Chapter 6

Complex Zernike moments and Hough transform based solution for image retrieval

6.1. Introduction

In the previous chapter, a novel Hough transform based local descriptor is proposed. A widespread set of experiments reveal the effectiveness of the proposed local descriptor. Besides, the proposed local descriptor is combined with Zernike moments (ZMs) magnitude component to improve the image retrieval rate even further. Due to their rotation invariance property, the magnitude components of ZMs have been used as global features in many applications [121-124]. The phase coefficients of the ZMs are left out in the process of feature comparison among images. It is however, shown by [53-55,100] that the phase coefficients carry much more useful information than the magnitude. Therefore, feature matching process without involving ZMs phase would be incomplete and hence ineffective. The magnitude of ZMs remains unchanged under image rotation, but this phenomenon is not true for phase coefficients; thus, prohibiting the use of phase coefficients in shape description. The phase coefficient undergoes a shift, which depends on the rotation angle. An approach is adopted by several researchers [53-55] to first estimate the rotation angle between a query image and the training image and then make a correction in the phase coefficients. If the query image is the rotated version of training image then their phase coefficients become similar after phase correction. If the images are entirely distinct, then their phases will be different. This approach is conceptually perfect but the estimation of rotation angle is a tedious task.

In this chapter, we propose an alternate approach to correct the real and imaginary components of ZMs of the query image (which is assumed to be rotated version of the training image) and then compare them for feature matching. Unlike [53-55], the proposed approach eliminates the step of estimation of rotation angle between query and training images and then correcting phase coefficients. We use improved corrected real and imaginary components of ZMs as global image features. The proposed strategy not only provides a better feature representation process, it also provides twice the number of global features (real and imaginary components) rather than ZMs magnitude only features. This is an attractive proposition because the low-order moments are less affected by image noise and the computation time is also reduced. The similar concept of complex ZMs two
corrected components is adopted by Singh et al. [100] for face and character recognition. However, in our methodology, we apply this approach to determine its accuracy on image retrieval. In addition, the proposed two component solution is combined with proposed histograms of radii of circular arcs based local descriptor (which is described in the previous chapter in Section 5.2.2.) to build a hybrid system. The discriminative power of global and local features is evaluated by three measures: 1) Euclidean distance 2) City block distance, and 3) Bray-Curtis distance. The retrieval performance of the proposed system is compared with various major regions, and contour based shape descriptors. For performance evaluation, important parameters are taken into consideration, such as 1) geometric transformations (rotation, scale, and translation), 2) photometric transformations (noise, blur, partial occlusion, distortion, JPEG compression, etc.). For this purpose, we perform experiments on various standard databases to establish the effectiveness of the proposed system. The major contributions of this chapter include the following:

- To propose a new and improved region based descriptor based on ZMs real and imaginary components, which is robust to geometric and photometric transformations.
- To combine two corrected components of ZMs based global features and histograms of radii of circular arcs based local features to develop an effective hybrid system.
- To utilize an effective classifier that can discriminate images more effectively and coalesce both global and local features to provide high retrieval rate as compared to other generally used classifiers.

6.2. Two corrected components based ZMs features

In the proposed approach, we use both real and imaginary components of ZMs individually rather than computing their magnitude. We assume the query image to be a rotated version of the training image and find the phase difference between them. The phase difference is used to correct the real and imaginary components of ZMs of the query image. If the two images are similar but the rotated version of each other, then the corrected components of the query image will be the same as that of the training image. If the two images are dissimilar, then the corrected ZMs of the query image will be distinct from ZMs of training image. In the proposed solution, we eliminate the step of rotation
angle estimation and directly correct the real and imaginary components of the query image. Mathematically, the complete process is explained as follows

Let \( Z_{pq} \) and \( Z'_{pq} \) be ZMs of the original and rotated images, respectively, with order \( p \) and repetition \( q \), then the two moments are related by

\[
Z'_{pq} = Z_{pq} e^{-j\alpha}, \quad (6.1)
\]

where \( \alpha \) is the angle of rotation. Since ZMs magnitudes are rotation invariant therefore, we have

\[
|Z'_{pq}| = |Z_{pq} e^{-j\alpha}| = |Z_{pq}|, \quad (6.2)
\]

ZMs phase coefficients are related by

\[
\psi'_{pq} = \psi_{pq} - q\alpha \quad (6.3)
\]

or

\[
q\alpha = \psi_{pq} - \psi'_{pq}, \quad (6.4)
\]

where

\[
\psi'_{pq} = \tan^{-1}\left(\frac{I(Z'_{pq})}{R(Z'_{pq})}\right), \quad \psi_{pq} = \tan^{-1}\left(\frac{I(Z_{pq})}{R(Z_{pq})}\right), \quad (6.5)
\]

where \( \psi'_{pq} \) and \( \psi_{pq} \) are the phase coefficients of the rotated and original image respectively, and \( I(\cdot) \) and \( R(\cdot) \) are the imaginary and real components of ZMs, respectively. Using Eq. (6.1), let \( Z'^c_{pq} \) be the corrected ZMs derived from the rotated version of ZMs as follows

\[
Z'^c_{pq} = Z'_{pq} e^{jq\alpha}, \quad (6.6)
\]

or

\[
R(Z'^c_{pq}) + jI(Z'^c_{pq}) = \left(R(Z'_{pq}) + jI(Z'_{pq}) \right) \times \left(\cos(q\alpha) + j\sin(q\alpha)\right), \quad (6.7)
\]

where \( q\alpha \) is determined using Eq. (6.4). The relation given by Eq. (6.7) is decomposed as
If the query image is similar to the training image, then

$$R(Z_{pq}^c) = R(Z_{pq}') \cos(q\alpha) - I(Z_{pq}') \sin(q\alpha)$$

$$I(Z_{pq}^c) = R(Z_{pq}') \sin(q\alpha) + I(Z_{pq}') \cos(q\alpha)$$  \hspace{1cm} (6.8)

From Eq. (6.8), we obtain two corrected invariant real and imaginary components of ZMs, which we use as global features for images in the proposed system. The real and imaginary components are corrected by the phase difference between the original and rotated images. Thus, in the proposed system, we use real and imaginary components of ZMs individually rather than using the single features set obtained through ZMs magnitude. This has two major advantages over the conventional ZMs magnitude only feature and the recent approaches comprising of both the magnitude and the phase. 1) we obtain twice the number of features at low orders of ZMs as compared to the existing approaches, which is very useful as low order moments are robust to image noise. 2) we do not need to estimate the rotation angle whose estimation is a complex and computation intensive task. We consider an image from MPEG-7 database from “fork” class and rotate it at various angles as depicted in Table 6.1. The average mean square error (MSE) is estimated using

$$MSE = \frac{1}{F} \sum_{i=0}^{F} \left( R(Z_i^c) - R(Z_i) \right)^2 + \frac{1}{F} \sum_{i=0}^{F} \left( I(Z_i^c) - I(Z_i) \right)^2,$$  \hspace{1cm} (6.10)

where \((Z_i^c)\) and \((Z_i)\) are the corrected and original real and imaginary components, respectively, and \(F\) represents total number of moments upto \(p_{\text{max}} = 12\) (see Table 6.3). The results are given in Table 6.1, which demonstrates small MSE for the corrected and original real and imaginary components for all the rotated versions of the original image. On the other hand, while considering dissimilar images the MSE increases to high values as displayed in Table 6.2. Therefore, the proposed ZMs based solution is capable of discriminating similar and dissimilar images.
<table>
<thead>
<tr>
<th>Rotated Image</th>
<th><img src="image1.png" alt="Image" /></th>
<th><img src="image2.png" alt="Image" /></th>
<th><img src="image3.png" alt="Image" /></th>
<th><img src="image4.png" alt="Image" /></th>
<th><img src="image5.png" alt="Image" /></th>
<th><img src="image6.png" alt="Image" /></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotation angle</td>
<td>0°</td>
<td>72°</td>
<td>108°</td>
<td>180°</td>
<td>270°</td>
<td>324°</td>
</tr>
<tr>
<td>Average MSE</td>
<td>0</td>
<td>0.022</td>
<td>0.316</td>
<td>0.244</td>
<td>0.299</td>
<td>0.313</td>
</tr>
</tbody>
</table>

Table 6.1 Average MSE with respect to the image given in second column with rest of the images rotated at various angles

<table>
<thead>
<tr>
<th>Dissimilar images</th>
<th><img src="image7.png" alt="Image" /></th>
<th><img src="image8.png" alt="Image" /></th>
<th><img src="image9.png" alt="Image" /></th>
<th><img src="image10.png" alt="Image" /></th>
<th><img src="image11.png" alt="Image" /></th>
<th><img src="image12.png" alt="Image" /></th>
</tr>
</thead>
<tbody>
<tr>
<td>Average MSE</td>
<td>0</td>
<td>69.92</td>
<td>140.24</td>
<td>64.52</td>
<td>40.67</td>
<td>111.89</td>
</tr>
</tbody>
</table>

Table 6.2 Average MSE with respect to the image given in second column with rest of the images

6.2.1. Feature dimensionality

The selection of appropriate and optimal number of features is an important task for an effective image retrieval system. A small number of features do not provide satisfactory results, while the high number of features prone to “overtraining” and reduce the computation efficiency. In addition, higher order ZMs are vulnerable to image noise and numerical instability. Therefore, three orders of moments 10, 12, and 16 are common in image retrieval applications. Normally, the maximum order of moments $p_{\text{max}}$ is taken to be 12, which is a good tradeoff between the performance and computation complexity. The ZMs with $p_{\text{max}} = 12$ generates 41 moments given in Table 6.3. The moments $Z_{0,0}$ and $Z_{1,1}$ are excluded from the features set as $Z_{0,0}$ indicates an average gray value of image and $Z_{1,1}$ is the first order moment, which is zero if the centroid of the image falls on the centre of the disc. Moments with $q = 0$ are also discarded from the features set as they do not possess significant phase information.
6.3. Classification technique

An effective classifier is one, which can preserve the discrimination powers of features and match them appropriately. In existing methods, the Euclidean distance (ED) measure, called $L_2$ norm is used the most frequently. In ED, distances in each dimension are squared before summation, which puts more emphasis on those features for which the dissimilarity is large. Therefore, to overcome this issue we suggest city block (CB) distance, also called $L_1$ norm, and Bray-Curtis (BC), also called Sorensen’s distance measure [90,93]. The BC measure normalizes the feature values by dividing the summation of absolute differences of corresponding feature vectors by the summation of their absolute sums. We analyze the performance of the proposed system using the three distance measures ED, CB, and BC, and analyze that the BC similarity measure outperforms the rest of the measures. The ED, CB and BC measures for the proposed region based descriptor are given as:

\[
d_{r}^{ED}(Q,D) = \sqrt{\sum_{i=0}^{F} \left[ (R(Z_{r}^{Q}) - R(Z_{r}^{D}))^2 + (I(Z_{r}^{Q}) - I(Z_{r}^{D}))^2 \right]},
\]

(6.11)

\[
d_{r}^{CB}(Q,D) = \sum_{i=0}^{F} \left[ |R(Z_{r}^{Q}) - R(Z_{r}^{D})| + |I(Z_{r}^{Q}) - I(Z_{r}^{D})| \right],
\]

(6.12)

\[
d_{r}^{BC}(Q,D) = \frac{\sum_{i=0}^{F} \left[ |R(Z_{r}^{Q}) - R(Z_{r}^{D})| + |I(Z_{r}^{Q}) - I(Z_{r}^{D})| \right]}{\sum_{i=0}^{F} \left[ |R(Z_{r}^{Q}) + R(Z_{r}^{D})| + |I(Z_{r}^{Q}) + I(Z_{r}^{D})| \right]},
\]

(6.13)
where $Z^Q_i$ and $Z^D_i$ are the ZMs features of the query and training images, respectively, and $F = 40$. The ED, CB, and BC measures for the proposed local descriptor, i.e., HRC (which is described in Chapter 5 in Section 5.2.2.) are given as

$$d^{ED}_c(Q, D) = \sqrt{\sum_{i=0}^{H-1} [p_i(Q) - p_i(D)]^2},$$

(6.14)

$$d^{CB}_c(Q, D) = \sum_{i=0}^{H-1} |p_i(Q) - p_i(D)|,$$

(6.15)

$$d^{BC}_c(Q, D) = \frac{\sum_{i=0}^{H-1} |p_i(Q) - p_i(D)|}{\sum_{i=0}^{H-1} |p_i(Q) + p_i(D)|},$$

(6.16)

where $p_i(Q)$ and $p_i(D)$ represent the feature vectors of the query and training images, respectively, and $H$ is the number of features, which is 10 for contour based features. Since we consider both global and local features to describe the shape, the above corresponding similarity measures are combined to compute the overall similarity given as

$$D^{ED}(Q, D) = w_c d^{ED}(Q, D) + w_r d^{ED}(Q, D),$$

(6.17)

$$D^{CB}(Q, D) = w_c d^{CB}(Q, D) + w_r d^{CB}(Q, D),$$

(6.18)

$$D^{BC}(Q, D) = w_c d^{BC}(Q, D) + w_r d^{BC}(Q, D),$$

(6.19)

where $w_c$ and $w_r$ represent the weight factors of the contour based and region based similarity measures, respectively. In our experiments, we assume that the region and contour based features contribute equally, therefore, we set $w_c = w_r = 0.5$ for simplicity.

For evaluating the classification performance of above mentioned classifiers, we experiment on Kimia-99 and MPEG-7 shape databases to measure the retrieval accuracy of the proposed system using precision and recall $(P-R)$ curves. It is observed from Fig. 6.1(a) and 6.1(b) that the BC similarity measure performs better than the two other measures. The performance of CB is slightly superior to ED. Keeping this analysis in view, we use BC similarity measure in rest of the experiments.
6.4. Experimental study and performance evaluation

Our goal is to present the user a subset of most relevant images that are similar to the query image. In our system, the global features are extracted using the corrected real and imaginary components of ZMs and local features are extracted using the histograms of radii of circular arcs/curves (HRC), where curves are obtained by using Hough transform. The similarity value is obtained by combining the global and local features using BC classifier. The similarity values are sorted in increasing order, and the images with closest similarities are presented to the user as the most relevant images to the query image. For performance evaluation of the proposed system, we compare it against three contour based descriptors, such as FD [35], WLD [44], and CPDH [43], and three region based descriptors MI [47], ZMD [96] where ZMs magnitude is considered, and GFD [61]. Besides, we compare the proposed ZMs descriptor with complex Zernike moments (CZM) [53], optimal similarity [54], and adjacent phase [55] approaches. These approaches consider ZMs phase coefficients along with ZMs magnitude for image matching and retrieval, and we refer to them ZMMPD in rest of the chapter. The experiments are performed on an Intel Pentium Core 2 Duo 2.10 GHz processor with 3 GB RAM. All algorithms are implemented in VC++ 9.0 under Microsoft Windows environment. We consider eight databases to review the system performance under various conditions, which include rotation, scale, translation, partial occlusion, noise, blur, JPEG compression, distortion, etc. The retrieval accuracy is measured in terms of precision and recall curves.
6.4.1. Performance comparison and experiment results

In order to assess the system retrieval performance for each kind of image, all the images in the database are served as query. We present the performance of the proposed system for eight databases (their respective description can be found in Appendix A), in three sub-sections in which first section demonstrates image subject invariance analysis, second section presents geometric invariance analysis and in the third section analysis of robustness to photometric transformations is presented. The retrieval accuracy is presented by the $P-R$ curves. As mentioned earlier, we compare the proposed solution with ZMMPD and other contour and region based methods. The results are given as follows:

6.4.1.1. Image subject invariance analysis

In order to analyze the performance of the proposed descriptors more apparently, we compare the performance of the proposed contour based descriptor, i.e., HRC with other existing contour based descriptors FD, CPDH, and WLD and the proposed region based descriptor, i.e., corrected ZMs with other region based descriptors MI, ZMD, GFD, and ZMMPD individually. Subject invariance means that the performance of the system is analyzed for Kimia-99 and MPEG-7 databases, where the images within the same class represent large variations. Initially, the performance of the proposed descriptors is examined for Kimia-99 database, which includes distorted and partial occluded shapes, and the results are given in Fig. 6.2(a) for ZMMPD, which shows that CZM perform worst than other methods. The proposed corrected ZMs overpower other existing methods followed by optimal similarity and adjacent phase methods. When both ZMs and HRC based features are taken into account the performance of the proposed system is highly improved. The $P-R$ curves for comparison of the proposed HRC with contour based techniques are presented in Fig. 6.2(b), which apparently represent that the proposed HRC supersedes other contour based descriptors, and the combination of both proposed ZMs+HRC highly improves the retrieval accuracy. The performance of the region based techniques is given in Fig. 6.2(c), which also reveals that the proposed ZMs outperform rest of the region based techniques, and MI represents the overall worst performance. However, the proposed approach ZMs+HRC outperform all the existing region based descriptors. The next analysis is performed over MPEG-7 database and the performance of ZMMPD is given in Fig. 6.3(a), which apparently reveals the superiority of the proposed
ZMs over other methods followed by optimal similarity, adjacent phase, and CZM approaches. The $P-R$ curves for contour and region based descriptors are given in Fig. 6.3(b) and 6.3(c), respectively. It is observed that the proposed HRC and proposed ZMs perform better than other contour and region based descriptors and the proposed combined approach overcomes other methods and still preserves its superiority.

Fig. 6.2 Comparison of the proposed descriptors performance with (a) ZMMPD, (b) contour based descriptors, and (c) region based descriptors for Kimia-99 database
6.4.1.2. Geometric invariance analysis

This analysis is performed to review the effects of geometric transformations such as rotation, scale, and translation, i.e., whether the proposed system is invariant to such
changes in the query image. We compare the performance of the proposed descriptor (ZMs+HRC) with other regions, and contour based descriptors. The $P-R$ curves are shown in Fig. 6.4(a) for rotation invariance test. It is observed that the performance of WLD is the worst among all the methods. The proposed descriptor supersedes to rest of the methods and achieves 100% rotation invariance followed by FD and MI. This is due to the inherent property of ZMs to be rotation invariant by virtue of which the $P-R$ curves of ZMD overlaps with that of the proposed descriptor. The $P-R$ curves for the scale invariance test are depicted in Fig. 6.4(b), which reveals that the proposed descriptor along with the region based descriptors ZMD and MI are highly scale invariant and achieve 100% accuracy for this test, and their $P-R$ curves coincide with each other. The worst performance is given by WLD. The third test is performed for analyzing translation effect, and the results are given in Fig. 6.4(c), which shows that the highest accuracy is achieved by CPDH followed by MI, FD, and the proposed descriptor. Nevertheless, the performance of the proposed descriptor is far superior to GFD and ZMD. Therefore, we see that the proposed descriptor is invariant to geometric transformations.
6.4.1.3. Analysis of robustness to photometric transformations

In this analysis, the robustness of the proposed descriptor is evaluated, i.e., system is analyzed for photometric transformations such as noise, blur, and JPEG compression. The $P-R$ curves for noise test are given in Fig. 6.5(a), which shows that the proposed descriptors are highly robust to noise and achieves 100% retrieval accuracy for noise affected images, whereas the contour based descriptors gives poor performance. Figure 6.5(b) depicts the results for blur test in which MI gives the worst performance, and the proposed system still preserves its superiority. The $P-R$ curves for JPEG-compression test are given in Fig. 6.5(c), which shows that proposed descriptor attains the highest retrieval accuracy followed by CPDH, whereas MI gives the worst performance among all the descriptors.
Fig. 6.5 Performance analysis of the proposed descriptors for robustness to (a) noise (b) blur and (c) JPEG-compressed images

6.4.2. Top correct retrieval performance

The top ten retrieval results by the proposed system for Kimia-99 database are given in Fig. 6.6. The top retrievals for this database are displayed because it contains variations in images of several kinds, such as partial occlusion, rotation, different styles, poses, etc. All nine types of images are served as queries and it is observed that the proposed system is capable of retrieving relevant images from the database and only one irrelevant image appears for hand query image at rank 10.
Fig. 6.6 The top retrieval results by the proposed system for Kimia-99 database, according to the relevance to the query image.
6.5. Conclusion

In this chapter, we provide a novel solution to image retrieval system in which ZMs based global features and HRC based local features are utilized. The proposed corrected two components of ZMs approach eliminates the step of estimation of rotation angle in order to correct the phase coefficients of ZMs and to make it rotation invariant. Infact, it directly uses the relationship among the original and rotated phase coefficients of ZMs to compute $q\theta$. The corrected real and imaginary components of ZMs are used individually rather than using them to compute magnitude. Thus, the corrected real and imaginary components of ZMs are used as feature vectors representing the global aspect of images. Since the magnitude of ZMs combines both real and imaginary components, thus making it less sensitive to changes in the image. The proposed descriptor allows them to be used separately for the feature matching process. Thus, for the same moment order, we have twice the number of features than that provided by the ZMs magnitude only approach. On the other hand, the existing approaches that involve both the magnitude and the phase require the estimation of rotation angle, which is laborious to compute and the estimated rotation angle is prone to error. For local features, the histograms of normalized radii of curves are used. Both global and local features are combined by using the Bray-Curtis similarity measure to compute the overall similarity among images. The experimental results reveal that the proposed ZMs and ZMs+HRC descriptors outperform existing recent region and contour based descriptors. The vast analyses also reveal that the proposed system is robust to geometric and photometric transformations.