PAPER *Special Section on Machine Vision Applications*

Precise Pupil Contour Detection Based on Minimizing the Energy of Pattern and Edge

Mayumi YUASA†, **Osamu YAMAGUCHI**†, *and* **Kazuhiro FUKUI**†, *Members*

SUMMARY We propose a new method to precisely detect pupil contours in face images. Pupil contour detection is necessary for various applications using face images. It is, however, difficult to detect pupils precisely because of their weak edges or lack of edges. The proposed method is based on minimizing the energy of pattern and edge. The basic idea of this method is that the energy, which consists of the pattern and the edge energy, has to be minimized. An efficient search method is also introduced to overcome the underlying problem of efficiency in energy minimization methods. "Guide patterns" are introduced for this purpose. Moreover, to detect pupils more precisely we use an ellipse model as pupil shape in this paper. Experimental results show the effectiveness of the proposed method.

key words: *pupil contour detection, pattern recognition, edge detection, energy minimization*

1. Introduction

Detection of feature points (eyes, nostrils, etc.) in face images is necessary for various applications using face images such as face recognition and gaze detection. Eye pupils are considered to be the most important feature among those facial feature points. It is difficult to detect pupils precisely in a face image. This difficulty is attributable to two factors: contours of pupils consist of weak edges and edges are often lacking along the contours.

Many pupil detection methods have been reported. Some are based on edge and shape information, for example, using deformable templates [1] or Hough transform [2]. A method using view-based pattern matching has also been reported [3]. However, there are problems respecting the robustness and accuracy of those methods. Recently, approaches have been proposed in which both appearance and shape are used. An active appearance model [4] is one of those approaches. An active appearance model, however, is often defined rather strictly, and therefore, it is inappropriate for pupils since their appearance changes dramatically depending on the individual, face direction and lighting conditions.

We propose a new method to precisely detect pupil contours in faces. The method is based on minimizing the energy that consists of the pattern energy and the edge energy. Edge detection and pattern matching have complementary characteristics. Table 1 shows features of those methods. Pattern matching is subject to a problem concerning accuracy and edge detection is subject to a problem concerning robustness. Therefore, if used in combination, pattern matching and edge detection can compensate for each other's deficiencies.

Table 1 Characteristics of edge detection and pattern matching.

Method	Edge detection	Pattern matching
Scale	Local	Global
Location accuracy	Good	Poor
Noise robustness	Poor	Good

In [5] a method to extract facial feature points based on a combination of pattern matching and shape extraction has been proposed. It is a robust and flexible system compared with the other methods mentioned above. But in [5] pattern matching and edge detection are used only separately, i.e., the edge detection was used only as the pre-processing for the pattern matching.

A more generalized method is proposed in this paper. It integrates pattern matching and edge detection in a unique framework of energy minimization and therefore it is a more robust and flexible method. An efficient search method using "guide patterns" is also introduced to overcome the optimization problem inherent in energy minimization methods.

If pupils are detected accurately, they can be used for face recognition for example. In view-based face recognition [6], the performance is highly dependent on the accuracy of the detection of face location and size. Accurate location is thus necessary for view-based face recognition. In gaze detection the problem is more serious. In view-based gaze detection (for example [7]), in addition to accurate locations of the pupils, shape information is also needed. The proposed method has a possibility of being used for gaze detection since it can handle an ellipse model as a pupil shape.

In Sect. 2, we present the basic concept of our method. The relationship between pattern and shape is discussed in Sect. 3. In Sect. 4, we introduce the concept of "guide patterns." In Sect. 5, the details of the pupil contour detection method with ellipse models are described. In Sect. 6, the experimental result and the discussion based on it are presented. Conclusions are

Manuscript received March 28, 2003.
Manuscript revised July 31, 2003.

†The authors are with Corporate Research and Development Center, Toshiba Corporation, Kawasaki-shi, 212–8582 Japan.

given in the final section, Sect. 7.

2. Basic Concept

We use the energy minimization scheme to integrate two types of information. Various energy minimization methods have been proposed for use in image recognition. For example, an active contour model [8] is one of the well-known methods. Employment of pattern energy, however, has not been discussed sufficiently so far. Figure 1 shows a conceptual diagram of the proposed method. The integrated energy is defined in Eq. (1).

$$E(\mathbf{X}) = \alpha E_p(\mathbf{X}) + \beta E_e(\mathbf{X}) \tag{1}$$

The integrated energy consists of the pattern energy $E_p(\mathbf{X})$ and the edge energy $E_e(\mathbf{X})$. α and β are weighting coefficients. $\mathbf{X}$ is a feature vector which describes the object shape and location. For example, if the object shape is described by the coordinates of its contour, the set of coordinates can be used for $\mathbf{X} = (x_1, x_2, \ldots, x_n)$. In the case of a parametric shape object, dimension of the feature vector can be reduced by using their parameters.

The pattern energy $E_p(\mathbf{X})$ is the similarity between a pattern normalized by shape $\mathbf{X}$ and a reference pattern normalized by a correct shape $\mathbf{X}_c$ (Fig. 2).

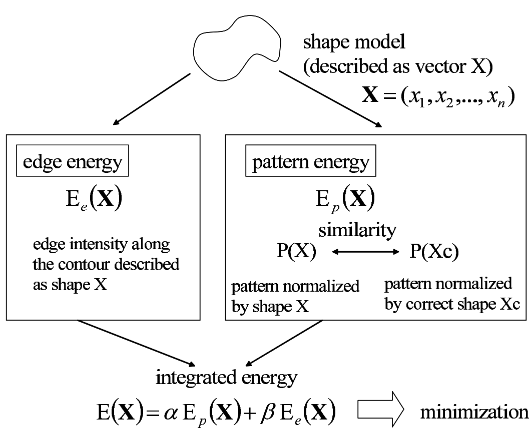


Fig. 1 Conceptual diagram of the proposed method.

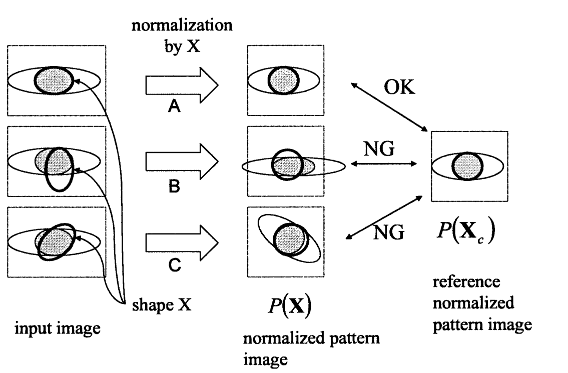


Fig. 2 Examples of the images normalized by shape model.

We discuss the relationship between pattern and object shape in Sect. 3. The edge energy $E_e(\mathbf{X})$ is based on the edge intensity along the contour of shape $\mathbf{X}$. The energy originated in edge is subject to a problem in that the energy is incorrect when the position is far from the correct position. The pattern energy does not dramatically change as the edge energy does. Examples of the pattern and the edge energy maps are shown in Fig. 3. The edge energy has an acute peak, which is important for determining the shape precisely.

3. Relationship between Pattern and Shape

In this section, we briefly summarize the relationship between pattern and shape of contour.

The contour shape of an object is closely related to the pattern through a normalization process. Suppose, for instance, that an ellipse shape is normalized to be a circle shape (A in Fig. 2). If the contour is detected correctly, the normalized pattern is similar to that of reference patterns which are normalized by the contour shape. On the contrary, if the contour is not detected correctly (B and C in Fig. 2), the normalized pattern is dissimilar to that of reference patterns. Therefore, the pattern similarity between an input pattern and the reference pattern is able to describe the accuracy of detected contours. In other words, the similarity of patterns changes as the normalization parameter (i.e. object shape and location) changes. Therefore, the normalization parameter can be derived from the change of the pattern.

There are other approaches to deal with the pattern, i.e. the appearance of the object parametrically. For example, the method using parametric eigenspaces [9] is proposed. It uses the direction or the pose of the object as the parameters for patterns. We use the object contour shape as parameters instead of them.

4. Guide Patterns

There is always a search problem in energy minimization. If the initial point is far from the correct one, the energy minimization does not work correctly. To overcome this problem, we introduce "guide patterns," which plays a role in navigating to the correct point.

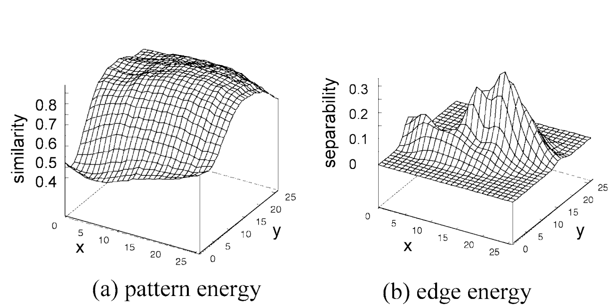


Fig. 3 Examples of the pattern and the edge energy maps (x and y are some of the parameters of $\mathbf{X}$).

Figure 4 shows examples of the guide patterns along the path from the initial position to the correct position in 2-dimensional space. The guide pattern $P(\mathbf{X}_c+\mathbf{dX})$ is a set of pattern images that are normalized by slightly shifted shape ($\mathbf{X}=\mathbf{X}_c+\mathbf{dX}$)).

Figure 5 shows examples of pupil image patterns when the shape of a pupil is described as a circle. The guide pattern can suggest the most likely direction in energy minimization procedure, because the shift parameters are known in advance. The direction is the opposite of the shift parameter used in normalization. If $Sim(\mathbf{X},\mathbf{X}_c+\mathbf{dX})$, which is the similarity between a pattern at $\mathbf{X}$ and a pattern at $\mathbf{X}_c+\mathbf{dX}$, is greater than $Sim(\mathbf{X},\mathbf{X}_c)$, one can guess that $\mathbf{X}$ is near $\mathbf{X}_c+\mathbf{dX}$. Therefore, to get close to the correct point $\mathbf{X}_c$, it has to be translated in $-\mathbf{dX}$ direction.

This idea is similar to the parametric eigenspace method for navigating a robot to the target position. Maeda et al. [10] adopted the strategy of estimating the current position of the robot by learning the images along the path to the target.

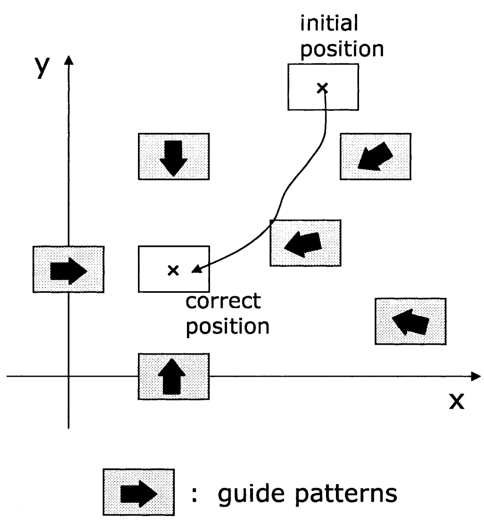


Fig. 4 Guide patterns along the path to the correct position in 2-D. (Guide patterns are described as boxes with arrows.)

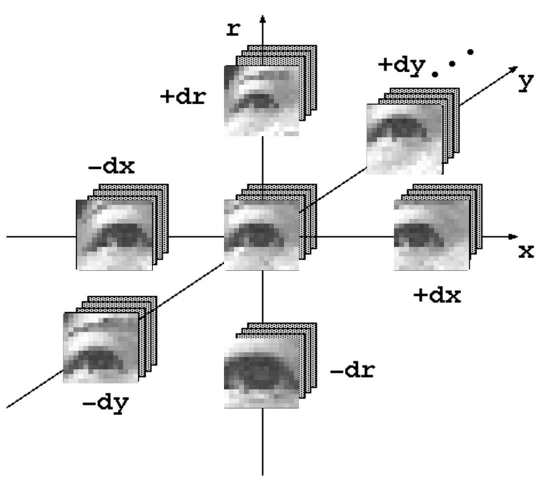


Fig. 5 Examples of the images normalized with shifted parameters.

5. Pupil Contour Detection by the Proposed Method

5.1 Elliptical Shape Model and the Definition of Energy

In this section we describe the pupil contour detection method based on the energy minimization. The contour of a pupil is approximated by an ellipse model. For the generalized feature vector, $\mathbf{X}$, we use five ellipse parameters that are the location of the center point, (x,y), major and minor axes, a and b, and the rotational angle, θ.

$$\mathbf{X}=(x,y,a,b,\theta) \tag{2}$$

To calculate the similarity for pattern energy $E_p(\mathbf{X})$, we use the subspace method [12]. The subspace method, which can absorb pattern deformation, is more robust than the simple correlation method. Therefore, it can be used for pupils that change their appearances depending on the individual or lighting conditions. The correct pattern subspace is made from the images normalized by correct ellipse parameters. The elliptical shape is normalized to be a circular shape (Fig. 2). The feature vector used for the subspace method is created by raster scan of the normalized image. When the normalized image is $n\times n$, the feature vector has $n\times n$ dimensions. In subspace method, the feature vector can be expressed as linear subspaces spanned by multiple eigenvectors. The bases of the subspace are calculated by principle component analysis. Figure 6 illustrates the subspace method. The similarity is defined by the angle between the subspace D and the input vector g. The squared cosine of the angle is defined as

$$\cos^2\theta=\frac{1}{||g||}\sum_{n=1}^{N}(g,\phi_n)^2 \tag{3}$$

where ϕ_n is the bases of D and N is the dimension of the subspace.

For the edge energy $E_e(\mathbf{X})$, we use separability [5]. Separability is based on linear discriminant analysis†. It represents the degree of difference between two regions and is used to detect edges or contours. Separability is defined by η in Eq. (4).

$$\eta=\frac{{\sigma_b}^2}{{\sigma_T}^2} \tag{4}$$

$${\sigma_b}^2=n_1(\overline{P_1}-\overline{P_m})^2+n_2(\overline{P_2}-\overline{P_m})^2$$

†Linear discriminant analysis is used for various applications. The threshold selection method known as Otsu's method [11] is the example.

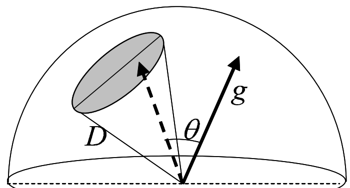


Fig. 6 Subspace method. (The similarity is defined by the angle θ between the subspace D and the input vector g.)

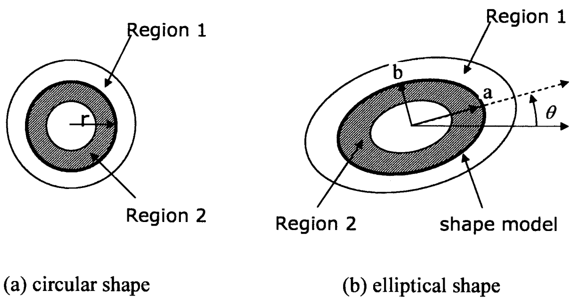


Fig. 7 Separability filters.

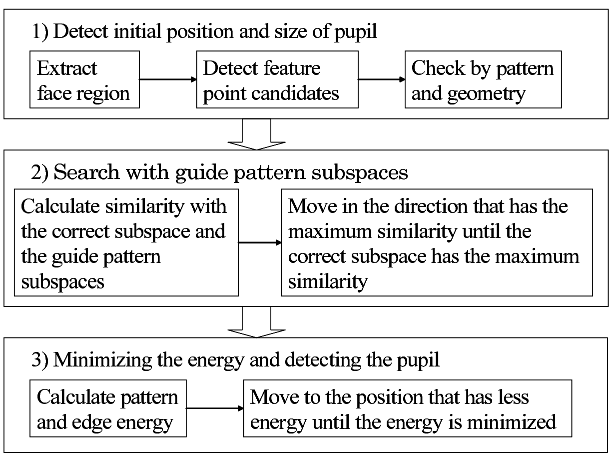


Fig. 8 Framework of the proposed pupil contour detection method.

$$\sigma_T{}^2 = \sum_{i=1}^{N} \left(P_i - \overline{P_m}\right)^2$$

where P_i is an image intensity at pixel i, $\overline{P_1}$, $\overline{P_2}$, $\overline{P_m}$ are the mean values of the image intensity in region 1, 2, 1+2, and n_1, n_2, N are the numbers of pixels in each region. Figure 7 (a) shows an example of the circular separability filter [5]. We use an ellipse separability filter to calculate the edge intensity. Figure 7 (b) shows an example of the ellipse separability filter. The shape of the filter represents an elliptical model.

The framework of the proposed method is shown in Fig. 8. The details of each step are as follows.

5.2 Detection of Initial Location and Size

The facial parts detection method proposed in [5] is used to detect the initial location and the size of the pupils as a circle.

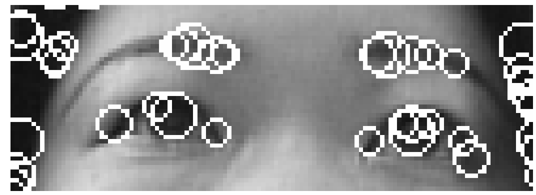


Fig. 9 Examples of the feature point candidates.

First, the face region is detected in an input image. We use a face detection method based on pattern matching using the subspace method [12].

And then candidate feature points are detected by separability filters with circular shape (Fig. 7 (a)). For each pixel in the detected face region, the separability is calculated while changing the position and radius of the filter. The feature point candidates are selected among the local maximum peaks of separability value. Figure 9 shows the examples of the detected feature point candidates. Finally the candidate points are verified by pattern matching [5].

5.3 Search with Guide Pattern Subspaces

In order to avoid detecting wrong contours and to reduce processing time, we use "guide pattern search," which is carried out according to the following procedure.

"Guide pattern subspaces" are used instead of simple guide patterns. The guide pattern subspaces are generated from the images normalized by slightly shifted known parameters that describe an elliptical shape.

First, the similarity is calculated with the correct pattern subspace and the guide pattern subspaces which are created in advance. The correct pattern subspace is made from the images normalized by correct parameters. The guide pattern subspaces, on the other hand, are made from the images normalized by the parameters at a known distance from the correct one. We use two guide pattern subspaces for each ellipse parameter (Fig. 5). For instance, $+dx$ and $-dx$ are created for the parameter x.

The position is translated in the opposite direction to that of the shift parameter that has the maximum similarity. The search is continued until the similarity with the correct pattern subspace becomes greater than that with any guide pattern subspace.

For example, if the similarity between an input image and the subspace cropped at the location of $+dx$ from the correct location is the largest among the guide subspaces, one can guess that the input image is close to $+dx$ from the correct location. Therefore, the position has to be translated in $-dx$ direction.

5.4 Minimizing the Energy and Detecting the Pupil Contours

Pupil contours are detected by minimizing the total energy which consists of pattern energy and edge energy (Eq. (1)). The weighting coefficients α and β are discussed in Sect. 6.3. The pattern energy $E_p(\mathbf{X})$ and the edge energy $E_e(\mathbf{X})$ are described in Eqs. (5) and (6).

$$E_e(\mathbf{X}) = -Sep(\mathbf{X}) \tag{5}$$

$$E_p(\mathbf{X}) = -Sim(\mathbf{X}, \mathbf{X}_c) \tag{6}$$

$Sep(\mathbf{X})$ is the separability value at $\mathbf{X}$, which was described as η in Eq. (4). $Sim(\mathbf{X}, \mathbf{X}_c)$ is the similarity between an input pattern and the correct pattern subspace. Negative signs are added in order to incorporate them into the energy minimization scheme.

The position is moved step by step to the location that has less energy from the initial position. This process is continued until the current position has the minimum energy in the neighboring area.

6. Experimental Result

We show the experimental result obtained by the above-mentioned method. Figure 10 shows examples of correctly detected pupil contours. The proposed method was applied to real face images. Evaluation was performed for six people. We used 20 images for each person. Both left and right pupil images are used for the evaluation. The correct pupil contours used for evaluation were extracted manually. We have used 1600 sample images of eight people to generate both the correct pattern subspace and the guide pattern subspaces. The data of the test people are not contained in the above sample images. 225 dimensions for the normalized image size of 15×15 pixels are used in the experiment. The method of normalization is described in Sect. 3. Figure 5 shows the examples of the normalized images.

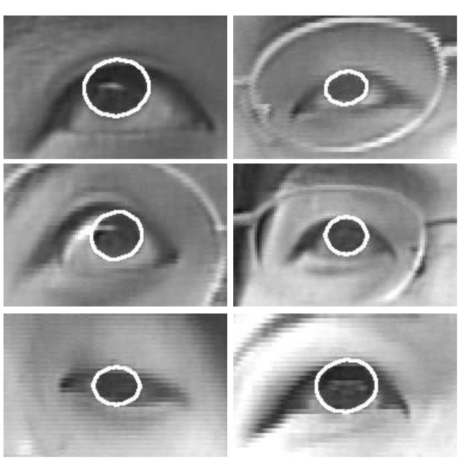


Fig. 10 Examples of detected pupil contours.

The results show that almost all pupil contours were detected correctly.

The experimental system consists of a personal computer (Intel Pentium III 933 MHz), a video camera and an image capture board (Matrox Meteor-II). The process works in approximately real time from capturing to displaying the result for one person in an image of 640×480 size.

The estimation of contour accuracy was performed based on the difference of the area inside the detected contour from that of the correct contour. The difference of area includes both inner and outer region of the correct contour. The area difference is normalized by the area in the correct contour in order to be invariant to changes in the pupil size. Figure 12 shows the definition of "normalized area difference (NAD)." This value is appropriate in the present case, because it includes not only the information on the difference in location but also on that in shape. For example, when the shape of the object is circle, if the detected contour has the same size and shape, and only the center point differs by half the radius, the normalized area difference is about 63%. If the threshold t for NAD is decided by applications, whether the detection is successful or not is able to be decided. For example, when the threshold t for NAD is 30%, which corresponds to about 25% of the radius in location of the center point, the detection rate in this experiment is about 98%.

Figure 11 shows the examples of incorrectly detected pupil contours. Some causes of failure are considered. One is that an initial parameter is far from a correct one. Other causes are the influence of the face direction and strong edge of eyelashes.

In Sect. 6.1, a comparison with the previous method is presented. Then we discuss the features of the proposed method in more detail. Dependency

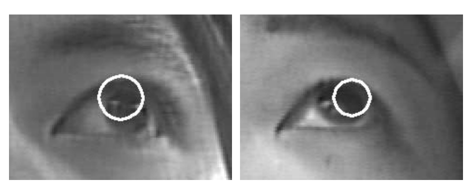


Fig. 11 Examples of incorrectly detected pupil contours.

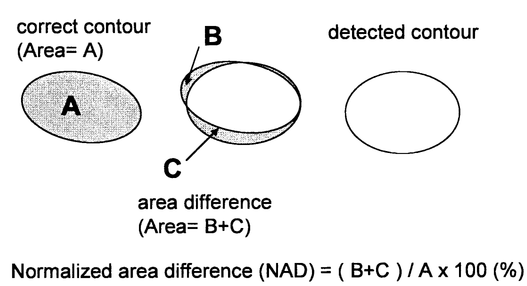


Fig. 12 Definition of normalized area difference (NAD).

on initial parameters (Sect. 6.2), the weighting coefficients (Sect. 6.3), and the effect of guide pattern search (Sect. 6.4) are considered.

6.1 Comparison with the Previous Method

The experimental results compared with the results of the previous method reported in [5] are presented.

The estimation using the mean value of the normalized area difference for six people as mentioned above is shown in Fig. 13. The error of detected contours is much smaller than that in the case of the previous method. The errors are almost half of those in the case of the previous method. The error of center point is also used for evaluation. In this case the normalization is carried out using the radius that has the same area with the correct contour of elliptical shape. The evaluation by the normalized errors of center points is also shown in Fig. 14. This graph also shows that the proposed method is more effective than the previous method.

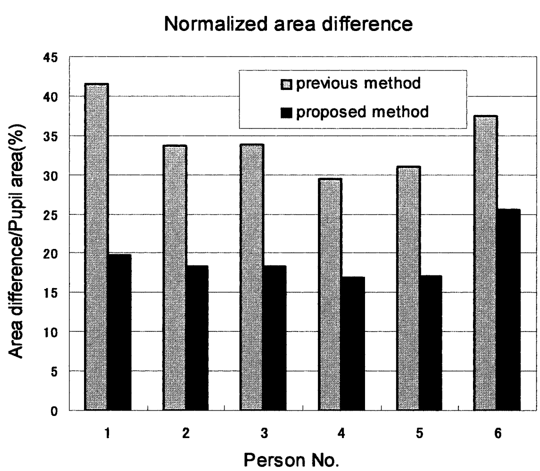


Fig. 13 Estimated normalized area differences from correct contour.

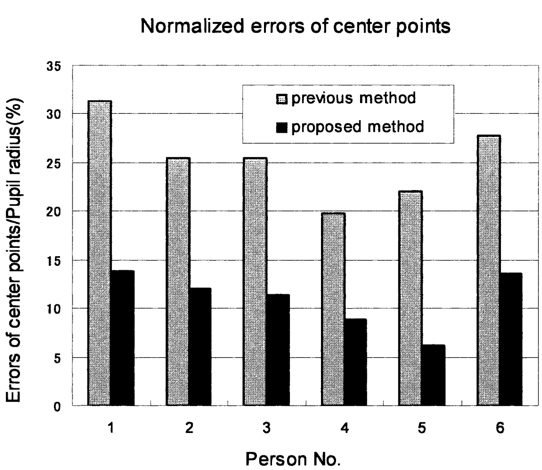
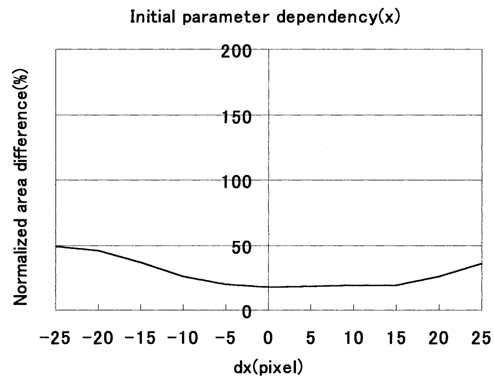


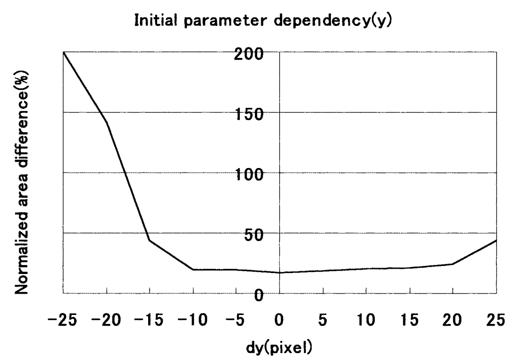
Fig. 14 Estimated normalized errors of center points from correct contour.

6.2 Dependency on Initial Parameters

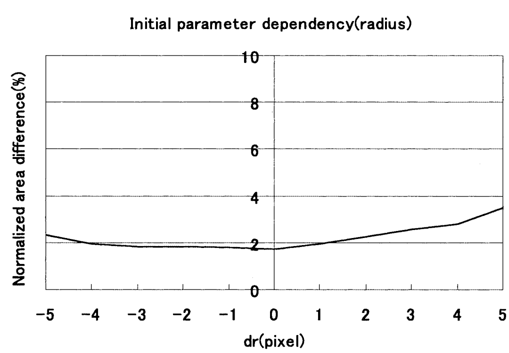
The evaluation is performed using only the mean values of six people in Sect. 6.2 to 6.4. Figure 15 shows the initial parameter dependency evaluated by normalized area differences. To evaluate initial parameter dependency, we set initial parameters artificially instead of using the process described in Sect. 5.2. For Fig. 15 (a), (b) and (c) the initial parameters are set at dx, dy and dr from the parameters of the correct contour, re-



(a)



(b)



(c)

Fig. 15 Initial parameter dependency evaluated by normalized area differences. Initial parameters for (a), (b) and (c) are dx, dy and dr from correct parameters, respectively.

spectively, where x means the horizontal location, y is the vertical one and r represents the size of the pupil. These results show that the detection accuracy does not largely depend on the initial radius, which represents the size of the pupil. For the horizontal direction represented by x, the allowed range of the initial parameter is approximately within twice the pupil radius†. (In this experiment, the radius of the pupils is about 10 pixels.) On the other hand, for the vertical direction, as the initial position moves in the upper direction, the error rapidly becomes large. This means that the eyebrows are at the upper position near the pupils. The eyebrows are similar to the pupils when they are normalized to small images. Therefore if the initial position is near the eyebrow, it sometimes falls in the false local minimum of the energy. Such an initial position, however, is not detected at the pre-processing stage described in Sect. 5.2, because the eyebrows are learned as error patterns corresponding to pupil patterns [5].

6.3 Weighting Coefficients

The weighting coefficients α and β in Eq. (1) are decided experimentally. Figure 16 is the graph of the normalized area differences when the ratio of the weighting coefficients α and β changes. The value is normalized as the sum of α and β becomes 1. In this experiment, the initial parameters are the same as those in Sect. 6.1. The best performance occurs when α is around 0.5. When α is 0, using the edge energy only, or 1, using the pattern energy only, the performance deteriorates. It shows the effectiveness of using both edge energy and pattern energy. We have decided to use $\alpha = 0.5$ and $\beta = 0.5$ from this experiment.

6.4 Effect of Guide Patterns

We also describe the effect of introducing guide patterns. Table 2 shows the mean computational cost, which corresponds to the processing time in one step of the guide pattern search (Sect. 5.3). The processing time in one step of the energy minimization (in Sect. 5.4) is about 10 times as long as that of the guide pattern search. The processing time for pattern matching using subspace method is almost the same in regard to guide pattern search and energy minimization when the number of reference subspaces is the same. However, the processing time for making normalized images differs, because energy minimization needs $n+1$ cropping procedure for searching n-neighbors for each step (we use 10-neighbors), whereas guide pattern search needs only one cropping for each step.

In this experiment, the initial parameters are the same as those in Sect. 6.1. This result shows that both accuracy and efficiency have been improved by using guide patterns.

The evaluation using various initial parameters is also performed. Figure 17 shows the initial parameter dependency of normalized area difference in horizontal direction††. For this evaluation we use only horizontal direction because the error in the initial position tends to be in horizontal direction as shown in Fig. 9. The computational cost is also shown in Fig. 18. These results show that both accuracy and efficiency have been much more improved by using guide patterns when the initial parameter is far from the correct one.

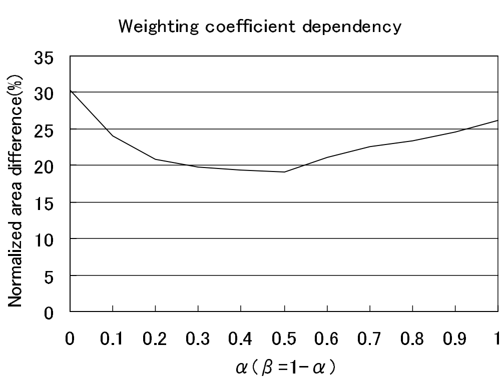


Fig. 16 Weighting coefficient dependency (evaluated by normalized area difference).

Table 2 Computational cost for with and without guide pattern search.

With/without guide pattern search		with	without
Normalized area difference (%)		19.1	20.0
Mean number of steps	guide pattern search (s1)	1.26	0
	energy minimization (s2)	4.68	4.99
Computational cost (s1+s2×10)		48.0	49.9

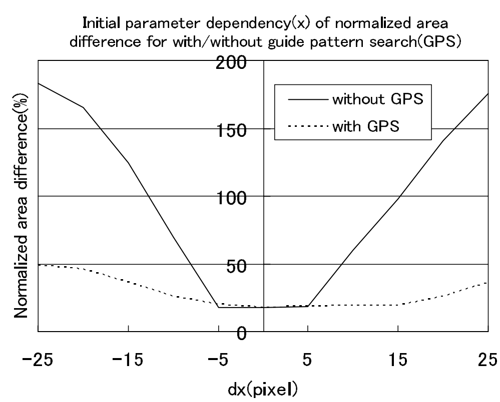


Fig. 17 Initial parameter dependency(x) of normalized area difference for with/without guide pattern search.

†Slight asymmetry on either side originates in the deviation of the image data, such as lighting conditions and face directions, used in the experiment.

††The initial parameter is artificially changed. So the result is not equal to that in Table 2.

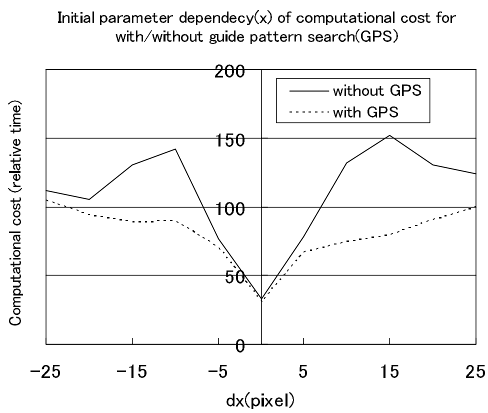


Fig. 18 Initial parameter dependency(x) of computational cost for with/without guide pattern search.

7. Conclusion

We have proposed a precise pupil detection method based on minimizing the energy of pattern and edge. An ellipse model is introduced for precise detection. We have also proposed the guide pattern subspace that helps to minimize the energy smoothly and stably. The method has been verified through experimental results.

In future work we intend to apply this algorithm to a large number of samples and to extend it to treat other facial parts such as lip contours.

References

[1] A.L. Yuille, P.W. Hallinan, and D.S. Cohen, "Feature extraction from faces using deformable templates," Int. J. Comput. Vis., vol.8, no.2, pp.99–111, 1992.
[2] G. Chow and X. Li, "Towards a system for automatic facial feature detection," Pattern Recognition, vol.26, no.12, pp.1739–1755, 1993.
[3] A. Pentland, B. Moghaddam, and T. Starner, "View-based and modular eigenspaces for face recognition," 1994 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '94), pp.84–91, June 1994.
[4] T.F. Cootes, G.J. Edwards, and C.J. Taylor, "Active appearance models," Proc. 5th European Conf. on Computer Vision, vol.2, pp.484–498, 1998.
[5] K. Fukui and O. Yamaguchi, "Facial feature point extraction method based on combination of shape extraction and pattern matching," Systems and Computers in Japan, vol.29, no.6, pp.49–58, 1998.
[6] O. Yamaguchi, K. Fukui, and K. Maeda, "Face recognition system using temporal image sequence," Proc. 3rd IEEE Int. Conf. on Automatic Face and Gesture Recognition, pp.318–323, April 1998.
[7] S. Baluja and D.A. Pomerleau, "Non-intrusive gaze tracking using artificial neural networks," Advances in Neural Information Processing Systems (NIPS) 6, 1994.
[8] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," Int. J. Comput. Vis., vol.1, no.4, pp.321–331, 1988.
[9] H. Murase and S.K. Nayar, "Visual learning and recognition of 3-D objects from appearance," Int. J. Comput. Vis., vol.1, no.14, pp.5–24, 1985.
[10] S. Maeda, Y. Kuno, and Y. Shirai, "Active navigation vision based on eigenspace analysis," Proc. 1997 Int. Conf. on Intelligent Robots and Systems (IROS '97), pp.1018–1023, Sept. 1997.
[11] N. Otsu, "A threshold selection method from gray-level histograms," IEEE Trans. Syst., Man. Cybern., vol.SMC-9, no.1, pp.62–66, 1979.
[12] E. Oja, Subspace methods of pattern recognition, Research Studies Press, England, 1983.

Mayumi Yuasa received her B.S. and M.S. degrees from Osaka University, in 1990 and 1992, respectively. In 1992, she joined Toshiba Corporation. She currently works at Multimedia Laboratory, Toshiba Corporate Research and Development Center.

Osamu Yamaguchi received his B.E. and M.E. degrees from Okayama University, in 1992 and 1994, respectively. In 1994, he joined Toshiba Corporation. He is currently a research scientist at Multimedia Laboratory, Toshiba Corporate Research and Development Center. He is a member of IPSJ.

Kazuhiro Fukui received his B.E. and M.E. degrees from Kyushu University, in 1986 and 1988, respectively. In 1988, he joined Toshiba Corporation. He is currently a senior research scientist at Multimedia Laboratory, Toshiba Corporate Research and Development Center.