Robust Motion Detection and Tracking of Moving Objects using HOG Feature and Particle Filter

Bhuyain Mobarok Hossain *bhuyain001@gmail.com*
Faculty of Engineering / Systems Innovation Engineering
Tokushima University
Tokushima, 770-8506, Japan

Stephen Karungaru *karunga@is.tokushima-u.ac.jp*
Faculty of Engineering / Systems Innovation Engineering
Tokushima University
Tokushima, 770-8506, Japan

Kenji Tereda *terada@is.tokushima-u.ac.jp*
Faculty of Engineering / Systems Innovation Engineering
Tokushima University
Tokushima, 770-8506, Japan

Abstract

Detection and tracking of moving objects has gained significant importance due to intense technological progress in the field of computer science dealing with video surveillance systems. Human motion is generally nonlinear and non-Gaussian and thus many algorithms are not suitable for tracking. One of the applications to maintain universal security is crowd control. The main problem of video surveillance is continuous monitoring with regard to crime prevention. For security monitoring of live surveillance systems, target identification and tracking strategies can automatically send warnings to monitoring officers. In this paper, we propose a robust tracking of a specified person using the individuals' feature. The proposed method to determine automatic detection and tracking combines Histogram of Oriented Gradient (HOG) feature detection with a particle filter. The Histogram oriented Gradient features are applied to single detection window for the identification of human area, after we use particle filters for robust specific people tracking using color and skin color based on the characteristics of a target individual. We have been improving the implementation, evaluation system of our proposed methods. In our systems, for experiments, we choose structured crowded scenes. From our experimental results, we have achieved high accuracy detection rates and robust motion tracking for specific targets.

Keywords: Video Surveillance System, Robust Tracking, HOG feature, Particle Filter.

1. INTRODUCTION

Human identification and tracking have been focused on for many years as per the perspectives of various areas of application such as video surveillance, activity analysis, intelligent control and human-computer interaction [1]. At present Visual surveillance is an essential research in the field of computer vision, in a video-based surveillance system. In video surveillance systems the most important research is human detecting and tracking in the crowded areas. The crowded scenes include public places such as airports, bus station, concerts, subway, religious festivals, football matches, railway stations, and shopping market; where many people gather is a challenge for those interested in safety and security systems. In this case, for video surveillance we need to see a specific place or area in the crowded area. One of its applications is in crowd control to maintain the general security in public places. In the crowded area some people commit crime and go to and from a location severally. Our system aims at tracking such persons.

Crowded scenes can be divided into two categories: Structure crowded scene and unstructured crowded scene. In structured scene, people are moving one direction like near elevators, and stations. Unstructured scenes are very complex because the movement of the crowd and direction of movement is random and overlap over time [2]. The objective of crowded research is that the result be applied frequently and instantly by other objects. Visual excuse, ambiguity in crowded scenes, complex behaviors and scene imagery make analysis difficult. In our research, we have focused on the criminal problems, detecting and tracking. [3]. Figure1, shows structure and unstructured crowded scenes. In our research, we choose a structured crowd scene for experiments.

FIGURE 1: Shows (a) Unstructured and (b) Structure Crowded Scene.

2. RELATED WORK

Moving objects detection and tracking is a very important research field today in video surveillance systems. Recently a lot of methods proposed for detecting and tracking specific target in crowded areas in the field of computer vision. Bhuyain et. al proposed detecting and tracking specific person in crowded areas. In their method for detecting and tracking specific person, they apply Histogram oriented Gradient and particle filter [3]-[5]. Agarwal and Cai provided an analysis of movement, tracking and recognition of human activities [4] M. M. Naushad Ali et al proposed robust tracking algorithm based on particle filter. Multiple moving human detection and tracking using frame difference and morphological operation in a video sequence is discussed in [5]. Dalal & Triggs have proposed an excellent method for human detection. Its dense network method uses histograms oriented of gradients (HOG), calculated at 16 x 16 pixel size blocks to represent the detection window. This representation proved to be strong enough to classify humans using SVM linearity [6] Qi, Z.Ting, et al is using Particle Filter to track people detected advance with other techniques to enhance results and solve contour information problems [7]. Bristow and S. Lucey have provided the tracking system targets for certain objects, such as cars or pedestrians, and may require the use of primary knowledge about the shape of the target object [8]. Myo Thida et al. presented a macroscopic model, microscopic model, and crowd event detection [9]. Li Teng, et al. provided an analysis survey which describes the behavior of crowd area and their activities [10]. Hu et al. have also presented motion detection, tracking and behavioral understanding [11]. Ruiyue Xu implemented multiple human detections and tracking based on head detection for real-time video surveillance [12]. Malik Souded et al, was first concerned about the creation of tracking targets and he tried many approaches, from object detection in still images, to the configuration by motion detecting and assembling, through models utilized on whole or parts of objects [13]. Di Xie et al proposed a powerful new visual surveillance system that contributed two main ingredients: human head detection and goal tracking [14]. Mohammed Lahraichi, Khalid Housni, Samir Mbarki have proposed moving objects that are represented by points or delimiter areas discovered before or during the trace, and correspondence between objects is done across frames by creating a relationship between points or area features [15]. D Reid provided tracking for moving objects, measuring motion parameters and obtaining a visual record of a moving object as an important area of application in image processing [16].

3. PROPOSED METHOD

In the crowded areas moving objects detecting and tracking is a very important research field today. In the video surveillance systems, manual continuous monitoring by an officer is impossible. Therfore, automated human detection and tracking can improve security systems in the video surveillance system. Crowded scene has a lot of people with different characteristics and features such as hair colors, heights, weights, clothes, skin colors, and so on. Our proposed method is divided into two parts: human detection and tracking a specific person. Before detecting and tracking targets, we set target features, already defined by the operator, for example, blue shirts and wearing masks and glasses. Those features are used for tracking specific target. In our proposed method for robust human detection and tracking we apply histogram oriented gradient with SVM to detecting multi human in the video. When SVM successfully detects multi human area, we apply the particle filter for tracking specific target persons in the crowded area. Our proposed method is described in the following flowchart.

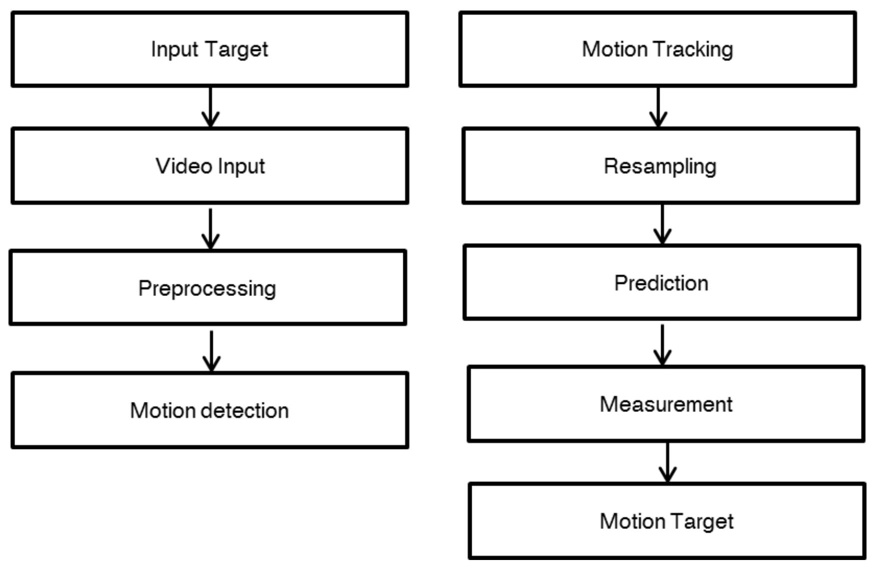

FIGURE 2: Motion Proposed Method Flowchart.

3.1 Pre-processing

Preprocessing means processing of raw data to prepares for another processing. We use color space for image processing in the first step of our process. We have a lot of color spaces like, RGB, HSV, YIQ, YUV, CMYK, YCbCr, etc. For skin color detection HSV and YCbCr color spaces are better compared to the RGB color space. In our method we use the HSV color space in order to detect and track specific person. For this reason, first we transformed RGB color space to HSV color space, because the RGB color space is not robust: it is affected by the environmental condition when illumination changes. The HSV color space is useful for detection and tracking to extract the color information of the target such as skins, clothes, hairs, and shoes. In the flowchart, at setting the features of target step, we set the parameters of the targets information as hue and brightness selecting from HSV color space. Figure 3, shows an example skin color detection result.

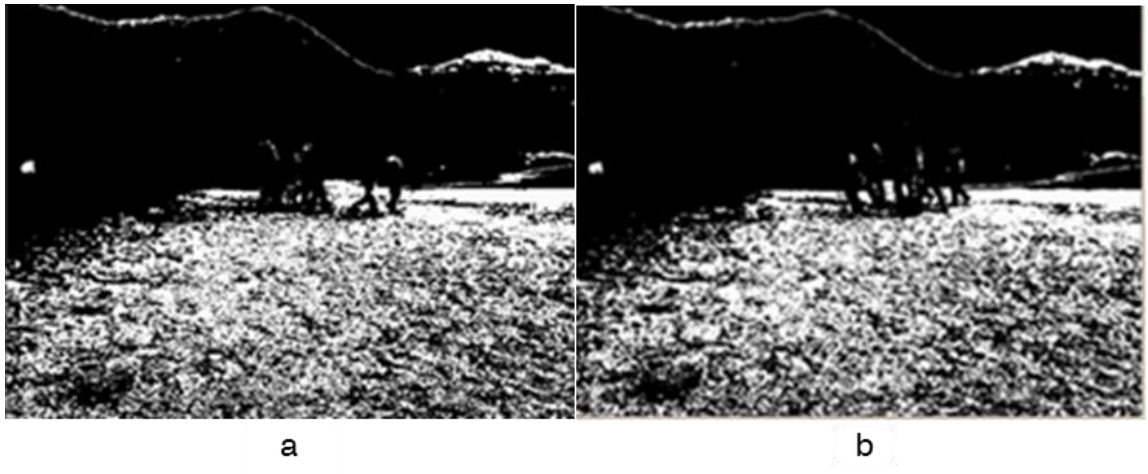

FIGURE 3: Skin color detection.(a) People are moving for from camera and go one direction (b) People are moving far from camera come back one direction.

3.2 Human Detection

The histogram of oriented gradients (HOG) is a great feature descriptor used in computer vision and image processing for the purpose of object detection. As stated before, the technique counts occurrences of histogram oriented gradient in localized parts of an image. This method is similar to the histogram trend direction, fixed attribute conversion descriptors, and shape contexts, but differs in that it is calculated on a dense network of cells with equal spacing and uses overlapping local contrast to improve accuracy [17]. A lot of methods have for learning the HOG feature. But in this work we choose the SVM because we want classify into only inside datasets moving human or not. A Support Vector Machine (SVM) can be linear or non-linear. A Support Vector Machine (SVM) is a discriminatory classifier formally defined by a separate contradiction. SVM is efficiently applied to a HOG specifier for human identification [17]. For the HOG descriptor, the important features in the scene are calculated using horizontal and vertical kernels using gradient magnitude and orientation features. SVM parameters have already been learned especially for human detection before starting the algorithm; it reduces the arithmetic time of the learning process [6, 17]. In this method, we us HOG detection by setting parameters for block size, cell size and gamma correction. After successfully detecting the human area, they are all surrounded by a rectangle including the specific target.

3.3 Particle Filter for Motion Target Tracking

A particle filter is a technique without optimization for performing Bayesian recursion using Monte Carlo approximations. The particle filter processing has several steps, re-sampling, predicting and updating weights, and measuring on the basis of probability function. Our main motivation of using a particle filter is to track one or several targets [3]. We applied the particle filter on the results of the HOG-SVM. Our particle filter applies the Bayesian filter by sampling Monte Carlo, which is the representation of the inferior density of a set of random particles with their associated weights as well as the calculation of the estimates [19]-[20].

In our approach, the Bayesian filter is used to measure the targets of previous and background density functions with a set of features. Then, the generated particles proceed to the specific targets of the individual's set of characteristics. To set up the particle filters, we use the number of particle numbers, the number of scanning range for the probability function as the parameters to make gaps random. When we successfully track targets, particles gather around the target area.

4. EXPERIMENTS

We proposed our method for robust detecting and tracking in crowded scene, a robust motion specific person, that's our main objective for experiments. In this experiment we use only structure crowded scene for our object. For robust detecting and tracking we assumed structure crowded situation where the occlusion of human body occurred due to the moving person. The occlusion of human body is a big and difficult problem in crowded scenes as we mentioned previously. For experiments we create our datasets outside of laboratory to evaluate our method.

The specification of the experiment environment is shown in Table 1.

Computer	Intel Core i7,Memory 8GB
Camera	Logical USB Camera
Resolution	640x340,30 fbs
Software	Visual C++, OpenCV

TABLE 1: Specification of Experiment Environment.

4.1 Experimental Setup

We used two datasets for our experiments. Dataset1 has a lot of feature like car, bicycle. Dataset 1 we use for cheek if our SVM classify only moving human or not. Dataset 2 we use for detecting and tracking specific person robustly by motion. All people have different feature like, black color pant, black color shirt, red color shirts, blue color shirts, and white color. Environmental conditions of these datasets are natural. In this dataset targets person wear blue color shirts and black color pants. We set the target as person. We specified that the target's appearance as skin color and blue shirts.

The datasets used in the experiments are shown in table 2.

Dataset	Total No.frames	Time duration
Dataset1	2493 frames	73 sec
Dataset2	4800 frames	1.62 sec

TABLE 2: Test Datasets for Experiments.

4.2 Parameter Discussion

In our experiment, we use two methods for our specific goal. The histogram oriented gradient (HOG) and particle filter. In our system first, we use Histogram oriented gradient (HOG) and SVM for detection human in video surveillance system. In our system we set HOG feature parameter like block size (16x16) and cell size parameter (8x8). Many methods exist for learning HOG feature for example, neural network and SVM, etc. In this work we choose the linear SVM because we want to classify into only 2 classes, human have or not. When we find a human area on our system we run a particle filter to track certain people moving robustly.

To set up a particle filter, we use the number of particles, the variance of random particles, and the scan range of the probability function as parameters. In the experiments, the number of syntheses of particle filter parameters has been set to 200, optional to set the particle at 8.0 and randomize scanning range of 20 second 20 pixels set. Particle filter weight 1.0 and radius 1.0 and color of particle filter green dots.

4.3 Experimental Results

To evaluate our system, we apply our proposed method to the datasets. HOG feature results are surrounded by a blue rectangle in the image which area have human. Figure 4 shows experimental for human detection. First, we experiment on dataset1 which datasets have a lot of features like, car, human, motorbike, bicycle, etc. We will check in this datasets our system classify only human or not. Second, we use dataset2 for tracking robust motion specific person in the crowded area.

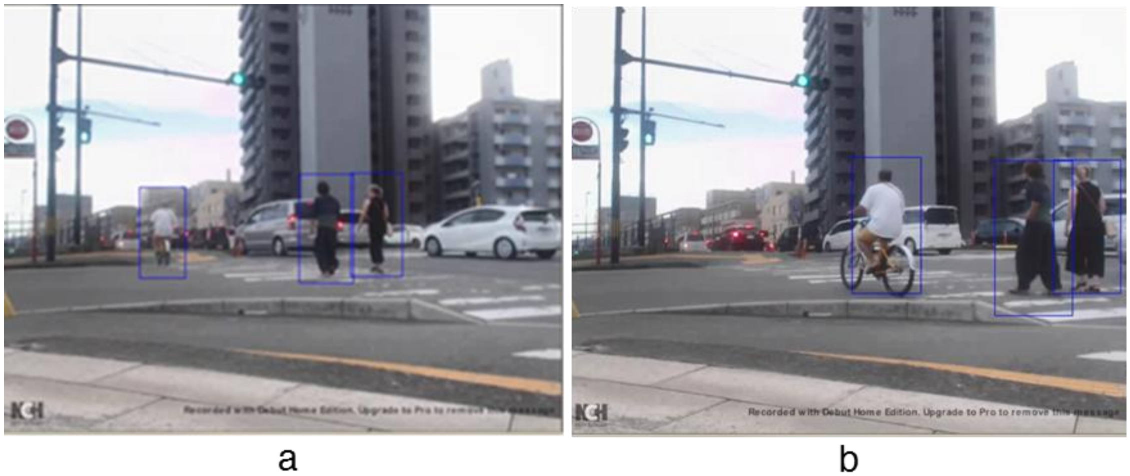

FIGURE 4: SVM classify the human only and not any other object with multiscale (a) human far from camera (b) human near of camera.

In our system we saw Histogram oriented Gradient (HOG) and A Support Vector Machine (SVM) detecting moving person perfectly. But when we experiment dataset 2, we still had an occlusion problem. To solve this occlusion problem, we combined Histogram oriented Gradient (HOG) and particle filter. Where there is occlusion problem our system, particle filter can track the moving specific person. Then we combined HOG feature and particle filter to detecting and tracking robust specific person perfectly.

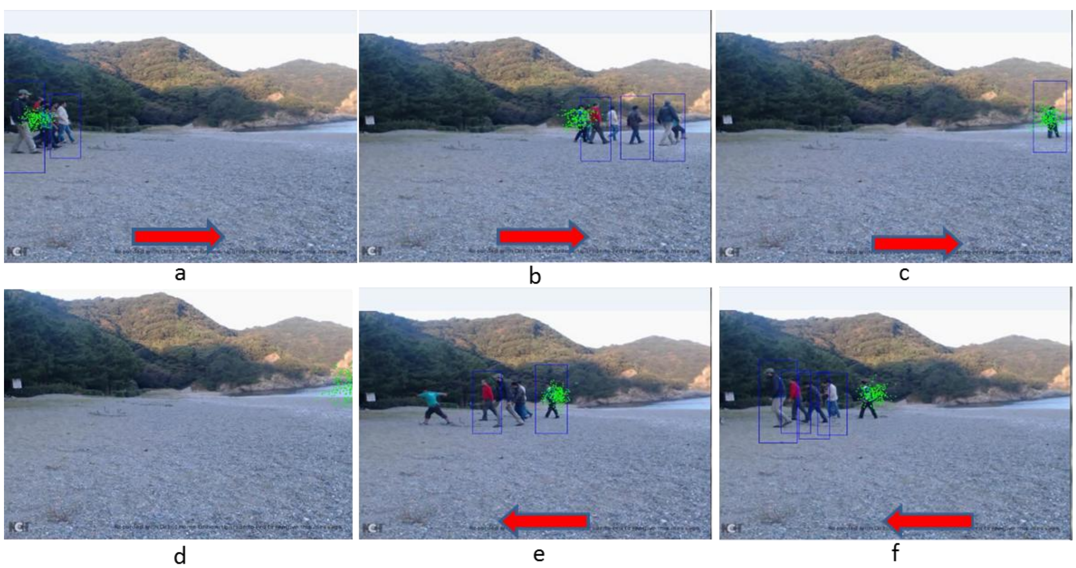

FIGURE 5: Detecting and Tracking robustly a specified person running on structured crowd scene experimental (a) Image particle filter (b) Particle filter start tracking (c) Particle Filter continue Tracking. (d) People are out of frame and Particle filter are initial, (e) Particle filter tracking specific person again, (f), Particle filter tracking specific.

The experimental results show moving human are successfully detected by the HOG feature and surrounded by a blue rectangle in the image. Particle filters tracks the specific person strongly.

In our experiment dataset2 present our HOG-SVM classifying only human or not and dataset 1 present HOG-SVM and particle detecting and tracking specific person robustly.

Table 3 shows the results of success and failure rate of the detecting and tracking specific robust moving.

Dataset	HOG feature detection [%]		Particle Filter [%]	
	Success	**Failure**	**Success**	**Failure**
Dataset 1	90.00	10.00	99.00	1.00
Dataset 2	95.00	5.00	99.00	1.00

TABLE 3: Evaluation of Proposed Method.

These are calculated as follows:

Success = [Number of success detection frames / Total number of frames] * 100 [%],

Failure = 100 [%] – Success.

In our experimental results, we have succeeded in detecting more than 90% of the human detection in the crowded area, and tracking a specific person more than 98% in our data set. The HOG detection failure rate was high due to blockage, tracking the target particle filter rather than detecting the HOG. From these results, we confirmed that our proposed method combines HOG-SVM and particle filter with a strong target feature towards the occlusion problem.

5. CONCLUSION

In this paper, we proposed an innovative method combining the detection of the HOG feature and the particle filter to track the robust motion specific target people. Based on the evaluation results using the data set including the closure problem, our proposed method can be applied to different scenes where many people cross.

We proposed two guidelines for the future expansion of work: Improvements to the detection and tracking algorithms of the current system for protecting the elderly and the socially vulnerable people, and the extension system to create smart applications of high-level robust motion tracking in crowded scene.

6. REFERENCES

[1] A. E. Dubois and A. Mitiche, “Real-time system for high-level video representation: application to video surveillance“, Proc. of SPIE Int. Symp. on Electronic Imaging, Conf. on Visual Comm. and Image Proc. (VCIP), Santa Clara, USA, 5022 (2003), 530-541.

[2] Bhuyain Mobarok Hossain, Stephen Karungaru, Kenji Terada, Akinori Tsuji,”Human Detection and Tracking Using HOG Feature and Particle Filter in Video Surveillance System,” International Journal of Advanced Intelligence Volume 9, Number 3, pp.397-407, January, 2018.

[3] Bhuyain Mobarok Hossain, Stephen Karungaru, Kenji Terada, Akinori Tsuji,”Real time Specified person Tracking in a Crowded Scene Using Particle Filter,” The International Conference on Electrical Engineering (ICEE2017),July,4-7,2017 (Weihei, China).

[4] M. Wegmuller, J. P. von der Weid, P. Oberson, and N. Gisin, “High resolution fiber distributed measurements with coherent OFDR,” in Proc. ECOC’00, 2000, paper 11.3.4, p. 109.

[5] Aggarwal, J.K., Cai, Q., “Human motion analysis: a review“ Computer Vision and Image Understanding ,Volume 73, Issue 3, 1 March 1999, Pages 428-440.

[6] M. M. Naushad Ali,M. Abdullah-Al-Wadud, et al," Moving Object Detection and Tracking using Particle Filter," Applied Mechanics and Materials," ISSN: 1662-7482, Vols. 321-324, pp 1200-1204,

[7] N. Dalal and B. Triggs. "Histograms of orient-ed gradients for human detection". Conference on Computer Vision and Pattern Recognition (CVPR), 2005.

[8] Bristow and S. Lucey,"Why do linear svms trained on HOG features perform so well?," CoRR, vol. abs/1406.2419, 2014.

[9] S. Munder, C. Schnoerr, and D. M. Gavrila. Pedestrian Detection and Tracking Using a Mixture of View-Based Shape-Texture Models. IEEE Transactions Intelligent Transportation Systems, 9(2):333{343, 13, 2008.

[10] Li, Teng, et al. "Crowded scene analysis: A survey." IEEE transactions on circuits and systems for video technology 25.3: 367-386, 2015.

[11] Hu, W., Tan, T., Wang, L., Maybank, S.: A survey on visual surveillance of object motion and behaviors. IEEE Trans Rev. 34(3),pp 334–352 , 2004.

[12] Ruiyue et al," Multiple human detection and tracking based on head detection for real-time video surveillance," Journa lMultimedia Tools and Applications archive Volume 74 Issue 3, Pages 729-742, February 2015.

[13] Malik SOUDED et al," An Object Tracking in Particle Filtering and Data Association Framework, Using SIFT Features, "International Conference on Imaging for Crime Detection and Prevention(ICDP)(2011).

[14] Di Xie, Lu Dang, Ruofeng Tong," Video Based Head Detection and Tracking Surveillance System," 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2012).

[15] Mmohammed Lahraichi, Khalid Housni, Samir Mbarki," Visual Tracking Using Particle Filter Based on Gabor Features," International Journal of Intelligent Engineering & System, Vol.11, No.4, January 15, 2018.

[16] D. Reid, "An algorithm for tracking multiple targets," IEEE Trans. on Automation and Con-trol, Vol. AC-24, December 1979, p84-90.

[17] Qi, Z., Ting, R. Husheng, F. Jinlin, Z.: "Particle filter object tracking based on Harris-SIFT feature matching". Proc. Eng. 29, 924? 929 (2012).

[18] Bahera H. Nayef," A Comparison between Linear and Non-Linear Machine Learning Classifiers," Journal of Al-Nahrain University, Vol.19 (2), June, 2016, pp.145-153.

[19] Haofu Liao, "Human detection based on Histograms Oriented Gradients and SVM". EECE-7373, Pattern Recognition, SPRING, April 23, 2013.

[20] Yizheng Cai, Robust Visual Tracking for Multiple Targets. B.E., Zhejiang University, 2003 ECCV: 9th European Conference on Computer Vision, Graz, May 7-13, 2006 proceedings, part IV (pp.107-118).