Analysis of Moment Invariants and its Uses for the Retrieval of Trademark Images

by

Lau Ka Lai

A Thesis

Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Philosophy

in the Department of Electronic and Information Engineering

The Hong Kong Polytechnic University

Hong Kong SAR, China

2003

Abstract

Derived from the general theory of moments, moment invariants have been frequently used as features for shape recognition and classification. Moments can provide characteristics of an object and uniquely represent its shape. Moment invariants are proven to be invariant under object translation, scaling and rotation. They are useful because they define a simple and fast set of seven elements.

In this research, an intensive study has been carried out to analyze the physical meaning of Hu moment invariants. From the second-order moments, the concept of best-fit ellipse has been formulated. This formulation is confirmed by both theoretical analysis and experimental work. The ratio of major and minor axes from the best-fit ellipse can be applied to describe the global characteristic of the object.

An analysis of quantization effects due to scaling and rotation of both regular and irregular objects is presented. The scaling errors for all approaches are large when the scaling factor is smaller than 0.5. Moreover, the rotational errors are big for the objects rotated other than the multiples of 90^{o}. Besides, this error analysis has been applied to object searching using a threshold selection scheme.

A new shape descriptor, called Multi-Layer Shape Descriptor (MLSD), making use of recursive sub-division technique on the basis of the best-fit ellipse concept has also been proposed. This local descriptor is invariant to translation, scaling and rotation. In our practical work, two-stage mechanism has been used to design a trademark image retrieval system. The first step acquires the global characteristic, i.e. the ratio of major and minor axes of the best-fit ellipse; while the second step obtains the local characteristics, i.e. MLSD. The system enhances 13.96% of the retrieval accuracy and improves the computational efficiency by

reducing the average number of comparison to 55.55% of that required by the other approaches.

Acknowledgements

I would like to take this opportunity to express my sincere gratitude to those people who have provided me valuable comments during my research.

First of all, I would like to thank my supervisor, Prof. W. C. Siu, for his feasible research guidance, which has made this thesis possible. His continuous support and inspiration made my research studies enjoyable. I considered myself very fortunate to be given an opportunity to work with Prof. Siu.

Special thanks are given to Dr. N. F. Law for her invaluable advice in the ideation of the shape descriptor. Her rich knowledge in mathematical analysis contributed a lot to this work. I would also like to thank our group members, Dr. Y. L. Chan, Mr. H. K. Cheung, Mr. C. P. Chow, Mr. K. T. Fung, Mr. K. C. Hui, Mr. W. L. Hui and Mr. W. H. Wong, Mr. Z. Y. Wang and Ms. H. L. Wong for their genial support and encouragement over my work.

Thanks also go to all members of the Centre of Multimedia Signal Processing who, in various ways, made my research studies a pleasant and memorable experience.

I would also express my gratitude to the Department of Electronic and Information Engineering and the Centre of Multimedia Signal Processing for providing me a comfortable working environment, and the Hong Kong Polytechnic University for their financial support to carry out my research work.

Last but not the least, I am deeply grateful to my family, for their love, trust and support throughout all these years.

Statement of Originality

The following contributions reported in this thesis are claimed to be original:

1. The interpretations of second-order moments give elucidation of the first and second elements of Hu moment invariants. The first invariant element measures the total spread relative to the area square. The second invariant element measures the degree of elongation of a best-fit ellipse on the shape. (Chapter 3, Section 3.2)
2. The interpretations of third-order moments give elucidation of the third and fourth elements of Hu moment invariants. Useful properties have also been found for objects with symmetric along $y = \pm x/\sqrt{3}$ and $y = \pm\sqrt{3}x$, when the third invariant is used. For the fourth invariant element, when fitting a best-fit ellipse on the object, the term of $a_f{}^2 + b_f{}^2$ is formed to be constant. (Chapter 3, Section 3.3)
3. An analysis of the quantization error due to scaling and rotation is presented. Both analytical expressions in continuous and discrete domains and experimental results shown that the Dudani moment invariants tend to have a larger error. (Chapter 4, Sections 4.3 and 4.4)
4. An error analysis is applied to object searching. It provides a way of choosing the threshold and is useful for ensuring that the two-stage approach is invariant to translation, scaling and rotation. (Chapter 4, Section 4.5)
5. Useful features are extracted from Multi-Layer Shape Descriptor

(MLSD), which are invariant to translation, scaling and rotation. We have also shown that the local characteristics can improve the retrieval accuracy of trademark image retrieval system by 13.96%. (Chapter 5, Section 5.2)

6. The improved two-step mechanism using both global and local characteristics gives 55.55% saving in terms of the average number of comparison for each image, as compared with the other approaches. (Chapter 5, Section 5.3)

Table of Contents

Abstract i
Acknowledgements iii
Statement of Originality iv
Table of Contents vi
List of Figures ix
List of Tables xi
List of Publications on which this Research is Based xii

Chapter 1 Introduction 1

1.1. Introduction 1
1.2. Text-based Image Retrieval 2
1.3. Content-based Image Retrieval 2
1.4. Information on Trademark Images 4
1.5. Feature Extraction on Shape 5
1.5.1. Region-based Geometric Techniques 6
1.5.2. Boundary-based Geometric Techniques 9
1.6. Current Shape-based Trademark Image Retrieval System 10
1.6.1. Query By Image Content (QBIC) 10
1.6.2. Automatic Retrieval of Trademark Images by Shape ANalysis (ARTISAN) 11
1.6.3. Using Negative Shape Features for Logo Similarity Matching 11
1.7. Organization of this Thesis 12

Chapter 2 Moment-based Shape Description Techniques 14

2.1. Introduction 14
2.2. Hu Moment Invariants 15
2.2.1. Definitions 15
2.2.2. Invariant Properties of Rotation, Translation and Scaling 20
2.3. Affine Moment Invariants 23
2.3.1. Affine Transformation 23
2.3.2. Definitions 24
2.4. Dudani Moment Invariants 28
2.4.1. Definitions 28
2.5. Improved Moment Invariants 29

2.5.1. Definitions 29
2.5.2. Invariant Properties of Rotation, Translation and Scaling 31
2.6. Sluzek's Moment-based Shape Descriptor 36
2.6.1. Definition 37

Chapter 3 Interpretations on Second and Third-order Moments 39

3.1. Introduction 39
3.2. Interpretations on Second-order Moments 41
3.2.1. First Moment Invariant Element: Total Spread Relative to the Area Square 41
3.2.2. Second Moment Invariant Element: Degree of Elongation of a Best-fit Ellipse on the Shape 42
3.2.3. Experimental Analysis 44
3.3. Interpretations on Third-order Moments 50
3.3.1. Third Moment Invariant Element 50
3.3.2. Fourth Moment Invariant Element 54
3.3.3. Experimental Analysis 54
3.4. Chapter Conclusion 59

Chapter 4 Improved Scheme for Object Searching using Moment Invariants 60

4.1. Introduction 60
4.2. Moments And Quantization Effects 62
4.2.1. Definition of Moment Invariants in Continuous Domain 62
4.2.2. Quantization Problem 62
4.3. Quantization Due to Scaling 63
4.3.1. Regular (rectangular) Objects 63
4.3.1.1. Hu Moment Invariants 63
4.3.1.2. Affine Moment Invariants 64
4.3.1.3. Dudani Moment Invariants 65
4.3.1.4. Improved Moment Invariants 65
4.3.1.5. Experimental Analysis 66
4.3.2. Irregular Objects 68
4.4. Quantization Due to Rotation 70
4.4.1. Regular (rectangular) Objects 70

4.4.1.1. Experimental Analysis 71
4.4.2. Irregular Objects 72
4.4.2.1. Experimental Analysis 72
4.5. Use of Error Analysis in Object Retrieval 74
4.6. Chapter Conclusion 78

Chapter 5 Multi-layer Shape Descriptor (MLSD) and Improved Mechanism with Both Global and Local Characteristics 80

5.1. Introduction 80
5.2. Multi-Layer Shape Descriptor (MLSD) 82
5.2.1. Recursive Sub-division of an Object on the Basis of the Best-fit Ellipse Concept Definitions 82
5.2.2. Parameters Invariant to Rotation, Scaling and Translation 84
5.3. Similarity Matching with both Global and Local Characteristics 86
5.3.1. The Framework of the Trademark Image Retrieval System 86
5.3.2. Detailed Descriptions of the Trademark Image Retrieval System 87
5.4. Experimental Results 90
5.5. Chapter Conclusion 94

Chapter 6 Conclusions 95

6.1. Conclusions of the Present Work 95
6.2. Future Work 97
6.2.1. Emptiness Problem 97
6.2.1.1. Filled objects 98
6.2.1.2. Trademark objects 100
6.2.2. Analysis on Moment-based Approaches with Different Deformations 104
6.2.3. Feature Extraction on Color 106

References 108

List of Figures

Figure 1-1	A traditional object searching model	3
Figure 1-2	Examples of three types of trademarks: (a) Character-in-mark, (b) Device-mark, and (c) Composite-mark	5
Figure 1-3	Taxonomy of shape description techniques	6
Figure 1-4	(a) Original logo image, (b) its version with border, (c) Negative component and (d) other components	12
Figure 2-1	(a) Unoccluded square, (b-c) Occluded square by circles with 25% and 100% of unoccluded square's area respectively and located at the centroid	36
Figure 2-2	(a) Unocculeded triangle, (b-c) Occluded triangle by circles with 25% and 100% of unoccluded triangle's area respectively and located at the centroid	37
Figure 3-1	Four rectangles with different af and bf in best-fit ellipse	43
Figure 3-2	(a) Equilateral triangle T1, (b-f) T2-T6, and (g) Square T7 with their best-fit ellipses	46
Figure 3-3	Three leaf-shaped objects with (a) $k=1$, (b) $k=1.5$, and (c) $k=2$ and Three horse-shaped objects with (d) $k=1$, (e) $k=1.5$, and (f) $k=2$	48
Figure 3-4	Three plots for leaf-shaped objects with constant width $b=108$	49
Figure 3-5	Three plots for horse-shaped objects with constant width $b=119$	49
Figure 3-6	Three plots for (a) Leaf-shaped and (b) Horse-shaped objects with elongation factor $k = 4.5$	50
Figure 3-7	A filled triangular as the object	51
Figure 3-8	Examples of filled triangular objects with symmetry along $y = \pm x/\sqrt{3}$	53
Figure 3-9	Examples of filled triangular objects with symmetry along $y = \pm\sqrt{3}x$	53
Figure 3-10	Five examples of filled triangular objects with five rotations	55
Figure 3-11	Examples of filled triangular objects	56
Figure 3-12	Examples of filled triangular objects	57
Figure 3-13	Examples of filled triangular objects	58
Figure 4-1	(a) Binary image in the continuous domain and (b) Sampled version of the binary image in the discrete domain	62

Figure 4-2	(a) Rectangular object and (b) its scaled version, (c) Horse-shaped objects and (d) its scaled version	63
Figure 4-3	Plot of error against scaling factor for rectangular objects for (a) Hu and Affine moment invariants and (b) Dudani and Improved moment invariants	67
Figure 4-4	Plot of error against scaling factor for irregular objects for (a) Hu and Affine moment invariants and (b) Dudani and Improved moment invariants	69
Figure 4-5	(a) Rectangular object and (b) its rotated version, (c) Horse-shaped objects and (d) its rotated version	70
Figure 4-6	Plot of error against the angle of rotation for rectangular objects for (a) Hu and Affine moment invariants and (b) Dudani and Improved moment invariants	71
Figure 4-7	Plot of error against the angle of rotation for irregular objects for (a) Hu and Affine moment invariants and (b) Dudani and Improved moment invariants	73
Figure 4-8	A two-stage object searching model	75
Figure 4-9	Some example objects in the test database	76
Figure 5-1	The object after first sub-division	83
Figure 5-2	The object after second sub-division	84
Figure 5-3	A quadtree representation of object with n-level decompositions	84
Figure 5-4	Example image with global characteristic	87
Figure 5-5	A flow diagram of the trademark image retrieval system	90
Figure 5-6	Example images in Set 1 and Set 2	91
Figure 6-1	Three irregular filled objects	98
Figure 6-2	Four trademark objects	100
Figure 6-3	A rectangular object with width of edge w_1 and w_2	101
Figure 6-4	Actual values of four approaches while varying width of edge of rectangular objects	103
Figure 6-5	Actual values of the (a) First element and (b) Second element of Improved moment invariants on the elliptical and rectangular objects	104
Figure 6-6	(a) Original object, objects after (b) Cylindar-horizontal 45%, (c) Cylinder-vertical 45%, (d) Perspective-horizontal 45%, (e) Perspective-vertical 45%, (f) Pinch 45, (g) Punch 45, (h) Skew-horizontal 45, (i) Skew-horizontal -45, (j) Skew-vertical 45, and (k) Skew-vertical -45	105
Figure 6-7	Extraction of color blob histogram	107

List of Tables

Table 3-1	Results for four rectangles with constant $b=93$	45
Table 3-2	Results for four triangles with $b = a / 2$ and $c=104$	46
Table 3-3	Results of seven trapeziums with $a = b+d$ and $c=100$	47
Table 3-4	Experimental readings of the triangular objects shown in Fig. 3-10(a-e)	55
Table 3-5	Experimental readings of the triangular objects shown in Fig. 3-11(a-d)	56
Table 3-6	Experimental readings of the triangular objects shown in Fig. 3-12 (a-f)	57
Table 3-7	Experimental readings of the triangular objects shown in Fig. 3-13 (a-d)	58
Table 4-1	Results of different selections	78
Table 5-1	Retrieval accuracy of Set 1	92
Table 5-2	Retrieval accuracy of Set 2	93
Table 6-1	Actual values of Hu moment invariants with different scaling factors	99
Table 6-2	Actual values of Hu moment invariants with different rotational angles	99
Table 6-3	Actual values of Hu moment invariants for Fig. 6-2(a-d)	100

List of Publications on which this Research is Based

International Journal Papers

1. K. L. Lau, W. C. Siu and N. F. Law, "Best-fit Ellipse Concept from Second Order Moments," Paper is submitted to Journal of Information Technology and Applications.
2. K. L. Lau, W. C. Siu and N. F. Law, "Multi-Layer Shape Descriptor Invariants under Rotation, Scaling and Translation for Object Searching," Paper completed to be submitted for publication in a journal.

International Conference Papers

1. K. L. Lau, W. C. Siu and N. F. Law, "Best-fit Ellipse Concept from Second Order Moments," *IEEE International Conference on Information Technology and Applications, ICITA 2002*, Bathurst, Australia, pp.167-5, 25-29 November 2002.
2. K. L. Lau, W. C. Siu and N. F. Law, "Improved Scheme for Object Searching using Moment Invariants," *Advances in Multimedia Information Processing - The 3rd IEEE Pacific-Rim Conference on Multimedia, PCM 2002*, Huinchu, Taiwan, pp. 783-790, 16-18 December 2002.

Chapter 1

Introduction

1.1. Introduction

In recent year, the amount of digital image collections has been increasing rapidly. These image collections play an important role for describing, representing, and storing for various applications, like military reconnaissance, weather forecasting, criminal investigation, biomedical imaging, finger printing, art galleries, trademark images identification, etc.

Applications that consider image retrieval as a principal activity are both numerous and diverse. In order to access such huge amount of information, there should be good means to browse, search, and retrieval the images. Two major

approaches are investigated to cope with this problem. One is the traditional text-based image retrieval while the other is content-based image retrieval. The former retrieval approach is to interpret each image manually with text labels and attributes (such as filenames, captions, and keywords). These textual descriptions try to describe the contents of images. This approach is to extract certain characteristics from the images directly. These are based on the characteristics of visual image contents.

1.2. Text-based Image Retrieval

Text-based image retrieval [1-3] is originated from late 1970's. This approach uses the textual features to annotate and retrieve images; however it has insuperable problems. The major problem is human intervention. Creating annotations by hands can become hopeless and very time-consuming, because of the huge amount of database and error-prone problem. The next problem is that it is a highly subjective task for indexing the contents of images manually. Different people may have different perceptibilities over the same image. Moreover, in some cases, the keyword becomes complex and it is very difficult to describe relevant contents of images by just simple phrases. Besides, if the textual features of the image database are to be worldwide shared (such as images from World Wide Web [4-6]), the linguistic barriers and the disordered source of images cannot be practically provided by an effective declaration.

1.3. Content-based Image Retrieval

In early 90's, the emergence of large-scale image database [7-9] has emphasized the difficulties of manual annotation. To overcome the outstanding need for effective and efficient indexing, content-based image retrieval (CBIR) [10-14] was proposed and it has been under intensive research from then on. Content-based

image retrieval is to investigate the techniques of image annotation and retrieval based on the visual and semantic content instead of text labels and attributes of the images. This kind of techniques is however more challenging to investigate and implement, because it is not obvious how to extract such relevant content from images, which would in addition be suitable for searching and retrieval.

Object recognition [15-17] is the process of identification of shapes, forms or configurations by automatic means, with minimal interaction with a human. The basic structure of a traditional object searching model is shown as Fig. 1-1. For the image archival part, each input image is scanned and its feature vector is extracted. These feature vectors will be stored into an image database. Once a query image is submitted for searching those relevant images in the database, the image searching part begins. The feature of the query image is extracted and then used to match with all features vectors in the database. After the matching process, the resultant relevant images are displayed to the user.

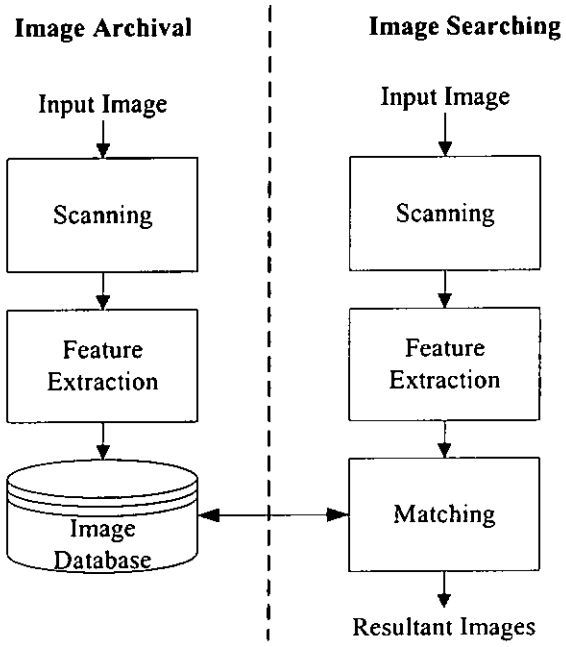


Figure 1-1 A traditional object searching model

1.4. Information on Trademark Images

Trademarks are specially a word, personal names, letters, numerals, figurative elements or designed marks that identify companies, products, and services. They also aim at distinguishing the source of goods or services of one party from those of others. A successful trademark can build up the proprietor's business goodwill. It not only can bring profitable interest, but also help to promote the proprietor's business. A registered trademark gives the exclusive legal right to use, license or sell it for the goods and services for which it is registered. The imitation of registered trademark is illegal. However, there are so many trademarks around the world and how to avoid designing a trademark similar to an existing one becomes an important problem. Owing to the large intake of applications, the registry [18-20] takes about 12 months for a trademark registration normally. Therefore, there is a great demand for an automatic trademark retrieval system based on shape feature.

In general, trademarks are disjoint complex pattern consisting in various natural or geometric objects (such as sun, mountains, circle or triangle etc) and text. Trademarks can be divided into three types: character-in-mark, device-mark, and composite-mark. Some examples are shown in Fig.1-2. For those character-in-mark trademarks, it contains only characters or words. The device-mark trademarks have graphical or figurative elements only, while the composite-mark trademarks combine the characteristics of both character-in mark and device-mark trademarks. That is the composite-mark one contains both characters and graphical elements.

Figure 1-2 Examples of three types of trademarks: (a) Character-in-mark, (b) Device-mark, and (c) Composite-mark

1.5. Feature Extraction on Shape

Many techniques of shape description and recognition have been proposed [21-30] since the early 1990's. The taxonomy [31-32] of different shape description techniques is shown in Fig.1-3. Usually they are classified as Region- or Boundary- based techniques. Region-based techniques take into account internal details and boundary details. On the other hand, the boundary-based techniques use only the contour or border of an object.

Techniques of both classes can be sub-classified into spatial and transform domains. The spatial domains techniques measure the appearance of the objects. They are used when the description is made from the area of the object inside its boundary. For the transform domains techniques, they are based on mathematical transforms such as Fourier- [33-35] and Wavelet- [36-39] transforms. In this thesis, the main scope of discussions is on the geometric techniques.

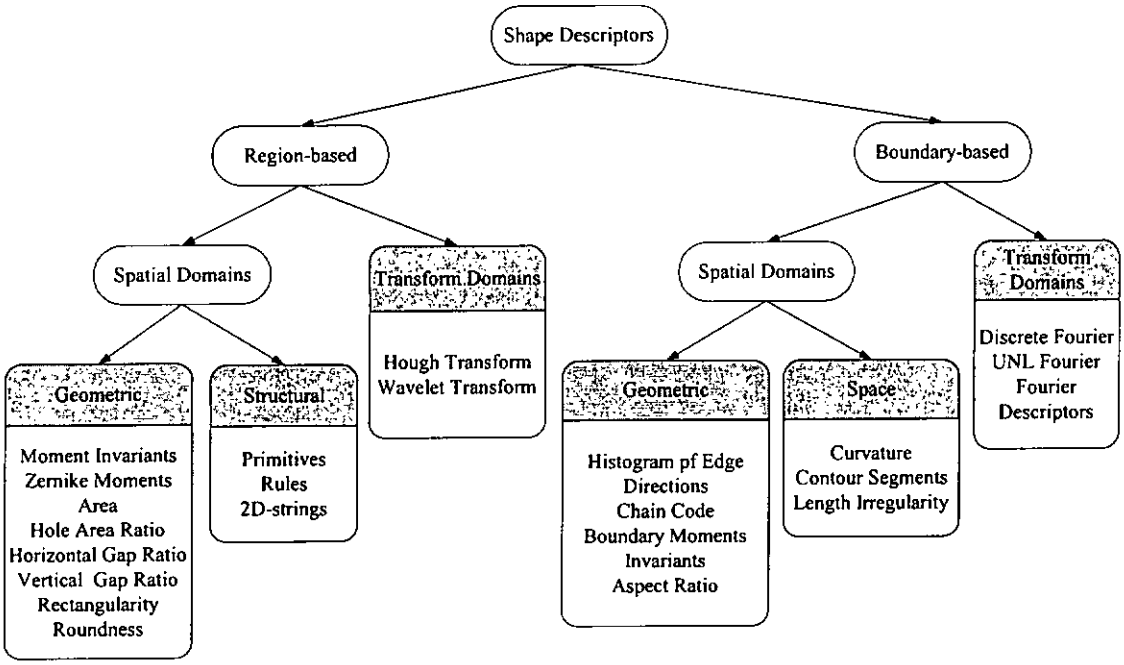


Figure 1-3 Taxonomy of shape description techniques

1.5.1. Region-based Geometric Techniques

Region-based geometric techniques [40, 42-47] apply the mathematical properties derived from the objects. They take consideration on the interior details within a complete contour of the object. The calculations of this kind of techniques are usually performed for binary images.

A. Moment Invariants

Moment Invariants is the most popular shape description technique. Hu [40] proposed the fundamental derivations of the descriptor, which is invariant under translation, rotation and scaling. The $(p+q)^{th}$ order moments of a continuous density distribution function $f(x,y)$ are defined as,

$$m_{pq} = \iint_R x^p y^q f(x,y) dxdy \tag{1.1}$$

where $p,q = 0,1,2,.....$

The central moments of $f(x,y)$ are defined as,

$$\mu_{pq} = \iint_R (x-\bar{x})^p (y-\bar{y})^q f(x,y)dxdy \tag{1.2}$$

where $\bar{x} = m_{10}/m_{00}, \bar{y} = m_{01}/m_{00}$.

Eqns. 1.3 to 1.9 show a set of seven Hu moment invariants [40]. These moment invariants are derived from the second- and the third-order moment only.

$$\phi_1 = (\mu_{20} + \mu_{02}) / \mu_{00}^2 \tag{1.3}$$

$$\phi_2 = \left((\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2\right) / \mu_{00}^4 \tag{1.4}$$

$$\phi_3 = \left((\mu_{30} - 3\mu_{12})^2 + (\mu_{03} - 3\mu_{21})^2\right) / \mu_{00}^5 \tag{1.5}$$

$$\phi_4 = \left((\mu_{30} + \mu_{12})^2 + (\mu_{03} + \mu_{21})^2\right) / \mu_{00}^5 \tag{1.6}$$

$$\phi_5 = \left(\begin{array}{l} (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})\left[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2\right] + \\ (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})\left[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2\right] \end{array} \right) \Big/ \mu_{00}^{10} \tag{1.7}$$

$$\phi_6 = \left(\begin{array}{l} (\mu_{20} - \mu_{02})\left[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2\right] + \\ 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) \end{array} \right) \Big/ \mu_{00}^7 \tag{1.8}$$

$$\phi_7 = \left(\begin{array}{l} (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})\left[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2\right] + \\ (3\mu_{12} - \mu_{03})(\mu_{21} + \mu_{03})\left[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2\right] \end{array} \right) \Big/ \mu_{00}^{10} \tag{1.9}$$

Several modifications and improvements in calculation have been proposed. For example Wong, Siu et al. [41] expressed the set of moment invariants, Flusser and Suk [42-43] investigated a new set of descriptor using the affine transform and Dudani et al. [44] applied to aircraft identification.

B. Zernike Moments

Zernike moments have been proposed by Khotanzad and Hong [45-47] for object recognition. The Zernike moments are derived from radial polynomials and formed a complete orthogonal set over the interior of the unit circle. The radial polynomials vector $R_{nm}(\rho)$ is defined as,

$$R_{nm}(\rho) = \sum_{k=0}^{\frac{n-|m|}{2}} (-1)^k \frac{(n-k)!}{k!(\frac{n+|m|}{2}-k)!(\frac{n-|m|}{2}-k)!} \rho^{n-2k} \tag{1.10}$$

where n $\in$ **N** $\cup$ {0}, m $\in$ **Z**, $|m| \leq$ n and n-$|m|$ is even.

The Zernike moments of order n with repetition m is defined as,

$$A_{nm} = \frac{n+1}{\pi} \iint_{x^2+y^2 \leq 1} f(x,y) V_{nm}^*(\rho,\theta) dxdy \tag{1.11}$$

where $\{V_{nm}(x,y)\}$ is the set of Zernike polynomials and is defined as,

$$V_{nm}(x,y) \cong V_{nm}(\rho,\theta) = R_{nm}(\rho) e^{-jm\theta} \tag{1.12}$$

C. Rectangularity

Rectangularity [48-49] measures how well an object fills in its minimum enclosing rectangle. It is independent of the translation, rotation and scaling of the objects. The definition of rectangularity of an object is,

$$Rectangularity = \frac{A_{Object}}{A_{\min R}} \tag{1.13}$$

where A_{Object} is the area of the object and A_{minR} is the area of its minimum rectangle. Hence a square or rectangle obtains value 1 and other shapes values less than one.

D. Roundness/ Circularity

Roundness [48-49], so-called circularity, measures the ratio of the perimeter and area of the object. It is also independent of the translation, rotation and scaling of the objects. The definition of roundness of object is,

$$Roundness = \frac{P_{Object}^2}{A_{Object}} \tag{1.14}$$

where P_{Object} is the length of the perimeter of the region and A_{Object} is the area of the object. Han et al [28] proposed to use the inverse of the roundness and

multiply of 4π to normalize the feature. So, a circle obtains value 1 and other shapes values less than one.

1.5.2. Boundary-based Geometric Techniques

Boundary-based geometric techniques [50-52, 57] are derived from the boundaries of the objects. Some of them are required the techniques of segmentation or edge-detection.

A. Histogram of Edge Directions

The histogram of edge directions [29-30, 53] is a popular boundary-based shape descriptor. All images are preprocessed before extracting the edges and their directions. Canny operator [54], Sobel operator [55] or Gaussian derivatives [56] can be applied in this edge detecting preprocessing. Then the directions of the edges are quantized and the required histogram is form. Moreover, Jain et al [29] proposed a way to make the histogram much smoother and less sensitivity to rotation. The smoothed histogram is defined as,

$$H_e^s(i) = \frac{\sum_{j=i-k}^{i+k} H_e(j)}{2k+1} \tag{1.15}$$

where H_e is the original histogram of edge directions and k is the degree of smoothness.

B. Boundary Moment Invariants

Improved moment invariants [57] are the modified version of Hu moment invariants [40] and they are applied to the shape boundary rather than the region. The main difference is on the normalization factor. The seven moment invariants for improved moment invariants, denoted as ψ_1 through ψ_7, become,

$$\psi_1 = (\mu_{20} + \mu_{02}) / \mu_{00}^{3} \tag{1.16}$$

$$\psi_2 = \left((\mu_{20} - \mu_{02})^2 + 4\mu_{11}^{2}\right) / \mu_{00}^{6} \tag{1.17}$$

$$\psi_3 = \left((\mu_{30} - 3\mu_{12})^2 + (\mu_{03} - 3\mu_{21})^2\right) / \mu_{00}^{\,8} \tag{1.17}$$

$$\psi_4 = \left((\mu_{30} + \mu_{12})^2 + (\mu_{03} + \mu_{21})^2\right) / \mu_{00}^{\,8} \tag{1.18}$$

$$\psi_5 = \left(\begin{array}{l} (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})\left[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2\right] + \\ \quad (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})\left[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2\right] \end{array} \right) \Big/ \mu_{00}^{\,16} \tag{1.19}$$

$$\psi_6 = \left(\begin{array}{l} (\mu_{20} - \mu_{02})\left[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2\right] + \\ \quad 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) \end{array} \right) \Big/ \mu_{00}^{\,11} \tag{1.20}$$

$$\psi_7 = \left(\begin{array}{l} (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})\left[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2\right] + \\ \quad (3\mu_{12} - \mu_{03})(\mu_{21} + \mu_{03})\left[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2\right] \end{array} \right) \Big/ \mu_{00}^{\,16} \tag{1.21}$$

1.6. Current Shape-based Trademark Image Retrieval System

Many famous shape-based trademark image retrieval systems [9, 31, 58-64] have been built. The best known are Query By Image Content (QBIC), Automatic Retrieval of Trademark Images by Shape Analysis (ARTISAN) and Matching using negative shape feature.

1.6.1. Query By Image Content (QBIC)

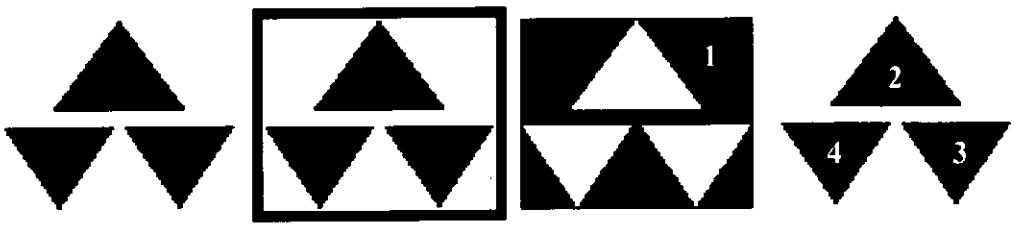
The shape features of Query By Image Content were presented by Flickner et al. [65-66]. Binary images in the database are firstly segmented so as to extract the objects. Then the object is represented by geometric shape features, which includes area, roundness, invariant moments, etc. The similarity distance between two images is obtained by using a weighted Euclidean distance. Based on these kinds of techniques, a trademark retrieval system is developed. Each trademark image is segmented to form representative components. Then the shape features for each component are stored in a database. In fact, the user can submits a query by selecting one of the representative components. The size of the database contains 1000 trademark images.

1.6.2. Automatic Retrieval of Trademark Images by Shape ANalysis (ARTISAN)

The development of the Automatic Retrieval of Trademark Images by Shape Analysis [67-68] aims at providing a prototype of shape retrieval for trademark images. The approach is firstly to extract region boundaries of a shape and represented them by straight-line and circular-arc approximately. Then the shape features of aspect ratio, roundness, relative area, directedness, straightness, etc are extracted and stored into the database. ARTISAN was experimented with more than 10000 trademark images.

1.6.3. Using Negative Shape Features for Logo Similarity Matching

Soffer and Samet [69-70] defined a representation of negative symbols, which is based on their interior with the shapes considered as holes. The authors aim at artificially adding 4-pixels width border around the logos and creating a negative symbol. Each logo is segmented into its constituent components and each region is given a number. The component labeled "1" in Fig. 1-4 is the negative component resulting from adding the border. Shape features are extracted for each component. There are four global shape descriptors (first invariant moment, circularity, eccentricity and rectangularity) and three local shape descriptors (horizontal gaps per total area, vertical gaps per total area and radio of hole area to total area). This matching was experimented with 130 images.



(a) (b) (c) (d)

Figure 1-4 (a) Original logo image, (b) its version with border, (c) Negative component and (d) other components

1.7. Organization of this Thesis

Hu introduced an object recognition approach to describe the shape information from the binary image based on the general linear transform. It is intended to propose a set of seven moment invariants to characterize shape feature of images, which are invariant under translation, rotation and scaling. Several moment-based approaches have been investigated for handling different aspects of the recognition. This study aims at providing some meaningful interpretations of the fundamental Hu moment invariants. From the properties of lower order moments, an improved two-stage mechanism is developed so as to improve the efficiency of the retrieval system. Our study also evaluates the quantization/ digitization effects on rotation and scaling for these four moment-based approaches. In order to have a better shape descriptor, a multi-layer shape descriptor is proposed. The dissertation is organized as follows.

Chapter 2 explores the theoretical background and mathematical details of five moment-based approaches (they are Hu moment invariants, Affine moment invariants, Dudani moment invariants, Improved moment invariants and Sluzek's moment-based shape descriptor). Chapter 3 presents the interpretations of second and third-order Hu moments. Chapter 4 shows the analysis of quantization effect

due to scaling and rotation. A proposed improved scheme for object searching using moment invariants is also described. Chapter 5 gives a detailed description of the proposed Multi-Layer Shape Descriptor (MLSD) and the improved mechanism with both global and local characteristics. Chapter 6 concludes this thesis by summarizing out contributions in our study and suggesting future works.

Chapter 2

Moment-based Shape Description Techniques

2.1. Introduction

The abilities of shape descriptors to provide similar shape features for objects with different rotations, scaling and locations are essential considerations in the field of object recognition. Two-dimensional moment invariants have been developed to extract the features of objects and their invariance properties of general shapes have been taken into account. The usefulness of these invariants in image analysis and object identification has been testified in many monographs. They are also applied in practical applications.

The fundamental derivations of moment invariants were first investigated by

Hu [40]. A set of seven moment invariants equations, usually referred to as the Hu moment invariants, is developed. Afterward, a large amount of publications, which focus on a wide range of aspects (such as the enhanced approaches of the moments [42-44, 57, 72], the application-based developments [21, 61, 73], the error analysis [74-75], etc) are released. In this chapter, a simple mathematical elaboration of the fundamental theory (i.e. Hu Moment Invariants) is presented. We first show the derivations of Raw Moments, Central Moments, and Normalized Central Moments. Then the elaborations on four moment-based approaches are discussed. They are Affine moment invariants [42-43], Duandi moment invariants [44], Improved moment invariants [57] and Sluzek's moment-based approach [72].

2.2. Hu Moment Invariants

Hu moment invariants [40] are simple among all moment functions, with the kernel function defined as a product of pixel coordinates. The main advantage is that image coordinate transformations can be easily expressed and analyzed in terms of the corresponding transformations in the moment space. They are invariant with respect to image plane transforms. The computations of these moments for images are simple. Hu moment invariants are geometric moments, sometimes referred to as Cartesian moments, or regular moments.

2.2.1. Definitions

An image can be considered as a two-dimensional density distribution $f(x,y)$, where the function value denotes the intensity at the pixel location (x,y). Let R be the image region of the x-y plane. A general definition of moment functions ξ_{pq} of order $(p+q)$, of the function $f(x,y)$ can be defined as,

$$\xi_{pq} = \iint_R \Psi_{pq}(x,y) f(x,y) dxdy \tag{2.1}$$

where moment kernel $\Psi_{pq}(x,y)$ is a continuous function of (x,y) in R, p and q are the indices of the coordinates x, y respectively.

A. Raw Moments

Raw moments are defined with the kernel $\Psi_{pq}(x,y) = x^p y^q$ (c.f. eqn. 2.1). The two-dimensional $(p+q)^{\text{th}}$ order moments of a continuous density distribution function $f(x,y)$ are defined as,

$$m_{pq} = \iint_R x^p y^q f(x,y) dxdy \tag{2.2}$$

where $p,q = 0,1,2,.....$

Uniqueness theorem: If the intensity function $f(x,y)$ is piece-wise continuous and bounded in the region R, the moment sequence is uniquely determined by $f(x,y)$ and conversely.

Existence theorem: If the intensity function $f(x,y)$ is piece-wise continuous and bounded in the region R, the moment sequence is uniquely determined by $f(x,y)$ and conversely.

In practice, moments have to be calculated on digital images and discrete summations have to be used for approximation instead of the continuous integral. Raw moments for a digital image are defined as,

$$m_{pq} = \sum_{y=1}^{M} \sum_{x=1}^{N} x^p y^q f(x,y) \tag{2.3}$$

where $p,q = 0,1,2,....., M$ and N are the horizontal and vertical dimensions, respectively, of the intensity function $f(x,y)$.

Considering the order up to third moments, they are shown as follows,

Zero-order moment: $m_{00} = \sum_{y=1}^{M} \sum_{x=1}^{N} f(x,y)$

First-order moment: $m_{10} = \sum_{y=1}^{M}\sum_{x=1}^{N} xf(x,y)$; $m_{01} = \sum_{y=1}^{M}\sum_{x=1}^{N} yf(x,y)$

Second-order moment: $m_{20} = \sum_{y=1}^{M}\sum_{x=1}^{N} x^2 f(x,y)$; $m_{02} = \sum_{y=1}^{M}\sum_{x=1}^{N} y^2 f(x,y)$;

$$m_{11} = \sum_{y=1}^{M}\sum_{x=1}^{N} xyf(x,y)$$

Third-order moment: $m_{30} = \sum_{y=1}^{M}\sum_{x=1}^{N} x^3 f(x,y)$; $m_{03} = \sum_{y=1}^{M}\sum_{x=1}^{N} y^3 f(x,y)$;

$$m_{21} = \sum_{y=1}^{M}\sum_{x=1}^{N} x^2 yf(x,y); \quad m_{12} = \sum_{y=1}^{M}\sum_{x=1}^{N} xy^2 f(x,y)$$

B. Central Moments

By definition, the moment of order zero, m_{00}, represents the total intensity of an image. For a binary image, this term gives the geometrical area of the image region.

The first-order moments, m_{10} and m_{01}, provide the intensity moment about the y-axis and x-axis of the images respectively. The centroid, $(\bar{x}, \bar{y})$, is defined as,

$$\bar{x} = \frac{m_{10}}{m_{00}} \tag{2.4}$$

$$\bar{y} = \frac{m_{01}}{m_{00}} \tag{2.5}$$

For a binary image, the centroid gives the geometrical center of the image region. It is often convenient to evaluate the moments with the origin of the reference system shifted to the centroid of the image. This transformation makes the moment computation independent of the position of the image reference system. The moments computed with respect to the centroid is called central moments, and is defined as,

$$\mu_{pq} = \iint_R (x-\bar{x})^p (y-\bar{y})^q f(x,y)dxdy \tag{2.6}$$

where $p,q = 0,1,2,.....$

Central moments for a digital image are approximated by double summations and are defined as,

$$\mu_{pq} = \sum_{y=1}^{M}\sum_{x=1}^{N}(x-\bar{x})^p(y-\bar{y})^q f(x,y) \tag{2.7}$$

where $p,q = 0,1,2,\ldots..$, M and N are the horizontal and vertical dimensions, respectively, of the intensity function $f(x,y)$=1.

From the definition, the central moments can be expressed in terms of raw moments. We have,

$$\mu_{00} = m_{00} \tag{2.8}$$

$$\mu_{10} = \mu_{01} = 0 \tag{2.9}$$

$$\mu_{20} = m_{20} - \bar{x}m_{10} \tag{2.10}$$

$$\mu_{02} = m_{02} - \bar{y}m_{01} \tag{2.11}$$

$$\mu_{11} = m_{11} - \bar{y}m_{10} \tag{2.12}$$

$$\mu_{30} = m_{30} - 3\bar{x}m_{20} + 2\bar{x}^2 m_{10} \tag{2.13}$$

$$\mu_{03} = m_{03} - 3\bar{y}m_{02} + 2\bar{y}^2 m_{01} \tag{2.14}$$

$$\mu_{21} = m_{21} - 2\bar{x}m_{11} - \bar{y}m_{20} + 2\bar{x}^2 m_{01} \tag{2.15}$$

$$\mu_{12} = m_{12} - 2\bar{y}m_{11} - \bar{x}m_{02} + 2\bar{y}^2 m_{10} \tag{2.16}$$

Eqns. 2.17-2.23 show a set of seven moment invariants proposed by Hu [40]. These moment invariants are invariant under both translation and rotation and are derived from the second- and the third-order moment only.

$$I_1 = \mu_{20} + \mu_{02} \tag{2.17}$$

$$I_2 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^{\ 2} \tag{2.18}$$

$$I_3 = (\mu_{30} - 3\mu_{12})^2 + (\mu_{03} - 3\mu_{21})^2 \tag{2.19}$$

$$I_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{03} + \mu_{21})^2 \tag{2.20}$$

$$\begin{aligned} I_5 = {} & (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})\left[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2\right] + \\ & (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})\left[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2\right] \end{aligned} \tag{2.21}$$

$$\begin{aligned} I_6 = {} & (\mu_{20} - \mu_{02})\left[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2\right] + \\ & 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) \end{aligned} \tag{2.22}$$

$$\begin{aligned} I_7 = {} & (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})\left[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2\right] + \\ & (3\mu_{12} - \mu_{03})(\mu_{21} + \mu_{03})\left[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2\right] \end{aligned} \tag{2.23}$$

Image features characterized by I_1 to I_7 are independent of its position and orientation. They are, however, not scaling invariant. Let us discuss normalized central moments which are to achieve scale invariant.

C. Normalized Central Moments

The normalized central moments replace μ_{pq} with η_{pq} to achieve scaling invariant. The definition of η_{pq} is as follow,

$$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^{(p+q+2)/2}} \tag{2.29}$$

Then, the seven Hu moment invariants, which are invariant with respect to translation, scale and rotation of an image, denoted as ϕ_1 through ϕ_7, become.

$$\phi_1 = \frac{I_1}{\mu_{00}^{\ 2}} \tag{2.30}$$

$$\phi_2 = \frac{I_2}{\mu_{00}^{\ 4}} \tag{2.31}$$

$$\phi_3 = \frac{I_3}{\mu_{00}^{5}} \tag{2.32}$$

$$\phi_4 = \frac{I_4}{\mu_{00}^{5}} \tag{2.33}$$

$$\phi_5 = \frac{I_5}{\mu_{00}^{10}} \tag{2.34}$$

$$\phi_6 = \frac{I_6}{\mu_{00}^{7}} \tag{2.35}$$

$$\phi_7 = \frac{I_7}{\mu_{00}^{10}} \tag{2.36}$$

2.2.2. Invariant Properties of Rotation, Translation and Scaling

A. Rotation Invariant

If an image is rotated by an angle θ, the transformation of its pixel coordinates becomes,

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{2.37}$$

where x', y' are the new coordinates and x, y are the original coordinates.

It is always desirable to derive a set of moment terms which are independent of the rotation angle θ. Using second-order moments as an example, their expressions for a rotated image are,

$$m'_{20} = \left(\frac{1+\cos\theta}{2}\right) m_{20} - (\sin 2\theta) m_{11} + \left(\frac{1-\cos\theta}{2}\right) m_{02} \tag{2.38}$$

$$m'_{02} = \left(\frac{1-\cos\theta}{2}\right) m_{20} + (\sin 2\theta) m_{11} + \left(\frac{1+\cos\theta}{2}\right) m_{02} \tag{2.39}$$

$$m'_{11} = \left(\frac{\sin 2\theta}{2}\right) m_{20} + (\cos 2\theta) m_{11} - \left(\frac{\sin 2\theta}{2}\right) m_{02} \tag{2.40}$$

From eqns. 2.38-2.39, it is easily seen that $(m_{20}' + m_{20}') = (m_{20} + m_{20})$ is independent of θ, hence is rotation invariant. Note that, by definition, the raw moments given in eqns. 2.2-2.3 are rotation invariant.

B. Translation Invariant

Consider the following linear shifting of image coordinates from (x,y) to (x',y'),

$$x' = x + a \tag{2.41}$$

$$y' = y + b \tag{2.42}$$

where x, y are the original coordinates and x', y' are the new coordinates.

The raw moments of these two coordinates are defined as,

$$m_{pq} = \iint_R x^p y^q f(x,y)dxdy \tag{2.43}$$

$$m'_{pq} = \iint_R x'^p y'^q g(x',y')dx'dy' \tag{2.44}$$

where m_{pq} and m'_{pq} are the raw moments in x, y and x', y' coordinates respectively.

For zeroth-order moment: $m'_{00} = \iint_R g(x',y')dx'dy' = \iint_R f(x,y)dxdy = m_{00}$. It is because the intensity functions in region R of the images are the same (i.e. g(x',y') = f(x,y)) and eqns. 2.41-2.42 imply that $dx'=dx$ and $dy'=dy$.

For first-order moment:

$$m'_{10} = \iint_R x'g(x',y')dx'dy' = \iint_R (x+a)f(x,y)dxdy = m_{10} + a(m_{00});$$

$$m'_{01} = \iint_R y'g(x',y')dx'dy' = \iint_R (y+b)f(x,y)dxdy = m_{01} + b(m_{00}).$$

Then the centroid $(\bar{x}', \bar{y}')$ in x', y' coordinates becomes,

$$\bar{x}' = \frac{m'_{10}}{m'_{00}} = \frac{m_{10} + am_{00}}{m_{00}} = \bar{x} + a \tag{2.45}$$

$$\bar{y}' = \frac{m'_{01}}{m'_{00}} = \frac{m_{01} + bm_{00}}{m_{00}} = \bar{y} + b \tag{2.46}$$

Therefore, the central moments in x', y' coordinates becomes,

$$\mu'_{pq} = \iint_R (x' - \bar{x}')^p (y' - \bar{y}')^q g(x',y')dx'dy'$$

$$= \iint_R (x+a-\bar{x}-a)^p (y+b-\bar{y}-b)^q f(x,y)dxdy = \mu_{pq} \qquad (2.47)$$

From eqn. 2.47, it is seen that the central moments given in eqns. 2.6-2.7 is invariant to translation.

C. Scaling Invariant

A transform of the image pixel coordinates by a uniform scaling factor k, is given by,

$$x' = kx\,; \quad y' = ky \qquad (2.24)$$

The above transformation also leads to the following expression for the scaled area:

$$dx'dy' = k^2 dxdy \qquad (2.25)$$

The moments of the scaled image can now be expressed in terms of the moments of the original image as

$$m'_{pq} = k^{p+q+2} m_{pq} \qquad (2.26)$$

From the above equation, we also get,

$$m'_{00} = k^2 m_{00} \qquad (2.27)$$

Eliminating the unknown scaling factor k from eqns 2.19-2.20, we have

$$\frac{m'_{pq}}{(m'_{00})^{(p+q+2)/2}} = \frac{m_{pq}}{(m_{00})^{(p+q+2)/2}} \qquad (2.28)$$

Thus, we may define the term η_{pq} as,

$$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^{(p+q+2)/2}} \qquad (2.29)$$

It is invariant under scaling variation of an image. Scale invariance can also be alternatively achieved by other scale-normalization schemes.

2.3. Affine Moment Invariants

Affine transformation [71] is a class of linear 2-D geometric transformation which maps the pixel intensity values located at position (x, y) in an input image into new pixel values at (x', y') in an output image by linear combination of translation, scaling and rotation.

2.3.1. Affine Transformation

The general affine transformation in homogeneous coordinates is usually defined as,

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = R \begin{bmatrix} x \\ y \end{bmatrix} + T \tag{2.30}$$

where R is matrix for scaling and rotation and T is the matrix for translation. If *det(R) = 1* the transformation is orientation preserving. If *det(R) = -1* the transformation is orientation reversing.

For pure scaling, $R = \begin{bmatrix} r_{11} & 0 \\ 0 & r_{22} \end{bmatrix}$ and $T = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

For pure rotation, $R = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$ and $T = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

For pure translation, $R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $T = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$.

A particular example combining scaling, rotation and translation is shown as follows,

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} r_{11} & 0 \\ 0 & r_{22} \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x - \bar{x} \\ y - \bar{y} \end{bmatrix}$$

$$= \begin{bmatrix} r_{11}\cos\theta(x-\bar{x}) + \sin\theta(y-\bar{y}) \\ -r_{22}\sin\theta(x-\bar{x}) + \cos\theta(y-\bar{y}) \end{bmatrix} \tag{2.31}$$

By separating the above equation, we have,

$$x' = -r_{11}(\bar{x}\cos\theta + \bar{y}\sin\theta) + (r_{11}\cos\theta)x + (r_{11}\sin\theta)y \tag{2.32}$$

$$y' = r_{22}(\bar{x}\cos\theta - \bar{y}\sin\theta) - (r_{22}\sin\theta)x + (r_{22}\cos\theta)y \tag{2.33}$$

Then, they can be written as,

$$x' = a_0 + a_1 x + a_2 y \tag{2.34}$$

$$y' = b_0 + b_1 x + b_2 y \tag{2.35}$$

when a_0, a_1, a_2, b_0, b_1 and b_2 are the coefficients.

The affine transformation can be decomposed into six one-parameter transformations. They are:

Translation properties: (1): $\begin{aligned} x' &= x + \alpha \\ y' &= y \end{aligned}$,

(2): $\begin{aligned} x' &= x \\ y' &= y + \beta \end{aligned}$.

Scaling property: (3): $\begin{aligned} x' &= \omega \cdot x \\ y' &= \omega \cdot y \end{aligned}$.

One-axis scaling property: (4): $\begin{aligned} x' &= \delta \cdot x \\ y' &= y \end{aligned}$.

Skew properties: (5): $\begin{aligned} x' &= x + t \cdot y \\ y' &= y \end{aligned}$,

(6): $\begin{aligned} x' &= x \\ y' &= t' \cdot x + y \end{aligned}$.

2.3.2. Definitions

The affine moment invariants [42-43] consist of four second- and third- order moments parameters. The approach aims at performing invariance under the general affine transformation (eqns. 2.34-2.35). Any function, which is

invariant under the above six one-parameter transformations, will be then invariant under the general affine transformation. Three special theorems about invariants are formed after satisfying all conditions of the transformations. They are,

I: Apolar:

$$\frac{1}{2}\sum_{i=0}^{k}(-1)^i\binom{k}{i}\mu_{i,k-i}\mu_{k-i,i} \tag{2.36}$$

after normalization is invariant if k is even.

II: Hankel determinant:

$$\begin{vmatrix} \mu_{0k} & \mu_{1,k-1} & \cdots & \mu_{q-1,k-q+1} \\ \mu_{1,k-1} & \mu_{2,k-2} & \cdots & \mu_{q,k-q} \\ \vdots & & & \\ \mu_{u-1,k-u+1} & \mu_{u,k-u} & \cdots & \mu_{u+q-2,k-u-q+2} \\ \mu_{0k'} & \mu_{1,k'-1} & \cdots & \mu_{q-1,k'-q+1} \\ \mu_{1,k'-1} & \mu_{2,k'-2} & \cdots & \mu_{q,k'-q} \\ \vdots & & & \\ \mu_{v-1,k'-v+1} & \mu_{v,k'-v} & \cdots & \mu_{v+q-2,k'-v-q+2} \\ \vdots & & & \end{vmatrix} \tag{2.37}$$

after normalization is invariant if $k = q + u - 2$, $k = q + v - 2$... are orders of moments.

III: Discriminant of polynomial:

$$\sum_{i=0}^{k}\binom{k}{i}\mu_{i,k-i}x^{k-i} \tag{2.38}$$

after dividing by μ_{0k} and normalization is invariant.

Note that the discriminant of a general polynomial, which is defined as $f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n$, is equal to,

$$(-1)^{n(n-1)/2}\begin{vmatrix} a_0 & a_1 & \cdots & a_n & 0 & 0 & \cdots & 0 \\ 0 & a_0 & \cdots & a_{n-1} & a_n & 0 & \cdots & 0 \\ & & & \vdots & & & & \\ 0 & \cdots & 0 & a_0 & a_1 & a_2 & \cdots & a_n \\ b_0 & b_1 & \cdots & b_{n-1} & 0 & 0 & \cdots & 0 \\ 0 & b_0 & \cdots & b_{n-2} & b_{n-1} & 0 & \cdots & 0 \\ & & & \vdots & & & & \\ 0 & \cdots & 0 & b_0 & b_1 & b_2 & \cdots & b_{n-1} \end{vmatrix} \tag{2.39}$$

where $b_0, b_1, \ldots, b_{n-1}$ are coefficients of $f'(x)$, i.e. $b_j = (n-j)a_j$.

A. The First Second-order Invariant Parameter

The first invariant parameter, which is second-order based, is derived by theorem I and denoted as ζ_1 after normalization,

$$\zeta_1 = \frac{1}{2}\left((-1)^0\binom{2}{0}\mu_{02}\mu_{20} + (-1)^1\binom{2}{1}\mu_{11}\mu_{11} + (-1)^2\binom{2}{2}\mu_{20}\mu_{02}\right)\Big/\mu_{00}^{\ 4}$$

$$= \frac{1}{\mu_{00}^{\ 4}}\left(\mu_{02}\mu_{20} - \mu_{11}^{\ 2}\right) \tag{2.40}$$

where the normalization factor is μ_{00}^{γ}; $\gamma = \sum_{j=1}^{numElement}(p_j + q_j)/2 + numElement$ where *numElement* is the number of elements involved in each term. For example, we choose the term $\mu_{02}\mu_{20}$ to calculate γ. This term has two elements (i.e. μ_{02} and μ_{20}), so *numElement* is 2. Thus, γ = ((0+2)+(2+0))/2+2 = (2(2+0)/2)+2 = 4. Note that if we use other term to calculate γ, the normalization factor remains the same.

B. The Second Third-order Parameter

The second third-order based invariant parameter is derived by theorem III and denoted as ζ_2 after normalization. The polynomial for third-order moment is $f(x) = \mu_{03}x^3 + 3\mu_{21}x^2 + 3\mu_{12}x + \mu_{30}$ and its discriminant is as follows,

$$-\begin{vmatrix} \mu_{03} & 3\mu_{12} & 3\mu_{21} & \mu_{30} & 0 \\ 0 & \mu_{03} & 3\mu_{12} & 3\mu_{21} & \mu_{30} \\ 3\mu_{03} & 6\mu_{12} & \mu_{21} & 0 & 0 \\ 0 & 3\mu_{03} & 6\mu_{12} & 3\mu_{21} & 0 \\ 0 & 0 & 3\mu_{03} & 6\mu_{12} & 3\mu_{21} \end{vmatrix}$$

Then the corresponding second invariant parameter ζ_2 is

$$\zeta_2 = \frac{1}{\mu_{00}^{10}}(\mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{30}\mu_{12}^3 + 4\mu_{03}\mu_{21}^3 - 3\mu_{21}^2\mu_{12}^2) \quad (2.41)$$

where the normalization factor is μ_{00}^{γ}; $\gamma = (2(3+0)+2(0+3))/2 + 4 = 10$.

C. The Third Second- and Third-order Invariant Parameter

The third invariant parameter is from both second- and third order moments. From the Hankel determinant (i.e. theorem II), we have,

$$\begin{vmatrix} \mu_{02} & \mu_{11} & \mu_{20} \\ \mu_{03} & \mu_{12} & \mu_{21} \\ \mu_{12} & \mu_{21} & \mu_{30} \end{vmatrix}$$

Then the corresponding third invariant parameter ζ_3 is

$$\zeta_3 = \frac{1}{\mu_{00}^{7}}(\mu_{20}(\mu_{21}\mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12}) + \mu_{02}(\mu_{12}\mu_{30} - \mu_{21}^2)) \quad (2.42)$$

where the normalization factor is μ_{00}^{γ}; $\gamma = ((2+0)+(2+1)+(0+3))/2+3 = 7$.

D. The Fourth Second- and Third-order Invariant Parameter

The fourth invariant parameter consists of three second-order and two third-order moments. It is defined as,

$$\begin{aligned} \zeta_4 = \frac{1}{\mu_{00}^{11}}(&\mu_{20}^3\mu_{03}^2 - 6\mu_{20}^2\mu_{11}\mu_{12}\mu_{03} - 6\mu_{20}^2\mu_{02}\mu_{21}\mu_{03} + 9\mu_{20}^2\mu_{02}\mu_{12}^2 + \\ &12\mu_{20}\mu_{11}^2\mu_{21}\mu_{03} + 6\mu_{20}\mu_{11}\mu_{02}\mu_{30}\mu_{03} - 18\mu_{20}\mu_{11}\mu_{02}\mu_{21}\mu_{12} - \\ &8\mu_{11}^3\mu_{30}\mu_{03} - 6\mu_{20}\mu_{02}^2\mu_{30}\mu_{12} + 9\mu_{20}\mu_{02}^2\mu_{21}^2 + \\ &12\mu_{11}^2\mu_{02}\mu_{30}\mu_{12} - 6\mu_{11}\mu_{02}^2\mu_{30}\mu_{21} + \mu_{02}^3\mu_{30}^2) \end{aligned} \quad (2.43)$$

where the normalization factor is μ_{00}^{γ}; $\gamma = (3(2+0)+2(0+3))/2+5 = 11$.

2.4. Dudani Moment Invariants

Dudani et al. [44] derived another set of seven second- and third-order moment-based measures that are invariant under translation, rotation and the distance of the object from the camera. They are proposed for use in automatic aircraft identification.

2.4.1. Definitions

To perform the properties of translation and rotation, Dudani moment invariants bring along the functions in eqns. 2.17-2.23, which are proposed by Hu [40]. Moreover, it is known that invariance property of the distance of object from the camera is a kind of scaling property. Normalization by the radius of gyration of a planar pattern is used for achieve scaling invariant.

A. Central Moments

The central moments derived by Hu [40] is used for undertaking the rotation and translation invariance and is defined in eqn. 2.6.

B. Normalization

For the scaling invariant property, Dudani moment invariants make use of the characteristic of distance along optical axis. When an object moves along the optical axis of the camera, the first-order effect on the image is just a change in its size. The second-order effect diminishes as the distance of the object from the camera increases. The radius of gyration β of a planar pattern is then defined as,

$$\beta = \sqrt{\mu_{20} + \mu_{02}} \tag{2.44}$$

The radius of gyration is directly proportional to the size of the image and inversely proportional to the distance of the object along the optical axis.

Therefore, the radius of gyration β can be used to normalize I_1 through I_7 to achieve scaling invariant. Thus, the following seven moments M_1 through M_7 can be found as,

$$M_1 = \frac{I_1}{\beta m_{00}} = \frac{\sqrt{I_1}}{m_{00}} \tag{2.45}$$

$$M_2 = \frac{I_2}{\beta^4} \tag{2.46}$$

$$M_3 = \frac{I_3 m_{00}}{\beta^6} \tag{2.47}$$

$$M_4 = \frac{I_4 m_{00}}{\beta^6} \tag{2.48}$$

$$M_5 = \frac{I_5 m_{00}^{\ 2}}{\beta^{12}} \tag{2.49}$$

$$M_6 = \frac{I_6 m_{00}}{\beta^8} \tag{2.50}$$

$$M_7 = \frac{I_7 m_{00}^{\ 2}}{\beta^{12}} \tag{2.51}$$

2.5. Improved Moment Invariants

Improved moment invariants [57] are the modified version of Hu moment invariants [40] and they are applied to the shape boundary rather than the region.

2.5.1. Definitions

A. Raw Moments

The modified raw moments are defined with the kernel $\Psi_{pq}(x,y) = x^p y^q$ (c.f. eqn. 2.1). The two-dimensional $(p+q)^{\text{th}}$ order moments of a continuous density distribution function $f(x,y) = 1$ are defined as,

$$m_{pq} = \int_C x^p y^q ds \tag{2.52}$$

where $p,q=0,1,2,\ldots..$, $\int_C$ is a line integral along the curve C and $ds = \sqrt{(dx)^2 + (dy)^2}$.

B. Central Moments

The modified central moments is defined as,

$$\mu_{pq} = \int_C (x-\bar{x})^p (y-\bar{y})^q ds \tag{2.53}$$

where $p,q = 0,1,2,\ldots..$

The modified raw moment of order zero, m_{00}, represents the total intensity of an image. For a binary image, this term gives the geometrical area of the image region. The first-order moments, m_{10} and m_{01}, provide the intensity moment about the y-axis and x-axis of the images respectively. The centroid $(\bar{x}, \bar{y})$ is defined as,

$$\bar{x} = \frac{m_{10}}{m_{00}} \tag{2.54}$$

$$\bar{y} = \frac{m_{01}}{m_{00}} \tag{2.55}$$

The central moments for a digital image are approximated by double summations as,

$$\mu_{pq} = \sum_{(x,y)\in C} (x-\bar{x})^p (y-\bar{y})^q \tag{2.56}$$

where $p,q = 0,1,2,\ldots..$, M and N are the horizontal and vertical dimensions, respectively, of the intensity function $f(x,y)$.

C. Normalized Central Moments

In order to achieve scaling invariance, the modified normalized central moments, η_{pq}, is defined as follow,

$$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^{p+q+1}} \tag{2.57}$$

Then, seven Improved moment invariants, which are invariant with respect to translation, scale and rotation of an image, denoted as ψ_1 through ψ_7, become,

$$\psi_1 = \frac{I_1}{\mu_{00}^{\ 3}} \tag{2.58}$$

$$\psi_2 = \frac{I_2}{\mu_{00}^{\ 6}} \tag{2.59}$$

$$\psi_3 = \frac{I_3}{\mu_{00}^{\ 8}} \tag{2.60}$$

$$\psi_4 = \frac{I_4}{\mu_{00}^{\ 8}} \tag{2.61}$$

$$\psi_5 = \frac{I_5}{\mu_{00}^{\ 16}} \tag{2.62}$$

$$\psi_6 = \frac{I_6}{\mu_{00}^{\ 11}} \tag{2.63}$$

$$\psi_7 = \frac{I_7}{\mu_{00}^{\ 16}} \tag{2.64}$$

2.5.2. Invariant Properties of Rotation, Translation and Scaling

A. Rotation Invariant

A rotation of an image by an angle θ has an associate pixel coordinate transformation given by,

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{2.65}$$

where x', y' are the new coordinates and x, y are the original coordinates.

If the improved moment invariants achieve the rotation invariant, eqn. 2.66 should be satisfied.

$$\psi_i' = \psi_i \tag{2.66}$$

where $1 \le i \le 7$.

We are going to prove $\psi_1' = \psi_1$ and $\psi_2' = \psi_2$ as examples. $\psi_j' = \psi_j$, for $3 \le j \le 7$ can be similarly proved by trigonometric identities. The modified central moments of the rotated image is as follows,

$$\mu_{pq}' = \int_{C'} (x')^p (y')^q \, ds' \tag{2.64}$$

By substituting $x' = x\cos\theta - y\sin\theta$ and $y' = x\sin\theta + y\cos\theta$ into eqn. 2.64, then we have,

$$\begin{aligned} \mu_{pq}' &= \int_{C'} (x')^p (y')^q \, ds' \\ &= \int_{C} [x\cos\theta - y\sin\theta]^p [x\sin\theta + y\cos\theta]^q \, ds \end{aligned} \tag{2.67}$$

where $ds' = ds$

For zeroth-order central moment: $\mu_{00}' = \mu_{00}$. It is because the length of the boundary of the rotated curve and the original curve are the same.

Proof of $\psi_1' = \psi_1$: From eqn. 2.58, we have,

$$\psi_1' = \frac{\mu_{20}' + \mu_{02}'}{\mu_{00}'^{\,3}} \tag{2.68}$$

By multiplying $(\mu_{00}')^3 = (\mu_{00})^3$ into both sides,

$$\begin{aligned} (\mu_{00}')^3 \psi_1' &= (\mu_{00}')^3 \frac{\mu_{20}' + \mu_{02}'}{\mu_{00}'^{\,3}} \\ &= \int_{C} [x\cos\theta - y\sin\theta]^2 + [x\sin\theta + y\cos\theta]^2 \, ds \\ &= \int_{C} [x^2\cos^2\theta - 2xy\cos\theta\sin\theta + y^2\sin^2\theta + \\ &\quad x^2\sin^2\theta + 2xy\cos\theta\sin\theta + y^2\cos^2\theta] ds \\ &= \int_{C} [x^2(\cos^2\theta + \sin^2\theta) + y^2(\cos^2\theta + \sin^2\theta)] ds \end{aligned}$$

$$= \int_C x^2 + y^2 ds = \mu_{20} + \mu_{02} = (\mu_{00})^3 \frac{\mu_{20} + \mu_{02}}{\mu_{00}^3} = (\mu'_{00})^3 \psi_1 \qquad (2.69)$$

Hence,

$$\psi'_1 = \psi_1$$

Proof of $\psi'_2 = \psi_2$: Similarity from eqn. 2.59, we have,

$$\psi'_2 = \frac{(\mu'_{20} - \mu'_{02})^2 + 4\mu'^2_{00}}{\mu'^6_{00}} \qquad (2.70)$$

Multiplying $(\mu'_{00})^6 = (\mu_{00})^6$ into both sides,

$$(\mu'_{00})^6 \psi'_1 = (\mu'_{00})^6 \frac{(\mu'_{20} - \mu'_{02})^2 + 4\mu'^2_{11}}{\mu'^6_{00}}$$

$$= \left[\int_C [x\cos\theta - y\sin\theta]^2 - [x\sin\theta + y\cos\theta]^2 ds\right]^2 + 4\left[\int_C [x\cos\theta - y\sin\theta][x\sin\theta + y\cos\theta] ds\right]^2$$

$$= \left[\int_C (x^2\cos 2\theta + y^2\cos 2\theta - 2xy\sin 2\theta)\, ds\right]^2 + \left[\int_C (x^2\sin 2\theta - y^2\sin 2\theta - 2xy\cos 2\theta)\, ds\right]^2$$

$$= (\mu_{20}\cos 2\theta - \mu_{02}\cos 2\theta - 2\mu_{11}\sin 2\theta)^2 + (\mu_{20}\sin 2\theta - \mu_{02}\sin 2\theta - 2\mu_{11}\cos 2\theta)^2$$

$$= \begin{aligned}&(\mu_{20}^2\cos^2 2\theta + \mu_{02}^2\cos^2 2\theta + 4\mu_{11}^2\sin^2 2\theta - 2\mu_{20}\mu_{02}\cos^2 2\theta \\ &\quad - 4\mu_{11}\mu_{20}\sin 2\theta\cos 2\theta + 4\mu_{11}\mu_{02}\sin 2\theta\cos 2\theta)^2 + \\ &(\mu_{20}^2\sin^2 2\theta + \mu_{02}^2\sin^2 2\theta + 4\mu_{11}^2\cos^2 2\theta - 2\mu_{20}\mu_{02}\sin^2 2\theta \\ &\quad + 4\mu_{11}\mu_{20}\sin 2\theta\cos 2\theta - 4\mu_{11}\mu_{02}\sin 2\theta\cos 2\theta)^2\end{aligned}$$

$$= \begin{aligned}&\mu_{20}^2(\cos^2 2\theta + \sin^2 2\theta) - 2\mu_{20}\mu_{02}(\cos^2 2\theta + \sin^2 2\theta) + \\ &\mu_{202}^2(\cos^2 2\theta + \sin^2 2\theta) + 4\mu_{11}(\cos^2 2\theta + \sin^2 2\theta)\end{aligned}$$

$$= (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2 = (\mu_{00})^6 \frac{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}{\mu_{00}^6} = (\mu'_{00})^6 \psi_2 \quad (2.71)$$

Hence,

$$\psi'_2 = \psi_2$$

B. Translation Invariant

Consider the following linear shifting of the image coordinates from (x,y) to (x',y'),

$$x' = x + a \quad (2.72)$$

$$y' = y + b \quad (2.73)$$

where x, y are the original coordinates and x', y' are the new coordinates.

The raw moments of these two coordinates are defined as,

$$m_{pq} = \int_C x^p y^q ds \quad (2.74)$$

$$m'_{pq} = \int_{C'} x'^p y'^q ds' \quad (2.75)$$

where m_{pq} and m'_{pq} are the modified raw moments in x, y and x', y' coordinates respectively.

For zeroth-order moment: $m'_{00} = \int_{C'} ds' = \int_C ds = m_{00}$. It is because the length of the boundary in region C of the images are the same and eqns. 2.72-2.73 imply that $ds' = ds$.

For first-order moment: $m'_{10} = \int_{C'} x' ds' = \int_C x + a \, ds = m_{10} + a(m_{00})$; $m'_{01} = \int_{C'} y' ds' = \int_C y + b \, ds = m_{01} + b(m_{00})$.

Then the centroid $(\bar{x}', \bar{y}')$ in x', y' coordinates becomes,

$$\bar{x}' = \frac{m'_{10}}{m'_{00}} = \frac{m_{10} + a m_{00}}{m_{00}} = \bar{x} + a \quad (2.76)$$

$$\bar{y}' = \frac{m'_{01}}{m'_{00}} = \frac{m_{01} + bm_{00}}{m_{00}} = \bar{y} + b \tag{2.77}$$

Therefore, the central moments in x', y' coordinates becomes,

$$\mu'_{pq} = \int_{C'} (x' - \bar{x}')^p (y' - \bar{y}')^q ds'$$

$$= \int_{C} (x + a - \bar{x} - a)^p (y + b - \bar{y} - b)^q ds = \mu_{pq} \tag{2.78}$$

From eqn. 2.78, it is seen that the central moments given in eqns. 2.53-2.56 are invariant to translation.

C. Scaling Invariant

A transform of the image pixel coordinates by a uniform scale factor k, is given by,

$$x' = kx \; ; \quad y' = ky \tag{2.79}$$

It is supposed that the curve, C', is obtained by homogeneously rescaling the coordinates by a factor k,

$$C' = kC \tag{2.80}$$

Then, the modified central moments becomes,

$$\mu'_{pq} = \int_{C'} x'^p y'^q ds'$$

$$= \int_{C} (kx)^p (ky)^q d(ks)$$

$$= k^{p+q+1} \int_{C} (x)^p (y)^q d(s) = k^{p+q+1} \mu_{pq} \tag{2.81}$$

For zeroth-order moment: $\mu_{00} = \int_{C} ds = |C|$ and $\mu'_{00} = \int_{C'} ds' = |C'| = k|C|$.

Moreover, for any $k > 0$, we have,

$$\eta'_{pq} = \frac{\mu'_{pq}}{(\mu'_{00})^{p+q+1}} = \frac{k^{p+q+1} \mu_{pq}}{k^{p+q+1} (|C|)^{p+q+1}} = \frac{\mu_{pq}}{(\mu_{00})^{p+q+1}} \tag{2.82}$$

It is invariant under scaling variation of an image.

2.6. Sluzek's Moment-based Shape Descriptor

Sluzek [72] proposed a new method of shape characterization. This shape descriptor aims at identifying and inspecting of two-dimensional objects. Instead of using noise sensitive high order moments [40, 76], a set of second order-based invariants can be obtained by occluding the object by varying the size of the circle. This set of descriptors is related to the symmetry of the shape. It tries to have more position-invariant descriptors and to improve the resolution of these descriptors.

Fig. 2-1 and 2-2 demonstrate the idea of the occluded square and triangle respectively. The reason for choosing circle as the occluding mask is based on the fact that circle is the geometrical shape that is naturally and perfectly invariant to rotation in two-dimensional space. The circle is located at the centroid of the object and is used to mask out the central part of the shape. The area of a circle can be between 0% to maximum percentage. The criterion of getting the maximum percentage is that at least a part of region R of object is visible after occluded a circle.

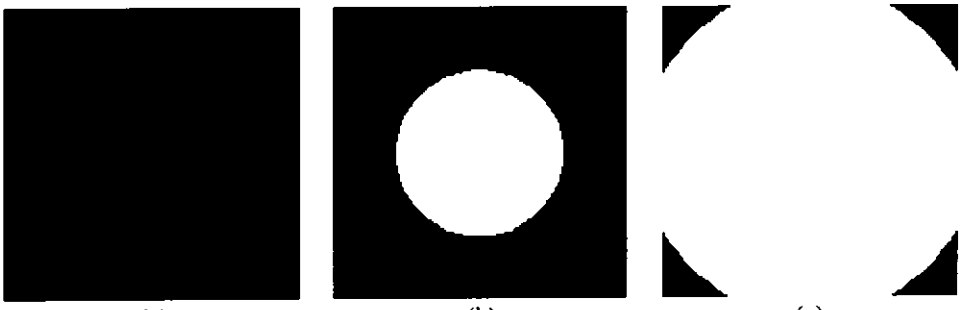


Figure 2-1 (a) Unoccluded square, (b-c) Occluded square by circles with 25% and 100% of unoccluded square's area respectively and located at the centroid

Figure 2-2 (a) Unocculeded triangle, (b-c) Occluded triangle by circles with 25% and 100% of unoccluded triangle's area respectively and located at the centroid

2.6.1. Definition

Sluzek's shape descriptor is an enhanced version of Hu moment invariants [1]. It employees the central moments (c.f. eqns. 2.6-2.7) of Hu moment invariants and only the moments of order two are involved. They are,

$$\phi_1 = \frac{\mu_{20} + \mu_{02}}{\mu_{00}^{\ 2}} \tag{2.83}$$

$$\phi_2 = \frac{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^{\ 2}}{\mu_{00}^{\ 4}} \tag{2.84}$$

$$\phi_8 = \frac{1}{4}\left[(\phi_1^{\ 2} - \phi_2\right] = \frac{\mu_{20}\mu_{02} - \mu_{11}^{\ 2}}{\mu_{00}^{\ 4}} \tag{2.85}$$

Let C(α) be a family of circles defined by the equation:

$$(x - \bar{x})^2 + (y - \bar{y})^2 = \alpha\, m_{00}/\pi \tag{2.86}$$

where $0<\alpha<\alpha_{max}$; α_{max} is the maximum ratio between the area of circle and the area of the object.

Therefore, for each α, it can obtain functions $\phi_1(\alpha)$, $\phi_2(\alpha)$, $\phi_8(\alpha)$, etc. These functions are called m-invariant function of region. They are invariant to scaling, translation and rotation of the object, since these functions are

derived from Hu moment invariant and the occluded mask is a circle.

Then the shape descriptors, so called m-invariant descriptors, can be obtained by,

$$\{\phi_x(\alpha_1), \quad \phi_x(\alpha_2), \quad \cdots \quad \phi_x(\alpha_n),\} \tag{2.87}$$

To compare two region R_1 and R_2, the similarity measure is done by,

$$Q(\phi_x(\alpha), R_1, R_2) = \frac{V(R_1, \phi_x(\alpha)) + V(R_2, \phi_x(\alpha))}{|\phi_x(R_1) - \phi_x(R_2)|} \tag{2.88}$$

where $V(R, \phi_x)$ is the magnitude of deviation of a given R and the selected invariant ϕ_x, $\phi_x(R_1)$ and $\phi_x(R_2)$ are the theoretical values of the invariants for R_1 and R_2.

Chapter 3

Interpretations on Second and Third-order Moments

3.1. Introduction

One of the widely used shape descriptors is a set of moment invariants derived by Hu [40]. It consists of a set of seven moment invariants to characterize image features. The approach has been extended to a larger set by Wong and Siu [41] and different forms of moment invariants [42-44, 57]. The performance of different shape descriptors on the objects, such as the reduced chain code, Fourier descriptors, Moment Invariants and Zernike moments, has been addressed by a few researchers [46-47, 53]. From their analysis, it is shown that high accuracy and

fast retrieval rates are achievable for describing the objects by calculating the moment invariants. In order to provide a meaningful description of shape, it is important to understand the meaning of these elements. However, there are few pertinent studies. Leu [74] elaborated the invariants in terms of spreadness and elongation from the maximum and minimum moments of inertia. But their formulations have not shown the elucidation of second-order based moment invariants in a simple and meaningful way. Also, no analysis on these two terms for different kinds of shapes are given. Therefore, the objective of this paper is to make an interpretation of the first two second order-based Hu moment invariants so that they can be used effectively in different applications. Recently, a new method [72] of shape characterization was proposed. Instead of using noise sensitive high order moments [40, 75], a set of second order-based invariants can be obtained by varying the size of the circle. The circle is used to mask out the central part of the shape. This set of descriptors is related to the symmetry of the shape. Following the same line of thought, an ellipse can be used to study the meaning of the second order-based invariants. It is found that the first invariant measures the total spread of the shape relative to its area square while the second invariant measures the degree of elongation of a best-fit ellipse on the shape. We can obtain the same values of third invariant for those objects with symmetric along $y = \pm x/\sqrt{3}$ and $y = \pm\sqrt{3}x$. For the fourth invariant, when fitting a best-fit ellipse on the object, there are constants on the term of $a_f{}^2 + b_f{}^2$.

This chapter is organized as follows. In Section 3.2, we give the interpretations on second-order moments. It includes the theological and experimental analysis. In Section 3.3, the studies of third and fourth invariants are shown.

3.2. Interpretations on Second-order Moments

Hu moment invariants consist of a set of seven second and third-order moment invariant elements. They are designed to be independent to translation, scaling and rotation. In this section, we discuss the study of the elements theoretically and experimentally. Firstly, we show the interpretation of second order moments. They are the first and second moment invariant elements (ϕ_1 and ϕ_2).

3.2.1. First Moment Invariant Element: Total Spread Relative to the Area Square

To understand the meaning of the first moment invariant $\phi_1 = \frac{\mu_{20} + \mu_{02}}{\mu_{00}^2}$ in eqn. 2.30, we first consider the meaning of μ_{00}, μ_{20} and μ_{02}. We can see that,

- μ_{00} measures the area of the object.
- μ_{20} measures the spread of the object with respect to $\bar{x}$ in x-direction.
- μ_{02} measures the spread of the object with respect to $\bar{y}$ in y-direction.

Thus,

- μ_{20}/μ_{00}^2 measures the spread of the object in x-direction normalized by the area square.
- μ_{02}/μ_{00}^2 measures the spread of the object in y-direction normalized by the area square.

Therefore, the first moment invariant ϕ_1 measures the total spread relative to the area square. Note that if μ_{20}/μ_{00}^2 is equal to μ_{02}/μ_{00}^2, this means that the spread in x-direction is equivalent to the spread in y-direction.

3.2.2. Second Moment Invariant Element: Degree of Elongation of a Best-fit Ellipse on the Shape

To study $\phi_2 = \frac{((\mu_{20} - \mu_{02})^2 + 4\mu_{11}^{\ 2})}{\mu_{00}^{\ 4}}$ in eqn. 2.31, let us consider an ellipse rotated by θ_f degree anti-clockwisely. Let us denote μ'_{pq} and μ_{pq} respectively as the central moments for the ellipse with respect to the original x', y' and the rotated x, y axes. Then we have,

$$\mu'_{20} = \cos^2\theta_f \mu_{20} + \sin^2\theta_f \mu_{02} \tag{3.1}$$

$$\mu'_{02} = \sin^2\theta_f \mu_{20} + \cos^2\theta_f \mu_{02} \tag{3.2}$$

$$\mu'_{11} = \sin\theta_f \cos\theta_f (\mu_{20} - \mu_{02}) \tag{3.3}$$

Using eqn.3.1-3.3, the orientation can be found as shown below,

$$\tan 2\theta_f = 2\mu_{11}/(\mu_{20} - \mu_{02}) \tag{3.4}$$

Let a_f and b_f be the major and the minor axes rotated by θ_f degree respectively. We have,

$$\mu_{20} = (\pi \cdot a_f^3 b_f)/4 \tag{3.5}$$

$$\mu_{02} = (\pi \cdot a_f b_f^3)/4 \tag{3.6}$$

$$\mu_{11} = 0 \tag{3.7}$$

Using eqn.3.5-3.7, ϕ_2 for an ellipse can be expressed as,

$$\phi_2 = \pi^2 a_f^2 b_f^2 (a_f^2 - b_f^2)^2 / 16\mu_{00}^{\ 4} \tag{3.8}$$

For ϕ_2 to be zero, a_f needs to be equal to b_f. Moreover, ϕ_2 increases when the difference between a_f and b_f increases. Thus, ϕ_2 measures the degree of elongation of the ellipse. For shapes other than an ellipse, we could consider fitting an ellipse onto the shape at its centroid with (from eqns.3.9-3.10)),

$$a_f = \left(\frac{4}{\pi}\right)^{\frac{1}{4}} \left(\frac{\mu_{20}{}^3}{\mu_{02}}\right)^{\frac{1}{8}} \tag{3.9}$$

$$b_f = \left(\frac{4}{\pi}\right)^{\frac{1}{4}} \left(\frac{\mu_{02}{}^3}{\mu_{20}}\right)^{\frac{1}{8}} \tag{3.10}$$

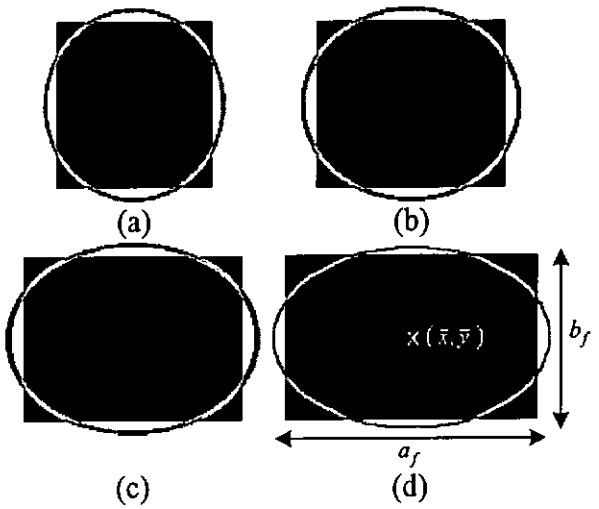


Figure 3-1 Four rectangles with different a_f and b_f in best-fit ellipse

If the difference between a_f and b_f is large, the shape is elongated, i.e., ϕ_2 is large. If they are the same, the best-fit ellipse becomes a circle, i.e., $\phi_2 = 0$, as in Fig. 3-1(a). The idea of fitting a best-fit ellipse onto the images is illustrated in Fig. 3-1.

The pseudo-code for the best-fit ellipse method is as follows,

```
define and obtain a_f, b_f and angle θ_f   for the object
for (i = 0;i < width;i++){
    for (j = 0;j < height;j++){
        xx = i - mx;                    // mx: x-coordinate of centroid
        yy = j - my;                    // my: y-coordinate of centroid
        v = [(xx cos(ang) + yy sin(ang))/a]^2 + [(yy cos(ang) + xx sin(ang))/b]^2
        if (0.95<v<1.05)
            set the pixel as the ellipse pixel
    }
}
```

3.2.3. Experimental Analysis

To confirm the findings from the previous section, six different shapes including squares, rectangles, triangles, trapeziums, and two irregular (leaf- and horse-) shapes were used for the experimental analysis. We have found that the ratio of the major and minor axes of the best-fit ellipse gives a strong discrimination power and is able to provide better shape resolutions and insensitive to Gaussian noises.

Squares

Let us firstly discuss the case of a filled square with length a. μ_{00}, μ_{20}, μ_{02} and μ_{11} in eqn. 2.7 are calculated for the filled square object. It is found that, $\mu_{00} = a^2$, $\mu_{20} = a^4/12$, $\mu_{02} = a^4/12$ and $\mu_{11} = 0$. Then we have $\mu_{20}/\mu_{00}^2 = \mu_{02}/\mu_{00}^2 = 1/12$. Therefore, ϕ_1 = 1/6, which is a constant. As μ_{20} = μ_{02} and μ_{11} = 0, ϕ_2 is always zero for the square.

Rectangles

We then consider the case of a filled rectangle with length a and width b. From eqn.2.7, it can be shown that, $\mu_{00} = ab$, $\mu_{20} = a^3b/12$, $\mu_{02} = ab^3/12$ and μ_{11} = 0. Then we have $\mu_{20}/\mu_{00}^2 = a/12b$ and $\mu_{02}/\mu_{00}^2 = b/12a$. Therefore, $\phi_1 = (a^2 + b^2)/12ab$ and $\phi_2 = (a^2 - b^2)^2/144a^2b^2$. Consistent with the above part of squares, when the length, a, and the width, b, are the same (i.e. $a = b$), the rectangle becomes a square, and ϕ_1 becomes a constant while ϕ_2 is zero.

Table 3-1 Results for four rectangles with constant b=93

	ϕ_1	$\phi_2 \times 10^{-3}$	a_f	b_f	a_f/b_f
a=93	0.167	0.0	53.1	53.1	1.000
a=112	0.170	0.001	63.9	53.1	1.204
a=130	0.176	0.003	74.2	53.1	1.398
a=148	0.185	0.006	84.5	53.1	1.591

Table 3-1 shows that ϕ_1, ϕ_2 and the ratio a_f/b_f increase with respect to an increase of a as shown in Fig. 3-1 (a) to Fig. 3-1 (d). Moreover, it is shown that the best-fit ellipse for Fig. 3-1 (a) with $a_f/b_f = 1$ is the closest to the circle.

Triangles

We further discuss the case of a filled triangle with vertex (0,0), (b,c) and (a,0). Using eqn. 2.7, it can be proved that, $\mu_{00} = ac/2$, $\mu_{20} = ac(a^2+b^2-ab)/36$, $\mu_{02} = ac^3/36$, $\mu_{11} = ac^2(2b-a)/72$.

For $\phi_2 = \dfrac{(a^2+b^2-ab-c^2)^2 + c^2(2b-a)^2}{81a^2c^2}$ to be zero, we have $a = 2b$ and $c = \sqrt{3}b$. This is indeed an equilateral triangle. If an ellipse with parameters in eqns. 3.9-3.10 (i.e. $a_f = \left(\dfrac{4}{\pi}\right)^{\frac{1}{4}}\left(\dfrac{c^2}{a^2+b^2-ab}\right)^{\frac{1}{8}}$, $b_f = \left(\dfrac{4}{\pi}\right)^{\frac{1}{4}}\left(\dfrac{a^2+b^2-ab}{c^2}\right)^{\frac{1}{8}}$ and $\tan 2\theta_f = \dfrac{c(2b-a)}{a^2+b^2-ab-c^2}$) is fitted onto this equilateral triangle, it is found that $a_f = b_f = 0.38931a$. In other words, an equilateral triangle with zero ϕ_2 has a best-fit ellipse with $a_f = b_f$, thus confirming the theoretical study.

In our experimental work, a set of four triangles was tested. Table 3-2 shows the results obtained. As a increases from 120 to 180, ϕ_1 increases due to an increase in the total spread of the triangles. Also both ϕ_2 and a_f/b_f increase as the shape becomes elongated. The best-fit ellipse is the closest to a circle when ϕ_2 tends to 0 as shown in Fig. 3-2a.

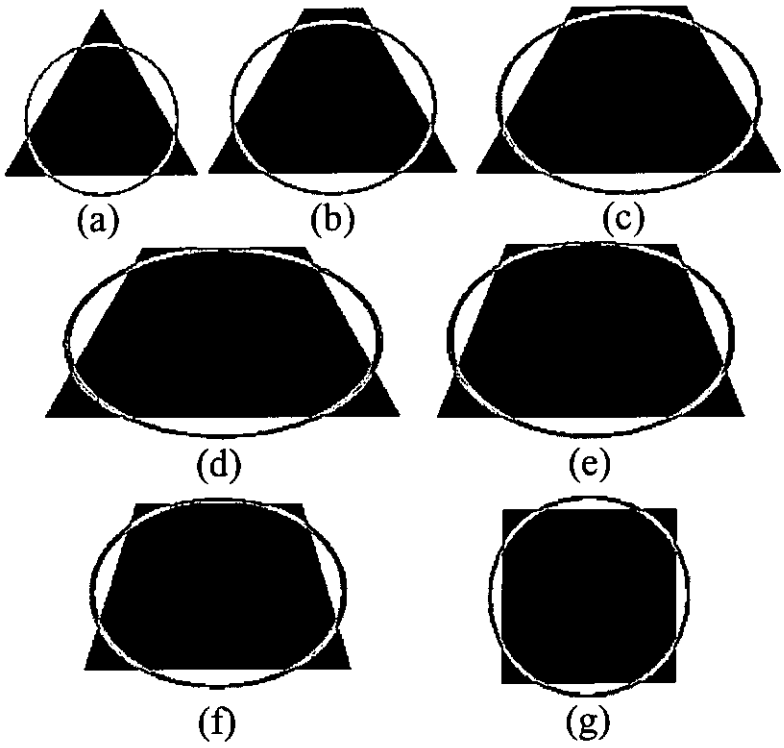


Figure 3-2 (a) Equilateral triangle T1, (b-f) T2-T6, and (g) Square T7 with their best-fit ellipses

Table 3-2 Results for four triangles with $b = a/2$ and $c=104$

	ϕ_1	$\phi_2 \times 10^{-3}$	a_f	b_f	a_f/b_f
a=120	0.193	0.004	47.0	46.5	1.011
a=140	0.195	0.857	54.6	47.0	1.164
a=160	0.199	2.748	61.5	47.0	1.309
a=180	0.206	5.548	68.6	47.0	1.459

Trapeziums

The case of a filled trapezium with vertices (0,0), (b,c), (d,c) and (a,0) was considered. Using eqn. 2.7, it can be proved that, $\mu_{00} = c(a+d-b)/2$,

$$\mu_{20} = \frac{c\left[(d^2+a^2)(d+a)-b^3\right]}{12} - \frac{c(a^2+d^2+ad-b^2)^2}{18(a+d-b)},$$

$$\mu_{02} = \frac{c^3(3d+a-3b)}{12} - \frac{c^3(a+2d-2b)^2}{18(a+d-b)}, \text{ and}$$

$$\mu_{11} = \frac{c^2\left[3(d-a)^2+8a(d-a)+6a^2-3b^2\right]}{24} - \frac{c^2(a+2d-2b)(a^2+d^2+ad-b^2)}{18(a+d-b)}.$$

For $\phi_2 = 0$ (i.e.$\mu_{20} = \mu_{02}$ and $\mu_{20} = 0$), we have $a = b+d$ and $c^2 = \frac{3d^2(2d^2 - 2ad + a^2)}{2d^2 - a^2 + 2ad}$. Let us consider two special cases. The first case is when $d = b$, i.e., a triangle. We found that $a = 2b$ and $c = \sqrt{3}b$, which implies an equilateral triangle. The best-fit ellipse is a circle for $\phi_2 = 0$. The second case is when $a = d$ and $b = 0$, i.e., a square. We found $c = a = d$, in which the best-fit ellipse is again a circle.

Table 3-3 shows the experimental results for a set of seven trapeziums. T1 is a triangle while T7 is a square. T4 is the trapezium with the biggest width (Fig. 3-2d). ϕ_1 increases from T1 to T4 and then decreases from T5 to T7 due to a change in the total spread. ϕ_2 and a_f / b_f show a close correlation. ϕ_2 becomes 0 when a_f / b_f tends to 1. T4 has the largest ϕ_2 and a_f / b_f since the degree of elongation is the biggest among all cases in this set.

Table 3-3 Results of seven trapeziums with $a = b+d$ and $c=100$

		ϕ_1	$\phi_2 \times 10^{-3}$	a_f	b_f	a_f/b_f
T1	b=58, d=58	0.182	0.014	46.1	45.2	1.019
T2	b=58, d=93	0.186	1.054	62.0	52.0	1.193
T3	b=58, d=128	0.190	4.643	79.0	54.2	1.457
T4	b=58, d=158	0.200	9.822	95.0	55.2	1.722
T5	b=40, d=140	0.184	4.593	82.3	55.9	1.472
T6	b=30, d=130	0.177	2.543	75.5	56.3	1.341
T7	b=0, d=100	0.167	0.0	57.1	57.1	1.0

Irregular Leaf-Shaped Objects

For experiments with objects of irregular shapes, an irregular leaf shaped object with length a_o and width b as shown in Fig. 3-3a. is used. A set of 8 objects is generated by varying the length a, i.e. $a = ka_o$ where k = elongation factor

$=1+n/2$ for $0 \le n \le 7, n \in Z$ (Fig. 3-3b and 3-3c). Fig 3-4 shows the results for ϕ_1, ϕ_2 and a_f / b_f. As expected, ϕ_1 increases due to an increase in the total spread. ϕ_1 and a_f / b_f both increase as the degree of elongation increases with k. However, a_f / b_f has a much larger discrimination power since its curve in Fig. 3-4 has a gradient larger than that of ϕ_2 and is thus able to provide better shape resolutions. As to the robustness of the best-fit ellipse shape descriptor a_f / b_f, due to the 'smoothing' effect of fitting an ellipse to a general shape, a_f / b_f is less sensitive to the effect of digitization and noise than ϕ_2.

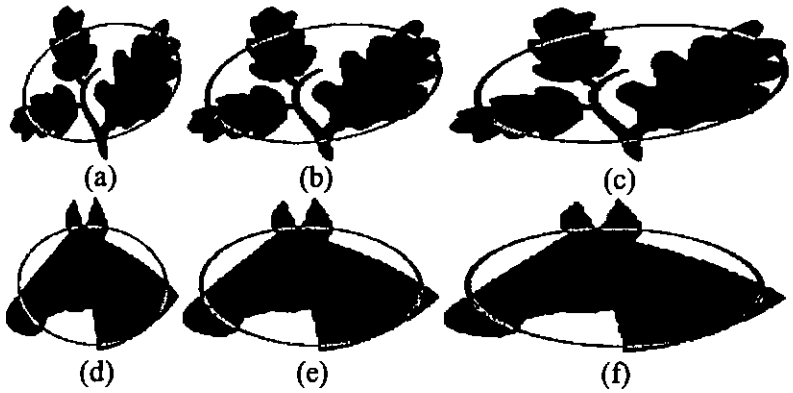


Figure 3-3 Three leaf-shaped objects with (a) $k=1$, (b) $k=1.5$, and (c) $k=2$ and Three horse-shaped objects with (d) $k=1$, (e) $k=1.5$, and (f) $k=2$

Irregular Horse-Shaped Objects

Another set of experiments was to make use of a horse shaped object with length a_o and width b as shown in Fig. 3-3d. They were also generated by varying length a, i.e. $a = ka_o$ where $k = 1 + n/2$ for $0 \le n \le 7, n \in Z$ (Figs. 3-3e and 3-3f). Fig. 3-5 shows the results for ϕ_1, ϕ_2 and a_f / b_f. ϕ_1 increases due to an increase in the total spread. ϕ_2 and a_f / b_f both increase as the degree of elongation increases with k.

Irregular Leaf-Shaped and Horse-Shaped Objects with Gaussian Noise

To show the sensitivities of ϕ_1, ϕ_2 and a_f / b_f against the noise, we have tried to

add the Gaussian noise with SNR = 0dB to 25dB to the same sets of objects (k = 4.5) in the previous two sections. The Gaussian noise is with zero mean. Figs. 3-6(a) and 3-6(b) give the percentages of ϕ_1, ϕ_2 and a_f / b_f deviated from the actual values for leaf-shaped and horse-shaped objects respectively. They were obtained by taking the percentage difference between the values of ϕ_1, ϕ_2 and a_f / b_f of the original binary object (c.f. Fig. 3-3) and that of the objects with added noises. In Fig. 3-6, it is seen that ϕ_1 and ϕ_2 suffer from a larger percentage of variation than a_f / b_f for both leaf-shaped and horse-shaped objects, especially for SNR = 0dB.

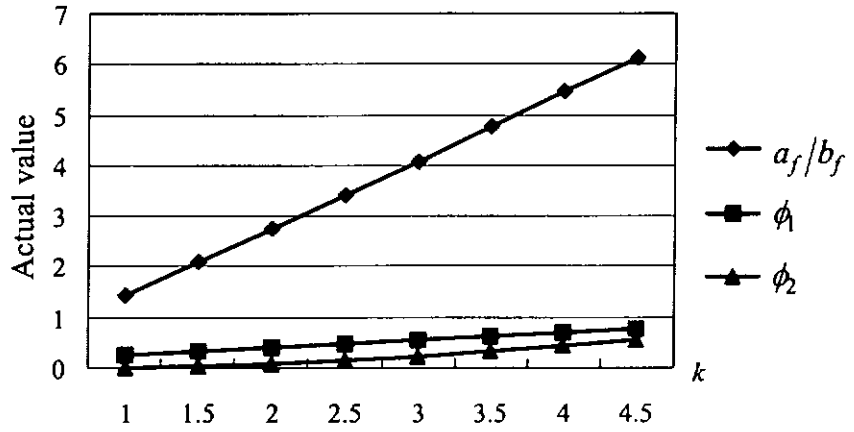


Figure 3-4 Three plots for leaf-shaped objects with constant width *b=108*

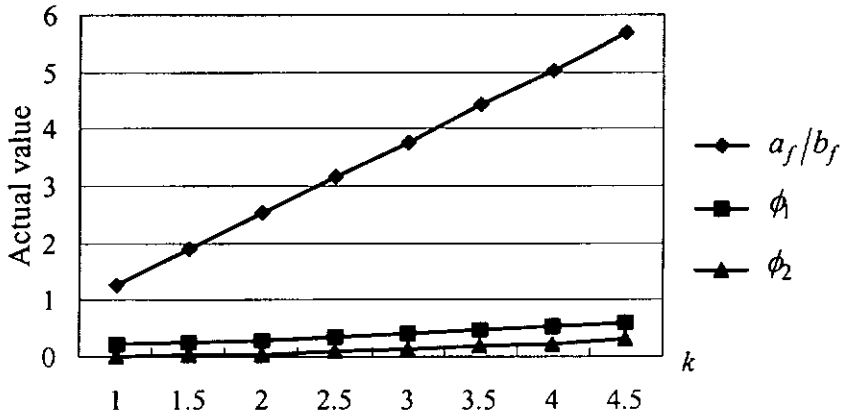


Figure 3-5 Three plots for horse-shaped objects with constant width *b=119*

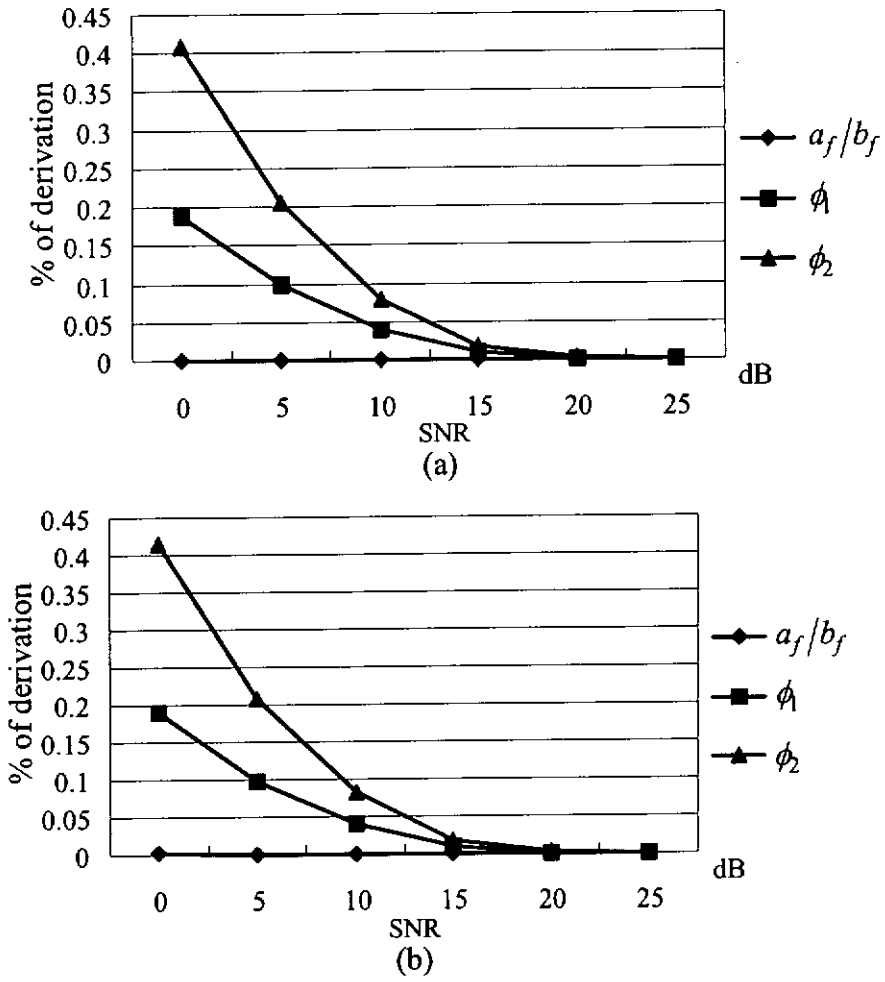


Figure 3-6 Three plots for (a) Leaf-shaped and (b) Horse-shaped objects with elongation factor $k = 4.5$

3.3. Interpretations on Third-order Moments

The interpretation of third order moments is as shown below. They are the third and fourth moment invariant elements (ϕ_3 and ϕ_4).

3.3.1. Third Moment Invariant Element

To understand the meaning of the third moment invariant $\phi_3 = \frac{(\mu_{30} - 3\mu_{12})^2 + (\mu_{03} - 3\mu_{21})^2}{\mu_{00}^5}$ in eqn. 2.32, we first consider the conditions under which ϕ_3 becomes zero. It can be seen that ϕ_3 is zero when the following two cases happen at the same time.

$$(\mu_{30} - 3\mu_{12})^2 = 0 \tag{3.11}$$

$$(\mu_{03} - 3\mu_{21})^2 = 0 \tag{3.12}$$

Then, we try to think about the measurement directions of μ_{30}, μ_{12}, μ_{03} and μ_{21}. From eqn. 2.6, their expressions are shown as below,

$$\mu_{30} = \iint (x - \bar{x})^3 f(x, y) dxdy \tag{3.13}$$

$$\mu_{12} = \iint (x - \bar{x})^1 (y - \bar{y}) f(x, y) dxdy \tag{3.14}$$

$$\mu_{03} = \iint (y - \bar{y})^3 f(x, y) dxdy \tag{3.15}$$

$$\mu_{21} = \iint (x - \bar{x})^2 (y - \bar{y}) f(x, y) dxdy \tag{3.16}$$

μ_{30} and μ_{12} measure x-direction, while μ_{03} and μ_{21} measure y-direction. Although it is difficult to have a measurement meaning on the parameters of $\mu_{30} - 3\mu_{12}$ or $\mu_{03} - 3\mu_{21}$, it can be expected that the third moment invariant is highly related to the symmetry.

Let us use a triangle as an example to illustrate the theory. We discuss the case of a filled triangular as shown in Fig. 3-7.

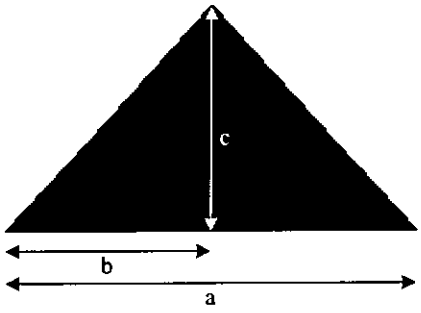


Figure 3-7 A filled triangular as the object

μ_{30}, μ_{12}, μ_{03} and μ_{21} in eqn. 2.7 are calculated for the filled triangular object. It is found that,

$$\mu_{00} = \frac{ac}{2} \tag{3.17}$$

$$\mu_{30} = \frac{ac(a+b)(2a-b)(a-2b)}{540} \tag{3.18}$$

$$\mu_{12} = \frac{ac^3(2b-a)}{540} \tag{3.19}$$

$$\mu_{03} = \frac{ac^4}{270} \tag{3.20}$$

$$\mu_{21} = \frac{ac^2(-a^2+2b^2-2ab)}{540} \tag{3.21}$$

For $(\mu_{30} - 3\mu_{12})^2 = 0$ in eqn. 3.11, we have,

$$ac(a+b)(2a-b)(a-2b) - 3ac^3(2b-a) = 0 \tag{3.22}$$

$$ac(a-2b)[(a+b)(2a-b) + 3c^2] = 0 \tag{3.23}$$

By the definition of the triangular object as shown in Fig. 3-7, we have $a > b$ and $c \neq 0$. Therefore, for $(a+b)(2a-b) = 0$, a is equal to $2b$.

If $a = 2b$, $\mu_{30} = \frac{ac(a+b)(2a-b)(a-2b)}{540} = 0$ and $\mu_{12} = \frac{ac^3(2b-a)}{540} = 0$ happen. It results that the object is symmetrical in the x-direction, since μ_{30} equals $3\mu_{12}$. Otherwise, if $a \neq 2b$, both μ_{30} and $3\mu_{12}$ do not equal to zero and there is no other choice of a, b and c that could make $(\mu_{30} - 3\mu_{12})^2 = 0$.

In fact, if we want $(\mu_{30} - 3\mu_{12})^2 = 0$ for the triangular object, the object should be symmetry along $y = \pm x/\sqrt{3}$. Fig. 3-8 gives two examples.

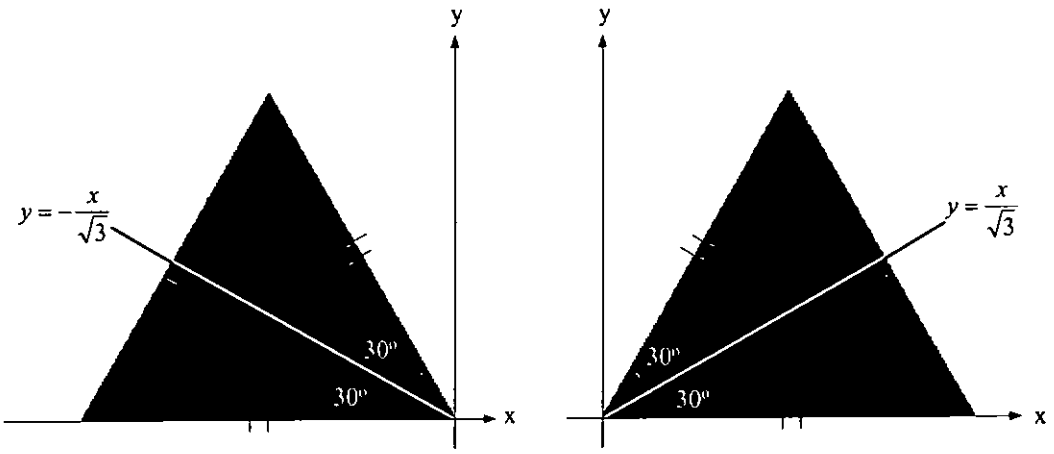


Figure 3-8 Examples of filled triangular objects with symmetry along $y = \pm x/\sqrt{3}$

For $(\mu_{03} - 3\mu_{21})^2 = 0$ in eqn. 3.12, we have,

$$c^2 = \frac{3(-a^2 + 2b^2 - 2ab)}{2} = \frac{3[(2b + a)(b - a) - ab]}{2} \tag{3.24}$$

We have $a > b$ and $c \neq 0$ by the definition of the triangular object as shown in Fig. 3-7. Therefore, both $b - a$ and $-ab$ are negative. It is impossible that c^2 is equal to a negative number. So, there is not any choice of a, b and c that could make $(\mu_{03} - 3\mu_{21})^2 = 0$.

In fact, if we want $(\mu_{03} - 3\mu_{21})^2 = 0$ for the triangular object, the object should be symmetry along $y = \pm\sqrt{3}x$. Fig. 3-9 gives two examples.

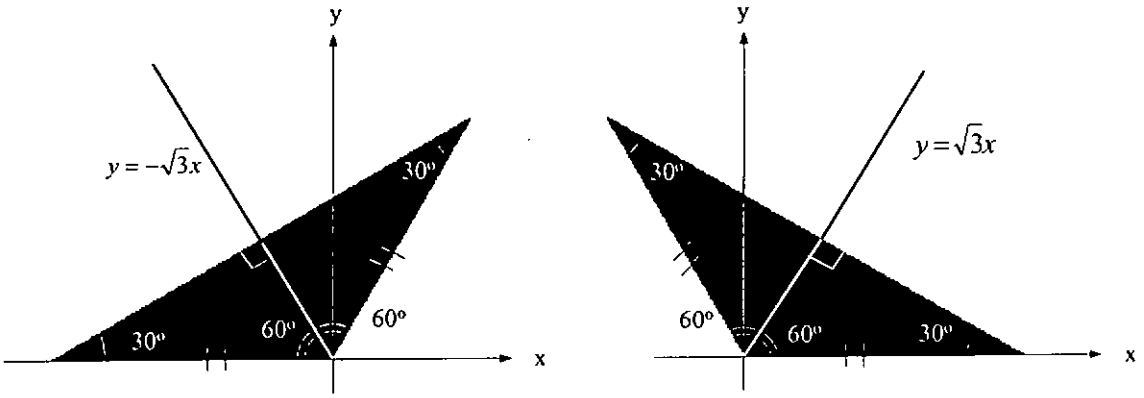


Figure 3-9 Examples of filled triangular objects with symmetry along $y = \pm\sqrt{3}x$

3.3.2. Fourth Moment Invariant Element

The expression of the fourth moment invariant $\phi_4 = \frac{(\mu_{30} + 3\mu_{12})^2 + (\mu_{03} + 3\mu_{21})^2}{\mu_{00}^5}$ in eqn. 2.33 is similar to that of the third moment invariant ϕ_3 in eqn. 2.32. We first recall the measurement directions of μ_{30}, μ_{12}, μ_{03} and μ_{21}. From eqn. 2.7, μ_{30} and μ_{12} give measurement in the x-direction, while μ_{03} and μ_{21} give measurement in the y-direction. It is difficult to have a measurement meaning on the parameters of $\mu_{30} + 3\mu_{12}$ or $\mu_{03} + 3\mu_{21}$, but when fitting the best-fit ellipse onto the object, it is found that we got the constant $a_f^2 + b_f^2$ term.

3.3.3. Experimental Analysis

Experimental analysis of the third moment invariant

Set 1: rotate 20°, 40° and 60° anti-clockwise for $(\mu_{30} - 3\mu_{12})^2 = 0$

A set of four triangular objects is generated as shown in Fig. 3-10. Fig. 3-10 (b), 3-10 (c), 3-10 (d) and 3-10 (e) are the rotated version of Fig. 3-10 (a). The rotational angles of Fig. 3-10 (b), 3-10 (c), 3-10 (d) and 3-10 (e) are 30°, 60°, 90° and 120° respectively. Table 3-4 illustrates the experimental readings of the third moment invariant.

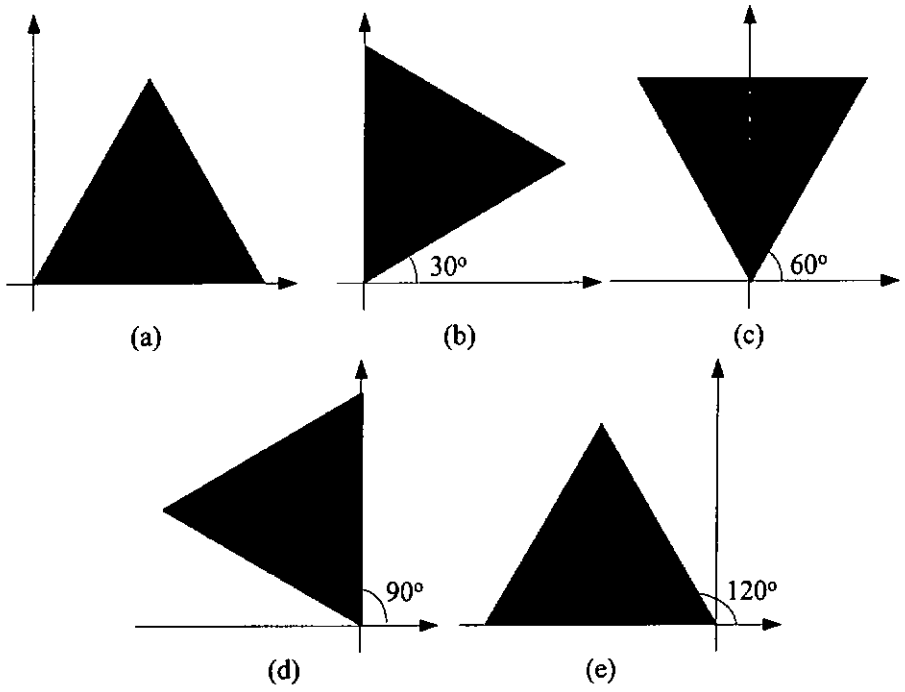


Figure 3-10 Five examples of filled triangular objects with five rotations

Table 3-4 Experimental readings of the triangular objects shown in Fig. 3-10(a-e)

Figure 3-10	Measures (in pixel)				$\frac{\mu_{30}}{3\mu_{12}}$	Third moment invariant ϕ_3
	Base *a*	*b*	Height *c*	Rotate		
(a)	200	100	173	0°	0.986946	0.0046
(b)	200	100	173	30°	-0.332531	0.0046
(c)	200	100	173	60°	-1.085169	0.0046
(d)	200	100	173	90°	-0.332531	0.0046
(e)	200	100	173	120°	0.986946	0.0046

From Table 3-4, it is observed that the third moment invariant parameters ϕ_3 of five triangular objects are the same, while $\mu_{30}/3\mu_{12}$ increases from Fig. 3-10 (a) to (d) and then decreases from Fig. 3-10 (d) to (e). It shows that if the object is symmetric along $y = \pm x/\sqrt{3}$, $(\mu_{30} - 3\mu_{12})^2$ tends to zero.

Set 2: rotate 20°, 40° and 60° anti-clockwise for $(\mu_{03} - 3\mu_{21})^2 = 0$

A set of four triangular objects is generated as shown in Fig. 3-11. Fig. 3-11

(b), 3-11 (c), and 3-11 (d) are the rotated version of Fig. 3.11 (a). The rotational angles of Fig. 3-11 (b), 3-11 (c), and 3-11 (d) are 20°, 40° and 60° respectively. Table 3-5 illustrates the experimental readings of the third moment invariant.

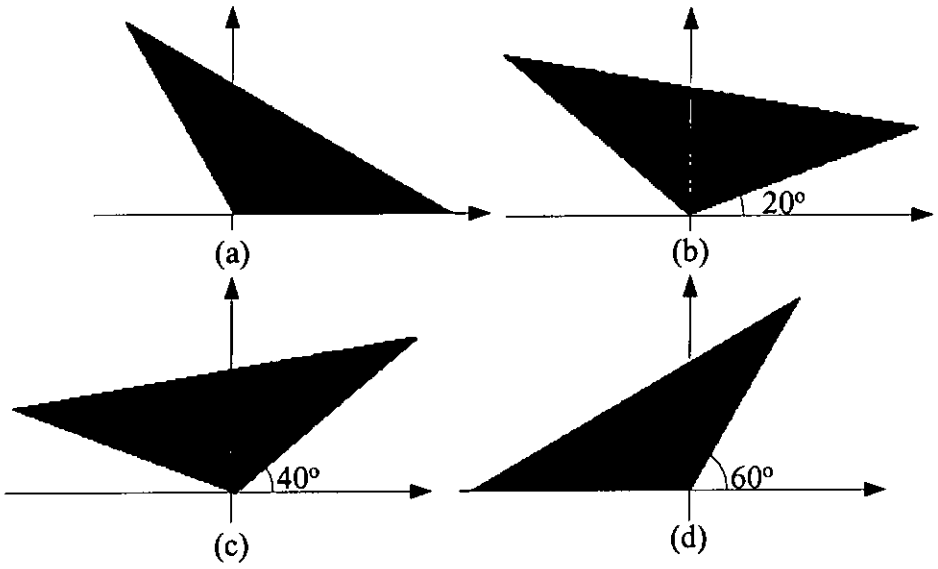


Figure 3-11 Examples of filled triangular objects

Table 3-5 Experimental readings of the triangular objects shown in Fig. 3-11(a-d)

Figure 3-11	Measures (in pixel)				$\frac{\mu_{03}}{3\mu_{21}}$	Third moment invariant ϕ_3
	Base a	b	Height c	Rotate		
(a)	230	115	177	0°	1.010491	0.0082
(b)	230	115	177	20°	-0.008209	0.0082
(c)	230	115	177	40°	-0.008742	0.0082
(d)	230	115	177	60°	1.010491	0.0082

The results from Table 3-5 are similar as that from Table 3-4. It can be seen that the third moment invariant parameters, ϕ_3, of four triangular objects are the same. Moreover, $\mu_{03}/3\mu_{21}$ increases from Fig. 3-11 (a) to (c) and then decreases from Fig. 3-11(c) to (d). In fact, it is consistent with the theoretical study that if the object is symmetric along $y = \pm\sqrt{3}x$, $(\mu_{03} - 3\mu_{21})^2$

tends to zero.

Experimental analysis of the fourth moment invariant

Set 1: constant lower base *a* but different *b* and height *c* for $\mu_{30}+3\mu_{12}=0$

A set of six triangular objects is used to analyze the parameter of $\mu_{30}+3\mu_{12}=0$ of the fourth moment invariant as shown in Fig. 3-12.

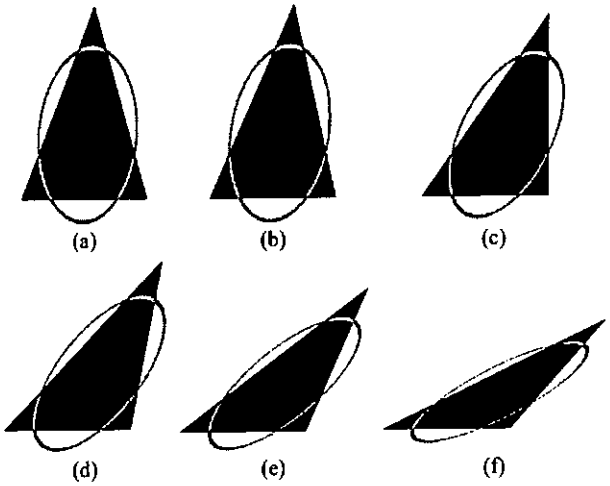


Figure 3-12 Examples of filled triangular objects

Table 3-6 Experimental readings of triangular objects as shown in Fig. 3-12 (a-f)

Figure 3-12	Measures (in pixel)			a_f	b_f	$a_f^2+b_f^2$	Fourth moment invariant ϕ_4
	Base *a*	*b*	Height *c*				
(a)	180	105	270	121.464	69.912	19641.191	0.00068861
(b)	180	120	268	121.752	69.426	19643.519	0.00068711
(c)	180	180	255	124.470	64.422	19642.975	0.00069908
(d)	180	225	234	127.746	57.672	19645.100	0.00077350
(e)	180	270	201	131.598	48.168	19638.190	0.00086360
(f)	180	315	149	135.810	34.614	19642.485	0.00119621

As shown in Table 3-6, $a_f^2+b_f^2$ remains unchanged for these six triangular objects, while the fourth moment invariant, ϕ_4, is increasing. It shows that when fitting a best-fit ellipse onto these objects, the sum of the

square of major principle axis a_f^2 and the square of minor principle axis b_f^2 is a constant.

Set 2: constant lower base *a* but different *b* and height *c* for $\mu_{03} + 3\mu_{21} = 0$

A set of four triangular objects is used to analyze parameters of $\mu_{03} + 3\mu_{21} = 0$ of the fourth moment invariant as shown in Fig. 3-13.

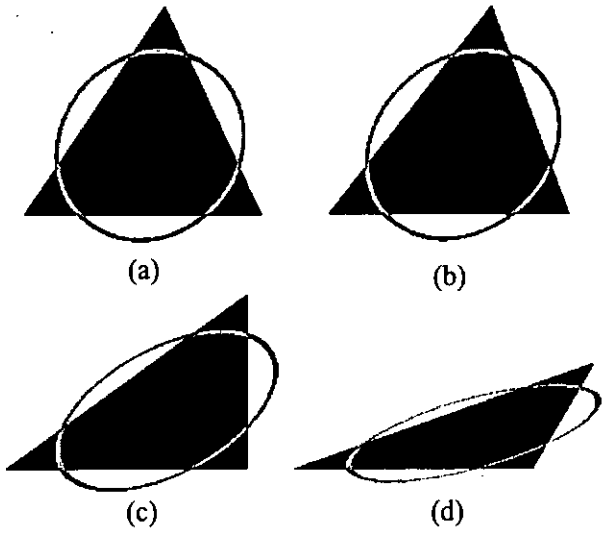


Figure 3-13 Examples of filled triangular objects

Table 3-7 Experimental readings of triangular objects as shown in Fig. 3-13 (a-d)

Figure 3-13	Measures (in pixel)			a_f	b_f	$a_f^2 + b_f^2$	Fourth moment invariant ϕ_4
	Base *a*	*b*	Height *c*				
(a)	180	105	155	73.371	66.619	54.564	0.00001152
(b)	180	120	153	76.523	62.974	54.564	0.00004550
(c)	180	180	128	88.012	45.558	54.564	0.00067670
(d)	180	225	78	95.726	25.650	54.564	0.00665928

Similar as Table 3-6, it shows in Table 3-7 that $a_f^2 + b_f^2$ remains constant for these four triangular objects, while the fourth moment invariant ϕ_4 is increasing. This implies that when fitting a best-fit ellipse onto these objects, the sum of the square of major principle axis a_f^2 and the square of minor

principle axis b_f^2 is a constant.

3.4. Chapter Conclusion

This chapter provides some interpretations of the second and third order-based moment invariants using the best-fit ellipse formulation. The theoretical analysis shows that the first invariant measures the total spread of the shape relative to its area square while the second invariant measures the degree of elongation of a best-fit ellipse on the shape. Six different shapes, including squares, rectangles, triangles, trapeziums, and two irregular (leaf- and horse-) shaped objects, and objects with Gaussian noise were used to confirm the findings. Moreover, we have also used a triangle as an example to elaborate some theories of the third and fourth invariants. It is found that the values of the third invariant for those objects with symmetry along $y = \pm x/\sqrt{3}$ and $y = \pm\sqrt{3}x$ remains constant. For the fourth invariant, when fitting a best-fit ellipse on the object, $a_f^2 + b_f^2$ becomes a constant.

Chapter 4

Improved Scheme for Object Searching using Moment Invariants

4.1. Introduction

Multimedia is the use of computers to process and /or transmit media information including text, audio, image and video. One of the major areas of interest in multimedia is multimedia information retrieval/ searching. For these kinds of applications, shape can capture the prominent elements of an object. The shape descriptor is not only able to do object identification but also can give a complete and real representation of it. Using a set of moment invariants is one of the most popular approaches employing shape descriptors that has been widely used for

object identification [15-17], character recognition [41], etc. The problem is often formulated using a continuous image function and the moments are defined using continuous integrals. It has been shown that this continuous formulation can ensure the moment invariants are independent of translation, scaling, and orientation of objects. In practice, moments have to be calculated on digital images and the discrete summations have to be used instead of the continuous integral. Then, the moment invariants calculated might not be truly invariant to scale, rotation and translation.

This chapter presents an analysis of quantization effects on several popular moment-based approaches. They are Hu moment invariants [40], Affine moment invariants [42-43], Dudani moment invariants [44] and Improved moment invariants [57]. Teh [76] and Salama [77] used only rectangular objects to illustrate error. This chapter extends their work by considering both regular (rectangular) and irregular objects.

We propose to apply this error analysis to object search applications. In particular, the analysis is useful for deciding the value of the threshold so that all scaled and rotated objects can be retrieved successfully.

This chapter is organized as follows. In Section 4.2, we discuss the moment invariants in the continuous and discrete domains and the quantization problem. Section 4.3 then provides some scaling properties of moment-based approaches. Sections 4.4 and 4.5 summarize the quantization effects of four moment-based approaches due to scaling and rotation respectively. Section 4.6 describes the use of error analysis in object retrieval.

4.2. Moments And Quantization Effects

4.2.1. Definition of Moment Invariants in Continuous Domain

As shown in Chapter 2, the central moments μ_{pq} in continuous domain (eqn 2.6) is defined as $\mu_{pq} = \iint_R (x-\bar{x})^p (y-\bar{y})^q f(x,y)dxdy$ where $p,q = 0,1,2,\ldots$.

4.2.2. Quantization Problem

For a digital image, the double integrals in μ_{pq} have to be replaced by the double discrete summations for approximation. Eqn. 2.7 is shown the definition of central moments μ_{pq} in discrete domain, which is $\mu_{pq} = \sum_{y=1}^{M}\sum_{x=1}^{N}(x-\bar{x})^p (y-\bar{y})^q f(x,y)$ where $p,q = 0,1,2,\ldots$

Using eqn. 2.7 to calculate μ_{pq}, the moment invariants defined in eqns. 2.17-2.23 might not be strictly invariant when the image is rotated or scaled. To illustrate this point, let us consider Fig. 4-1. It is clearly seen that errors always occur due to sampling and quantizing of the continuous image for digital computation.

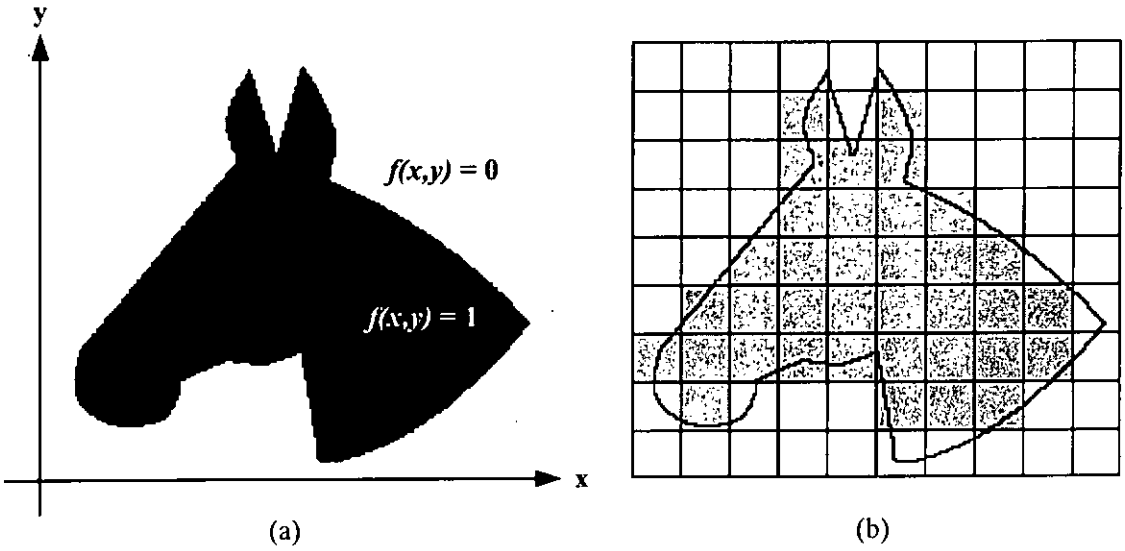


Figure 4-1 (a) Binary image in the continuous domain and (b) Sampled version of the binary image in the discrete domain

4.3. Quantization Due to Scaling

4.3.1. Regular (rectangular) Objects

We first consider the case of a regular (rectangular) object with length a and width b as shown in Fig. 4-2 (a). The scaled version of Fig. 4-2 (a) with a scaling factor of r is shown in Fig. 4-2 (b).

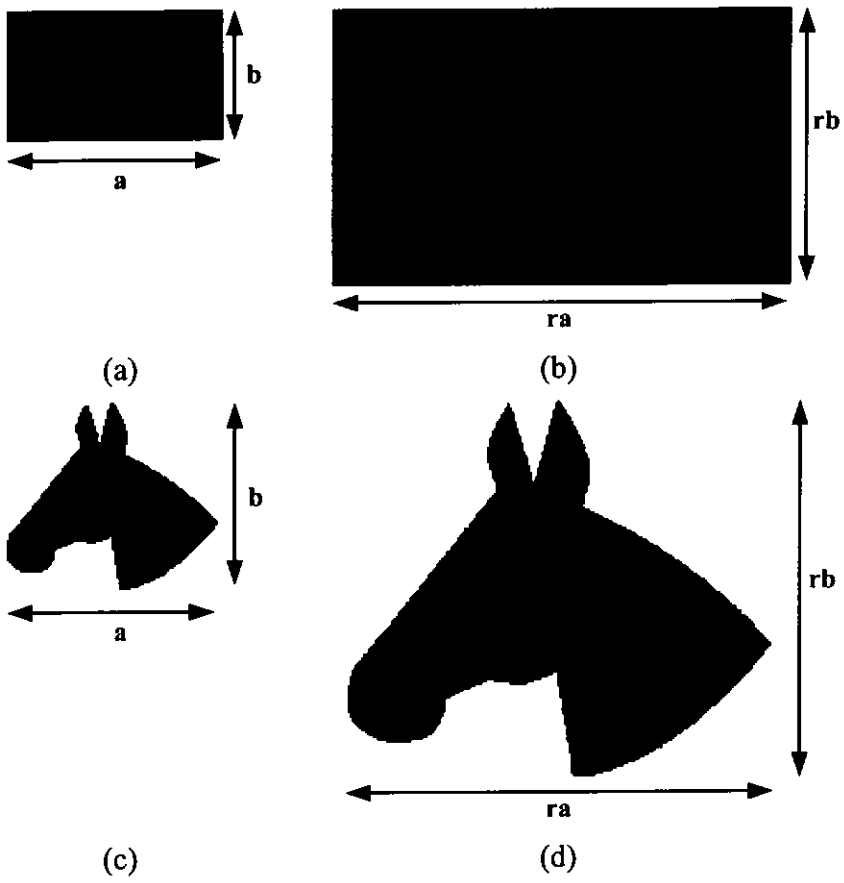


Figure 4-2 (a) Rectangular object and (b) its scaled version, (c) Horse-shaped objects and (d) its scaled version

4.3.1.1. Hu Moment Invariants

Expressions for ϕ_1 to ϕ_7 in eqns. 2.30-2.36 are now obtained for both continuous integral (eqn. 2.6) and discrete summation (eqn. 2.7) cases. We have,

CASE 1: CONTINUOUS INTEGRAL CASE 2: DISCRETE SUMMATION

$$\phi_1 = \frac{a^2+b^2}{12ab} \quad (4.1) \qquad \phi_1 = \frac{a^2+b^2-2/r^2}{12ab} \quad (4.4)$$

$$\phi_2 = \left(\frac{a^2-b^2}{12ab}\right)^2 \quad (4.2) \qquad \phi_2 = \left(\frac{a^2-b^2}{12ab}\right)^2 \quad (4.5)$$

$$\phi_3 \text{ to } \phi_7 = 0 \quad (4.3) \qquad \phi_3 \text{ to } \phi_7 = 0 \quad (4.6)$$

The quantization effect due to scaling has a bigger effect on ϕ_1 than ϕ_2. The difference comes from a scaling factor, r, in the numerator. When r tends to 0, the numerator of ϕ_1 in eqn. 4.4 will be significantly smaller than that in eqn. 4.1. Therefore, the difference becomes significantly large. If the scaling factor r is large enough, the value, i.e.-$2/r^2$, becomes very small and can be ignored. Then, the difference of ϕ_1 will be very small and tend to zero.

4.3.1.2. Affine Moment Invariants

Same as Section 4.3.1.1., the rectangular objects are used to study the scaling effect on ζ_1 to ζ_4 (c.f. Fig. 4-2 a-b). ζ_1 to ζ_4 in eqns. 2.40-2.43 are calculated for both cases. It is found that,

CASE 1: CONTINUOUS INTEGRAL CASE 2: DISCRETE SUMMATION

$$\zeta_1 = \frac{1}{144} \quad (4.7) \qquad \zeta_1 = \frac{1}{144} - \frac{(a^2+b^2)}{144r^2a^2b^2} + \frac{1}{144r^4a^2b^2} \quad (4.9)$$

$$\zeta_2 \text{ to } \zeta_4 = 0 \quad (4.8) \qquad \zeta_2 \text{ to } \zeta_4 = 0 \quad (4.10)$$

The difference of ζ_1 of the Affine moment invariants in these two cases is significantly large. We can see that ζ_1 in continuous case is a constant. Consider the discrete case, ζ_1 involves three parameters (length a, width b, and scaling factor r). When there is a change in the scaling factor r (for fixed a and b), ζ_1 becomes large as the scaling factor increases. Therefore, the error in ζ_1 would be large when the scaling factor tends to 0.

4.3.1.3. Dudani Moment Invariants

Again, expression for M_1 to M_7 in eqns. 2.45-2.51 can be found for both cases. We have,

CASE 1: CONTINUOUS INTEGRAL

$$M_1 = \left(\frac{a^2+b^2}{12ab}\right)^{\frac{1}{2}} \quad (4.11)$$

$$M_2 = \left(\frac{a^2-b^2}{a^2+b^2}\right)^2 \quad (4.12)$$

$$M_3 \text{ to } M_7 = 0 \quad (4.13)$$

CASE 2: DISCRETE SUMMATION

$$M_1 = \left(\frac{a^2+b^2-2/r^2}{12ab}\right)^{\frac{1}{2}} \quad (4.14)$$

$$M_2 = \left(\frac{a^2-b^2}{a^2+b^2-2/r^2}\right)^2 \quad (4.15)$$

$$M_3 \text{ to } M_7 = 0 \quad (4.16)$$

The scaling factor is present in the numerator of M_1 (eqn. 4.14), while it is present in the denominator of M_2 (eqn. 4.15) for discrete cases. The error in M_2 is larger than that in M_1, since the denominator in M_2 involves a power of 2 and the nominator in M_1 takes square root. Besides, as r increases, the value of M_1 decreases, while M_2 increases.

4.3.1.4. Improved Moment Invariants

As shown from the definition in section 2.5, it is mentioned that the Improved moment invariants are used to describe a shape boundary rather than its region. Therefore, we have to transform the objects (Fig. 4-2 a-b) into the boundary-based objects by detecting their edges with one-pixel width.

ψ_1 to ψ_7 of eqns. 2.58-2.64 for both cases becomes,

CASE 1: CONTINUOUS INTEGRAL

$$\psi_1 = \frac{a^2+b^2}{12a^2b^2} \quad (4.17)$$

$$\psi_2 = \left(\frac{a^2-b^2}{12a^2b^2}\right)^2 \quad (4.18)$$

$$\psi_3 \text{ to } \psi_7 = 0 \quad (4.19)$$

CASE 2: DISCRETE SUMMATION

$$\psi_1 = \frac{a^2+b^2-2/r^2}{12r^2a^2b^2} \quad (4.20)$$

$$\psi_2 = \left(\frac{a^2-b^2}{12r^2a^2b^2}\right)^2 \quad (4.21)$$

$$\psi_3 \text{ to } \psi_7 = 0 \quad (4.22)$$

It is easily seen that there are big differences for both the numerator and denominator of ψ_1. Both parts involve the scaling factor r. Therefore, the error in ψ_1 would be large even if the scaling factor showed a small change. Moreover, when considering ψ_1 and ψ_2 from the discrete summation case, it is seen that the error in ψ_2 is smaller than the error in ψ_1. This is because ψ_2 contains the scaling factor at the denominator only.

4.3.1.5. Experimental Analysis

In our experimental work, we generated a set of rectangular objects with $a=120$ pixels, $b=60$ pixels, and let scaling factor r change between 0.05 and 1.75. Fig. 4-3 shows a plot of errors against the scaling factor.

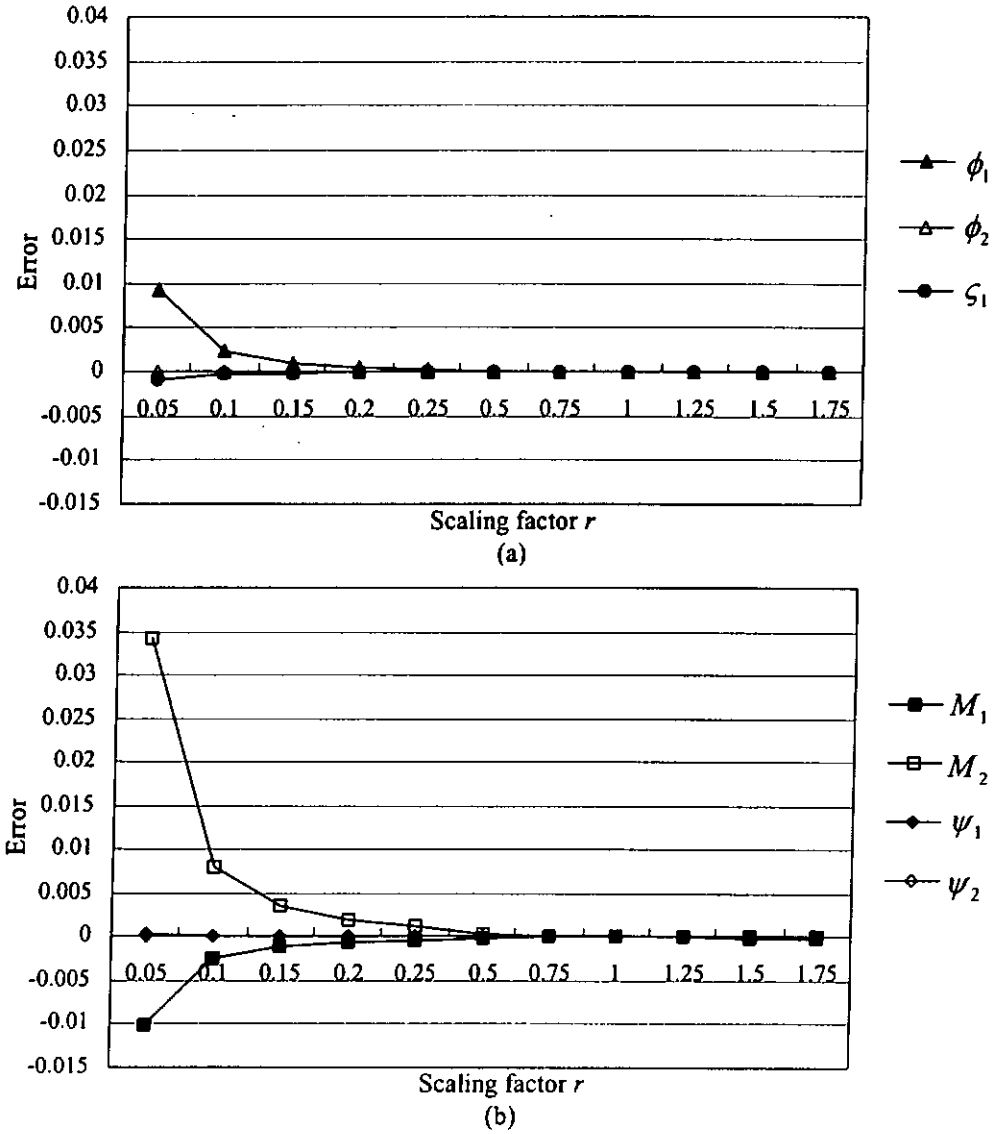


Figure 4-3 Plot of error against scaling factor for rectangular objects for (a) Hu and Affine moment invariants and (b) Dudani and Improved moment invariants

From this figure, we can observe that all four moment-based approaches show a similar trend. The errors tend to zero as the scaling factor *r* increases, but the errors become large when *r* decreases. When the scaling factor reaches 0.5, the errors are essentially equal to zero. In other words, if the size of the target object is half of the original size, the error due to scaling would not be significant using either of these four approaches. However, when the scaling factor is as small as 0.05, it can be seen that Dudani moment invariants are

affected heavily by the scaling.

In summary, concerning the overall sensitivity of these four approaches, it is shown that

- the most sensitive measure of the digitization/quantization error due to scaling on rectangular objects is Dudani moment invariants (M_1 and M_2), followed by Hu moment invariant (ϕ_1 and ϕ_2), and Affine moment invariants (ζ_1), whilst the Improved moment invariants (ψ_1 and ψ_2) give the least scaling error.
- When r is small, i.e. the size of the object is small, all four approaches show a large digitization/ quantization error.

4.3.2. Irregular Objects

In practical applications such as object searching or classification, we cannot expect the shape of most of the objects to be regular. In fact, in most cases, they are irregular objects. Hence, it is essential to perform an analysis on irregular objects. Fig. 4-4 shows the results of a set of irregular horse-shaped objects with $a=187$ pixels, $b=167$ pixels as shown in Fig. 4-2c, and the scaling factor r ranging between 0.05 and 1.75.

Note that we consider the first four parameters only. All other parameters can be ignored since their values are much smaller than 10^{-7}.

Comparing Fig. 4-3 and 4-4, there is no big difference in the general trend. This indicates that the four moment-based approaches are robust to quantization and digitization error when r ranges from 0.5 to 1.75. However, the Improved moment invariants approach suffers a higher error due to the scaling of irregular objects as compared to regular objects.

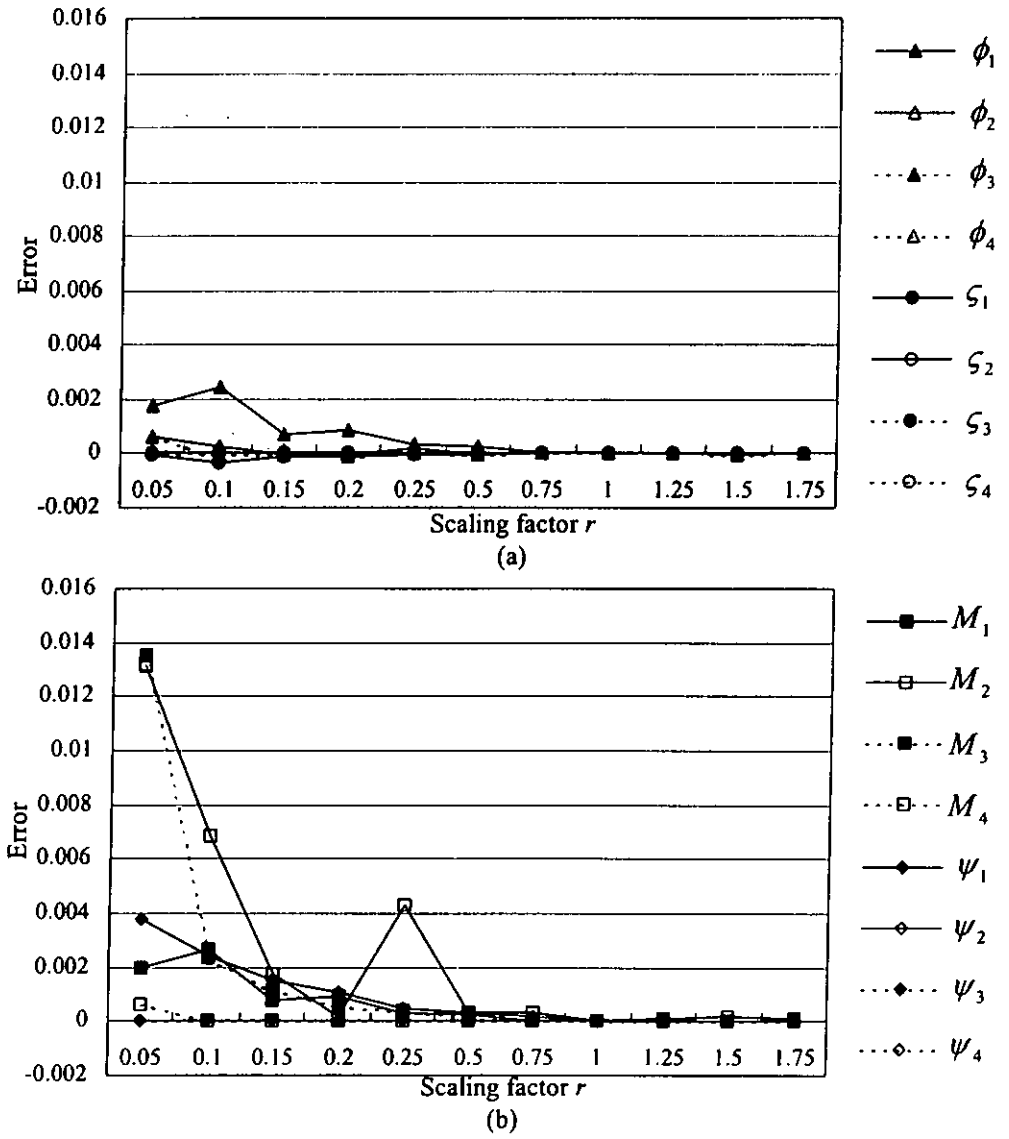


Figure 4-4 Plot of error against scaling factor for irregular objects for (a) Hu and Affine moment invariants and (b) Dudani and Improved moment invariants

Moreover, it can also be seen that

- the most sensitive measure to the error due to scaling on irregular objects is the Dudani moment invariants (M_1 to M_4) and this is followed by the Improved moment invariant (ψ_1 to ψ_4).
- Hu (ϕ_1 to ϕ_4) and Affine moment invariants (ζ_1 to ζ_4) have the smallest error due to scaling.

In general, for object searching applications, there are many objects with

different sizes. To retrieve similar objects, the descriptor should be able to achieve scaling invariance. From the above studies, it is found that all four moment-based approaches follow similar trends in error due to scaling. The error decreases as r increases.

4.4. Quantization Due to Rotation

4.4.1. Regular (rectangular) Objects

The previous section presents the analysis of error due to scaling. In this section, we give results of our studies on quantization effects due to rotation. Fig. 4-5 shows the rectangular and irregular objects and their rotated version. Let the angle of rotation be θ,

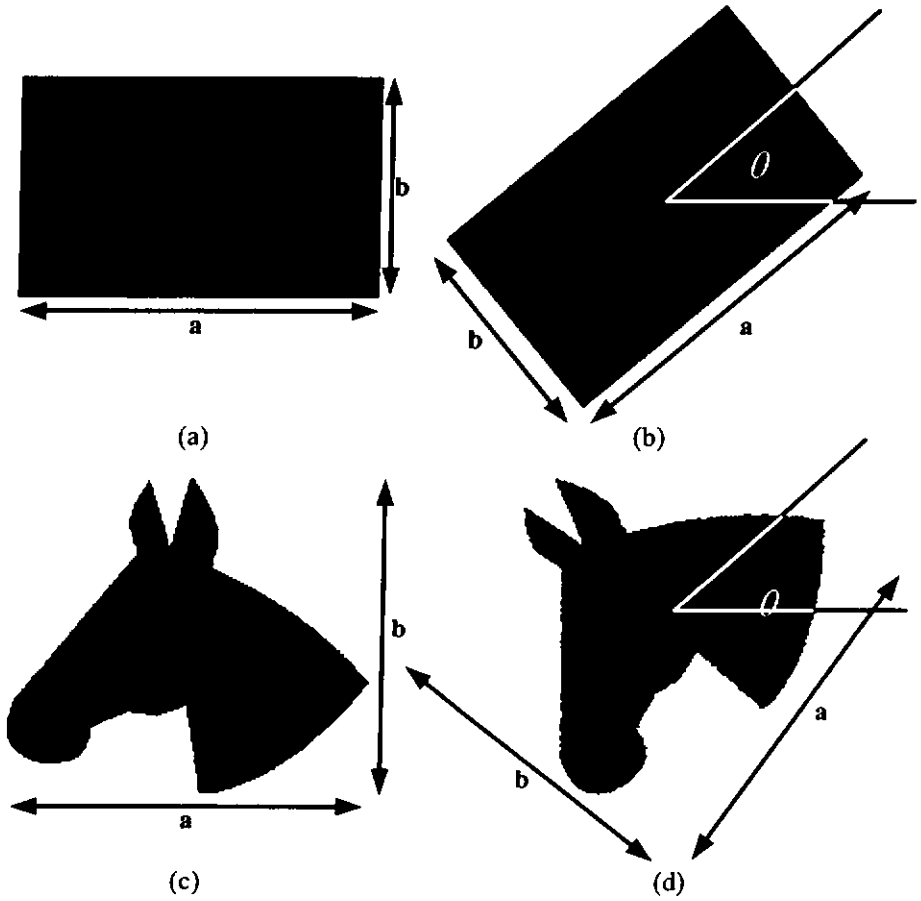


Figure 4-5 (a) Rectangular object and (b) its rotated version, (c) Horse-shaped objects and (d) its rotated version

4.4.1.1. Experimental Analysis

Fig. 4-6 shows the error curves of a set of rectangular objects, with $a=120$ pixels and $b=60$ pixels, and θ ranging from 0° to 180°.

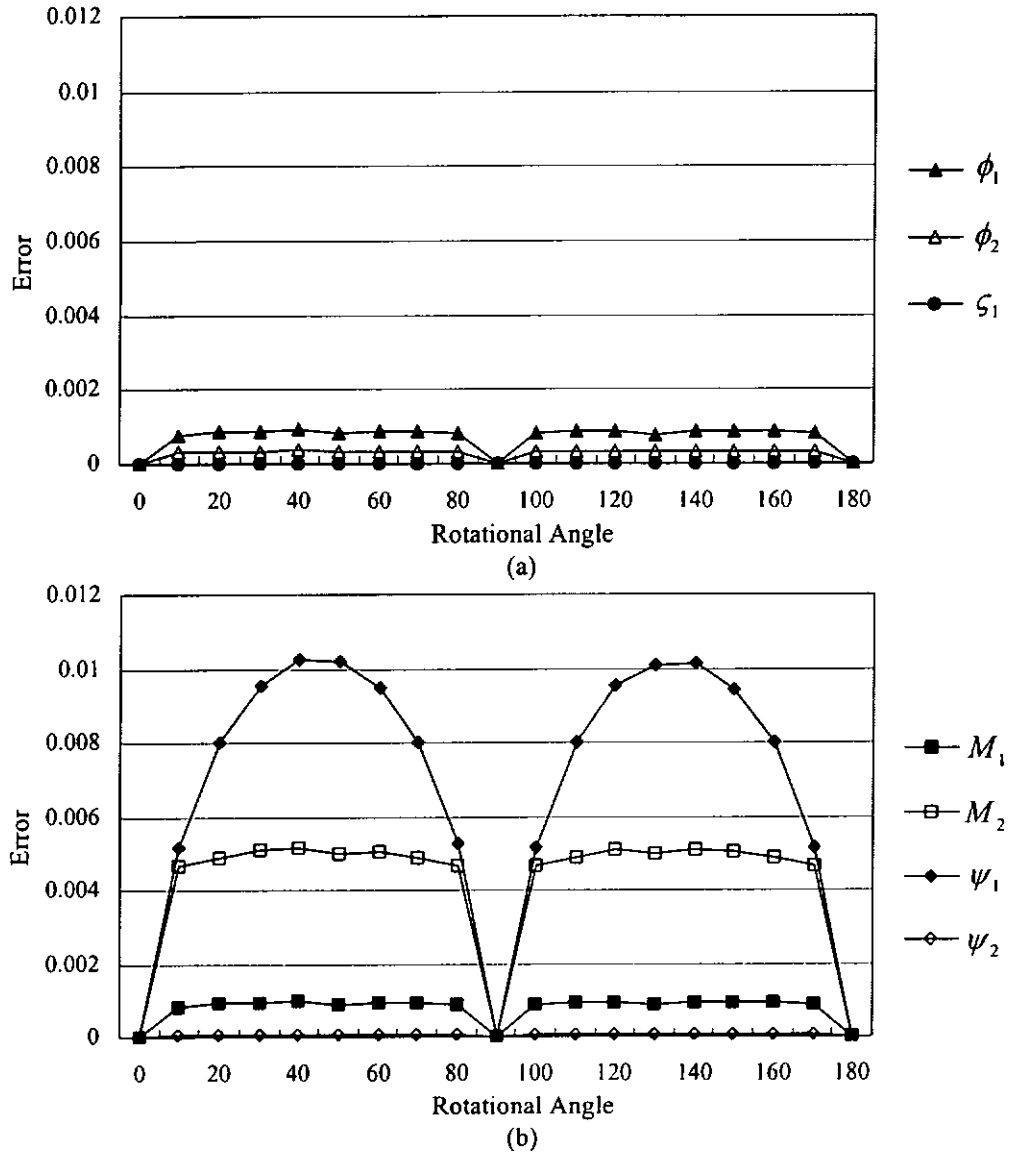


Figure 4-6 Plot of error against the angle of rotation for rectangular objects for (a) Hu and Affine moment invariants and (b) Dudani and Improved moment invariants

It can be seen from Fig. 4-6 that they form an "M" shape. There is no error at angles where $\theta = 0°$, 90° and 180°. Other than these three angles, ψ_1 suffers the largest rotational error at angles between 40° and 50° and between

130° and 140°. In fact, the second and the third most sensitive measures of error due to rotation are M_1 and M_2.

In summary, we have found that,

- the most sensitive measure of error due to rotation on rectangular objects is the Improved moment invariants (ψ_1 and ψ_2). It is then followed by Dudani moment invariants (M_1 and M_2).
- Hu (ϕ_1 and ϕ_2) and Affine moment invariants (ζ_1 and ζ_2) achieve the smallest error due to rotation.

4.4.2. Irregular Objects

4.4.2.1. Experimental Analysis

A set of irregular objects with $a=187$ pixels and $b=167$ pixels for θ ranges from 0° to 180° is examined.

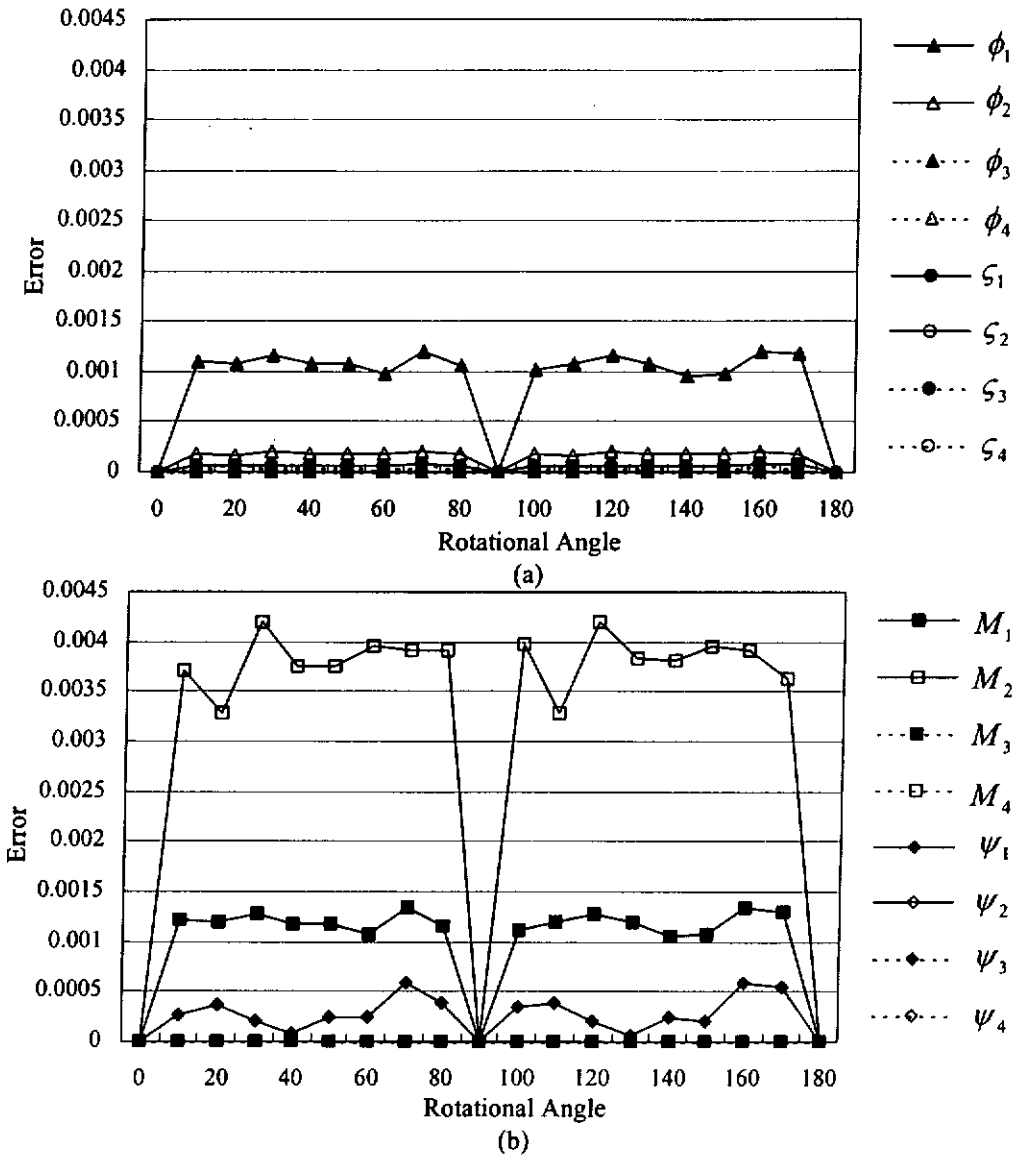


Figure 4-7 Plot of error against the angle of rotation for irregular objects for (a) Hu and Affine moment invariants and (b) Dudani and Improved moment invariants

Similar to the observations in Section 4.4.1.1, all curves in Fig. 4-7 show an "M" shape. None of the four approaches has error at the angles where $\theta = 0^{\circ}$, 90° and 180°. Apart from these three angles, the most sensitive measure of error due to rotation is M_2. The second and the third most sensitive measures are M_1 and ϕ_1.

Note also that

- the most sensitive measure of error due to rotation on irregular

objects is the Dudani moment invariants (M_1 to M_4). It is then followed by Hu moment invariants (ϕ_1 to ϕ_4).

- Improved (ψ_1 to ψ_4) and Affine moment invariants (ζ_1 to ζ_4) achieve the smallest error due to rotation.

In practice, we often find objects with different orientations. For example, we can have a triangle in an upward direction in an image, and a triangle in an upside down position in another image. However, the descriptor should be able to achieve rotation invariance. As found in the above study, all four moment-based approaches show similar behavior for errors due to rotation. The errors only take place at angles other than 0°, 90° and 180°. If the rotation is exactly at 0°, 90°, 180°, 270° or 360°, it is expected that no rotation error will occur.

To sum up, in order to avoid errors due to a digitization/ quantization effect on scaling, the ratio of the width/length between the examining object and the original object should not be less than 0.5. As the ratio decreases, the scaling error will be significantly increased. Therefore, for the same class of objects, we need to ensure that the size of the objects is within a certain range. On the other hand, it is usually impossible to restrict the object rotation to only 0°, 90° or 180° in order to avoid rotational errors.

4.5. Use of Error Analysis in Object Retrieval

In the previous two sections, we have proved that these four moment-based approaches suffer from different degrees of digitization/ quantization error due to scaling and rotation. In this section, we will take these into account in the design of the object searching system so that the accuracy of the matching process can be improved.

Let us review a two-stage searching algorithm as shown in Fig. 4-8. In the image searching part, both global and local characteristics of an input image are used in the matching process. The first step tries to acquire the global information, which aims to produce a fast listing from the whole database to select those objects with similar global characteristics; while the second step obtains the local information about the object. The purpose is to have a detailed matching between the query object and the short-listed objects. The advantage of using a two-stage algorithm is that there is no need to do the matching on the whole database for every query submission. The global information about the object is extracted first to avoid the requirement for a detailed matching. This can reduce the computational complexity, resulting in an efficient and effective object searching algorithm. However, to achieve a good accuracy for retrieving scaled and rotated objects, a threshold selection is crucial.

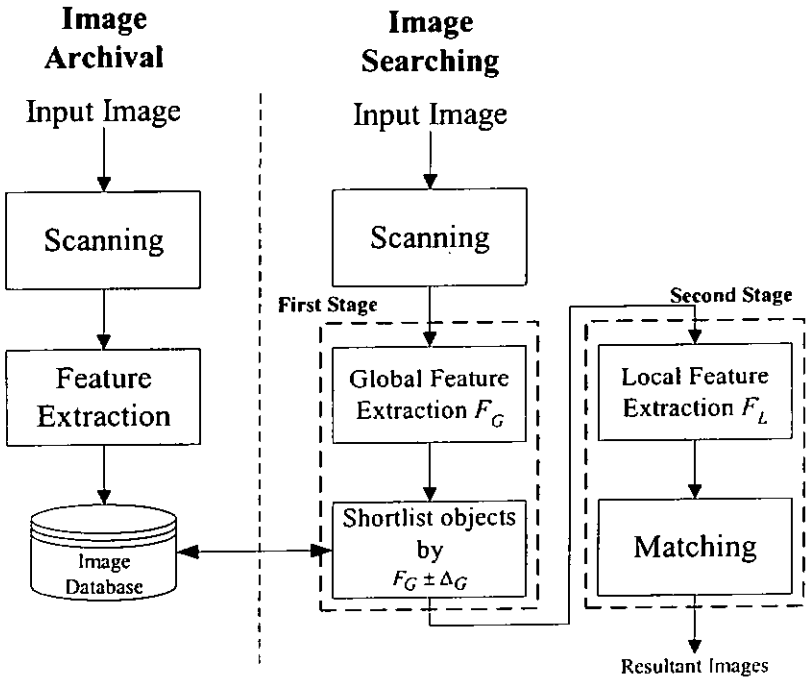


Figure 4-8 A two-stage object searching model

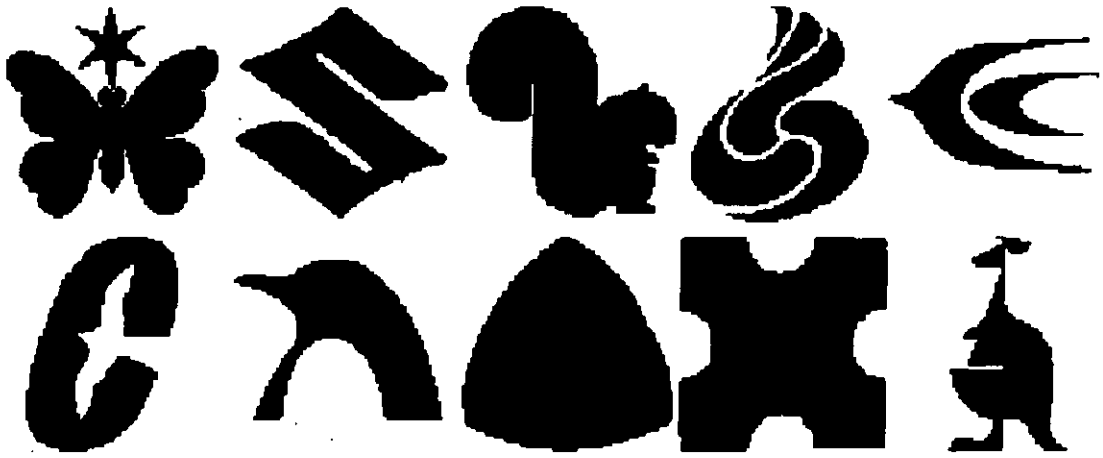


Figure 4-9 Some example objects in the test database

We have tested this selection method for the approach with the biggest quantization effect: Dudani moment invariants approach. Let F_G be the feature vector of the first stage operation, i.e.,

$$F_G = (M_1 \quad M_2 \quad \cdots \quad M_7)^T \tag{4.23}$$

and Δ_G be the error limit, i.e.,

$$\Delta_G = (\lambda_1 \quad \lambda_2 \quad \cdots \quad \lambda_7)^T \tag{4.24}$$

where λ_i is the maximum error due to the quantization effect on corresponding M_i.

We search relevant objects within the range, $F_G \pm \Delta_G$, so that a range of possible objects can be retrieved. If Δ_G equals zero, only one object will be selected at this stage, which is not desirable. If Δ_G is set to a large value, we cannot guarantee low computational complexity and too many irrelevant objects may be retrieved. In order to achieve a high retrieval rate and at the same time to have a low computational complexity, the threshold, Δ_G, can be obtained by using the analysis in this chapter, which takes consideration of the quantization effect of scaling and orientation.

A set of the experimental databases was setup: 1000 images of 100 classes containing digitally scaled and rotated objects. Each class contained 5 scaling factors (0.05, 0.15, 0.25, 0.5, and 0.75) and 5 rotational angles (0°, 23°, 45°, 68° and 90°). Fig. 4-9 shows some examples of the test objects. Two queries were submitted to the system. These two objects are the rectangular object in Fig. 4-2a and the horse-shaped object in Fig. 4-2c. For the first query with the rectangular object, Figs. 4-3 and 4-6 show that the highest error of M_1 for Dudani approach is 0.012; while M_2 is 0.035 (defined as Set 1). Then we put $\Delta_G = \begin{pmatrix} 0.012 \\ 0.035 \end{pmatrix}$, while M_3 - M_7 are zero. For the second query with the horse-shaped object, the highest errors of the first four elements of Dudani approach are 0.003, 0.014, 0.014 and 0.001 (defined as Set 5) respectively, which can be deduced from Fig. 4-4 and 4-7. Then we put $\Delta_G = \begin{pmatrix} 0.003 \\ 0.014 \\ 0.014 \\ 0.001 \end{pmatrix}$, while λ_5 - λ_7 are ignored as $M_5 - M_7$ are smaller than 10^{-7}. Furthermore, we also tested the cases for selecting bigger Δ_G. Sets 2, 3 and 4 were obtained by multiplying Set 1 with 2, 5 and 10 respectively. Similarly, sets 6, 7 and 8 were obtained by multiplying Set 5 with 2, 5 and 10 respectively. Table 4-1 shows the results of these eight sets of selection.

Table 4-1 Results of different selections

Δ_G	Number of object retrieved in the same class	Number of object retrieved in the other classes
Rectangular objects		
Set 1	10	0
Set 2	10	7
Set 3	10	94
Set 4	10	589
Horse-shaped objects		
Set 5	10	0
Set 6	10	35
Set 7	10	193
Set 8	10	287

From Table 4-1, it is seen that when we set Δ_G to the highest error (i.e. Set 1 and 5), it can guarantee that all scaled and rotated objects in the same class are selected and no object from other classes is included in the retrieval. However, when Δ_G is set to a larger value (i.e. Set 4 and 8), it is seen that many objects from other classes are obtained in the retrieval result. In fact, it also indicates that if the threshold is wrongly selected, the retrieval result will include objects from other classes and thus the computational complexity for the matching process increases significantly. Therefore, the error analysis provides a way of choosing the threshold and is useful for ensuring that the two-stage approach is invariant to translation, scaling and rotation.

4.6. Chapter Conclusion

This chapter has presented an analysis of quantization error for four moment-based approaches. They are Hu moment invariants, Affine moment invariants, Dudani moment invariants and Improved moment invariants. Analytical expressions in continuous and discrete domains have been derived and experimental results have

also been found for regular (rectangular) and irregular objects undergoing scaling and rotation changes.

We have found that the error due to scaling increases as the object size decreases. The Dudani moment invariants tend to have a larger error than that of the other three moment-based approaches. Moreover, the error due to rotation increases as the object is rotated with an angle other than $\theta = 0^{o}$, $90^{o}$and 180^{o}.

In order to ensure the shape descriptor is invariant to translation, scaling and rotation, we can apply the analysis of the quantization effects to scaling and rotation for selecting the threshold for an object searching system. From our experimental work, it is proved that all scaled and rotated objects can be retrieved correctly by setting the threshold to the value of the error.

Chapter 5

Multi-layer Shape Descriptor (MLSD) and Improved Mechanism with Both Global and Local Characteristics

5.1. Introduction

The best-fit ellipse concept of a binary object from second order Hu moments is discussed in Chapter 3. The theoretical study shows that the first invariant measures the total spread of the shape relative to its area square while the second invariant measures the degree of elongation of a best-fit ellipse on the shape. Moreover, they are confirmed by the experimental analysis using triangles,

trapeziums and irregular leaf shapes. In this chapter, to apply these findings practically, a Multi-Layer Shape Descriptor (MLSD) invariant to rotation, scaling and translation and an improved two-step mechanism are proposed. For MLSD, a given binary object is divided into four sub-regions by two principal axes of the best-fit ellipse, which is located at the center of mass. Each sub-region can be then subdivided into four sub-regions in the same way. The node of a quadtree representation is corresponding to each region of the object. Three kinds of parameters invariant to rotation, scaling and translation are calculated for each node. Then the shape descriptor is represented as a vector of all parameters. In order to achieve an effective and efficient retrieval system, we propose an improved two-step mechanism, which considers both global and local characteristics of the object. All objects in the database are sorted by the global characteristics, i.e. the ratio of major and minor principle axes of the best-fit ellipse. Then, for those objects having similar ratio, the similarity distance is calculated by the Euclidean distance of their feature vectors. Furthermore, the experimental results conforming to MPEG-7 core experiments are also presented.

This chapter is organized as followed. In Section 5.2, we present a proposed Multi-Layer Shape Descriptor (MLSD). A detailed elaboration of the feature vector, which is invariant to rotation, scaling and translation, is given also. In Section 5.3, the similarity matching with both global and local characteristics is presented. The global characteristic of a shape is obtained from the best-fit ellipse concept, whereas the local characteristic is obtained from the MLSD. In Section 5.4, we give the experimental results to indicate that our shape descriptor and improved two-step mechanism gives superior performance over other moment-based approaches in terms of retrieval accuracy and computational speed.

5.2. Multi-Layer Shape Descriptor (MLSD)

We have made a discussion on using the concept of best-fit ellipse from its realization using second-order based moment invariant in the Chapter 3. Let us have a look on the MPEG-7 eXperimentation Model (XM). In this model, the bounding box descriptor is obtained by fitting a tightest rectangular box. The tightest box is designed such that it fully contains the object and the sides of the box are parallel to the principal axes of the object. Moreover, another shape descriptor by Sluzek [72] tries to describe the prototype object by a family of shapes. These shapes are created by occluding the object by circles of different radius located in the object's center of mass. Our proposed multi-layer shape descriptor (MLSD) applies the best-fit ellipse concept. MLSD extends and enhances the circle-based bounding box descriptor proposed by Sluzek. The main enhancement is on the fitting element, which becomes ellipses rather than a rectangular box [78-79] or circles [72].

5.2.1. Recursive Sub-division of an Object on the Basis of the Best-fit Ellipse Concept

For a given object O, we first define a_o and b_o as the major and minor axes of the best-fit ellipse respectively. These two unique axes can represent the object. The shape of an object is classified using the following procedures:

Step1. The major and minor axes of the best-fit ellipse of the object, a_o and b_o, are obtained as,

$$a_o = \left(\frac{4}{\pi}\right)^{\frac{1}{4}} \left(\frac{\mu_{20}^3}{\mu_{02}}\right)^{\frac{1}{8}} \quad \text{and} \quad b_o = \left(\frac{4}{\pi}\right)^{\frac{1}{4}} \left(\frac{\mu_{02}^3}{\mu_{20}}\right)^{\frac{1}{8}}$$

and then, the orientation is defined as,

$$\tan 2\theta_o = \frac{2\mu_{11}}{\mu_{20} - \mu_{02}}$$

Step2. The shape is divided into four sub-regions by two principal axes corresponding to the two axes of best-fit ellipse, a_o and b_o, at the centre of mass of the shape. Fig. 5-1 demonstrates the case of an original object after the first division. The original object is divided into four regions: A, B, C and D.

Step3. For each sub-region, the major and minor axes of the best-fit ellipse calculation and the sub-division processes of steps1 and 2 are performed. Fig. 5-2 shows the case of the original object after the second division. Each region of A, B, C and D is divided into four sub-regions of AA-AD, BA-BD, CA-CD and DA-DD respectively.

Step 4. From this division concept, each region can be represented using the quadtree structure as in Fig. 5-3. Each node represents a region. The parameters which are invariant to rotation, scaling and translation of each node are extracted. A shape descriptor is represented as a vector of all parameters obtained.

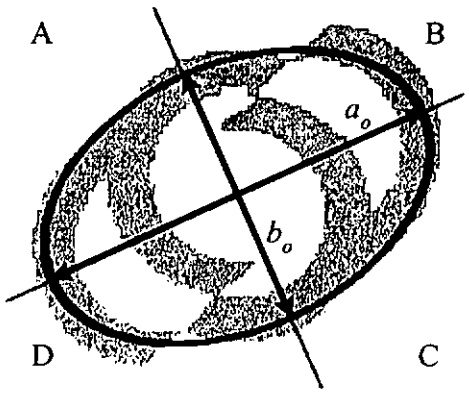


Figure 5-1 The object after first sub-division

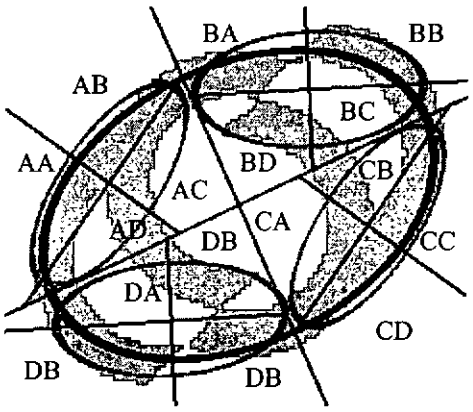


Figure 5-2 The object after second sub-division

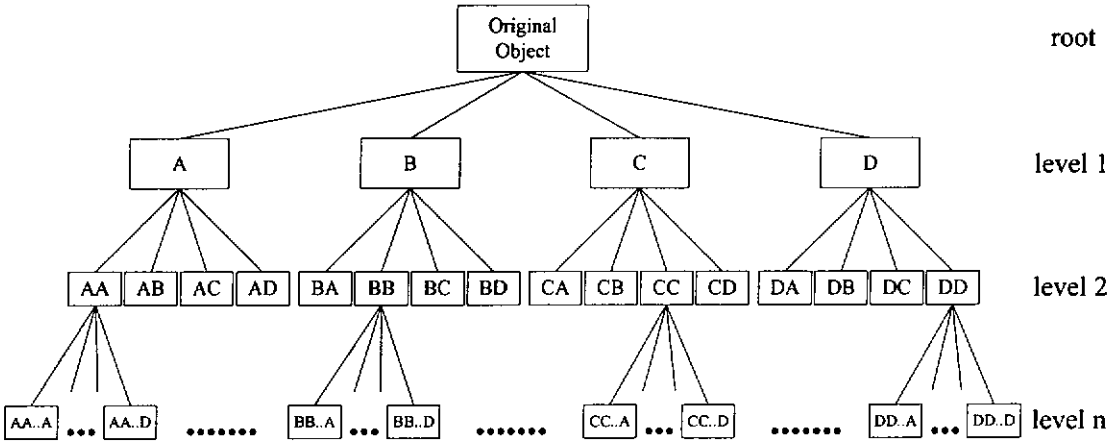


Figure 5-3 A quadtree representation of object with n-level decompositions

5.2.2. Parameters Invariant to Rotation, Scaling and Translation

Three parameters, namely *axes_ratio*, *pixel_insideBFE_ratio* and *pixel_outsideBFE_ratio* which are invariant to rotation, scaling and translation are obtained for each node. For a node i, let N_w be the total number of pixels belonging to the whole object.

- *axes_ratio*: the ratio of major axis value to the minor axis value of node i. The term *axes_ratio* of node i is expressed as,

$$axes\_ratio(i) = \frac{a_i}{b_i} \tag{5.1}$$

- *pixel_insideBFE_ratio*: the ratio of pixel inside the ellipse to total pixels in a region *r*. The term *pixel_insideBFE_ratio* for region *r* is expressed as,

$$pixel\_insideBFE\_ratio(r) = \frac{P_{inside\_r}}{N_r} \tag{5.2}$$

- *pixel_outsideBFE_ratio*: the ratio of pixel outside the ellipse to total pixels of a region *r*. The term *pixel_insideBFE_ratio* for region *r* is expressed as,

$$pixel\_outsideBFE\_ratio(r) = \frac{P_{outside\_r}}{N_r} \tag{5.3}$$

For each node of the quadtree, it consists seven parameters. Let us consider the root node as an example. Their parameters are as follows.

(1). Principal axes ratio of the object: $\frac{a_{root}}{b_{root}}$; Orientation: θ_{root}

(2). x-coordinate of center of mass: $\bar{x} = m_{10}/m_{00}$; y-coordinate of center of mass: $\bar{y} = m_{01}/m_{00}$

(3). $pixel\_insideBFE\_ratio(root) = \frac{P_{inside\_w}}{N_w}$;

$pixel\_outsideBFE\_ratio(root) = \frac{P_{outside\_w}}{N_w}$

(4). $pixel\_insideBFE\_ratio(A) = \frac{P_{inside\_A}}{N_A}$;

$pixel\_outsideBFE\_ratio(A) = \frac{P_{outside\_A}}{N_A}$

(5). $pixel\_insideBFE\_ratio(B) = \frac{P_{inside\_B}}{N_B}$;

$pixel\_outsideBFE\_ratio(B) = \frac{P_{outside\_B}}{N_B}$

(6). $pixel\_insideBFE\_ratio(C) = \frac{P_{inside\_C}}{N_C}$;

$pixel\_outsideBFE\_ratio(C) = \frac{P_{outside\_C}}{N_C}$

(7). $pixel\_insideBFE\_ratio(D) = \frac{P_{inside\_D}}{N_D}$;

$pixel\_outsideBFE\_ratio(D) = \frac{P_{outside\_D}}{N_D}$

5.3. Similarity Matching with both Global and Local Characteristics

In the previous section, we have described the shape descriptor (MLSD) of an object, which makes use of second-order moments. In this section, we will draw the attention to a practical application, i.e. trademark image retrieval system.

5.3.1. The Framework of the Trademark Image Retrieval System

Let us firstly review the searching model of a traditional retrieval method. Fig. 1-1 shows a block diagram of a generic image archival and searching model. For the image archival part, each input image is scanned and its feature vector is extracted. These feature vectors will be stored into an image database. Once a query image is submitted for searching those relevant images in the database, the image searching part begins. The feature of the query image is extracted and then used to match all feature vectors in the database. After the matching process, the resultant relevant images are displayed to the user.

In fact, when considering this image searching part, it is easily seen that the process is the same for all query image submission and the computational complexity is high, especially in the matching process. All images in the image database are needed to be accessed for every inputted image. Therefore, when the size of the database is large, accessing all images for searching causes heavy computational load and is time consuming.

In order to achieve an effective and efficient retrieval system, the matching process should be modified so as not to load all feature vectors from

the whole database for every query image submission. In other words all database images should be accessed only once, i.e., at the time when the user submits the first query. To do this, we can consider extracting the global characteristic of each object. The selection of such global information should be done in a simple and fast way. Therefore, the idea of the best-fit ellipse, i.e. the ratio of major and minor principle axes of best-fit ellipse in the object, is a sensible choice. It is because the ratio can describe the object globally with simple and fast calculations. Fig. 4-8 demonstrates the framework of the two steps searching model for image retrieval application. Moreover, Fig. 5-4 presents several example images with its best-fit ellipses. From the major and the minor axes, the shapes can be categorized into different groups according to their aspect ratio. Thus, it is clear that the global characteristic of each image can be described by using the concept of the best-fit ellipse.

Figure 5-4 Example image with global characteristic

5.3.2. Detailed Descriptions of the Trademark Image Retrieval System

The trademark image retrieval system aims to retrieve a group of relevant images in a trademark image database. The retrieval approach consists of two

stages. The first stage tries to acquire the global characteristic; while the second stage obtains the local characteristic. The global information employs the property of second order moments. That is the ratio of principle axes of best-fit ellipse in the objects. For the local information, it measures the *axes_ratio*, *pixel_insideBFE_ratio* and *pixel_outsideBFE_ratio* between the query image and the selected images.

The flow diagram of the trademark image retrieval system is shown in Fig. 5-5. The detailed description of the retrieval system is as follows:

Step1. The user inputs a query image.

Step2. After the scanning, its feature vector is extracted. It is assumed that the major and minor principal axes of the best-fit ellipse on the query object are a_f and b_f respectively and their calculations are defined in eqns.3.9-3.10.

Step3. Every time when the user submits a query, we check whether it is the first submission. If it is so, a list is generated to index all database images in ascending order of their global characteristics. This characteristic is done by calculating the ratio of major and minor principle axes of the best-fit ellipse on the object. The advantages of using such a ratio are not only simplicity and rapid calculation but also its ability in defining the global property of a shape, i.e., the elongation.

Step4. If the system finds that the first query has been already submitted, then it will bypass Step 3 and go to Step 5.

Step5. Only a sub-set of images is selected for the matching process. Of course, we look for images that match the global properties of

the query image. This image set is chosen by pre-defining a threshold Δ_{GI}. If the ratio of the principal axes of the object in a database image is within $a_q/b_q \pm \Delta_{GI}$, the database image is inserted into the selected sub-set for measuring the local characteristic.

Step6. The MLSD feature vectors between the query image and each image in this set are calculated.

Step7. At last, the similarity score is measured by the Euclidean distance between the query feature vector and the image feature vectors in the set. For n level of decomposition, the total number of node is

$$TotalNode = \sum_{i=0}^{n} 4^i \tag{5.4}$$

The similarity score is defined as,

$$Similarity\ Score = \sum_{i=0}^{TotalNode} \sum_{j=2}^{6} (q_{ij} - s_{ij}) \tag{5.5}$$

The resultant images are obtained according to this similarity score and displayed to the user.

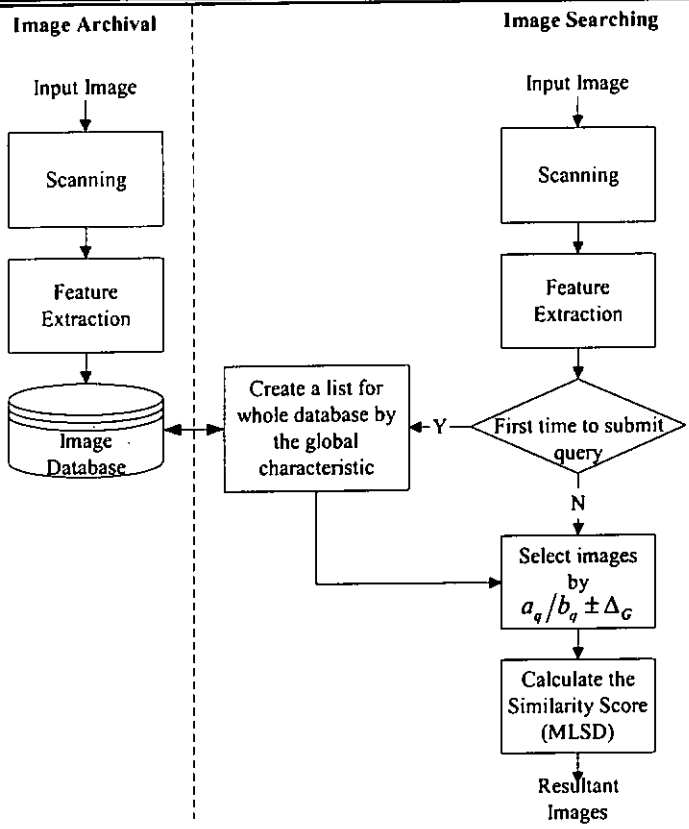


Figure 5-5 A flow diagram of the trademark image retrieval system

In summary, the proposed framework is designed using the best-fit ellipse concept and an improved mechanism for image matching. Global information about the shape is extracted from the parameters of the best-fit ellipse while unnecessary image matching is avoided. These can reduce the computational complexity, resulting in an efficient and effective image searching algorithm. Moreover, local information is extracted from the sub-division of the shape. These can further improve the retrieval accuracy of the algorithm.

5.4. Experimental Results

We have tested the proposed shape descriptor (MLSD) after the first sub-division and four moment-based approaches (they are Hu moment invariants, Affine moment invariants, Dudani moment invariants and Improved moment invariants) with the core experiment for shape descriptions in MPEG-7. In this core

experiment, it contains 2810 trademark images (Set 2), which are disjoint binary regions and categorized into 10 groups. In order to evaluate the rotation and scaling properties, another set (Set 1), which contains 2000 trademark images, was set up. They were generated by selecting 20 images from each group. Then, each selected image was scaled with 5 scaling factors on both width and length. The scaling factors were 0.25, 0.50, 0.75, 1.25 and 1.50. Moreover, each selected image was also rotated with 5 rotational angles. They were 0°, 23°, 45°, 68° and 90°. Fig. 5-6 shows some samples of those trademark images.

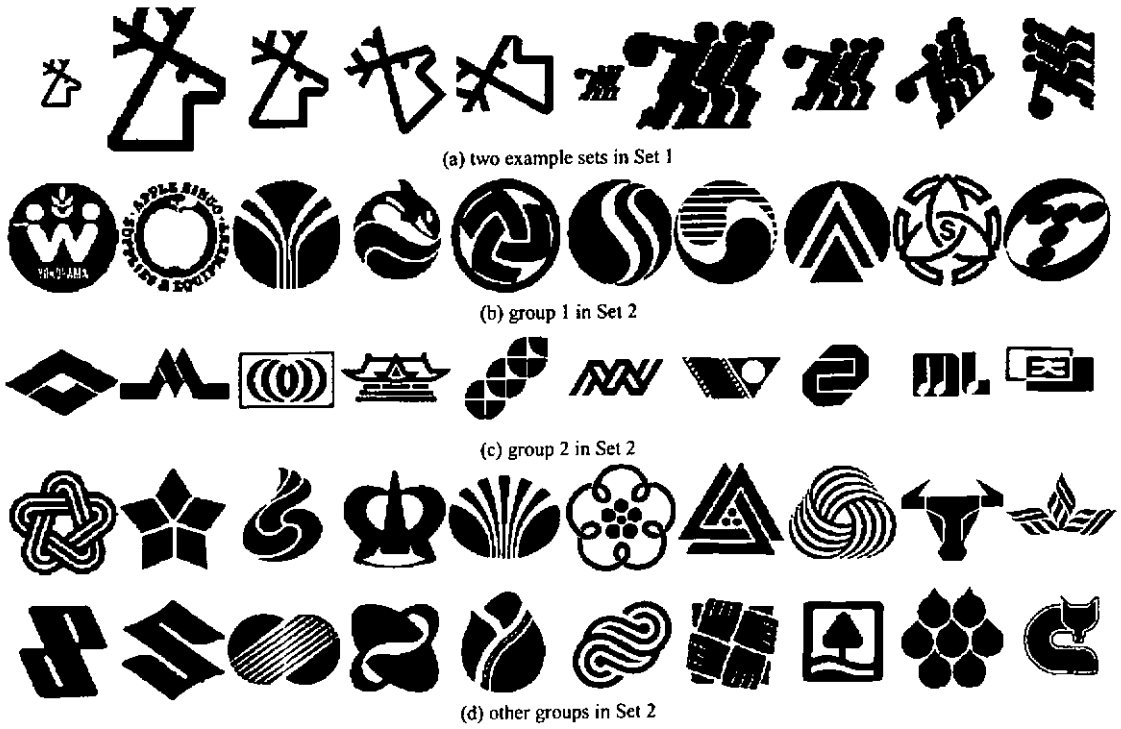


Figure 5-6 Example images in Set 1 and Set 2

In our experiments, the retrieval accuracy is defined as the ratio of the number of relevant images in the N-top scores of the retrieved image set divided by N, the number of relevant images in the database. Tables 5-1 and 5-2 show the comparison of the retrieval accuracy between four moment-based approaches and our proposed shape descriptor (MLSD) with improved two-step mechanism for Set

1 and Set 2 images respectively.

Table 5-1 Retrieval accuracy of Set 1

Shape descriptor	Δ_G	Retrieval accuracy	Average number of comparison for each image (percentage saved)
Hu	-	79.57	2000 (0%)
Affine	-	56.39	2000 (0%)
Dudani	-	80.99	2000 (0%)
Improved	-	32.11	2000 (0%)
MLSD	-	81.95	2000 (0%)
MLSD+ Improved mechanism	0.01	84.38	201 (89.95%)
	0.02	**85.59**	**309** (84.55%)
	0.04	85.45	460 (80.75%)
	0.06	85.44	550 (77.00%)
	0.08	85.42	603 (69.85%)
	0.1	85.23	654 (67.30%)
	0.5	83.64	1348 (32.60%)
	1	82.86	1735 (13.25%)

From Table 5-1, it is seen that the proposed MLSD can achieve the highest retrieval accuracy compared with other four moment-based approaches. The retrieval system is not only able to improve the accuracy, but also speeds up the searching and saves as far as 84.55% of the computation. Thus, it indicates that the global characteristic, i.e. the ratio of major and minor principle axes of best-fit ellipse in the object, can successfully describe the shape at an early stage so as to give quantity matching in the database and therefore to speed up the retrieval system.

Table 5-2 Retrieval accuracy of Set 2

Shape descriptor	Retrieval accuracy										Average number of comparison for each image (percentage saved)
	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10	
Hu	35.07	31.54	44.79	19.22	18.38	38.80	15.42	12.61	14.52	37.75	2810 (0%)
Affine	26.46	22.78	18.97	12.73	14.87	30.23	13.86	10.87	14.74	31.82	2810 (0%)
Dudani	32.22	18.91	46.76	16.62	17.86	34.98	15.54	12.61	15.19	38.78	2810 (0%)
Improved	29.64	18.61	25.52	12.47	12.22	29.32	13.74	10.00	14.07	32.65	2810 (0%)
MLSD	33.67	30.45	20.59	20.65	16.50	39.22	16.87	12.17	20.30	35.11	2810 (0%)
MLSD+ Improved mechanism ($\Delta_G = 0.2$)	**39.83**	**30.38**	**45.79**	**24.16**	**21.28**	**45.42**	**18.43**	**13.49**	**20.37**	**38.02**	**1250** (55.55%)

Table 5-2 shows a possible saving of 55.55% after using the improved mechanism. Our proposed MLSD with improved two-step mechanism can significantly reduce the computational complexity, while the retrieval accuracies are much better in almost all of the groups.

To evaluate a shape descriptor for MPEG-7, the descriptor is required to be invariant to rotation, scaling and translation and to be able to measure perceptual similarity of shapes. Other than the invariant properties, the proposed MLSD achieves the advantages of small memory for storing the descriptor, small computation for descriptor extraction and hierarchical represented with quadtree structure. Moreover, we applied the MLSD into a general application, i.e. trademark image retrieval. From the experimental results, it is shown that the retrieval system with considering both global and local characteristics can perform good retrieval accuracy and fast searching. Besides, no drastic performance degradation will happen with increasing size of the database.

5.5. Chapter Conclusion

We have presented a retrieval system with a new multi-layer shape descriptor (MLSD) invariant to rotation, scaling and translation and an improve two-step mechanism. From the elaboration of the best-fit ellipse concept, it is found that the first invariant measures the total spread of the shape relative to its area square while the second invariant measures the degree of elongation of a best-fit ellipse on the shape. On the other hand, MLSD describes the object by dividing it into four sub-regions by using two principal axes of the best-fit ellipse, which is located at the center of mass. Each sub-region is then sub-divided into four sub-regions in the same way. Then the node of a quadtree representation is used to represent the corresponding region of the object. Three kinds of parameters invariant to rotation, scaling and translation are calculated for each node. Besides, an improved two-step mechanism with considering both global and local characteristics of the object is proposed to speed up the matching process. From the experimental results, it is shown that MLSD can achieve better retrieval rate than other four moment-based approaches and the improved two-step mechanism not only further improves the retrieval accuracy but also avoids unnecessary image matching. It reduces computational complexity, resulting in an effective and efficient image retrieval algorithm.

Chapter 6

Conclusions

6.1. Conclusions of the Present Work

In this thesis, a practical important tool, the Moment Invariants, for pattern recognition has been studied. The moment invariants are the popular and useful image features for shape recognition. It gains the advantages of simple and fast calculations as well as the shape features extracted are invariant to rotation, scaling and translation of the objects. Moreover, several moment-based approaches (Hu moment invariants, Affine moment invariants, Dudani moment invariants, Improved moment invariants and Sluzek's moment-based approach), which are developed to extend the recognition ability for different kinds of objects, are

introduced.

In the initial part of this work, the basic theories and derivations of these five moment-based approaches are presented. We have investigated how the features extracted from these moments can perform the rotation, scaling and translation invariance.

In Chapter 3, we have discussed the interpretations of second and third-order based Hu moment invariants. From both theoretically and experimentally analysis, it is found that the first invariant measures the total spread of the shape relative to its area square while the second invariant measures the degree of elongation of a best-fit ellipse on the shape. Useful properties have also been found for objects with symmetric along $y = \pm x/\sqrt{3}$ and $y = \pm\sqrt{3}x$, when the third invariant is used. For the fourth invariant element, when fitting a best-fit ellipse on the object, the term of $a_f^2 + b_f^2$ is formed to be constant..

In Chapter 4, we have given an analysis of quantization error due to scaling and rotation for four moment-based approaches. They are Hu moment invariants, Affine moment invariants, Dudani moment invariants and Improved moment invariants. We have found that the error due to scaling increases as the object size decreases. The Dudani moment invariants tend to have a larger error than that of the other three moment-based approaches. Moreover, the error due to rotation increases as the object is rotated with an angle other than $\theta = 0°$, 90°and 180°.

In order to ensure that the shape descriptor is invariant to translation, scaling and rotation, we can apply the analysis of the quantization effects to scaling and rotation for selecting the threshold for an object searching system. From our experimental work, it is proved that all scaled and rotated objects can be retrieved correctly by setting the threshold to the value of the error.

In Chapter 5, we have presented a retrieval system with a new multi-layer shape descriptor (MLSD) invariant to rotation, scaling and translation and an improve two-step mechanism. The MLSD has been used to describe the object by dividing it into four sub-regions by two principal axes of the best-fit ellipse, which is located at the center of mass. Each sub-region is then sub-divided into four sub-regions in the same way. Then the node of a quadtree representation is used to represent the corresponding region of the object. Three kinds of parameters invariant to rotation, scaling and translation are calculated for each node. Besides, an improved two-step mechanism considering both global and local characteristics of the object is proposed to speed up the matching process.

From the experimental results, it is shown that the MLSD can achieve better retrieval rate than other four moment-based approaches and the improved two-step mechanism not only further improves the retrieval accuracy but also avoids the unnecessary image matching. This can reduce the computational complexity, resulting in an effective and efficient image retrieval algorithm.

6.2. Future Work

Based on the current work reported in the thesis, further work can be carried out in several directions. Let us elaborate some of these directions.

6.2.1. Emptiness Problem

Moment invariants are a kind of region-based shape descriptors. They cover both contour and internal details of an object. Let us consider two types of objects (filled objects and trademark objects) and investigate the classification ability of the moment invariants within these two types.

6.2.1.1. Filled objects

The characteristic of filled objects is that it is no need to consider the internal details of an object. We can evaluate the filled objects as a whole shape. Three irregular filled objects are used for this experiment (Fig. 6-1). For each type of objects, a set of four objects is generated by varying length a and width b, i.e. $a=ka_o$ and $b=kb_o$ where $k=0.5,\ 1,\ 2$ *and* 3. Moreover, another five objects, which were generated by varying the rotational angle, θ, at 0°, 45°, 90°, 135° and 180° are used. Table 6-1 and Table 6-2 show the exact results of Hu moment invariants with these two sets of objects. Table 6-1 tries to show the property of scaling invariance, while Table 6-2 tries to show the property of rotational invariance. Moreover, it is easily seen that these three objects can be distinguished themselves even only using second order based moment invariants, i.e. ϕ_1 and ϕ_2. Therefore, it indicates that calculating the objects by the second order based Hu moment invariants can achieve certain invariance properties, when the filled objects are being examined.

Figure 6-1 Three irregular filled objects

Table 6-1 Actual values of Hu moment invariants with different scaling factors

Figure 6-1	Rotational angle	ϕ_1	ϕ_2	ϕ_3	ϕ_4	$\phi_5(\times10^{-7})$	$\phi_6(\times10^{-5})$	$\phi_7(\times10^{-7})$
(a)	0.5	0.276996	0.010226	0.002186	0.000567	-5.6	-3.3	-2.8
	1	0.275193	0.009710	0.002214	0.000545	-5.2	-3.6	-3.0
	2	0.275221	0.009710	0.002214	0.000545	-5.2	-3.6	-3.0
	3	0.275226	0.009710	0.002214	0.000545	-5.2	-3.6	-3.0
(b)	0.5	0.551339	0.227978	0.033342	0.003605	-392.0	-171.2	53.6
	1	0.551518	0.228330	0.033608	0.003621	-399.0	-172.8	26.8
	2	0.551556	0.228330	0.033608	0.003621	-399.0	-172.8	26.86
	3	0.551563	0.228330	0.033608	0.003621	-399.0	-172.8	26.8
(c)	0.5	0.347064	0.036229	0.017482	0.002149	110.0	40.9	-72.6
	1	0.345212	0.035022	0.016652	0.001911	85.0	35.6	-66.4
	2	0.345238	0.035022	0.016652	0.001911	85.0	35.6	-66.4
	3	0.345243	0.035022	0.016652	0.001911	85.0	35.6	-66.4

Table 6-2 Actual values of Hu moment invariants with different rotational angles

Figure 6-1	Rotational angle	ϕ_1	ϕ_2	ϕ_3	ϕ_4	$\phi_5(\times10^{-7})$	$\phi_6(\times10^{-5})$	$\phi_7(\times10^{-7})$
(a)	0°	0.275221	0.009710	0.002214	0.000545	-5.2	-3.6	-3.0
	45°	0.273921	0.009524	0.002183	0.000530	-4.9	-3.5	-3.0
	90°	0.275032	0.009670	0.002226	0.000545	-5.2	-3.5	-3.0
	135°	0.274136	0.009612	0.002180	0.000521	-4.7	-3.5	-2.9
	180°	0.275032	0.009670	0.002226	0.000545	-5.2	-3.5	-3.0
(b)	0°	0.551556	0.228330	0.033608	0.003621	-399.0	-172.8	26.8
	45°	0.548230	0.225702	0.032723	0.003530	-377.4	-167.2	38.4
	90°	0.550786	0.227550	0.033533	0.003615	-397.0	-172.2	29.1
	135°	0.548428	0.225778	0.032798	0.003536	-378.8	-167.5	38.9
	180°	0.551272	0.228066	0.033586	0.003619	-398.2	-172.7	24.3
(c)	0°	0.345238	0.035022	0.016652	0.001911	85.0	35.6	-66.4
	45°	0.342763	0.034094	0.016271	0.001840	79.1	33.9	-62.3
	90°	0.344805	0.034819	0.016658	0.001939	86.3	36.0	-68.5
	135°	0.342482	0.033740	0.016167	0.001817	75.9	33.2	-62.6
	180°	0.344805	0.034819	0.016658	0.001939	86.3	36.0	-68.5

6.2.1.2. Trademark objects

There is a big difference between filled objects and trademark objects themselves. Trademark objects contain both the contour and the internal details. Let us consider four trademark objects (Fig. 6-2). They are similar in shape, but their values obtained from Hu moment invariants (Table 6-3) are very different. As shown in Table 6-3, we cannot categorize these objects into the same group especially Figs. 6-2a and 6- 2b. In view of the object itself, generally these two objects are the same, but the difference is on the width of edges. The width of edge of Fig. 6-2a is "thicker" than that of Fig. 6-2b. Therefore, it indicates that there is a limitation on moment invariants in trademark object identification.

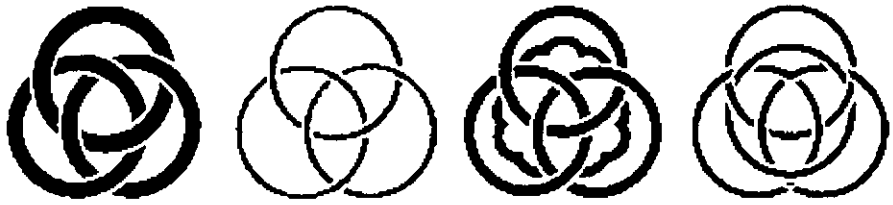


Figure 6-2 Four trademark objects

Table 6-3 Actual values of Hu moment invariants for Fig. 6-2(a-d)

Figure 6-2	ϕ_1	$\phi_2(\times 10^{-6})$	ϕ_3	$\phi_4(\times 10^{-7})$	$\phi_5(\times 10^{-10})$	$\phi_6(\times 10^{-10})$	$\phi_7(\times 10^{-10})$
(a)	0.336555	105.9	0.001073	1.6	0	-13.0	0
(b)	1.039531	379.6	0.034821	5897.4	-26564.0	113778.0	2927.0
(c)	0.400986	2.6	0.004155	65.1	-10.0	1.0	2.0
(d)	0.612406	1.7	0.010449	101.4	-8.0	-81.0	-32.0

In order to elaborate the problem clearly, a simple rectangular object (Fig. 6-3) with the width of edges w_1 and w_2 is used as an example to show the results of some preliminary theoretical and experimental studies.

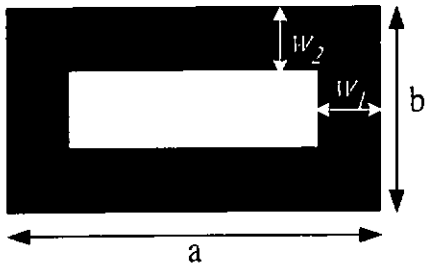


Figure 6-3 A rectangular object with width of edge w_1 and w_2

The second order based Hu moment invariants (ϕ_1 and ϕ_2) are calculated for both continuous integral and discrete summation cases. The required equations are shown as below.

Case 1: Continuous integral

$$\phi_1^{CR} = \frac{a^3b + ab^3 - (a-w_1)^3(b-w_2) - (a-w_1)(b-w_2)^3}{12(aw_2 + bw_1 - w_1w_2)^2} \tag{6.1}$$

$$\phi_2^{CR} = \frac{[a^3b - ab^3 - (a-w_1)^3(b-w_2) + (a-w_1)(b-w_2)^3]^2}{144(aw_2 + bw_1 - w_1w_2)^4} \tag{6.2}$$

Case 2: Discrete summation

$$\begin{aligned}\phi_1^{DR} = \Bigg[& \frac{2ab}{3}[(2a-1)(4a-1) + (2b-1)(4b-1)] - \\ & 2(aw_2 + bw_1 - w_1w_2)[2a(a-1) + 2b(b-1) + 1] + \\ & \frac{a-w_1}{3}[2(6b-1)(b-w_2) + 4w_2^2(w_2-3b) + 8b^2(3w_2-2b))] + \\ & \frac{b-w_2}{3}[2(6a-1)(a-w_1) + 4w_1^2(w_1-3a) + 8a^2(3w_1-2a))]\Bigg] \\ & \hline 16(aw_2 + bw_1 - w_1w_2)^2\end{aligned} \tag{6.3}$$

$$\begin{aligned}\phi_2^{DR} = \Bigg[& \frac{2ab}{3}[(2a-1)(4a-1) - (2b-1)(4b-1)] - \\ & 4(aw_2 + bw_1 - w_1w_2)[a(1-a) - b(1-b)] - \\ & \frac{a-w_1}{3}[2(6b-1)(b-w_2) + 4w_2^2(w_2-3b) + 8b^2(3w_2-2b))] + \\ & \frac{b-w_2}{3}[2(6a-1)(a-w_1) + 4w_1^2(w_1-3a) + 8a^2(3w_1-2a))]\Bigg] \\ & \hline [4(aw_2 + bw_1 - w_1w_2)]^4\end{aligned} \tag{6.4}$$

As shown in eqns. 6.1 to 6.4, the expressions in discrete summation case

requires a larger computation complexity as compared with that in continuous integral case. It also indicates that the influence of the widths w_1 and w_2 in the discrete summation is higher. Therefore, the second order based moment invariants cannot perform well when the objects contain emptiness part, like hole, and their widths of edge are difference.

In order to have a deeper analysis on the aspects that the object contains emptiness part, we further simplified the rectangular example by equalizing the w_1 and w_2 (i.e. $w = w_1 = w_2$) and did some re-calculation. The following two equations have been formulated.

Case 1: Continuous integral

$$\phi_1^{CR} = \frac{-1}{6} + \frac{(a-b)^2}{12w(a+b-w)} \tag{6.5}$$

$$\phi_2^{CR} = \left[\frac{a-b}{12w} - \frac{a-b}{12(a+b-w)} + \frac{ab(a-b)}{6w(a+b-w)^2}\right]^2 \tag{6.6}$$

Case 2: Discrete summation

$$\phi_1^{DR} = \frac{-1}{6} + \frac{(a-b)^2}{12w(a+b-w)} - \frac{a+b}{24w(a+b-w)^2} \tag{6.7}$$

$$\phi_2^{DR} = \left[\frac{a-b}{6w} - \frac{(a-b)(a+b)}{12w(a+b-w)} + \frac{ab(a-b)}{6w(a+b-w)^2}\right]^2 \tag{6.8}$$

In discrete summation case, it is seen that w presents in the denominators of ϕ_1^{DR} and ϕ_2^{DR} only. It indicates that if parameter w is large enough, the effect on different widths of edges in the object can be ignored. It is consistent with the findings from the examples of filled objects in the previous section. Therefore, the second order based moment invariants perform better in identifying the filled objects. On the other hands, if the parameter w is decreasing and lastly reaches 1, the problem of emptiness becomes significant.

It is because both ϕ_1^{DR} and ϕ_2^{DR} contain the parameter, w, in the denominators. So, the difference of the second order based moment invariants between objects with 15 pixels width of edge and 1 pixel width of edge is big. This idea is confirmed by the experimental results as shown in Table 6-3. Hence, we refer this as emptiness problem. This problem occurs when the objects contain holes and their widths of edges are different.

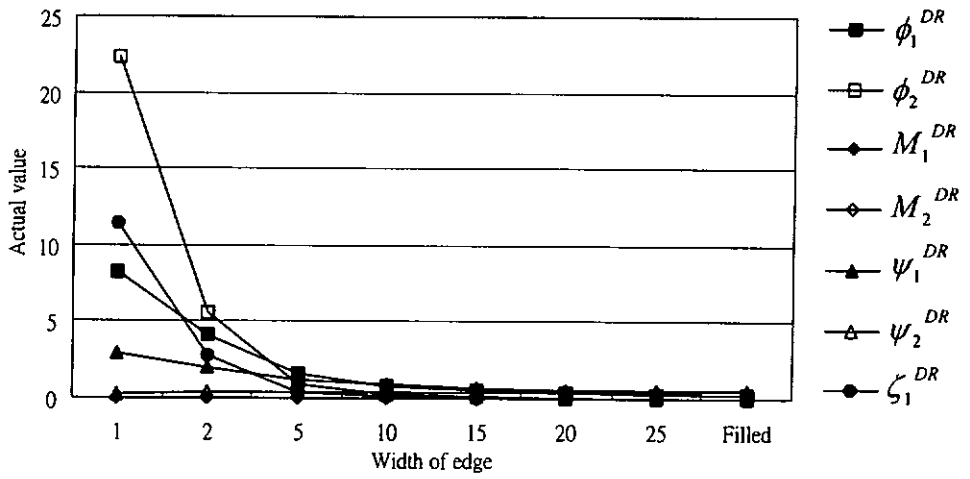


Figure 6-4 Actual values of four approaches while varying width of edge of rectangular objects

Moreover, this problem occurs not only in Hu moment invariants but also the other moment-based approaches, which have been discussed in the pervious section. Fig. 6-4 shows the actual values of four moment-based approaches in second order while varying width of edge of rectangular objects. Besides, since improved moment invariants aims to handle the contour objects, the edge detection is applied to two sets of elliptical and rectangular (same as in Fig. 6-4) objects with varying width of edge. It is also found that this approach fails to distinguish these two different groups of objects. From Fig. 6-5, it is shown that the actual values are overlapped. For the first element of improved moment invariants, the values for elliptical objects range 0.004859 to 0.019511

as the width of edge increases from 1 to filled, while the values for rectangular objects range 0.005223 to 0.020834. The second element gives similar result. That is the values for rectangular objects range 0.0842×10^{-4} to 1.3006×10^{-4}, while the values for rectangular objects range 0.0894×10^{-4} to 1.3758×10^{-4}. As a result, the emptiness problem happens on all four moment-based approaches and it cannot be resolved by simply using edge detection only.

Therefore, we are interested in applying the normalization technique on MLSD to solve this emptiness problem.

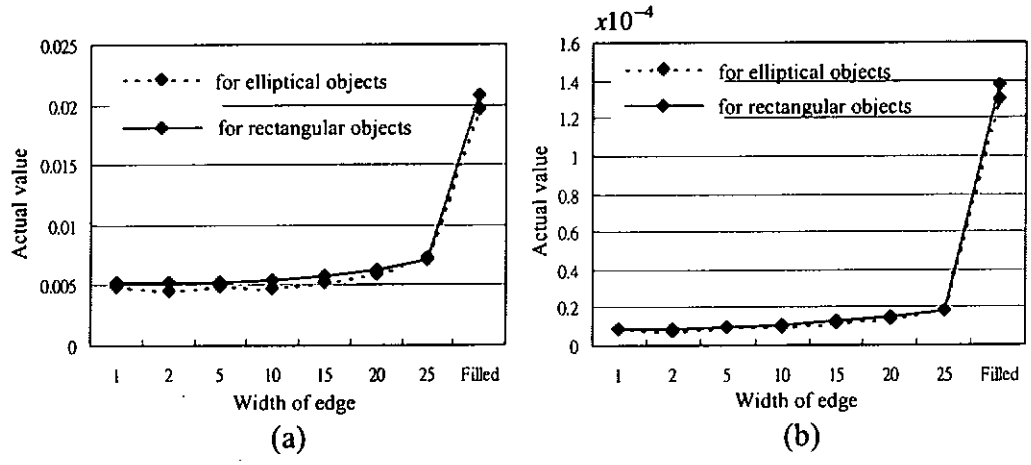


Figure 6-5 Actual values of the (a) First element and (b) Second element of Improved moment invariants on the elliptical and rectangular objects

6.2.2. Analysis on Moment-based Approaches with Different Deformations

In Chapter 4, we have presented the error analysis of four moment-based approaches under the scaling and rotational changing of the objects. The studies show how well four approaches can perform under the quantization error due to scaling and rotation. In fact, there are a number of deformations, such as cylinder-horizontal, cylinder-vertical, perspective-horizontal, perspective-vertical, pinch, punch and skew. Fig. 6-6 shows some sample objects after deformations.

We can carry out the experimental work to evaluate and realize the discriminative power of these approaches for the deformed objects.

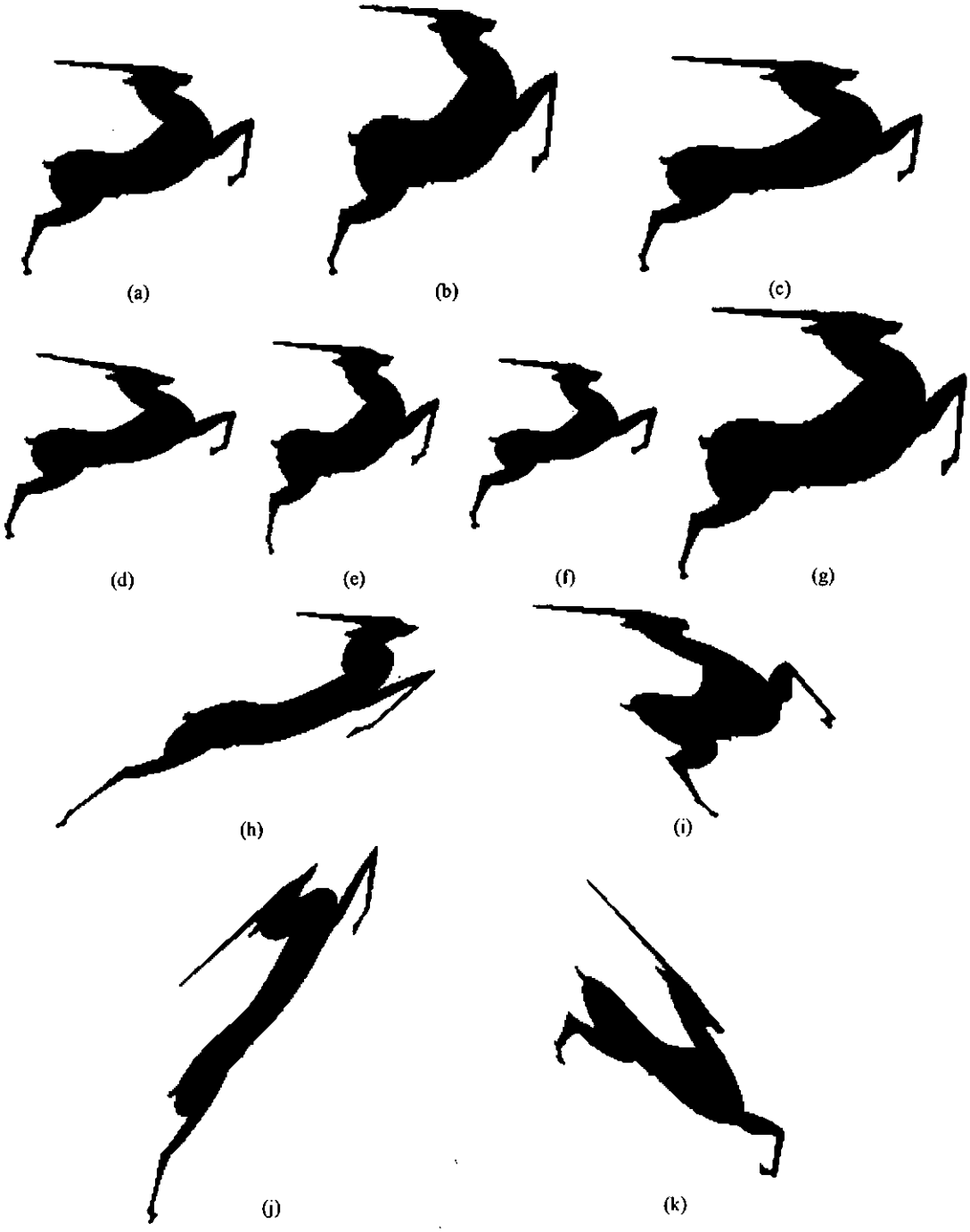


Figure 6-6 (a) Original object, objects after (b) Cylindar-horizontal 45%, (c) Cylinder-vertical 45%, (d) Perspective-horizontal 45%, (e) Perspective-vertical 45%, (f) Pinch 45, (g) Punch 45, (h) Skew-horizontal 45, (i) Skew-horizontal -45, (j) Skew-vertical 45, and (k) Skew-vertical -45

6.2.3. Feature Extraction on Color

Color is another key feature of images. It is a visual feature, which is immediately perceived when looking at an image.

Qien et al. [80] proposed a new method, called color blob histogram, for image indexing and retrieval that is based on pixel statistics from varying spatial scales. The blob histogram employs a structuring element to determine the frequency distribution of pixels locally in the image and to detect local groups of pixels with uniform color attributes. The frequency distribution and relative sizes of such groups are summarized into a table termed blob histograms. By embedding spatial information, color blob histograms are able to distinguish images that have the same color pixel distribution but contain objects with different sizes or shapes, without the need for segmentation. Using isotropic structuring elements, blob histograms are invariant to rotations and translations of the objects in an image. The conventional color histogram encodes the frequency distribution of single pixels with uniform color, while color blob histogram encodes relative size and frequency distribution of groups of pixels with uniform color.

The colors in the image are quantized into M different colors $c_0, c_1, c_2, .., c_{M-1}$. The procedure of calculating blob histogram includes: (i) to find the histogram at scale s which consists of multiple bins for each color c_m and (ii) to store each bin $hs(m,l)$ the number of locations at which the structuring element contains a certain percentage of pixels that are color c_m. The percentage axis of the blob histogram is quantized into L different values $r_0, r_1, r_2, .., r_{L-1}$. The blob histogram has MxL bins. The histogram is computed by visiting each location in the image, retrieving the colors of all pixels

contained in the structuring element *Es* overlaid on that location, computing the percentage of pixels of each color, and incrementing the histogram-bins *hs(m,l)* corresponding to the percentage of pixels with a certain color.

For example as shown in Fig.6-7: Eight different colors and three different fraction ranges (*M=8* and *L=3*). The structuring element is square and has a size of 9 by 9 pixels. At this location, the structuring element contains 14 pixels with color C1 (about 17% of the pixels), 24 pixels with color C3 (about 30%) and 43 pixels with color C7 (about 53%). In this case, the structuring element is counted three times. It represents a group of pixels of a different size in each time. The blob histogram *hs(m,l)* can be normalized by the number of structuring elements.

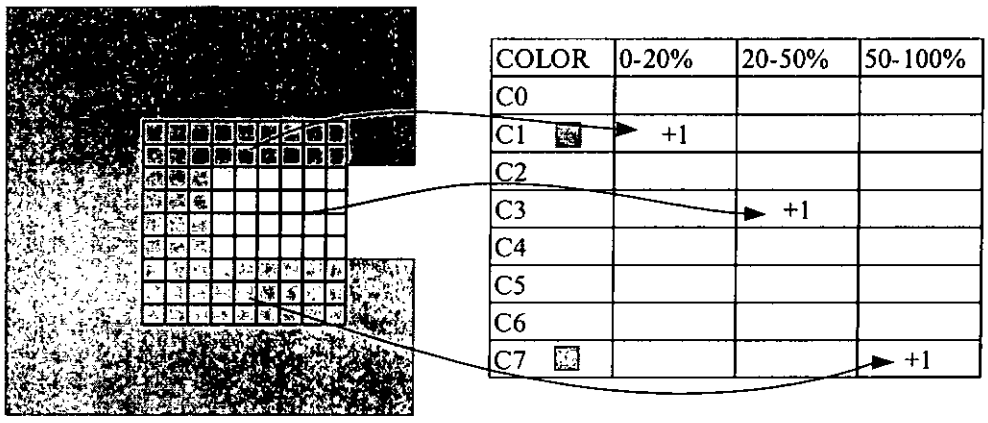


Figure 6-7 Extraction of color blob histogram

Note that the shape of a structuring element is not restricted. If we use isotropic structure elements, the computed blob histograms are invariant to rotations and translations of the objects in an image. Moreover, when *L* equals 1, the blob histogram becomes a conventional color histogram. Hence, our next suggestion of further work is to modify the structuring element by using different shaped or sized elements or to re-develop the blob histogram by using an even-distributed fraction ranges, *L*.

References

[1] H. Tamura, N. Yokoya, "Image Database Systems: A Survey," *Pattern Recognition*, 17, 1, 1984, pp. 29-43.

[2] S. K. Chang, A. Hsu, "Image Information Systems: Where do we go from here?," *IEEE Transactions on Knowledge and Data Engineering*, 4, 5, 1992, pp. 432-442.

[3] H. R. Turtle, W. B. Croft, "A Comparison of Text Retrieval Models," *The Computer Journal*, 35, 3, 1992, pp. 279-290.

[4] S. Sclaroff, L. Taycher, M. L. Cascia, "ImageRover: A Content-Based Image Browser for the World Wide Web," *Proceedings of IEEE Workshop on Content-based Access of Image and Video Libraries*, 6, 1997, pp. 2-9.

[5] D. Eichmann, "The RBSE Shpider- Balanching Effective Search against Web Load," *Proceedings of the 1st International Conference on the World Wide Web*, 1994, pp. 113-120.

[6] O. McBran, "Gencl and WWWW: Tools for taming the Web," *Proceedings of the 1st International Conference on the World Wide Web*, 1994.

[7] F. Rabitti, P. Stanchev, "An Approval to Image Retrieval from Large Image Databases," *ACM SIGIR Conf. in R & D Information Retrieval*, 1987, pp. 284-295.

[8] B. Huet, E. R. Hancock, "Shape Recognition from Large Image Libraries by Inexact Graph Matching," *Pattern Recognition Letters*, 20, 11-13, 1999, pp. 1259-1269.

[9] A. K. Jain and A. Vailaya, "Shape-based Retrieval: A Case Study with Trademark Image Databases," *Pattern Recognition*, 31, 9, 1998, pp. 1369-1390.

[10] Y. Rui, T. S. Huang, S. F. Chang, "Image Retrieval: Past, Present, And Future," *International Symposium on Multimedia Information Processing*, 1996.

[11] Y. Rui, T. S. Huang, S. F. Chang, "Image Retrieval: Current Techniques, Promising Directions, and Open Issues," *Journal of Visual Communication and Image Representation,* 10, 1, March 1999, pp. 39-62.

[12] M. Markkula, M. Tico, B. Sepponen, et al., "A Test Collection for the Evaluation of Content-Based Image Retrieval Algorithms – A User and Task-Based Approach," *Information Retrieval*, 4 (3/4), 2001, pp. 275-294.

[13] H. Muller, W. Muller, D. McG. Squire, et al., "Performance Evaluation in Content-Based Image Retrieval: Overview and Proposals," *Pattern Recognition Letters*, 22, 5, 2001, pp. 593-601.

[14] M. Koskela, J. Laaksonen, E. Oja, "Comparison of Techniques for Content-Based Image Retrieval," *Proceedings of Infotech Oulu International Conference on Information Retrieval, IR 2001*, 2001.

[15] J. Zhang, X. Zhang, H. Krim, "Invariant Object Recognition by Shape Space Analysis," *International Conference on Image Processing, ICIP 1998, Proceedings.*, 3, 3, 1998, pp.581-585.

[16] K. B. Eom, "Shape Recognition Using Spectral Features," *Pattern Recognition Letters*, 19, 2, 1998, pp. 189-195.

[17] G. Neil, K. M. Curits, "Shape Recognition Using Fractal Geometry," *Pattern Recognition*, 30, 12, 1997, pp. 1957-1969.

[18] The World Intellectual Property Organization (WIPO), http://www.wipo.org/.

[19] Intellectual Property Department, The Government of the Hong Kong Special Administrative Region, http://www.info.gov.hk/ipd/eng/index.htm.

[20] The Intellectual Property Services Centre (IPSC), The Hong Kong Productivity Council (HKPC), http://www.ipsc.org.hk/engindex.html.

[21] J. L. Shih, L. H. Chen, "A New System for Trademark Segmentation and Retrieval," *Image and Vision Computing*, 19, 13, 1, 2001, pp. 1011-1018.

[22] S. Loncaric, "A Survey of Shape Analysis Techniques," *Pattern Recognition*, 31, 8, 1, 1998, pp. 983-1001.

[23] R. C. Veltkamp, "Shape Matching: Similarity Measures and Algorithms," *International Conference on Shape Modeling and Applications, SMI 2001*, 2001, pp. 188-197.

[24] R. Singh, N. P. Papanikolopoulos, "Planar Shape Recognition by Shape Morphing," *Pattern Recognition*, 33, 10, 2000, pp. 1683-1699.

[25] W. H. Lin, J. S Lee, C. H. Chen, Y. N. Sun, "A New Multiscale-Based Shape Recognition Method," *Signal Processing*, 65, 1, 1998, pp. 103-113.

[26] T. Özuğur, Y Denizhan, E. Panayirci, "Feature Extraction in Shape Recognition Using Segmentation of the Boundary Curve," *Pattern Recognition Letters*, 18, 10, 1997, 1049-1056.

[27] E. E. Stannard, D. Pycock, "Recognizing 2-D Shapes from Incomplete Boundaries," *IEE Colloquium on Applied Statistics pattern Recognition*, 1999/063, 1999, pp.12/1-12/6.

[28] K. A. Han, J. C. Lee, C. J. Hwang, "Image Clustering Using Self-Organizing Feature Map with Refinement," *IEEE International Conference on Neural Networks*, 1, 1995, pp. 465-469.

[29] A. K. Jain, A. Vailaya, "Image Retrieval Using Color and Shape," *Pattern Recognition*, 29, 8, 1996, pp. 1233-1244.

[30] A. Vailaya, Y. Zhong, A. K. Jain, "A Hierarchical System for Efficient Image Retrieval," *The 13th International Conference on Pattern Recognition*, 1996, pp. 356-360.

[31] B. M. Mehtre, M. S. Kankanhalli, W. F. Lee, "Shape Measures for Content Based Image Retrieval: A Comparison," *Information Processing & Management*, 33, 3, 1997, pp. 319-337.

[32] M. Safar, C. Shahabi, X. Sun, "Image Retrieval by Shape: A Comparative Study," *IEEE International Conference on Multimedia and Expo, ICME 2000*, 1, 2000, pp. 141-144.

[33] G. Bachman, L. Narici, et al., *Fourier and Wavelet Analysis*, New York, Hong Kong: Springer, 2000.

[34] H. Kauppinen, T. Seppanen, M. Pietikainen, "An Experimental Comparison of Autoregressive and Fourier-based Descriptors in 2-D Shape Classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17, 2, 1995, pp. 201-207.

[35] L. F. Estrozi, A. G. Campos, et al., "Comparing Curvature Estimation Techniques," *Sba Controle Automao*, 00, 00 / Jan., Fev., Mar, Abril de 0000.

[36] C. K. Chui, *An Introduction to Wavelets*, San Diego, CA: Academic Press, 1992.

[37] J. D. Lee, "A New Scheme for Planar Shape Recognition Using Wavelets," *Computers & Mathematics with Application*, 39, 5-6. 2000, pp. 201-216.

[38] K. C. Hung, "The Generalized Uniqueness Wavelet Descriptor for Planar Closed Curves," *IEEE Transactions on Image Processing*, 9, 5, 2000, pp. 834-845.

[39] M. I. Khalil, M. M. Bayoumi, "A Dyadic Wavelet Affine Invariant Function for 2D Shape Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23, 10, 2001, pp. 1152-1164.

[40] M. K. Hu, "Visual Pattern Recognition by Moment Invariants," *IRE Trans Inform Theory*, IT-8, 1962, pp. 179-187.

[41] W. H. Wong, W. C. Siu and K. M. Lam, "Generation of Moment Invariants and their Uses of Character Recognition," *Pattern Recognition Letters*, 16, 1995, pp. 115-123.

[42] J. Flusser, T. Suk, "A Moment-based Approach to Registration of Images with Affine Geometric Distortion," *IEEE Trans. on Geoscience and Remote Sensing*, 32, 2, March 1994, pp. 382-387.

[43] J. Flusser, T. Suk, "Pattern Recognition by Affine Moment Invariants," *Pattern Recognition*, 26, 1, 1993, pp. 167-174.

[44] S. A. Dudani, K. J. Breeding, R. B. Mcghe, "Aircraft Identification by Moment Invariants," *IEEE Trans. on Computers*, C-26, 1, Jan. 1977, pp. 39-45.

[45] A. Khotanzad, Y. H. Hong, "Invariant Image Recognition by Zernike Moments," *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 12(5), 1990, pp. 489-497.

[46] M. Zhenjiang, "Zernike Moment-Based Image Shape Analysis and its Application," *Pattern Recognition Letters*, 21, 2, 2000, pp. 169-177.

[47] W. Y. Kim, Y. S. Kim, "A Region-Based Shape Descriptor Using Zernike Moments," *Signal Processing: Image Communication*, 16, 1-2, 2000, pp. 95-102.

[48] K. R. Castleman, *Digital Image Processing*, Prentice Hall Inc., Englewood Cliffs, N.J., 1996.

[49] M. Sonka, V. Hlavac, R. Boyle, *Image Processing, Analysis and Machine Vision*, Pacific Grove, CA: PWS Pub., 1999.

[50] H. Freeman, A. Saghar, "Generalized Chain Codes for Planar Curves," *Proceedings of the 4th International Joint Conference on Pattern Recognition*, 1978, pp. 701-703.

[51] L. Gupta, M. D. Srinath, "Contour Sequence Moments for the Classification of Closed Planar Shapes," *Pattern Recognition*, 20, 3, 1987, pp. 267-272.

[52] J. Livarinen, A. Visa, "Shape Recognition of Irregular Objects," *In D. P. Casasent (Ed.), Intelligent Robots and Computer Vision XV: Algorithms, Techniques, Active Vision, and Materials Handling, Proc. SPIE 2904*, 1996, pp. 25-32.

[53] D. McG. Squire, "Generalization Performance of Factor Analysis Techniques Used for Image Database Organization" Technical Report, Vision, Centre Universitaire D' Infromatique Groupe Vision, Universite De Geneve, 98.01, January 30, 1998.

[54] J. Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8, 6, 1986, pp. 679-698.

[55] K. K. Pingle, "Visual Perception by Computer," *Automatic Interpretation and Classification of Images*, Academic Press, New York, 1969, pp. 277-284.

[56] D. Marr, E. Hildreth, "Theory of Edge Detection," *Proceedings of Royal Society London Series B*, 207, 1980, pp. 187-217.

[57] C. C. Chen, "Improved Moment Invariants for Shape Discrimination," *Pattern Recognition*, 26, 5, 1993, pp. 683-686.

[58] J. P. Eakins, J. M. Boardman, M. E. Graham, "Similarity Retrieval of Trademark Images," *IEEE Multimedia*, 5, 2, 1998, pp. 53-63.

[59] G. Ciocca, R. Schettini, "Similarity Retrieval of Trademark Images," *Proceedings in International Conference on Image Analysis and Processing*, 1999, pp. 915-920.

[60] S. Alwis, J. Austin, "Searching Image Database Containing Trademarks," *IEE Colloquium on Neural Networks in Interactive Multimedia Systems*, 1998/446, 1998, pp. 2/1-2/5.

[61] S. Ravela, R. Manmatha, "On Computing Global Similarity in Images," *Proceedings in 4th IEEE Workshop on Applications of Computer Vision, WACV' 98*, 1998, pp. 82-87.

[62] J. Neurnann, H. Samet, A. Soffer, "Integration of Local and Global Shape Analysis for Logo Classification," *Pattern Recognition Letters*, 23, 12, 2002, pp. 1449-1457.

[63] Y. S. Kim, W. Y. Kim, "Content-Based Trademark Retrieval System Using Visually Salient Features," *Proceedings on IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997, pp. 307-312.

[64] F. Nabhani, T. Shaw, "Performance Analysis and Optimization of Shape Recognition and Classification using ANN," *Robotics and Computer-Integrated Manufacturing*, 18, 3, 2002, pp. 177-185.

[65] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, et al., "Query by Image and Video Content: The QBIC System," *IEEE Computer*, 28, 1995, pp. 23-32.

[66] W. Niblack, R. Barber, W. Equitz, et al., "The QBIC Project: Querying Images by Content Using Color, Texture, and Shape," *SPIE 1908, Storage and Retrieval for Image and Video Databases*, 1993, pp. 173-187.

[67] J. P. Eakins, K. Shields, J. Boardman, "ARTISAN A Shape Retrieval System Based On Boundary Family Indexing," *SPIE 2670, Storage and Retrieval for Still Image and Video Databases IV. Proceedings*, 1996.

[68] J. P. Eakins, J. M. Boardman, K. Shields, "Retrieval of Trademark Images by Shape Feature the ARTISAN Project," *IEE Colloquium on Intelligent Image Databases*, 1996, pp. 9/1-9/6.

[69] A. Soffer, H. Samet, "Negative Shape Features for Image Database Consisting of Geographic Symbols," *The 3rd International Workshop in Visual Form Capri.*, 1997, pp. 1-13.

[70] A. Soffer, H. Samet, "Using Negative Shape Features for Logo Similarity Matching," *The 4th International Conference on Pattern Recognition*, 1, 1998, pp. 571-573.

[71] A. M. Tekalp, *Digital Image Processing*, Chapter 9: Methods Using Point Correspondences, Prentice Hall PTR, Prentice Hall Inc., Upper Saddle River, N.J. 07458, 1996.

[72] A. Sluzek, "Identification and Inspection of 2-D Objects Using New Moment-Based Shape Descriptors," *Pattern Recognition Letters*, 16, 1995, pp. 687-697.

[73] L. A. Torres-Mendez, J. C. Ruiz-Suarez, L. E. Sucar, G. Gomez, "Translation, Rotation, Scale-Invariant Object Recognition," *IEEE Transaction on Systems, Man. and Cybernetics – Part C: Application and Review*, 30, 1, 2000, pp. 125-130.

[74] Jia Guu Leu, "Computing A Shape's Moments From Its Boundary," *Pattern Recognition*, 27, 10, 1991, pp. 949-957.

[75] T. M. Hupkens, J. DE Clippeleir, "Noise and Intensity Invariant Moments," *Pattern Recognition Letters*, 1995, pp. 371-376.

[76] C. H. Teh, R. T. Chin, "On Digital Approximation of Moment Invariants," *Computer Vision, Graphics, and Image Processing*, 33, 1986, pp. 318-326.

[77] G. I. Salama, A. L. Abbott, "Moment Invariants and Quantizaiton Effects," *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, 1998, pp. 157-163.

[78] S. Jeannin, MPEG-7 Visual Part of eXperimentation Model Version 1.0, Doc. ISO/IEC JTC1/SC29/WG11 N2695, 47th MPEG Meeting, Seoul, March 1999.

[79] H. K. Kim, J. D. Kim, "Region-based shape descriptor invariant to rotation, scale and translation," *Signal Processing: Image Communication*, 16, 2000, pp. 87- 93.

[80] Richard J. Qian, Peter J. L. van Beek and M. Ibrahim Sezan, "Image Retrieval Using Blob Histograms," *IEEE International Conference on Multimedia and Expo, ICME 2000*, 1, 2000, pp. 125 -128.