Stereo Matching Using Belief Propagation

Paper Reading for CIS II Course Project

Xiang Xiang

Mentors: Dr. Dan Mirota, Dr. Greg Hager and Dr. Russ Taylor

Computer Science, JHU

March 14, 2013

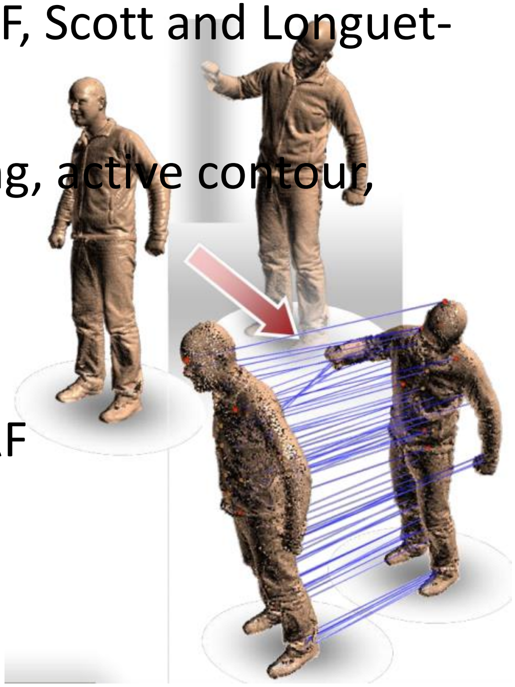
Outline

- Background
- **Paper Review: Approach**
- Discussion

Matching

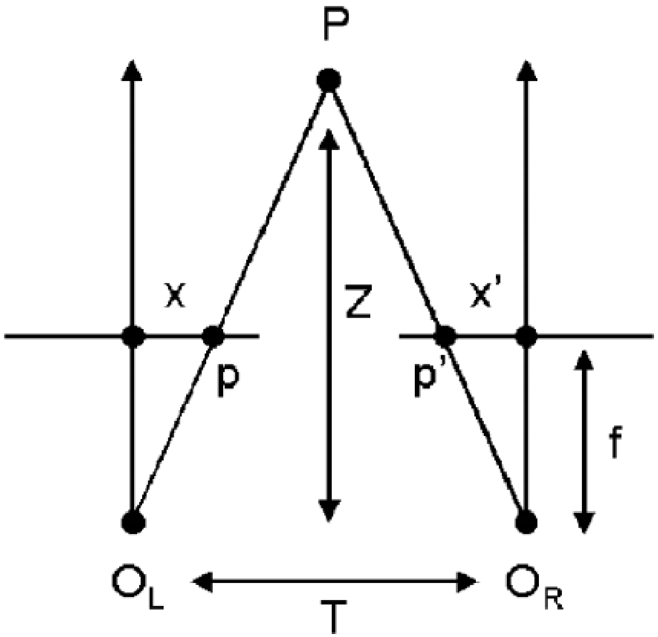
Matching, correspondence, alignment, registration, …

- String matching: theoretic CS, brute force, KMP.
- Color matching: color space, perception theory
- **Point** (cloud) **matching**: image registration, ICP, kd-tree
- **Feature** (point) **matching**: NNDR, kd-tree BBF, Scott and Longuet-Higgins
- Shape matching: distance transform, chamfer matching, active contour, level set, shape context
- Surface matching: 3D shape matching, rendering
- **Stereo matching**: block matching, DP, MRF
- Temporal matching: HMM, DWT



Geometry (1/2)

Assumption: parallel optic axis

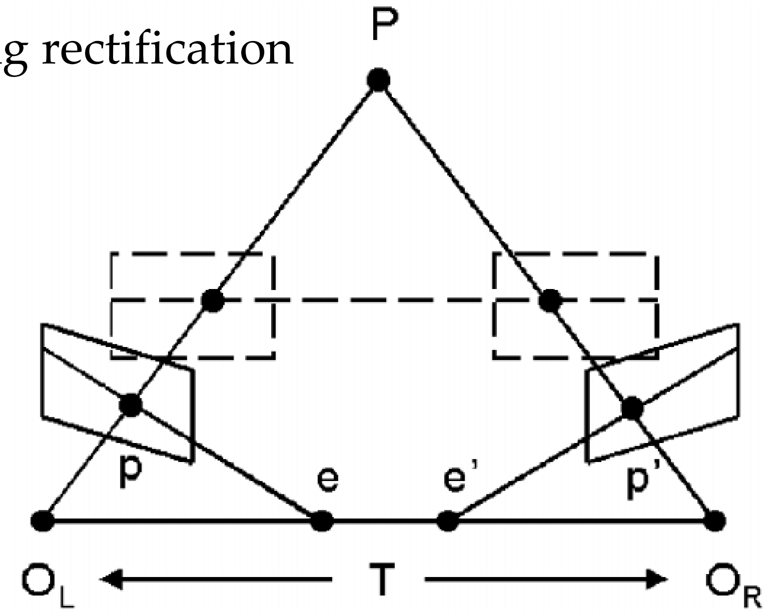


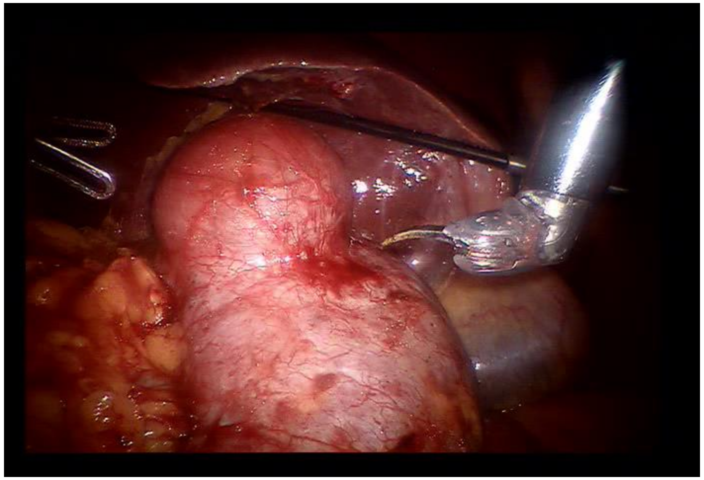
$$d = x - x'$$

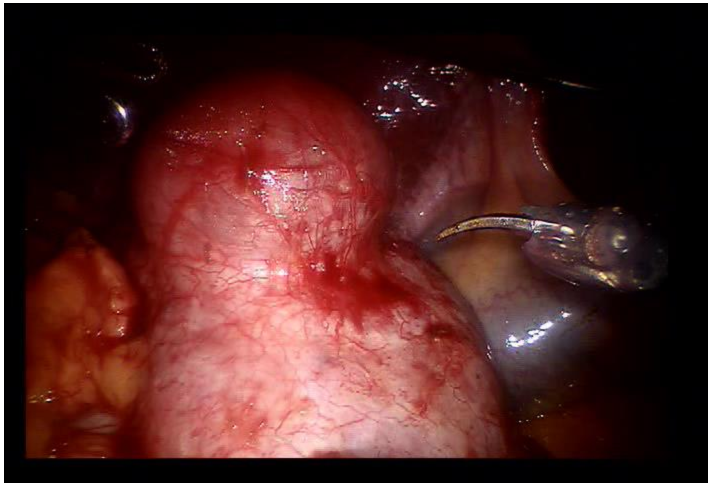
$$Z = f\frac{T}{d}$$

Geometry (2/2)

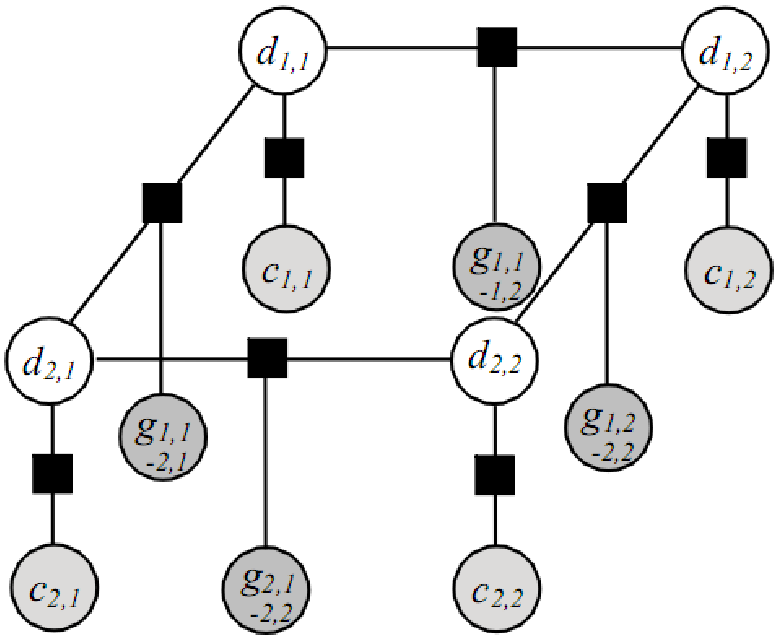
Loosen assumption using rectification







Markov Random Fields



Stereo algorithms generally perform:
1. matching cost computation;
2. cost (support) aggregation;
3. disparity computation and optimization;
4. disparity refinement.

Joint probability of the MRF:

$$P(\mathcal{D},\mathcal{C},\mathcal{G}) = \prod_{p\in\mathcal{P}} \Phi(d_p, \mathbf{c}_p) \prod_{(p,q)\in\mathcal{N}} \Psi(d_p, d_q, g_{pq})$$

Disparity optimization: M.L.E by (Loopy) Belief Propagation

Max Likelihood in Belief Propagation (BP) == Min Energy in Graph cuts

$$E(f) = E_{data}(f) + E_{smooth}(f)$$

$$= \sum_{p \in \mathcal{P}} U(d_p, \mathbf{c}_p) + \sum_{(p,q) \in \mathcal{N}} V(d_p, d_q, g_{pq})$$

$$= \sum_{p \in \mathcal{P}} -\log \Phi(d_p, \mathbf{c}_p) + \sum_{(p,q) \in \mathcal{N}} -\log \Psi(d_p, d_q, g_{pq})$$

$$P(\mathcal{D}, \mathcal{C}, \mathcal{G}) = \prod_{p \in \mathcal{P}} \Phi(d_p, \mathbf{c}_p) \prod_{(p,q) \in \mathcal{N}} \Psi(d_p, d_q, g_{pq})$$

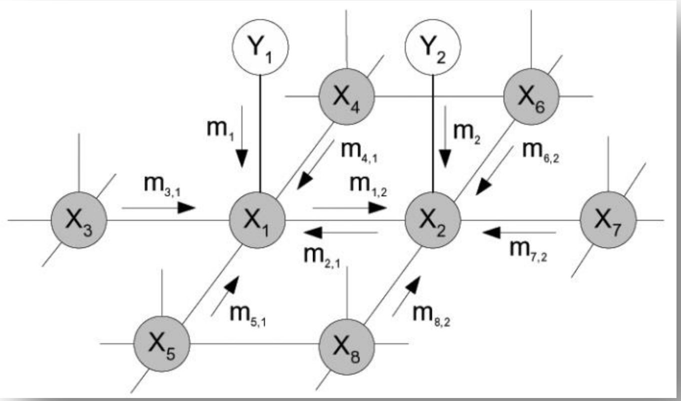
Taking negative log-likelihood:

$$-\log P(\mathcal{D}, \mathcal{C}, \mathcal{G}) = \sum_{p \in \mathcal{P}} -\log \Phi(d_p, \mathbf{c}_p) + \sum_{(p,q) \in \mathcal{N}} -\log \Psi(d_p, d_q, g_{pq})$$

Paper Review: Approach

Stereo Matching Using Belief Propagation

PAMI, 25(7),2003 ECCV 2002



Jian Sun — *XJTU, China (MSRA, TR 35)*

Nan-Ning Zheng — *XJTU, China (XJTU, F-IEEE)*

Heung-Yeung Shum — *MSRA, China (MS VP, F-IEEE)*

Min Energy => Max Likelihood => Max a Posterior

Bayesian Model

D: Disparity field
L: Line process field
O: Occlusion field

$$P(D, L, O|I) = \frac{P(I|D, L, O)P(D, L, O)}{P(I)}$$

Prior (independence assumption)

$$P(D, O, L) = P(D, L)P(O)$$

Markov property

$$P(D, L, O) \propto \prod_s \prod_{t \in N(s)} \exp(-\varphi_c(d_s, d_t, l_{s,t})) \prod_s \exp(-\eta_c(o_s))$$

$l_{s,t}$. $l_{s,t}$ is the line variable between d_s and d_t

$$\varphi_c(d_s, d_t, l_{s,t}) = \varphi(d_s, d_t)(1 - l_{s,t}) + \gamma(l_{s,t})$$

Likelihood (independence assumption)

$$P(I|D, O, L) = P(I|D, O)$$

Likelihood (i.i.d. assumption)

$$P(I|D, O) \propto \prod_{s \notin O} \exp(-F(s, d_s, I))$$

Combine all

$$(P(D, O, L|I) \propto \prod_{s \notin O} \exp(-F(s, d_s, I)) \prod_s \exp(-\eta_c(o_s)) \prod_s \prod_{t \in N(s)} \exp(-(\varphi(d_s, d_t)(1 - l_{s,t}) + \gamma(l_{s,t})))$$

$F(s, d_s, I)$ is the matching cost function of pixel s with disparity d_s given observation I

$$F(s, d_s, I) = \min\{\overline{d}(s, s', I)/\sigma_f, \overline{d}(s', s, I)/\sigma_f\}$$

MAP ≈ MLE

$$(P(D,O,L|I) \propto \prod_{s \notin O} \exp(-F(s,d_s,I)) \prod_{s} \exp(-\eta_c(o_s)) \prod_{s} \prod_{t \in N(s)} \exp(-(\varphi(d_s,d_t)(1-l_{s,t}) + \gamma(l_{s,t})))$$

$$P(\mathcal{D},\mathcal{C},\mathcal{G}) = \prod_{p \in \mathcal{P}} \Phi(d_p, \mathbf{c}_p) \prod_{(p,q) \in \mathcal{N}} \Psi(d_p, d_q, g_{pq})$$

M.A.P.

$$\max_{D,L,O} P(D,L,O|I) =$$

$$\max_{D} \left\{ \max_{O} \prod_{s} \exp(-(F(s,d_s,I)(1-o_s) + \eta_c(o_s)o_s)) \max_{L} \prod_{s} \prod_{t \in N(s)} \exp(-(\varphi(d_s,d_t)(1-l_{s,t}) + \gamma(l_{s,t}))) \right\}$$

Relaxation

$$max_{O} \prod_{s} \exp(-(F(s,d_s,I)(1-o_s^a) + \eta_c(o_s^a)o_s^a))$$

$$= \exp(-\min_{O} \sum_{s} (F(s,d_s,I)(1-o_s^a) + \eta_c(o_s^a)o_s^a))$$

Robust estimator

$$\rho_d(d_s) = \min_{o_s^a}(F(s,d_s,I)(1-o_s^a) + \eta_c(o_s^a)o_s^a)$$

$$\rho_p(d_s,d_t) = \min_{l_{s,t}^a}(\varphi(d_s,d_t)(1-l_{s,t}^a) + \gamma(l_{s,t}^a))$$

$$\max_{D,L,O} P(D,L,O|I) = \max_{D} P(D|I) =$$

$$\max_{D} \left\{ \prod_{s} \exp(-\rho_d(d_s)) \prod_{s} \prod_{t \in N(s)} \exp(-\rho_p(d_s,d_t)) \right\}$$

Approximate MAP by BP

Simplify notation

$$\prod_s \exp(-\rho_d(d_s)) \prod_s \prod_{t \in N(s)} \exp(-\rho_p(d_s, d_t))$$

$$\prod_{p \in \mathcal{P}} \Phi(d_p, \mathbf{c}_p) \prod_{(p,q) \in \mathcal{N}} \Psi(d_p, d_q, g_{pq})$$

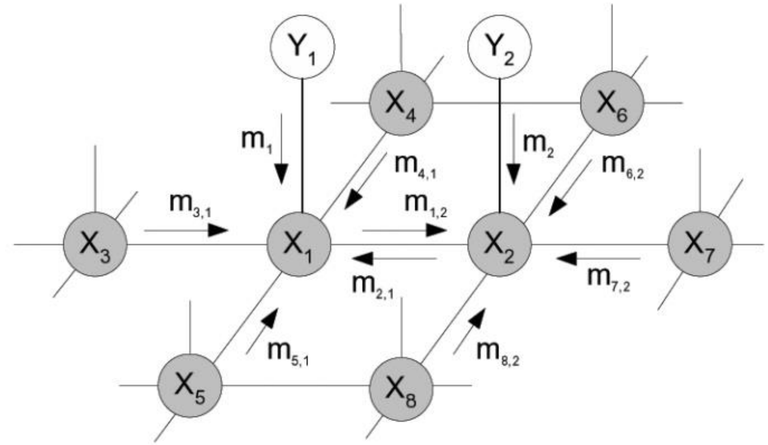
$$\prod_s \psi_s(x_s, y_s) \prod_s \prod_{t \in N(s)} \psi_{st}(x_s, x_t)$$

$$\psi_{st}(x_s, x_t) = m_{st}(x_s, x_t) = m_{st}(x_t)$$

$$\psi_s(x_s, y_s) = m_s(x_s, y_s) = m_s(x_s)$$

Power of Bayesian model:
incorporating multi-priors

Multi-view extension:
Modifying matching cost *F*



Loopy BP by Max-Product

1. Initialize all messages $m_{st}(x_t)$ as uniform distributions and messages $m_s(x_s) = \psi_s(x_s, y_s)$.

2. Update messages $m_{st}(x_t)$ iteratively for i = 1:T

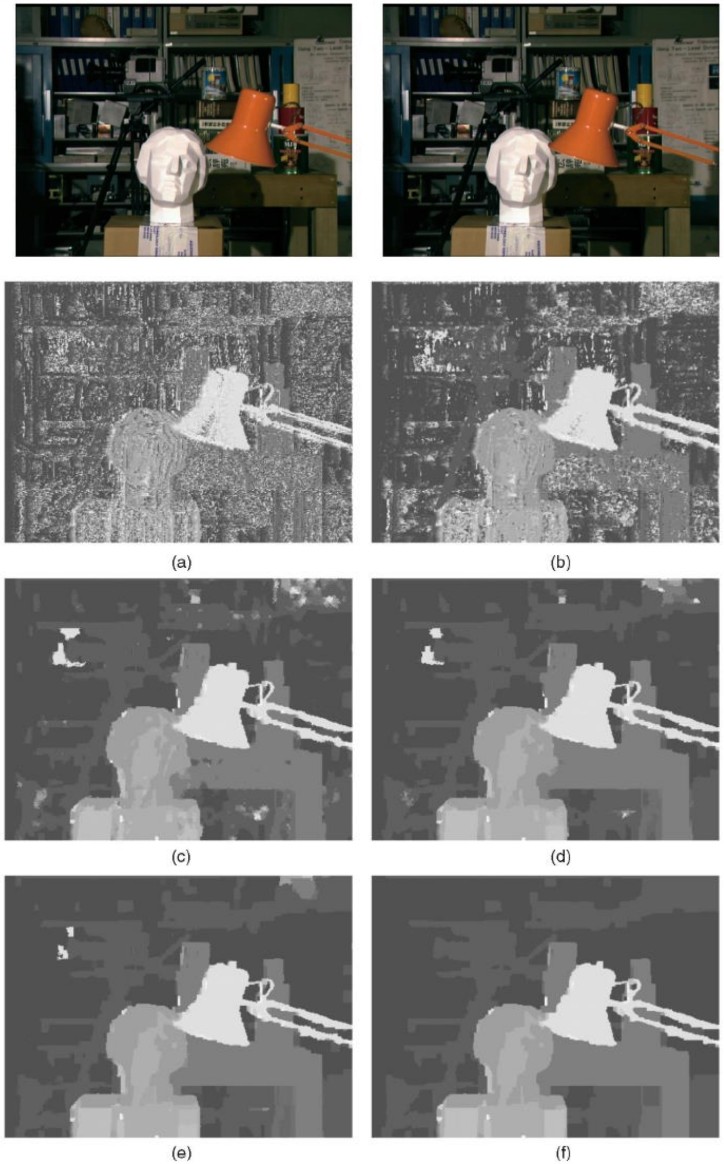
$$m_{st}^{i+1}(x_t) \leftarrow \kappa \max_{x_s} \psi_{st}(x_s, x_t) m_s^i(x_s) \prod_{x_k \in N(x_s) \backslash x_k} m_{ks}^i(x_s).$$

3. Compute beliefs

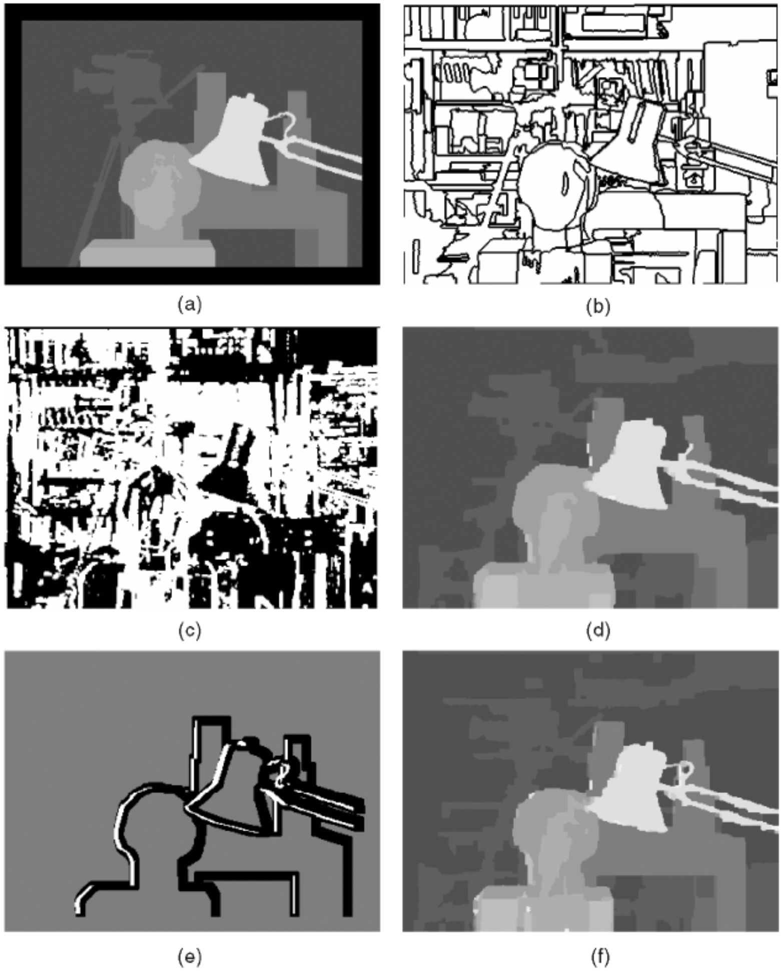
$$b_s(x_s) \leftarrow \kappa m_s(x_s) \prod_{x_k \in N(x_s)} m_{ks}(x_s)$$

$$x_s^{MAP} = \arg\max_{x_k} b_s(x_k).$$

Results



(a) 0
(b) 1
(c) 8
(d) 16
(e) 32
(f) 64



(a) Ground truth. (b) Image segmentation result. (c) Textureless regions. (d) Max-product result without segmentation. (e) Discontinuity (white) and occlusion (black) regions. (f) Max-product result with segmentation.

Discussion

Reference

[1] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via Graph cuts. In IEEE ICCV, 1999.

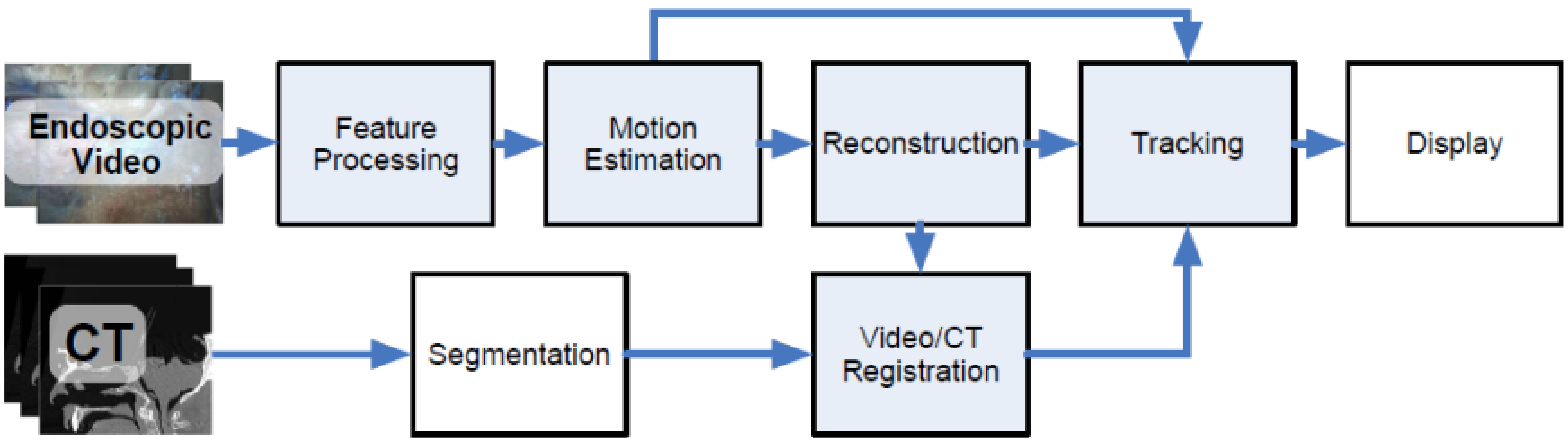
[2] M. Z. Brown, D. Burschka, and G. D. Hager. Advances in computational stereo. IEEE T-PAMI, 25(8), 2003.

[3] J. Sun, N. Zheng, and H. Shum. Stereo Matching Using Belief Propagation. *IEEE T-PAMI*, 25(7),2003.

[4] G. Puerto-Souza and G. Mariottini. A Fast and Accurate Feature-Matching Algorithm for Minimally Invasive Endoscopic Images, IEEE T-MI, iss. 99, Jan. 2013.

[5] D. Scharstein and C. Pal. Learning conditional random fields for stereo. In IEEE CVPR, 2007.

[6] D. Mirota, H. Wang, R. Taylor, M. Ishii, G. Gallia and G. Hager. A System for Video-Based Navigation for Endoscopic Endonasal Skull Base Surgery. IEEE T-MI, 31(4), 963-976, 2012.



Thank you! Comments!