Usage of HoG (Histograms of Oriented Gradients) Features for Victim Detection at Disaster Areas

Yücel Uzun[1], Muhammet Balcılar[2], Khudaydad Mahmoodi[3], Feruz Davletov[4], M. Fatih Amasyalı[5], and Sırma Yavuz[6]

[123456]Yildiz Technical Univeristy, Computer Eng. Dept. İstanbul, Turkey
yuceluzun@windowslive.com, muhammet@ce.yildiz.edu.tr, khudaydad222@gmail.com, fdavletov@gmail.com, mfatih@ce.yildiz.edu.tr, sirma@ce.yildiz.edu.tr

Abstract

Employing robot teams at disaster areas requires usage of autonomous navigation methods. Moreover, autonomous navigation requires autonomous victim detection. Human skin color based victim detection methods may not be robust since the variations of the lightening conditions at disaster areas. Histograms of Oriented Gradients (HoG) were presented as an alternative way of human detection. In literature, HoG based methods proved their efficiency on the datasets including upright humans. But, the victims have very large variation of poses at a disaster area. In this work, the efficiency of HoG based methods was investigated on a dataset including very different poses and lightening conditions. We have reached 95% success on automatic victim detection problem.

1. Introduction

At disaster areas, the usage of the robot teams for the purpose of victim detection and rescue was became very popular research area. This area includes several problems such as autonomous robot navigation [1], task management [2], communication between robots, self-localization, map building [3], real time processing of huge sensor data, victim detection. Since 2006, RoboCup Virtual Robot Search and Rescue league is organized to address these problems [4]. Moreover, DARPA has recently announced Robotic Challenge to develop ground robots capable of executing complex tasks in dangerous, degraded, human-engineered environments [5].

USARSim, is a simulation environment used in RoboCup Virtual Robot Search and Rescue competitions [6]. In this competition, finding victims in an unknown environment using a robot team is aimed. Autonomous navigation methods are employed because manual navigation of several robots is not effective. Automatic victim detection is very important when autonomous navigation is used. Because, a robot can pass a victim without detecting him/her if it has not a automatic victim detection capability. In USARSim environment, only camera sensor can be used to detect a victim.

In real disaster areas, there are very similar problems to the simulation environments especially if there are more number of robots then the robot operators. Because of this reason, the solutions for the simulation environment can be easily adapted the real disaster areas.

Human skin color is very useful cue in human detection. But, the color-based methods are not robust in disaster areas because of its diverse lightening conditions. Histograms of Oriented Gradients (HoG) are an alternative way of human detection [7]. In literature, HoG based methods proved their efficiency on the datasets including upright humans. But, the victims have very large variation of poses at a disaster area. In this work, HoG based victim detection system was developed. The efficiency of HoG features was investigated for the large variations of the victim poses.

2. Previous works

Object detection is an old area of image processing. Plenty of surveys about object detection [8, 9, 10] and early stages of the human detection exist [11].

There are various approaches for human detection. Papgeorgiou and Poggio [12] used Haar wavelet based application, combined with a polynomial SVM. Gavrila and Philomen [13] extracted edge images and matched them to a set of learned exemplars using chamfer distance. In several studies part-based approaches developed. [14, 15, 16]

Dalal and Triggs [7] used the HoG representation. This method is widely used for following studies [17, 18].

3. Overview of the Method

In the USARSim environment, methods for the variable problems are expected to work accurate and reliable. In addition they have to work fast in order not to block vital issues like the communication or control of the robots. We developed a simpler algorithm based on existing HoG methods because of the performance concerns.

In the first step of this method, color image is converted to grayscale image. To avoid the lightning differences, histogram of the image is equalized.

After the histogram equalization, in order to find the gradients of the image, Sobel Operator was applied with the following 3x3 kernels;

$$G_x = \begin{matrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{matrix} \quad G_y = \begin{matrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{matrix}$$

With the derivative values of the image, the weighted vote and the edge orientation are calculated. While $\mathrm{G} = \sqrt{\mathrm{G_x}^2 + \mathrm{G_y}^2}$ gives the weighted vote of the pixel, $\theta = \mathrm{atan}\left(\frac{\mathrm{G_y}}{\mathrm{G_x}}\right)$ gives the edge orientation.

In the next step, pixels are separated to the bins based on their orientation. The image is divided into the fixed size blocks and with the summation of the weighted values of the each pixel in bins, histogram of the block is calculated.

In this method, overlapping blocks and sliding window structures were avoided. The overlapping blocks increases not only amount of the calculations but also used memory space. It's possible to miss a victim with sliding window. Unlike the classical pedestrian detection, in the USARSim environment the subjects stand in many various shapes and positions. Besides, the approach angle of the robot to the victim creates variations. If victim does not fit in the window, then it is likely to miss.

4. Datasets

We created a dataset to test our detector. In the USARSim environment, the subjects have a wide range of variations in pose, appearance, illumination and background. Moreover, the approach angle of the robot to the subject creates lots of variations. During the creation of the dataset, these situations had been considered. Our current database contains 1460 negative and 943 positive images.

Fig. 1. Some sample images from our database. The subjects have a wide range of variations in pose, appearance, illumination and background.

5. Experimental results

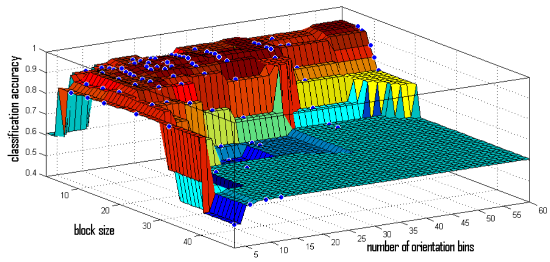
HoG features are calculated for the sub-blocks of the image. The sub-block size is an effective parameter. Moreover the number of orientation bin is another effective parameter of HoG features. We have tried several configurations for both of parameters.

Our experiments consist of 100 (block size, number of orientation bin) pairs. We have 100 different HoG representation. For each image on our dataset, these 100 representations are calculated. In other words, we have formed 100 datasets. Each dataset is divided into equal sized training and test sets. For each representation a binary classification task is employed.

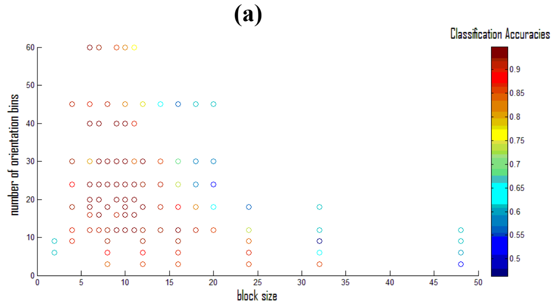
In literature, HoG features were generally classified with Support Vector Machines (SVMs) [7]. We have tried several classifiers including decision trees, multilayer perceptions etc. on our datasets. We found that the SVMs have better performance than all the other classifiers. Our results are parallel with the literature. So, we decided the usage of SVMs as our classifier. The hyper-parameters of our SVMs are as follows: kernel: 2nd degree polynomial, C=1.

In Figure 2 (a) and (b), the classification accuracies are given for 100 different HoG representations of our test sets. In Figure 2(a), the classification accuracy surface is shown. Each blue point represents a HoG representation (block size-number of orientation bin pair). The X coordinates show the sub-block size, Y coordinates show the number of orientation bin, and Z coordinates show the test classification accuracy when the X block-size and Y number of orientation bin is used.

In Figure 2 (b), each circle shows a different HoG representation. The circle colors show the classification accuracy. The best classification accuracy (95%) is obtained with 10*10 sub-blocks and 24 orientation bins. The experimented ranges can be seen at the figures.



(a)



(b)

Fig. 2. Classification accuracies of 100 different HoG features

When we look at the Figure 2(a) and (b), the best block-sizes are in the rage of [6-15]. Very small and very big sub-blocks produce very bad results as expected. For the number of orientation bins, only very small number of orientation bins produces bad results. After the 10bins, very good results can be obtained with different block-sizes.

Figure 3 shows same samples which are correctly classified.

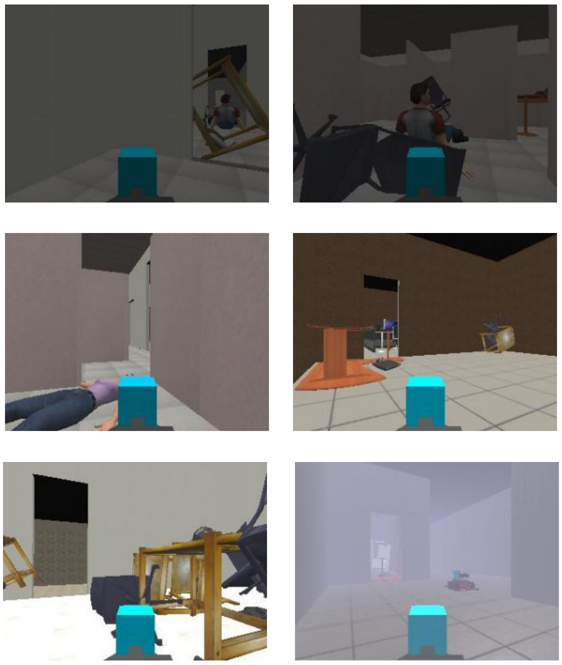


Fig. 3. Some examples that are correctly classified.

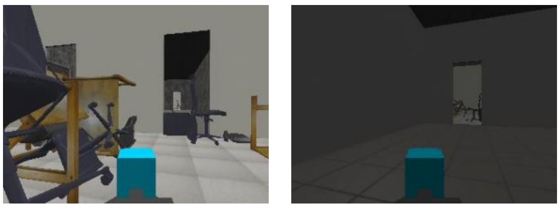


Fig. 4. False positive examples.

Figure 4 shows two false positive examples. False positives usually occur in the messy office environments. Edge orientation of the chairs, monitors etc. very close to the edge orientation of the subjects and this creates confusion.

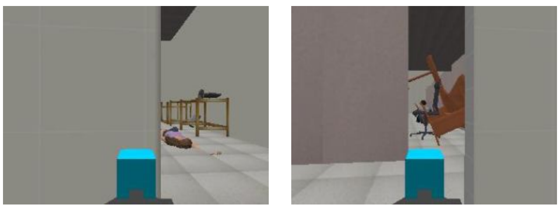


Fig. 5. False negative examples.

Figure 5 shows two false negative examples. False negatives generally occur if the subject is too far or not completely appears in image.

6. Conclusions

Urban Search and Rescue is a growing area of robotic research. The RoboCup Federation organizes Virtual Robots competition since 2006. In order to successfully compete in this competition, teams need to field multi-robot solutions that cooperatively explore and map an environment while searching for victims. When the number of the robots grows, the manual navigation is not a usable strategy. Therefore, the teams need develop autonomous navigation methods. But, autonomous navigation requires autonomous victim detection while searching victims. In literature, several human detection solutions were presented. Histogram of Oriented Gradients (HoG) serves a robust solution for varying background and lightening conditions. But, there is no extensive validation of HoG features for the different human poses. In a disaster area, the victims can be found with very different poses. In this work, we investigated the efficiency of HoG features for the different victim poses and obtained 95% success on automatic victim detection.

7. References

[1] Amigoni, F.; Basilico, N.; Quattrini Li, A. (2012). “How Much Worth Is Coordination of Mobile Robots for Exploration in Search and Rescue?”, *Proceedings of the RoboCup International Symposium*, Mexico City, Mexico, June 24 2012.

[2] Gerkey B, Mataric M (2004), “A formal analysis and taxonomy of task allocation in multi-robot systems”, *International Journal of Robotics Research*, 23:939–954

[3] D. Hahnel , R. Triebel , W. Burgard and S. Thrun "Map building with mobile robots in dynamic environments", *Proc. IEEE Int. Conf. Robotics Automation*, pp.1557 -1563 2003

[4] http://wiki.robocup.org/wiki/Rescue_Simulation_League

[5]www.darpa.mil/Our_Work/TTO/Programs/DARPA_Robotics_Challenge.aspx

[6] Carpin, S., Lewis, M., Wang, J., Balakirsky, S., Scrapper, C.: “USARSim: A robot simulator for research and education”, *In: Proc. ICRA*. (2007) 1400-1405

[7] Navneet Dalal and Bill Triggs, “Histograms of Oriented Gradients for Human Detection”, *International Conference on Computer Vision & Pattern Recognition* (CVPR '05) 1 (2005) 886—893.

[8] J. Komala Lakshmi, and Dr.M.Punithavalli, "A Survey on Performance Evaluation of Object Detection Techniques in Digital Image Processing", *IJCSI Inter-national Journal of Computer Science Issues*, Vol. 7, Issue 6, November 2010 ISSN (Online): 1694-0814.

[9] Bahadır Karasulu, "Review And Evaluation Of Well-Known Methods For Moving Object Detection And Tracking In Videos", *Journal Of Aeronautics And Space Technologies,* July 2010, Volume 4, Number 4, (11-22)

[10] Peter M. Roth and Martin Winter, "Survey Of Appearance-Based Methods For Object Recognition", Technical Report, ICG–TR–01/08

[11] D. M. Gavrila, “The visual analysis of human movement: A survey”, *CVIU*, 73(1):82–98, 1999.

[12] C. Papageorgiou and T. Poggio, “A trainable system for object detection”, *Internation Journal of Computer Vision (*IJCV), 38(1):15–33, 2000.

[13] D. M. Gavrila and V. Philomin, “Real-time object detection for smart vehicles”, *In Proceedings of the Conference on Computer Vision and Pattern Recognition*, Fort Collins, Colorado, USA, pages 87–93, 1999.

[14] P.Felzenszwalband, D.Huttenlocher, “Pictorial structures for object recognition”, *International Journal of Computer Vision* (IJCV), 61(1):55–79, 2005.

[15] S. Ioffe and D. Forsyth, “Probabilistic methods for finding people”, *International Journal of Computer Vision* (IJCV), 43(1):45–68, 2001.

[16] K. Mikolajczyk, C. Schmid, and A. Zisserman, “Human detection based on a probabilistic assembly of robust part detectors”, *European Conference on Computer Vision* (ECCV), 2004.

[17] Qiang Zhu, ShaiAvidan, Mei-Chen Yeh, Kwang-Ting Cheng, “Fast Human Detection Using a Cascade of Histograms of Oriented Gradients”, *Computer Vision and Pattern Recognition,* 2006 IEEE Computer Society Conference

[18] Navneet Dalal, Bill Triggs, Cordelia Schmid, “Human Detection Using Oriented Histograms of Flow and Appearance”, *ECCV* (2) 2006: 428-441