Shape Based Image Retrieval Using Fused Features

Maria Abro[1,*], Shahnawaz Talpur[1], Nouman Qadeer Soomro[2] and Nazish Aslam Brohi[1]

[1]Department of Computer System Engineering Mehran University of Engineering and Technology Jamshoro, Pakistan, abromaria102@gmail.com, shahnawaz.talpur@faculty.muet.edu.pk, nazishsarparrah@gmail.com

[2]Mehran UET Kairpur Campus Sindh, Pakistan, noumansoomro@muetkhp.edu.pk

Abstract

For content-based image retrieval, the shape is one of the most important discriminatory elements. The form captures most of the perceptual information of the observed objects on images in many applications, while colour and texture can often be omitted without affecting the performance of the retrieval. Unfortunately, there may be significant changes in shape, such as deformation, scaling, changes in orientation noise, and partial concealment. Accurate shape description remains, therefore, a challenging technical issue. The study performs experimental analysis to identify the problem. The adoption of the MPEG-7 and KIMIA-99 standard has significant importance to simplify the image retrieval process. The Fourier Descriptors, Moment-Based Features, Hierarchical Centroids and Histogram of Oriented Gradients have been applied for extraction of images from datasets. The fusion of features has been done by Discriminant Correlation Analysis and Direct Concatenation of features it has been evident that by fusion of features we obtained approximately 90% accurate and better results.

*Corresponding author. Email: abromaria102@gmail.com

1. Introduction

The shape is one of the essential elements of an image and plays a very important role in content-based search methods. Compared to texture or colour, the entire object can be represented by a shape, but at the same point, it requires a number of parameters to be represented explicitly. A better representation of the shape should be compact and retain the most important features of an object and the main requirements are compact features, improved accuracy of recovery, general application, low computational time and robustness of noise for good representation. By comparing and storing the understandable representation that major transition information can be achieved by representing the smart shape descriptor. Consequently, finding efficient and meaningful descriptors of the shape is a challenge in the recognition and retrieval of shape. To address these issues, we propose image retrieval based on shape using more than one method of extraction of features to improve accuracy. Without significantly increasing the cost of computing.

The study shows that by fusing several descriptors together the performance increases, by fusing two strategies, direction concatenation of features and discriminant correlation analysis. For fusion the applied techniques are, Fourier descriptors: Fourier descriptors are a one the unique way of representing the boundary of a shape [33], Hierarchal centroids: Finding the centre of the mass divides the original image into sub-images. The first step in shape-based image retrieval is to find the centre of the mass of the object in question [30], Histogram of Oriented Gradients: The gradients of an image are useful because the magnitude of the gradients is always large around the corners and edges, and the edges and corners are very important for any shape of an object [32], Moment based feature: The basic moments can be the simple weighted average of the image intensities, standard deviation or the variance [35].

2. Related work

In the area of shape representation, there are two major categories one is contour based approaches and the other is region-based approaches. Where the region approaches [1] have their region based strategies and methods simply depends on the Global features obtained from the shape and, every pixel inside a shape is utilized to get the shape representation for example, moment invariants [2] [3] and geometric invariants, multi-scale Fourier description [4], Legendre moments [5] and Zernike moments [6]. Since region descriptors cannot catch many critical shape points of interest which are the significant factor for recognizing comparative objects and have minor deformation or occlusion and they are robust to noise.

Contour-based strategies and methods mark out an object's shape by its boundary information. Among them, we have found boundary shape methods including Splines [7] [8], chain code [9], polygons approximation [10] and curvature scale space [11] used in object representation. We will survey the latest approaches like multiscale technique [12] [13] shape context [14] [15], Fourier descriptor [16] [17] [18], contour flexibility [19], and so forth [20] [21]. We are currently focusing on contour descriptors which are not the same as descriptors inside of the shape [22]. We would combine the shape descriptors to make it more effective, as we have spectral [23] [24] and moment based [25]. To make shapes gradually smooth we would use Mokhtarian and Mackworth [26] method and contour based multi scale description techniques like CSS (curvature scale space) [27] and MCC (multi scale convexity concavity) [22]. The performance is depending on the quality of the shape decomposition, so we would use convex shape decomposition [28] [29].

3. Proposed Methodology

In this work, we evaluate the performance of several features i.e. Fourier descriptors [16], Hierarchical Centroids [30], Moment-based descriptor [31] and Shape Context Descriptors [32]. Besides individual evaluations, we show that by fusing several descriptors together the performance increases, for this we tested two fusion strategies: direction concatenation of features and discriminant correlation analysis. In this section, we briefly explain the features and fusion strategies.
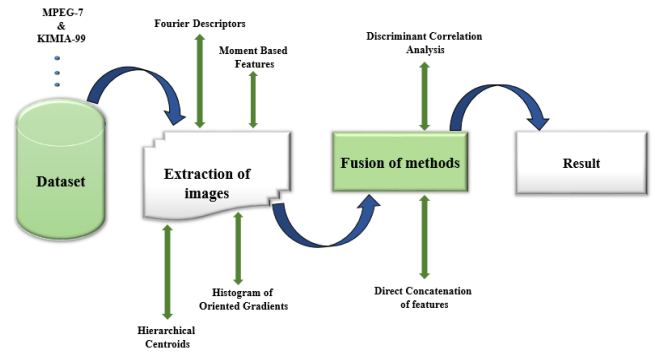

Figure 1

3.1 Fourier Descriptors

Fourier descriptors are a one the unique way of representing the boundary of a shape. For every boundary point/pixel (x, y), Fourier transform is applied, and it is converted to a complex number $(x + yi)$. Applying the inverse Fourier transform on these complex numbers, original shape of object can be recovered. The computation of Fourier descriptors can be broken down into following steps:

I. Compute the edges of shape and find the edge coordinates.

II. Iterate over each edge coordinates (x, y) clockwise and convert each coordinate to complex valued vector $(x + yi)$. For example: $(5,6) \rightarrow (5 + 6i)$.

III. Compute the discrete Fourier transform of the complex-valued vector using the equation below:

$$FD = DFT[z[m]] = \frac{1}{N} \sum_{m=0}^{N-1} [m] e^{-\frac{j2\pi mk}{N}}$$
$$(k = 0, ..., N-1) \qquad (1)$$

Where $z[m]$ complex valued vector of each pixel, and N is the total number of edge pixels.

The resultant Fourier descriptor inherits the several properties of Fourier transformation i.e. Translation, Scaling, and Rotation invariances. The translation invariance is that no matter where the edge coordinate exists in an image, the descriptor value will remain same. The scaling invariance is no matter how large or small is the object, the descriptor should remain the same, and for example an apple shape can be large or small, depending upon the camera location from where the image is captured. Similarly, the rotation is no matter how the object is oriented in an image, for example an image is captured of a chair in standing position and another image is captured where the chair is fallen on the ground, the

Fourier descriptor remains the same because the shape of the chair does not change after being fallen at 90º. The more details of Fourier Descriptors can be found in [33] [16].

3.2 Hierarchical Centroids

The approach of hierarchical centroids is dividing the original image into sub-images by finding the centre of the mass. In case of shape-based image retrieval, the first step is to find the centre of the mass of the given object. The hierarchical centroids algorithm [30] takes input an image and computes the x-coordinate of the centre of the object. Based on the x-coordinate of the centre, it further divides the image into sub-images and this process continues recursively until the desired level of centroids are found. The resultant centroids of the image are stored together to form a descriptor. These features play a vital role in identifying the shape of any object, because each object may have unique shape. To further bring the robustness into the descriptor, the original image is rotated at $90^{\circ}, 180^{\circ}$ and 270° and descriptor each transformed image is computed separately.

3.3 Moment based Features

The use of image moments has been extensively studied in the past [34] [31] [35] [36] to extract the useful of image properties from the image intensities. The basic moments can be the simple weighted average of the image intensities, standard deviation or the variance etc., and often referred as raw moments. In this research, we employ the seven invariant moments that are scale, rotation, and translation invariants.

3.4 Shape Context based Features

The basic shape features can be the edges and contours of the shape of the object in an image, but in this work we exploit the more powerful descriptor that has been popular descriptor since last decade, known as Histogram of Oriented Gradients (HoG) [32]. The image gradients (first order derivatives in x-y directions) of any image useful because the gradients magnitude is always large around the corners and edges, and for any shape of any objects the edges and corners are very important. HoG computes the histogram of these oriented gradients from small windows like $(8 * 8), or (16 * 16)$ and accumulating these together forms the descriptor.

3.5 Fusion Strategies

In this work, we compare two fusion strategies and show the importance of fusing different complementary features together by achieving state-of-the-art accuracy on the two standard datasets MPEG-7 [37] Kimia-99 [38]. The first fusion strategy is direct concatenation of features, in which all four types of descriptors discussed earlier are directly concatenated and Euclidean distance is used to evaluate the performance. In second strategy, the Discriminant Correlation Analysis (DCA) is employed [39] [40].

a. Direct Concatenation of Features

In fusion, the one of the basic strategies is to combine all the features together without and processing, in many cases this kind of fusion works better than individual, and our results also show that combining features together gives better performance as compared to individual features.

b. Discriminant Correlation Analysis

DCA is an effective feature fusion strategy the exploits from the correlation of features of inter and intra class. The method tries to maximize the inter class correlation and at the same time also minimizes the intra class correlation, so that the features of each class lies within the class. The benefit of this method is twofold: it considers the class structure and secondly, its computational complexity is very low, which is required in practical applications like shape-based image retrieval.

4. Experiments and Results

In this section, we explain all the experiments followed results and findings. We assess the performance of individual features and also fusion of all the possible combinations using two fusion strategies explained in Section 3.5. To start with, the dataset used in these experiments are discussed below:

4.1 Datasets

In order to analyse and highlight the importance of fusion of features, two different datasets are used: MPEG-7 [37] and Kiami-99 [38]. MPEG-7 dataset consists of 70 classes with 20 images for each class. Kimia99 dataset only contains 9 classes with 11 images in each class. The both datasets have been widely used to measure the performance of shape-based image retrieval. The examples shapes of both MPEG-7 and Kimia99 dataset are shown in Figure 1 and Figure 2 respectively.
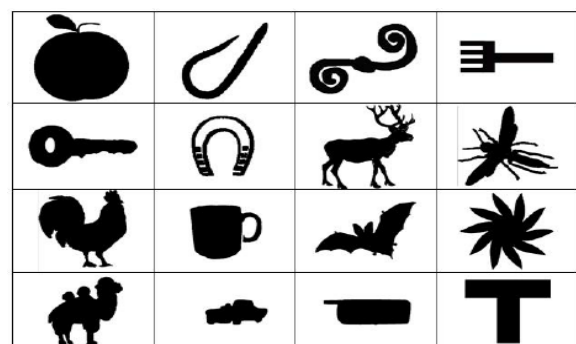

Figure 2. Example shapes taken from MPEG-7 dataset.

Figure 3. Examples shapes from Kimia99 dataset.

4.2 Results and Discussion

In the proposed technique we present the results of four different features i.e. Fourier descriptors (FD), Hierarchical Centroids (HC), Moment based features (MF), Shape Context Features (SCF) and also evaluate two features fusion strategies: Direct Concatenation (DC) of features, and Discriminant Correlation Analysis (DCA).

a. Dataset MPEG-7

In this step the image1, image2 image3 and image5 are representing the (initial) query image of a 'Dog' that will be taken to retrieve the most similar images from the dataset MPEG-7 with the help of FD, HC, MF and SCF. Other five images are the retrieved images after extraction.

Figure 4. Image1 retievals using Fourier descriptors

Figure 5. Image2 retievals using Moment Based Features

Figure 6. Image3 retievals using Hierarchical Centroids Descriptors

Figure 7. Image4 retievals using Histogram Of Oriented Gradients

After obtaining the results in the previous step as showns in the images we will fused the features of all the techniqes to aquire the better accurary. Similarly as mentioned above query image will be taken to retrieve the most similar images from the dataset MPEG-7 with fused techniques.

Figure 8. Final resultant fused image(Dataset MPEG-7)

b. Dataset KIMIA-99

In the same way as above: image1, image2 image3 and image5 are representing the (initial) query image of a 'Cat' will be taken to retrieve the most similar images from the dataset KIMIA-99 with the help of FD, HC, MF and SCF. Other five images are the retrieved images.

Figure 9. Image1 retievals using Fourier descriptors

Figure 10. Image2 retievals using Moment Based Features

Figure 11. Image3 retievals using Hierarchical Centroids Descriptors

Figure 12. Image4 retievals using Histogram Of Oriented Gradients

And again after getting the results in the previous step as showns in the images we will fused the features of all the techniqes to aquire the better accurary. Similarly as mentioned above query image will be taken to retrieve the most similar images from the dataset KMIA-99 with fused techniques.
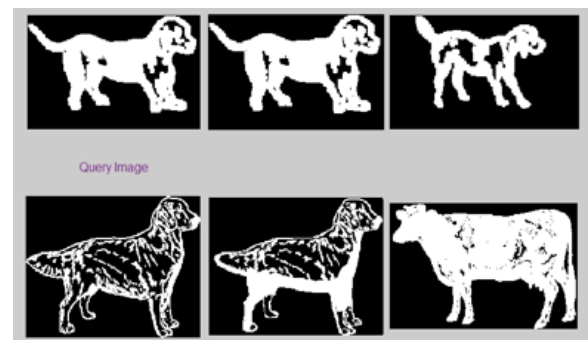

Figure 13. Final resultant fused image (Dataset KIMIA-99)

4.3 Evaluation Metrics

For the problem of shape-based image retrieval, Bull-Eye Score is the standard evaluation metrics. To compute the bull eye score, every shape in the dataset is compared with all the other shapes, and the number of shapes from the same class among the top 40 retrieval is counted. The final bull eye retrieval rate is computed using the equation below:

$$Bull\,Eye\,Score = \frac{correctly\;retrieved}{classes * No.\,of\,images} \quad (2)$$

Where correctly retrieved is the total number of shapes from the same class and denominator represents the highest possible number.

Alongside the bull eye retrieval score it also compute and report the Nearest Neighbour (NN) accuracy. Specifically, we report five matching retrieval accuracies.

In Table 1, the bull eye score of individual features is reported. The Shape Context features clearly outperforms all others, and the performance of moment-based features is very low. This behaviour of results was already expected, because the moment-based features are not very good in capturing the shape of the object.

Table 1. Bull Eye Score on MPEG-7 Dataset for each descriptor separately, where FD, HC, MF and SCF represents Fourier Descriptors, Hierarchical Centroids, Moment Features, and Shape Context Features respectively.

Features	Bull Eye Score (%)
FD	64.72
HC	62.77
MF	34.26
SCF	65.62

The results of fusion of different combination of features is provided in Table 2. The results of DC fusion strategy directly correlate with the results of individual features, as can be observed from the fusion of FD and HC, that bull eye score of DC approach is slightly higher than their individual performance, and similar trend can be observed for all the other combination of features. But this trend is violated for DCA strategy. When two different features that captures two different properties of the images, the bull eye score is very low. For example, the FD captures the boundary of the shape, and HC focuses on the centers of the mass of the object, both computes two different properties of any shape, due to their complementary nature, the correlation between both is very low, and hence the DCA performance is also low. Whereas, the performance of the DCA strategy is higher, where features are somehow correlated with each other. As can be observed from the results, when SCF is combined with itself and DCA is applied the bull eye score jumps to 80.55% from 65.62%.

Table 2. Bull Eye Score on MPEG-7 dataset for both fusion Strategies: DC (Direct Concatenation) and DCA (Discriminant Correlation Analysis) is presented for different combination of features.

Features	Bull Eye Score (%)	
	DC	DCA
FD + HC	65.29	34.36
FD + MF	66.82	53.62
FD + SCF	65.93	27.70
HC + MF	63.58	57.85
HC + SCF	66.90	40.32
MF + SCF	65.67	53.07
FD + HC + MF	65.90	62.84
FD + HC + SCF	67.17	36.75
FD + MF + SCF	65.96	67.01
HC + MF + SCF	66.92	35.74
FD + MF + HC + SCF	67.19	41.79
FD + FD	64.72	64.72
HC + HC	62.77	50.02
MF + MF	34.26	39.10
SCF + SCF	65.62	80.55

Besides bull eye score, the top 1 and top 5 nearest neighbour based accuracy is also reported in Table 3 and Table 4 respectively. The top 1 retrieval rate for all the features is 100%, this shows that the features we evaluated in this study are well suited for the problem of shape-based image retrieval. But the top 5 retrieval rate suggests that these features alone are not sufficient for a robust image retrieval system.

Table 3. Top 1 Nearest Neighbor (NN) accuracy is reported for all the features on MPEG Dataset.

Features	Top 1 NN Accuracy (%)
FD	100.0
HC	100.0
MF	100.0
SCF	100.0

Table 4. Top 5 Nearest Neighbor (NN) accuracy is reported on MPEG dataset for all the features.

Features	Top 5 NN Accuracy (%)
FD	85.26
HC	82.47
MF	47.01
SCF	87.21

The top 5 nearest neighbour accuracy on MPEG-7 dataset for both fusion strategies is tabulated in Table 5. All the possible combination of features against both fusion strategies are evaluated. Unlike the bull eye score, the top 5 retrieval accuracies are higher. The best results that are achieved is 91.96% for the fusion of SCF with itself, and all the accuracies are higher than the results obtained by individual features, that highlights the importance and need of fusion of different features. As it is stated earlier, that different features capture the different properties of the shapes of an object in an image, so fusing these different properties together always increase the overall performance of the system.

In Figure 3, the qualitative results on few images of MPEG-7 dataset is illustrated. The top 5 retrievals of DCA fusion strategy of SCF features with itself is presented. The results clearly suggest the fusion of Shape Context features with itself is a good choice for the problem of shape-based image retrieval. The DCA, finds the correlation between the features and projects the features to a new space where these are more separable.

Table 5. Top 5 Nearest Neighbor (NN) Accuracy on MPEG-7 dataset for both fusion Strategies: DC (Direct Concatenation) and DCA (Discriminant Correlation Analysis) is presented for different combination of features.

Features	Top 5 NN Accuracy (%)	
	DC	DCA
FD + HC	84.52	62.88
FD + MF	85.32	71.18
FD + SCF	87.38	57.06
HC + MF	83.61	73.60
HC + SCF	87.88	70.64
MF + SCF	87.26	68.40
FD + HC + MF	85.42	82.24
FD + HC + SCF	87.98	66.94
FD + MF + SCF	87.62	83.52
HC + MF + SCF	87.92	53.14
FD + MF + HC + SCF	88.02	69.54
FD + FD	85.26	85.26
HC + HC	82.47	78.40
MF + MF	47.01	49.60
SCF + SCF	87.22	91.96

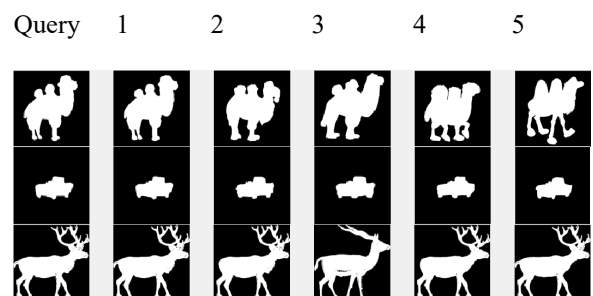| Query	1	2	3	4	5

Figure 14. Qualitative results of DCA fusion of SCF with itself.

4. Conclusion

In this study, we evaluate the performance of four different descriptors: Fourier, Hierarchical Centroids, Moment-based, Shape Context features for the task of shape-based image retrieval. Moreover, two different fusion strategies: Direct Concatenation of features and Discriminant Correlation Analysis, for different combinations of features are evaluated and assessed. The results clearly show that performance of separate features is lower than the fusion of different features and separate features are not enough for building a robust and accurate image retrieval system. Because each feature captures a different property of the objects in the images, combining these different complementary features together increases the overall performance of the system.

References

[1] Aubert, G., Barlaud, M., Faugeras, O. and Jehan-Besson, S., 2003. Image segmentation using active contours: Calculus of variations or shape gradients? *SIAM Journal on Applied Mathematics*, *63*(6), pp.2128-2154.

[2] Hu, M.K., 1962. Visual pattern recognition by moment invariants. *IRE transactions on information theory*, *8*(2), pp.179-187.

[3] Flusser, J., 2000. On the independence of rotation moment invariants. *Pattern recognition*, *33*(9), pp.1405-1410.

[4] Direkoğlu, C. and Nixon, M.S., 2011. Shape classification via image-based multiscale description. *Pattern Recognition*, *44*(9), pp.2134-2146.

[5] Teague, M.R., 1980. Image analysis via the general theory of moments. *JOSA*, *70*(8), pp.920-930.

[6] Ma, Z.M., Zhang, G. and Yan, L., 2011. Shape feature descriptor using modified Zernike moments. *Pattern Analysis and Applications*, *14*(1), pp.9-22.

[7] Cohen, F.S., Huang, Z. and Yang, Z., 1995. Invariant matching and identification of curves using B-splines curve representation. *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society*, *4*(1), pp.1-10.

[8] Laiche, N. and Larabi, S., 2014, January. 2D shape matching based on B-spline curves and dynamic programming. In *2014 International Conference on Computer Vision Theory and Applications (VISAPP)* (Vol. 1, pp. 484-491). IEEE.

[9] Liu, Y.K., Wei, W., Wang, P.J. and Žalik, B., 2007. Compressed vertex chain codes. *Pattern Recognition*, *40*(11), pp.2908-2913.

[10] Mokhtarian, F., Abbasi, S. and Kittler, J., 1997. Efficient and robust retrieval by shape content through curvature scale space. In *Image Databases and Multi-Media Search* (pp. 51-58).

[11] Mokhtarian, F. and Bober, M., 2013. *Curvature scale space representation: theory, applications, and MPEG-7 standardization* (Vol. 25). Springer Science & Business Media.

[12] Alajlan, N., El Rube, I., Kamel, M.S. and Freeman, G., 2007. Shape retrieval using triangle-area representation and dynamic space warping. *Pattern recognition*, *40*(7), pp.1911-1920.

[13] Belongie, S., Malik, J. and Puzicha, J., 2002. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (4), pp.509-522.

[14] Kokkinos, I., Bronstein, M.M., Litman, R. and Bronstein, A.M., 2012, June. Intrinsic shape context descriptors for deformable shapes. In *2012 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 159-166). IEEE.

[15] Zhang, D. and Lu, G., 2002, August. Generic Fourier descriptor for shape-based image retrieval. In *Proceedings. IEEE International Conference on Multimedia and Expo* (Vol. 1, pp. 425-428). IEEE.

[16] Persoon, E. and Fu, K.S., 1986. Shape discrimination using Fourier descriptors. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (3), pp.388-397.

[17] Bartolini, I., Ciaccia, P. and Patella, M., 2005. Warp: Accurate retrieval of shapes using phase of fourier descriptors and time warping distance. *IEEE transactions on pattern analysis and machine intelligence*, *27*(1), pp.142-147.

[18] Xu, C., Liu, J. and Tang, X., 2008. 2D shape matching by contour flexibility. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *31*(1), pp.180-186.

[19] Attalla, E. and Siy, P., 2005. Robust shape similarity retrieval based on contour segmentation polygonal multiresolution and elastic matching. *Pattern Recognition*, *38*(12), pp.2229-2241.

[20] Shu, X. and Wu, X.J., 2011. A novel contour descriptor for 2D shape matching and its application to image retrieval. *Image and vision Computing*, *29*(4), pp.286-294.

[21] Gorelick, L., Galun, M., Sharon, E., Basri, R. and Brandt, A., 2006. Shape representation and classification using the poisson equation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (12), pp.1991-2005.

[22] Zhang, D. and Lu, G., 2003. A comparative study of curvature scale space and Fourier descriptors for

shape-based image retrieval. *Journal of Visual Communication and Image Representation*, *14*(1), pp.39-57.

[23] Tieng, Q.M. and Boles, W.W., 1997. Recognition of 2D object contours using the wavelet transform zero-crossing representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *19*(8), pp.910-916.

[24] Papakostas, G.A., Karakasis, E.G. and Koulouriotis, D.E., 2010. Novel moment invariants for improved classification performance in computer vision applications. *Pattern Recognition*, *43*(1), pp.58-68.

[25] Mokhtarian, F. and Bober, M., 2013. *Curvature scale space representation: theory, applications, and MPEG-7 standardization* (Vol. 25). Springer Science & Business Media.

[26] Loncaric, S., 1998. A survey of shape analysis techniques. *Pattern recognition*, *31*(8), pp.983-1001.

[27] Liu, H., Liu, W. and Latecki, L.J., 2010, June. Convex shape decomposition. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 97-104). IEEE.

[28] Ren, Z., Yuan, J., Li, C. and Liu, W., 2011, November. Minimum near-convex decomposition for robust shape representation. In *2011 International Conference on Computer Vision* (pp. 303-310). IEEE.

[29] Park, J., Govindaraju, V. and Srihari, S.N., 2000. OCR in a hierarchical feature space. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(4), pp.400-407.

[30] Yan, F., Mei, W. and Chunqin, Z., 2009, August. SAR image target recognition based on Hu invariant moments and SVM. In *2009 Fifth International Conference on Information Assurance and Security* (Vol. 1, pp. 585-588). IEEE.

[31] Dalal, N. and Triggs, B., 2005, June. Histograms of oriented gradients for human detection.

[32] Yang, L., 2007, October. An overview of distance metric learning. In *Proceedings of the computer vision and pattern recognition conference.*

[33] Kim, W.Y. and Kim, Y.S., 2000. A region-based shape descriptor using Zernike moments. *Signal processing: Image communication*, *16*(1-2), pp.95-102.

[34] Kim, H.K. and Kim, J.D., 2000. Region-based shape descriptor invariant to rotation, scale and translation. *Signal Processing: Image Communication*, *16*(1-2), pp.87-93.

[35] Flusser, J. and Suk, T., 2006. Rotation moment invariants for recognition of symmetric objects. *IEEE Transactions on Image Processing*, *15*(12), pp.3784-3790.

[36] Chang, S.F., Sikora, T. and Purl, A., 2001. Overview of the MPEG-7 standard. *IEEE Transactions on circuits and systems for video technology*, *11*(6), pp.688-695.

[37] M. M. S. d. Souza, "Kimia-99 Dataset," 8 April 2016. [Online]. Available: https://github.com/mmssouza/kimia99. [Accessed 15 March 2019].

[38] Haghighat, M., Abdel-Mottaleb, M. and Alhalabi, W., 2016. Discriminant correlation analysis: Real-time feature level fusion for multimodal biometric recognition. *IEEE Transactions on Information Forensics and Security*, *11*(9), pp.1984-1996.

[39] Haghighat, M., Abdel-Mottaleb, M. and Alhalabi, W., 2016, March. Discriminant correlation analysis for feature level fusion with application to multimodal biometrics. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1866-1870). IEEE.

[40] "MPEG-7 Dataset," 8. [Online]. Available: http://www.dabi.temple.edu/~shape/MPEG7/dataset.html [Accessed 15 March 2019].