

A low cost architecture for high performance face detection



Weina Zhou^{a,b,*}, Huafeng Wu^a, Xiaoyang Zeng^b

^a Shanghai Maritime University, China

^b Fudan University, China

ARTICLE INFO

Article history:

Available online 9 June 2015

Keywords:

Face detection
AdaBoost
High resolution
High performance
Low cost

ABSTRACT

Face detection has been playing an important role in numerous fields in recent years, and is considered to be a promising technology in the future. However, low cost implementation is still a difficulty due to huge computation of detection algorithm, especially when the detection is required to be applied in embedded systems. In this paper, a new architecture is proposed based on an efficient face detection algorithm. The architecture has three contributions. The first is a specially designed frame buffer which improves the data access efficiency and provides a big data throughput. The second is a new integral image refresh method, which can renew the integral image in one clock cycle to provide feature values for classification timely. The third contribution is a 4-stage pipeline structure for feature calculation, which improves the classification speed by 3 times almost without any increase in hardware area. The experiment results show that the architecture consumes less hardware and power resources while retaining a high-level of detection capability in processing 1024×1024 images.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Face detection is the process of determining the locations and sizes of human faces in a source image or video sequence. Face detection is now playing an important role in a wide range of applications [1–6], such as access control, video communication, digital camera, human computer interactions, monitoring and surveillance. As computer science develops, higher performance of detection is much urgently needed. That is to say, faces are required to be detected in real time with higher accuracy rates.

Thanks to the powerful processors and efficient face detection algorithms like the scheme proposed by Viola and Jones [7], faces now always can be detected at a real-time speed with a high accuracy rate in use of computers. However, to consider applications in portable handheld devices, the cost performance requirements are much higher than when using computers.

Huge computation is needed to ensure a favorable performance in face detection due to factors such as rotation, pose, illumination and scale, especially in processing high resolution images. The computation always requires too much power and hardware resources for a common embedded system to provide.

Existing face detection embedded systems [8–14] with real-time frame rates can be mainly classified into two kinds. The first kind reduces the resolution of input images or decreases the bits of pixels to cut down the data amount for processing, or uses oversimplified algorithms to detect the face. However, this method would result in a decline in accuracy and can only be used in favorable conditions or specific environment. The other kind of systems aims at constructing an efficient circuit structure to reduce power and hardware resources meanwhile maintaining excellent accuracy and speed. But their power or hardware consumption is still too much for embedded use.

Based on AdaBoost-based face detection algorithm, this paper proposes a new integral image (each location of the image holds the sum of all the pixels to the left and above of that location) refresh method and a 4-stage-pipeline architecture to accelerate the detection process and the transfer of the pixel data. A dual-port frame buffer is also designed to partially reserve the input data. The architecture proposed can process images with size up to 1024×1024 , and can achieve a 95% face detection accuracy rate. What is more, when it runs at 100 MHz, it consumes much less power and hardware area than previous works.

The remaining of this paper is organized as follows. Section 2 gives introduction to AdaBoost-based face detection algorithm and existing hardware for detection. Section 3 describes the proposed architecture in detail, and Section 4 goes to evaluate the performance of the architecture. Section 5 gives the conclusions.

* Corresponding author at: Room 409, College of Information Engineering Building, Haigang Avenue 1550, Pudong, Shanghai 201306, China. Tel.: +86 02138282800.

E-mail address: wnzhou@shmtu.edu.cn (W. Zhou).

2. Related work

This section firstly describes the face detection based on AdaBoost learner using haar-like features, and then introduces the existing hardware for face detection.

2.1. Face detection based on AdaBoost

The AdaBoost-based face detection algorithm is considered as one of the most efficient object detection algorithms, which can achieve remarkably better performance in terms of detection accuracy and speed. Moreover, it is suitable for hardware implementation, since it is simple and regular in computation, and is extended in architecture.

The AdaBoost-based face detection algorithm is distinguished by three contributions, that is, AdaBoost algorithm, integral image and cascade structure.

The AdaBoost-based face detection algorithm utilizes AdaBoost as a learning algorithm to select a number of visual features for producing efficient and accurate classifiers. The most popular features used in AdaBoost algorithm are the “haar-like” features. They are fixed-size images which contain a small number of black and white rectangles. Edge characteristics or line characteristics in an image can be detected by them. The computation of a haar-like feature involves calculating the sum of the pixel values in the white rectangles of the feature minus the sum of pixel values in the black rectangles.

To speed up the feature computation, “integral image” is proposed to alternatively represent the input image. In integral image, each pixel location holds the sum of all the pixels in the region between the origin and that location. Fig. 2.1 shows the three kinds of haar-like features, the definition of integral image and the rectangle calculation with integral image values. As shown in Fig. 2.1, the integral image value in position p is the sum of all the pixel values in the grid region. Thus, the sum of rectangle D could simply be gained by one addition and two subtractions of the four corner points of the rectangle.

To achieve accurate classification, adequate selected haar-like features are always adopted to constitute a strong classifier to detect faces from the image. The strong classifier can be divided into several “stages”, which are arranged in a cascade structure to improve the detection efficiency. In a cascade structure, the first stage is always simple and is constituted by a few haar-like features, and stages afterwards are generally slightly more complex

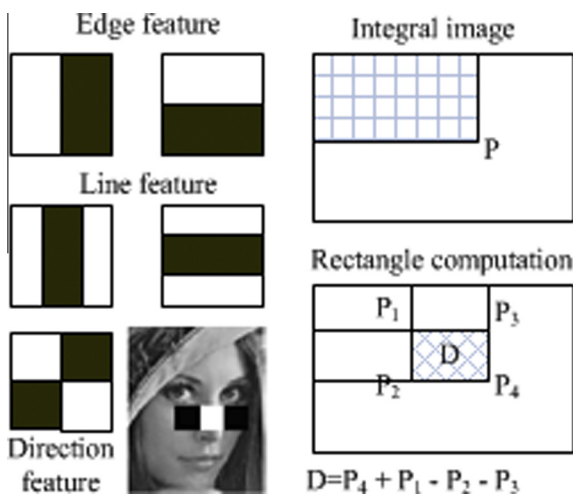


Fig. 2.1. Basic concepts in the AdaBoost-based face detection algorithm.

than the last one. Because a majority of background usually can be excluded by the judgment of first several stages and is not needed for further feature calculating, the computing is mainly concentrated on object regions reducing the total computation amount greatly.

In the original AdaBoost-based face detection algorithm [7], the detection uses features starting at 24×24 pixels, and faces larger than this size are detected by enlarging the features by a factor. However, when the feature size becomes larger, data access becomes sparser, this may cause a cache miss or require large cache memory for a specialized processor to achieve fast memory access. Therefore, it is a better choice to downscale the original image to detect faces larger than 24×24 in hardware implementation, which has been successfully used in some hardware [13].

To compensate the light variations, the variance (Var) is used to rectify the feature threshold (t_0) given in the training set for face judgment [12]. The compensated threshold (t) could be obtained by Eq. (2.1), which dynamically takes care of any lighting variations encountered during the detection stage and improves the overall accuracy. And the variance (Var) in Eq. (2.1) could be calculated by Eq. (2.2).

$$t^2 = t_0^2 \cdot Var \quad (2.1)$$

$$Var = \left(\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} p_{ij}^2 / (M \cdot N) \right) - \left(\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} p_{ij} / (M \cdot N) \right)^2 \quad (2.2)$$

In Eq. (2.2), p_{ij} is the gray value of the pixel in the position (i, j) . M, N are width and height of the search window, both are 24 in the design. Therefore, $\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} p_{ij}$ and $\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} p_{ij}^2$ can be directly obtained from integral image and squared integral image (each pixel location holds the square sum of all the pixels to the left and above of that location) respectively, since they are just the values in the position $(M-1, N-1)$ according to the definition of integral image and squared integral image.

The flow chart of the detection scheme is shown in Fig. 2.2, and the detection procedure can be described as follows:

1. Obtain the pixels of the search window.
2. Compute the integral image and square integral image of the window.
3. Calculate the variance of this window for threshold adjustment in the classification of following stages.
4. Do the feature classification with cascade structure. The classification will terminate once it fails in any stage of the cascade structure. If the window passes all the stages, the position and the size of the search window will be reserved.
5. Judge whether all the sub-windows has been classified. If yes, go to step 6. If no, repeat step 1–5 for classification of the next neighboring window.
6. Downscale the image and repeat step 1–5.

2.2. Existing hardware for face detection

Although the AdaBoost-based face detection scheme has faster and better performance than many other schemes, it still requires too much computational cost for embedded use. As mentioned in previous section, two kinds of architectures have been proposed.

The architectures proposed by Chen, Hori and Hanai are the typical examples of the first kind. Chen et al. [8] came up with a 0.64 mm^2 real time face detection design with low consumption in power, but its detection accuracy declines to 81.57%. It reduced the resolution of the input image to 160×120 and only the highest 4 bits of the pixel value were stored in memory. The architecture

Download English Version:

<https://daneshyari.com/en/article/461386>

Download Persian Version:

<https://daneshyari.com/article/461386>

[Daneshyari.com](https://daneshyari.com)