Neurocomputing 228 (2017) 106–118

Contents lists available at ScienceDirect

Neurocomputing

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

Body part boosting model for carried baggage detection and classification

Wahyono, Joko Hariyono, Kang-Hyun Jo*

School of Electrical Engineering, University of Ulsan, 93 Daehak-ro, Nam-gu, Ulsan 680-749, Republic of Korea

ARTICLE INFO

Keywords:
Carrying baggage detection and classification
Part model
Body part-baggage spatial model
Video surveillance
Scalable histogram of oriented gradient
Parallelogram Haar-like feature
Support vector machine
Mixture model
Boosting machine

ABSTRACT

In the automatic video surveillance system, the detection of a human carrying baggage is a potentially important objective for security and monitoring purposes in the public spaces. This paper introduces a new approach for detecting and classifying baggage carried by a human on the images. It utilizes the spatial information of the baggage in reference to the body of the human carrying it. A human-baggage detector is modeled by the body parts of a human, including the head, torso, leg, and baggage parts. The feature descriptors are extracted for each part based on its characteristics and these features are further trained using a support vector machine (SVM) classifier. A mixture model is built specifically for the baggage part due to a significant variation in shape, size, color, and texture. The boosting strategy constructs a strong classifier by combining a set of weak classifiers which are obtained by training the body part. The proposed method has been extensively evaluated using the public datasets. The experimental results confirm that the proposed method is viable for a state-of-the-art in the carried baggage detection and classification system.

1. Introduction

In the last decade, automatic video surveillance systems (AVSS) have gained significant attention from the computer vision research community. The detection of carried baggage is one of important tasks in AVSS, which is used for security and monitoring purposes in public areas. Conceptually, the carried baggage detection can be achieved by observing the changes in human appearances caused by carried baggage, allowing humans with carried baggage to be distinguished from those without. In practice, the human-baggage detector is viable for this task, however, the task is inherently difficult due to the wide range of the baggage types and methods of carrying.

Several approaches have been proposed in literature for the detection of baggage that has been abandoned [1,2], and that which is actively being carried [3–6]. Tian et al. [1] proposed a method to detect abandoned and removed objects using foreground subtraction and analysis. In their approach, the background is modeled via three Gaussian mixtures and texture information in order to handle changing light conditions. The static regions obtained through background subtraction are then analyzed using a region growing method. These regions are further classified as either abandoned or removed objects based on established heuristic rules. In some cases, however, this method produced a high number of false alarms due to imperfect background subtraction. To overcome this problem, Fan et al. [2] proposed a relative attributes schema to prioritize alerts by ranking candidate regions. However, in the real implementation, identifying the owner of the abandoned baggage is very important. Therefore, as a prior stage, the system should be capable to detect the person carrying baggage. The authors from [4–6] proposed the same concept to detect the human carrying baggage by utilizing a sequence of moving humans to create a spatial temporal template. This was aligned against view-specific exemplars generated offline in order to obtain the best matched. Carried baggage was finally detected from the temporal protrusion. Tzanidou et al. extended the framework adding a color information [7] to increase accuracy rates. The system can also classify the baggage type based on its position in reference to the human body [4]. However, this method assumes that the parts of the carried objects protrude beyond the body silhouette and therefore the method cannot detect non-protruding carried objects. Authors from [6] conducted further analysis for estimating whether a protruding pixel belongs to a carried object by collecting data samples consisting of some features of the protruding pixels. The data was then analyzed using a fuzzy classification method and the fuzzy membership degrees were combined into a binary Markov Random field. The protrusion problem can also be solved via the method from [3] by utilizing a ratio color histogram. Using the assumption that the colors of carried objects differ from the clothes, a good accuracy was achieved. However, this method is dependent on an event where the bag is being transferred or abandoned. The assumption of observing the person before and after the changes in carrying status is application specific and cannot be used as a general carried baggage detector. Another suggested method, introduced by reference [8], utilized an image-based crowd analysis

* Corresponding author.
E-mail addresses: wahyono@islab.ulsan.ac.kr (Wahyono), joko@islab.ulsan.ac.kr (J. Hariyono), acejo@ulsan.ac.kr (K.-H. Jo).

http://dx.doi.org/10.1016/j.neucom.2016.10.038
Received 21 December 2015; Received in revised form 6 October 2016; Accepted 15 October 2016
Available online 02 November 2016

using a Finite Time Lyapunov Exponents (FTLE). This method increased the performance of the state-of-the-art in detecting people carrying baggage. However, it does not classify the baggage into a certain type. Baggage classification can be useful for analyzing the owner of unattended baggage, for example, staff officers found unattended baggage in a certain locations. Fortunately, this type of baggage was known to be either handbag, luggage, or backpack; therefore, when the staff members attempted to find the baggage owner on the recorded CCTV video, the searching domain could be reduced to humans carrying baggage of a specific type based on abandoned baggage.

This paper addresses a novel approach for detecting and classifying carried baggage. It utilizes the strong spatial information of baggage relative to the body parts. Instead of constructing a model for the entire human region, it builds a model for each body part [9] including the bag part according to any possible placements, referred to as a body part-baggage spatial model. Overall, this paper offers the following major contributions:

- Introducing a human-baggage model which is based on the spatial information of the body parts and the relationship between parts for detecting and classifying the carried baggage.
- Training for features of the baggage part using a mixture model for solving strong baggage variation problems (location, size, shape, color, and texture).
- Combining the Scalable Histogram of Oriented Gradient (SHOG) and parallelogram Haar-like features (PHF) for handling scale and skew problems within the part regions.
- Observing a boosting machine for interpolating a final score of each part detection in order to verify the existence of the baggage in the human region.
- Utilizing an integral image and an approximate radian basis function (RBF) kernel for reducing the complexity of the computation process.
- Collecting a new dataset of human-baggage regions.

An initial version of our model has been introduced in [10]. Compared to our previous work, there are several major extensions that merit highlighting: 1) Combining the Scalable Histogram of Oriented Gradient and parallelogram Haar-like features for handling the skewed and scale problem within regions. 2) Utilizing an approximate RBF kernel on the SVM classification to reduce computational complexities. 3) Describing more details regarding the SVM classifier and the boosting strategy used in our implementation. 4) Conducting comprehensive experiments to evaluate the processing time. 5) Adding a new baggage dataset collected from Google images. 6) Comparing the proposed model to the state-of-the-art in carried baggage detection method. 7) Conducting extensive experiments for more body partitions.

This paper is divided into a number of sections. Section 2 presents our part model. Section 3 explains the feature descriptors used in our framework. Section 4 describes a training strategy using the boosting machine on multiple classifiers. Experimental settings and results will be explained in Section 5. Lastly, Section 6 concludes our works.

2. The proposed model

2.1. An overview

This section presents the details of our model for detecting human-baggage objects in an image. The main diagram of the proposed method is depicted in Fig. 1. The first step of the method is to the extract region candidates of a human carrying baggage from a given input image via the sliding window with a multiscale region search. To simplify, we use a full body of human detector which is trained by the support vector machine on a scalable histogram of oriented gradient features. It is observed that both human-with-baggage and human-without-baggage regions might be detected as candidates. Thus, the body part-baggage spatial model is proposed to detect whether or not the human is carrying baggage. Our model is based on the spatial information of a bag on the human body. A body proportional schema is used to estimate spatial parameters. Specifically, the human body is modeled into several body parts including the bag, which is placed in any possible position. Since the baggage may vary in shape, size, color, and texture, a mixture model is used. The final decision involves reviewing the various arrangements of parts based on their spatial relationship. It is performed separately by evaluating the score of each part model.

2.2. Human body parameter

Using a human body proportional model [11], as shown Fig. 2(a), it is deducted that in average the height and width of a person are $H = 8\,h$ and $W = 2\,h$, respectively, where h is the height of the head, such that $h = H/8$. A bend line, B, is defined as the center of the body on the vertical axis, and a vertical line, C, is denoted as the center of the body on the horizontal axis that traverses the centroid of the body. Let T be the position of the top of the head on the image, and L be the left most location of the body, then $B = T + 4\,h$ and $C = L + h$. These parameters are used to create our spatial model of human-baggage that is described in the next subsection.

2.3. Body part-baggage spatial model

The general idea of the spatial model is adopted from [4], by placing the bag in a certain location according to the body proportion and the position of the camera. Our spatial model is constructed based on three major categories of bag, 1) backpack or hand bag, 2) tote or duffle bag, and 3) rolling luggage. As shown in Fig. 2, spatial models of bags define the set of conditions for checking whether the bag exists in a front view direction. After observing our training dataset, it is found that most backpacks are located near or on top of the bend line with an average height equal to $3\,h$. Most handbags and rolling bags are located at the bottom of the bend line, in which the average height of handbag and rolling bag are $2h$ and $4h$, respectively.

The body part-baggage spatial model verifies whether a human is carrying baggage, after the human region candidates have been detected using a sliding window, as described in Section 2.1. For instance, if our model identifies the location depicted in Fig. 2(b) as a bag with the highest probability value, then the bag is classified as a backpack; if not, then the bag is placed in a different location from one of the other categories. If no baggage is identified in any of the spatial models, it is concluded that the human region does not contain any baggage. In the other words, the region is assigned as a human without a baggage region.

This model has been designed by dividing regions into several parts, taking into account the possible placements of baggage on the human body. If the coordinates of the full body region relative to the input image are known, then the coordinates of each part could be determined based on the proposed spatial model. Hereafter, each sub-region is verified by a part model which is described in the following subsection.

2.4. Body part model

A part model is a schema that seek to detect a small number of features and their relative positions to then determine whether the object of interest is present in the image. Many object detection and recognition problems have been successfully implemented using part-based models, such as face [13], human [14], and general object detection [9], with incredible results. Recently, [15] introduced the utilization of a part model for deformable objects tracking in a vision