

OBTAINING 3-D SHAPE FROM SILHOUETTE INFORMATIONS INTERPOLATED BY PHOTOMETRIC STEREO

Changsuk Cho, Haruyuki Minamitani

Minamitani Lab. Institute of Biomedical Engineering
Faculty of Science and Technology, Keio University
3-14-1, Hiyoshi, Kohoku-Ku, Yokohama, Japan

ABSTRACT

The idea of photometric stereo is to serially vary the direction of incident illumination on the state of the holded view point, but it can locally reconstruct only the front surface to a view point. However the silhouette informations obtained by plural viewing directions provide the entire 3D structure of an object surface, but the 3D information of the region whose edge is unexposed to the viewing direction cannot be obtained. Therefore the two methods make up the demerits each other. In this paper, the possibility of complete three dimensional shape reconstruction was investigated using photometric stereo and silhouette informations. For the entire 3D reconstruction, a chicken bone is used as an experimental object and the reconstructed result is presented and discussed.

INTRODUCTION

In accordance with the development of computerized automatic manufacturing, more and more increases the need for sensing and reconstructing 3 dimensional structure. Many traditional method for 3D sensing and reconstruction thereby have been developed and presented in the field of computer vision and sensing, such as the methods scanning slit ray beam [1], interpreting the patterns projected [2][3] and using stereo pair images by 2 cameras on the basis of principle of triangulation [4][5].

The methods using slit ray beam have the merit of which can produce precise result but it needs the complex and expensive equipments comparing the proposed method of this paper and has the difficulty to match the discreted slit lines on a very indented irregular surface. The methods interpreting the projected patterns such as binary encoded light pattern, grid pattern and the moire method offer the stable reconstructed results but the resolution depends on the width of a projected pattern or a contour line, especially in the reconstruction using images obtained CCD camera the precision goes down. It is caused by the spacing between patterns or contours, which is needed for interpretation. Also it has the difficulty of interpretation caused by the discrete lines. On

the other hand, the stereo pair matching by two camera is not a completed method because its running cost and the unstable matching algorithm.

The photometric stereo method [6][7][8] can offer the continuous depth informations of surface as pixel unit, so that it can level up the precision locally in accordance with the magnification of lenz of CCD camera, and needs the simple equipments comparing other methods. But its demerit is that it has the possibility to get a twisted result due to specular reflection and the cumulation process for depth conversion. Otherwise the silhouette information can offer a stable and precise result only to the edges exposed to viewer [9]. It means that the twisted result obtained by the photometric stereo can be corrected by the silhouette informations. It is the reason we use the photometric stereo and silhouette information.

SILHOUETTE

The silhouette information from a view point includes partly the 3D information of an object as an edge information of 2D projected image. Therefore the silhouette informations obtained by plural viewing directions include 3D informations reconstructed by the edges exposed to the viewing points.

The line of sight passing through the edge of silhouette is othgonal to the surface normal of an object if it has smooth surface. Assuming a point $(X(\theta, y_e), Z(\theta, y_e))$ represents an edge point of the silhouette rotated with viewing direction of azimuth θ on an height y_e , and (x_e, z_e) represent an edge point of silhouette on image plane, the point can be computed as

$$\begin{aligned} X(\theta, y_e) &= x_e \cos(\theta) + z_e \sin(\theta), \\ Z(\theta, y_e) &= z_e \cos(\theta) - x_e \sin(\theta). \end{aligned} \quad (1)$$

Then the function of line of sight $LF(x, z : y_e, \theta)$ to the point $(X(\theta, y_e), Z(\theta, y_e))$ can be computed as,

$$\begin{aligned} LF(x, z : y, \theta) &= \{(x_{lf}, z_{lf}) | x_{lf} = x, \\ & z_{lf} = (-1/\tan(\theta))(x - X(\theta, y_e)) + Z(\theta, y_e)\}. \end{aligned} \quad (2)$$

where x represents all x values on image plane.

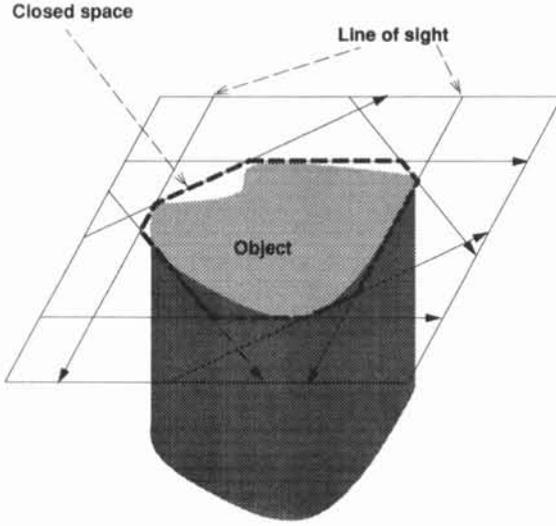


Fig.1. Line of sight function and the space closed by the line of sight functions.

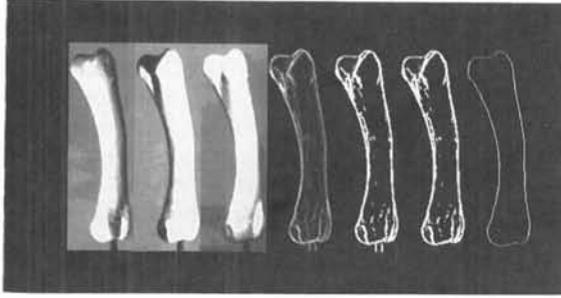


Fig.2. The closed space of slice of the experimental object at each sampled height y .

The silhouette informations obtained from plural viewing directions provide the restraints of inequality function for the 3-D shape reconstructed by photometric stereo as shown in Fig.1.

Let suppose the closed space $\Phi(y)$ composed of the set of lines of sight on a height y on the condition that the object was projected from plural viewing direction. And let $P(x, z : y, \theta)$ represent the point of the 3-D shape reconstructed by photometric stereo on a height y on a rotating degree θ , then vall the $P(x, z : y, \theta)$ must satisfy the following condition:

$$\Phi(y) \ni P(x, z : y, \theta). \quad (3)$$

Fig.2 shows the closed space $\Phi(y)$ of the experimental object projected by plural viewing directions with 10 degree interval on the sampled heights y .

LOCAL 3D SURFACE

Fig.3 shows the geometry of photometric stereo. The relation among the intensity I obtained by CCD camera

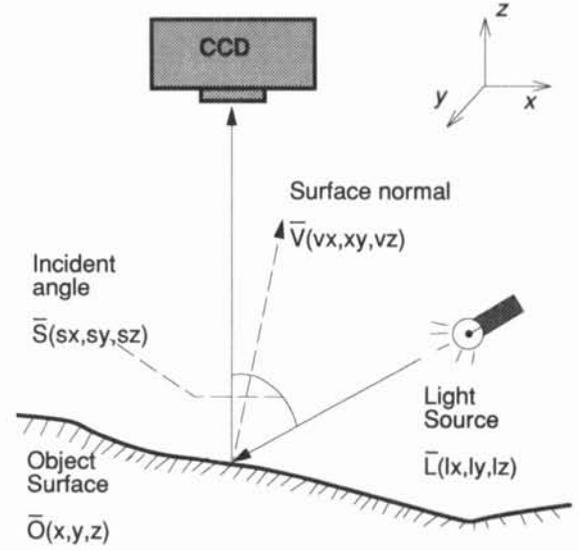


Fig.3. The geometry of photometric stereo.

and the normal vector \bar{V} on a point of object surface and the incident light vector \bar{S} regarding to an illumination direction was introduced in the reference [8][6] as shown in the following. Here \bar{V} , \bar{S} are supposed as unit vectors.

$$I = R(\bar{S} \cdot \bar{V}), \quad (4)$$

where R represents the reflectance factor of the object surface. In the case of 3 views, (4) can be written as

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = R \begin{bmatrix} s_{x1} & s_{y1} & s_{z1} \\ s_{x2} & s_{y2} & s_{z2} \\ s_{x3} & s_{y3} & s_{z3} \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} \quad (5)$$

where x, y, z represent x - y - z coordinates, and 1,2,3 the counter of illuminations. Denoting $[S]$ the incident light matrix, \bar{I} the intensity feature vector, (5) becomes

$$\bar{I} = R[S]\bar{V}. \quad (6)$$

Assuming that the incident vector is known, the reflectance factor and the surface normal can be calculated as

$$R = |[S]^{-1}\bar{I}|, \quad (7)$$

$$\bar{V} = R^{-1}([S]^{-1}\bar{I}). \quad (8)$$

The depth value z can be obtained by integrating tangent vectors of the surface normal[6][7]. Considering a surface normal (v_x, v_y, v_z) at the point (a, b) on the image plane, then the tangent vector is given as

$$T(a, b) = (-v_x/v_z, -v_y/v_z), \quad (9)$$

and the depth $Z(a, b)$ can be obtained as

$$Z(a, b) = \int^b \int^a T(x, y) dx dy \quad (10)$$

If the parallel light source is used as incident light, it is easy to obtain the incident light angle information[6][8].

But it is difficult to use parallel light to the big object, we used the point light source. For obtaining the incident light angle information, the estimation algorithm of the incident light angle of point light source is used[7].

The incident light angle of point light source can be computed by the relation between the surface orientation and the light source orientation. Let \vec{O} represent the surface coordinate vector with elements (x, y, z) and \vec{L} represent the light source coordinate vector with elements (l_x, l_y, l_z) , then the incident light angle can be computed by computing the unit vector \vec{S} . The x and y can be obtained from image plane, but the depth information z is unknown. Because of the unknown z , we estimated the incident light angle by iterative process. Let t represent the iteration number for estimation of the incident light angle, then this process can be shown as follows.

- STEP 0: Setting the value of $Z(x, y)$ to be zero.
- STEP 1: Computing the incident light angle.
 $\vec{S}_t = (\vec{L} - \vec{O}_{t-1})$.
- STEP 2: Computing the surface normal.
 $\vec{V}_t = R^{-1}([\vec{S}]_t \vec{I})$.
- STEP 3: Depth conversion of equation (10).
- STEP 4: IF $\{ |Z(x, y)_{t-1} - Z(x, y)_t| < \varepsilon, \varepsilon \rightarrow 0 \}$
 THEN stop ELSE go to STEP 1.

COMBINING SURFACES

A silhouette information at a viewing direction can be extracted by differentiating the average image of the three images digitized at each incident angle of illumination. Fig.4 shows the extracted silhouette. At each viewing direction, a local 3-D shape of surface is reconstructed using photometric stereo.

Let $\Omega(\theta_i)$ represent the local 3-D data set reconstructed by the photometric stereo at viewing degree θ_i . The unnecessary background information of the $\Omega(\theta_i)$ can be removed by silhouette information as shown in Fig.5.

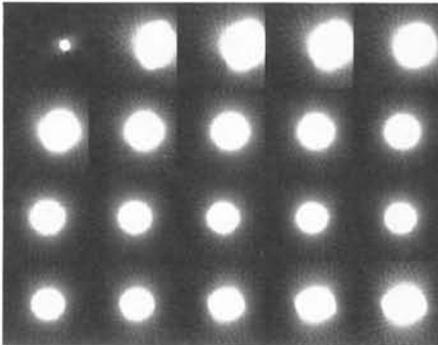


Fig.4. Extraction of silhouette.

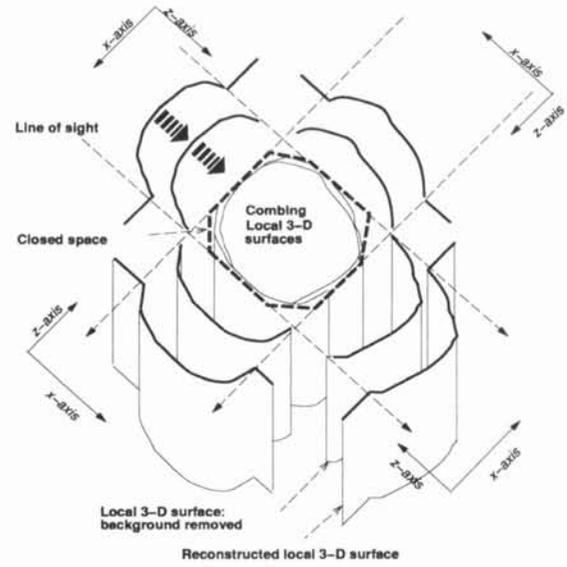


Fig.5. Combining the local 3-D shapes reconstructed by photometric stereo and silhouettes.

The each $\Omega(\theta_i)$ must be rotated mathematically at its viewing direction θ_i for coincidence of x - y - z coordinates. If $\hat{\Omega}(\theta_i)$ symbolizes the data set whose background is erased, and $\hat{\Omega}_r(\theta_i)$ the rotated data set, $\hat{\Omega}_r(\theta_i)$ is computed as follow:

$$\hat{\Omega}_r(\theta_i) = \{(x_r, y_r, z_r) | x_r = x \cos \theta_i + z \sin(\theta_i), \\ y_r = y, z_r = z \cos \theta_i - x \sin \theta_i, x, y, z \in \Omega(\theta_i)\} \quad (11)$$

$\hat{\Omega}_r(\theta_i)$ are positioned for being abutted to the closed space. Let $\hat{\Omega}_a(\theta_i)$ symbolize this data set positioned. A point on the object surface has plural estimates reconstructed locally as seen in Fig.5, the weighted average of (12) which considers the center points of viewed image as more important is used for integrating the estimates. The Ω_{com} in (12) symbolizes the complete 3-D data set reconstructed and (x_c, z_c) represent the point of tangency between the closed space and the $\hat{\Omega}_a(\theta_i)$.

$$D_x = \max(X(\theta, y_e)) - \min(X(\theta, y_e)) \\ D_z = \max(Z(\theta, y_e)) - \min(Z(\theta, y_e)) \\ w_i = 1 - \frac{\sqrt{(x_r - x_c)^2 + (z_r - z_c)^2}}{\sqrt{D_x^2 + D_z^2} / 2}, x_r, z_r \in \hat{\Omega}_a(\theta_i) \\ \Omega_{com} = \left\{ (\hat{x}, \hat{y}, \hat{z}) | \hat{x} = \frac{\sum^i w_i x_i}{\sum^i w_i}, \hat{y} = y, \hat{z} = \frac{\sum^i w_i z_i}{\sum^i w_i}, \right. \\ \left. x, y, z \in \hat{\Omega}_a(\theta_i) \right\} \quad (12)$$

where i is the number of sampled viewing direction.

RESULT

The surface of bone was painted by white water color for reducing the specular reflection[7]. As incident illumination, a 40 watt tungsten bulb whose glass is transparent is used. Fig.6 shows the experimental setup. The

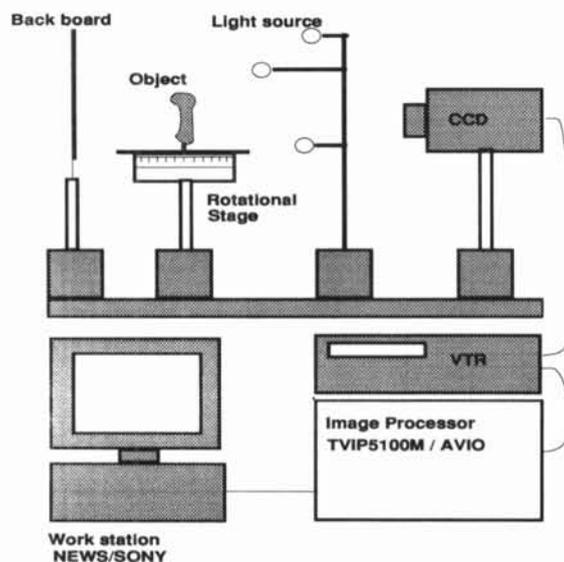


Fig.6. The experimental setup.

object mounted on the rotational plane was serially illuminated by the bulb at three incident angles on each viewing direction. We sampled 36 viewing directions at 10 degree interval from all direction on the rotational plane and digitized three images by CCD camera varying serially the direction of incident illumination on the state of the holded view point on each viewing direction. The distance between the camera and the object is approximately 60cm and height of the bone is about 6.8cm. We assume the projection is othogonal for symplifying the geometry because the perspective projection can be approximated as an othographic projection on the condition that the distance between the viewing point and the object is large in relation to the object size.

The complete 3-D shapes of bone reconstructed by the proposed algorith in Fig.7 show reasonable results and confirmed the possiblity of complete 3-D reconstruction by the proposed method. The proposed method can be applied to the field of artificial organ as like bone and model for molding. For the application, use of replica is effective for erasing the error from specular reflection. We are enhancing the proposed algorithm to manage the object having several isolated silhouettes and investigating the relation between the precision of the reconstructed result and the number of viewing direction.

REFERENCES

- [1] K.Nakano, Y.Watanabe, S.Kanno, "Extraction and recognition of 3-dimensional information by projecting a pair of slit-ray beams." *Proc.IEEE Int.Conf. Pattern Recognition*, Rome,Italy,736-738,1988.
- [2] M.Tanaka, Y.Kawaura, K.Omura, M.Sasaki, T.Iwa, Y.Munemoto, K.Matsuura, M.Horita, S.Taniguchi,

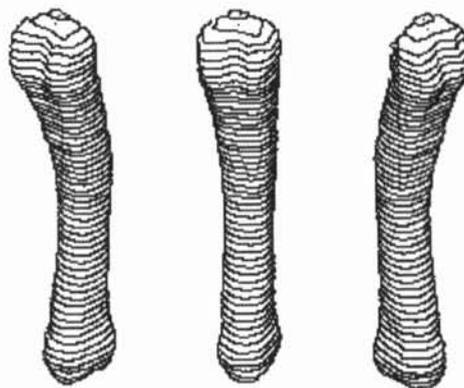


Fig.7. The complete 3-D shape of bone reconstructed by the proposed method.

- "TV endoscopic 3-dimensional object profilometry with projection and grid pattern." *Gastroentero. Endoscopy*(in Japanese),vol.32,1119-1132,1990.
- [3] P.Vuylsteke,A.Oosterlinck, "Range image acquisition with a single binary-encoded light pattern." *IEEE Trans.PAMI*,vol.12,148-163,1990.
- [4] E.Badiqué,Y.Komiya,N.Ohyama,T.Honda, J.Tsujiuchi,"Use of color image correlation in the retrieval of gastric surface topography by endoscopic stereopair matching." *Appl.Opt.*,vol.27,941-948,1988.
- [5] N.H.Kim,A.C.Bovik,S.J.Aggarwal, "Shape description of biological objects via stereo light microscopy." *IEEE Trans. Syst. Man & Cybern.*, vol.20, 475-489, 1990.
- [6] E.N.Coleman Jr.,R.Jain, "Obtaining 3 dimensional shape of textured and specular surface using 4 source photometry." *CGIP*, vol.18,309-328,1982.
- [7] C.S.Cho, H.Minamitani, "A new photometric method using 3 point light sources." *IEICE Trans.*, vol.E76-D, 898-904, 1993.
- [8] R.J.Woodham, "Photometric method for determining surface orientation from multiple images." *Opt.Engin.*, vol.19,139-144,1980.
- [9] J.Y.Zheng,F.Kishino, "Recovering 3D models from silhouette sequence and detecting unexposed regions." *IEICE Trans.*(in japanese), vol.J76-D-2, 1114-1122,1993.