A Multi-Layer MRF Model for Object-Motion Detection in Unregistered Airborne Image-Pairs

Cs. Benedek[1,2] T. Szirányi[1,2] Z. Kato[3] J. Zerubia[4]

[1]Distributed Events Analysis Research Group
Computer and Automation Research Institute, Hungary

[2]Faculty of Information Technology
Pázmány Péter Catholic University, Hungary

[3]Institute of Informatics, University of Szeged, Hungary

[4]ARIANA joint Research group INRIA/CNRS/UNSA, France

International Conference on Image Processing, 2007

Outline

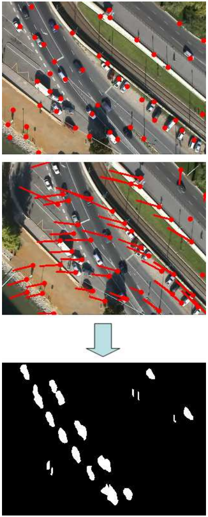
Outline

Introduction

- Object motion detection:
  - image registration for camera motion compensation
  - frame differencing
- Registration techniques
  - Feature correspondence
  - Global 2D transform
    - Pixel correspondence based homography estimation
    - Global correlation methods (Fourier)
  - Plane + parallax models
    - Sparse parallax with epipole estimation
    - Dense parallax and several moving objects

Introduction

- Object motion detection:
    - image registration for camera motion compensation
    - frame differencing
- Registration techniques
    - Feature correspondence
    - Global 2D transform
        - Pixel correspondence based homography estimation
        - Global correlation methods (Fourier)
    - Plane + parallax models
        - Sparse parallax with epipole estimation
        - Dense parallax and several moving objects



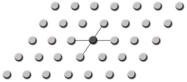
Introduction

- Object motion detection:
  - image registration for camera motion compensation
  - frame differencing
- Registration techniques
  - Feature correspondence
  - Global 2D transform
    - Pixel correspondence based homography estimation
    - Global correlation methods (Fourier)
  - Plane + parallax models
    - Sparse parallax with epipole estimation
    - Dense parallax and several moving objects

Notations

- S: input lattice (2D grid of pixels)
  - s: a given pixel

Lattice S



- Images to be compared
  - $\mathcal{X}_1$, $\mathcal{X}_2$: images over S
  - $\mathcal{X}_2^{\dagger}$: registered second frame
- Segmentation: assigning fg or bg labels to the pixels
  - fg: foreground i.e. object displacement
  - bg: background

Outline

Feature definition 1/2

Gray level difference of the registered images $\mathcal{X}_1$ and $\mathcal{X}_2^{\dagger}$

- Gray level difference

$$d(s) = x_2^{\dagger}(s) - x_1(s).$$

- Probabilistic interpretation
  - Background: Gaussian distribution

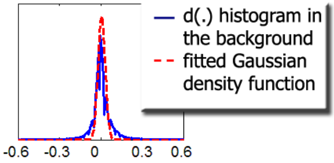
$$P(d(s)|\mathrm{bg}) = N(d(s), \mu, \sigma)$$

  - Foreground: uniform density

$$P(d(s)|\mathrm{fg}) = u_d$$

- Labeling s by the ML estimate:

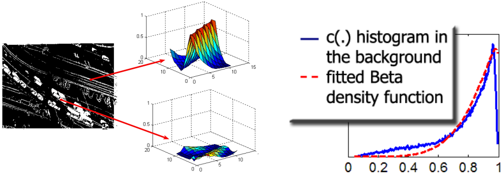
$$\mathrm{argmax}_{\psi \in \{\mathrm{fg},\mathrm{bg}\}} P(d(s)|\psi)$$

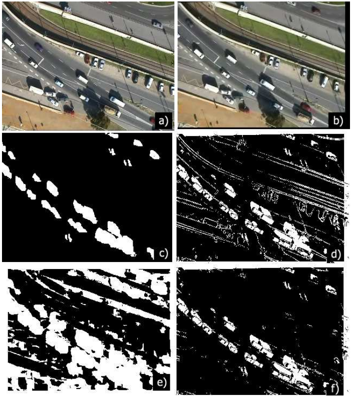


Feature definition 2/2

Local correlation peak value

- Local correlation peak value: $c(s)$





  - Background: Beta distribution

$$P(c(s)|\mathrm{bg}) = \mathcal{B}(c(s), \beta_1, \beta_2)$$

  - Foreground: uniform density

$$P(c(s)|\mathrm{fg}) = u_c$$

- AND operation for results based on $d(.)$ and $c(.)$ features

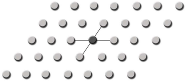
Outline

Graph-based interpretation of the problem

Pixel lattice

- Input lattice: S (2D grid of pixels)
  - s: a given pixel

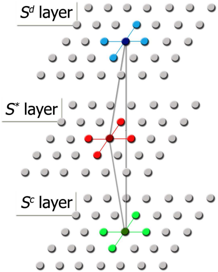
Lattice S



Graph representation

- Graph G: sites are arranged into three layers: S^d, S^c, S^*.
- To each pixel s a unique site is assigned at each layer:
  - $s \rightarrow \{s^d, s^c, s^*\}$
  - e.g. s^d: a site in S^d

Graph G

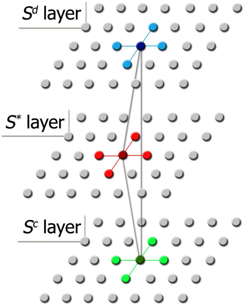


Labeling model on the three-layer graph

- ω: a labeling operator on the sites of G with $\{\mathrm{fg}, \mathrm{bg}\}$ labels
  - fg: foreground i.e. object displacement
  - bg: background

- Semantic of the labeling
  - S^d layer: $\omega(s^d)$ is based on the $d(s)$ observation
  - S^c layer: $\omega(s^c)$ is based on the $c(s)$ observation
  - S^* layer: $\omega(s^*)$ is directly influenced by the $\omega(s^d)$ and $\omega(s^c)$ labels but not by the observations

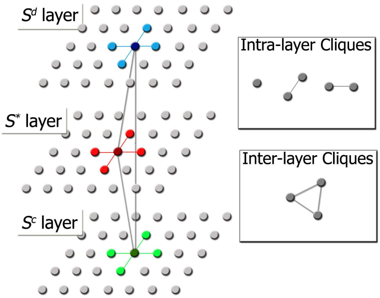
Graph G



Connections in the graph

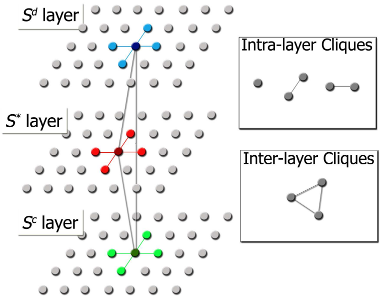
- Connections in G: direct dependencies in the labeling of the neighboring site
  - Let C be an arbitrary clique.
  - V_C clique potential: 'low', if the labeling of C is semantically correct.

- Singletons:
  - $\omega(s^d)$ / $\omega(s^c)$ labels should be consistent with the local observations $d(s)/c(s)$
- Doubleton (inter-layer) cliques:
  - Smooth segmentation is expected at each layer
- Inter-layer cliques
  - "AND constraint" between $\omega(s^d)$, $\omega(s^c)$ and $\omega(s^*)$ should usually hold



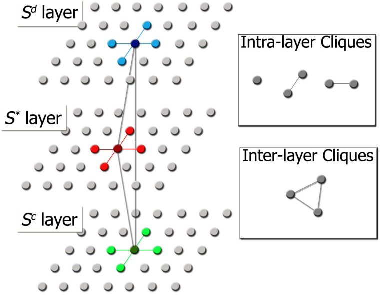
Connections in the graph

- Connections in G: direct dependencies in the labeling of the neighboring site
  - Let C be an arbitrary clique.
  - V_C clique potential: 'low', if the labeling of C is semantically correct.
- Singletons:
  - $\omega(s^d)$ / $\omega(s^c)$ labels should be consistent with the local observations $d(s)/c(s)$
- Doubleton (inter-layer) cliques:
  - Smooth segmentation is expected at each layer
- Inter-layer cliques
  - "AND constraint" between $\omega(s^d)$, $\omega(s^c)$ and $\omega(s^*)$ should usually hold



Connections in the graph

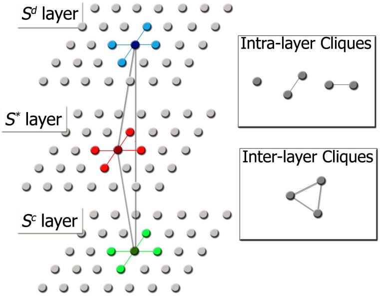
- Connections in G: direct dependencies in the labeling of the neighboring site
  - Let C be an arbitrary clique.
  - V_C clique potential: 'low', if the labeling of C is semantically correct.
- Singletons:
  - $\omega(s^d)$ / $\omega(s^c)$ labels should be consistent with the local observations $d(s)/c(s)$
- Doubleton (inter-layer) cliques:
  - Smooth segmentation is expected at each layer
- Inter-layer cliques
  - "AND constraint" between $\omega(s^d)$, $\omega(s^c)$ and $\omega(s^*)$ should usually hold



Connections in the graph

- Connections in G: direct dependencies in the labeling of the neighboring site
  - Let C be an arbitrary clique.
  - V_C clique potential: 'low', if the labeling of C is semantically correct.

- Singletons:
  - $\omega(s^d)$ / $\omega(s^c)$ labels should be consistent with the local observations $d(s)/c(s)$
- Doubleton (inter-layer) cliques:
  - Smooth segmentation is expected at each layer
- Inter-layer cliques
  - "AND constraint" between $\omega(s^d)$, $\omega(s^c)$ and $\omega(s^*)$ should usually hold



MRF labeling model

- Global labeling: $\underline{\omega} = \left\{\omega(s^i) | s \in S, i \in \{d, c, *\}\right\}$.
- Observation process $\mathcal{F} = \{f(s) | s \in S\}$, where $f(s) = [d(s), c(s)]$
- MAP estimation of the optimal global labeling:

$$\widehat{\underline{\omega}} = \text{argmax}_{\underline{\omega} \in \Omega} P(\underline{\omega} | \mathcal{F})$$

  where Ω denotes the set of all the possible global labelings.
- **(Hammersley-Clifford theorem)**: $P(\underline{\omega} | \mathcal{F})$ can be factorized into individual terms whose domains are the cliques of the graph.

$$P(\underline{\omega} | \mathcal{F}) = \frac{1}{Z} \exp\left(-\sum_{C \in \mathcal{C}} V_C(\underline{\omega})\right)$$

  where C is an arbitrary clique and V_C is the potential of C.

MRF labeling model

- Global labeling: $\underline{\omega} = \left\{ \omega(s^i) | s \in S, i \in \{d, c, *\} \right\}$.
- Observation process $\mathcal{F} = \{f(s) | s \in S\}$, where $f(s) = [d(s), c(s)]$
- MAP estimation of the optimal global labeling:

$$\widehat{\underline{\omega}} = \mathrm{argmax}_{\underline{\omega} \in \Omega} P(\underline{\omega} | \mathcal{F})$$

  where Ω denotes the set of all the possible global labelings.
- **(Hammersley-Clifford theorem)**: $P(\underline{\omega} | \mathcal{F})$ can be factorized into individual terms whose domains are the cliques of the graph.

$$P(\underline{\omega} | \mathcal{F}) = \frac{1}{Z} \exp \left(- \sum_{C \in \mathcal{C}} V_C(\underline{\omega}) \right)$$

  where C is an arbitrary clique and V_C is the potential of C.

MRF labeling model

- Global labeling: $\underline{\omega} = \left\{\omega(s^i) | s \in S, i \in \{d, c, *\}\right\}$.
- Observation process $\mathcal{F} = \{f(s) | s \in S\}$, where $f(s) = [d(s), c(s)]$
- MAP estimation of the optimal global labeling:

$$\underline{\widehat{\omega}} = \mathrm{argmax}_{\underline{\omega} \in \Omega} P(\underline{\omega} | \mathcal{F})$$

where Ω denotes the set of all the possible global labelings.

- **(Hammersley-Clifford theorem)**: $P(\underline{\omega}|\mathcal{F})$ can be factorized into individual terms whose domains are the cliques of the graph.

$$P(\underline{\omega}|\mathcal{F}) = \frac{1}{Z} \exp\left(-\sum_{C \in \mathcal{C}} V_C(\underline{\omega})\right)$$

where C is an arbitrary clique and V_C is the potential of C.

MRF labeling model

- Global labeling: $\underline{\omega} = \left\{\omega(s^i)|s \in S, i \in \{d, c, *\}\right\}$.
- Observation process $\mathcal{F} = \{f(s)|s \in S\}$, where $f(s) = [d(s), c(s)]$
- MAP estimation of the optimal global labeling:

$$\widehat{\underline{\omega}} = \mathrm{argmax}_{\underline{\omega} \in \Omega} P(\underline{\omega}|\mathcal{F})$$

  where Ω denotes the set of all the possible global labelings.

- **(Hammersley-Clifford theorem)**: $P(\underline{\omega}|\mathcal{F})$ can be factorized into individual terms whose domains are the cliques of the graph.

$$P(\underline{\omega}|\mathcal{F}) = \frac{1}{Z} \exp\left(-\sum_{C \in \mathcal{C}} V_C(\underline{\omega})\right)$$

  where C is an arbitrary clique and V_C is the potential of C.

Singleton potentials

- Observation dependent (a posteriori) terms of the energy function
- $V_{\{s^d\}}\left(\omega(s^d)\right) = -\log P(d(s)|\omega(s^d))$
  - e.g. $P(d(s)|\omega(s^d) = \mathrm{bg})$ is the probability that background generates the observed $d(s)$ value
- $V_{\{s^c\}}\left(\omega(s^c)\right) = -\log P(c(s)|\omega(s^c))$
- $V_{\{s^*\}}\left(\omega(s^*)\right) = 0$
  - no direct links with the observation at S^* layer

- Modeling the probabilities:

	bg	fg
$d(s)$	Gauss	uniform
$c(s)$	Beta	uniform

Doubleton potentials

- Doubleton cliques: smoothing priors of the segmentation within each layer.
- The potential of an intra-layer clique $C_2 = \{s^i, r^i\} \in \mathcal{C}_2$, $i \in \{d, c, *\}$:

$$V_{C_2} = \begin{cases} -\delta^i & \text{if } \omega(s^i) = \omega(r^i) \\ +\delta^i & \text{if } \omega(s^i) \neq \omega(r^i) \end{cases}$$

  for a constant $\delta^i > 0$.

Inter-layer clique potentials

- Inter-layer cliques: a pixel is likely to be generated by the background process, if at least one corresponding site has the label 'bg' in the S^d and S^c layers.
- Background indicator function

$$I_{\mathrm{bg}} : S^d \cup S^c \cup S^* \to \{0,\ 1\},$$

where

$$I_{\mathrm{bg}}(q) = \begin{cases} 1 & \text{if } \ \omega(q) = \mathrm{bg} \\ 0 & \text{if } \ \omega(q) \neq \mathrm{bg}. \end{cases}$$

- The potential of an inter-layer clique $C_3 = \{s^d, s^c, s^*\}$:

$$V_{C_3}(\underline{\omega}_{C_3}) = \begin{cases} -\rho & \text{if } \ I_{\mathrm{bg}}(s^*) = \max\left(I_{\mathrm{bg}}(s^d), I_{\mathrm{bg}}(s^c)\right) \\ +\rho & \text{otherwise}, \end{cases}$$

with $\rho > 0$.

Outline

Test datasets and reference methods

- Database: 83 image pairs from three test sets
- Manually generated ground truth masks
- Metrics: F-measure (harmonic mean of precision and recall of foreground detection)
- Reference methods:
  - Gray level difference based MRF
  - Method of Farin and With, ICIP 2005[1]
  - Supervised affine matching

[1] D. Farin and P. With, "Misregistration Errors in Change Detection Algorithms and How to Avoid Them," in *Proc. International Conference on Image Processing (ICIP)*, vol. 2, pp. 438-441, Genoa, Italy, Sept. 2005.

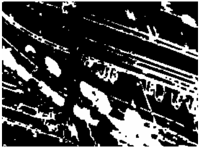
Results

Image 1

Image 2

Ground truth

Only gray level difference



Supervised affine matching

Farin's method

3-layer MRF

Results

Image 1

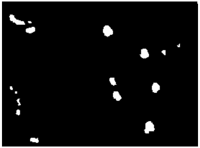
Image 2

Ground truth

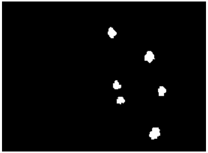
Only gray level difference

Supervised affine matching

Farin’s method



3-layer MRF



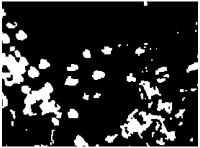
Results

Image 1

Image 2

Ground truth

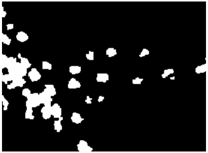
Only gray level difference



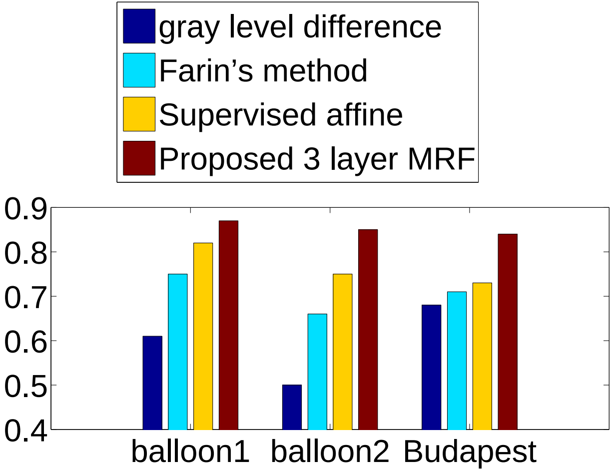
Supervised affine matching

Farin's method

3-layer MRF



Results considering the F-measure



Questions. . .

Thank You for Your Attention!