Performance Study of an Improved Legendre Moment Descriptor as Region-Based Shape Descriptor[1]

V. P. Dinesh Kumar and T. Tessamma

Department of Electronics, Cochin University of Science and Technology, Kerala, Cochin, 682 022 INDIA

e-mail: dineshvp@cusat.ac.in, tess@cusat.ac.in

Abstract—This paper reports a novel region-based shape descriptor based on orthogonal Legendre moments. The preprocessing steps for invariance improvement of the proposed Improved Legendre Moment Descriptor (ILMD) are discussed. The performance of the ILMD is compared to the MPEG-7 approved region shape descriptor, angular radial transformation descriptor (ARTD), and the widely used Zernike moment descriptor (ZMD). Set B of the MPEG-7 CE-1 contour database and all the datasets of the MPEG-7 CE-2 region database were used for experimental validation. The average normalized modified retrieval rate (ANMRR) and precision-recall pair were employed for benchmarking the performance of the candidate descriptors. The ILMD has lower ANMRR values than ARTD for most of the datasets, and ARTD has a lower value compared to ZMD. This indicates that overall performance of the ILMD is better than that of ARTD and ZMD. This result is confirmed by the precision-recall test where ILMD was found to have better precision rates for most of the datasets tested. Besides retrieval accuracy, ILMD is more compact than ARTD and ZMD. The descriptor proposed is useful as a generic shape descriptor for content-based image retrieval (CBIR) applications.

Key words: CBIR, MPEG-7, Zernike Moments, Legendre Moments, Angular Radial Transformation, region-based shape descriptor.

DOI: 10.1134/S1054661808010045

1. INTRODUCTION

Content-based image retrieval (CBIR) has emerged as an important area in computer vision and multimedia computing. Extending traditional databases to include pictorial information creates many new challenges and applications for database systems. As the demand for digital images increases, the need to store and retrieve images in an intuitive and efficient manner arises. CBIR focuses on intuitive and efficient methods for retrieving images from databases solely based on the content contained in the images.

Object shape features provide a powerful clue to object identity and functionality and can be used for object recognition [1]. A good descriptor captures characteristic shape features in a concise manner, and it should be invariant to scaling, rotation, translation and to various types of shape distortions. Loncaric [2] presents a comprehensive review of the methods reported for shape analysis and image retrieval. For images there are two notions of similarity: region-based and contour-based similarity. In region-based techniques, descriptors are derived using all the pixel information within a shape region, whereas contour-based descriptors express the shape properties of the object outline. MPEG-7 supports both notions of similarity, and the Angular Radial Transform Descriptor (ARTD) [3] and curvature scale-space Descriptor (CSSD) [4] have been adopted by MPEG-7 as the region-based shape descriptor and contour-based shape descriptor, respectively [5, 6]. A recent study conducted by Dengsheng and Guojun [7] reports that the Zernike Moment Descriptor (ZMD) has superior performance as both a region-based and a contour-based shape descriptor. The Angular Radial Transform (ART) also belongs to the broad Zernike moment family [1]. It is shown that orthogonal Legendre Moments (LM) have better image representation capability with fewer coefficients (low order), compared to Zernike Moments based on the reconstruction error calculation [8, 9]. In addition, Legendre Moments are computationally less complex than Zernike Moments. However, there is not much research reported in the literature on the usage of LM in image retrieval, except for the work proposed by Mandal et al., in which a feature vector is formed using LM and Wavelet Transform coefficients [10]. This could be due to their lack of inherent invariance to shape transformations. Dinesh et al. have demonstrated that the Improved Legendre Moment Descriptor (ILMD) can be derived from orthogonal Legendre moments with some preprocessing steps, which improve invariance to geometric transformations [11]. The accuracy of ILMD as a contour-based shape descriptor is established in that report. The representation accuracy and invariance properties of ILMD as a region-shape descriptor are compared with ZMD and ARTD in this study.

[1] The text was submitted by the authors in English.

Received May 5, 2007

2. ORTHOGONAL MOMENT-BASED SHAPE DESCRIPTORS

The orthogonal moment, which is the projection of an image function onto an orthogonal polynomial, provides compact image representation due to the lack of information redundancy. They have been utilized as pattern features in a number of applications to achieve invariant recognition of two-dimensional image patterns [12]. The moments are used as a shape descriptor by selecting an appropriate order sufficient to accurately represent the shape and by computing their invariants to various shape transformations. The reconstruction of the original image from the descriptors is not expected.

2.1. Zernike Moment Descriptor (ZMD)

The ZMD is widely used as a region-based shape descriptor [7]. The complex Zernike moments [13] for discrete image intensity function $F(x, y)$ of order n with repetition m are defined as

$$A_{nm} = \frac{n+1}{\pi}\sum_x\sum_y F(x, y)V^*_{nl}(x, y), \quad x^2 + y^2 \le 1, \tag{1}$$

where $n = 0, 1, 2, \ldots, \infty$ and m takes on positive and negative integer values subject to the conditions $n - |m| =$ even, $|m| \le n$. The symbol * denotes the complex conjugate. The Zernike polynomials

$$\begin{aligned} V_{nm}(x, y) &= V_{nm}(\rho\cos\theta, \rho\sin\theta) \\ &= R_{nm}(\rho)\exp(jm\theta), \end{aligned} \tag{2}$$

where ρ is the length of the vector from the origin to (x, y) pixel and θ is the angle between vector ρ and x-axis in the counterclockwise direction. The Zernike polynomials are a complete set of complex-valued functions orthogonal on the unit disc $(x^2 + y^2) \le 1$. The radial polynomial $R_{nm}(\rho)$ is defined as

$$R_{nl}(\rho) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \times \frac{(n-s)!}{s!\left(\frac{n+|m|}{2}-s\right)!\left(\frac{n-|m|}{2}-s\right)!}\rho^{n-2s}. \tag{3}$$

The Zernike moments are computed with the centroid chosen as the origin of the unit disc and the distance from the origin to the farthest pixel as the Zernike basis function radius for geometric invariance. The first 36 moments of order 0…10 are computed and normalized with the mass of the image [14].

2.2. Angular Radial Transformation Descriptor (ARTD)

The 2-D Angular Radial Transform (ART) is the MPEG-7 proposed region-based shape descriptor [1]. It belongs to the broad Zernike Moment family and provides a compact and efficient way to express pixel distribution within a 2-D object region. The set of the orthogonal moment basis is defined on a unit disc in polar coordinates (ρ, θ). From each shape, a set of ART coefficients F_{nm} is extracted as follows:

$$\begin{aligned} F_{nm} &= \langle V_{nm}(\rho, \theta), f(\rho, \theta)\rangle \\ &= \int_0^{2\pi}\int_0^1 V^*_{nm}(\rho, \theta)f(\rho, \theta)\rho\, d\rho\, d\theta, \end{aligned} \tag{4}$$

where $f(\rho, \theta)$ is an image intensity function in polar coordinates and $V_{nm}(\rho, \theta)$ is the ART basis function of order n and m. The basis functions are separable along the angular and radial directions and are defined as follows:

$$V_{nm}(\rho, \theta) = \frac{1}{2\pi}\exp(jm\theta)R_n(\rho), \tag{5}$$

where $R_n(\rho) = \begin{cases} 1, & n = 0, \\ 2 \quad \cos(\pi n\rho), & n \ne 0. \end{cases}$

As the ART and Zernike moments belong to the same family, they have similar transformation invariance properties. The magnitude of the complex ART coefficients normalized using the image mass is used as the shape descriptor. Since the ART basis function takes the unit disk as their domain, all shapes are normalized into a unit circle of fixed radius of 128 pixels. The first 35 moment coefficients of order $n = 10$ and $m = 10$ are used as the ART shape descriptor [1].

2.3. Improved Legendre Moment Descriptor (ILMD)

The Legendre moments are based on Legendre polynomials, which form an orthogonal basis set in the interval [–1, 1] [16]. The pth order Legendre polynomial is defined as follows:

$$\begin{aligned} P_p(x) &= \frac{1}{2^p p!}\frac{d^p}{dx^p}(x^2-1)^p, \\ x \in [-1, 1] &= \sum_{j=0}^{p} a_{pj}x^j. \end{aligned} \tag{6}$$

The values of a_{pj} can be expressed as

$$a_{pj} = (-1)^{(p-j)/2}\frac{(p+j)!}{2^p((p-j)/2)!((p+j)/2)!j!}, \tag{7}$$

where $p - j =$ even. For orthogonality to exist in the moments, the image function is defined over the same interval as the basis set. The Legendre moment of order $(p + q)$ for a discrete image function $F(x, y)$ is defined as

$$\lambda_{pq} = \frac{(2p+1)(2q+1)}{4}\sum_x\sum_y P_p(x)P_q(y)F(x, y), \quad (8)$$

where $p, q \varepsilon Z$.

From the experimental studies [11], it was found that the LM descriptor did not provide invariance to geometric transformations. In order to provide LMD invariance to geometric transformations and shape distortions, some preprocessing steps were applied before the computation of Legendre moments to obtain the Improved Legendre Moment Descriptor (ILMD) [11]. The steps are the following:

• Compute the major axis of the shape and rotate the image so as to align the major axis parallel to the x-axis.

• Determine the bounding rectangle (minimum sized rectangle to contain the shape) of the image.

• Map the pixel coordinates in the bounding rectangle to the range [–1, 1].

• Compute the Legendre Moment coefficients of order 1 to 5.

• Normalize the coefficients using the zeroth order geometric moment.

Use the absolute value of the Legendre Moment coefficients as the ILMD shape descriptor.

It was experimentally found that the first 20 coefficients provide accurate representation of the shape, and, hence, the ILMD was computed of order 1 to 5.

3. EXPERIMENTAL RESULTS AND DISCUSSION

The region-based shape descriptors can be applied to both the contour shape and the region shape database. Set B of the MPEG-7 contour shape database and CE-2 region shape database were used for experimental validation. CE-2 has been organized by MPEG-7 into six datasets, e.g., Set A1, A2, A3, A4, B, and the whole database. CE-2 is designed to test the region shape descriptor's behavior under different shape variations. The details of each dataset used are given below.

• Set B of CE-1 has 1400 shapes, which have been classified into 70 classes. Each class has 20 similar member shapes. Set B is for testing of similarity-based retrieval or for testing shape descriptors' robustness to various arbitrary shape distortions, including rotation, scaling, arbitrary skew and stretching, defection, indentation, and other variations.

• Set A1 of CE–2 consists of 2881 shapes from the whole database; it is for the test of scale invariance. The 100 shapes in Set A1 are organized into 20 groups (5 similar shapes in each group). In our experiment, all 100 shapes from the 20 groups are used as queries to test the retrieval.

• Set A2 of CE-2 consists of 2921 shapes from the whole database; it is for the test of rotation invariance. The 140 shapes in Set A2 are organized into 20 groups (7 similar shapes in each group). In our experiment, all 140 shapes from the 20 groups are used as queries to test the retrieval.

• Set A3 of CE-2 consists of 3101 shapes from the whole database; it is for the test of rotation/scale invariance. The 330 shapes in Set A3 are organized into 30 groups (11 similar shapes in each group). In our experiment, all 330 shapes from the 30 groups are used as queries to test the retrieval.

• Set A4 of CE-2 consists of 3101 shapes from the whole database; it is for the test of robustness to perspective transform. The 330 shapes in Set A4 are organized into 30 groups (11 similar shapes in each group). In our experiment, all 330 shapes from the 30 groups are used as queries to test the retrieval.

• Set B of CE-2 consists of 2811 shapes from the whole database; it is for the subjective test. The 682 shapes in Set B are manually sorted out into 10 classes by MPEG-7. In our experiment, all 682 shapes from the 10 classes are used as queries to test the retrieval.

• The whole database of CE-2 consists of 3621 shapes, 651 shapes of the 3621 shapes are organized into 31 groups (21 similar shapes in each group). For the 21 similar shapes in each group, there are 10 perspective transformed shapes, 5 rotated shapes, and 5 scaled shapes. The 31 groups of shapes reflect overall shape operations, and they test the overall robustness of a shape descriptor. In our experiment, all 651 shapes from 31 groups are used as queries to test the retrieval.

Each shape in the individual dataset of the two databases is indexed using the three described region shape descriptors. A good retrieval result in response to a visual feature-based query would be good indicator for the expressiveness of the descriptor. In the experiments, the so-called query-by-example (QBE) paradigm has been employed. In QBE, the respective descriptor values are extracted from a query image and then matched to the corresponding descriptors of images contained in the database. The distance (or dissimilarity) between

ANMRR Results

Database	ZMD	ARTD	ILMD
CE1–B	0.4489	0.4128	**0.3431**
CE2–A1	0.1585	0.1364	**0.0504**
CE2–A2	**0.0154**	**0.0154**	0.0193
CE2–A3	0.0019	**0.0003**	0.0011
CE2–A4	**0.1060**	0.1733	0.1119
CE2–B	0.3867	0.3756	**0.3621**
CE2–whole	0.2314	0.2239	**0.2172**

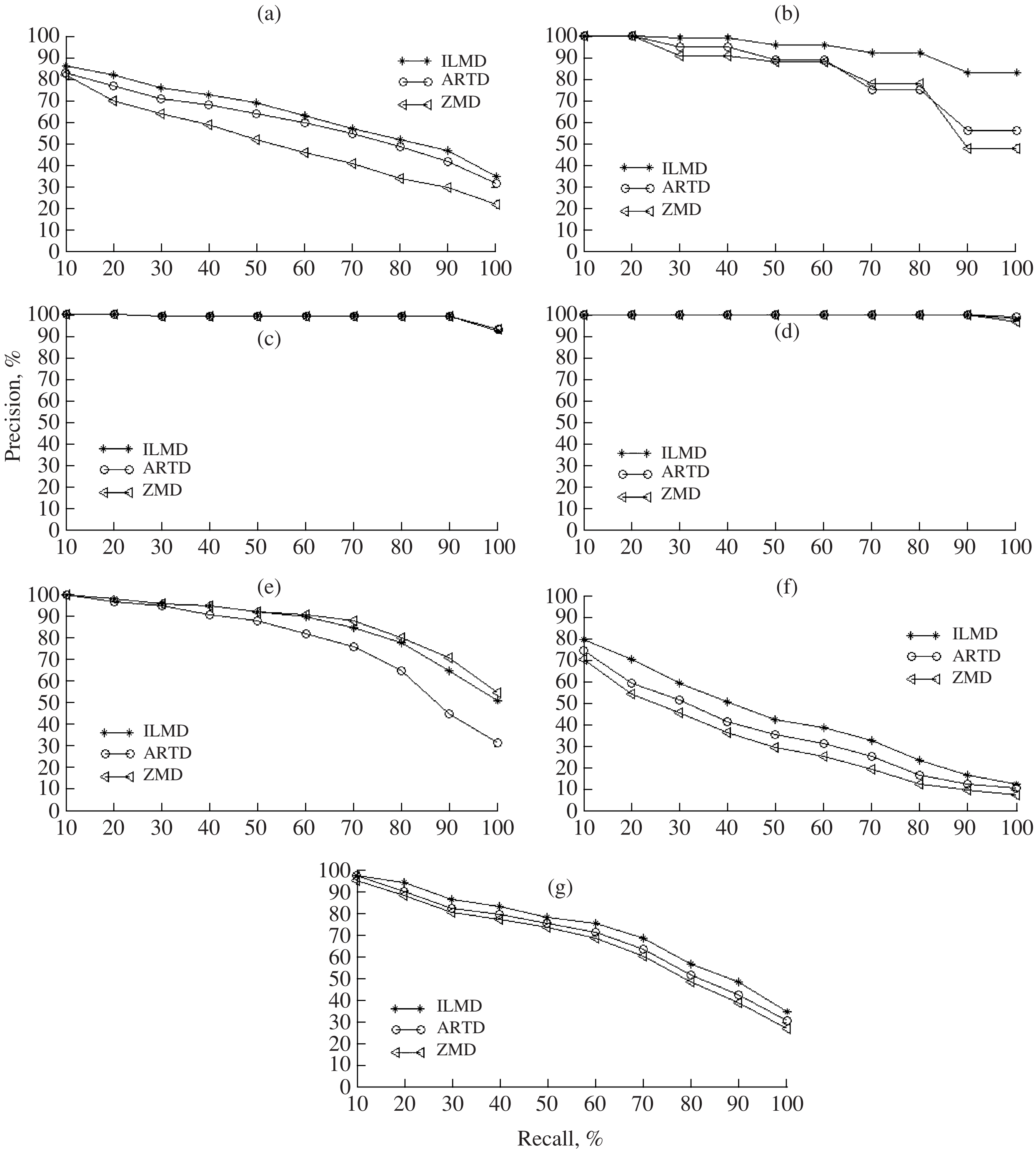

Fig. 1. Average precision-recall of the three region shape descriptors on the MPEG-7 shape database. (a) Average precision-recall of 1400 retrievals using three region shape descriptors on Set B of the MPEG-7 contour shape database CE-1. (b) Average precision-recall of 100 retrievals using three region shape descriptors on Set A1 of the MPEG-7 region shape database CE-2. (c) Average precision-recall of 140 retrievals using three region shape descriptors on Set A2 of the MPEG-7 region shape database CE-2. (d) Average precision-recall of 330 retrievals using three region shape descriptors on Set A3 of the MPEG-7 region shape database CE-2. (e) Average precision-recall of 330 retrievals using three region shape descriptors on Set A4 of the MPEG-7 region shape database CE-2. (f) Average precision-recall of 682 retrievals using three region shape descriptors on Set B of the MPEG-7 region shape database CE-2. (g) Average precision-recall of 651 retrievals using three region shape descriptors on the MPEG-7 region shape database CE-2.

two shapes described by the descriptor is calculated using an L-2 norm. The commonly employed Precision-Recall pair measurement was used to evaluate the descriptor performance [15]. Precision P, a measure of accuracy, is defined as the ratio of the number of retrieved relevant shapes r to the total number of retrieved shapes n; i.e., $P = r/n$. Recall that R is a measure of robustness, which is defined as the ratio of the

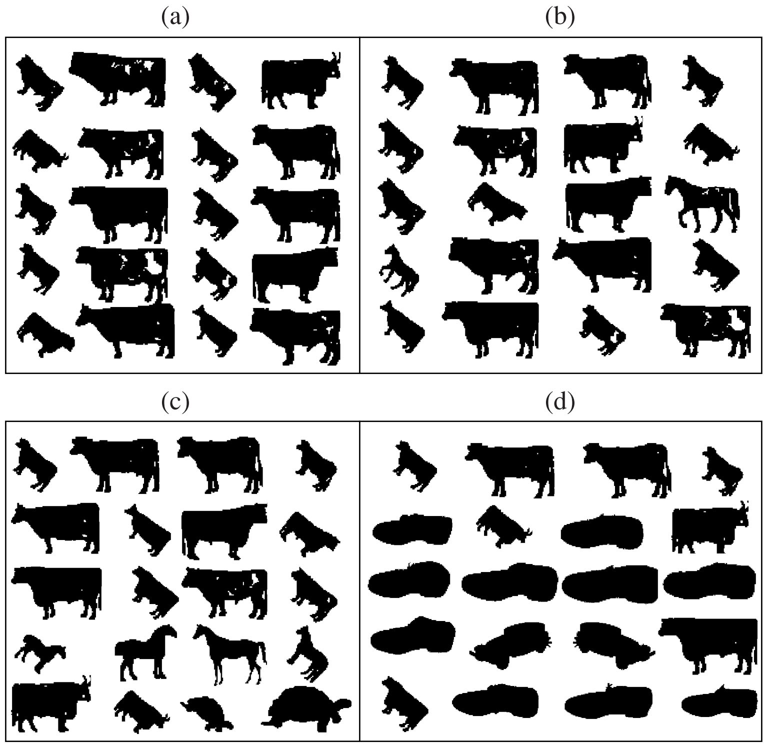

Fig. 2. An example result of the Query-By-Example test using Set B of the MPEG-7 contour shape database CE-1. (a) Ground truth set of "cattle" class. Retrieval results of (b) ILMD, (c) ARTD, and (d) ZMD. Left—top image of set (b–d) is the query image.

number of retrieved relevant images r to the total number m of relevant shapes in the whole database; i.e., $R = r/m$. In order to be objective in the comparisons, a quantitative measure called the Average Normalized Modified Retrieval Rate (ANMRR) [16] is also employed.

The ANMRR results are detailed in the table. For the CE-1 Set B database, the ILMD has a comparatively lower score than ARTD, and ARTD has a lower score than ZMD. For the CE-2 Set A2 (rotation invariance), ARTD and ZMD have equal performance and are marginally better than ILMD. For the CE2 Set A4, ZMD was slightly better than ILMD and ILMD was better than ARTD. The subjective test (CE2—Set B) and whole database showed better accuracy of ILMD, and ARTD was better than ZMD. This shows the overall superior performance of ILMD over ARTD and ZMD.

From the analysis of precision charts in Fig. 1, it can be seen that the ILMD outperforms ARTD and ZMD for most of the datasets tested. The only dataset for which ZMD was better than ILMD was SetA4 of CE-2, for perspective transformations (Fig. 1e). Furthermore, the performance of ARTD was better than ZMD in most of the cases. Overall, ILMD showed superior accuracy in comparison with ARTD and ZMD.

Figure 2 lists an example of QBE retrieval results using the cattle image of CE-1 Set B. Figure 2a shows the ground-truth data of the class. The ground-truth images not only differ by geometric transformations, but also of different types of cattle and with varying body shades. Figures 2b–2d show the first 20 retrieval results of ILMD, ARTD, and ZMD, respectively, in increasing order of distance, and the same query image

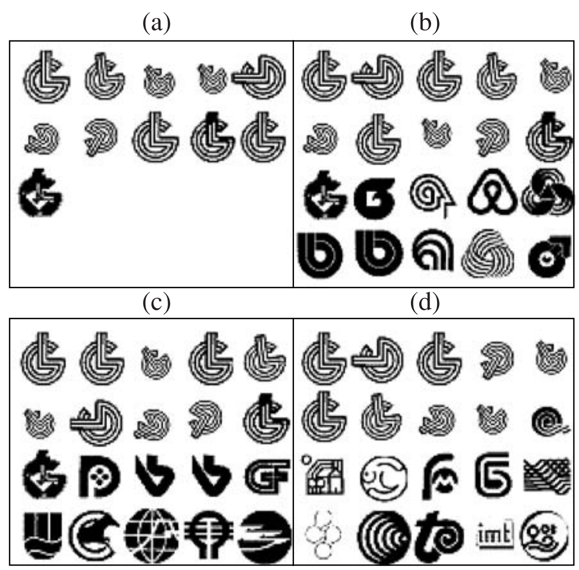

Fig. 3. An example result of the Query-By-Example test using the MPEG-7 region shape database CE-2. (a) Ground-truth set. Retrieval results of (b) ILMD, (c) ARTD, (d) ZMD. Left—top image of set (b–d) in the query image.

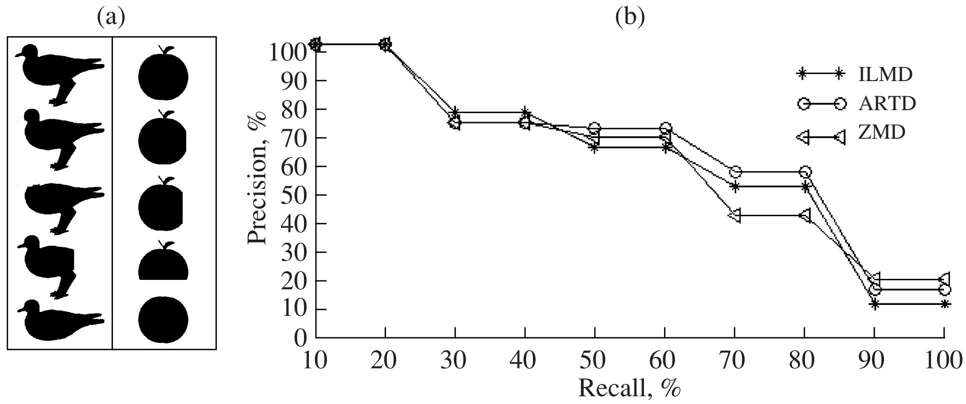

Fig. 4. (a) Example of two classes from occlusion database. (b) Average precision-recall of 50 retrievals using three shape descriptors using the occlusion database.

(left-top image) was used in all cases. It can be noted that the ILMD retrieved maximum number of images from the ground-truth and the two cases that differed are of horses ((3, 4) and (4, 1) of (b) in the (row, column) order). The ARTD has 6 misses ((4th row, 1–4 columns) and (5 row, 3–4 columns)), and ZMD has 12 misses. Figure 3 shows the first 20 retrieval results of a query image from the whole database CE-2. The ILMD retrieved all matching shapes, and the additional shapes retrieved showed greater similarity to the query image than that of the other descriptors.

To specifically test the invariance to occlusions, a database was constructed using 10 images from the CE-1 Set B. Each image was subjected to varying levels of occlusion to form a set of five images and the total database size was 50. Figure 4a shows example of two classes from the occlusion database. Each image was used as a query image and precision-recall measurement was carried out. Figure 4b shows the average precision-recall of 50 retrievals. It can be noted that ILMD has better performance for low recall rates, while performance of ZMD is slightly better than other descriptors for 100% recall rate.

Overall, performance of ILMD was better than the other two descriptors and performance of ILMD was better than ZMD. As ILMD is proved to be a better contour-based shape descriptor than ARTD and ZMD, it is proposed as a better general-purpose shape-descriptor. The ILMD is also more compact as its size is only 20, while that of ARTD and ZMD are 35 and 36, respectively.

4. CONCLUSIONS

A novel region-based shape descriptor, the Improved Legendre Moment Descriptor (ILMD), based on orthogonal Legendre moments, is reported in this paper. The preprocessing steps for invariance improvement of the descriptor is discussed, and experiments are conducted for validating the accuracy of the proposed descriptor with the state-of-the-art region-based descriptors like the angular radial transformation (ART) descriptor and the Zernike moment descriptor. Set B of the MPEG-7 CE-1 contour database and all the datasets of the MPEG-7 CE-2 region database were used for experimental validation. The average normalized modified retrieval rate (ANMRR) and precision-recall pair were employed for benchmarking the performance of the candidate descriptors. The ILMD has lower ANMRR values than ARTD for most of the datasets, and ARTD has a lower value compared to ZMD. This indicates that overall performance of the ILMD is better than ARTD and that of ARTD is better than ZMD. This result is confirmed by the precision-recall test, and the ILMD is found to have better precision rates for most of the datasets tested. Besides retrieval accuracy, ILMD is more compact than ARTD and ZMD. The proposed descriptor is useful as a generic shape descriptor for content-based image retrieval (CBIR) applications.

ACKNOWLEDGMENTS

This research is partly funded by the Council for Scientific and Industrial Research (CSIR), India. The authors would like to thank Dr. Miroslaw Bober, Visual Information Laboratory, Mitsubishi Electric, Guildford, U.K., for providing the MPEG–7 shape database for their experiments.

REFERENCES

1. M. Bober, "MPEG-7 Visual Shape Descriptors," IEEE Trans. Circuits Systems Video Technol." **11** (6), 716–719 (2001).

2. S. Loncaric, "A Survey of Shape Analysis Techniques," Pattern Recognit. **31** (8), 983–1001 (1998).

3. W.-Y. Kim and Y.-S. Kim, "A New Region-Based Shape Descriptor," *ISO/IEC MPEG99/M5472* (Maui, Hawai, 1999).

4. M. Bober, "Shape Descriptor Based on Curvature Scale Space," *MPEG-7 Proposal P320* (Lancaster U.K., 1998).

5. J. M. Martinez, R. Koenen, and F. Pereira, "MPEG-7: The Generic Multimedia Content Description Standard," IEEE Multimedia (Part 1) **9** (2), 78–87 (2002).

6. J. M. Martinez, "Standards—MPEG-7: Overview MPEG-7 Description Tools," IEEE Multimedia (Part 2) **9** (3), 83–93 (2002).

7. D. Zhangn and G. Lu, "Evaluation of MPEG-7 Shape Descriptors Against Other Shape Descriptors," Multimedia Systems **9**, 15–30 (2003).

8. R. Mukundan, S. H. Ong, and P. A. Lee, "Image Analysis by Tchebichef Moments," IEEE Trans. Image Processing **10** (9), 1357–1364 (2001).

9. H. El Fadili, Kh. Zenkouar, and H. Qjidaa, "Lapped Block Image Analysis via the Method of Legendre Moments, EURASIP J. Appl. Signal Processing **9**, 902–913 (2003).

10. M. K. Mandal, T. Aboulnasr, and S. Panchanathan, "Image Indexing Using Moments and Wavelets," IEEE Trans. Consumer Electron. **42** (3), 557–565 (1996).

11. V. P. Dinesh Kumar, J. Lijesh, K. Roopesh, and T. Tessamma, "Improved Legendre Moments Descriptor as Contour-Shape Descriptor," in *Proc. of AMSE Conf. on Modeling and Simulation* (Minsk, Belarus, 2004).

12. The Cho-huak and T. Roland Chin, "On Image Analysis by the Methods of Moments," IEEE Trans. Pattern Anal. Machine Intell. **10** (4), 496–513 (1988).

13. M. R. Teague, Image Analysis via the General Theory of Moments," J. Opt. Soc. Am. **70** (8), 920–930 (1980).

14. B. Ye and J.-X. Peng, "Invariance Analysis of Improved Zernike Moments," J. Opt., Ser. A: Pure Appl. Opt. **4**, 606–614 (2002).

15. A. Del Bimbo, "Visual Information Retrieval," Ed. by M. Kaufmann (San Francisco, USA, 1999), pp. 56–57.

16. B. S. Manjunath, Ph. Salembier, and T. Sikora, "Introduction to MPEG-7: Multimedia Content Description Interface" (John Wiley and Sons, 2003).

Dinesh Kumar V.P. was born in 1974 and took his M.Sc. degree in Electronics Science from Cochin University of Science and Technology (CUSAT) in 1996. He worked as a senior software engineer in Dharma Systems (P) Ltd., India, until 2001 and was involved in the design and development of the SQL engine for relational database management systems. He is now pursuing his PhD degree in Digital Electronics at the Department of Electronics, CUSAT, as a senior research fellow of the Council of Scientific and Industrial Research (CSIR), India. His area of research is content-based image retrieval and computer applications. He is a member of IEEE, and he has published over 10 papers in international journals or conference collections.

Dr. Tessamma Thomas was born in 1955 and took her M. Tech in Electronics and Communications Systems in 1979 from Cochin University of Science and Technology (CUSAT), Kerala, India, and the PhD in the field of speech processing from CUSAT in 1996. She is a professor of digital electronics at the Department of Electronics, CUSAT, India. Her area of research is computer vision, content-based image retrieval, medical image processing, and speech processing. She is a life member of the Society for Biotechnology, India, and has published over 45 research papers in international journals and conference collections.