ECSE 6650 - Computer Vision

Final Project: 3-D Face Reconstruction Using Passive Stereo

Jie Zou, Peng-Jui Ku, Luyun Chen

December 17, 2001

1. Introduction

Three-dimensional facial reconstruction is an important research topic in computer vision. It consists of recreating a 3D human face model from two stereo images captured from different viewpoints. Compared to the 3D reconstruction of other types of object, a face is primarily characterized by several distinctive facial features such as eyebrows, eyes, nose, mouth, and face contour. It is important to accurately render these features in order to obtain a successful facial reconstruction.

In this final project, we apply the same 3D reconstruction approach implemented in Project #2 to analyze the performance of passive stereo technique regarding to this specific problem. In addition, we modify the existing algorithms to solve some practical issues encountered during the rectification and correlation processes. Section 2 describes the reason of these modifications and the theory behind them. It is followed by the illustration of experimental setup and results in section 3. The conclusion is summarized in section 4.

2. Mathematic discussion of passive stereo

2.1 Camera calibration

Conventional camera calibration methods rely on two or more images of a calibration pattern: that is, a 3-D object of known geometry, possibly located in known position in space and generating image features, which can be located accurately. The accuracy of calibration depends on the accuracy of the measurements of the calibration pattern.

The full perspective projection camera model equation is:

$$l\begin{pmatrix} c \\ r \\ 1 \end{pmatrix} = P\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} p_1^T & p_{14} \\ p_2^T & p_{24} \\ p_3^T & p_{34} \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = WM\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} s_x f r_1 + c_0 r_3 & s_x f t_x + c_0 t_z \\ s_y f r_2 + r_0 r_3 & s_y f t_y + r_0 t_z \\ r_3 & t_z \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix},$$

where l is a scale factor, (c,r) is the image point, and (x,y,z) is the corresponding 3D point. P is called the homogeneous projection matrix. p_i^T's are 3×1 vectors and p_{i4}'s are scalars. $W = \begin{pmatrix} fs_x & 0 & c_0 \\ 0 & fs_y & r_0 \\ 0 & 0 & 1 \end{pmatrix}$ is often referred to as the intrinsic matrix and $M = \begin{pmatrix} R & T \end{pmatrix}$ as exterior matrix. f is the focal length and s_x, s_y are scale factors. r_1, r_2, and r_3 are the row vectors of the rotation matrix R, and $T = \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}$ is the translation matrix. c_0 and r_0 are the coordinates of the principal point in pixels relative to image frame.

From the above equation, for each pair of 2D-3D points, we have:
$p_1{}^T M_i + p_{14} - c_i p_3{}^T M_i - c_i p_{34} = 0$,
$p_2{}^T M_i + p_{24} - r_i p_3{}^T M_i - r_i p_{34} = 0$
where, (c_i, r_i) is the image point and $M_i = (x, y, z)^T$ is the corresponding 3D point.

For N points, we can setup a system of linear equations:

$$AV = \begin{pmatrix} M_1 & 1 & \vec{0} & 0 & -c_1 M_1 & -c_1 \\ \vec{0} & 0 & M_1 & 1 & -r_1 M_1 & -r_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ M_N & 1 & \vec{0} & 0 & -c_N M_N & -c_N \\ \vec{0} & 0 & M_N & 1 & -r_N M_N & -r_N \end{pmatrix} \begin{pmatrix} p_1^T \\ p_{14} \\ p_2^T \\ p_{24} \\ p_3^T \\ p_{34} \end{pmatrix} = 0, \qquad (2\text{-}1)$$

where A is a $2N \times 12$ matrix depending only on the 3-D and 2-D coordinates of the reference points, and V is a 12×1 vector.

In general, the rank of A is 11 if $N > 6$ and if these points are not coplanar. In order to yield a linear solution, we can impose one of the normal constraints, $\|p_3\|^2 = 1$, then the problem is converted to a constrained linear least-squares problem. That is to minimize $\|AV\|^2$ subject to $\|p_3\|^2 = 1$.

$$e^2 = \|AV\|^2 + l\left(\|p_3\|^2 - 1\right)$$

Decomposing A into two matrices B and C, and V into Y and Z, where

$$C = \begin{pmatrix} -c_1 M_1 \\ -r_1 M_1 \\ \vdots \\ -c_N M_N \\ -r_N M_N \end{pmatrix}, \; Y = \begin{pmatrix} p_1 \\ p_{14} \\ p_2 \\ p_{24} \\ p_{34} \end{pmatrix}, \text{ and } Z = p_3,$$

We have:

$$e^2 = \|AV\|^2 + l\left(\|p_3\|^2 - 1\right) = \|BY + CZ\|^2 + l\left(\|Z\|^2 - 1\right)$$
$$= (BY + CZ)^T (BY + CZ) + l\left(Z^T Z - 1\right)$$

Taking partial derivatives of e^2 with respect to Y and Z, and setting them to zeros yield:

$$Y = -\left(B^T B\right)^{-1} B^T C Z$$
$$C^T \left(I - B\left(B^T B\right)^{-1} B^T\right) C Z = l Z$$

The solution to Z is the eigenvector of matrix $C^T \left(I - B\left(B^T B\right)^{-1} B^T\right) C$. Given Z, we can solve for Y.

Substituting Y into $\|BY + CZ\|^2$ leads to $\|BY + CZ\|^2 = l$. This proves that solution Z corresponds to the eigenvector of the smallest positive eigenvalue of matrix $C^T \left(I - B\left(B^T B\right)^{-1} B^T\right) C$.

After P is calculated, we can easily establish the following equations:

$$\begin{aligned}
r_3 &= p_3 \\
t_z &= p_{34} \\
c_0 &= p_1^T p_3 \\
r_0 &= p_2^T p_3 \\
s_x f &= \pm\|p_1 \wedge p_3\| = \pm\sqrt{p_1^T p_1 - c_0^2} \\
s_y f &= \pm\|p_2 \wedge p_3\| = \pm\sqrt{p_2^T p_2 - r_0^2} \\
t_x &= (p_{14} - c_0 t_z)/(s_x f) \\
t_y &= (p_{24} - r_0 t_z)/(s_y f) \\
r_1 &= (p_1 - c_0 r_3)/(s_x f) \\
r_2 &= (p_2 - r_0 r_3)/(s_y f),
\end{aligned} \tag{2-2}$$

All the intrinsic and extrinsic parameters are thus solved.

Let $\hat{R}$ be the estimated rotation matrix. It is usually not orthonormal. We can find an orthonormal matrix R that is closest to $\hat{R}$ via SVD. Let $\hat{R} = UDV^T$. If we change D to identity matrix I, we have a new rotation matrix $R = UIV^T$. R is the orthonormal matrix that is the closest to $\hat{R}$.

2.2 Relative rotation and translation between two cameras

After we have done the camera calibration, we get rotation matrices for both left and right cameras R_L and R_R, and translation matrices for both left and right cameras T_L and T_R. We can easily get the relative rotation matrix R and translation matrix T between two cameras.

$$R = R_L R_R^T \text{ and } T = T_L - RT_R \qquad (2\text{-}3)$$

2.3 Rectification

The idea of the rectification is to transform left and right image planes into another two planes, which are coplanar and parallel to the base line. This arrangement has the advantage of limiting the search of corresponding points from a 2D image space to a 1D scan line; that is, the corresponding point for any points on the left image can be found on the same row in the right image.

If the focal lengths (in pixel unit) for left and right cameras are different, we need two steps for rectification. The first step is to construct two new images with the same focal length (in pixel unit), and the second one is to construct a rotation matrix to rotate the cameras such that the conjugate epipolar lines are collinear and parallel to the base line. These two steps are explained below. Since the rectification procedure for both left and right images are almost the same except the rectification rotation matrices are different, we will not distinguish the left and right images in the algebra formulation, but the difference in rectification rotation matrices will be explicitly specified.

If the focal lengths (in pixel unit) of left and right cameras are different, we need to construct a new image with new focal length (in pixel unit), and we set the new focal length for both left and right cameras to be the same. From perspective projection, we have $l\begin{pmatrix} c \\ r \\ 1 \end{pmatrix} = WM\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$, where $\begin{pmatrix} c \\ r \\ 1 \end{pmatrix}$ is the pixel coordinate, and $\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$ is the 3D coordinates.

So we have $\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = lM^{-1}W^{-1}\begin{pmatrix} c \\ r \\ 1 \end{pmatrix} = l'M^{-1}W'^{-1}\begin{pmatrix} c' \\ r' \\ 1 \end{pmatrix}$, where W' is the new intrinsic matrix and $\begin{pmatrix} c' \\ r' \\ 1 \end{pmatrix}$ is the new pixel coordinate. Hence, $\begin{pmatrix} c' \\ r' \\ 1 \end{pmatrix} = l W' W^{-1} \begin{pmatrix} c \\ r \\ 1 \end{pmatrix}$. Notice that this l is different from above l. It is l / l'. However, since all the l 's are just scale factors, it is not necessary to distinguish them.

The second step is to construct a rotation matrix to rotate the cameras such that the conjugate epipolar lines are collinear and parallel to the base line. We need to construct a triple of mutually orthogonal unit vectors e_1, e_2, and e_3. The first vector e_1 is given by the epipole, since the image center is in the origin, e_1 coincides with the direction of translation, and $e_1 = \frac{T}{\|T\|}$. Second vector e_2 must be orthogonal to e_1, so $e_2 = \frac{1}{\sqrt{T_x^2 + T_y^2}} \left[-T_Y, T_{x,} 0 \right]^T$. The third unit vector is unambiguously determined as $e_3 = e_1 \times e_2$. The rectification matrix is thus defined by $R_{rect} = \begin{pmatrix} e_1^T \\ e_2^T \\ e_3^T \end{pmatrix}$. We set the rectification rotation matrix for left image as $R_l = R_{rect}$, and rotation matrix for right image as $R_r = R_{rect} R$. Note that in this step, the image point is $\begin{pmatrix} c' \\ r' \\ 1 \end{pmatrix}$, and the intrinsic matrix is W'. For each image point $\begin{pmatrix} c' \\ r' \\ 1 \end{pmatrix}$, we can compute the rectified image point $\begin{pmatrix} c'' \\ r'' \\ 1 \end{pmatrix}$ by multiplying $\begin{pmatrix} c' \\ r' \\ 1 \end{pmatrix}$ with $W' R W'^{-1}$, so that $s \begin{pmatrix} c'' \\ r'' \\ 1 \end{pmatrix} = W' R W'^{-1} \begin{pmatrix} c' \\ r' \\ 1 \end{pmatrix}$, where W' is the intrinsic matrix and R is the rectification matrix. This equation can apply to both left and right images. From the first step, we have $\begin{pmatrix} c' \\ r' \\ 1 \end{pmatrix} = l W' W^{-1} \begin{pmatrix} c \\ r \\ 1 \end{pmatrix}$. This leads to

$$s \begin{pmatrix} c'' \\ r'' \\ 1 \end{pmatrix} = W' R W'^{-1} \begin{pmatrix} c' \\ r' \\ 1 \end{pmatrix} = W' R W'^{-1} l W' W^{-1} \begin{pmatrix} c \\ r \\ 1 \end{pmatrix} = l W' R W^{-1} \begin{pmatrix} c \\ r \\ 1 \end{pmatrix}.$$

Since s and l are just scale factors, we replace them by only one scale factor and obtain

$$s\begin{pmatrix} c'' \\ r'' \\ 1 \end{pmatrix} = W'RW^{-1}\begin{pmatrix} c \\ r \\ 1 \end{pmatrix}, \qquad (2\text{-}4)$$

where $\begin{pmatrix} c'' \\ r'' \\ 1 \end{pmatrix}$ is the rectified pixel coordinates and $\begin{pmatrix} c \\ r \\ 1 \end{pmatrix}$ is the original pixel coordinates; W is the original camera intrinsic matrix, and W' is the new intrinsic matrix; R is the rectification rotation matrix. For left camera, it is R_{rect}, and for right camera, it is $R_{rect}R$.

2.4 Establishing correspondences

After producing two rectified images from the previous step, the search for correspondences is reduced to a 1-D search.

One method to establish correspondence is SSD (Sum of Squared Differences) correlation method. For each left image pixel, its correlation with a right image pixel is determined by using a small correlation window of fixed size (e.g. 5*5 pixels or 7*7 pixels), in which we compute the SSD of pixel intensities:

$$c(\vec{d}) = \sum_{k=-W}^{W}\sum_{l=-W}^{W}\Psi(I_l(i+k, j+l), I_r(i+k-d_1, j+l-d_2)) \qquad (2\text{-}5)$$

where $(2W+1)$ is the width of the correlation window.
I_l and I_r are the intensities of the left and right image pixels respectively.
$[i, j]$ are the coordinates of the left image pixel.
$\vec{d} = [d_1, d_2]^T$ is the relative displacement between the left and right image pixels.
$\Psi(u,v) = -(u-v)^2$ is the SSD correlation function.

Since we are using the rectified images, the corresponding point of a left rectified image pixel has to be on the same row in the right rectified image. Therefore, d_2 is always equal to 0. The correlation-matching algorithm consists of calculating the correlation values pixel by pixel by varying d_1, which slides the correlation window from the left to the right in the right image along this row. The corresponding point is determined by the pixel that has the highest correlation value. The disparity between the two corresponding points is the distance d_1 that separates them.

The performance of correlation method can be jeopardized by occlusions (points with no counterpart in the other image) and spurious matches (false corresponding pairs created by noise). Two important constraints that reduce the effects of both phenomena are the left-right consistency constraint, which means that we only accept corresponding pairs found matching left-to-right and right-to-left, and epipolar constraint, which has been used in rectification.

2.5 Reconstruction

In the previous steps, we have determined the correspondences between the left and right rectified image points based on the correlation method. We can now reconstruct the 3-D points algebraically as follows.

Let $\begin{pmatrix} c_l' \\ r_l' \\ 1 \end{pmatrix}$ and $\begin{pmatrix} c_r' \\ r_r' \\ 1 \end{pmatrix}$ be the left and right rectified image points respectively. The original pixel coordinates can be recovered from the rectified coordinates: $\begin{pmatrix} c_l' \\ r_l' \\ 1 \end{pmatrix} = W_l' R_{rect} W_l^{-1} \begin{pmatrix} c_l \\ r_l \\ 1 \end{pmatrix}$ and $\begin{pmatrix} c_r' \\ r_r' \\ 1 \end{pmatrix} = W_r' R_{rect} R W_r^{-1} \begin{pmatrix} c_r \\ r_r \\ 1 \end{pmatrix}$. The 3D coordinates can be solved through the perspective projection equation: $l \begin{pmatrix} c \\ r \\ 1 \end{pmatrix} = WM \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$. Combining the above equations, we have:

$$l_l \begin{pmatrix} c_l' \\ r_l' \\ 1 \end{pmatrix} = W_l' R_{rect} W_l^{-1} \begin{pmatrix} c_l \\ r_l \\ 1 \end{pmatrix} = W_l' R_{rect} W_l^{-1} W_l M_l \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = W_l' R_{rect} M_l \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}, \text{ and}$$

$$l_r \begin{pmatrix} c_r' \\ r_r' \\ 1 \end{pmatrix} = W_r' R_{rect} W_r^{-1} \begin{pmatrix} c_r \\ r_r \\ 1 \end{pmatrix} = W_r' R_{rect} R W_r^{-1} W_r M_r \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = W_r' R_{rect} R M_r \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}.$$

There are five unknowns $(x,\ y,\ z,\ l_l,\ l_r)$ and 6 equations; the solution can be obtained by the least-squares method.

Let $P_l = W_l' R_{rect} M_l$ and $P_r = W_r' R_{rect} R M_r$

$$P_l \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} - ?_l \begin{pmatrix} c_l \\ r_l \\ 1 \end{pmatrix} = \begin{pmatrix} P_{l11} & P_{l12} & P_{l13} & P_{l14} \\ P_{l21} & P_{l22} & P_{l23} & P_{l24} \\ P_{l31} & P_{l32} & P_{l33} & P_{l34} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} - ?_l \begin{pmatrix} c_l \\ r_l \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$P_r\begin{pmatrix}x\\y\\z\\1\end{pmatrix} - ?_r\begin{pmatrix}c_r\\r_r\\1\end{pmatrix} = \begin{pmatrix}P_{r11} & P_{r12} & P_{r13} & P_{r14}\\P_{r21} & P_{r22} & P_{r23} & P_{r24}\\P_{r31} & P_{r32} & P_{r33} & P_{r34}\end{pmatrix}\begin{pmatrix}x\\y\\z\\1\end{pmatrix} - ?_r\begin{pmatrix}c_r\\r_r\\1\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$$

The combination of the two equation systems yields

$$\begin{pmatrix}P_{l11} & P_{l12} & P_{l13} & -c_l & 0\\P_{l21} & P_{l22} & P_{l23} & -r_l & 0\\P_{l31} & P_{l32} & P_{l33} & -1 & 0\\P_{r11} & P_{r12} & P_{r13} & 0 & -c_r\\P_{r21} & P_{r22} & P_{r23} & 0 & -r_r\\P_{r31} & P_{r32} & P_{r33} & 0 & -1\end{pmatrix}\begin{pmatrix}x\\y\\z\\l_l\\l_r\end{pmatrix} = \begin{pmatrix}-P_{l14}\\-P_{l24}\\-P_{l34}\\-P_{r14}\\-P_{r24}\\-P_{r34}\end{pmatrix} \qquad (2\text{-}6)$$

The least-squares solution of the linear system $AX = B$ is given by

$$X = (A^T A)^{-1} A^T B$$

3. Experimental procedure and results

In this section, we present the major steps of the experimental procedure that leads to the 3D reconstruction of a human face. Each step is followed by the experimental result.

3.1 Camera calibration

We use the pictures shown in Figure 1 to perform camera calibration. We manually measure the 3D coordinates of the corners of each square, and we also manually measure the 2D pixel coordinates of the corners of each square in left and right images.

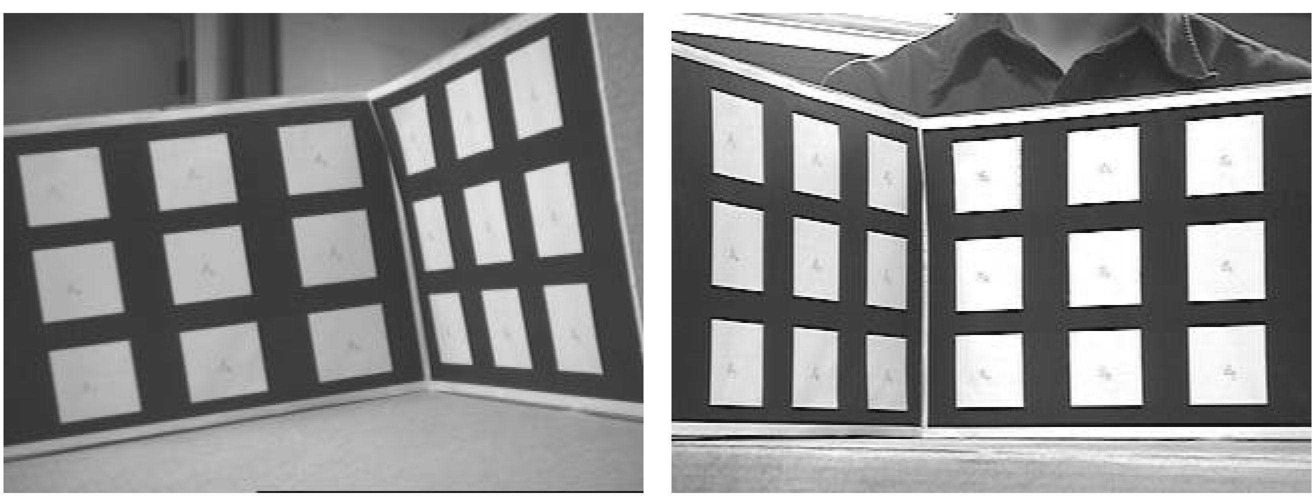

Figure 1: Camera calibration patterns

Using the procedure introduced in section 2.1 and equations (2-1) and (2-2), we get the calibration results. They are listed in Table 1 in the format of intrinsic matrix W and external matrix M.

Table 1 Camera calibration results

$$W_L = \begin{pmatrix} 663.10 & 0 & 139.66 \\ 0 & -602.19 & 121.67 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_L = \begin{pmatrix} -0.9558 & 0.2528 & -0.1505 & 9.3397 \\ -0.1398 & 0.0598 & 0.9884 & -8.5706 \\ -0.2588 & -0.9657 & 0.0218 & 85.5288 \end{pmatrix}$$

$$W_R = \begin{pmatrix} 818.68 & 0 & 114.54 \\ 0 & -803.28 & 166.65 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_R = \begin{pmatrix} -0.5105 & 0.8599 & -0.0024 & 1.5299 \\ 0.0413 & 0.0273 & 0.9988 & -5.2457 \\ -0.8589 & -0.5098 & 0.0495 & 110.0386 \end{pmatrix}$$

3.2 Relative rotation and translation between two cameras

Following the formula (2-3), we computed the relative orientation and translation matrices between the two cameras:

$$R = \begin{pmatrix} 0.7056 & -0.1829 & 0.6846 \\ 0.1204 & 0.9830 & 0.1385 \\ -0.6983 & -0.0153 & 0.7156 \end{pmatrix} \text{ and } T = \begin{pmatrix} -68.0329 & -18.8390 & 7.7704 \end{pmatrix}.$$

3.3 Rectification

We understand that in the intrinsic matrix, $s_x f$ and $s_y f$ only change the focal length, so they scale the images, while c_0 and r_0 shift (translation) the images. We only need to set $s_x f$ and $s_y f$ of the new intrinsic matrices for left and right cameras the same in order to make the rectified left and right images in the same scale, and we can freely choose c_0 to horizontally shift the rectified images such that we can view the most part of the scene.

After several tries, we set the new intrinsic matrices as:

$$W_L' = \begin{pmatrix} 663.10 & 0 & 209.66 \\ 0 & -602.19 & 121.67 \\ 0 & 0 & 1 \end{pmatrix}, \; W_R' = \begin{pmatrix} 663.10 & 0 & -390.34 \\ 0 & -602.19 & 121.67 \\ 0 & 0 & 1 \end{pmatrix}.$$

Actually, we use the left camera's focal length, and we choose the values of c_0 in order to view the most part of the scene.

Following the procedure introduced in section 2.4 and equation (2-4), we obtain the rectified left and right images for both calibration pattern and human face as shown in

Figure 2. The first row shows the rectified images of calibration pattern, the second row shows the original images of human face, and the last row shows the rectified images of human face. We can easily verify that the corresponding points are all on the same row.

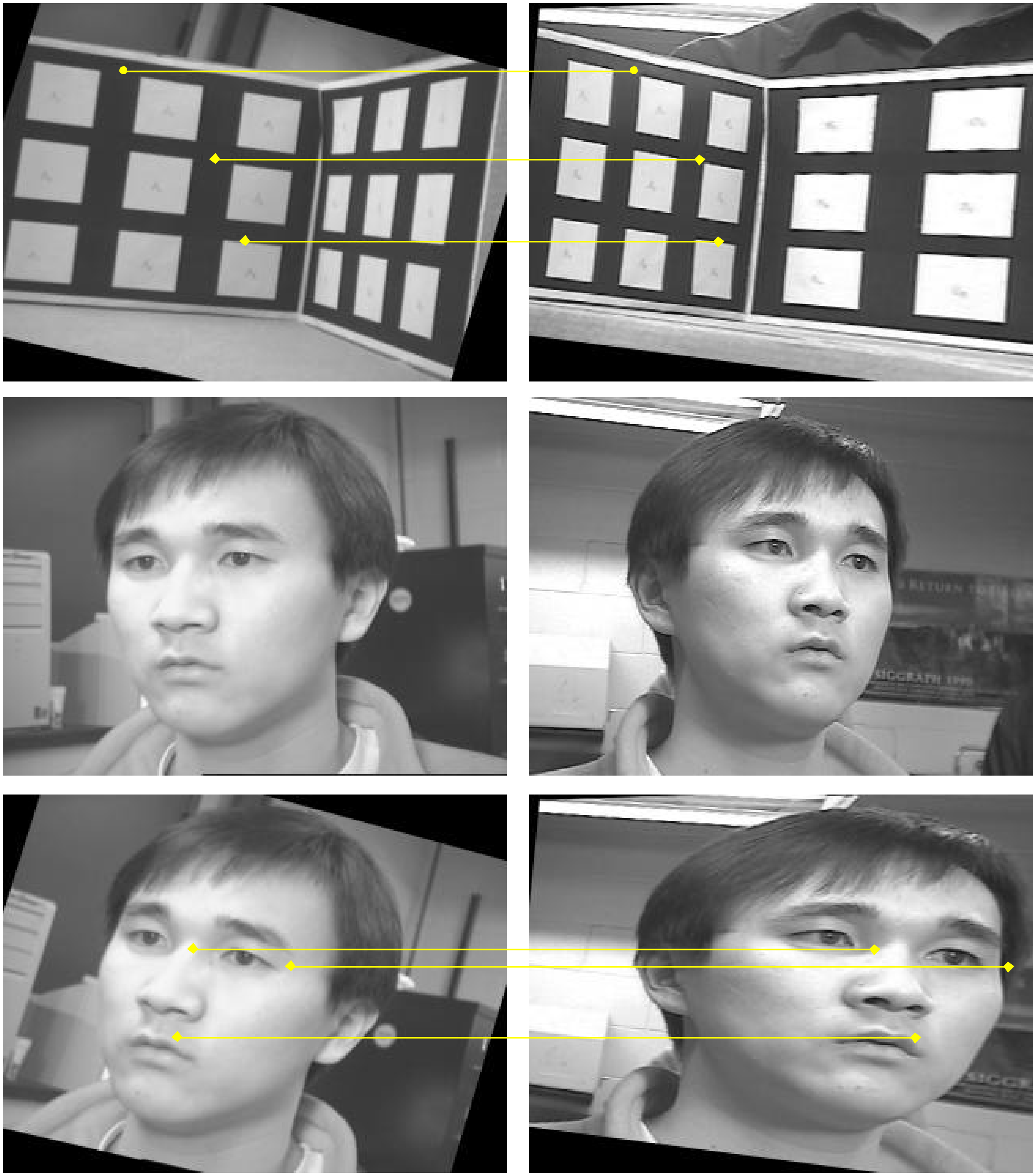

Figure 2: Rectification results

3.4 Establishing correspondences

The SSD correlation method is used to establish correspondences with a window size of 5-by-5 pixels. Two constraints are imposed to reduce the number of spurious matching points. The first constraint is defined on the correlation value, which consists of rejecting lowly correlated points if the correlation value falls below a certain threshold. The second one is based on the left-right consistency constraint. Only corresponding pairs found matching left-to-right and right-to-left are accepted. In practice, a discrepancy of one or two pixels is tolerated to satisfy the consistency constraint. Once the correlated pair is located on the left and right images, we compute the disparity between the two matching points. If a corresponding cannot be found, the disparity value is simply set to 0. Figure 3 illustrates disparity maps under different threshold settings. The top three figures represent the results obtained by defining the correlation threshold at –1,000, -10,000 and –100,000 (from left to right). The bottom two disparity maps are the results generated from 0-pixel (left image) and 1-pixel (right image) tolerance in consistency constraint without any threshold on correlation value. Notice that the face features become more discernable by decreasing the correlation threshold value. This is due to the large differences in brightness and contrast between the original face images. If the correlation threshold is set too high, some correct matching pairs may be rejected, which results in a loss of information during the 3D reconstruction process. On the other hand, we observe that disparity map generated with 0-pixel consistency tolerance includes less matching points than the one with 1-pixel tolerance. It has the advantage of containing fewer noises in the image.

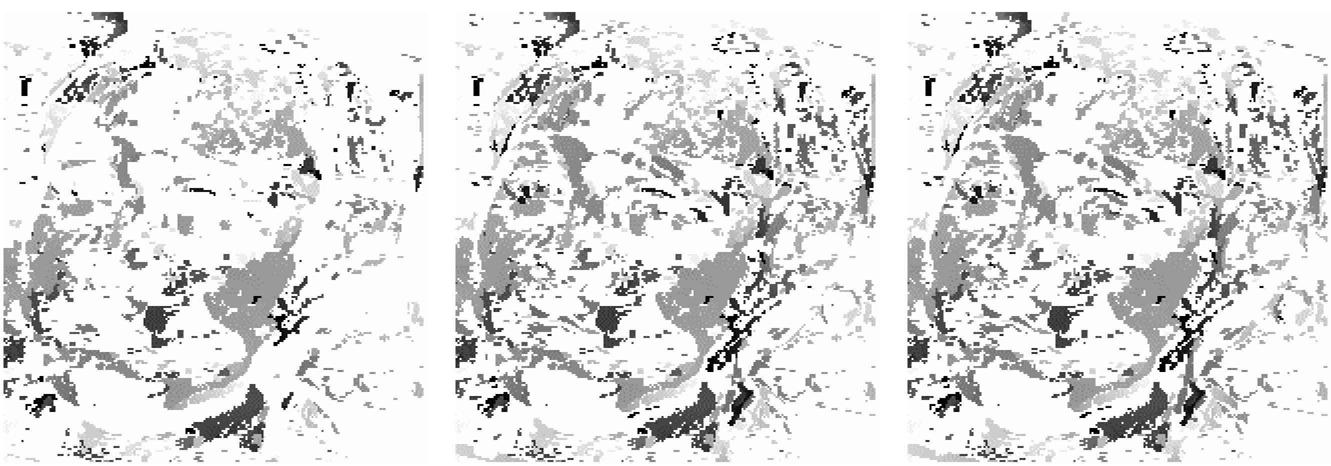

Figure 3a. Disparity maps with different correlation thresholds
(From left to right: -1,000, -10,000, and -100,000, with 1 pixel consistency constraint)

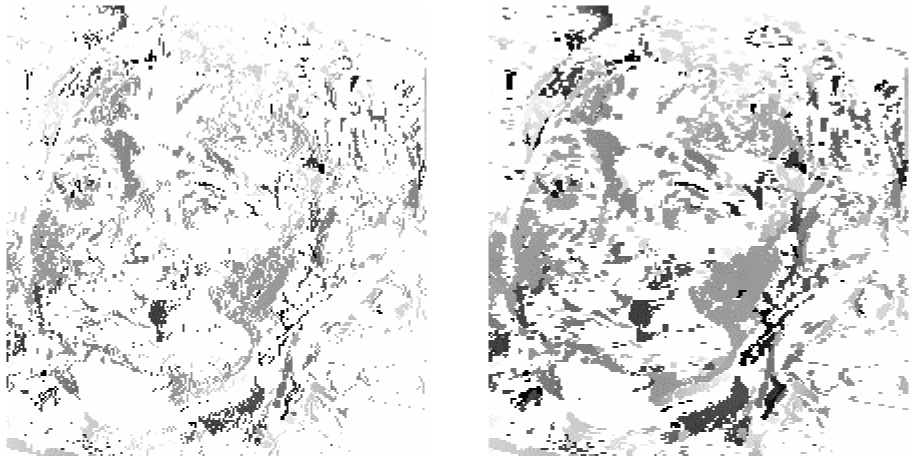

Figure 3b. Disparity maps with 0 (left) and 1 (right) pixel tolerance in consistency constraint without correlation threshold

3.5 Reconstruction

Following the equation (2-6), we reconstructed the 3D shape of the human face as shown in Figure 4. The 3D face was obtained using the corresponding points found with 0-pixel consistency constraint and no correlation threshold. Due to the noises in the background, which superpose with the points in the front, it is difficult to recognize the human face quite clearly. Moreover, the correlation method is unable to work very well with significant differences in brightness and contrast between the left and right images. Nevertheless, we can roughly distinguish some face features such as face contour, eyes, nose ridge, and mouth.

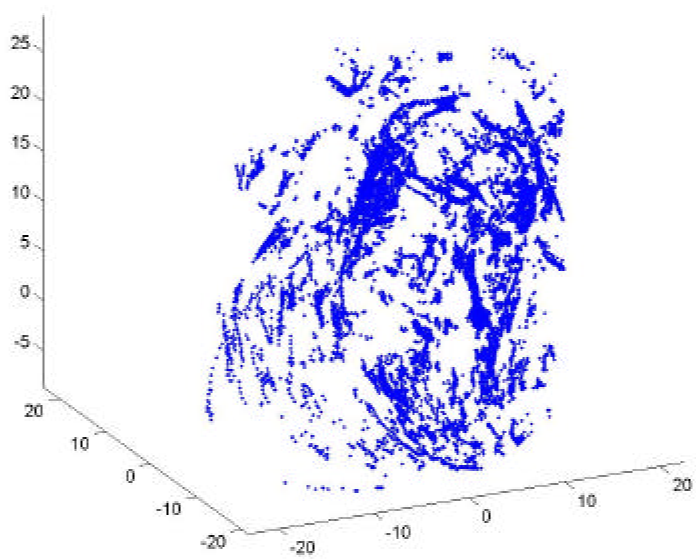

Figure 4 Three-dimensional reconstruction of the human face

4. Summary

In this project, we have applied the 3D reconstruction technique to restore the 3D scene of a human face from two images taken from different viewpoints. The major steps consisted of camera calibration, computation of relative orientation and position between the two cameras, and image rectification in order to facilitate the point matching procedure with the correlation-based method. The disparity maps were displayed to identify the face features from the images. The 3D coordinates of the face were algebraically reconstructed from the corresponding points on left and right rectified images.

The main contribution of this project was the generalization of the rectification method, which can be applied to rectify images taken by cameras with different intrinsic parameters. The mathematical formulation has been derived and the rectification results proved that the method was successful. On the other hand, we tried to improve the correlation performance by adding left-right consistency constraint in addition to correlation threshold. Unfortunately, the differences in brightness and contrast were too significant so that the addition of the consistency constraint was not sufficient to completely eliminate the noises. The disparity maps and reconstruction results were therefore not optimal. However, we were still able to recognize some face features from the 3D plot.