Journal of Applied Sciences

ISSN 1812-5654

Journal of Applied Sciences 6 (4): 795-799, 2006
ISSN 1812-5654

Fast Template-based Face Detection Algorithm Using Quad-tree Template

Chang-Woo Park and Mignon Park
Precision Machinery Research Center, Korea Electronics Technology Institute,
401-402 B/D 193, Yakdae-Dong, Wonmi-Gu, Puchon-Si, Kyunggi-Do, 420-140, Korea

Abstract: Machine recognition of faces from still and moving images is emerging as an active research area spanning several research fields such as image processing, pattern recognition, computer vision and communication. In addition, human face detection, recognition and identification techniques have numerous applications such as human interface based systems and real-time video systems for surveillance and security, etc. In this study, we suggest an algorithm for face detection from 2D color images using the quad-tree deformable template matching method. The deformable template matching method is used to find the face area and locate facial features. In order to make the algorithm insensitive to luminance, we convert the conventional RGB coordinates into the normalized CIE coordinates. The face detection procedure can speed-up under limited conditions and we also modify the algorithm for fast detection using the quad-tree hierarchical template of the face. The quad-tree template can easily and quickly sizable and the comparison between the template and the image is weighted according to the deviation of color. The template comparison results with the plain texture template and the quad-tree template are shown in this study.

Key words: Face detection, template, quad-tree template, machine recognition

INTRODUCTION

Face recognition technique from still and motion images is an active research subject and related to various fields such as image processing, pattern recognition, computational vision, man-machine interface and machine intelligence. It also has a lot of practical and industrial applications. FRT(Face Recognition Technology) from still images or image sequences can be appled to recognize the faces in passports, credit cards, driver's licences, criminal's montages, surveillance camera system, visual identification system, etc.

Face recognition problem can be divided into 3 parts (Chellapa *et al*., 1995): detection of the facial areas from the input image, facial feature extraction and classification, recognition and identification of the face. This study is focused on face detection and facial feature extraction using deformable template matching technique (Ji *et al*., 1996).

The template-based face detection is an simple and useful method to detect face regions and locate facial features. It is easy to implement into a hardware system from the software algorithm then to make the system fast. However the face detection by template matching have difficulty in searching face regions of various sizes and directions under various luminance conditions though this technique has relatively high performance of face detection and facial feature extraction within limited variation of the facial images. Face region detection methods using spatio-temporal characteristics and features-obtained by the image processing such as edge images, motion energy, color PDF(probability density function) and maximum likelihood (Pentland, 1995) are efficient for face detection under various input conditions. So it is expected to be efficient and show good performance to combine the two methods.

Figure 1 shows the whole procedures of the face detection system. The inputs can be a image sequence or a still image and the candidates of the face-like region are selected through image processing and image feature extraction methods. From the candidates, the initial parameters of the face template are estimated. Then the locations of the face and its features are matched to one of the templates slightly deformed from the initial template.

This study focuses on the template matching technique-the right part of the procedure in Fig. 1. Ji *et al*. (1996) suggested the face detection algorithm using the face template with shape and texture. The template is made up of the WFM (Wire Frame Model) and the texture of the mean face and it can be deformed by affine transform and triangle texture rendering. The

Corresponding Author: Dr. Chang-Woo Park, Precision Machinery Research Center, Korea Electronics Technology Institute, 401-402 B/D 193, Yakdae-Dong, Wonmi-Gu, Puchon-Si, Kyunggi-Do, 420-140, Korea
Tel: 032-621-2854 Fax: 032-621-2866

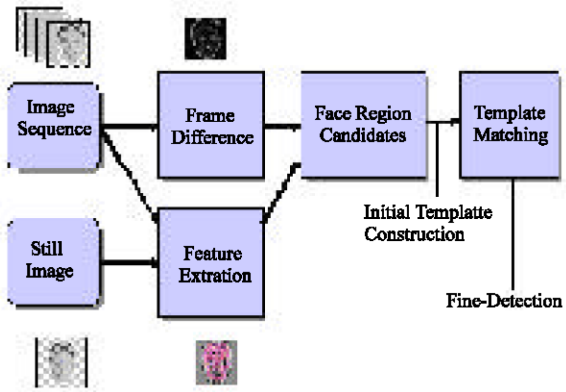


Fig. 1: The face detection procedure using image feature extraction and template matching

comparison and detection algorithm is simple and fast but the rendering takes relatively much computation time. In this study, the template-based detection algorithm is modified to have better performance and less computation time. Using the quad-tree style template, the rendering can be much faster and the comparison between the face region and the template is weighted according to the amount of color changes of the area.

COLOR NORMALIZATION

Figure 2 shows the template-based face detection procedures suggested in this study. The input is a color image of a face. The luminance of the image cannot be consistent and this luminance change cause serious mistakes of the template detection results. In order to make the algorithm insensitive to luminance change, the input must be normalized. The proposed algorithm uses the chromaticity of the normalized CIE component that is obtained from the RGB (Red Green Blue) component of the face image. The detection using the normalized CIE coordinates can remove influence of luminance change and it is expected to have robust identification performance.

The color information can be represented with three component. The RGB coordinate system has been widely applied as the color representation scheme. Another color coordinate system, CIE(X, Y, Z axis) has also three color components and one of them, Y represents the luminance. The input image in EGB coordinates can be converted into the CIE coordinates by applying the following matrix transform formula:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 2.7690 & 1.7518 & 1.1300 \\ 1.0000 & 4.5907 & 0.0601 \\ 0.0000 & 0.0565 & 5.5943 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad (1)$$

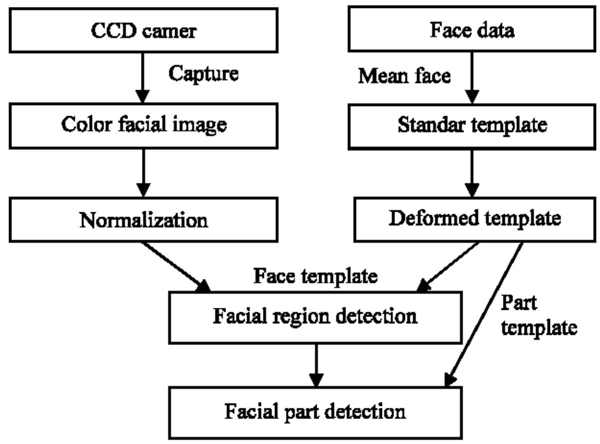


Fig. 2: The template-based face detection algorithm

In machine vision problems, optical conditions of the environment have much influence on the performance of the vision system. The luminance change is a major factor of the influence and one of commonly experienced problems. If we can eliminate the influence of luminance change, the identification performance is expected better. X, Y and Z values in the CIE coordinates can be normalized as follows:

$$\begin{aligned} &D = X + Y + Z \\ &x = X/D, y = Y/D, z = Z/D \end{aligned} \qquad (2,3)$$

Now the normalized values of x, y, z have only the chromaticity components of the pixel. Since x+y+z=1, the chromaticity components can be represented with only x and y values.

TEMPLATE-BASED FACE DETECTION

The face detection algorithm using the deformable face template with shape and texture (Ji *et al*., 1996) is very simple and easy to implement into hardware processing systems for fast detection. The template is made up of the WFM (Wire Frame Model) and the texture of the mean face. Figure 3 shows the standard WFM and template and template deformation. The WFM is matched with the mean face texture for some facial features such as eyes, eyebrows, nose, mouth, etc. The WFM is deformed by affine transform or scaling, rotation and translation and then the deformed template texture is constructed by the triangular texture rendering from the mean face. The mean facial image is obtained by adjusting 3-D facial shape model (WFM) on face images of 50 male (Aizawa *et al*., 1989) and averaging their vertices and textures.

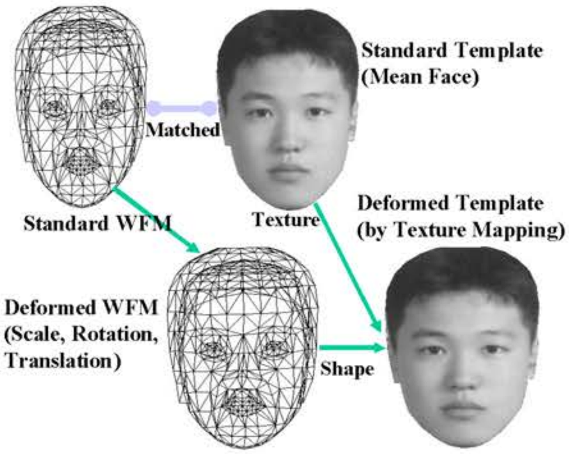


Fig. 3: WFM (Wire Frame Model) and deformation of the facial template

The constructed textures are used for face region detection as the face templates. Therefore, the color and luminance of the template can be adjusted by manipulating the initial texture or the mean face. Since the color and luminance of the template are changable, this method can overcome the various sizes and light conditions of the input image.

After the initial WFM is made up from image processing and feature extraction results, the face region including some facial features such as eyes, eyebrows, nose and mouth is detected using template deformation and template matching method. The detection method is as follows:

$$E(x, y) = \sum |I(x + a, y + b) - T(a, b)| \tag{4}$$

where E is the absolute sum of the pixel value difference between the template T and the image I for all pixels in the template region Rt. The template texture and the input are color-normalized images. The translated position (x+a, y+b) with minimum is selected the best fitting position where the face region is detected.

The best fitting template among the templates deformed from the initial one can be used as the base template for part detection. The templates for each facial parts such as eyes, eyebrows, nose and mouth are constructed using the template of the detected face region. That is, the eye parts of the face template directly make the eye templates, the nose part makes the nose templates and so on. The facial part templates can also be deformed by the affine transform of the WFM and texture rendering. The facial part detection is performed the same way as the face region detection.

The sizes and positions of the resulting WFM are including the detected face and facial parts location. That means the resulting template texture and WFM are matched for each vertices onto their position whithin the input face.

QUAD-TREE FACE TEMPLATE

The template matching algorithm is simple and fast but the texture rendering to construct deformed templates takes relatively much computation time. In this study, the template rendering algorithm is modified employing the quad-tree structure to have better performance and less computation time under some limited conditions.

We assume that the direction of the face in the input is little tilted so that the only parameters for face detection is the size and position of the face. In that case, the rendering procedure only requires scaling of rectangular images. The slight difference caused of direction changes can be detected in the facial part detection. Using the quad-tree style template, the procedure can be much faster and the comparison between the face region and the template is weighted according to the amount of color change of the area.

Figure 4 shows the quad-tree template. The quad-tree template starts from the parent node which has no parent and four child areas. The parent node covers the whole area of the template and each child node covers the one-fourth area of its parent. Each node has its node values which is the mean pixel values(color or luminance) of its area. The derivation of the children from one node depends on the standard deviation in the area. If the deviation is over a certain threshold, 4 new children are derived from the node.

The initial quad-tree template is constructed from the initial texture template. Once the initial quad-tree template is set, it can be easily and quickly deformed in various sizes and resolutions. When a 100×100 pixel image is stretched into a 150×150 pixel image, 22500 pixel value calculations for the new image are needed. If the quad-tree representation of the 100×100 image has a little color change and about 5000 nodes in the bottom nodes of the tree, it needs only 5000 pixel values which will be duplicated into the 150×150 image.

In the face region detection, the small features such as eye shape and color are not very important because the specific features can vary from person to person. That means that we can use relatively low-resolution templatefor the face region detection and high-resolution template for the specific part detection. Figure 4 shows that multi-resolution template construction from the quad-tree image.

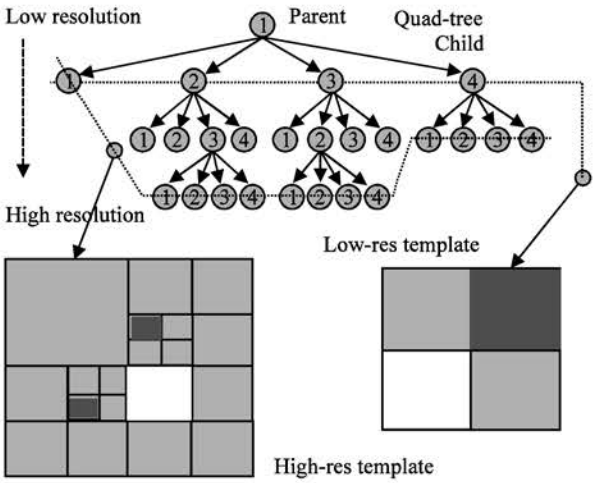


Fig. 4: Quad-tree representation of a 2D image and the multi-resolution face templates

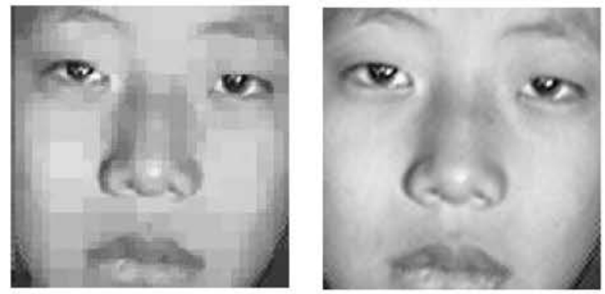
EXPERIMENTAL RESULTS

The input images for the face identification are captured by a color CCD camera and a video overlay board. The original images are 512×480×3 (512×480 pixels and 3 color elements in RGB coordinate system). In these images, 150×150 areas of the mean face including face features are selected as the initial face template. The normalized CIE needs x and y elements per pixel. Therefore the template is converted into a 150 X 150 X 2 image.

Figure 5 shows an example of the quad-tree template. The quad-tree image (a) is obtained from the original image (b) with the standard deviation threshold of 80. The eye parts are high-resolutional and the skin around cheek is low-resolutional.

Figure 6 is the fitness value (inverse of the error function) in template matching around the face image of the template itself. The fitness has maximum at (0, 0). (a) is the result using the conventional WFM rendered template and (b) is using the quad-tree template. The best fitting position is the top of the 3D graph and the value is rapidly decreasing in both cases.

Figure 7 is the fitness value in template matching around the different face image. (a) is the result using the conventional WFM rendered template and (b) is using the quad-tree template. The best fitting position is not as clear as of the above results because the specific parts of different people are not the same. Also the decay of the fitness is not as rapid in both cases. Figure 6 and 7 shows that the quad-tree template shows similar or a little better performance in face region detection. And the computation time of the quad-tree is much shorter than when using the WFM rendering template.



(a) (b)

Fig. 5: An example of (a) the quad-tree representation of (b) the original image. (up to std. dev. = 80)

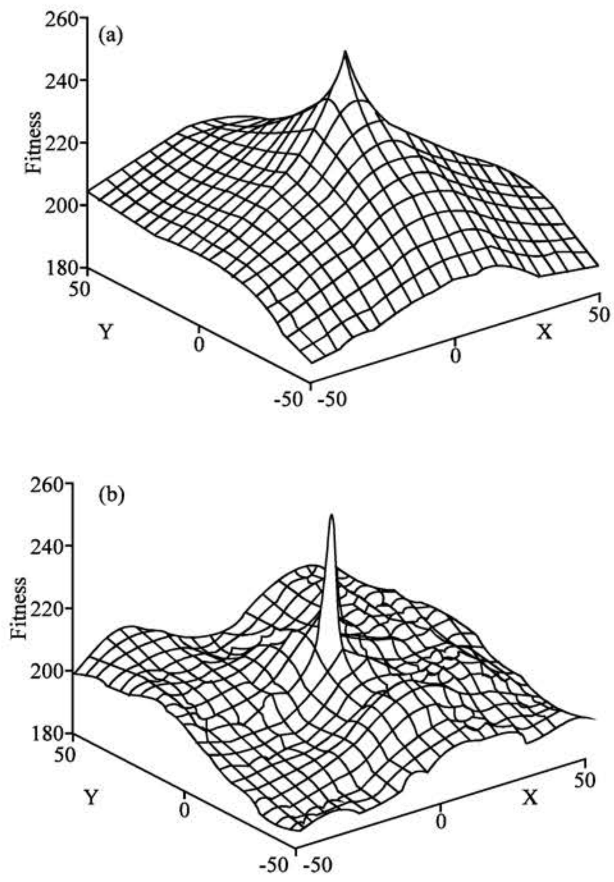


Fig. 6: Fitness values of the template matching on the template image itself (a) WFM rendered template (b) the quad-tree template

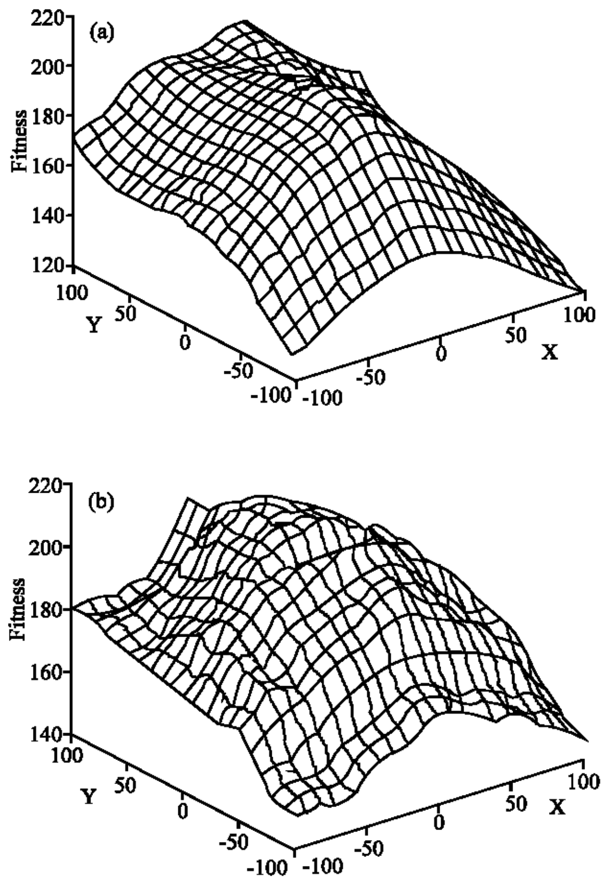


Fig. 7: Fitness values of the template matching on the different input image (a) WFM rendered template (b) the quad-tree template

CONCLUSIONS

This study suggested the face detection algorithm using the quad-tree template matching which is fast can this study suggested the face setection alogrithrn using the Qued-tree templete matching is fast can easily construct multi-resolutional template. The template matching detection is simple and efficient way when the face size and position are estimatable. The technique is also applied to the facial part detection. Since the texture template can easily influenced by luminance conditions, the input image of the detection and the template texture are converted into normalized-CIE coordinate system.

The fitness variation graph of the template matching procedure show that the template matching is efficient for face detection because it has peak at the fitting position and decay from the peak around. The result of the quad-tree template also shows similar and better performances.

The template matching algorithm is simple and easily implementable into practical systems. However the performance of the system can lower for various inputs such as direction and size changes, multi-face inputs, etc. So the preprocessing to obtain initial parameters of the template are required.

REFERENCES

Aizawa, K., H. Harashima and T. Satio, 1989. Model-based analysis synthesis image coding system for a person's face. Signal Process. Image Com., 1: 139-152.

Chellappa, C., W. Wilson and S. Sirohey, 1995. Prolog to human and machine recognition of faces: A Survey. Proceedings of The IEEE., 83: 704-740.

Ji, S., E. Kim, Y. Kim, M. Park and C. Choi, 1996. Automatic adjustment of 3D facial shape model using deformable template and fuzzy ellipsoidal model. Proc. of ITC-CSCC '96, Seoul, Korea, pp: 484-487.

Pentland, A., 1995. Machine understanding of human action. MIT Media Lab. Perceptual Computing Section Technical Report, No. 350, pp: 1-9.