2S-03

Incremental Training of a Human Detector for Office Environment

Tatsuya Hanyu† and Qiangfu ZHAO

Department of Computer Science, University of Aizu, Fukushima, Japan
E-mail: {m5211141,qf-zhao}@u-aizu.ac.jp

Abstract: To design a good human detector, we may collect a huge number of data, and train the detector off-line. However, even if the training data set is very large, it may not contain enough information for some particular environment, and the obtained model may not work well. In this paper, we study incremental learning of a support vector machine-based human detector in an office environment, and investigate the "growth process" of the detector. Experimental results show that it is possible to obtain a good human detector customized to a certain environment with less data via incremental learning.

Keywords: human detection; support vector machine; incremental learning.

1. INTRODUCTION

To design a good human detector, we may collect a huge number of data, and train the detector off-line. However, even if the training data set is very large, it may not contain enough information for some particular environment, and the obtained model may not work well. In this paper, we study incremental learning of a support vector machine-based human detector in an office environment, and investigate the "growth process" of the detector. Experimental results show that it is possible to obtain a good human detector customized to a certain environment with less data via incremental learning.

In fact, human detection must be performed in real time. Therefore, high efficiency and high speed are required. As mentioned above, training a detector offline with a huge training set often results in a very large and inefficient system. Updating this large-scale system in real time also takes time. But starting with a relatively small training set, trying to add the necessary data in real time can solve the problem. In this study, this type of training process is called incremental training.

2. THE PROPOSED INCREMENTAL TRAINING METHOD

In the following, we discuss the incremental method for improving the accuracy of the SVM-based human detector using new images. In the process of human detection with the human detector, we get a number of sub-images that do not contain human at all. These sub-images produce false positive errors, and they are "near-miss" data useful for improving the current detector. These data are added to the training set as negative examples.

In this study, we design an SVM detector using the following method.

1. Prepare a small training set $\omega_{training}$ and a test set $\omega_{testing}$.
2. Construct an SVM from $\omega_{training}$.
3. If $\omega_{testing}$ is empty, stop;
4. Take(without return) a test image from $\omega_{testing}$, and detect the human using the SVM;
5. Calculate the accuracy of human detection. If the accuracy is less than a given threshold T, return to step 3.
6. Add sub-images corresponding to the false points (both negative and positive) to $\omega_{training}$, and return to step 2.

In the experiments, the number of positive training data is 1,000, and the number of negative training data is the same. The number of testing images is 100. In practice, we may need many images to get a good human detector. In this experiment, we just want to examine the trend of the incremental training process. Therefore, we prepared only 100 images that include only one person in each image and examined how the detection rate changes.

In the proposed method, it is possible to efficiently generate the needed training data, so that the accuracy of human detection can be improved efficiently.

† is the presenter of this paper.

Fig. 1 Example images (pedestrian image and backgraund image)

3. EXPERIMENT

3.1. Database

We prepared 1000 positive human images and 1000 background images for training. For example, Fig. 1 shows part of the images.

For each training image, the height is 36 pixels, and the width is 18 pixels. All training images are gray scale images. And we prepared 100 testing images. These 100 images are photos taken in the same laboratory space. They have bigger sizes than the training images, but they contain only one (the same) person. And also we prepared the test data obtained after meanshift-based filtering. Meanshift image filtering method is one of image filtering method. An example is shown in Fig. 2.

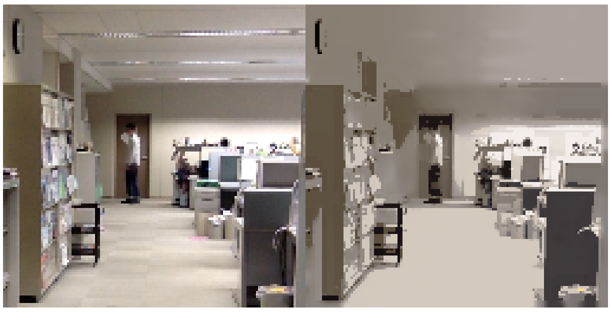

Fig. 2 Meanshift Image Filtering (left image is original image; right image is meanshift filtering image

Fig. 3 Results Images

Fig. 4 Results Meanshift

We tested this method because, We did not know whether HOG method was enough or not. Thus we also research that method.

3.2. Experiments

In addition, we prepared 1000 human images and 1000 background images to train an SVM detector off-line. That is, we use only 2000 images for off-line training. The reason why we construct the SVM using only 2000 training data is because we need to add new data, and we want to add only needed data.

The results of this experiment are shown in Fig. 3. In the figures, "before training" means the database ω has not been updated; and "after training" means the database ω has been updated with needed data obtained from a newly observed image. The result of incremental learning "grow process" are discribes in Fig. 3. At first, we tried human detection with the SVM obtained off-line. And the result are left image of "image1" in Fig. 3. That SVM could not detect human well. From the results, we can see many images that are just background images are recognized as human. Look at Fig. 3 before training, all pointers are judged as human by the SVM.

And we compare the each accuracy of original image experiments and meanshift image experiments. One of this experiment is describe at Fig. 4. The upper line shows the accuracy trend that used normal image for test image of proposed method. And lower side line shows the accuracy trend of meanshift image filtering one. Those results are not much change. But normal image experiment gave a slightly better result.

4. DISCUSSION

We thought meanshift result is better, but the result was normal image result was better. We think this because HOG method can exclude noise data. Also the meanshift method must cut the necessary information of feature data.

5. CONCLUSION

Based on Experiments parts, we think the proposed method (Incremental Training of a Human Detector for Office Environment) improved that accuracy of the constructed SVM. And meanshift image filtering method is not match for this experiments. If we use more high resolution image for this experiments, of course the meanshift image filtering method must be usable.

REFERENCES

[1] Vapnik, Vladimir Naumovich, and Vlamimir Vapnik. Statistical learning theory. Vol. 1. New York: Wiley, 1998.

[2] T, Mitsui, 'Human Detection by Two Stages Adaboost with Joint HOG' http://www.vision.cs.chubu.ac.jp/ yuu/CV/paper/SSII2008.pdf

[3] scikit-learn 0.19.1 http://scikit-learn.org/