2011 3rd International Conference on Signal Processing Systems (ICSPS 2011)
IPCSIT vol. 48 (2012) © (2012) IACSIT Press, Singapore
DOI: 10.7763/IPCSIT.2012.V48.1

An Algorithm for Accurate Face Detection in Complex Background and Illumination

Lihua Huang , Pingan Mu, Shuguang Dai , Qichang Li+

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai, China

Abstract. Face detection is the basis for face recognition, which makes it critical to improve its detection efficiency. In order to enhance the robustness and accuracy of face detection, this dissertation proposes a detection algorithm, which combines YCbCr color model with neural network classifier theory, realizing face detection with high accuracy in complex background and illumination. The experimental results show that this algorithm's accuracy is up to 95%.

Keywords: face detection; YCbCr model; neural network; SNOW classifier; LabVIEW

1. Introduction

Face recognition is a technology making use of visual features of face and computer technology for identification information [1]. Due to its advantages of not disturbing tested individuals, face recognition has been widely used in fields like identification, video surveillance, access control systems, residential safety, and information security [2]. Face detection as basis of face recognition, its accuracy directly affects following operations. Nowadays, face detection algorithms are mainly divided into three categories: skin-color-based model [3], facial-feature-based model and statistical-based model [4]. In facial-feature-based model, first, obtain facial features, skin gray values texture and other features of face, and then determine whether it's face or not based on priori knowledge. This algorithm needs many face features and involves in image understanding, which is a difficult issue in image processing, thus limiting its further development. Skin-color-based model is based on clustering of skin pixels, so it can quickly locate the skin region. However, in this model, the skin region which doesn't contain a face and the region which is similar to skin but not skin will greatly affect the accuracy. Statistical-based model treats human face as a class, training a classifier with large numbers of face and non-face samples, and then classifies all the possible regions in the image. This model is so much like two classification problem in statistical pattern recognition. It is efficient for images with one or multiple faces in complex background, but computation is too large.

2. Candidate human face region locating in complex background and illumination

As is discussed earlier, statistical-based model has higher location accuracy and robustness, but large computation. Since computation is proportional to data size, to overcome this problem, this paper proposes locating candidate face region firstly based on skin-color-model algorithm, and then accurately locating face from candidate region. Skin color models include RGB, YCbCr, HSL and so on. Research shows that distribution of skin color in YCbCr model is the most concentrated and clustering [5]. Regardless of race, skin color, in YCbCr model, the Cb and Cr value are always kept in a specific range. So YCbCr model is the best choice for face detection. First, capture images with a camera, and then transform images from RGB model into YCbCr model according to the following (1) [6].

+ Corresponding author. *E-mail address*: hlh818@163.com

$$\begin{cases} Y = 0.299R + 0.587G + 0.114B \\ Cb = -\frac{R}{6} - \frac{G}{3} + \frac{B}{2} + 128 \\ Cr = \frac{R}{2} - \frac{5G}{12} - \frac{B}{12} + 128 \end{cases} \tag{1}$$

Among Y, Cb, Cr of face pixels, Cb and Cr meet (2).

$$\begin{cases} 108 \le Cb \le 123 \\ 135 \le Cr \le 156 \end{cases} \tag{2}$$

Create a binary image according to (3), with the same size as original image.

$$f(x,y) = \begin{cases} 1, 108 \le Cb \le 123 \,\&\, 135 \le Cr \le 156 \\ 0, other \end{cases} \tag{3}$$

It is found there are many noise points in the binary image, directly affecting detection accuracy. Thus it’s necessary to make some morphological operations such as corrosion, swelling and others to the binary image. Among all the connected domains, candidate face region must be the largest. After filtering smaller regions, a more integral facial candidate region is formed. Multiply the obtained binary image with Y Plane image in YCbCr model to get candidate face gray image. Because non-skin regions are quickly removed, it greatly reduces the following calculation. In fig 1 are results of candidate human face region in a variety of backgrounds and illumination.

FIG. 1 HUMAN FACE CANDIDATE REGION

3. Face region locating

3.1. SNOW algorithm.

SNOW (Sparse Network of Winnows) for face detection is initially proposed by Roth [7]. The basic theory of this training system is to use pre-defined samples features to train a linear sparse network, and then get a classifier using the trained network. Each node in linear network directly corresponds with the sample features.

3.2. SNOW algorithm

Uses Winnow update rule to training network. [8] Simply speaking, Winnow rule is increasing or decreasing the weights of network nodes according to pre-defined positive or negative samples. In the face detection process, face samples are positive, non-face samples negative. The main training purpose is to determine the weights of network through continuous adjustment of weights and threshold until the total error meets the accuracy requirement. Let $A_t = \{a_1, \ldots, a_m\}$ be a set of features for positive samples, θ_t threshold, $\omega_{t,i}$ the weight of the ith feature to the target node t. If and only if

$$\sum_{a_i \in A^t} \omega_{t,i} > \theta_t \tag{4}$$

The sample is active. Otherwise, modify weights according to training rules.

In addition, the rule has two parameters, a promotion parameter $\alpha > 1$, and a demotion parameter $0 < \beta < 1$. When the sample is face, but $\sum_{a_i \in A_t} \omega_{t,i} \le \theta_t$, let

$$\forall a_i \in A_t, \omega_{t,i} \leftarrow \alpha \bullet \omega_{t,i} \tag{5}$$

When the sample is non face, but $\sum_{a_i \in A^t} \omega_{t,i} > \theta_t$, let

$$\forall a_i \in A_t, \omega_{t,i} \leftarrow \beta \bullet \omega_{t,i} \tag{6}$$

SNoW system generating the functions connecting inputs to output is based on how the inputs are represented. [8] In order to reduce computation, input to the system is collection of input nodes, instead of intensity value of each pixel.[7] Each collection is a Boolean representation. Since each intensity value of each pixel is linked with a weight, the distribution of weights associated with that pixel can approximate nonlinear functions of the intensity value.

After training with numerous face and non-face samples, a classifier is formed. When inputting an image containing human face, the network will locate face. The traditional method mainly uses image intensity value as a feature to train the network. As possible intensity value of each pixel is 0~255, the input matrix is quite large. For example, a 19x19 pixel image (361 total pixels) with 256 intensity values, has a 361x256 input matrix for snow representation, among which only 361 elements are active. The process takes up much resource, and works with low efficiency and poor robustness. To overcome these shortcomings, this paper redefines the sample feature and classifying algorithm.

3.3. An improved algorithm based on SNOW classifier.

In the new algorithm, first do image segmentation in training samples, and then map them to a new range. The steps are as follows:

First, calculate the local mean brightness of sample images.

Second, take the mean brightness as threshold, do segmentation on sample images. Name this binary image B (x, y).

$$B(x,y) = \begin{cases} 1 & f(x,y) \geq \dfrac{\sum_{i,j \leq 19} f(i,j)}{19 \times 19} \\ 0 & other \end{cases} \tag{7}$$

Last, map B(x,y) to B'(x,y) according to (8)

$$B'(x,y) = B(x,y) + 2\times B(x+1,y) + 4\times B(x,y+1) + 8\times B(x+1,y+1) \tag{8}$$

According to above algorithm, possible intensity value of each pixel is 0~15, thus making size of input matrix to the SNOW system decreased greatly, which means its size has been changed from 361x400 to 361x16. Obviously, the algorithm has greatly reduced computation.

This dissertation took MIT face database as training samples to train a classifier. MIT face database was created by the MIT Media Lab, containing 2429 face samples, 4548 non-face samples. [9] Each sample is gray image with size of 19X19, which is the smallest size that can be recognized as a face (see fig 2).

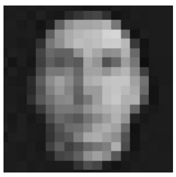


FIG. 2 A TRAINING FACE SAMPLE

Since computation is proportional to image size, the best way to save time is to reduce image size. Therefore, it's better to resample all test images into size of 100X100. Experiment results show that it does not affect accuracy of face locating. At last, the face is found and marked in the original image.

First, compute feature value of every training sample, and then train classifier. Do preprocessing on test images, such as image smoothing, illumination compensation, and so on. Find the candidate face region based on YCbCr model, compute features, and find human face according to the classifier. What's more, selecting the appropriate step for searching is also very important. In this dissertation, the step in x-direction and y-direction are both one pixel. Algorithm block diagram is as below:

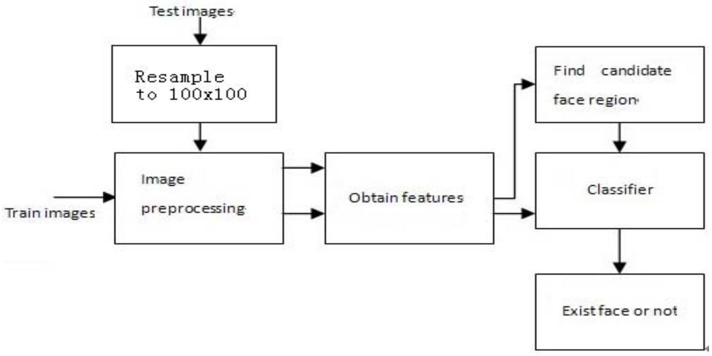
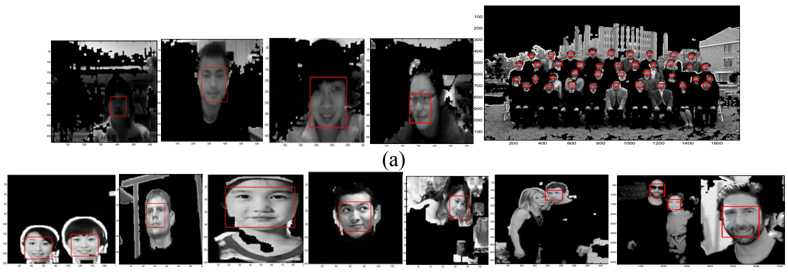


FIG. 3 IMAGE PROCESSING FLOW CHART

4. Experimental results

In the experiment, some images to test are acquired by a camera built-in a dell notebook, with image size: 640X480, and resolution: 130 million pixels; another 150 images from the Internet. Collect 150 images under different backgrounds, light, with a single face or double faces or more. Use NI LabVIEW VISION module with matlab programming language to test and verify algorithm. Download different images on the Internet 150. Experimental results show that among all 300 images, 285 of them are accurately located face. The accuracy was 95%.



(a)

(b)

FIG. 4 RESULTS OF FACE DETECTION (A) IMAGES FROM INTERNET (B) IMAGES TAKEN BY THE CAMERA

FIG. 5 FACE DETECTION IN SOURCE IMAGE

Compared fig 4 with fig 5，it is obvious that when there is no preprocessing on source image like finding candidate human face region, it is very likely to take non-face as face.

TABLE 1 COMPARISON OF TWO ALGORITHMS

Criteria	Algorithm	images			total
		One face (100 images)	Two faces (100 images)	More than 2 faces (100 images)	300
Consuming Time (average time)	SNOW	258ms	266ms	281ms	/
	Improved SNOW	171ms	156ms	270ms	/
Accuracy (300 images in all)	SNOW	97	93	85	275(91.6%)
	Improved SNOW	99	97	89	285(95%)

5. Acknowledgement

This research is supported by Leading Academic Discipline Project of Shanghai Municipal Education Commission, Project Number: J50505

6. References

[1] Yan Yan, Zhang Yujin, Advancement in video-based face recognition, Computer Science, 2009, No. 5

[2] Zhang Cuiping, Su Guangda, face recognition technology review, Journal of Image and Graphics, 2000 , No 11

[3] Ai Haizhou, Liang Luhong, Xu Guang, Zhang Bo, face detection based on templates and skin color, Journal of Software, 2001 12

[4] Liu Meng Fei, face recognition technology, a master's degree thesis, Wuyi University, 2009

[5] Guo Junbin, Guo Xiaosong, Pu Pengcheng, rapid positioning of human eyes based on the relationship between skin color and geometric, computer measurement and control, 2009,17 (3)

[6] Xu Feng, Zhao Huaixun, Tu Qiang, face detection based on skin color and eye positioning, computer systems applications, 2010, No. 2 19

[7] D. Roth, M. Yang, and N. Ahuja, “A snow-based face detector,”in In Advances in Neural Information Processing Systems 12

[8] Evan C. Smith, A SNoW-Based Automatic Facial Feature Detector, MPLab Technical Report TR 2001.06

[9] http://cbcl.mit.edu/software-datasets/FaceData2.html