a commercially viable intra-oral scanner is available, there should be no need for making dental impressions, a process that is uncomfortable for patients and unreliable due to human error. The intra-oral scanner will also
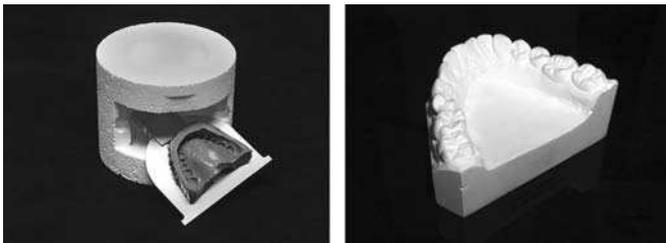
Figure 9.3. At left, a dental impression. At right, a dental plaster model. Copyright © Hytec, Inc., Los Alamos, NM, USA, www.hytecinc.com. Reproduced with permission.

eliminate the current process in which the orthodontist mails dental impressions to Align. This will save time and improve record keeping.

9.4 The Six-stage Process

The current Invisalign process comprises six stages:

1. The orthodontist prepares a diagnosis and treatment data, including X-rays, photographs, an impression of dental arches, and a wax bite showing the relationship between the upper and lower dental arches. The materials are sent to Align with a prescription describing the desired positions and movement of the patient's teeth.

2. The orthodontic data are used to construct a treatment based on the patient's dentition. The dental impressions are scanned with the CT system to create digital replicas. The 3-D digital model and the orthodontist's prescription are sent via the Internet to Align's virtual treatment facilities in the United States and Mexico.

3. Hundreds of software operators and clinic staff in virtual treatment centers prepare computer-simulated treatment and visualizations using Align's proprietary software. Operators create a 3-D visual plan that simulates tooth movement in 2-week increments during the treatment period. Align provides free software that allows orthodontists to visualize treatment over the Internet. Orthodontists can review the treatment and request adjustments if necessary. They can also invite the patient to see the course of treatment. When the orthodontist approves the proposed treatment, manufacturing begins.

4. Align software replicates the patient's tooth positions at each 2-week stage of the treatment. The molds are fabricated directly from 3-D computer models using stereolithography machines. Multiple molds are made at the same time, providing high production throughput. A set of 6-month treatment molds for upper and lower teeth arches are sent to the manufacturing center.

5. The aligners are manufactured by robots, using a machine that places pressure-forming polymeric sheets over each mold. Finished aligners are trimmed by laser cutters and packaged for each patient. Following final inspection, they are sent to the prescribing orthodontist.

6. After 6 months of treatment with the invisible aligners, new impressions of the upper and lower teeth arch are made and replicated into digital models. The new digital models are compared with the previously simulated models at the 6-month stage. If they are the same, the manufacturing center will make the aligners for the next 6 months. If they differ, the new models are used to guide the re-computation of treatment for the next 6 months. The changes are recorded and aligners are manufactured for the next 6 months

based on the new information. This procedure continues until the end of the treatment, which is typically between 10–18 months.

A major part of software development at Align is aimed to reduce the virtual treatment work done by operators. Align has created software that automates time-consuming tasks such as identifying different teeth, calculating a clinically acceptable final tooth position, deriving intermediate states between the start and end of treatment, reconstructing an approximation of the gum lines for treatment, recomputing the treatment plan in midcourse if necessary, and recording and tracking all changes.

In manufacturing, Align implemented software that automatically trims aligners on computer-aided numerical control (CNC) machines. Originally, that work was done manually. The company also implemented software that rapidly shells 3-D digital models to save material when producing molds on stereolithography machines.

9.5 Achievement

In 5 years, Align achieved what experts in the field thought impossible. The company used reverse engineering and other technologies to differentiate itself in the marketplace and compete effectively against giants in the medical device industry. In recognition of its achievement, Align received the 2002 Medical Design Excellence Award for advances in product innovation, design, engineering, and delivery.

Align's unique process has made it one of the largest manufacturers of mass-customized products in the world. More than 100,000 patients have entered treatment with Invisalign, and the company has produced more than 5 million custom aligner devices.

9.6 Digital Dentistry Becomes Reality

For two decades, digital dentistry meant simply capturing an image of teeth with a dental camera. The skilled work of making crowns and bridges was entirely in the hands of craftsmen at dental labs. During the last couple of years, this scenario has changed. The German company KaVo has seen the future and it is fully digital. KaVo has developed a complete computer-based system with stand-alone components that automate the process of preparing inlays, onlays, crowns, bridges, veneers, and frameworks in an almost unlimited range of materials for dental laboratories.

Dental lab work has traditionally involved a great deal of craftsmanship to create a decent-quality appliance from an impression. Faced with increased demand and higher labor costs, dental labs outsource work to places with lower

Figure 9.4. The KaVo Everest system includes a 3-D scanner, reverse engineering software, a 5-axis milling machine, a sintering machine, computer, desk, and material. Copyright © KaVo Dental Corp., Lake Zurich, IL, USA, www.kavousa.com. Reproduced with permission.

Figure 9.5. Everest 3-D Scanner. Copyright © KaVo Dental Corp., Lake Zurich, IL, USA, www.kavousa.com. Reproduced with permission.

labor costs, and often incur losses in productivity, profitability, and quality. KaVo designed its Everest system to change all that.

The Everest process starts with a 3-D scanner that can capture a model of a tooth stump in about four minutes with an accuracy of 20 μm. The scanner can be operated easily with one hand and provides good coverage of complex shapes. It looks like a microwave oven with a turntable that tilts. A single-button push sets the scanner in motion. It captures data from 15 different angles and produces 15 partial point clouds that are automatically aligned and merged into a single point cloud with 1:1 precision compared to the original. The scanner is easy to calibrate, and an adjustment once a week is sufficient to ensure high precision.

The second step is to process the order through Everest control software, which is automated and pre-installed in the computer connected to the scanner. There are four modules in the control software: scan, surface, CAD, and CAM.

The scan module forms the user interface to the scan unit. With a few menu-guided steps, the operator can control the scan process for both stone and wax

models. For direct scanning of production copings, KaVo uses special gypsum, which permits reflection-less exposure to guarantee reliable measurement of shiny surfaces. Scan data are automatically passed to the surface module.

The surface module imports two data sets. One set is from the tooth stump, also called a die, and another is from the wax model of the oversized crown. The first step is to align the 15 different scans of die relative to the surface of the crown and combine the die and the crown into the same orientation. This forms the basis of the mathematical surface calculation that is performed automatically. A preparation line is automatically detected on the die, and the oversized
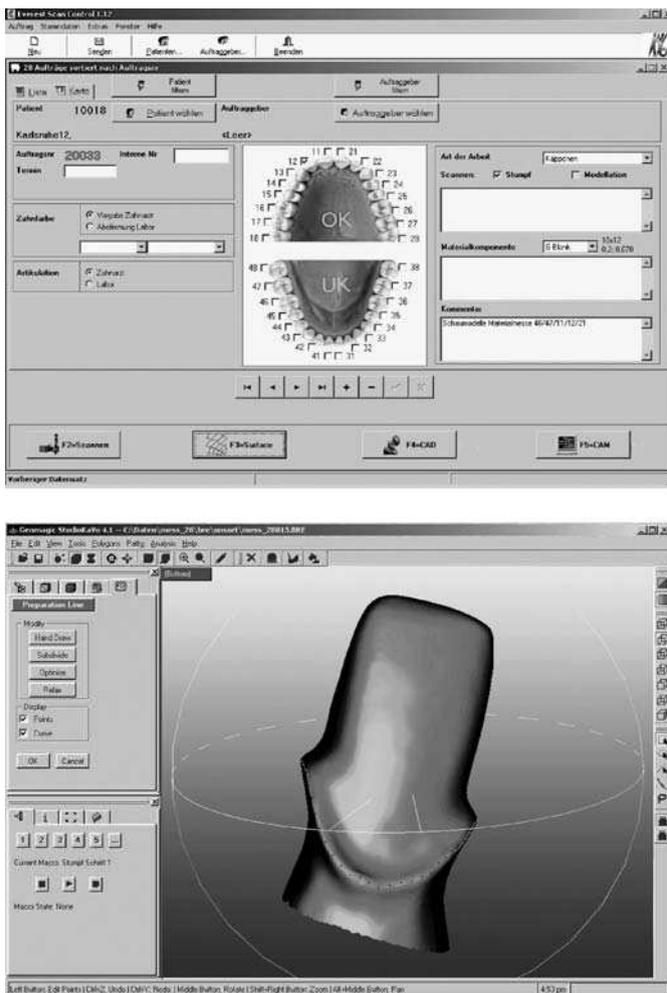
Figure 9.6. *On top*, the Everest software interface. *On bottom*, a screen shot of the surface module. Copyright © KaVo Dental Corp., Lake Zurich, IL, USA, www.kavousa.com. Reproduced with permission.

crown is cut according to the line. The die and the crown are then oriented and automatically adjusted so that there is no undercut. All of the operations in the surface module can be run automatically or interactively for operator-assisted modifications. Bridges are prepared in a similar process. Once the crown or bridge surface is prepared, the file is passed on to the CAD module.

The CAD module enables operators to design the final copings or bridges on the computer screen. The software automatically positions the digital model and trims these to the gingival. The model can also be individually positioned or scaled. The finished design is now ready for the CAM module.

The CAM module automatically calculates cutting data and takes into account the properties of the material that will be used for the final product. It also computes the blank size and avoids collisions automatically.

Everest Engine is a desktop 5-axis numerical control milling machine. No cutting or grinding task is geometrically too complex for the system. After the crown or other dental appliance is manufactured, the material is sintered, if required. on a microprocessor-controlled thermal unit that operates without supervision.

In an age of compromised quality due to outsourcing, KaVo's Everest system makes it cost-effective to keep production of dental devices in the experienced hands of dental technicians. All work can be conducted on-site, providing greater control over production schedules and quality.

The KaVo Everest system is a complete reverse engineering system that combines all four technological advances described in the introduction: new 3-D scanners, automated reverse engineering software, a fabrication machine that produces one-of-a-kind custom devices in large quantities, and new medical-grade fabrication materials.

Like Invisalign, KaVo is an example of how a visionary company can position itself as a market leader by being the first in its industry to adopt state-of-the-art technology. Others will follow, once the leaders pave the road.

9.7 Hearing Instruments Meet the Digital Age

Estimates of the total population suffering from hearing loss in industrialized countries run as high as 10%. Of that, only 20% currently owns hearing instruments. The hearing instrument market is estimated at approximately \$2 billion at the manufacturers' level and \$5 billion at the retail level. The more compelling number is not the size of the market today, but the potential market if all people with hearing loss purchased hearing instruments. In the United States alone, 24 million people suffer from significant hearing loss. Custom-fit hearing instruments are popular in the UnIted States and have 80% of the market. Some of these instruments are so advanced that they could be called hearing computers–the earliest wearable computers.

Custom hearing instruments need highly individualized shells. For decades, making custom shells for hearing instruments was done by sculptors. The process
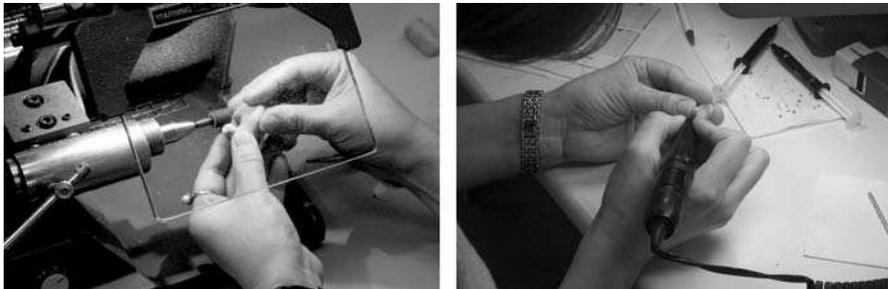
Figure 9.7. Traditional process using manual sculpting and drilling. Copyright © GN ReSound Corp., Bloomington, MN, USA, http://205.242.218.80%20/home.aspx. Reproduced with permission.
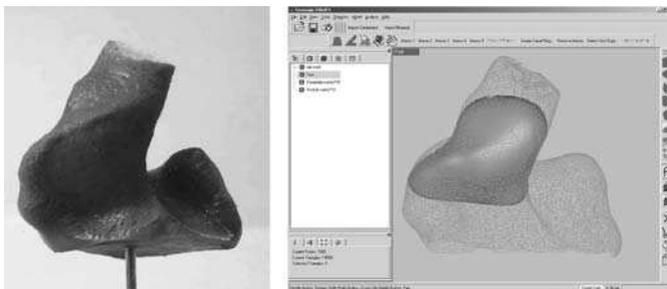
Figure 9.8. At left, an ear impression. At right, the digital shell created using Geomagic eShell software, a specialized reverse engineering application. Copyright © GN ReSound Corp., Bloomington, MN, USA, http://205.242.218.80%20/home.aspx. Reproduced with permission. A color reproduction of this figure can be seen in the Color Section (pages 219–230).

was labor-intensive and could not be easily scaled to increase production efficiency. The sculptors had different levels of skill, making it difficult to control quality.

The scenario is changing quickly, however, thanks to new reverse engineering and rapid manufacturing systems, and innovative use of these systems by hearing instrument manufacturers such as GN ReSound, one of the world's largest hearing instrument manufacturers. The company tested digital methods in the early 1990s, but technologies and materials were not ready for production use. As new digital technologies became available, GN ReSound was among the first to adapt them to its production process.

For GN ReSound, the new digital process not only means increased manufacturing efficiency, it holds the promise of reaching the holy grail for hearing instrument companies: designing products attractive enough for the 80% of people who need hearing instruments, but refuse to buy them for reasons of aesthetics, comfort, or cost.

GN ReSound has benefited from a quick adaptation curve in reverse engineering and rapid manufacturing. The company began developing its process in

December 2001 and 18 months later was producing 1,500 units a month, with plans to double that during the following 3 months. The initial benefit is a 75% reduction in labor to manufacture hearing instrument shells. GN ReSound's current process goes like this:

1. An impression of the inner ear is made in the traditional way.
2. The impression is trimmed and measured by a specialized 3-D scanner.
3. A rubber ear mold is made to test the fit of the finished device.
4. Scan files from the impression are loaded into reverse engineering software to design the hearing instrument shell. This is the virtual sculpting step that replaces the old manual processes. The software simulates the traditional process and also provides visual feedback and automation, making the sculpting process easier and reducing human error. The shell exterior is free-form and must match a person's ear canal. The shell interior is produced using some traditional mechanical design procedures for operations such as creating a predefined hole at the canal tip for receivers, designing a vent along the profile of the shell, adding and removing a local region, and eliminating undercuts. The software creates a watertight STL file for rapid manufacturing, and a virtual ear mold to visualize the fit between the ear mold and the designed shell.
5. Shells are made with either a selective laser sintering (SLS) or a stereolithography (SLA) system that creates a physical part by heating and hardening powder or liquid material, layer-by-layer. The completed part is then sandblasted, tumbled, and matched to the original impression.
6. The fit of the components is verified. The device is checked for protrusion in the rubber ear mold.
7. The components of the hearing aid are assembled.

The digital reverse-engineering process used by GN ReSound is in its early stages, essentially a digital process that simulates the traditional process. Major improvements in the future are expected to come from further automation of the sculpting process, more efficient finishing and machining, and implementing new designs that are not possible in the traditional process.

Figure 9.9. An SLA system and the hearing instrument shells it produces. Copyright © GN ReSound Corp., Bloomington, MN, USA, http://205.242.218.80%20/home.aspx. Reproduced with permission.

Further automation of the sculpting process is an immediate priority in 2004. Although sculpting is performed on a computer, it is still "human-driven" according to Russ Schreiner, GN ReSound's project director. GN ReSound is working on embedding the logic behind sculpting–considered more of an art form than a process–into its reverse engineering software.

"Once that is accomplished", said Schreiner, "software can be written to perform most, if not all, tasks automatically". Schreiner also expects to take better advantage of digital manufacturing capabilities. With a different design, for example, faceplate assembly can be eliminated. The faceplate can be assembled with the shell and fabricated as one piece. "Rapid manufacturing machines allow you to produce features that are not possible with conventional molding or casting", he said. "We expect to design features in the shell that enhance product performance and facilitate more efficient manufacturing."

A major step forward for the industry, according to Schreiner, is eliminating the ear impression. "Making the ear impression is an unreliable process and not appreciated by customers", he said. A few companies are working on scanners small enough to go into the ear canal and make a direct scan. This would take much of the human error out of capturing shape information and enable complete electronic transfer of orders, rather than delivering orders with impressions by next-day mail.

Along with improving the sculpting process, digital capture of inner ear geometry will help deliver better fit. GN ReSound's digital processes have helped improve efficiency in producing "remakes", where a new shell is required due to a fit-related problem. The need for remakes is a perennial problem in the hearing instrument industry. GN ReSound now has all the digital information on file and does not have to go through the process of making a new impression. Increased accuracy from the digital process should reduce the remake rate. An additional benefit of an all-digital process, said Schreiner, is that it will be less expensive to make shells in multiple styles and sizes for the customer to try out before manufacturing. This will help ensure the best fit and feel.

"The new process has proven that it can be just as good if not better compared to the traditional process in overall fit and performance of the product", said Schreiner. "The real breakthrough will come when we make new products that are attractive to those who need hearing instruments but would not consider them in the past."

9.8 Reverse Engineering–A Better Knee Replacement

Knee replacement surgery is performed on more than 250,000 Americans annually, yet less technology is brought to bear on this issue than on the development of common household appliances. "More engineering analysis goes into the washing machine in your home than into the artificial knee joints implanted in

Figure 9.10. Motion data capture and a 3-D model of a knee joint. Copyright © University of Florida, Computational Biomechanics Laboratory, Gainesville, FL, USA, http://eng.ufl.edu. Reproduced with permission.

people", said Dr. B.J. Fregly, assistant professor in the Department of Mechanical and Aerospace Engineering at the University of Florida.

Fregly is addressing that problem by taking technologies traditionally used for mechanical engineering and extending them to the human body. Though the study is still in the early stages, Fregly and his team believe that a greater understanding of the factors that contribute to joint failure will lead to longer lasting artificial knee designs and better surgical procedures.

Problems related to wear typically limit the life span of knee implants to 15 or 20 years–possibly less for more active patients. Because younger patients are being diagnosed with knee joint problems, the restricted life span and functional limitations of artificial knees are becoming an increasing concern.

Fregly, whose previous work has been published in the *Journal of Biomechanics* and the *Journal of Biomechanical Engineering*, leads a group of researchers who are evaluating and testing artificial knee designs by simulating a real-world environment specific to individual patients. His team begins the process by placing reflective markers on the patient's skin and clothing during treadmill walking and stair-rise activities. The patient's motion data are captured by LCD devices and stored in a computer database.

Dr. Scott Banks of the Biomotion Foundation in West Palm Beach, Fla., augments the motion capture data with a dynamic X-ray procedure called fluoroscopy. Fluoroscopy accurately measures knee joint motion for natural or artificial knees. When studying artificial knees, Banks obtains CAD models of the implant components from the manufacturer. He then uses custom software to match the 3-D CAD models to each 2-D fluoroscopic image as though he were orienting an object to a photograph of its shadow. The image-matched components are used to quantify the 3-D motion of the patient's knee under real-life loading conditions, such as walking and climbing stairs.

The process for studying natural knees is more difficult, since CAD models of the bones are not readily available. To create models of natural knees, Fregly's team uses CT scans, which produce static 2-D image slices of a patient's leg. The
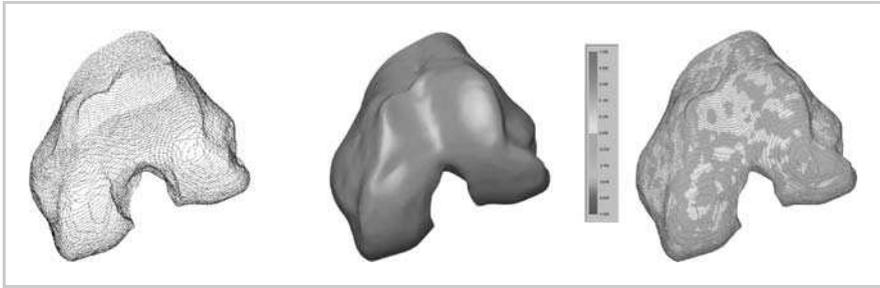
Figure 9.11. Scanned point cloud data, a polygonal model, and a model showing stress analysis. Copyright © University of Florida, Computational Biomechanics Laboratory, Gainesville, FL, USA, http://eng.ufl.edu. Reproduced with permission. A color reproduction of this figure can be seen in the Color Section (pages 219–230).

CT data are imported into image-processing software, where the 2-D slices are stacked to create a 3-D point cloud model.

Researchers then use reverse engineering software to generate accurate 3-D computer models automatically from the point cloud data. The team uses polygonal models for shape-related tasks such as image matching. The polygonal models are then converted into highly accurate mathematical surface models for contact stress analysis.

Once the researchers have developed contact stress predictions from the movement data, the final comprehensive wear model is created with help from Dr. Greg Sawyer, a University of Florida friction and wear specialist. Combining accurate knee motion data with contact stress predictions creates a wear model that pinpoints the exact places where an artificial knee is likely to fail.

Eventually, Fregly and Dr. Rafi Hafka, an optimization specialist in the Mechanical and Aerospace Engineering department, will be able to fine-tune a full-body model to reproduce movement data collected from the patient before surgery. The surgeon and research team will then perform the surgery on the computer model to predict which surgical parameters or implant designs will produce the best outcome for a particular patient. Patients will be followed long-term and will be remeasured in the movement lab following surgery to assess the model's predictions.

By developing a better understanding of how and where stress and movement produce wear, Fregly, Banks, Sawyer, and Hafka hope that advances can be made to extend the life span and functionality of artificial knees.

9.9 The Quest for a Total Artificial Heart

It has been 20 years since surgeons placed the country's first artificial heart in 61-year-old Barney Clark. The Washington dentist lived just 112 days after doctors implanted the Jarvik-7 device. Since then, medical science has continued to
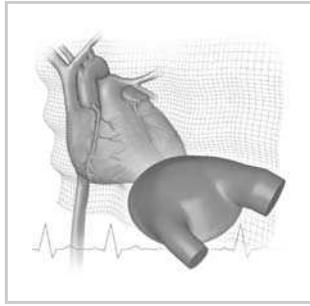
Figure 9.12. In the foreground, a digital model of the Cleveland Clinic's total artificial heart. Copyright © Cleveland Clinic Foundation, Cleveland, OH, USA, www.clevelandclinic.org. Reproduced with permission.

strive for an alternative for those with heart disease. Today, there are still at least 30,000 more people waiting for transplants each year than there are hearts from donors. Medical institutions such as the Cleveland Clinic Foundation are working to create a total artificial heart (TAH) that will bring life to patients without access to donors.

The Cleveland Clinic's most ambitious project is the MagScrew total artificial heart, being developed in collaboration with Foster-Miller Inc., which could eventually be implanted permanently. "Unlike LVADs, which are designed to aid one ventricle of the native heart, our TAH is a device that replaces the sick natural ventricles and native valves of a patient", said Stephan Weber, senior research engineer in the Department of Biomedical Engineering at the Cleveland Clinic Foundation. "Applications for this system would exist for patients with end-stage heart failure who would not survive until a suitable heart donor is located for transplant or for patients who are not eligible for heart transplantation."

When implanted, the inlets to the artificial heart are connected to the natural atria; the outlets connect to the aorta (left) and pulmonary artery (right). Eventually, the Clinic's TAH is fully implanted with no wires passing through the skin for a power source. Implanted with the heart are the controller, internal battery, compliance chamber, and an energy transfer system. One of the most critical issues with the TAH is the way the blood flows through the device. "Certain flow patterns have been proven to cause blood damage", said Weber. "Features have been added to our model to improve the flow pattern through our pump."

Cleveland Clinic researchers determined that the best way to ensure a proper flow pattern was to reverse engineer the complex geometries of the blood pump housings in the TAH. "We needed this model to be as accurate as possible to ensure proper blood flow within the chamber", Weber said. Weber turned to a 3-D modeling and CAD service bureau to supply a surface model and a 3-D solid model representation of the chamber.

A plastic model of the internal volume of one of the blood pump chambers was used to make the computer models. The unusual shape makes it difficult to scan and generate an accurate 3-D model for the device. Engineers began by

painting the pump housing white and scanning it with a white-light scanner. Reverse engineering software was then used to convert the resulting scanned data into a 3-D solid model almost identical to the original device.

"The precision of the software helped us over one of the largest hurdles between the existing device and a TAH that can be manufactured accurately", says Weber. "The process saved both time and money. We had attempted to measure the geometry of the volume using a CMM machine, but this took too much time and the resulting model was not even close to being accurate enough."

Weber thinks that the reverse engineering approach brings the Cleveland Clinic a little closer to making its total artificial heart a reality. For heart patients, it could eventually mean the difference between life and death.

9.10 Moving Toward Mass Customization

Henry Ford's first great contribution to America was the Model T, which rolled off the assembly line at the rate of one every 24 seconds. At the time, it was an amazing display of industrial efficiency. Such production line automation has made things more affordable, but at a cost of one-size-fits-all.

Today it's not the raw power of machines that is shaping our economy, but the subtle power of knowledge. Reverse engineering technology with its computer-based applications and software is erasing the edict that customizaion must carry a high price. The medical device industry is in the forefront of the next industrial revolution: mass customizaion–the ability to produce individualized products with the same efficiency and cost as mass production.

Customized design and individualized services will reshape the way products are developed and marketed. Advances in reverse engineering technology hold the promise of bringing the price of personalized goods within the budgets of most consumers. The progression to customizaion arises from the free market's relentless drive to bring what we buy closer to what we want. In one industry after another, companies are using reverse engineering and digital manufacturing to customize products for individual consumers. They are doing it because new technology makes it practical and competition makes it imperative. Nowhere is this more evident than in the medical device market, where products must be made individually for each customer.

When the technology becomes more mature and price/performance reaches a point where mass customizaion can be applied to consumer products, we will see personalized products and services moving from medical treatment to life enhancement. Take hearing instruments, for example. The digital process will move from hearing aids to hearing computers, from medical devices to life-enhancing products. Imagine a hearing computer designed for people with perfect hearing, with functions that include wireless communication, noise reduction, FM transmission, sound enhancement, and an MP3 player, all included in a custom shell that fits comfortably in a person's ear canal.

Imagine that instead of making your feet conform to fit a pair of shoes, you could get shoes that conformed perfectly to your feet. Reverse engineering and digital manufacturing can make this happen and in the process, create new jobs where customers are, instead of outsourcing manufacturing to the other half of the earth. Why should we pay the costs of shipping, inventory, unwanted products, and loss of jobs in our homeland for shoes that do not even fit?

As industrial nations face increasing competition in the world market, keeping skilled laborers employed has been a key social issue. In the transition from craftsmanship to digital processes, GN ReSound has been able to keep its sculptors and teach them how to use computers instead of knives. Not only have the sculptors kept their jobs, there has been increased job satisfaction because of the following factors:

1. The digital process is a cleaner and healthier environment without the dust and noise of the traditional process.
2. Career advancement is achieved by the transition from manual laborer to knowledge worker.
3. Job security comes from the efficiency of the digital process and the requirement for production to be close to customers.

Companies adopting reverse engineering are shifting the business paradigm from manufacturing-centric production to consumer-centric customization. The revolution that happened to one- and two-dimensional computing is about to repeat itself in the three-dimensional domain.

In the one-dimensional realm, the step from analogue to digital signals greatly expanded the capacity of phone, music, and other forms of communication. In the two-dimensional realm, the leap from the typewriter to word- and image-processing forever changed the way information is expressed and disseminated. With fast and accurate 3-D scanners and automated reverse engineering software, we now can digitize body parts and physical objects in their true forms. This offers a convincing alternative to a decade of blank-screen design proliferated by CAD systems.

By making it faster and simpler to customize, reverse engineering will enable manufacturers to deliver exactly what we want at a competitive price. Just as mass production was the hallmark of the Industrial Age in the twentieth century, mass customizaion promises to dominate the Knowledge Age in the twenty-first century.

Acknowledgments

Chapter 10

Legal Aspects of Reverse Engineering

Brian C. Beherens and Reuven R. Levary, Saint Louis University, USA

10.1 Introduction

With the current eruption of technological development, technology is advancing at a rapid rate. The legal world is beginning to realize a need for concomitant advancement. Laws must keep pace with technological development. The copyright laws that currently govern the development of computer software, for example, are still quite ambiguous. There is debate over the scope of the law as well as difficulty in applying the law. In the light of conflicting case law, it is troublesome for attorneys to give clients advice as to whether particular software development techniques violate copyright laws or carry other legal consequences (Behrens and Levary 1998).

As more and more cases involving computer software or computer programs are being decided, laws regarding copyright infringement in this area will become clearer. Both terms, software and computer programs, refer to the written language that instructs a computer which functions and actions to run. This does not make the application of copyright law to software easier today, however. Problems in deciding software infringement cases often arise because the courts have a difficult time understanding the computer technology discussed by litigants. Furthermore, judges are unclear about the methods and procedures that computer companies are using. As more cases are decided, courts are becoming more familiar with the concepts and issues that are being litigated and are thus beginning to make more informed decisions. Courts are beginning to realize how their decisions affect the computer industry as a whole (Cifuentes 1999, 2001).

Reverse engineering as it relates to computer software is coming under legal scrutiny. This will set the stage for the legal impact of reverse engineering in other disciplines. For this reason, this chapter will address various issues

surrounding computer software copyright law. In particular, it will examine the legal aspects of software copyright law, exploring its past, present, and future, while also explaining a relatively recent concept known as "reverse engineering". Section 10.2 will provide a basic overview of copyright law, including its history and the policy considerations behind copyright law. Section 10.3 will explain the methods that computer programmers use to develop new software that can interact with or enhance existing programs, or even compete with them directly. Section 10.4 will detail the most recent case decisions on reverse engineering and the applicable law: Sega Enterprises Ltd. v. Accolade, Inc. (Sega Enterprises 1992) and Atari Games Corp. v. Nintendo of America, Inc. (Atari Games 1992). Section 10.5 will relate the "fair use" statutory defense to a claim of copyright infringement. This section will explain how application of this statutory defense will enable the computer industry to progress instead of becoming bogged down in litigation. Finally, the Conclusion, Section 10.6, will argue that copyright law (as it relates to computer software) should be amended to define the rights of copyright holders better, so that it is more specific and useful to practitioners. Because computer software is a type of creative work that differs greatly from traditional literary works, it is extremely difficult to adapt and apply copyright law to protect it. For this reason, it will be argued that a new system of laws needs to be developed for this unique industry.

10.2 Copyright Law

When the original Copyright Act was drafted in 1909, traditional copyright law focused on protecting literary works (Copyright Act 1909). The Constitution grants Congress the power to enact laws to promote the progress of science and useful arts by securing for limited times to authors and inventors the exclusive right to their respective writings and discoveries. With the advent of computers and computer software, disputes arose over how to protect this new technology. The Copyright Office decided that computer programs were similar in nature to books, and thus began protecting these programs as "literary works" in 1964 (NCNT 1978). Some years later, major revisions in copyright law took place, resulting in the Copyright Act of 1976. Shortly thereafter, in 1978, a congressionally sponsored study by the National Commission on New Technological Uses of Copyrighted Works (CONTU) was finalized in a report. The report ultimately suggested that copyright law was indeed the best way to protect computer software intellectual property rights (NCNT 1978). In 1980, Congress incorporated CONTU's suggestions into the newly amended 1976 Act. To this day, computer software remains under the copyright category of literary works.

Unfortunately, however, CONTU's report and findings did not make analysis of software copyright infringement claims any easier. Software developers, attorneys, and judges alike have had a difficult time applying copyright law to computer software. The difficulty arises as a result of the nature of computer

software. Computer programs are similar to books and other literary works in that they are written, creative products of an intellectual nature. Additionally, computer software, like other written works, is also easily duplicated and reproduced (Nimmer and Krauthaus 1987).

However, computer programs are substantially different from traditional literary works in that words in computer programs are written in cryptic languages that only a computer can decipher rather than in human-readable form. Further, computer programs are more than just words–they are instructions that enable a computer to carry out specific functions and complex tasks. Programmers write software in high-level languages such as BASIC, FORTRAN, COBOL, Ada, Pascal, *etc.*, referred to as "source code", that closely resembles English. After the source code is completed, the programmer uses a compiler to convert the program into "object code", a language consisting of ones and zeros that the computer can understand. Because it cannot be read by humans, once a program is converted into object code, one cannot determine how a program functions without taking steps to reverse this nonreadable language back into source code.

This conversion process presents many legal questions. First, copyright law protects only an author's original *expression* of a given idea and does not protect "any idea, procedure, process, system, method of operation, concept, principle or discovery, regardless of the form in which it is described … or illustrated" (Copyright Act 1988). Therefore, to protect an author's idea from unauthorized use, a patent must be obtained. Increasingly, the argument is being made that computer programs are not simply the expression of an idea but a collection of functions and processes that are not protected by copyright law and thus may be copied. Herein lies the great debate: how rationally to separate the expression of ideas in a computer program from the functions that are carried out. Courts have long been tangling with the idea–expression dichotomy since the case of Baker v. Selden (Baker v. Selden 1879). Unfortunately, there appears to be no easy answer to this question, as Judge Learned Hand noted, "obviously, no principle can be stated as to when an imitator has gone beyond copying the 'idea', and has borrowed its 'expression' … decisions must therefore inevitably be ad hoc" (Peter Pan Fabrics, Inc. v.Martin Weiner Corp.1960)

Second, even if an idea or function carried out in a computer program is not protected by copyright law, is it a violation of copyright law to copy a computer program and to translate it from object code back to source code? These are questions that will arise more frequently as we move into an age where many everyday products incorporate computer chips in their design. For example, most of today's newer automobiles have some type of computer chip, which controls the automobile's engine performance, while storing other diagnostic codes in the chip's memory to better enable mechanics to assess any problem in need of maintenance. Courts will have to determine some type of framework for deciding these cases unless a new statutory scheme that would be drafted by Congress, such as a *sui generis* system often proposed by commentators. The

recent statutory changes in copyright law may not be enough to accommodate the rapidly changing world of computer technology. With these issues presented as a backdrop, computer technology and reverse engineering as well as the legal issues surrounding them will now be examined in greater detail.

10.3 Reverse Engineering

A significant amount of litigation involving computer software focuses around the issue of "reverse engineering". Unfortunately, although much litigation surrounds reverse engineering, the concept is probably one of the most misunderstood by the courts, legal counsel, and executives in the computer industry (Davis 1993). The United States Supreme Court defined reverse engineering as "a fair and honest means of starting with the known product and working backwards to define the process which aided in its development or manufacture" (Kewanee v. Bicron 1974). Additionally, in 1989, a United States District Court defined reverse engineering as "The process of starting with a finished product and working backwards to analyze how the product operates or how it was made" (Secure v. Time 1989). Reverse engineering is the method by which computer programmers study an existing program in machine-readable code by breaking it down into human-readable form to create a similar product or one that can be used in conjunction with the existing software. This process of converting a program from object code back into source code is known as "decompilation" or "disassembly".

Some computer software companies claim that reverse engineering is unlawful because it is far too easy to use the process to discover how an existing program operates and make slight modifications to it and then market the amended version as a new software product (Hamilton 1993). If a second program is developed in this manner that is substantially similar to the original and has more in common than just the functional operations, it is most likely a case of software copyright infringement. Keep in mind that copyright law protects the expression of an author's work, not the underlying functions of a program that can be protected only by patent law.

The more recent cases explored in Section 10.4 suggest that certain requirements must be met today before an altered computer program can be marketed as a "new" product (Hayes 1992; Nimmer and Krauthaus 1987). The truth is that not all reverse engineering efforts are illegal–most of the time they are not. Often, programmers use reverse engineering for purposes other than producing a competitive program. Reverse engineering is used for breaking down software to teach students how to write programs, to repair software that is malfunctioning (Davis 1993), to produce similar software that runs on a different computer system (*i.e.,* on an IBM or a Macintosh), to modify a program for use on one's own computer (Lewis v. Nintendo 1992), and to develop software that operates in conjunction with the original software.

10.4 Recent Case Law

The goal of copyright law is twofold. First, copyright law is intended to promote the freedom of information and advance ideas that may benefit society, and second, to protect the interests of authors in their creative work by giving them control over its dissemination and allowing them to reduce the exploitation of their work. Although these interests must be balanced, the Supreme Court has clearly noted that "the primary objective of copyright is not to reward the labor of authors, but to promote the Progress of Science and useful Arts" (Feist v. Rural 1991).

10.4.1 Sega Enterprises Ltd. v. Accolade, Inc.

Sega Enterprises Ltd. is a producer and marketer of home video game systems and game cartridges. Accolade, Inc. does not manufacture its own video game system but instead develops video games for use with video game systems made by other companies. Accolade wished to develop game cartridges for use in Sega's "Genesis" video game system but did not desire to obtain a license from Sega (Sega v. Accolade 1992). To produce games that were compatible with the Genesis system, Accolade purchased several Sega game cartridges and reverse engineered them to discover the games' source code and to discern how the games functioned. After obtaining this knowledge, Accolade was able to develop its own software for use on the Genesis system and began selling it own cartridges in 1990 (Sega v. Accolade 1992).

Because many companies began to produce games that competed with those developed by Sega, Sega developed another version of its video game system called "Genesis III". This new game system included a feature called the "Trademark Security System" (TMSS), which prevented any competitor's game cartridges from working on the new system unless they contained a certain code. This code was given only to those companies who had licensure from Sega. After Accolade discovered that its games did not run on the Genesis III system, it purchased game cartridges that were compatible with the system, used reverse engineering to obtain the new code, and implemented this code in its own games.

After discovering that Accolade had copied the code for use in its games, Sega filed suit against Accolade for copyright infringement, seeking to enjoin Accolade from reverse engineering any more of its games or from manufacturing any other games with the reverse engineered code. In granting the injunction in favor of Sega, the trial court reasoned that Accolade violated Sega's exclusive right to reproduce the object code by making copies of the original games for purposes of reverse engineering (Sega v. Accolade 1992).

The Ninth Circuit granted Accolade's request for an interlocutory appeal, while also dismissing the preliminary injunction (Sega v. Accolade 1992). On

appeal, Accolade set forth four main arguments: (1) intermediate copying is not a copyright infringement if the end product is not substantially similar to the copyrighted program; (2) disassembly of a computer program is not an infringement of copyright law when it is used to access the program's ideas and functions, pursuant to 17 U.S.C. 102(b); (3) copying a program is authorized by 17 U.S.C. 117, if it is an essential step in the utilization of the computer program in connection with a machine; and (4) disassembly of a computer program is lawful under the "fair use" exception of 17 U.S.C. 107 (Sega v. Accolade 1992). Ultimately, the Ninth Circuit was not persuaded by the first three of Accolade's arguments, but agreed that disassembly was protected under the "fair use" exception to copyright law (Sega v. Accolade 1992).

In assessing Accolade's "fair use" defense, the Ninth Circuit was guided by the factors set out in section 107 (Copyright Act 1992) to determine whether the use was "fair". Section 107 sets forth the following four factors:

1. The purpose and character of the use, including whether such use is of a commercial nature or is for non-profit educational purposes;
2. The nature of the copyrighted work;
3. The amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
4. The effect of the use upon the potential market for or value of the copyrighted work.

In consideration of the first factor, "the purpose and character of the use", Sega argued that Accolade's sole intent in copying the code was for the commercial purpose of developing software to compete with that of Sega and was not therefore unfair. Although proof of a commercial infringer, the Ninth Circuit agreed with Accolade on this point that the direct purpose of copying the code was to ascertain the proper functional elements needed to produce software that would be compatible with the Genesis III system. The court pointed out that the functional aspects of the Genesis III-compatible games were not protected by copyright law. Additionally, the court noted that for purposes of public policy, encouraging competition within the industry was healthy, and thus copying a program to obtain a code for compatibility purposes was valid; This reasoning by the court was somewhat groundbreaking because it was unlike similar cases which held that copying software to produce other compatible software is not a legitimate reason for infringing copyright law.

In analyzing the second factor of the fair use defense, the "nature of the copyrighted work", the court also found this to be in favor of Accolade. In explaining the difficulty of analyzing this factor with regard to computer software, the court pointed out that most literary works are in human-readable form, whereas the ideas and functions of a computer program cannot be discerned without disassembling the program. In this case, the code needed for compatibility purposes was hidden within Sega's software. The public policy goal of market competition is to be fostered, the court explained, and "if disassembly of copyrighted object

code is per se an unfair use, the owner of the copyright gains a de facto monopoly over the functional aspects of his work–aspects that were expressly denied copyright protection by Congress" (Sega v. Accolade 1992). The court continued, asserting that "In order to enjoy a lawful monopoly over the idea or functional principle underlying a work, the creator of the work must satisfy the more stringent standards imposed by the patent laws" (Sega v. Accolade 1992). The court suggested that because computer programs contain functional aspects that cannot be protected by copyright law and the only way to examine the functions of a program is to copy it, software should be afforded a lower degree of protection than traditional literary works (Sega v. Accolade 1992).

The third factor of the fair use defense, the amount of the original work copied, was without question in favour of Sega because Accolade had copied the entire code. However, the court felt this factor should receive very little weight because Accolade had copied the code to understand its function, not to produce new software based only on this portion of the copyrighted program. Overall, the court decided that this factor did not preclude use of the fair use defense (Sega v. Accolade 1992).

In assessing the fourth and final factor of fair use analysis, the effect the copying has on the value of the copyrighted work, the Ninth Circuit took into account that most consumers who own video game systems, such as the Genesis III system, purchase numerous games. This was an important aspect in this case because if a competitive product is derived from a copyrighted work and the new product severely diminishes the market value and/or sales of the original product, the probability of succeeding with a fair use defense is slim. In Sega, however, the court referred once again to the public policy of encouraging the free exchange of ideas and market competition and noted that reverse engineering of software encouraged this activity. Because a new product will normally reduce the sales of an existing product, simply arguing that a decrease in sales should warrant a finding of unfair use would be oversimplified and would not carry out the true intention of the Copyright Act. Simply stated, the courts desire competition, not monopolies. Thus, the court found this factor in favour of Accolade as well. The importance of the Sega Court's holding is succinctly represented in its opinion:

Where disassembly is the only way to gain access to the ideas and functional elements embodied in a copyrighted computer program and where there is a legitimate reason for seeking such access, disassembly is a fair use of the copyrighted work, as a matter of law (Sega v. Accolade 1992).

10.4.2 Atari Games Corp. v. Nintendo of America, Inc.

It is no coincidence that the holding of Atari Games Corp v. Nintendo of America, Inc. (Copyright Act 1992) and the decision in the Sega case were similar. Atari was decided shortly before the Sega Court issued its opinion, and thus Atari

established the framework and rationale for the subsequent Sega decision. Like Sega v. Accolade, this case was physically located in the Ninth Circuit, but because it involved a violation of patent law, the Court of Appeals for the Federal Circuit had jurisdiction over the appeal. In this case, Nintendo of America (Nintendo) filed a lawsuit against Atari Games Corporation (Atari) for copyright and patent infringement. Nintendo was a manufacturer of a video game system called the Nintendo Entertainment system (NES), and Atari desired to produce compatible game cartridges for use with the NES. Nintendo had designed the NES so that the video games system contained a complex electronic key and lock type system, which allowed only game cartridges with this specific electronic key to function on the NES. To obtain this "key" or object code, a game manufacturer had to set up a licensing agreement with Nintendo, and then purchase special cartridges with the electronic key from Nintendo at a substantial cost. If a game was not on one of these special cartridges, the game would be locked out and would not run on the NES.

Atari attempted to reverse engineer one of these cartridges to decipher the code for the key-lock system, by trying to "peel" the computer chip, which contained the code. "Peeling" is a process whereby engineers gradually peel a computer chip's layers back and microscopically examine the "1"s and "0"s that make up the object code. This process is very costly and time-consuming and does not often produce meaningful results (Davis, 1993). Failing to have success with this method, Atari attempted to monitor the electronic signals passed back and forth between the NES and the game cartridge to discover a pattern that might reveal the code. This method was also unsuccessful.

Atari then fraudulently obtained a copy of the NES's source code from the Copyright Office by misrepresenting that it needed the code for copyright litigation purposes between Atari and Nintendo. Atari then used this source code to develop its own program that unlocked the NES and allowed Atari's games to run on the system. After Nintendo filed suit against Atari for copyright infringement, the district court entered a preliminary injunction against Atari.

On appeal, Atari argued that it had used a different microprocessor, different instructions, and a completely different computer language in designing its program and therefore there could be no copyright infringement. Atari also asserted the fair use defense that the Federal Circuit ultimately agreed with in theory. The court asserted that "an author cannot acquire patent-like protection by putting an idea, process, or method of operation in an unintelligible format and asserting copyright infringement against those who try to understand that idea, process, or method of operation" (Atari v. Nintendo 1992). The court continued its support of the fair use defense, explaining, "the Copyright Act permits an individual in rightful possession of a copy of a work to undertake necessary efforts understand the work's ideas, processes and methods of operation. This permission appears in the fair use exception to copyright exclusivity" (Atari v. Nintendo 1992).

However, because Atari had obtained a copy of Nintendo's source code by an illegal means, by misrepresentation to the Copyright Office, the court declined to grant Atari its fair use defense (Atari v. Nintendo 1992). It is important to note that this case points out one of the characteristics of the fair use defense–to be entitled to its protection, one must have "clean hands" to assert the defense. In other words, anyone asserting the defense must show that it acquired a copy of the program or code by an authorized means (Atari v. Nintendo 1992). Atari probably would have won the case had it merely been able to obtain the code by reverse engineering the game cartridges it had lawfully obtained. Unlike the Sega Court, the Atari Court did not assess in any great detail the four factors to be considered in the fair use defense. What is clear, however, is that the Atari Court wholeheartedly sponsored the fair use defense in computer software copyright infringement cases, to give the public greater access to the program's underlying functions and ideas and to promote the advancement of technology. This opinion ultimately was very important because it set forth the reasoning that the Sega Court adopted in its opinion.

10.5 The Fair Use Statutory Defense

10.5.1 History and Changing the Law

Although the Sega and Atari decisions left many aspects of computer software copyright infringement unclear (Soma *et al.*,1994), they did make enormous strides in advancing the law in this area. The four factors used to analyze whether a certain reverse engineering effort should be allowed to assert the fair use defense are not perfectly clear, and one court may interpret the results of applying the factors very differently from another court. The four-step analysis, however, is a step in the right direction to providing for a more certain framework for deciding future cases.

Though many commentators question this case-by-case analysis and ponder whether the United States should develop a *sui generis* system for deciding computer copyright infringement cases similar to that adopted in the European Community, it is important to note that as more and more cases are decided in the future, the law will become clearer as well.

10.5.2 What Do We Know About Proper Reverse Engineering

Until some sort of legislative scheme is developed to deal with the discrepancies in our current system, how should an attorney advise clients who are contemplating reverse engineering a product? Although there is currently no certainty in this area, recent case law has made it prudent to recognize several caveats when undertaking a reverse engineering project. First, attorneys should advise

their clients to obtain an "authorized" copy of software. One of the hurdles that prevented Atari from successfully asserting the fair use defense to its software infringement charge was the fact that it had "unclean hands" because it obtained a copy of Nintendo's source code by misrepresenting to the Copyright Office that it needed the code for litigation in which they were involved. The Federal Circuit advised others that "to invoke the fair use exception, an individual must possess an authorized copy of a literary work" (Atari v. Nintendo 1992). To avoid "tainting" your fair use defense, a computer company should obtain a software program, which it desires to reverse engineer from the open market, just as any other consumer would. Additionally, one must be careful not to violate any licensing agreements signed directly with a computer company. For example, when a company such as Sega licenses other software companies to use its code to produce compatible games for use on the Sega video game system, the company often must promise not to attempt to reverse engineer any of Sega's codes for use in other games. Courts desire to uphold such agreements and will probably view any reverse engineering in such circumstances as tainting a later fair use defense. Remember, the recent trend of court decisions is to accept the fair use defense and allow computer companies to reverse engineer software to advance technology and create market competition for software, but not at the expense of violating a more important public policy–voiding fraud and misrepresentations while encouraging fair play.

Second, attorneys should advise clients to be sure that there is no other means to obtain the information regarding the software other than by reverse engineering the program. The Sega Court specifically pointed out this caveat by stating "we conclude that where disassembly is the only way to gain access to the ideas and functional elements embodied in a copyrighted computer program … disassembly is a fair use of the copyrighted work, as a matter of law" (Atari v. Accolade 1992). Some companies publish certain codes to allow for the development of compatible software. Therefore, if the source code for a particular function or process is obtainable by any other legal means, all of these avenues should be explored before reverse engineering is undertaken.

Third, if reverse engineering is used out of necessity, engineers should be advised to reverse engineer only the portions of the original program that are necessary to decipher the precise functional elements needed for the new program. Courts become weary of reverse engineering efforts that use portions of the original program to the extent that the newly developed software is similar in expression to the original. In Lotus Development Corp. v. Paperback Software International (Copyright Act 1990), the court held that Paperback infringed on Lotus copyrighted software by devising their spreadsheet software to have a look and feel similar to the Lotus popular 1-2-3-spreadsheet program. Rather than copy only the functional elements of the 1-2-3 program, Paperback had essentially copied much of Lotus' aesthetic qualities as well (Copyright Act 1990). It is difficult to say when a court would consider two programs to be too similar, but the Sega Court attempted to give a few examples where reverse engineering may

have resulted in more than necessary portions of software being copied. Unfortunately, however, these examples are somewhat vague and cannot be readily applied to everyday situations with any confidence (Hayes 1992). Thus, to be on the safe side, avoid copying program components related to the expression of the program or its aesthetic qualities.

Fourth, computer companies should be sure to divide their reverse engineering efforts into two groups of engineers/programmers–one group to reverse engineer the computer program and the other group to develop the new software. This method of ensuring "clean hands" is recommended so that if a software developer is later charged with software copyright infringement, the company can produce records to show that the newly developed program was completed without directly copying the original software code. Specifically, this process begins with the first group of programmers reverse engineering the original software into the source code, so that it can be read by human programmers. Next, the first group will explain in a written journal or log the functions of the original program as well as the ideas the program uses, without describing the expressive content of how the program will look to the user. This journal is then given to the second group of programmers who attempt to design a computer program that emulates these same functions and ideas. If the second group cannot communicate directly with the original group of engineers, this helps to insulate the development process from any direct copying of the original program's code. This process should ensure that the end product of the newly developed program will look somewhat different from the original program that was reverse engineered. The new program can then be marketed directly against the original program with little fear of violating the original's copyright protection (McCabe 1994).

Fifth, despite all of the previously mentioned precautions, computer software companies should take steps to conduct research on the product to be reverse engineered to ensure that patent law does not provide protection for the particular process or function that is to be reverse engineered and used in a new program. Keep in mind that although copyright law cannot protect functions and ideas, patent law does. Because of the rigorous requirements that must be met to get a patent on a certain function or process, very few computer programs are patented. Nevertheless, it is important to protect yourself from any patent infringement lawsuits by inquiring into this area before reverse engineering and marketing a new program.

Last, a prudent software developer should consult with an intellectual property attorney to be sure that it will be in compliance with the latest case law. In this rapidly progressing area of law, new cases are being ruled on everyday. With the uncertainty that exists today, the law could shift drastically in a very short period of time.

10.6 Conclusion

Justice O'Connor suggested that there is a major public benefit to reverse engineering in that it encourages inventors and businesses to devise similar products (Bonito v. Thunder 1989). Marketing similar products will encourage the original product's inventors to continue to improve their products to stay ahead of the competition while encouraging healthy competition and price pressures in the open market. It appears as though recent cases such as Sega and Atari have adopted this same viewpoint. Early case law seemed to lean toward overprotection of software copyrights, but the trend now is to promote the primary goal of the Copyright Act: to promote the progress of science and the useful arts.

Favoring free access to the functions and ideas not protected by patent law can only promote further creativity in an industry that is progressing with warp speed. As a result, the fair use defense should be given more weight by the courts in the future to enable U. S. computer companies to compete on a level playing field with overseas competitors, who are subject to fewer restrictive regulations. The Sega and Atari decisions represented great strides in developing the law in this area, but more clarity is needed so that computer software manufacturers have a better idea of what they can and cannot do when developing new programs. Whether this takes the form of new legislation or stronger case law, more clarification will enable U. S. computer companies to compete better in global markets.

For those companies wishing to safeguard their newly developed software, patent law may be the best protection. As mentioned earlier, however, patenting of software entails a rigorous maze of formalities and leaping through many complex hoops, which may render this protection infeasible when time, money, and effort are weighed against benefits. In many instances, by the time a patent on software is received, the software is already outdated. Similarly, the emerging trend of case law discussed above makes copyrighting of software infeasible as well. The most favorable way of protecting software without restricting society's access to new technology in this scientific field is to adopt a *sui generis* system much like that adopted by the European Community. A *sui generis* system could combine the low cost and speed of copyright law with the protection of patent law. By balancing the favorable characteristics of copyright and patent law, software companies should continue to reap the rewards of producing new technology without fear that new programs will be quickly copied by competitors. Society, however, would eventually have access to the new processes and techniques used in new software.

With case law determining whether or not computer software will be protected, the law is lagging behind the computer industry as it sprints toward the future. Ultimately, the law will have to catch up to be useful to those in the industry. The only way to accomplish this is to adopt a *sui generis* system specifically tailored to this area of law, so that all competitors in the industry will know the rules before they play the game.

Chapter 11

Barriers to Adopting Reverse Engineering

Kiran Jude Fernandes, University of York, UK

This chapter presents a study of the barriers to adopting reverse engineering technology. Previous literature suggests that various factors play a role in adopting technology; however, there is little research into the factors affecting the adoption of RE technology in particular. This chapter investigates the forms of barriers that affect the adoption of RE technology in manufacturing firms. A three-phase factor analysis approach (FAA) is used to investigate these "critical" factors.

11.1 Background

With increasing competition from foreign companies during the last decades, more and more U. K. manufacturing companies are being forced to look outside of the United Kingdom to survive. Manufacturing made up one-sixth of the U.K. economy in 2005 and was responsible for around two-thirds of all U. K. exports. It generated around 3.5 million jobs directly and millions more through the supply chain and related services. It was estimated that there were about 4 million manufacturing businesses in the United Kingdom at the start of 2003. The vast majority of these (99%) had less than 50 employees. It is for this reason that the author in this chapter particularly concentrates on analyzing these firms.

The manufacturing sector is a knowledge-intensive and volatile business cluster. Companies operating in such clusters, especially in the private tendering area have to deal with high information needs and display a high degree of complexity. These complex operating conditions faced by manufacturing companies are compounded by additional challenges such as global competition and high labor costs, especially in countries such as the United Kingdom, where the where the cost of labor has more than doubled in the last decade. To be adaptable to these factors, companies have to develop new and innovative products in a very short

period of time, which can often be costly and risky. As shown in previous chapters, RE is one method that can help companies achieve products in a very short period of time and thus gain an advantage over competitors. Ferreira and Alves document the advantages of using RE as a core business strategy within foundry technology (Ferreira and Alves 2003). Even though researchers have shown the advantages of using RE within manufacturing processes and have documented results similar to the findings of Ferreira and Alves, very few companies actually adopt RE as a core technology within their processes. Researchers state that huge capital cost, installation time, and high risk are the main reasons for not adopting RE. Where these technologies have been implemented successfully, organizations in the construction industry have seen savings in total project cost.

Unfortunately, in the area of adopting RE technology, most of the research available focuses on the cost, technical aspects, project size, and the extent of design reuse, and there is little research that considers factors at managerial and strategic levels. Given this knowledge gap, this research aims to study factors that affect the adoption of RE technology within manufacturing companies; especially because costs of new RE systems are low and reliability is high. A three-phase factor analysis approach is proposed as part of the main research agenda. The outcome of the study is based on empirical evidence that can help small manufacturing enterprises (SMEs), which might wish to adopt RE technology by overcoming potential obstacles and hence reduce the risk of failure during implementation. For other industries intending to implement RE technology, the results and findings from this study can be used as a case study for adopting RE in the near future. Furthermore, academia can use the findings of this study as a basis to initiate related studies in the RE area.

11.2 The Research Model

Table 11.1 below summarizes the results from prior research which enterprises must consider while adopting new technological system. These systems could include new information and communications technology (ICT) systems or in this case RE technology. To understand these factors in detail, the author has used a multidimensional model based on organizational, project related, and environmental factors. Therefore, the study will focus on these three dimensions that determine adoption, rather than exclusively on attributes of the technology itself. This would include factors identified by Rogers (1995) and proponents of the technology adoption model (Davis *et al.* 1989; Moon and Kim 2001).

The successful adoption of new technologies, also known as a type of innovation, has been a major contributor to the improved commercial performance of organizations (Utterback 1996). An innovative technology can be defined as "any idea, practice or material artifact perceived as new by the relevant unit of adoption" (Zaltman, 1973). In this case, the technology is RE, and the unit of adoption is the SME. Hence, the innovation under examination is not the fundamental RE technology, but rather its adoption by a particular unit.

Table 11.1. Factors affecting the adoption and implementation of new technology

Researchers	Factors
Hoffman *et al.* (1996); Mitropoulos and Tatum (1999)	Assistance of consultants
Hoffman *et al.* (1996); Maxwell and Westerfield (2002)	Champion in the organization
Hoffman *et al.* (1996); Woiceshyn (2000)	Coordination of organizational resources
Gatignon and Robertson (1989)	Degree of business competition
Chen (2003); Rogers (1995)	Internal needs
Mason *et al.* (1981); Speier and Vankatesh (2002); Karahanna *et al.* (1999); Higgins and Hogan (1999)	Participation of users
Gatignon and Robertson (1989)	Selection of vendors
Galbraith (1956); Bajwa and Lewis (2003)	Size of organization
Teo *et al.* (1997), Srinivasan *et al.* (2002)	Skills of the team
Teo *et al.* (1997); Hambrick and Mason (1984); Mitropoulos and Tatum (1999); O'Neill *et al.*(1998)	Top management support

Wejnert (2002) found that a wide variety of factors influence the diffusion of a technology. These factors, individually and collectively, can have a potentiating or mitigating effect on its diffusion. Their impact can vary, according to the characteristics of the technology, and the context into which it is being implemented. Organizational dimensions include factors such as size of the company (Galbraith 1956; Bajwa and Lewis 2003), top management support (Teo *et al.* 1997; Hambrick and Mason 1984; Mitropoulos and Tatum 1999; O'Neill *et al.* 1998), existence of a champion (Hoffman *et al.* 1996; Maxwell and Westerfield 2002), and internal needs (Chen 2003; Rogers 1995).

The size of an organization is an influential factor on its ability to adopt new technologies, as found by a number of authors (Galbraith 1956; Bajwa and Lewis 2003). Brandyberry (2003) found no link between organizational size and its ability to adopt new technologies, whereas other studies have found that larger organizations adopt more new technologies than smaller ones (Bajwa and Lewis 2003). However, this was only the case with those technologies whose adoption was resource intensive, are complex to use, and require large amounts of technical support.

The role played by senior managers has been identified as a major contributing factor in the behavior of organizations adopting new technologies. In their seminal work on upper echelons, Hambrick and Mason (1984) suggest "organizational outcomes–both strategies and effectiveness–are viewed as reflections of

the values and cognitive bases of powerful actors in the organization". Senior managers take two different approaches to the adoption of new technologies (Mitropoulos and Tatum 1999). These authors found that when technologies were perceived as strategically important, then a rational approach was taken to analyzing new technologies. For nonstrategic technologies, a behavioral process was followed that was characterized by managers focusing on the minimization of the consequences of failure, opportunism, and experimentation. O'Neill *et al.* (1998) established that managers within organizations continued to use strategies (the adoption of a technology could be an example of one) that exhibited levels of performance that were less than favorable. This is due to their being trapped by incentives and the structure of the situation. This again points to managers adopting nonrational approaches to the adoption of new technologies.

Speier and Venkatesh (2002) found that new technologies, which were positively adopted in the short term by employees of organizations, could be rejected in the longer term. This was perceived (by the employees) to be the case because they negatively affected the ability of persons to undertake their professional activities. Karahanna *et al.* (1999) also discovered a change in the pre- and postadoption beliefs and attitudes that affect the adoption of a technology. Potential adopters are highly influenced by social norms, defined as what relevant others (human actors) expect a person to do with respect to a technology, whereas post- adoption users of a technology are influenced only by perceived usefulness and image. Higgins and Hogan (1999) suggested that fostering a sense of user ownership and cross-functional collaboration were both factors that significantly influenced the successful adoption of new technologies by organizations.

With the support and commitment from top management, it is easier to secure the required capital, human resources, and other related internal resources in the adoption and development process. Companies can cope with the unexpected obstacles with ease during the implementation of RE technology by securing top management support and explicitly defining the intended goals and objectives. A champion actively supports and promotes the project and provides information, material resources, and political support. Champions are important to RE as well as to other information and communications technology (ICT) projects because RE is widely recognized as an ICT technology. Maxwell and Westerfield (2002) found that the presence of "opportunistic entrepreneurs", a type of champion, within organizations had a direct effect on the level of use of innovative technologies by that organization.

Chen (2003) points out that "internal needs" are central for implementing information technology in an organization. Rogers (1995) also found that the ability of an innovation to offer "relative advantage" over another solution, with respect to meeting the internal needs of an individual or organization, is a critical factor in whether or not it is adopted. In addition to this, several researchers (Galbraith 1956) concluded that the larger an organization, the more resources and capital that can be allocated to adopt new forms of technology.

Project dimensions include factors such as the skill of the project team (Teo *et al.* 1997; Srinivasan *et al.* 2002), coordination of resources (Hoffman *et al* 1996; Woiceshyn 2000), level of user participation (Mason *et al.* 1981; Speier and Vankatesh 2002; Karahanna *et al.* 1999; Higgins and Hogan 1999), and external help (Gatignon and Robertson 1989). It can be seen that careful evaluation and planning is critical in implementing a RE system because of the expense involved. The skills of the resident IT team have a major influence on the outcome of the RE adoption task. It has been the author's experience that a highly skilled and motivated project implementation team is better equipped to manage and solve the technical problems of a RE installation. Such a team can also ensure better user involvement and training, which is critical to the success of any ICT project. Srinivasan *et al.* (2002) advocate that the ability of an organization to sense and respond to new technological possibilities is a distinct capability that is needed to exploit them successfully. This is, of course, based on the assumption that a thorough needs analysis has already been done before the procurement by the project team. A motivated user is more likely to increase and meet the expectations and requirements of the customer. Hoffmann *et al.*(1996) identify money, people, and time as the three pillars of a successful technology implementation project. These resources can help the project team meet all stated milestones to expectations. Once tasks are identified, the project time line is influenced by the amount of time and people assigned to the implementation task. RE comparatively is a new technology and hence many organizations do not have the necessary skill set to implement this on their own. In such instances, Hoffman (1996) states that assistance from external consultants can be beneficial to the success of the project. Mitropoulos and Tatum (1999) develop this proposition by suggesting that consultants are one of the most reliable sources of information for organizations to access, when seeking to understand the performance of a technology and the potential problems faced in implementing it.

Where organizations make available resources to their employees to fund the adoption of new technologies that have commercial benefits, it has been shown to have a motivating effect on employees. The consequence of this is that a virtuous circle is formed that propagates a detailed and highly incentivized approach to analyzing new technological opportunities (Woiceshyn 2000). The converse of this was also found to be true, *i.e.*, where resources were low, a viscous circle was created with low levels of motivation to adopt new technologies.

The third dimension on environmental factors includes the degree of business competition and the selection of vendors (Gatignon and Robertson 1989). Gatignon (1989) states that there is a strong correlation between the degree of business competition and the adoption of new technology. In addition to this, the selection of vendors is critical to the implementation strategy. Companies typically are more comfortable using technology from reliable vendors rather than new or nonapproved suppliers. The outcome of this literature forms the basis of the research hypothesis and is listed in Table 11.2 and shown in Figure 11.1:
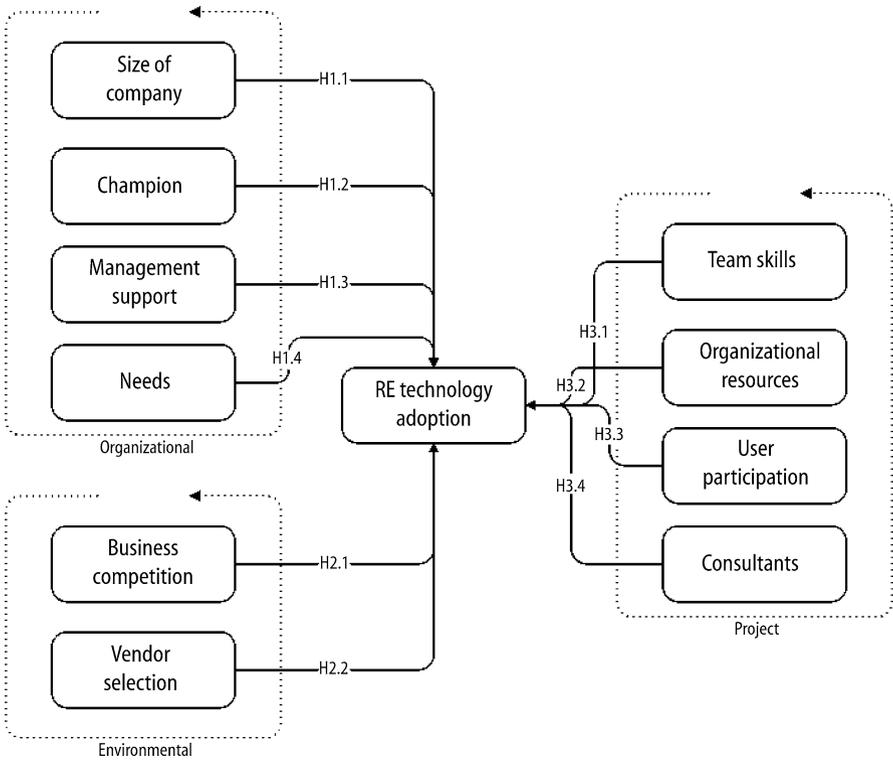
Figure 11.1. Research model

Table 11.2. Research hypothesis

Dimension	Hypothesis	Hypothesis details
Organizational	H1.1	The size of the organization affects the adoption of RE technology.
	H1.2	The existence of champions affects the adoption of RE technology.
	H1.3	Top management support affects the adoption of RE technology.
	H1.4	The organization's internal needs affect the adoption of RE technology.
Environmental	H2.1	The degree of business competition affects the adoption of RE technology.
	H2.2	The selection of the vendor affects the adoption of RE technology.

Table 11.2. (continued)

Dimension	Hypothesis	Hypothesis details
Project	H3.1	The skill of the project team affects the adoption of RE technology.
	H3.2	The coordination of organizational resources affects the adoption of RE technology.
	H3.3	The participation of end users affects the adoption of RE technology.
	H3.4	Assistance from consultants affects the adoption of RE technology.

11.3 Research Methodology

The factor analysis approach reported in this chapter employs a three-phased approach to study the critical factors that affect the adoption and/or implementation of RE technology within SMEs. The first phase, the factor determination phase, is concerned with the generation of factors, which might affect the adoption of RE technology. This phase, described in the previous section, depends on
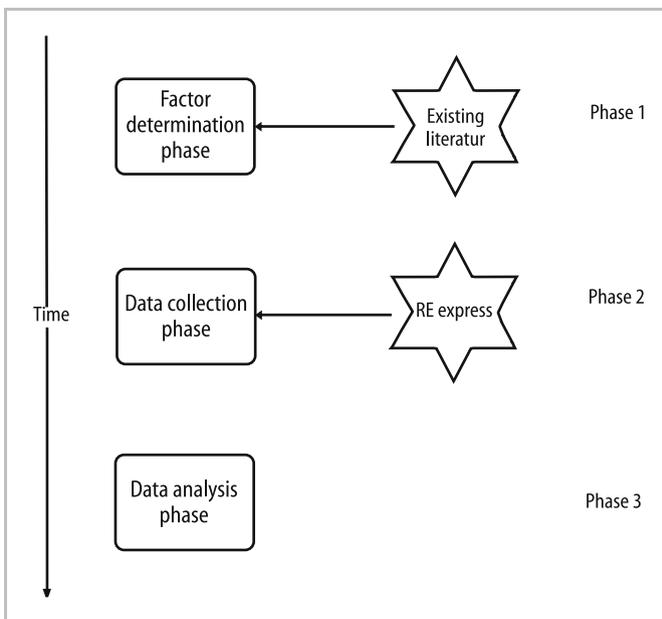
Figure 11.2. Factor analysis approach

existing literature. The outcome of this phase is the research model shown in Figure 11.1. The data collection phase is the second phase of the factor analysis approach. Data from leading experts were collected using a questionnaire to validate the research model. More detail on this phase is given in the next section. The final phase of this approach is the analysis phase, where the data collected were used to test the hypothesis established in the research model. Figure 11.2 shows the factor analysis approach.

11.4 Factor Analysis Approach

As can be seen from Figure 11.2, this study proposes a three-phased factor analysis approach to determine the critical factors that effects the adoption of RE technologies within SMEs. The contributions of this research are at both theoretical and practical levels. At a theoretical level, this study contributes a new methodology to elicit and validate critical factors for adopting RE technologies. This method can be easily adapted to any general factor validation process. This method relies both on strong theoretical foundations (factor determination phase) and empirical evidence (data collection phase) and hence can be considered a semiformal method, as opposed to traditional survey techniques. On a practical level, this study contributes 'knowledge of RE' to companies who are in the process of implementing RE technologies within their operations. This can also form a good reference for other construction companies to establish and develop strategies for implementing RE within key business processes. The three phases of this approach are as follows:

11.4.1 Factor Determination Phase

Section 11.2 of this study explained in detail the key factors that have been determined from the literature. The authors used a multidimensional model to represent the organizational, environmental, and project dimensional factors. Using such a dimensional model allows clustering the factors under appropriate dimensions, and also facilitates a logical data collection and analysis study. The 10 factors cover most of the major aspects in a SME from three angles or dimensions. The outcome of this phase is a multidimensional model as shown in Figure 11.1.

11.4.2 Data Collection

The outcome of the factor determination phase was a hypothetical research model. This multidimensional model lists factors that affect the adoption of RE technology within manufacturing SMEs. This study adopted a qualitative approach using semistructured interviews with experts in the field of RE. The data

collection was carried out by the author, who is a qualified CMM operator, and both editors analyzed the data. Face-to-face interviews approximately 50 minutes long were conducted with each expert in an off-site setting. The number of interviewees was determined during data collection at the point where no new insights were being obtained from interviewees and saturation was reached. Power (2003) referred to the difficulty of conducting fieldwork in auditing, and in an analysis of a number of field studies suggested that many of these studies were actually "research by stealth" which had been conducted on "the margins of more orthodox work" (p. 380) to legitimate the research. In carrying out this research and in communications with the firms, the researchers discussed the study openly and did not disguise the research topic or link it to research on any other topic. The participating experts accepted the need for research in this area. Semi-structured interviews were conducted using an interview guide, which was divided into a number of sections, each section containing general questions and potential probes, and was based on examples of interview guides in the literature (*e.g.*, Lillis 1999; Patton 1990). The interview guide approach enables the interviewer to have a list of areas for questioning but at the same time allows probing of questions in more depth depending on the answers provided by the interviewee. Patton (1990) described this approach as, … a process in which the interviewer focuses her questions on some limited number of points. She may range quite widely around a point, but this would be done only as a means of getting the required information on that particular point.

The interview guide approach "… helps minimize bias through the prespecification of non-directive questions and probes" (Lillis 1999, p. 87) and the preconditioning reduces the tendency to resort to unplanned, nonneutral probes during the interview. Patton (1990) recommended that … good questions should, at a minimum, be open-ended, neutral, singular, and clear (p. 295). These principles were used in the design of questions.

Before commencing, each interviewee was assured of the confidentiality of the study, and their permission was obtained to record the interview. Because of the sensitivity of some of the interview questions, it was emphasized that no other individual would have access to the tapes and that the anonymity of participants would be fully respected. All interviewees gave permission to tape the interviews and no interviewee asked for the tape to be turned off at any stage. A relationship of trust was established by the common professional ground and the assurance of confidentiality. Throughout the interviews, the researcher found no evidence that the questions were perceived as overly confrontational in any way, and interviewees seemed relaxed and willing to answer questions honestly.

While recognizing that no single research method can be truly objective, it is important to document all the steps taken to increase objectivity as far as possible and to be constantly alert for subjectivity at data collection and analysis stages (Patton 1990). Previous field research has been criticized for failing to attend to such research criteria as validity and reliability. Several steps were taken to limit bias and increase objectivity both during the interview and in

analyzing the interview data. First, the interview guide was used to ensure consistent and complete coverage of all the themes in each interview. Second, as well as taping the interviews, notes were also taken during the interview to keep a record of important points that needed clarification or that appeared to conflict with other points. Following each interview, notes were made by the researcher (as recommended by Patton 1990) on such issues as the level of rapport built up, interviewee's reaction to questions, initial impressions, and any other points about the context of the interview which the researcher considered should be noted. Third, transcripts of the interviews were reviewed by a colleague (who was not involved in the research) with particular emphasis on objectivity and freedom from bias in questions. Fourth, to limit bias in data analysis, a structured analytical method was used with both researchers involved in analyzing the data.

11.5 Findings

This section will seek to explain why some of the variables have an impact on a company's adoption of RE, and others not. Some prior studies have indicated that the larger the organization, the more resources and capital can be secured to adopt the new technology. This study focused on a single industry. The results indicate that the size of the company does not affect the decision whether or not to adopt RE technology within core processes. There are a number of factors, located within the design of the study, which could explain the results generated. First, the absence of factors found by Bajwa and Lewis (2003) that contribute to large organizations being more successful at adopting new technologies than smaller ones (resource intensity, complexity of use, and the requirement of a large amount of technical support) could account for the role of size. Moreover, it reflects the diversity of opinion within the literature in this area, which is polarized regarding this matter. Another factor that needs to be considered when examining these results is the fact that 80% of the sample was pre-, rather than postadoption of this technology. For example, the importance of an integrated project team may come to light only during the actual process of implementation, when it would be necessary to address the requirements of multiple internal stakeholders. It could also reflect Karahanna *et al.*'s (1999) observations that the criteria of managers, pre- and postadoption, undergo a change. Preadopters are highly influenced by social norms, defined as what relevant others (human actors) expect a person to do with respect to a technology, whereas postadoption users of a technology are influenced only by perceived usefulness and image. Hence, the views of other senior managers, or members of external communities of practice, could have a higher weighting placed on their views than those of a cross-functional team.

The characteristics of the technology itself should also be considered a factor influencing the results. For example, the RE technology could possess a low level of what Rogers (1995) refers to as compatibility. This can be defined as "the degree to which an innovation is perceived as consistent with the existing values, past experiences, and needs of potential adopters". This would mean that there would be a minimum need to consult with multiple stakeholders, as represented in a project team. This would reflect the findings that the project team was not a contributing factor to the adoption of this technology. Other factors in Roger's framework could explain other parts of the findings. For example, a low level of complexity, defined as "the degree to which an innovation is perceived as relatively difficult to understand and use", could account for the lack of a perceived need for consultants as actors in the adoption process.

Champions in an organization can appreciate the contribution of adopting new technology and can actively persuade the associates, subordinates, and top management as to the ideas and benefits of the new technology. This study concurs with the theoretical research model and concludes that a champion is essential to adopting RE within the targeted construction sector.

In many researches, top management support has been a key factor affecting the adoption of new technology. Empirical evidence in this study states that top management support can help to secure the capital, labor, and related resources, which determine the success of implementing RE within an organization. In addition, this research clearly states that the internal needs of a company are a key factor affecting the company's adoption of new RE technology.

Enterprises have always tried hard to raise their competitive advantage by adopting new technologies within processes. The findings of this study reiterate this fact. Prior research supports the importance of the role played by members of the project team in adopting new technology. However, this study shows that there is not sufficient evidence to prove this fact. The reason for this might be due to the cost of the RE systems.

Because the cost of RE systems is quite high, the decision probably remains with the financial experts, rather than the technical experts of a company. Such complex projects are time-consuming, require huge amounts of capital, and often have a high degree of risk. The study indicates that coordination of resources is a key contributor to the successful implementation of RE technology within a company.

The participation of end users in the adoption of a process can assist in better defining needs explicitly and increasing the opportunities of having a successful technology adoption. This research hypothesis is verified by the outcomes of this empirical study. The purpose of employing external consultants is to acquire their professional experience and knowledge, which can be used to facilitate the implementation and adoption of the technology. This study finds that having assistance from external consultants is not an important factor in the adoption of RE technology within the targeted audience.

11.6 Conclusions and Recommendations for Further Research

With the sudden rise in the use of RE technology in recent years, users must be careful that this technology might not be a perfect solution for all types of reverse engineering activities. This study provides the user with many adoption-related factors, which should be carefully considered before adoption attempts are made. The results of this study indicate an exhaustive list of factors in the organizational, environmental, and project dimensions, which should be taken into account. Key areas that require particular attention are a champion within a company, top management support, internal needs, degree of business competition, coordination of organizational resources, and participation of users.

Color Section
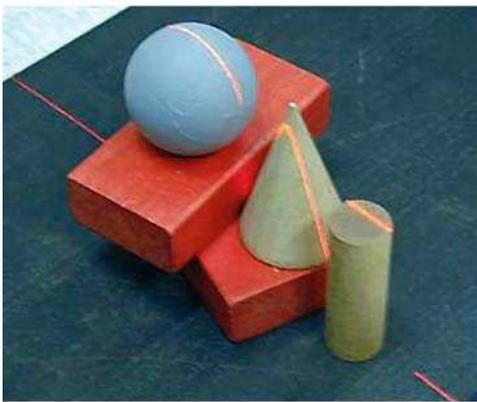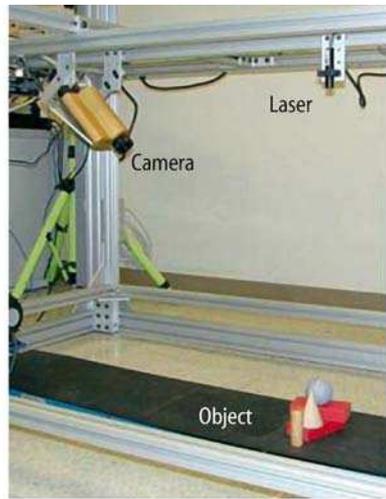
Figure 2.10. Examples of a sheet-of-light system. (**a**) The laser sheet casts a line on the objects under measurement. (**b**) This system is the IVP Ranger Scanner.
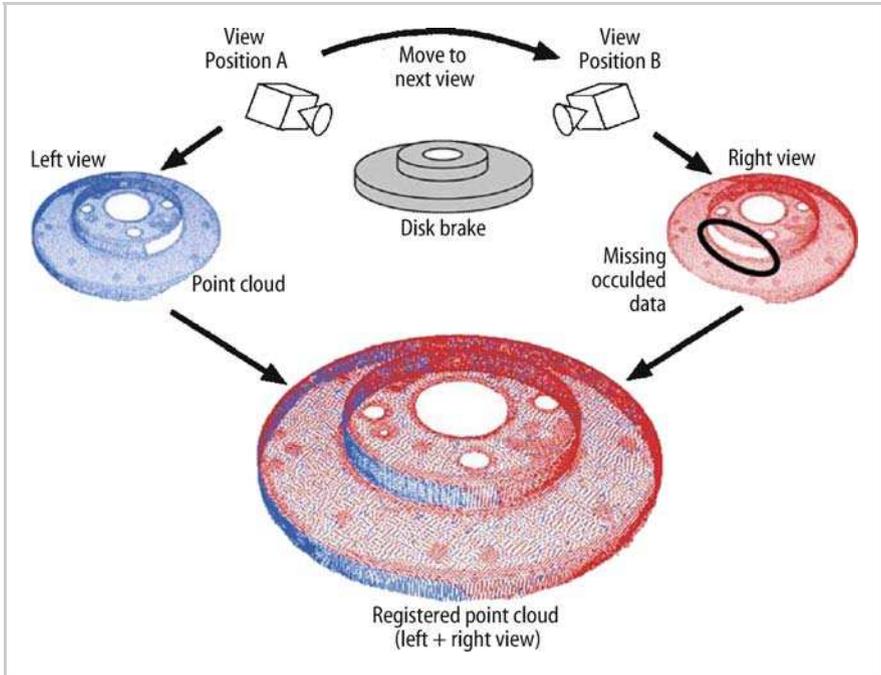
Figure 2.14. Multiple view scanning and registration overcome occlusions. Part of the brake is excluded in each of the views above. When these views are registered together, one view fills the occluded regions of the other view.
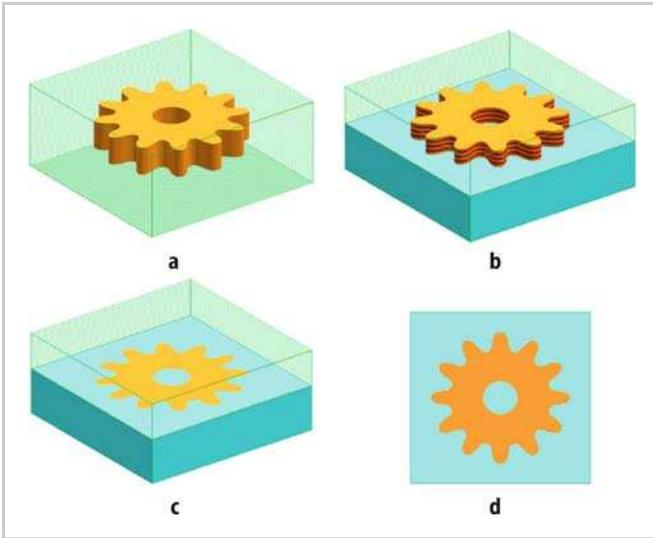
Figure 3.11. A procedure for data acquisition in a destructive RE system. (**a**) The part and matrix combination is embedded in a contrasting colored plastic matrix. (**b**, **c**) The part is machined layer-by-layer to expose the cross-sectional images. (**d**) The newly exposed surface of the part.
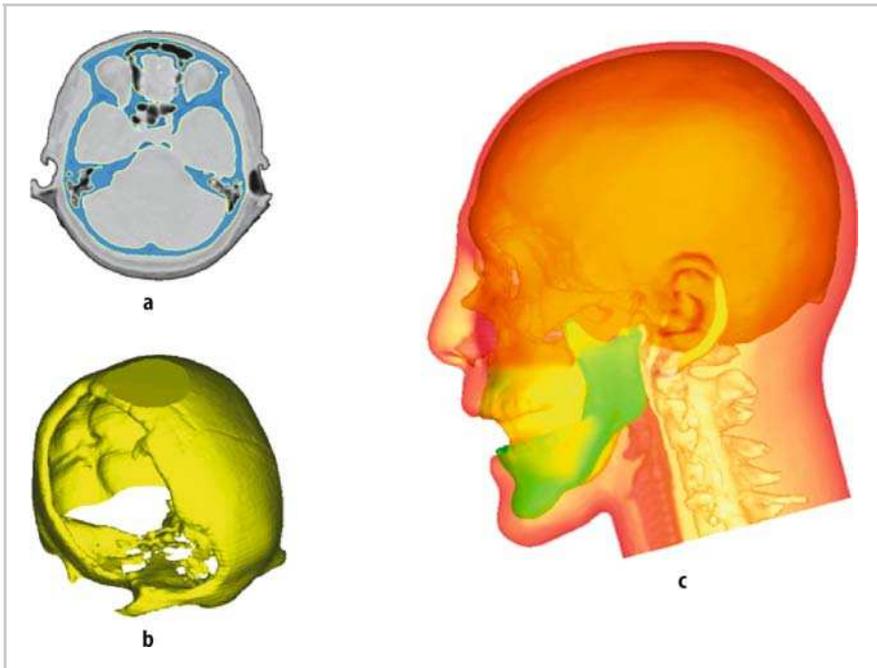
Figure 3.14. (**a**) A skull defined by the thresholding technique. (**b**, **c**) 3-D models of the cranial defect skull (**b**), and other anatomical structures of the head (**c**).
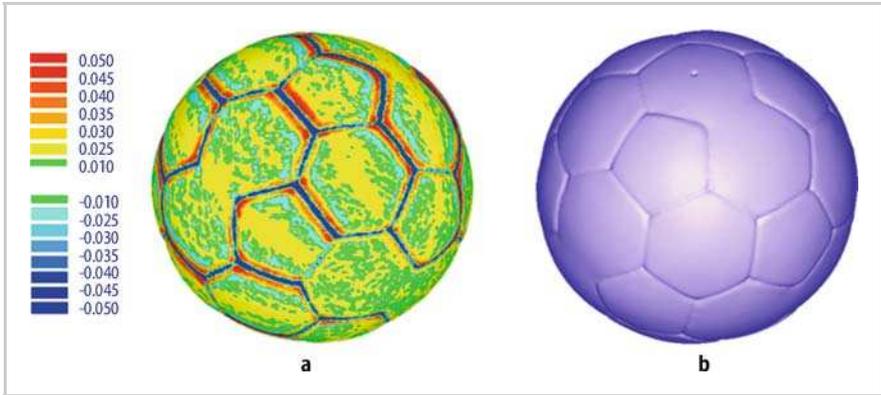
Figure 3.18. An error map (**a**) between an original polygon football model and (**b**) a resulting NURBS model.
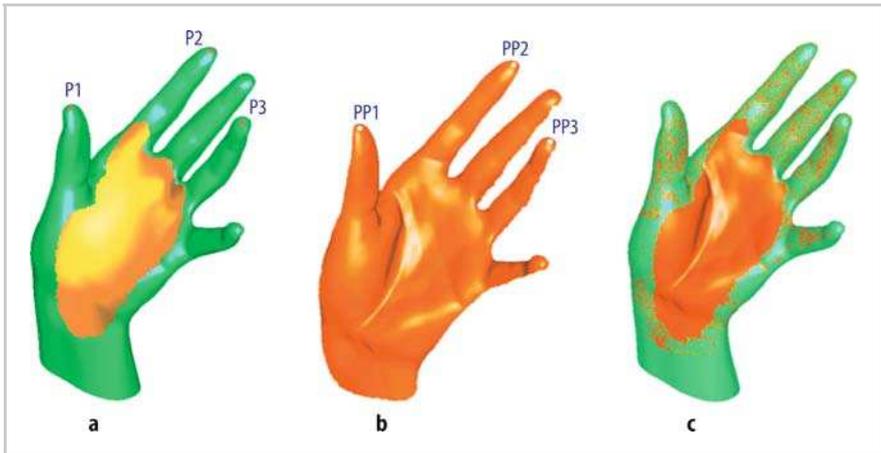
Figure 3.19. (**a**) The first point cloud with three landmark points (P1, P2, and P3). (**b**) The second point cloud with three landmark points (PP1, PP2, and PP3) that correspond to those of the first point cloud. (**c**) The resulting point cloud after registration.

Figure 5.2. The basic RP process (FDM)
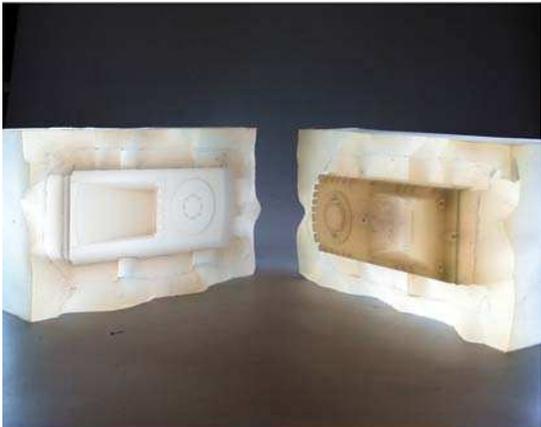
Figure 5.17. Vacuum casting of a talking barcode scanner
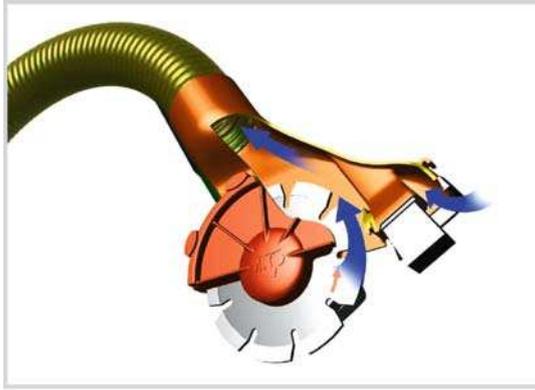
Figure 5.18. A Typical RM application
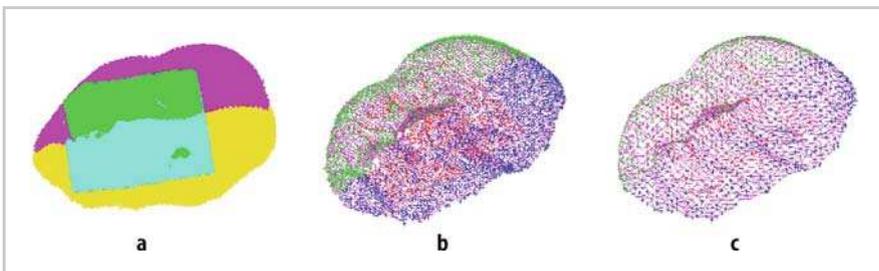
Figure 6.1. The point cloud data of a face mask and its triangular meshes. (**a**) Cloud data set. (**b**) Triangular mesh model with 5410 vertices. (**c**) Triangular mesh model with 1441 vertices. Reprinted from *Computer-Aided Design,* Vol 33, Sun W, Bradley C, Zhang Y, Loh H, Cloud data modelling employing a unified, non-redundant triangular mesh, pp. 188–189, Copyright (2001) with permission from Elsevier.
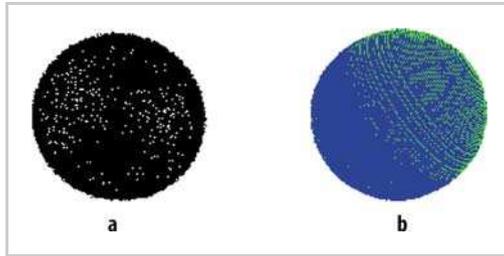
Figure 6.8. The original cloud data and the RP model in case study one. (**a**) The original cloud data. (**b**) The RP model. Reprinted from *Computer-Aided Design,* Vol 36, Wu Y, Wong Y, Loh H, Zhang Y., Modelling cloud data using an adaptive slicing approach, pp. 232–239, Copyright (2004) with permission from Elsevier.

Figure 7.4. The completed digital model of the SB2 engine block with rotating assembly. Copyright © Richard Childress Racing Enterprises, Inc., Welcome, NC, USA, www.rcrracing.com. Reproduced with permission.

Figure 7.8. Visualized results from CFD analysis generated by CEI's EnSight software. Copyright © Advantage CFD, Brackley, Northamptonshire, UK, www.advantage-cfd.co.uk. Reproduced with permission.

Figure 7.12. Color maps visualize the deviation between the scanned sheet metal part and its corresponding reference CAD model

Figure 7.13. An HTML report of a sheet metal part and a PDF report of an intake turbine
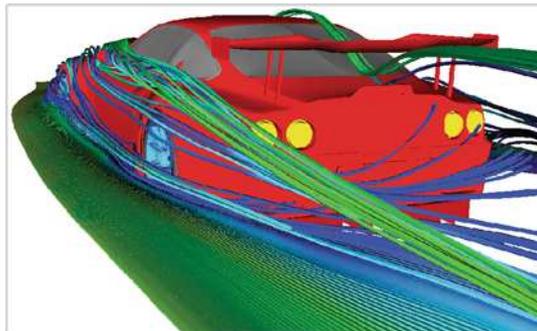
Figure 8.8. X-38 model
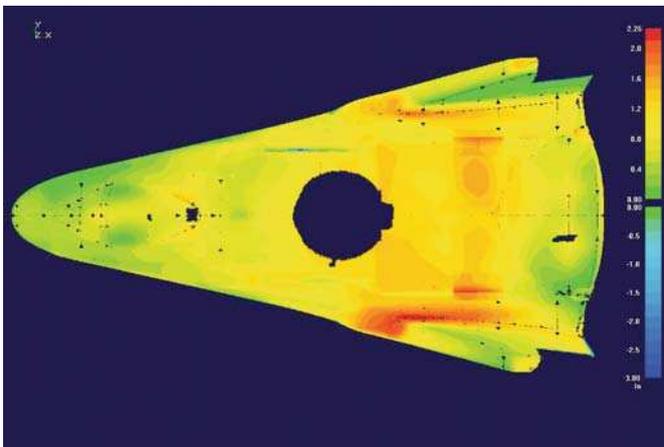
Figure 8.16. Top view of as-built compared to as-designed. Copyright © Aerospace Manufacturing Technologies, Inc., Arlington, WA, USA, http://www.amtnw.com/. Reproduced with permission.
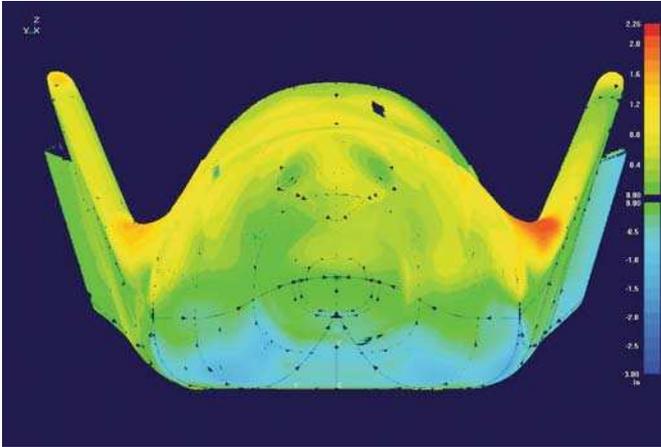
Figure 8.17. Rear view of as-built compared to as-designed. Copyright © Aerospace Manufacturing Technologies, Inc., Arlington, WA, USA, http://www.amtnw.com/. Reproduced with permission.
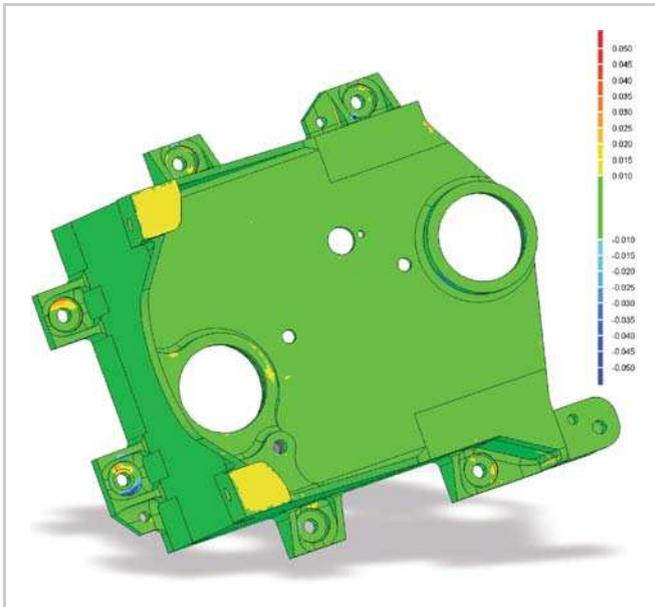
Figure 8.21. Comparison model detailing deviations between the CAD model and the scanned part
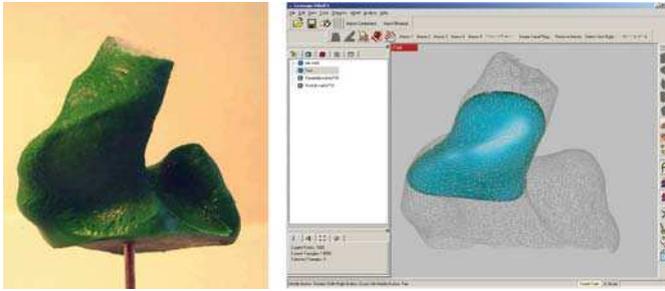
Figure 9.8. At left, an ear impression. At right, the digital shell created using Geomagic eShell software, a specialized reverse engineering application. Copyright © GN ReSound Corp., Bloomington, MN, USA, http://205.242.218.80%20/home.aspx. Reproduced with permission.
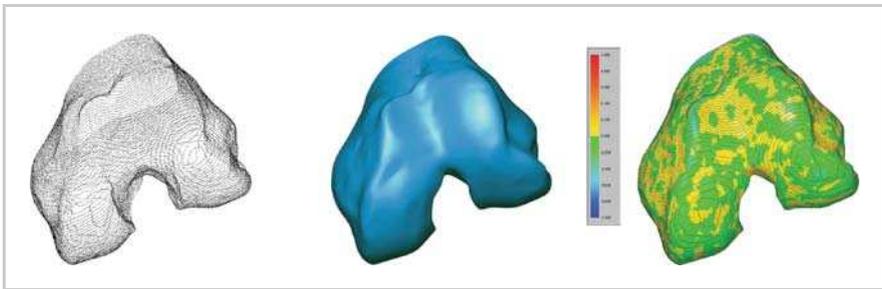
Figure 9.11. Scanned point cloud data, a polygonal model, and a model showing stress analysis. Copyright © University of Florida, Computational Biomechanics Laboratory, Gainesville, FL, USA, http://eng.ufl.edu. Reproduced with permission.

References

1. Abella R, Daschbach J, McNichols R(1994) Reverse engineering applications. Comput Ind Eng 26 (2):381–385

2. Akira H (2001) Microfocus X-ray computed tomography and its industrial applications. Anal Sci. Jpn Soc Anal Chem 17:123–125.

3. Alain B (1999) A review of state-of-the-art reverse engineering. Proc TCT (Time Compression Technol) Eur Conf 1999, Nottingham Royal Moat House, UK, pp. 177–188

4. ASTM (1992). Standard guide for computed tomography (CT) imaging, ASTM, Philadelphia

5. Atari Games Corp. v. Nintendo of America, Inc. (Fed. Cir. 1992) 975 F.2d, pp. 832–876

6. Bajwa D, Lewis L (2003) Does size matter? An investigation of collaborative information technology adoption by US firms, J InformTechnol Theory Appl 5 (1):29–46.

7. Baker v. Selden. 101 US, pp. 99, 1879

8. Beall JL, Gordon IT, Gournay JP, Kirkland BL, White JV (1996) Analysis of porosity in lower Ismay phylloid algal packstone using high-resolution computed X-ray tomography. Am Assoc Petrol Geol Annu Meet Abstr Vol. 5, No. 13.

9. Behrens B, Levary R (1998) Some practical legal aspects of software reverse engineering. Commun ACM 41 (2):27–29

10. Bellian JA, Kerans C, Jennette DC (2005) Digital outcrop models: Applications of terrestrial scanning lidar technology in stratigraphic modelling, J Sediment Res 75 (2):166–176.

11. Bernardini F, Bajaj CL, Chen J, Schikore DR (1999) Automatic reconstruction of 3D CAD models from digital scans. Int J Comp Geom Appl 9 (4&5):327–369

12. Bichsel M, Pentland AP (1992) A simple algorithm for shape from shading. Proc IEEE CVPR Conf Champaign, Ill, pp. 459–465.

13. Böhler W, Heinz G, Marbs (2001) The potential of non-contact close range laser scanners for cultural heritage recording, CIPA Int Symp Proc, Potsdam, Germany

14. Bonito Boats, Inc. v. Thunder Craft Boats, Inc. (1989) 489 US., p 141

15. Bradley C, Vickers G (1993) Free form surface reconstruction for machine vision rapid prototyping. Opt Eng 32: 2191–2200

16. Brandyberry A (2003) Determinants of adoption for organizational innovations approaching saturation, Eur J Innovation Manage 6 (3):150–158

17. Bremer A (2003) Technische Productinformatie 2. Het snel vervaardigen van modellen (Rapid Prototyping). Delftse Universitaire Pers, Delft.

18   Brooks RA, Di Chiro G (1976) Principles of computer assisted tomography (CAT) in radiographic and radioisotopic imaging. Phys Med Biol 21 (5):689–732

19   Bruno D, Nick B (2004) An overview of avalanche photodiodes and pulsed lasers as they are used in 3D laser radar type applications. Proc SPIE. 5434:187–195

20   Campbell RJ, Flynn PJ (2001) A survey of free-form object representation and recognition techniques. Comput Vision Image Understanding 81(2):166–210

21   Caspi D, Kiryati N, Shamir J (1998) Range imaging with adaptive color structured light. IEEE T Pattern Anal 20 (5):470–480

22   Chang HL, Charles AC, Avi C (2000) Does face recognition rely on encoding of 3-D surface? Examining the role of shape-from-shading and shape-from-stereo. Perception 9 (6):729–743

23   Chen C, Hung Y, Chiang C, Wu J (1997) Range data acquisition using colour structured lighting and stereo vision. Image Vision Comput 15:445–456

24   Chen CH, Kak AC (1987) Modelling and calibration of a structured light scanner for 3-D robot vision. Proc IEEE Int Conf Robot Automat, Raleigh, NC, pp. 807–815

25   Chen M (2003) Factors affecting the adoption and diffusion of XML and Web services standards for E-business systems, Int J Human Comput Stud 58 (3):259–279

26   Chen YH, Liu C (1997) Robust segmentation of CMM data based on NURBS. Int J Adv Manuf Technol 13:530–534

27   Chivate P, Jablokow A (1993) Solid-model generation from measured point data. Comput Aided Des 25: 587–600

28   Choi B, Shin H, Yoon Y, Lee J (1998) Triangulation of scattered data in 3D space. Comput Aided Des 20: 239–248

29   Cifuentes C (2001) Reverse engineering and the computing profession. Computer 34 (12):136–138

30   Cifuentes C (1999) The impact of copyright on the development of cutting-edge reverse engineering technology. Proc 6th Working Conf Reverse Eng, Atlanta, GA. IEEE-CS Press, pp. 66–76

31   Cignoni P, Montani C, Scopigno R (1998) A fast divide and conquer Delaunay triangulation algorithm in E^d. Comput Aided Des 30:333–341

32   Copyright Act 17 U.S.C. (1988) 102 (b), p 107

33   Copyright Act 740 F.Supp. (D. Mass. 1990), pp. 37–39

34   Copyright Act 975 F.2d (Fed. Cir. 1992), p 832

35   Copyright Act of 1909. ch 320, 35 Sat., p 1075

36   Creath K, Wyant JC (1992) Moiré and fringe projection techniques In: Malacara D (ed) Optical shop testing, 2nd edn. Wiley, New York

37   Curless B, Levoy M (1996) A volumetric method for building complex models from range images. Proc ACM SIGGRAPH, pp. 303–312

38   David FR (2001) An introduction to NURBS: With historical perspective. Morgan Kaufmann, San Francisco:

39   Davis F, Bagozzi R, Warshaw P (1989) User acceptance of computer technology: A comparison of two theoretical models. Manage Sci 35 (8):436–451

40   Davis GG III (1993) Scope of protection of computer-based works: Reverse engineering, clean rooms and de-compilation 370 PLI/PAT, pp. 115

41   DICOM standard–Digital imaging and communications in medicine: http://medical.nema.org

42   Dolenc A, Mäkelä I (1994) Slicing procedures for layered manufacturing techniques. Comput Aided Des 26:119–126

43   Donald G, Mitchell C, Cohen MS (2004) MRI principles. Elsevier Science, Philadelphia, PA

44   Dupuis P, Oliensis J (1992) Direct method for reconstructing shape from shading. IEEE Proc. Comput Vision Pattern Recognition pp. 453–458

45   Eck M, Hoppe H (1996) Automatic reconstruction of B-splines surfaces of arbitrary topological type. Proc ACM SIGGRAPH, pp. 325–334

46   Fang T, Piegl L (1995) Delaunay triangulation in three dimensions. IEEE Comput Graphics Appl 15:62–69.

47   Faugeras O (1993) Three dimensional computer vision–a geometric viewpoint. MIT Press, Cambridge, MA

48   Feist Publications, Inc. v. Rural Tel. Serv. Corporation (1991) 499 US., pp. 340–349

49   Ferreira J, Alves N (2003) Integration of reverse engineering and rapid tooling in foundry technology. J Mater Process Tech 142 (2):374–382

50   Finkel E (1991) Copyright protection for computer software in the nineties. Santa Clara Comput High Tech Law J 7:202

51   Friedman D, Landes W, Posner R (1991) Some economics of trade secret law. J Econ Perspectives 5 (1):61–72

52   Galbraith J (1956) American capitalism (Rev. edn.). Houghton Mifflin, Boston

53   Galt A (Renishaw plc) (1997) Scanning technology–the missing link. Prototyping Technol Int, pp. 121–124.

54   Gatignon H, Robertson T (1989) Technology diffusion: An empirical test of competitive effects, J Marketing 53 (1):35–49

55   Gondrom S, Schropfer S (1999) Digital computed laminography and tomosynthesis – functional principles and industrial applications. In: Int. Symp. Comput Tomography Ind Appl Image Process Radiol, Berlin

56   Gondrom S., Zhou I, Maisl M, Reiter H, Kroning M, Arnold W (1999) X-ray computed laminography: An approach of computed tomography for applications with limited access. Nucl Eng Des 190:141–147

57   Grin P, Narasimhan L, Yee S (1992) Generation of uniquely encoded light patterns for range data acquisition, Pattern Recognition 25(6):609–616

58   Guo B (1997) Surface reconstruction: From points to splines, Comput Aided Des 29(4):269–277

59   Hambrick D, Mason P (1984) Upper echelons: The organization as a reflection of its senior managers, Acad Manage Rev 9(2):193–206

60   Hamilton D (1993) US. Criticizes Japan on panel software. Wall Street Journal, B5, November 10, 1993

61   Hayes D (1992) The legality of disassembly of computer programs. Comput/Law J.1:14

62   Hieu L, Vander Sloten J, Bohez E, Khanh L, Binh P, Oris P, Toshev Y, Zlatov Y (2005) Medical rapid prototyping applications and methods, Assembly Automat J 25(4):284–292

63   Hieu LC, Bohez E, Vander Sloten J, Oris P, Phien HN, Vatcharaporn E, Binh PH (2002) Design and manufacturing of cranioplasty implants by 3-axis CNC milling. Technol Health Care 10(5):413–423

64   Hieu LC, Bohez E, Vander Sloten J, Phien HN, Vatcharaporn E, Binh PH, and Oris P (2003) Design for medical rapid prototyping of cranioplasty implants. Rapid Prototyping J 9(3):175–186

65   Higgins S, Hogan P (1999) Internal diffusion of high technology industrial innovations: An empirical study. J Bus Ind Marketing 14(1):61–75

66   Hoff W, Ahuja N 1989) Surfaces from stereo: Integrating feature matching, disparity estimation, and contour detection. IEEE T Pattern Anal 11(2):121–136

67   Hoffman D, Novak T (1996) Marketing in hypermedia computer-mediated environments: Conceptual foundations. J Marketing.60:50–68,

68   Hoffman R, Jain K (1987) Segmentation and classification of range images. IEEE T Pattern Anal 9:608–620

69   Hoppe H, DeRose T, Duchamp T, McDonald J, Stuetzle W (1992) Surface reconstruction from unorganised points. Comput Graphics (Proc SIGGRAPH), pp. 71–78

70   Horn B, Hilden H, Negahdaripour S (1998) Closed-form solution of absolute orientation using orthonormal matrices. J Opt Soc Am A (Opt Image Sci) 5(7):1127–1135

71   Horn BKP (1990) Height and gradient from shading. Int J Comput Vision 5(1):

72   Horn BKP, Brooks MJ (1989) Shape from shading. MIT Press, Cambridge, MA

73   Hsieh Y, Drake S, Riesenfeld R (1993) Reconstruction of sculptured surfaces using coordinate measuring machines. Adv Des Automat, 2:35–46

74   Hsu DY (1998) Spatial error analysis: A unified application-oriented treatment. Wiley-IEEE Press, Piscataway, NJ

75   Integrated Vision Products (2000) MAPP 2500 Ranger User Manual, Integrated Vision Products, Sweden

76   Iovea M, Marinescu A, Rizescu C, Georgesc GH (1994) Using computerized tomography in non-destructive testing of the solid materials. Annu Symp Inst Solid Mech, Romanian Academy, Bucharest, Romania, pp. 357–364

77   Iovea M, Marinescu A, Chitescu P, Sava T (1994) Three-dimensional method of representation in industrial computerized tomography. The 6th Joint PS-APS Int Conf Phys Comput, Lugano, Switzerland

78   Joaquim S, Jordi P, Joan B (2004) Pattern codification strategies in structured light systems. Pattern Recognition 37(4):827–849

79   Johns RA, Steude JS, Castanier LM, Roberts PV (1993) Nondestructive measurements of fracture aperture in crystalline rock cores using X-ray computed tomography. J Geophys Res 98:1889–1900

80   Johnson A, Herbert M (1990) Three-dimensional map generation from side-scan sonar images. J Energy Resourc Technol Trans ASMET 112:96–102.

81   Johnson AE, Hebert M (1996) Seafloor map generation for autonomous underwater vehicle navigation. Auton Robot 3(2&3):145–168

82   Kafri O, Glatt I (1978) The physics of moiré metrology. Wiley, New York

83   Kak AC, Malcolm S (2001) Principles of computerized tomographic imaging. Society of Industrial and Applied Mathematics, Philadelphia, PA

84   Karahanna E, Straub D, Chervany N (1999) Information technology adoption across time: A cross-sectional comparison of pre-adoption and post-adoption beliefs, MIS Q 23(2):183–212

85   Ketcham RA, Carlson WD (2001) Acquisition, optimization and interpretation of X-ray computed tomographic imagery: Applications to the geosciences. Comput Geosci, 27:381–400

86   Kewanee Oil Corporation v. Bicron Corp. (1974) 416 US, pp. 470–476

87   Kimmel R, Siddiqi K, Kimia BB, Bruckstein AM (1995) Shape from shading–level set propagation and viscosity solutions. Int J Comput Vision 16 (2):107–133

88   Kiyasu S, Hoshino H, Yano K, Fujimura S (1995) Measurement of the 3-D shape of specular polyhedrons using an m-array coded light source. IEEE Trans Instrum Meas 44, (3):775–778

89   Klette R, Schlüns K, Koschan A (1998) Computer vision–three-dimensional data from images, Springer, Singapore

90   Kulkarni P, Dutta D (1995) Adaptive slicing for parametrisable surfaces for layered manufacturing. In: Proc ASME Des Automat Conf, Boston, MA, pp. 211–217

91   Lai J, Lu C (1996) Reverse engineering of composite sculptured surfaces. Int J Adv Manuf Technol 12:180–189

92   Lawson C (1977) Software for C^1 surface interpolation Mathematical Software III. Academic, Oxford, UK

93   Lee KM, Kuo CCJ (1993) Shape from shading with a linear triangular element surface model. IEEE T Pattern Anal 15(8):815–822

94   Lee I (2000) Curve reconstruction from unorganised points. Comput Aided Geom D 17:161–177

95   Lewis Galoob Toys, Inc. v. Nintendo of America, Inc. (9th Cir. 1992) 780 F. Supp. 1283 (N.D. Cal. 1991), aff'd. 964 F. 2d, p 965

96   Lichti DD, Harvey BR (2002) The effects of reflecting surface properties on time-of-flight laser scanner measurements. Proc 95th CIG Annu Geomatics Conf, Ottawa, Can

97   Lillis A (1999) A framework for the analysis of interview data from multiple field research sites. Account Fin, 39 (1):79–105

98   Lin R (2001) Adaptive slicing for rapid prototyping. In: Gibson I (ed) Software solutions for rapid prototyping. Professional Engineering, London and Bury St Edmunds, UK, pp. 129–154

99   Liu G, Wong Y, Zhang Y, Loh H (2002) Error-based segmentation of cloud data for direct rapid prototyping. Comput Aided Des 35:633–645

100   Mark AB, Richard CS (2003) MRI: Basic principles and applications, Wiley, Wilmington, DE

101   Mason R, Mitroff I (1981) Challenging strategic planning assumptions: Theory, cases, and techniques, Wiley, New York

102   Maxwell J, Westerfield D (2002) Technology entrepreneurism characteristics related to the adoption of innovative technology. SAM Adv Manage J 67(1):9–21

103   McCabe P (1994) Reverse engineering computer software: A trap for the unwary? Comput Law Assoc 9:1.

104   Michael WV, Robert HK, Donald EG, Robert LB, Larson JW, Neil PS (1985) Three-dimensional surface reconstruction for industrial computed tomography. Appl Opt 24(23):4075–4082

105   Mitropoulos P, Tatum C (1999) Technology adoption decisions in construction organizations. J Constr Eng M, September/October:330–338

106   Moon J, Kim Y (2001) Extending the TAM for a World-Wide-Web context. Inform Manage 38:217–230

107   Morano RA, Ozturk C, Conn R, Dubin S, Zietz S, Nissanov J (1998) Structured light using pseudorandom codes. IEEE T Pattern Anal.20(3):322–327

108   Motavalli S, Shamsaasef R (1996) Object-oriented modelling of a feature based reverse engineering system. Int J Comp Integ M 9(5):354–368

109   National Commission on New Technological Uses of Copyright Words, (1978) Final Report

110   Nimmer R, Krauthaus P (1987) Copyright and software technology infringement: Defining third party development rights. 62 Ind Law J, pp. 21

111   O'Neill H, Pouder R, Buchholtz A (1998) Patterns in the diffusion of strategies across organizations: Insights from the innovation diffusion literature, Acad Manage Rev 21(1):98–114

112   Oliensis J, Dupuis P (1993) A global algorithm for shape from shading. Proc Int Conf Comput Vision, pp. 692-701

113   Ould M (1995) Business processes. Wiley, Indianapolis, IN

114   Page D, Koschan A, Sun Y, Abidi M (2003) Laser-based imaging for reverse engineering. Sens Rev, Special Issue on Machine Vision and Laser Scanners 23(3): 223–229

115   Pagès J, Salvi J, García R, Matabosch C (2003) Overview of coded light projection techniques for automatic 3D profiling. IEEE Int Conf RobotAutomat, ICRA Taipei, Taiwan, pp. 133–138

116   Pallek D, Baumann PH, Bütefisch KA, Kompenhans J (1999) Application of moiré interferometry for model deformation measurements in large scale wind tunnels. Euromech 406:167–170

117   Park J, DeSouza GN (2005) 3D modelling of real-world objects using range and intensity images. In: Patnaik S, Jain LC, Tzafestas G, Bannore V (eds) Innovations in machine intelligence and robot perception, Springer-Verlag, London, UK

118   Park J, DeSouza G, Kak A (2001) Dual-beam structured-light scanning for 3-D object modelling, 3rd Int Conf 3-D Digital Imaging Modelling (3DIM '01), 28 May– 01 June, Quebec City, Can, pp. 65–72

119   Patton M (1990) Qualitative evaluation and research methods, 2nd edn. Sage, Beverly Hills, CA

120   Paulus MJ, Gleason SS, Sari-Sarrafa H, Johnson D.K, Foltz CJ, Austin DW, Easterly ME, Michaud EJ, Dhara MS, Hunsicker MS, Wall JW, Schell M (2000) High-resolution x-ray CT screening of mutant mouse models. SPIE Int Symp Biomed Opt, San Jose, 3921:270–279

121   Peter Pan Fabrics, Inc. v. Martin Weiner Corp. (2d Cir. 1960) 274 F.2d, pp. 487–489

122   Pham D, Dimov S (2001) Rapid manufacturing. The technologies and applications of rapid prototyping and rapid tooling. Springer-Verlag, London

123   Pham DT, Dimov SS (2001) Rapid manufacturing. Springer Verlag, London, New York, Heidelberg

124   Pitman J (1992) Probability. Springer, Berlin

125   Power M (2003) Auditing and the production of legitimacy. Account Org Soc, 28(4):379–394

126   Reid GT, Rixon RC, Stewart H (1998) Moiré topography with large contour intervals. Proc. SPIE 814:307–313

127   Reinhard K, Ryszard K, and Karsten S (1999) Shape from shading and photometric stereo methods. In: Jaehne B,. Haussecker H, Geissler P (eds) Handbook of computer vision and applications, Vol. 2, Signal processing and pattern recognition. Academic, San Diego, pp. 532–590

128   Reverse Engineering Hardware – CGI Inc.: www.reverse-eng.com

129   Reverse Engineering Hardware – FARO Technologies: www.faro.com

130   Reverse Engineering Hardware – Immersion Corporation: www.immersion.com

131   Reverse Engineering Hardware – Mitutoyo Corporation: www.mitutoyo.com

132   Reverse Engineering Hardware – Renishaw Inc.: www.renishaw.com

133   Reverse Engineering Hardware – Roland DGA Corporation: www.rolanddga.com

134   Reverse Engineering Hardware – Surphaser: www.surphaser.com

135   Reverse Engineering Hardware and Software – Paraform – Metris: www.metris.com

136   Reverse Engineering Software – Copy CAD – DelCAM PLC: www.copycad.com

137   Reverse Engineering Software – Geomagic Studio – Geomagic: www.geomagic.com

138   Reverse Engineering Software – GSI studio: www.geometrysystems.com

139   Reverse Engineering Software – Magics RP and Mimics – Materialise NV: www.materialise.com

140   Reverse Engineering Software – Polyworks: www.innovmetric.com

141 Reverse Engineering Software – Rapid Form – INUS Technology, Inc. – www.rapidform.com

142 Rocchini C, Cignoni P, Montani C, Pingi P, Scopigno R (2001) A low cost scanner based on structured light, Computer Graphics Forum (Eurographics 2001 Conf. Proc.) 20(3):299–308

143 Rogers, E (1995) Diffusion of innovations, 4th edn. Free Press, New York

144 Rooks SM, Benhabib B, Smith KC (1995) Development of an inspection process for ball-grid-array technology using scanned-beam X-ray laminography. IEEE Trans. CPMT-A 18(4):851–861

145 Sabourin E, Houser S, Bohn J (1997) Accurate exterior, fast interior layered manufacturing. Rapid Prototyping J,3:44–52

146 Salvi J, Batlle J, Mouaddib E (1998) A robust-coded pattern projection for dynamic 3D scene measurement, Pattern Recognition Lett 19:1055–1065

147 Sea-Image Corporation: www.sea-image.com, and Austeck Pty Ltd: www.austeck.com

148 Secure Services Technology, Inc. v. Time & Space Processing (E.D. Va. 1989) 722 F. Supp., pp. 1354–1361

149 Sega Enterprises Ltd. v. Accolade, Inc. (9th Cir. 1992) 977 F. 2d, pp. 1392–1395; 1510–1528

150 Sekimoto K, Kamagami N, Hisamitsu Y, Ono K, Nagata K (2004) Development of 3D laser radar for traffic monitoring. IHI Eng Rev 37(2):71–74

151 Shi M (2002) Surface reconstruction using an edge-based growing method. Masters thesis, National University of Singapore, Singapore

152 Soma JT *et al.* (1994) Software interoperability and reverse engineering. Rutgers Comput Tech Law J 20:189

153 Speier C, Venkatesh V (2002) The hidden minefields in the adoption of sales force automation technologies, J Marketing 66(3):98–111

154 Srinivasan R, Lillien G, Rangaswamy A (2002) Technology opportunism and radical technology adoption: An application to E-business, J Marketing 66:47–60

155 Stewart C, Bu (2000) Computed tomography. McGraw-Hill Appleton & Lange, New York, NY

156 Suk M, Bhandarkar SM (1992) Three-dimensional object recognition from range images. Springer-Verlag, Tokyo

157 Sun C (2002) Fast stereo matching using rectangular subregioning and 3D maximum-surface techniques. Int J Comput Vision 47(1–3):99–117

158 Sun W, Bradley C, Zhang Y, Loh H (2001) Cloud data modelling employing a unified, non-redundant triangular mesh. Comput Aided Des 33:183–193

159 Suzuki M, Kanaya M (1988) Application of moiré topography measurement methods in industry. Opt Lasers Eng, 8(3&4):171–188

160 Szymon R, Olaf HH, Marc L (2002) Real-time 3D model acquisition, ACM Trans Graphics (TOG). Proc 29th Annu Conf Comput Graphics Interact Tech 21(3): 438–446

161 Tamas V, Ralph RM, Jordan C (1997) Reverse engineering of geometric models–an introduction. Comput Aided Des 9(4):255–268

162 Teo, T, Tan M, Buk K (1997) A contingency model of Internet adoption in Singapore Int J Electron Comm 22:57–69

163 Thompson D (1995) The optimization of part orientation for solid freeform manufacture. Master's Thesis, University of Texas, pp. 38–44

164 Thompson WB, Owen JC, de St. Germain HJ (1999) Feature-based reverse engineering of mechanical parts. IEEE T Robot Automat 15:57–66

165   Toshiba IT & Control Systems Corporation. http://www.toshiba-itc.com; and Micro Photonics Inc.: http://www.microphotonics.com.

166   Tyberg J, Bohn J (1998) Local adaptive slicing. Rapid Prototyping J 4:118–127

167   Utterback J (1996) Mastering the dynamics of innovation. Harvard Business School, Cambridge, MA

168   Valle S, Zanzi L, Lentz H, Braun HM (2000) Very high resolution radar imaging with a stepped frequency system. Proc 8ᵗʰ Int Conf Ground Penetrating Radar, GPR 2000, Gold Coast, Australia, pp. 464–470

169   Várady T, Martin RR, Cox J (1997) Reverse engineering of geometric models - an introduction. Comput Aided Des 29 (4):255–268

170   Weir D, Milroy M, Bradley C, Vickers G (1996) Reverse engineering physical models employing wrap-around B-spline surfaces and quadrics. Proc Inst Mech Eng B 210:147–157

171   Wejnert B (2002) Integrating models of diffusion of innovations: A conceptual framework, Annu Rev Sociol 28:297–325

172   Wikipedia, URL: http://en.wikipedia.org/wiki/Reverse_engineering, accessed on 11/28/2005

173   Wohlers T (2004) Rapid prototyping, tooling & manufacturing state of the industry. Annu Worldwide Progr Rep. Wohlers Associates, Fort Collins, CO

174   Woiceshyn J (2000) Technology adoption: Organizational learning in oil firms, Org Stud, 21(6):1095–1118

175   Wu Y, Wong Y, Loh H, Zhang Y (2004) Modelling cloud data using an adaptive slicing approach. Comput Aided Des 36:231–240

176   Yang M, Lee E (1999) Segmentation of measured point data using a parametric quadric surface approximation. Comput Aided Des 31:449–457

177   Yau H, Haque S, Menq C (1995) Reverse engineering in the design of engine intake and exhausts ports. Manuf Sci Eng, 64:139–148

178   Zaltman G, Duncan R, Holbeck J (1973) Innovation and organizations. Wiley, London, pp. 10

179   Zerr B, and Stage B (1996) Three-dimensional reconstruction of underwater objects forms a sequence of sonar images. Proc Int Conf Image Process 3:927–930

Index