1

Introduction

- Do you see that cloud, that's almost in shape like a camel?
- By the mass, and 't is like a camel, indeed.
- Methinks, it is like a weasel.
- It is backed like a weasel.
- Or, like a whale?
- Very like a whale.

W. Shakespeare, *Hamlet*

Analysis and understanding of shapes is one of the most fundamental tasks in our interaction with the surrounding world. When we see a picture, we *understand* it by recognizing the depicted objects and relating them to concepts we have learned throughout our lives. An evidence to such an understanding is the fact that we can translate a picture into a higher-level semantic description and communicate this description to another person. Hearing about a "beautiful house made of rosy brick, with geraniums in the windows and doves on the roof" [131], most of us would create a vivid mental image of it in our imagination, that is, in some sense perform the inverse process of synthesis of the picture from its semantic description.

In our everyday experience, we are often unaware of how extremely complex the shape analysis performed by the brain is, because it is done mostly subconsciously, without involving the higher level of cognition. Just imagine how different and difficult our lives would be if this ability were gone. Waking up in the morning, we would meet a lot of new people – their faces would seem unfamiliar, because we recognize humans by the visual features of their faces. Going into the bathroom would become an adventure of epic proportions, as we would be unable to decide whether the object we pick up is a razor or a toothbrush, because such a decision requires the analysis of the geometric shape of these objects. Even continuing reading this book would become impossible, because the shapes of the letters would lose their meaning, and as a result, we would lose our ability to read. Of course, this apocalyptic scenario is exaggerated, as we do not rely only on visual information in our lives, and other senses could somehow compensate for its loss. Yet, for most of us, the most significant information about the surrounding world comes from vision, such that, rephrasing an old proverb, a picture is worth a thousand odors or touches.

In the era of computers, attempts to imitate the ability of the human visual system to understand shapes gave birth to the fields of computer vision and pattern recognition. When we say "shape" in this context, we usually imply a visual representation of the object rather than the object itself. Because

Figure 1.1. Cave paintings are the one of the earliest evidences of the interest of our prehistoric fathers in the understanding of shapes and their use as a method of communicating information. Shown here are animal shapes from Salle des Taureaux (Grotte de Lascaux, France), dating back to around 15,000 BC.

what we see is actually a two-dimensional picture perceived by our eyes, a common way to think of a shape is of a two-dimensional projection of the three-dimensional object. However, computers "see" the world differently from us humans. A computer's "eye" can be an ultrasonic sensor introduced into the human body by an intravascular catheter, or a hyperspectral camera mounted on a Martian rover crawling through the canyons of the red planet. Such sensors may provide information normally imperceivable by humans, such as sub-millimeter accurate measurement of object dimensions, or colors in spectral ranges unseen by our eyes. Accordingly, in some computer vision applications, besides the common two-dimensional shapes, we can encounter shapes represented as three-dimensional point clouds, triangular meshes, or parametric or implicit surfaces.

Usually, some model that relates the shapes to the underlying objects is assumed. We call the model of shapes used throughout this book the *non-rigid world*. In this world, objects have a certain degree of flexibility by virtue of their natural properties. Consequently, we may find a great variety of shapes produced as a result of deformations of a non-rigid object. Being able to analyze the properties of such shapes and describe their behavior is the key to understanding the non-rigid world.

Figure 1.2. The legendary Russian ballerina Maya Plisetskaya and a stretching Siberian tiger are living examples of non-rigid objects we will encounter in this book (tiger photo reproduced by courtesy of Malene Thyssen).

1.1 Similarity of non-rigid shapes

In the epigraph we chose to open the book, we quote a renowned scene from the third act of Shakespeare's *Hamlet*, in which the Prince of Denmark and Polonius argue about the shape of clouds they see from the windows of the Elsinor Castle.[1] Speaking in modern language, the topic of the two noblemen's dialogue is *non-rigid shape similarity*, or how to compare shapes that are susceptible to deformations. Clouds are only one example of such shapes; we see plenty of other non-rigid shapes in the world surrounding us at all scales.

Generally speaking, in the problem of shape similarity we are looking for a quantitative measure of "distance" between two shapes: if this distance is small, we conclude that the shapes are similar. The dispute between Hamlet and Polonius teaches us that the definition of similarity may be subjective: depending on the criterion used for similarity, one can recognize the same shape of a cloud as a camel, a weasel, or a whale.

In the non-rigid world, this problem is especially acute due to the fact that an object can assume many forms as a result of its deformations. The same object deformed in different ways may result in shapes that are apparently dissimilar. As an illustration, we resort to the example of a non-rigid object

Figure 1.3. Depending on our definition of similarity, we may recognize hand postures used in the Rock, Paper, Scissors game either as the "objects" they mean to imitate (top) or as postures of a hand (bottom).

literally available in our hands: the human hand. Most of the readers are probably familiar with *Rock, Paper, Scissors*, a hand game played by children in many countries (Figure 1.3). In this game, at count, the players simultaneously change their hands into any of three "objects": rock (represented by a closed fist), scissors (two extended fingers) or paper (open palm). The definition of shape similarity in this example is ambiguous, as we find disagreement about how to consider our objects. As children playing the Rock, Paper, Scissors game, we recognize the hand shapes as the objects they intend to mimic. As adults, we say that all these "objects" are nothing but deformations of the same human hand. The reason for this difference is due to the fact that in the first case, we consider the postures of the hands as stand-alone rigid shapes, whereas in the second case, we consider them as deformations of a non-rigid shape.

Defining similarity of non-rigid shapes, we are looking for properties that distinguish between what really characterizes the object and what can be at-

Figure 1.4. Clay figures deformed nearly isometrically (top) and non-isometrically (bottom).

tributed to its deformations. Such properties are called *deformation-invariant*, and a similarity criterion based on these properties is called *deformation-invariant similarity*. In the human hand example, the length of the fingers, which always remains the same no matter how we articulate them, is one of the deformation-invariant characteristics.

Saying that a property is deformation-invariant, we need to specify what type of deformation is considered. A human hand sculpted out of clay will deform differently than will a real hand of flesh and bones: for example, the length and the width of the fingers of a clay hand may change almost without any restriction. Moreover, because a piece of clay can be torn apart, pierced, and bent, we can create almost any shape out of it (Figure 1.4). In this case, it will be hard to find any properties invariant under such a wide class of deformations.

If, however, we restrict ourselves to deformations similar to those of a human hand, we are in a much better situation. In geometric jargon, we call such deformations *articulations* or more generally *isometries*, and say that they preserve the *intrinsic geometry* of the shape. It appears that deformations of many natural objects can be modeled as isometries. Human and animal bodies deform approximately isometrically: we can flex our hands and legs to some extent but cannot stretch or shrink them. Of course, as in any model, there is a certain degree of inaccuracy in this assumption. At the same time, the benefit of limiting our discussion to the class of isometric deformations is that it leads to a well-defined geometric criterion of similarity based on the comparison of intrinsic geometry, which we refer to as the *intrinsic similarity*.

Such a similarity is invariant to isometries and therefore allows us to compare shapes no matter how they are bent.

1.2 Correspondence problems

Another important class of problems in shape analysis is known under the generic name of *correspondence problems*. When we put on a glove, we unconsciously solve a problem of great complexity: how to wear the glove such that it best fits the hand. For this purpose, we obviously need to insert each finger into its corresponding location in the glove. By saying "corresponding location," we implicitly assume that there is a natural correspondence between the hand and the glove – the thumb goes into the thumb of the glove, the index finger into the index finger, and so on. However, an automatic computation of such a "natural" correspondence is by no means trivial. Because both the hand and the glove are non-rigid objects, the correspondence must be independent of the deformations, that is, deformation-invariant.

The fact that we encounter deformation invariance again in this context suggests that correspondence and similarity problems are intimately related. In the glove fitting example, we can regard the "easiness" of putting on the glove as a criterion of similarity: if in order to fit the glove we make an effort to stretch it significantly (thus changing the intrinsic geometry), this means that the two objects are dissimilar (Figure 1.5). Had the jury in the United States court spoken in our terms, it would formulate the reason to acquit O. J. Simpson in his controversial murder trial as "intrinsic dissimilarity.[2]" The same criterion can be used for the definition of correspondence: we would like the fitting to be performed the easiest way. Trying to fit the thumb into the index finger of the glove is not an easy job. Therefore, like in similarity problems, we can speak about *intrinsic correspondence* between two objects and apply similar numerical tools for its computation.

Similarity and correspondence are two archetype problems used for shape *analysis* and *synthesis* applications. Many applications in the non-rigid world can be seen through the glasses of these two problems. As an example, we consider the problem of face recognition, dealing with the question of how to distinguish between the faces of two different people. Because of the flexibility of facial tissues and our ability to express a wide range of emotions, the face is a non-rigid object. Therefore, face recognition falls into the category of non-rigid shape similarity problems. Modeling facial expressions as deformations of the facial surface and using intrinsic geometric similarity criteria, we can distinguish between features resulting from expressions and those characterizing the person's identity, or in other words, make our face recognition *expression-invariant*.

Looking at the same application from a different perspective, we can ask the question how to find the same facial features in two different expressions

Figure 1.5. Glove fitting is an example of non-rigid shape similarity and correspondence problems. Correct correspondence (left) does not require significant stretching of the glove. Incorrect correspondence (right) results in a significant deformation.

of a face. This is exactly the problem of finding a deformation-invariant correspondence between two non-rigid shapes. The knowledge of the correspondence between faces allows us to perform different manipulations of them, including exaggeration of expressions, expression-invariant texture mapping, and *morphing* between faces.

In general, looking at the non-rigid world from the perspective of correspondence problems, we find applications more related to computer graphics and geometry processing (dealing mostly with synthesis), rather than to computer vision and pattern recognition (traditionally dealing with analysis problems).

1.3 A landscape of problems

To conclude our brief introduction to the forthcoming journey to the non-rigid world, let us overview the landscape of related fields and try to position the problems we will encounter in this book. In a broad sense, our problems belong to the realm of computer vision. As we mentioned, computer vision deals with extracting information about objects surrounding us from their visual representation. Traditionally, the most commonly used representation

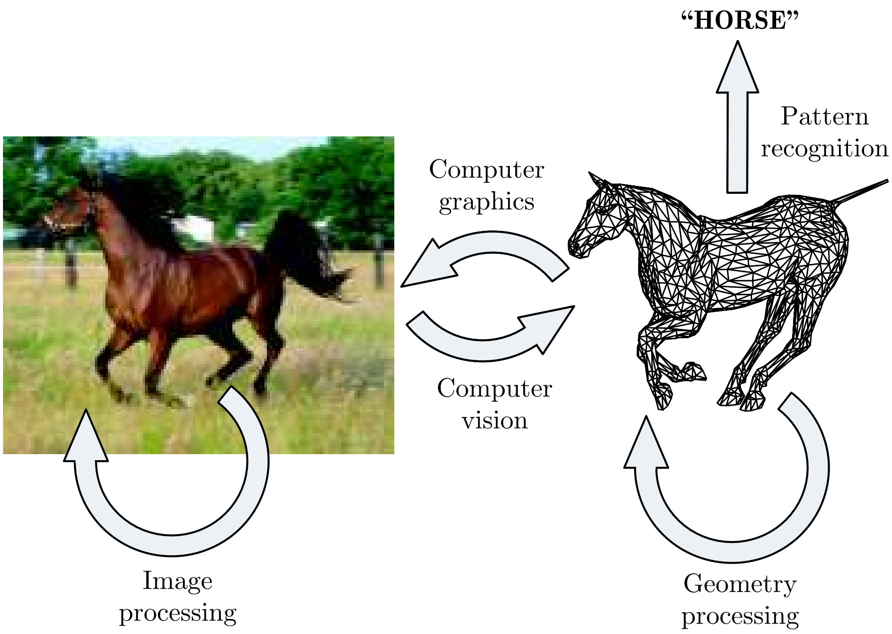

Figure 1.6. Conceptual representation of the relations between the fields of computer vision, computer graphics, image processing, pattern recognition, and geometry processing.

of the world is a two-dimensional image and the objects are geometric models, therefore, computer vision can be thought of as a "geometry from image" problem. A classic example is reconstructing the geometry of a scene from images taken by multiple cameras. The converse problem is addressed in the field of *computer graphics*: how to realistically and aesthetically render the description of the world, or how to produce an image from a geometric model. *Geometry processing* works with geometric models, trying to improve their quality or make their handling easier. *Image processing*, at the other end, operates on the images themselves, getting an image as the input and producing its "better" version as the output (Figure 1.6).

This division, once so clear to a specialist in each of the above fields, is becoming less obvious and less relevant in our days. For example, research results from the past decade have revealed a profound relation between geometry and image processing.[3] Considering images as geometric objects and operating on them using geometric tools created a revolution in image processing. The opposite process is taking place in the geometry processing community: it appears that by representing geometric objects as images, many efficient and powerful methods can be borrowed from image processing and adapted for geometry processing.

A similar situation seems to be happening in the analysis of non-rigid shapes: as we heavily rely on the model of the non-rigid world, our problems and methods reach far beyond "classic" computer vision. The tools we will use in this book range from theoretical geometry, used to model non-rigid objects, to methods used in manifold learning and artificial intelligence, which we employ for invariant representation of intrinsic geometry. In the next chapters, we will meet many disciplines and approaches brought together by problems of non-rigid shapes analysis. Such a diversity, in our opinion, makes research in this fascinating new field interesting and attractive.

Notes

[1]This citation from the Bard also appears as an epigraph in the Ph.D. thesis of Dragomir Anguelov [8].

[2]The moment at which the prosecution asked O. J. Simpson to put on a glove that allegedly had been used at the crime scene was the turning point of the trial, as the glove appeared too tight for Simpson to put on. The defense reflected this fact in the phrase "If it doesn't fit, you must acquit." As a result, the jury acquitted Simpson from murder charges.

[3]This relation is addressed in the book *Numerical Geometry of Images* [225].