MVA2011 IAPR Conference on Machine Vision Applications, June 13-15, 2011, Nara, JAPAN

A face identification method of non-native animals for intelligent trap.

Mayumi KOUDA, Masakazu MORIMOTO†, Kensaku Fujii
Division of Computer Engineering, University of Hyogo
Shosya 2167, Himeji, 671-2280, JAPAN
†morimoto@eng.u-hyogo.ac.jp

Abstract

In this paper, we develop a face identification method to distinguish target non-native alien animals from other native animals using camera captured images. When a camera recognizes that target animal walked in the cage, it traps the animal in the cage. Here, we set raccoon as target non-native animal, and detect its face region by using HOG features. However, the raccoon face detector often confused by raccoon dog, which is a native animal to be preserved. So, after detecting raccoon face candidates, we distinguish them by several features and SVM. Some experimental results show that we can completely distinguish raccoon and raccoon dog from camera captured images.

1. Introduction

In recent years, various non-native animals are becoming wild and quickly multiplying in Japan. As a result, it not only harms local ecosystem but also damages to historical structures and farm products seriously. Therefore, many traps are set up to eliminate these non-native animals. The trap cage is a most common trap to catching target destructive animals, but it often causes a problem that precious native animals get caught in the trap and hurt, sometimes results in death.

To solve this problem, we introduce an intelligent trap cage system which can distinguish target destructive animals from other native animals by attached camera. It captures only a target animal and releases or disperses other animals. To realize this system, we develop a face detection and recognition method for wild animals. A bait and camera are placed at back side of cage. The camera captures wild animal's face when the animal comes to cage entrance.

In this paper, we select *raccoon* as a target non-native animal. Various damages caused by raccoons are increase rapidly. Moreover, some researchers insist that the number of raccoons might increase to 100 times in next 15 years. Therefore, it's urgent that extermination of wild raccoon. However, as shown in figure 1, the appearance of *raccoon* is very similar to *raccoon dog* ("Tanuki" in Japanese), which is designated as a protected native species. So, we have to distinguish these two animals.

Our intelligent raccoon trap first raccoon's face by using HOG features and SVM classifier from captured image. However, several raccoon dog's faces are also detected by the detector. So, face discrimination between raccoon and raccoon dog is done by SVM by using other features of input images.


(a) Raccoon


(b) Raccoon dog

Figure 1. Raccoon and raccoon dog.

2. Raccoon face detection

To detect raccoon's face, we use Histograms of Oriented Gradients (HOG) features and Support Vector Machine (SVM) classifier.

2.1. Histograms of Oriented Gradients (HOG)

HOG is a feature extraction method which represents shape of object [1]. It is robust over illumination change and local geometric change. HOG divides the image window into small spatial regions of size 5 x 5 pixels (*.cells.*), for each cell accumulating a local histogram of gradient magnitude and gradient directions over the pixels of the cell (Fig 2). Cells are normalized in the somewhat larger spatial region of size 3 x 3 cells (*.block.*).

Gradient magnitude m and gradient direction θ are calculated by the following equation.

$$m(x,y) = \sqrt{f_x(x,y)^2 + f_y(x.y)^2} \tag{1}$$

$$\theta(x,y) = \tan^{-1}\frac{f_y(x.y)}{f_x(x,y)} \tag{2}$$

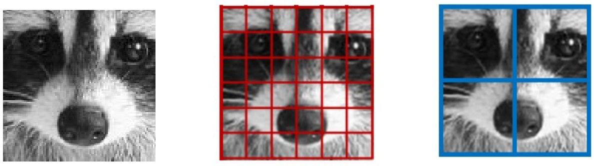
$$\begin{cases} f_x(x,y) = L(x+1,y) - L(x-1,y) \\ f_y(x,y) = L(x,y+1) - L(x,y-1) \end{cases} \tag{3}$$

Here, L is luminance of a pixel. θ is 0° to 360°, but it is changed into 0° to 180° here. θ is divided every 20 degrees. m and θ make a 9-bin histogram for each cell.

Then cells are normalized in each block from following equation.

$$v = \frac{v}{\sqrt{\left(\sum_{i=0}^{k} v(i)\right)^2 + \varepsilon}} \qquad (\varepsilon = 1) \tag{4}$$

Here, v is HOG feature, k is the number of HOG features in the block, ε is a coefficient for preventing division by zero problems.



(a) Input image (b) cell (5 x 5 pixel) (c) block (3 x 3 cell)

Figure 2. Image division for HOG.

2.2. Support Vector Machine (SVM)

SVM is a machine learning method for two class pattern recognition problem based on maximum margin strategy introduced by Vapnik[2]. SVM is a partial case of kernel-based methods, it maps feature vectors into higher-dimensional space using some kernel function, and then it builds an optimal hyperplane that fits into the training data. The solution is optimal in a sense that the margin between the separating hyperplane and the nearest feature vectors from the both classes (in case of 2-class classifier) is maximal.

SVM can handle non-linear hypotheses by introducing a kernel function such as polynomial kernel, RBF kernel and sigmoid kernel. In our method we use RBF kernel. RBF kernel nonlinearly maps samples into a higher dimensional space unlike the linear kernel, and it can handle the case when the relation between class labels and attributes is nonlinear. Furthermore, the linear kernel and the sigmoid kernel behave like RBF for certain parameters.

2.3. Raccoon face detector

To learn raccoon face features, we cut out raccoon's face region manually from 139 raccoon images included in caltech256[3]. Then we shift its clipping position in eight directions and rotate them from -45 to 45 degrees with 15 degrees interval. As a result, we get 3753 sheets as positive images. Also, we collect 7000 negative images from caltech256 clutter images. After that, all learning images are shrunk to 30x30 pixels and extract features by HOG. After normalizing features for every block, we perform learning by Support Vector Regression (SVR). It outputs degree of raccoon face similarity from input face candidates.

At detection stage, we set initial size of search window to 10x10 pixels and enlarge it 1.1 times after every raster scan to get raccoon face similarity. After all scanning, the window with highest similarity is selected as a detected face. If there aren't enough similar windows, the detector doesn't return detection result.

2.4. Experimental results of face detection

To estimate our race detector's accuracy, we collect *raccoon*, *raccoon dog*, *dog* and *cat* images from Web. Dog images are used from database images of caltech256. A raccoon face detection result is shown in Table 1. It shows that face detection accuracy by front face image of *raccoon* achieves 88.9%. Figure 3 shows some examples of raccoon face detection result. Figure 3(a) shows a correct detection result and figure 3(b) shows incorrect one. In fig.3(b), body pattern that looks like raccoon's face was detected.

By applying sideways face images of *raccoon*, it reduces until 67.7%. On the other hand, when we input faces images of other unalike species, *dog* and *cat*, face detection rate becomes low just as our wished. However, when we input face images of *raccoon dog*, it responds about 30%. These results show that, when we want to catch target raccoon only, we have to confirm whether the animal is actually raccoon or not at successive stage.

Table 1. Results of face detection.

	Number of inputs	Detected faces	Detection rate[%]
Raccoon (front face)	36	32	88.9
Raccoon (sideways)	31	21	67.7
Raccoon dog (front face)	39	12	30.8
Raccoon dog (sideways)	45	14	31.1
Dog	103	6	5.8
Cat	50	2	4.0

(a) Correct detection (b) Incorrect detection

Figure 3. Example of face detection result.

3. Face discrimination between raccoon and raccoon dog

Here, we discriminate face images between raccoon and raccoon dog by using SVM. At this stage, we use alternate features to distinguish animal faces. To evaluate discrimination accuracy, here we use 151 images, 67 raccoon images and 84 raccoon dog images.

3.1 Feature extraction for discrimination

As shown below, here we employ three types of feature extraction method from facial image, Pixel value, Discrete Cosine Transform (DCT) and Principal Component Analysis (PCA). We also apply some feature selection before discrimination step.

Figure 4. Raccoon's face images.

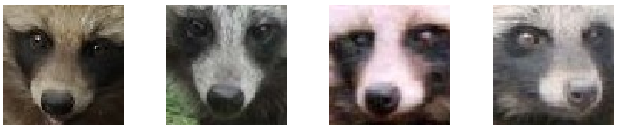


Figure 5. Raccoon dog's face images.

Pixel value: The simplest input vector is using pixel value. We normalize pixel value, as shown in Eq. 1.

$$V_{PIX}(x,y) = \frac{f(x,y) - p_{min}}{p_{max} - p_{min}} \quad (1)$$

where, p_{min}, p_{max} represent minimum and maximum value of the whole training image pixels, respectively.

Discrete Cosine Transform (DCT): The DCT is one of methods which transform input signals into frequency domain representation [4]. Most of the signal information tends to be concentrated in a few low-frequency components of the DCT. The 2-D DCT is given by the formula:

$$F(u,v) = \frac{2C(u)C(v)}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cdot \cos\frac{(2x+1)u\pi}{2N} \cos\frac{(2y+1)v\pi}{2N} \quad (2)$$

$$\text{where}, 0 \leq x,y,u,v \leq N-1, C(n) = \begin{cases} \frac{1}{\sqrt{2}} & (n=0) \\ 1 & (n \neq 0) \end{cases}$$

The base image of 2-D DCT of N= 8 is shown in fig. 6. The ingredient of u = v = 0 expresses the direct-current ingredient. It expresses a high frequency ingredient as u and v become large.

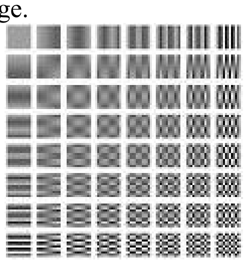


Figure 6. Base image of 2-D DCT (N=8).

Principal Component Analysis (PCA): PCA is the method of expressing many variables using δ principal component [5]. There is distribution in the feature space of input data as shown in Fig. 7. 1st principal component is a straight line which data distribution serves as the maximum through the center of gravity of data. 2nd principal component is a straight line which is perpendicular to 1st principal component and passes along the center of gravity of data. *d*-th principal component is derived similarly.
1st principal component have most amount of information. Below, the amount of information decreases as it becomes 2nd principal component, ..., *d*-th principal component. Therefore, the information on original data is expressed even if it does not use until d-th principal component. The information on original data is expressed by using from 1st principal component to δ-th principal component (δ<d). Thereby, we are compressible into the feature of a low dimension.

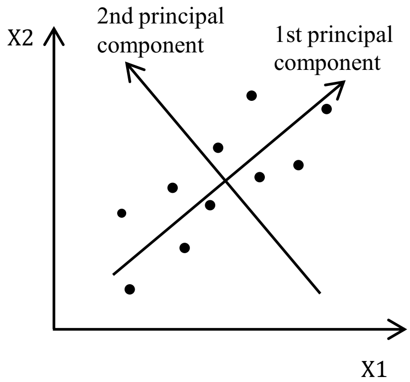


Figure 7. Principal component.

3.2 Feature selection

To improve discrimination accuracy, we apply some feature selection to sort out effective features for identification from whole *d* features[6]. The number of combinations of features is 2^d. Since the combination of the feature will increase if d becomes large, it is difficult to verify all the combination. Therefore, the problem is whether to select the effective feature efficiently. In this paper, we use wrapper method. The wrapper method selects subset of the features using trial discrimination by learning dataset.

The searching method has two types, forward stepwise selection and backward stepwise selection. Forward stepwise selection is search method of adding the most effective feature one by one. Backward stepwise selection is search method of deleting the most unnecessary feature one by one.

3.3 discrimination

At discrimination stage, the input face images are shrunk to 8x8 pixels before extracting feature vectors. This procedure reduces vector dimension and variation of each input. Then we try three types of feature extraction method from facial image, Pixel value, DCT and PCA and two kinds of feature selection method, forward and backward wrapper method. Then, SVM is used again for discrimination.

3.4 Experimental results of discrimination

To evaluate discrimination accuracy, we apply 10-fold cross validation to discriminate raccoon and raccoon dog. Table 2 shows result of each feature extraction method with best feature selection. By applying feature selection, DCT achieve 100% discrimination for our dataset. The verbose results of feature selection for PCA and DCT, forward and backward methods are shown in Fig. 8. The DCT features with forward wrapper method achieve 100% discrimination rate during 18-20 components.

Some images restored from the selected 18 components of DCT are shown in Fig. 9. These figure show that by selecting features, it emphasizes facial difference between these two species, raccoon's horizontal black line around eyes and raccoon dog's vertical white line between eyes.

Table 2. Result of discrimination.

	Pixel value	DCT	PCA
Discrimination rate[%]	92.7	100	96.0

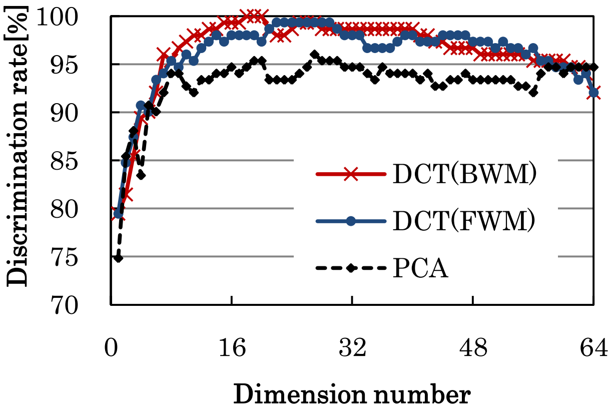


Figure 8. Discrimination rate of raccoon.

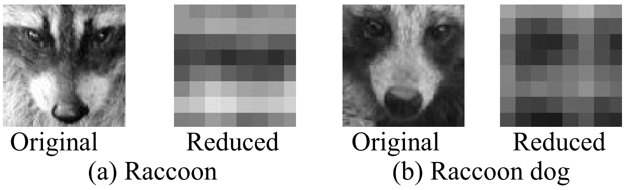


Figure 9. Effects of feature selection.

4 Conclusions

In this paper, we develop an animal face identification method to distinguish target non-native alien animal, raccoon, to realize intelligent trap cage. Our raccoon face detector using HOG and SVM detects raccoon face accurately, but it also detect raccoon dog at some level. Therefore, succeeding face discriminator using DCT features and SVM distinguish raccoon and raccoon dog stably. In future work, we are going to improve face detection accuracy and speeding up of face detection.

References

[1] Tomokazu Mitsui, Hironobu Fujiyoshi, “Object Detection by Joint Features based on Two-Stage Boosting,” Symposium on Sensing via Image Information, 2008.

[2] Takio Kurita, “Introduction to Support Vector Machine”

[3] Caltech256: http://www.vision.caltech.edu/Image_Datasets/Caltech256/

[4] N. Ahmed, T. Natarajan, K. R. Rao, “Discrete cosine transform,” IEEE Transactions on Computers, vol. C-23, pp. 90-93, 1974.

[5] I. T. Jolliffe, ”Principal Component Analysis” ,2002

[6] Kohavi, R, John, G, “Wrapper for feature subset selection, ” Artificial Inteligence, pp. 273-324,1997.