Hard hat detection in video sequences based on face features, motion and color information

Shan Du

Computer Research and ...

Hard Hat Detection in Video Sequences based on Face Features, Motion and Color Information

Shan Du, Mohamed Shehata and Wael Badawy
IntelliView Technologies Inc.
808-55 Avenue NE, Calgary, Alberta, Canada T2E 6Y4
{du, shehata, badawy}@intelliview.ca

***Abstract*—Human's safety in construction areas is vital. A hard hat is required to enter a construction area. Stopping a person who is not wearing a hard hat entering a construction area is very important. Video-based surveillance to detect hard hat is a new solution to this safety problem. This paper brings different video processing techniques together to construct a framework for fast and robust hard hat detection in construction areas. The proposed system can detect face and hard hat in real time.**

Keywords- Face detection; Haar-like features; hard hat detection; motion detection; color-based.

I. INTRODUCTION

A hard hat is required to enter a construction area to protect human's life. Stopping a person who is not wearing a hard hat entering a construction area is very important. Video-based surveillance to detect hard hat is a new solution to this safety problem. This is a new topic in computer vision and pattern recognition. In this paper, we describe a combined machine learning and image processing approach for hard hat detection in video sequences. There are three main parts of this framework. The first is the person's face detection based on Haar-like face features [1][2]. The second is the motion detection and skin color detection used to reduce the false alarms of faces. The third is the hard hat detection using the color information above the face regions.

The remaining of this paper is organized as follows: Section II presents the details of the proposed method. Section III describes our experiment setup and results. Section IV concludes the paper.

II. THE PROPOSED FRAMEWORK

The proposed framework is shown in Figure 1. It consists of three main components. The first one is the face detection based on Haar-like features. In this stage, we detect all possible face regions. The second one is the motion and color filtering. Before face detection, we first detect motions in video sequences. If there is no motion or little motion in the video, we do nothing further. If there is motion, we obtain the motion regions and apply face detection only on these regions to avoid scanning the whole image. This can save the computation time and remove false alarms on the background. After face detection, there may still have some false alarms of faces. We use face skin color to filter out the non-face blocks. The third one is hard hat color detection.

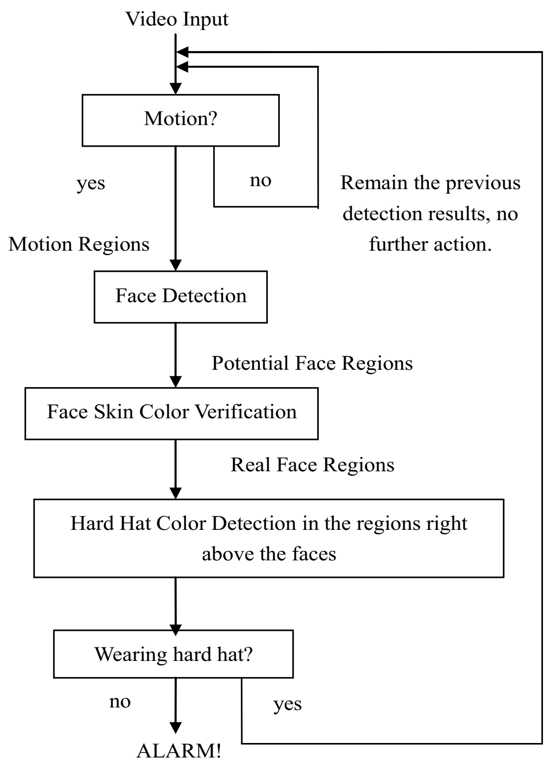


Figure 1. The proposed framework.

We test the regions right above faces to detect if there is a hard hat or not.

A. *Human Face Detection*

Face detection is a two-class (face/non-face) classification problem. As the face manifold is highly complex, due to the variations in facial appearance, lighting, expressions, and other factors, face classifiers that achieve good performance are very complex.

The learning-based approach constitutes the most effective one for constructing face/non-face classifiers. Viola *et al.* introduced a rapid object detection scheme based on a boosted cascade of simple features [1]. Lienhart *et al.* introduced a novel set of 45° rotated Haar-like features which significantly enriched the basic set of simple Haar-like features and can be calculated very efficiently[2]. The real-time speed and good performance of them can be explained

978-1-61284-840-2/11/$26.00 ©2011 IEEE

by the following: First, AdaBoost learning algorithms are used for learning of highly complex classifiers. AdaBoost methods learn a sequence of easily learnable weak classifiers, and boost them into a single strong classifier via a linear combination of them. Second, the real-time speed is achieved by the use of integral images for rapid computation of Haar-like features. Moreover, the use of cascade structures further speeds up the computations. The face detection technique used in this paper is based on Haar-like features introduced by Viola *et al.* and then improved by Lienhart *et al*.

We use fourteen Haar-like feature prototypes shown in Figure 2 that include four edge features, eight line features and two center-surround features. The sum of the pixels which lie within the white rectangles is subtracted from the sum of pixels in the black rectangles. These prototypes are scaled independently in vertical and horizontal direction in order to generate a rich, over-complete set of features.

Given a feature set and a training set of positive and negative samples, any machine learning algorithms can be used to learn a classification function. In this paper, AdaBoost method [3] is used to both select a small set of features and train the classifier. In its original form, the AdaBoost learning algorithm is used to boost the classification performance of a simple learning algorithm (e.g., it might be used to boost the performance of a simple perception). It does the final classification by combining a collection of weak classification functions to form a stronger classifier.

There are over 117,000 rectangle features associated with each 24×24 sub-window, a number far larger than the number of pixels. The main challenge is to find a very small number of features that can be combined to form a strong classifier. In support of this goal, the weak learning algorithm is designed to select the single rectangle feature which best separates the positive and negative examples. For each feature, the weak learner determines the optimal threshold classification function, such that the minimum number of examples are misclassified.

The conventional AdaBoost procedure can be easily interpreted as a greedy feature selection process. Consider

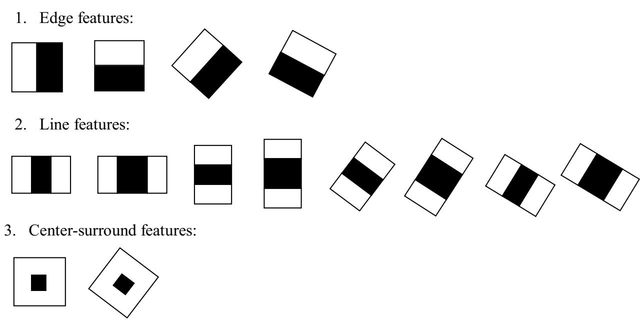


Figure 2. 14 Haar-like features.

the general problem of boosting, in which a large set of classification functions are combined using a weighted majority vote. The challenge is to associate a large weight with each good classification function and a smaller weight with poor functions. AdaBoost is an aggressive mechanism for selecting a small set of good classification functions which nevertheless have significant variety. Drawing an analog between weak classifiers and features, AdaBoost is an effective procedure for searching out a small number of good "features" which nevertheless have significant variety.

A cascade of classifiers is used to achieve increased detection performance while radically reducing the computation time. A cascade of classifiers is a degenerated decision tree where at each stage a classifier is trained to detect almost all objects of interest while rejecting a certain fraction of the non-object patterns. Instead of finding faces, the algorithm discards non-faces. It is faster to discard a non-face than to find a face.

The key insight is that smaller and therefore more efficient boosted classifiers reject many of negative sub-windows while detecting almost all positive instances. Simpler classifiers are used to reject the majority of sub-windows before more complex classifiers are called upon to achieve low false positive rates.

The cascade structure is shown in Figure 3.

B. Motion and Color Filtering

Before face detection, we first detect motions in video sequences. If there is no motion or little motion in the video, we do nothing further; just remain the previous detection results. If there is motion, we obtain the motion regions and apply face detection only on these regions to avoid scanning the whole image. This can save the computation time and filter out some false alarms on the background. After face detection, there may also have some false alarms of faces (Figure 6). We use face skin color to filter out the non-face blocks. Unlike other face detection methods that use skin color information as a pre-processing step, we use skin color to do post-processing.

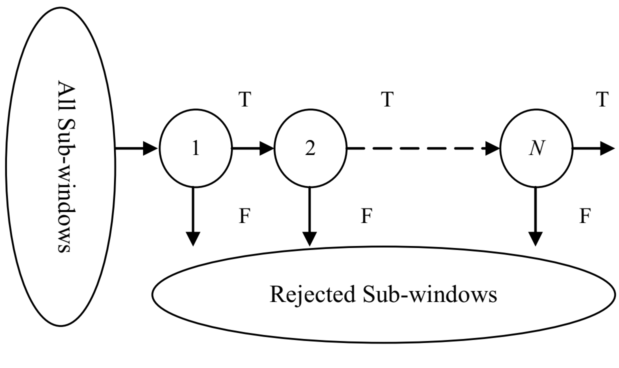


Figure 3. Cascade of classifiers.

C. Hard Hat Detection

The hard hat detection is based on the color information in the regions right above each real face region. For both the face color detection and the hard hat color detection, we use the YCbCr [4][5] and HSV [6] color spaces. To examine these regions, we first segment them using some morphological operations as shown in Figure 7(b). This is because these regions not only consist of hair or hard hat (Figure 7(c)), but also may consist of forehead and background (Figure 7(c)). Then we need to skip the forehead part to test if there is a hard hat covering the hair.

III. Experiment Setup and Results

In this section, we test our scheme using 5 benchmark video sequences that have human faces inside, namely Akiyo, Carphone, Claire, Foreman, and Grandma. Among them, Foreman is the only sequence that has hard hat.

The face detection results are shown in Figure 4 and Figure 5. In Figure 6, we also show some results of face detection false alarms. It is obvious that the false alarms look somewhat like a face. Since the face detection algorithm use only the gray-scale information, it is hard to differentiate these false alarms from faces. But using color information, it is easy to filter them out. Table 1 shows the improved performance of adding motion and color information to the original AdaBoost face detection method [1][2]. We can increase the detection rate as well as decrease the false alarms.

To detect hard hats, we examine the regions right above the faces. If the color inside them is similar to a hard hat, we believe there is a hard hat, otherwise, no hard hat.

The above-face regions are shown in Figure 7(a). To examine these regions, we first segment them using some morphological operations as shown in Figure 7(b). This is because these regions not only consist of hair or hard hat (Figure 7(c)), but also may consist of forehead and background (Figure 7(c)). Then we need to skip the forehead part to test if there is a hard hat covering the hair.

Finally, for every 10 frames, we fuse the 10 results using majority voting to make the final decision. In this case, we can eliminate the occasional false positives or false negatives. Figure 8 shows the output information of the proposed system (the title above each image).

Figure 4. Face detection results. The sub-image at the left-upper corner shows the region right above the face (in this sequence, it is the hard hat region).

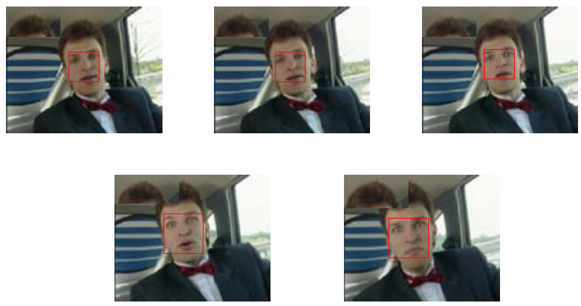


Figure 5. Face detection results. The sub-image at the left-upper corner shows the region right above the face (in this sequence, it is the hair).

IV. Conclusions

In this paper, we proposed a hard hat detection approach in video sequences using Haar-like face features, motion and color Information. This is a new solution to secure the construction area safety. This is a new topic in computer vision and pattern recognition. The proposed method has three main parts. The first is the person's face detection based on Haar-like face features. The second is the motion detection and skin color detection used to reduce the false alarms of faces. The third is the hard hat detection using the color information above the face regions. The proposed system can detect face and hard hat in real time.

Acknowledgment

The authors would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) and Alberta Ingenuity Fund (AIF) for their support.

References

[1] P. Viola and M. J. Jones, "Robust real-time face detection," International Journal of Computer Vision, vol. 57, no. 2, pp. 137-154, 2004.

[2] R. Lienhart and J. Maydt, "An extended set of Haar-like features for rapid object detection," Proc. of IEEE International Conference of Image Processing, vol. 1, pp. 900-903, 2002.

[3] Y. Freund and R. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," Journal of Computer and System Sciences, 1997.

[4] R. L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face detection in color images," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 24, no. 5, pp. 696-706, 2002.

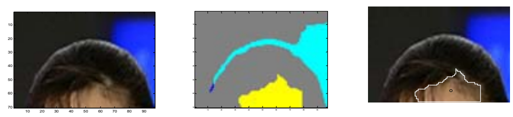
[5] S. Gundimada, L. Tao, and V. Asari, "Face detection technique based on intensity and skin color distribution," Proc. of International Conference on Image Processing, pp. 1413-1416, 2004.

[6] K. Sabottka and I. Pitas, "Segmentation and tracking of faces in color images," Proc. of International Conference on Automatic Face and Gesture Recognition, pp. 236-241, 1996.

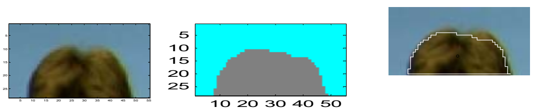
Figure 6. False alarms of face detection.

Table 1. Increased face detection rate and decreased false alarms by using motion and skin color information.

Sequence Name	Total Frame Number	Increased True Positives (motion only)	Decreased False Positives (motion only)	Increased True Positives (motion and skin color)	Decreased False Positives (motion and skin color)
Akiyo	300 frames	181	4	181	4
Carphone	382 frames	47	1	47	2
Claire	494 frames	43	1	43	1
Foreman	298 frames	0	10	0	30
Grandma	870 frames	246	1	246	1



(a) (b) (c)



(a) (b) (c)

(a) (b) (c)

Figure 7. Segmentation of the above-face regions.

Figure 8. The output information of the proposed system (the title above each image).