

Uzun Mesafeli Hareketli Görüntülerde Yayaların Biyometrik Özelliklerine Göre Yaşlarının Sınıflandırılması

Ömer Faruk İNCE, İbrahim Furkan İNCE, Jang Sik PARK

Department of Electrical and Electronics Engineering, Kyungsoong University, Korea, Republic Of
farukince21@ks.ac.kr

(Geliş/Received:02.05.2016; Kabul/Accepted:17.04.2017)

DOI: 10.17671/gazibtd.309264

Özet— Video tabanlı insan tespiti, günümüzde bir hayli yaygın olan çalışmalardan biridir ve bu konu hakkında birçok çalışmalar ve tasarılar mevcuttur. Daha da ayrıntılı bir çalışma için; nesne tespit edildikten sonra, nesnelere sınıflandırılabilir veya takip edilebilir. Yetişkin ve çocuk sınıflandırması, sosyal güvenliğin sağlanması için, özellikle de günümüzde artan vakaları göz önünde bulundurduğumuzda, pek yararlı olabilir. Yapılan çalışmanın gayesi, videolardaki görüntülerden insanları yetişkin ve çocuk olarak iki sınıfa ayırmaktır. Öncelikle insan tespiti için Haar sınıflandırıcı kullanılmıştır. Bir sonraki adımda ise, kafa ve tüm vücut uzunluğu kullanılarak biyometrik bir oran çıkarımı yapılmıştır. Bu orana göre de tespit edilen insanın yetişkin veya çocuk olduğu belirlenmiştir. Sonuçlarımız göstermektedir ki, yetişkin sınıflandırmasındaki doğruluk payı %74.7 ve çocuk sınıflandırmasındaki doğruluk payı ise %68.1'dir.

Anahtar Kelimeler— Biyometrik Özellikler, Yaş Sınıflandırması, Yaya Algılama

Using Biometric Features on Long Distance Videos for Accurate Pedestrian Age Classification

Abstract—Pedestrian detection is popular topic currently, and there are various studies on image and video based pedestrian detection. After detection, it can be more specified as classification, tracking, etc. Classification of child and adult is useful for the social security, because the number of social crimes is increasing. The goal of this study is classification of pedestrian in two categories as child and adult. In this study, Haar cascade classifiers are used. At first, the full body and head area of the pedestrian are detected. Next, biometric ratio between height and head of pedestrian is calculated. Then, age is classified by comparing the biometric ratio between height of head and whole body. Experimental results showed that proposed biometric feature has a significant discriminant value for distinguishing the child from the adult on long distance videos. The accuracy rates are 74.7% for adult, and 68.1% for child.

Keywords— Biometric Features, Age Classification, Pedestrian Detection.

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. One of the most significant stages of object detection is extracting features. For accurate real-time object detection, features have to be robust, differential, and easy to calculate [1]. Classification of child and adult is useful for the social security, because the number of social crimes is increasing. For this classification, several studies have been introduced. However, those studies are based on near distance [2]. For far distance classification, biometric information is very applicable. The ratio between the heights of head and body is a good information for age

classification, because the size of head becomes optimum before puberty.

We developed a new method for age classification with biometric information. We used the length ratio between head and full-body. The statistical information is used to decide the threshold for classification.

In this paper, general overview and previous studies on age classification is referred in Section II. The proposed new approach for age classification is explained in detail, in Section III. Following that, experimental results with considerations are shown in Section IV. Lastly, we end the paper with conclusion in Section V.

2. AGE CLASSIFICATION

2.1. Haar-like Features

Paula Viola and Michael Jones [2] proposed an effective object detection method which uses Haar-like feature based cascade classifiers.

Basically, in this machine learning work, cascade function is trained with data as positive and negative images. Then it can be used for detecting objects in digital images.

First of all, algorithm needs various data, so a lot of positive and negative images should be collected to train classifier. After that, useful features are needed to be extracted from those images. In this point, Haar features [3] involve the process. Haar features can be considered as convolutional kernels, and to obtain every single feature, we need to subtract sum of pixels under white rectangle from sum of under black rectangle.

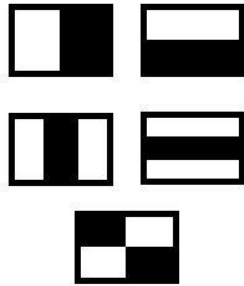


Figure 1. Haar-like features

To detect a good number of features, we have to use all possible sizes and locations of each kernel. That means, we will face a calculation issue which requires long time to calculate. Even though, we have 24×24 size window, it has more than 160000 number of features. For every single feature, sum of pixels are needed to be calculated. Also, for decreasing the time of calculation process, integral images were introduced. By using integral image, it doesn't matter how large the number of pixels are. It just simplifies the calculation, and calculates much faster.

In despite of everything, there is a big problem which is obtaining useful features. In other words, most of possible features are irrelevant. Here the answer is adaboost.

In general, boosting is a statistical method that uses the original distribution of positive and negative examples to compute simple rules also called weak classifiers and combines them to create a stronger classifier. Adaboost is most commonly used for binary classification, but with minor modifications it can also deal with multiple classes [4].

First of all, we apply all of those features, on training images. After that, algorithm finds useful features according to minimum error rate. Then, it can classify the object by using those features. Of course, there will be some errors or misclassifications. That is why we focus on selecting minimum error rates. Each image has equal probability at the beginning and after each steps are done, weight of misclassified images are increased. This process will be done iteratively until needed accuracy or error rate is obtained.

Final classifier is the last step of the algorithm. Sum of these weak classifiers are called final classifier. We call it weak, because it cannot classify images by itself. But when weak classifiers get together, they form strong classifier.

2.2. Overview to Previous Studies on Age Classification

Up to now, there are not so many studies 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. On a limited database, 100% accuracy is not applicable in real world applications. To complete similar age group classification, another technique which is based on craniofacial development theory was proposed [5]. But the approach is frustrated again due to limited data sets. With this way, it's very difficult to get robustness while we're dealing with real life conditions which has limitless circumstances.

There is also another method which is proposed by W.B. Horng, C.P. Lee and C.W. Chen [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. It can only detect babies. Despite everything, results were satisfactory and to find features, the analysis is pretty complex.

Lanitis [7] started up Active Appearance Models. To estimate the age, combined shape and texture parameters are dismissed by using classifiers. Simple quadratic fitting, shortest distance classifier, and Neural network are compared with age estimation accuracies of these classifiers. Geng [8] suggested static modeling on ageing patterns. It was based upon comparing input image with those patterns which consist of multiple images of various ages for each person.

3. PROPOSED VIDEO BASED CHILD AND ADULT CLASSIFICATION WITH BIOMETRY

Since studies on adult and child classification were introduced, different types of techniques are used. Some of them just preferred to use short distance detection

techniques, some of them used facial features to detect people according to their ages. In this study, our aim is different than the previous studies. We tried to classify people in long distance video scenes. For example, while one of previous studies was working on short distance classification (such as based on facial features) with webcams, this study focuses on detecting and classifying people on the street by using CCTV. Working on long distance can be considered as harder to implement, but it is definitely more useful for daily applications which are benefits of people.

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 below:

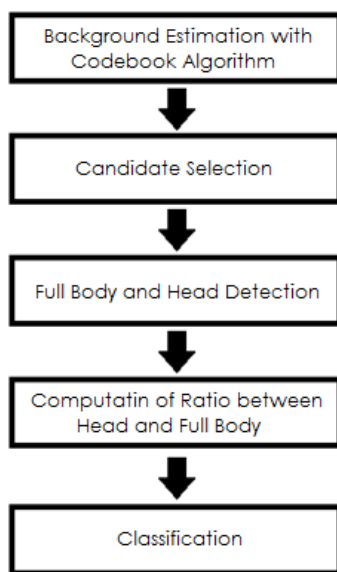


Figure 2. Block diagram of the proposed method

As it is seen from the block diagram, first, we need to separate foreground object from the scene by using codebook background subtraction. It is necessary to detect pedestrian faster and more accurate.

After first step, it is easy to detect pedestrian in the scene with using cascade classifiers based on adaboost algorithm. In the implementation part, more detail information, how cascade classifiers detect pedestrian, will be given.

Secondly, we use same way, which is using cascade classifiers to detect head of pedestrian. The only difference is that in the second step we were focusing to find full body in the whole scene. However in third step, we are trying to detect head of pedestrian inside full body.

After having full body information and head information for pedestrian, algorithm can be considered for classification. Here, the main idea of this study steps in.

There are medical facts that child and adult can be separated from each other by using biometry. In this study, the calculation of biometric ratio can be done by proportioning the length of full body to the length of head of the person, that is what we call computation of ratio. More details are given in the implementation part.

After we obtain threshold ratio by using biometry, we compare our ratio with threshold value, to see if pedestrian is child or adult.

3.1. Biometry for Classification

So far, we referred the proposition of previous algorithms for age classification and main framework of the proposed method. Now we are focusing on the last, but the most important step of the framework. That is classification of child and adult in digital videos.

In a digital video, there are various objects. Detecting human or human's head in the video, is applicable while ROI (Region of Interest) is human or humans' head. For implementation, we used Haar cascade classifier. After we detected the body, our new region of interest would be human body. After that, we needed to detect head of human. In this stage, we employed another Haar cascade classifier. By doing these steps, we finished first three steps which is shown in block diagram. For the third step, we needed to calculate biometry which is head and body ratio by equation:

$$\bar{O} = \lambda_H / \lambda_B \quad (1)$$

where λ_H and λ_B are height of head and body respectively, and \bar{O} means the ratio of biometry. There is a reason for using this ratio. It is medical fact that human's head size reaches optimum level before they become teenager. From this point of view, we can guess that this ratio can be discriminative between children and adults. We cannot measure people's heights from pixels in digital images. The reason is that it can be deceptive according to distance between pedestrian and camera. So, we need to solve this measurement problem. If we apply relative measurement which is related to proportioning head to body length by using classifiers, we would solve that problem. By this idea, we can classify that pedestrian is child or adult. If head over body ratio is bigger than 0.31, we can then classify that pedestrian is child. If that ratio is smaller than 0.31, that means pedestrian is adult. This is the main goal of the study. Also, with this proposition we can get low computation cost, and gain time. The reason is that we just focus on ratio, so we are not dependent to distance. Until now, the main idea of the proposed algorithm is explained, and threshold is mentioned. Now, we are focusing on how to calculate threshold value. To calculate threshold value λ_H / λ_B is measured on different videos which are taken in different mediums, with different positions.

Three main areas are assigned manually for observing relative measurement changes, and 4 different mediums are used to calculate optimum threshold value. Since it is calculated manually and four of all mediums have different conditions, the distance between bounds of areas might be slightly different than each other. Even though mediums, camera angles and distance are different, this situation gives better consideration points for obtaining optimum threshold value.

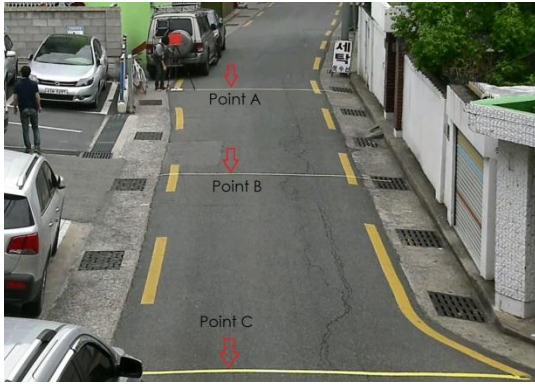


Figure 3. Separation points in the scene

These three areas are shown in figure, and the main idea is to see whether or not distance effect the ratio. For this reason, videos are collected and measurement is applied on both children and adults.



Figure 4. Separation points in the scene

The ratio λ_H / λ_B is calculated by counting pixel values manually. First, pixel values for height of head and then height of full body is calculated. All of values for children and adults are collected and their average values are calculated. The calculation rates for each groups in different mediums are given in the table.

Table 1. Ratios in different positions in the scene

Point	Medium 1			Medium 2		
	A	B	C	A	B	C
Child	0.30	0.32	0.32	0.32	0.33	0.32
Adult	0.24	0.23	0.23	0.27	0.26	0.26
Point	Medium 3			Medium 4		
	A	B	C	A	B	C
Child	0.31	0.31	0.33	0.32	0.32	0.30
Adult	0.28	0.27	0.28	0.28	0.27	0.29

According to table, average child and adult threshold values are 0.316 and 0.263 respectively. In the table, it is possible to see that adult's ratios are less stable than children. The main reason is that some of adults are really tall, and some of them are short. That issue affects the average ratio. However, the children's heights are quite

close to each other. Therefore, children's threshold value is taken as a reference threshold value. To be sure for exact threshold value, this calculation result is also verified by trial and error technique.

4. EXPERIMENTAL RESULTS

In this chapter, experimental results of this study will be focused on. First of all, four different videos are used to see how accurate this algorithm is. Also some of those videos' sizes are 1920×1080, and the rest of it is 1280×720. The aim was working on different media to get better idea for accuracy. Therefore, videos were taken in different conditions in Korea. Besides, each video has different view angle, illumination and distance from the target objects. OpenCV v2.4.9 via C++ on Microsoft Visual Studio 2010 is used for testing.

To classify pedestrians, we used two different Haar cascade classifiers which are for full body detection and head detection. 800 positive and 1600 negative images are used for full body XML file. Also HS.xml file, that Modesto Castrillon-Santana has copyrights, is used for head detection [9].

In this study, objects which are full body and head, are represented with different color rectangles. For full body detection, green color rectangle is used. However, when it comes to head detection, things are getting complicated. According to proposed algorithm, first we detect full body, then we focus on detecting head, after that we apply relative biometric measurement. If the ratio is bigger than threshold, that means the object is child. So, head rectangle will be shown as blue rectangle. If the ratio is smaller than threshold, that means the object is adult. So, head rectangle will be shown as red rectangle.

Results are given in two main categories: results of detecting and classification child and adult, and consideration on performance and accuracy rates of tested videos.



Figure 5. Child adult classification in different mediums

Results of four different videos are shown above. Two of them are adult detection, and the other two are child detection.

The reason of working on different media, is to see the real performance rate of this study. The first video is taken in subway, the second one is taken from street and the last two videos were taken from same street which is different street than the second video's medium. All media are in Busan, South Korea. Also, detection continuity is shown in the Figure 6.

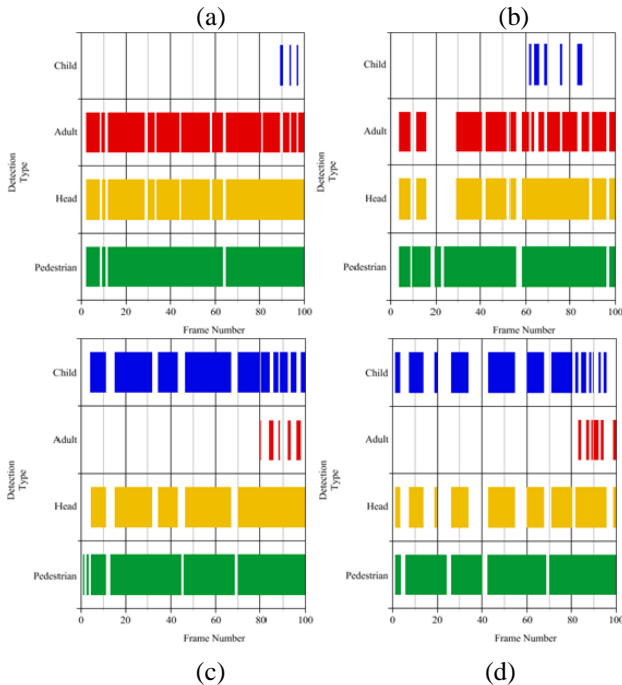


Figure 6. Detection continuity of experimented videos:
(a) Video 1, (b) Video 2, (c) Video 3, (d) Video 4

4.1. Considerations

Four different video results are shown in last section. According to the results, it is possible to say that classification for child and adult is working properly. However, numerical results are needed to see how accurate the proposed video based child and adult classification, is. Based on this, performance rates of each pedestrian, head, child and adult detection are given in the table within details, also average rates are calculated as well.

Table 2. Accuracy rate for detection and classification

	Full Body	Head	Adult	Child
Video 1 (Adult)	94.5%	90.8%	87.3%	2.7%
Video 2 (Adult)	89%	70.4%	62.1%	9.6%
Average (Adult)	91.75%	82.1%	74.7%	6.1%
Video 3 (Child)	93.7%	82.7%	6.6%	81.7%
Video 4 (Child)	92.5%	62%	8.6%	54.5%
Average (Child)	93.1%	72.35%	7.6%	68.1%

Detection and classification processes are implemented in narrow distance area where is between 2 meters and 10 meters away to camera. As the object goes away from camera, it is difficult to detect full body and head of human. That's why we work on narrower distance, and numbers of frames are not much.

Even though same algorithm is worked on those videos, the accuracy difference is relatively different. To answer that, there are couple reasons. Primarily, the biggest problem is having problem with detecting head. Because, once we find the head in the full body continuously, classification works successfully. However, failing for detecting head, decreases the total accuracy. For example, in the first video, woman is walking in subway station and since there is no other moving object, and her clothes' colors are totally different than background, the medium is totally proper to make pedestrian detection. Also, she is pretty close to the camera. For that reason, continuous head detection is possible and results of first video support this theory.

Same as the first two videos, the problem is same for the last two videos. That is difficulty of having continuous head detection. The XML file that we got to use on head detection, is trained for frontal angles. For theory, it can be useful, but in practice pedestrians always change the direction of their heads which causes decrements in the accuracy rate. Otherwise, this study's accuracy rate would be much better. Making a new proper XML file which is trained for every kind of head position, may solve this problem permanently.

5. CONCLUSION

Object detection is applicable for certain objects such as pedestrians, vehicles, or any other objects. Up to now, there are so many approaches and applications in various areas of computer vision about object detection.

In the process of time, children appear in daily life more than they have been. For example, children are under danger more than adults. Also criminal results show that child abduction rates are getting bigger and bigger every day. We can increase examples for sure. So we aimed to create an algorithm for child and adult classification for the benefits of children.

In this study, we proposed a new classification method to classify adult and child in videos with biometry. The absolute size of object is difficult to measure, since its size changes as the distance of its position changes in the scene. So, we used the information of head and full-body ratio instead of the exact size. Even though it cannot estimate the heights of pedestrian exactly, it showed that this approach is useful for adult and child classification. We tested four different videos which are taken in unequal conditions.

Experimental results showed that the proposed algorithm can classify pedestrian with 74.7% for adult, and 68.1%

for child. According to the performance graphs, continuous head detection, without any loss, is required. The reason is that, if we lose head detection, we will automatically lose the ratio calculation since we don't have any head information. So, creating more successful Haar cascade classifier for head, could be useful for better accuracy. In the future study, we will focus on improving the accuracy and expanding the application area such as night time age classification.

ACKNOWLEDGMENT

This work was supported by the Leading Human Resource Training Program of Regional Neo Industry through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (No. 2016H1D5A1910536) and Brain Busan 21.

REFERENCES

- [1] K. Lee, C. H. Lee, S. A. Kim, and Y. H. Kim, "Fast Object Detection Based On Color Histograms and Local Binary Patterns", presented at the TENCON 2012, IEEE Region 10 Conference, Cebu, November, 2012.
- [2] P. Viola, M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features", Computer Vision and Pattern Recognition, 2001.
- [3] Internet: OpenCV Dev Team, "*penCV Dev Team, and Pattern Recogn*", http://docs.opencv.org/3.0-beta/doc/py_tutorials/py_objdetect/py_face_detection/py_face_detection.html, 2014.
- [4] G. R. Rakate, S. R. Borhade, P. S. Jadhav, 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, 2012.
- [5] Y. H. Kwon and N. da Vitoria Lobo, "Age classification from facial images", Computer Vision and Image Understanding, 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. 183-191, 2001.
- [7] A. Lanitis, C. Draganova, and C. Christodoulou, "Comparing Different Classifiers For Automatic Age Estimation", IEEE Transactions on Systems, Man and Cybernetics, vol. 34, no. 1, pp. 621-628, 2004.
- [8] X. Geng, Z. Zhou, Y. Zhng, G. Li, and H. Dai., "Learning from facial aging patterns for automatic age estimation", inProc. ACM Multimedia 2006, pp. 307-316, 2006.
- [9] M. Castrillon-Santana, O. Deniz- Suarez L., Anton -Canalis, and J. Lorenzo - Navarro, "Face and Facial Feature Detection Evaluation", VISAPP 2008, 2008.