SIMULTANEOUS ESTIMATION OF SUPER-RESOLVED IMAGE AND 3D INFORMATION USING MULTIPLE STEREO-PAIR IMAGES

Kazuto Kimura, Takayuki Nagai

Department of Electronic Engineering
The University of Electro-Communications
1-5-1 Chofugaoka Chofu-shi,
Tokyo 182-8585 Japan

Hiroto Nagayoshi, Hiroshi Sako

Central Research Laboratory
Hitachi Ltd.
1-280 Higashikoigakubo Kokubunji-shi,
Tokyo 185-8601 Japan

ABSTRACT

This paper examines a simultaneous estimation of both high resolution (high accuracy) 3D information and high-resolution image from multiple low-resolution stereo-pair images. The proposed method is based on the MAP (Maximum A Posteriori) framework and these two relevant problems (estimation of both high-resolution 3D and intensity images) are incorporated into a unified cost function. Then the solution is obtained iteratively by a fixed point algorithm, which yeilds 3D information while the high-resolution image is fixed, and vice versa. Each solution can be obtained by solving a linear equation. It can be expected that the proposed method improves accuracy of both 3D information and high-resolution image simultaneously. Some experimental results are shown to validate the proposed method.

Index Terms— Super-resolution, 3D reconstruction, MAP

1. INTRODUCTION

Recovery of 3D information using degenerate 2D images has been one of the most important research topics in the area of computer vision. Moreover, the camera-based 3D measurement has many applications in areas of computer graphics and industrial machine vision. A number of approaches for the 3D reconstruction problem have been proposed in the past. One of the most famous methods is stereo processing, which uses multiple images taken from different view points to reconstruct 3D structure. In general, the problem of 3D reconstruction using multiple images is reduced to the point correspondence problem. Therefore, the algorithm for finding corresponding points in multiple images is the key issue for 3D reconstruction and quality of the resulting 3D information largely depends on the performance of the algorithm. At the same time, the resolution of input images is also a vital factor in accuracy of resulting 3D information. The method called Super Resolution (SR) of images aims to produce a high-resolution image from a set of low-resolution images by recovering or inferring plausible high-frequency image content. Many SR methods have been proposed in the past[1]. Typical approaches try to reconstruct a high-resolution image using the sub-pixel displacements of several low-resolution images. As far as far-field images are concerned, a single registration parameter to the overall image can be assumed. In fact, many conventional SR methods as in [2] utilize this assumption, which dramatically reduces the number of parameters to be estimated. This is not plausible, however, for near-field images since the displacement between corresponding pixels depends on the distance between the camera and the target point in 3D space. Therefore, 3D information is very important for the SR problem. This fact clearly shows that the problems of 3D reconstruction and image resolution enhancement are closely related to each other, nonetheless, these two issues have been considered separately.

This paper proposes a simultaneous estimation method of both high-resolution (accuracy) 3D information and high-resolution intensity image from multiple low-resolution stereo-pair images. The proposed method is based on the MAP (Maximum A Posteriori) framework and these two relevant problems (3D information and intensity image estimation) are incorporated into a unified cost function. Then we propose an iterative fixed point algorithm, which yields 3D information while the high-resolution image is fixed, and vice versa. Each solution can be obtained by solving a linear equation. It can be expected that the proposed method improves accuracy of both 3D information and high-resolution image simultaneously.

Related works include many multi-frame image SR methods and 3D estimation from multiple images. In [2], MAP based image SR method has been proposed. However the authors assume the use of far-field images such as aerial imagery. Hence a single registration parameter is introduced and no 3D information is considered. Two relevant methods, which simultaneously estimate high-resolution image and depth information, have been proposed. One of them uses defocus cue[3] and the other one utilizes photometric cue[4]. These methods yield accurate high-resolved image and depth information, however, defocus cue requires multiple images of the same scene taken with different focal lengths. To apply photometric cue to this problem, multiple images has to be taken under some different illumination conditions. In [5], authors have proposed the method of improving the depth estimation accuracy and image quality at the same time. However, they utilize simple averaging for improving image quality, which results in marginal improvement. On the other hand, we formulate and solve the problem using the MAP framework, which is a very different idea from that of [5].

2. SIMULTANEOUS ESTIMATION

2.1. Problem statement

Here we use multiple stereo-pair images (sequence) taken by a calibrated stereo camera. It is also assumed that 3D geometry of cameras is known. This is a feasible assumption since the algorithm is targeted at robotic systems, in which the camera position is accurately controlled. Optical flow can be combined with a rotary encoder to obtain pixel basis camera motion. In the following subsections, we consider the problem of estimating super-resolved image and 3D depth from a sequence of stereo-pair images $\boldsymbol{O}_i = [\boldsymbol{O}_i^L \ \boldsymbol{O}_i^R]$ as shown in Fig.1.

2.2. MAP formulation

The problem here is to estimate super-resolved image $\hat{\boldsymbol{X}}$ and super-resolved 3D depth $\hat{\boldsymbol{Z}}$ from a sequence of low-resolution stereo images $\boldsymbol{O}_i$. This problem can be formulated as

$$\hat{\boldsymbol{X}}, \hat{\boldsymbol{Z}} = \underset{\boldsymbol{X},\boldsymbol{Z}}{\operatorname{argmax}} P(\boldsymbol{X}, \boldsymbol{Z}|\boldsymbol{O}_0, \boldsymbol{O}_1, \cdots, \boldsymbol{O}_{N-1}), \qquad (1)$$

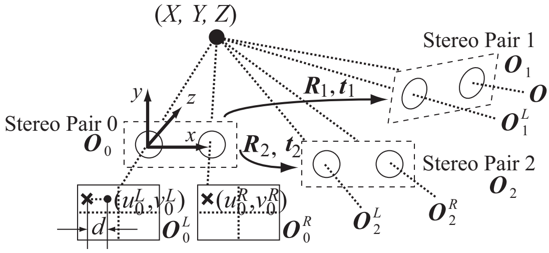

Fig. 1. An example of stereo camera location.

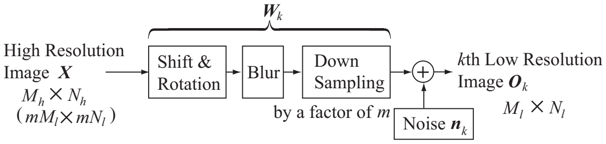

Fig. 2. The model for low-resolution image generation.

where N represents the number of stereo-pair. The left image $\boldsymbol{O}_0^L$ of the stereo-pair $\boldsymbol{O}_0$ is chosen to the reference frame, which is the base of the super-resolved image $\hat{\boldsymbol{X}}$. Now the problem can be written as the following minimization

$$\hat{\boldsymbol{X}}, \hat{\boldsymbol{Z}} = \underset{\boldsymbol{X},\boldsymbol{Z}}{\mathrm{argmin}} \left\{ -\sum_i \log P(\boldsymbol{O}_i^L|\boldsymbol{X},\boldsymbol{Z}) - \sum_i \log P(\boldsymbol{O}_i^R|\boldsymbol{X},\boldsymbol{Z}) - \log P(\boldsymbol{X}) - \log P(\boldsymbol{Z}) \right\}, \tag{2}$$

where conditional independence among $\boldsymbol{O}_i$'s and independence between $\boldsymbol{X}$ and $\boldsymbol{Z}$ are assumed. The likelihood $P(\boldsymbol{O}_i^L|\boldsymbol{X},\boldsymbol{Z})$ represents how likely the low-resolution left image of the i-th stereo-pair is input for given super-resolved image $\boldsymbol{X}$ and 3D depth $\boldsymbol{Z}$. Let the additive noise $\boldsymbol{n}_k$ in Fig.2 be Gaussian, we then have

$$\log P(\boldsymbol{O}_i^L|\boldsymbol{X},\boldsymbol{Z}) \propto \left(\boldsymbol{O}_i^L - \boldsymbol{W}(\boldsymbol{Z},\boldsymbol{R}_i,\boldsymbol{t}_i,m)\boldsymbol{X}\right)^T \times \left(\boldsymbol{O}_i^L - \boldsymbol{W}(\boldsymbol{Z},\boldsymbol{R}_i,\boldsymbol{t}_i,m)\boldsymbol{X}\right), \tag{3}$$

where $\boldsymbol{W}(\cdot)$ is a matrix that operate rotation $\boldsymbol{R}_i$, shift $\boldsymbol{t}_i$, blurring, and downsampling by a factor of m as shown in Fig.2. It should be noted that the matrix $\boldsymbol{W}(\cdot)$ contains 3D information $\boldsymbol{Z}$ implicitly and is abbreviated as $\boldsymbol{W}_i^L$. For the right image, $P(\boldsymbol{O}_i^R|\boldsymbol{X},\boldsymbol{Z})$ can be written in exactly the same way as the left one. The third term and fourth one in the right-hand side of Eq.(2) represent prior information on the image and the 3D structure, respectively. We use the smoothness constraint for this priori knowledge. Finally, the problem to be solved can be written as

$$\hat{\boldsymbol{X}}, \hat{\boldsymbol{Z}} = \underset{\boldsymbol{X},\boldsymbol{Z}}{\mathrm{argmin}} \left\{ \sum_{i=0}^{N-1} (\boldsymbol{O}_i^L - \boldsymbol{W}_i^L\boldsymbol{X})^T(\boldsymbol{O}_i^L - \boldsymbol{W}_i^L\boldsymbol{X}) + \sum_{i=0}^{N-1} (\boldsymbol{O}_i^R - \boldsymbol{W}_i^R\boldsymbol{X})^T(\boldsymbol{O}_i^R - \boldsymbol{W}_i^R\boldsymbol{X}) + \lambda_X \boldsymbol{X}^T\hat{\boldsymbol{C}}_X\boldsymbol{X} + \lambda_Z \boldsymbol{Z}^T\hat{\boldsymbol{C}}_Z\boldsymbol{Z} \right\}, \tag{4}$$

where λ_X and λ_Z represent weight for image smoothness and weight for depth smoothness, respectively. The matrices $\hat{\boldsymbol{C}}_X$ and $\hat{\boldsymbol{C}}_Z$, which impose the smoothness constraints, are composed of the laplacian kernel. We also set a weight for each pixel according to its discontinuity. This idea is known as the line-process and is introduced to prevent from obtaining overly smooth result. To minimize Eq.(4), we take an iterative fixed point algorithm, which gives 3D information $\boldsymbol{Z}$ while the high-resolution image $\boldsymbol{X}$ is fixed, and vice versa. It is obvious that Eq.(4) is in a quadratic form with respect to $\boldsymbol{X}$, and the solution is given by solving a linear equation accordingly. In contrast, Eq.(4) is not in a quadratic form with respect to $\boldsymbol{Z}$, since the matrices $\boldsymbol{W}_i^L$ and $\boldsymbol{W}_i^R$ contains 3D information $\boldsymbol{Z}$ intricately. In the following subsection, a linearization of Eq.(4) with respect to $\boldsymbol{Z}$ is examined.

2.3. Estimation of SR 3D information

This subsection describes the method for estimating super-resolved 3D depth for fixed $\boldsymbol{X}$. The initial values for super-resolved image and 3D depth are given by interpolation of the low-resolution input images and stereo matching of the input stereo-pair, respectively. It is worth noting that 3D depth $\boldsymbol{Z}$ and disparity of the reference stereo-pair $\boldsymbol{d}$ are connected through the equation $d = Bf/Z$, where f and B represent the focal length and the baseline length, respectively. . Thus estimating $\boldsymbol{Z}$ is equivalent to estimating $\boldsymbol{d}$. The rest of the paper considers disparity $\boldsymbol{d}$ of the reference stereo-pair $\boldsymbol{O}_0$ instead of $\boldsymbol{Z}$. Now let the rotation matrix and shift vector be

$$\boldsymbol{R}_i = \begin{bmatrix} a_0^i & a_1^i & a_2^i \\ b_0^i & b_1^i & b_2^i \\ c_0^i & c_1^i & c_2^i \end{bmatrix}, \quad \boldsymbol{t}_i = \begin{bmatrix} t_0^i \\ t_1^i \\ t_2^i \end{bmatrix}. \tag{5}$$

(U_i^L, V_i^L) represents a pixel coordinate on the left image that is a rotated and shifted version of the reference SR image by $\boldsymbol{R}_i$ and $\boldsymbol{t}_i$. The coordinate (U_i^L, V_i^L) can be written using the pixel coordinate (U, V) on the reference SR image and its disparity d (the subscript is omitted for notational convenience) as

$$U_i^L = \frac{B(a_0^iU + a_1^iV + a_2^if) + t_0^id}{B(c_0^iU + c_1^iV + c_2^if) + t_2^id}, \tag{6}$$

$$V_i^L = \frac{B(b_0^iU + b_1^iV + b_2^if) + t_1^id}{B(c_0^iU + c_1^iV + c_2^if) + t_2^id}. \tag{7}$$

Here we concentrate on the error of i-th left image in Eq.(4). It can be rewritten in a pixel basis as

$$\boldsymbol{O}_i^L(u,v) - \sum_U\sum_V \boldsymbol{W}_i^L(U_i^L, V_i^L, u, v)\boldsymbol{X}(U,V), \tag{8}$$

where $\boldsymbol{O}_i^L(u,v)$ represents a pixel value of $\boldsymbol{O}_i^L$, which is a lexicographical ordering vector of the image, at a location (u,v). $\boldsymbol{X}(U,V)$ also denotes a pixel value of $\boldsymbol{X}$ at (U,V). On the other hand, $\boldsymbol{W}_i^L(U_i^L, V_i^L, u, v)$ represents a component of the matrix $\boldsymbol{W}_i^L$ at $(U_i^L + M_hV_i^L, u + M_lv)$, where $U_i^L + M_hV_i^L$ and $u + M_lv$ denote column index and row index, respectively. We assume that the blurring process is approximated by Gaussian kernel with known variance σ^2. Then we have

$$\boldsymbol{W}_i^L(U_i^L, V_i^L, u, v) = \frac{1}{2\pi\sigma^2}\exp\left\{-\frac{(um - U_i^L)^2 + (vm - V_i^L)^2}{2\sigma^2}\right\}, \tag{9}$$

where m represents that the resolution of the SR image is m times larger than that of input low-resolution image in each direction ($M_h = mM_l$, $N_h = mN_l$). Let the current estimate of the disparity be

d_0 and Δd be the difference between the current estimate and the true disparity d, that is $d = d_0 + \Delta d$. First order Taylor series expansion of $\boldsymbol{W}_i^L(U_i^L, V_i^L, u, v)$ about d_0 gives $\boldsymbol{G}(U, V, u, v) + \boldsymbol{K}(U, V, u, v)\Delta d$, where

$$\boldsymbol{G}(U,V,u,v) = \frac{1}{2\pi\sigma^2} \exp\left\{ -\frac{(um - \hat{U}_i^L)^2 + (vm - \hat{V}_i^L)^2}{2\sigma^2} \right\},$$

$$\boldsymbol{K}(U,V,u,v) = \boldsymbol{G}(U,V,u,v)\{2\rho_u^i(um - \hat{U}_i^L) + 2\rho_v^i(vm - \hat{V}_i^L)\},$$

and

$$\begin{aligned} \hat{U}_i^L &= \frac{\alpha}{\gamma}, \quad \hat{V}_i^L = \frac{\beta}{\gamma} && (10) \\ \rho_u^i &= \frac{t_0^i\gamma - t_2^i\alpha}{\gamma^2}, \quad \rho_v^i = \frac{t_1^i\gamma - t_2^i\beta}{\gamma^2} && (11) \\ \alpha &= B(a_0^iU + a_1^iV + a_2^if) + t_0^id_0, && (12) \\ \beta &= B(b_0^iU + b_1^iV + b_2^if) + t_1^id_0, && (13) \\ \gamma &= B(c_0^iU + c_1^iV + c_2^if) + t_2^id_0. && (14) \end{aligned}$$

It should be noted that the indices of $\boldsymbol{G}(\cdot,\cdot)$ and $\boldsymbol{K}(\cdot,\cdot)$ are (U, V) instead of using $(\hat{U}_i^L, \hat{V}_i^L)$ for convenience of explanation. $\hat{U}_i^L$ and $\hat{V}_i^L$ can be obtained from U and V using Eq.(10). Finally we have

$$\begin{aligned} &\boldsymbol{O}_i^L(u,v) - \sum_U \sum_V \boldsymbol{W}_i^L(U_i^L, V_i^L, u, v)\boldsymbol{X}(U,V) \\ &\approx \boldsymbol{O}_i^L(u,v) - \sum_U \sum_V \boldsymbol{G}(U,V,u,v)\boldsymbol{X}(U,V) \\ &\quad - \sum_U \sum_V \boldsymbol{K}(U,V,u,v)\boldsymbol{X}(U,V)\Delta d. \qquad (15) \end{aligned}$$

It is worth noting that Eq.(15) is linear with respect to Δd, which is the variable to be obtained. The above approximation results in the following square error function

$$\|\boldsymbol{O}_i^L - \boldsymbol{G}\boldsymbol{X} - \boldsymbol{K}\boldsymbol{X}_d\Delta\boldsymbol{d}\|^2 = \Delta\boldsymbol{d}^T\boldsymbol{Q}_i^L\Delta\boldsymbol{d} - 2\boldsymbol{\epsilon}_i^L\Delta\boldsymbol{d} + \delta{\boldsymbol{O}_i^L}^2,$$

where

$$\begin{aligned} \boldsymbol{Q}_i^L &= \boldsymbol{X}_d^T\boldsymbol{K}^T\boldsymbol{K}\boldsymbol{X}_d, \quad \boldsymbol{\epsilon}_i^L = \delta{\boldsymbol{O}_i^L}^T\boldsymbol{K}\boldsymbol{X}_d, \\ \delta\boldsymbol{O}_i^L &= \boldsymbol{O}_i^L - \boldsymbol{G}\boldsymbol{X}, \\ \boldsymbol{X}_d &= diag\{\boldsymbol{X}(0,0) \cdots \boldsymbol{X}(M_h - 1, N_h - 1)\}. \end{aligned}$$

$\boldsymbol{G}$ and $\boldsymbol{K}$ are matrices whoes components are $\boldsymbol{G}(U, V, u, v)$ and $\boldsymbol{K}(U, V, u, v)$, respectively. For right images, similar approximation can be made. Finally, the minimization problem in Eq.(4) with respect to the disparity $\boldsymbol{d}$ can be written as

$$\begin{aligned} \hat{\boldsymbol{d}} = \underset{\boldsymbol{d}}{\text{argmin}} &\left\{ (\boldsymbol{d} - \boldsymbol{d}_0)^T\boldsymbol{Q}(\boldsymbol{d} - \boldsymbol{d}_0) \right. \\ &\left. -2\boldsymbol{\epsilon}^T(\boldsymbol{d} - \boldsymbol{d}_0) + \lambda_Z\boldsymbol{d}^T\hat{\boldsymbol{C}}_Z\boldsymbol{d} \right\}, \qquad (16) \end{aligned}$$

where

$$\boldsymbol{Q} = \sum_{i=0}^{N-1}(\boldsymbol{Q}_i^L + \boldsymbol{Q}_i^R), \quad \boldsymbol{\epsilon} = \sum_{i=0}^{N-1}(\boldsymbol{\epsilon}_i^L + \boldsymbol{\epsilon}_i^R).$$

Therefore, setting the partial derivative of the above equation with respect to $\boldsymbol{d}$ to zero yields optimal $\hat{\boldsymbol{d}}$ by

$$\hat{\boldsymbol{d}}^{(n)} = \left(\boldsymbol{Q} + \lambda_Z\hat{\boldsymbol{C}}_Z\right)^{-1}\left(\boldsymbol{Q}\hat{\boldsymbol{d}}^{(n-1)} + \boldsymbol{\epsilon}\right). \qquad (17)$$

2.4. Estimation of SR image

Here, we focus on the minimization of Eq.(4) with respect to the super-resolved image $\boldsymbol{X}$ for a fixed disparity $\boldsymbol{d}$. By fixing disparity $\boldsymbol{d}$, Eq.(4) becomes a quadratic form with respect to $\boldsymbol{X}$. Therefore $\hat{\boldsymbol{X}}$ can be obtained by solving the following linear equation

$$\begin{aligned} \hat{\boldsymbol{X}} = &\left(\sum_{i=0}^{N-1}({\boldsymbol{W}_i^L}^T\boldsymbol{W}_i^L + {\boldsymbol{W}_i^R}^T\boldsymbol{W}_i^R) + \lambda_X\hat{\boldsymbol{C}}_X \right)^{-1} \\ &\times \left(\sum_{i=0}^{N-1}({\boldsymbol{W}_i^L}^T\boldsymbol{O}_i^L + {\boldsymbol{W}_i^R}^T\boldsymbol{O}_i^R) \right). \qquad (18) \end{aligned}$$

And then, the disparity $\hat{\boldsymbol{d}}$ is re-estimated using the method described in the above subsection for the fixed $\hat{\boldsymbol{X}}$. These two processes are alternately iterated until convergence and a super-resolved image and a 3D depth are obtained.

2.5. Weight for smoothness constraint

How to set the weights λ_X and λ_Z in Eqs.(17) and (18) is an important issue. We will focus our discussions on λ_X since the same idea is applicable to λ_Z.

Since the noise in Fig.2 is assumed to be Gaussian, λ_X can be considered as the ratio between mean square error of images $E_i(\boldsymbol{X})$ and smoothness constraint error $E_x(\boldsymbol{X}) = \boldsymbol{X}^T\hat{\boldsymbol{C}}_X\boldsymbol{X}/(M_hN_h)$. Therefore, we approximate the weight using the estimated super-resolved image $\hat{\boldsymbol{X}}$ that is obtained in the previous iteration as $\lambda_X = E_i(\hat{\boldsymbol{X}})/E_x(\hat{\boldsymbol{X}})$. It seems that changing the weight according to the mean square error at each iteration is plausible. At the first stage of iteration, the estimated super-resolved image contains a large error that results in a large weight to smoothness constraint (large λ_X). On the other hand, weight for the square error of images gets higher (small λ_X) as the estimated super-resolved image $\hat{\boldsymbol{X}}$ gets closer to the solution $\boldsymbol{X}$.

3. EXPERIMENTS

The proposed algorithm was tested on 8 low-resolution stereo-pair images (16 images in total) shown in Fig.3(a), which were generated by a computer graphics software. We carried out 4 times super-resolution in each direction($m = 4$). Figure 3(b) shows the correct high-resolution image. The reference input image is shown in Fig.3(c). The result of cubic spline interpolation is given in Fig.3(d), which clearly shows the lack of high frequency component. The method in [2] was directly applied to the input images and the result in Fig.3(e) was obtained. In this case, the assumption of a single registration parameter is not valid and, thus, the result is not good enough. We also used the disparity, which was obtained by the stereo matching (Normalized Cross Correlation:NCC) for the interpolated stereo-pair, instead of using a single registration parameter in [2] (stereo+[2]). Subjectively better result has been obtained as shown in Fig.3(f), however, some serious matching errors invoked large errors in the super-resolved image. This result plainly indicates the importance of the simultaneous estimation of super-resolved 3D information and intensity image. Figure 3(g) shows the result of the proposed method. In contrast to other methods, the proposed method gives sharp super-resolved image. In fact, the mean absolute error between estimated super-resolved image and correct one for the proposed method is the smallest as shown in Tab.1, which objectively shows validity of the proposed method. Furthermore, as shown in Figs.4(a)-(c), it can be confirmed that the proposed method estimate accurate 3D information from multiple low-resolution stereo-pair images. The result is far better than that of the stereo matching using the two interpolated stereo-pair images as shown in Fig.4(c).

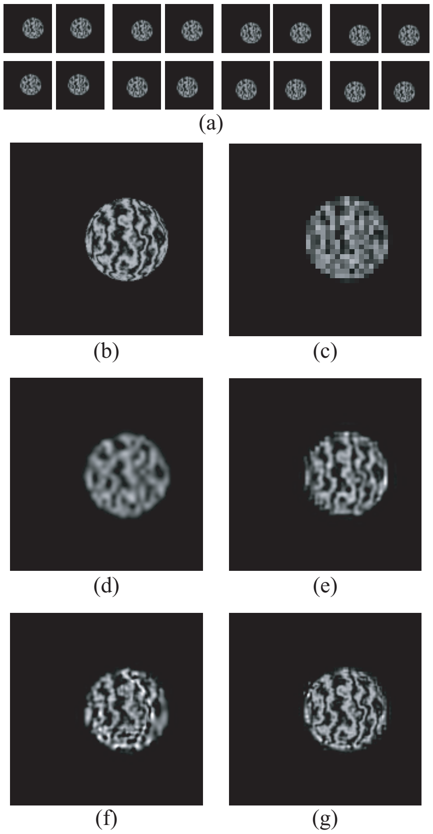

Fig. 3. Simulated results. (a)The input low-resolution stereo-pair images. (b)True high-resolution image. (c)The reference low-resolution image. (d)Interpolated by cubic spline. (e)The result of [2]. (f)Stereo+[2]. (g)Proposed method.

Table 1. Mean absolute error between the estimated super-resolved image and the ground truth.

	Interpolation	MAP	Stereo+MAP	Proposed
MAE	5.7	4.9	5.4	2.9

Here we show the result on real images. We fixed a stereo camera on an XY-stage and moved it 2cm at a time in each direction. The input stereo-pair images shown in Fig.5(a) were taken from 8 different positions in the same way as the simulation. The target object is a plane with texture(some Chinese characters on check). Four times super-resolution ($m = 4$) was carried out in this experiment. Figures 5(b) and (c) show the results. Figure 5(b) is the result of interpolation using cubic spline. The result of the proposed method is illustrated in Fig.5(c). One can see that the proposed method has yielded far better result compared to that of the interpolation. From Fig.6, it is also confirmed that the smooth plane has been obtained by the proposed method.

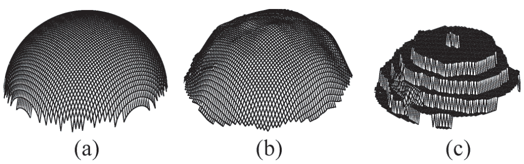

Fig. 4. Estimated 3D information. (a)True 3D information. (b)Proposed method. (c)Stereo matching(NCC).

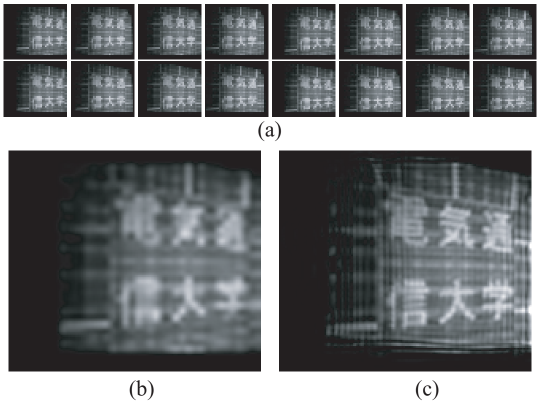

Fig. 5. The result on real input images. (a)The input low-resolution stereo-pair images. (b)The result of cubic spline interpolation. (c)The result of the proposed method.

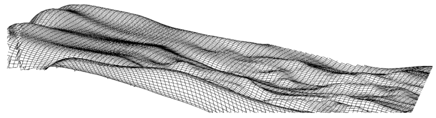

Fig. 6. Estimated 3D information for the real image.

4. CONCLUSION

In this paper we have proposed a simultaneous estimation method of both high-resolution (highly accurate) 3D information and high-resolution image from multiple stereo-pair images. The proposed method is based on the MAP (Maximum A Posteriori) framework and these two relevant problems (high accuracy 3D information and high-resolved image estimation) are incorporated into a unified cost function. Then the solution is obtained iteratively by a fixed point algorithm, which gives 3D information while the high-resolution image is fixed, and vice versa. Validity of the proposed method has been confirmed through some experimental results using both computer generated and real images.

5. REFERENCES

[1] S.C.Park, M.K.Park and M.G.Kang, "Super-Resolution Image Reconstruction: A Technical Overview", *IEEE Signal Processing Magazine*, pp.21–36, May 2003.

[2] R.C.Hardie, K.J.Barnard and E.E.Armstrong, "Joint MAP Registration and High-Resolution Image Estimation Using a Sequence of Undersampled Images", *IEEE Trans. on Image Processing*, vol.6, pp.1621–1633, Dec. 1997.

[3] DD.Rajan and S.Chaudhuri, "Simultaneous Estimation of Super-Resolved Scene and Depth Map from Low Resolution Defocused Observations", *IEEE Trans. on Pattern Anal. and Machine Intell.*, vol.25, pp.1102–1117, Sep. 2003.

[4] M.V.Joshia and S.Chaudhuri, "Simultaneous Estimation of Super-Resolved Depth Map and Intensity Field Using Photometric Cue", *Computer Vision and Image Understanding*, vol.101, issue 1, pp.31–44, Jan. 2006.

[5] K.Ikeda, M.Shimizu and M.Okutomi, "Simultaneous Improvement of Image Quality and Depth Estimation using Stereo Images", IPSJ SIG Technical Report, 2005-CVIM-149, pp77–82, May 2005 (in japanese).