# Robust Visual Motion Analysis:
## Piecewise-Smooth Optical Flow and Motion-Based Detection and Tracking


**Ming Ye**

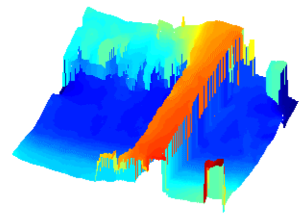
Electrical Engineering, University of Washington

mingye@u.washington.edu


## Structure From Motion

Rigid scene + camera translation

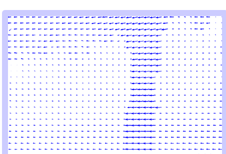
Estimated horizontal motion



Depth map

## What Is Visual Motion

- **2D image velocity**
  - 3D motion projection
  - Temporal correspondence
  - Image deformation

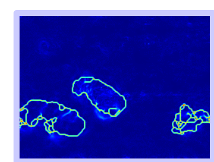


- **Optical flow**
  - An image of 2D velocity
  - Each pixel $V_{s=(x,y)} = (u_s, v_s)$
  - $(x,y,t) \Leftrightarrow (x+u,y+v,t+1)$

## Scene Dynamics Understanding

Estimated horizontal motion

- **What're moving? How?**
  - Surveillance
  - Event analysis
  - Video compression



Motion smoothness

## Target Detection and Tracking

A tiny airplane --- only observable by its distinct motion

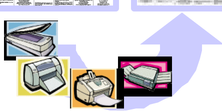
Tracking results

## Research Areas

- Structure from motion
- Scene dynamics analysis
- Object detection and tracking
- Video compression
- Image/video enhancement
- Image-based rendering

- Visual motion estimation

## Image Distortion Measurement

- Image deformation
  - Measure it. Remove it.
  - Image-based rendering

## Outline

- Optical flow estimation
  - Background
  - A local method with error analysis
  - A Bayesian approach with global optimization
- Motion-based detection and tracking

# Optical Flow Estimation

# Previous Work (1/2)

- **Brightness conservation**
  - Matching-based $I(x,y,t) = I(x+u, y+v, t+1)$
  - Gradient-based $I_x u + I_y v + I_t = 0$ (OFC)
- **Flow smoothness**
  - Local parametric $AV = b:$ $\begin{pmatrix} I_{x_1} & I_{y_1} \\ \vdots & \vdots \\ I_{x_N} & I_{y_N} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = -\begin{pmatrix} I_{t_1} \\ \vdots \\ I_{t_N} \end{pmatrix}$ [Lucas-Kanade 81] [Haralick-Lee 83]
  - Global optimization

$$\operatorname{argmin} \sum_{\text{all sites } s} \{(I_{x_s} u_s + I_{y_s} v_s + I_{t_s})^2 + l \sum_{n \in N_s^4} [(u_s - u_n)^2 + (v_s - v_n)^2]\}$$

[Horn-Schunck 81]

# Basics

- **Template matching**

- **Assumptions:**
  - Brightness conservation
  - Flow smoothness
- **Difficulties:**
  - Aperture problem (local information insufficient)
  - Outliers (motion boundaries, abrupt image noise)

# Previous Work (2/2)

- **Handle motion discontinuities & Outliers**
  - Robust statistics **[Black–Anandan 96]**

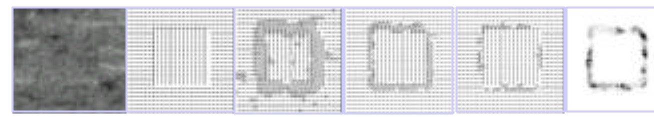
$$\arg\min \sum_{\text{all sites } s} \{r(I_{x_s} u_s + I_{y_s} v_s + I_{t_s}, s_B) + l \sum_{n \in N_s^4} [r(u_s - u_n, s_S) + r(v_s - v_{n_j}, s_S)]\}$$

  - Many others
- **Higher-level methods**
- **Problems:**
  - Gradient calculation
  - Global formulation: $s_B, s_S, l$ values?
  - Computational complexity

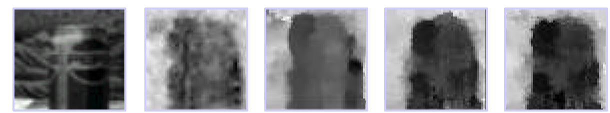
# Two-Stage Robust Optical Flow Estimation with Error Propagation

**A Local Approach**

# Results



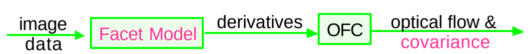
Sampled by2: True LS-LS LS-R R-R Confidence



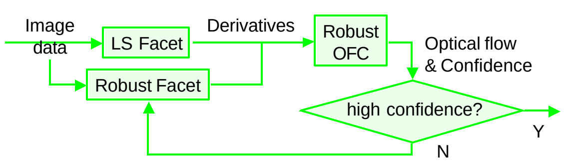
Horizontal flow: M-OFC LS-LMedS LS-R R-R

# Method

- **2-stage regression (LS)** [Haralick-Lee 83, Ye-Haralick 98]

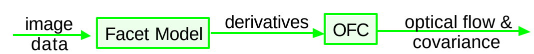



- **Previous: robust OFC only**
- **2-stage-robust adaptive scheme** [ Ye-Haralick 00]




# Error Analysis

- **Covariance propagation** [Haralick 96]
  - (Approx.) linear system + small errors



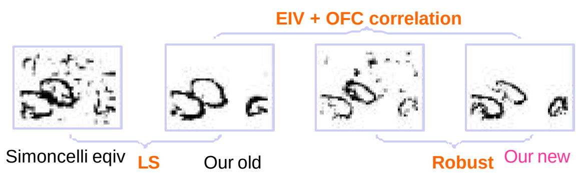

- **Previous work**

| | Image noise var. | EIV | OFC corr. |
|---|---|---|---|
| Simoncelli 91 | No | No | No |
| Szeliski 89 | Yes | No | No |
| Nagel 94 | No | Yes | Yes |
| Ye-Haralick 98 | Yes | Yes | Yes |

- **New: reject outliers first**

## Results

- **A simple motion boundary detector**




- **Error analysis: why bother**
  - Accurate uncertainty is just as important
  - Uncertainty is anisotropic, varies from site to site

## Problem Statement

*Assuming only brightness conservation and piecewise-smooth motion, find the optical flow to best describe the intensity change in three frames.*

# Estimating Piecewise-Smooth Optical Flow with Global Matching and Graduated Optimization

**A Bayesian Approach**

## MAP/MRF Formulation

- **Maximum A Posterior Criterion:**

$$\tilde{V} = \operatorname{argmax}_V P(V/D) = \operatorname{argmax}_V \underbrace{P(D|V)}_{\text{Likelihood}}\underbrace{P(V)}_{\text{Prior}}$$

- **Prior: Markov Random Fields**
  - Neighborhood system: 8-connected $N_s^8$, pairwise
  - Gibbs distribution equivalent

$$P(V) = \exp(-E_s(V))/Z, \quad E_s(V) = \sum_{n \in N_s^8} r(|V_s - V_n|, s_{S_s})$$

- **Likelihood: exponential**
- **Global optimization problem**

## Global Energy Design

- **Global energy** $E = \sum_{\text{allsitess}} E_B(V_s) + E_S(V_s)$
- **Matching error** $E_B(V_s) = r(e_W(V_s), s_{B_s})$
  - Warping error $e_W(V_s) = \min(|I^-(V_s) - I_s|, |I^+(V_s) - I_s|)$
  - 3-Frame Matching Without aliasing, all pixels in a frame are visible in the previous or the next frame.
- **Smoothness error** $E_S(V_i) = \frac{1}{8} \sum_{n \in N_s^8} r(|V_s - V_n|, s_{S_s})$

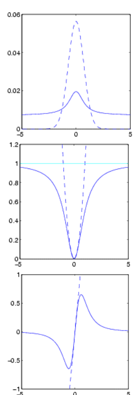
## Advantages

- **Compare with [Black-Anandan 96]**

$$\arg\min \sum_{\text{allsitess}} \{ r(I_{x_s} u_s + I_{y_s} v_s + I_{t_s}, s_B) + l \sum_{n \in N_s^4} [ r(u_s - u_n, s_S) + r(v_s - v_{n_j}, s_S) ] \}$$

| | Proposed | Black-Anandan 96 |
|---|---|---|
| Brightness constr | Matching-based | Gradient-based |
| Scales $s_B, s_S$ $l$ | Local adaptive | Rigid+tuning |
| Contral para | Constant | Tuning |

## Error Function $r(x, s)$:




- **A distribution with fatter tails**
- **An error norm less drastic than L2**
  - Robust against outliers
  - Simultaneous segmentation
    - Smoothness outliers = motion discontinuities
- **Use Geman-McClure for redescending & normalization**

$$r(x, s) = \frac{x^2}{s^2 + x^2} \qquad y(x, s) = r'(x, s) = \frac{2xs}{(s^2 + x^2)^2}$$

## Solution Technique

- **Largescale nonconvex problem**
  - Statistical relaxation: slow
  - Graduated NonConvexity: LS initialization, scales control annealing
- **Our strategy**
  - Fastest descent
  - 3-step graduated optimization
  - Two sub-optimal formulations
    - Provide robust initial estimates
    - Gradually learn the local parameters

## I: OFC-Based Local Regression

- **Lucas-Kanade constraint:** $AV = b$
- **High-breakdown criterion (LMS/LTS)**
- **Fast deterministic algorithm**
  - Least-squares (LS) initial estimate
  - Propagate using an LMS-LS procedure
  - Adaptive outlier resistance
  - Faster, more stable accuracy
- **Estimate scales** $s_{B_s}, s_{S_s}$ **from inliers**

## II: OFC-Based Global Optimization

- **Given** $V, s_{B_s}, s_{S_s}$ **, find** $\Delta V$ **to minimize**

$$E(\Delta V) = \sum_{\text{all sites } s} \{ r(e_B(\Delta V_s), s_{B_s}) + \frac{1}{8} \sum_{n \in N_s^8} r(|V_s + \Delta V_s - V_n - \Delta V_n|, s_{S_s}) \}$$

- **Solution: Successive Over Relaxation**

$$u_{\text{new}} = u_{\text{old}} - w \frac{1}{T(u_{\text{old}})} \frac{\partial E}{\partial u_{\text{old}}}, \quad T(u) = \frac{I_x^2}{s_B^2} + \frac{8}{s_S^2}$$

  - Adaptive step size
  - Initial has dominantly high-freq errors
  - Fast convergence

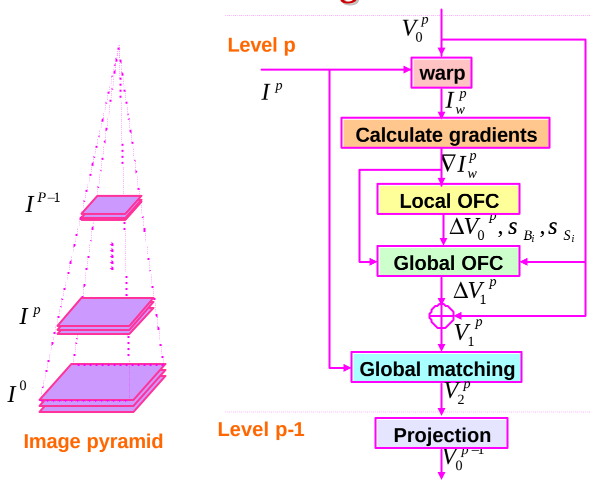
## III: Minimizing the Global Energy

- **Given** $V_{\text{initial}}$
- **Calculate** $s_{B_s}, s_{S_s}$
- **Fastest descent by propagation**
  - Generate candidates: $V_c \in \{V_i, i \in N_s^8; \overline{V_i}\}$
  - Replace $V_s$ by $V_c$ if global energy $E$ drops

## Hierarchical Process

- **Handle large motions (>2 pixels/frame)**
- **Limitations:**
  - Sub-sampling, warping and projection errors
  - May become the accuracy bottleneck
- **Step III directly works on the image data and is less sensitive to such errors**

## Overall Algorithm




## Experiments

## Advantages

- **Best of Everything**
  - Local OFC
    - High-quality initial flow estimates
    - Robust local scale estimates
  - Global OFC
    - Improve flow smoothness
  - Global Matching
    - The optimal formulation
    - Correct errors caused by poor gradient quality and hierarchical process
- **Results: fast convergence, high accuracy, simultaneous motion boundary detection**

## Quantitative Measures

- **True:** $V_0=(u_0,v_0)'$ **, estimate:** $V=(u,v)'$
- **Our error measure**

$$e=(|u-u_0|,|v-v_0|)\big|_{\text{all sites}}$$

  - Cdf curve of $e$, Average: $\bar{e}$
- **Barron's angular error** [Barron 94]

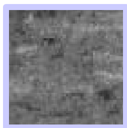
$$e_\angle=\overline{\Delta q}\,(^\circ),\ \Delta q=\arccos\frac{(V_0',1)\cdot(V',1)'}{|(V_0',1)|\cdot|(V',1)|}$$

- **Error magnitude:**

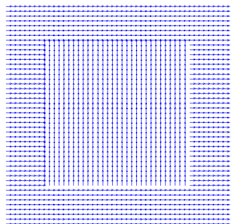
$$e_{|\,|}=\overline{|\Delta V|}\,(\text{pixels}),\ |\Delta V|=|V-V_0|$$

## TS: Translating Squares

- **Homebrew, ideal setting, test performance upper bound**



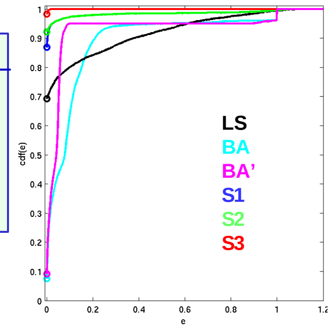
**64x64, 1pixel/frame**



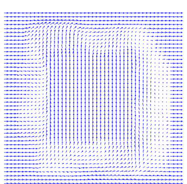
**Groundtruth (cropped), Our estimate looks the same**

## TS: Quantitative Comparison

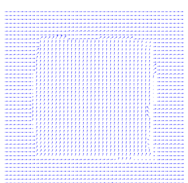
| | $e_\angle(^\circ)$ | $e_{\lvert\cdot\rvert}(\mathrm{pix})$ | $\bar{e}(\mathrm{pix})$ |
|---|---|---|---|
| **LS** | **6.14** | **0.151** | **0.0925** |
| **BA** | **8.04** | **0.209** | **0.120** |
| **BA'** | **5.88** | **0.149** | **0.0815** |
| **S1** | **1.09** | **0.0266** | **0.0180** |
| **S2** | **1.09** | **0.0264** | **0.0179** |
| **S3** | **1.15e-2** | **3.50e-4** | **2.23e-4** |



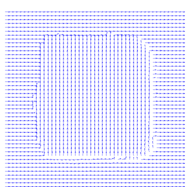

## TS: Flow Estimate Plots



**LS**



**BA**



**S1 (S2 is close)**

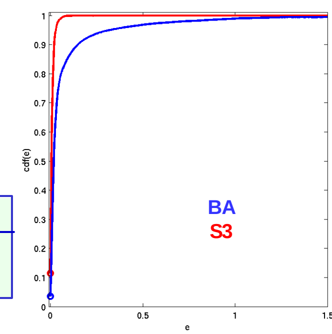
S3 looks the same as the groundtruth.

- S1, S2, S3: results from our Step I, II, III (final)

## TT: Translating Tree

**150x150 (Barron 94)**

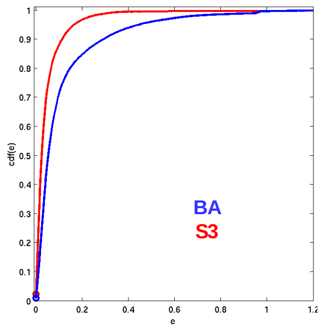
| | $e_\angle(^\circ)$ | $e_{\lvert\cdot\rvert}(\mathrm{pix})$ | $\bar{e}(\mathrm{pix})$ |
|---|---|---|---|
| **BA** | **2.60** | **0.128** | **0.0724** |
| **S3** | **0.248** | **0.0167** | **0.00984** |



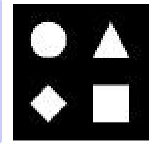
## DT: Diverging Tree

150x150 (Barron 94)

| | $e_\angle(^\circ)$ | $e_{\lvert\bullet\rvert}(\text{pix})$ | $e(\text{pix})$ |
|---|---|---|---|
| BA | 6.36 | 0.182 | 0.114 |
| S3 | 2.60 | 0.0813 | 0.0507 |

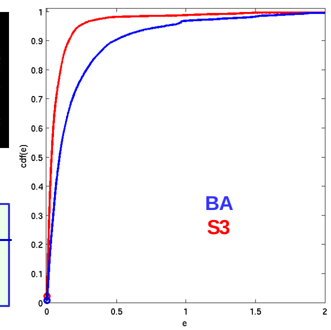



## DTTT: Motion Discontinuities



TT + DT + cookie-cutters

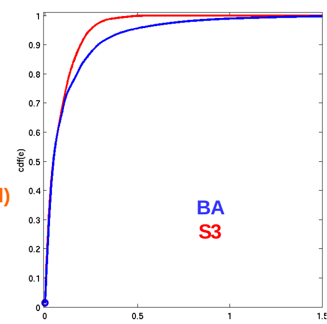
| | $e_\angle(^\circ)$ | $e_{\lvert\bullet\rvert}(\text{pix})$ | $\bar{e}(\text{pix})$ |
|---|---|---|---|
| BA | 10.9 | 0.331 | 0.204 |
| S3 | 4.03 | 0.132 | 0.0807 |




## YOS: Yosemite Fly-Through

316x252 (Barron, cloud excluded)

| | $e_\angle(^\circ)$ | $e_{\lvert\bullet\rvert}(\text{pix})$ | $\bar{e}(\text{pix})$ |
|---|---|---|---|
| BA | 2.71 | 0.185 | 0.118 |
| S3 | 1.92 | 0.120 | 0.0776 |

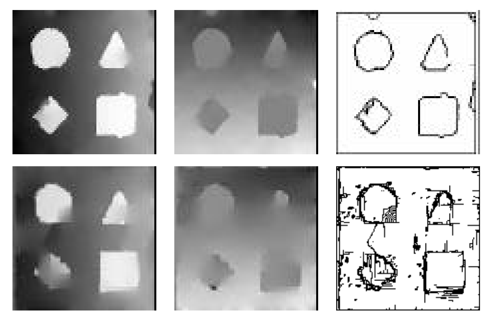



## DTTT

- u-, v-components as intensity images

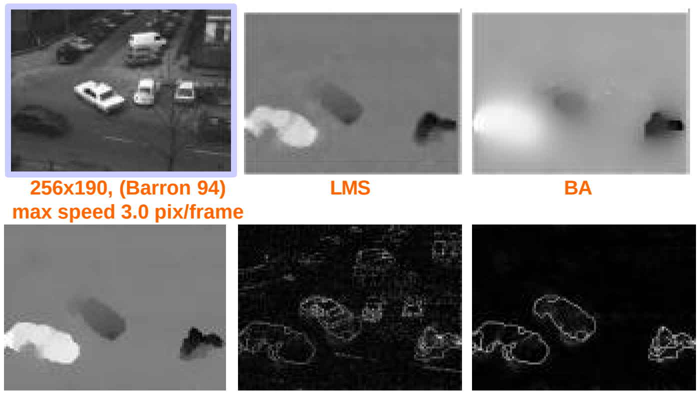
Ours:

BA:



u v Boundary

## TAXI: Hamburg Taxi



256x190, (Barron 94)
max speed 3.0 pix/frame

LMS

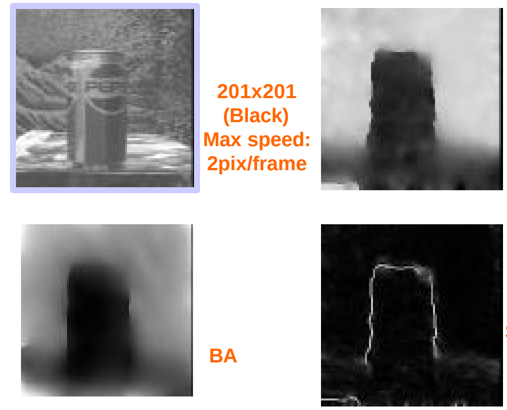
BA

Ours

Error map

Smoothness error

## Pepsi Can



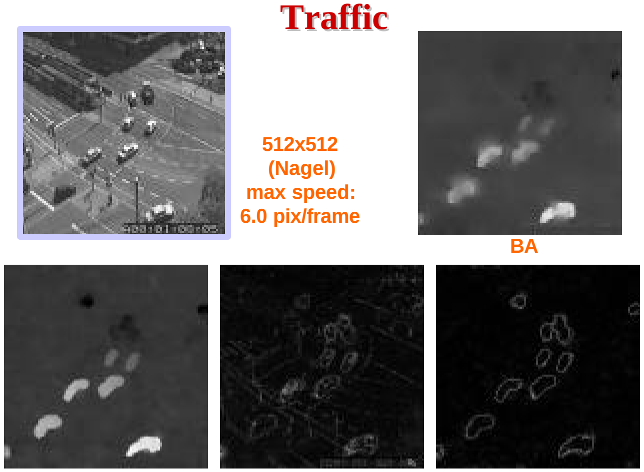
201x201
(Black)
Max speed:
2pix/frame

Ours

BA

Smoothness
error

## Traffic



512x512
(Nagel)
max speed:
6.0 pix/frame

BA

Ours

Error map

Smoothness error

## FG: Flower Garden

360x240 (Black)
Max speed: 7pix/frame

BA

LMS

Ours

Error map

Smoothness error

# Conclusion and Discussion

## Contributions (2/2)

- **Results**
  - High accuracy
  - Fast convergence
  - By product: motion boundaries
- **Significance**
  - Foundation for higher-level (model-based) visual motion analysis
  - Methodology applicable to other low-level vision problems

## Contributions (1/2)

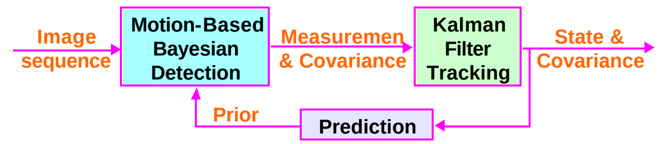
- **Formulation**
  - More complete design, minimal parameter tuning
    - Adaptive local scales
    - Strength of two error terms automatically balanced
  - 3-frame matching to avoid visibility problems
- **Solution: 3-step optimization**
  - Robust initial estimates and scales
  - Model parameter self-learning
  - Inherit merits of 3 methods and overcome shortcomings

## Future Work

- **Applications**
  - Non-rigid motion estimation (medical, human)
  - Higher-level visual motion analysis
    - Motion segmentation, model selection
    - Occlusion reasoning
    - Layered / contour-based representation
  - Warping w/ discontinuities
- **Refinement**
  - Bayesian belief propagation (BBP)
  - Better global optimization (BBP, Graph cuts etc)

# A Motion-Based Bayesian Approach to Aerial Point-Target Detection and Tracking

## The Algorithm



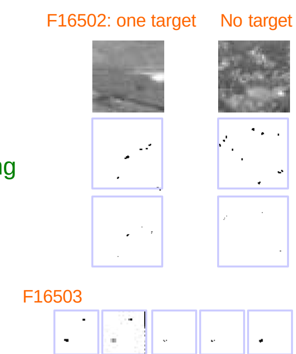

- State variable: 2D position and velocity
- Track initialization, termination and maintenance

## The Problem

- UAV See And Avoid System
- Point target detection and tracking

## Motion-Based Bayesian Detection

- **Background motion:**
  - Parametric optical flow
- **Object candidates:**
  - Fitting outliers
  - Motion: 3x3 SSD + fitting
- **Independent motion**
  - $\chi^2$ test
- **Bayesian mode**
  - Augment candidate set
  - Validate/update motion



# Experiments

- **1800-frame data:**
  - One target 1x2-3x3
  - Clutter (ground objects)
  - Camera wobbling
  - Low image quality
- **Results**
  - Target in track since 2nd frame
  - No false detection
  - Error: mean=0.88, sd=0.44 pixels
- **Show demo**

# Publications

**Patent Pending**

- "Document Image Matching and Annotation Lifting", with Marshall Bern and David Goldberg, *US Patent Application* (filed by Xerox Corp.), September 2001.

**Book Chapter**

- 1. Ming Ye and Robert M. Haralick, "Image Flow Estimation Using Facet Model and Covariance Propagation", *Vision Interface : Real World Applications of Computer Vision (Machine Perception and Artificial Intelligence Book Series Vol. 35)*, (Ed.) M. Cherietand Y. H. Yang, World Scientific Pub Co., pp. 209-241, Jan. 2000.

**Submission/Preparation**

- 2. Ming Ye, Robert M. Haralick and Linda G. Shapiro, "Estimating Piecewise-Smooth Optical Flow with Global Matching and Graduated Optimization", (submitted to) *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Feb. 2002.
- 3. "A motion-based Bayesian approach to aerial point target detection and tracking" (in preparation).

**Conference Papers**

- 4. Ming Ye, Robert M. Haralick and Linda G. Shapiro, " Estimating Optical flow Using a Global Matching Formulation and Graduated Optimization ", (accepted to) *I6th International Conference on Image Processing*, 2002.
- 5. Ming Ye and Robert M. Haralick, "Local Gradient Global Matching Piecewise-Smooth Optical Flow", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 712-717, 2001.
- 6. Ming Ye, Marshall Bern and David Goldberg, " Document Image Matching and Annotation Lifting", *Proc. International Conference on Document Analysis and Recognition*, pp. 753-760, 2001.
- 7. Ming Ye and Robert M. Haralick, "Two-Stage Robust Optical Flow Estimation", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 623-8, 2000.
- 8. Ming Ye and Robert M. Haralick, "Optical Flow From A Least-Trimmed Squares Based Adaptive Approach", *Proc. 15th International Conference on Pattern Recognition*, Vol. 3, pp. 1052-1055, 2000.
- 9. S. Aksoy, M. Ye, M. Schauf, M. Song, Y. Wang, R. M. Haralick, J. R. Parker, J. Pivovarov, D. Royko, S. Sun and S. Farneback, "Algorithm Performance Contest", *Proc. 15th International Conference on Pattern Recognition*, Vol. 4, pp. 870-876, 2000. ICPR'00
- 10. Ming Ye and Robert M. Haralick, "Image Flow Estimation Using Facet Model and Covariance Propagation", *Proc. Vision Interface'98*, pp. 51-58, 1998.