Research Journal of Applied Sciences, Engineering and Technology 8(13): 1613-1617, 2014
DOI:10.19026/rjaset.8.1141
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: September 07, 2014 Accepted: October 01, 2014 Published: October 05, 2014

Research Article

Object Recognition Based on Wave Atom Transform

[1]Thambu Gladstan and [2]E. Mohan
[1]Department of ECE, Shri JJT University, Jhunjhunu, Rajasthan, India
[2]P.T. Lee Chengalvaraya Naicker College of Engineering and Technology, Oovery, Kanchipuram, Tamilnadu, India

Abstract: This study presents an efficient method for recognizing object in an image based on Wave Atom Transform (WAT). Object recognition is achieved by extracting the energies from all coefficients of WAT. The original image is decomposed by using the WAT. All coefficients are considered as features for the classification process. The extracted features are given as an input to the K-Nearest Neighbor (K-NN) classifier to recognize the object. The evaluation of the system is carried on using Columbia Object Image Library Dataset (COIL-100) database. The classification performance of the proposed system is evaluated by using classification rate in percentage, which is achieved by varying the angle between the views.

Keywords: Classification rate, feature extraction, KNN classifier, object recognition, wave atom transform

INTRODUCTION

Object recognition in a real world atmosphere is still a very demanding task. In computer vision, object recognition is the task of finding a given object in an image or video sequence. Large amount of objects in images are recognized by human with little effort, despite the fact that the image of the objects may vary by different viewpoints, different sizes/scale or even when they are translated or rotated. This task is still a challenge for computer vision systems. The same object can have many different appearances if viewed from different angles or under different illumination conditions.

LITERATURE REVIEW

A confirmation and rotation invariant object recognition is proposed in Kim *et al*. (2012). By using DoG filter and local adaptive binarization, a binary image reserving spotless object margins is achieved. An object region from surroundings is extracted with compensated edges that reserves geometry information of object. Neural network is used to recognize the object. In order to approximate the pose of recognized object, which is handled by a robot standard shape model, represented object class, is used.

A new data set with images that contain many instances of different object categories and propose an expert model that captures the contextual information among more than a hundred object categories using a tree arrangement is proposed in Choi *et al*. (2012). This mold incorporates global image features, dependencies between object categories and outputs of local detectors into one probabilistic framework. This context model improves object recognition performance and provides a coherent detail of a scene, which enables a responsible image querying system by multiple object categories.

A novel method for object category recognition by civilizing the popular Bag-of-Words (BoW) methods from the following two aspects is proposed in Wang *et al*. (2010). Firstly, to extract a series of high level visual features which develop both the local spatial co occurrence between low level visual words and the global spatial layout of the object parts. To obtain the global spatial features, a fast method is proposed to generate the semantic meaningful object parts by exploiting the geometric position distribution of the local salient regions. At last, multi-kernel learning framework is proposed to adaptively integrate extracted features in an optimal way.

Bayesian approach of dynamically selecting camera parameters in order to distinguish a given object from a finite set of object classes is proposed in Huber *et al*. (2012). Gaussian process regression is applied to learn the likelihood of image features given the object classes and camera parameters. In doing so, the object recognition task can be treated as Bayesian state assessment problem. For improving the recognition accuracy and speed, the selection of suitable camera parameters is formulated as a sequential optimization problem. A new video surveillance object recognition algorithm is presented in Wu and Xiao (2010), in which

Corresponding Author: Thambu Gladstan, Department of ECE, Shri JJT University, Jhunjhunu, Rajasthan, India

improved invariant moments and length-width ratio of object are extracted as shape feature, while color histograms of object are utilized as color feature. By combining both shape and color the object recognition is achieved.

Closed 2D curves are parameterized and Fourier descriptors are used in Arbter *et al*. (1990) to produce a set of normalized coefficients which are invariant under affine transformations. The method is verified on silhouettes of aircraft. Since the shapes of airplanes are more or less planar when seen from large distances, they give rise to affine transformations when rotated in 3D.

MATERIALS AND METHODS

Wave atom transform: Wave atom transform is presented by Demanet in 2009. The transformation, obeying the parabolic scaling law (Laurent and Lexing, 2009), can be considered a variant of 2D wavelet packets. Wave atom transform have two very significant properties. First one is the ability to adapt to arbitrary local directions of a pattern. The second one is the ability to sparsely represent anisotropic patterns aligned with the axes. Wave atoms offer sharp frequency localization than other wave packets. It also has significant sparse expansion for oscillatory functions when compared with wavelets, curvelets and Gabor atoms.

The forms of wave packets, known as wavelets, Gabor, ridgelets, curvelets and wave atoms, are created using two parameters, which are α and β. These variables symbolize decomposition and directional ability for all wave forms. α and β values are 1/2 for wave atoms and Fig. 1 shows wave packet's support in space and in frequency plane. Here, α corresponds to the multi scale structure of the transform and β corresponds to directional selectivity.

Actually, wave atoms are built from tensor products of 1D wave packets. One-dimensional wave packets can be represented as $\psi^j_{m,n}(X)$, where j, $m \geq 0$ and $n \in Z$ frequency restrictions are $\pm\omega_{j,m} = \pm\pi\, 2^j\, m$ with $C_1\, 2^j \leq m \leq C_2\, 2^j$ space restrictions is defined as:

$$X_{j,n} = 2^j n \tag{1}$$

Two-dimensional wave atoms φ_μ ($X1$, $X2$) are constructed with subscript $\mu = (j, m, n)$, where $m = (m_1, m_2)$, $n = (n_1, n_2)$ 2D orthonormal basis is written as follows:

$$\varphi^+_\mu(X_1, X_2) = \psi^j_{m1}\left(X_1 - 2^{-j} n_1\right)\psi^j_{m2}\left(X_2 - 2^{-j} n_2\right) \tag{2}$$

$$\varphi^-_\mu(X_1, X_2) = H\psi^j_{m1}\left(X_1 - 2^{-j} n_1\right)H\psi^j_{m2}\left(X_2 - 2^{-j} n_2\right) \tag{3}$$

where, H is Hilbert transform. The wave atom tight frame is formed by combination of (2) and (3):

$$\varphi^{(1)}_\mu = \frac{\varphi^+_\mu + \varphi^-_\mu}{2}, \varphi^{(2)}_\mu = \frac{\varphi^+_\mu - \varphi^-_\mu}{2} \tag{4}$$

KNN classifier: The K-Nearest Neighbor algorithm (K-NN) is a method for classifying objects based on closest training examples in the feature space. K-NN is a type of instance-based learning where the function is only approximated locally and all computation is deferred until classification.

In K-NN, an object is classified by a majority vote of its neighbors; with the object being assigned to the class most common amongst its k nearest neighbors (k is a positive integer, typically small). If $k = 1$, then the object is simply assigned to the class of its nearest neighbor. The neighbors are taken from a set of objects for which the correct classification is known. This can be thought of as the training set for the algorithm, though no explicit training step is required.

A measure in which to determine the distance between two scenarios is established and simply pass through the data set, one scenario at a time and compare it to the testing scenario. Let our training data set is represented as a matrix $D = N{*}P$, containing P scenarios $s^1,\ldots,s^p$ where each scenarios s^i contains N features $s^i = \{S^i_1,\ldots,S^i_N\}$. A vector o with length P of output values $o = \{o^1,\ldots,o^p\}$ accompanies this matrix, listing the output value o^i for each scenario S^i.

It should be noted that the vector o can also be seen as a column matrix; if multiple output values are desired, the width of the matrix may be expanded. KNN can be run in these steps:

- Store the output values of the M nearest neighbors to query scenario q in vector $r = \{r^1,\ldots r^M\}$. By repeating the following loop M times
- Go to the next scenario S^i in the data set, where the current iteration within the domain is $\{1,\ldots P\}$:
  - If q is not set or $q<d\ (q, S^i)$: $q \leftarrow d\ (q, S^i)$, $t \leftarrow o^i$
  - Loop until reach the end of the data set (i.e., i = P)
  - Store q into vector c and t into vector r
- Calculate the arithmetic mean output across as follows:

$$\bar{r} = \frac{1}{M}\sum_{i=1}^{M} r_i \tag{5}$$

- Return $\bar{r}$ as the output value for the query scenario q

PROPOSED ALGORITHM

The proposed method consists of two stages namely feature extraction stage and classification stage. The two stages are described below.

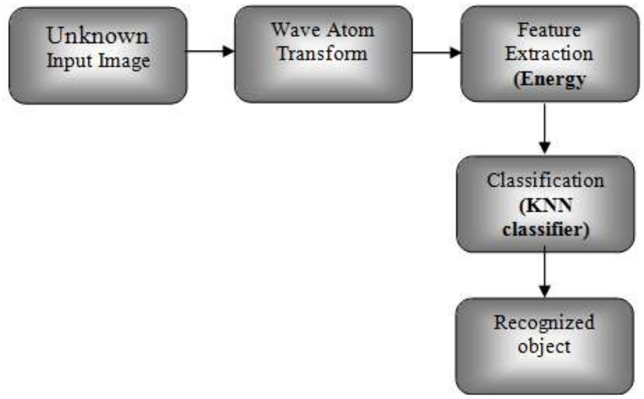

Fig. 1: Block diagram of the proposed method

Table 1: Classification rate of the proposed system by varying the angles between views

No. of training views	36	18	12	8	6
Angle b/w views	10°	20°	30°	45°	60°
Obj_1	91.67	85.19	75.00	60.94	37.88
Obj_2	100	100	100	100	100
Obj_3	100	98.15	95.00	76.56	56.06
Obj_4	100	100	100	100	100
Obj_5	100	100	100	100	100
Obj_6	100	94.44	80.00	59.38	43.94
Obj_7	100	100	100	100	100
Obj_8	100	85.19	71.67	51.56	33.33
Obj_9	100	100	100	100	100
Obj_10	100	100	100	96.88	89.39
Obj_11	100	100	100	96.88	89.39
Obj_12	100	100	100	98.44	84.85
Obj_13	100	98.15	98.33	85.94	80.30
Obj_14	100	100	100	100	93.94
Obj_15	97.22	88.89	76.67	54.69	40.91
Obj_16	100	96.30	88.33	67.19	66.67
Obj_17	100	100	100	100	100
Obj_18	100	100	100	100	100
Obj_19	100	87.04	73.33	56.25	36.36
Obj_20	100	100	100	100	100
Obj_21	97.22	90.74	68.33	51.56	37.88
Obj_22	100	100	96.67	95.31	75.76
Obj_23	100	96.30	85.00	73.44	54.55
Obj_24	100	100	100	100	100
Obj_25	100	100	100	100	100
Obj_26	100	100	100	100	100
Obj_27	100	87.04	75.00	59.38	40.91
Obj_28	100	100	100	100	83.33
Obj_29	100	100	100	100	96.97
Obj_30	100	100	100	100	100
Obj_31	90.56	74.07	60.00	48.44	25.76
Obj_32	100	100	100	100	100
Obj_33	100	100	100	100	100
Obj_34	100	100	100	100	100
Obj_35	100	100	100	100	100
Obj_36	100	100	100	100	95.45
Obj_37	100	90.74	81.67	53.13	36.36
Obj_38	100	94.44	90.00	81.25	71.21
Obj_39	100	98.15	96.67	85.94	66.67
Obj_40	100	100	98.33	90.63	84.85
Obj_41	100	100	100	100	89.39
Obj_42	100	98.15	95.00	89.06	71.21
Obj_43	100	100	100	95.31	80.30
Obj_44	100	88.89	76.67	59.38	42.42
Obj_45	100	100	98.33	96.88	86.36
Obj_46	100	94.44	85.00	71.88	48.48
Obj_47	100	100	100	100	100

Table 1: Continue

No. of training views	36	18	12	8	6
Obj_48	100	100	98.33	81.25	69.70
Obj_49	100	100	100	100	100
Obj_50	100	100	100	100	100
Obj_51	100	100	93.33	79.69	66.67
Obj_52	100	100	100	100	98.48
Obj_53	100	100	100	100	100
Obj_54	100	96.30	90.00	71.88	42.42
Obj_55	100	100	100	98.44	78.79
Obj_56	100	100	100	100	100
Obj_57	100	98.15	93.33	84.38	71.21
Obj_58	100	100	100	100	100
Obj_59	100	100	96.67	85.94	80.30
Obj_60	100	100	100	82.81	69.70
Obj_61	100	98.15	95.00	98.44	89.39
Obj_62	100	98.15	95.00	78.56	72.73
Obj_63	100	100	98.33	100	100
Obj_64	100	100	100	100	100
Obj_65	94.44	87.04	81.67	62.50	40.91
Obj_66	100	100	100	100	100
Obj_67	41.67	44.44	36.67	20.31	15.15
Obj_68	100	94.44	78.33	62.50	59.39
Obj_69	88.89	74.07	53.33	42.19	31.82
Obj_70	100	100	100	100	100
Obj_71	100	100	100	100	100
Obj_72	100	100	100	100	100
Obj_73	100	100	100	100	100
Obj_74	100	100	100	100	93.94
Obj_75	100	100	98.33	96.88	89.39
Obj_76	94.44	77.78	63.33	48.44	37.88
Obj_77	100	98.15	98.33	89.06	83.33
Obj_78	100	100	91.67	75.00	65.15
Obj_79	100	90.74	85.00	62.50	40.91
Obj_80	100	98.15	91.67	82.81	78.79
Obj_81	100	98.15	93.33	76.56	66.67
Obj_82	97.22	98.15	93.33	78.13	72.73
Obj_83	100	100	100	100	100
Obj_83	100	100	100	100	100
Obj_84	58.33	53.70	45.00	40.63	25.76
Obj_85	100	100	100	92.19	78.79
Obj_86	100	100	100	100	100
Obj_87	100	100	100	100	100
Obj_88	100	100	100	100	100
Obj_89	100	100	100	100	100
Obj_90	100	100	100	100	100
Obj_91	100	96.30	90.00	78.13	60.61
Obj_92	100	100	91.67	78.13	72.73
Obj_93	100	100	100	100	100
Obj_94	100	100	100	100	100
Obj_95	100	100	100	100	100
Obj_96	94.44	88.89	81.67	51.56	34.85
Obj_97	100	100	100	100	95.45
Obj_98	83.33	75.93	71.67	62.50	42.42
Obj_99	100	100	100	100	98.48
Obj_100	100	94.44	86.67	75.00	50.00
Average	98.29	96.07	92.57	85.93	78.15

Feature extraction stage: Feature extraction is an essential pre-processing step to pattern recognition and machine learning problems. In the proposed system, the wave atom coefficients are used as features to recognize the objects. The original image is decomposed by using the WAT. All coefficients are considered as features for the classification process. This method is applied to all the training object images and the feature vectors are stored in the database called as Object-Base. The energy of each directional sub-band of the image *I* is calculated by using the formula in (2):

$$Energy_e = \frac{1}{RC}\sum_{i=1}^{R}\sum_{j=1}^{C}\left|I_e(i,j)\right|$$

where,
$I_e(i,j)$: The pixel value of the eth sub-band
R, C : Width and height of the sub-band, respectively

Classification stage: The classifier used to classify the object is KNN classifier. The unknown object is selected and its features are extracted. Here all the coefficients of WAT are considered as the feature of the unknown object. Both the extracted features and the stored database of the training images are used as an input to the K-Nearest Neighbor (K-NN) classifier to recognize the object. The classification performance of the proposed system is evaluated by using classification rate in percentage.

RESULTS AND DISCUSSION

The evaluation of the system is carried on using Columbia Object Image Library dataset (COIL-100) database. The size of image in the database is 128×128 and the number of objects in the data base is 100 and 72 images are available per object; one at every 5° of rotation. By varying the angles by 10, 20, 30, 45, 60°, respectively between views the training and testing data are formed. The classification rate in percentage is tabulated and shown in Table 1.

From Table 1, it is clear that 100 give the better classification accuracy. Higher the degree lowers the classification rate. From Table 1, it is clear that 10° give the better classification accuracy. Higher the degree lowers the classification rate.

CONCLUSION

In this study, an efficient method for recognizing object in an image based on WAT is presented. Ninety eight percent of classification rate is achieved by the proposed method by varying the angle between the views. The experimental results show that 100 give the better classification accuracy when compared with the other degrees. Higher the degree lowers the classification rate.

REFERENCES

Arbter, K., W.E. Snyder, H. Burkhardt and G. Hirzinger, 1990. Application of affine-invariant Fourier descriptors to recognition of 3-d objects. IEEE T. Pattern Anal. Mach. Intell., 12(7): 640-647.

Choi, M.J., A.B. Torralba and A.S. Willsky, 2012. A tree-based context model for object recognition. IEEE T. Pattern Anal. Mach. Intell., 34(2): 240-252.

Huber, M.F., T. Dencker, M. Roschani and J. Beyerer, 2012. Bayesian active object recognition via gaussian process regression. Proceeding of the 15th International Conference on Information Fusion (FUSION), pp: 1718-1725.

Kim, K., J. Kim, S.S. Kang, J. Kim and J. Lee, 2012. Object recognition for cell manufacturing system. Proceeding of the 9th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), pp: 512-514.

Laurent, D. and Y. Lexing, 2009. Wave atoms and time up scaling of wave equations. J. Numer. Math., 113(1): 1-71.

Wang, M., Y. Wu, G. Li and X.D. Zhou, 2010. Object recognition via adaptive multi-level feature integration. Proceeding of the 12th International Asia-Pacific on Web Conference (APWEB), pp: 253-259.

Wu, J. and Z.T. Xiao, 2010. Video surveillance object recognition based on shape and color features. Proceeding of the Image Signal Processing Conference (CISP), pp: 451-454.