# Binocular and Multiple View Stereo using Tensor Voting

Philippos Mordohai and Gérard Medioni

A NATIONAL SCIENCE FOUNDATION
ENGINEERING RESEARCH CENTER

## Research Goal

Derivation of scene descriptions from Binocular and Multiple View Stereo in terms of:

- Surfaces
- Curves
- Junctions
- Depth and Orientation Discontinuities

## Role in IMSC

Inexpensive and reliable model generation for:

- Graphics
- Animation
- Haptics

Models of people objects and environments for:

- tele-conferencing
- Immersipresence
- e-commerce
- cartography

## Research Approach

- Initial matching using multiple methods
  - Normalized cross-correlation
  - Interval matching
- Data Representation by second-order symmetric non-negative definite tensors
- *Smoothness* enforcement in the design of the voting fields
- Information propagation by Tensor Voting
- Use of *saliency* to determine correct matches
- Group points into 3-D surfaces to obtain color information for scene surfaces
- Use color information to validate matches and generate matching hypotheses for unmatched pixels

## Uniqueness & Related Work

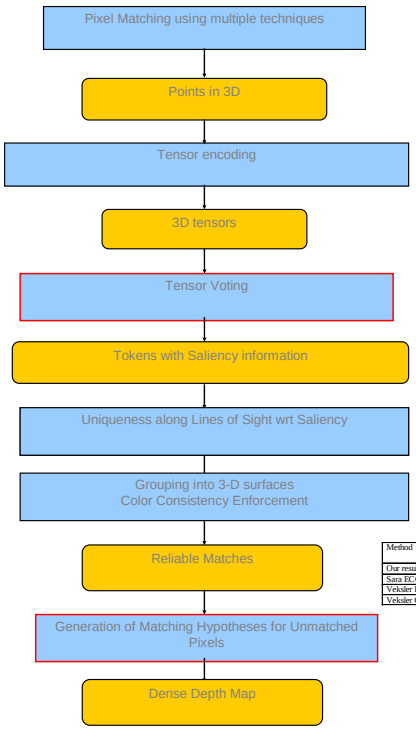
- Global Optimization Approaches
  - Processing in 2-D (even 1-D)
  - High computational complexity
  - Many unable to handle more than two views
- Methods using Image Segmentation
  - Image Segmentation is harder than Stereo in some cases
- Our Approach
  - Processing in 3-D
  - Integration of multiple matching methods
  - Manageable computational cost due to locality of Tensor Voting
  - Linear wrt number of images
  - Grouping based on binocular cues to infer color distributions

## 5-Year Plan

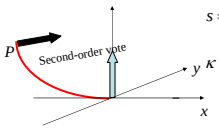
Short Term:
- Integration of curvature

Mid Term:
- Multiple-resolution structure extraction
- More efficient data structures

Long Term:
- Multiple-scale Tensor Voting
- Multiple-viewpoint texture-mapping



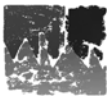
Multiple matches for every feature are allowed at this stage




$$s = \frac{l\theta}{\sin\theta}$$

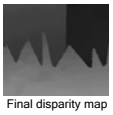
$$\kappa = \frac{2\sin\theta}{l}$$

$$DF(s,\kappa,\sigma) = e^{-\left(\frac{s^2+c\kappa^2}{\sigma^2}\right)}$$



Set of Reliable Matches

| Method | Tsukuba | | Sawtooth | | Venus | | Map | |
|---|---|---|---|---|---|---|---|---|
| | Error | Density | Error | Density | Error | Density | Error | Density |
| Our results | 1.18% | 74.5% | 0.27% | 78.4% | 0.20% | 74.1% | 0.08% | 94.2% |
| Sara ECCV02 | 1.4% | 45% | 1.6% | 52% | 0.8% | 40% | 0.3% | 74% |
| Veksler IJCV02 | 0.38% | 66% | 1.62% | 76% | 1.83% | 68% | 0.22% | 87% |
| Veksler CVPR03 | 0.36% | 75% | 0.54% | 87% | 0.16% | 73% | 0.01% | 87% |



Final disparity map

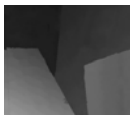
Final error map

Venus

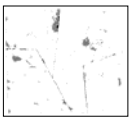
Cones

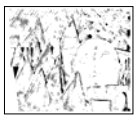
Left Image

Disparity map



Error map





| Dataset | Unoccluded | | Untextured | | Discontinuities | |
|---|---|---|---|---|---|---|
| | Error | Rank | Error | Rank | Error | Rank |
| Tsukuba | 2.19% | 10 | 0.92% | 5 | 11.93% | 11 |
| Sawtooth | 0.53% | 4 | 0.0% | 1 | 4.91% | 6 |
| Venus | 0.36% | 1 | 0.16% | 2 | 5.00% | 4 |
| Map | 0.33% | 9 | - | - | 4.69% | 10 |

| Dataset | Our result | SSD | Dynamic Programming | Graph Cuts |
|---|---|---|---|---|
| Cones | 5.57% | 17.8% | 17.1% | 12.6% |
| Teddy | 9.10% | 26.5% | 30.1% | 29.3% |

## Publications

**Journals:**

M.S. Lee, G. Medioni, and P. Mordohai, "Inference of Segmented Overlapping Surfaces from Binocular Stereo", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 24, no. 6, pp. 824-837, June 2002.

**Conference Proceedings:**

P. Mordohai and G. Medioni, “Stereo using Monocular Cues within the Tensor Voting Framework”, in Proc. European Conference on Computer Vision, Prague, Czech Republic, 2004.

P. Mordohai and G. Medioni, “Dense Multiple View Stereo with General Camera Placement using Tensor Voting”, submitted to the Second International Symposium on 3-D Data Processing, Visualization and Transmission, 2004.

P. Mordohai, G. Medioni, and M.S. Lee, "Inference of Segmented Overlapping Surfaces from Binocular and Multiple-View Stereo", Third Workshop on Perceptual Organization in Computer Vision, 8th IEEE International Conference on Computer Vision, Vancouver, British Columbia, Canada, 2001.

P. Mordohai and G. Medioni, "Perceptual Grouping for Multiple View Stereo using Tensor Voting", in Proc. International Conference on Pattern Recognition, Quebec City, Quebec, Canada, vol. 3, pp. 639-644, 2002.