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

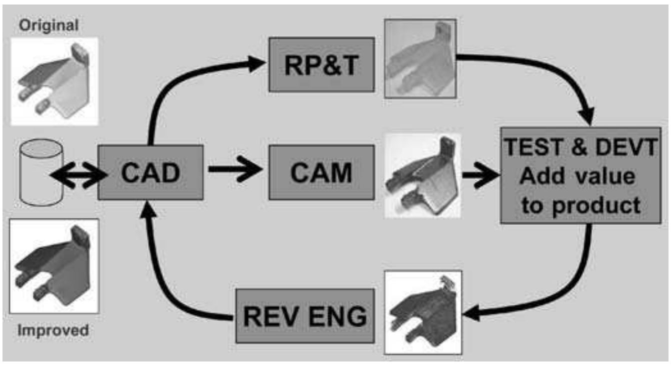

Figure 1.1. Product development cycle

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

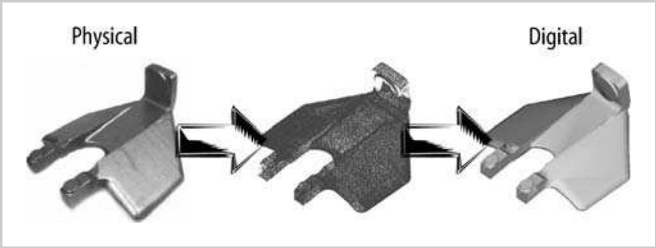

Figure 1.2. Physical-to-digital process

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:

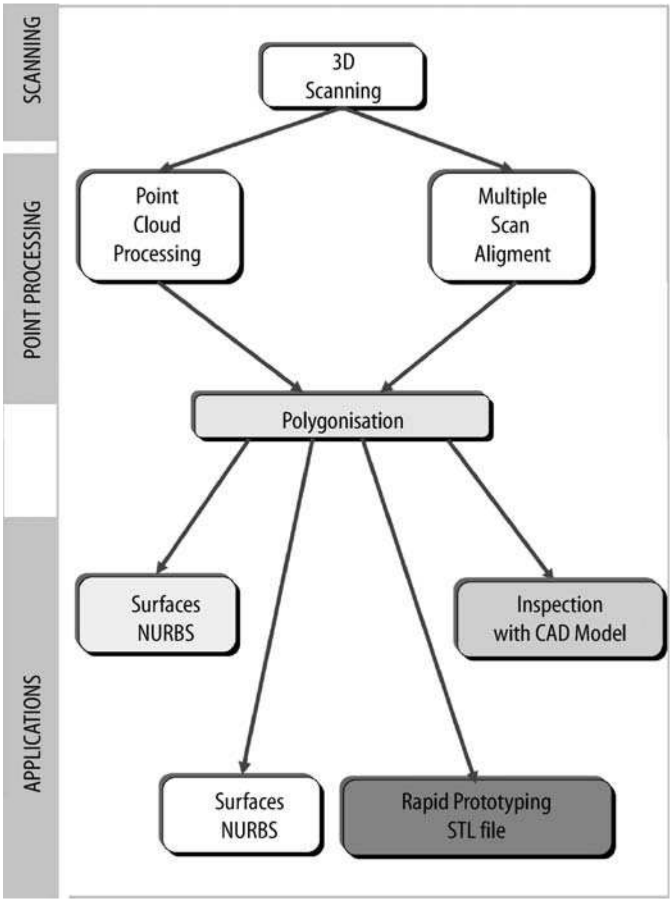

Figure 1.3. Reverse engineering – the generic process

- 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

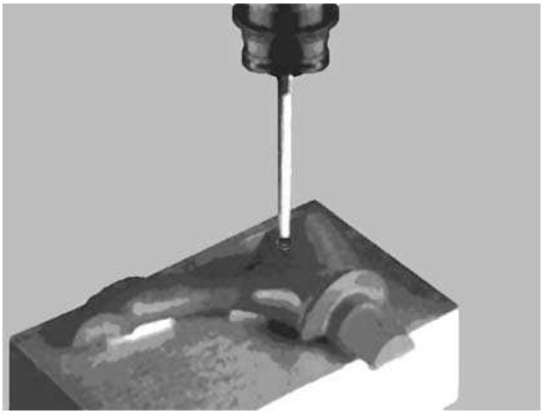

Figure 1.4. Contact scanning touch probe. Originally published in *Rapid Prototyping Casebook*, McDonald, J.A., Ryal, C.J. and Wimpenny, D.I., 2001, © John Wiley and Sons Limited. Reproduced with permission.

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.

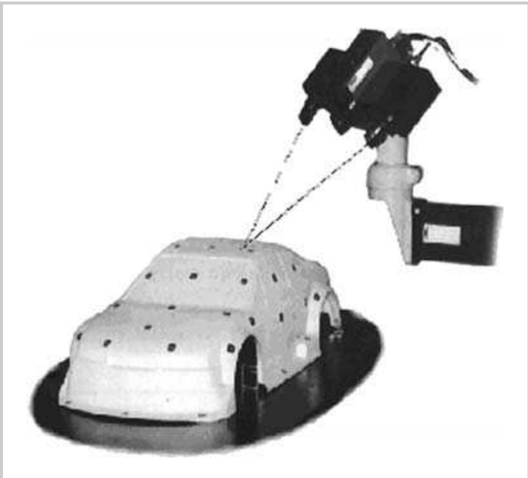

Figure 1.5. Optical scanning device. Originally published in *Rapid Prototyping Casebook*, McDonald, J.A., Ryal, C.J. and Wimpenny, D.I., 2001, © John Wiley and Sons Limited. Reproduced with permission.

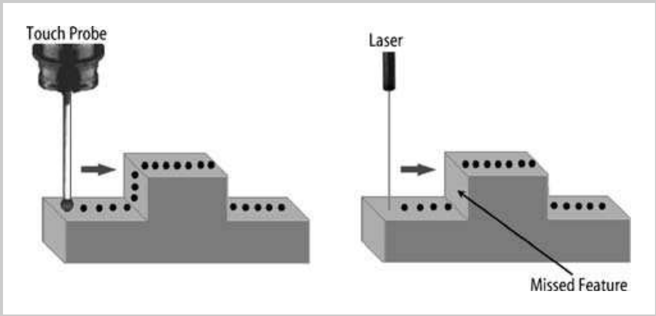

Figure 1.6. Vertical faces–touch probe versus a laser. Originally published in *Rapid Prototyping Casebook*, McDonald, J.A., Ryal, C.J. and Wimpenny, D.I., 2001, © John Wiley and Sons Limited. Reproduced with permission.

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.

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.

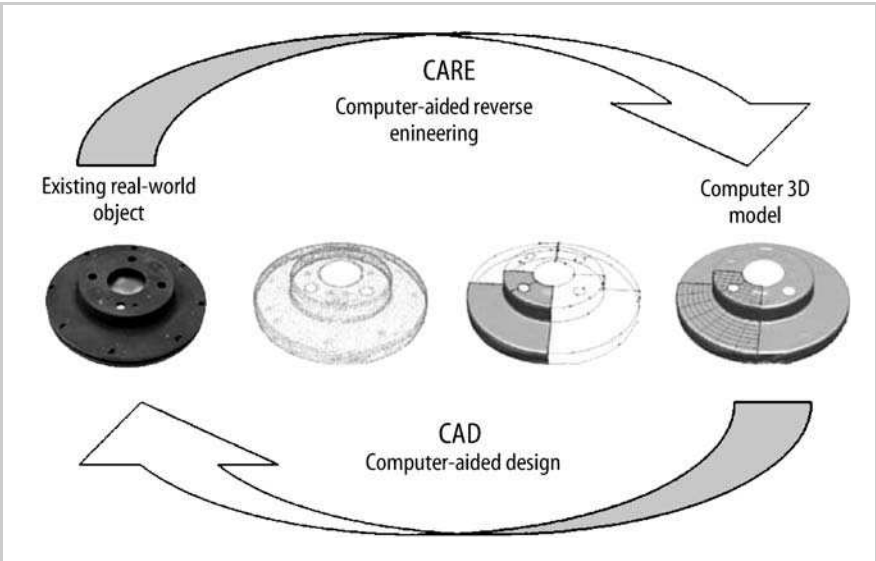

Figure 2.1. Computer-aided reverse engineering (CARE) process

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

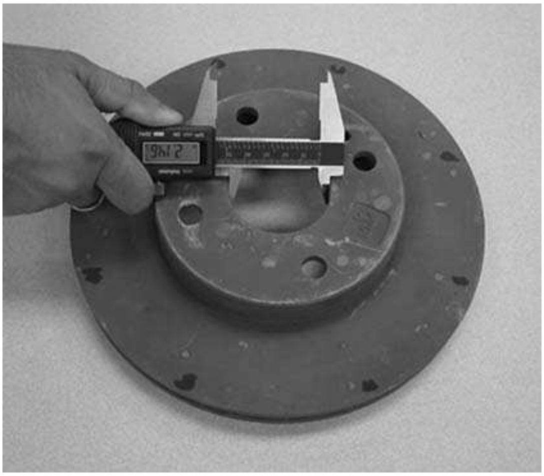

Figure 2.2. Measuring the disk brake using a caliper

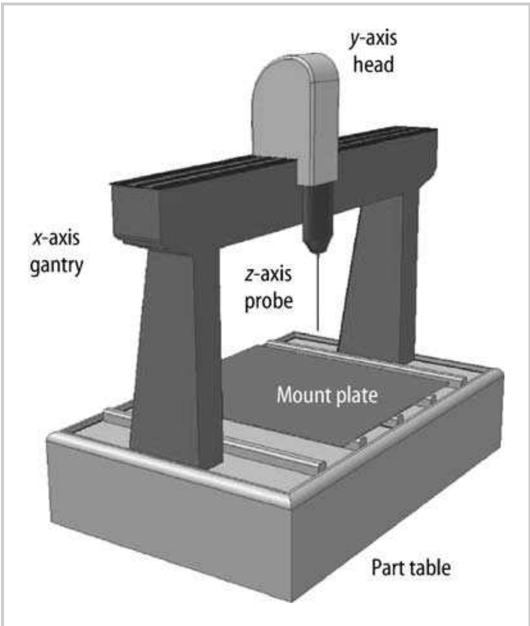

Figure 2.3. Conceptual view of a CMM that illustrates the major components of most systems

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

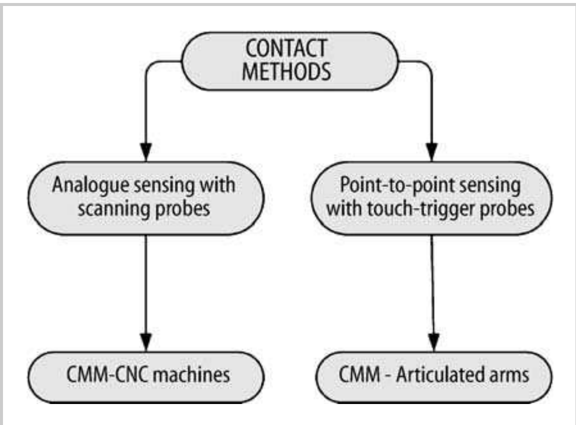

Figure 3.1. RE hardware classification–contact methods

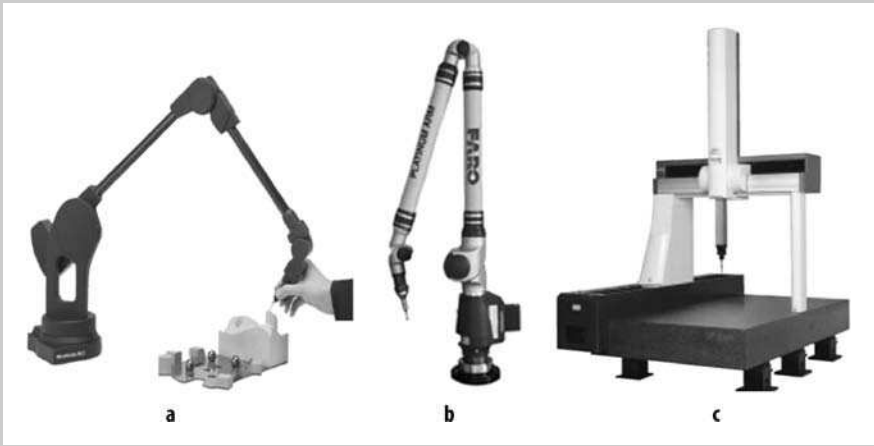

Figure 3.2. (**a**) MicroScribe MX Articulated Arm from Immersion Corporation (Immersion, 2005). (**b**) Faro Arm–Platinum articulated arm from FARO Technologies (FARO, 2005). (**c**) Mitutoyo CMM machine–CRA Apex C model (Mitutoya, 2005).

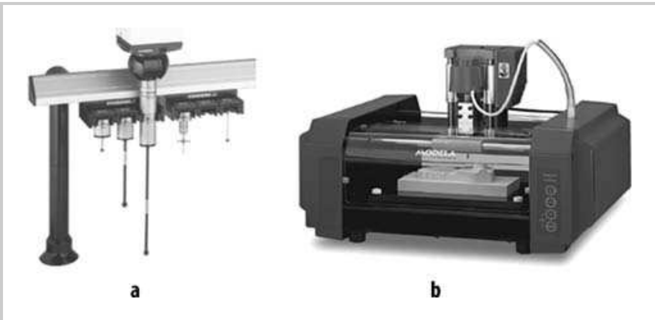

Figure 3.3. (**a**) SP25M scanning probes from Renishaw Inc (2005). (**b**) Roland DGA Corp. MDX-15/20 scanning and milling machine, using the Roland Active Piezo Sensor for 3-D scanning (Roland, 2005).

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 × 203 × 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 × 100 × 60		
		MDX-20	200 × 150 × 60		
Point-to-point sensing with a touch- trigger probe, CMM	Mitutoyo	Euro-C-121210	1205 × 1205 × 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;

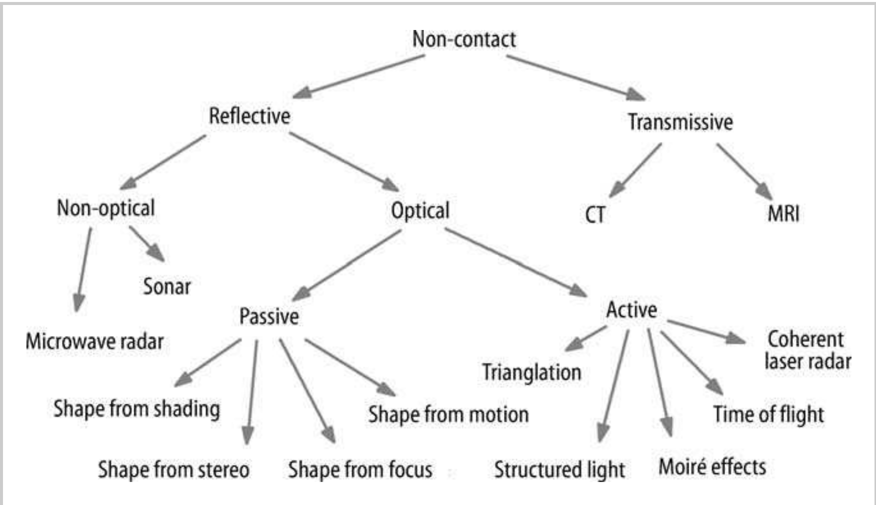

Figure 3.4. RE hardware classification–noncontact methods

(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

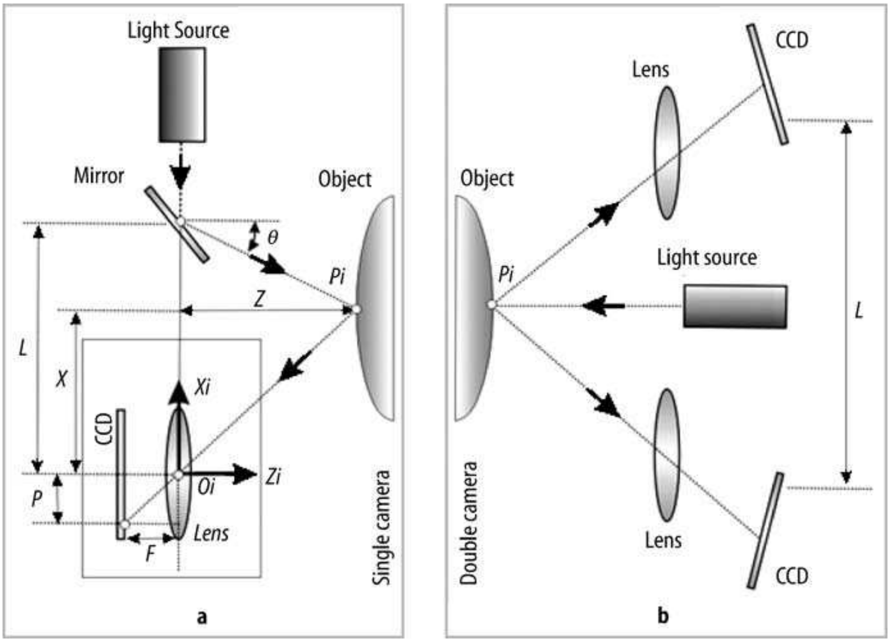

Figure 3.5. Triangulation methods: (**a**) single and (**b**) double camera arrangement

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.

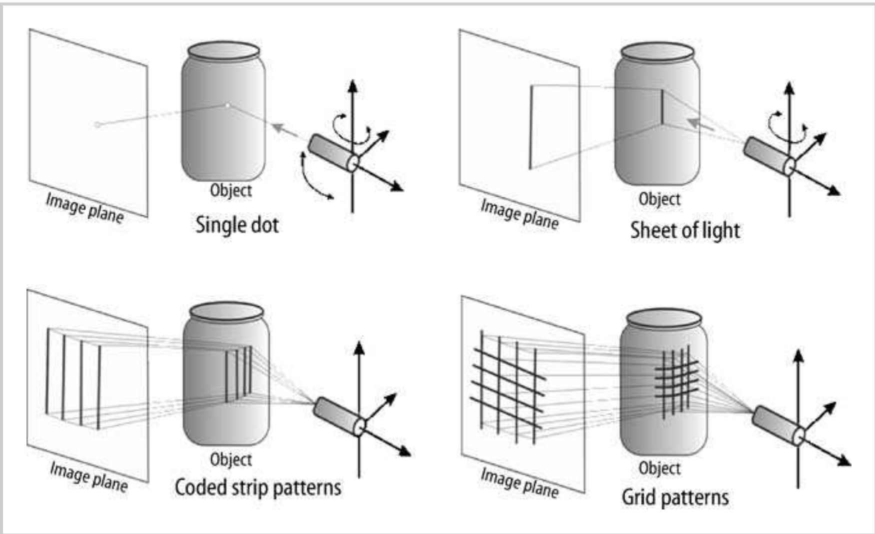

Figure 3.6. Different light patterns used in structured-light techniques

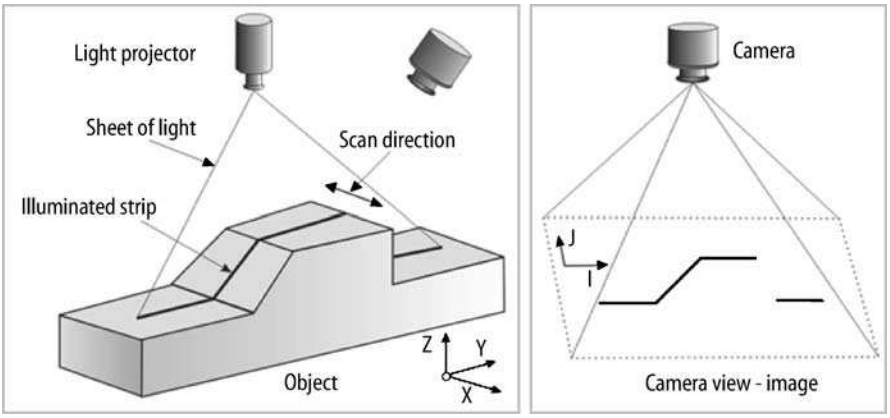

Figure 3.7. 2-D image acquisition in the structured-light technique using a sheet-of-light pattern

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.

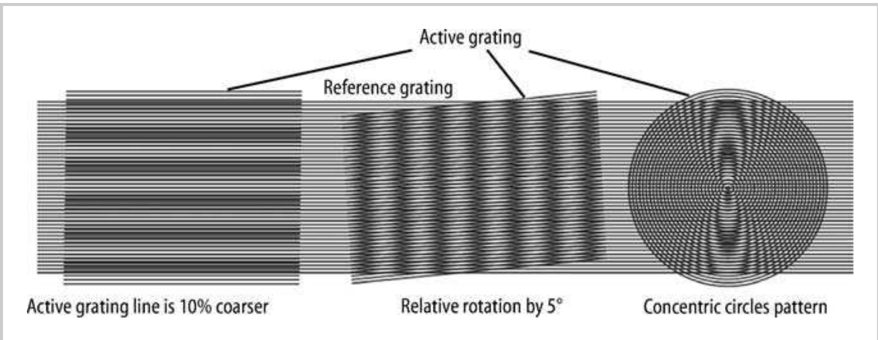

Figure 3.8. Formation of moiré fringes

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

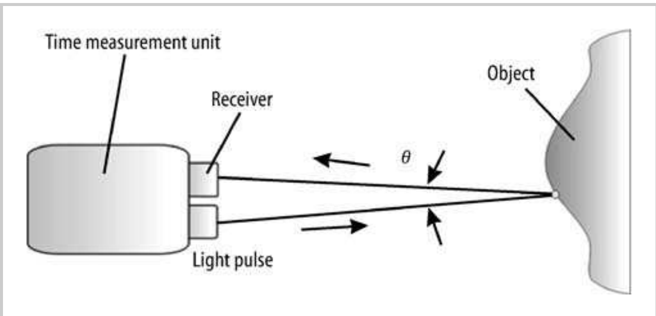

Figure 3.9. Principle of TOF scanners

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.

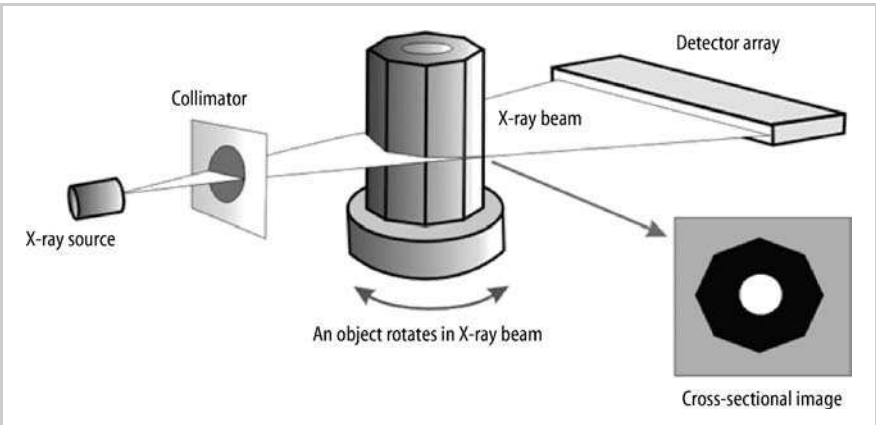

Figure 3.10. Working principle of the CT scanner

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.

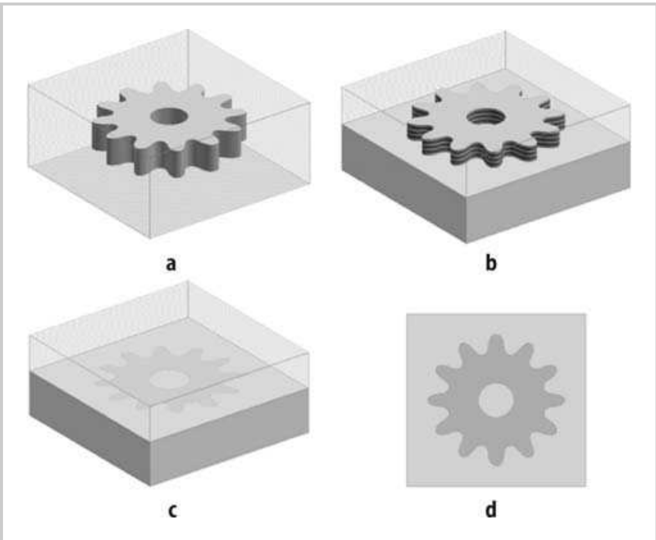

Figure 3.11. A procedure for data acquisition in a destructive RE system. (**a**) The part and matrix combination is embedded in a contrasting colored plastic matrix. (**b**, **c**) The part is machined layer-by-layer to expose the cross-sectional images. (**d**) The newly exposed surface of the part. A color reproduction of this figure can be seen in the color section (pages 219–230).

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

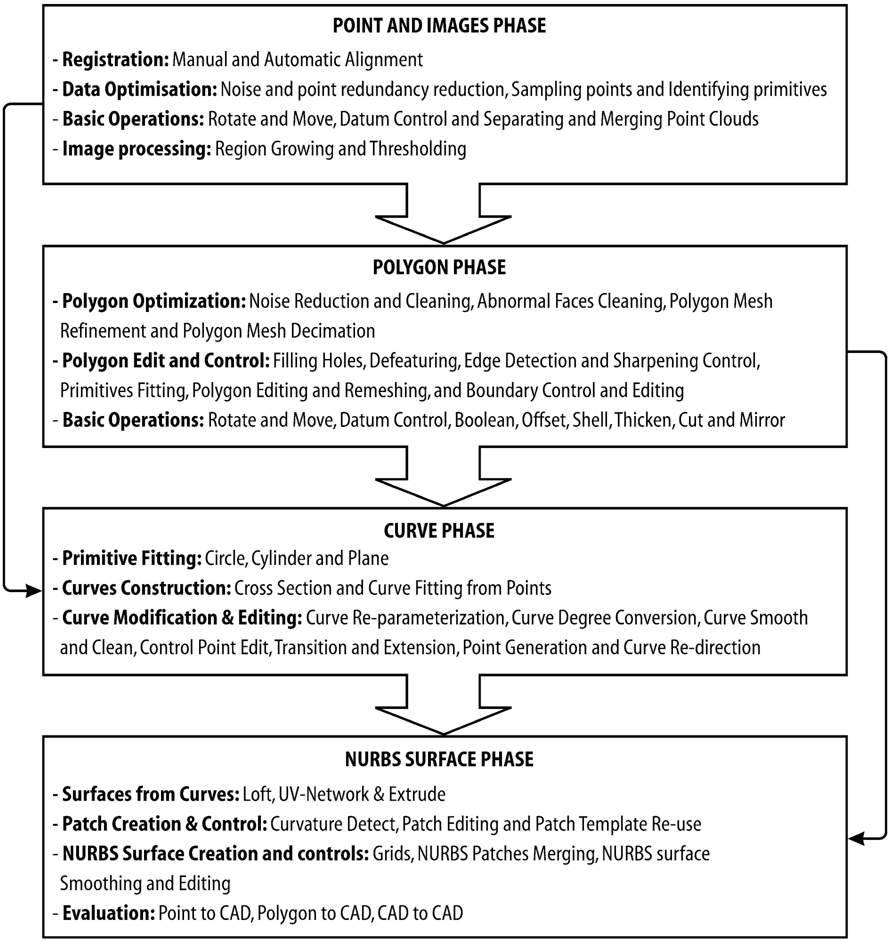

Figure 3.12. Four phases of the RE data processing chain with fundamental RE operations

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

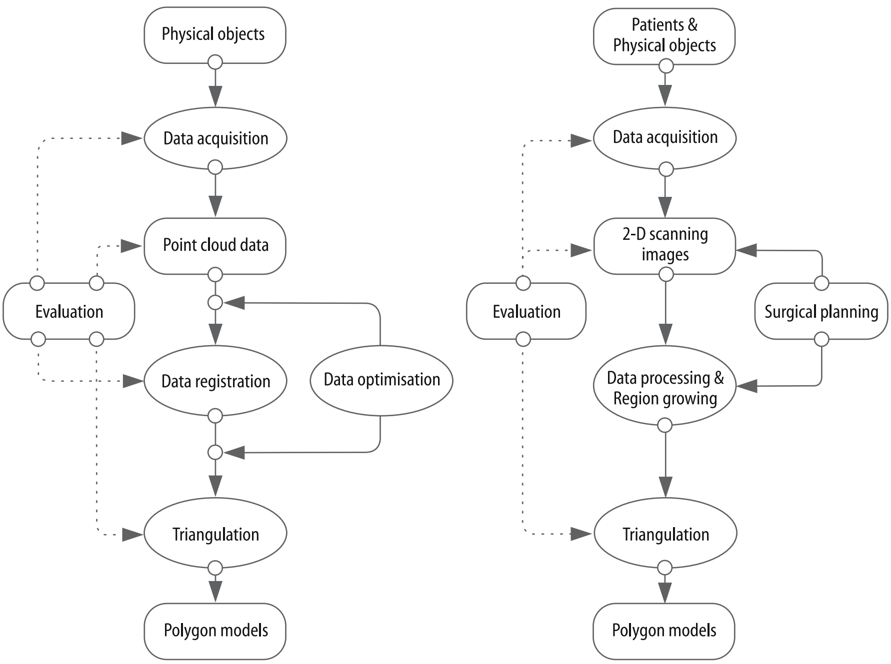

Figure 3.13. A flowchart for transforming RE scan data into 3-D triangle mesh models. (**a**) Point cloud data as the input. (**b**) 2-D scan images as the input.

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).

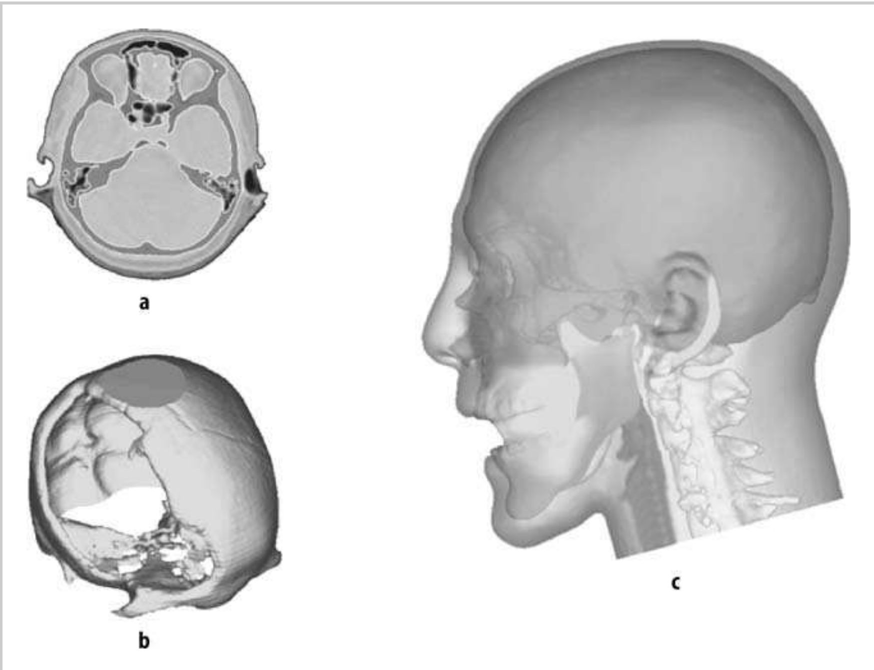

Figure 3.14. (**a**) A skull defined by the thresholding technique. (**b**, **c**) 3-D models of the cranial defect skull (**b**), and other anatomical structures of the head (**c**). A color reproduction of this figure can be seen in the Color Section (pages 219–230).

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.

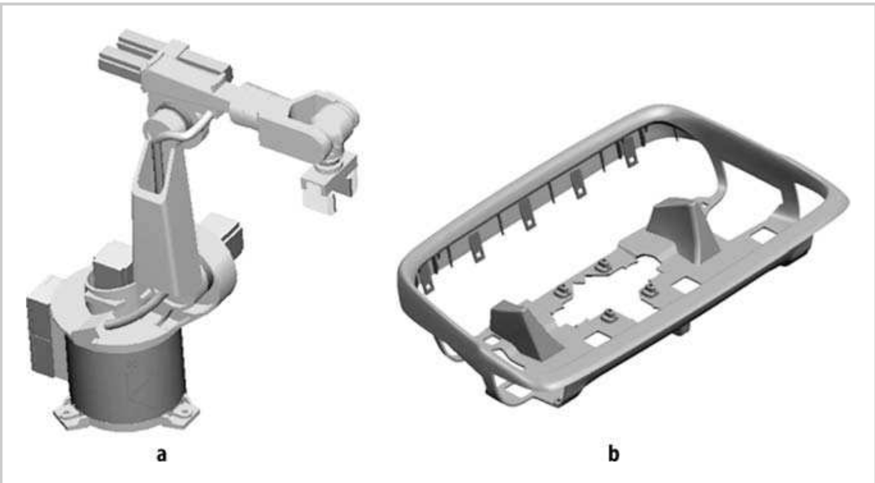

Figure 3.15. (**a**) An RE model of a mechanical part with simple geometries (KUKA robot). (**b**) An RE model of a complex part with free-form surfaces.

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.

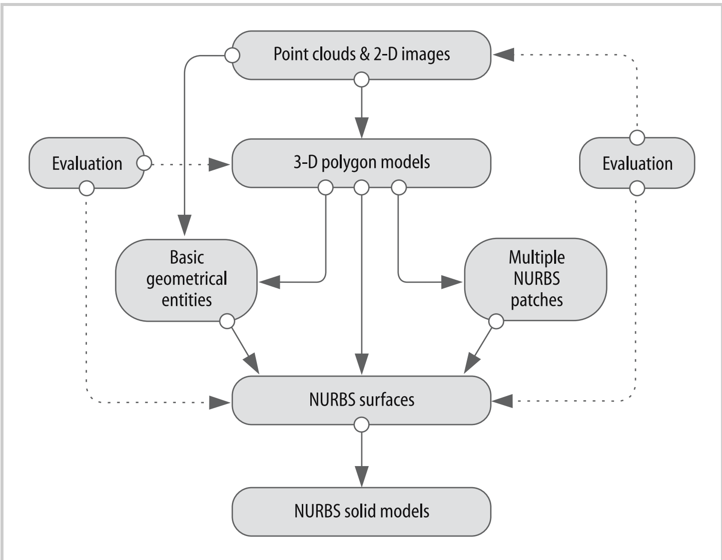

Figure 3.16. Different RE approaches for NURBS construction

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

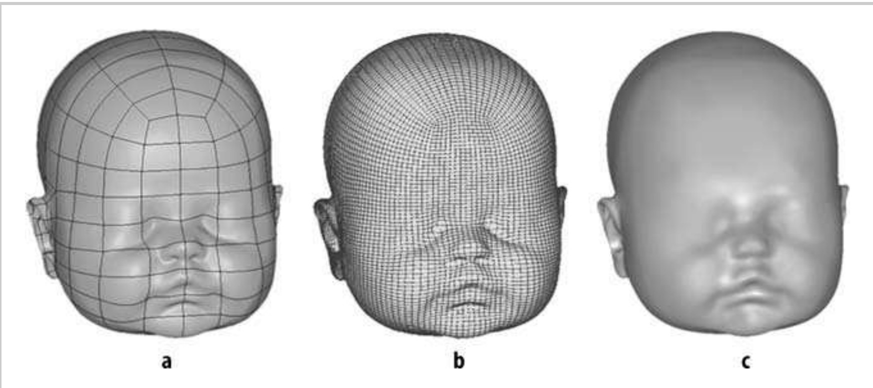

Figure 3.17. (**a**) An infant's head model with a NURBS patch layout. (**b**) NURBS patches with UV grids of control points. (**c**) A NURBS solid model of the head.

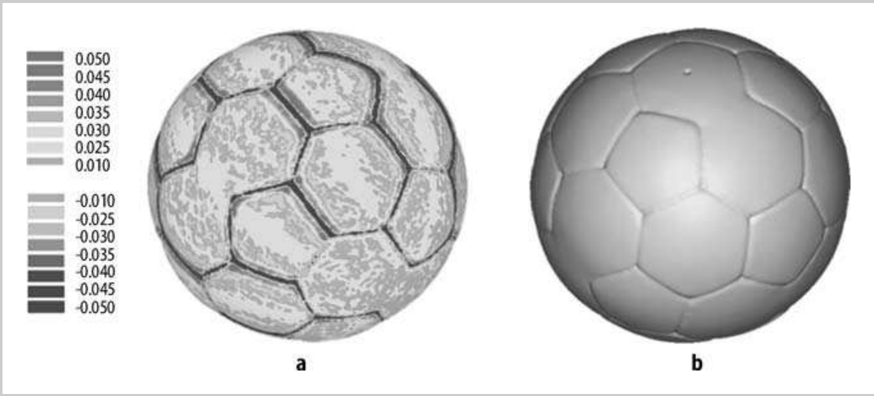

Figure 3.18. An error map (**a**) between an original polygon football model and (**b**) a resulting NURBS model. A color reproduction of this figure can be seen in the Color Section (pages 219–230).

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