CHILD AND ADULT CLASSIFICATION USING BIOMETRIC FEATURES BASED ON VIDEO ANALYTICS

Omer Faruk Ince[1,2], Ibrahim Furkan Ince[1,2], Jang-Sik Park[1,2]
Jong-Kwan Song[1,2] and Byung-Woo Yoon[1,2]

[1]Department of Electrical Engineering
[2]Department of Electronic Engineering
Kyungsung University
309, Suyeong-ro, Nam-gu, Busan 48434, Korea
{ farukince21; furkanince; jsipark; jsong; bwyoon }@ks.ac.kr

ABSTRACT. *As the number of social insecurity in regard to social crimes is on its rise, it requires a CCTV camera a higher accuracy in detecting the objects including pedestrians for efficient work of catching criminals. As the importance of the function of pedestrian detection is socially agreed upon, more studies on image and video based pedestrian detection have been conducted. In terms of that, the goal of this study is classification of pedestrian in two categories as a child and an adult. In this study, Haar cascade classifiers are used. This method first detects a full body and a head. Then, it measures the biometry given the relative proportioning length of a full body and a head. Moving average algorithm is used to obtain threshold ratio. Experimental results show the accuracy 100% for children and 64.5% for adults.*
Keywords: Adaboost, Haar-like feature, Local binary patterns (LBP), Probabilistic Data Association Filter (PDAF), Child and adult classification

1. **Introduction.** Object detection in visual platforms has been studied for a long time. It can be used for pedestrian detection, vehicle detection, gender classification, age classification, etc. The main purpose of object detection is pinpointing the location of target in the scene. For accurate real-time object detection, features need to be robust, differential, and easy to calculate [1]. Classification of child and adult is useful for the social security in that it makes it much easier to catch criminals based on the detected physical traits. For classification, several studies of classification between child and adult have been introduced. Those studies are based on cameras using near distance which is between 20 and 50 cm [2], such as working on only facial features or certain part of body. Since the video surveillance systems mostly use CCTV cameras, previous studies are not adequate enough to prevent criminal cases. For long distance classification which is between 2 meters and 10 meters away from the CCTV camera, biometric information is very applicable. So, a new method for age classification with biometric information is proposed. We used the ratio between lenght of head and full-body. The statistical information which is moving average is used to decide the threshold for classification in different spots in the scene.

The organization of this paper is as follows. In Section 2, methods which are used to detect and track the object, are briefly mentioned. In Section 3, the proposed algorithm is explained. In Section 4, child and adult classification is given with the description of datasets and experiments. The last section depicts the considerations for this approach.

2. Reletad Approaches.

2.1. Methods that are used for pedestrian detection and tracking.
In this study, several different methods are used to detect and track the pedestrian in the scene. Adaboost using Haar-like features and local binary pattern (LBP) features are applied to detecting full body and head of person, respectively. Then, Probabilistic Data Association Filter (PDAF) had a role to track the person in the scene.

Rapid object detection is introduced by Viola and Jones [2], and is applied to the pedestrian detection. Their detector shows highly speedy detection with the simple Haar-like features that can be calculated in an efficient way by using integral images and Adaboost classifiers in a cascade structure [3].

Boosting is a statistical method that employs the original distribution of positive and negative examples to compute simple rules (weak classifiers) and combines them into creating a stronger classifier [4].

Local binary patterns (LBP) algorithm is a kind of descriptor that is used for various classifications in computer vision. Basically, it focuses on reference point and the neighbors, and then LBP is combined with the histogram of oriented gradients (HOG) descriptor.

Probabilistic Data Association Filter (PDAF) is a statistical approach to the problem of data association uncertainty. This uncertainty occurs when the sensor devices read (correct or false) measurements whose origin is not the intended target; or the signal from the target is very weak and to detect the target, the detection threshold should be lowered, which can cause background signals and sensor noise to be detected, yielding false readings. It is mostly useful to track ballistic missiles and some other objects.

2.2. Overview to previous studies on age classification.
Up to now, there are not so many researches published on the prediction of people's age from digital image or video. In the first instance, Kwon and Lobo [5] started predication of people's age from digital facial images and they described a method to classify people into three groups as babies, young adults, and senior adults. There is also another method which is proposed by Horng et al. [6]. This method put an extra group into account as middle aged adults to Kwon & Lobo's approach. The biggest restriction in this method is that they cannot detect children and younger adults who are under 40 years old. Lanitis et al. [7] started up active appearance models. To estimate the age, combined shape and texture parameters are dismissed by using classifiers. Reddy et al. [8] suggested a model using grey level co-occurence matrix (GLCM) based on diagonal local binary patterns (LBP) .

Similar to the proposed study, another method which uses 3D depth cameras to enable children's safety applications came out [9]. The classification method is based on human skeletal features which are collected using a 3D depth camera.

3. Proposed Video Based Child and Adult Classification with Biometry.
Various types of techniques are approached to classify adults and children. Most of them were conducted based on near distance classification, such as using facial features. In this study, the proposed aim is different from the previous studies. The proposed algorithm works on long distance video scenes using CCTV cameras. Working on long distance is considered harder than any other methods to implement, but it is definitely more useful for daily applications. The reason is that CCTV cameras are everywhere and it does not cost a lot to be used as video surveillance system.

Starting from this point of view, first of all, requirements for pedestrian detection from long distance scene should be determined. After that, classification requirements for child and adult can be discussed. Block diagram of the proposed algorithm can be shown as Figure 1.
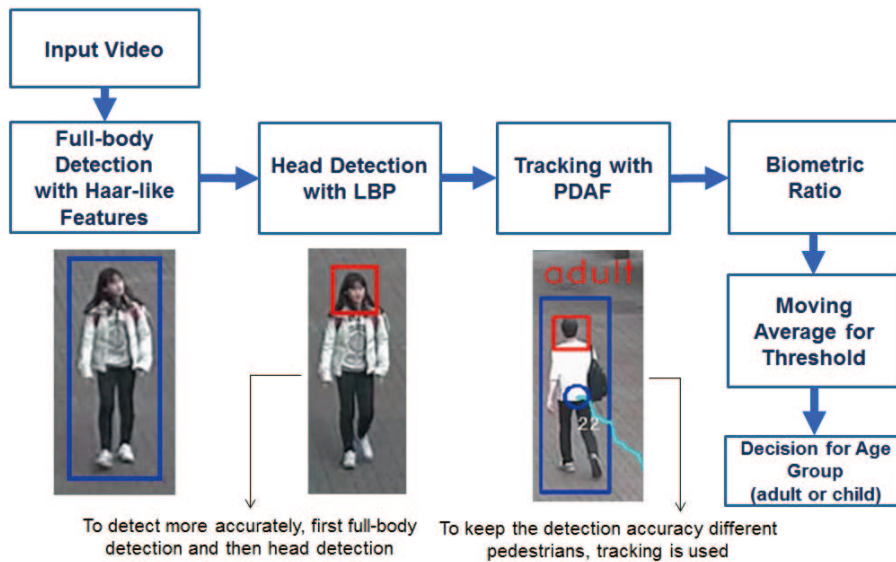
FIGURE 1. Block diagram of the proposed method

It is a wellknown fact that, head reaches its full size until a person becomes a teenager; stated in other words, head grows faster than other parts of the body. So it could be useful feature for classification of pedestrian in two categories as adult and child. If biometric ratio is applied between head and full body, it is significant to assign if pedestrian is adult or child.

In a still image, there are various objects captured from tested videos. Adaboost algorithm based Haar cascaded classifiers are used to detect the pedestrian's full body in the scene. The first step is detecting the pedestrian body. Then, new region of interest would be pedestrian body for head detection. So, head will be detected in this particular body. LBP algorithm is used to detect head. One of the advantages of applying head detection is to improve the accuracy of pedestrian detection, because sometimes computer can detect some other objects as pedestrian. For the next step, biometric calculation for head and body ratio is required; it is shown in Equation (1).

$$r = l_H/l_B \tag{1}$$

where l_H and l_B represent the length of head and body, respectively. Also, r means the ratio of proportion.
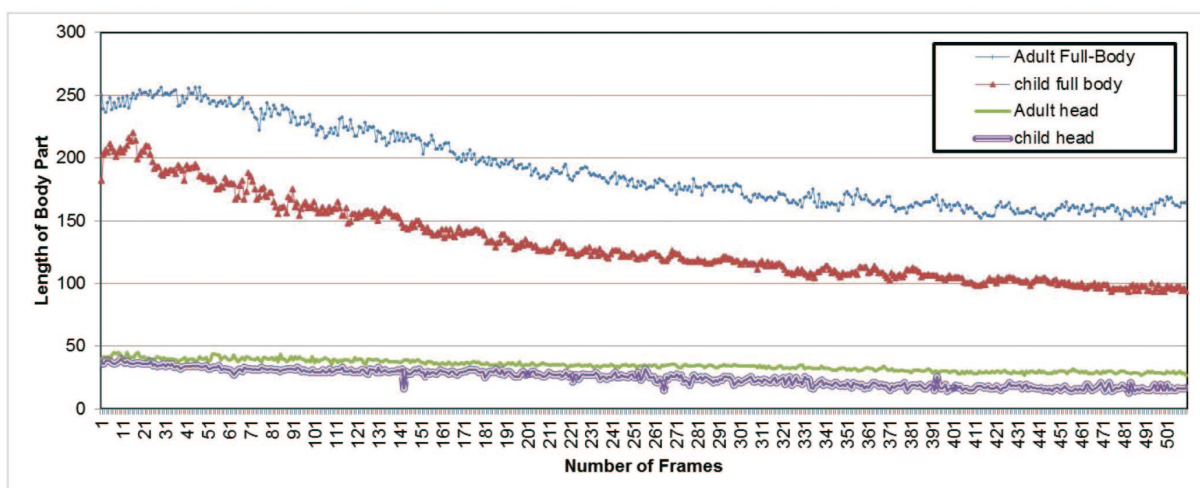

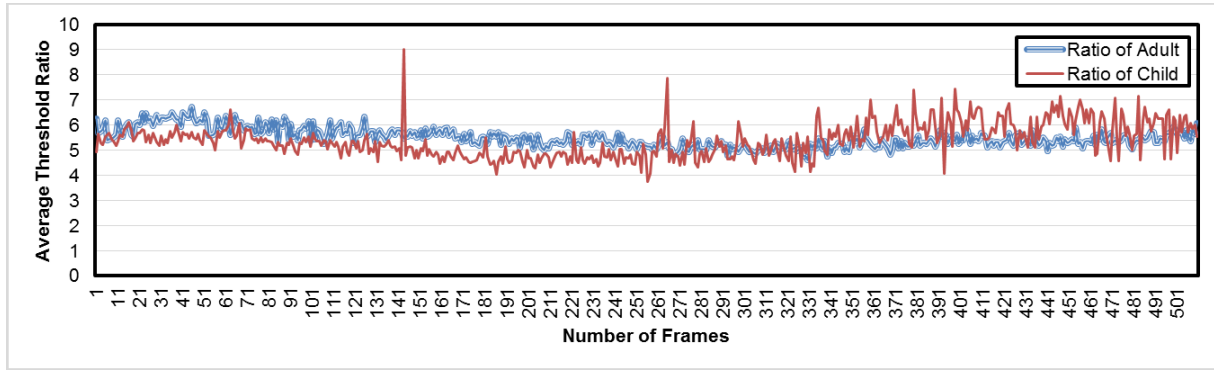
FIGURE 2. Length of body parts over frames

FIGURE 3. Average threshold ratio over number of frames

In some cases there might be more than one pedestrian in the scene, and it is difficult to identify those pedestrians in the next frame. Therefore, Probabilistic Data Association Filter is used to track pedestrians. By means of that, it is easy to recognize identity of pedestrian in the next frame.

Instead of assigning the ultimate threshold value, moving average algorithm is applied to obtaining different threshold values, because the biometric ratio for adults and children changes according to the position of pedestrian in the scene. Extracting the graph of average ratio over number of frames, is decisive for assigning the ultimate threshold values for certain spots in the scene.

Moving average algorithm can be assumed as low-pass filter. To calculate the threshold values in different spots in the scene, equation below is used.

$$h_{avg} = a * h_{avg} + (1 - a) * h \tag{2}$$

Here, 'h' represents the length of the detected area. For the head and the body, the same equation is used. Also, 'a' is the degree of the weighting decrease, a constant smoothing factor, and it should be between 0 and 1. In this study 'a' is 0.9.

Significance of threshold value is required for comparison with obtained biometric ratios. Consequently, certain amount of adults and children are manually examined to calculate the threshold. It is clear that a child and an adult can be classified using comparison between threshold and obtained biometric ratio.

4. **Experimental Environment.** In this chapter, experimental results of this study will be focused on. Videos are made by Kyungsung University Electronic Engineering Department, and taken in Kyungsung University campus, Busan, Korea. For implementation OpenCV v3.0.0 via C++ on Microsoft Visual Studio 2013 are used. Number of images, which are used for both day and night time full-body and head detection is shown in Table 1.

Among the 59 different test subjects, in both day and night time conditions, the accuracy rates are 64.5% for adults, and 100% for children.

The reason of working on different light conditions, is to see the real performance rate of this study. It is clear to see that day time detection rates are better than night time detection. The reason of lower adult classification accuracy is because of some grown up's

TABLE 1. Number of images used for training

Day Time	Full-Body	Head	Night Time	Full-Body	Head
Positive	3041	3022	Positive	2046	2006
Negative	8040	8040	Negative	6082	8040
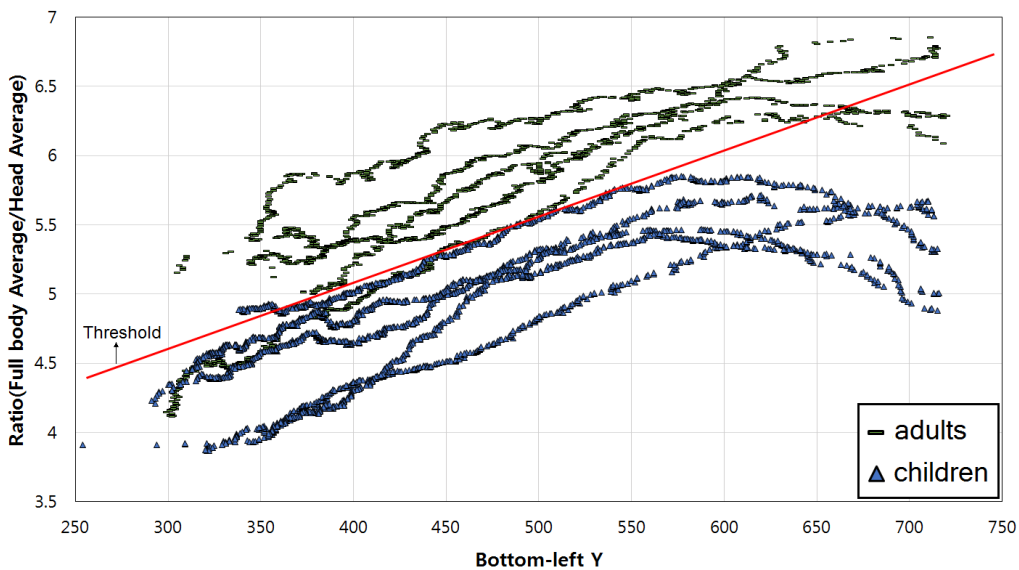
(a)
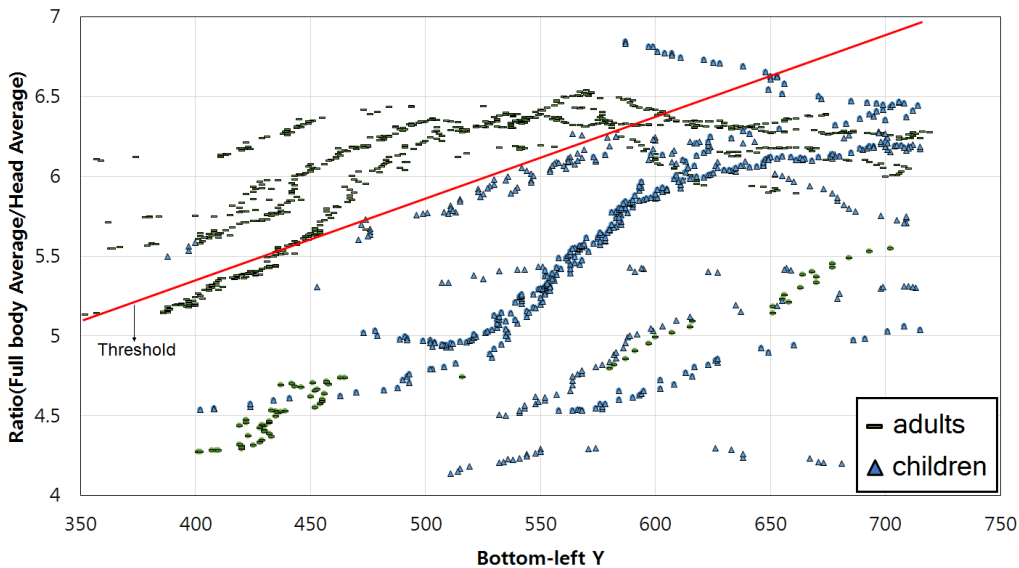
(b)

FIGURE 4. Change of threshold ratios according to the position: (a) day time, (b) night time

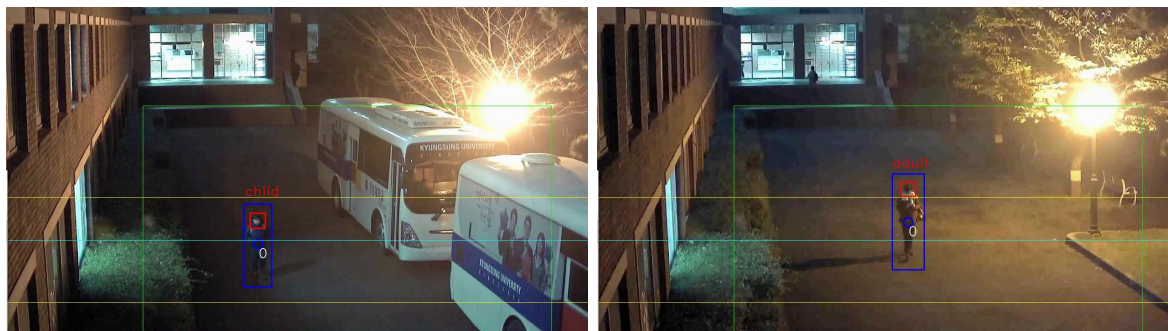FIGURE 5. Child adult classification in day time

FIGURE 6. Child adult classification in night time

physical appearance's being out of average. It is easy to see from the threshold chart, that some adult's threshold value is under the reference threshold line.

Since most of studies are based on facial features, there are not many approaches to compare the results with the proposed method. However, the proposed study can be compared with the study in Section 2.2 [9]. These two approaches work on full body features, but while their method can classify children with 97% accuracy using 3D depth camera, the proposed method can classify children with 100% accuracy using CCTV camera. Also, since the 3D depth cameras have high cost to be used for video surveillance systems, it is a not useful method for daily life conditions.

5. **Considerations.** In this study, a new approach is proposed. Biometric features are applied on many purposes. Therefore, biometry is used to classify children and adults from long distance. Since the number of social crimes is growing, this approach has been invented as one of the ways to deter crimes. As the object goes away from camera, it is difficult to keep correct threshold for subject. So, different threshold values are applied according to the position of the person in the scene. To do that, average learning method is performed. Also using tracking techniques helped to increase identification in the next frame. In light of this information, classification rate is 64.5% for adults, and 100% for children. However, there are still undesirable results for the adult classification accuracy. In Figure 4, it is clear to see that some adult's biometric ratios are under the threshold value while some children's biometric ratios are above the threshold value. One of the main reasons is that some people's body features are out of the average ratio of age group they are member of. For the future, higher accuracy for adult detection is required.

REFERENCES

[1] K. Lee, C. Lee, S.-A. Kim and Y.-H. Kim, Fast object detection based on color histograms and local binary patterns, *TENCON IEEE Region 10 Conference*, Cebu, Philippines, pp.19-22, 2012.

[2] P. Viola and M. Jones, Rapid object detection using a boosted cascade of simple features, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001.

[3] G. R. Rakate, S. R. Borhade, P. S. Jadhav and M. S. Shah, Advanced pedestrian detection system using combination of Haar-like features, Adaboost algorithm and edgelet-shapelet, *IEEE Conference on Computational Intelligence and Computing Research*, Coimbatore, India, pp.18-20, 2012.

[4] P. Viola and M. J. Jones, Robust real-time face detection, *International Journal of Computer Vision*, vol.57, pp.137-154, 2004.

[5] Y. H. Kwon and N. da V. Lobo, Age classification from facial images, *Computer Vision and Image Understanding*, vol.74, pp.1-21, 1999.

[6] W. B. Horng, C. P. Lee and C. W. Chen, Classification of age groups based on facial features, *Tamkang Journal of Science and Engineering*, vol.4, no.3, pp.183-192, 2001.

[7] A. Lanitis, C. Draganova and C. Christodoulou, Comparing different classifiers for automatic age estimation, *IEEE Trans. Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol.34, no.1, pp.621-628, 2004.

[8] A. M. Reddy, K. SubbaReddy and V. V. Krishna, Classification of child and adulthood using GLCM based on diagonal LBP, *International Conference on Applied and Theoretical Computing and Communication Technology*, Davangere, pp.857-861, 2015.

[9] C. Basaran, H. J. Yoon, H. K. Ra, S. H. Son, T. Park and J. Ko, Classifying children with 3D depth cameras for enabling children's safety applications, *Proc. of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp.343-347, 2014.