Fusion of Empirical Wavelet Features for Object Recognition

Murugan S[1], Dr Anjali Bhardwaj[2], Dr Ganeshbabu TR[3]

[1]Research scholar, [2]Associate Professor, [3]Professor

[1,2]Department of ECE, Maharishi University of Information Technology, Lucknow, India

[3]Department of ECE, Muthayammal Engineering College, Rasipuram, India

Abstract— *In this paper, an approach to recognize object efficiently is presented based on empirical wavelet features. In many computer vision applications, object recognition is required and it is a challenging task due to size and orientation of objects in the image. The proposed approach uses Empirical Wavelet Transform (EWT) to extract the characteristic of objects in an image. From the components of EWT, energy and entropy features are extracted. Then K-nearest neighbor classifier is used to recognize the object in the given image. The results show that the fusion of energy and entropy features provides better classification accuracy of 99.81% where the energy and entropy features provide 98.42% and 98.97% respectively on the benchmark object database named Columbia Object Image Library Dataset (COIL-100).*

Keywords— Object recognition, Empirical wavelet transform, energy features, entropy features, KNN classifier.

I. INTRODUCTION

Humans can easily recognize objects of varying size, shape, and orientations. However, it is a challenging task for computers. To ease object recognition, many automated approaches are developed recently. Some of them are addressed in this section. Learning strategy that models membership functions of the fuzzy attributes of surfaces is employed using GA [1] for object recognition. The objective function aims at enhancing recognition performance in terms of maximizing the degree of discrimination among classes. It is composed of three stages: retrieval and feature extraction of number of local parts from each model object, modeling the objects by feature vectors and similarity measurement.

A group-sensitive multiple kernel learning technique is used for object recognition to accommodate the inter-class correlation and intra-class diversity in [2]. A midway representation between the individual images and object category is obtained. An optimization model to concurrently perform kernel dictionary learning and prototype selection is discussed in [3]. The representation matrix is implemented to ensure that only a few samples are actually used to reconstruct the dictionary. So a convergent algorithm is employed to resolve the formulated non-convex optimization problem.

Context model based object recognition is discussed in [4]. It gives an efficient model that captures the information for more than a hundred object categories using a tree structure. It improves the performance of the system and also a coherent interpretation of a scene is obtained. Data driven un-falsified control is implemented for solving the drawbacks in visual servoing for object recognition in [5]. It recognizes an object through matching image features. Supervisory visual servoing is implemented until an accord between the model and the selected features is achieved, so that model recognition and object tracking are done successfully.

Multiple kernel learning (MKL) is an approach for selecting and combining kernels functions for a given recognition task. For solving the optimization problems, the state of MKL including different formulations and algorithms are discussed in [6] which focus on their applications to object recognition. Partial object recognition based on the corner point effective mapping is discussed in [7]. The features are extracted by using the corner point analysis. Then, neural network is used for recognition.

A self adaptive module is discussed in [8] for object recognition. It consists of one selector and four passes. Among the four passes, two are direct passes; one residual pass and one maxout pass with different receptive fields and depths. And the selector is designed to help the user to choose reasonable output. A prototype robot is designed for pick and place an object in [9]. Image processing concepts are used for recognition using arduino and MATLAB.

In this paper, an object recognition approach is presented based on EWT and KNN classifier. The organization of the paper is as follows: The mathematical background of EWT is given in Section 2 and the next section presents the proposed object recognition system. The results obtained by the proposed system using KNN classifier are discussed in section 4 and the conclusion is made in the final section.

II. EMPIRICAL WAVELET TRANSFORM

Unlike in Fourier and wavelet transform, the basis filters of EWT are not predefined and are a signal dependent method [10]. It is based on the information content in the given image or signal. The Fourier spectrum in the range 0 to π is segmented into *M* number of parts. Band pass filters in each segment defines the empirical wavelets. Littlewood-Paley and Meyer's wavelets are used as a bandpass filters with the empirical scaling function $\xi_m(W)$ and the empirical wavelets $\varsigma_m(W)$ can be described as

$$\xi_m(W)=\begin{cases}1 & if\,|W|\le(1-\lambda)\omega_m\\ \cos\left[\frac{\pi}{2}F(\lambda,\omega_m)\right] & if\,(1-\lambda)\omega_m\le|W|\le(1+\lambda)\omega_m\\ 0 & otherwise\end{cases}\tag{1}$$

$$\xi_m(W)=\begin{cases}1 & if\,(1+\lambda)\omega_m\le|W|\le(1-\lambda)\omega_{m+1}\\ \cos\left[\frac{\pi}{2}F(\lambda,\omega_{m+1})\right] & if\,(1-\lambda)\omega_{m+1}\le|W|\le(1+\lambda)\omega_{m+1}\\ \sin\left[\frac{\pi}{2}F(\lambda,\omega_m)\right] & if\,(1-\lambda)\omega_m\le|W|\le(1+\lambda)\omega_m\\ 0 & otherwise\end{cases}\tag{2}$$

where $F(\lambda,\omega_m)$ and $F(\lambda,\omega_{m+1})$ can be expressed as

$$F(\lambda,\omega_m)=F\left(\frac{1}{2\lambda\omega_m}\left(|W|-(1-\lambda)\omega_m\right)\right)$$

$$F(\lambda,\omega_{m+1})=F\left(\frac{1}{2\lambda\omega_{m+1}}\left(|W|-(1-\lambda)\omega_{m+1}\right)\right)\tag{3}$$

The $F(z)$ satisfies the following criteria,

$$F(z)=\begin{cases}0 & if\ z\le 0\\ 1 & if\ z\ge 1\\ F(z)+F(1-z)=1 & \forall z\in[0,1]\end{cases}\tag{4}$$

The EWT decomposition on 2D images [11] is described as follows. Let *x* denotes the image and the EWT decomposition consists of the following steps;

Step 1: Compute 1D Fourier transform of each row r of x; X(r ; Ω) and columns c of x; X(Ω ; c) and calculate the *mean* row and column spectrum magnitudes as follows:

$$X_R=\frac{1}{N_{Rw}}\sum_{r=0}^{N_{Rw}}X(r,\Omega)$$

$$X_c=\frac{1}{N_{C1}}\sum_{C=0}^{N_{C1}}X(\Omega,c)\tag{5}$$

Where number of rows and columns are denoted by N_{RW} and N_{Cl} respectively.

Step 2: Perform boundaries detection on X_R and X_C and build the corresponding filter bank $\left\{\xi_1^R, \left\{\varsigma_m^R\right\}_{m=1}^{N_R}\right\}$ *and* $\left\{\xi_1^C, \left\{\varsigma_m^C\right\}_{m=1}^{N_C}\right\}$ respectively. N_R and N_C are the number of *mean* row and column sub-band respectively.

Step 3: Filter *x* along the rows $\left\{\xi_1^R, \left\{\varsigma_m^R\right\}_{m=1}^{N_R}\right\}$ which provides (N_R+1) output images.

Step 4: Filter (NR+1) output images along the columns with $\left\{\xi_1^C, \left\{\varsigma_m^C\right\}_{m=1}^{N_C}\right\}$ this provides (N_R+1) (N_C+1) sub-band images.

In this study, EWT is used as a feature extraction technique. EWT is used for the diagnosis of glaucoma in medical image processing [12] and several extensions for the 1D adaptive wavelet frames to 2D signals (images) for EWT is explained in [11].

III. PROPOSED SYSTEM

The main aim of the proposed system is to recognize object in an image efficiently. It is composed of two computational blocks; feature or information extraction, and classification. An excellent feature extraction step will improve the accuracy of any recognition or classification system. This block produces set of salient features that represents the information required for the next stage. Figure 1 shows the feature or information extraction stage of the proposed EWT based object recognition system.

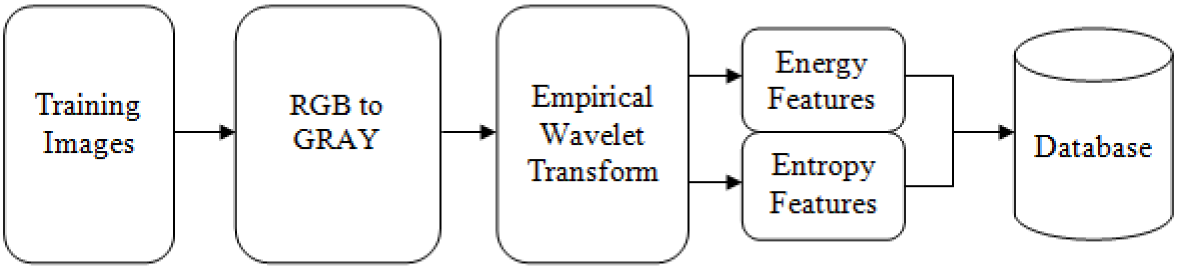

FIGURE 1 FEATURE OR INFORMATION EXTRACTION STAGE OF THE PROPOSED EWT BASED OBJECT RECOGNITION SYSTEM

EWT is a signal dependent decomposition technique and widely used in image processing applications such as medical image classification in [12]. However, EWT have not been studied for object recognition. Hence, the proposed system uses EWT as feature extraction technique. From the components of EWT, energy and entropy features are extracted and fused to form the feature vector. The energy and entropy features are defined in the following equations:

$$EWT_k = \frac{1}{RC}\sum_{i=1}^{R}\sum_{j=1}^{C}\left|x_k(i,j)\right| \tag{6}$$

where x_k is the k^{th} component of EWT decomposed image. *R* and *C* are the height and width of the image.

$$\text{Shannon entropy} = -\sum_i C_i^2 \log\left(C_i^2\right) \tag{7}$$

$$\text{Log Entropy} = \sum_i \log\left(C_i^2\right) \tag{8}$$

$$\text{Sure Entropy} = \left|C_i\right| \le \varepsilon \rightarrow e(s) = \sum_i \min\left(C_i^2, \varepsilon^2\right) \tag{9}$$

where C_i is the coefficients of a particular component *i* with log(0) = 0 and ε is a positive threshold. It is obtained using the principle of Steins unbiased risk estimate [13]. These features belongs to the same objects are extracted, grouped and used in the classification stage. Figure 2 shows the classification stage of the proposed EWT based object recognition system.

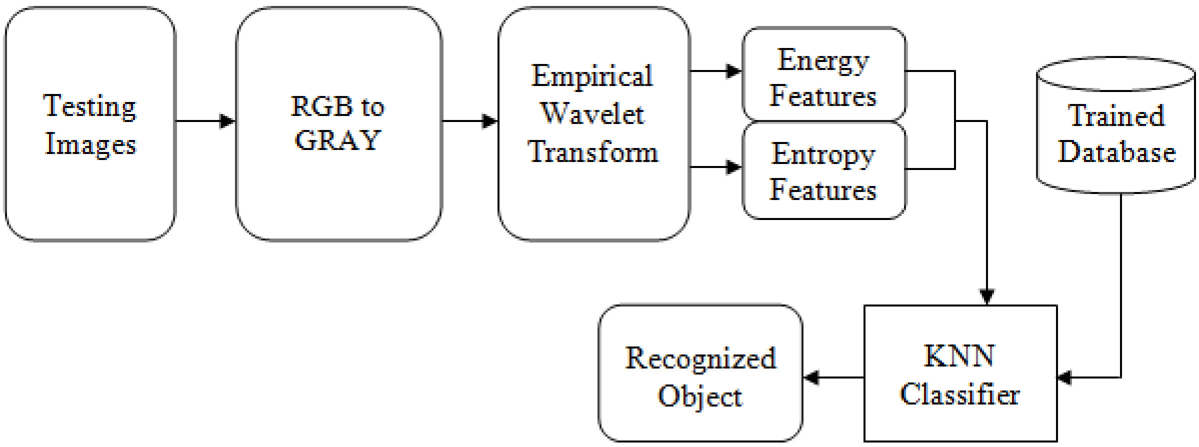

FIGURE 2 CLASSIFICATION STAGE OF THE PROPOSED EWT BASED OBJECT RECOGNITION SYSTEM

Classification is the final stage of the proposed system. The same technique used for extracting features of training images is applied to the testing image. KNN classifier is used for the classification. It is instance based classifier and hence there is no need for separate training stage. For testing, the database obtained from the feature extraction phase is given as one of the input the classifier. K-nearest neighbor classification is performed by finding K nearest neighbors in the feature space defined by the given training feature database. Each neighbor votes on the classification of the unknown object. Each vote may be counted equally or more priority may be given to votes of the closest neighbors. It computes the Euclidean distance between the testing objects features with the database. The identity of object which has the minimum distance is retuned by the classifier. The Euclidean distance measures calculation is as follows: Let us consider $u=(x_1,y_1)$ and $v=(x_2,y_2)$ are two points. The Euclidean distance between these two points is given by

$$Euclidean\ distance\ (u,v)=\sqrt{(x_1-y_1)^2+(x_2-y_2)^2} \tag{10}$$

If the points have n-dimensions such as $u=(x_1,x_2,x_3,\ldots x_n)$ and $v=(y_1,y_2,y_3,\ldots y_n)$ then the generalized Euclidean distance formula between these points is

$$Euclidean\ distance\ (u,v)=\sqrt{(x_1-y_1)^2+(x_2-y_2)^2+\ldots+(x_n-y_n)^2} \tag{11}$$

IV. RESULTS AND DISCUSSION

To measure the performance of the proposed object recognition system, COIL-100 [14] database images are taken. Figure 3 shows the sample objects in the COIL database. It consists of 100 objects of 128 x128 pixels resolution.

FIGURE 3 COIL DATABASE

Each object has 72 images captured using CCD colour camera with a 25 mm lens at every 5 degrees of rotation. Hence, 7200 images are available for the analysis. For training and testing purpose, the database is divided into two sets. It is based on the turn table rotation. In this study, six predefined turntable rotations such as 10, 20, 30, 45, 60, and 90 are used for making training images and their corresponding testing images are tested by the proposed system. The classification accuracy is used to analyze the performance of the system. It is defined by

$$Accuracy(\%) = \frac{Number\,of\ correctly\,classified\,objects}{total\,number\,of\ objects\,tested} x100 \tag{12}$$

The accuracy of each object is computed using the above formula and average classification accuracy of 100 objects is obtained. Tables 1 to 3 show the performances of the EWT based object recognition system using energy, entropy, and fusion of both features respectively.

TABLE 1
PERFORMANCE OF THE EWT BASED OBJECT RECOGNITION SYSTEM USING ENERGY FEATURES

EWT Decomposition Level	Average accuracy in percentage (%)					
	10^0	20^0	30^0	45^0	60^0	90^0
2	86.64	77.72	71.85	65.63	56.77	50.44
3	86.78	78.50	73.00	66.19	59.79	52.25
4	89.31	80.65	75.98	69.23	62.92	57.37
5	94.92	88.24	82.13	73.81	67.52	60.25
6	98.42	93.50	88.68	82.41	75.71	68.31

TABLE 2
PERFORMANCE OF THE EWT BASED OBJECT RECOGNITION SYSTEM USING ENTROPY FEATURES

EWT Decomposition Level	Average accuracy in percentage (%)					
	10^0	20^0	30^0	45^0	60^0	90^0
2	87.19	78.09	72.18	65.94	57.08	50.74
3	87.33	78.87	73.33	66.50	60.09	52.54
4	89.86	81.02	76.32	69.55	63.23	57.66
5	95.47	88.61	82.47	74.13	67.82	60.54
6	98.97	93.87	89.02	82.72	76.02	68.60

TABLE 3
PERFORMANCE OF THE EWT BASED OBJECT RECOGNITION SYSTEM USING THE FEATURE FUSION OF ENERGY AND ENTROPY FEATURES

EWT Decomposition Level	Average accuracy in percentage (%)					
	10^0	20^0	30^0	45^0	60^0	90^0
2	88.03	78.65	72.68	66.41	57.53	51.18
3	88.17	79.43	73.83	66.97	60.55	52.99
4	90.69	81.57	76.82	70.02	63.68	58.10
5	96.31	89.17	82.97	74.59	68.27	60.99
6	99.81	94.43	89.52	83.19	76.47	69.04

It is observed from tables 1 to 3 that the fusion approach produces 99.81% accuracy which is higher than the accuracy of energy and entropy features. The average accuracy obtained by the proposed approach using energy and entropy features is 98.42% and 98.97% respectively. For all cases, the maximum average accuracy is obtained at higher level EWT decomposition i.e., at 6th level. Table 4 shows the individual objects accuracy obtained by the EWT based object recognition system using fusion approach

TABLE 4
INDIVIDUAL OBJECTS ACCURACY OBTAINED BY THE EWT BASED OBJECT RECOGNITION SYSTEM USING FUSION APPROACH

#Object	Accuracy (%)	#Object	Accuracy (%)	#Object	Accuracy (%)	#Object	Accuracy (%)
1	100.00	26	100.00	51	100.00	76	100.00
2	100.00	27	100.00	52	100.00	77	100.00
3	100.00	28	100.00	53	100.00	78	100.00
4	100.00	29	100.00	54	100.00	79	100.00
5	100.00	30	100.00	55	100.00	80	100.00
6	100.00	31	100.00	56	100.00	81	100.00
7	100.00	32	100.00	57	100.00	82	100.00
8	100.00	33	100.00	58	100.00	83	100.00
9	100.00	34	100.00	59	100.00	84	97.22
10	100.00	35	100.00	60	100.00	85	100.00
11	100.00	36	100.00	61	100.00	86	100.00
12	100.00	37	100.00	62	100.00	87	100.00
13	100.00	38	100.00	63	100.00	88	100.00
14	100.00	39	100.00	64	100.00	89	100.00
15	100.00	40	100.00	65	100.00	90	100.00
16	100.00	41	100.00	66	100.00	91	97.22
17	100.00	42	100.00	67	94.44	92	100.00
18	100.00	43	100.00	68	100.00	93	100.00
19	100.00	44	100.00	69	94.44	94	100.00
20	100.00	45	100.00	70	100.00	95	100.00
21	97.22	46	100.00	71	100.00	96	100.00
22	100.00	47	100.00	72	100.00	97	100.00
23	100.00	48	100.00	73	100.00	98	100.00
24	100.00	49	100.00	74	100.00	99	100.00
25	100.00	50	100.00	75	100.00	100	100.00
Average							99.81

It is inferred from table 4 that among the 100 objects in the COIL database only 5 objects are misclassified. Also, the accuracy of each object is over 94%. It is concluded that for better classification of different objects in the COIL database, the 6th level energy and entropy features are selected by the proposed system.

V. CONCLUSION

In this paper, an approach for the recognition 100 objects in the COIL-100 database is presented using EWT and KNN. As EWT gives a better approximation of images than DWT, it produces an excellent performance for object recognition. From the EWT decomposed images, energy and entropy features are extracted. They are fused together and given to the KNN classifier for classification. Experimental results show that the proposed fusion approach produces 99.81% accuracy. Also, it is clearly observed that the fusion of energy and entropy features gives the highest accuracy than their individual counterpart.

REFERENCES

[1] Liu, Y. H., Lee, A. J and Chang, F., "Object Recognition using Discriminative Parts", Computer Vision and Image Understanding, pp. 854-867, 2012.

[2] Yang, Jingjing, et al., "Group-Sensitive Multiple Kernel Learning for Object Recognition", IEEE Transactions on Image Processing, vol. 21, no. 5, pp. 2838-2852, 2015.

[3] Liu, Huaping, and Fuchun Sun, "Online Kernel Dictionary Learning for Object Recognition", IEEE International Conference on Automation Science and Engineering, pp. 268-273, 2016.

[4] Choi, Myung Jin, Antonio Torralba, and Alan S. Willsky, “A Tree-Based Context Model for Object Recognition", IEEE transactions on pattern analysis and machine intelligence, vol. 34, no. 2, pp. 240-252, 2012.

[5] Jiang, Ping, et al., "Unfalsified Visual Servoing for Simultaneous Object Recognition and Pose Tracking", IEEE Transactions on Cybernetics, vol. 46, no. 12, pp. 3032-3046, 2016.

[6] Bucak, Serhat S., Rong Jin, and Anil K. Jain., "Multiple Kernel Learning for Visual Object Recognition: A Review", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, no. 7, pp. 1354-1369, 2014.

[7] Poonam and Sharma, M., "A Corner Feature Adaptive Neural Network Model for Partial Object Recognition", 4th International Conference on Reliability, Infocom Technologies and Optimization, pp. 1-6, 2015.

[8] Wang, Zhenyang, Zhidong Deng, and Shiyao Wang., "SAM: A Rethinking of Prominent Convolutional Neural Network Architectures for Visual Object Recognition", IEEE International Joint Conference on Neural Networks, pp. 1008-1014, 2016.

[9] Sahu, Uma, et al., "LIBO: The Grasping Robot Using Object Recognition", International Conference on Electrical, Electronics, and Optimization Techniques, pp.1-6, 2016.

[10] Gilles, J., “Empirical Wavelet Transform”, IEEE Transactions on Signal Processing, vol. 61, no. 16, pp. 3999–4010, 2013.

[11] Gilles, J., Tran, G., and Osher, S., “2D Empirical Transforms: Wavelets, Ridgelets, and Curvelets Revisited”, SIAM Journal on Imaging Sciences, vol. 7, no. 1, pp. 157–186, 2014.

[12] Maheshwari, Shishir, Ram Bilas Pachori, and Rajendra Acharya, U., "Automated Diagnosis of Glaucoma using Empirical Wavelet Transform and Correntropy Features Extracted from Fundus Images", IEEE Journal of Biomedical and Health Informatics, vol.99, pp. 2168-2194, 2016.

[13] Donoho, D.L. and Johnstone, I.M., “Adapting to Unknown Smoothness via Wavelet Shrinkage”, Journal of the American Statistical Association, vol. 90, no. 432, pp. 1200-1224, 1995.

[14] COIL Database: http://www.cs.columbia.edu/CAVE/software/softlib/coil-100.php