

PAPER

A New Shape Description Method Using Angular Radial Transform

Jong-Min LEE[†], *Student Member and* Whoi-Yul KIM^{†a)}, *Nonmember*

SUMMARY Shape is one of the primary low-level image features in content-based image retrieval. In this paper we propose a new shape description method that consists of a rotationally invariant angular radial transform descriptor (IARTD). The IARTD is a feature vector that combines the magnitude and aligned phases of the angular radial transform (ART) coefficients. A phase correction scheme is employed to produce the aligned phase so that the IARTD is invariant to rotation. The distance between two IARTDs is defined by combining differences in the magnitudes and aligned phases. In an experiment using the MPEG-7 shape dataset, the proposed method outperforms existing methods; the average BEP of the proposed method is 57.69%, while the average BEPs of the invariant Zernike moments descriptor and the traditional ART are 41.64% and 36.51%, respectively.

key words: ART, phase, rotation invariance, image retrieval, CBIR

1. Introduction

Due to the tremendous increase in digital images, the development of accurate and efficient content-based image retrieval (CBIR) systems has become more important. Efficient image description is one of the key components in CBIR systems. Since an object's shape is a critical piece of information about an image, many researchers have aimed to develop an efficient shape descriptor [1]–[11]. Shape descriptors can be categorized into two types [8]: contour-based descriptors including the curvature scale space [5], generic Fourier descriptor (GFD) [9], and shape context [7]; and region-based descriptors including the Zernike moment descriptor (ZMD) [6], Legendre moments [10], and angular radial transform descriptor (ARTD) [11]. Since there is no general feature that would work best for every kind of image, choosing appropriate features is important in designing an image retrieval system. Contour-based descriptors such as shape context, which exhibits state-of-the-art performance in shape description, are more popular than region-based descriptors. However, they have several limitations including the fact that 1) the shape contour is unavailable in many cases, and 2) shape content is more important than contour features [12]. These limitations can be overcome by using region-based methods.

Among the existing region-based shape descriptors, ZMD has been widely used to describe the shape of a region because Zernike moments (ZMs) outperform other mo-

ments, such as geometrical moments, Legendre moments, and pseudo-Zernike moments, in terms of sensitivity to noise, redundancy, and reconstruction error [13], [14].

Rotational invariance is a desirable property for image descriptors. Traditional ZMD uses only the magnitudes of ZMs to achieve rotational invariance [6], [15]. However, since the phase is more informative than the magnitude in describing an image, several image description methods that use ZM phases have been introduced [16]–[18]. The phases of ZMs are not invariant with rotation, although they carry important information about an image. Therefore, comparing two images using the phases of two ZMs requires an additional process to align the rotation angle between the two images [16], [18]. In order to overcome this drawback, Li et al. proposed the invariant Zernike moment descriptor (IZMD) [17], which employs a phase correction scheme to obtain rotational invariance. The IZMD outperformed ZMD and GFD in the experiment described in [17].

In this paper we propose a new shape description method composed of a rotationally invariant angular radial transform descriptor (IARTD) and its distance function. The IARTD consists of the magnitude and aligned phase of the angular radial transform (ART) coefficients. The proposed method is motivated by the IZMD; however, we not only expand the phase correction scheme of the IZMD to ART, but also modify it in order to produce a more compact feature vector. The distance function of the IARTD is defined by combining the differences in the magnitudes and phases to achieve better image retrieval performance compared to the distance function of the IZMD. As demonstrated in our experiments, the proposed shape description method outperforms the IZMD in image retrieval under various image transformations; moreover, the dimension of the IARTD is smaller than that of the IZMD.

We will briefly review the fundamentals of ZMs and ART in Sect. 2, and those of IZMD in Sect. 3. The proposed shape description method is introduced in Sect. 4. Experimental data and the criteria for performance comparison are then presented in Sect. 5. The results and analysis of the comparisons are given in Sect. 6, and we conclude the paper in Sect. 7.

2. Fundamentals of Zernike Moments and Angular Radial Transform

The ZMs and ART are complex moments calculated using different orthogonal basis functions. The basis functions are

Manuscript received September 21, 2011.

Manuscript revised January 13, 2012.

[†]The authors are with the Hanyang University, Seoul, 133–791, Korea.

a) E-mail: wykim@vision.hanyang.ac.kr

DOI: 10.1587/transinf.E95.D.1628

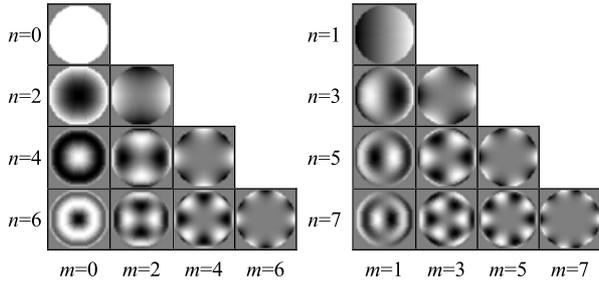


Fig. 1 Real parts of Zernike basis functions up to order 7.

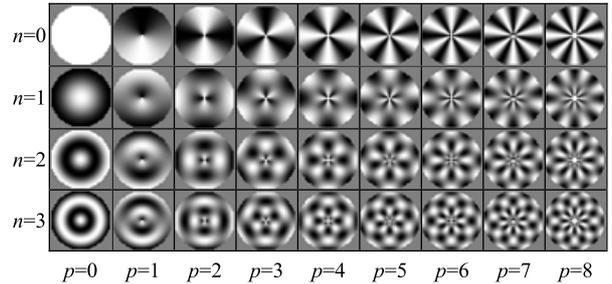


Fig. 2 Real parts of ART basis functions up to $n = 3$ and $m = 8$.

defined on a unit disk in polar coordinates.

2.1 Zernike Moment

ZMs of order n with repetition m are defined as

$$Z_{nm} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^1 f(\rho, \theta) V_{nm}^{ZM*}(\rho, \theta) \rho d\rho d\theta, \quad (1)$$

where $f(\rho, \theta)$ is the input image function, $V_{nm}^{ZM}(\rho, \theta)$ is the complex Zernike basis function, and $*$ denotes the complex conjugate. The Zernike basis functions are defined as

$$V_{nm}^{ZM}(\rho, \theta) = R_{nm}^{ZM}(\rho) \exp(jm\theta), \quad |\rho| \leq 1, \quad (2)$$

where

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

In (3), the order n is a non-negative integer, and the repetition m is an integer satisfying $n-|m| = (\text{even})$ and $|m| \leq n$. The real parts of Zernike basis functions are visualized up to $n = 7$ in Fig. 1.

2.2 ART

ART possesses additional pattern description capability that takes into account complexities in both radial and angular directions compared to ZMs [11].

$$F_{np} = \int_0^{2\pi} \int_0^1 f(\rho, \theta) V_{np}^{ART*}(\rho, \theta) \rho d\rho d\theta, \quad (4)$$

where F_{np} is an ART coefficient of order n and repetition p . p is a non-negative integer. $V_{np}^{ART}(\rho, \theta)$ is the ART basis function that is separable along the angular and radial directions as follows:

$$V_{np}^{ART}(\rho, \theta) = A_p(\theta) R_n^{ART}(\rho). \quad (5)$$

In order to achieve the rotation invariance of the magnitude, an exponential function is used for the angular basis function,

$$A_p(\theta) = \frac{1}{2\pi} \exp(jp\theta). \quad (6)$$

Depending on the types of radial basis functions, two different transforms can be defined; these are ART with cosine (ART-C) and sine (ART-S) radial basis functions.

$$\text{ART-C} : R_n^C(\rho) = \begin{cases} 1 & n = 0 \\ 2 \cos(\pi n \rho) & n \neq 0, \end{cases} \quad (7)$$

$$\text{ART-S} : R_n^S(\rho) = \begin{cases} 1 & n = 0 \\ 2 \sin(\pi n \rho) & n \neq 0. \end{cases} \quad (8)$$

The real and imaginary parts of ART-C used in this paper are shown in Fig. 2.

In previous research, ART performed better than ZMs in describing an image [8], [11].

2.3 Phase and Rotation Angle

In order to compare the phases of two sets of ZMs for two images, the rotation angle between the images should be considered. The rotation angle α between two images can be defined using phases of ZMs and ART coefficients, respectively. For ZMs, the phase differences between Z_{nm} from an original image and Z'_{nm} from a rotated image are given by

$$\Theta_{nm} \equiv \arg\left(\frac{Z'_{nm}}{Z_{nm}}\right) = m\alpha, \quad 0 < \Theta_{nm} \leq 2m\pi, \quad (9)$$

or

$$\begin{aligned} \Phi_{nm} &= (\varphi'_{nm} - \varphi_{nm}) \bmod (2\pi) \\ &= (m\alpha) \bmod (2\pi), \quad 0 < \Phi_{nm} \leq 2\pi, \end{aligned} \quad (10)$$

where φ_{nm} are phases of ZMs. More details of the fundamentals and properties of ZMs are presented in [17], [18]. Since ZMs and ART are defined using similar angular basis functions, α can be defined similarly using phases of ART coefficients with only trivial modification.

3. Invariant Zernike Moments Descriptor

IZMD was proposed by Li et al. to obtain better image description capability than traditional ZMD while maintaining rotational invariance [17]. The scale and translation invariance are obtained by the pre-normalizing image proposed by Teague [9]. The influence introduced by a rotation can simply be removed by a phase correction. To construct IZMD from the ZMs of an image, phases of different orders and

repetitions are combined to form the complex valued rotation invariant as follows:

$$Z'_{nm} = Z_{nm} e^{-jm\varphi_{n_0,1}}; n_0 \in \left\{ 1, 3, \dots, \left(2 \times \left\lfloor \frac{n_{\max}}{2} \right\rfloor - 1 \right) \right\} \quad (11)$$

From (11), the phase correction is derived as

$$\varphi'_{nm} = \varphi_{nm} - m\varphi_{n_0,1} = \varphi'_{nm} - m\varphi'_{n_0,1}; \quad (12)$$

φ^r is the phase of ZM calculated from the rotated image, and φ' is the corrected phase. A more detailed derivation of (12) can be found in [17]. In order to ensure the phase correction is robust to noise, n_0 is set to 3. Therefore, (12) becomes

$$\varphi'_{nm} = \varphi_{nm} - \varphi_{3,1} = \varphi'_{nm} - \varphi'_{3,1}. \quad (13)$$

IZMD is a feature vector that consists of magnitudes and corrected phases of the ZMs, excluding Z'_{00} and Z'_{11} . This means that Z'_{n0} , where $n \geq 2$ and $n = (\text{even})$, are used to construct a feature vector. This is inappropriate because the phases of ZMs do not have valuable information when repetition m is zero; values of imaginary parts are always zero. Similarly, the phase of Z'_{31} is always zero due to the phase correction.

The distance function of the IZMD is defined using distances of ZM magnitudes and phases as follows:

$$D_{IZMD} = w_{ang} D_{ang} + w_{mag} D_{mag}; w_{ang} + w_{mag} = 1, \quad (14)$$

$$D_{ang} = \frac{1}{N} \sqrt{\sum_{i=1}^N \left(\frac{(\varphi_i^{query} - \varphi_i^{stored})}{\pi} \right)^2};$$

$$0 \leq (\varphi_i^{query} - \varphi_i^{stored}) \leq \pi,$$

$$D_{mag} = \frac{1}{N} \sqrt{\sum_{i=1}^N \left(\frac{(|Z_i^{query}| - |Z_i^{stored}|)}{\max(|Z_i^{query}|, |Z_i^{stored}|)} \right)^2},$$

where i is the index of components and the superscripts, and *query* and *stored* represent values that were calculated from a queried image and a stored image, respectively. Li et al. assumed that the sum of magnitude and phase differences provides roughly equal contributions to the similarity measure, and set $w_{ang} = w_{mag} = 0.5$ [17]. As mentioned above, however, ZMs where m is zero are residual, and $Z'_{3,1}$ is always zero while the magnitudes are non-zero. Therefore, a distance calculated by (14) is affected more by differences in magnitudes than by differences in phases.

4. Proposed Method

In this section we present the details of the new shape description method, IARTD, and its distance function. The overall process of IARTD extraction is shown in Fig. 3.

The proposed shape descriptor IARTD consists of magnitudes and corrected phases of ART coefficients. Scale and translation invariances are obtained via the pre-normalization process proposed by Teague [10]. The main differences between the IZMD and the IARTD are 1) the

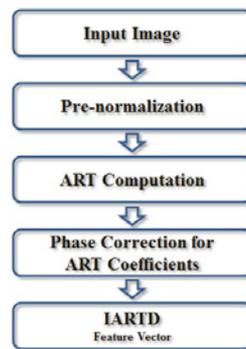


Fig. 3 Extraction of IARTD.

shape feature description (ZMs and ART), 2) the phase correction scheme, and 3) the distance function. Since ART was already presented in Sect. 2.2, in the following section we describe the steps of the new phase correction scheme and the similarity measurement of the proposed method.

4.1 Phase Correction Scheme

In order to remove the impact of rotation from phases of ART coefficients, we use the modified phase correction method of the IZMD. Unlike the repetition m of ZMs, ART basis functions are defined for every order n and repetition p since the repetition p of ART is independent of order n . Therefore, we modified the phase correction scheme of IZMD as shown in (15).

$$\phi'_{np} = \phi_{np} - \phi_{n,1} = \phi'_{np} - \phi'_{n,1}, \quad (15)$$

where ϕ is the phase of an ART coefficient. In the case of the IZMD, only one moment phase at $n = 3, m = 1$ is used as the anchor value for phase correction. In the proposed method, the coefficients at $p = 1$ of each order are used to correct the phases of ART coefficients of the same order. The proposed phase correction scheme can reduce the dimension of the IARTD by as much as $(n + 1) \times 2$ without loss of descriptive power. Finally, the IARTD is defined as

$$F'_{np} = F_{np} e^{-jm\varphi_{n,1}}. \quad (16)$$

4.2 Feature Vector Construction

The proposed shape descriptor IARTD is a feature vector consisting of magnitudes and corrected phases of ART coefficients. In order to avoid producing biased distance as in the IZMD, we define the IARTD as in (17), excluding residual coefficients that have no valuable phase information.

$$IARTD = \{|F'_{np}|, \phi'_{np}\}; n \geq 0, p \geq 2. \quad (17)$$

F'_{n0} and F'_{n1} are not included in the feature vector since they provide no valuable phase information; the imaginary part of F'_{n0} is always zero, and the phase of F'_{n0} also always becomes zero from the phase correction (15).

To use ART for an image feature, the maximum order and the maximum repetition should be defined appropriately. These two values depend on the actual resolution of the unit disk. A performance comparison of a number of different IARTD features is presented in Sect. 6. The experimental results show that the IARTD, with a maximum order of up to two and a maximum repetition of up to ten, can have sufficient shape representation power for the experimental data.

4.3 Distance Function

The magnitude of a vector represents its length and the phase represents the direction. In order to obtain good image retrieval results using both the distances of the magnitude and the phase, we define the distance of two IARTDs as in (18).

$$D_{IARTD} = \begin{cases} \frac{1}{C} \sum_n \sum_{p \geq 2} \| |F_{np}^{query}| - |F_{np}^{stored}| \| \times \phi_{np}^{diff} & C \neq 0 \\ 0 & C = 0 \end{cases}, \quad (18)$$

$$\phi_{np}^{diff} = \frac{|\phi_{np}^{query} - \phi_{np}^{stored}|}{\pi}; \quad 0 \leq \phi_{np}^{diff} \leq 1,$$

ϕ_{np}^{diff} is within the range of $[0, \pi]$ because it is an angle-based distance. C is the number of cases in which both $|F_{np}^{query}| - |F_{np}^{stored}|$ and ϕ_{np}^{diff} are nonzero. According to (18), the immediate distance $|F_{np}^{query}| - |F_{np}^{stored}| \times \phi_{np}^{diff}$ is small when both $|F_{np}^{query}| - |F_{np}^{stored}|$ and ϕ_{np}^{diff} are small. On the other hand, the immediate distance becomes zero when $|F_{np}^{query}| - |F_{np}^{stored}|$ or ϕ_{np}^{diff} is zero, regardless of whether the other value is small or large. This may lead to an inappropriate distance. Therefore, we do not use the immediate distances for which values are zero to calculate a distance of two IARTDs except for when all immediate distances are zero. In our experiment, we verify that the proposed distance function outperforms (14) in terms of image retrieval results for the experimental data.

5. Dataset and Evaluation Criteria

5.1 Experimental Image Dataset

In order to evaluate the image retrieval performance of the proposed shape descriptor, we used images from the MPEG-7 CE-1 shape dataset [19] in the experiment. The MPEG-7 shape dataset consists of various images that can be used to efficiently evaluate the performance of a shape descriptor in an image retrieval context. The following is a brief introduction of the dataset. Figure 4 shows some samples of original images from the CE-1 A1 and A2 datasets.

- Scale test dataset: The CE-1 A1 dataset consists of 70 groups with 420 images. Each group has one original image and five images with different scales (0.1, 0.2,

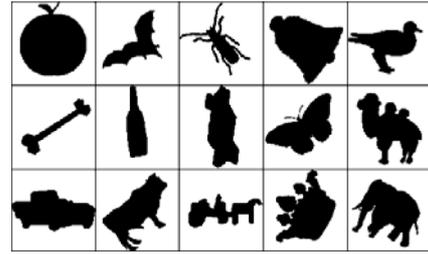


Fig. 4 Examples of original images from the CE-1 A1 and A2 datasets.

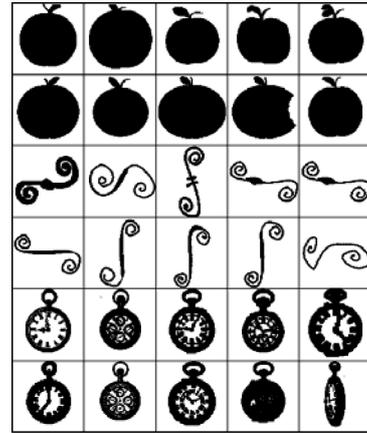


Fig. 5 Examples from the CE-1 B dataset.

0.25, 2, 3).

- Rotation test dataset: The CE-1 A2 dataset consists of 70 groups with 420 images. Each group has one original image and five images with different orientations (9°, 36°, 45°, 90°, 150°).
- Deformation test dataset: The CE-1 B dataset consists of 70 groups with 1400 images. Each group has one original image and nineteen arbitrarily deformed images. Figure 5 shows examples of deformed images in the CE-1 B dataset.
- Noise test dataset: To evaluate the pixel noise robustness of the shape description methods, we generated three impulse noise test datasets (B3, B5, and B7) by adding different degrees of impulse noise to the CE-1 B dataset in a uniform distribution pattern. Impulse noise comprises 3% (B3), 5% (B5), and 7% (B7) of the number of pixels in an image. Figure 6 shows examples of the three noisy datasets.

5.2 Measurement of Retrieval Performance

We employed Bull’s eye performance (BEP) [19] to compare the image retrieval performance of the shape descriptor. BEP was measured in terms of the number of correct retrievals among the top $2N$ retrievals, where N is the number of shapes that are relevant (or similar) to the query in the DB. In the experiments, we measured BEP (i.e., the number of correct retrievals/ N) by calculating the percentage of all images in the DB. We present the average BEP values from

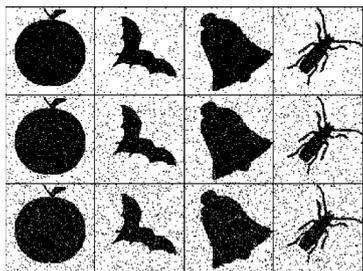


Fig. 6 Examples of noisy DBs (top row: B3, middle row: B5, bottom row: B7).

all query results to compare the performance of the shape descriptors in the next section.

6. Experimental Results

We performed a number of experiments to evaluate the performance of the proposed shape description method and to compare it with the ARTD and IZMD for the experimental dataset. The experiment consisted of three parts. We first identified the optimal parameters for the IARTD. Then we verified the effectiveness of the proposed phase correction scheme. In the third part, we compared the image retrieval performances of the IARTD with that of the traditional ARTD and IZMD. The experiments were designed to measure the robustness of the proposed IARTD and its associated distance function to scaling, rotation, deformation, and pixel noise.

We used ten for the maximum order of the IZMD based on the experimental results reported by Li [16]. Consequently, the dimension of the IZMD was 68. To verify that the proposed measurement of similarity (18) provides better image retrieval results than (14), we also performed image retrieval using IZMD-Eq17; shapes were described by the IZMD, but (18) was used for the distance function and for IARTD-Eq14. In addition, shapes were described by the IARTD, but (14) was used as the distance function. The size of an input image was normalized to 101×101 via the pre-normalization of all descriptors.

6.1 IARTD with Different Orders and Repetitions

In order to determine the effect of the values of the maximum order n_{max} and maximum repetition p_{max} on the performance of the IARTD, we performed experiments using the CE-1 B dataset. We first fixed p_{max} to 10, and then compared the retrieval results of different n_{max} values ranging from 1 to 4. The results are shown in Fig. 7.

When the max order was greater than two, the retrieval performance was not remarkably improved (Fig. 7), which means the IARTD with a max order of 2 could yield the best tradeoff between the computation cost and retrieval performance. We then performed similar experiments for the max repetition p with a fixed max order ($n_{max} = 2$).

The retrieval performance did not vary as drastically in

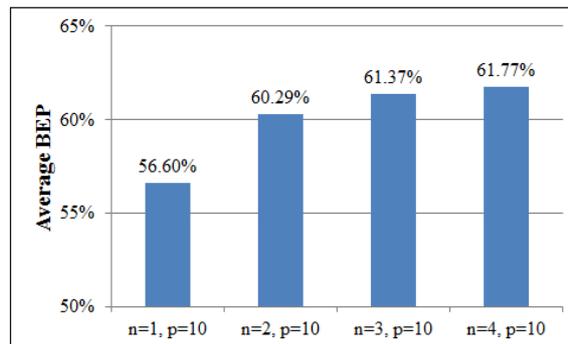


Fig. 7 BEP performance of the IARTD on deformed image DB (CE-1 B) for different max order (the p_{max} is fixed at 10).

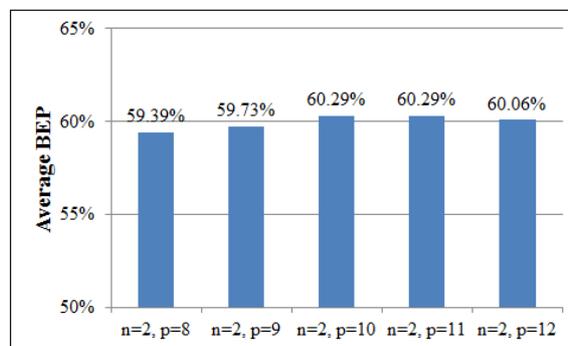


Fig. 8 BEP performance of the IARTD on deformed image DB (CE-1 B) for different max repetition (n_{max} is fixed at 2).

Table 1 BEPs (%) of IARTDs obtained using different phase correction schemes.

	A1	A2	B	B3	B5	B7	Avg.
F_{01}	96.33	99.90	56.27	31.42	29.59	27.34	56.81
F_{11}	96.38	99.76	55.38	26.23	25.73	25.13	54.77
F_{21}	95.86	99.43	56.26	24.41	22.86	22.29	53.18
F_{n1}	96.81	100.00	60.30	31.03	30.97	27.03	57.69

Fig. 8 as in Fig. 7. However, evidence for selecting an appropriate value for p_{max} was more obvious. When p_{max} was greater than 10, the retrieval performance was either not improved ($p_{max} = 11$) or even more degraded ($p_{max} = 12$). Therefore, we set $n_{max} = 2$ and $p_{max} = 10$ for the remaining experiments based on these results. Consequently, the dimension of the IARTD was $((2+1) \times 10) - (2+1) \times 2 = 54$, which is smaller than the dimension of the IZMD.

6.2 Image Retrieval Results for Different Phase Correction Schemes

To verify the efficiency of the proposed phase correction scheme (15), we compared the image retrieval performances of the IARTDs obtained using (13) and (15). Table 1 lists BEP values obtained using these two different phase correction schemes. The BEPs in the first three rows were obtained

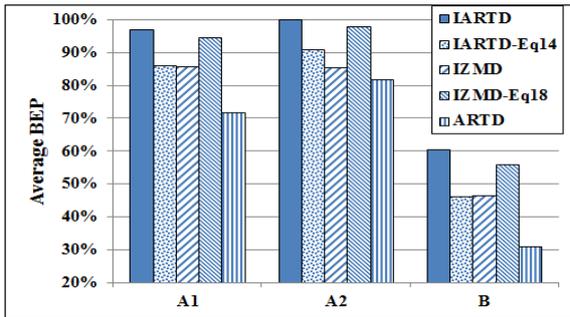


Fig. 9 BEP performances of five shape description methods for MPEG-7 CE-1 A1, A2, and B DBs.

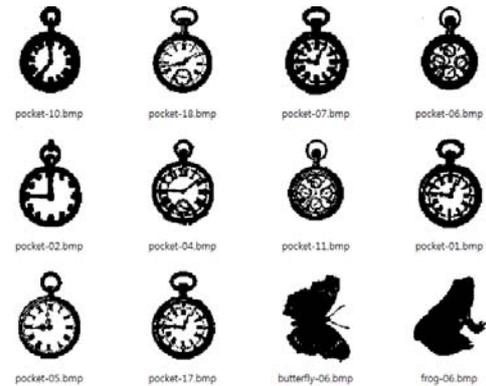
using (13) with the anchor values in the left-most column, and the BEPs in the bottom row were obtained using the proposed phased correction scheme (15). The IARTDs generated using (13) include the magnitudes and phases of F_{n1} in the feature vector.

As listed in Table 1, in most cases the IARTD obtained from the proposed phase correction scheme performed the best (F_{n1}). The highest BEPs in each column are represented in bold. When (13) was used for phase correction with different anchor values (F_{01} , F_{11} , F_{21}), the highest BEP was obtained with F_{01} . However, when we calculated the average BEPs of each anchor value using the six BEPs in Table 1, the average BEP of the F_{n1} was found to be 57.69%, which is slightly higher than 56.81% for F_{01} . Therefore, the proposed phase correction scheme enables more compact description of a shape without loss of descriptive power.

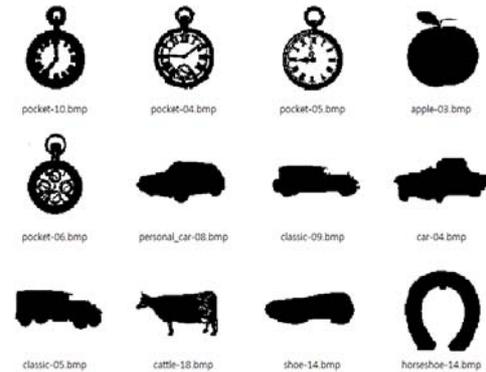
6.3 Image Retrieval Results on Scale, Rotation, and Deformation Test Datasets

Figure 9 shows the average BEPs for the image retrieval results on the scale, rotation, and deformation test datasets. All the description methods yielded excellent performance results when handling changes caused by image scaling and rotation. However, image deformation significantly degraded the performance of all the description methods. Among the five image description methods, the proposed IARTD always performed the best in terms of BEP for the three datasets. As expected, all four shape description methods that use phase to describe shape performed better than the traditional ARTD.

Two observations can be made from the graph in Fig. 9. First, the IARTD performs better than IZMD-Eq18, while IARTD-Eq14 performs better than IZMD. Therefore, we can conclude that the IARTD has better shape descriptive ability and discriminability than the IZMD, even though the feature vector dimension of the IARTD is smaller than that of the IZMD. Second, combined with the distance function, the IARTD performs better than IARTD-Eq14, while IZMD-Eq18 performs better than the IZMD. The results demonstrate that the proposed distance function (18) provides better image retrieval results than (14). In the case of the IZMD, IZMD-Eq18 outperforms IZMD due to



(a) Image retrieval result by IARTD (BEP: 89.54%)



(b) Image retrieval result by IZMD (BEP: 73.71%)

Fig. 10 Examples of image retrieval results of the IARTD and IZMD on deformation test DB (up to rank 12, $N = 19$).

the biased distance function. The average BEPs over the three datasets for IARTD, IZMD, and ARTD were 85.70%, 72.61% and 61.56%, respectively.

Figure 10 shows some sample retrieval results from the deformation test dataset using the IARTD and IZMD. The displayed results are ranked in ascending order according to their distances from the query. The upper-left image is the queried image and the rest are retrieved images. We obtain more number of images that belong to the group of the queried image with the IARTD. The BEPs of the IARTD and IZMD for the queried image were 89.54% and 73.71%, respectively.

The IARTD outperforms the ARTD and IZMD, but it also suffers from lower performance compared to the other types of shape descriptors for CE-1 B, which contains deformed images. For example, the average image retrieval performance of the IARTD was 60.03%, while the shape context, which provides state-of-the-art performance in image retrieval, yielded 76.51% BEP for the same dataset [7]. We discuss the performance degradation and noise robustness of the IARTD in Sect. 6.4.

6.4 Image Retrieval Results for Noise Test Datasets

In order to evaluate the robustness to noise of the five shape

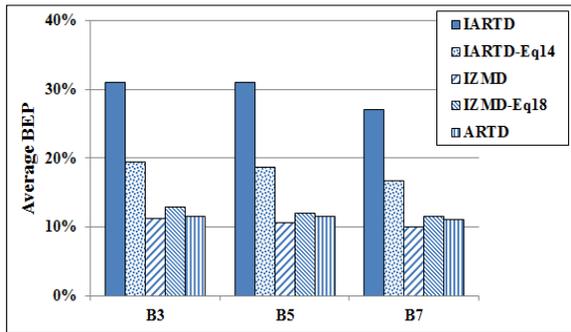


Fig. 11 BEP performances of five shape description methods for the three noise test DBs.

description methods, we also carried out BEP performance comparisons of the three noise test datasets. Figure 11 illustrates the performance comparison of the description methods under three different degrees of impulse noise.

The IARTD always outperformed other description methods throughout the experiments. Additionally, there was a significant difference in performance between the IARTD-Eq14 and IZMD. The two ZM-based methods, IZMD and IZMD-Eq18, resulted in significantly degraded performance for the three noise test datasets. From this observation, we can conclude that the feature vector of the IARTD is more robust to noise than that of the IZMD. The average BEPs of the IARTD, IZMD, and ARTD for the three noise test datasets were 29.68%, 10.68%, and 11.47%, respectively.

The IARTD yielded better image retrieval performance than the IZMD and ARTD as evidenced by the experiment described in Sects. 6.3 and 6.4. However, the performances of the five shape description methods compared in this paper were poor for images that contained noise. The performance degradation of the IARTD and IZMD was caused by phase correction schemes that tend to fail in aligning the phase components when images contain noise. Since creating a binary silhouette of an object from real-world images with a segmentation algorithm remains a difficult problem in the field of computer vision and pattern recognition [20], shape descriptors should be robust to noise to some extent. Therefore, the noise robustness of IARTD should be improved prior to its use in real applications. We intend to further improve the noise robustness of the IARTD in future work.

7. Conclusion

Efficient image description methods are critical for effective image retrieval. In this paper we proposed a new shape description method involving a shape descriptor and a distance function. The proposed shape descriptor, IARTD, consists of the magnitudes and the corrected phases of ART coefficients and is invariant to rotation. Due to its phase correction scheme, the IARTD maintains a neat description without loss of descriptive power. The distance function is defined by combining two differences: magnitudes and phases. Experiments on different datasets derived primarily from the

standard MPEG7 CE-1 shape dataset supported the robustness of the proposed shape description method under various image transformations and its superior image retrieval performance compared to that of the traditional ARTD and IZMD methods. The average BEP for all experimental images tested using the proposed method was 57.69%, while the average BEPs of the IZMD and traditional ARTD were 41.65% and 36.51%, respectively.

References

- [1] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Inf. Theory*, vol.8, no.2, pp.179–187, Feb. 1962.
- [2] E. Persson and K. Fu, "Shape discrimination using Fourier descriptors," *IEEE Trans. Syst., Man Cybern.*, vol.SMC-7, no.3, pp.170–179, March 1977.
- [3] W.I. Grosky and R. Mehrotra, "Index-based object recognition in pictorial data management," *Comput. Vis. Graph. Image Process.*, vol.52, no.3, pp.416–436, Dec. 1990.
- [4] Y.Y. Tang and C.Y. Suen, "Extraction of peripheral shape features in Chinese character recognition," *Proc. 12th IAPR Int. Conf. Comput. Vis. Image Process.*, vol.2, pp.377–379, Oct. 1994.
- [5] F. Mokhtarian and A.K. Mackworth, "A theory of multiscale, curvature based shape representation for planar curves," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.14, no.8, pp.789–805, Aug. 1992.
- [6] W.Y. Kim and Y.S. Kim, "A region-based shape descriptor using Zernike moments," *Signal Process., Image Commun.*, vol.16, pp.95–102, 2000.
- [7] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.24, no.4, pp.509–522, April 2002.
- [8] M. Bober, "MPEG-7 visual shape descriptors," *IEEE Trans. Circuits Syst. Video Technol.*, vol.11, no.6, pp.716–719, June 2001.
- [9] D. Zhang and G. Lu, "Shape-based image retrieval using generic Fourier descriptor," *Signal Process., Image Commun.*, vol.17, no.10, pp.825–848, Nov. 2002.
- [10] M. Teague, "Image analysis via the general theory of moments," *J. Optical Soc. Am.*, vol.70, pp.920–930, Aug. 1980.
- [11] W.-Y. Kim and Y.-S. Kim, "A new region-based shape descriptor," *ISO/IEC MPEG99/M5472*, Maui, Hawaii, Dec. 1999.
- [12] D. Zhang and G. Lu, "Review of shape representation and description techniques," *Pattern Recognit.*, vol.37, no.1, pp.1–19, 2004.
- [13] R.J. Prokop and A.P. Reeves, "A survey of moment-based techniques for unoccluded object representation and recognition," *Graphical Models and Image Processing*, vol.54, no.5, pp.438–460, Sept. 1992.
- [14] C.H. The and R.T. Chin, "On image analysis by the method of moments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.10, no.4, pp.496–513, July 1998.
- [15] A. Khotanzad and Y.H. Hong, "Invariant image recognition by Zernike moments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.12, no.5, pp.489–497, May 1990.
- [16] J. Revaud, G. Lavoué, and A. Baskurt, "Improving Zernike moments comparison for optimal similarity and rotation angle retrieval," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.31, no.4, pp.627–637, April 2009.
- [17] S. Li, M.-C. Lee, and C.-M. Pun, "Complex Zernike moments features for shape-based image retrieval," *IEEE Trans. Syst. Man Cybern.*, vol.39, no.1, pp.227–237, Jan. 2010.
- [18] Z. Chen and S.K. Sun, "A Zernike moment phase-based descriptor for local image representation and matching," *IEEE Trans. Image Process.*, vol.19, no.1, pp.205–219, Jan. 2010.
- [19] S. Jeannin and M. Bober, "Description of core experiments for MPEG-7 motion/shape," *Technical Report ISO/IEC JTC 1/SC29/WG 11/N2690*, MPEG-7, Seoul, March 1999.

- [20] H. Zhang, J.E. Fritts, and S.A. Goldman, "Image segmentation evaluation: A survey of unsupervised methods," *Comput. Vis. Image Understand.*, vol.110, no.2, pp.260–280, 2008.



Jong-Min Lee received M.S. degrees in Information and Communication Engineering from the graduate school of communication & information at Hanyang University, Seoul, Korea in 2001. He is now a doctoral candidate in the Division of Electrical and Computer Engineering at Hanyang University. His research interests include object segmentation, shape representation, and pattern recognition.



Whoi-Yul Kim received the Ph.D. degree in Electronics Engineering from Purdue University, W.L., IN, USA in 1989. From 1989 to 1994, He was with the Erick Johanson School of Engineering and Computer Science at the University of Texas at Dallas. He joined Hanyang University in 1994 where he is now a professor in the Department of Electronics and Computer Engineering. His research interests include visual surveillance, face tracking and identification, motion analysis, face recognition and MPEG-7 applications, where he contributed to the development of the MPEG-7 visual descriptors.