Computers in Industry 100 (2018) 267–277

Contents lists available at ScienceDirect

Computers in Industry

journal homepage: www.elsevier.com/locate/compind

An intelligent vision-based approach for helmet identification for work safety

Hao Wu[a], Jinsong Zhao[a,b,*]

[a] State Key Laboratory of Chemical Engineering, Department of Chemical Engineering, Tsinghua University, Beijing, China
[b] Beijing Key Laboratory of Industrial Big Data System and Application, Tsinghua University, Beijing, 100084, China

ARTICLE INFO

Keywords:
Accident prevention
Image recognition
Helmet identification
Work safety

ABSTRACT

Helmets are essential equipments to protect workers from danger during inspection and operation. Considering that some workers would not always obey the regulation, video surveillance systems covering the whole factory and supervisors are needed to monitor whether workers are wearing helmets or not. However, with a large number of surveillance screens, it is difficult to identify any helmet violation behavior during any time, which can lead to severe accidents. With the rapid development of image recognition technologies, computer vision-based inspections have been one of the most important industrial application areas. In this paper, an intelligent vision-based approach for helmet identification is proposed. This approach focuses on monitoring whether workers are wearing helmets or not, at the same time, identifying the colors of helmets. A color-based hybrid descriptor composed of local binary patterns (LBP), hu moment invariants (HMI) and color histograms (CH) is proposed to extract features of helmets with different colors (red, yellow and blue). Then a hierarchical support vector machine (H-SVM) is constructed to classify all features into four classes (red-helmet, yellow-helmet, blue-helmet and non-helmet). This approach is tested on our data set and the average accuracy of helmet identification is 90.3%.

1. Introduction

Struck by falling object is a common hazard across all industry sectors and has resulted in many cases of workplace injury and fatality. For example, there were more than 8000 accidents in Poland in 2012 involving injuries resulting from an impact of falling objects and nearly half of that number happened at industrial worksites [27]. In Netherland, falling object accidents constitute 12% of the reported 12,500 accidents which have occurred between 1998 and 2004 [37]. The percentage of fatal accidents owing to struck by falling objects in the United States is 5–6% according to the statistics of the Occupational Safety and Health Administration (OSHA) [37]. In order to protect workers or visitors from being struck by falling objects, helmet is an effective safety precaution [19,36]. Therefore, it is regulated to wear helmets in certain construction and manufacturing occupations, oil fields, refineries, and chemical plants [3,5,13,20].

However, due to the discomfort, some workers would not always obey the regulations [63]. It is reported that 47.3% victims did not have personal protective equipment or wear it correctly when danger arose [1]. In order to ensure that workers obey the regulations and wear helmets on work sites, video surveillance system covering factories and supervisors are needed. However, with a large number of surveillance screens, it is a tough job to identify any helmet violation behavior during any time. Therefore, there is an urgent demand to develop an intelligent vision-based system to carry out this surveillance task.

Image recognition technologies have made significant progress in recent years. They have been applied in face recognition [23,25,56,61], vehicle detection [22,53,65], pedestrian detection [18,42,55]. With the development of this field, automated visual inspections have been used widely in industrial applications [2,8,9,49,52]. For helmet identification, numerous related studies have been conducted. Wen et al. [12] firstly proposed a method of modified hough transform for helmet identification in ATM's surveillance systems. Then Chiverton [21] studied an approach based on histogram of oriented gradients (HOG) descriptor and support vector machines (SVM). Its accuracy of 85% shows the validity of applying image recognition to helmet identification. Park et al. [31] developed an approach based on Hue-Saturation-Value (HSV) color histograms for construction worker detection. Waranusast et al. [48] used K-Nearest Neighbors (KNN) to classify the shape and color information extracted from images for motorcycle helmet detection and the accuracy is 74%. Recently, Dahiya et al. [28] conducted a comparison among three widely used feature descriptors, HOG, Scale-

* Corresponding author at: Room 350, Department of Chemical Engineering, Tsinghua University, Beijing, 100084, China.
E-mail addresses: h-wu16@mails.tsinghua.edu.cn (H. Wu), jinsongzhao@tsinghua.edu.cn (J. Zhao).

https://doi.org/10.1016/j.compind.2018.03.037
Received 2 August 2017; Received in revised form 9 December 2017; Accepted 27 March 2018
0166-3615/

Invariant Feature Transform (SIFT) and local binary patterns (LBP) using SVM classifier, for bike-riders helmet detection. Results show that HOG descriptor combined with linear SVM possesses the best performance for this topic.

While the above mentioned researches utilize only one descriptor for feature extraction, Silva et al. [46] suggested a hybrid descriptor, the combination of Circle Hough Transform (CHT), HOG and LBP, for motorcycle helmet detection. The hybrid descriptor trained by Random Forest algorithm returns 94.23% accuracy rate. The results also indicate that the combination of edge (HOG), texture (LBP) and geometric (CHT) lead to better performance than single descriptor, because it contains more information of images. Silva et al. [47] also studied the effect of combining Wavelet Transform (WT), HOG and LBP. While combined with the multi-layer perceptron (MLP), the reported accuracy is 91.37%.

The previous approaches with combining descriptors aim to extract features from grayscale images that originate from color images. However, grayscale images only contain one-dimensional data while color images have three dimensions. This means that the approaches based on grayscale images will lose color information, which for human eyes, is the most significant information. Besides, actually there are more than one helmet color and it is necessary to identify the color of each helmet. In different factories, different helmet colors have different meanings. For example, in chemical factories, operators wear red helmets and safety supervisors wear yellow, while blue helmets are usually seen on other work sites. In order to identify the helmet colors, extracting color features from images is necessary. Rubaiyat et al. [1] combined color-based and CHT to solve this problem. However, the average accuracy of identification reached only 80.48% and it is difficult for practical application. Therefore, there is a need to develop a new approach, which can identify helmet colors with high accuracy.

This paper proposes a new technology framework with a color-based hybrid descriptor for helmet identification. Based on collected videos, motion detection and pedestrian detection are applied to extract head image segments. Then the color-based hybrid descriptor, combining LBP (texture), hu moment invariants (HMI) (geometry) and color histograms (CH) (color) together are used to transform image segments into feature vectors. Then a hierarchical support vector machine (H-SVM) is utilized to classify the head image segments into four classes (red-helmet, yellow-helmet, blue-helmet and non-helmet). The proposed approach can not only monitor whether workers are wearing helmets or not, but also able to identify the helmet colors with higher accuracy.

The rest of this paper is organized as follows: Section 2 describes the framework of the intelligent helmet identification. Section 3 shows the algorithm of feature extraction and classification. Section 4 presents experimental results for pedestrian detection and helmet identification, and finally Section 5 summarizes this paper.

2. Intelligent vision-based approach framework

This section presents the framework of the intelligent vision-based approach for helmet identification. Fig. 1 shows the work flowsheet of the approach, which mainly contains three steps: motion detection, pedestrian detection and helmet identification.

2.1. Motion detection

Videos are collected and saved from the surveillance camera as dataset, which is a sequence of consecutive image frames (1920 pixels × 1080 pixels). Volunteers are asked to walk around within the scope of the cameras, either with helmets on or off. In order to eliminate most of static image frames, the first step is to capture the segments of moving objects. In this field, early used methods are based on Gaussian Mixture Model (GMM) [15,35,41]. Besides, in order to improve the adaptation speed of GMM, KadewTraKuPong and Bowden [39] proposed the Adaptive Mixture of Gaussians. However, a study performed by Zivkovic and Van Der Heijden [66] showed that KNN has better performance on simple static scenes. Since our videos are collected from an outdoor static scene, KNN is selected as the basic method here. The motion detection process is illustrated in Fig. 2a–d. First, mean filtering operation (see Fig. 2a) is applied on the original images for noise elimination. After that, background subtraction based on KNN is utilized to detect the moving objects (see Fig. 2b). From this figure, we find some gray regions, of which the corresponding gray values are small. Therefore, binarization is implemented on Fig. 2b and the moving region is marked with white color in Fig. 2c. At the final step, in order to find the whole region of moving objects, the closing operation of morphology is used to complement blank regions among white marked pixels and to eliminate spots (see Fig. 2d). Thus, the motion regions can be marked correctly from videos and the image segments of moving objects are sliced with size of 320 pixels × 160 pixels.

2.2. Pedestrian detection

To capture the images of workers, human images should be distinguished from non-human images, that is, pedestrian detection. HOG is applied for this field. It was firstly proposed by Dalal and Triggs in 2005 [34] and reportedly to give near-perfect results and to outperform existing feature sets for pedestrian detection [38]. First, HOG descriptor transforms images of moving objects into feature vectors. Linear SVM classifier is then used to train the transformed feature vectors. Through the combination of the HOG descriptor and the SVM classifier, the images of moving objects are classified into two classes, human and non-human classes. The image segments of workers are stored for helmet identification in the next step.

2.3. Helmet identification

Since workers have been detected by pedestrian detection in Section 2.2, the head regions of human image segments should be extracted first for the color identification. In former motion detection in Section 2.1, humans body regions are marked with white color. While in this step, a new yellow rectangle is built and the top point of the white-marked region is identified and set as the middle point on the top line of the rectangle. Here the head region is sliced by the yellow rectangle with size of 50 pixels × 50 pixels (see Fig. 3).

Then a color-based hybrid descriptor is proposed for feature extraction of the head image segments. It combines LBP, HMI and CH, to extract feature sets from the head image segments. In order to identify the helmets, the head image segments are classified into two classes (helmet and non-helmet) by SVM classifier. Since helmet-color identification is a multi-class problem, a H-SVM is constructed with the combination of three linear SVM classifiers. The head image segments are classified into four classes (red-helmet, yellow-helmet, blue-helmet, non-helmet). The detailed helmet identification algorithm is described in the next section.

3. Algorithm of pedestrian detection and helmet identification

In the intelligent vision-based approach, there are two steps based on image recognition technologies, pedestrian detection and helmet identification. The former is to detect workers and the latter is to identify helmets. Generally, each process of image recognition contains two parts to identify specific objects: feature extraction and feature classification.

In feature extraction, an image frame is transformed into a feature vector using descriptors. Due to the rapid development in this field, many descriptors are available now to detect the objects. Three main categories of descriptors are shown as follows:

(a) Shape-based: HMI [32], Fourier descriptor (FD) [17], Wavelet