A Refractive Camera for Acquiring Stereo and Super-resolution Images

Chunyu Gao and Narendra Ahuja
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
Email: cgao@uiuc.edu, ahuja@vision.ai.uiuc.edu

Abstract

We propose a novel depth sensing system composed of a single camera, and a transparent plate which is placed in front of the camera and rotates about the optical axis of the camera. The camera takes a sequence of images as the plate rotates, which provide the equivalent of a large number of stereo pairs. Compared with conventional multi-camera stereo systems, the use of a single camera for capturing stereo pairs helps improve the accuracy of detecting correspondences. The availability of the large number of stereo pairs also reduces matching ambiguities, even for objects with low texture. By using both the images and the estimated depth map, we show that the proposed system is also capable of generating super-resolution images. Experimental results on reconstructing 3D structures and recovering high-resolution images are presented.

1. Introduction

Stereo is one of the most widely explored approaches to estimating scene depth. It usually refers to spatial stereo, wherein two cameras, separated by a baseline, simultaneously capture stereo image pairs. The spatial disparity in the images of the same scene is utilized to estimate the scene depth. More than two cameras are also used to capture the disparity information across multiple views. Lee and Kweon presented an approach that uses a bi-prism to capture stereo images on the left and right halves of a single CCD sensor [11]. Nene and Nayar used different types of mirrors, including planar, ellipsoidal, hyperboloidal and paraboloidal mirrors, to capture stereo image pairs on a single sensor [12].

An alternative to such spatial stereo methods is temporal stereo wherein a single camera is relocated to a set of viewpoints to capture two or more images sequentially. This loses the parallel imaging capability and therefore the ability to handle fast moving objects, but it reduces the number of cameras as well as eliminates the frequent need for photometric calibration of multiple cameras for feature matching. Instead of relocating the camera, Kim et. al., proposed shifting the viewpoint by imaging off a rotating mirror [9].

An alternative to capturing the disparity by changing viewpoints, as in the approaches above, has been investigated by Adelson and Wang [1]. A lenticular array is placed in front of the sensor which changes the light path as a function of the object depth. The depth information is extracted by analyzing a set of sub-images. Non-disparity based approaches to depth estimation include depth from defocus [15] and depth from focus [6]. These methods also use a single camera to acquire multiple images over time but they estimate depth from changes in image blur rather than triangulation, and therefore, will not be reviewed here.

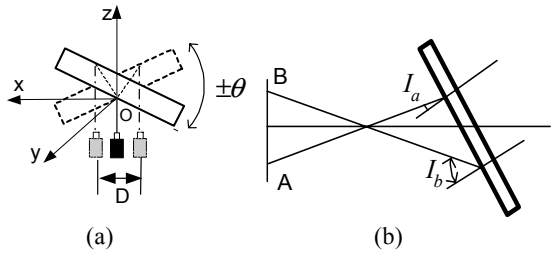


Fig 1. (a) Rotating the plate by ± θ results in a small baseline stereo system. (b) The incident angle varies with the pixel location.

This paper presents an approach to depth estimation that places a rotating, tilted, transparent, parallel plate in front of a conventional camera to achieve depth based disparities among a large number of images acquired as the plate rotates. The use of a parallel plate for a single-camera stereo was introduced by Nishimoto and Shirai [13], although in a different context. They acquire a stereo pair of images by using two orientations of the plate, as shown in Fig. 1a. They assume that tilting the plate about the y-axis by the two angles $\pm\theta$ results in two equivalent cameras whose effective positions are along the x-axis and symmetric

about the origin, thus forming a small baseline stereo system. However, as is known from optics, the lateral shift introduced by a transparent plate is a function of incident angle, which in our case varies with the tilt angle of the plate as well as the position of the feature point on the sensor. As illustrated in Fig 1b, the incident angle of point *B* is much bigger than that of point *A*. As a result, the shift of point *B* introduced by the plate is also much bigger than the shift of point *A*. Thus, the assumption of [13] is only approximately correct when the camera FOV is small. A general model was recently proposed for an arbitrary FOV [7], and validated by capturing stereo images by varying the tilt angle. However, this method is of limited practical value since each time the tilt angle changes, the calibration procedure has to be re-executed.

In this paper, we introduce a new dimension of control; we rotate the plate around the optical axis of the camera while keeping its tilt angle fixed and the rotation axis constantly aligned with the optical axis. This method avoids the repeated calibration required by the method of [7], is much easier to implement, and provides many more stereo images since many plate orientations can be produced during a continuous smooth rotation. As a consequence of the larger number of images, as well as the large tilt angles for each orientation made possible by rotation, the approach we propose here improves the accuracy of depth estimation. Further, the large number of resulting images, in conjunction with the estimated depth map, facilitates generation of super-resolution images. Super-resolution imaging using a transparent plate has also been investigated by Ben-Ezra et. al [2]. However, the difference between our system and theirs is quite fundamental. They place the plate in front of the sensor plane, perturb the plate orientation to achieve image shifts, use the shifts to achieve super-resolution, but have no depth sensing capability. Achieving super-resolution functionality as an integral part of depth estimation is another major contribution of this paper, and another distinction between our system and other single camera stereo systems.

The rest of this paper is organized as follows. In Sec. 2, we briefly review the mathematical model described in [7]. Then in Sec. 3, the rotating plate system is introduced, advantages of rotating the plate over tilting are highlighted, and a 4-point technique to align the plate rotation axis with the optical axis is presented. Techniques used to estimate depth and reconstruct super-solution images are discussed in Sec. 4. Finally, experimental results are presented in Sec. 5.

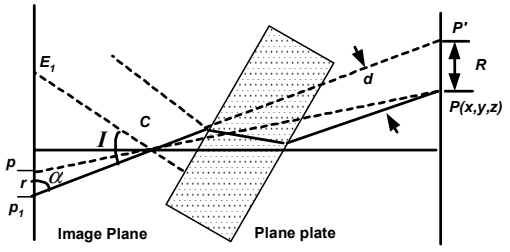


Fig 2: Projection of a 3D point through a plate

2. Mathematical model

2.1 Imaging Geometry

We briefly review the mathematical model of a static plate-camera system presented in [7]. Without loss of generality, the camera is assumed to be a pinhole which can be simplified as a projection center along with its intrinsic parameters (focal length *f*, image center (u_0, v_0) and distortion factor *k*). A parallel planar plate is placed in front of the camera. The imaging properties of the plate and its pose can be fully specified by four parameters: two intrinsic parameters, namely refractive index *n* of the plate material and thickness *t*, and two extrinsic parameters, namely the 2D coordinates of a special point on the image plane where the normal vector of the plate through the optical center intersects with the image plane. This special point is referred to as the essential point for a given plate pose.

Besides representing the plate pose, the usage of the essential point could limit the corresspondence search to 1D. As shown in Fig 2, an object point $P(x,y,z)$ is imaged at point $p_1(u_1, v_1)$ on the sensor plane through the projection center *C* and the plate. $p(u,v)$ is the image of *P* through the projection center *C* without the plate, hereafter referred to as the reference point. E_1 is the essential point under a certain pose (pose 1). The line passing through these three points is defined as the essential line of the image point p_1 (or *p,* or 3D point *P*) corresponding to pose 1. If the plate is reoriented to a new pose (pose 2) with the essential point E_2, then the image of the 3D point *P* through the plate can be found on the new essential line E_2p. Therefore, the correspondence search is limited on the essential line.

2.2 Pixel Shifts and Depth

It is well known from optics that a light ray passing through a planar plate encounters a lateral displacement. For a camera-plate system, this results in a location shift in the image of a scene point. The shifts depend on the intrinsic plate properties, the plate pose and the 3D positions of the objects. For a given object point, it is assumed that the point is shifted by the plate in a plane parallel to the image plane. As shown in Fig 2, the 3D point *P* is shifted by displacement *R* to point *P′*. The image points of *P* and *P′* through the projection center *C* are *p* and p_1, respectively. The displacement of *P* in image space is the distance between *p* and p_1

and is denoted as r. If both R and r are known, then the depth of P can be calculated by triangulation:

$$z = f * \frac{R}{r} \qquad R = \frac{d}{\sin a} \tag{1-2}$$

$$\cos\alpha = \frac{\overline{p_1C} \cdot \overline{p_1E_1}}{\left|\overline{p_1C}\right|\left|\overline{p_1E_1}\right|} \quad \cos I = \frac{\overline{p_1C} \cdot \overline{E_1C}}{\left|\overline{p_1C}\right|\left|\overline{E_1C}\right|} \tag{3-4}$$

$$d = t \sin I \left[1 - \sqrt{\frac{1 - \sin^2 I}{n^2 - \sin^2 I}} \right] \tag{5}$$

where: α is the angle between line p_1E_1 and p_1C; d is the lateral displacement of the chief ray of P, given by equation (5); I is the angle between line p_1C and line E_1C which is the incident angle of the chief ray; n is refractive index and t is thickness of the plate [19]. r is the shift of reference point p in image space, or disparity, caused by the plate.

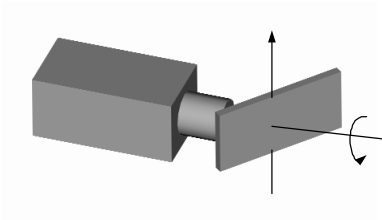


Fig.3. Concept illustration: Rotating a tilted plate for depth acquisition and super-resolution.

3. A Rotating Plate System

In order to sense depth via triangulation, at least two images captured through the plate under two different poses are necessary. In [7], multiple stereo images are captured by manipulating the plate tilt angle. We will show in this section that rotating a plate with a fixed tilt angle leads to a more calibration-friendly and higher-performance system than the use of multiple [7] or two [13] angles.

3.1 Rotation vs. Tilt

Fig. 3 illustrates the addition of the rotating plate in our prototype system. The rotation axis of the plate, hereafter, referred to as plate axis, is aligned with the optical axis of the camera. This seemingly simple modification, of adding rotation, leads to the following significant advantages.

(1) *More stereo images*. Each plate pose has an associated essential point on the image plane. The image trajectory of essential points as the plate rotates forms a circle since the plate axis is perpendicular to the image plane and parallel to the camera optical axis. We will refer to it as the essential circle. The center of the circle coincides with the center of the sensor; and its radius depends on the tilt angle of the plate. If the images captured are indexed by the rotation angle of the plate, then the rotating plate can be completely represented by the plate intrinsic parameters (refractive index n and thickness t), radius of the essential circle and the index angles. Once these parameters are obtained via a single calibration procedure, this new representation enables the system to capture many more stereo images using a continuous smooth rotation and the essential points for these images can be easily computed given current index angle. On the other hand, in [7] each plate pose requires a careful calibration process, and only these calibrated poses can be used for depth estimation. This practically limits the number of plate poses that can be employed. The ability of acquiring large number of images not only significantly reduces matching ambiguity and improves depth accuracy, it also facilitates super-resolution.

(2) *Mismatch rejection and more accurate depth maps*. As shown in Fig 4, the rotating plate system searches for correspondences in many directions. As a consequence, false match in one search direction could be easily rejected in other searching directions. Thus, our approach can improve the accuracy of correspondences and reject the ambiguities. On the contrary, the tilting plate system only searches for correspondences in either horizontal or vertical direction which could lead to mismatches for the object with low texture.

(3) *Better depth resolution*. For the proposed approach, the plate is rotated about the camera axis with a fixed tilt angle which can be set as a large fixed value instead of being varied from a large value to a small value as for the tilt-only case. The larger the tilt angle is, the larger is the spatial disparity, and consequently, the higher the resolution of the depth estimates.

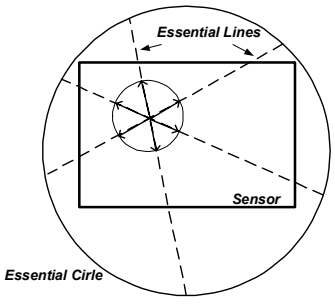


Fig 4: The rotating plate system searches for correspondences in many directions. Dash-lines are the essential lines; arrows show search directions; essential points are on the big circle.

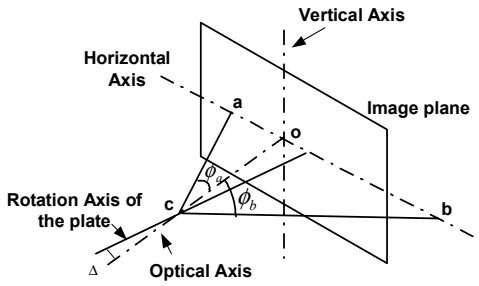


Fig. 5. Estimation of approximate plate orientation error using essential point pairs.

3.2 Calibration

The advantages accruing from plate rotation are contingent on the plate axis being aligned with the camera optical axis. Since under alignment the essential points of the rotating plate lie on the essential circle, ensuring alignment becomes quite simple: iteratively estimate the deviation of the essential points from a circle, and adjust the orientation of the plate axis to eliminate this error. Below, we introduce a four-point alignment method which uses four essential points, in pairs having opposite locations, to estimate alignment error between the plate axis and the camera optical axis. The four points are chosen such that they are evenly distributed on the sensor plane.

Aligning the plate axis with the camera optical axis can be done in one single rotation if its rotation angle and corresponding rotation axis are known. However, practically it is hard to achieve such rotation directly since the rotation axis could be in any arbitrary direction. Instead to estimate this rotation angle, we use two sub-rotations about two well-defined axes to approximate the alignment rotation. We decompose the rotation into three sub-rotations about the three orthogonal axes. For convenience, we choose the three camera axes (Fig 5) as sub-rotation axes. Since our target is to align the plate axis with one of the sub-rotation axes, namely, the camera optical axis, the rotation about the optical axis is redundant, and two rotations about the horizontal and vertical axis would suffice. Our strategy is to use two pairs of essential points to estimate these two sub-rotations. The estimated value is an approximation of the orientation difference between the plate axes and the camera axis, and is referred to as orientation error of the plate axes. Fig. 5 shows a pair of essential points *a* and *b* on the horizontal axis. If the plate axis is in the plane *abc*, then it must bisect the angle $\phi_a + \phi_b$. Thus the orientation error Δ of the plate axis is $(\phi_a - \phi_b)/2$. However, the plate axis is often off the plane. In this case, $(\phi_a - \phi_b)/2$ provides an approximation of the sub-rotation about the vertical axis. Similarly, the sub-rotation about the horizontal axis can be estimated using a pair of essential points on the vertical axis. By iteratively estimating and using the approximation of orientation errors to adjust the orientation of the plate axis, eventually, the plate axis can be aligned with the optical axis with arbitrarily high accuracy. The main steps of this alignment procedure are as follows:

(a) Select 4 plate poses and use the plate calibration method proposed in [7] to find the essential point for each pose.
(b) Calculate the distances between four essential points and image center. If the distances are not equal, compute the angle differences between each pair of essential points and estimate orientation error of the plate axis.
(c) Adjusting the plate orientation to compensate for the orientation errors.
(d) Repeat step (a) through (c) until the distance differences are within a small threshold.

During the alignment process, the radius of the essential circle and the intrinsic parameters of the plate are obtained as a by-product. Thus the calibration of these parameters is integrated with the alignment of the plate axis and no further calibration is necessary.

4. Depth Acquisition and Super-resolution

4.1 Finding Correspondences

For the normal stereo matching problem, search for the correspondences is limited to along the epipolar lines. Based on the mathematical model of Sec. 2.1, the

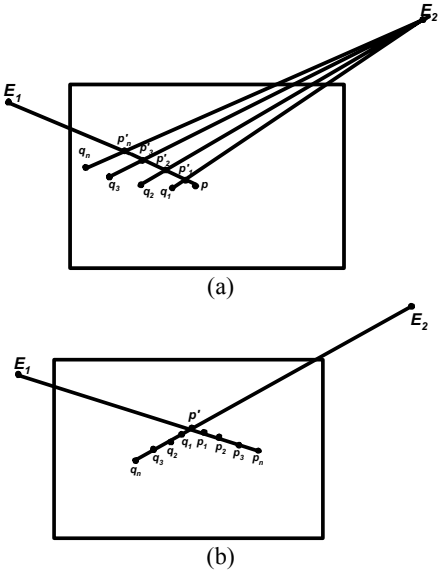


Fig.6: Two correspondence search strategies. (a) The correspondence search starts with points of input images; the search involves multiple essential lines. (b) The correspondence search starts with reference points; the search involves only two essential lines.

use of the essential point and essential line can limit the search to 1D. Our search can start with points in the first image and find matches in the rest of the images. However, this strategy involves multiple essential line pairs as illustrated in Fig 6a: given a point p in the first image, one can first calculate the positions of a set of potential reference points p_1', p_2', p_3', …, p_m' on the essential line E_{1p} using mapping equation (1~5), then calculate the positions of potential matching points q_1, q_2, q_3, …, q_m, on a set of essential lines E_2q_1, E_2q_2, E_2q_3, ..., E_2q_m. The reason for involving multiple essential lines is that multiple reference points are introduced into the search process. An alternative solution is to start the search with the reference points. The advantage of doing so is that the essential lines are well defined. As illustrated in Fig 6b, given a reference point p', searching for matches is limited to the essential line pair E_1P', E_2P', i.e. only two essential lines are involved. Another advantage of searching through the reference points is that the mapping from reference points to corresponding image points can be pre-calculated for a certain depth range because the mapping process depends only on the system geometry which is available from the calibration process.

Once these potential match points are located by searching through the essential lines, the intensity values of these pixels can be interpolated using bilinear interpolation or the nearest neighbor. The best match can be found by comparing the matching cost function such as SAD, or SSD. In our implementation, we use SSSD-in-inverse-distance [14]. The depth map is estimated using graph cuts method proposed in [16].

4.3 Super-resolution

So far, we have discussed depth recovery. In some applications, the texture of 3D objects is also required, e.g., for rendering. Using a large set of images and associated depth map, a high-resolution image can be generated.

Multi-frame super-resolution restoration idea was first presented by Tsai and Huang [20]. Since then, many algorithms have been developed [3, 18]. In this paper, we follow the restoration framework proposed by Elad and Feuer [5]. Given m low-resolution images I_1, I_2, ...I_m, the image formation process of I_k can be formulated as:

$$\vec{L}_k = D_k B_k W_k \vec{H} + \vec{N}_k \tag{6}$$

where: $\vec{L}_k$ is the low-resolution image I_k in vector form; $\vec{H}$ is the high-resolution representation in vector form; W_k is geometrical warping matrix; B_k is image blurring caused by the sensor; D_k *is* the decimation matrix; N is additive image noise.

In our case, the m low-resolution images correspond to those taken through the plate; and the geometrical warping matrix W_k can be obtained from the mapping between the reference image and the captured images. The blurring matrix B_k , which depends on the quality of camera lens and the plate, is unknown. In our experiment, we found that the blurring matrix has a minor effect on estimation, so we set it to the identity matrix. The decimation matrix D_k is related to the decimation ratio between the high-resolution image and *k-th* captured images; we set this ratio to 4:1. Grouping m equations corresponding to the m low-resolution images yields:

$$\begin{bmatrix} \vec{L}_1 \\ \vdots \\ \vec{L}_m \end{bmatrix} = \begin{bmatrix} D_1 B_1 W_1 \\ \vdots \\ D_m B_m W_m \end{bmatrix} \vec{H} + \begin{bmatrix} \vec{N}_1 \\ \vdots \\ \vec{N}_m \end{bmatrix} \tag{7}$$

In a compact form:

$$Y = MX + N \tag{8}$$

This is a classic image restoration problem which can be solved by many algorithms such as ML estimator, MAP estimator or POCS (Projection on Convex Sets) [21, 8, 10]. In our implementation, ML estimator is used which estimates the unknown image X by maximizing the conditional probability density function $P(Y/X)$. Since recovery of X is ill conditioned, a regularization term is necessary in the maximization. The additive noise is assumed to be zero-mean Gaussian with autocorrelation G. Then the estimate of the unknown image is:

$$\hat{X} = \arg\max \{[Y - MX]^T G^{-1} [Y - MX] + \lambda [\gamma X]^T V [\gamma X]\} \tag{9}$$

where $\hat{X}$ is the estimated higher resolution image; V is a weigh matrix and γ is the Laplacian [10]. $\hat{X}$ can be

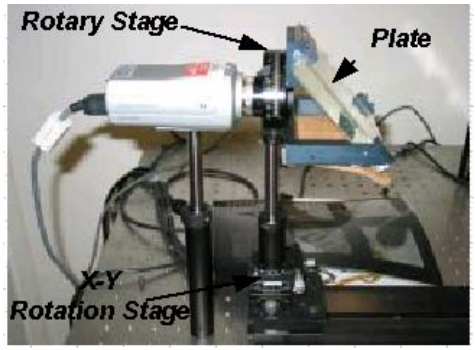


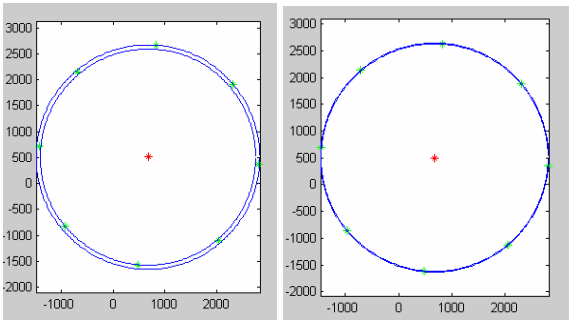
Fig 7. A snapshot of the experimental setup: an XY rotation stage is used to align the plate axis with the camera axis.

solved by differentiating the equation (9).

5. Experimental Results

5.1 Experimental Setup and System Alignment

In our experiment, a Sony DFX900 color camera equipped with a 12mm lens was used to capture images. An ordinary glass plate with thickness 13mm and refractive index around 1.6 was mounted on a rotary stage and tilted approximately 45º with respect to the image plane. The rotation axis of the plate is aligned with the camera view axis using the method described in the Sec. 3.2. A snapshot of our experimental setup is shown in Fig 7. The essential points corresponding to a set of fixed rotation angles are obtained from calibration process [7]. A total of 8 points are used to estimate the plate parameters, while four of them are used to estimate the orientation errors. Fig 8 shows the location of these points before and after alignment. After alignment is achieved, the average distance from the essential points to the image center is 2130.32 pixels. The standard deviation is 4.57 pixels which is only 0.21% of the average distance.

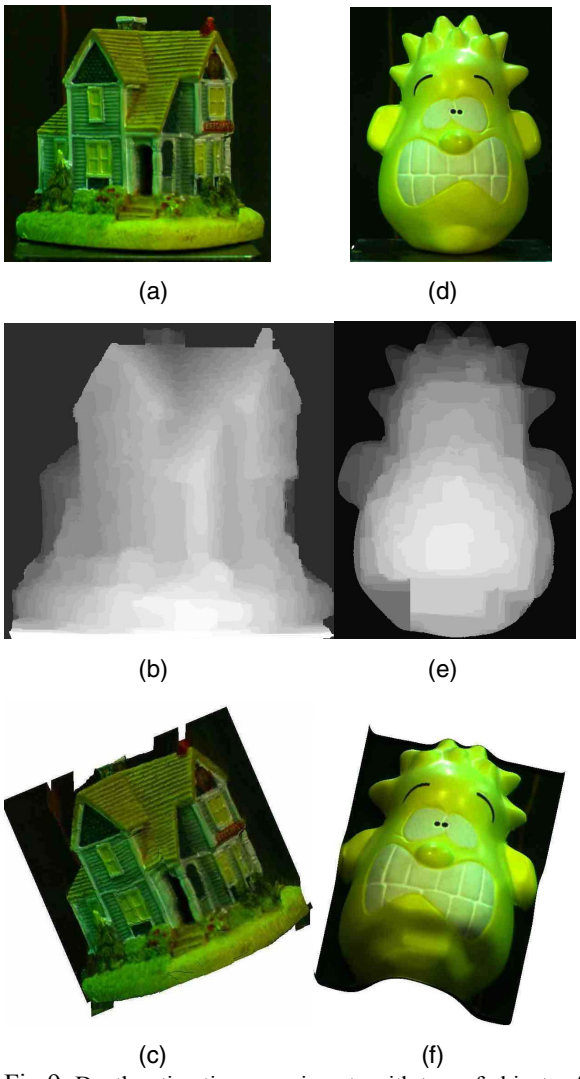


(a) (b)

Fig 8. Distribution of essential points on the image plane before and after the plate rotation axis is aligned with the camera optical axis. (a) Before: all the essential points lie between two concentric circles. The radius of the big circle is the largest distance between the essential points and the image center, and the radius of the small circle is the smallest distance. The center of the circles is the image center. (b) After: the two circles overlap.

5.2 Depth Sensing

A total of 36 plate poses, evenly distributed within the range of 0-360º, were used to recover depth. The coordinates of essential points of these poses are computed from the essential circle and their index angles. For the depth sensing experiment, we test our system with two objects: a doll house and a monster head. The experimental results of the House object are shown in the first column of Fig 9. 9a is one of the input images taken through the plate, 9b shows the recovered depth map, and 9c shows a new view generated from the reconstructed house model. Since the house model has rich texture, the window size used



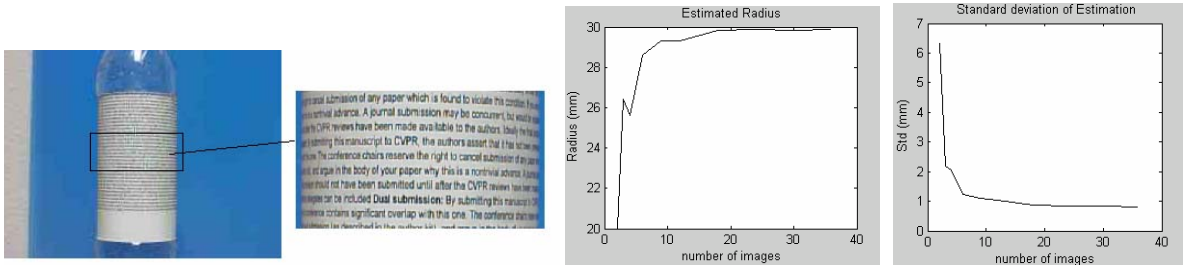
(a) (d)

(b) (e)

(c) (f)

Fig 9. Depth estimation experiments with two of objects. A "House" object with a rich texture is shown in column 1. A "Head" object with low texture is shown in column 2. Row 1 shows one of input images; Row 2 shows the recovered depth maps; Row 3 shows the reconstructed 3D models.

to calculate the cost is just 1x1, i.e. a single pixel. The advantage of using a single pixel is that depth discontinuities are well preserved. The ambiguity caused by depth discontinuity is minimized. The second column of Fig 9 shows the experimental results for the "Head". Figs 9d~9f show one of input images, the recovered depth map and a new view of the reconstructed head model, respectively. This result indicates that the proposed method performs well for objects having a limited amount of texture. A 5x5 window was used in this experiment. The depth range of the objects is around 350-450mm which is captured at a resolution of 1mm and displayed using the intensity range of 0-255 (9b and 9e). Most of the depth variations are visible via shading and edges.



(a) (b) (c)

Fig 10. (a). A small portion of input image is used to evaluate the depth accuracy. (b). Plot of estimated radius of the cylinder using different numbers of input images; (c). Standard deviation of the estimation error for different numbers of input images.

5.3 Depth Accuracy vs. Number of Input Images

Capturing a large number of stereo images is a major advantage of the proposed system over other single camera stereo systems. In the last experiment, we show that with a large set of stereo images, the system successfully recovers depth for a diverse set of objects. However, capturing and processing images increase the required user effort and computation. Therefore, it is useful to study how the number of images affects depth accuracy. To do so, we chose a cylindrical shape object (a bottle) covered with texture as test object. The diameter of the cylinder is 60mm. A total of 36 images were captured through the plate in the same way as done in the depth sensing experiment. By selecting different sample intervals, we formed 10 groups of image subsets containing 2, 3, 4, 6, 9, 12, 18, 24, 30 and 36 images respectively. For each subset, the images are chosen in the way that their corresponding rotation angles are widely distributed within the range of 0-360º. Different depth maps are estimated using each image subsets. The radius of the cylinder is estimated using these depth maps. Fig.10b shows the plot of the estimated radius vs. the number of samples. The plot clearly shows that more input images yield better depth accuracy, although the incremental gain diminishes with the number of images; in our case, when the number of images is more than 18, the additional images have negligible effect on accuracy. Fig.10c shows the standard deviation of the estimation error for different numbers of images.

5.4 Reconstructing Super-resolution Image

The super-resolution imaging capability was evaluated qualitatively using the same object as used for the depth sensing accuracy experiment. All 36 input images are used to estimate the depth of the paper-wrapped surface and construct the high-resolution image. One of 36 input images is shown in Fig 11. The sub-image in the small window is selected to compare the image quality of different image reconstruction methods. The results are shown in Fig 12. Fig 12a shows the up-sampled version using nearest neighbor interpolation. Fig 12b shows the up-sampled version using bilinear interpolation. Fig 12c shows the ML-reconstructed super-resolution image, whose quality can be seen to be significantly better than the up-sampled versions using the nearest neighbor and bilinear interpolation methods, but still not as good as the ground truth (Fig 12d) as is to be expected.

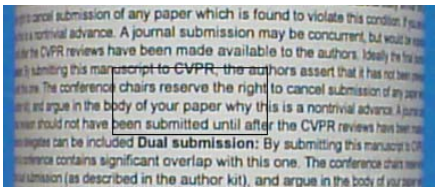


Fig 11: One of input images for constructing super-resolution image. The sub-image in the small window is used to evaluate image quality of the reconstructed image

6. Conclusion

In this paper, we have proposed a single camera stereo system which is capable of recovering depth and generating a super-resolution image. The system includes a rotating parallel planar plate with rotation axis aligned with the camera optical axis. Such alignment significantly simplifies system calibration and enables the system to capture a large number of stereo images during a continuous smooth rotation, compared with other plate-camera systems. The large number of stereo images help reduce matching ambiguities, even for objects with low texture, and improve the accuracy of the depth estimates. In future, we plan to focus on the quantitative evaluation of depth resolution and range, as well as on integration of depth estimation and the super-resolution image reconstruction which could further improve the system performance.

Acknowledgments

The support of the National Science Foundation under grant NSF IBN 04-22073 is gratefully acknowledged.

References

[1] Edward H. Adelson and John Y.A. Wang, "Single Lens Stereo with a Plenoptic Camera", *In IEEE Transactions on PAMI*, 14(2), February 1992.

[2] Moshe Ben-Ezra, Assaf Zomet, and Shree K. Nayar, " Jitter Camera: High Resolution Video from a low Resolution Detector", *In IEEE Conference on CVPR*, 135-142, 2004

[3] S. Borman and R. Stevenson. "Spatial resolution enhancement of low-resolution images sequences: a comprehensive review with directions for future research". *Tech. Report*, Univ. of Notre Dame, 1998.

[4] U. R. Dhond and J. K. Aggarwal. "Structure from stereo – a review", *IEEE Trans. on systems, Man, and Cybern.,* 19(6): 1489-1510, 1989

[5] M. Elad and A. Feuer. "Restoration of Single Super-resolution Image from Several Blurred, Noisy and Down-sampled Measured Images". *IEEE Trans. on Image Processing*, 6(12):1646–1658, 1997.

[6] K. Engelhardt and G. Hausler. "Acquisition of 3D data by focus sensing". *Applied optics*, 27:4684-4689, 1988

[7] Chunyu Gao and Narendra Ahuja, "Single camera stereo using planar parallel plate", Pattern Recognition, 17th International Conference on (ICPR'04) Volume 4, pp. 108-111 08, 2004, Cambridge UK.

[8] A. K. Jain, "Fundamentals in Digital Image Processing". Englewood Cliffs, NJ: Prentice-Hall, 1989.

[9] Hyongsuk Kim, Chun-Shin Lin, Jaehong Song, and Heesung Chae "Distance Measurement Using a Single Camera with a Rotating Mirror", *International Journal of Control, Automation, and Systems,* vol. 3, no. 4,pp. 542-551, December 2005.

[10] R. L. Lagendijk and J. Biemond "Iterative Identification and Restoration of Images". Boston,MA Kluwer, 1991.

[11] Doo Hyun Lee and In So Kweon, "A novel stereo camera system by a biprism", *IEEE Trans. on Robotics and Automation*, 16(5), Oct. 2000.

[12] S. Nene and S. Nayar. Stereo with mirrors. *In Proceedings of the 6th International Conference on Computer Vision*, Bombay, India, January 1998

[13] Y. Nishimoto, and Y, Shirai. "A feature-based stereo model using small disparities," *Proc. of the IEEE International Workshop on Industrial Applications of Machine Vision and Machine Intelligence*. Seiken Symposium, 192-196, Tokyo 2-5 Feb. 1987

[14] M. Okutomi and T. Kanade. "A multiple-baseline stereo." In *IEEE Transs on Pattern Analysis and Machine Intelligence*, 15(4):353–363, 1993.

[15] A. P. Pentland, "A new sense for depth of field," *In IEEE Trans. on Pattern Analysis and Machine Intelligence*, 9:523-531, 1987.

[16] S. Roy. "Stereo without epipolar lines: A maximum flow formulation." *International Journal of Computer Vision*, 34(2/3):147–161, 1999.

[17] Daniel Scharstein and Richard Szeliski, "A taxonomy and evaluation of dense two frame stereo correspondence algorithm", *International Journal of Computer Vision,* 47:7-42, April 2002

[18] C. A. Segall, R. Molina and A. K. Katsaggelos, " High Resolution Images from a Sequence of Low Resolution and Compressed Observations: A Review," *IEEE Signal Processing Magazine,* 20(3):37-48, May 2003

[19] Warren J. Smith. "Modern optical Engineering," McGraw-Hill, 3rd edition, NY, 2000

[20] R. Y. Tsai and T. S. Huang, "Multiple frame image restoration and registration," in *Advances in Computer Vision and Image Processing*, vol. 1, T. S. Huang, Ed. Greenwich, CT: JAI, pp. 317–339, 1984.

[21] D. C. Youla, "Generalized image restoration by the method of alternating orthogonal projections," *IEEE Trans. Circuits System*, vol. CAS-25, pp. 694–702, 1978.

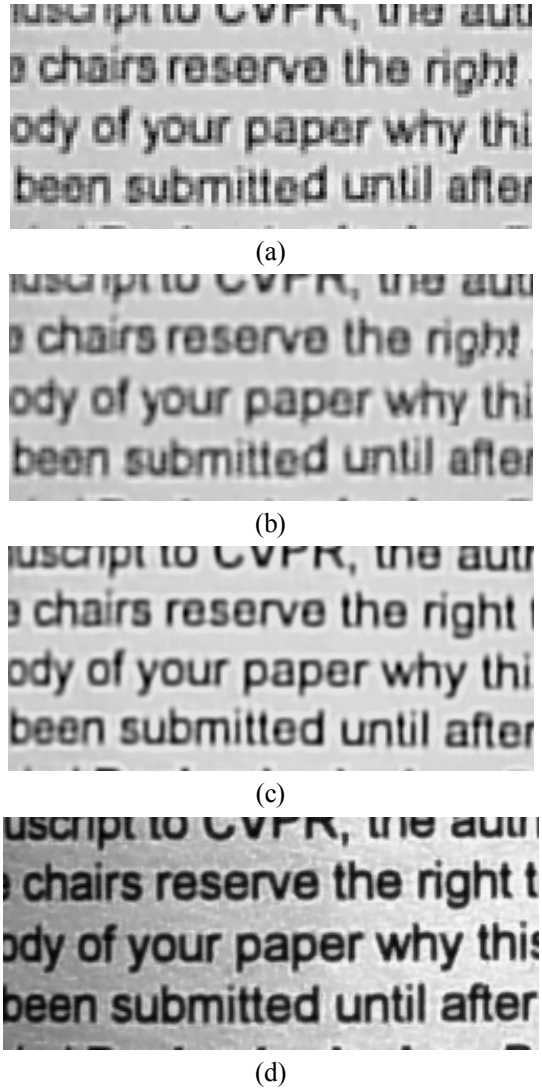


Fig 12 The images up-sampled by different methods: (a). Nearest Neighbor Interpolation; (b). Bilinear Interpolation; (c) Super-resolved image using the proposed ML method. (d). Ground truth image taken through a zoom lens for qualitative comparison.