Viewpoint invariant reconstruction of visible surfaces using first-order regularisation

J.-H. Yi

The application of a first-order regularisation technique to the problem of reconstruction of visible surfaces is described. The approach is a computationally efficient first-order method that achieves approximate invariance. The results indicate that the proposed method for surface reconstruction performs well on sparse noisy range data.

Introduction: Surface reconstruction is necessary to derive a complete representation of a surface from sparse noisy sets of geometric information such as depth and orientation. The reconstruction result must be invariant with respect to viewpoint, i.e. to rotations and translations of the surfaces being reconstructed. In this Letter a novel, computationally efficient form of a first-order energy functional for surface reconstruction that achieves approximate invariance is proposed. Invariant reconstruction in the context of regularisation has been investigated by approximating an invariant energy function. Second-order models that are capable of invariant reconstruction are investigated in [1, 2]. However, a convex approximation to the first-order invariant form has not been previously reported. This is an important case to consider because first-order methods are more computationally efficient than current high-order methods.

Our algorithm consists of three steps: an initial reconstruction, partial derivative estimates from the initial reconstruction result, and a second reconstruction which uses the estimated derivatives. The estimated derivatives are inserted as constants into an approximately invariant energy functional (second reconstruction) which is then convex. The importance of the estimated derivatives is that they improve the performance of the second reconstruction with respect to invariance. The use of estimated derivatives that are not invariant does not yield truly invariant reconstruction. However, a significant improvement in invariance can be achieved in the second reconstruction by using this non-invariant information.

Viewpoint invariant reconstruction: Depth constraint data in an explicit form, $z(x, y)$, is used in our work given sparse range data z_{ij}. The perpendicular distance $|z - c|\cos\phi$ between the surface, z, and the constraint surface, c, is invariant where $\cos\phi$ is the surface slant [1]. The squared distance, $(z - c)^2\cos^2\phi$, which is also invariant, is used in our work. The data compatibility term $E(z, D)$ is as follows:

$$E(z,D) = \sum_{i,j}(z_{i,j}-c_{i,j})^2\cos^2\phi = \sum_{i,j}(z_{i,j}-c_{i,j})^2\frac{1}{1+z_x^2+z_y^2} \qquad (1)$$

where z_x and z_y are the first-order derivative at the measurement location (i, j). Estimates $\hat{z}_x$ and $\hat{z}_y$ of z_x and z_y are inserted as constants in the computation. This agrees with the research result by Ikeuchi and Kanade [3] where they reported a noise model of typical light-stripe range finder. According to their findings, the larger the angle between the surface normal and the illuminator direction of a light stripe, the larger uncertainty exists in the sensed z value.

We present a first-order stabilising function that is both convex and approximately invariant. We make a convex approximation to the following first-order stabilising function that is non-convex and invariant.

$$E_p(Z) = \iint_\Omega\left(\sqrt{1+z_x^2+z_y^2}-1\right)dxdy \qquad (2)$$

$\Omega \subset \Re$ denotes the image domain. A convex approximation to eqn. 2 that has been commonly used is as follows:

$$E_p(Z) = \iint_\Omega(z_x^2+z_y^2)\,dxdy \qquad (3)$$

which assumes $z_x \simeq 0$ and $z_y \simeq 0$; i.e. the integrand, $z_x^2 + z_y^2$ of eqn. 3 is the first-order Taylor expansion of the integrand, $\sqrt{(1 + z_x^2 + z_y^2)} - 1$ of eqn. 2, at $z_x^2 + z_y^2 = 0$. The approximation error of this first-order expansion is not ignorable when the slope $z_x^2 + z_y^2$ is large, i.e. surfaces of an image are steep. Our goal is to reduce this approximation error. Therefore, our approximation is to use a first-order expansion of the Taylor series at $g = \hat{g}$ instead of at $g = 0$. The approximation becomes

$$\sqrt{1+z_x^2+z_y^2}\Big|_{at\ z_x^2+z_y^2=\hat{g}} \simeq \sqrt{1+\hat{g}}+\frac{1}{2\sqrt{1+\hat{g}}}(z_x^2+z_y^2-\hat{g}) \qquad (4)$$

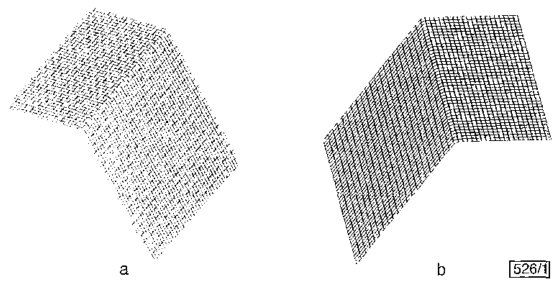

Fig. 1 *Three-dimensional displays of images*
a Data1
b Data2

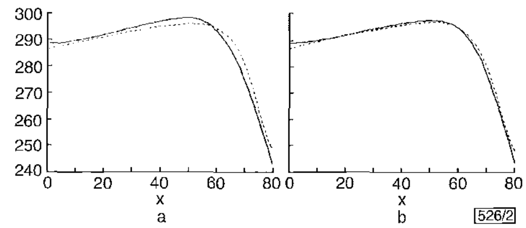

Fig. 2 *Invariance performance in the case of 90% sparseness and* $\lambda = 3.0$
a Normal fit (90% sparseness, V/A = 0.7070)
b Invariant fit (90% sparseness, V/A = 0.3340)
——— data1
·········· data2 rotated

Discrete equations: For simplicity, we will use the following forward finite difference to approximate the continuous surface.

$$\hat{z}_x = \frac{1}{h_x}(z_{i,j+1}-z_{i,j}) \quad \text{and} \quad \hat{z}_y = \frac{1}{h_y}(z_{i+1,j}-z_{i,j}) \qquad (5)$$

The image domain $\Omega \subset \Re^2$ is tessellated into rectangular subdomains with sides of h_x and h_y in the x and y directions, respectively. Nodes are located at subdomain corners where they are shared by adjacent subdomains. Combining the data compatibility measure and the stabilising function discretised using the finite difference (eqn. 5), we obtain

$$E^\lambda(Z) = \sum_{i,j} l_{i,j}\frac{1}{1+\hat{g}_{i,j}}(z_{i,j}-c_{i,j})^2 + \lambda^2\sum_{i,j}\frac{1}{\sqrt{1+\hat{g}_{i,j}}}\left\{\frac{1}{h_x^2}(z_{i,j+1}-z_{i,j})^2+\frac{1}{h_y^2}(z_{i+1,j}-z_{i,j})^2\right\} \qquad (6)$$

where $\hat{g}_{i,j}$ is the estimate of $z_x^2 + z_y^2$ at the location of node (i, j) and $l_{i,j}$ is zero where no data is provided at the location of (i, j). Constant terms are ignored. The resulting SOR (successive over-relaxation) updating equations for inside pixels are as follows. If there is data at node (i, j), i.e. $l_{i,j} = 1$:

$$z_{i,j}^{n+1} = z_{i,j}^n - \frac{\omega}{a+b\lambda^2}\Bigg[(a+b\lambda^2)z_{i,j}^n - \frac{1}{1+\hat{g}_{i,j}}c_{i,j} - \lambda^2\Bigg(\frac{z_{i,j-1}^{n+1}}{h_x^2\sqrt{1+\hat{g}_{i,j-1}}}+\frac{z_{i-1,j}^{n+1}}{h_y^2\sqrt{1+\hat{g}_{i-1,j}}} + \frac{z_{i,j+1}^{n}}{h_x^2\sqrt{1+\hat{g}_{i,j}}}+\frac{z_{i+1,j}^{n}}{h_y^2\sqrt{1+\hat{g}_{i,j}}}\Bigg)\Bigg] \qquad (7)$$

otherwise, i.e. $l_{i,j} = 0$:

$$z_{i,j}^{n+1} = z_{i,j}^{n} - \frac{\omega}{b}\left[b z_{i,j}^{n} - \left(\frac{z_{i,j-1}^{n+1}}{h_x^2\sqrt{1+\hat{g}_{i,j-1}}} + \frac{z_{i-1,j}^{n+1}}{h_y^2\sqrt{1+\hat{g}_{i-1,j}}} + \frac{z_{i,j+1}^{n}}{h_x^2\sqrt{1+\hat{g}_{i,j}}} + \frac{z_{i+1,j}^{n}}{h_y^2\sqrt{1+\hat{g}_{i,j}}} \right)\right] \quad (8)$$

where

$$a = \frac{1}{1+\hat{g}_{i,j}}$$

$$b = \frac{1}{h_x^2\sqrt{1+\hat{g}_{i,j-1}}} + \frac{1}{h_y^2\sqrt{1+\hat{g}_{i-1,j}}} + \frac{1}{h_x^2\sqrt{1+\hat{g}_{i,j}}} + \frac{1}{h_y^2\sqrt{1+\hat{g}_{i,j}}}$$

Experimental results: We have compared the invariance performance of our method, eqns. 7 and 8 (named 'invariant fit'), with that of the commonly used first-order method based on eqn. 3 (named 'normal fit'). To test for invariance, we use 'data1' and 'data2' as shown in Fig. 1. 'Data1' and 'data2' have two inclined planes of which the slopes are tan15° and tan75°. 'Data2' is obtained by rotating 'data1' by 60° about the y-axis. The Gaussian noise $N(0, 1)$ is added to 'data1' and 'data2' in the direction of the surface normal vector. To show the invariance property, dense and sparse 'data1' and 'data2' that are noisy are reconstructed and the reconstructed result for 'data2' is rotated back into correspondence with 'data1'. The difference between the two reconstructed surfaces is examined by computing the volume, V, between them divided by the average surface area, A, of the two reconstructions [4]. This measure provides the average distance between surfaces.

Fig. 2 illustrates the results of invariance tests by comparing one slice of the reconstruction results. According to our observations, when the data is dense, the reconstructed surface follows the data closely for both reconstructions. However, the difference in the invariance property between 'normal fit' and 'invariant fit' was still visible. As the sparseness of input image increases (i.e. when there are fewer data points), the invariance performance of 'normal fit' gets much worse than that of 'invariant fit' as shown in Fig. 2 where 90% of the pixels are missing. Invariance of 'invariant fit' (V/A = 0.3340) is much better than 'normal fit' (V/A = 0.7070).

Conclusion: In contrast to previous work, our approach is the first first-order computationally efficient method to achieve approximate invariance. It works especially well on sparse noisy range data. These facts make the proposed method for surface reconstruction more attractive than other existing methods.

Acknowledgment: This work was supported by grant No.1999-2-515-001-5 from the Basic Research Programme of the Korea Science & Engineering Foundation.

28 September 2000
Electronics Letters Online No: 20001389
DOI: 10.1049/el:20001389

J.-H. Yi (*School of Electrical and Computer Engineering, Sungkyunkwan University, Suwon, Korea*)

E-mail: jhyi@ece.skku.ac.kr

References

1 BLAKE, A., and ZISSERMAN, A.: 'Invariant surface reconstruction using weak continuity constraints'. Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, Miami, Florida, June 1986, pp. 22–26

2 STEVENSON, R.L., and DELP, E.J.: 'Viewpoint invariant recovery of visual surfaces from sparse data', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1992, **14**, pp. 897–909

3 IKEUCHI, K., and KANADE, T.: 'Modeling sensors and applying sensor model to automatic generation of object recognition program'. Proc. DARPA Image Understanding Workshop, 1987, pp. 697–710

4 YI, J., and CHELBERG, D.M.: 'An invariant performance measure for surface reconstruction using the volume between two surfaces', *IEEE Trans. Syst. Man Cybern. A*, 1998, **28**, (5), pp. 601–612

Protocol sequences for collision channel without feedback

V.C. da Rocha Jr.

A novel construction technique of protocol sequences for the users of a collision channel without feedback is introduced. This construction is a generalisation of the Massey-Mathys construction, with the practical advantage that the resulting cycle length N is typically much shorter, being at most equal to that obtained by the Massey-Mathys construction. The sender of each successfully transmitted packet is identified by a generalised version of the decimation decoding technique.

Introduction: The collision channel without feedback is a channel model proposed by Massey [1 – 3] for the situation where M users share a common communication channel but, because of unknown offsets among their clocks, are unable to transmit their data packets in a time-sharing mode and, because a feedback link is not available, can never determine these time offsets. Furthermore, because of the lack of a feedback link, the users can never be sure of the outcome of their individual packet transmissions.

The aim of this Letter is the introduction of a new construction technique of protocol sequences for the users of a collision channel without feedback. This construction shares with the Massey-Mathys construction [3] the fact that both are specific for the slot-synchronised case and are applicable to duty factor vectors with rational components. We review Massey's construction, which we refer to as Construction (i), and introduce our new construction technique and refer to it as Construction (ii). The cycle length N of the protocol sequences obtained by Construction (ii) is at most equal to that obtained by Construction (i), being typically much shorter, which can be an advantage in practical applications. We then present a generalisation of the technique known as decimation decoding [3], for identifying the sender of each successfully transmitted packet. The coding problem for reconstructing the packets lost in collisions, when these new protocol sequences are used, is exactly the same as in the classical case and thus can be handled by the usual procedures available in the literature, e.g. [3, 4]. Following Massey [3] we impose in the sequel the restriction that all M users align their packet transmissions to fall within time slots on their local clocks, and hence also within time slots on the receiver's clock, since in this case the time offsets δ_i, $1 \le i \le M$, are integer multiples of the slot length.

Classical technique: For the sake of completeness we shall describe next the only technique published so far [3] for the synthesis of any duty factor vector $\mathbf{p} = (p_1, p_2, ..., p_M)$ with only rational components, i.e. $\mathbf{p} = (q_1/q, q_2/q, ..., q_M/q)$, where $q_1, q_2, ..., q_M$ are non-negative integers and q is a positive integer assumed to be chosen as small as possible. In what follows we shall refer to this construction as Construction (i). A special protocol matrix S_{Mq} is constructed for the given $\mathbf{p}$ using an intermediary matrix with q-ary components. The intermediary matrix A_{Mq} is the $M \times q^M$ matrix whose jth column is the M place radix-q representation of the integer $q^M - j$, with least significant digit at the top. The desired protocol matrix S_{Mq} is obtained by mapping, within the ith row of A_{Mq}, the q-ary digits $q - 1, q - 2, ..., q - q_i$ to 1s and mapping the q-ary digits $q - q_i - 1, ..., 1, 0$ to zeros. We remark that this technique, although quite general, leads in many situations to rather long cycle lengths $N = q^M$, which is not usually a desirable feature in practice. Fortunately, as we show in the following Section, there are many situations where this difficulty can be overcome. For example, for $\mathbf{p}$ = (3/6, 2/6, 1/6) the resulting protocol sequence cycle length is N = 6^3 = 216 when Construction (i) is employed, while the value N = 36 results when instead we use Construction (ii), introduced in the following Section.

New construction of protocol sequences: We observe that the constraint adopted in Lemma 1 of [3] of using duty factor vectors $\mathbf{p} = (p_1, p_2, ..., p_M)$ which are probability vectors, i.e. vectors $\mathbf{p}$ such that $\Sigma_{i=1}^{M}\, p_i = 1$, is important for achieving rate points on the outer boundary of the capacity region but duty factor vectors $\mathbf{p}$ which are not probability vectors, although suboptimal in terms of rate, may be interesting in applications where a shorter cycle length may be desirable.