Virtual mirror based on 3D shape reconstruction and real-time face tracking

Takao Makino*[a], Toshiya Nakaguchi[a], Norimichi Tsumura[a], and Yoichi Miyake[b]
[a] Department of Information and Image Sciences, Chiba University,
Yayoi-cho, Inage-ku, Chiba, 263-8522, Japan
[b] Research Center for Frontier Medical Engineering, Chiba University,
Yayoi-cho, Inage-ku, Chiba, 263-8522, Japan

ABSTRACT

A simple 3D virtual mirror interface system is proposed for acquisition and reconstruction of facial pattern, which enables a real time tracking of 3D facial pattern. The movement of lip in the face, which can be estimated based on mean shift and Kalman filter is introduced to recognize the rotation and parallel shift of face. Modified Photometric Stereo Method (PSM) is also proposed to acquire the 3D shape without the influence of deficiency in shadow area of the object. It is shown that the proposed system is effective to acquire and reconstruct for the real-time tracking of 3D facial pattern image than traditional approaches.

Keywords: Virtual mirror, Photometric Stereo method, Real time tracking, Mean shift, Kalman filter

1. INTRODUCTION

Human face gives various kinds of information such as feeling, age, physical condition and so on. Therefore it is important to acquire and reconstruct the facial pattern precisely for various kinds of applications. Our goal is to create the electronic cosmetics systems with real-time tracking of 3D facial pattern. A virtual mirror system[1] for electronic cosmetics has been used with mechanical half mirror as shown in Figure 1. The system was not familiar for consumer since it can acquire and reproduce only two-dimensional image and expensive mechanical half mirror is used. In this paper, we propose a simple virtual mirror interface system without using half mirror as shown in Figure 2. The system is consisted of movie camera, three light sources placed different angles, monitor and personal computer. The 3D image is acquired and reconstructed by using Photometric Stereo Method (PSM)[2,3]. The conventional PSM uses multiple light sources to take multi-angle pictures for the calculation of normal vector and reflectance of the object. However, in this method, the fatal error is arisen to reconstruct the 3D image at low reflection area such as hair and iris. Therefore, we

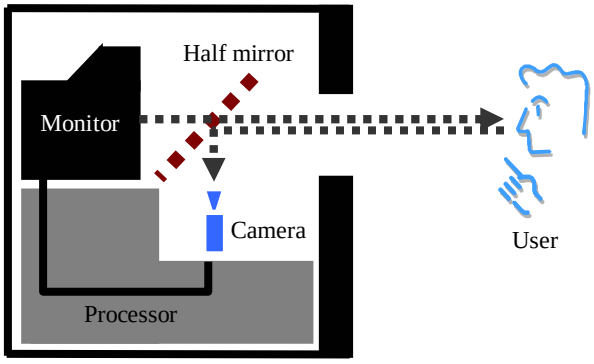


Figure 1. Typical virtual mirror interface

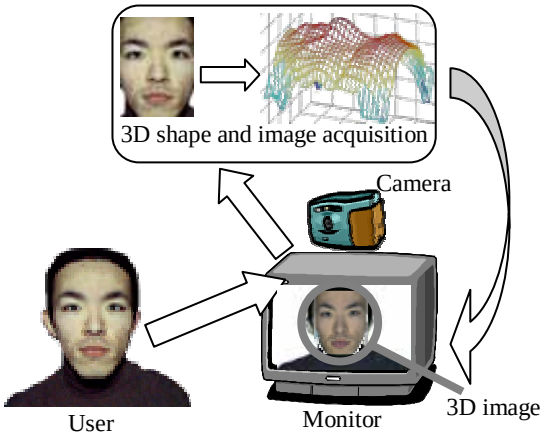


Figure 2. Overview of our virtual mirror interface

* makino@graduate.chiba-u.jp; phone +81-43-290-3261

propose a new method to decrease this fault of conventional PSM and construct the real-time 3D facial pattern tracking system to recognize the rotation and parallel shift of face.

Block diagram of this interface is shown in Figure 3. The 3D human face is acquired by PSM in pre-processing, and then the face tracking is performed on each frame image taken by video camera.

2. 3D FACE SHAPE AND IMAGE ACQUISITION

Our interface employs the PSM to acquire the 3D face image. It obtains surface normal vector and reflectance of the object by using more than 2 light sources. However, this method has a shortcoming that black area on the face, such as the pupil of the eye and the eyebrow, occurs the fatal error of the calculation. Therefore we propose the modified method to avoid the fatal error.

2.1. Photometric Stereo Method

PSM estimates surface normal and reflectance data by using at least 3 images taken by different directions of light source as follows,

$$r(x,y) = \left\| \mathbf{M}^{-1}\mathbf{I}(x,y) \right\|, \tag{1}$$

$$\mathbf{n}(x,y) = \mathbf{M}^{-1}\frac{1}{r(x,y)}\mathbf{I}(x,y), \tag{2}$$

where

$$\mathbf{M} = \begin{bmatrix} \mathbf{m}_1 & \mathbf{m}_2 & \cdots & \mathbf{m}_l \end{bmatrix}^T,$$

$$\mathbf{m}_k = \begin{bmatrix} m_{kx} & m_{ky} & m_{kz} \end{bmatrix}^T,$$

$$\mathbf{I}(x,y) = \begin{bmatrix} \frac{i_1(x,y)}{t_1} & \frac{i_2(x,y)}{t_2} & \cdots & \frac{i_l(x,y)}{t_l} \end{bmatrix}.$$

Here, $l(\geqq 3)$, $r(x,y)$, $\mathbf{m}_k$, $\mathbf{n}(x,y)$, $\boldsymbol{i}_k(x,y)$ and t_k denote number of light sources, surface reflectance on the image grid, a directional vector of k-th light source, a normal vector of the surface, image intensity acquired on k-th light source and k-th light intensity, respectively. Normal vector map is shown in Figure 4(a) and the reflectance image is shown in Figure 4(b). By integrating normal vector distribution, we can reconstruct the 3D shape of human face.

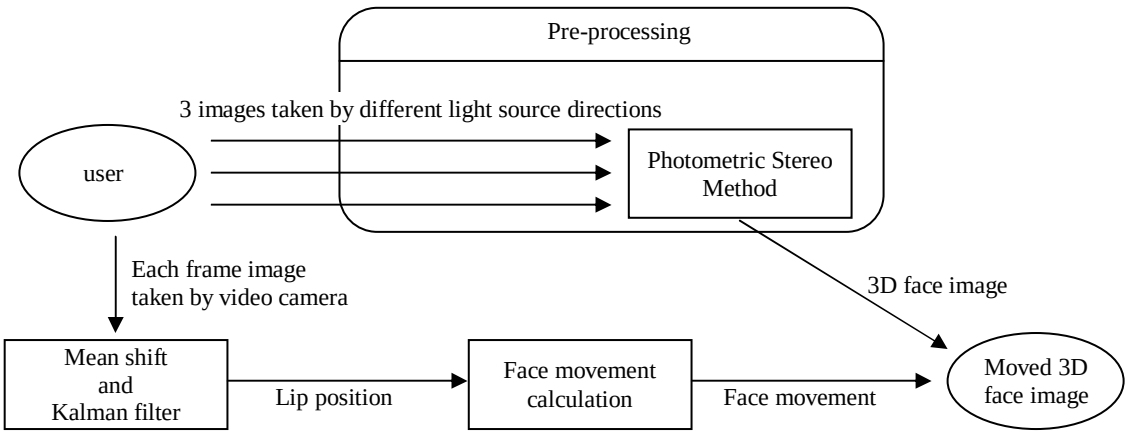
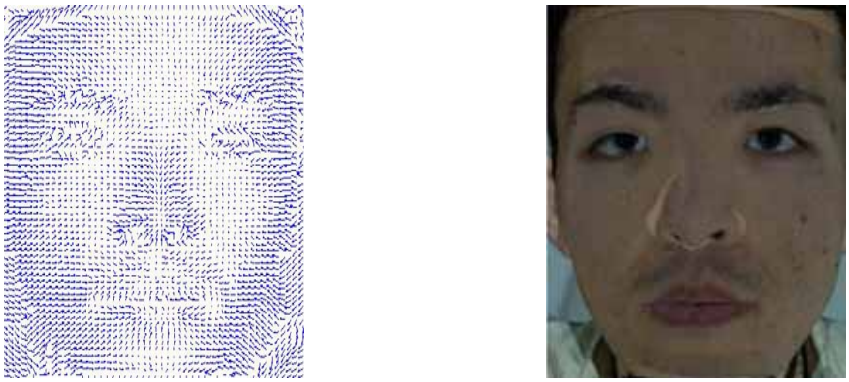


Figure 3. Block diagram of 3D virtual mirror interface



(a) (b)

Figure 4. (a)Normal vector map , (b)the reflectance image

(a) (b)

Figure 5. Reconstructed 3D shapes, (a)conventional PSM, (b)modified PSM

2.2. Modified Photometric Stereo Method

However the conventional PSM has a serious problem that correct normal vectors can not be obtained on the black area. Therefore, the 3D shape that reconstructed by integration of normal vector has a fatal error in and around the black area. Reconstructed 3D shape by the conventional PSM is shown in Figure 5(a). In this case, the integration of normal vector is calculated only horizontally, then the calculation error occurred by the black area is accumulated at the left and right edge of the face. To solve this problem, we propose the method that the integration is performed both horizontally and vertically ignoring the black area and the 3D shape of the black area is interpolated linearly by means of surrounding points. Reconstructed 3D shape by the modified PSM is shown in Figure 5(b). It is shown that the calculation error is drastically reduced.

3. FACE TRACKING

In this section, the face tracking method using mean shift and Kalman filter[6,7,8,9] is described. It has been proven to be efficient and relatively robust to the rapid movement of the object. In this method, the lip position is estimated by mean shift tracker with Kalman filtering. In case that the tracker loses the lip, the size of the predicted area with Kalman filter is temporarily expanded and iterate mean shift tracker until it converges.

3.1. Mean Shift Tracking

In this study, the mean shift tracking based on color histogram model is used. The center and size of the target object are found via mean shift algorithm. At first, the initial region where mean shift calculation is performed is decided arbitrary. This region called search window. Next, the image is converted to the color probability distribution image via a color

histogram of the target object. Using target color histogram, the color probability distribution image which bright pixels correspond to the target color is obtained. In the search window, the mean position of the target object is computed and the search window is centered at that position. The mean position can be obtained as follows,

$$x_c = \frac{M_{10}}{M_{00}}, \; y_c = \frac{M_{01}}{M_{00}}, \tag{3}$$

where

$$M_{00} = \sum_x \sum_y I(x, y),$$
$$M_{10} = \sum_x \sum_y xI(x, y),$$
$$M_{01} = \sum_x \sum_y yI(x, y).$$

Here, (x_c,y_c), $I(x,y)$, x and y denote the mean search window position, the pixel (probability) value in the position (x,y) in the image and the range over the search window, respectively. Position of the target object is tracked by iterating this mean position calculation until center of the search window converges.

3.2. Kalman Filter Tracking

The Kalman filter[11] is a set of mathematical equations to estimate the state of a process, in a way that minimizes the mean of the squared error. The equations for the Kalman filter fall into two groups: "time update", and "measurement update". The time update equations predict the target state which is characterized by its position and velocity, while the measurement update equations correct the predicted target state using an actual measurement at that frame. The time update equations are as follows,

$$\mathbf{x}^-_{k+1} = \mathbf{A}\mathbf{x}_k, \tag{4}$$
$$\mathbf{P}^-_{k+1} = \mathbf{A}\mathbf{P}_k\mathbf{A} + \mathbf{Q}_k, \tag{5}$$

where $\mathbf{x}^-_k$, $\mathbf{x}_k$, $\mathbf{A}$, $\mathbf{P}^-_k$, $\mathbf{P}_k$ and $\mathbf{Q}_k$ denote a priori state estimate at frame k, a posteriori state estimate at frame k, a state transition matrix, a priori estimate error covariance, a posteriori estimate error covariance and process noise covariance matrix, respectively. The measurement update equations are as follows,

$$\mathbf{x}_k = \mathbf{x}^-_k + \mathbf{K}_k\left(\mathbf{z}_k - \mathbf{H}\mathbf{x}^-_k\right), \tag{6}$$
$$\mathbf{P}_k = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}\right)\mathbf{P}^-_k, \tag{7}$$
$$\mathbf{K}_k = \mathbf{P}^-_k\mathbf{H}\left(\mathbf{H}\mathbf{P}^-_k\mathbf{H} + \mathbf{R}_k\right), \tag{8}$$

where $\mathbf{z}_k$, $\mathbf{H}$, $\mathbf{I}$ and $\mathbf{R}_k$ denote a measurement vector at frame k, the measurement matrix, the identity matrix and measurement noise covariance matrix, respectively. The state estimate is updated in equation(6) finally. Given some initial conditions(e.g. initial estimates for $\mathbf{x}_k$ and $\mathbf{P}_k$), the filter estimates the state at some frame and then obtains feedback in the form of noisy measurements.

3.3. Combining Mean Shift and Kalman Filter

An integrated method, mean shift and Kalman filter has been proposed[9,10]. The measurement vector is derived based on mean shifts, while the prediction of the next target position is computed by a Kalman filter. In case the mean shift tracker loses the object, the size of search window is temporarily expanded and mean shift tracking is iterated. Block diagram of this tracking is shown in Figure 6.

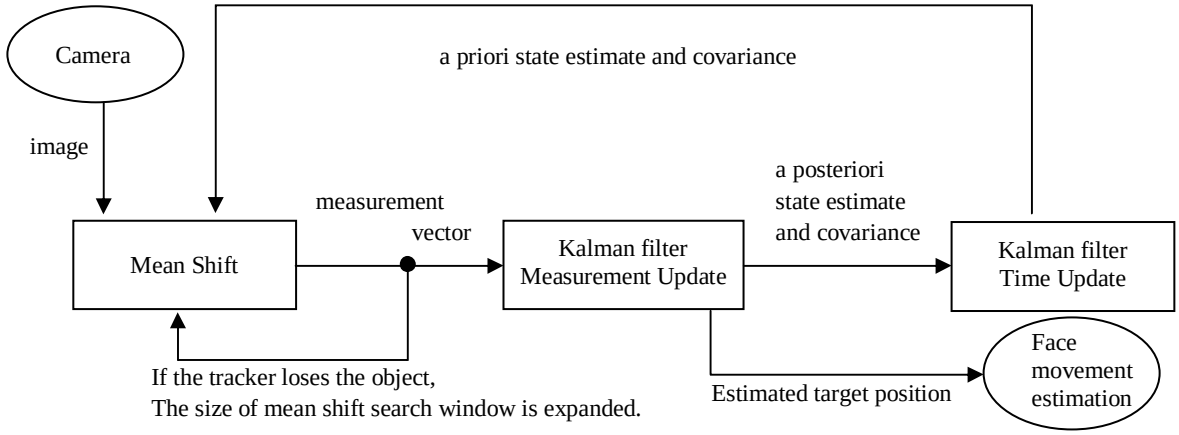


Figure 6. Block diagram of integrated tracking mean shift and Kalman filter

4. FACE MOVEMENT ESTIMATION

After estimating the lip position, eyes positions and the neck position are estimated by using the color based segmentation. The face movement is estimated by these features.

To predict the face rotation and parallel shift, we use the lip, eyes and the neck positions. They are predicted by using both YIQ and HSV color spaces with positional relationship to the lip as shown in Figure 7. First, candidate positions for eyes are extracted by I-component of YIQ and V-component of HSV. Eyes positions are the nearest symmetrical candidates to the lip, except candidate positions on cavity of nose. The neck position is center of the neck width, and the neck width is estimated as the width of the skin color area under the lip position. The distance between the neck position and the lip position is half of the distance between the lip position and the center point of eyes.

From the distance between the lip and the neck, the roll rotation angle of the face is calculated as shown in Figure 8. Where *d*, *w* and *r* denote the depth of the face, the width of the neck and the distance between the lip and the neck. Depth of the face *d* can be calculated as follows

$$d = \frac{3w}{2} \ . \tag{9}$$

The rotation angle of the face θ can be given as

$$\theta = \sin^{-1}(\frac{2r}{3w}) \ . \tag{10}$$

The distance of parallel shift on x and y-axis can be estimated by the distance of the neck movement. The neck also is used to estimate parallel shift on z-axis. In the case that the neck width increased, we assume that the face approached the camera, while if the neck width decreased, the face separated from the camera.

.

5. EXPERIMENT

Geometry of the virtual mirror interface is shown in Figure 8. Here we use three light sources (Speed light Nikon SB-30) and the camera(Sony DFW-VL500). In front of the lens of the camera and three lights, a polarizing plate is attached to avoid surface reflective light. Figure 10 shows the acquired images from three light sources for the modified PSM. For the PSM, direction vectors of the light sources and the camera are measured in advance. After reconstructing 3D facial shape and image, face tracking starts during the experiment. For the movie acquisition, the image processing and the real-time 3D image rendering, DirectShow[12], Intel's OpenCV library[13], and the OpenGL 3D graphics library[14] is employed. Execution results are shown in Figure 11. Right half on the screen represents input image from the camera

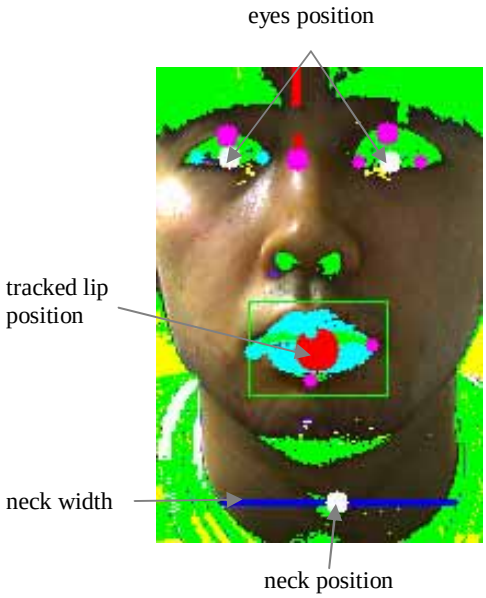


Figure 7. Eyes and the neck prediction

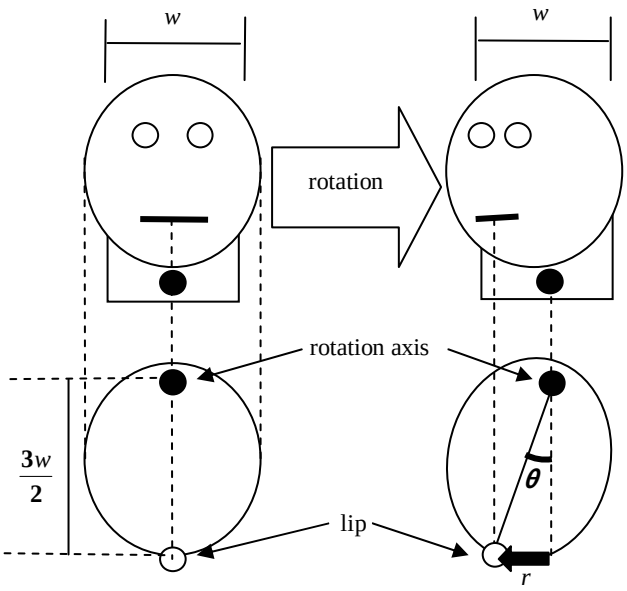


Figure 8. Calculation of rotation angle

with the result of eyes and the neck prediction, and another represents the virtual mirror with processed image. From Figure 11(a), (b), we can see that the face tracking is successfully working. However, this interface can't reproduce the human's hair because of its complicated shape and that PSM's essential problem. We found that the interface is evaluated by six people as Figure 11(c), any individually dependent problems are not observed.

6. CONCLUSION AND DISCUSSION

In this paper, we proposed the simple 3D virtual mirror interface which consists of multiple light sources and a camera in addition to the set of the personal computer. The 3D face shape and image are reconstructed by using the modified PSM which can reconstruct 3D shape on black area accurately. Mean shift and Kalman filter based real-time face tracking with the lip position works successfully without any individual dependency.

We consider that this interface can be applied to the electronic cosmetics systems that can simulate make-up, plastic surgery and color appearance under various illuminations. Therefore, to implement more realistic 3D virtual mirror interface, the hair and user's expression must be reproduced, and accuracy of face reconstruction must be improved. Additionally, exact color reproduction of the face is also required.

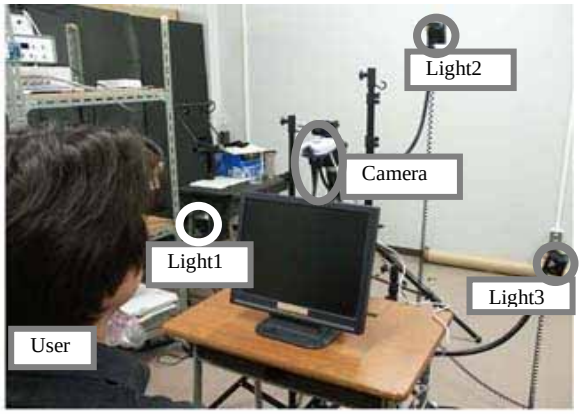
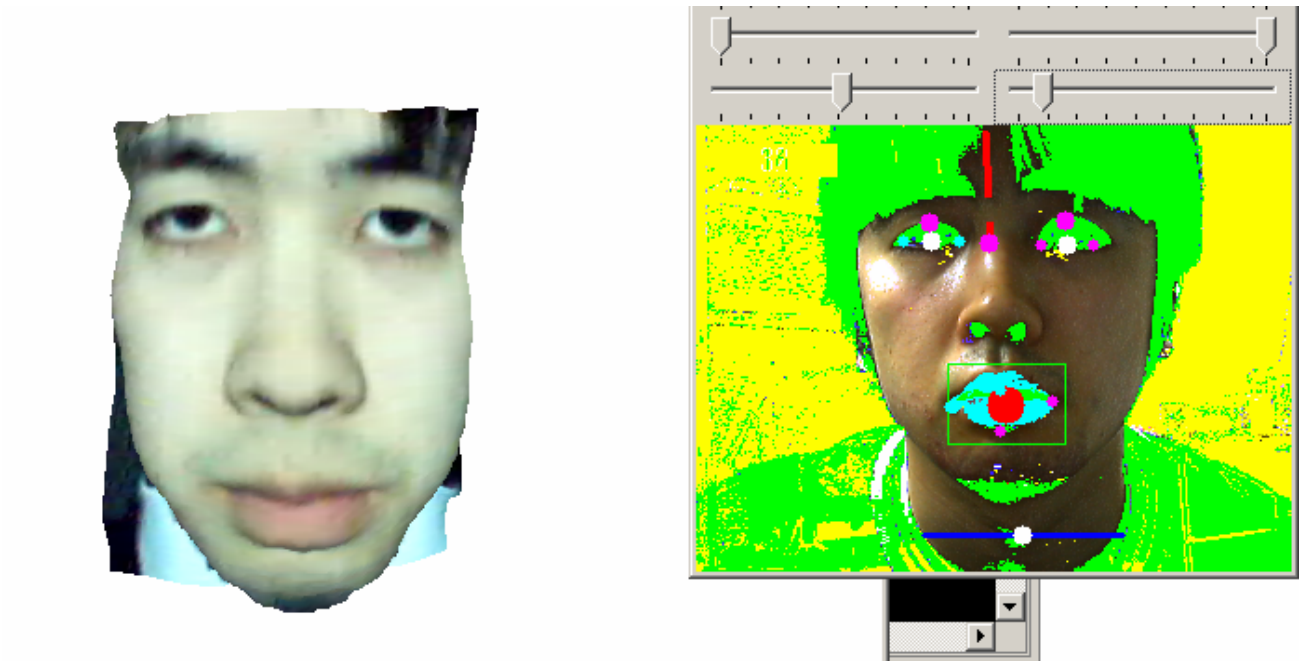
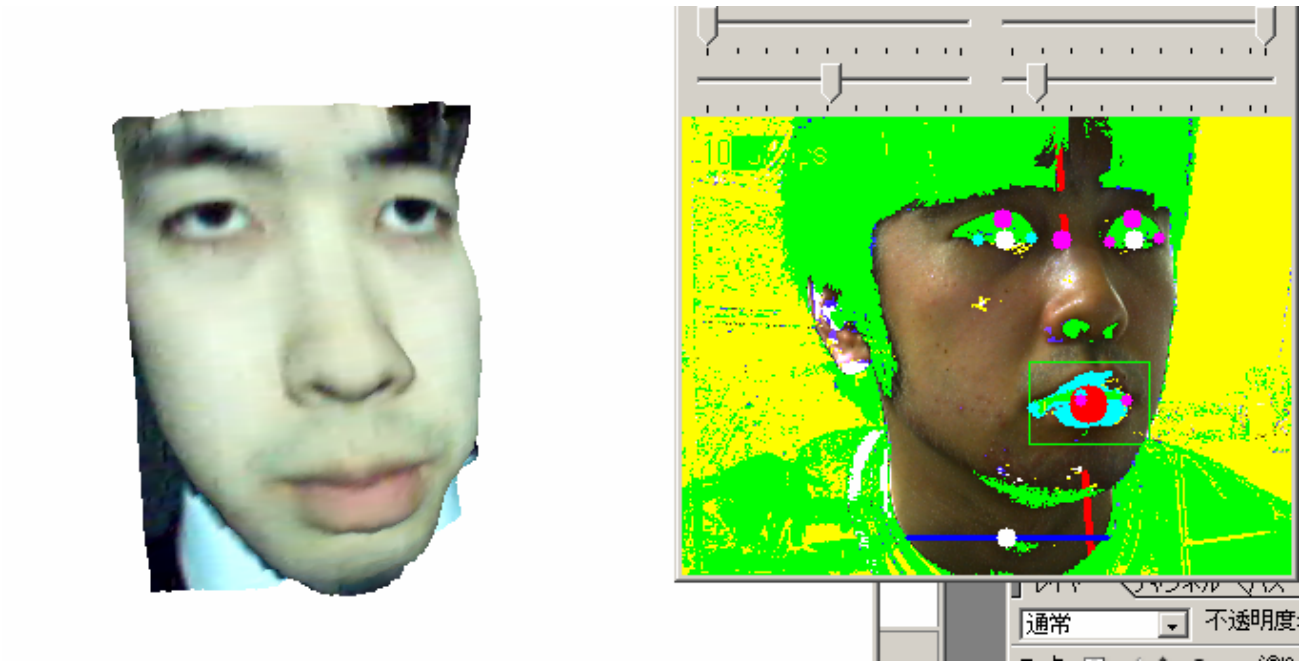


Figure 9. Geometry of the virtual mirror interface

Figure 10. Acquired images from three light sources

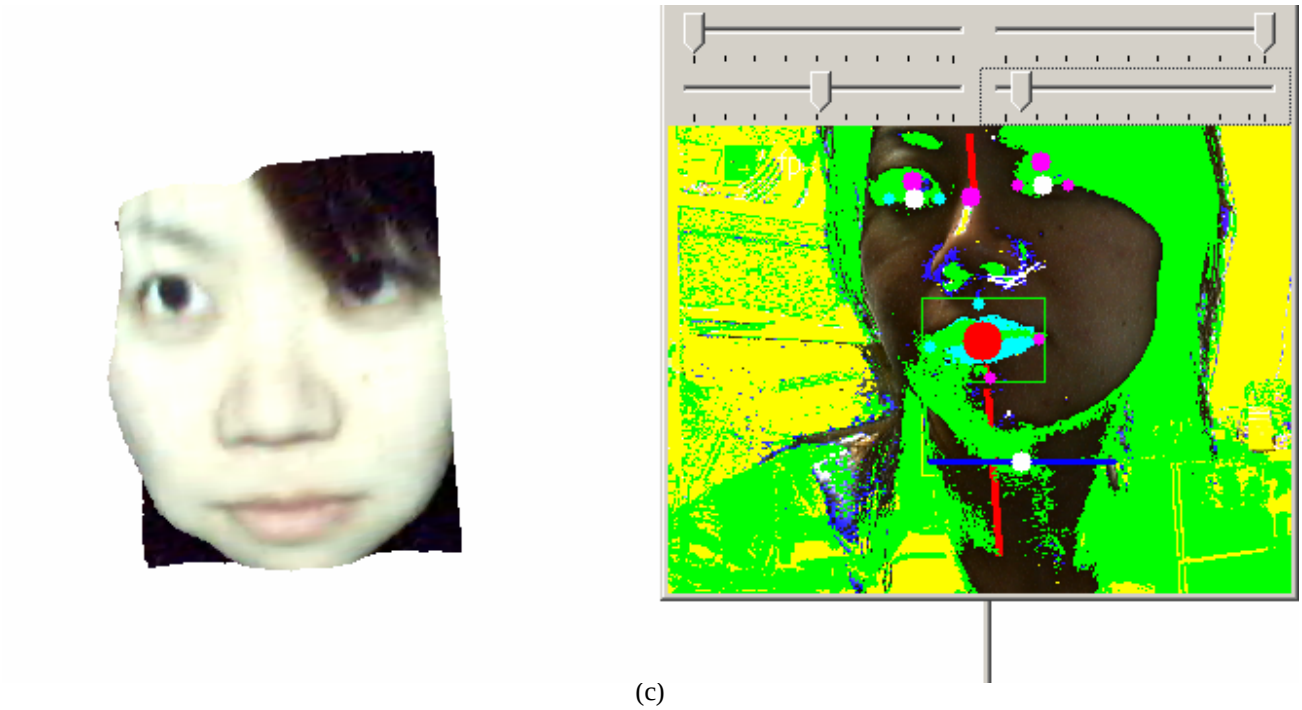


(a)



(b)

Figure 11. (a),(b)Face tracking is successfully working, (c)evaluated by the other user



(c)

Figure 11. (a),(b)Face tracking is successfully working, (c)evaluated by the other user

REFERENCES

1. T.Darrell, G.Gordon, J.Woodfill, M.Harville, *A Virtual Mirror Interface using Real-time Robust Face Tracking*, In Proceedings of the Third International Conference on Face and Gesture Recognition, April 14-16, 1998.
2. R. J. Woodham,*Photometric method for determining surface orientation from multiple images*, Optical Engineering 19, pp.139-144, 1980.
3. Norimichi Tsumura, Minh Nguyet Dang, Takao Makino and Yoichi Miyake, *Estimating the directions to light sources using images of eye for reconstructing 3D human face*, Color Imaging Conference 2003: 77-81
4. Rowley, H., Baluja, S., and Kanade, T, *Neural Network-Based Face Detection*, Proc. IEEE Conf. on Computer Vision and Pattern Recognition, San Francisco, CA, pp. 203-207. 1996.
5. Jean-Christophe Terrillon, Martin David, and Shigeru Akamatsu, *Automatic detection of human faces in natural scene images by use of a skin color model and of invariant moments*, in International Conference on Face and Gesture Recognation, number 3, pp.112-117.IEEE,April 1998.
6. Xiaozhou Wei, Zhiwei Zhu, Lijun Yin, Qiang Ji, *A Real Time Face Tracking And Animation System*,first IEEE Workshop on Face Processing in Video, in conjunction with IEEE International Conference on Computer Vision and Pattern Recognition (CVPR'04).
7. Zhiwei Zhu, Qiang Ji, Kikuo Fujimura, Kuangchih Lee, *Combining Kalman Filtering and Mean Shift for Real Time Eye Tracking Under Active IR Illumination*, International Conference on Pattern Recognition, 2002.
8. Qiang Ji, Xiaojie Yang, Real-Time Eye, Gaze, and Face pose Tracking for Monitoring Driver Vigilance, Real-Time Imaging, Volume 8 , Issue 5 (October 2002), Pages: 357 - 377.
9. Yoichi Kuroda, Kazuyuki Morioka, Joo-Ho Lee1, Hiroyuki Yashima1, Hideki Hashimoto, *Modified Hybrid Tracking Algorithm Using Mean Shift and Kalman Filter*, 2004 RISP International Workshop on Nonlinear Circuit and Signal Processing (NCSP'04) Hawaii, USA, Mar. 5-7, 2004.
10. D.Comaniciu, V.Ramesh, *Mean Shift and Optimal Prediction for Efficient Object Tracking*, IEEE Int. Conf. Image Processing (ICIP'00), Canada, Vol. 3, 70-73, 2000
11. Greg Welch and Gary Bishop, *An Introduction to the Kalman Filter*, University of North Carolina at Chapel Hill technical report TR 95-041
12. http://www.microsoft.com/windows/directx/default.aspx
13. http://www.intel.com/research/mrl/research/opencv/
14. http://www.opengl.org/