Chapter 1

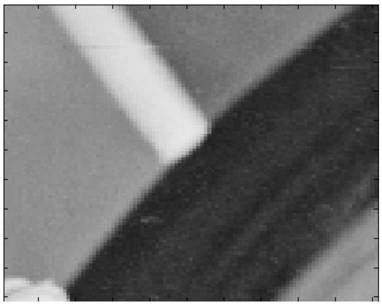
Introduction

Traditionally, the raw data of an image are often recorded in array form, with each array element representing the achromatic and/or chromatic information at a corresponding image pixel. In many applications, a direct use of this form may face three major drawbacks: bulkiness in data size, inefficiency in feature detection, and inflexibility in image content manipulation. The first drawback comes from the large number of pixels involved in an image. The second drawback comes from the fact that the spatial features, like boundaries and smooth regions, are not explicitly specified. The third drawback comes from the fact that we need to perform pixel-wise manipulation to change the color or shape of an object in an image. To find a more effective way to represent images, plenty of methods have been proposed in the literature [1]-[38]. For example, boundary-based methods, like chain code and signature [3]-[8], have been proposed to describe an image in terms of object boundaries. Region-based methods, like quadtree decomposition [9]-[11], have been proposed to describe an image in terms of smooth regions. Transform-based methods, like DCT (Discrete Cosine Transform) and DFT (Discrete Fourier Transform) [17]-[19], have been used to describe an image in terms of its transform coefficients. Multi-resolution methods, like Gaussian pyramid and wavelet decomposition, have been proposed to describe an image in hierarchical forms [20]-[26]. Each representation has its strong points and its suitable applications. In recent years, new methods are still emerging, trying to offer new ways to effectively represent images. For example, in [31], an image is represented in an edge-based approach by parametrically modeling relevant image surface variations. An approximation of the original image can be reconstructed under the framework of regularization theory. In [32], image is represented using irregular distributed samples in space. In [33], an image is represented by Gaussian Markov random field parameters in a multi-resolution way. In [34], the components obtained by multi-scale differential operators are used to represent images. In [35], multi-scale edges, accompanied with a suitable wavelet model, are used to decompose an image. In [36], a set of block pattern models that satisfy certain image variation constraints are used to represent images.

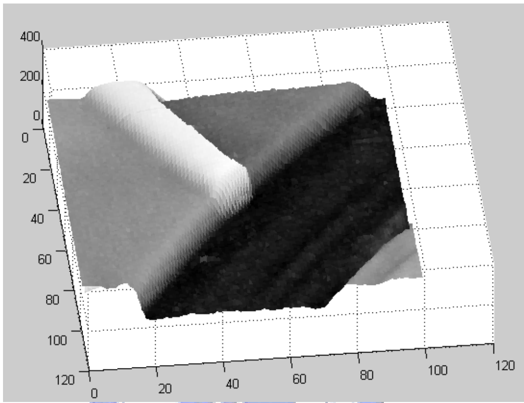
In this thesis, we propose an alternative image representation, which is based on high-curvature points on image surface. Given an intensity image, I(x,y), the surface

$$S(x,y) = (x,y,I(x,y)) \tag{1.1}$$

is called the "image surface". Fig 1.1(a) shows an image section clipped from the well-known image "lena". The corresponding image surface of Fig. 1.1(a) is shown in Fig. 1.1(b), where the altitude axis indicates the intensity value at each pixel. Over this image surface, flat regions map to smooth parts of the image, while cliff regions map to object boundaries in the image. Hence, an effective way to represent the image surface can also be used to effectively represent the original image. Furthermore, the image surface not only offers an analogous way to represent the intensity image, but also indicates an intuitive way to manipulate the image. As will be demonstrated in this thesis, many image processing applications, like image enhancement, image editing image feature detection, may actually be imagined as some kinds of manipulations over the image surface.



(a)



(b)

Fig. 1.1. Illustration of image surface (a) Subimage extracted from the upper-right area of “Lena”. (b) Corresponding image surface of (a).

Here, we first discuss the concept of the proposed representation scheme from the viewpoint of a 1-D intensity profile. Fig. 1.2(a) shows an intensity profile extracted from a real image. Fig. 1.2(b) shows the edge points extracted from this profile. Traditionally, these edge points are used as the major features of the profile. However, based merely on these edge points, the reconstruction of the original image is not an easy task at all. In this thesis, we adopt the use of some other feature that can effectively represent an intensity profile. In the use of the new image feature, we imagine an intensity profile as an elastic string stretched by a few pulleys, as shown in Fig. 1.2(c). Conceptually, these pulleys locate at highly curved places, and the radius of each pulley is inversely proportional to the local curvature of the profile. The location and size of these pulleys offer sufficient information to describe the outline of the profile. As shown in Fig. 1.2(d), both step edges and ridges can be well represented in this manner.

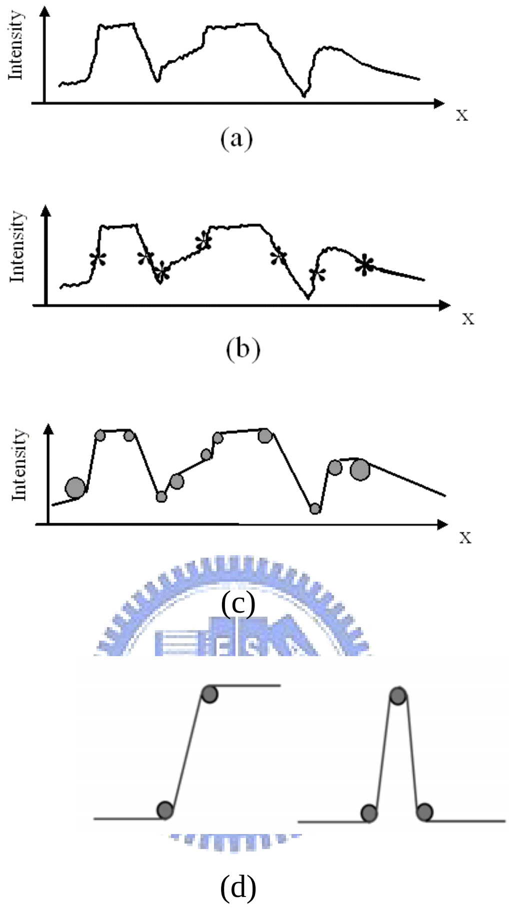


Fig. 1.2. Concept of the proposed method in one-D profile. (a) Original profile. (b) Edge points on the profile (marked as “*” signs). (c) Profile emulation using an elastic string and a few pulleys (verge points). (d) Representations for step edges (left) and ridges (right).

The concept of using pulleys to represent intensity profiles can be easily extended to represent 2-D images, as will be explained in Chapter 3. These extracted feature points are called “verge point” in this thesis. With this verge-point representation, image data can be greatly reduced. A visually similar approximation of the original image can also be reconstructed from these verge points via iterative linear interpolation, as will be explained in Chapter 3. By manipulating these verge points, various applications, like image compression, image feature detection, image enhancement, and image editing, can also be achieved. These functions will be described in Chapter 4 and Chapter 5. The overview of the proposed image representation scheme is presented in Fig. 1.3.

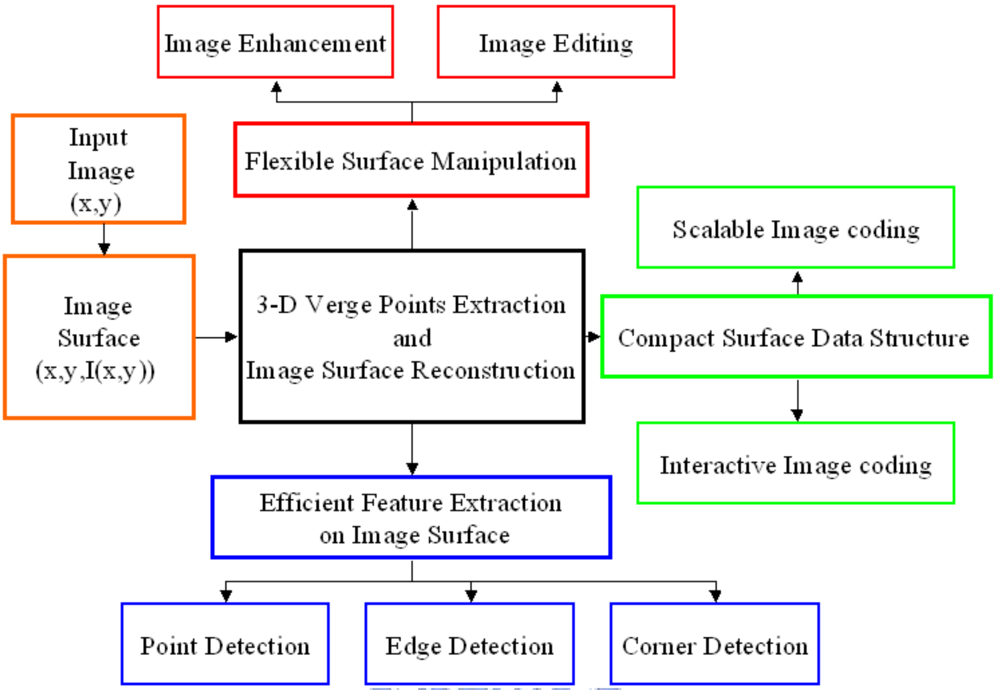


Fig. 1.3 Overview of the proposed image representation scheme

The concept of this new representation is first seen in [37]. In [37], two one-dimensional morphological operators, dilation operator and erosion operator, are used to detect verge points on image surfaces. Then, in [38], a two-dimensional Hessian operator is adopted to replace the one-dimensional operators proposed in [37]. In this thesis, based on the prototypes developed in [37] and [38], we further derive a comprehensive and versatile representation framework. In this new representation framework, the following accomplishments have been achieved.

(1) We have reinvestigated verge points extraction and image surface reconstruction.

- We have revisited the one-dimensional case using one-dimensional differential estimators. Analyses of 1-D feature extraction and parameter reduction are discussed.
- We have comprehensively reinvestigated the two-dimensional verge point detectors and developed some theoretic analyses over these operators.
- We have discussed the influences of different scales on the detection of verge points.
- The selection of curvature threshold has been explored. Noise variance in an image is estimated and then used for automatic curvature threshold determination.

(2) To construct a compact form to store verge points, we develop some hierarchical data structures.

- To efficiently represent these extracted image features, a new B-spline curve approximation scheme, which decomposes a feature curve into shape component and color component, is proposed. This decomposition improves both compactness and flexibility.
- Parameter selections in image encoding procedure, such as B-spline approximation allowance and intensity quantizer, are discussed.
- By arranging the data structure in different hierarchical forms, images can be transmitted with the capability of interactivity and scalability. The interactive transmission allows users to select “feature of interest” easily. The scalable coding features include
    - Spatial scalability.
    - SNR scalability
    - Shape scalability

(3) Based on the proposed representation features, some crucial elements for image understanding or object tracking can be efficiently detected. These fundamental features include

- Points.
- Lines.
- Edges.
- Corners.
- Smooth regions.

(4) To verify the flexibility of the proposed verge-point representation, we have developed a framework for image editing by properly manipulating these verge points. These manipulations include

- Contrast enhancement.
- Sharpness enhancement.
- Object-based intensity editing.
- Object-based shape editing.

1.3 Thesis Organization

This thesis is structured as follows. In Chapter 2, reviews of related works in the literature are introduced. The extraction of verge point and the reconstruction of image surface are described in Chapter 3. In Chapter 4, we discuss the diversified data structures for verge point representation. By arranging the verge points in different hierarchical forms, a scalable and interactive image transmission can be achieved. In Chapter 5, based on the proposed image representation, we present the process of detecting fundamental elements, such as edges and corners, for image understanding.

Moreover, image editing based on the manipulation of verge points is also presented. Finally, in Chapter 6, we conclude this thesis.