WaSS: A Novel Hybrid Method for Object Recognition Using Wavelet based Statistical and Structural Approaches

G.Karuna[1]
Research Scholar
GRIET, Hyderabad

Dr.B.Sujatha[2]
Dept. of CSE
GIET, Rajahmundry

Dr.P.Chandrasekhar Reddy[3]
Professor
Dept. of ECE
JNTUH, Hyderabad

***Abstract* — Object recognition is one of the most important tasks in computer vision domain. In this paper, a novel method is proposed to recognize objects using shape information based on statistical and structural approaches. For the extraction of shape information, first decompose the original image, then real complement approach using rotations are proposed. Wavelets rearrange the shape of an object for reaching a desirable size. In addition, a set of statistical features are constructed, which can be used for object recognition. We have applied this method to representation and recognition of Flavia and Swedish leaf datasets. Experiments show that a combined statistical and structural approach is superior to other state-of-the-art methods.**

Keywords— Object recognition, shape, statistical features, real complement approach.

I. INTRODUCTION

Shape is an intrinsic feature for image understanding, which is stable to illumination and variations in object color and texture. Because of these advantages, shape is widely considered for object recognition. In particular, with the recent advance in contour detection proposed by Arbelaez et al. in [1], shape based object recognition in natural image is becoming more practical and attracts more attention in computer vision community. Main challenges in shape based object recognition include deformation, occlusion and viewpoint variation of objects. Various shape descriptors have been proposed to address these challenges, e.g., [2, 3, 4, 5]. Therefore, shape descriptors should be capable to deal with these problems in order to guarantee intra-class compactness and inter-class separability.

Object recognition has become one of the classic problems in Computer Vision. It is commonly divided into different sub-problems, such as segmentation, feature extraction, object representation, detection or classification, which are tackled with different techniques or from different points of view. In our case, we are interested in two of these problems: techniques related to object representation and, mainly, to feature extraction. In this sense, the description and identification of objects can be done using different visual cues such as shape, color or texture. Among them, shape is probably one of the most widely considered. A lot of methods of object recognition were explored based on shapes of objects [6–10].

Reviewing the literature, many shape descriptors, capable to deal with some of the problems, have been proposed. A survey on shape recognition can be found in [11]. These descriptors can be broadly classified into two kinds of categories: statistical and structural approaches. Statistical approaches use a feature vector derived from the image to describe the shape, for example, the curvature scale space (CSS) descriptor [12] uses the external contour for coding the shape, Zernike moments [13], is Shape Context [14], which is based on the relation between shape pixels SIFT descriptor [15] uses local information and has been mainly applied for object recognition. Statistical techniques used as classifiers within the classification task include those based on similarity (e.g., template matching, k-nearest neighbor), probability (e.g., Baye's rule), boundaries (e.g., decision trees, neural networks), and clustering (e.g., k-means, hierarchical). The quantitative nature of statistical pattern recognition, however, makes it difficult to discriminate among groups based on the morphological (i.e., shape based or structural) sub patterns and their interrelationships embedded within the data. This limitation provided the impetus for the development of a structural approach to pattern recognition.

Structural approaches use syntactic grammars to discriminate among objects belonging to different groups based upon the arrangement of their morphological (i.e., shape based or structural) features. The description task of a structural pattern recognition system is difficult to implement because there is no general solution for extracting structural features. In this approach, pattern to be recognized can be decomposed into simpler components called primitives and the relationships among them which constitute the data. For example, morphological operations are used to characterize shapes and textures. A structural pattern recognition system typically includes feature extractors to identify instances of morphological characteristics of the data which, in turn, are used as the basis for classification using syntactic grammars.

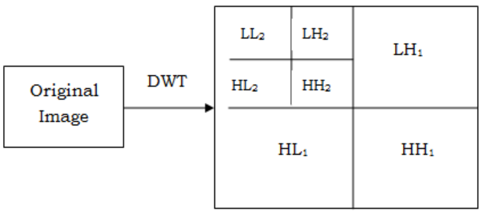
In this paper, we propose a novel hybrid method for object recognition using Wavelet based Statistical and Structural (WaSS) approaches. The proposed method represents a novel combination of two well renowned methods which allows for extraction of a concise set of shape descriptors that enable discrimination between objects with fine shape differences. Usually, structural features can be used with either a statistical or structural classifier. Statistical features cannot be used with a structural classifier because they lack relational information, however statistical information can be associated with structural primitives and used to resolve ambiguities during classification or embedded directly in the classifier itself.

The rest of this paper is organized as follows. In Section 2 we discuss Discrete Wavelet Transform (DWT). In Section 3 we present our novel methodology for exact representation of shapes and also describe our methodology to quantify differences between shape models for object classification based on shape descriptors. Section 4 describes the experimental design and the results of experimental evaluations. Finally, the work is concluded in Section 5.

II. DISCRETE WAVELET TRANSFORM

A wavelet is a mathematical function which divides a given function into different scale components. Actually the image is decomposed into four sub-bands and critically sub-sampled by applying Ist level DWT as shown in Fig.1. LH1, HL1 and HH1 represent the finest scale wavelet coefficients i.e., detail images while the sub-band LL1 corresponds to coarse level coefficients (approximate image). The 2nd level of wavelet coefficients are achieved by decomposing sub-band LL1 and sub-sampling it as shown in Fig.1. If further decompositions are needed, LL2 will be used. This process is continued until a final scale is reached. The values of the coefficients in approximation and detail sub-band images are the essential features uniquely characterize the shape of the object, which are shown here as useful for shape analysis and discrimination.

Fig.1: DWT at second level decomposition.

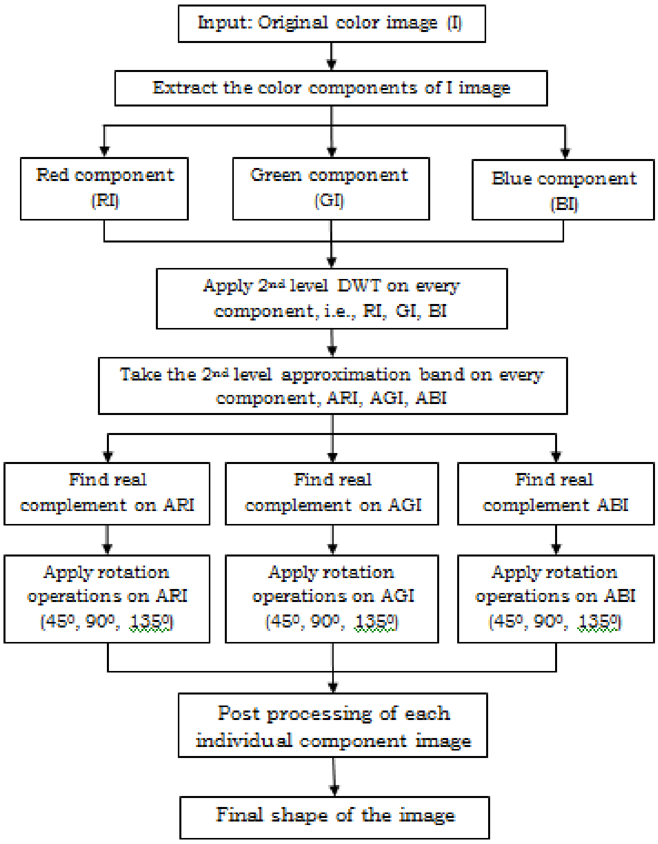


By applying DWT, the image is detached into multi-resolution representation with the aim of retaining the least coefficient possible without losing useful image information. The retained information is useful for shape analysis, namely, shape representation and classification, which can be discussed in the next section.

III. EXACT REPRESENTATION OF SHAPE BY WAVELET BASED REAL COMPLEMENT METHOD

The wavelet based real complement method using rotations is proposed for extracting edges from color images. The flowchart of the proposed method is given in Fig. 2. The process of real complement approach is discussed in Section 3.1, and rotation operations are discussed in Section 3.2.

Fig.1: Proposed shape representation using real complement approach.



A. *Real Complement Approach*

The real complement approach works like a low pas filter strongly attenuating the lower image frequencies. This method is based on image complement and the steps involved as follows:

Step 1: Read the original image (I) of size m×n.
Step 2: Find the complement of the image (I_c).
Step 3: Find the image difference between original and complement images i.e. I_1=I-I_c.
Step 4: Apply global threshold on the difference image $I_{1,}$ resulting image is I_2.
Step 5: Obtain the real complement of the image by differencing with I and I_2 i.e. I_3=I-I_2.

Image differencing method makes the background uniform and is depending on the intensity/ gray scale of the background pixels of the image. They categorized the images into three cases based on background information. The first case is background to very bright, i.e. the background of I_1 remains almost same since the complement of bright intensity values and is subtracted to from its original image leads to nearest intensity range. The second case is background to slightly dark i.e. background of the image is slightly dark and is analogous to first case but nearer to its original intensity range. The third case is background to intermediate range i.e. if the difference is a positive value, a similar gray value is retain otherwise it becomes black. Finally, real complement approach produces binary information of the image achieved by using thresholding method.

B. *Rotation Operations*

Literature reveals several edge detectors based on filters, derivatives, gradients, transforms etc. A simple method to detect edges has been introduced based on image rotations. The proposed method uses rotation operations for efficient detection of edges. Here the image is rotated based on different values of θ where θ ϵ {45∘, 90∘, 135∘}. The process of rotation operations are three steps.

Step 1: Read the original image (I) of size m×n.
Step 2: Apply rotation operations based on θ values.
Step 3: Obtain edge pixels by image differencing method.

In general, an image is represented as a matrix and is rotated in two types. The first type of rotation is to rotate on entire image and the second type of rotation is to rotate on blocks of the image and of size is 3×3 and 5×5 each. In the second type of rotation, two basic parameters are required for rotation of blocks of the image namely direction and the angle. The direction may be clockwise or anti-clockwise and the angles are represented by θ. The following figures Fig. 3.1(a), 3.1(b), 3.1(c), 3.1(d) shows the representation of image rotations after rotation of θ. In the present paper/thesis uses both cases i.e. on rotation on entire image and block wise image. Bilinear and Bicubic interpolation are the two approaches used to rotate the images. The function imrotate of the Matlab package is used to perform these interpolations. By observing the results obtained by two cases, block wise rotations are well formed compared to entire image rotation. This way of rotation of images guarantees that variation of pixels from neighboring data that means sharp variations and minute transitions can be captured effectively.

Fig. 3: Representation of image rotations after rotation of θ.

A	B	C
D	X	E
F	G	H

(a) Block of size 3×3

D	A	B
F	X	C
G	H	E

(b) 45^0 rotation

F	D	A
G	X	B
H	E	C

(c) 90^0 rotation

G	F	D
H	X	A
E	C	B

(d) 135^0 rotation

The following algorithm represents the procedure for finding rotated images of θ.

Step 1: Read the Original image (I) of size m×n.
Step 2: Apply 45^0 rotations in clockwise direction on 3×3 convolution of original image denoted as I_1.
Step 3: Find the difference between I and I_1 i.e I_2= I-I_1.
Step 4: Apply 45^0 rotations in anti-clockwise direction on 3×3 convolution of original image denoted as I_3.
Step 5: Find the difference between I and I_3 i.e I_4= I-I_3.
Step 6: To obtain edge pixel, add I_5=I_2+I_4.
This process is repeated for rotations of 90 and 135 degrees.

IV. EXPERIMENTAL RESULTS

In this section, we are going to show the performance of the results is presented on a Flavia shape dataset [16] and Swedish leaf dataset [17]. First of all, build an exact shape representation using wavelet based morphological structural approach, which can be implemented by real complement approach. In addition, we utilize DWT method to add spatial layout information of edge fragments in to our shape representation. By observing the results of standard dataset, the proposed real complement approach with rotation operations carry out an exact shape of the object based on edges. But some of the images don’t form an exact shape it requires additional processing and also some of them are not connected edges. For removing unwanted edge pixels pruning is the best method. After pruning morphological erosion is required for exact shape of the image.

For object recognition a set of statistical shape features are evaluated. Then we use, leave-one-object-out cross-validation [18] test for the correctness of objects. For evaluating object recognition/classification performance three different situations are considered:

type-1: half training and half testing (without rotation) i.e randomly select half of the shapes in each class for training and the rest shapes for testing in each round;

type-2: 75% training and 25% testing (with rotation angle of 0^0, 45^0, 90^0, 135^0 and, 180^0)

type-3: leave-one-out, for each shape, i.e., we use all shapes except the current one for training and use the current one for testing;

By considering the three situations (type-1, type-2, type-3), a K-NN classifier is used for finding average classification accuracy. In addition, standard derivation of classification accuracies is reported in Table 1. We compare classification accuracy of our WaSS method with other shape-based recognition methods [3, 17, 19, 20] in Table 2. The proposed WaSS out performs [3, 17, 19, 20] by over 3% when using type-1 classification i.e., half for training and half for testing. In addition, very recent literatures [21, 22], is compared with WaSS, show good performance than the most recent results in [21, 22] as shown in Table 3.

Table 1: Classification accuracy of the proposed WaSS method on Flavia [16] and Swedish leaf dataset [17]

Classification type	Dataset	Recognition/classification accuracy (%)	Standard Deviation
Type-1	Flavia	98.28	3.8
	Swedish	96.50	5.6
Type-2	Flavia	93.80	13.5
	Swedish	94.55	139
Type-3	Flavia	90.75	21.2
	Swedish	91.65	19.6

Table 2: Comparison of classification accuracy of WaSS method with other shape recognition methods on Swedish leaf dataset [17].

Method	Classification accuracy (%)
Moment+ Area+ Curvature [17]	82
Fourier [3]	89.6
SC+DP [3]	88.12
IDSC+DP [3]	94.13
IDSC+ Morphologic Strategy [19]	94.80
Shape tree [20]	96.28

Table 3: Comparison of classification accuracy of WaSS method with most recent methods on Swedish leaf dataset [17].

Method	Classification accuracy (%)
Shape Context [21]	96.9
Bag of Contours [22]	97.6

V. CONCLUSIONS

In this paper, we present a novel shape representation called wavelet based real complement approach for object recognition. To the best of our knowledge, this is the first paper that introduces the idea of WaSS for object recognition. The proposed method consists of two steps. First, represent the shape of the object. Next, construct the feature vector for classification of objects based on extracted shapes. Finally compute the classification accuracy of the standard datasets. In the experiments, we have extensively tested the performance of WaSS; all these experimental results on shape bench marks show that WaSS is able to achieve the state-of-the-art performance; moreover the proposed method is very effective in recognizing objects with large viewpoint changes while only a small number of training images are required to build an effective object recognition system.

Acknowledgment

The authors would like to express their gratitude to chairman K.V.V. Satya Narayana Raju, MLC and chairman, K.Sasikiran Varma, Managing Director GIET for providing advanced research lab facilities through Srinivasa Ramanujan Research Forum GIET.

References

[1] P. Arbelaez, M. Maire, C. Fowlkes, J. Malik, “Contour detection and hierarchical image segmentation”, IEEE Trans. Pattern Anal. Mach. Intell., Vol.33, pp: 898–916, 2011.

[2] S. Belongie, J. Malik, J. Puzicha, “Shape matching and object recognition using shape contexts, IEEE Trans. Pattern Anal. Mach. Intell,. Vol.24, pp: 509–522, 2002.

[3] H. Ling, D. Jacobs, “Shape classification using the inner-distance”, IEEE Trans. Pattern Anal. Mach. Intell., Vol.29, pp: 286–299, 2007.

[4] F. Mokhtarian, S. Abbasi, J. Kittler, etal., “Efficient and robust retrieval by shape content through curvature scale space”, Ser. Softw. Eng. Knowl. Eng., Vol.8, pp: 51–58, 1997.

[5] N. Alajlan, I. ElRube, M. Kamel, G. Freeman, “Shape retrieval using triangle-area representation and dynamic space warping”, Pattern Recognition, Vol.40, pp: 1911–1920, 2007.

[6] M.R. Daliri, V. Torre, “Classification of silhouettes using contour fragments”, Computer Vision and Image Understanding, Vol. 113, pp: 1017–1025, 2009.

[7] M.R. Daliri, V. Torre, “Shape recognition based on kernel-edit distance”, Computer Vision and Image Understanding, Vol. 114, pp: 1097–1103, 2010.

[8] X. Bai, X. Yang, L.J. Latecki, W.Liu, Z.Tu, “Learning context-sensitive shape similarity by graph transduction”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.32, Iss.5, pp: 861–874, 2010.

[9] X. Bai, B.Wang, X.Wang, W.Liu, Z.Tu, “Co-transduction for shape retrieval”, IEEE Transactions on Image Processing, Vol. 21, pp: 2747–2757, 2012.

[10] X. Shu, X.Wu, “A novel contour descriptor for 2D shape matching and its application to image retrieval”, Image and Vision Computing, Vol. 29, pp:286–294, 2011.

[11] D. Zhang, G. Lu, “Review of shape representation and description techniques”, Pattern Recognition, Vol.37, pp: 1–19, 2004.

[12] F. Mokhtarian, A. Mackworth, “Scale-based description and recognition of planar curves and two-dimensional shapes”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.8 pp: 34–43, 1986.

[13] A. Khotanzad, Y. Hong, “Invariant image recognition by zernike moments”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.12, pp: 489–497, 1990.

[14] S. Belongie, J. Malik, J. Puzicha, “Shape matching and object recognition using shape contexts”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.24, pp:509–522, 2002.

[15] D. Lowe, “Distinctive image features from scale-invariant key points”, International Journal of Computer Vision, Vol.60, pp: 91–110, 2004.

[16] Wu G.S., Bao F.S., Xu E.Y., Wang Y.X., Yi-Fan Chang Y.F., Xiang Q.L., “A Leaf Recognition Algorithm for Plant classification Using Probabilistic Neural Network”, IEEE International Symposium on Signal Processing and Information Technology, 2007.

[17] O.Söderkvist, “Computer vision classification of leaves from Swedish trees”, (Ph.D. thesis), Linköping, 2001.

[18] B.Leibe, B.Schiele, “Analyzing appearance and contour based methods for object categorization”, in: IEEE Conference on Computer Vision and Pattern Recognition, Vol.2, pp:II–409, 2003.

[19] R. Hu, W.Jia, Y.Zhao, J.Gui, “Perceptually motivated morphological strategies for shape retrieval, Pattern Recognition, pp: 105–111, 2012.

[20] P.Felzenszwalb, J.Schwartz, “Hierarchical matching of deformable shapes”, in: IEEE Conference on Computer Vision and Pattern Recognition, pp:1–8, 2007.

[21] X.Wang, X.Bai, W.Liu, L.J.Latecki, “Feature context for image classification and object detection”, in: IEEE Conference on Computer Vision and Pattern Recognition, pp.961–968, 2011.

[22] Xinggang Wang, Bin Feng, Xiang Bai, Wenyu Liu, Longin Jan Latecki, “Bag of contour fragments for robust shape classification” Pattern Recognition, Vol.47, pp:2116–2125, 2014.

[1]G. Karuna is currently working as Assistant Professor in Computer Science and Engineering Department, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, A.P., India. She is working towards her Ph.D. at Jawaharlal Nehru Technological University, Hyderabad, A.P., India. She received her M.Tech. in Computer Science and Engineering from Jawaharlal Nehru Technological University, Hyderabad, India in 2010. She has eight years experience of teaching undergraduate students and post graduate students. She has published 5 research papers in National and International journals. Her research interests are in the areas of Image Processing, Image Watermarking, Cryptography and Network Security.

[2]Dr.B.Sujatha received the B.Tech. degree from JNT University, Kakinada in 1997 and received her M.Tech. (Computer Science & Engineering) from Andhra University in 2002. She received her Ph.D in t he area of Digital Image Processing in Computer Science & Engineering from University of Mysore in 2013. She is having 12 years of teaching experience. Presently she is working as an Professor in GIET, Rajahmundry. She has published 21 research publications in Inter National Journals. She is a member of SRRF-GIET, Rajahmundry. Her research interest includes Image Processing and Pattern Recognition.

[3]Dr.P.Chandrasekhar Reddy, working as Professor in coordination, Department of ECE, JNTUH, Hyderabad. He received his Ph.D in the area of Computer Networks in Computer Science & Engineering from JNTU, Hyderabad. He has twenty years of experience in teaching and guided 10 Ph.D students and more than 100 M.Tech. students for various projects. He has published 30 National and 26 International Journals. His area of interests are Image Processing, Computer Networks and Network Security.