2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)
July 6-11, 2014, Beijing, China

Rapid Face Detection Using an Automatic Distributing Detector Based on Fuzzy Logic

Wanjuan Song,Wenyong Dong,Jian Zhang

***Abstract*—To improve the efficiency of a face detector, this paper presents an automatic distributing detector (ADD) based on the fuzzy theory to improve the performance of face detection. The main contributions lie in:1) A new Haar-like feature representation based on the fuzzy membership function is proposed, 2)The entropy of feature set is employed as choice criteria to select weak classifiers , 3) The AdaBoost algorithm is used to train weak classifiers, and 4)The distributor which can dynamically select stronger classifiers is constructed. The experiment results show that the proposed method not only determines rapidly the sub-window which contains the human face, but also tune the classifier dynamically to adaptive new samples. The accuracy and speed of our method are also promoted comparison with the state-of-art detectors. On the other hand, as for the image sub-window which is like face, according to the value of membership function, distributor can dynamically select the remaining stronger classifiers to determine. This detector can effectually improve detection speed and has better detection performance.**

I. INTRODUCTION

Face detection is an important and fundamental issue in pattern recognition and computer vision which mainly based on statistical models[1]-[3]. Recently the researches about face detection mostly focused on face localization in the pixel image. Many factors may affect the detection result, such as face expression, surrounding light and so on.

In the recent years, there emerged many face detection algorithms. Some of them were based on statistical learning method. The content involved every facet of the pattern recognition, such as Principal Component Analysis(PCA)[4], Artificial Neural Network(ANN)[5],[6],[14], Bayes decision rule[7], Support Vector Machines(SVMs)[8],[9].In these methods, Rowley's ANN[14] was the most typical. It not only improved the efficiency of face detection, but also provided a way which can use statistical theory to deal with the problems in face detection.

Up to now,Viola's method[10] was the most effective method in face detection , which used AdaBoost [11] based on Harr-like features. This method provided a new image representation called integral image. It can evaluate rectangle features very rapidly. Viola constructed his classifiers by selecting a small number of important features using AdaBoost. Cascade detector was the vivid description for Viola's system. In the condition of detection performance similar with [14], its speed was roughly 15 times faster than [14]. And it had been implemented in real time.

In Viola's system, cascade detector screened the sub-windows layer by layer, which leads to two problems. The first problem is the slower speed of the detector, the other one is the lower utilization rate of the features. So there are two aspects which could be improved in the detection efficiency and the utilization rate of the features.

This paper introduces how the fuzzy set theory applies to face detection. The ADD built with this theory and use Harr-like rectangles as features. The feature's membership is selected as detecting parameter. During the training process, the ADD will select appropriate weak classifiers by AdaBoost learning algorithm. Finally these weak classifiers will construct some stronger classifiers. According to the similar degree between image sub-windows and human face, the ADD can dynamically select stronger classifiers to determine whether the sub-window is a face. This determination method for sub-windows was adapted in ADD instead of the screening sub-windows layer by layer mentioned in Viola's method [10]. In this point, the image features are fully utilized. Just because the ADD mustn't select every stronger classifier to determine a face sub-window, it can improve the detection efficiency. There are three key contributions. 1) The fuzzy set theory is applied in the face training and face detection. It can improve detecting speed. 2) The weak classifiers are constructed by selecting an important feature using AdaBoost and feature set's entropy. 3) Building a distributor which can dynamically select stronger classifier to determine a sub-window. In the detection process, the ADD evaluates sub-window's average membership in a stronger classifier. If this membership is less than threshold value, this sub-window will be deeming as non-face. On the contrary, the ADD will send parameters to distributor, and it will dynamically select the remaining stronger classifiers to determine whether this sub-window is a face.

The following sections of the paper discuss how the fuzzy set theory is used in face detection, implementation of the detector and the experiment results. Section 2 will introduce some basic knowledge about fuzzy set theory which can apply to face detection. Section 3 will focus on face detection based on fuzzy set theory, including Harr-like features and integral image, the method of how to select weak classifiers, and how to construct distribution face detector. Section 4 is the experiment results and analysis. Finally, Section 5

Wanjuan Song is with the Computer School, Wuhan University, Wuhan city, China (email:key_swj@163.com). She is presently with Co-Innovation Center of Elementary Education Information Technology Services of Hubei.
Wenyong Dong is with the Computer School, Wuhan University, Wuhan city, China (Corresponding author dwy@whu.edu.cn).
Jian Zhang is with the Computer School, Wuhan University, Wuhan city, China. He is presently with the College of Sport Engineering and Information Technology,Wuhan Sports University.
Manuscript received Dec 20, 2013. This work was in part supported by the NSF of China under Grant Nos. 61170305 and 60873114.

summarizes the paper.

II. FUZZY SET THEORY

A. Membership Function

Usually, there are two types of membership function: S function and π function. S function is a monotone increasing function with the value range of 0 to 1. Three parameters (a, b, c) can confirm this function. Zadeh defined a normative S function, which is

$$\mu_{AS}\left(x_i;a,b,c\right)=\begin{cases}0, & x_i \le a\\ 2[(x_i-a)/(c-a)]^2, & a\le x_i\le b\\ 1-2[(x_i-c)/(c-a)]^2, & b\le x_i\le c\\ 1, & x_i\ge c\end{cases} \tag{1}$$

A π function, whose function image is like a π, can be defined by S function:

$$\mu_{A\pi}(x_i;a,c,a')=\begin{cases}\mu_{AS}(x_i;a,b,c), & x_i\le c\\ 1-\mu_{AS}(x_i;c,b',a'), & x_i\ge c\end{cases} \tag{2}$$

Note x_k as the kth character, $(x_k)_{av}$, $(x_k)_{\max}$, $(x_k)_{\min}$ as average value, max value, min value, we have

$$c=(x_k)_{av} \tag{3}$$

$$b'=c+\max\left\{\left|(x_k)_{av}-(x_k)_{\max}\right|,\left|(x_k)_{av}-(x_k)_{\min}\right|\right\} \tag{4}$$

$$b=2c-b', a=2b-c, a'=2b'-c \tag{5}$$

B. The Entropy of a Fuzzy Set

The entropy of a fuzzy set describes how the elements belonging one set are in a group, which is defined as

$$H(A)=\frac{1}{n\ln 2}\sum_{i=1}^{n}S_n(\mu_A(x_i)) \tag{6}$$

where

$$S_n(\mu_A(x_i))=-\mu_A(x_i)\ln\mu_A(x_i)-(1-\mu_A(x_i))\ln(1-\mu_A(x_i)) \tag{7}$$

For π function, if the set's entropy value is small, it means more elements in a group and the easier classifying. However, S function is converse to π function.

III. FACE DETECTION BASED ON FUZZY SET THEORY

In Viola's paper[10], a new image representation called the "Integral Image" was introduced and a method based on AdaBoost to train classifiers was given. In this paper, we (draw some experience from Viola's) also use Harr-like features as our detection system's features, calculate membership as every feature's parameter, select appropriate weak classifiers by AdaBoost learning algorithm, and then construct a distribution face detector. In the following we will introduce our method.

A. Harr-like features and integral image

The value of a Harr-like feature is the difference between the sum of the pixels within white and black rectangular regions (see Fig1). It can reflect the gray transformation of a sub-window.

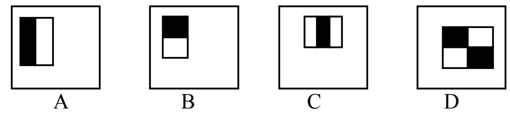

Fig.1. Harr-like features. Two-rectangle features are shown in (A) and (B). (C) shows a three-rectangle feature, and (D) is a four-rectangle feature.

To rapidly calculate the rectangle features' value, Viola designed an intermediate representation for the image called the integral image.

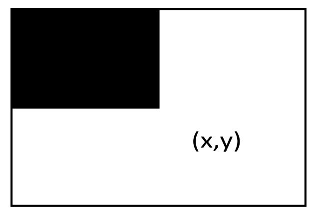

Fig.2. The value of the integral image at point (x, y) is the sum of all the pixels above and to the left.

The integral image at location (x, y) contains the sum of the pixels above and to the left of (x, y), inclusive:

$$ii(x,y)=\sum_{x_i\le x,y_i\le y} i(x_i,y_i) \tag{8}$$

Where $ii(x,y)$ is the integral image and $i(x_i,y_i)$ is the original image (see Fig 2). Using the following pair of recurrences:

$$s(x,y)=s(x,y-1)+i(x,y) \tag{9}$$

$$ii(x,y)=ii(x-1,y)+s(x,y) \tag{10}$$

The integral image can be computed in one pass over the original image. (Where $s(x,y)$ is the cumulative row sum.)

In the detection process, any rectangular sum can be computed rapidly by the integral image. Detector scans original image only one time. Since the two-rectangle features defined above involve adjacent rectangular sums, they can be computed in six array references, eight in the case of the three-rectangle features, and nine for four-rectangle features (see Fig3). The value of integral image at location 1 is the sum of the pixels in rectangle A. The value at location 2 is A + B, at location 3 it is A + C, and at location 4 it is A + B + C + D. The sum within D can be computed as 4 + 1 – (2 + 3).

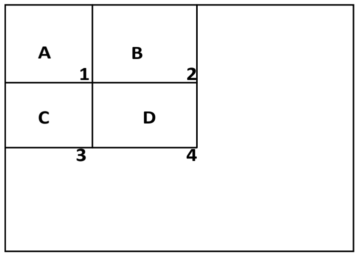

Fig.3.The sum of the pixels within rectangle D can be computed with four array references.

B. Weak Classifier

For a 19×19 pixels face image, the exhaustive set of rectangle features is quite large. Washing out some improper rectangle features, there leave 73984 candidates can be selected. First, we train these features on the positive examples to get their distribution. According to the result, π function (formula (2)) is more suitable for our system. And then we get every feature's max value, min value, average value and entropy. Finally, we train these features on the negative examples and select appropriate features by the following formula:

$$h_j(x) = \begin{cases} 1 & if \quad \mu_{j\pi}(x) > 1 - H_{aj} \\ 0 & \text{others} \end{cases} \tag{11}$$

where h_j is the jth feature, x is a 19×19 negative image, $\mu_{j\pi}(x)$ is x's membership within the jth feature, and H_{aj} is the jth features' entropy.

A single rectangle feature which best separates the negative examples can be selected by formula (11). For the examples which are classified incorrectly, the training system will change these examples' weight, and then repeat above steps. So every time we will get a weak classifier. And its membership threshold value is $1 - H_{aj}$.

C. Distribution Face Detector

In Viola's paper, a training algorithm for building a cascaded detector is proposed [10], [13]. We will use a similar algorithm to build a distribution detector; the difference is our distribution detector employs membership μ as a parameter. Its framework is :

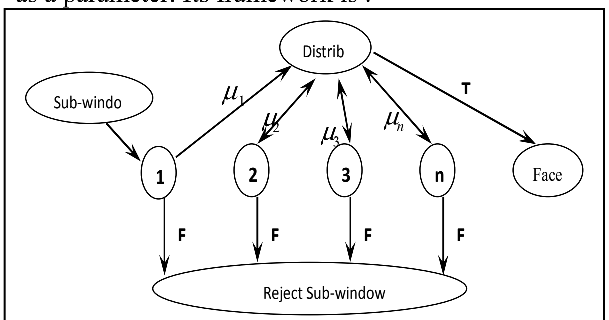

Fig.4. Distribution face detector

This detector is composed of some stronger classifiers and a distributor. Every stronger classifier includes some weak classifiers. The frontal stronger classifier is simple and includes several weak classifiers. So it can quickly wash out the most negative sub-windows. The behind stronger classifier includes more weak classifiers, and has more powerful classifying ability. In order to wash out the most negative sub-windows, the frontal stronger classifiers' membership μ_i' are equal to the max of their weak classifiers'. And the behind stronger classifier must judge sub-windows exactly, so their membership μ_i' are equal to the min of their weak classifiers'. When a sub-window is entered, the first stronger classifier will evaluate its membership by each of its weak classifiers, and get the average value μ_i . If $\mu_i \leqslant \mu_i'$, the sub-window will be washed out, otherwise it will return μ_i to the distributor, and according to μ_i ,distributor will select behind stronger classifier to detect this sub-window, until the result is concluded. The detection algorithm is

TABLE I
DETECTION ALGORITHM

x is the sub-window
i = 1
μ_i' is membership threshold value of the ith stronger classifier
$f(i, \mu_i(x))$ is the distributor function
where i ≤ n
evaluate the membership $\mu_i(x)$
if $\mu_i(x) \leq \mu_i'$
Sub-window is not a face, return false
i = $f(i, \mu_i(x))$
Sub-window is a face, return true

where

$$f(i, \mu_i(x)) = i + \left\lceil (n-i)\left(\frac{(\mu_i(x) - \mu_i')}{1 - \mu_i'}\right) \right\rceil \tag{11}$$

According to sub-window's membership, distributor can dynamically select the remaining stronger classifiers to determine whether a sub-window is a face. Usually if sub-window is like a face, its membership value $\mu_i(x)$ will be near to 1, and according to formula (12), distributor will select the behind stronger classifier. Through the algorithm described in table 1, our detector will have some characters, which are: For the sub-window which is not like a face, the detector rapidly rejects it by the front simple stronger classifiers; for the sub-window which is little like a face, the detector selects more remaining stronger classifiers to evaluate it from the details; for the sub-window which is much like a face, the detector selects behind stronger classifiers which have stronger classifying ability to determine whether it is a face.

D. Cascade Detector

Viola's cascade detector [10] looks like a filter, where every stronger classifier rejects the number of negative sub-windows (see fig5).

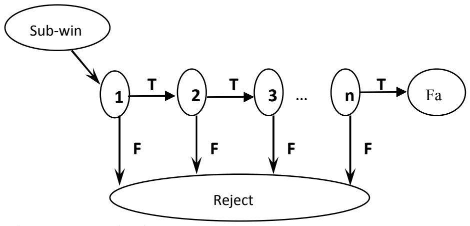

Fig.5. Cascade detector

If a face sub-window enters this detector, every classifier must check it. But our distribution detector can dynamically select the stronger classifiers to detect sub-windows. If the membership of this sub-window is much high than first classifier's, according to formula (12), some of classifiers will be skipped, and distributor will select the behind stronger classifier to check it. In the view of utilization of the characters, our detector is more efficient than Visla's.

IV. Experiments and Results

A. Training

We have trained our system on the MIT+CMU training set[14]. This training set includes 2429 positive images and 4548 negative images. The size of these images is 19 by 19 pixels. Every positive image comprises most important characters of frontal face except hair and ears. A part of training examples are shown in fig6.


Fig.6.A part of frontal upright face images used for training.

First, through the preconditioning, system selects 73984 candidate rectangle features from a 19 by 19 pixels face model. Then after training, it selects 2500 features from those candidates. All the selected features were sorted by classifying ability in descending order. The first ten selected features are shown in Table II.

The first feature measures the difference in intensity between the region of the eyes and a region across the upper cheeks. The feature capitalizes on the observation that the eye region is often darker than the cheeks.

At the first feature all the positive images' value spread from -7300 to 1400. They are shown in fig8.

In this fig the horizontal axis is the feature value of positive image, and the vertical axis is positive image's count. This image is like a π, this is why we select π function (formula (2)) to calculate feature's membership.

TABLE II
THE FIRST TEN SELECTED FEATURES

FEATURE TYPE	BEGINX	BEGINY	ENDX	ENDY	MINV	MAXV	AVGV	HA
B	0	2	16	9	-7258	1372	-2904.8118	0.259197
B	12	2	17	9	-2808	272	-1131.1577	0.220802
C	5	0	12	16	-440	4008	1561.3767	0.189701
B	2	3	16	8	-4188	604	-1596.03	0.251247
A	5	1	9	9	-1772	467	-655.6641	0.237739
C	6	0	11	15	-251	2427	889.6567	0.173752
B	2	3	5	8	-1156	387	-440.2783	0.191847
B	8	7	11	13	-310	1529	476.8958	0.282581
C	8	0	17	15	-4221	1711	-1285.1976	0.19327
C	6	0	11	15	-251	2427	889.6567	0.173752

While feature type is the Harr-like feature(see fig1) . (BeginX, BeginY) and (EndX, EndY) show the rectangle's begin point and end point. MinV, MaxV and AvgV show the positive images' min value, max value and average value at this rectangle feature. HA is this rectangle feature's entropy. For the first rectangle feature it is shown in fig7.


Fig.7. The first rectangle feature

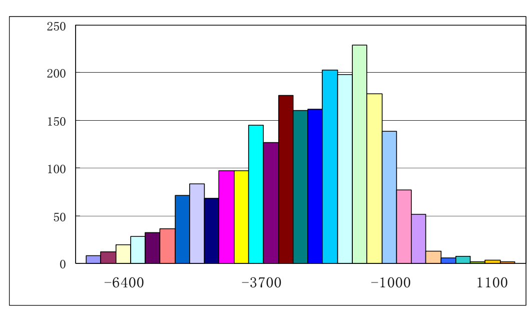

Fig.8.The distribution of positive images at the first feature

All the negative images' value spread from -10000 to 7000 at this feature. It is shown in fig 9.

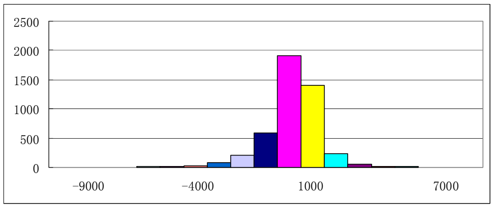

Fig.9.The distribution of negative images at the first feature

The comparison between positive images and negative images is shown in fig10.

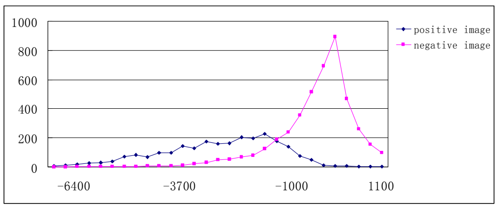

Fig.10.The comparison between positive images and negative images

In fig10 there is much different distribution between positive images and negative images. They have different grouping fields. Their distribution of membership at the first feature is shown in fig11.

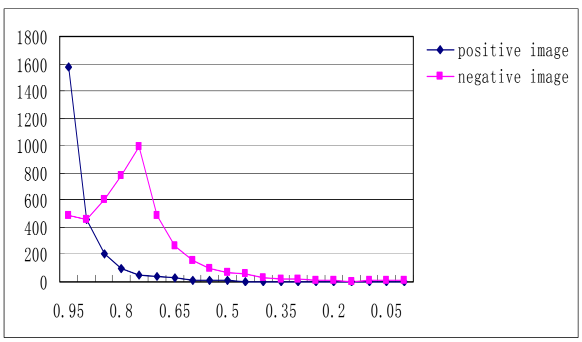

Fig.11.The distribution of membership

The horizontal axis is membership value, and the vertical The horizontal axis is membership value, and the vertical axis is image's count.

From the fig11, we can see that for the positive images most memberships are near 1. And as the descending of membership value, the count of positive images falls off sharply. This pattern is perfect for a fuzzy set which can be used to describe the degree how an image is similar to human face. But for the negative images they have much different curve. So the membership has ability to distinguish the face images from non-face images, and at the same time it has the character of fuzzy set.

Using these selected 2500 features build a detector which has 43 stronger classifiers. Each stronger classifier classifies sub-windows by their membership threshold value μ_i'. If sub-window's membership is less than μ_i', the detector will deem it as non-face, and rejects it; If membership is greater than μ_i', detector sends parameters to distributor. In terms of parameters' value, distributor selects appropriate stronger classifier from remaining.

B. Detection

We tested our system on the MIT+CMU frontal face test set[14]. This set consists of 130 images with 507 frontal faces. The size of initial scanning window is 30 by 30, starting scale is 1.25 and step size is 1.5. The max size of scanning window is 240×240. And this is the same with Viola's. The results are shown in table Ⅲ.

TABLE Ⅲ
DETECTION RESULTS ON MIT+CMU FRONTAL TESTING SET

Detector \ False Alarm	5	10	17	31	50	62	95	193	422
Our Method	75.7%	78.7%	79.4%	83.9%	86.1%	86.8%	89.3%	90.3%	92.2%
Viola-Jones	-	78.3%	-	85.2%	88.8%	-	90.8%	-	93.7%
Rowley	-	83.2%	-	86%	-	-	89.2%	-	89.9%

Comparing with Viola's, in the situation of same false alarms, detector's performance only declines a little. But it is better than Rowley'. This result is better suit for our expected, but not the best. So we analyze the false detection. We find that most of false alarms emerge from the low-resolution images. When the resolution of image is less than 250 thousand pixels and there are more than eight faces emerging in this image, the count of false alarm increases rapidly.

So we tested our system on another image set. This image set includes 117 our own photos with 327 frontal faces. And the resolution of these images is more than 500 thousand pixels and there are less than ten faces emerging in one image. The detection rate is 96.94%. This is obviously encouraging for us.

So we can get a conclusion that a face's resolution is the key for our detector. If the resolution of a face is lower, this face will be blurry. And for the fuzzy set, a blurry face can belong to face set in a way, but at the same time it can belong to other sets in other ways. So the classifiers are hard to get a right judgment.

Viola has tested his cascade detector on a 700 Mhz Pentium III processor, and the pixel of video image is 384×288. It can process a image in about 0.067 seconds [10]. In order to comparing with Viola's, we also test our system on a 700 Mhz Pentium III processor. In the same pixel of video image, its time is about 0.062. The speed is faster than Viola's.

For testing our detector in real-time, we run it on a 2.27GHz Intel(R) Core(TM) i3 CPU, and the pixel of video image is 384×288. It can process an image in about 0.051 seconds. But there is a problem referred above. That is low-resolution image will impact the detection rate. And now most video cameras' resolution is 640×480. So for getting higher detection rate and suiting the normal video camera, we tested our detector on the same computer, but the pixel of video image is 640×480. And for improving the detecting speed in the higher resolution, the scanning window's starting scale and step size must be changed. The size of initial scanning window is 30 by 30, starting scale is 1.35 and step size is 2. The max size of scanning window is 240×240. The detecting speed is about 0.076 seconds. It means the detector can process 13 frames in a second.

Fig.12. Some detection results of frontal faces

V. CONCLUSIONS

This paper introduces a new face detection method based on fuzzy set theory, and describes how a face detector is built by Haar-like features, integral image and AdaBoost learning algorithm. According to similar degree between image sub-window and human face, this detector can dynamically select stronger classifier to scan the sub-window. The results show a face detector based on fuzzy set theory can effectually improve detection speed.

It is a new attempt of applying fuzzy set theory to face detection. Furthermore, this paper only describes how this theory is applied to the frontal face detection. For the profile and rotation face detection, there is much work waiting for us. The preliminary results suggest that this detector can work with them. Cascade detector is widely applied in pattern recognition [15], [16]. The constructed way of the classifiers could be discussed. But for getting better result; further research is needed to be done in our future work.

REFERENCES

[1] Cha Zhang , Zhengyou Zhang. A Survey of Recent Advances in Face Detection. Technical Report, Microsoft Research, MSR-TR-2010-66, June 2010.

[2] Krestinin, I.A., Seredin, O.S. Excluding Cascading Classifier for Face Detection. Proc. of the 19th Int. Conf. on Computer Graphics and Vision, 380－381 (2009)

[3] P. Kakumanu, S. Makrogiannis, N. Bourbakis. A Survey of Skin-Color Modeling and Detection Methods. Pattern Recognition Volume 40, Issue 3, March 2007, Pages 1106–1122

[4] Hazem M. El-Bakry. New Fast Principal Component Analysis for Real-Time Face Detection. Machine Graphics & Vision International Journal ,Volume 18 Issue 4, April 2009 , Pages 405-425

[5] R. Feraud , O.J. Bernier , Jean-Emmanuel Viallet , et al. A Fast and Accurate Face Detector Based on Neural Networks. IEEE Trans. Pattern Analysis and Machine Intelligence , 2001 , 23 (1) : 42-53

[6] Chiunhsinu Lin. Face Detection in Complicated Backgrounds and Different Illumination Conditions by Using YCbCr Color Space and Neural Network, Pattern Recognition Letters, vol.28, issue 16, pp.2190-2200, 2007.

[7] Bo Wu, Ram Nevatia. Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of Edgelet based Part Detectors. International Journal of Computer Vision Volume 75, Number 2 (2007), 247-266.

[8] P.Wang and Q. Ji. Multi-view face detection under complex scene based on combined svms. In Proc. of ICPR, 2004.

[9] K. Hotta. View independent face detection based on combination of local and global kernels. In International Conference on Computer Vision Systems, 2007.

[10] Viola P, Jones M.J. Robust Real-time Face Detection. International Journal of Computer Vision. 57(2).237-157. 2004

[11] T.-K. Kim and R. Cipolla. MCBoost: Multiple classifier boosting for perceptual co-clustering of images and visual features. In Proc. of NIPS, 2008. 11

[12] L. A. Zadeh, Fuzzy sets, Information and Control, vol. 8, pp.338-353,1965

[13] Jun Nishimura. Tadahiro Kuroda. Versatile Recognition Using Haar-like Feature and Cascaded Classifier[J]. IEEE Sensors Journal,2010,(5):10

[14] H. A. Rowley. Neural network-based human face detection : [Ph. D. dissertation]. Pittsburgh , USA : Carnegie Mellon University , 1999

[15] Tian.H and Duan.Z .A novel multiplex cascade classifier for pedestrian detection ,Pattern recognition letters,vol.34.no.14,pp.1987-1693,2013

[16] Wen. J,B. Xiong.YS and Wang. SL . A novel two-stage weak classifier selection approach for adaptive boosting for cascade face detector. Neruocoputing,vol.116,pp..122-135.2013