Feature Fusion in Improving Object Class Recognition

[1]Noridayu Manshor, [2]Amir Rizaan Abdul Rahiman, [3]Mandava Rajeswari and [3]Dhanesh Ramachandram

[1]Department of Computer Science, Faculty of Computer Science and Information Technology, Universiti Putra Malaysia
[2]Department of Multimedia, Faculty of Computer Science and Information Technology, Universiti Putra Malaysia
[3]Computer Vision Research Group, School of Computer Science, Universiti Sains Malaysia

Abstract: Problem statement: Extraction of features in object class recognition researches previously gives attention to local features as discriminative features. This is because local features have invariant properties that are robust to viewpoints, translation and rotation. However this feature still has a limitation to represent high-level representation of objects. The problem will occur if the object is too small and do not have strong local features. **Approach:** This study proposes the combination of different features with local features for improving performance of object class recognition. The objective of this study is to address the problem of building object class representation based on these different features. The different features are sourced from boundary-based shape features. The dataset used consists of segmented objects with unrestricted poses and sizes from publicly image database. Both types of features are combined using feature fusion approach by concatenating those features in a new single feature vector. This new feature vector is trained by Support Vector Machine (SVM) to predict of unknown object class. **Result/Conclusion:** Experimental result show the inclusion of more than one type of features yields improvements of object class recognition compared to using single feature.

INTRODUCTION

An image can be easily understood by human but everyone has different views in describing an image. In Fig. 1, human may describe the image based on the scenery or surrounding, such as "city" or "outdoor" image. This image may also be recognized based on objects contained in the image such as "car" and "trees". Those objects are identified based on their features such as shapes and colors. In computer vision research, there are various features introduced by researchers to support an object recognition system that is able to capture similar concepts as understood by humans.

Features of objects in an image can be extracted through their shapes, colors and sizes. In addition, the objects can be seen in different range of views, for example front view, side view or rear view. To relate these visual features into a higher level of conceptual representation that is closer to human understanding, it is sufficient to identify the category or class of object, known as Object Class Recognition. To achieve this, the semantic gap between the simplicity of visual features and the richness of user semantics needs to be reduced (Hare *et al*., 2006). At this juncture, much efforts in related research attempt to map an object within an image to a suitable class, which is also referred to as a "concept".

Earlier researchers introduce local features to identify objects with different variability in terms of poses and sizes. Local features refer to features that are extracted based on interest points detected on the object generated by region detector.

Fig. 1: Sample of dataset from Graz02 http://www.emt.tugraz.at/~pinz/data/GRAZ_02/

Interest points capture information from its neighbors and invariant under scales, translation and rotation. What makes local features appealing is their ability to represent the variability of object classes of different scales, orientation, sizes or poses. Opelt *et al*. (2006a) use many local features such as Scale Invariant Transform Features (SIFT), subsampled gray values, basic intensity moments and moment invariants as input to the boosting classifier in recognizing object classes. The author concludes that the classification performance of combinatory local features yield higher accuracy as compared to solely using the SIFT feature. However, many object classes such as "cups", "horses" or "cows" are better described by shape features as compared to local features. For example, "cup" objects have limited local features, for example fixed color or shade. This will make it difficult to discriminate among the classes and in turn result in poor recognition results.

Furthermore, local features focus on the local information of objects without considering other properties such as shape. This causes a problem for the computer to recognize objects that have limited or plain local features (Mansur and Kuno, 2007). Shape features are often used as a replacement of, or complement to local features in several works, such as in (Opelt *et al*., 2006b; Yu *et al*., 2007; Shotton *et al*., 2008). Due to the richness of information, shape is an important part of the semantic content of images and it should be the main feature in recognizing object classes (Yang *et al*., 2008). Several researchers concentrate on local shape information such as shape context and area. Other shape features are based on contour fragments (Shotton *et al*., 2008), which represent the partial shape of objects. However, the contour fragment cannot guarantee the actual shape of the object. In addition, it may be affected by high resolution noise and small details may disappear in low resolution noise.

Minority applications concentrate on full contour of the object's shape such as in face recognition (Su *et al*., 2003) and medical image retrieval (Arun and Menon, 2009; Jeong and Radke, 2007; Schaefer *et al*., 2009), where such objects have roughly restricted shape poses. Nevertheless, natural images can consist of objects with different poses. For example, Fig. 2 shows images of car class in different poses such as rear, front and side poses. Although, they have different poses, these images can be categorized into similar class. However, the representation of shape for an object changes once the poses of the object changes. To overcome this limitation, several papers take the advantageous from local features in combining with shape features to contribute to the improvement of

object class recognition (Mansur and Kuno, 2007; Opelt *et al*., 2006b; Zhang *et al*., 2005).

Fig. 2: Several car poses

This study proposes the boundary-based shape features that describe the entire contour of object class to be combined with local features. Comparing with the color and the texture, the shape is described after the objects in the image have been segmented. Moreover, the shape features are capable to represent the entire object, hence can be interpreted by human vision. Good recognition accuracy requires an effective shape features that are as similar as possible to the interpretation of human perceptual (Yang *et al*., 2008). The advantages of these features are that they can be robustly extracted from the image. They are insensitive to surface features such as texture, color and also invariant to lighting conditions. Furthermore, the shape of objects may be easily encoded.

However, the object's shape extracted by boundary-based features may lead a problem of ambiguity in recognition process. This is because for a natural image, single pose of the object is insufficient in identifying the actual objects. Hence, this study will consider numerous poses of objects in resolving the ambiguity problem. Then, the local features is used to discriminate the object that cannot be distinguished using shape features since it is ability to resolve the problem of detecting objects in various poses, scales and rotations.

To predict the class of unlabelled objects based on visual features, feature fusion approach is used. Feature fusion is a method of combining multiple features in a new single feature vector (Oliveira and Nunes, 2008; Sun *et al*., 2009; Ali, 2007). This is a simple approach where the features are mapped into one feature space. A Support Vector Machine (SVM) classifier is used to train this new feature vector due to its ability to generalize and to support high-dimensional and non-linear data for classification.

MATERIALS AND METHODS

Feature fusion is a straightforward method that will form the input to the classifier. At present, researchers are facing difficulties in determining the combination of methods that could produce optimal results (Kludas *et al*., 2008; Dimitrovski *et al*., 2011). In our study, the features are sourced from two different types of

features, which are boundary-based shape features and local features. The former type of feature is based on the outline of segmented objects while the latter are based on the interior information of objects. The motivation in this study is to demonstrate that combination of different features is able to produce better recognition performance as compared to using single type of feature. The feature fusion framework proposed by our study is illustrated in Fig. 3. Each feature is computed separately and is represented as feature vector. All feature vectors of boundary-based shape features and local features are then concatenated into a single vector. The vector can be defined as the following Eq. 1:

$$O_{fi} = [FD_1, FD_2, FD_3, ..., FD_j]$$
$$O_i = [EFD_1, EFD_2, EFD_3, ..., EFD_j]$$
$$O_i = [MI_1, MI_2, MI_3, ..., MI_j]$$
$$O_i[SIFT_1, SIFT_2, SIFT_3, ..., SIFT_j]$$
$$OC_i = [FD_1, ..., FD_j, EFD_1, ..., EFD_j,$$
$$MI_1, ..., MI_j, SIFT_1, ..., SIFT_j]$$

(1)

where, O_i is a object class, FD, EFD, MI represents the boundary-based shape features and SIFT represents the local features. OC_i corresponds to a new feature vector that resulted from concatenation of all feature vectors, FD, EFD, MI and SIFT.

Feature extraction: Feature extraction is divided into two types of features; boundary-based shape features and local features. The boundary-based shape features used in this study are Fourier Descriptors (FD), elliptical FD (EFD) and Moment Invariants (MI) are extracted from a segmented dataset. For local features, Scale Invariant Feature Transform (SIFT) is adapted to cooperate with the shape features.

Boundary-based shape features: This boundary-based shape features are based on silhouettes of the segmented objects. The primary factors that are taken into consideration include invariance under translation, rotation, reflection and scaling (Gonzalez et al., 2004). These features are employed due to its accurate in shape representation and can be easily normalized (Zhang and Lu, 2004). Previously, FD and EFD features have not been used in the object class recognition research. FD and EFD are widely used in medical image processing (Arun and Menon, 2009; Jeong and Radke, 2007; Schaefer et al., 2009; Reig-Bolano et al., 2010).

FD: FD values are produced by the Fourier transformation of a given image that represents the shape of the object in frequency domain (Gonzalez et al., 2004). Based on frequency analysis, the Fourier coefficients can be used to describe shape of an object. These shape descriptors are normalized in order to make them independent from translation, rotation and scale. Higher frequency descriptors will generate detailed shape of an object, whereas lower frequency descriptors will create rough shape from the original object. In shape description, the Fourier transform theory may be applied in many different ways. In this study, the boundary (outline) of the object is treated as a layer in a complex plane (Zahn and Roskies, 1972), with row and column co-ordinates of each point on the boundary, $B(k) = [x(k), y(k)]$, $k = 0,1,....,K-1$ can be expressed as a complex number as denoted by Eq. 2:

$$b(k) = x(k) + jy(k)$$

(2)

where, j is the sqrt(-1). The boundary point starts at an arbitrary point, (x_o, y_o) and are traced around the boundary in counterclockwise direction at a constant speed. The result is a sequence of coordinates that are represented by complex numbers. Figure 4 shows an example of object boundary.
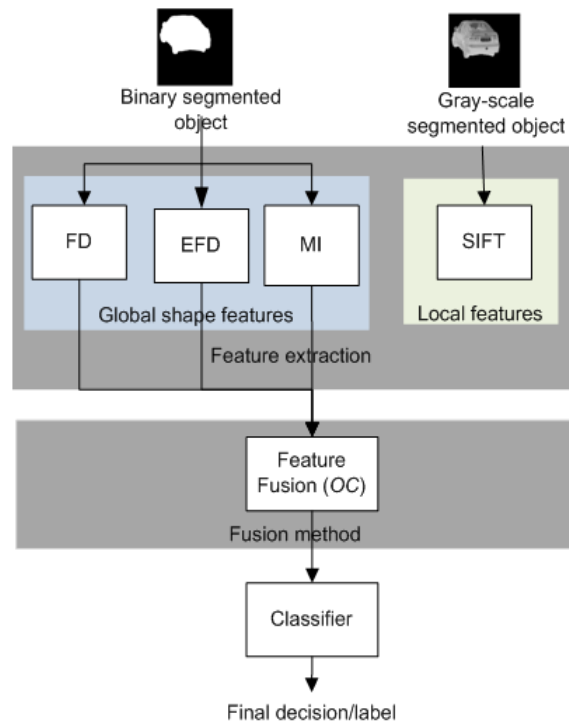

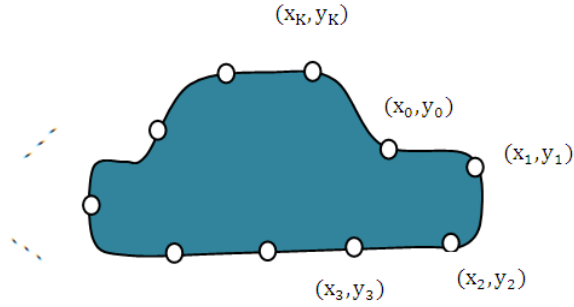Fig. 3: The feature fusion approach framework

78

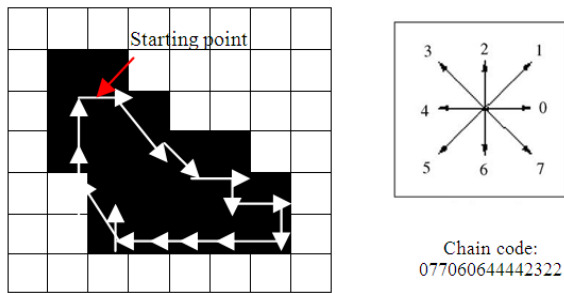Fig. 4: Boundary points extracted for car object

Fig. 5: (a) Shape extracted based on contour; (b) Differential chain code sequence at starting point generated from (a) using 8-connected

Dealing with discrete images, the Discrete Fourier Transform (DFT) is applied. The DFT of b(k) is defined as Eq. 3:

$$DFT(u) = \sum_{k=0}^{K-1} b(k) e^{\frac{j2\pi uk}{K}} \qquad (3)$$

where u = 0, 1, 2,…,K-1. The complex coefficient DFT(u) are called Fourier Descriptors of the boundary that gives the shape of an object. The inverse of Fourier transform of these coefficients will restore b(k) where k = 0, 1, 2,…,K-1 as shown in Eq. 4:

$$b(k) = \frac{1}{K} \sum_{u=0}^{K-1} DFT(u) e^{\frac{j2\pi uk}{K}} \qquad (4)$$

The inverse Fourier Descriptors is computed by specifying number the descriptors in order to obtain a closed spatial curve.

EFD: Similar to FD, EFD is applied to the closed contour of object based on the boundary information. The closed contour is defined with differential chain code, represented as a point coordinate of closed contour. Figure 5 illustrates the example contour of a binary image with its chain code generated from this image.

Based on Fig. 5, the length (dt_i) of element (v_i) of the chain code is given by Eq. 5:

$$dt_i = 1 + \left(\frac{\sqrt{2}-1}{2} \right)(1-(-1)^{v_i}) \qquad (5)$$

Therefore, for the whole number of element in a contour, the length is Eq. 6:

$$t_n = \sum_{i=1}^{n} dt_i \qquad (6)$$

The following equations present the projection of each v_i, on X and Y-axis, respectively Eq. 7:

$$dx_i = sign(6-v_i) * sign(2-v_i),$$
$$dy_i = sign(4-v_i) * sign(v_i) \qquad (7)$$

For all elements in the chain, p, the projection on X and Y-axis will be Eq. 8:

$$x_p = \sum_{i=1}^{p} dx_i, y_p = \sum_{i=1}^{p} dy_i \qquad (8)$$

EFD is calculated from the sum of elliptical harmonics. In identifying the closed contour points, K and N harmonics are considered. Kuhl and Giardina (1982) use four Fourier coefficients, a_n, b_n, c_n and d_n in each harmonic. Equation 9 presents these four coefficients. These harmonics and their corresponding coefficients are used to produce coordinates that define ellipses that fit within the object's outline to represent the object's shape.

$$a_n = \frac{T}{2n^2\pi^2} \sum_{i=1}^{k} \frac{dx_i}{dt_i} \left[\cos\frac{2n\pi t_i}{T} - \cos\frac{2n\pi t_{i-1}}{T} \right]$$

$$b_n = \frac{T}{2n^2\pi^2} \sum_{i=1}^{k} \frac{dx_i}{dt_i} \left[\sin\frac{2n\pi t_i}{T} - \sin\frac{2n\pi t_{i-1}}{T} \right]$$

$$c_n = \frac{T}{2n^2\pi^2} \sum_{i=1}^{k} \frac{dy_i}{dt_i} \left[\cos\frac{2n\pi t_i}{T} - \cos\frac{2n\pi t_{i-1}}{T} \right]$$

$$d_n = \frac{T}{2n^2\pi^2} \sum_{i=1}^{k} \frac{dy_i}{dt_i} \left[\sin\frac{2n\pi t_i}{T} - \sin\frac{2n\pi t_{i-1}}{T} \right]$$

$$(9)$$

Moment Invariants (MI): MI are shape features that have been succesfully used in pattern recognition

research such as in aircraft recognition (Sarfraz, 2006), object class recognition (Yuan and Hui, 2008), face recognition (Nabatchian *et al*., 2008) and handwriting recognition (Ramteke and Mehrotra, 2008). This features can be extracted from the boundary and interior region of an object. In this research, MI values are extracted from the segmented objects based on boundary points based on Hu (1962) who propose seven expressions to be calculated from normalized central moments that are invariant to object scales, translations and rotations. Hence, MI features used in this research is able to represent different geometrical features in input objects. MI may also be applied for disjoint shapes that cannot be supported by FD (Chen, 2003).

Local features: In object class recognition, each object will have a unique representation. However, this is difficult because the same object may be interpreted using many poses. One of the disadvantages concerning the chosen shape feature is the silhouette information, which may be insufficient and ambiguous. Similar silhouettes are often corresponds to different objects from different viewpoints. To overcome this, the study also considers using local features, SIFT to be combined with boundary-based shape features. SIFT is the best local features to recognize various objects in different views and scales, including blurry images as well as images with changes in lighting and translation (Mikolajczyk and Schmid, 2005; Lowe, 1999).

SIFT feature extraction employs the bag of keypoints approach (Csurka *et al*., 2004) that is based on vector quantization of the SIFT features extracted from the object. The difference-of-Gaussian is applied to identify the interest points of an object. The dimension of object's local features is based on the number of interest points generated by the region detector, which is usually 128-dimensional extracted from multiple interest points of the object patches. Once a multi-dimensional feature set has been extracted from an object, a clustering algorithm is performed to generate the visual vocabulary. In order to construct a bag of keypoints as the feature vector, the number of patches assigned to each cluster is calculated and the learning algorithm is applied to train this feature vector. The category of test data can be determined based on the model designed.

RESULTS
    The goal of this study is to investigate whether a fusion of different types of features in a single feature vector improves the performance in recognizing object

classes. Comparison of the proposed work is carried out against the recognition results from a single feature. For the purpose of this experiment, the Graz02 dataset is used because the objects included in the dataset are more realistic and are not limited by changes of poses, size, lighting, translation or illumination.

    Empirically, 40 descriptors of FD and 28 EFDs are used in this study. This number accurately describes the shape of objects. Figure 7 presents the error rates using SVM classifier for each object class for different number of FDs. The error rate improves slightly between 30 and 40 number of FDs. However, the error rate was increased to 0.02% in recognition error using more than 40 FDs. Therefore in this study, we state that 40 FDs present optimal descriptors for each object class in terms of accuracy.

    Following previous research (Opelt *et al*., 2006a), the SIFT features are clustered using *K*-Means algorithm with K = 100. The new feature vector as mentioned in Eq. 1, OC represents a single feature space with total 175-dimension and is then trained using the SVM binary classifier in order to model each class. For recognition, all features are extracted from testing data and the trained model is used to predict the final object class. The Radial Basis Function (RBF) kernel is applied with gamma, γ and cost, C parameters acquired using 10-cross validation approach. The size of training data and testing data is adopted from Opelt *et al*. (2006a). For SVM training, 150 positive samples and 150 negative samples are used. The total of testing sample is 150, where 75 are positive sample and 75 are negative samples. The negative samples consist of the remaining two concepts.
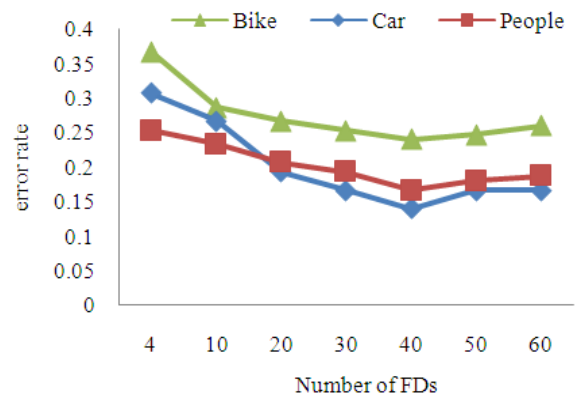

Fig. 7: Error rates for different number of FDs for bike, car and people class

Table 1: ROC rates of feature fusion compared to the results of single feature using Graz02 (Opelt *et al*., 2006a) dataset

Class	Global shape features Feature			Local feature SIFT	fusion FF
	FD	EFD	MI	SIFT	FF
Bike	0.873	0.889	0.853	0.564	0.954
Car	0.970	0.860	0.651	0.401	0.971
People	0.966	0.943	0.833	0.665	0.983

Table 2: Comparison of ROC equal rates with other works

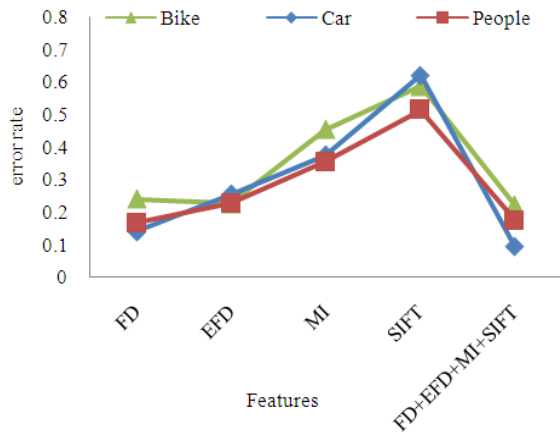|	Ours	Hegazy and Denzler (2008)	Opelt *et al*. (2006a)
Bike	0.954	0.747	0.778
Car	0.971	0.813	0.705
People	0.983	0.813	0.812

Fig. 8: Error rate of single features and combination features using feature fusion approach

All experiments in this study are evaluated using Receiver-Operating-Characteristic curve (ROC) for presenting results in recognition as shown in Table 1 This evaluation method can be a good measurement for recognition performance since it takes into account the difference between errors on positive and negative examples (Rakotomamonjy, 2004). The combination of shape-based features with local features (FF) yields improvement of classifications performance as opposed to using single feature only. From this result, Feature Fusion (FF) exceeds the ROC rates in all classes. Combination of boundary-based shape features and SIFT features has further increased the performance for bike class by 8.1% and car class by 0.1%, whereas the

performance of the people class about 1.7%. On contrary, boundary-based shape features alone averaged up to 85% of objects are correctly classified. However, for local feature, the performances on bike and car classifiers were not promising, which is only 56.4% and 40.1% correct classification, respectively. In Fig. 8, we present the error rates using SVM classifier for each object class for different features to give a clearer picture about the recognition performance of single feature with FF (FD+EFD+MI+SIFT). From this figure, the error rate was improved slightly using FD+EFD+MI+SIFT whereas, the error rate has increased significantly in recognition error when using SIFT solely.

Table 2 shows the result and comparison to the state-of-the art approaches in Opelt *et al*. (2006a) and Hegazy and Denzler (2008) which combined the variety of local features solely. We observe that combining global and local features improve the classification state-of-art features more than 15% for all object classes.

DISCUSSION

In recognizing objects, shape and local features play an important role in producing a successful recognition system to reduce the classification error. However, from our observation, the object classes in Graz02 (Opelt *et al*., 2006a) dataset do not have much information on their local features, hence it is not able to increase the recognition performance even though *SIFT* local features are robust to scales, viewpoints and illumination. Also, we also observed that shape features have high influence on the final decision in feature fusion approach even when local features of objects are limited.

CONCLUSION

This study proposed a combination of boundary-based shape features and local features using feature fusion technique with a binary SVM classifier. The experimental results on this challenging dataset show that the performance of feature fusion improved the classification accuracy as compared to using single feature. However, some drawbacks noted in this approach include the high-dimensional feature vector and contradictory information when there are too many different features combined. From this method, it is hard to identify which features are exactly relevant or impactful to the resulting accuracy since all features are represented in one feature space. To overcome this weakness and to reduce the computational time, further

research will explore the use of decision fusion methods in aggregating different types of features.

REFERENCES

Ali, J.M., 2007. Content-based image classification and retrieval: A rule-based system using rough sets framework. Int. J. Intell. Inform. Technol., DOI: 10.4018/jiit.2007070103

Arun, K.S. and H.P. Menon, 2009. Content based medical image retrieval by combining rotation invariant contourlet features and fourier descriptors. Int. J. Recent Trends Eng., 2: 35-39. http://ijrte.academypublisher.com/vol02/no02/ijrte02023539.pdf

Chen, Q., 2004. Evaluation of OCR algorithms for images with different spatial resolutions and noises. M.A. Sc., Thesis, University of Ottawa, http://www.worldcat.org/title/evaluation-of-ocr-algorithms-for-images-with-different-spatial-resolutions-and-noises/oclc/70812352

Csurka, G., C. Dance, L. Fan, J. Willamowski and C. Bray, 2004. Visual categorization with bags of keypoints. Proceedings of the Pattern Recognition and Machine Learning in Computer Vision Workshop, (PRMLCVW'04), Grenoble, France, http://videolectures.net/lmcv04_dance_vcbk/

Dimitrovski, I., D. Kocev, S. Loskovska and S. Džeroski, 2011. Hierarchical annotation of medical images. Patt. Recog., 44: 2436-2449. DOI: 10.1016/j.patcog.2011.03.026

Gonzalez, R.C., R.E. Woods and S.L. Eddins, 2004. Digital Image Processing using MATLAB. 3st Edn., Prentice Hall Upper Saddle River, NJ. http://www.amazon.com/Digital-Image-Processing-Using-MATLAB/dp/0130085197

Hare, J.S., P.A.S. Sinclair, P.H. Lewis, K. Martinez and P.G.B. Enser *et al*., 2006. Bridging the semantic gap in multimedia information retrieval: Top-down and bottom-up approaches. Proceedings of the 3rd European Semantic Web Conference (ESWC-06), June, 12 LNCS 4011, Springer Verlag. http://eprints.soton.ac.uk/262737/

Hegazy, D. and J. Denzler, 2008. Generic Object Recognition Using Boosted Combined Features. Proceedings of the 2nd International Conference on Robot Vision, (ICRV'08), Springer-Verlag, pp: 355-366. DOI: 10.1007/978-3-540-87442-3_36

Hu, M.K., 1962. Visual pattern recognition by moment invariants. IRE Trans. Inform. Theory, 8: 179-187. http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=1057692&url=http%3A%2F%2Fieeexplore.ieee.org%2Fiel5%2F4547527%2F22787%2F01057692.pdf

Jeong, Y. and R.J. Radke, 2007. Reslicing axially sampled 3D shapes using elliptic Fourier descriptors. Med. Image Anal., 11: 197-206. http://www.bibsonomy.org/bibtex/2762cb6348265f374ebf168bd35bd76a3/dblp

Kludas, J. and E. Bruno, 2008. Information fusion in multimedia information retrieval. Adaptive Multimedial Retrieval: Retrieval, User and Semantics, (RRUS'08), Book, pp: 147-159. DOI: 10.1007/978-3-540-79860-6_12

Kuhl, F.P. and C.R. Giardina, 1982. Elliptic Fourier features of a closed contour* 1. Comput. Graphics Image Process., 18: 236-258. DOI: 10.1016/0146-664X(82)90034-X

Lowe, D.G., 1999. Object recognition from local scale-invariant features. Proceedings of the International Conference on Computer Vision (ICCV'99), 2: 1150-1157. DOI: 10.1109/ICCV.1999.790410

Mansur, A. and Y. Kuno, 2007. Integration of multiple methods for robust object recognition. Proceedings of the SICE Annual Conference, Sept. 17-20 Saitama Univ, pp: 1990-1995. DOI: 10.1109/SICE.2007.4421313

Mikolajczyk, K. and C. Schmid, 2005. A performance evaluation of local descriptors. IEEE Trans. Patt. Anal. Mach. Intell., 27: 1615-1630. DOI: 10.1109/TPAMI.2005.188

Nabatchian, A., E. Abdel-Raheem and M. Ahmadi, 2008. Human face recognition using different moment invariants: A comparative study. IEEE, 3: 661-666. DOI: 10.1109/CISP.2008.479

Oliveira, L. and U. Nunes, 2008. On integration of features and classifiers for robust vehicle detection. Proceedings of the 11th International IEEE Conference on Intelligent Transportation Systems, Beijing, China. DOI: 10.1109/ITSC.2008.4732545

Opelt, A., A. Pinz, M. Fussenegger and P. Auer, 2006a. Generic object recognition with boosting. IEEE Trans. Patt. Anal. Mach. Intell., 28: 416-431. DOI: 10.1109/TPAMI.2006.54

Opelt, A. and A. Pinz *et al*., 2006b. Fusing shape and appearance information for object category detection. In BMVC. Edinburgh, UK, pp: 117-127.

Rakotomamonjy, A., 2004. Optimizing area under ROC curve with SVMs. ROC Anal. Artif. Intell., 71-80.

Ramteke, R.J. and S.C. Mehrotra, 2008. Recognition of handwritten devnagari numerals. Int. J. Comput. Process. Oriental Languages, pp: 9-16. http://bisite.usal.es/archivos/dcai%20reig.pdf

Reig-Bolano, R., P. Marti-Puig, S. Rodriguez, J. Bajo and V. Parisi-Baradad *et al*., 2010. Otoliths identifiers using image contours EFD. Distribut. Comput. Artif. Intell., 79: 9-16. http://www.springerlink.com/content/88726hhj1l54 1303/?MUD=MP

Sarfraz, M., 2006. Object recognition using moments: Some experiments and observations. Proceedings of the International Conference on Geometric Modeling and Imaging (GMAI), IEEE Computer Society. DOI: 10.1109/GMAI.2006.39

Schaefer, G., M. Zavisek and T. Nakashima 2009. Thermography based breast cancer analysis using statistical features and fuzzy classification. Patt. Recog., 42: 1133-1137. DOI: 10.1016/j.patcog.2008.08.00

Shotton, J., A. Blake and R. Cipolla, 2008. Multi-scale categorical object recognition using contour fragments. IEEE Trans. Patt. Anal. Mach. Intell., 30: 1270-1281. PMID: 18550908

Su, Z., H. Zhang and S.Ma, 2003. Relevance feedback in content-based image retrieval: Bayesian framework, feature subspaces and progressive learning. IEEE Trans. Image Process., 12: 924-937.PMID: 18237966

Sun, X. and M. Chen *et al*., 2009. Action recognition via local descriptors and holistic features. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Yang, M. and K. Kpalma *et al*., 2008. A survey of shape feature extraction techniques. Patt. Recog. Techn. Technol. Appli., 626-626.

Yu, X. and Y. Li *et al*., 2007. Object detection using shape codebook. Proceedings of the British Machine Vision Conference (BMVC'07), pp: 1-10.

Yuan, R. and W. Hui, 2008. Object identification and recognition using multiple contours based moment invariants. IEEE.

Zahn, C.T. and R.Z. Roskies, 1972. Fourier descriptors for plane closed curves. IEEE Trans. Comput., 21: 269-281.DOI: 10.1109/TC.1972.5008949

Zhang, D. and G. Lu, 2004. Review of shape representation and description techniques. Patt. Recog., 37: 1-19. http://vis.uky.edu/~cheung/courses/ee639_fall04/re adings/shapeReview.pdf.

Zhang, W. and B. Yu *et al*., 2005. Object class recognition using multiple layer boosting with heterogeneous features. Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), June, 20-25, IEEE Xplore Press, pp: 323-330. DOI: 10.1109/CVPR.2005.251

4.12.2012