Object Classification Based on Contourlet Transform in Outdoor Environment

Nguyen Thanh Binh(✉)

Faculty of Computer Science and Engineering,
Ho Chi Minh City University of Technology, Ho Chi Minh City, Vietnam
ntbinh@cse.hcmut.edu.vn

Abstract. Classification of objects is an important task in computer vision. In the case that the objects are occlusion or outdoor environment, classification of objects is a challenging problem. The primary goal of this paper is to classify the object into two classes: human and car in an outdoor environment. In order to detect object classification, most of existing methods separated detecting object region from pre-defined background model. Here, we propose a method to implement classification of human and car in outdoor environment using contourlet transform combined with support vector machine as a classifier for classification of objects. The proposed method tested on standard dataset like PEST2001 dataset. For demonstrating the superiority of the proposed method, we have compared the results with the other recent methods available in literature.

Keywords: Object classification · Contourlet transform · Support vector machine

1 Introduction

Classification of objects is an important task in computer vision, where we classify human and non-human objects in real scene [1]. There are two tasks for image understanding: object detection and classification in the past decades. The object classification aims to predict the existence of objects within images while the object detection is localizing the objects [2]. Any object classification algorithm is to develop a method having capability to interpret the objects into different groups. Object classification algorithm must work under real-time constraints and must be robust in variation in natural conditions, different sizes of human objects, etc [10]. Feature selection and machine learning are the key components in any classification algorithm. Most of object classification algorithms developed base on Machine learning methods [3].

In the past, many algorithms have been built to object classification. Lowe [4] used Scale Invariant Feature Transform as a feature descriptor for object recognition. Lu [5] proposed a visual feature for object classification based on binary pattern. Dalal [6] proposed Histogram of oriented Gradient (HoG) as a feature descriptor for object detection. Cao [7] proposed a method by extending the HoG to boosting HoG feature.

P.C. Vinh et al. (Eds.): ICTCC 2014, LNICST 144, pp. 341–349, 2015.
DOI: 10.1007/978-3-319-15392-6_32

All the methods discussed above have local advantages or disadvantages depending on the features they have used [10].

Yu [8] proposed wavelet method for visual classification. This method uses real valued discrete wavelet transform. Real valued wavelet transform has three major problems: lack of shift sensitivity, poor directionality and lack of strong edge detection [11]. This drawback affects the process of feature selection. To increase the ability to identify objects, we use contourlet transform to overcome these problems.

In this paper, we propose a method to implement classification human and car in an outdoor environment using contourlet transform combined with support vector machine as a classifier for classification of objects. The proposed method was tested on a standard dataset like PEST2001 dataset. For demonstrating the superiority of the proposed method, we have compared the results with the other recent methods by Lu [5] and Renno [9]. We use three different performance metrics: average classification accuracy, true positive rate (recall), and predicted positive rate (precision) for this comparison.

The rest of the paper is organized as follows: in section 2, we described the basic of feature selection, contourlet transform and details of the support vector machine classifier; the proposed method is presented in section 3; the results of proposed method are given in section 4 and conclusions in section 5.

2 Background

2.1 The Contourlet Transform

Real valued wavelet transform suffers from three major problems: lack of shift sensitivity, poor directionality and lack of strong edge detection. Do [12] proposed a solution to overcome these problems by contourlet transform (CT).

Contourlets constitute a new family of frames that are designed to represent smooth contours in different directions of an image. A contourlet is easily applied in image processing because its representation is a fixed transform [12, 16]. The contourlet not only inherits the main qualities of wavelet transform, such as multi-scale and time-frequency information, but also captures direction characteristics. It holds the geometrical formation of images and implements a true sparse representation of images. The contourlet allows for a different number of directions at each scale and aspect ratios. This feature allows an efficient contourlet-based approximation of a smooth contour at multiple resolutions. The discrete contourlet transform is a multiscale and directional decomposition using a combination of Laplacian pyramid (LP) and directional filter bank (DFB) [12, 16].

The idea of the contourlet construction [12] is: let $a_0[n]$ be the input image, the output after the LP step is I bandpass images $b_i[n]$, i =1, 2,..., I and a lowpass image $a_I[n]$. Each bandpass image $b_i[n]$ is decomposed by an ℓ_i -level DFB into 2^{ℓ_i} bandpass directional images $c_{i,k}^{(l_i)}[n]$, for k = 0, 1,…, 2^{ℓ_i}-1 .

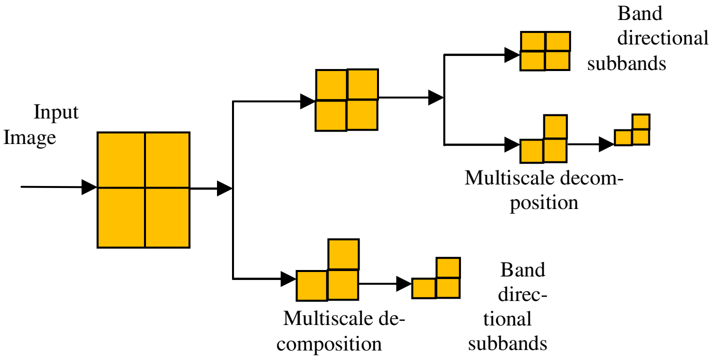

Fig. 1. Contourlet decomposition

In the discrete contourlet transform, the multiscale and directional decomposition steps are decoupled. Therefore, we have different numbers of directions at different scales. Contourlet decomposition proceeds through two main steps: first, LP multiscale decomposition is performed; then directional filter bank decomposition is used to link point discontinuity to linear structures. In more detail, an image is decomposed into a low pass image and bandpass images by the LP decomposition. Each bandpass output is further decomposed by the DFB step. The output of the DFB step consists of smooth contours and directional edges. In this paper, each directional subband at each level consists of 2^n element, where n is a positive integer. Figure 1 shows a contourlet decomposition [16]. Human and car object classification are a problem where the objects may present in translated as well as rotated form among different scenes. Contourlet transform has the time-frequency-localization and multiscale properties of wavelets. It offers a high degree of directionality and anisotropy. Therefore, the properties of contourlet transform will be useful for classification of human and car object.

2.2 Support Vector Machine Classifier

Support vector machines (SVM) include associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis in machine learning. SVM can efficiently perform a non-linear classification, implicitly mapping their inputs into high-dimensional feature spaces.

SVM is a popular classifier. The classifier objects are into two categories: object and non-object data [13]. In here, we detect two types: human and car object.

An n-dimensional object x has n-coordinates.

$$x = (x_1, x_2, x_3, \ldots\ldots, x_n),$$

where, each $x_i \in R$ for i=1, 2, 3,….,n .

Each object x_j belongs to a class $y_j \in \{-1,+1\}$. Consider a training set T of m patterns together with their classes,

$$T = \{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$$

and a dot product space S, in which the objects are embedded: $x_1, x_2, ..., x_m \in S$.

Any hyperplane in the space S can be written as:

$$\{x \in S \mid w.x + b = 0\}, w \in S, b \in R$$

The dot product w.x is defined by [10]:

If there exists at least one linear classifier defined by the pair (w, b) which correctly classifies all objects as shown in Figure 2 then a training set of objects is linearly separable [10]. The linear classifier is represented by the hyperplane H (w.x+b=0) and defines a region for class +1 and another region for class -1 objects.

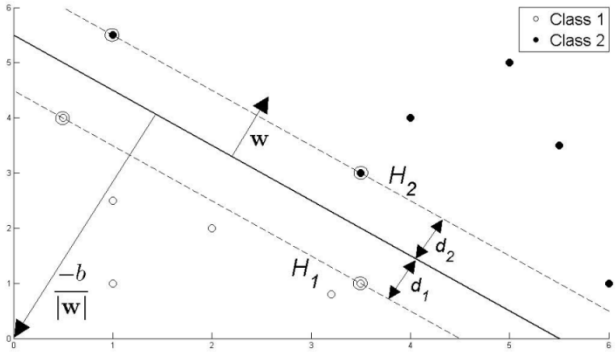

Fig. 2. Linear classifier [10] defined by the hyperplane H

After training, the classifier is ready to predict the class membership for new objects, different from those used in training. The class of object x_k is determined with the equation [10]:

$$class(x_k) = \begin{cases} +1 & if \quad w.x_k + b > 0 \\ -1 & if \quad w.x_k + b < 0 \end{cases}$$

2.3 Feature Selection

Feature selection is to select a subset of input variables with no predictive information by eliminating features. It can significantly improve the comprehensibility of the resulting classifier models. A feature is a function of one or more measurements computed so that it quantifies some significant characteristics of objects [15]. In any object classification algorithm, the selection of appropriate feature is very important. We see that the performance of classifier will increase if the correct feature is selected for classification algorithm. In our proposed work for human and car classification, we have used contourlet transform coefficients as a feature set. We have taken combination of two different features – contourlet transform and support vector machines. A brief description of these two features and why they are useful for human object classification are described in subsection 2.1 and 2.2 respectively.

3 The Proposed Method

In this section, we propose a method for object classification. The proposed method uses contourlet transform coefficient as a feature evaluation set and support vector machine as a classifier for classification of data into two categories: human object and car object. Steps of the proposed method are described as figure 3:

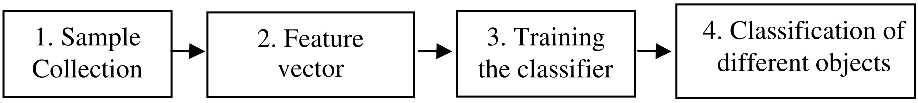

Fig. 3. The process of the proposed method

Fig. 4. Sample images with car and human objects of PEST2001 dataset

Firstly, we collect sample images for training and testing the classifier. In here, we have taken PETS2001 dataset [14] images for training and testing purpose. We have created our own dataset that consists of 500 images (300 images for training and 200 images for testing). We have shown some car and human objects of PEST2001 dataset [14] in figure 4.

The images in PEST2001 dataset are of different size. To reduce complexity, we should require normalization of these images. The collected images are scale normalized to 256 x 256 pixel dimensions. We also converted to the gray level images from the RGB color space.

Secondly, we compute feature vectors. In the proposed method, image frames are decomposed into complex wavelet coefficients using contourlet transform. After applying contourlet transform, we get coefficients in form of two filters: low-pass filter image and high-pass filter image as shown in figure 1. The value of high-pass filtered image is used as feature values of different images, because high-pass filtered image provides detailed coefficient of images, which is in form of complex values. We have skipped the value of the low-pass filtered image, because the low-pass filter image provides the approximation of coefficients of the image, which is in form of real values.

Thirdly, we train the classifier using feature values as the same algorithm in [10], which we have got in step 3. We have used SVM classifier, in which we have assign value '0' for car object data and value '1' for human object data by detecting car and human from image to image. By using feature value of images and assigning value of data, SVM classifier trained for classification. Detailed information of SVM classifier is given in section 2.

Finally, we are to classify the test data into one of the two categories: car and human object. For this process, we compute the feature vector of image using step 3 of the proposed method, then this computed feature value is supplied into SVM classifier, where SVM classifier analyzes this feature value by previously trained data and gives the result of two value '0' and '1', where '0' indicates car object data and '1' indicates human object data. The same process will be repeated for all test data.

4 Experimental and Evaluation

In this section, we apply the procedure described in section 3 and achieved a superior performance in our object classification experiments as demonstrated in this section. For performance evaluation, we compare the results of the proposed method based on combined contourlet transform (CT) with SVM with the methods: method proposed by Lu [5] and Renno [9].

The quality of car and human object is increasing by comparison with the value of average classification accuracy, True positive rate (TPR) (Recall), and Predicted positive rate (PPR) (Precision). The proposed method has been tested on PEST2001 person dataset [14]. We have evaluated the proposed method for multiple levels of contourlet transform coefficients (L = 1, 2, …, 6).

The different performance metrics, such as Average classification accuracy (ACA), True positive rate (TPR) (Recall) and Predicted positive rate (PPR) (Precision), are depended on four values: True Positive (TP), TN (True Negative), FP (False Positive) and False Negative (FN), where [10]:

+ TP is the number of images, which are originally positive images and classified as positive images.

+ TN is the number of images, which are originally negative images and classified as negative images.

+ FP is the number of images, which are originally negative images and classified as positive images.

+ FN is the number of images, which are originally positive images and classified as negative images.

All above three performance metrics are defined in [10]. In here, we review parameters following:

+ ACA is defined as the proportion of the total number of prediction that was correct: $\mathrm{ACA}=\frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}}$

+ TPR is defined as the proportion of positive cases that were correctly classified as positive: $\mathrm{TPR\,(Recall)}=\frac{\mathrm{TP}}{\mathrm{FP+FN}}$

+ PPR is defined as the proportion of the predicted positive cases that were correct: $\mathrm{PPR(Precision)}=\frac{\mathrm{TP}}{\mathrm{FP+TP}}$

Now, we have experimented on PEST2001 dataset. Here, we report the results as shown in figure 5.

Fig. 5. Car and human classification in PEST2001 dataset (image in RGB color space)

Table 1 shows the value of TPR, PPR and ACA of proposed method with other method.

From table 1, one can observe that the proposed method gives better than performance results at higher levels of contourlet transform in comparison to other methods [6, 9], as a feature, for human and car object classification.

5 Conclusions

In the present work, our aim is to classify objects into two types of classes: human and car. We develop a method for object classification in real scenes using contourlet transform as a feature set. Contourlet allows for a different number of directions at each scale and aspect ratios. This feature allows an efficient contourlet to have based approximation of a smooth contour at multiple resolutions. Human and car object classification is a problem where the objects may present in translated as well as rotated form among different scenes. Contourlet transform has the time-frequency-localization

Table 1. Performance Measure Values TPR, PPR and ACA

Methods Name	TPR (Recall) (%)	PPR (Precision) (%)	ACA (%)
The Proposed method with CT (Level-1) as a feature	88.00	88.00	88.00
The Proposed method with CT (Level-2) as a feature	91.00	89.87	89.01
The Proposed method with CT (Level-3) as a feature	93.00	92.17	92.50
The Proposed method with CT (Level-4) as a feature	94.00	94.89	95.05
The Proposed method with CT (Level-5) as a feature	94.00	94.89	95.05
The Proposed method with CT (Level-6) as a feature	95.00	96.02	96.50
Method used by Lu [5]	91.00	86.14	87.00
Method used by Renno [9]	90.00	85.34	86.07

and multiscale properties of wavelets. It offers a high degree of directionality and anisotropy. Therefore, the properties of contourlet transform will be useful for classification of human and car objects.

The proposed approach first trains SVM classifier by using contourlet coefficients of data as a feature set and then classifies testing data into one of the two categories: human and car objects. The proposed method is compared with other methods proposed by Lu [5] and Renno [9]. Experiments show that the proposed method gives better classification results at higher levels of contourlet transform and provide better results than other methods. The proposed method can detect human objects in a complex background.

References

1. Hu, W., Tan, T.: A Survey on Visual Surveillance of Object Motion and Behaviors. IEEE Transaction on System, Man, and Cybernetics **34**(3), 334–352 (2006)
2. Song, Z., Chen, Q., Huang, Z., Hua, Y., Yan, S : Contextualizing object detection and classification. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1585–1592 (2011)
3. Wang, L., Hu, W., Tan, T.: Recent Developments in Human Motion Analysis. Pattern Recognition **36**(3), 585–601 (2003)
4. Lowe, D.: Object recognition from local scale invariant features. In: Proceeding of 7th IEEE International Conference on Computer Vision, pp. 1150–1157 (1999)
5. Lu, H., Zheng, Z.: Two novel real-time local visual features for omnidirectional vision. Pattern Recognition **43**(12), 3938–3949 (2010)
6. 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)
7. 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)
8. Yu, G., Slotine, J.J.: Fast wavelet-Based visual classification. In: Proceeding of IEEE International Conference on Pattern Recognition (ICPR), pp. 1–5 (2008)

9. 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, pp. 1–8 (2007)
10. Khare, M., Binh, N.T., Srivastava, R.K.: Dual tree complex wavelet transform based human object classification using support vector machine. Journal of Science and Technology, Vietnam Academy of Science and Technology **51**(4B), 134–142 (2013)
11. Binh, N.T., Khare, A.: Image Denoising, Deblurring and Object Tracking, A new Generation wavelet based approach. LAP LAMBERT Academic Publishing (2013)
12. Do, M.N., Vetterli, M.: The Contourlet Transform: An Efficient Directional Multiresolution Image Representation. IEEE Transactions on Image Processing **14**, 2091–2106 (2005)
13. Noble, W.S.: What is Support Vector Machine. Nature Biotechnology **24**(12), 1565–1567 (2006)
14. PEST2001 Dataset. http://www.cvg.rdg.ac.uk/PETS2001/pets2001-dataset.html (last Accessed November 11, 2014)
15. Castleman, K.R.: Digital Image Processing. Prentice Hall, Englewood Cliffs (1996)
16. Thanh Binh, N., Dien, T.A.: Object detection and tracking in contourlet domain. In: Hung, N.M., Suzuki, J., Tung, N.T., Vinh, P.C. (eds.) ICCASA 2012. LNICST, vol. 109, pp. 192–200. Springer, Heidelberg (2013)