Carried Baggage Detection and Classification Using Part-Based Model

Wahyono and Kang-Hyun Jo(✉)

Intelligent Systems Laboratory, Graduate School of Electrical Engineering,
University of Ulsan, Ulsan 680-749, Korea
wahyono@islan.ulsan.ac.kr, acejo@ulsan.ac.kr

Abstract. This paper introduces a new approach for detecting carried baggage by constructing human-baggage detector. It utilizes the spatial information of baggage in relevance to the human body carrying it. Human-baggage detector is modeled by body part of human, such as head, torso, leg and bag. The SVM is then used for training each part model. The boosting approach is constructing a strong classification by combining a set of weak classifier for each body part. Specify for bag part, the mixture model is built for overcoming strong variation of shape, color, and size. The proposed method has been extensively tested using public dataset. The experimental results suggest that the proposed method can be alternative method for state-of-the art baggage detection and classification algorithm. AQ1

Keywords: Carried baggage detection and classification · Part-based model · Video surveillance · HOG · Boosting SVM · Mixture model

1 Introduction

In the last decade, automatic video surveillance (AVS) system has more attention from the computer vision research community. Detecting of carried object is one of important parts of AVS. This task is potentially important objective for security and monitoring in public space. However, the task is inherently difficult due to the wide range of baggage that can be carried by a person and the different ways in which they can be carried. In the literature, there have been several approaches proposed for detecting baggage that abandoned by the owners [1, 2] or still being carried [3–5] by them. Tian et al. [1] proposed a method to detect abandoned and removed object using background subtraction and foreground analysis. In their approach, the background is modeled by three Gaussian mixture that combining with texture information in order to handle lighting change conditions. The static region obtained by background subtraction is then analyze using region growing and is classified as abandoned or removed object by some rules. However, in some cases this method produces many false alarm due to imperfect background subtraction. To overcome this problem, Fan et al. [2] proposed relative attributes schema to prioritize alerts by ranking candidate region. However, in real implementation to know who is the owner of abandoned baggage is very important. Therefore, as prior process, the system should capable to detect the person who carried baggage. The authors from [3, 4] proposed same concept to detect

D.-S. Huang and K. Han (Eds.): ICIC 2015, Part III, LNAI 9227, pp. 1–8, 2015.
DOI: 10.1007/978-3-319-22053-6_31

carried object by people. They utilized the sequence of human moving to make spatial temporal template. It was then aligned against view-specific exemplar generated offline to obtain the best match. Carried object was detected from the temporal protrusion. The author in [4] extend the framework such that the system can also classify the baggage type based on the position in relevance to the human body carrying it. However, the method assumes that parts of the carried objects are protruding from the body silhouettes. Due to its dependence on protrusion, the method cannot detect non protruding carried object. The protruding problem can be solved by method from [5]. This method utilized ratio color histogram. Using assumption of the color of carried object is different with clothes, it will achieve good result in accuracy. However, this method is dependence on event where the bag being transferred or left. The assumption of observing the person before and after the change in carrying status is application specific and cannot be used as a general carried object detector.

This paper proposes a novel approach for detecting people carrying baggage. Our approach utilizes the strong connection between baggage and body parts. Instead of constructing model for entire of object, our method build model for each part [8] of body including the bag part according to possible placement. Overall, this paper offers the following major contributions; (1) Part-based model schema and the relationship among them specific for detecting person carrying baggage and classifying the baggage based on our spatial model. (2) Bag part mixture model for solving strong variation problem of baggage (e.g., location, size, shape, and color).

2 System Model

This section presents the detail of our model for detecting human-baggage object on the image. Our model based on spatial information of bag on the human body.

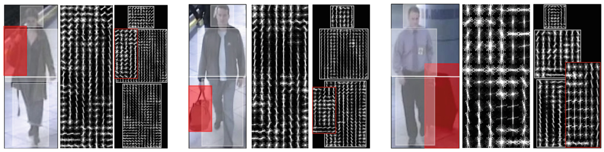
2.1 Human Body Parameter and Baggage Spatial Model

Using human body proportional model, as shown Fig. 1(a), it can be deducted that in average the height and weight of a person are $H = 8h$, $W = 2h$, respectively, where h is the length of head. Therefore $h = H/8$. Bend line B is defined as the center of the body in vertical axis, vertical line C is denoted as center of body in horizontal axis that traverse the centroid of body. Let define T be the position of the top of the head in the image, and L be the most left location of body in the image, then $B = T + 4h$ and $C = L + h$. These all parameters are used for making our spatial model of human-baggage that is described in detail in the next subsection.

The general idea of spatial model is adopted from [4], by placing the bag in certain location according of the body proportion and the viewing direction of the person. Our spatial model is divided in into three major categories of bag, (1) backpack or hand bag, (2) tote bag or duffle bag and (3) rolling luggage. As shown in Fig. 1(b)–(d), spatial models of bag define the set of conditions for checking whether the bag exists or not in front view direction. For instance, if our part model detect that location depicted in Fig. 1(b) as bag with high probability value, then the bag is classified as a backpack; if

(a) (b) (c) (d)

Fig. 1. Human-baggage spatial model of front view direction. The category of bag is divided into three major categories according to body proportions (a); (b) backpack or hand bag, (c) tote or duffle bag, and (d) rolling luggage.



(a) Backpack (b) Duffle bag (c) Rolling Luggage

Fig. 2. Visualizations of HOG features from some models on i-Lids data [11]. For each category, first image is input image, the second image is the initial root filter for a human-bag model, and the last image is part filter model from the first image.

not, then it is placed the bag in location of other categories. In addition, if there is no bag identified in all spatial models, it is concluded the human is not carrying any bag.

2.2 Part-Based Model

Many object detection and recognition problems have been successfully implemented using part based model, such as face [6], human [7], and general object detection [8] with incredible result. In our work, the human-baggage is modeled based on observations of the part models and their relative position among them. In this paper, each human-baggage model is divided into four part models; head part, torso part, leg part and bag part, as shown in Fig. 2. Height of head, torso, and leg part are h, $3h$ and $4h$, respectively, while the height of bag part is varying according to our model. The feature vector is then extracted from full body and part model. The framework described here is independent of the specific choice of feature. In our implementation, the histogram of oriented gradient (HOG) [9] feature is used for description the model.

The part detection is represented by $m = \Phi_i(x, y, l, p)$, where specifying an anchor position (x, y) relative to full body in the l^{th} level of the pyramid scale image and part p. The score of part interpolation based on hybrid of boosting Support Vector Machine (SVM) [10] is defined by the formulation as follows:

$$J(m) = \sum_{i=1}^{n} \omega_i \Phi_i(x, y, l, p) \quad (1)$$

where ω_i is the weighting, $\omega_i > 0$ and Φ_i is the output probability of SVM classification of the component p, n is the number of component.

2.3 Mixture Model

Detecting baggage is not easy due to variation of color, size and appearance. Consequently, the bag in certain location may exhibit more intra-class variation for by a single bag part model. Thus, specific for bag part, mixture model is used for handling this problem. More formally, let define a distribution f is a mixture of K component distribution $f_1, f_2, \ldots, f_K$ if

$$f(x) \sum_{k=1}^{K} \lambda_k f_k(x, \theta_k) \quad (2)$$

with the λ_k being the mixing weights, $\lambda_k > 0$, $\sum_k \lambda_k = 1$, x is feature vector of observations, and θ_k specify as parameter vector of the k^{th} component. The overall parameter vector of the mixture model is thus $\theta = (\lambda_1, \ldots, \lambda_K, \theta_1, \ldots, \theta_K)$. Our goal now is to estimate these all parameters using Expectation-Maximization (EM).

2.4 Detection

Let M be the number of possible model (including human object that do not carry any baggage) learned in our framework. The final score of object hypothesis being human-baggage object is the maximum value among the score of each model that formulated as Eq. (3). In addition, if the value of J_{final} is less than a fixed threshold, the object hypothesis is classified as other objects.

$$J_{final} = \max_{m_1, \ldots, m_M} (J(m_1), J(m_2), \ldots, J(m_M)) \quad (3)$$

2.5 Training

A hybrid technique of boosting SVM implemented by [10] is used for training our human-baggage detection system. The boosting technique introduces its ability to extract high discriminative features to construct strong classifier from set of weak classifier. In our model, set of weak classifiers is built according to a set of part models. The standard SVM technique is used to learning of partial part model. The details of the standard SVM can be found in [12]. The SVM is applied to learn part model as a weak

(a) Category 1 (b) Category 2 (c) Category 3

Fig. 3. Some samples used for training. Category 1 includes backpack and handbag, category 2 consists of tote bag and duffle bag, and category 3 contains rolling luggage. First, second and third row for each category are representing viewing direction from front, side and back viewing direction. Thus, in total, nine models of human-baggage are built.

classifier. The boosting technique is then used for interpolating the full body detection. In practical data, the weighting of weak classifiers are automatically decided by our algorithm. The result training indicates that the head part has the largest weighting. Also, the head part is used to distinguish the view direction of object. The output probability of SVM classification is computed by

$$\Phi(x) = P(y = 1|x) = \frac{1}{1 + \exp(-h(x))} \tag{4}$$

where $h(x)$ is the signed distance of feature vector x to the margin of the SVM model.

3 Experiments

Our model was tested on human-baggage dataset. It was collected from small subset of INRIA [9], Caltech [3] and our own images. Since our work was focused on human-baggage detection, only human carrying baggage images were selected. The summary of our dataset is shown in Table 1. Our dataset is divided into two groups, training and testing groups. Each group is classified into 3 categories manually for creating ground truth. The training group contains 338 human-baggage object consisted of 132, 111, and 95 data for category 1, 2, and 3, respectively. The testing group contains 202 human-baggage object distributed as 78, 67, and 57 data for category 1, 2, and 3, respectively. Data from category 1 and 2 is resized to be 128 × 64 pixels resolution, while for category 3 is more wider becomes 128 × 72 pixels resolution. Figure 3 shows several samples of our dataset used in our implementation.

Fig. 4. Some typical detection results. The human-baggage object is detected as red bounding box. The baggage is then classified into category 1, 2 and 3 which are represented by blue, yellow and green bounding boxes, respectively (Color figure online).

Our model was evaluated for classified object into human with or without baggage. The human carrying baggage is set to be positive samples. Human without baggage is set to be negative samples. For first evaluation, 338 positives samples and 1,352 negative sample were used. Our method achieves detection rate of 77.02 %. Since, as our knowledge, there is no method researching this specific task, we do not compare our method with others yet. However, we have tried to just use original HOG and SVM on full body [9], but the result is not promising around 45.62 %. Next, our method was evaluated to classify image into three baggage categories. Table 2 summaries the evaluation result on training dataset. Our method obtains classification rate as much as 76.51 %, 77.47 %, 80 % for each category, respectively. In average, it achieves true classification rate of 77.21 %. Table 3 depicts the evaluation result on testing dataset.

Table 1. Dataset specification

Baggage category	#Training	#Testing
1. Backpack/handbag	131	78
2. Tote bag/duffle bag	111	67
3. Rolling luggage	95	57
Total	338	202

Table 2. Evaluation on training data

		Detection		
		C1	C2	C3
Ground truth	C1	101	14	7
	C2	9	86	16
	C3	5	14	76

Table 3. Evaluation on testing data

		Detection		
		C1	C2	C3
Ground truth	C1	46	23	9
	C2	12	39	16
	C3	7	15	35

Our method achieves 59.40 %, which for category 1, 2 and 3 are 58.97 %, 58.2 %, 61.40 %, respectively. Last, our method was evaluated on full image databases. Sliding window mechanism in any position and scale are used to detect human-baggage region. In practical, the same human-baggage region is usually detected with several times with overlapped bounding boxes. Therefore, it is necessary to combine the overlapped regions for unifying detection and rejecting misdetections. Some typical results of our method are shown in Fig. 4.

4 Conclusion

In this paper, part-based model for detecting and classifying baggage carried by human was introduced. First, the human region was modeled into four parts, head part, body part, leg part and baggage part. The model utilized the location information of baggage relative to human body. Histogram of oriented gradient (HOG) features were extracted on each part. The features were then trained using boosting support vector machine (SVM). Gaussian mixture model was also applied for modeling the baggage part for handling variation of baggage size, shape and color. After conducting extensive experiment, our method achieves 77.02 % and 59.40 % for detection and classification rate, respectively. However, our method has some limitations for detecting and classifying baggage carried by human. First, it may fail to detect multiple baggage carried by the same person. The additional model should be considered in our future work for handling this problem. Second, our method fail to detect overlapping human-baggage region. Increasing the number of part body can be one of the solutions for solving this issue.

References

1. Tian, Y.L., Feris, R., Liu, H., Humpapur, A., Sun, M.-T.: Robust detection of abandoned and removed objects in complex surveillance video. IEEE Trans. SMC Part C **41**(5), 565–576 (2011)
2. Fan, Q., Gabbur, P., Pankanti, S.: Relative attributes for large-scale abandoned object detection. In: International Conference on Computer Vision (2013)
3. Damen, D., Hogg, D.: Detecting carried objects from sequences of walking pedestrians. IEEE Trans. PAMI **34**(6), 1056–1067 (2012)
4. Tzanidou, G., Edirishinghe, E.A.: Automatic baggage detection and classification. In: IEEE 11th ICISDA (2011)
5. Chuang, C.-H., Hsieh, J.-W., Chen, S.-Y., Fan, K.-C.: Carried object detection using ratio histogram and its applications to suspicious event analysis. IEEE Trans. Circ. Syst. Video Technol. **19**(6), 911–916 (2009)
6. Tzimiropoulos, G., Pantic, M.: Gauss-Newton deformable part models for face alignment in-the-wild. In: ICIC 2014, Taiyuan, China, 3 Aug 2014
7. Hoang, V.-D., Hernandez, D.C., Jo, K.-H.: Partially obscured human detection based on component detectors using multiple feature descriptors. In: Huang, D.-S., Bevilacqua, V., Premaratne, P. (eds.) ICIC 2014. LNCS, vol. 8588, pp. 338–344. Springer, Heidelberg (2014)

8. Felzenszwalb, P.F., Girshick, R.B., McAllester, D., Ramanan, D.: Object detection with discriminatively trained part based models. IEEE Trans. PAMI **32**(9), 1627–1645 (2010)
9. Dalal, N., Triggs, B.: Histogram of oriented gradients for human detection. In: Conference on Computer Vision and Pattern Recognition, pp. 886–893 (2005)
10. Hoang, V.-D., Le, M.-H., Jo, K.-H.: Hybrid cascade boosting machine using variant scale blocks based HOG features for pedestrian detection. Neurocomputing **135**, 357–366 (2014)
11. Home Office Scientific Development Branch: Imagery library for intelligent detection systems (i-LIDS). In: The Institution of Engineering and Technology Conference on Crime and Security, pp. 445–448 (2006)
12. Chang, C.-C., Lin, C.-J.: LIBSVM: a library for support vector machine. ACM Trans. Intell. Syst. Technol. **2**, 1–27 (2011)
13. Dollar, P., Wojek, C., Schiele, B., Perona, P.: Pedestrian detection: a benchmark. In: International Conference on Computer Vision and Pattern Recognition (2009) AQ2