MTA SZTAKI

Object detection in remote sensing applications

Csaba Benedek

Distributed Events Analysis Research Laboratory (DEVA)
Institute for Computer Science and Control (MTA SZTAKI)
email: benedek.csaba@sztaki.mta.hu
http://web.eee.sztaki.hu/remotesensing/

Summer School on Image Processing
Veszprém, Hungary, 2013

MTA SZTAKI

Content

Introduction

Region level change detection for remote sensing images

Building change detection

Marked Point Process model

Traffic monitoring from aerial Lidar point clouds

Hierarchical PCB Inspections

MTA SZTAKI

Content

Introduction

Region level change detection for remote sensing images

Building change detection

Marked Point Process model

Traffic monitoring from aerial Lidar point clouds

Hierarchical PCB Inspections

MTA SZTAKI

Introduction

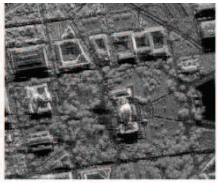
- Remote sensing
  - aerial/satellite measurement and data analysis
- Applications
  - cartography, monitoring of agricultural activity, detection of pollution and environmental crimes, management of urban area expansion

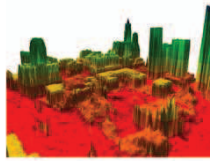
Introduction - remote sensing

- Various data sources and modalities

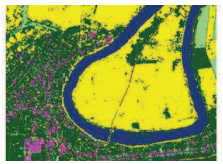
Aerial image

Satellite image


Radar (SAR) image


Lidar 3D mesh

- Progress: quick increase of the available data resolution
  - Low level approaches: region based analysis, pixel level features
  - Object information: data based and prior geometric descriptors



Geographic Information Systems (GIS) - critical issues

- **Automation:** Current GIS systems require deep human intervention and do not enable a fully automated processing
  - huge databases - slow processing
- **Data heterogeneity:** integration of various data sources
  - space borne, airborne and terrestrial RS data
  - optical imagery, Lidar, radar, etc.
- **Various spatial scales of the data:**
  - Hierarchical content interpretation
- **Various temporal scales of the data:**
  - Change detection - application dependent goals and challenges
- **Adaptation:**
  - Quick evolution of data quality - constant adaptation needed

MTA SZTAKI

Geographic Information Systems (GIS) - critical issues

- **Automation:** Current GIS systems require deep human intervention and do not enable a fully automated processing
  - huge databases - slow processing
- **Data heterogeneity:** integration of various data sources
  - space borne, airborne and terrestrial RS data
  - optical imagery, Lidar, radar, etc.
- **Various spatial scales of the data:**
  - Hierarchical content interpretation
- **Various temporal scales of the data:**
  - Change detection - application dependent goals and challenges
- **Adaptation:**
  - Quick evolution of data quality - constant adaptation needed

MTA SZTAKI

Geographic Information Systems (GIS) - critical issues

- **Automation:** Current GIS systems require deep human intervention and do not enable a fully automated processing
  - huge databases - slow processing
- **Data heterogeneity:** integration of various data sources
  - space borne, airborne and terrestrial RS data
  - optical imagery, Lidar, radar, etc.
- **Various spatial scales of the data:**
  - Hierarchical content interpretation
- **Various temporal scales of the data:**
  - Change detection - application dependent goals and challenges
- **Adaptation:**
  - Quick evolution of data quality - constant adaptation needed

Geographic Information Systems (GIS) - critical issues

- **Automation:** Current GIS systems require deep human intervention and do not enable a fully automated processing
  - huge databases - slow processing
- **Data heterogeneity:** integration of various data sources
  - space borne, airborne and terrestrial RS data
  - optical imagery, Lidar, radar, etc.
- **Various spatial scales of the data:**
  - Hierarchical content interpretation
- **Various temporal scales of the data:**
  - Change detection - application dependent goals and challenges
- **Adaptation:**
  - Quick evolution of data quality - constant adaptation needed

MTA SZTAKI

Geographic Information Systems (GIS) - critical issues

- **Automation:** Current GIS systems require deep human intervention and do not enable a fully automated processing
  - huge databases - slow processing
- **Data heterogeneity:** integration of various data sources
  - space borne, airborne and terrestrial RS data
  - optical imagery, Lidar, radar, etc.
- **Various spatial scales of the data:**
  - Hierarchical content interpretation
- **Various temporal scales of the data:**
  - Change detection - application dependent goals and challenges
- **Adaptation:**
  - Quick evolution of data quality - constant adaptation needed

Csaba Benedek (DEVA Laboratory) Object detection in remote sensing applications 8 July 2013 6 / 102

MTA SZTAKI

Content

Introduction

Region level change detection for remote sensing images

Building change detection

Marked Point Process model

Traffic monitoring from aerial Lidar point clouds

Hierarchical PCB Inspections

MTA SZTAKI

Region level change detection

- Large (many years) time differences $\rightarrow$ different seasons, illumination conditions, vegetations etc.
- Aimed solution:
  - featuring *relevant* changes versus irrelevant differences in remotely sensed image pairs:
    - new built-up regions, building operations, planting of trees, fresh plough-land, groundwork before building-over etc
- Input – preliminary registered orthophotos:

Image 1 from 2000

Image 2 from 2005

Shortcommings of conventional frame differencing

- Structural changes may be confused in the color-difference domain – several missing/false changes in the difference map

Image 1

Image 2

Colored difference image (CDI)

Change mask overlaid by thresholded CDI

Ground Truth change mask

MTA SZTAKI

Shortcommings of conventional frame differencing

- Structural changes may be confused in the color-difference domain – several missing/false changes in the difference map

Image 1

Image 2

Colored difference image (CDI)

Change mask overlaid by thresholded CDI

Ground Truth change mask

MTA SZTAKI

Proposed improvements

- Instead of purely color based investigations, **simultaneous usage of radiometric and spatial structure features** for comparison
- Instead of frame differencing **general statistical co-occurrence model** for the pixel values of the different time layers
  - dealing with *unknown* nonlinear radiometric effects (seasonal vegetation changes, color adjustment in post processing)

MTA SZTAKI

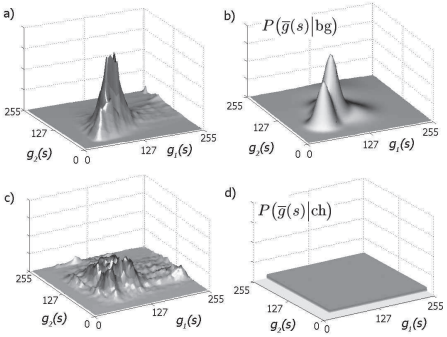
Proposed improvements

- Instead of purely color based investigations, **simultaneous usage of radiometric and spatial structure features** for comparison
- Instead of frame differencing **general statistical co-occurrence model** for the pixel values of the different time layers
  - dealing with *unknown* nonlinear radiometric effects (seasonal vegetation changes, color adjustment in post processing)

Global statistics of color co-occurrences

- Feature vector of pixel s is pair of intensity/color values of s in the two images: $\overline{g}(s) = [g_1(s), g_2(s)]^T$, $g_1(s) \in G_1$, $g_2(s) \in G_2$
- Global statistics in changed/background regions:
  - Class 'background': $P(\overline{g}(s)|\mathrm{bg}) =$

$$= \sum_{i=1}^{K} \kappa_i \cdot \eta(\overline{g}(s), \overline{\mu}_i, \Sigma_i)$$



  - Class 'change': $P(\overline{g}(s)|\mathrm{ch}) =$

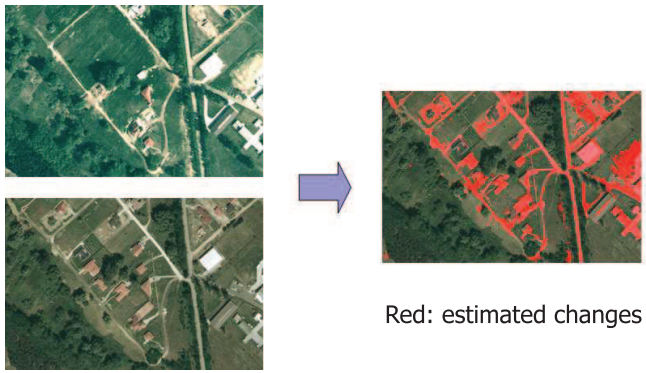
$$= \begin{cases} \frac{1}{(b_1-a_1)\cdot(b_2-a_2)}, & \text{if } \overline{g}(s) \in \Gamma \\ 0 & \text{otherwise} \end{cases}$$

  - $\overline{g}(s) \in \Gamma$ iff $a_1 \le g_1(s) \le b_1$ and $a_2 \le g_2(s) \le b_2$

Performance of the Color Statistics Feature

- Result of the color based ML pixel classification

$$\phi_g(s) = \mathrm{argmax}_{\psi \in \{\mathrm{ch,bg}\}} P(\overline{g}(s)|\psi)$$



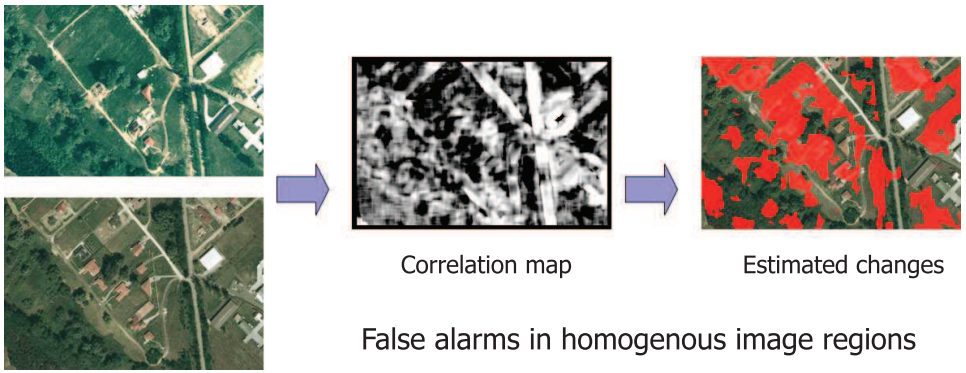
Red: estimated changes

False alarms in **textured** image regions

Local block correlation

- Feature $c(s)$: normalized cross correlation between the rectangular neighborhoods of pixel s in G_1 resp. G_2 images.
- Correlation based ML pixel classification

$$\phi_c(s) = \mathrm{argmax}_{\psi \in \{\mathrm{ch,bg}\}} P(c(s)|\psi)$$

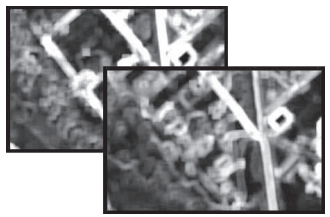


Correlation map Estimated changes

False alarms in homogenous image regions

Feature integration

- Feature selection based on local contrast:
  $\overline{\nu}(s) = [\nu_1(s), \nu_2(s)]^T$
  - $\nu_i(s)$, $i \in \{1, 2\}$: variance of the gray levels over the $v \times v$ neighborhood of s in G_i

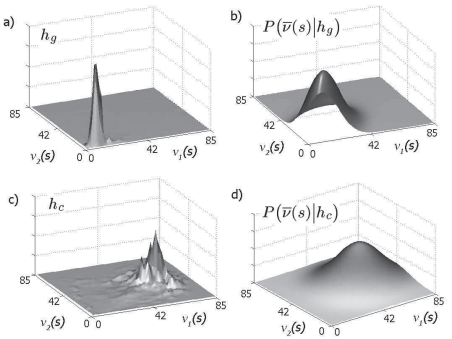


$\nu_1(.)$ $\nu_2(.)$

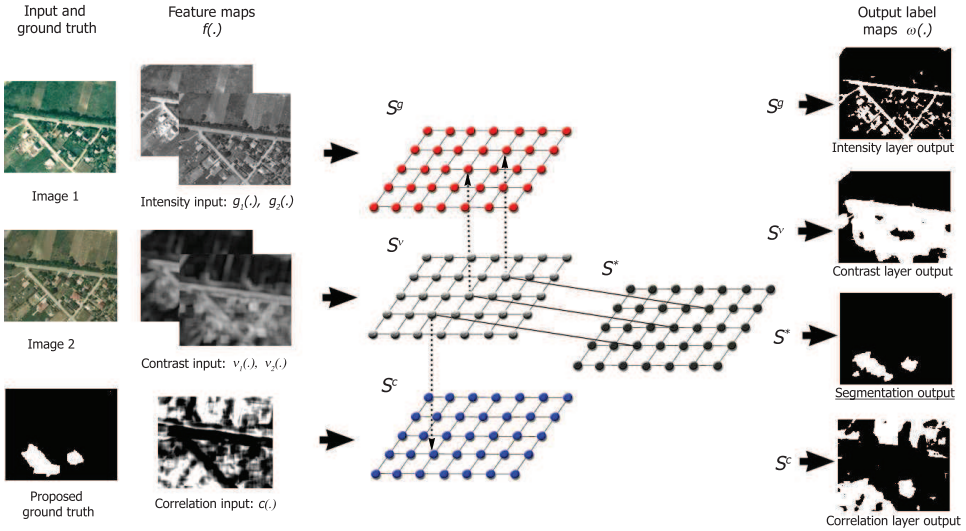
- Gaussian models for the reliability of the g/c features:

$$P(\overline{\nu}(s)|h_g) = \eta\left(\overline{\nu}(s), \overline{\mu}_g, \overline{\overline{\Sigma}}_g\right)$$

$$P(\overline{\nu}(s)|h_c) = \eta\left(\overline{\nu}(s), \overline{\mu}_c, \overline{\overline{\Sigma}}_c\right)$$



CXM model structure



Probability modeling in CXM

- Global labeling: $\underline{\omega} = \{\omega_s | s \in S\}$
- Observation process: $\mathcal{F} = \{f_s | s \in S\}$
- MAP estimation of the optimal global labeling in space Ω:

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

- Data-prior probability decomposition:

$$P(\underline{\omega}|\mathcal{F}) \propto \underbrace{\prod_{s \in S} P(f_s|\omega_s)}_{P(\mathcal{F}|\underline{\omega})} \cdot \underbrace{\frac{1}{Z} \exp\left(-\sum_{C \in \mathcal{C}} V_C(\omega_C, \omega_C^{\mathcal{A}})\right)}_{P(\underline{\omega})}$$

- ω_C: set of labels inside clique C: $\omega_C = \{\omega(q) | q \in C\}$
- $\omega_C^{\mathcal{A}}$: the set of node labels pointed by the address nodes of C:

$$\omega_C^{\mathcal{A}} = \{\tilde{\omega}(a) \big| a \in \mathcal{A} \cap C, \omega(a) \neq \text{nil}\}$$

- $\mathcal{A}$: set of address nodes and $\tilde{\omega}(a) = \omega(\omega(a))$ for $a \in \mathcal{A}$

Labeling optimization

- MAP estimation of the optimal global labeling $\hat{\underline{\omega}}$:

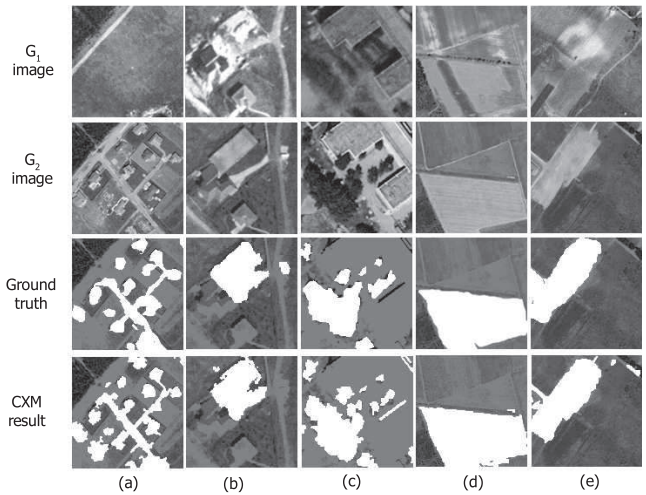
$$\hat{\underline{\omega}} = \arg\min_{\underline{\omega} \in \Omega} \left\{ \sum_{s \in S} -\log P\left(\overline{g}(s) | \omega(s^g)\right) + \right.$$

$$+ \sum_{s \in S} -\log P\left(c(s) | \omega(s^c)\right) + \sum_{s \in S} -\log P\left(\overline{v}(s) | \omega(s^v)\right) +$$

$$\left. + \sum_{i;\{s,r\} \in \mathcal{C}_2} V_{C_2}\left(\omega(s^i), \omega(r^i)\right) + \sum_{s \in S} V_{C_3}\left(\omega(s^*), \tilde{\omega}(s^v)\right) \right\}$$

- Cliques and clique potentials:
  - Singletons: data – label consistency
  - Intra-layer connections: smooth label maps V_{C_2}
  - Inter-layer interactions: label fusion V_{C_3}
- Optimization by simulated annealing
- Output: labeling of the S^* layer.

Ground truth generation



Change prototypes considered for ground truth generation (a) new built-up regions (b) building operations (c) planting of trees (d) fresh plough-land (e) groundwork before building over

CXM comparison to frame differencing

Image 1

Image 2

Ground Truth change mask

Change mask overlaid by thresholded CDI

CXM result

CXM comparison to frame differencing

Image 1

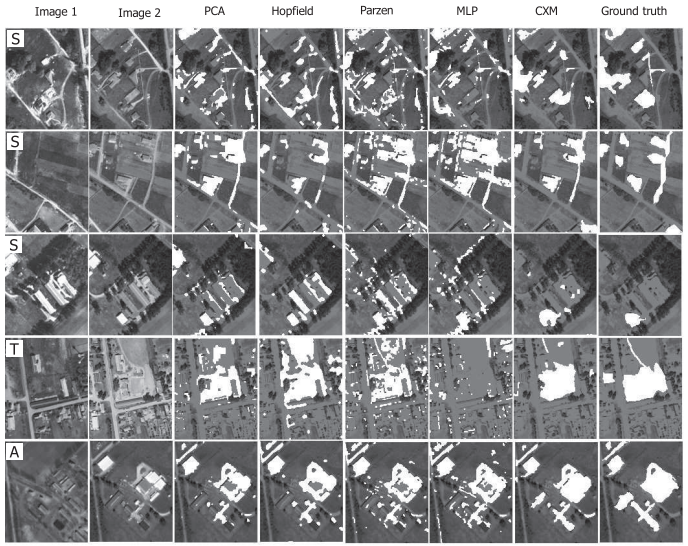
Image 2

Ground Truth change mask

Change mask overlaid by thresholded CDI

CXM result

CXM comparison to state-of-the-art methods



Content

Introduction

Region level change detection for remote sensing images

Building change detection

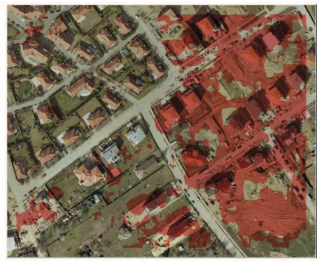
Marked Point Process model

Traffic monitoring from aerial Lidar point clouds

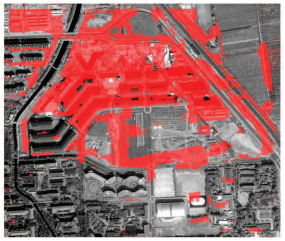
Hierarchical PCB Inspections

CXM with HR images & plenty of changes

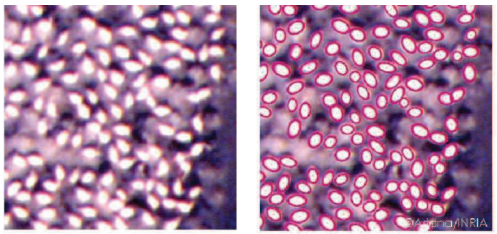
Object level change classification is necessary



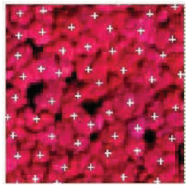
- Plenty of changes: object level change classification is necessary



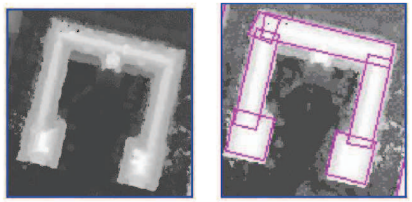
Corresponding demos of the Ariana Project



Flamingos (optical images)



Trees (infrared)



Buildings from Digital Elevation models

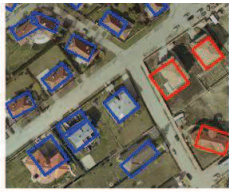
Penguins (new project)

Concept of the work

- Purely focusing on buildings and building changes
- Combining object extraction and change detection in a unified probabilistic model
  - General similarity information supports to separate changed and unchanged areas
  - Two-view information may increase the detection accuracy of *unchanged* buildings
  - Search for building changes can be focused on the *changed* areas

Problem definition

- Building = set of rectangular building segments.
- **The input**: pair of co-registered grayscale or color remote sensing (aerial or satellite) images, taken at different time instances
- **The output**: in each image we provide the size, position and orientation parameters of the detected building segments, with giving information which objects are new, demolished, modified/rebuilt and unchanged

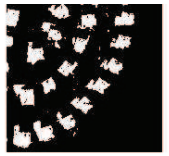


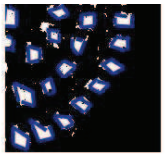
Detection results (changes with red)

Object extraction: bottom-up vs. inverse approaches

- Bottom-up approach: extracting *primitives* (roof blobs, edges, corners etc.) and constructing the buildings from them

Input


'Building filter'


CCA and poligon fitting

- Advantage:
  - can be relatively fast
- Drawbacks:
  - construction of appropriate primitive filter may be difficult/inaccurate (above a simple color filter is used)
  - challenging to incorporate a priori information (e.g. shape, size)
  - challenging to model object interactions

Object extraction: direct vs. inverse approaches

- Inverse approach:
  - assigning a fitness value to each possible configuration

LOW fitness value

HIGH fitness value

- Advantages:
  - complex object appearance models can be used
  - easy to incorporate a priori information (e.g. only rectangles)
  - easy to model object interactions (e.g. penalize intersection, favour similar orientation)
- Drawbacks:
  - high computational need (search in a high dimension population space)
  - local maxima or false maxima of the fitness function

Proposed solution

Marked Point Process model with Multiple Birth and Death Dynamics

- The algorithm
  1. Start with an empty population
  2. Iterate object birth and death steps till convergence is obtained in the global population
     - *Birth*: we propose multiple objects and add them to the population.
     - *Death*: we remove some probably incorrect objects
- Important properties:
  - *Birth*: exploration is driven by low level features (bottom-up)
  - *Death*: *inverse* approach in object description
  - Stochastic process both for object proposition and acceptance (unlike in conventional hypothesis generation-acceptance techniques)
  - Simulated annealing framework to ensure convergence

Proposed solution

Marked Point Process model with Multiple Birth and Death Dynamics

- The algorithm
  1. Start with an empty population
  2. Iterate object birth and death steps till convergence is obtained in the global population
     - *Birth*: we propose multiple objects and add them to the population.
     - *Death*: we remove some probably incorrect objects
- Important properties:
  - *Birth*: exploration is driven by low level features (bottom-up)
  - *Death*: *inverse* approach in object description
  - Stochastic process both for object proposition and acceptance (unlike in conventional hypothesis generation-acceptance techniques)
  - Simulated annealing framework to ensure convergence

Feature selection for Multiple Birth and Death

- Essential questions:
  - How can we generate relevant objects efficiently in the birth process?
    - $\rightarrow$ Low level features.
  - How can be ensured that the adequate objects survive the death step?
    - $\rightarrow$ Object level features.

Feature types

- Low level or region level features
  - characterize a given pixel and its neighborhood
    - e.g. color, texture, correlation for a rectangular image region
  - calculated once over the whole image lattice
  - only used to estimate where the objects are *likely* to be located, and how they *can* look like
    - local non-uniform probability distributions for birth and parameters
    - theoretically: do not effect the final solution just the speed of convergence
    - practically: necessary to obtain a good solution in reasonable time
- Object level features
  - characterizes a given object candidate,
    - fitness value for each possible oriented rectangle
  - should be calculated in each iteration for each proposed object
  - accuracy is crucial: acceptance of an object depends on these features

MTA SZTAKI

Feature types

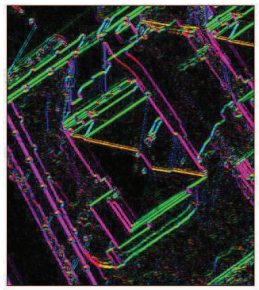
- Low level or region level features
  - characterize a given pixel and its neighborhood
    - e.g. color, texture, correlation for a rectangular image region
  - calculated once over the whole image lattice
  - only used to estimate where the objects are *likely* to be located, and how they *can* look like
    - local non-uniform probability distributions for birth and parameters
    - theoretically: do not effect the final solution just the speed of convergence
    - practically: necessary to obtain a good solution in reasonable time
- Object level features
  - characterizes a given object candidate,
    - fitness value for each possible oriented rectangle
  - should be calculated in each iteration for each proposed object
  - accuracy is crucial: acceptance of an object depends on these features

Csaba Benedek (DEVA Laboratory) Object detection in remote sensing applications 8 July 2013 33 / 102

Gradient Statistics Feature

- Regions of buildings contain sharp edges in **perpendicular** directions

Input image



Gradient image
-*intensity*: absolute value of the gradient
-*hue*: angle of the gradient

Weighted Gradient Orientation Density (WGOD)

- Notations:
  - ∇g_s: gradient vector at pixel s (Prewitt Operator), with absolute value $||\nabla g_s||$ and angle ϑ_s
  - $R(s,l)$: rectangular neighborhood of pixel s with side length l.
  - $b_1 \dots b_K$: K bins after discretization of the $[-\pi/2, \pi/2]$ domain
- Weighted gradient orientation histogram:

$$h_s[b_k] = \frac{1}{N_s} \sum_{r \in R(s,l)} ||\nabla g_r|| \cdot \mathbb{I}_{[\vartheta_r \in b_k]}$$

where

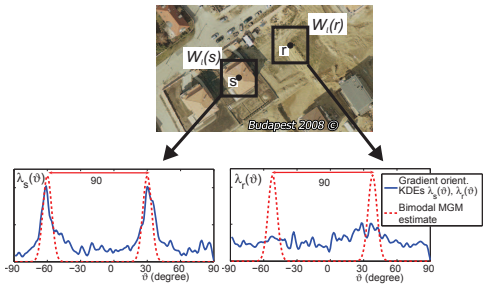
$$N_s = \sum_{r \in R(s,l)} ||\nabla g_r||$$

- Weighted Gradient Orientation Density (WGOD) around pixel s:

$$\lambda_s(\vartheta) = h_s[b_k]|_{\vartheta \in b_k}$$

Weighted Gradient Orientation Density (WGOD)

- $\lambda_s(\vartheta)$ local gradient orientation density around s pixel
- Perpendicular edges - peaks in 90° distance
- Feature:
  - matching a bi-modal Gaussian Mixture (bGM) to the WGOD
  - norm. cross correlation of WGOH and bGM



- Extracted features:
  - c_s^{val}: maximal correlation value at pixel s
  - m_s corr. peak location: angle shift for the maximal correlation

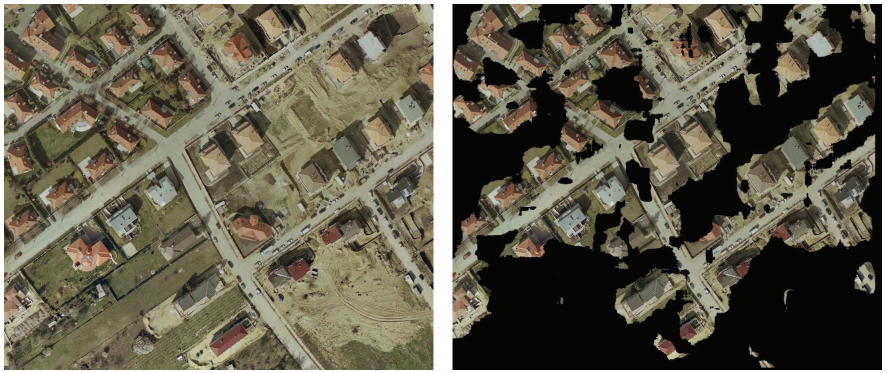
Result of gradient-orientation based building region detection

- Obtaining a "birth map" $P_b(s) = f_{\text{birth}}(c_s^{\text{val}})$
  - c_s^{val}: maximal correlation value at pixel s
  - $\forall s \in S$: $P_b(s)$ is the probability that s is the center of a building segment
  - $f_{\text{birth}}(c_s^{\text{val}})$ is a monotonously increasing function of c_s^{val} (high correlation - higher probability)
- Obtaining a "estimated mean orientation map": m_s
  - if there is a building segment centered at s its orientation θ can be considered as $m_s + \eta_s$
    - η_s: zero-mean Gaussian random variable, with deviation σ_θ
    - orientation of a building candidate in s should be generated according to $f_\theta(.,s) = \eta(., m_s, \sigma_\theta)$ distribution

Result of gradient-orientation based building region detection

- Obtaining a “birth map” $P_b(s) = f_{\text{birth}}(c_s^{\text{val}})$
  - c_s^{val}: maximal correlation value at pixel s
  - $\forall s \in S$: $P_b(s)$ is the probability that s is the center of a building segment
  - $f_{\text{birth}}(c_s^{\text{val}})$ is a monotonously increasing function of c_s^{val} (high correlation - higher probability)
- Obtaining a “estimated mean orientation map”: m_s
  - if there is a building segment centered at s its orientation θ can be considered as $m_s + \eta_s$
    - η_s: zero-mean Gaussian random variable, with deviation σ_θ
    - orientation of a building candidate in s should be generated according to $f_\theta(., s) = \eta(., m_s, \sigma_\theta)$ distribution

‘Birth map’ for building detection

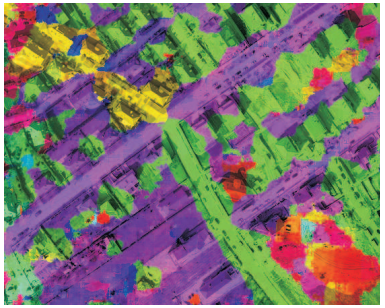


Thresholded correlation value map: regions of possible building segments

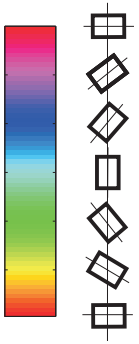
Local dominant gradient direction map

Input image:

Dominant gradient direction map m_s:

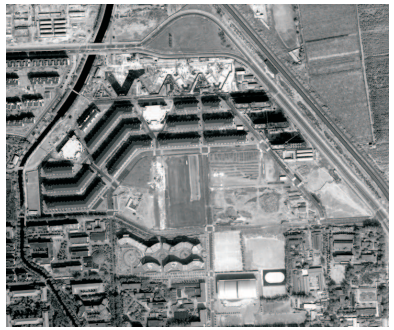


Color code of the estimated orientation:

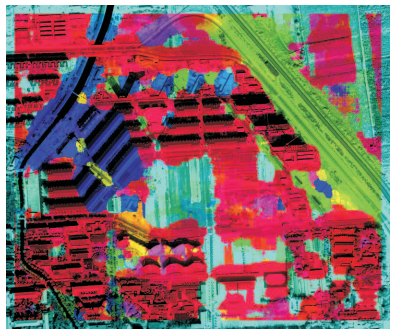


Local dominant gradient direction map

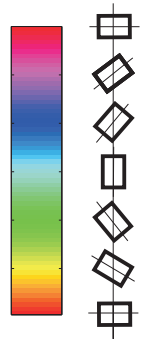
Input image:



Dominant gradient direction map m_s:



Color code of the estimated orientation:

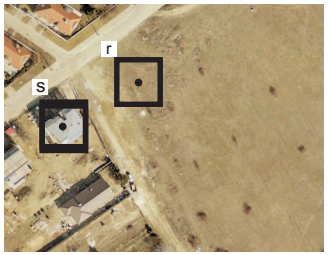


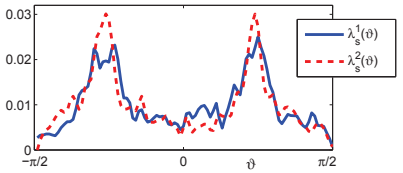
Low level region comparison based on gradient statistics

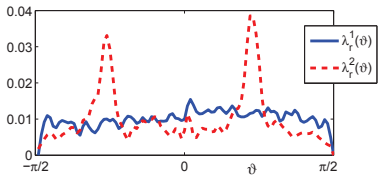
- $\lambda_s^i(.)$ $i \in \{1,2\}$: low level texture descriptor
  - gradient magnitude-orientation statistics in the neigborhood of pixel s in the ith image
  - appropriate for estimating the "building-membership"
  - independent of coloring effects (image sensor) and illumination
  - robust regarding parallax and registration errors
- Comparing $\lambda_s^1(.)$ and $\lambda_s^2(.)$: low level region similarity checking based on "building-focused" features
  - dissimilarity measure: Batthacharya distance

$$b(s) = -\log \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{\lambda_s^1(\vartheta) \cdot \lambda_s^2(\vartheta)} d\vartheta$$

Region comparison based on gradient statistics

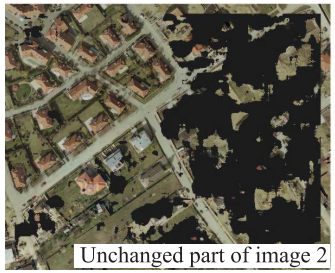


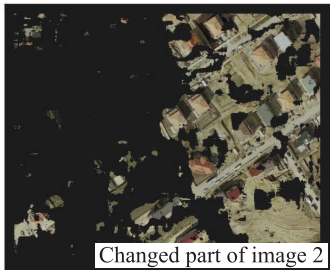




Comparing the $\lambda(.)$ functions in the two image layers regarding two selected pixels. s corresponds to an unchanged point and r to a built-up change.

Result of low level comparison





Shadow feature

- Buildings are next to their shadows (if available)
- Shadow color model (can be noisy) + information about the shadow direction

Shadow filter (color)

Near-to-shadow areas (considering the shadow direction)

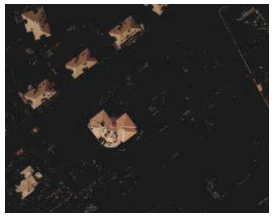
Roof color filtering

- Red roof-color can be filtered in illumination invariant color representations
  - *H* channel in HSV, *a* channel in CIE L*a*b* color spaces
- Usually not all the roofs are detected
- Left: *a* chrominance channel. Right: highlighted roof points

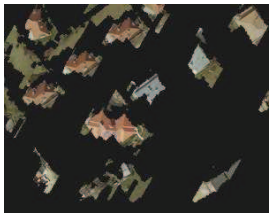
Low level built-in region estimation

(a) Input image

(b) Gradient statistics



(c) Roof color



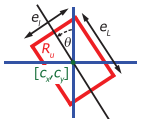
(d) Shadow based

Conclusions of the low level experiments

- Gradient feature
  - general (does not need color or shadows)
  - roughly locates the built-up regions
  - proposes orientation for the building candidates
  - results several false alarm
- Shadow feature
  - accurate localization, if the following conditions hold
    - presence of shadows
    - availability of a reasonable shadow filter
- Color feature
  - accurate footprints
  - needs color images
  - results several missed alarms (‘non-red’ roofs)

Object level features

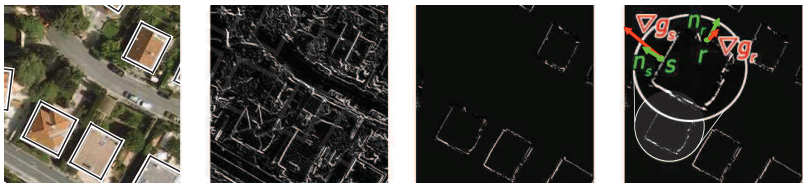
- u: rectangular building segment with parameters:
  - c_x, c_y: center coordinates
  - e_L, e_l: side lengths
  - θ: orientation



- $\varphi(u)$ energy term (negative fitness): evaluates the hypothesis that u is a building in the image
  - if $\varphi(u) < 0$, u is *attractive*
    - adding u to the population may increase the ‘global fitness’ of the population
  - sample features - subterms: gradient ($\varphi_{\mathrm{gr}}(u)$), roof color ($\varphi_{\mathrm{cr}}(u)$) and shadow ($\varphi_{\mathrm{sh}}(u)$)

Feature examples: gradient below building edges

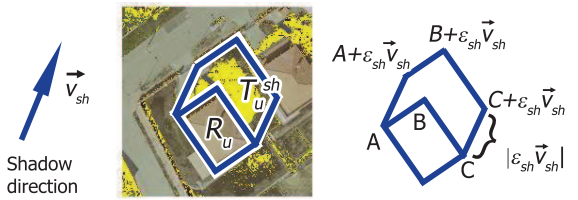
- Below the edges of a good rectangle candidate u we must find pixels which have
  - high (absolute) gradients
  - the orientation of the gradient is close to the normal vector of the closest rectangle side (θ or $\theta + \frac{\pi}{2}$).



$$\varphi_{\mathrm{gr}}(u) = \mathcal{Q}\left(\int_{s\in u} \nabla g_s \cdot n_s ds\right)$$

- $\mathcal{Q}$: monotonously decreasing, nonlinear function

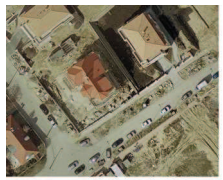
Shadow feature



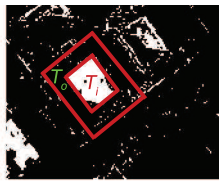
- Assumption: given the shadow direction vector and a (noisy) cast shadow map:
  - $\mu_{\mathrm{sh}}(s)$: shadow membership of pixel s
  - R_z: inside area, T_u^{sh}: outside area in the shadow direction

$$\chi_i = \int_{R_u} 1 - \mu_{\mathrm{sh}}(s) ds, \qquad \chi_o = \int_{T_u^{\mathrm{sh}}} \mu_{\mathrm{sh}}(s) ds$$

Color roof feature



Red roof

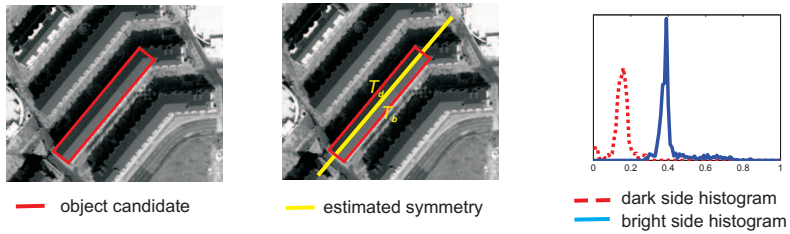


Color mask

- Assumption: given the (noisy and incomplete) roof-color map:
  - $\mu_{\mathrm{c}}(s)$: roof-color membership of pixel s
  - T_i: inside area, T_o: outside area

$$\mathcal{C}_i = \int_{T_i} \mu_{\mathrm{c}}(s) ds, \qquad \mathcal{C}_o = \int_{T_o} 1 - \mu_{\mathrm{c}}(s) ds$$

Two side roof feature

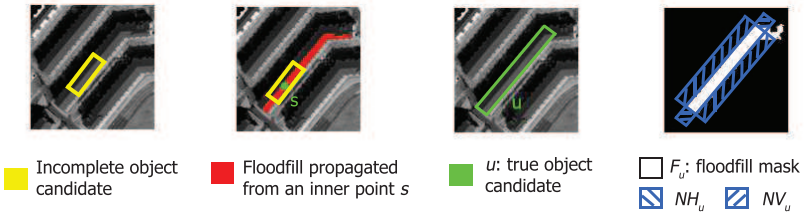


- Homogenous roofs: kurtosis describes the peakness of the histograms
  - g_s: gray value of pixel s

$$\kappa_d = \frac{\sum_{T_d} g_s^4}{\left(\sum_{T_d} g_s^2\right)^2} \qquad \kappa_b = \frac{\sum_{T_b} g_s^4}{\left(\sum_{T_b} g_s^2\right)^2}$$

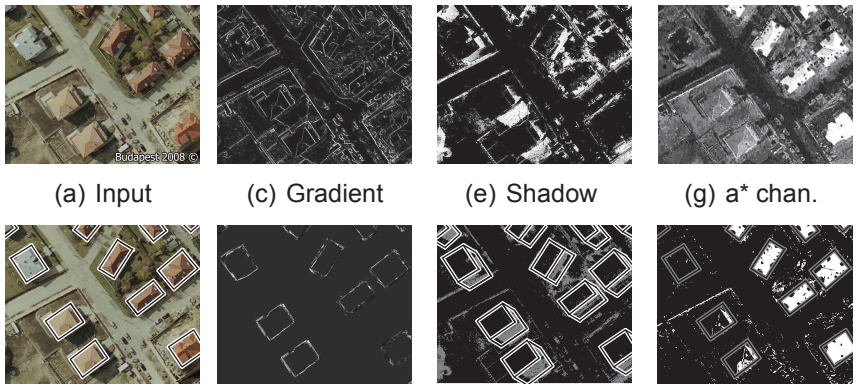
$$\varphi_\kappa(u) = \max\left[\mathcal{Q}(\kappa_d(u), d_d^\kappa, D_d^\kappa), \mathcal{Q}(\kappa_b(u), d_b^\kappa, D_b^\kappa)\right]$$

Two side roof feature



- Kurtosis+shadow may result in incomplete roofs
  - roof parts are also bi-homogenous
- Extracting the maximal homogenous blobs
  - f_s: set of pixels resulted by floodfill propagation from pixel s
  - $F_u = \cup_{s \in u} f_s$ floodfill mask of object candidate u
  - NH_u, NV_u: horizontal and vertical neighboring rectangles of u
    - Intersection of F_u with NH_u and NV_u must be *small*.

Integration of features



(a) Input (c) Gradient (e) Shadow (g) a* chan.

(b) Ground Truth (d) Grad. feat. (f) Shad. feat. (h) Color mask

Different features in an image part

Integration of features

- Prototype definition:
  - many feature constraints in AND relationship (MAX operator in the negative fitness domain)
- Using different prototypes parallelly:
  - the prototype energy terms in OR relationship (MIN operator in the negative fitness domain)
- Example: (gradient AND shadow) OR roof color:

$$\varphi(u) = \min\left\{ \max\{\varphi_{\mathrm{gr}}(u), \varphi_{\mathrm{sh}}(u)\}, \varphi_{\mathrm{rc}}(u)\right\}.$$

Content

Introduction

Region level change detection for remote sensing images

Building change detection

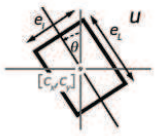
Marked Point Process model

Traffic monitoring from aerial Lidar point clouds

Hierarchical PCB Inspections

Marked Point Process Model (MPP)

- Robust Bayesian tool for extraction of object configurations



- Point process: arbitrary number of point wise entities in the 2-D (or 3-D) space
  - e.g.: vehicle centers

$$\overline{p} = \{p_1, p_2, \ldots, p_n\},\ p_i \in K \subset \mathbb{R}^2$$

- Markers: geometric description of the objects e.g. rectangle: width e_L, height e_l, orientation θ

$$M = \left]-\frac{\pi}{2}, \frac{\pi}{2}\right] \times [L_{\min}, L_{\max}] \times [l_{\min}, l_{\max}]$$

- Marked object = point + markers
  - for rectangles: five parameters

$$u_i \in K \times M$$

Notations

- $\mathcal{H}$: space of u objects (rectangles)
- Object configuration $\omega = \{u_1, \ldots, u_n\}$
- Ω configuration space (H – Borel set in $\mathcal{H}$):

$$\Omega = \bigcup_{n=0}^{\infty} \Omega_n, \quad \Omega_n = \{\{u_1, \ldots, u_n\} \subset H^n\}$$

- $\sim$: neighbourhood relation in $\mathcal{H}$
- neighbourhood of u in ω:

$$\mathcal{V}_u(\omega) = \{v \in \omega | u \sim v\}$$

- MAP estimation of the optimal object configuration $\hat{\omega}$:

$$\hat{\omega} = \arg\max_{\omega \in \Omega} p(\omega|\mathcal{D})$$

Marked Point Process approach

$p(\omega|\mathcal{D})$ follows Gibbs distribution:

$$p(\omega|\mathcal{D}) = P_{\mathcal{D}}(\omega) = \frac{1}{Z} \cdot \exp\left[\sum_{u \in \omega} A(u, f_u(\mathcal{D})) + \gamma \cdot \sum_{\substack{u,v \in \omega \\ u \sim v}} I(u, v)\right]$$

where $f_u(\mathcal{D})$ is a feature vector for object u, while $A(u, f_u(\mathcal{D}))$ and $I(u, v)$ are singleton/interaction potentials, and:

Unary potentials

- Characterizes the fitness value of a given a building segment $\{c_x, c_y, e_L, e_l, \theta, \xi\}$ as a function of the local image data in both images $\mathcal{D}_1$ and $\mathcal{D}_2$, but independently of other object of the population.

$$A(u, f_u(\mathcal{D})) = \mathbb{I}_{[\xi(u) \neq 2]} \cdot \varphi(u, \mathcal{D}_1) + \mathbb{I}_{[\xi(u) \neq 1]} \cdot \varphi(u, \mathcal{D}_2) + \\ + (-1)^{\mathbb{I}_{[\xi(u) \neq *]}} \int_u b(s) ds$$

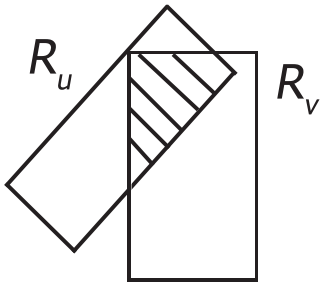
where

- $\xi(u) \in$ {1,2,*}: image index flag (*: unchanged, 1: only in first the image, 2: only in the second image)
- $\varphi(u, \mathcal{D}_1)$ and $\varphi(u, \mathcal{D}_2)$ are the building fitness values in images $\mathcal{D}_1$ and $\mathcal{D}_2$
- $b(s)$ is the low level dissimilarity function

Interaction potentials

- Penalize overlap of objects sharing a time layer:

$$I(u,v) = \mathbb{I}_{[\xi(u)\simeq\xi(v)]} \cdot \frac{\mathrm{Ter}(R_u \cap R_v)}{\mathrm{Ter}(R_u \cup R_v)}$$



  - where $\xi(u) \simeq \xi(v)$ **iff**

$$\xi(u) = \xi(v), \text{ or } \xi(u) = *, \text{ or } \xi(v) = *$$

  - allow overlap of modified/re-built buildings

$$\xi(u) \neq \xi(v); \quad \xi(u), \xi(v) \in \{1, 2\}$$

Results – Budapest images

Source: András Görög

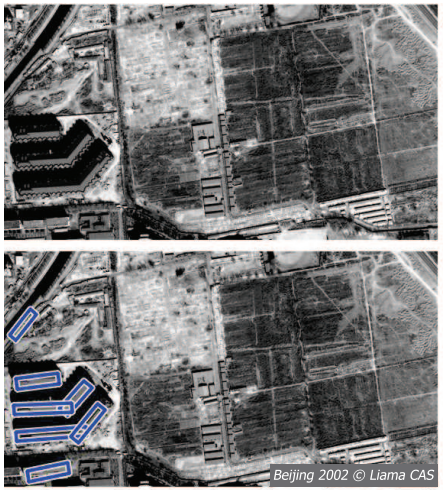
detected unchanged buildings

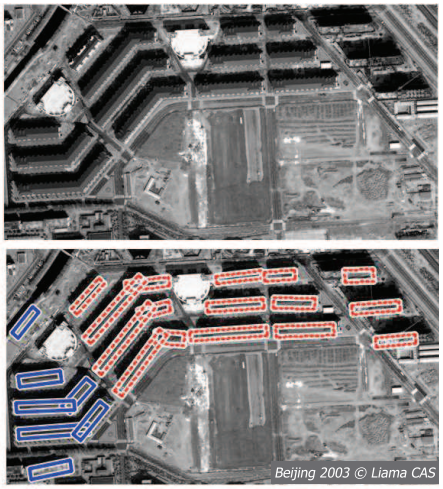
detected new/demolished buildings

Iterative optimization - Budapest

Results – Beijing Quickbird Satellite Images

Source: LIAMA Laboratory, Beijing, China



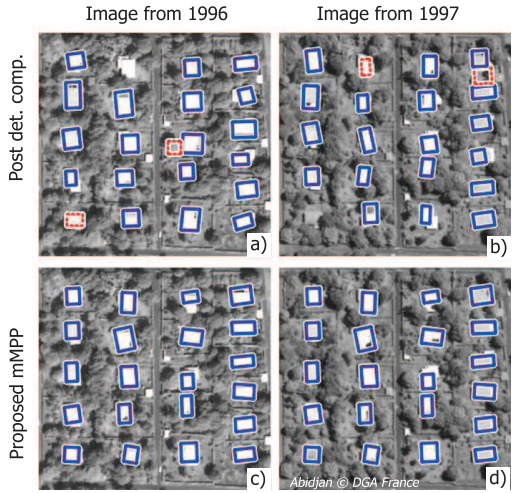


detected unchanged buildings

detected new/demolished buildings

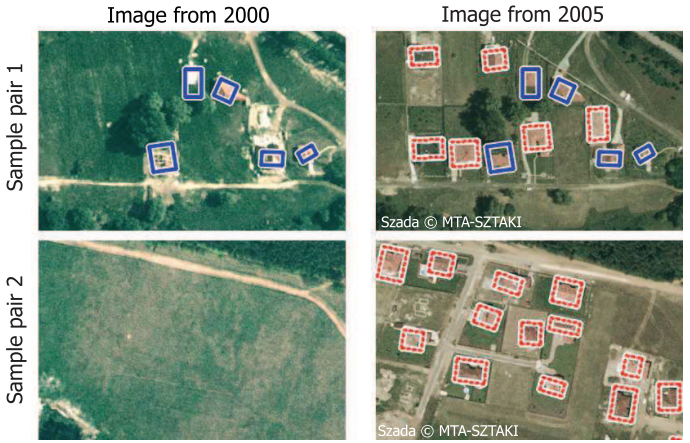
Results – Abidjan Satellite Images

Source: French Defense Agency (DGA)



Results – SZADA aerial image pair

Source: Hungarian Institute of Geodesy Carthography and Remote Sensing (FÖMI) Budapest



Results - comparison to reference techniques

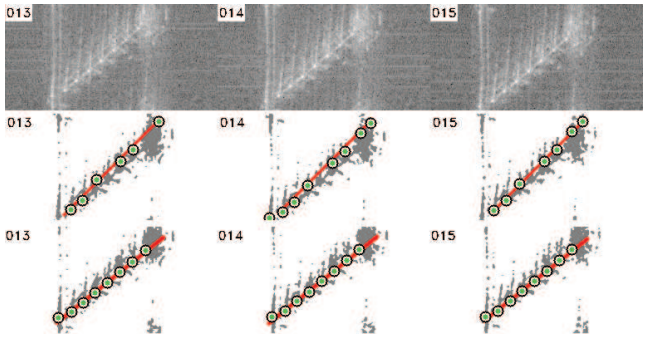
(a) Edge based (b) Segm-Merge (c) MPP (d) GT

(e) Edge based (f) Segm-Merge (g) MPP (h) GT

Tracking ship targets in ISAR image sequences

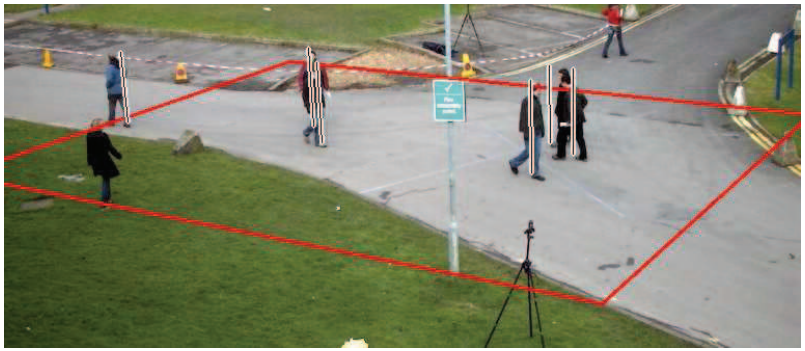
ISAR: Inverse Synthetic Aperture Radar

- Pattern recognition in noisy radar image sequences
- Ship axis + permanent scatterers (e.g. containers)



Cs. Benedek and M. Martorella: "Moving Target Analysis in ISAR Image Sequences with a Multiframe Marked Point Process Model", **IEEE Trans. on Geoscience and Remote Sensing**, in press 2013

Pedestrian detection and height estimation from multiple camera views - 3DMPP



Á. Utasi and Cs. Benedek: “A Bayesian Approach on People Localization in Multi-Camera Systems”, **IEEE Trans. on Circuits and Systems for Video Technology**, vol 23, no. 1, pp. 105-115, 2013

Content

Introduction

Region level change detection for remote sensing images

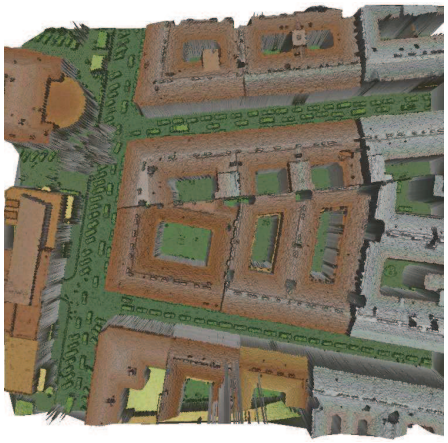
Building change detection

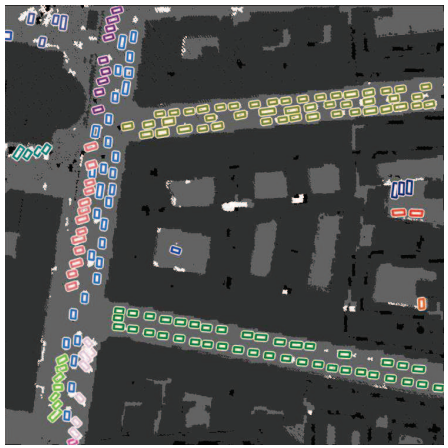
Marked Point Process model

Traffic monitoring from aerial Lidar point clouds

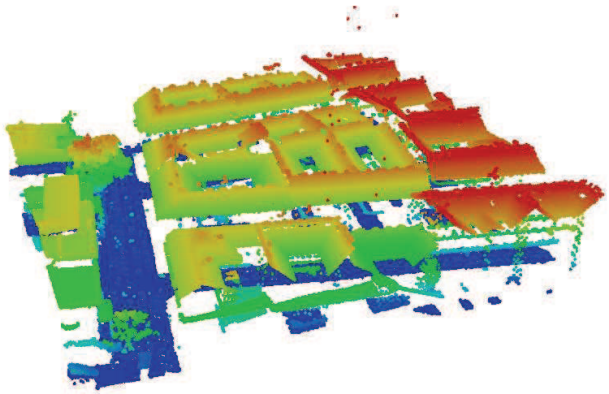
Hierarchical PCB Inspections

Traffic monitoring from aerial Lidar point clouds





Traffic monitoring from aerial Lidar point clouds



Point cloud segmentation

Classes: vehicle, ground, roof, vegetation, clutter

- Features used by the segmentation algorithm
  1. Point density - removal of sparse clutter regions
     - noise and reflections from walls
  2. Echo number - filtering vegetation
     - trees and bushes cause multiple laser returns
  3. RANSAC-based plane estimation - terrain identification
     - piecewise planar ground is expected
  4. Large connected *flat* regions - extraction of building roofs
     - floodfill based on point proximity and surface normal features
  5. Vehicle candidates: outlier points which fit prior height distributions of the possible vehicles.
- MRF based spatial filtering
  - class probability models & graph cut based optimization

MTA SZTAKI

Point cloud segmentation

Classes: vehicle, ground, roof, vegetation, clutter

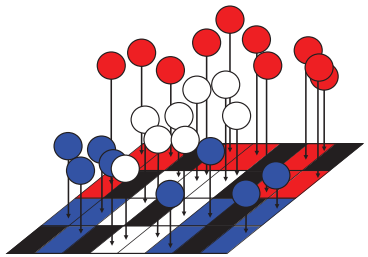
- Features used by the segmentation algorithm
  1. Point density - removal of sparse clutter regions
     - noise and reflections from walls
  2. Echo number - filtering vegetation
     - trees and bushes cause multiple laser returns
  3. RANSAC-based plane estimation - terrain identification
     - piecewise planar ground is expected
  4. Large connected *flat* regions - extraction of building roofs
     - floodfill based on point proximity and surface normal features
  5. Vehicle candidates: outlier points which fit prior height distributions of the possible vehicles.
- MRF based spatial filtering
  - class probability models & graph cut based optimization

MTA SZTAKI

Result of Point cloud segmentation

MTA SZTAKI

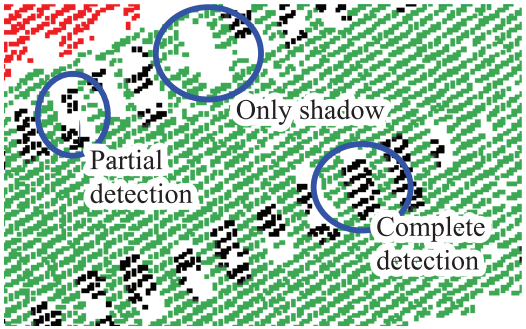
Point cloud projection to the ground plane



- Projecting the labels of *ground*, *roof* and *vehicle* points to a ground lattice
- Multiple points at a pixel: precedence for *vehicle* point candidates
- Undefined labels – no label interpolation

- Notations:
  - S: pixel lattice on the ground, $s \in S$: a single pixel
  - $\nu(s)$: label of pixel s, $\nu(s) \in \{\text{vehicle}, \text{background}, \text{undefined}\}$
    - background = ground + roof
  - $g(s)$: intensity value of the point assigned to pixel s
    - $g(s) = 0$ if $\nu(s) = \text{undefined}$

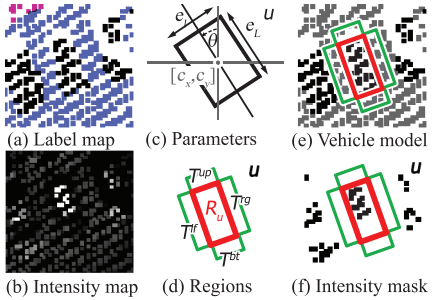
Challenges of vehicle detection in the label map



- Missing or partially detected vehicles in the label map
- 'Shadow-vehicles' due to occlusion: only holes in the ground
- Regular alignment of vehicles: similar size and orientation

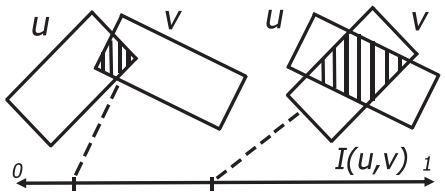
- Proposed solution: population level traffic description
  - Prior information about vehicle geometry and interactions

Data terms: Features for vehicle detection

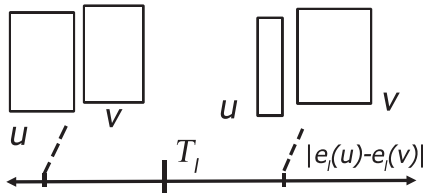


- *vehicle evidence* feature
  - we expect several pixels classified as vehicle within the rectangle of u
- *intensity* feature
  - additional evidence for image parts containing high intensity regions
- *internal background* feature
  - within the rectangle only very few background pixels may occur
- *external background* feature
  - a vehicle surrounded by background regions ($\geq$3 sides)

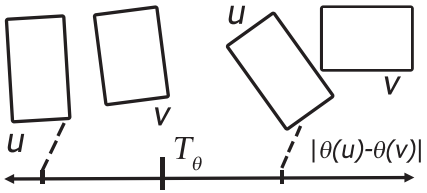
Prior terms: overview



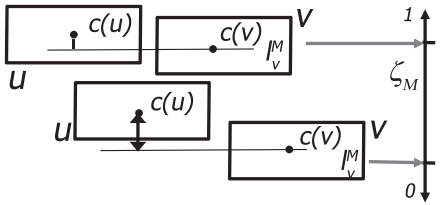
Non-overlapping φ_p^{ov}



Width similarity φ_p^{l}

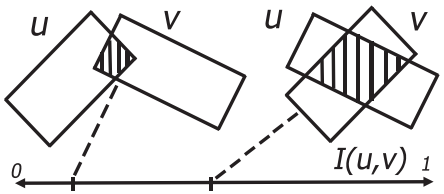


Weak alignment φ_p^{θ}

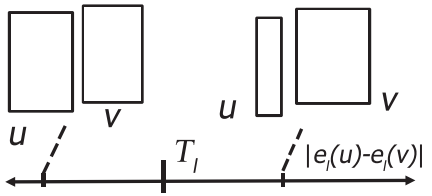


Strong alignment φ_p^{at}

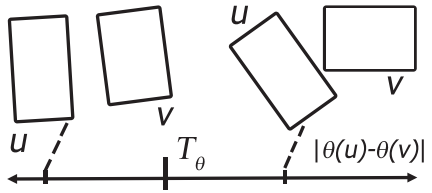
Prior terms: overview



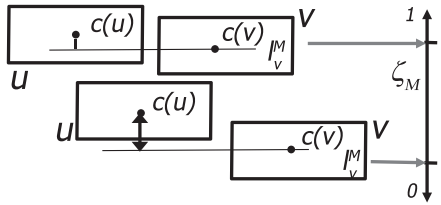
Non-overlapping φ_p^{ov}



Width similarity φ_p^{l}

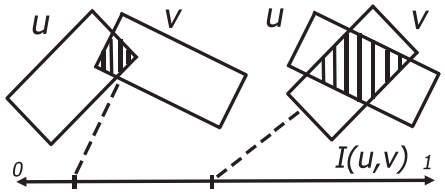


Weak alignment φ_p^{θ}



Strong alignment φ_p^{at}

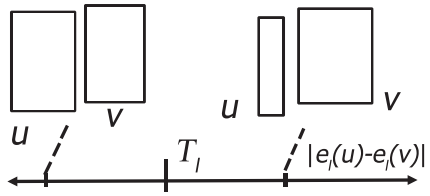
Prior terms: Non-overlapping



- Goal: avoid multiple vehicles in overlapping positions
- Overlapping coefficient $I(u,v)$ between two object rectangles:

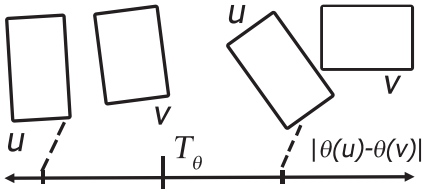
$$I(u,v) = 2 \cdot \frac{|R_u \cap R_v|}{|R_u| + |R_v|}$$

Prior terms: width similarity



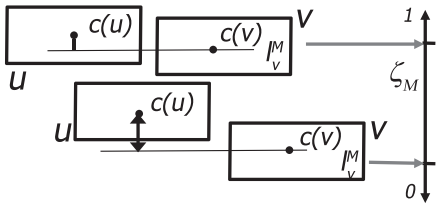
- Goal: avoid to merge contacting vehicles into the same object candidate
- We penalize rectangles with significantly different width (e_l) parameters in local neighborhoods

Prior terms: weak alignment



- Goal: favor objects with similar orientation in a local neighborhood.
- Penalizing object pairs based on the absolute difference of the orientations
- Only small weight factor to allow irregularly oriented cars

Prior terms: strong alignment



- Goal: strongly favor vehicles forming regular lines or rows
  - e.g. parking or standing in front of a traffic lamp
- Feature: reward if $c(u)$ is close either to the major or to the minor axis of v
  - Not a necessary criterion, however if satisfied, one can be less strict regarding other energy terms

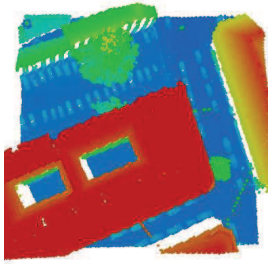
Integration of the energy components

- Global energy calculation

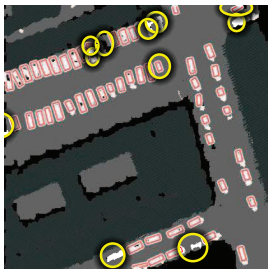
$$\Phi_{\mathcal{D}}(\omega) = \sum_{u \in \omega} \Psi_{\mathcal{D}}(u, \mathcal{N}_u)$$

- Neighborhood-energies
  - Generally: vehicles should pass each feature constraint
  - Detecting 'shadow vehicles': if the strong alignment criterion is satisfied, only internal and external background features are considered
  - Further prior terms: additive terms in the global energy function

Qualitative experiments

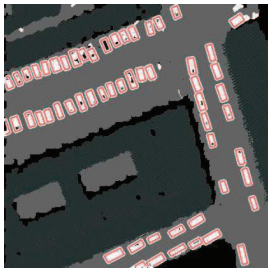


- Data: dense urban area from Budapest, Hungary
- Ground truth: manual generation with and interactive tool
- DEM-PCA primary errors by cars partially hidden by trees



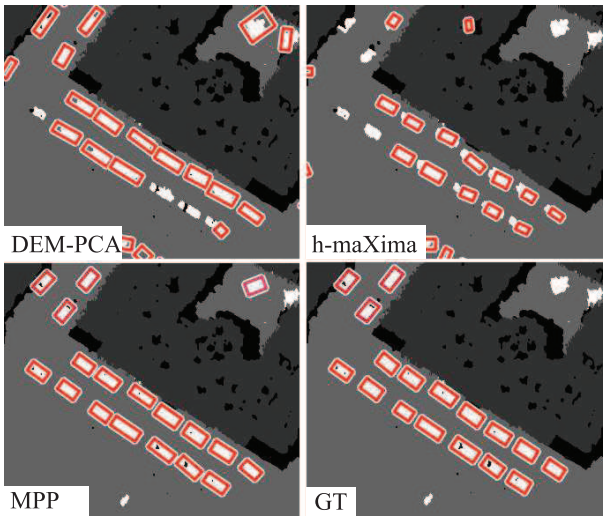
DEM-PCA result

MPP results



Ground Truth

Qualitative experiments



Content

Introduction

Region level change detection for remote sensing images

Building change detection

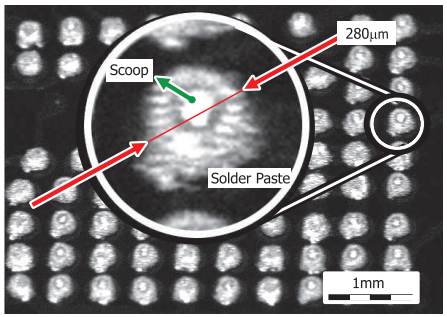
Marked Point Process model

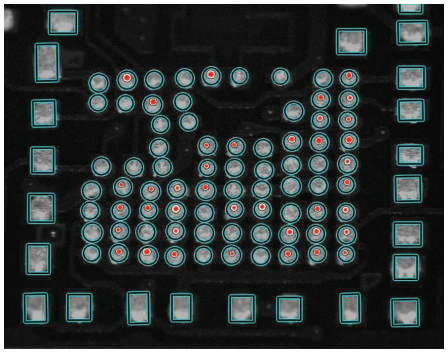
Traffic monitoring from aerial Lidar point clouds

Hierarchical PCB Inspections

Printed Circuit Board inspection

Solder paste extraction and scooping detection





Cs. Benedek, O. Krammer, M. Janóczki and L. Jakab: "Solder Paste Scooping Detection by Multi-Level Visual Inspection of Printed Circuit Boards", **IEEE Trans. on Industrial Electronics**, vol. 60, no. 6, pp. 2318-2331, 2013

Introduction

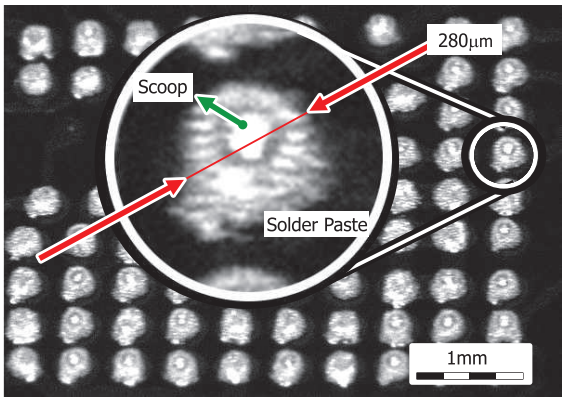
Automatic optical inspection (AOI)

- Main steps of the process
  - Manufacturing - artifacts in PCBs and consequences
  - Imaging system limitations - resolution, defocus blur
  - Machine vision, recognition
- Evolution of the industrial cameras
  - Consequences of the increased resolution and image quality:
    - Circuit element (e.g. solder paste) extraction may become trivial
    - Fine effects are observable and may be tractable with state-of-the-art machine vision algorithms
- Reflow soldering
  - used for mechanical fastening and electrical joining of surface mounted components
  - several soldering defects can be attributed to solder paste stencil printing

Introduction

Scooping in optical PCB images

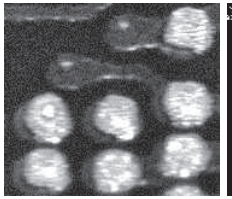
- Scooping - one of the major printing defects
  - deposited solder paste has a concave profile and its volume is less then it is intended to be based on the stencil aperture volume



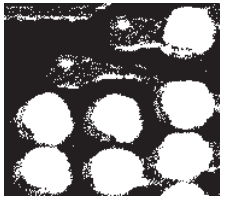
10μm resolution image of a PCB, a magnified solder paste, with a scoop.

Scooping detection - challenges

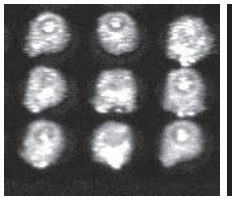
Limitations of simple morphological approaches



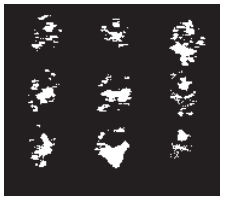
(i) Noisy input



(j) Paste mask

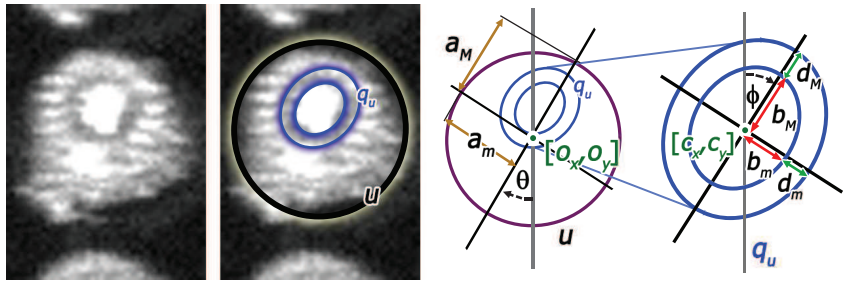


(k) Pastes scoops



(l) Scoop mask

Inverse model



- Model elements:
  - u Solder Paste (SP): ellipse
  - q_u Scoop Object (SO) $\in$ { two concentric ellipses, NULL }
  - $\mathcal{H}$: object space (SP+SO)
- Output of the detector: ω
  - configuration of a finite number of u 'parent' and q_u 'child' objects:

$$\omega = \{u_1, \ldots, u_n\}, \quad n \in \mathbb{N}^+ \quad \forall i\ u_i \in \mathcal{H}$$

Configuration energy and optimization problem

- Configuration space Ω:

$$\Omega = \bigcup_{n=0}^{\infty} \Omega_n, \quad \Omega_n = \{\{u_1, \ldots, u_n\} \in \mathcal{H}^n\}$$

- Gibbs distribution on the Ω space:

$$P(\omega) = \frac{1}{Z} \cdot \exp\left[-\Phi(\omega)\right]$$

where $\Phi(\omega)$ is the configuration energy.

- Maximum Likelihood configuration estimate:

$$\omega_{\mathrm{ML}} = \operatorname*{argmax}_{\omega \in \Omega} P(\omega) = \operatorname*{argmin}_{\omega \in \Omega} \Phi(\omega)$$

- Data term - prior term decomposition

$$\Phi(\omega) = \Phi_p(\omega) + \Phi_d(\omega)$$

Prior energy

Parent-level geometric interactions

- $\varphi_p^i(u, v)$: penalizes overlapping between different Solder Paste ellipses u and v

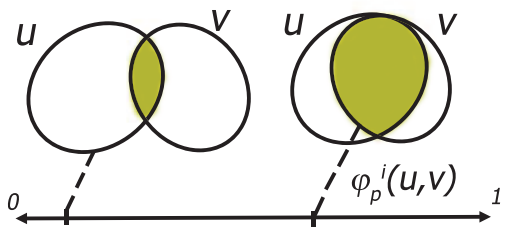


Figure: Intersection area at parent level

Prior energy

Parent-child geometric interactions

- $\varphi_p^c(u, q_u)$: penalizes overhang between a parent Solder Paste u and its child scoop q_u

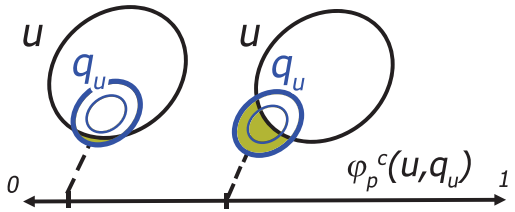


Figure: Scoops overhanging area w.r.t the parent SP object

- Prior term of the configuration energy:

$$\Phi_p(\omega) = \sum_{\substack{u,v \in \omega \\ u \sim v}} \varphi_p^i(u, v) + \sum_{u \in \omega} \varphi_p^c(u, q_u)$$

Data energy

Parent level dataterm

- High contrast between the solder paste ellipse and its neighborhood:

$$d_B(u) = 1 - \int \sqrt{\lambda_u^{\text{in}}(x) \cdot \lambda_u^{\text{out}}(x)}dx$$

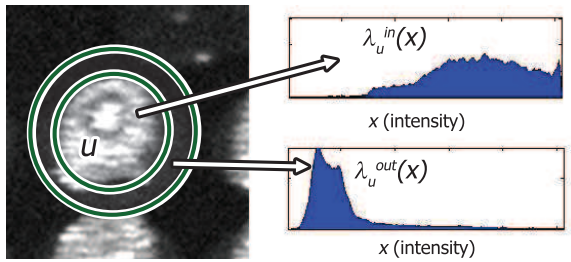


Figure: Utility of the $\lambda_u^{\text{in}}(x)$ and $\lambda_u^{\text{in}}(x)$ histograms for the parent data term calculation

Data energy

Child term

- Feature: contrast between the internal region, middle ring and neighborhood resp.

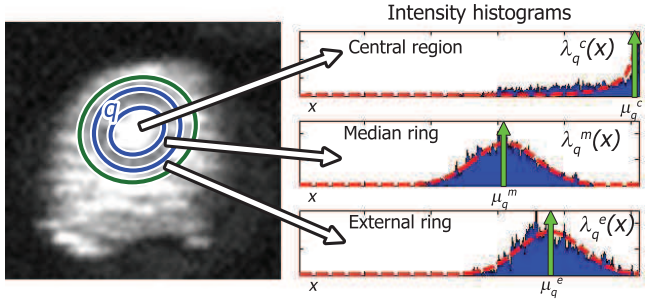


Figure: Child data term calculation

Reference Method

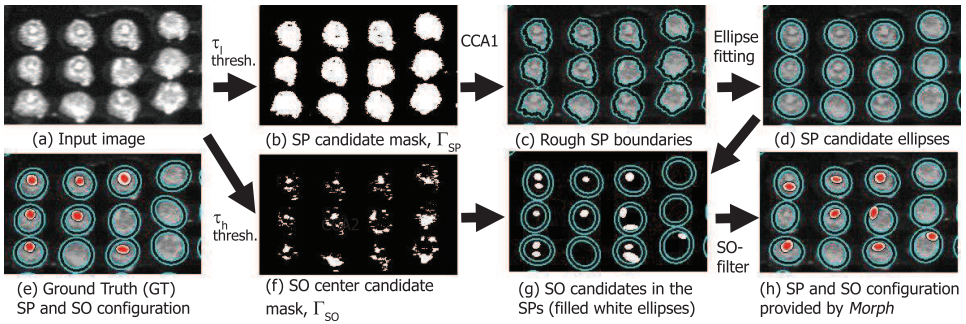
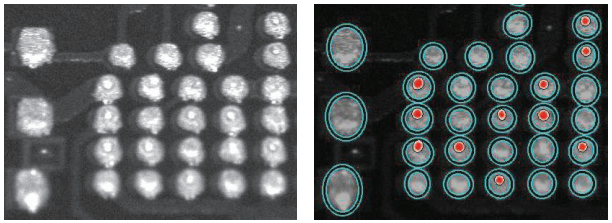
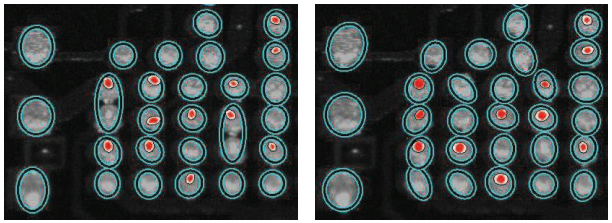


Figure: Demonstration of the consecutive steps of the *Morph* algorithm

Qualitative comparison

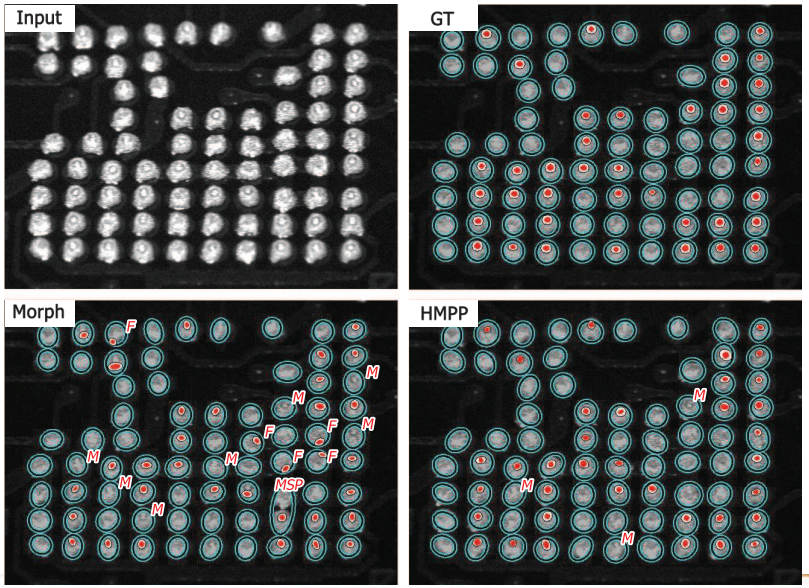


(a) Input image (b) Ground Truth



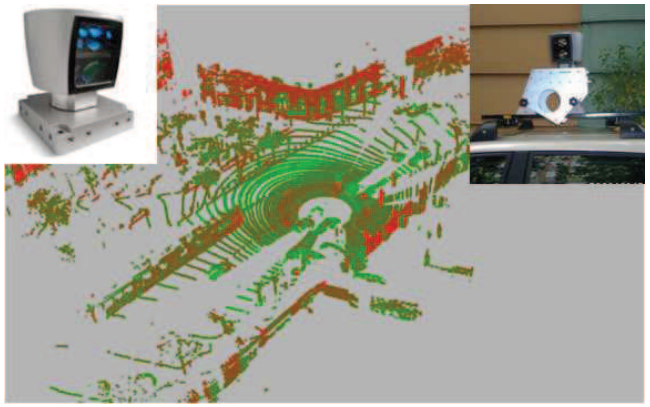
(c) Morph result (d) HMPP result

Results



Scene understanding from terrestrial LIDAR data

Velodyne HDL-64E camrea (see Google's self driving car)



i4D - integrated 4D project

Scene reconstruction in a 4D studio based on the Lidar output

Acknowledgment

- R&D projects supporting the work:
  - Hungarian Research Fund (OTKA #101598)
  - János Bolyai Research Scholarship of the Hungarian Academy of Sciences
  - i4D Internal Project of MTA SZTAKI
- Test data provider
  - Infoterra - Astrium GEO–Inf. Services, Hungary
  - Budapest University of Technology and Economics
  - City Council of Újbuda
- Thank you for your attention