Joint Conference of 1996 International Computer Symposium
December 19~21, Kaohsiung, Taiwan, R.O.C.

Face Detection Based on Edge and Intensity Information

Chun-Ying Chu and Shu-Yuan Chen @
Department of Computer Engineering and Science
Yuan-Ze Institute of Technology, Taiwan, R. O. C.
135 Far-East Rd., Nei-Li, Chung-Li
Taoyuan, 32026, Taiwan, R.O.C.

Abstract

This paper presents a new approach to locate human faces in a complex background, even when the number, sizes, and locations of faces are unknown. Furthermore, the proposed approach avoids both the overhead of explicitly constructing a priori face model and the full scans of the image with all possible sizes and locations of human faces. The method is based on edge and intensity information and built on hypothesis generation-and-verification paradigm. First, an automatic segmentation method is proposed to extract all the darker pixels exact to eyes. Then any two eye blocks can be used to form a face region. On the other hand, the face sides and cheeks are extracted from edge information and used for hypothesis generation and verification. In addition to these type of constraints related with face organs, the spatial disposal of face organs and the spatial relations of face patterns are also used to find out authentic human faces. Experimental results prove the practicality and feasibility of the proposed approach.

1. Introduction

Human face represents one of the most common patterns in our vision. Furthermore, the face is a unique feature of human being, even the faces of "identical twins" differ in some respects. Therefore, the automatic recognition of human faces is a significant problem in many applications such as criminal identification and security checks. And the first important step of automatic human face recognition is to detect faces in a given unknown picture. However, the task of automatic face detection in a complex background is difficult to cope with. In this paper, we propose a new method to locate the positions of faces in a complex background.

In [1, 2] a model-based approach is proposed to locate human faces in newspaper photographs. In [3], the goal of the system is to extract both the shapes and the locations of eyes and mouth from a front-view ID-type picture. Since the primary regions of eyes, eyebrows, and mouth should appear as relatively darker objects than their respective surroundings. All valley pixels are enough to include those regions. While in [4, 5], they can locate human faces in a complex background. In [4], the system is a hierarchical knowledge-based and consists of three levels. And the full scans of the image with all possible sizes and locations of human faces are acquired to detect faces without a priori knowledge. In [5], an efficient region-based algorithm to locate human faces is proposed. Face cheeks are used to find the possible positions and sizes of human faces. However, space and time overhead is needed to construct human mask model in [5].

In this paper, we propose a new approach to locate human faces using edge and intensity information. The basic idea of our approach is that the gray levels of eyes are generally darker than their surroundings, like that idea used in [3]. However, our problem is significantly different from that in [3], since they focus on the image with simple background and only one face. While we must locate faces in complex background with various number of faces.

To achieve the goal mentioned above, we propose an automatic segmentation method to obtain a gray plane composed of darker pixels. Moreover, the gray plane must correspond to all the eyes as exact as possible. Then, the blocks enclosing the compact darker pixels in the gray plane can be extracted as the primary features, eye blocks. On the other hand, edge plane can be attained by edge detection. Then, the secondary features, face sides and cheeks, are extracted from the edge plane. After that, face regions will be constructed by conjugating any two eye blocks. And according to facial context, some impossible face regions can be eliminated, if no face sides or cheeks are included in them. By this way, hypothesized face candidates can be generated. Finally, edge detection can be applied again on each hypothesized face candidate locally to obtain more subtle secondary features. Based on these subtle features, the verification is performed to locate the indeed human faces.

Some restrictions are made in this study to simplify the problem. They are described as follows.

1. The images only include the front views of faces.
2. Faces should not be occluded, in particular the eyes.
3. The sizes of faces should be greater than an absolute minimum value, say 32*32, otherwise they are too smaller to be easily detected [6].
4. The faces must be up-right, with almost negligible tilt.
5. The tolerant rotation angle is from - 15° to + 15°.

2. Feature extraction

The first step in feature extraction is to transfer gray-scale images into binary gray-representations including all the darker pixels for eyes, called gray planes, and binary edge-representations including all the edge pixels for face organs, called edge planes. Then the primary and secondary features are extracted from the gray planes and edge planes, respectively.

2.1. Primary feature extraction

The primary features are the eye blocks with darker gray levels than their surroundings. In general, the first local valley in histogram can be the cutting point to obtain all the valley pixels as darker objects. However, only employing histogram cutting is hard to obtain a correct gray plane composed of all the eyes as exact as possible. The reason is that the collection of all the eye blocks in a complex background may be sometimes over and sometimes under the set of all the valley pixels. This problem can be overcome by adaptively combining the techniques of histogram cutting and local segmentation. In other words, instead of directly collecting all the valley pixels as gray plane, we regard them as the bound limit of gray plane.

On the other hand, the concept of local segmentation is stated as follows. Intuitively, the concept of extracting darker pixels of eyes is similar to that of document segmentation because the extracted character/graphic pixels are dark pixels in general. Thus, we employ one global-based document segmentation technique, Lloyd's method presented in [8], in our approach. However, eyes usually are relatively darker objects in their respective surroundings, hence local thresholding rather than global thresholding is preferred in this problem. According to the above reasons, we will modify Lloyd's method to have local-thresholding capability. The modified version of Lloyd's method is referred to as local-based Lloyd's method.

More specifically, the modification is described as follows. At first, original Lloyd's method is applied on an entire image to segment the image into two sets. One of them with the darker mean is chosen as the globally darker part for the image. Since eyes are darker objects, they should be included in the globally darker part. Then, original Lloyd's method is employed again on such globally darker part to yield locally darker pixels. Consequently, repeated invocations of Lloyd's method result in more coherent images with darker pixels. That means the histograms of the resulting darker parts approaching uni-mode step by step. However, eye blocks may disappear if iterations are employed infinitely. The reason is that eyes may not be globally dark patterns. Thus, some stop criterion must be defined to obtain good gray plane.

To solve this problem, we combine the techniques of histogram cutting and local-based Lloyd's method and rename it as the adaptive-combination segmentation method. And from experimental results, we figure out the following fact. That is if the result of adaptive-combination segmentation approaches the histogram cutting result as close as possible, then it will be the best gray plane, most exact to all the eyes. Thus the following rules are used to select the best gray plane. Assume that the result of histogram cutting is B, and the second and third iteration results of local-based Lloyd's method are C1 and C2, respectively. In addition, the numbers of pixels of C1, C2 and B are N1, N2 and Nb, respectively.

1. If N1>N2>=Nb, the image C2 is selected for gray plane.
2. If N1>Nb>N2, the image B is selected for gray plane.
3. If Nb>N1>N2, the image C1 is selected for gray plane.

An example is shown in Fig. 2 to illustrate the three different conditions of the selection rules. The gray planes obtained as shown in the first, second and third images of Fig. 1(f) are the corresponding C2, C1, and B images, respectively.

After that, the CRLA (constraint run length algorithm)[6] is used to smear the gray plane followed by the labeling method [9] to find the rectangular blocks enclosing each connected component. Then, the resulting bounding boxes can be considered possible eye blocks. However, the gray level of eyes are not the same, hence the components in the gray plane may be fragmentary. Thus, the modified CRLA will be employed instead of the CRLA. The basic idea of the modified version is as follows. Besides two darker pixels are close to each other, their gray levels must also be close to each

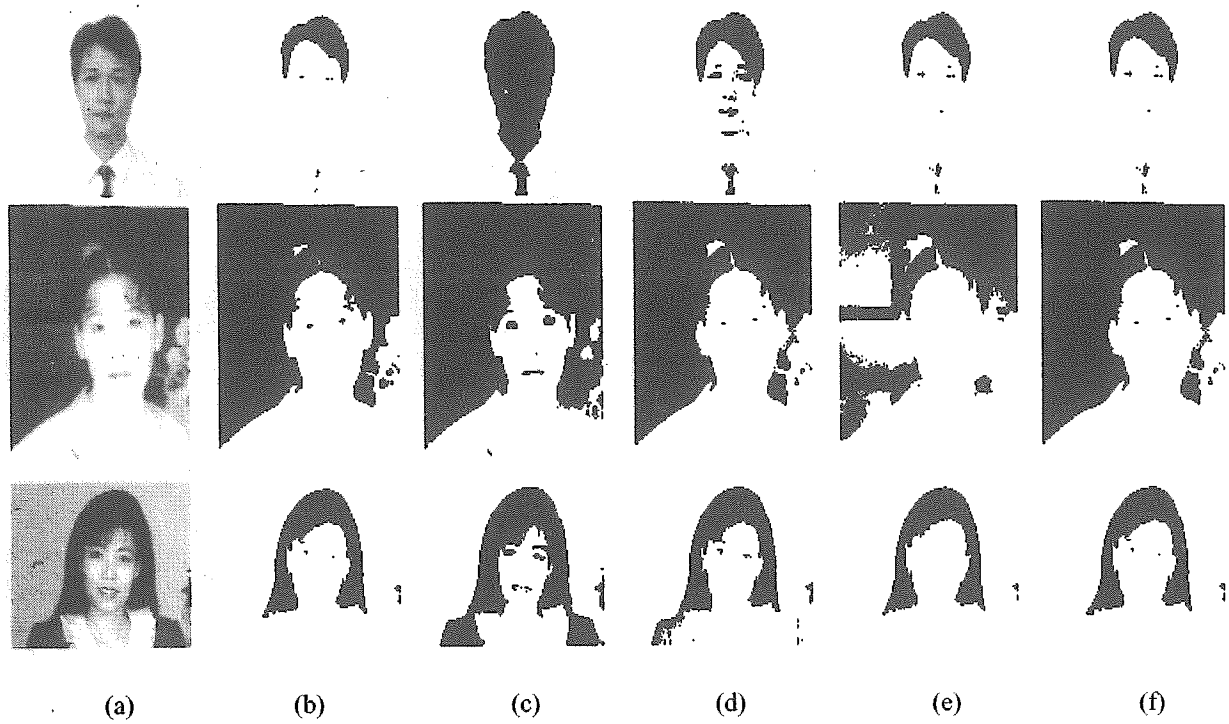


Fig. 1. Examples of the selection rules for gray planes.
(a) The source images; (b) the results of histogram cutting; (c), (d) and (e) the results of local-based Lloyd's method in the first, second, and third iterations, respectively; (f) the resulting gray planes.

other, then they can be merged.

2.2 Secondary feature extraction

The secondary features include face sides and face cheeks which are extracted from edge plane. Edge plane can be obtained by edge detection such as Sobel operator [9]. Then, face sides and cheeks can be extracted from the edge plane by the following ways.

(1) Face Sides

In general, most of the edges in the region of the left face side are in the directions larger than or equal to 90°. Similarly, edge directions in the region of the right face side are almost smaller than or equal to 90°. Thus, the distribution of edge directions corresponding to two face-side regions will have two specific forms, as shown in Figs. 2 and 3.

The distribution of edge directions in a region is defined as in [10] and can be obtained by the following way. Find out all the edge pixels covered by the region and compute their gradient-direction angles simultaneously. Then, the value of each gradient-direction angle is rounded to the nearest 22.5° and thus corresponds to one of the eight edge-direction categories equally spacing around 180° as depicted in Fig. 4. Furthermore, there are eight counters, one for each of the eight directions D_0 to D_7. These counters are used to record the number of pixels with the same directions. As a result, the distribution of edge directions in one region can be yielded.

After that, we will develop a method to measure how close the distribution of edge directions in a certain region meets the specific forms of face sides [11]. And this measure can be the criterion to detect face sides.

(2) Face Cheeks

In general, the cheeks of a face are almost white regions in edge plane.

3. Hypothesis generation

There are some characteristics in the human face patterns such as the attributes and spatial disposal of face organs. Thus, we can use these type of constraints to select the possible face candidates from all the face regions, each formed by a pair of eye blocks. In other words, we can eliminate some regions which are impossible to be human faces since they are not conformed to the properties of human faces.

In fact, the constraints included in this paper are the properties of face size, eye blocks, face sides, and face cheeks together with the facial context. Furthermore, under the assumption that human faces should not be occluded, some rules will be employed to eliminate overlapping situation.

First, the facial context implied by human face model is described in Section 3.1. Then the constraints on face organs and the rules for overlapping elimination are stated in Sections 3.2 and 3.3, respectively.

3.1 Human face model

As mentioned above, the spatial relations between face organs remain the same for most of faces. This property will be adopted to define the human face model [4] used in this paper. The model is shown in Figs. 5 through 7. Assume that in these figures the distance D is half of the distance between two eye blocks, and the basic rectangles are the quadrates with each side length as D. Then, a reasonable face region is a square with sides 4*D in both width and height. In addition, the reasonable face-side region is a rectangle with sides $3/2 * D$ in width and $5/2 * D$ in height, and its relative position with respect to the center of eye block is a specified constant, as shown in Fig. 6. Similarly, the reasonable size and relative position of each face-cheek region is specified as in Fig. 7.

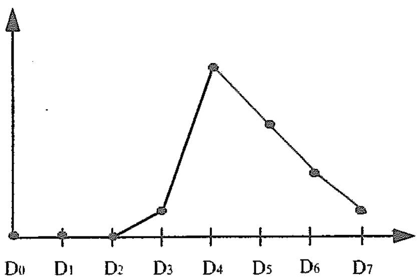


Fig. 3. Distribution of the edge directions in the region of left face side.

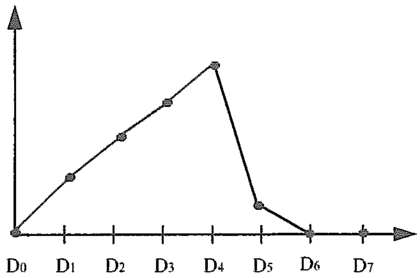


Fig. 4. Distribution of the edge directions in the region of right face side.

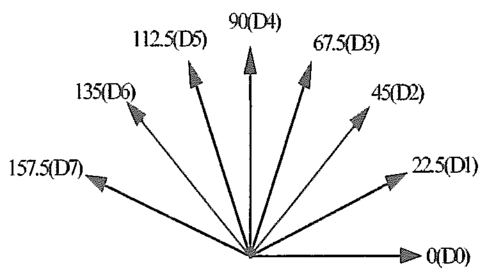


Fig. 4. Eight edge-direction categories.

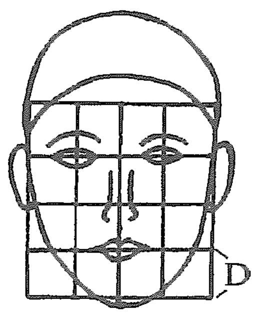


Fig. 5. The human face model.

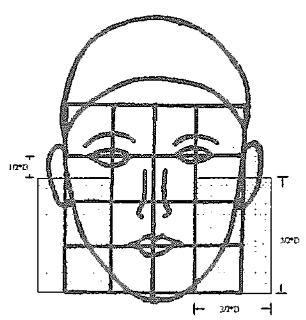


Fig. 6. The regions of face sides.

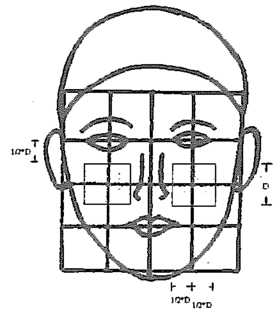


Fig. 7. The regions face cheeks.

3.2 Constraints for hypothesis generation

From the face model as shown in Figs. 5 through 7, some constraints are be derived to select possible face candidates. When patterns satisfy some of these constraints, they will obtain corresponding confidence values[11]. As a result, those with total amount of confidence values high enough will be regarded as hypothesized face candidates.

(1) Constraints on face size.

By the human face model as shown in Fig. 5, we can predict the sizes and locations of possible face candidates according to those of eye blocks. As shown in Fig. 5, the face size is 4*D x 4*D, thus some special values of D can be derived to be the constraints on face size.

When the faces are too small, they are hard to be detected. Usually, the meaningful images of human faces have the size of 32*32 at least [6]. Thus, we can obtain the first constraint that D can not be smaller than 8. The other constraint is that the boundary of face candidate can not be beyond the frontier of the image.

(2) Constraints on eye blocks.

Blockoverlap —

In general, the eyes of a human face don't overlap each other. Consequently, if the eye blocks overlap each other in one pair, this pair must be prevented from generating face candidates.

Area —

In general, the difference between areas of eyes in each pair should not be too large. Moreover, for each pair, the smaller the area of eye block the larger the allowable tolenrence between them. Thus, we will adaptively adjust the tolerance of area difference according to the size of eye block [11].

Angle —

Under the restriction that faces must be up-right with little amount of rotation angle, the direction of the line segment joining the two eye blocks is limited from -15 ° to +15° in this paper.

(3) Constraints on face sides.

Two face sides are the secondary features. When we find out a pair of darker blocks satisfying the properties of eyes, we want to search a pair of regions located at proper position and having specific form of edge-direction distribution as shown in Figs. 2 and 3. According to the human face model as shown in Fig. 6, the most possible face-side regions are located under eyes and near two sides of the face region formed by the pair of eyes. If two face sides actually exist, i.e. the above conditions are satisfied, the possibility of the pair of eye blocks being indeed eyes increases, and so does that of the face region being face candidate.

(4) Constraints on face cheeks.

Two face sides are the other secondary features. We wish to search two white regions in the surrounding of the pair of eye blocks as face cheeks. According to the human face model in Fig. 7, the most possible cheek regions are located under eyes. We don't strictly require that no edge pixel exists in the cheek regions. If we can find a pair of regions located at proper position with respect to eye blocks and not including too many pixels, then we can hypothesize one face candidate enclosing this pair of eye blocks.

3.3 Overlapping-candidate elimination

As mentioned above, the face region satisfying enough constraints can be considered a face candidate. Furthermore, each constraint will be given proper confidence value. Consequently, when two face candidates overlap each other, we can eliminate one of them according to the total amount of confidence values.

Let each face candidate have a flag with values "TRUE" or "FALSE" to indicate whether it can be survived when overlapping situation occurs. The basic idea of the selection criterion is that the flag of each candidate is retained the same until overlapping situation occurs. At that time, the one with the higher total amount of confidence values will enforce the lower one to be "FALSE". The details are included in [11].

4. Hypothesis verification

If the hypothesized candidates are really human faces, they should satisfy the characteristics and spatial disposal of face organs. On the other hand, the position and size of each hypothesized face can be determined, thus we can employ the edge detection again just on the hypothesized face candidate. Since this type of operation is restricted only to the local area, the resulting edge representation will be more accurate than that obtained from the entire image. Henceforth, the verification is achieved by checking the properties of the more subtle features. Moreover, after verification, the detected duplicate face patterns representing the same face should be reduced into one face only.

4.1 Constraints for hypothesis verification

In addition to the regions of eye and cheeks, the regions between two eyes are selected as the features for verification.

(1) Constraints on regions of eyes.

Because the gray values of eye blocks are different from their surrounding, the edge pixels are apparent in the boundary of eye blocks. Thus, there must exist some edge pixels around the two eye blocks.

(2) Constraints on regions between eyes.

There should be no edge pixels in the region between two eyes as shown in Fig. 5. However, some edge pixels related with nose may exist in this region. Thus some edge pixels detected in this region are allowed.

(3) Constraints on regions of cheeks.

From the human face model as shown in Fig. 7, it's impossible to contain edge pixels in the cheek regions of human face. However, a few edge pixels included in cheeks' regions are allowed. So we can allow a few edge pixels appearing in the regions.

4.2 Duplicate-face reduction

Sometimes, even the gray levels within one eye block are different. This fact results in broken eye blocks. In other words, one whole eye block may be divided into several parts. Therefore, after verification, there may exist duplicate face patterns representing the same human face. Nevertheless, such situation must be eliminated.

If two detected face patterns represent the same human face, their eye blocks should be close to each other. Hence, when two face patterns with the same total amount of confidence values overlap, we must check whether they are the same face. If it is true, the one with the smallest area will be chosen as the authentic human face pattern. Otherwise, spurious faces are included in the detection result. That is one factor leading to detection error in our experiment.

5. Experimental results

The proposed method is designed to locate human faces in a complex background. Test images are scanned from the AGFA StudioScan II Desktop scanner or taken by LOGITECH Footman Picture digital camera. Then, all the test images are processed by our method on 486-DX 66 personal computer. Fig. 8. illustrates the process of our method.

In this experiment, total 160 images containing various cases are tested. Table 5.1 summarize the experimental results of various cases.

6. Conclusions

In this paper, a novel approach based on edge and intensity information for face detection is proposed. Adaptive combination of the techniques of histogram cutting and local-based Lloyd's method leads the gray plane including all the eyes as exact as possible. Then, the gray plane is used to search the primary features, i.e. eye blocks, from which the face regions can be constructed. After that, hypothesis is generated and verified to locate human faces. All the hypothesis tests are based on the constraints related with the properties of face organs, the spatial disposal of face organs, and the spatial relation between face patterns.

Furthermore, our approach can process some special cases, such as wearing chapeau or beard. However, the case of wearing glasses can be handled only when the glasses are light-colored.

In summary, the proposed approach has the following merits:

1. Our method avoids the full scans of the image with all possible sizes and locations of human faces as acquired in [4]. Thus the processing time can be reduced in the proposed approach.

2. In our approach, human face model has been implicitly involved in hypothesis generation-and-verification rules, hence no space and time overhead is needed to construct human mask model as required in [5].

3. The proposed approach can detect more than one human faces in a complex background, unlike [3] focusing only one face in a simple background. Hence, the application fields of our method are diversity.

Further research may be directed to lifting the assumptions of our method, extending the proposed technique to locate human faces in color images and applying our approach to various human detection applications, etc.

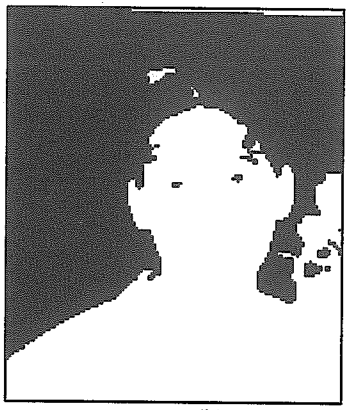
Table 5.1 Experimental result of 160 test image

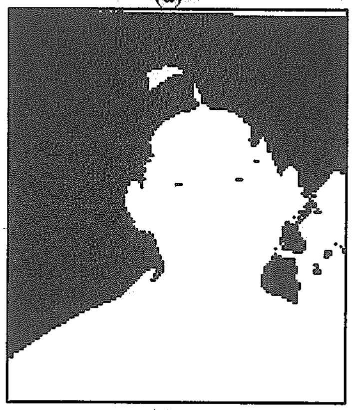
Cases	Number of test images	Average Extraction rate	Average number of spurious faces
One human face	60	0.92	0.07
Two human faces	35	0.87	0.14
Many human faces	5	0.85	0.40
No human face	15	—	1.13

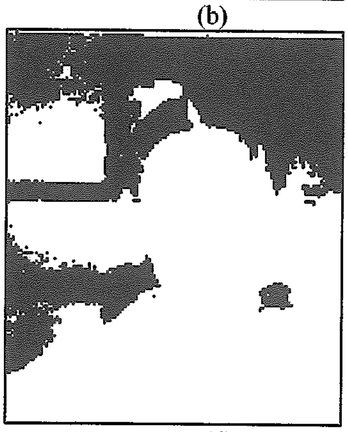
Special Cases	beard	15	0.65	0.33
	glasses	15	0.67	0.47
	chapeau	15	0.87	0.13
Total excluding special cases		115	0.89	0.24
Total		160	0.84	0.26

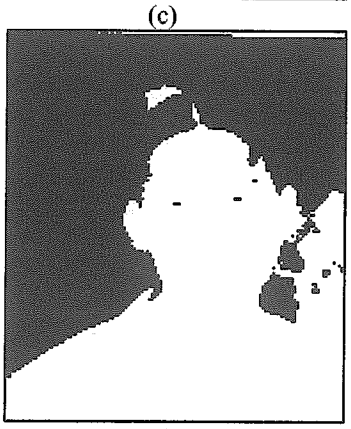
— : denotes none can be extracted.


(a)


(b)


(c)


(d)


(e)


(f)

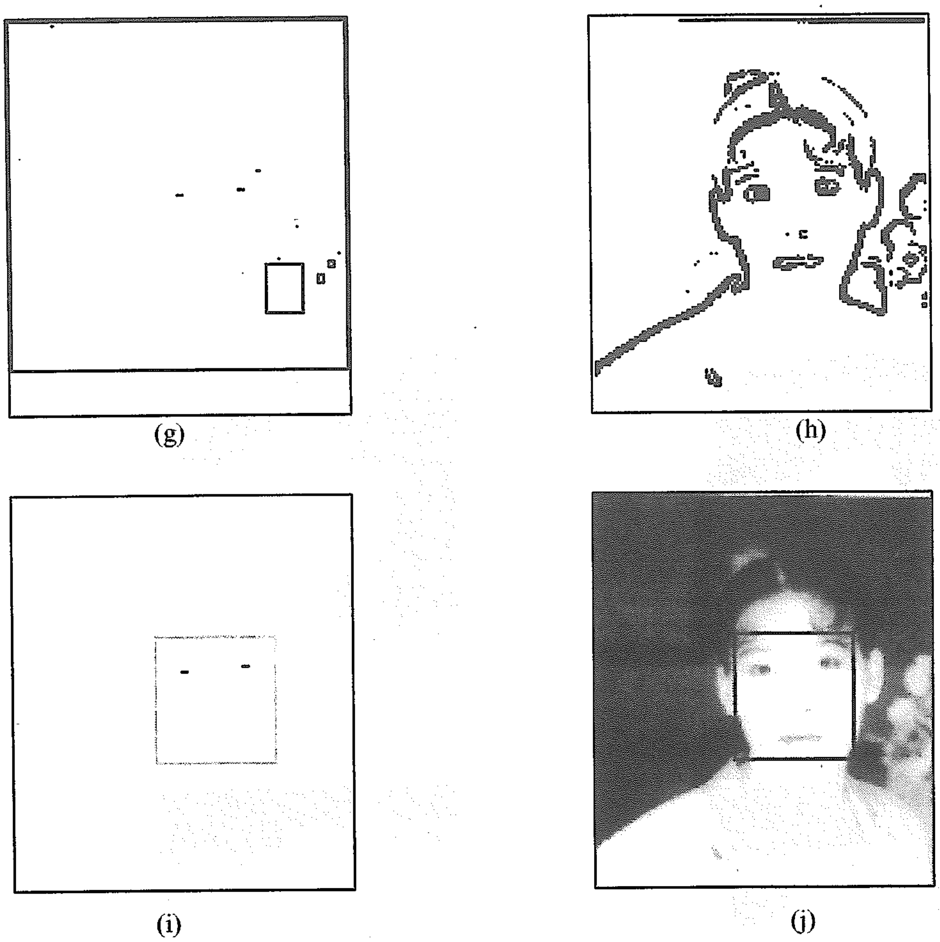


Fig. 8. The process of our method.
(a) The source image; (b) the result of histogram cutting; (c) and (d) the results of local-based Lloyd's method in the second and third iterations; (e) the resulting gray plane; (f) the result of smearing; (g) the result of block searching; (h) the resulting edge plane; (i) the face candidates; (j) the located human face.

References

[1] V. Govindaraju, et. al. "Locating human faces in newspaper photographs," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 549-554, 1989.

[2] V. Govindaraju, et. al. "A computational model for face location," in *Proc. Int. Conf. Computer Vision*, pp. 718-721, 1990.

[3] G. Chow and X. Li, "Towards a system for automatic facial feature detection," *Pattern Recognition*, vol. 26, pp. 1739-1755, 1993.

[4] G. Yang and T. S. Huang, "Human face detection in a complex background," *Pattern Recognition*, vol. 27, pp. 53-63, 1994.

[5] S. Y. Ho and D. H. Shi, "An efficient region-based algorithm for detecting human faces in a complex background," in *Proc. National Computer Symposium*, pp. 701-708, 1985.

[6] A. Samal and P. A. Iyengar, "Automatic recognition and analysis of human faces and facial expressions: a survey," *Pattern Recognition*, vol. 25, pp. 65-77, 1992.

[7] W. Y. Wong, R. G. Casey, and F. M. Wahl, "Document Analysis System," *IBM Journal of Research Development*, vol. 26, pp. 647-656, 1982.

[8] N. B. Venkateswarlu and R. D. Boyle, "New segmentation techniques for document image analysis," *Image and Vision Computing*, vol. 13, pp. 573-583, 1995.

[9] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 1992.

[10] H. Sun and W. Kwok, "Concealment of damaged block transformation coded images using projections onto convex sets," *IEEE Trans. Image Processing*, vol. 4, pp. 470-477, 1995.

[11] C. Y. Chu and S.Y. Chen, "Face detection based on edge and intensity information," *Master Thesis of the Institute of Electrical Engineering and Computer Engineering and Science*, Yuan Ze Institute of Technology, Taiwan, R.O.C. 1996.