# 1

# Introduction

## 1.1 Motivation

Shape analysis is an old topic that has been studied, for many centuries, by scientists from different boards, including philosophers, psychologists, mathematicians, biologists, and artists. However, in the past two decades, we have seen a renewed interest in the field motivated by the recent advances in 3D acquisition, modeling, and visualization technologies, and the substantial increase in the computation and storage power. Nowadays, 3D scanning devices are accessible not only to domain-specific experts but also to the general public. Users can scan the real world at high resolution, using devices that are as cheap as video cameras, edit the 3D data using 3D modeling software, share them across the web, and host them in online repositories that are growing in size and in number. Such repositories can include millions of every day objects, cultural heritage artifacts, buildings, as well as medical, scientific, and engineering models.

The increase in the availability of 3D data comes with new challenges in terms of storage, classification, and retrieval of such data. It also brings unprecedented opportunities for solving long-standing problems; First, the rich variability of 3D content in existing shape repositories makes it possible to directly reuse existing 3D models, in whole or in part, to construct new 3D models with rich variations. In many situations, 3D designers and content creators will no more need to scan or model a 3D object or scene from scratch. They can query existing repositories, retrieve the desired models, and fine-tune their geometry and appearance to suit their needs. This concept of context reuse is not specific to 3D models but has been naturally borrowed from other types of media. For instance, one can translate sentences to different languages by performing cross-language search. Similarly, one can create an image composite or a visual art piece by querying images, copying parts of them and pasting them into their own work.



*3D Shape Analysis: Fundamentals, Theory, and Applications,* First Edition.
Hamid Laga, Yulan Guo, Hedi Tabia, Robert B. Fisher, and Mohammed Bennamoun.

Second, these large amounts of 3D data can be used to learn computational models that effectively reason about properties and relationships of shapes without relying on hard-coded rules or explicitly programmed instructions. For instance, they can be used to learn 3D shape variation in medical data in order to model physiological abnormalities in anatomical organs, model their natural growth, and learn how shape is affected by disease progression. They can be also used to model 3D shape variability using statistical models, which, in turn, can be used to facilitate 3D model creation with minimum user interaction.

Finally, data-driven methods facilitate high-level shape understanding by discovering geometric and structural patterns among collections of shapes. These patterns can serve as strong priors not only in various geometry processing applications but also in solving long-standing computer vision problems, ranging from low-level 3D reconstruction to high-level scene understanding.

These technological developments and the opportunities they bring have motivated researchers to take a fresh look at the 3D shape analysis problem. Although most of the recent developments are application-driven, many of them aim to answer fundamental, sometimes philosophical, questions such as: *What is shape? Can we mathematically formulate the concept of shape? How to compare the shape of objects? How to quantify and localize shape similarities and differences?* This book synthesizes the critical mass of 3D shape analysis research that has accumulated over the past 15 years. This rapidly developing field is both profound and broad, with a wide range of applications and many open research questions that are yet to be answered.

## 1.2 The 3D Shape Analysis Problem

> *Shape is the external form, outline or surface, of someone or something as opposed to other properties such as color, texture, or material composition.*
> Source: Wikipedia and Oxford dictionaries.

Humans can easily abstract the form of an object, describe it with a few geometrical attributes or even with words, relate it to the form of another object, and group together, in multiple ways and using various criteria, different objects to form clusters that share some common shape properties. Shape analysis is the general term used to refer to the process of automating these tasks, which are trivial to humans but very challenging to computers. It has been investigated under the umbrella of many applications and has multiple facets. Below, we briefly summarize a few of them.

- **3D shape retrieval, clustering, and classification**. Similar to other types of multimedia information, e.g. text documents, images, and videos, the demand for efficient clustering and classification tools that can organize, automatically or semi-automatically, the continuously expanding collections of 3D models is growing. Likewise, users, whether they are experts, e.g. graphics designers who are increasingly relying on the reuse of existing 3D contents, or novice, will benefit from a search engine that will enable them to search for 3D data of interest in the same way they search for text documents or images.
- **Correspondence and registration**. This problem, which can be summarized as the ability to say which part of an object matches which part on another object, and the ability to align one object onto another, arises in many domains of computer vision, computer graphics, and medical imaging. Probably, one of the most popular examples is the 3D reconstruction problem where usually a 3D object is scanned by multiple sensors positioned at different locations around the object. To build the complete 3D model of the object, one needs to merge the partial scans produced by each sensor. This operation requires a correct alignment, i.e. registration, step that brings all the acquired 3D data into a common coordinate frame. Note also that, in many cases, 3D objects move and deform, in a nonrigid way, during the scanning process. This makes the alignment process even more complex. Another example is in computer graphics where a 3D designer creates a triangulated 3D mesh model, hereinafter referred to as the reference, and assigns to each of its triangular faces some attributes, e.g. color and material properties. The designer then can create additional models with the same attributes but instead of manually setting them, they can be automatically transferred from the reference model if there is a mechanism which finds for each point on the reference model its corresponding points on the other models.
- **Detection and recognition**. This includes the detection of low level features such as corners or regions of high curvatures, as well as the localization and recognition of parts in 3D objects, or objects in 3D scenes. The latter became very popular in the past few years with the availability of cheap 3D scanning devices. In fact, instead of trying to localize and recognize objects in a scene from 2D images, one can develop algorithms that operate on the 3D scans of the scene, eventually acquired using commodity devices. This has the advantage that 3D data are less affected than 2D images by the occlusions and ambiguities, which are inherent to the loss of dimensionality when projecting the 3D world onto 2D images. 3D face and 3D action recognition are, among others, examples of applications that have benefited from the recent advances in 3D technologies.

- **Measurement and characterization** of the geometrical and topological properties of objects on one hand and of the spatial relations between objects on the other hand. This includes the identification of similar regions and finding recurrent patterns within and across 3D objects.
- **Summarization and exploration** of collections of 3D models. Given a set of objects, one would like to compute a representative 3D model, e.g. the average or median shape, as well as other summary statistics such as covariances and modes of variation of their shapes. One would like also to characterize the collection using probability distributions and sample from these distributions new instances of shapes to enrich the collection. In other words, one needs to manipulate 3D models in the same way one manipulates numbers.

Implementing these representative analysis tasks requires solving a set of challenges, and each has been the subject of important research and contributions. The first challenge is **the mathematical representation** of the shape of objects. 3D models, acquired with laser scanners or created using some modeling software, can be represented with point clouds, polygonal soup models, or as volumetric images. Such representations are suitable for storage and visualization but not for high-level analysis tasks. For instance, scanning the same object from two different viewpoints or using different devices will often result in two different point clouds but the shape remains the same. The challenge is in designing mathematical representations that capture the essence of *shape*. A good representation should be independent of (or invariant to) the pose of the 3D object, the way it is scanned or modeled, and the way it is stored. It is also important to ensure that two different shapes cannot have the same representation.

Second, almost every shape analysis task requires a **measure that quantifies shape similarities and differences**. This measure, called dissimilarity, distance, or metric, is essential to many tasks. It can be used to compare the 3D shape of different objects and localize similar parts in and across 3D models. It can also be used to detect and recognize objects in 3D scenes. Shape similarity is, however, one of the most ambiguous concepts in shape analysis since it depends not only on the geometry of the objects being analyzed but also on their semantics, their context, the application, and on the human perception. Figure 1.1 shows a few examples that illustrate the complexity of the shape similarity problem. In Figure 1.1a, we consider human body shapes of the same person but in different poses. One can consider these models as similar since they are of the same person. One may also treat them as different since they differ in pose. On the other hand, the 3D objects of Figure 1.1b are only partially similar. For instance, one part of the centaur model can be treated as similar to the upper body of the human body shape, while the other part is similar to the 3D shape of a horse. Also, one can consider that the

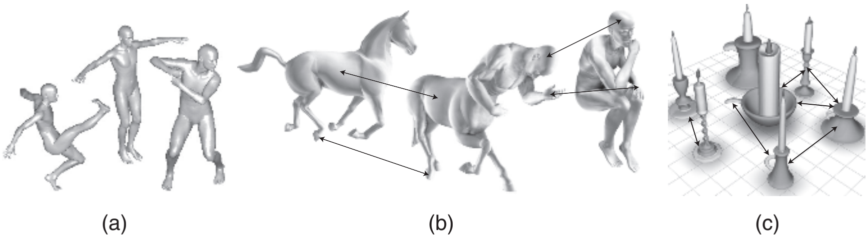


**Figure 1.1** Complexity of the shape similarity problem. (a) Nonrigid deformations. (b) Partial similarity. (c) Semantic similarity.

candles of Figure 1.1c are similar despite the significant differences in their geometry and topology. A two-year-old child can easily match together the parts of the candles that have the same functionality despite the fact that they have different geometry, structure, and topology.

Finally, these problems, i.e. representation and dissimilarity, which are interrelated (although many state-of-the-art papers treat them separately), are the core components of and the building blocks for almost every 3D shape analysis system.

## 1.3 About This Book

The field of 3D shape analysis is being actively studied by researchers originating from at least four different domains: mathematics and statistics, image processing and computer vision, computer graphics, and medical imaging. As a result, a critical mass of research has accumulated over the past 15 years, where almost every major conference in these fields included tracks dedicated to 3D shape analysis. This book provides an in-depth description of the major developments in this continuously expanding field of research. It can serve as a complete reference to graduate students, researchers, and professionals in different fields of mathematics, computer science, and engineering. It could be used for courses of intermediate level in computer vision and computer graphics or for self-study. It is organized into four main parts:

**The first part**, which is composed of two chapters, provides an in-depth review of the background concepts that are relevant to most of the 3D shape analysis aspects. It begins in Chapter 2 with the basic elements of geometry and topology, which are needed in almost every 3D shape analysis task. We will look in this chapter into elements of differential geometry and into how 3D models are represented. While most of this material is covered in many courses and textbooks, putting them in the broader context of shape analysis will help the reader appreciate the benefits and power of these fundamental mathematical tools.

Chapter 3 reviews the techniques that are used to capture, create, and preprocess 3D models. Understanding these techniques will help the reader, not only to understand the various challenges faced in 3D shape analysis but will also motivate the use of 3D shape analysis techniques in improving the algorithms for 3D reconstruction, which is a long-standing problem in computer vision and computer graphics.

**The second part**, which is composed of two chapters, presents a range of mathematical and algorithmic tools that are used for shape description and comparison. In particular, Chapter 4 presents the different descriptors that have been proposed in the literature to characterize the global shape of a 3D object using its geometry and/or topology. Early works on 3D shape analysis, in particular classification and retrieval, were based on global descriptors. Although they lack the discrimination power, they are the foundations of modern and powerful 3D shape descriptors.

Chapter 5, on the other hand, covers the algorithms and techniques used for the detection of local features and the characterization of the shape of local regions using local descriptors. Many of the current 3D reconstruction, recognition, and analysis techniques are built on the extraction and matching of feature points. Thus, these are fundamental techniques required in most of the subsequent chapters of the book.

**The third part of the book**, which is composed of three chapters, focuses on the important problem of computing correspondences and registrations between 3D objects. In fact, almost every task, from 3D reconstruction to animation, and from morphing to attribute transfer, requires accurate correspondence and registration. We will consider the three commonly studied aspects of the problem, which are rigid registration (Chapter 6), nonrigid registration (Chapter 7), and semantic correspondence (Chapter 8). In the first case, we are given two pieces of geometry (which can be partial scans or full 3D models), and we seek to find the rigid transformations (translations, scaling and rotations) that align one piece onto the other. This problem appears mainly in 3D scanning where often a 3D object is scanned by multiple scanners. Each scan produces a set of incomplete point clouds that should be aligned and fused together to form a complete 3D model.

3D models can not only undergo rigid transformations but also nonrigid deformations. Think, for instance, of the problem of scanning a human body. During the scanning process, the body can not only move but also bend. Once it is fully captured, we would like to transfer its properties (e.g. color, texture, and motion) onto another 3D human body of a different shape. This requires finding correspondences and registration between these two 3D objects, which bend and stretch. This is a complex problem since the space of solutions is large and requires efficient techniques to explore it. Solutions to this problem will be discussed in Chapter 7.

Semantic correspondence is even more challenging; think of the problem of finding correspondences between an office chair and a dining chair. While humans can easily match parts across these two models, the problem is very challenging for computers since these two models differ both in geometry and topology. We will review in Chapter 8 the methods that solve this problem using supervised learning, and the methods that used structure and context to infer high-level semantic concepts.

**The last part** of the book demonstrates the use of the fundamental techniques described in the earlier chapters in a selection of 3D shape analysis applications. In particular, Chapter 9 reviews some of the semantic applications of 3D shape analysis. It also illustrates the range of applications involving 3D data that have been annotated with some sort of meaning (i.e. semantics or labels).

Chapter 10 focuses on a specific type of 3D objects, which are human faces. With the widespread of commodity 3D scanning devices, several recent works use the 3D geometry of the face for various purposes including recognition, gender classification, age recognition, and disease and abnormalities detection. This chapter will review the most relevant works in this area.

Chapter 11 focuses on the problem of recognizing objects in 3D scenes. Nowadays, cars, robots, and drones are all equipped with 3D sensors that capture their environments. Tasks such as navigation, target detection and identification, object tracking, and so on require the analysis of the 3D information that is captured by these sensors.

Chapter 12 focuses on a classical problem of 3D shape analysis, which is how to retrieve 3D objects of interest from a collection of 3D models. Chapter 13, on the other hand, treats the same problem of shape retrieval but this time by using multimodal queries. This is a very recent problem that has received a lot of interest with the emergence of deep-learning techniques that enable embedding different modalities into a common space.

The book concludes in Chapter 14 with a summary of the main ideas and a discussion of the future trends in this very active and continuously expanding field of research.

Readers can proceed sequentially through each chapter. Some readers may want to go straight to topics of their interest. In that case, we recommend to follow the reading chart of Figure 1.2, which illustrates the inter-dependencies between the different chapters.

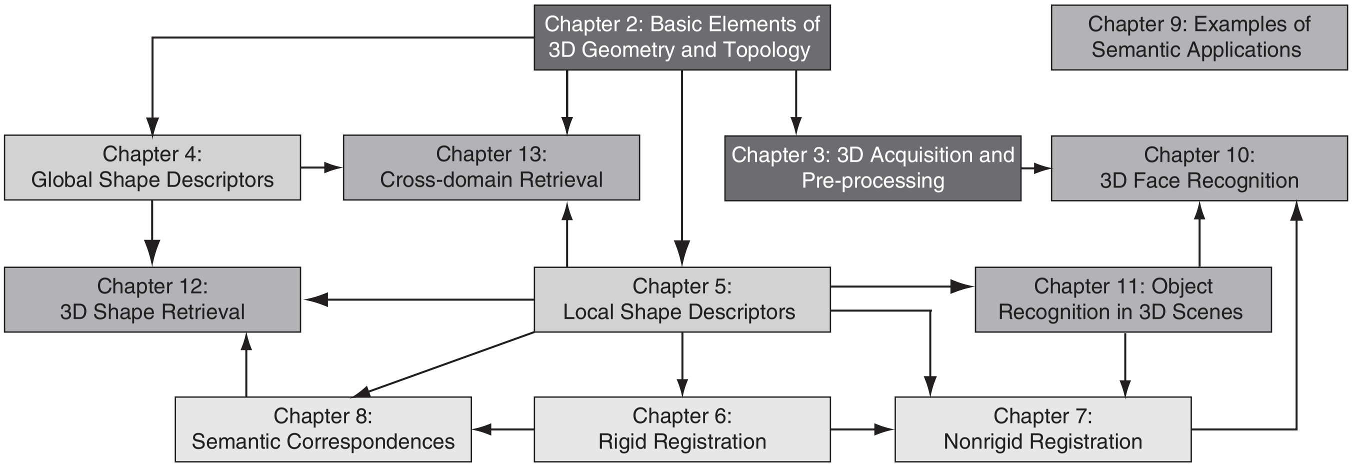


**Figure 1.2** Structure of the book and dependencies between the chapters.

## 1.4 Notation

Table 1.1 summarizes the different notations used throughout the book.

**Table 1.1** List of notations used throughout the book.

| Symbol | Description |
|---|---|
| $\mathbb{N}$ | Natural numbers |
| $\mathbb{R}$ | Real numbers |
| $\mathbb{R}_{>0}$ | Strictly positive real numbers |
| $\mathbb{R}_{\geq 0}$ | Nonnegative real numbers |
| $\mathbb{R}^2, \mathbb{R}^3, \dots, \mathbb{R}^n$ | 2D, 3D,…,$n$D Euclidean space, respectively |
| $\mathbb{S}^2$ | The spherical domain or a sphere in $\mathbb{R}^3$ |
| $D$ | A domain of a function. It is also used to refer to a parameterization domain |
| $\mathbb{C}$ | The complex plane, i.e. the space of complex numbers |
| $p, p_1, p_2, \dots$ | Points in $\mathbb{R}^2$ or $\mathbb{R}^3$ |
| $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2, \dots$ | Vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ |
| $\mathbf{n}$ | A normal vector |
| $\tilde{\mathbf{n}}$ | A unit normal vector |
| $\mathbf{v}$ | A tangent vector |
| $M$ | A triangular mesh |
| $V$ | The set (list) of vertices of a mesh |
| $T$ | List of (triangular) faces of a mesh |
| $t, t_1, t_2, \dots$ | Triangular faces of a mesh |
| $v, v_1, v_2, \dots$ | Vertices of a mesh |
| $m$ | The number of faces in a mesh |
| $n$ | The number of vertices in a mesh |
| $G$ | A graph |
| $V$ | The set of the nodes of a graph |
| $E$ | The list of edges in a graph |
| $v, v_1, v_2, \dots$ | Nodes of a graph |
| $e, e_1, e_2, \dots$ | Edges of a mesh or a graph |
| $\mathcal{N}(v)$ | The set of vertices (or nodes) that are adjacent or neighbors to the vertex (or node) $v$ in a mesh (or a graph) |
| $\lambda, \lambda_1, \lambda_2, \dots$ | Eigenvalues |
| $\Lambda, \Lambda_1, \Lambda_2, \dots$ | Eigenvectors |
| $f$ | A function |
| $X$ | A curve in $\mathbb{R}^n$ |

*(Continued)*

**Table 1.1** (Continued)

| Symbol | Description |
|---|---|
| $\Delta$ | A plane |
| $\mathcal{S}$ | A surface in $\mathbb{R}^3$. It is also used to denote the surface of a 3D object |
| $f$ | The function used to represent the surface of a 3D object |
| $\mathbf{x}$ | A descriptor |
| $\kappa$ | Curvature of a curve of a surface at a given point |
| $\kappa_n$ | Normal curvature of a surface at a given point |
| $\kappa_1$ | Minimum curvature of a surface at a given point |
| $\kappa_2$ | Maximum curvature of a surface at a given point |
| $H$ | Mean curvature of a surface at a given point |
| $K$ | Gaussian curvature of a surface at a given point |
| $s$ | Shape index of a surface at a given point |
| $\mathbf{e}$ | Principal direction of a surface at a given point |
| $\Delta$ | The Laplacian operator |
| $\nabla$ | The gradient operator |
| div | The divergence operator |
| Hess | The Hessian |
| $SO(3)$ | The space of rotations |
| $O$ | A rotation matrix, which is an element of $SO(3)$ |
| $\mathbf{t}$ | A translation vector, which is an element of $\mathbb{R}^3$ |
| $\Gamma$ | The space of diffeomorphisms |
| $\gamma$ | A diffeomorphism, which is an elements of $\Gamma$ |
| $\mathbf{t}$ | A translation vector in $\mathbb{R}^2$ or $\mathbb{R}^3$ |