EXTRACTION OF SURFACE ORIENTATION USING GRAY LEVEL DIFFERENCE STATISTICS

Mutsuhiro Terauchi, Hidegi Matsushima, Toshio Tsuji and Koji Ito

Faculty of Engineering, Hiroshima University
1-4-1 Kagamiyama, Higashi-hiroshima 724, Japan

INTRODUCTION

The Processes to reconstruct a 3D shape from a 2D image is one of the important problems in computer vision. In this paper we deal with the problem to extract the object surface orientation from a monocular view image, which is necessary in 3D reconstruction. Generally the process becomes ill-posed problem, because the 3D shape of an object is condensed onto the image by the projection. Therefore the solution of the orientation is not guaranteed to be unique, unless some supplement information is introduced about the object or the surface.

In case that there exists a coherent texture on the object surface, it is possible to utilize it for 3D reconstruction. In the last decade various methods are proposed in *Shape from Texture*. Assumed that a texel, which constitutes texture, is distributed equally on textured surface, Gibson proposed the procedure to estimate the orientation of the surface of the object using the distortion of texel density caused by perspective projection [1]. Ohta computed the orientation of the object surface by estimating vanishing points from the area of texel and its distortion caused by perspective projection [2]. These methods are based on a condition that the structure of texel has been already known. It is, however, not easy to extract a particular texel on the surface.

On the other hand, Witkin directed his attention to the direction of edge in texel. He estimated the surface orientation from the probability density function of edge directions under the parallel projection on condition that the directions of edges on the object surface are assumed to be equally distributed to all directions [3], [4]. Furthermore Aloimonos gave attention to 'edge element', which is components of texel, [5] and extracted the surface orientation from the distortion of the total length of edges in each local region on the image.

The above methods, however, are inapplicable to the texture in which no definite structure can be found. Then, it is required to estimate the surface orientation directly from the gray levels of the original image. This paper proposes the method to estimate the surface orientation from the difference statistics of the gray levels, which depends on the distance and direction on the image. On the basis of the assumption that any region on a textured plane in the 3D world has the same probability density function of the difference statistics, we formulate the relation between the distortion of the probability density function and the surface orientation under perspective projection. We then derive an algorithm for estimating the orientation of the textured plane from a static monocular view. Finally, some experimental results are shown.

Visual Form, Edited by C. Arcelli *et al.*
Plenum Press, New York

GEOMETRIC RELATIONS BETWEEN THE OBJECT PLANE AND THE IMAGE PLANE

Coordinate systems and perspective projection

The 3D coordinates system $O\text{-}XYZ$ as shown in Fig.1 is introduced to formulate the perspective projection, where its origin is assumed to be the optical center of the lens. The image plane $o\text{-}uv$ is placed orthogonal to Z axis (sight axis) and its origin o at $(0,0,f)$, where f is the focal length. The distance between the origin O and object plane P is set to E.The orthogonal coordinates plane $O'\text{-}UV$ is placed parallel to the image plane and its origin at $(0,0,E)$, where the axis U on the plane is parallel to the axis u on the image plane. Therefore the surface normal unit vector w is uniquely defined by α and β as $w = (\sin\alpha\cos\beta, \sin\beta, \cos\alpha\cos\beta)$. The equation of the plane P is described as follows,

$$bcX + dY - ac\,(Z - E) = 0 \quad , \tag{1}$$

where,

$$a = \cos\alpha,\ b = \sin\alpha,\ c = \cos\beta,\ d = \sin\beta\,.$$

A problem for estimating the orientation of the object surface is equivalent to computing the surface normal, i.e. the angle α and β.

When the focal length is set as $f = 1$, the relation between the object plane $P(U,V)$ and the image plane $I(u,v)$ results,

$$U = \frac{E\,(cu + bdv)}{ac - bcu - dv}\ , \quad V = \frac{Eav}{ac - bcu - dv}\ . \tag{2}$$

The relation (2) implies point wise correspondence between two on each plane. Next we consider the relation of two line segments on both planes.

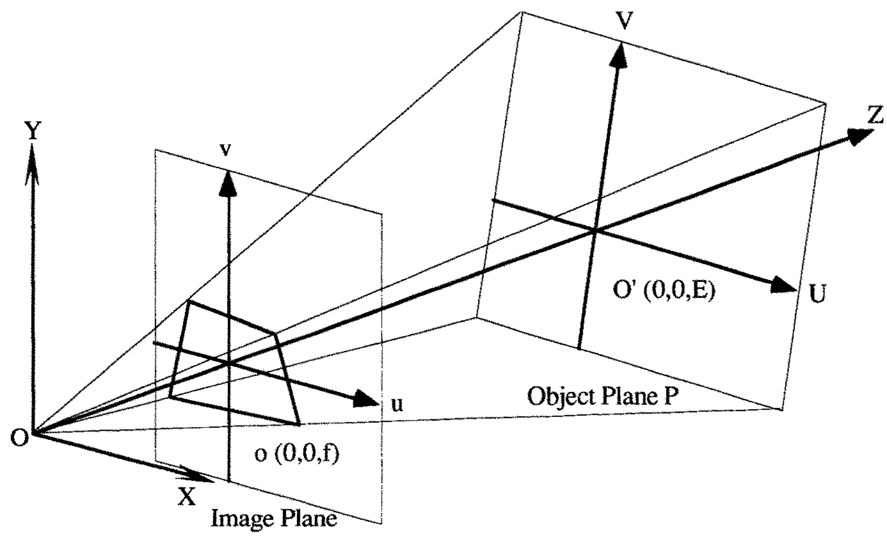


Fig. 1 The coordinate systems and the perspective projection.

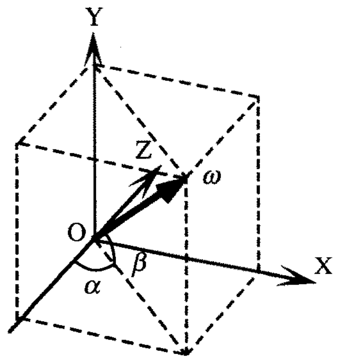


Fig. 2 The surface normal unit vector ω.

The length and direction of line segments on both planes

If a straight line on the object surface is projected onto the image under the perspective projection, then the length and direction are distorted. Let's here assume a line whose length is L, direction θ' and the terminal point at (U,V) on the object surface. Then the projected line is assumed to be the length l, direction θ and the terminal point at (u,v). The relations between L and l, θ' and θ are given as follows,

$$L = \frac{Eac}{\Big[(ac\text{-}bcu\text{-}dv)\{ac\text{-}bc(u + l\cos\theta) - d(v + l\sin\theta)\}}$$

$$\cdot\sqrt{\{(c\text{-}adv)\cos\theta + d(b{+}au)\sin\theta\}^2 + \{bv\cos\theta + (a\text{-}bu)\sin\theta\}^2}\cdot l\Big], \quad (3)$$

$$\theta = \tan^{-1}\left(\frac{bv\cos\theta + (a\text{-}bu)\sin\theta}{(c\text{-}adv)\cos\theta + d(b{+}au)\sin\theta}\right) \quad . \quad (4)$$

Next, as shown in Fig.3, let's give two parallel line segments with each terminal point (Um,Vm) and (Un,Vn) on the object surface plane, and also their projected lines with the terminal points (um,vm) and (un,vn) on the image plane. Generally, although $lm \neq ln$ and $\theta m \neq \theta n$ in terms of two lines on the image plane, the following relations are held from Eqs.(3) and (4).

$$[\{(c\text{-}adv_m)\cos\theta_m + d(b{+}au_m)\sin\theta_m\}^2 + \{bv_m\cos\theta_m + (a\text{-}bu_m)\sin\theta_m\}^2]^{1/2}\cdot l_m$$
$$\cdot[(ac\text{-}bcu_m\text{-}dv_m)\cdot\{ac\text{-}bc(u_m + l_m\cos\theta_m) - d(v_m + l_m\sin\theta)\}]^{-1}$$
$$= [\{(c\text{-}adv_n)\cos\theta_n + d(b{+}au_n)\sin\theta_n\}^2 + \{bv_n\cos\theta_n + (a\text{-}bu_n)\sin\theta_n\}^2]^{1/2}\cdot l_n$$
$$\cdot[(\mathrm{ac\text{-}bcu_n\text{-}dv_n})\cdot\{ac\text{-}bc(u_n{+}l_n\cos\theta_n) - d(v_n{+}l_n\sin\theta_n)\}]^{-1} \quad (5)$$

$$\frac{bv_m\cos\theta_m + (a\text{-}bu_m)\sin\theta_m}{(c\text{-}adv_m)\cos\theta_m + d(b{+}au_m)\sin\theta_m} = \frac{bv_n\cos\theta_n + (a\text{-}bu_n)\sin\theta_n}{(c\text{-}adv_n)\cos\theta_n + d(b{+}au_n)\sin\theta_n} \quad (6)$$

If such a pair of lines are found on the image plane, we can compute the surface normal angles α, β from Eqs.(5) and (6) using two pairs of lines. In this paper, however, we do not assume to use any structural information of texel such as lines or contours etc. Therefore we use the gray level difference statistics which is defined in terms of distance and direction on the image.

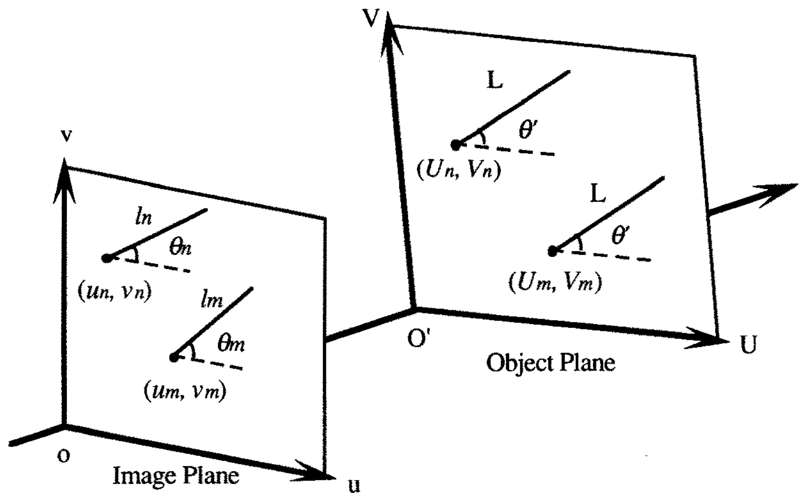


Fig. 3 The constraint conditions for the distance and direction on the image plane.

ESTIMATION OF SURFACE ORIENTATION USING GRAY LEVEL DIFFERENCE STATISTICS

Idea of metaphor

In our research the main purpose is to estimate the orientation of the object plane. In general an object plane is projected onto an image which is deformed under the perspective projection. Furthermore we assume the texture as the pattern represented by probability density function (PDS), here we introduce the gray level difference statistics which is the representation like as the polar coordinates. And to make it simple we call the PDS as "pattern." The pattern deformation in the image is caused by the deformation of the coordinates projected onto the image. If we could know how the coordinates or the pattern were deformed, we could obtain the orientation of the object surface using the geometry of projection. However the coordinates and the pattern on original object plane are not given. Therefore it is necessary to use more than two patterns and to make them related. The deformation of the pattern is according to the position on the image. In our study the probability density function of gray level difference statistics, i.e. the pattern is assumed to be the same everywhere on the object plane. By using the assumption and the geometry of the projection we can obtain the orientation of object plane.

Method of orientation extraction

The gray level difference statistics is a probability density function $P(K|l,\theta)$ of absolute value K of intensity difference between a point and another point with a distance l and a direction θ from its point. Let's assume that the characteristics of texture is able to be represented by the gray level difference statistics which depends on the distance and direction on the image [6].

Under the condition that the probability density function of gray level difference statistics is the same over all local regions on the object plane, the probability density functions Pm and Pn in local regions m and n hold the following relation.

$$P_m(K\,|L_m,\theta'_m) = P_n(K\,|L_n,\theta'_n) \quad , \tag{7}$$

when, $Lm = Ln$, $\theta'm = \theta'n$ on the object surface . On the other hand, when the object surface are projected onto image plane, the difference statistics is distorted by perspective projection. Therefore, Eq. (7) is not held in any local regions. However, on lm and ln, and θm and θn which satisfy the relations Eqs.(5) and (6), the density functions in the image are equal.

$$P_m(K \mid l_m, \theta_m) = P_n(K \mid l_n, \theta_n) \quad . \tag{8}$$

Therefore, the orientation of the surface plane can be computed if we can find lm, ln, θm and θn satisfying Eq.(8) using a couple of density functions in the local regions on the image. It is however too difficult to solve analytically these distances and directions, because the density function can not be anticipated in advance.

Here we introduce the following cost function:

$$J = \left| P_m(K \mid l_m, \theta_m) - P_n(K \mid l_n, \theta_n) \right| \quad . \tag{9}$$

Then we estimate the surface orientation by searching α and β which minimize the cost J.

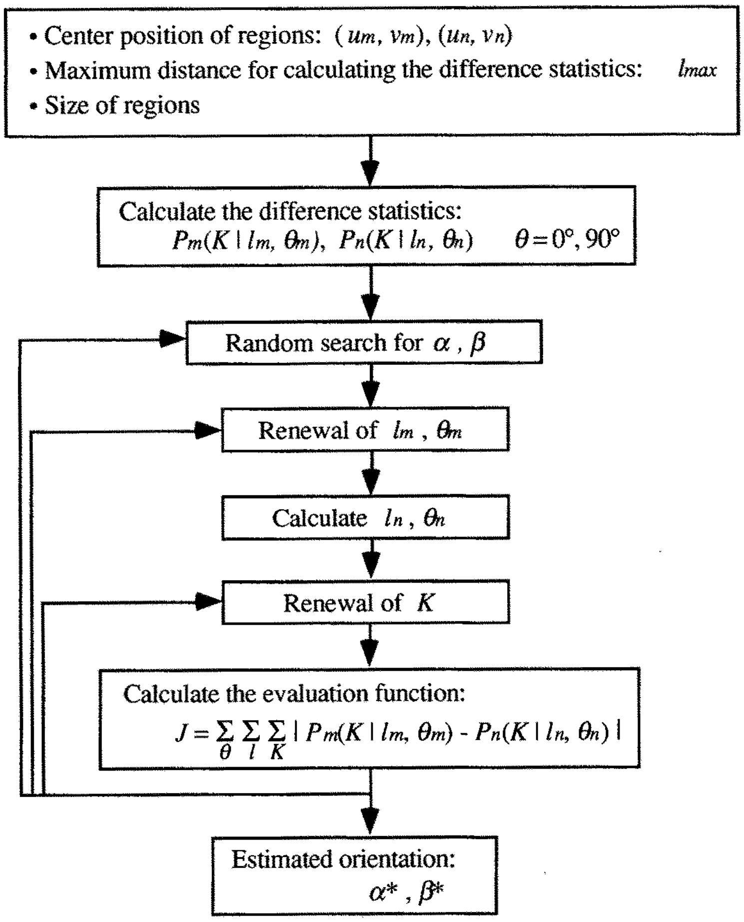


Fig. 4 The algorithm for extraction of a surface orientation.

Orientation Search Algorithm

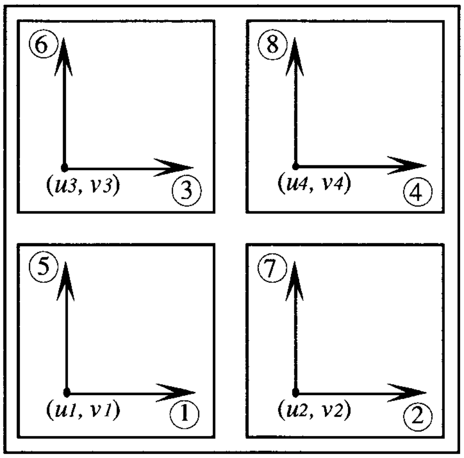
a) The outline of algorithm. The flow of algorithm for extracting surface orientation is shown in Fig.4. First of all, the area and center position of each local region and the maximum distance l_{max} for computing the difference statistics is computed are determined. Then GLDS (gray level difference statistics), distributions $P_m(K|l_m,\theta_m)$ and $P_n(K|l_n,\theta_n)$ for l_m and l_n ($\leq l_{max}$), and $0 \leq \theta_m$ and $\theta_n \leq \pi$ are computed in advance. Next, giving α, β, l_m and θ_m, the corresponding l_n and θ_n are obtained from Eqs.(5) and (6), and then the cost function J is computed from Eq.(9).

There have been, however left some problems in order to apply the method for extracting orientation of object surface using '***real image***'. So we explain how to solve them below.

b) local area matching. In general, an image is constructed of pixels which are sampled discretely, so that it is impossible to compute GLDS for all direction of θ. The obtainable directions are only $\theta = 0$, $1/4\pi$, $2/4\pi$, $3/4\pi$ for satisfiable number of data, i.e. neighboring closely and periodically each other. Therefore we arranged to choose some pairs of local regions as following alignment.

When a straight line on a object plane is projected onto an image, the projected line is still straight line. So the GLDS for the direction on the line which runs through the centers of each local region serves following equality $\theta = \theta_1 = \theta_2$ in Eq.(5). As a result Equation (5), (6) are simplified as shown in below by restricting the relative directions between two local regions to $\theta = 0$, $1/4\pi$, $2/4\pi$, $3/4\pi$ which hold above relation.

$$\frac{l_m}{(ac-bcu_m-dv_m)\cdot\{ac-bc\,(u_m+l_m\cos\theta_m) - d\,(v_m+l_m\sin\theta_m)\}} = \frac{l_n}{(ac-bcu_n-dv_n)\cdot\{ac-bc\,(u_n+l_n\cos\theta_n) - d\,(v_n+l_n\sin\theta_n)\}} \tag{10}$$



1) ① $P_1(K\,|\,l_1, 0°)$ ⟷ ② $P_2(K\,|\,l_2, 0°)$
2) ③ $P_3(K\,|\,l_3, 0°)$ ⟷ ④ $P_4(K\,|\,l_4, 0°)$
3) ⑤ $P_1(K\,|\,l_1, 90°)$ ⟷ ⑥ $P_3(K\,|\,l_3, 90°)$
4) ⑦ $P_2(K\,|\,l_2, 90°)$ ⟷ ⑧ $P_4(K\,|\,l_4, 90°)$

Fig. 5 The corresponding directions between local regions on an image.

As the directions to be computed have already given, one iteration loop for direction θ_m can be canceled, so the cost J is obtained only by renewing the distance l_m. In this experiment four local regions are set on the image as shown in Fig.5 on which GLDSs in two direction $\theta = 0,\ 1/2\,\pi$ are computed to get suficient information for estimation of the surface orientation $\alpha,\ \beta$.

c) interpolation of distribution matrix of difference statistics. The GLDS $P(K|l,\theta)$ can be obtained only in a pixel distance because of discrete image. Therefore, even if the distance l_m in Eq.(10) could be obtained, GLDS $P_n(K|l_n,\theta)$ can not be always computable. Then the matrix of GLDS $P = K[l][\theta]$ is interpolated using third order natural spline function about its length l.

d) random search method. In order to get the value α and β which minimize the cost J in Eq.(9) we may just compute the costs J for all α and β over a range $-1/2\,\pi \leq \{\alpha,\beta\} \leq 1/2\,\pi$, but it is too wasteful so that an random search algorithm which efficiently finds the optimal value is introduced as follows,

[algorithm]

1) randomly choose 20 sets of α,β in the range between $-1/6\,\pi$ and $1/6\pi$, and compute the cost for each set.
2) decide a center position $(\alpha,\beta)_{min}$ for search which minimizes the cost J out of 20 sets, and let the minimum cost be J_c.
3) compute costs J for 20 sets of α,β in the range s between $-10°$ and $10°$ around the center position $(\alpha,\beta)_{min}$, and let the minimum cost be J_{min}.
4) if $J_{min} < J_c$, then a set $\alpha,\ \beta$ which realizes the cost J_{min} is substituted into the center position $(\alpha,\beta)_{min}$ and jump to 3)
   else reduce the search range: $s = s\text{-}1°$, and jump to 3),
   where, not reduce the range less than $2°$; (the minimum range is $2°$).
5) if the center position were not be replaced in 10 times iteration, then the position is assumed to be an estimate (α^*,β^*).

RESULTS OF EXTRACTION OF ORIENTATION

Our method dedicated above is applied for generated textures and real ones in order to show its performance and effectivity.

Simulated texture Image

The periodical texture surface generated using *sin*-function is shown in Fig.6, in which the wave length is 16 (pixel unit) and the level of intensity is 8. The random texture surface generated using MRF (Markov Random Field) model is shown in Fig.7 [7], in which the level of intensity is 5. Both images are composed of 512×512 pixels, and assumed to have their origins at the left-down corner. We set four local regions centered at (100,100), (100,300), (300,100), (300,300) respectively, and which consist of 101×101 pixels in each image. The extent l_{max} up to which GLDS is to be computed is set as 10. Then a surface orientation is extracted from the image onto which periodical surface inclined ($\alpha=\beta=10°$) in 3D space is projected. The GLDS patterns in $0°$ direction are depicted in Fig.8, and the shape of cost function in Fig.9. The results of estimations are shown in Table 1. For the estimation for periodical texture pattern the orientations are extracted within $\pm1°$ error, and for MRF random texture within $\pm3°$ error.

Effect of center position of local region is also examined. The results of it are shown in Table 2, in which the projected image of inclined pattern of Fig.6 is used, where an area of each local region is 50×50. The table shows that error of orientation gets bigger, as the overlapped area increases. This is due to reduction of the number of excluding pixels; because overlapped area has the same statistic feature.

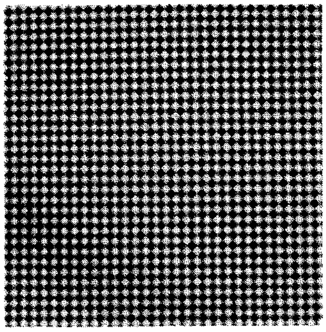


Fig. 6 The periodic pattern image.

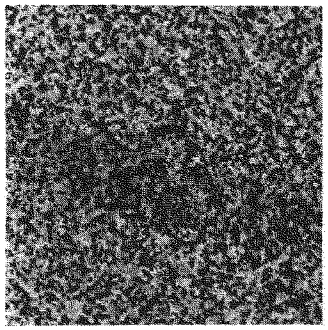


Fig. 7 The MRF textured image.

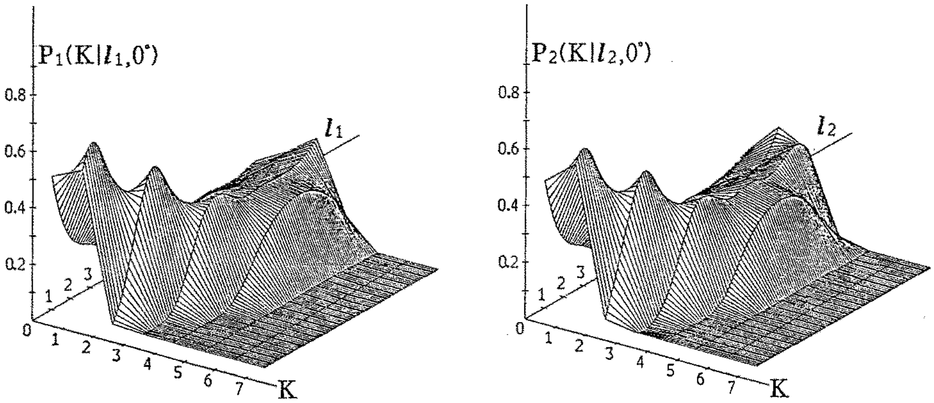


Fig. 8 An example of the gray level difference statistics of the inclined image surface ($\alpha = \beta = 10°$) of Fig. 6.

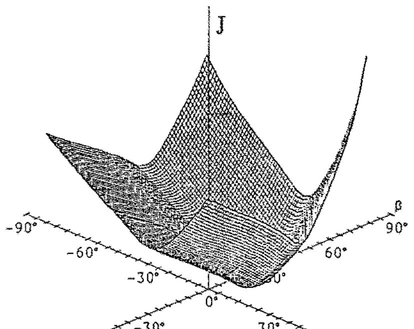


Fig. 9 The profile of the evaluation function of the inclied image surface ($\alpha = \beta = 10°$) of Fig. 6.

Table 1 The results of the estimation of the surface orientation (1)

Periodic pattern

True orientation α, β (deg.)	Estimates α^*, β^* (deg.)
10 , 10	10 , 11
20 , 0	20 , 0

MRF texture pattern

True orientation α, β (deg.)	Estimates α^*, β^* (deg.)
10 , 10	10 , 10
20 , 0	20 , 1
0 , 20	-3 , 22

region area : 101 x 101
center position : (100,100), (300,100), (100,300), (300,300)

Table 2 The results of the estimation of the surface orientation (2)

The center position of regions	Estimates α^*, β^* (deg.)
(100,100), (150,100), (100,150), (150,150)	11 , 7
(100,100), (200,100), (100,200), (200,200)	9 , 10
(100,100), (250,100), (100,250), (250,250)	8 , 11
(100,100), (300,100), (100,300), (300,300)	10 , 11
(100,100), (350,100), (100,350), (350,350)	10 , 10

region area : 101 x 101 True orientation: $\alpha = \beta = 10°$

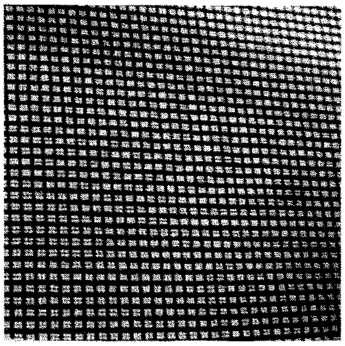


Fig. 10 Real textile image (1).
$\alpha = 20.9°, \beta = 20.2°$

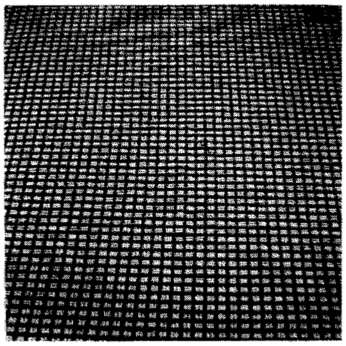


Fig. 11 Real textile image (2).
$\alpha = 0.0°, \beta = 48.1°$

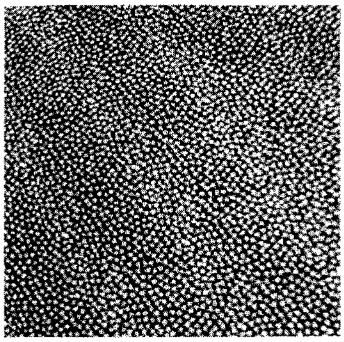


Fig. 12 Real textile image (3).
$\alpha = -18.0°, \beta = 24.2°$

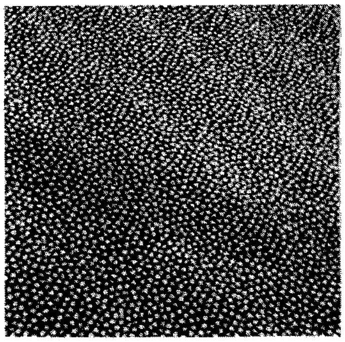


Fig. 13 Real textile image (4).
$\alpha = 13.1°, \beta = 37.6°$

Table 3 The results of the estimation of the surface orientation (3)

	True orientation α, β (deg.)	Estimates α^*, β^* (deg.)
Fig. 10	20.9 , 20.2	20 , 22
Fig. 11	0.0 , 48.1	0 , 47
Fig. 12	-18.0 , 24.2	-16 , 26
Fig. 13	13.1 , 37.6	7 , 38

center position : (-69,-70), (41, -70), (-69,40), (41,40)
region area : 81 x 81

Real texture Image

Images sampled by CCD video camera are also examined to estimate an orientation. The object plane is covered by textile fabrics. The primal image consists of 240 $\times$240 pixels and has the origin at the center and 256 intensity levels, but the number of levels is reduced to 8 using equal-probability quantizing algorithm [8]. The local centers are (-69,-70), (41,-70), (-69,40), (41,40) respectively, and each area has 81$\times$81 pixels. The images of the textiles is shown in Fig.10- 13. The orientations are reconstructed within an error $\pm 2°$ in spite of blurred image except for the image of Fig.13. For the image of Fig.13 the estimate of orientation gets accurate (α=13°,β=37°), when the area of local region is enlarged to 101 $\times$101. It means that exactness of the extraction depends upon the number and the character of data especially in statistical way.

As a result, it is shown that our method can extract the surface orientation from the projected image of textured surface plane using GLDS (gray level difference statistics).

CONCLUDING REMARKS

In this paper we proposed new method using GLDS for extracting a surface orientation from a single image. The method utilizes image intensity, which has not been used effectively. This enables us to obtain surface orientation directly without problems in edge extraction, texel estimation or other operations followed by it.

Our method is, however restricted for object surfaces which have uniform statistical feature about intensity of their texture. Furthermore an exactness of the estimation clearly depends upon positions and area of the local region of an input image. The above problems are left open. This study is initially aimed to reconstruct human body from a single image [9]. For the purpose the present method has to be arranged for curved surfaces.

ACKNOWLEDGMENT

This work was supported in part by the Scientific Research Foundation of Ministry of Education under Grant # 02750282. The first author wish to express his appreciation to the International Information Science Foundation for the support for his presentation.

REFERENCES

[1] J.J.Gibson, "The perception of the visual world," Houghton Mifflin, Boston (1950)

[2] Y.Ohta, K.Maenobu, T.Sakai, "A Method for obtaining plane surface orientation from texture under perspective projection," Proc. of IPSJ Workshop on CV,Vol.16-2, (1982)

[3] A.P.Witkin, "Recovering surface shape and orientation from texture," Artificial Intelligence, Vol.17, pp.17-45 (1981)

[4] J.R.Kender, "Shape from Texture: An aggregation transform that maps a class of texture into surface orientation," Proc. of IJCAI'79, pp.475-480 (1979)

[5] J.Aloimonos, "Shape from Texture," Biological Cybernetics, Vol.58, pp.345-360 (1988)

[6] J.S.Weszka, C.R.Dyer, A.Rosenfeld, "A comparative study of texture measures for terrain classification," IEEE Trans. System, Man & Cybern., Vol.SMC-6 (1976)

[7] S.Geman, D.Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," IEEE Trans. Pattern Anal. & Machine Intell., Vol.PAMI-6 (1984)

[8] R.M.Haralick, K.Shanmugam, I.Dinstein, "Textural features for image classification," IEEE Trans. System, Man & Cybern., Vol.SMC-3 (1973)

[9] H.Matsushima et al, "Extraction of Surface Orientation from Texture Using the Gray Level Difference Statistics," Trans. of IEICE, Vol.J73-D-II, pp.1993-2000 (1990) *in Japanese*.