コンピュータビジョンと
イメージメディア　103－2
（1997．1．23）

解の滑らかさの変動に適応する正則化
-正則化パラメータによる3D形状の知識表現-

宮島 耕治　　武川 直樹
NTTデータ通信株式会社　情報科学研究所
〒210 川崎市幸区堀川町66-2 興和川崎西口ビル9F
e-mail:miya@lit.rd.nttdata.co.jp

あらまし
本論文では，人間が物体の形状を認識する際に対象物体の事前情報を利用をしているという仮説と，人間が物体の解釈の可能性の中で最も単純な解釈を選ぶという仮説を用いた新しい3D形状の復元方法を提案する．

まず“物体の事前情報の利用”という仮説に基づいて，物体の事前情報として物体の形状を表現するsurface-and-contourモデルを提案し，これを正則化に適応する．このsurface-and-contourモデルに基づいて，物体の滑らかな形状と物体の境界部の位置に応じて正則化パラメータの値を画素毎に推定し，形状の滑らかさに応じた正則化を実現する．次に“物体の単純な記述”という仮説に基づき，物体形状の事前情報を表している画素毎の正則化パラメータをMDL基準によって評価する．これらの仮説に基づいて，正則化パラメータと物体の形状を同時に推定する方法を提案し，最後に提案手法をCG画像と実画像に適応した結果，およびその有効性について述べる．

Regularization Adapting for Various Smoothness of Solutions
-Knowledge representation of 3-D shape using regularization parameter-

Koji Miyajima and Naoki Mukawa
NTT DATA CORPORATION
Laboratory for Information Technology
KowaKawasaki-Nishiguchi Bld. 9F
66-2 Horikawa-cho, Saiwai-ku, Kawasaki 210 JAPAN

Abstract
In this paper we propose a new approach for estimating the 3-D shape of object. The approach introduces two hypotheses for human vision: the "use of prior information" hypothesis, in which people have prior knowledge before observations, and the "simple description" hypothesis, in which people choose the most simple description among many possible interpretations of objects. Based on first hypothesis, we propose a surface-and-contour model for representing an object's shape. The model is applied to the regularization formulation. We describe that the values of the regularization parameters given for each image pixel represent the location of smooth surfaces and contours. Based on the second hypothesis, we apply the minimum description length (MDL) principle to the regularization method. We then propose a method for estimating the regularization parameters and shapes of objects. We apply the method to estimate the shape of an object from two sequential images. We also provide experimental results on both CG and real sequential images.

1 Introduction

A fundamental problem of computer vision (CV) is estimating 3-D shapes of objects from 2-D images. The 3-D information from these images is degenerated in the imaging process. Therefore, in general it is impossible to obtain a stable and unique 3-D shape solution by only observing 2-D images. This is mainly because this inverse problem is an ill-posed problem. To transform the ill-posed problem into a well-posed problem, CV algorithms introduce prior-knowledge which humans use naturally as part of their common sense. By using the prior knowledge, CV algorithms can find solutions just as humans do.

Human vision (HV) can recognize the 3-D shapes of an object without the person being aware of using such prior knowledge. This is because people know through experience that the 3-D world is not chaotic but structured. HV automatically chooses a suitable solution which occurs with high possibility based on the prior knowledge. In other words, the object shape is empirically interpreted. Here, we call this the "use of prior information" hypothesis. Some prior information depends on physical properties in the real world. For example, the surface shape of an object will be smooth, the motion field of object surface will vary smoothly, light sources will probably exist above of the object, and surface will more likely be convex than concave. It is also widely believed that HV chooses the simplest solution among possible interpretations of the object shape. We call this the "simple description" hypothesis. Some principles of Gestalts psychology such as principles of proximity, similarity, common fate, good continuation and closure are all based on the hypothesis that HV recognizes objects as simple as possible. In CV, these hypotheses are explicitly described as the rule in the algorithm for obtaining the appropriate solution.

In this paper we attempt to estimate suitable 3-D surface shape of an object by applying the hypotheses of HV. More specifically, we treat 2.5-D or depth estimation problem. Based on the first hypothesis for HV, we model the shape of the object. In this paper the object shape is modeled so that they are composed of smooth surfaces connected with nodal lines in-between. In 2.5-D representation, the model can be interpreted as an object constructed of (1) smooth surfaces, (2) ridge lines which intersect the surfaces and (3) contour edges which separate the object from the background area. We refer to this model as the "surface-and-contour model."

Several models for representing 3-D objects have been proposed. Polyhedrons, super quadrics[6] and generalized cones[2] are widely used in the CV field. Although these models suit particular vision applications, they are inadequate for representing accurate object shapes in nature. We therefore employ the surface-and-contour model for more accurate estimation.

Next, let us consider the "simple description" hypothesis based on the surface-and-contour model. In HV, a person is able to estimate the locations of ridge lines and contour edges of an object. Moreover, using clues of shade and texture on the object surface, a person can estimate the object shape. However, it is difficult for CV to distinguish the ridge lines and contour edges from only the local information of images. The esimation problem for finding the location of ridge line and contour information and estimating the shape of the surface is a typical "chicken and egg problem." It is difficult to extract object regions and estimate surface shapes at the same time. After an object region is extracted correctly, it is easy to estimate the shape of the object. However, an object region can be correctly extracted only if the shape of the object has been estimated properly.

Two approaches have been proposed for solving this problem. One is to extract contour edges based on the local characteristics of the image [16],[14]. The other approach is to estimate contour edges and shapes of the surface at the same time based on the Markov Random Field (MRF) model[8]. Still there remains problem, because object regions are extracted based on local characteristics only, but not on global characteristics.

Based on the second hypothesis for HV, we must introduce the simplicity rule for recovering shape problem. Here, we attempt to estimate 3-D surface shapes and contour edges at the same time. The MDL principle is applied to accomplish this aim. As to the existing descriptions of an object shape such as polyhedrons, super quadrics and generalized cones, pioneering research efforts using MDL principle has been conducted. Each of these representation models is useful to particular applications. However, they are not inadequate to simulate the process of HV for estimating 3-D shape precisely.

In this paper, we propose a surface-and-contour model which is appropriate to shape recovery. We then propose the method for estimatiing the object shape using the MDL principle. This method is implemented as an extended version of the existing standard regularization.

The outline of the paper is as follows. In Section 2, we review standard regularization theory. We then present some attempts that have been made to improve the standard regularization method. In Section 3, we describe the two hypotheses of human vision which can be used in regularization formulation. We show that the regularization parameter represents the surfaces and contours of the object. In Section 4, the regularization parameter is determined so that it fits the region in terms of both smooth surfaces and contours. We propose the MDL-based regularization method. In Section 5, the proposed method is applied to 3-D shape estimation from the motion images. The experimental results clarify the advantages of the method.

2 Regularization and regularization parameter

2.1 Standard regularization theory

Standard regularization theory can be used in many kinds of CV applications. It enables ill-posed problems to be solved and determines unique and stable solutions from less reliable and incomplete observed data. Often it involves the assumption of smoothness as prior knowledge. The theory is considered to be an effective method for solving vision problems such as optical flow estimation, shape from shading, shape from motion and surface reconstruction. In standard regularization theory, the observation process is described as

$$\boldsymbol{y} = F\boldsymbol{z} \quad \text{(where } \boldsymbol{z} \text{ is unknown data, } F \text{ is linear operator).} \tag{1}$$

The inverse problem for estimating $\boldsymbol{z}$ from data $\boldsymbol{y}$ would come to a problem of how to minimize the following energy functional:

$$E = \| F\boldsymbol{z} - \boldsymbol{y} \|^2 + \alpha \| R\boldsymbol{z} \|^2. \tag{2}$$

In eqn.(2), $\| F\boldsymbol{z} - \boldsymbol{y} \|^2$ and $\| R\boldsymbol{z} \|^2$ are called the penalty functional and stabilizing functional, respectively. The stabilizing functional represents a constraint on solutions, e.g., smoothness. The term α, which represents the weight of the stabilizing functional, is called the regularization parameter.

In vision problems, the stabilizing functional represents smoothness of solutions and regularization parameters are given as *a priori* knowledge which represents the degree of smoothness of the object surface.

2.2 Prior knowledge in regularization

A disadvantage is often pointed out about standard regularization. Because the regularization parameter is a fixed value in a whole image, the solution may have errors in the region where the parameter does not fit. This is because the prior of shape model introduced in the stabilizing functional is not appropriate. To solve this, several studies have been proposed for extending the stabilizing functional term. Below we briefly review existing methods for handling both continuity and discontinuity on the object surface.

Li[13] proposed an adaptive method which is applicable to the discontinuous region. The method defines multiple stabilizing functionals based on the MRF model and uses the weighted sum of the functionals. However, the weights were not given explicitly. Geman and Geman[8] introduced an additional energy functional term based on the MRF model. The minimization process of the non-linear function is called "a line process." Discontinuity in the image can be detected based on the local features. The concept behind this approach can be applied to many areas of computer vision, and much research has been conducted on it. However, one disadvantage is that local features near edges are affected by noise, so some detected edges may be not appropriate. Terzopoulus[9] proposed a method for estimating a surface from sparse depth data, in which solutions are obtained by adaptively applying an appropriate weighting to a membrane model and a thin plate model based on the local features. Suitable weighting parameters are chosen on a pixel-by-pixel basis. However, a concrete method for choosing the weights of models was not discussed. Robert[16] proposed a stabilizing functional based on the second derivative of depth to preserve discontinuity in a disparity map from stereo images. Ghosal[14] proposed a stabilizing functional weighted by the gradient of the intensity of images. Using this method, optical flow is estimated in discontinuous regions. However, these methods change the weight of the stabilizing functional. Therefore, if noise is added in images, local features of the images are changed. Consequently, errors are generated in the estimated solution, because the weight is changed not only in discontinuous regions, but in continuous regions with noise as well.

The solution in these methods can be affected by local properties of the object such as noise, ridges and contours. The global structure of the object should be taken into account for the basis of the proper solution.

2.3 Simple description rule in regularization

Regarding the simple description hypothesis, Ayer[17] proposed an MDL-based optical flow estimation method in which solutions were adapted to a motion model. In this method, he assumes that there are multiple motion models in an image. The adequate number of the model is decided by minimizing the enconding length of the motion models. However, the 3-D structure is not taken into account, therefore it cannot be applied to 3-D structure estimation.

3 Use of the two HV hypotheses for regularization

This section presents a regularization using the hypotheses of HV. We showed in Section 1 that the surface-and-contour model is a natural representation of a 3-D object shape. We apply the surface-and-contour model to the regularization. In this paper, we assume that an object is constructed with (1) smooth surfaces, (2) ridge lines that intersect the surfaces, and (3) contours between the background and the surfaces or ridge lines.

To estimate the object shape, the strength of the constraint which represents the smoothness of solution should be adaptively changed according to the surface-and-contour model. In regularization, regularization parameter represents the degree of smoothness regarding the local shape of object. In other words, the regularization parameter indirectly represents the knowledge of the object shape such as surfaces, ridge lines and contours. Therefore, we must develop a criterion for determining the regularizaion parameter.

First, let us consider the prior information hypothesis. The evaluation functional (2) in regularization can be rewritten by the Bayesian approach. That is, the evaluation functional E in terms of the posterior probability distribution of $\boldsymbol{y}$ can be represented by

$$E = -logP(\boldsymbol{y}|\boldsymbol{z}, \alpha) + const. \tag{3}$$

The problem of minimuizing E is equivalent to that of maximizing $logP(\boldsymbol{y}|\boldsymbol{z}, \alpha)$ when the α is fixed. The term $-logP(\boldsymbol{y}|\boldsymbol{z}, \alpha)$ is the encoding description length in the evaluation function E. This encoding description length can be calculated by adding together the description lengths of the penalty functional and the stabilizing functional. However, in our problem, both the $\boldsymbol{z}$ and α are variables. Thus, α cannot be obtained through the single function of the description length.

Second, for obtaining regularization parameters we will describe an additional criterion. Based on the simple description hypothesis, representation of regularization parameters should be as simple as possible. Therefore, appropriate regularization parameters minimize the description length for representing the parameters. The regularization parameters and solution $\boldsymbol{z}$ can also be estimated so as to satisfy the both criteria at the same time.

4 3-D Object Estimatation using Regularization

This section presents a depth estimation algorithm using regularization. We then propose a method for obtaining regularization parameters based on the simple description rule.

4.1 Depth estimation algorithm

Let $\boldsymbol{T} = (U, V, W)^T$ be motion parameters of the camera and $\boldsymbol{\Omega} = (A, B, C)^T$ be rotation parameters of the camera. The velocity of vector $\boldsymbol{V} = (\dot{X}, \dot{Y}, \dot{Z})^T$ at a point $P(X, Y, Z)$ would then be represented as

$$\boldsymbol{V} = \boldsymbol{T} + \boldsymbol{\Omega} \times \boldsymbol{P}. \tag{4}$$

The point $P(X, Y, Z)$ is projected on the image plane. This point (x, y) is defined as

$$x = f\frac{X}{Z}, y = f\frac{Y}{Z}, \tag{5}$$

where f represents focal length.

Apparent motion on the image plane is then

$$\frac{\partial x}{\partial t} = f\left(\frac{\dot{X}}{Z} - \frac{X\dot{Z}}{Z^2}\right), \frac{\partial y}{\partial t} = f\left(\frac{\dot{Y}}{Z} - \frac{Y\dot{Z}}{Z^2}\right). \tag{6}$$

From (4), we obtain

$$\frac{\partial x}{\partial t} = f\left(\frac{U}{Z} + B - \frac{Cy}{f}\right) - x\left(\frac{W}{Z} + \frac{Ay}{f} - \frac{Bx}{f}\right), \quad (7)$$

$$\frac{\partial y}{\partial t} = f\left(\frac{V}{Z} + \frac{Cx}{f} - A\right) - y\left(\frac{W}{Z} + \frac{Ay}{f} - \frac{Bx}{f}\right). \quad (8)$$

Next let us consider the luminance constancy constraint on the image plane. Let I_x, I_y be a differential image of image $I(x, y, t)$ on x-y space, and I_t be a differential image of image $I(x, y, t)$ on temporal axis t. Assuming that the objects are rigid, we obtain the following constraint:

$$I_x u + I_y v + I_t = 0 \quad where \quad u = \frac{\partial x}{\partial t}, v = \frac{\partial y}{\partial t}. \quad (9)$$

From (7),(8) and (9), we can obtain the constraints represented by 3-D parameters as follows:

$$\begin{aligned} &\frac{f(I_x U + I_y V) - (xI_x + yI_y)W}{Z} \\ &= \left(I_y + \frac{y}{f}(xI_x + yI_y)\right)A - \left(I_x + \frac{x}{f}(xI_x + yI_y)\right)B \\ &\quad -(xI_y - yI_x)C - I_t. \end{aligned} \quad (10)$$

From (10), the penalty functional of regularization is

$$E_a = \int\int_\Omega (p\tau - q)^2 dxdy, \quad (11)$$

where

$$\begin{aligned} \tau &= 1/Z, \\ p &= f(I_x U + I_y V) - (xI_x + yI_y)W, \\ q &= \left(I_y + \frac{y}{f}(xI_x + yI_y)\right)A \\ &- \left(I_x + \frac{x}{f}(xI_x + yI_y)\right)B \\ &- (xI_y - yI_x)C - I_t. \end{aligned} \quad (12)$$

Integral is performed on the whole image region Ω. If we assume that the surface of the object is smooth, we can define the stabilizing functional as follows:

$$E_b = \int\int_\Omega \left(\left(\frac{\partial \tau}{\partial x}\right)^2 + \left(\frac{\partial \tau}{\partial y}\right)^2\right) dxdy. \quad (13)$$

By minimizing the sum of these evaluation functionals:

$$E = E_a + \alpha E_b, \quad (14)$$

we can estimate depth Z. We assume motion parameter and focal length of the camera are known. That is, $\boldsymbol{T}$, $\boldsymbol{\Omega}$, and f are given. We can resolve the minimization problem of (14) by solving Euler-Lagrange equation concerning τ.

Since $\partial E/\partial \tau = 0$, we can obtain the equation

$$(p^2\tau - pq) - \alpha\nabla^2\tau = 0. \quad (15)$$

After discretizing the equation, we can apply Gauss-Seidel method, and by approximation $-\nabla^2\tau \approx \tau - \bar{\tau}$, τ can iteratively be obtained as follows:

$$\tau^{n+1} = \frac{\alpha}{\alpha + p^2}\bar{\tau^n} + \frac{pq}{\alpha + p^2}, \quad (16)$$

where $\bar{\tau}$ is the local average of τ. Finally we can obtain depth Z, from τ.

4.2 Estimation of regularization parameter based on MDL principle

We will present a method for obtaining α in (14) using the MDL principle. The MDL principle achieves the mechanism of the simple description hypothesis in HV. We must evaluate the description length for describing the scene. First, we evaluate the description length for the penalty functional and stabilizing functional. The description lengths are represented by $L(E_a|\alpha)$ for the penalty functional and $L(E_b|\alpha)$ for the stabilizing functional. The terms $L(E_a|\alpha)$ and $L(E_b|\alpha)$ show the description lengths when the regularization parameter α is given. We represent description length of α as $L(\alpha)$. Then, we can calculate the total description length DL_E as the sum of the above description lengths:

$$DL_E = L(E_a|\alpha) + L(E_b|\alpha) + L(\alpha) + L(M) \quad (17)$$

In the equation, $L(M)$ shows the description length for index of α. Here, the regularization parameter α is a quantized value, thus the description length of the index α is constant. This constant value is empirically determined in advance. Therefore $L(M)$ in (17) can be ignored in the minimizing problem. In the following, we will show how to obtain each description length in (17).

First, let us consider $L(E_a|\alpha)$ and $L(E_b|\alpha)$. We represent $e_a(x_i, y_j)$ as the integrand of E_a, and $e_b(x_i, y_j)$ as the integrand of E_b:

$$e_a(x_i, y_j) = (q(x_i, y_j)\tau(x_i, y_j) - p(x_i, y_j))^2 \quad (18)$$

$$e_b(x_i, y_j) \quad = \quad \left(\frac{\tau(x_{i+1}, y_j) - \tau(x_{i-1}, y_j)}{2}\right)^2$$

$$+ \left(\frac{\tau(x_i, y_{j+1}) - \tau(x_i, y_{j-1})}{2}\right)^2 \quad (19)$$

where i and j indicate horizontal and vertical location of pixels in an image respectively.

The probability density functions $P_{e_a}(e_a|\alpha)$ about $e_a(x_i, y_j)$ and $P_{e_b}(e_b|\alpha)$ about $e_b(x_i, y_j)$ can be calculated if the regularization parameter $\alpha(x_i, y_j)$ is given for every image pixel (x_i, y_j). For simplicity, we represent $e_a(x_i, y_j)$, $e_b(x_i, y_j)$ and $\alpha(x_i, y_j)$ respectively as e_a, e_b and α.

After the α is given, description lengths about penalty functional and stabilizing functional can be represented by log likelihood of the probability density functions $P_{e_a}(e_a|\alpha)$ and $P_{e_b}(e_b|\alpha)$, respectively. Therefore, $L(E_a|\alpha)$ and $L(E_b|\alpha)$ can be represented as follows:

$$L(E_a|\alpha) = \sum_{x_i, y_j} (-log_2 P_{e_a}(e_a|\alpha)) \quad (20)$$

$$L(E_b|\alpha) = \sum_{x_i, y_j} (-log_2 P_{e_b}(e_b|\alpha)) \quad (21)$$

Here, we calculate $P_{e_a}(e_a|\alpha)$ and $P_{e_b}(e_b|\alpha)$ from histograms of $e_a(x_i, y_j)$ and $e_b(x_i, y_j)$, respectively.

Next, consider the description length $L(\alpha)$. Let P_α be the probability density function of α. We can evaluate the description length $L(\alpha)$ as follows:

$$L(\alpha) = \sum_{x_i, y_j} (-log_2 P_\alpha(\alpha)) \quad (22)$$

Thus, (17) can be rewritten as

$$DL_E = -\sum_{x_i, y_j} \left(log_2 P_{e_a}(e_a|\alpha) + log_2 P_{e_b}(e_b|\alpha) + log_2 P_\alpha(\alpha) \right) \quad (23)$$

In (23) the first term in the summation represents description lengths of errors between estimated and actual data. The second term represents the description lengths of the smoothness of solutions. The third term represents the description length of the regularization parameter which describes the object using the surface-and-contour model. Therefore, by minimizing DL_E, a solution of the object shape and parameter α can be found so that (1) the estimated solution has minimal errors to measured data, and (2) regularization parameter has minimal description.

4.3 Procedure for estimating regularization parameters

The problem remains to guide an algorithm for estimating the object shape and regularization parameters α based on the criteria in 4.2. The search space of the problem is very large, so much time is needed to find solution of object shape and α which minimize the description length (23). Currently, no efficient algorithms are known for finding a global minimum solution. Here, we introduce a method based on coarse-to-fine approach (Fig.1). The problem of convergence of the solution, however, needs further study.

In the coarse-to-fine approach, two sequential images are input and an initial regularization parameter α_0 (fixed value) is given. Based on (16), we estimate Z, the distance between the object surface and the camera. By reducing $\alpha(x_i, y_j)$ from a large initial value, α_0, to a small value for pixel by pixel, we estimate $\alpha(x_i, y_j)$ to minimize the description length (23) (Fig.2). The procedure is as follows:

Initial parameters set:

(i) Let two initial parameters α_0, α_1 be the initial parameters of $\alpha(x_i, y_j)$. Then estimate depths Z corresponding to each parameter α_0, α_1 where $\alpha_0 > \alpha_1$.

(ii) Initial division $k = 1$. Divide the image into small square regions, then calculate description lengths ${DL_E}^{k-1}$ and ${DL_E}^k$ corresponding to α_{k-1} and α_k in each region using (23).

(iii) In each region, compare ${DL_E}^{k-1}$ and ${DL_E}^k$. If ${DL_E}^{k-1} > {DL_E}^k$, divide the region into four small sub-regions, and give a new parameter α_{k+1} to each sub-region as $\alpha(x_i, y_j)$ so that $\alpha_{k+1} < \alpha_k$. This parameter is used for the next iteration. If ${DL_E}^{k-1} \leq {DL_E}^k$, give α_{k-1} to the region as $\alpha(x_i, y_j)$ for the next iteration.

(iv) Calculate the depth Z for the given $\alpha(x_i, y_j)$ in (iii). Then calculate the description length ${DL_E}^{k+1}$ in each region.

(v) Iteratively, put $k \leftarrow k + 1$ then go back to step (iii).

Finally, we can divide an image into very small regions. Through this process, by minimizing ${DL_E}^i$ in (23), we can determine $\alpha(x_i, y_j)$ for each small region.

5 Experimental Results and Discussions

To verify the performance of this method, we applied it to CG images in which the true shapes

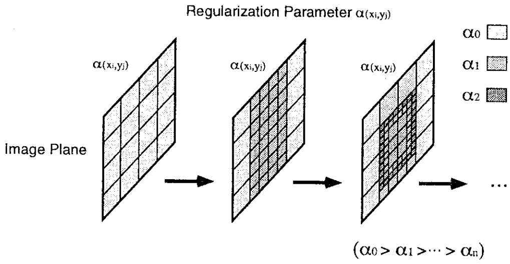

Figure 1: Outline of finding α based on coarse-to-fine approach

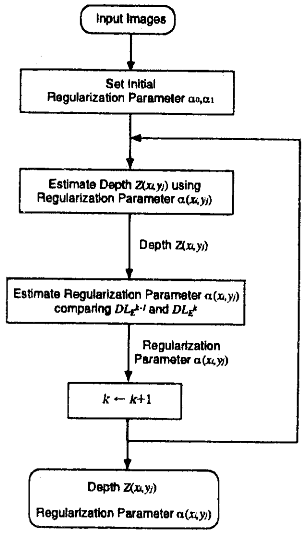

Figure 2: Flow chart of estimating depth and regularization parameters

of objects are known. Figures 3(a) shows image from a sequence of a cylinder moving vertically. Figures 3(b) shows the result of estimating regularization parameters. The intensity in Fig.3(b) indicates the values of the parameters, where the bright area shows the large α. Figure 3(c) and (d) show the result of estimating the shapes using the fixed α. Figure 3(e) shows the result of estimating the shapes using the proposed method.

As Figs.3(b) shows, in the discontinuous regions the values of the regularization parameters are lower than those in other regions, such as a boundary of the cylinder. Comparing figure 3(c), (d) and (e), we can see that the shape estimated by this method has fewer errors.

Next we applied this method to natural images. Figure 4(a) shows a image sequence in which a china vase moves. Figure 4(b) shows the result of estimating regularization parameters. Figure 4(c) and (d) show the result of estimating the shapes using the fixed α. Figure 4(e) shows the result of estimating the shape using the proposed method.

As Fig.4(b) shows, the values of the regularization parameters in the discontinuous regions are lower than those in other regions. We can see the shape estimated by this method has fewer errors.

In this section, we described shape estimation experiments based on the contour-and-surface model, which corresponds to HV. The proposed approach can be applied to other ill-posed problems in CV. We applied this approach to optical flow estimation. We obtained suitable regularization parameters corresponding to smooth and discontinuous regions and fewer errors in optical flow.

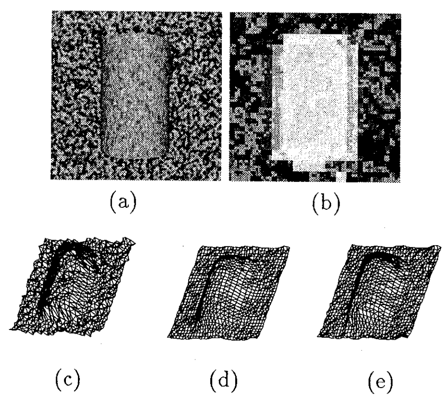

Figure 3: (a) Original sequence CG images, (b) Estimated regularization parameters, (c),(d) Depth plots of the estmated shapes using fixed $\alpha(\alpha = 1.0 \times 10^5$ and $\alpha = 4.0 \times 10^6)$, (e) Depth plots of the estmated shapes using proposed method$(1.0 \leq \alpha \leq 1.93 \times 10^7)$.

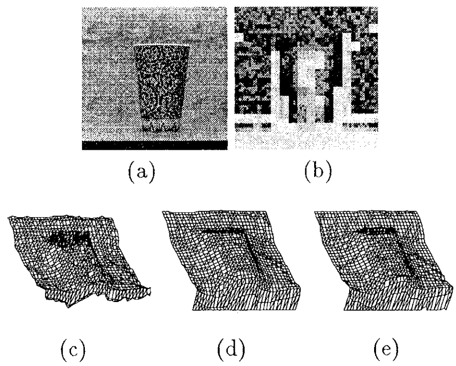

Figure 4: (a) Original sequence image, (b) Estimated regularization parameter, (c),(d) Depth plots of the estmated shapes using fixed $\alpha(\alpha = 2.5 \times 10^6$ and $\alpha = 1.0 \times 10^7)$, (e) Depth plots of the estmated shapes using proposed method($3.48 \leq \alpha \leq 7.71 \times 10^7$).

6 Conclusion

This paper proposed a new regularization-based approach which models the human vision process. For estimating the 3-D object shape, we apply the two hypotheses: the "use of prior information" hypothesis in which people have prior knowledge before observations, and the "simple description" hypothesis in which people choose the most simple description among many possible interpretations of objects.

We proposed a surface-and-contour model for describing the prior information of an object shape. In 2.5-D estimation, the location of surface and contours can be represented by regularization parameters since the parameters represent the degree of smoothness of the surface. To estimate the appropriate regularization parameters, we employ the MDL principle. We then showed that the energy functional in regularization is equivalent to the sum of two encoding description lengths which are calculated from the penalty functional and the stabilizing functional. We also showed that the description length of the regularization parameter must be minimized. We then proposed a concrete algorithm to obtain regularization parameters and object shape.

We applied the proposed method to shape reconstruction from images. In the experiments for CG images, we obtained regularization parameters which are suitable for estimating shape. The results show that our method is superior to the conventional method which uses a fixed regularization parameter. We also applied this method to natural scenes, and showed that the proposed method is effective for natural scenes. For future works, we will apply this approach to a wide variety of computer vision problems, such as shape from shading, shape from stereo, etc.

References

[1] P.H.Winston: "The Psychology of Computer Vision" McGraw-Hill Inc., 1975.

[2] D.Marr: "Vision" W.H.Freeman and Company, 1982.

[3] K.T.Spoehr and S.W.Lehmkuhle: "Visual Information Processing" W.H.Freeman and Company, 1982.

[4] R.J.Watt: "Visual Processing: Computational, Psychophysical, and Cognitive Research" Lawrence Erlbaum Associates Ltd., 1988.

[5] R.Szeliski: "Bayesian Modeling of Uncertainty in Low-level Vision" Kluwer Academic Publishers, 1989

[6] A.Barr: "Superquadrics and angle-preserving transformation", *IEEE Computer Graphics Application*, vol.18, pp.21-30,1981.

[7] A.Pentland "Perceptual organization and the representation of natural form",*Art. Intell.*, vol.28, pp.293-331,1986.

[8] S.Geman and D.Geman: "Stochastic relaxation Gibbs distributions and Baysian restoration of images", *IEEE Trans. on PAMI*, PAMI-6(6), pp.721-741,1984.

[9] D.Terzoponus: "Regularization of Inverse Visual Problems involving discontinuities", *IEEE Trans. on PAMI*, PAMI-8(4), pp.413-424,1986.

[10] B. Horn and B. Schunck: "Detemining Optical Flow", Artificial Intelligence **17**, pp.185-203,1981.

[11] Y.G.Leclerc: "Constructing simple stable description for image paritioning", *Int'l J. Computer Vision*, Vol.3, pp.73-102,1989.

[12] R.A.Barron and T.M.Cover: "Minimum complexity density estimation", *IEEE Trans. on IT*, IT-37, pp.1034-1054,1991.

[13] S.Z.Li: "On Discontinuity-Adaptive Smoothness Priors in Computer Vision", *IEEE Trans. on PAMI*, PAMI-17(6), pp.576-586,1995.

[14] S.Ghosal and P.Vanek: "A Fast Scalable Algorithm for Discontinuous Optical Flow Estimation" *IEEE Trans. on PAMI*, PAMI-18(2), pp.181-194,1996.

[15] M.Proesmans, L.V.Gool, E.Pauwels and A.Oosterlinck "Determination of Optical Flow and its Discontinuous using Non-Linear Diffusion", *ECCV'94*, pp.295-304,1994.

[16] L.Robert and R.Deriche: "Dense Depth Map Reconstruction: A Minimization and Regularization Approach which Preserves Discontinuities", *ECCV'96*, pp.439-451,1996.

[17] S.Ayer and H.S.Sawhney: "Layered Representation of Motion Video using Robust Maximum-Likelihood Estimation of Mixture Models and MDL Encoding", *ICCV'95*, pp.777-784,1995.

[18] S.X.Ju, M.J.Black and A.D.Jepson: "Skin and Bones: Multi-layer, Locally Affine, Optical Flow and Regularization with Transparency", *CVPR'96*, pp.307-314,1996.

[19] K. Kanatani: "Geometric Computation for Machine Vision", Oxford Univ. Press,U.K., 1993.