Human Object Classification in Daubechies Complex Wavelet Domain

Manish Khare[1], Rajneesh Kumar Srivastava[1], Ashish Khare[1(✉)], Nguyen Thanh Binh[2], and Tran Anh Dien[2]

[1] Image Processing and Computer Vision Lab, Department of Electronics and Communication, University of Allahabad, Allahabad, India
{mkharejk, rkumarsau}@gmail.com, ashishkhare@hotmail.com

[2] Faculty of Computer Science and Engineering, Ho Chi Minh City University of Technology, Ho Chi Minh, Vietnam
ntbinh@cse.hcmut.edu.vn, dientrananh@gmail.com

Abstract. Human object classification is an important problem for smart video surveillance applications. In this paper we have proposed a method for human object classification, which classify the objects into two classes: human and non-human. The proposed method uses Daubechies complex wavelet transform coefficients as a feature of object. Daubechies complex wavelet transform is used due to its better edge representation and approximate shift-invariant property as compared to real valued wavelet transform. We have used Adaboost as a classifier for classification of objects. The proposed method has been tested on standard datasets like, INRIA dataset. Quantitative experimental evaluation results show that the proposed method is better than other state-of-the-art methods and gives better performance.

Keywords: Object classification · Feature selection · Daubechies complex wavelet transform (DCxWT) · Adaboost classifier

1 Introduction

Classifying type of objects in real scene is a challenging problem for any computer vision system with many useful applications like object tracking, segmentation of moving object, smart surveillance etc. [1–3]. Correct video object classification is a key component for development of a smart surveillance system. For any video surveillance application, an object classification algorithm should hold following properties-

(i) object classification algorithm should perform under real time constraints.
(ii) object classification algorithm should be robust to the situations such as total or partial occlusions of object, color variation in human cloths, large variation in natural conditions, and varying lighting conditions etc.

To address the human object classification problem, different features and machine learning techniques have been used. Selection of effective features is a crucial step for

P.C. Vinh et al. (Eds.): ICCASA 2013, LNICST 128, pp. 125–132, 2014.
DOI: 10.1007/978-3-319-05939-6_13,

successful classification, Sialat et al. [4], in their pedestrian detection system, used Haar like features along with the decision tree. Viola et al. [5] used modified reminiscent of Haar basis function, for accomplishing object detection task. Lowe [6] used Scale Invariant Feature Transform (SIFT) as a feature descriptor for object recognition. Dalal and Triggs [7] proposed Histogram of oriented Gradient (HoG) as a feature descriptor for object detection, but HoG feature have disadvantages of having high dimensionality. Cao et al. [8] proposed a method by introducing an extension of the Histogram of oriented Gradient (HoG) features, known as boosting HoG feature. They used Adaboost scheme for boosting the HoG feature and SVM classifier for object classification. Lu et al. [9] proposed a visual feature for object classification based on binary pattern. These visual features are rotation invariant and exploit the property of pixel patterns.

All the methods discussed above depend on feature evaluation set therefore they have local advantages or disadvantage depending on the features they have used. Yu and Slotine [10] proposed a wavelet based method for visual classification. Method proposed by Yu and Slotine [10] uses real valued wavelet transform, but real valued wavelet transform is not suitable for video application because in case of video, object may be presented in translated and rotated form among different frames and coefficients of real valued wavelet transform corresponding to object region changes abruptly across different frames. Motivated by work of Yu and Slotine [10], we have proposed a new method for human object classification based on Daubechies complex wavelet transform (DCxWT) coefficients as a feature set. We have used Adaboost classifier for classifying human and non-human object classes. The DCxWT is having advantages of shift invariance and better edge representation as compared to real valued wavelet transform.

In the present work, our aim is to classify objects into two types of classes: human and non-human. We have experimented the proposed method at multiple levels of DCxWT and shown that performance of the proposed method is better at higher levels. The proposed method has been compared with the method using coefficients of discrete wavelet transform (DWT) as a feature set, and it has been shown that use of DCxWT as a feature set perform better than use of DWT coefficients as a feature set. We have compared the proposed method with other state-of-the-art methods proposed by Dalal and Triggs [7], Lu et al. [9], Renno et al. [11], and Chen et al. [12] in terms of average classification accuracy, True positive rate (Recall), True negative rate, False positive rate, False negative rate and Predicted positive rate (Precision).

The rest of the paper is organized as follows: Sect. 2 describes basics of used features (DCxWT), Sect. 3 describes Adaboost classifier and Sect. 4 describes the proposed method. Experimental results, analysis and comparison of the proposed method with other state-of-the-art methods are given in Sect. 5. Finally conclusion of the work is given in Sect. 6.

2 Feature Selection

Any object classification algorithm is commonly divided into three important components – extraction of features, selection of features and classification. Therefore, feature extraction and selection plays an important role in object classification.

We have used coefficients of Daubechies complex wavelet transform (DCxWT) as a feature for classification. A brief description of DCxWT is given in following subsection.

2.1 Daubechies Complex Wavelet Transform (DCxWT)

In object classification, we require a feature which remains invariant by shift, translation and rotation of object, because object may be present in translated and rotated form among different scenes. Due to its approximate shift-invariance and better edge representation property, we haved used DCxWT as feature set.

Any function *f*(*t*) can be decomposed into complex scaling function and mother wavelet as:

$$f(t) = \sum_k c_k^{j_0}\, \phi_{j_0,k}(t) + \sum_{j=j_0}^{j_{\max}-1} d_k^j\, \psi_{j,k}(t)$$

where, j_0 is a given low resolution level, $\left\{c_k^{j_0}\right\}$and $\left\{d_k^j\right\}$ are approximation coefficients $\left[\phi(u) = 2\sum_i a_i \phi(2u - i)\right]$ and detail coefficients $\left[\psi(t) = 2\sum_n (-1)^n \overline{a_{1-n}} \phi(2t - n)\right]$.

where $\phi(t)$and $\psi(t)$share same compact support [−L, L+1] and a_i's are coefficients. The a_i's can be real as well as complex valued and $\sum a_i = 1$.

Daubechies's wavelet bases $\{\psi_{j,k}(t)\}$ in one-dimension is defined through above scaling function $\phi(u)$ and multiresolution analysis of $L_2(\Re)$ [13]. During the formulation of general solution if we relax the condition for a_i to be real [14], it leads to complex valued scaling function.

DCxWT holds various properties [14], in which reduced shift sensitivity and better edge representation properties of DCxWT are important one.

3 Adaboost Classifier

Boosting a method to improve the performance of any learning algorithm, generally consist of sequentially learning classifier [15]. Adaboost itself trains an ensemble of weak learners to form a strong classifier which perform at least as well as an individual weak learner [11]. In our proposed work, we have used Adaboost algorithm which is firstly described by Viola and Jones [5] for their face detection system. Complete Adaboost algorithm for classifier is given below:

Adaboost algorithm for classifier-

- Given example images (I_1,J_1) (I_2,J_2),………… (I_n,J_n) where $J_i = 0,1$ for negative and positive example respectively.

  Initialize weights $W_{1,i} = \frac{1}{2n} \cdot \frac{1}{2p}$ for $j_i = 0,1$ respectively, where p and n are the number of positive and negative examples respectively.

- For t=1,2,……….T (Number of iterations)

1. Normalize the weights

$$W_{t,i} \leftarrow \frac{W_{t,i}}{\sum_{j=1}^{m} W_{t,j}}, \qquad W_t \text{ is the probability distribution}$$

2. For each feature, j, train a classifier h_j, which is restricted to using a single feature. The error (E_j) is evaluated with respect to W_t

$$E_j = \sum_i W_i \left| h_j(I_i) - J_i \right|$$

3. Choose the classifier, h_t with the lowest error E_t
4. Update the weights

$$W_{t+1,i} = W_{t,i} \beta_t^{1-\in_i}$$

where $\in_i = 0$, if example x_i is classified correctly
$\in_i = 1$, otherwise

$$\beta_t = \frac{E_t}{1 - E_t}$$

- The final strong classifier is

$$h(I) = \begin{cases} 1, & if \sum_{t=1}^{T} \alpha_t h_t(I) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha_t \\ 0, & otherwise \end{cases}$$

where, $\alpha_t = \log \frac{1}{\beta_t}$.

4 The Proposed Method

The proposed method is inspired by work of Yu and Slotine [10], which uses wavelet coefficients as a feature set. The proposed method uses Daubechies complex wavelet transform feature evaluation and Adaboost classifier for classification. For experimentation, we have considered two classes: human and non-human. Steps of the proposed method are as follows-

Step 1: Sample Collection- First step of the proposed method is to collect sample images for training the classifier. In our approach, we have taken INRIA dataset images for training purpose. The INRIA dataset [16] contains 2521 positive images and 1686 negative images of size 96 × 160. Positive images contains human object images and negative images contains any type of images in which human object is not present.

Step 2: Preprocessing of Images- The collected images are scale normalized to 256 × 256 pixel dimensions in order to reduce complexity. The scale normalized images in the RGB color space were converted to the gray level images.

Step 3: Feature Vector Calculation- For feature vector computation in the proposed method, image frames are decomposed into complex wavelet coefficients using Daubechies complex wavelet transform. After applying DCxWT, we get coefficients in four subbands namely – LL, LH, HL and HH. Values of LH, HL and HH subbands are used as feature values of different images. We skip value of LL subbands because LL subbands gives approximation coefficient of images and rest other LH, HL and HH gives detail coefficients of images.

Step 4: Human Object Classification- The calculated features are supplied into Adaboost classifier. This classifier separates, two types of images: human object image and non-human object image (do not contain any type of human). The test images have been categorized by this classifier.

5 Experimental Results

In this section, we provides experimental results of the proposed method and other state-of-the-art methods proposed by Dalal and Triggs [7], Lu et al. [9], Renno et al. [11], and Chen et al. [12] in terms of average classification accuracy, True positive rate(Recall), True negative rate, False positive rate, False negative rate and Predicted positive rate (Precision).

The proposed method for human object classification has been tested on standard dataset like – INRIA dataset [16]. INRIA dataset contains 2521 positive images and 1686 negative images of size 96 × 160. Here we present some representative images in Fig. 1.

Images shown in Fig. 1, have been taken from real scenes. By observing these images, one can observe that both frontal as well side views of human objects are

Fig. 1. Representative Human objects from INRIA dataset

taken into account. We have evaluated the proposed method for multiple levels of DCxWT coefficients (L-1,2……..7). Just to compare the performance of the proposed method, we have also experimented with multilevel DWT coefficients as a feature for human object classification.

The comparative study of the proposed method and other state-of-the-art methods [7, 9, 11, 12] are shown in Table 1 in terms of six different performance metrics – average classification accuracy, True Positive Rate (Recall), True Negative Rate (TNR), False Positive Rate (FPR), False Negative Rate (FNR) and Predicted Positive Rate (PPR). These performance metrics depends on four values – TP, TN, FP and FN, which are defined as: TP (True Positive) are number of images which are originally positive images and detected as positive images, TN (True Negative) are number of images which are originally negative images and detected as negative images, FP (False Positive) are number of images which are originally negative images and detected as positive images and FN (False Negative) are number of images which are originally positive and detected as negative images.

Average classification accuracy is the proportion of the total number of predictions that were correct. True Positive Rate (also known as Recall) is the proportion of

Table 1. Performance Measure Values

Methods name	TPR (Recall) (%)	TNR (%)	FPR (%)	FNR (%)	PPR (Precision) (%)	Average Accuracy (%)
The Proposed method with DCxWT (Level-1) as a feature	94	70	30	06	75.81	82
The Proposed method with DCxWT (Level-2) as a feature	94	75	25	06	78.99	84.50
The Proposed method with DCxWT (Level-3) as a feature	98	75	25	02	79.67	86.50
The Proposed method with DCxWT (Level-4) as a feature	98	80	20	02	83.05	89
The Proposed method with DCxWT (Level-5) as a feature	99	88	12	01	89.19	93.50
The Proposed method with DCxWT (Level-6) as a feature	100	94	06	00	94.34	97
The Proposed method with DCxWT (Level-7) as a feature	100	95	05	00	95.24	97.50
DWT (Level-1) as a feature	70	65	35	30	66.67	67.5
DWT (Level-2) as a feature	75	70	30	25	71.43	72.5
DWT (Level-3) as a feature	75	70	30	25	71.43	72.5
DWT (Level-4) as a feature	80	74	26	20	75.47	77
DWT (Level-5) as a feature	85	80	20	15	80.95	82.5
DWT (Level-6) as a feature	85	82	18	15	82.52	83.5
DWT (Level-7) as a feature	92	86	14	08	86.79	89
Method used Dalal and Triggs [7]	94	92	08	06	92.16	93
Method used by Lu et al. [9]	90	70	30	10	75.00	80
Method used by Renno et al. [11]	89	69	31	11	74.17	79
Method used by Chen et al. [12]	98	96	04	02	96.08	96

positive cases that were correctly identified. True Negative Rate is defined as the proportion of negatives cases that were classified. False Positive Rate is the proportion of negatives cases that were incorrectly classified as positive. False Negative Rate is the proportion of positives cases that were incorrectly classified as negative. Predicted Positive Rate (also known as Precision) is the proportion of the predicted positive cases that were correct. These values can be determined using following formulas:

$$Average\,Classification\,Accuracy = \frac{TP + TN}{TP + FN + TN + FP}$$

$$TPR(\mathrm{Recall}) = \frac{TP}{TP + FN}$$

$$TNR = \frac{TN}{TN + FP}$$

$$FPR = \frac{FP}{TN + FP}$$

$$FNR = \frac{FN}{TP + FN}$$

$$PPR(\mathrm{Precision}) = \frac{TP}{TP + FP}$$

From Table 1, one can observe that the proposed method gives better performance at higher levels of Daubechies complex wavelet transform in comparison to other discrete wavelet transform and state-of-the-art methods [7, 9, 11, 12] for human object classification.

6 Conclusion

In this paper, we have developed and demonstrated a new method for classification of human object in real scenes. The approach first train Adaboost classifier by using coefficients of Daubechies complex wavelet transform as a feature set, then classify different objects into one of two categories: human and non-human. We have also experimented the classification result by using discrete wavelet transform as a feature set. We compared the proposed method with other state-of-the-art methods proposed by Dalal and Triggs [7], Lu et al. [9], Renno et al. [11], and Chen et al. [12] in terms of average classification accuracy, TPR, TNR, FPR, FNR and PPR, and found that the proposed method perform better than other methods. The main advantage of the proposed method is that, the proposed method can detect human objects of different size whether too small or too large, the proposed method can detect human objects with complex background.

Finally it could be concluded that on the basis of observed results, the proposed method for human object classification have better average classification accuracy, better TPR, TNR, FPR, FNR and PPR values in comparison to other state-of-the art methods [7, 9, 11, 12].

Acknowledgement. This work was supported in part by Council of Scientific and Industrial Research (CSIR), Human Resource Development Group, India, Under Grant No. 09/001/(0377)/2013/EMR-I.

References

1. Hu, W., Tan, T.: A survey on visual surveillance of object motion and behaviors. IEEE Trans. Syst. Man Cybern. **34**(3), 334–352 (2006)
2. Khare, M., Srivastava, R. K., Khare, A.: Moving object segmentation in daubechies complex wavelet domain. Accepted for publication in Signal Image and Video Processing. Doi: 10.1007/s11760-013-0496-5 (2013)
3. Brooks, R.R., Grewe, L., Iyengar, S.S.: Recognition in wavelet domain: a survey. J. Electron. Imaging **10**(3), 757–784 (2001)
4. Sialat, M., Khlifat, N., Bremond, F., Hamrouni, K.: People detection in complex scene using a cascade of boosted classifiers based on Haar-like features. In: Proceeding of IEEE International Symposium on Intelligent Vehicles, pp. 83–87 (2009)
5. Viola P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proceeding of IEEE International Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 83–87 (2001)
6. Lowe, D.: Object recognition from local scale invariant features. In: Proceeding of 7th IEEE International Conference on Computer Vision, pp. 1150–1157 (1999)
7. Dalal N., Triggs, B.: Histograms of oriented gradients for human detection. In: proceeding of IEEE International Conference on Computer vision and Pattern Recognition, pp. 886–893 (2005)
8. Cao, X., Wu, C., Yan, P., Li, X.: Linear SVM Classification using boosting HoG features for vehicle detection in low-altitude airborne videos. In: Proceeding of IEEE International Conference on Image Processing, pp. 2421–2424. (2011)
9. Lu, H., Zheng, Z.: Two novel real-time local visual features for omnidirectional vision. Pattern Recogn. **43**(12), 3938–3949 (2010)
10. Yu, G., Slotine, J. J.: Fast wavelet-based visual classification. In: Proceeding of IEEE International Conference on Pattern Recognition (ICPR), pp. 1–5 (2008)
11. Renno, J. P., Makris, D., Jones, G. A.: Object classification in visual surveillance using Adaboost. In: Proceeding of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–8 (2007)
12. Chen, L., Feris, R., Zhai, Y., Brown, L., Hampapur, A.: An integrated system for moving object classification in surveillance videos. In: Proceeding of IEEE International Conference on Advanced Video and Signal based Surveillance, pp. 52-59 (2008)
13. Daubechies, I.: Ten lecture on Wavelets. Society for Industrial and Applied Mathematics (SIAM) (1992)
14. Clonda, D., Lina, J.M., Goulard, B.: Complex daubechies wavelets: properties and statistical image modeling. Sig. Process. **84**(1), 1–23 (2004)
15. Zhu, Z., Zou, H., Rosset, S., Hastie, T.: Multiclass Adaboost. Int. J. Stat. Interface **2**, 349–360 (2009)
16. INRIA Person Dataset. http://pascal.inrialpes.fr/data/human Last Accessed 29 September 2013