Decoration of Human Face in Video Image by Estimating Eye Positions with Particle Filter

Gefan Zhang Norikazu Ikoma Hideaki Kawano Hiroshi Maeda
Faculty of Engineering, Kyushu Institute of Technology, Kitakyushu, Fukuoka, Japan

***Abstract*— Decorating human face in video image by estimating eye positions with particle filter has been proposed. Estimation of eye positions is presented in a sequential manner by formulating a face motion in a state space model. The model consists of simplified two-dimensional face with two dimensional motion and rotation. The positions of the two eyes in an image frame play a key role in estimating the face posture. Real-time estimation of eye positions has been realized by particle filter algorithm with likelihood computation consisting of adequacy of face and eyes. Eye decoration works as a result for enlarging the areas of eyes by estimating pupil positions. Experiments with different conditions of face, such as static scene, rotation, parallel motion, and some other motions have conducted to demonstrate performance of the proposed method.**

I. Introduction

Thanks to the great development of visual technology, image processing in dynamic videos has become possible. In the earlier studies, image processing was mainly focused on the enhancement of colors in static images, such as in [3][4]. But recently advances of face recognition, and digital image processing are gradually used for decorating the objects in the images such as in [5][6][7][8]. With image processing technology, many digital products and applications about face decoration have debuted as [9][10].

Decoration of object in video image, as a novelty, is an interesting challenge for amusement, and is expected to be used in some applications. For instance, face decorations in video image by handy camera, real-time face decoration in smart phone, and be saved in video files. With such motivations, real-time face decoration in video image is chosen as our aim in this research.

Composition of this paper is as follows. First, eye positions need to be estimated in dynamic sequences based on a state-space model. Focusing on the eye positions in the static image, human face is approximately simplified as a plane in this research. Eye positions are used to form the state-space model. For the estimation of them, particle filter algorithm is applied in the model. Smooth change of the state over time is described by system equation of this random walk. For the calculation of likelihood, we focus on color of the pupils, in addition, face position which is got from face detection, is taken into consideration. Second, human face can be decorated according to the result of estimating eye positions. We mainly want to achieve three purposes: enlarge the eyes, make the face smaller and make the skin whiter. In this paper, only eye decoration has been implemented. From the result of estimating eye positions, eye decoration processing can be carried out.

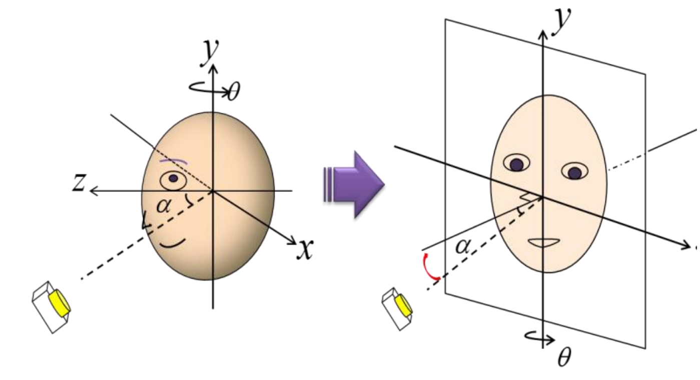
Experimental validation of our proposed method is based on estimation of eye positions, a real-time processing system is realized by using OpenCV and C language in different conditions of face, such as still, turning left and right, parallel motion, fore and after motion.

II. Model

Face posture and motion have been modeled first, relation between eye positions and their image will be described in a set of formulae. Based on the fact which has been written in the formulae, we construct a state-space model to represent its motion, and likelihood computation is also used in the state estimation by particle filter.

A. Simplification Face Model

Face model has been constructed at first in order to estimate the eye positions. Faces vary from each other, both face and head have their three-dimensional shapes at the same time. The three-dimensional face is treated approximately as a plate in this research. In various parts of the face, we focus on the two eyes. Therefore, there must be two eyes on the plate and representing the three-dimensional face which is shown in Fig.1. Three parameters θ, x, y have been used to facilitate our study which represent the angle of face rotation, the distance of parallel motions to two directions in the plate respectively. Angle between optical axis of the camera and normal line of the flat face α is fixed.



(a) Reality face (b) Flat face

Fig. 1 Simplified model of face

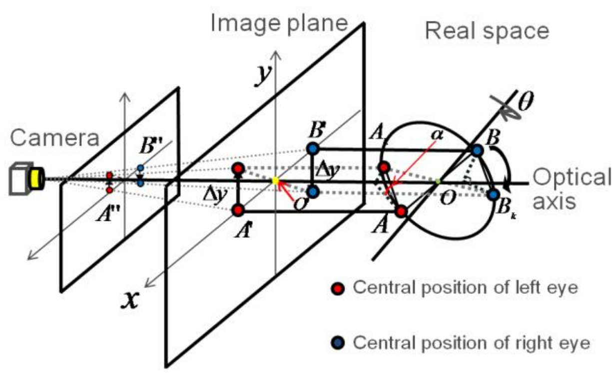


Fig. 3 Eye positions and their image

B. Eye Positions and Their Image

The two eyes play an important role in the estimation process. We calculate the positions of the two eyes and the region of face as Fig.2 shows. We choose the rectangle to represent the adequate model of each eye.

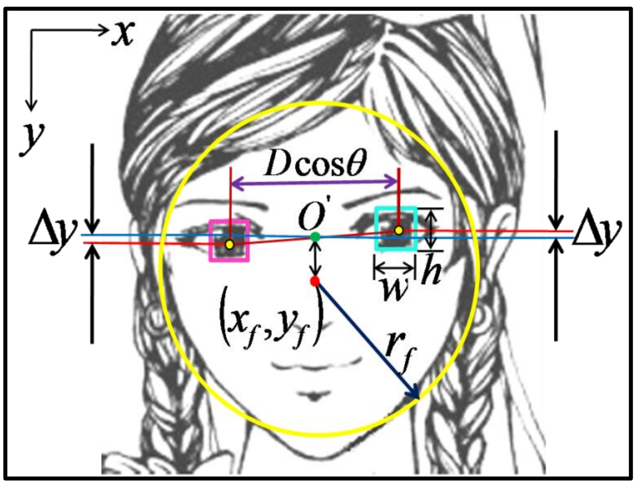


Fig. 2 Features on a face

Distance d_k between centers of the two eyes changes in two-dimensional image along with the rotation of face θ_k as time evolves. So that we have

$$d_k = D \cos \theta_k \quad (1)$$

where D is the distance between centers of the two eyes estimated in initial phase.

Vertical offset of the two eyes has been conducted from Fig.3.

It is shown in

$$\Delta y_k = \frac{d_k}{2} \sin \theta_k \sin \alpha \,. \quad (2)$$

Positions of the left and the right eye become as below:

$${}^l x_k = x_k - d_k/2, {}^l y_k = y_k - \Delta y_k \,, \quad (3)$$

$${}^r x_k = x_k + d_k/2, {}^r y_k = y_k + \Delta y_k \,. \quad (4)$$

C. State-Space Model

We construct a state-space model to represent face motion in a system model and likelihood computation for particle filter in an observation model. System model specifies time evolution of state variable $\boldsymbol{x}_k$, over time as

$$\boldsymbol{x}_k \sim f(\boldsymbol{x}_k | \boldsymbol{x}_{k-1}) \,. \quad (5)$$

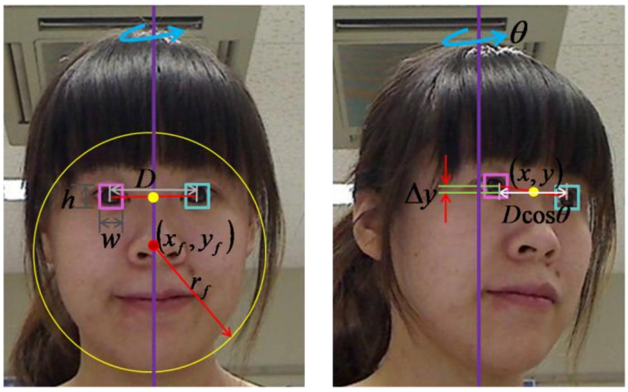
The observation model $\boldsymbol{I}_k$ specifies the process to obtain the observed image as

$$\boldsymbol{I}_k \sim h(\boldsymbol{I}_k | \boldsymbol{x}_k) \,. \quad (6)$$

1 System Model

System model has been used to describe the law of over time varying of the state vector. Random walk model is used when there are no rapid and substantial motions of the face. This is under the premise that the values of the state in current time are nearly the same as which in previous time. Both in the initial and tracking phases, these state variables and random walk models are used in the same way.

Fixed parameters are estimated in the initial phase of image processing. Later, during the tracking phase over-time changes of face posture are estimated. Therefore, state variables in these two phases are set as shown in Fig.4.



(a) Initial phase (b) Tracking phase

Fig. 4 Two phases of state estimation

For the initial phase, totally *m* frames were acquired under a prerequisite that movements between adjacent frames are small enough. Then after face detection in these *m* frames five parameters of the initial phase, position of the face *x* and *y*, height and width of the rectangle which represents pupil area *h* and *w*, distance between centers of the two eyes *D*, are calculated respectively, and selected as initial state variable $\boldsymbol{x}_k = (x_k, y_k, w_k, h_k, d_k)$.

$$x_k = x_{k-1} + v^x, \quad v^x \sim N(0, \tau_x^2)\,, \tag{7}$$

$$y_k = y_{k-1} + v^y, \quad v^y \sim N(0, \tau_y^2)\,, \tag{8}$$

$$w_k = w_{k-1} + v^w, \quad v^w \sim N(0, \tau_w^2)\,, \tag{9}$$

$$h_k = h_{k-1} + v^h, \quad v^h \sim N(0, \tau_h^2)\,, \tag{10}$$

$$d_k = d_{k-1} + v^d, \quad v^d \sim N(0, \tau_d^2)\,. \tag{11}$$

For the tracking phase, parameters of parallel movements *x* and *y*, the angle of face rotation, are taken as the state variable $\boldsymbol{x}_k = (x_k, y_k, \theta_k)$.

Smooth changes of system equations are

$$x_k = x_{k-1} + v^x, \quad v^x \sim N(0, \tau_x^2)\,, \tag{12}$$

$$y_k = y_{k-1} + v^y, \quad v^y \sim N(0, \tau_y^2)\,, \tag{13}$$

$$\theta_k = \theta_{k-1} + v^\theta, \quad v^\theta \sim N(0, \tau_\theta^2)\,. \tag{14}$$

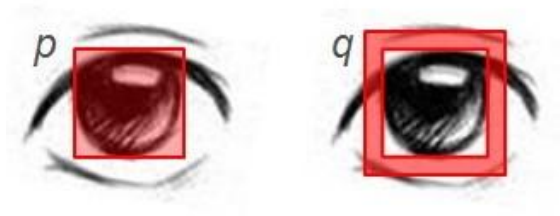
which can be written in conditional probability form $f(\boldsymbol{x}_k | \boldsymbol{x}_{k-1})$ as the system model.

2 Observation Model

The observation model plays the role in determining calculation methods of likelihood in image processing. For the state-space model, likelihood is the combination of eye likelihood and face likelihood

$$h(\boldsymbol{I}_k | \boldsymbol{x}_k) = \text{e\_ likelihood} \times \text{f\_likelihood}\,. \tag{15}$$

Likelihood of eye positions is calculated based on the color of the pupil in this research. The area around the pupil is light color, while the pupil area is dark color. Pupil color of general Asians is typically black or brown (dark brown). Based on this fact, calculation of likelihood in the plane is as follows: The shape of pupils is nearly round. However, rectangular region is used to approximately represent the pupil for simplifying the calculation of image processing which is shown in Fig.5. Dark pixels ratio in the rectangle is represented by p which is shown in Fig.5 (a), and dark pixel ratio in the area around the rectangle is q which is shown in Fig.5 (b).



(a) Rectangle for pupil (b) Area around the pupil

Fig. 5 Likelihood computation of eye

Using ratios of dark pixels in and around pupil areas (rectangle) *p* and *q*, the likelihood calculation of eye is computed in (15) for the state vector $\boldsymbol{x}_k = (x_k, y_k, \theta_k)$

$$e_{likelihood} = p(\boldsymbol{x}_k, \boldsymbol{I}_k) \times \{1 - q(\boldsymbol{x}_k, \boldsymbol{I}_k)\}\,. \tag{16}$$

Likelihood of face position is based on distance between two centers of the faces with its Gaussian distribution has deformed as

$$e_k^{(i)} = \sqrt{\left(x_k^{(i)} - x_k^f\right)^2 + \left(y_k^{(i)} - y_k^f\right)^2}\,, \tag{17}$$

$$f\_likelihood = \frac{1}{\sqrt{2\pi\tau^2}} exp\left\{\frac{-\left[e_k^{(i)}\right]^2}{2\tau^2}\right\}. \tag{18}$$

D. *Face Detection and Initialization*

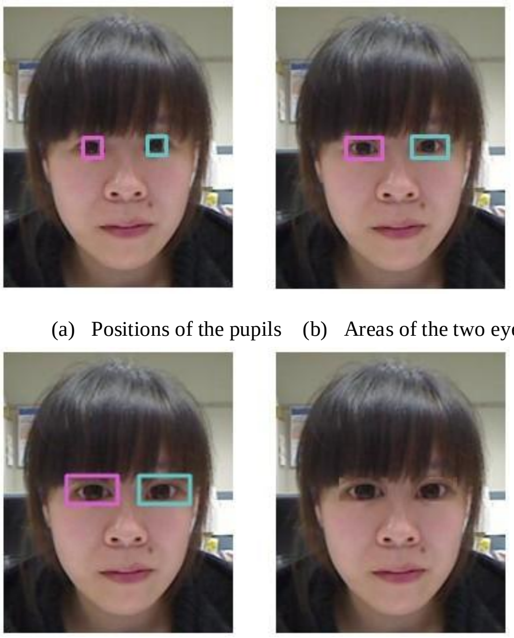
Face detector has been used to supplement the estimation of eye positions. We have used Haar-like cascade detector in [12] as the face detector. The result of face detection includes false and missing detection, so that it returns multiple candidates of the face, and we need to take into account these situations for the use of the face detector.

Taken into account the various errors such as false detection, missing, and multiple detections, we use a method to make the face detector a more robust one by aggregating the results obtained for several image frames, instead of using a single image frame result. Each pair of faces in all the samples, which were detected from these m frames, is selected to be compared. Firstly, for one pair, a similarity is set as half of the smaller radius between the two faces. Second, if the distance between centers of the two faces is smaller than the standard value, they could be seen as a match pair. With the same method we compared every pair of faces. After the comparison, we can get the most possible face which matches the most times with others. The parameters *x*, *y*, *r* from all the faces, which match with the most possible face have been added up respectively. Then the summation values are divided by the amount of matching faces and taken as the parameters of adequacy face.

Face detector has been also used for initializing the tracker with particle filter. If situations such as false or missing detection, happen in tracking phase, face detector needs to redetect the adequacy face which is used as initial value for estimating eye positions.

E. Eye Decoration

The process of face decoration can be done with the result of estimating face posture. Face decoration is the result of image processing obtained from eye positions estimation in Fig.6 (a). As the eye decoration, it means to calculate area of the eye from position of each pupil in Fig.6 (b), and then enlarge them which are shown in Fig.6(c), (d).



(a) Positions of the pupils (b) Areas of the two eyes

(c) Enlargement of the eyes (d) Result of eye decoration

Fig. 6 Processes of eye decoration

III. ESTIMATION OF STATE BY PARTICLE FILTER

Being a suitable method for state estimation of the proposed state space model particle filter has been proposed.

Particle filter [1] is one of promising approaches for this kind of applications since it guarantees the confidence by estimated probability distribution and time update procedure in [2] with certain number of particles is suit for real time applications.

Particles for the current time are generated according to the system model by referring each particle of previous time.

$$\tilde{\boldsymbol{x}}_k^{(l)} \sim f\left(\boldsymbol{x}_k \middle| \boldsymbol{x}_{k-1}^{(l)}\right), \;\; l = 1,2,\dots,M\,. \tag{15}$$

Then calculate likelihood by observation model, and let the weight $w_k^{(l)}$ being to have the likelihood value for each particle.

$$w_k^{(l)} = h\left(\boldsymbol{I}_k \middle| \tilde{\boldsymbol{x}}_k^{(l)}\right), \;\; l = 1,2,\dots,M\,. \tag{16}$$

Lastly, a procedure called “resampling” has been conducted by drawing with replacement from predicted particle set according to the probability proportional to the weight value $w_k^{(l)}$, thus obtaining a set of particles which approximate “filter distribution”.

IV. EXPERIMENTS

Experiments have been conducted in indoor environment with different conditions of face: 1) static scene; 2) rotation; 3) parallel motion; 4) depth motion. Video images were collected with a web camera. A real-time system has been implemented by C Language program with OpenCV 2.1 as in [11] .Members in our laboratory have been taken as the objects. Initial flame of the video images has been shown in Fig.7. For each condition, video images have employed. 500 particles are used in particle filter with parameters $\tau_x^2 = 10, \tau_y^2 = 10, \tau_\theta^2 = 15$. Some image frames with tracking result which is $k = 10, k = 20, k = 30, k = 40, k = 50$, have been shown in Fig. 8, 9, 10, 11.

Fig. 7 Typical image to decorate

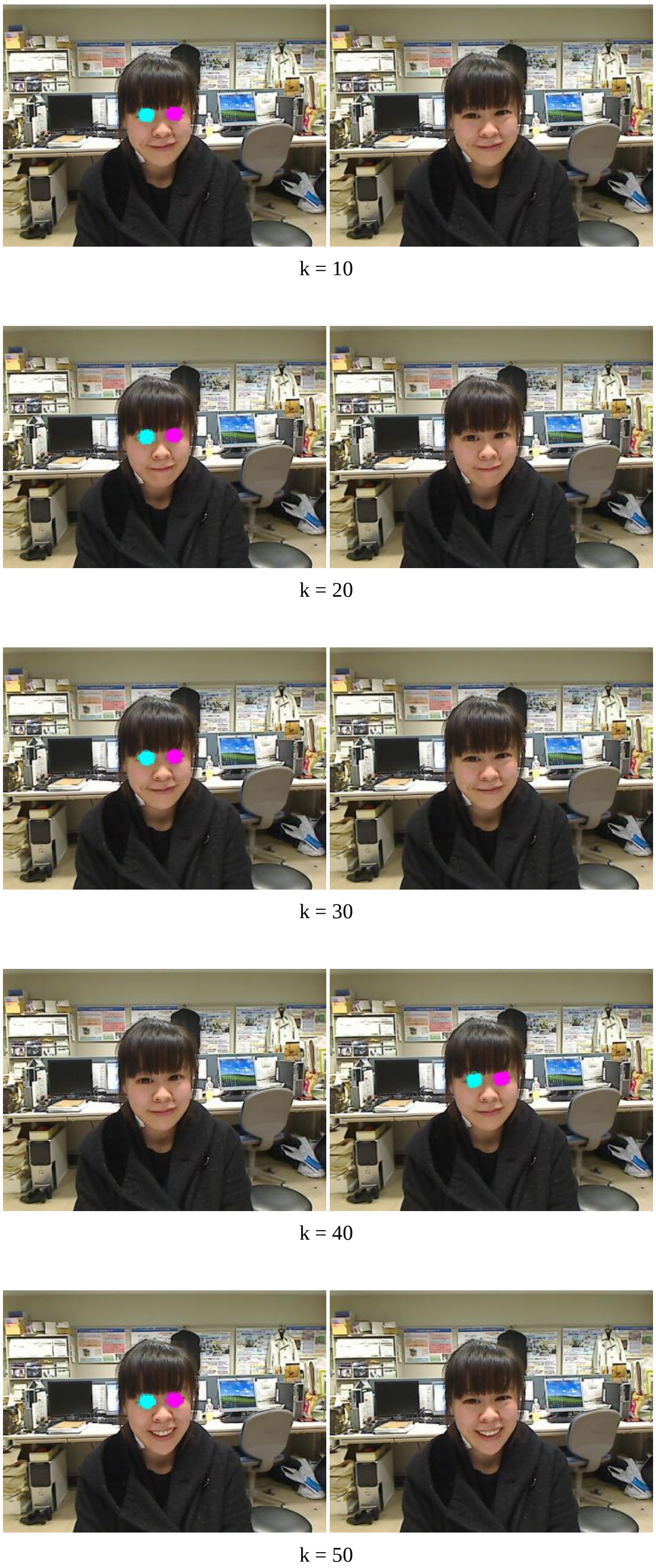


Fig.8　Face in the static scene

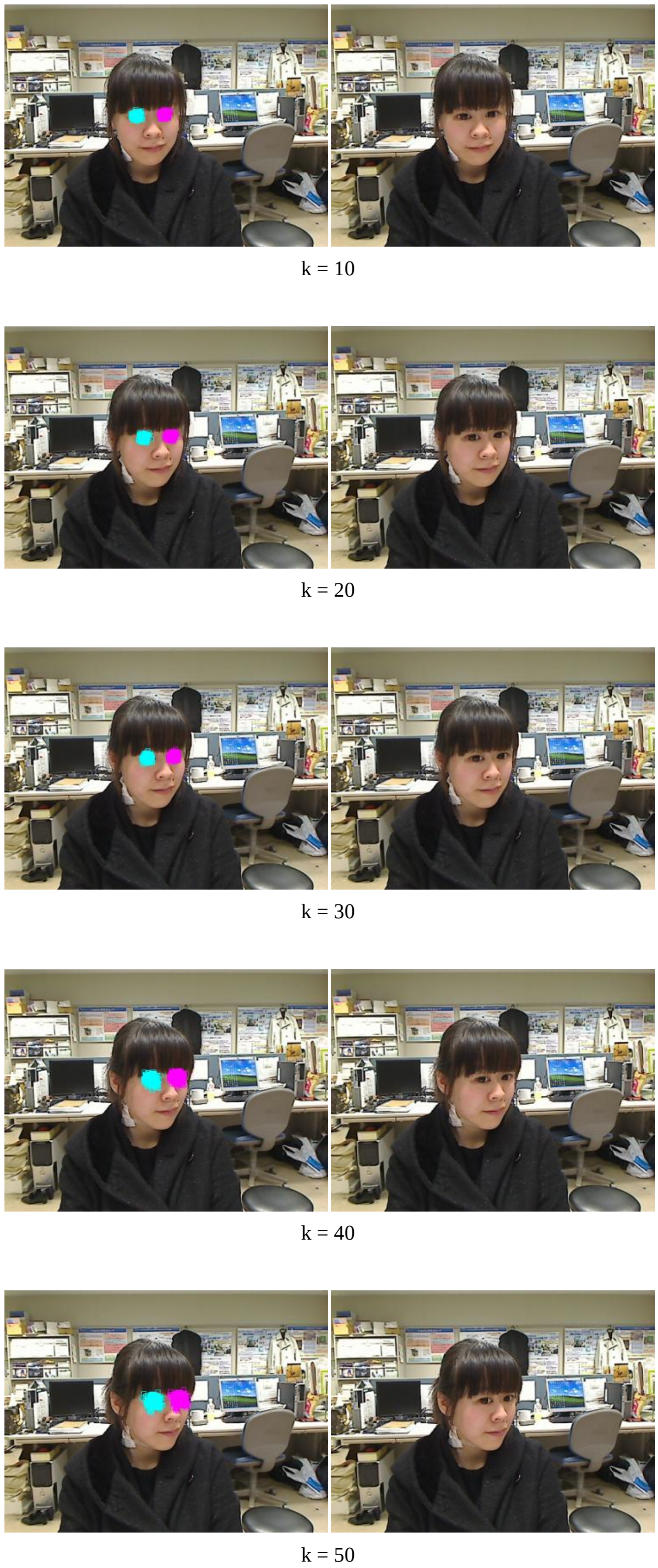
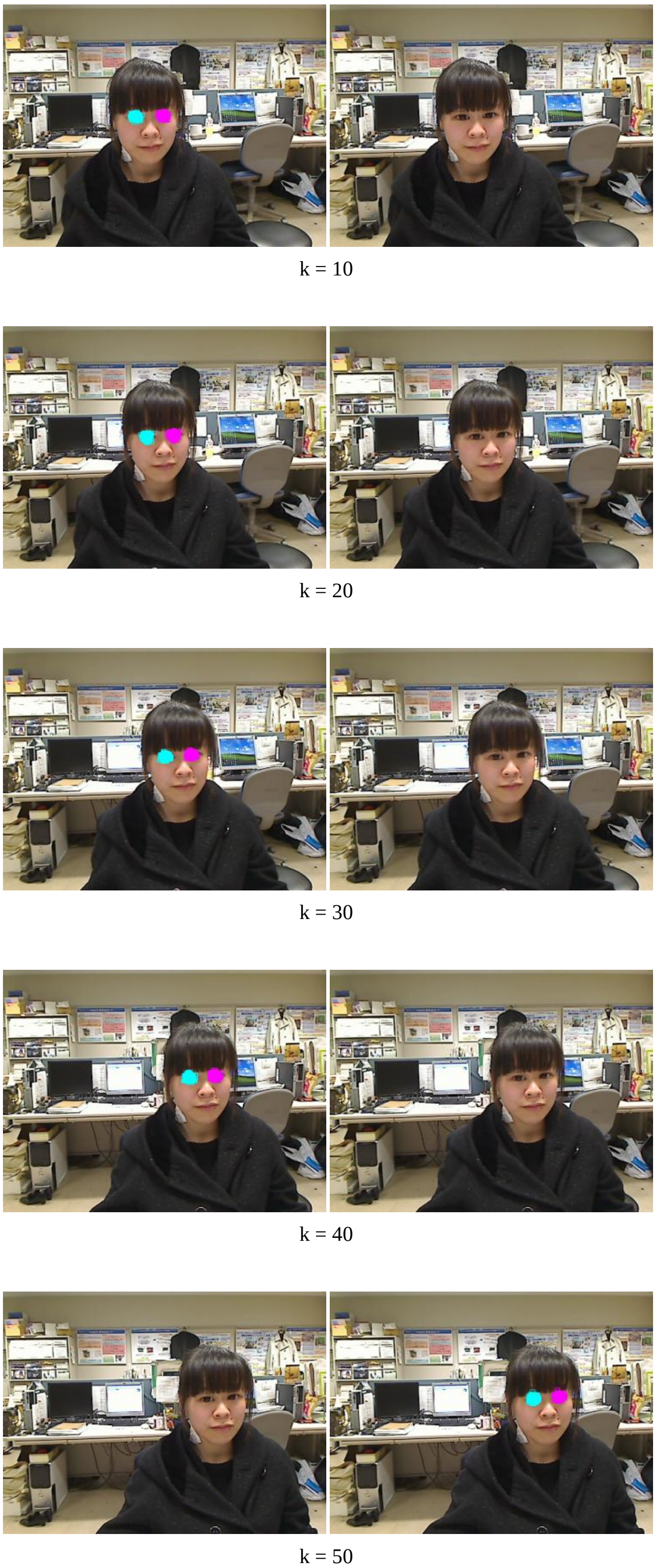


Fig.9　Face in the rotation



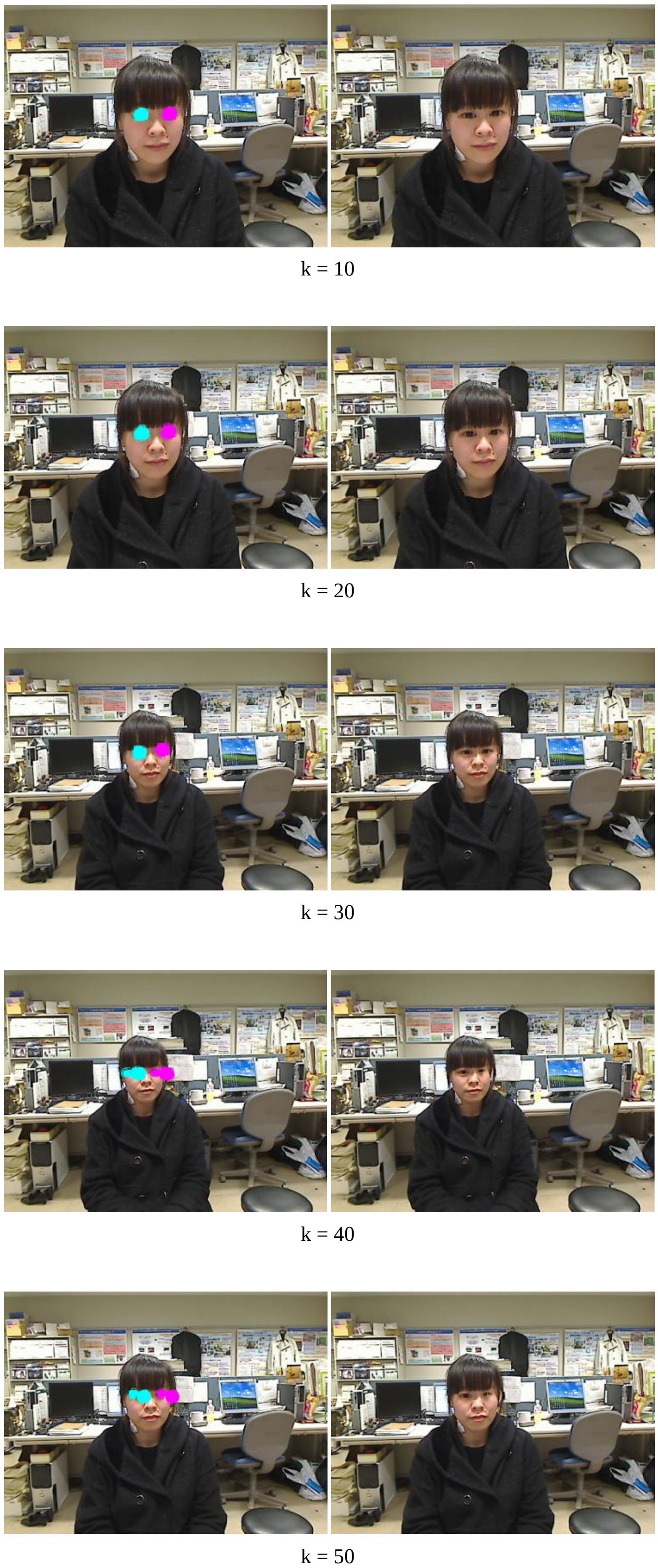


Fig.10　Face in the parallel motion

Fig.11　Face in the depth motion

V. CONCLUSION

We have proposed a method of face decoration by estimating eye positions in video image which have used particle filter algorithm. Experiments have been implemented in a real-time system with different conditions of face, such as 1) static scene, 2) rotation, 3) parallel motion, and 4) depth motion, have conducted to demonstrate performance of the proposed method. Experimental validation can be confirmed from performances in 1) 2) 3). Performance in 4) was difficult to achieve the result we wanted since it is out of our consideration. However, in practical situation, depth motion also exists. It is necessary to extend our proposed model in order to improve the performance of decoration.

In addition, color of the pupils, is used as likelihood evaluation and becomes a problem since its confusing color is similar to the eyebrows and fringes above the eyes. In this case, we limit range of the image to be processed by using the result of face detection. It may also be helpful to consider other factors besides color of the pupils. In future, we hope to further analyze a possibility that more than one person appear in video image by further enhancing our proposed method in order to make it more efficient.

ACKNOWLEGYMENT

This study was supported by JSPS KAKENHI Grant-in-Aid for Scientific Research (C) 22500158 and ISM cooperative research program (2011-ISM-CRP-2014).

REFERENCES

[1] A.Doucet, N.de Freitas, and N.Gordon, "Sequential Monte Carlo Methods in Practice", Springer, New York, 2001.

[2] Norikazu Ikoma, Yuduo Chen, Hideaki Kawano, Hiroshi Maeda, "Real-time Estimation of Face Posture by Eyes detection and tracking with Particle Filter", 4th International Symposium on Computational Intelligence and Industrial Applications ISCIIA, 2010.

[3] Miao Ma, Shengrong Ding, Yanfei Zhu, "A Self-Adaptive Method of Image Contrast Enhancement Based on Artificial Bee Colony Algorithm", International Conference on Circuit and Signal Processing, 2010.

[4] Yiqiang Chen, Jinjing Xie, Junfa Liu, "Multi-culture Facial Attractiveness Enhancement Based on Double Knowledge Transferring", International Journal of Information Technology Vol. 17 No. 1, 2011.

[5] Arakawa Kaoru, Nomoto Kohei, "A Face Beautification System Using Interactive Evolutionary Computing", IEIC Technical Report, 2005.

[6] Tajiri Fumio, Matsui Takashi, Arakawa Kaoru, Nomoto Kohei, "A Face Beautification System Using Interactive Evolutionary Computing", IEIC Technical Report, 2006.

[7] Tommer Leyvand, Daniel Cohen-Or, Gideon Dror, Dani Lischinski, "Data-driven enhancement of facial attractiveness", ACM SIGGRAPH paper, Los Angeles, California , 2008.

[8] EISENTHAL, Y., DROR, G., AND RUPPIN, E., "Facial attractiveness: Beauty and the machine", *Neural Computation 18*, 1, pp.119-142, 2006.

[9] Decoration products: http://panasonic.jp/dc/fx77/retouch.html

[10] Decoration products: http://www.bandainamcogames.co.jp/am/vg/eyemix/

[11] OpenCV 2.1 Reference: http://opencv.jp/opencv-2.1/c/

[12] P. Viola, M. Jones: "Robust Real-time Object Detection", International Journal of Computer Vision 57(2), pp.137-154, 2004.