

Vision-Based Modeling and Interaction

Zhengyou Zhang

Microsoft Research, Redmond, USA

zhang@microsoft.com

We strive to advance the state of the art in 3D computer vision, and develop flexible and robust techniques for ordinary users to gain 3D experience from an ordinary camera. In this talk, I will provide an overview of my research projects at Microsoft Research.

The first is on 3D object and face modeling from images taken by a free-moving camera.

Applications include product advertisement on the Web, virtual conference, and interactive games. We briefly cover the following topics:

- Camera calibration: determine a camera's internal parameters by just observing a planar pattern under unknown orientations.
- Stereo rectification: transform a pair of stereo images based on a geometric meaningful criterion such that the epipolar lines of the rectified images are aligned. This will considerably simplify the stereo matching process.
- Image matching: match points and curves across images within a probabilistic relaxation framework. We can obtain a much richer description of a scene using both points and curves.
- Progressive stereo matching: start with a few reliable point matches, then progressively add new matches during an iterative matching process based on two fundamental concepts: the disparity gradient limit principle and the least commitment strategy.
- 3D photo editing: Virtual or real objects can be inserted in the real images, while preserving the correct 3D information of the scene structure.
- Object modeling: A complete 3D model of an object is built by taking a number of snapshots around the object.
- Face modeling. With five mouse clicks, a 3D face model is built in a few minutes, and the model can be animated immediately. We have successfully built 3D face models for several hundred people including a few celebrities.

The second is on vision-based interaction.

Applications include human-computer interface and augmented reality. We briefly cover the following topics:

- Visual Screen: A prototype system which converts an ordinary screen into a touch screen.
- Visual Panel: A prototype system which converts a rectangular panel (e.g., an ordinary piece of paper) into a virtual mouse, keyboard and joystick.

References:

- [1] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330-1334, 2000.
- [2] Z. Zhang and C. Loop. Estimating the Fundamental Matrix by Transforming Image Points in Projective Space. *Computer Vision and Image Understanding*, 82(2):174-180, 2001.
- [3] C. Loop and Z. Zhang. Computing Rectifying Homographies for Stereo Vision. *IEEE Conf. Computer Vision and Pattern Recognition (CVPR99)*, Vol.I, pages 125-131, Colorado, June 1999.
- [4] Y. Shan and Z. Zhang. New Measurements and Corner-Guidance for Curve Matching With Probabilistic Relaxation. *International Journal of Computer Vision*, 46(2), 2001.
- [5] Z. Liu, Y. Shan and Z. Zhang. Expressive Expression Mapping with Ratio Images. In *SIGGRAPH*, pages 271-276, Los Angeles, August 12-17, 2001.
- [6] Z. Liu, Z. Zhang, C. Jacobs and M. Cohen. Rapid Modeling of Animated Faces From Video. *The Third International Conference on Visual Computing (Visual 2000)*, pages 58-67, September 2000, Mexico City.
- [7] Y. Shan, Z. Liu and Z. Zhang. Model-Based Bundle Adjustment with Application to Face Modeling. In *Proceedings of the 8th International Conference on Computer Vision (ICCV01)*, Vol. II, pages 644-651, July 2001, Vancouver, Canada.
- [8] Z. Zhang and Y. Shan. A Progressive Scheme for Stereo Matching. *Springer LNCS 2018: 3D Structure from Images - SMILE 2000*, M. Pollefeys et al. (eds.), pages 68-85, Springer-Verlag, 2001.
- [9] Z. Zhang. Image-Based Modeling of Objects and Human Faces. (Invited Paper) In *Proceedings of SPIE, Vol. 4309: Videometrics and Optical Methods for 3D Shape Measurement*, pages 1-15, San Jose, USA, January 2001.
- [10] Z. Zhang and Y. Shan. Visual Screen: Transforming an Ordinary Screen into a Touch Screen. *IAPR Workshop on Machine Vision Applications (MVA 2000)*, pages 215-218, Tokyo, Japan, November 2000.
- [11] Z. Zhang, Y. Wu, Y. Shan, and S. Shafer. Visual panel: Virtual mouse keyboard and 3d controller with an ordinary piece of paper. In *Workshop on Perceptive User Interfaces, ACM Digital Library*, November 2001. ISBN 1-58113-448-7