2012 IEEE International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS 2012) November 4-7, 2012

Crow Birds Detection using HOG and CS-LBP

Kidane Mihreteab
Nagaoka University of Technology
Niigata, Japan
mihreteab531@gmail.com

Masahiro Iwahashi
Nagaoka University of Technology
Niigata, Japan
iwahashi@vos.nagaokaut.ac.jp

Maki Yamamoto
Nagaoka University of Technology
Niigata, Japan
umiushi@vos.nagaokaut.ac.jp

***Abstract*—A robust image processing technique capable of detecting and localizing objects accurately plays an important role in many computer vision applications. In this paper, a feature based detector for birds is proposed. By combining Histogram of Oriented Gradients (HOG) and Center-Symmetric Local Binary Pattern (CS-LBP) as the feature set, detection of crows under various lighting conditions could be carried out. A dataset of crow birds with a wide range of poses and backgrounds was prepared and learned using linear Support Vector Machine (SVM). Experiments on different test images show that HOG and CS-LBP based descriptors can achieve 87% accuracy.**

Keywords-crow birds detection; HOG; CS-LBP; HOG CS-LBP detector

I. INTRODUCTION

Crows favor to gather in huge number on power lines and rooftops. These birds can pollute the environment as well disturbing the general peace in the vicinity of their gatherings. Hence, the ability to detect birds in images has very important applications in video surveillance, content-based image retrieval and aviation safety. However, detecting crow birds is a difficult task due to their small size and variations in background.

In comparison to human detection, less research is conducted for bird detection although it has a significant application in computer visions. Li *et al* [1] proposed a method based on optical flow and motion feature clustering for moving object extraction to detect birds from video. However, optical flow and motion feature clustering based methods are only applicable to detect moving objects. Kembhavi *et al* [2] employed initial pixel classification and pixelwise background model selection to track satin bowerbird from video, reporting an accuracy of almost 83%.

However, recently research on pedestrian detection [4, 5, 8, 9] show that sliding window classifiers are preeminent due to their good performance. In sliding window approach, the input image is scanned exhaustively from top left to the bottom right in different scales and positions. For each sliding window, certain feature sets are extracted and fed to the classifier, in which the classifier is trained ahead with a set of training data of the same type of features. The classifier will then determine whether the sliding window has an object or not.

Wang *et al* [5] applied an augmented feature by combining the Histogram of Oriented Gradients (HOG) features [4] with cell-structured Local Binary Pattern (LBP) feature [6] for detecting whole body of human. They reported the combination of HOG and cell-structured LBP feature gains more than 20% improvement over HOG detector. Zeng *et al* [9] combined multilevel HOG features with multilevel LBP features and they used Principal Component Analysis (PCA) [11] to reduce the dimension of the combined feature set to detect head-shoulder. Their results showed that the PCA based multilevel HOG-LBP acquired an accuracy of 89%. Zheng *et al* [8] also, proposed an effective dense Center-Symmetric Local Binary Pattern (CS-LBP) [3] and pyramid CS-LBP or Local Ternary Patterns (LTP) [13] for people detection. According to their report they achieved a detection rate over 80%. Hence, the detection performance acquired based on the above mentioned papers inspired us to the usage of sliding window classifier approach for bird or not bird decisions from still images.

This work aims to utilize feature extraction methods for crow birds' detection that is comparable in performance to the existing methods and with smaller feature dimension. We propose a discriminative descriptor feature by combining HOG and CS-LBP features. HOG [4] characterizes local object appearance and shape by the distribution of local intensity gradients and CS-LBP [3] which is a modified version of Local Binary Pattern (LBP) texture feature [12] and inherits both texture information and gradient based features. Thus, we believe that the combination of HOG and CS-LBP as one feature set can attain a better detection performance.

The main contributions of this paper are an augmented feature, which combines HOG with block-structured CS-LBP to extract local features from bird images. Secondly HOG and CS-LBP based bird detector achieve good detection performance with less feature vector dimension.

Experiments show that the proposed HOG CS-LBP based detector with linear Support Vector Machine (SVM) performs robustly on different bird images.

The rest of the paper is organized as follows. In section II details of HOG and CS-LBP feature extraction approach is presented. In section III experimental results along with discussions is given. Section VI concludes the paper.

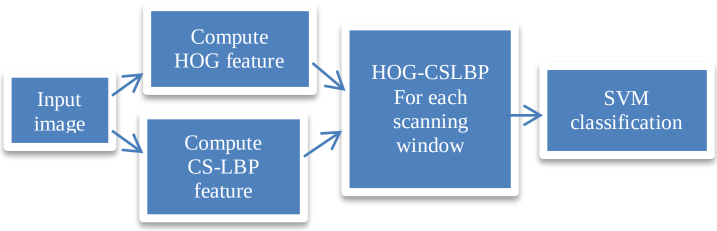

Figure 1. The block diagram of the proposed HOG CS-LBP detector

II. APPROACH

We use HOG and block-structured CS-LBP features to detect crow birds. This work is inspired by [4] and [8] which use dense Histogram of Oriented Gradients and Center-Symmetric Local binary pattern respectively. Fig. 1 shows the procedures of our crow birds detection algorithm based on combination of HOG and CS-LBP features.

A. *HOG feature extraction*

HOG feature [4] is evaluated based on well-normalized local histograms of image gradient orientations in a dense grid. HOG has been widely accepted as an excellent descriptor for capturing the edge or local shape information. Due to its great performance, it has been successfully applied for human detection in various applications.

To describe our bird class using HOG features, the descriptor properties are initially tuned to meet the requirements for our object of interest. The HOG feature extraction process is illustrated in Fig. 2.

For an input image of size 48x64 pixels, first we calculated the gradient magnitude of the input image using a simple 1-D [-1, 0, 1] masks without smoothing and with no gamma correction as in Fig. 2(b). Second we divide the gradient magnitude of the image into sixty four (64) blocks as shown in Fig. 2(c). Then, each block is further divided into four rectangular cells. Next, each pixel's gradient magnitude in each cell is voted into nine orientation bins based on the orientation of the gradient element centered on it. The 9 bins are evenly spaced between 0^0 -180^0.

Thirdly, the orientation histograms of the cells are concatenated to form block feature vector. Local contrast normalization of the block feature vector is done using L2-Hys normalization [4] to reduce the effect of local variations in illumination and foreground-background contrast. Next, the block feature vectors are concatenated to from feature vector of the window and the final HOG feature has a dimension of 2304 Fig. 2(d). This feature vector with 2304-D is the HOG descriptor of a bird image of window size 48x64 pixels.

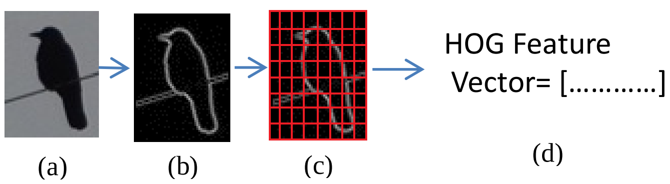

Figure 2. The HOG feature extraction process. (a) The input image. (b) The gradient magnitude of the input image. (c) The gradient magnitude division into 64 blocks. (d) The window HOG feature vector.

B. *CS-LBP feature extraction*

CS-LBP operator [3] was initially inspired by the Local Binary Pattern (LBP) operator [12]. CS-LBP is proposed to alleviate some drawbacks of the LBP operator such as, histograms of the LBP operator are long and it is not robust in flat images. In LBP operator [6, 7] the gray values of the neighboring pixels is compared with the gray value of the center pixel whereas in CS-LBP the gray level of the center-symmetric pairs of pixels are compared as shown in Fig. 3. The CS-LBP features are computed using (1) and (2):

$$CS-LBP_{p,r,t} = \sum_{i=0}^{N/2-1} S(g_i - g_{i+N/2})2^i \quad (1)$$

$$S(x) = \begin{cases} 1 \ \text{if}\ x \geq t, \\ 0 \ \text{otherwise} \end{cases} \quad (2)$$

where g_i and $g_{i+N/2}$ correspond to the gray values of center-symmetric pairs of pixels of N equally spaced pixels on a circle of radius r. Small value of t is used to threshold the gray level differences so as to increase the robustness of the CS-LBP feature on flat image regions. As we can notice from the computation of CS-LBP, it captures better the gradient information than the basic LBP due to the gray values difference while maintaining the LBP features. Thus, we constructed block-structured CS-LBP for crow birds' detection. In our experiment, we use the CS-LBP operator for a neighborhood of size 8 equally spaced on a circle of radius 1 and with $t = 0.02$.

The process of the CS-LBP feature extraction is shown in Fig 3. First, we normalize the gray level of the input image to reduce illumination effect. Then, we divide the gray-scale input image into 64 blocks to construct block-structured CS-LBP as shown in Fig. 3(a). Pixels in each block are voted into 16 bins according to the CS-LBP value calculated at each pixel in the block to create the CS-LBP descriptor Histogram. The L2-Hys normalization [4] is used for the feature histograms in a block. Finally, we compute the feature vectors for each block and then feature vectors of each block are concatenated into the final 1024-D window CS-LBP feature vector.

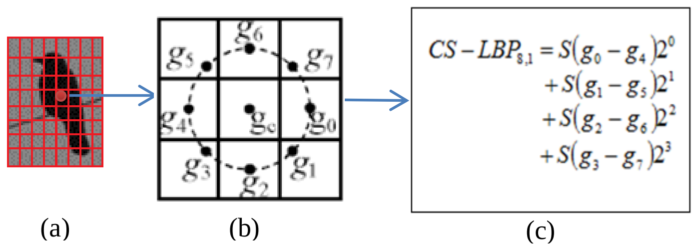

Figure 3. The CS-LBP feature extraction process. (a) The division of the input gray image into 64 blocks. (b) The 8 neighboring pixels on a circle of radius 1 of the pixel in the block. (c) The CS-LBP value of the pixel by comparing the gray value of the neighboring 8 pixels

C. Combination of HOG and CS-LBP

Based on HOG and CS-LBP features extracted, we augmented the feature vector, which combines the HOG feature with the block-structured CS-LBP feature. HOG for its robustness in capturing local shape or edge directions at different orientations and CS-LBP for its tolerance to illumination changes, robustness on flat images and its computational efficiency, we believe that the combination of these two feature sets can lead to a better bird feature capturing. Moreover, this extended feature set has less feature dimension in comparison to the state of the art algorithms. Such as, HOG-LBP feature descriptor and Multilevel HOG-LBP detector, though in the later detector PCA is used to reduce the dimension of the feature vector. This is because the computation of the CS-LBP reduces the dimension of the LBP calculation by only using the gray values of the center-symmetric pairs of pixels. The proposed HOG CS-LBP detector attained a better detection performance as compared to HOG-LBP and Multilevel HOG-LBP detectors. Table I. shows the detection results achieved using our HOG CS-LBP, HOG-LBP and Multilevel HOG-LBP combined feature vector descriptors on our bird dataset. Also, it shows the dimensions of the feature vectors for each scanning window for the above mentioned detectors, while the dimension of Multilevel HOG-LBP is before PCA is applied to reduce the dimension of the feature vector.

III. EXPERIMENTS

Large populations of crow birds daily gather on power lines and rooftops at Nagaoka station, a city located in Niigata prefecture Japan. These groups of birds can cause high noise and can be detrimental to the residents near this area. Therefore, a set of images is captured under different lighting conditions using Nikon D80 digital camera to generate experimental data. Datasets for training and testing are prepared from those captured images. Our HOG CS-LBP based detector is tested on these different images.

A. Datasets

We created crow bird's dataset from various images taken over a certain period of time. The crow birds in the images may appear in different pose or orientation. The positive training datasets are prepared by cropping from the gathered test images. The size of each positive sample used for our combined HOG CS-LBP detector is 48x64 pixels. And for our experiment we used 2000 positive samples and some of them are shown on Fig. 4. The negative examples are randomly subsampled and cropped from the same set of images as the positives, except they do not contain any bird. We used 8400 negative samples in our experiment.

Figure 4. Some positive samples from our crow birds data set.

TABLE I. COMPARING HOG-LBP, MULTILEVEL HOG-LBP AND OUR HOG CS-LBP DETECTION METHODS

Methods	Feature vector dimension	Detection rate
HOG-LBP	6080-D	79%
Multilevel HOG-LBP	7980-D	84%
HOG CS-LBP	3328-D	87%

B. Classifier

LIBSVM [14] is selected as our classification tool. We use linear SVM (c=1) for training and classifying on the crow birds dataset. As in [4] first, we trained a preliminary detector on the positive training windows and an initial set of negative windows. Second, the negative images are exhaustively scanned for false positives (hard examples). Then, the linear classifier is re-trained using this augmented training set (the cropped positive samples, initial negatives created by subsampling and hard examples). Finally, the input test image is scanned at all scales and positions for crow bird's detection. During this multi-scale detection a scale factor of 1.02 and pixel stride of 4x4 is used. Multiple overlapping detections are merged using [10].

C. Detection Evaluation

The evaluation methodology we used to quantify the detection performance of our detector is the detection rate versus false positive per-image (FPPI) curves [8]. Which plots detection rate along the y-axis and FPPI along the x-axis. We evaluate the detection performance of our HOG CS-LBP detector on various images. First, we evaluate the performance of the HOG-LBP and Multilevel HOG-LBP detectors on our test images. Then, we test our HOG CS-LBP based detector on the input images. The detection results are shown in Fig. 5. We achieved a detection rate of 92% at 0.02 FPPI on the tested images using our HOG CS-LBP feature detector. And the HOG-LBP feature detector achieved a detection rate of 84% at 0.02 FPPI while, PCA based Multilevel HOG-LBP attained a detection rate of 89% at 0.02 FPPI.

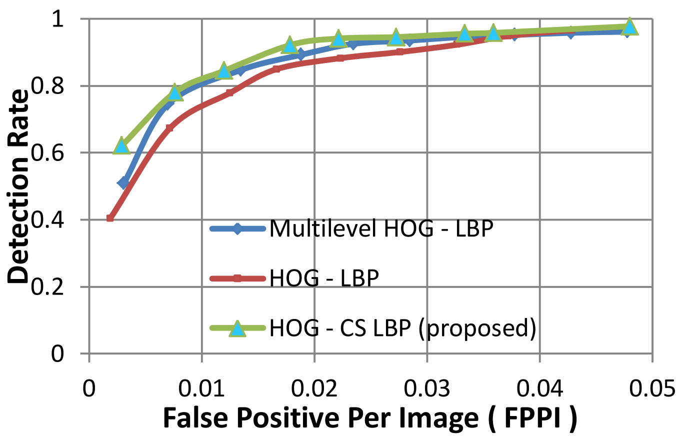

Figure 5. Detection rate versus false positive per image (FPPI) curves. Performance comparison between our HOG CS-LBP based detector, HOG-LBP and Multilevel HOG-LBP based detector on crow birds.

The detection rate of our detector at 0.005 to 0.04 FPPI is around 87% and that of the HOG-LBP and Multilevel HOG-LBP based detectors within the same range of FPPI is about 79% and 84% respectively.

In comparison to the HOG-LBP and Multilevel HOG-LBP feature detectors our proposed detector improves the detection rate significantly. Fig. 6 indicates the detection results we achieved on sample images using our detector under various lighting conditions.

Figure 6. Detection results of some sample test images under different lighting conditions

IV. Conclusion

In this paper, a combination of feature sets based on HOG and CS-LBP was presented. HOG for its ability for capturing local shape information and the block structured CS-LBP as a complement for texture information and gradient based features have resulted in achievement of better detection performance. Experimental results on the prepared crow dataset show that HOG CS-LBP feature combination based approach outperform the HOG-LBP based feature combination and attained slightly higher performance as compared to Multilevel HOG and LBP detector.

Future work on our detector includes dimension reduction of the block-structured CS-LBP by considering the uniform CS-LBP patterns only, and incorporating motion information using block matching or optical flow for real time applications.

References

[1] Y.Li, Y.Yin, and B.xu, "Bird objects detection and tracking on wild field circumstance",world congress on computer science and information engineering, pp.533-536,2009.

[2] A.kembhavi, R.Farrell, Y.Lou,D.Jacobs, R.Duaraiswami, and L.Davis,"Traking down under: folloeing the satin Boerbird",in proceedings of IEEE workshop on Applications of Computer vision, copper.mountain,co,USA,2008.

[3] M.Heikkila and C.Schmid,"Description of interest regions with Local binary patterns", Patt. Recogn., vol.42, no.3, pp.425-436,2009.

[4] N.Dalall and B.Triggs,"Histograms of oriented gradients for human dtection",in Proc. IEEE Conf. Comp. Vis. Patt. Recogn., pp.886-893,2005.

[5] X.Y.wang, X.Han and S.C.Yan,"An HOG-LBP human detector with partial occlusion handling", in ICCV,pp.32-39,2009.

[6] T.Ojala,M.Pietkainen and T.maenpaa,"multiresolution grayscale and rotation invariant texture analysis with local binary patterns", IEEE Trans. Pattern Anal.Mach.Intell,pp. 971-987,2002.

[7] T.Ahonen,A.hadid, and M.Peitikainen,"Face deescription with local binary patterns:application to face recognition",IEEE Trans. Pattern Anal.Mach.Intell,pp.2037-2041,2006.

[8] Y.Zheng,C.Shen,and R.Hartley,"Effective pedestrian detection using center-symmetric local binary/trinary patterns",in ACCV,pp.281-292,2010.

[9] C.Zeng and H.Ma,"Robust head-shoulder detection by PCA-based multilevel HOG-LBP detector for people counting",in international conference on Patt. Recogn.,pp.2069-2072,2010.

[10] D.Comanicu and P.Meer,"Meanshift analysis and applications",IEEE international conference on computer vision, pp.1197-1203,1999.

[11] I.T.Joliffe,"Principal component analysis", Springer-Verlag,1986,

[12] T.Ojala, M.Pietikainen, and D.Harwood,"A comparative study of texture measures with classification based on featured distribution",Pattern Recogn.,vol.29,n0.1,pp.51-59,1996

[13] X.Tan and B.Triggs,"Enhanced local texture feature sets for face recognition under difficult lighting conditions", IEEE Trans. Image Process., vol.19, no.6, pp.1635-1650.2010.

[14] C.C.Chang and C.J.Lin, "LIBSVM: a library for support vector machines,2001.