Marked point processes for object extraction in high resolution images: Application to Earth observation and cartography

J. Zerubia

Joint work with X. Descombes, C. Lacoste, M. Ortner, G. Perrin, R. Stoica, F. Lafarge

Ariana research group, http://www.inria.fr/ariana

Bayesian Approach

$$P(X \mid Y) = \frac{P(X)P(Y \mid X)}{P(Y)} \propto P(X)P(Y \mid X)$$

Y : observed data

X : unknown variable (objects, features, …)

$P(Y \mid X)$: likelihood

$P(X)$: prior

$P(X \mid Y)$: posterior

Estimated X : $X^* = \arg\max_X P(X \mid Y)$

Medium Resolution Data: Markov Random Fields

Prior: Markov Random Field on pixel values

Likelihood: conditional independence assumption

$$P(Y \mid X) = \prod_{s \in S} P(y_s \mid x_s)$$

No contextual information in the likelihood:

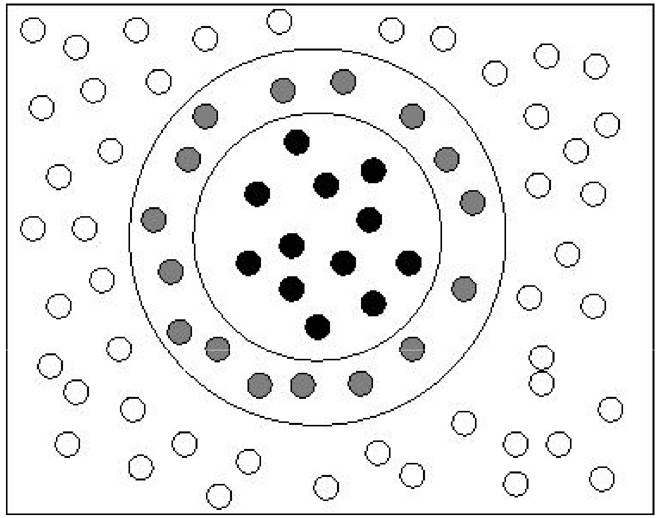
1 - uncorrelated noise

2 - no texture

Markov Random Fields

$$P(x_s \mid x_t, t \neq s) = P(x_s \mid x_t, t \in v_s)$$

v_s being the neighborhood of s



- Contextual Information Modeling
- Link with Statistical Physics: Gibbs Fields

From Context to Geometry

SPOT image © CNES

IKONOS image © Satellite imaging Corporation

IKONOS image © Satellite image Corporation

From Context to Geometry

How to extract structural information from HR images?

SPOT image © CNES

aerial image © IGN

High Resolution Data: From Pixels to Objects

- **Goal:** To model the observed scene as a configuration of objects (roads, rivers, buildings, trees):
  - To take into account data at a macroscopic scale.
  - To take into account the geometry of objects.
  - To take into account relations between objects (macro-texture).
  - Unknown number of objects (MRF on graph impossible).

Solution: Marked point processes

- **Stochastic modeling**: Set of objects in the scene $\equiv$ realization of a **marked point process**, **X**.
- **Optimization algorithm:** Monte Carlo sampler (e.g. RJMCMC) + simulated annealing.

Marked Point Processes

- A marked point process X on $\chi = \mathcal{P} \times \mathcal{M}$ is a point process on χ for which the point location is in $\mathcal{P}$ and the marks in $\mathcal{M}$.

- We define X by its probability density f w.r.t. the law $\pi_\nu(.)$ of a Poisson process known as the reference process ($\nu(.)$ is the intensity measure):

Sampling: Birth and Death Algorithm (Geyer/Moller-94)

- **Birth**: with probability ½, randomly propose a new point u in χ to be added to the current configuration x. Let y = x U {u}. Compute the acceptance ratio:

$$R_1(x,y) = \frac{f(y)}{f(x)}\frac{\nu(\chi)}{n(y)}$$

- **Death**: with probability ½, randomly propose a point v to be removed from x. Let y = x /{v}. Compute the acceptance ratio: $$R_2(x,y) = \frac{f(y)}{f(x)}\frac{n(x)}{\nu(\chi)}$$

- With probability α_i = min{1, R_i}, accept the proposition x_{t+1} = y, otherwise accept the proposition x_{t+1}=x.

Sampling: RJMCMC (Green-95)

- Generalization of Geyer/Moller-94
- Mixture of several proposition kernels:

$$Q(x,.) = \sum_m p_m(x) q_m(x,.) \quad \text{with} \quad \boldsymbol{Q}(\boldsymbol{x}, \boldsymbol{N}^f) \leq 1$$

- Convergence condition exists.

Sampling: RJMCMC

- **Algorithm**:

*At time **t**:*

1) *Select randomly a kernel* $\mathbf{q_m}$ *using the discrete law* $(\mathbf{p_m(x)})$

2) *Generate a new configuration* ***y*** *with respect to the selected kernel:*
   $\mathbf{y} \sim \mathbf{q_m}(\mathbf{x},.)$

3) *Compute the acceptance ratio:* $\mathbf{R_m}(\mathbf{x},\mathbf{y})$

4) *Compute the acceptance rate* α *:* $\alpha = \min(\mathbf{1},\mathbf{R_m}(\mathbf{x},\mathbf{y}))$

5) *With probability*
   - α *set:* $\mathbf{X_{t+1}} = \mathbf{y}$
   - *(1-*α*)* *set:* $\mathbf{X_{t+1}} = \mathbf{x}$

Optimization Algorithm

- **Goal:** To estimate a configuration maximizing **f(.)**

- **Simulated annealing:**

> *Successive simulations of* $\mathbf{f_t(x)\ n(dx)}$ *using a RJMCMC algorithm with:*
>
> $$\mathbf{f_t(x) = f(x)^{\frac{1}{T_t}}}$$
>
> *where* $\mathbf{(T_t)}$ *(≡ temperature) decreases towards zero.*

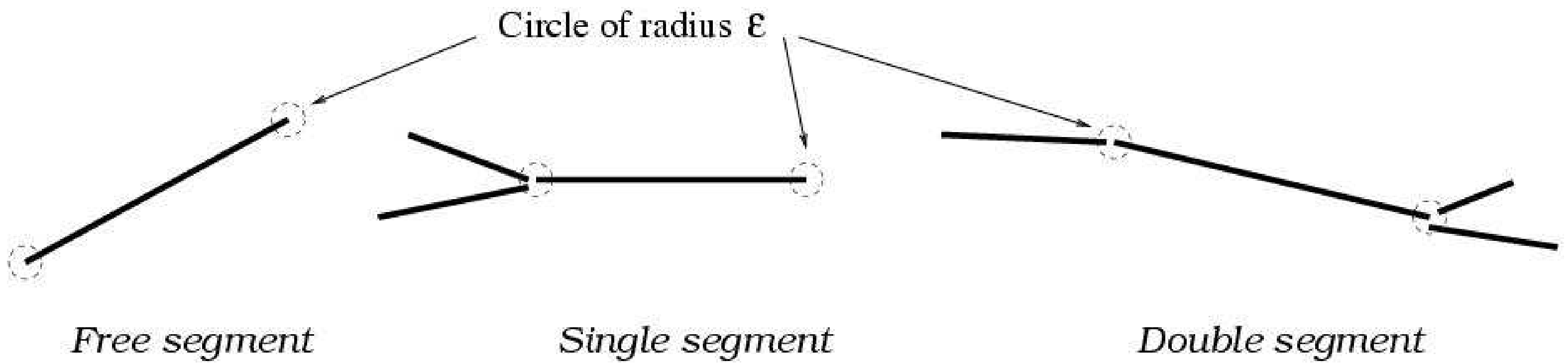
- Logarithmic decrease ⇒ global maximum.

- **In practice:** geometric decrease.

  At each step, $\mathbf{T_{t+1} = T_t \times c}$, where **c** is a constant close to 1.
  (c=0.99999 or c=0.999999 depending on the difficulty of the detection)

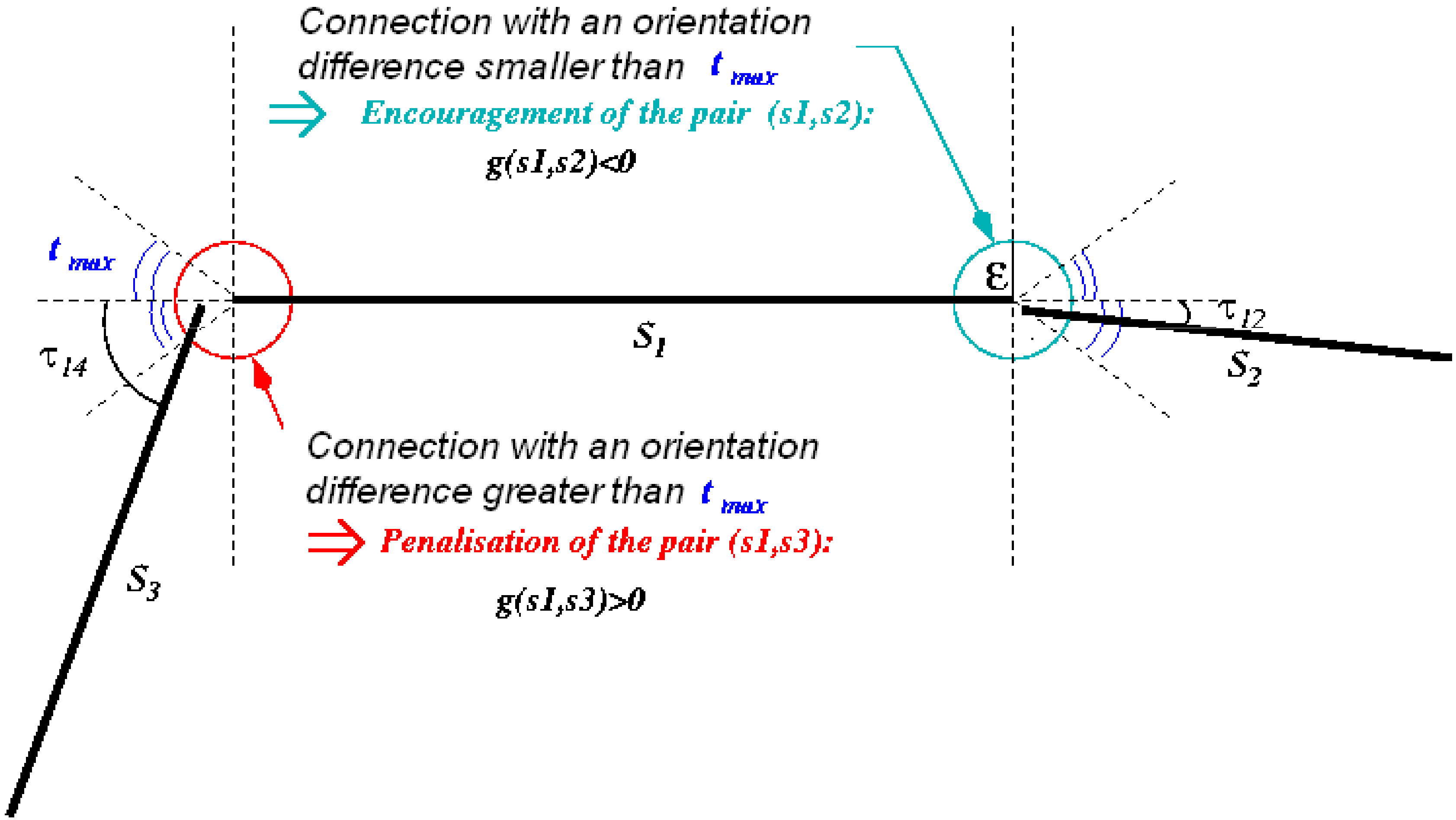
First example: Quality Candy Model for road network extraction

- Objects: segments
- Prior: models the connectivity and the curvature
- Data term



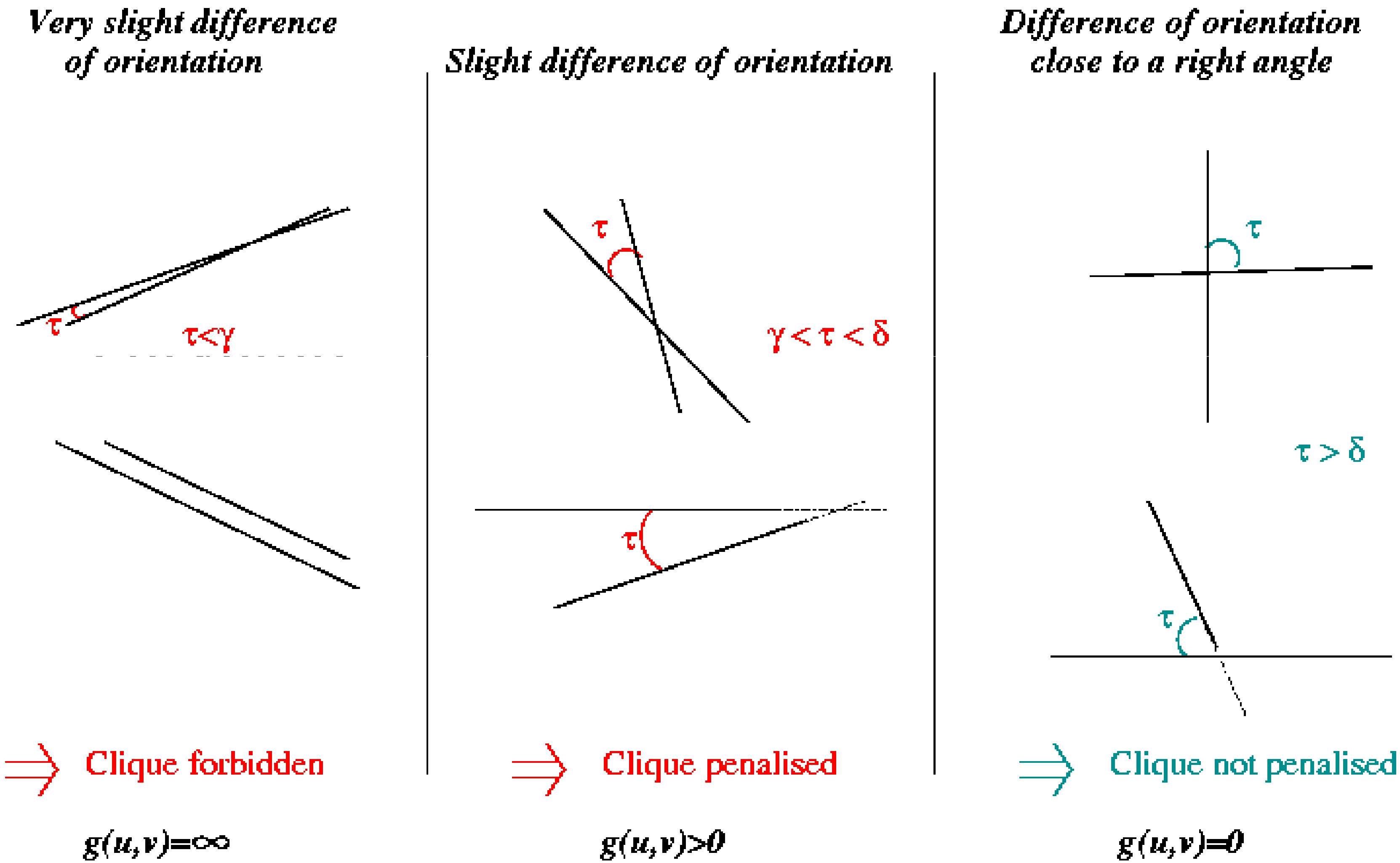
PhDs: R. Stoïca, C. Lacoste in collaboration with IGN and BRGM

First example: Quality Candy Model for road network extraction



PhDs: R. Stoïca, C. Lacoste in collaboration with IGN and BRGM

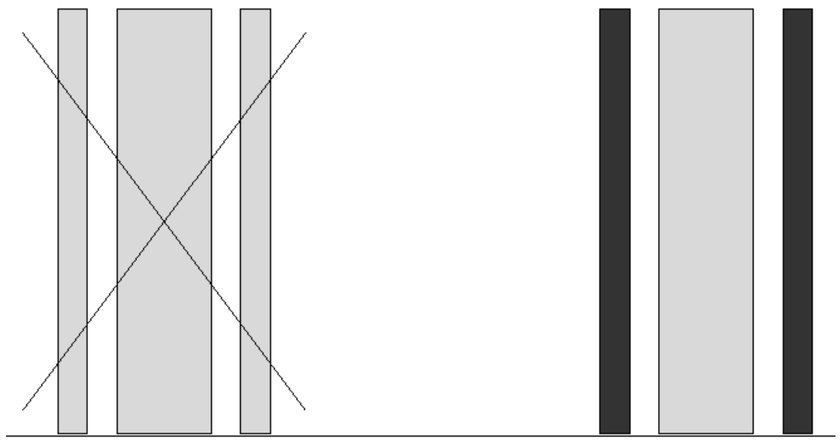
First example: Quality Candy Model for road network extraction



PhDs: R. Stoïca, C. Lacoste in collaboration with IGN and BRGM

First example: Quality Candy Model for road network extraction

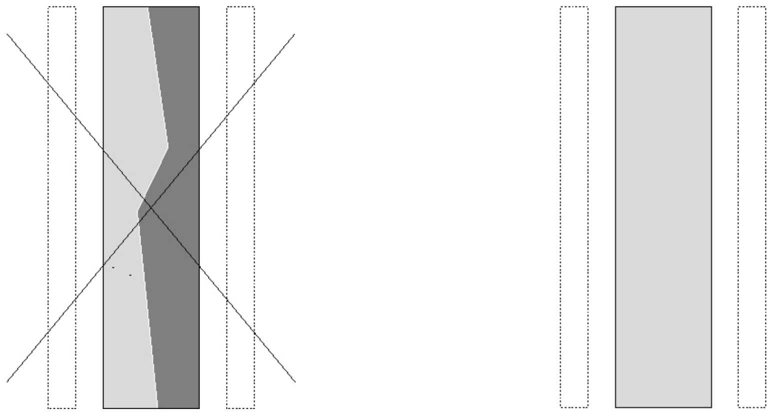
- Objects: Segments
- Prior: models the connectivity and the curvature
- First data term: t-test



PhDs: R. Stoïca, C. Lacoste in collaboration with IGN and BRGM

First example: Quality Candy Model for road network extraction

- Objects: Segments
- Prior: models the connectivity and the curvature
- Second data term: t-test



PhDs: R. Stoïca, C. Lacoste in collaboration with IGN and BRGM

Kernels of the RJMCMC algorithm

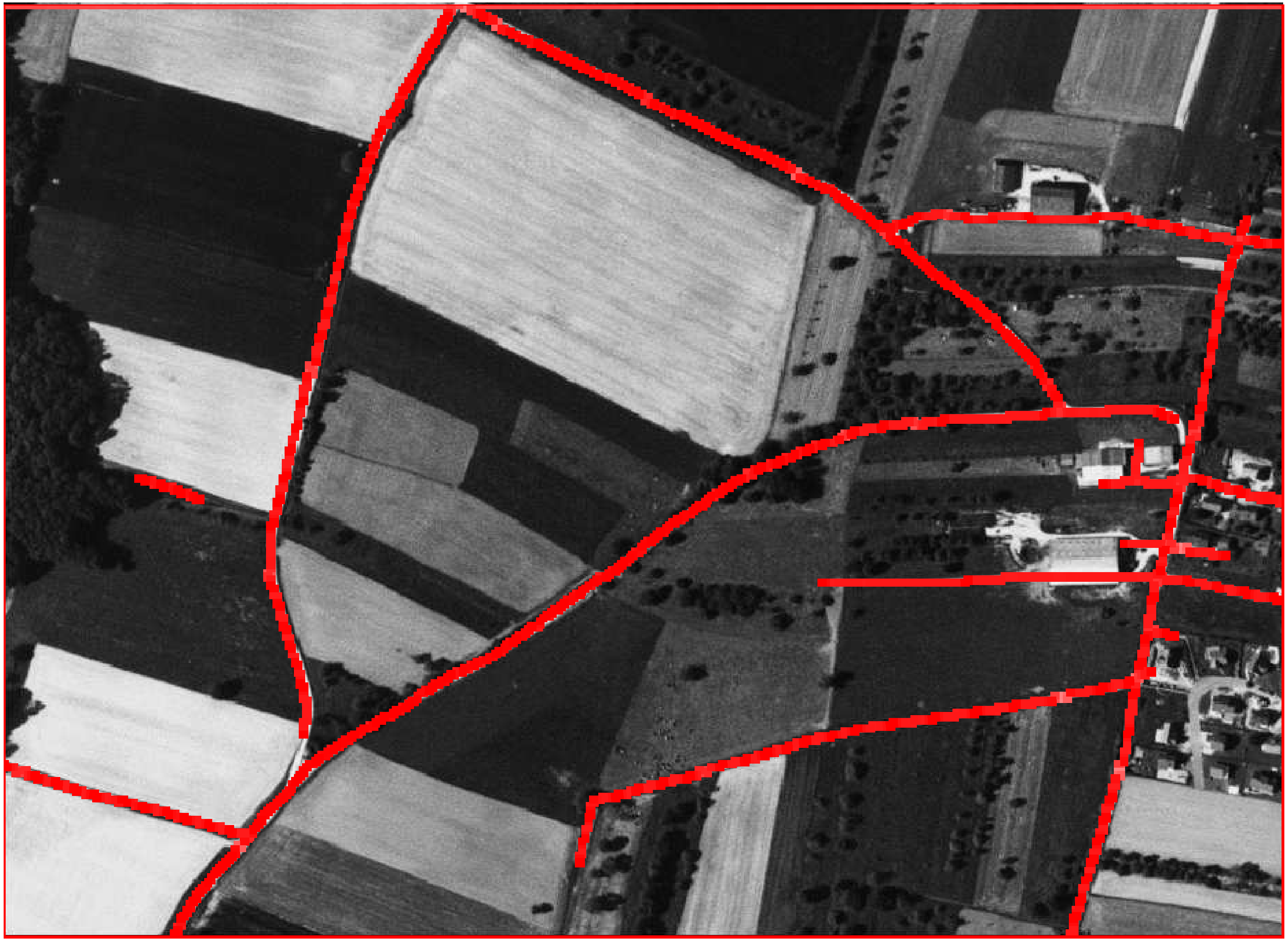
- Uniform birth and death
- Birth and death in a neighborhood
- Extension/contraction of a segment
- Translation of a segment
- Rotation of a segment

PhDs: R. Stoica, C. Lacoste

Results

PhDs: R. Stoica, C. Lacoste

Results



PhDs: R. Stoica, C. Lacoste

Results

PhDs: R. Stoica, C. Lacoste

Results

PhDs: R. Stoica, C. Lacoste

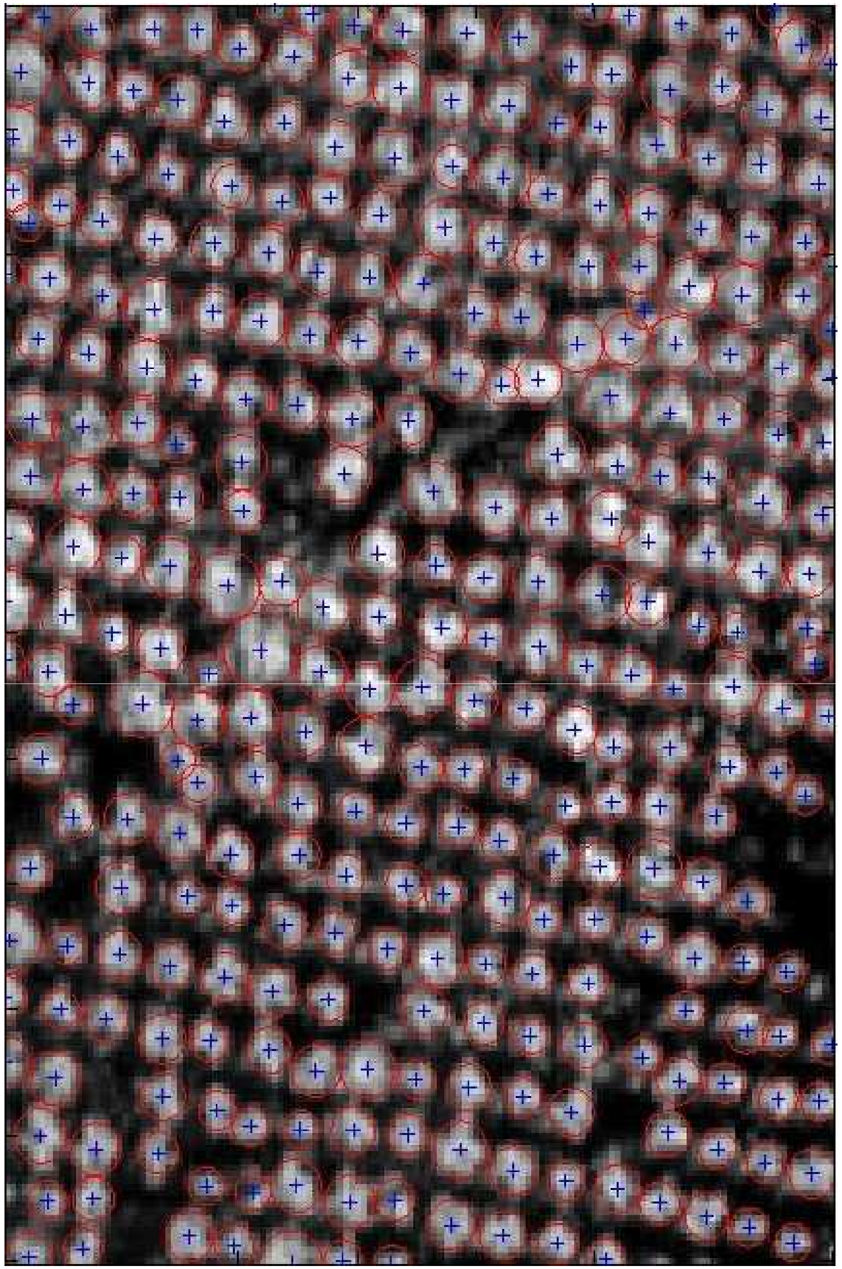
Second example: tree crown extraction
First method

- Object: disc
- Prior: non-overlapping trees
- Data: Gaussian likelihood

$$A_y(S(x)) = \prod_{p \in S(x)} p_{tree}(y_p) \prod_{p \notin S(x)} p_{notree}(y_p)$$

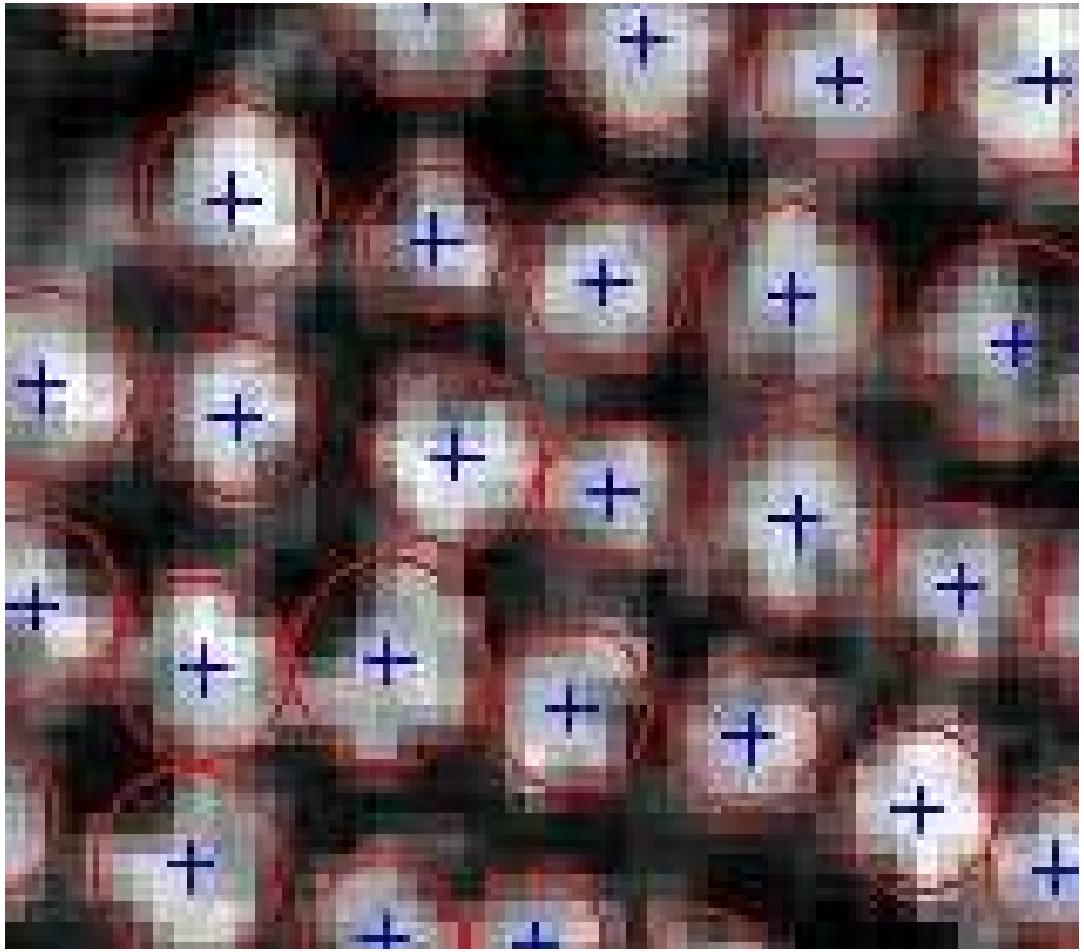
PhD: G. Perrin in collaboration with ECP

Result



PhD: G. Perrin in collaboration with ECP

Result

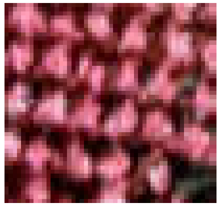


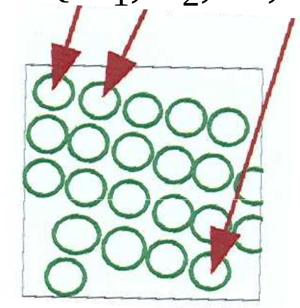
PhD: G. Perrin in collaboration with ECP

Second example: tree crown extraction
Second method

- Marks: ellipses or ellipsoids.

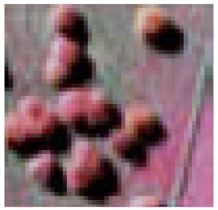
$X = \{X_1, X_2, \ldots, X_n\}$

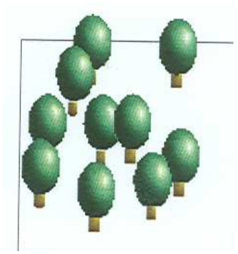




ELLIPSES: 2D model

Dense area: plantation (merged shadows)





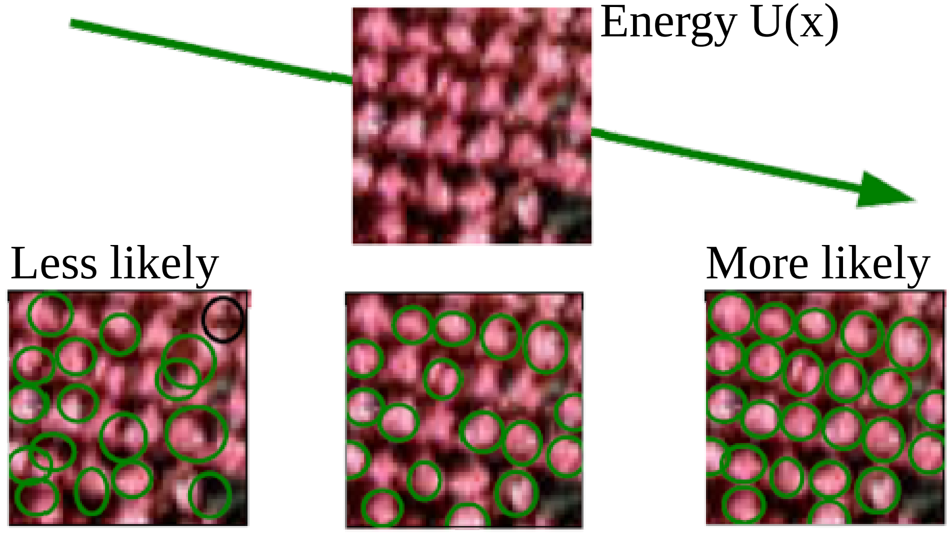
ELLIPSOIDS: 3D model

Sparse vegetation (drop shadows)

PhD: G. Perrin in collaboration with ECP

Density of the process

- Goal: design the density of the MPP in order to make tree configurations be the most likely configurations.
- Minimize the energy: $U(x) : f(x) = \frac{1}{Z} \exp\left(-U(x)\right)$
- Mathematical tools: RJMCMC algorithms + simulated annealing.



Poplars to be extracted with ellipses

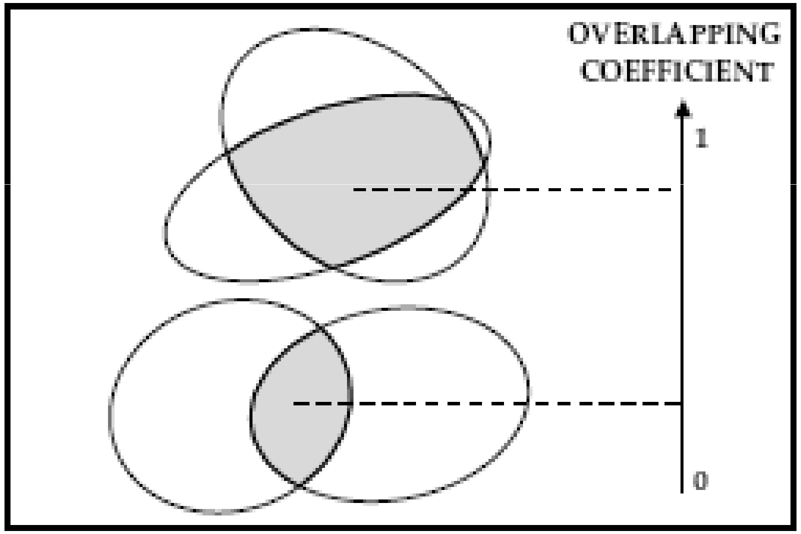
PhD: G. Perrin in collaboration with ECP

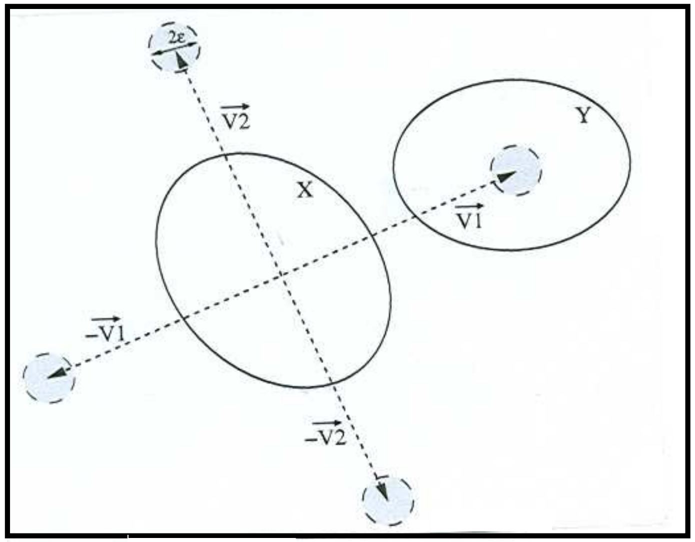
Energy of the model

- Regularizing term + data term:

$$U(x) = U_r(x) + U_d(x)$$

- $U_r(x)$: prior term = interactions between objects.



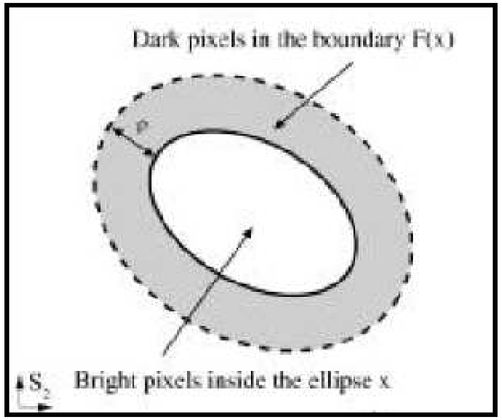


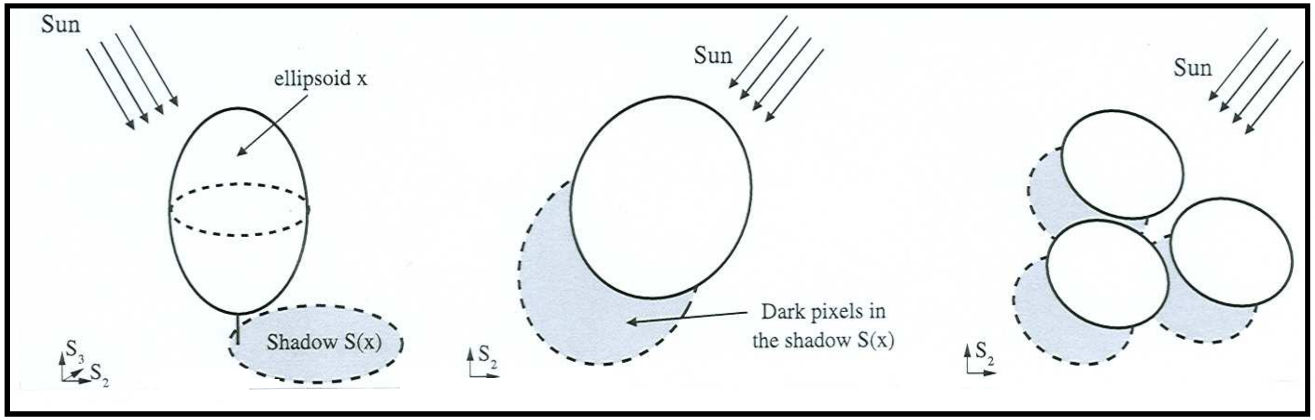
- $U_d(x)$: data term = fitting the object into the image.

$$U_d(x) = \gamma_d \sum_{x_i \in x} U_d(x_i)$$

Data energy term $U_d(x)$

- What is typical of the presence of a tree ?
  - high reflectance in the near infrared.
  - shadow.
  - neighborhood.
- In dense vegetation: merged shadows, shadow area = all around the tree.
- In sparse vegetation: drop shadows, shadow area = in the direction of the sunlight.

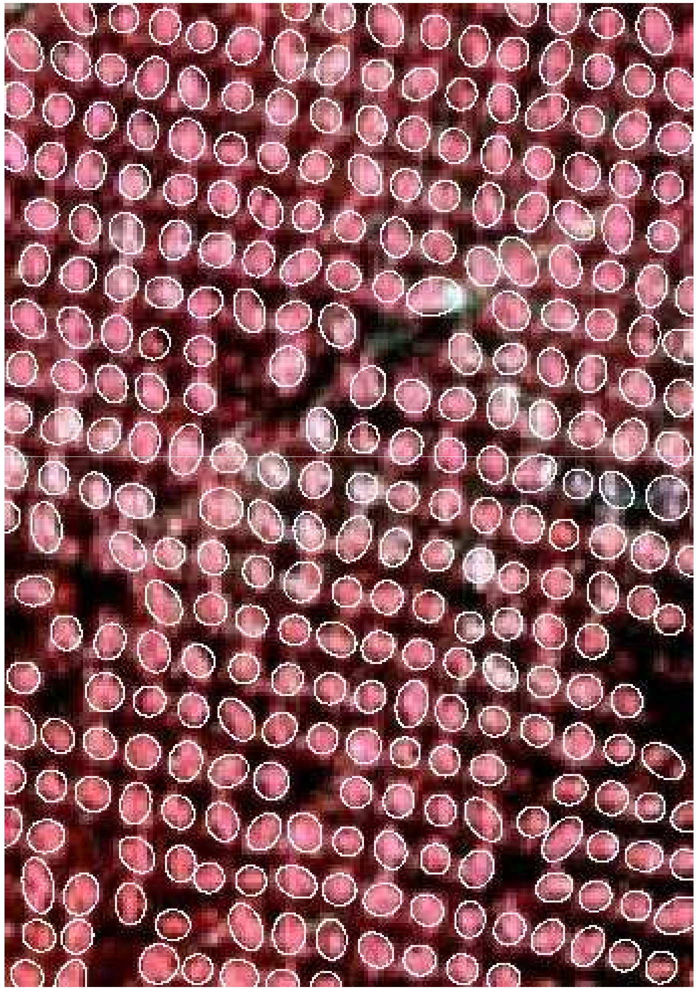




PhD: G. Perrin in collaboration with ECP

Results with the 2D model (1)

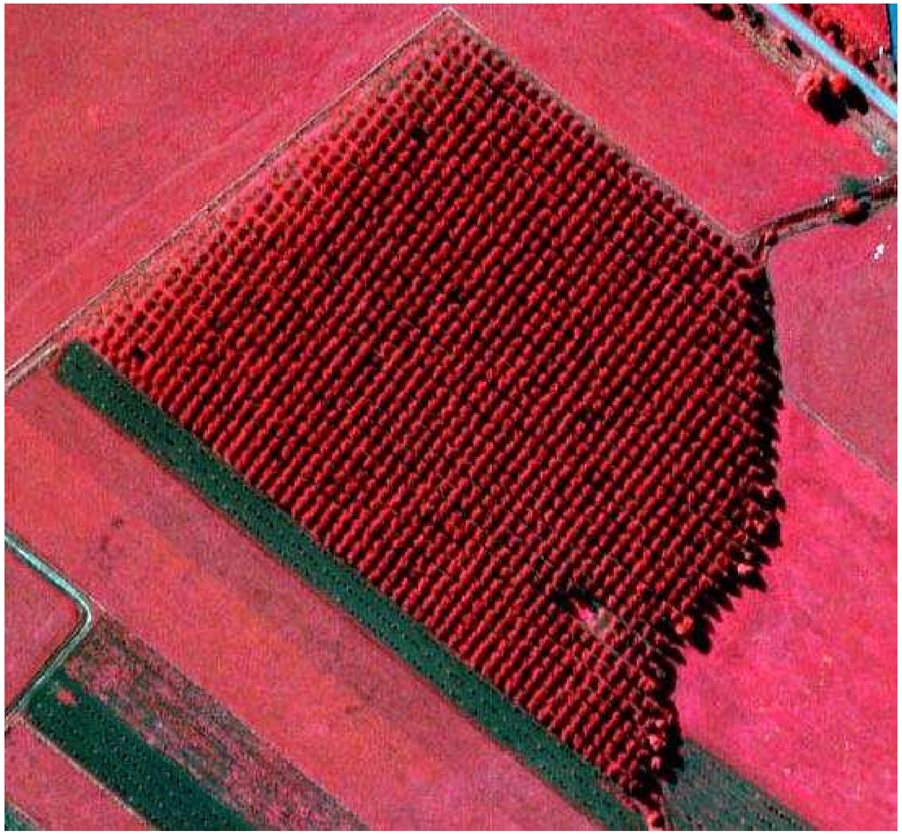
Poplar plantation. 1 ha ©IFN.



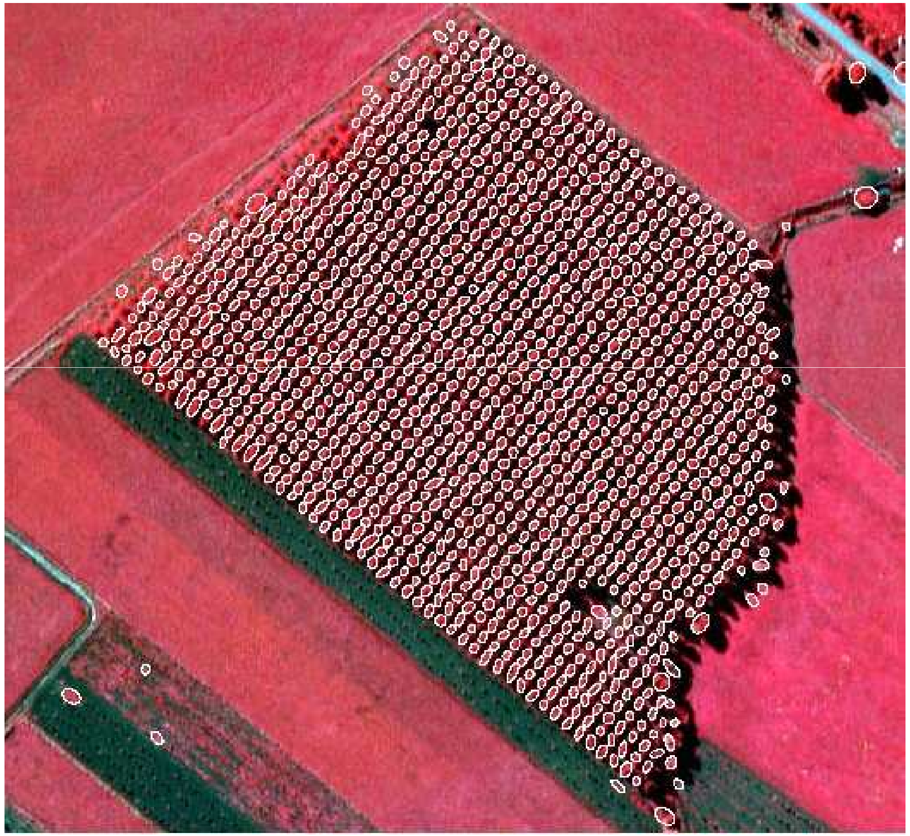
2D model extraction. © Ariana / INRIA

PhD: G. Perrin in collaboration with ECP

Results with the 2D model (2)



Poplar plantation. 7 ha ©IFN



2D model: more than 1300 objects. © Ariana / INRIA

PhD: G. Perrin in collaboration with ECP

Results with the 3D model (1)

• Application: sparse vegetation, trees on the borders of plantations, mixed height stands.

• Hypotheses: the position of the Sun is given, trees close to the nadir and at ground level (no deformation).

• Results: position, crown diameter, approximate height of the tree.

© IFN

© IFN

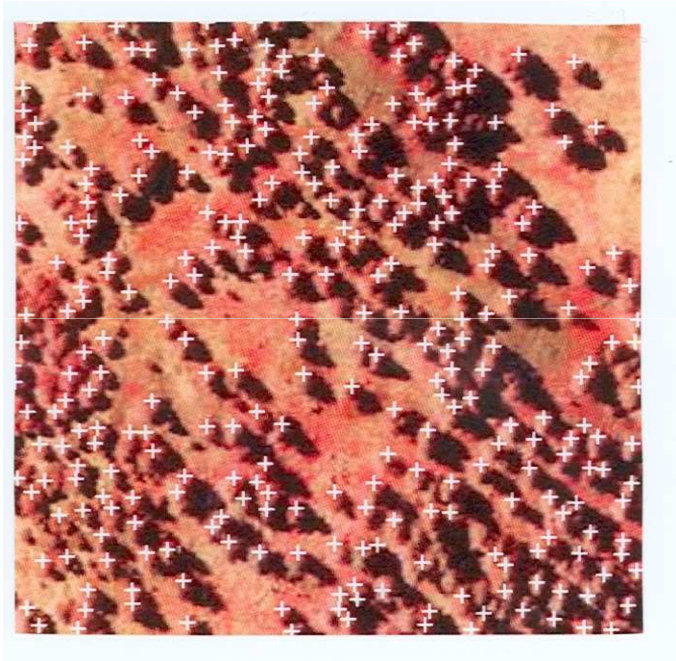
© IFN

PhD: G. Perrin in collaboration with ECP

Results with the 3D model (2)

• 3D model extraction in sparse vegetation.

2.5 ha (Alpes Maritimes) © IFN.

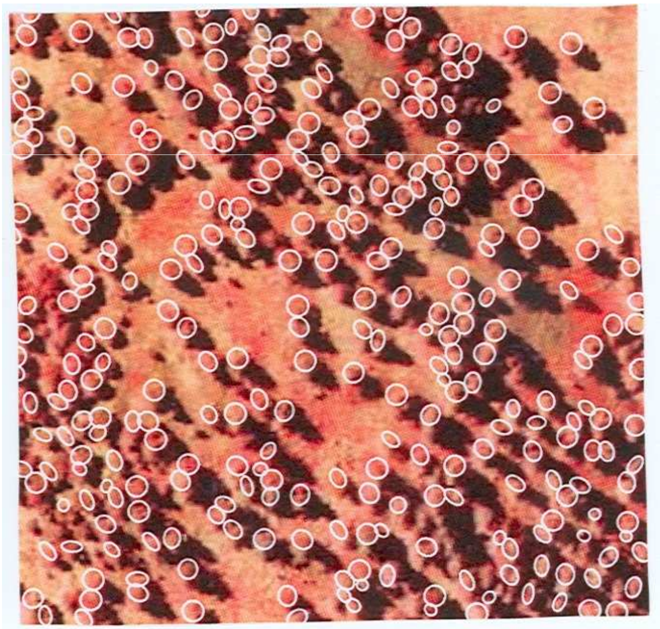


3D model extraction © Ariana / INRIA

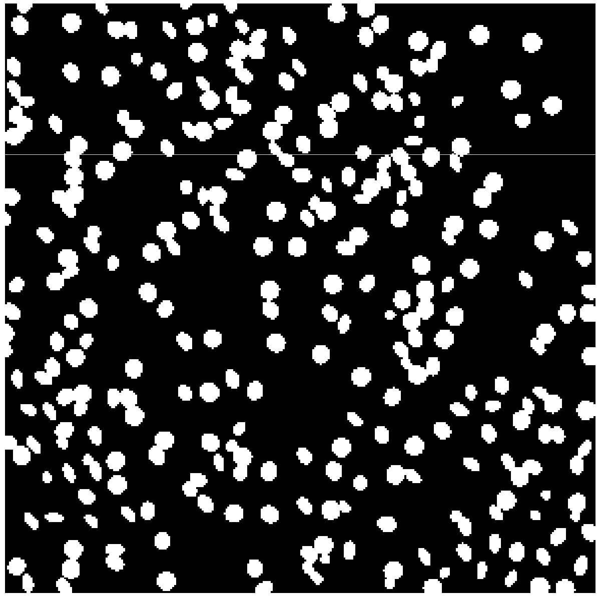
PhD: G. Perrin in collaboration with ECP

Results with the 3D model (3)

- Application: density of the sparse vegetation $\approx$ 19%.



3D model extraction. © Ariana / INRIA



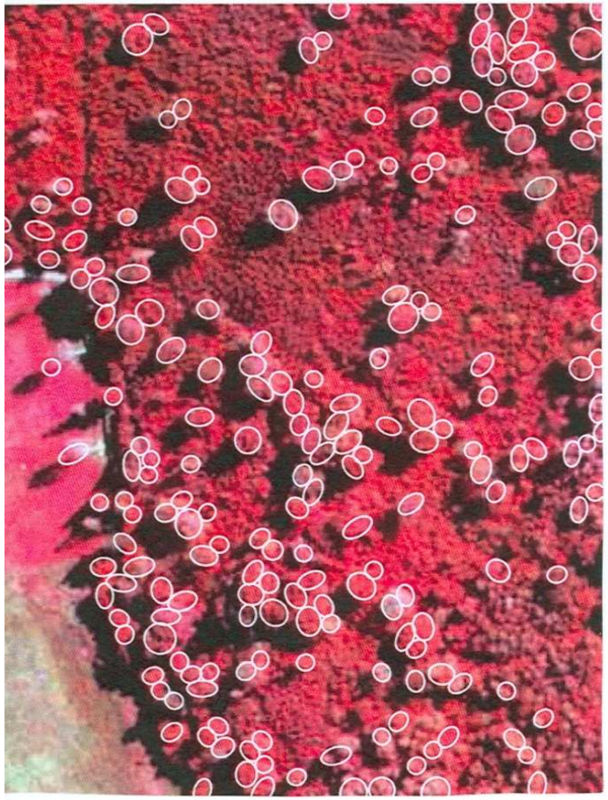
Binary image of the vegetation.

PhD: G. Perrin in collaboration with ECP

Results with the 3D model (4)

- Many objects.
- Information on the timber forest density $\approx 15\%$.

Mixed height stand (3 ha) © IFN.



3D model extraction © Ariana / INRIA

PhD: G. Perrin in collaboration with ECP

Third example: building extraction

Long-term goal: Creation of 3D urban databases

- public (urban planning, disaster recovery …)
- private (wireless telephony, movies …)
- military (operation training, missile guidance …)

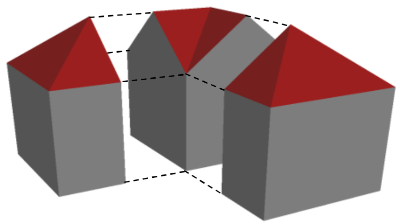
Third example: building extraction

Context

- spatial data (PLEIADES simulations)
- single type of data: a DEM
- automatic (without cadastral maps, without focalisation process)
- dense urban areas

Towards structural modeling

- adapted to data (object approach)
- good compromise generality / robustness
- modular



A building = an assembly of simple urban structures

2 stages: 2D extraction, then 3D reconstruction

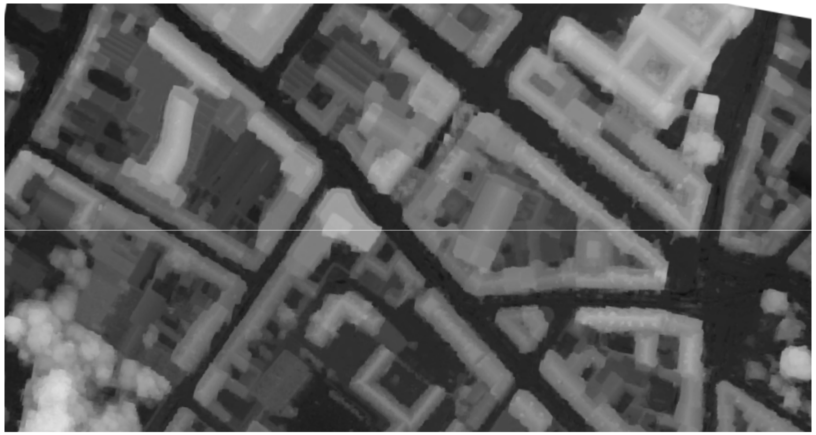
- computation is greatly reduced

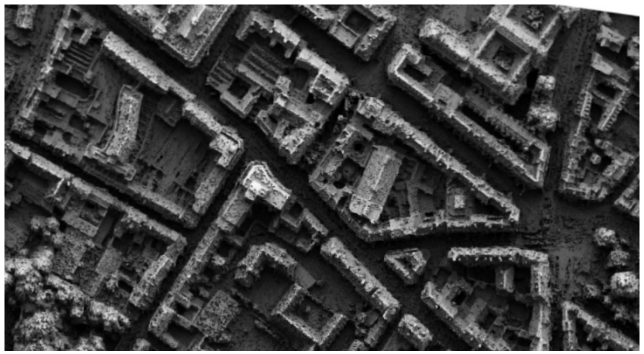
PhDs: M. Ortner, F. Lafarge in collaboration with IGN/CNES/DGA

Stereoscopy

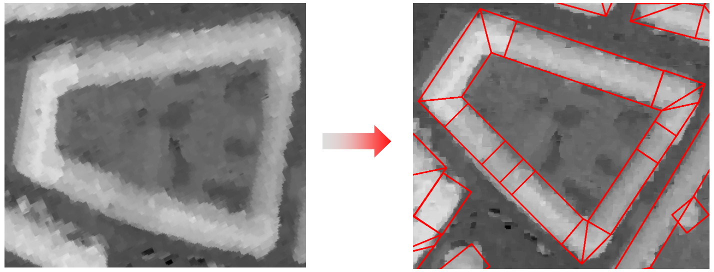
Pair of stereoscopic images

3D Information
example: Digital Elevation Model (DEM)
by [Pierrot-Deseilligny et al.,06]





Stage 1:
2D extraction of buildings



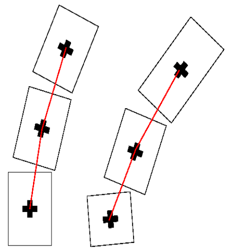
PhDs: M. Ortner, F. Lafarge in collaboration with IGN/CNES/DGA

2D extraction of buildings

Outlines of buildings by marked point processes [Ortner04]

- Energy minimization: $U = \rho U_{ext} + U_{int}$

- U_{ext} : data term
  - coherence between the location of a rectangle and discontinuities in the DEM



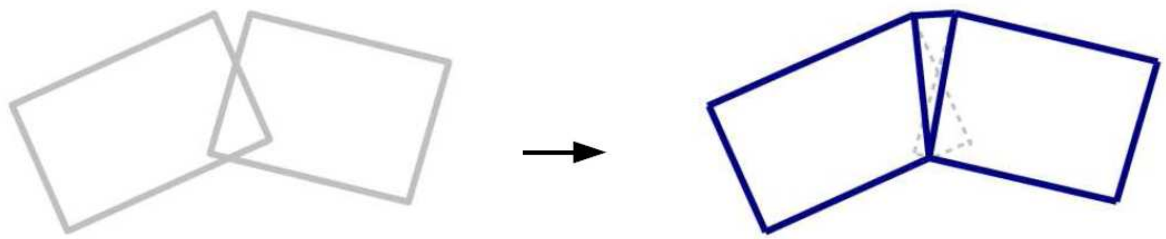
- U_{int} : regularizing term
  - introduction of prior knowledge about the object layout (alignment, paving, completion)

PhDs: M. Ortner, F. Lafarge in collaboration with IGN/CNES/DGA

2D extraction of buildings

Transformation of rectangles into structural supports [Lafarge07]

- transformation of rectangles into unspecified quadrilaterals which are ideally connected (without overlapping, with a common edge)

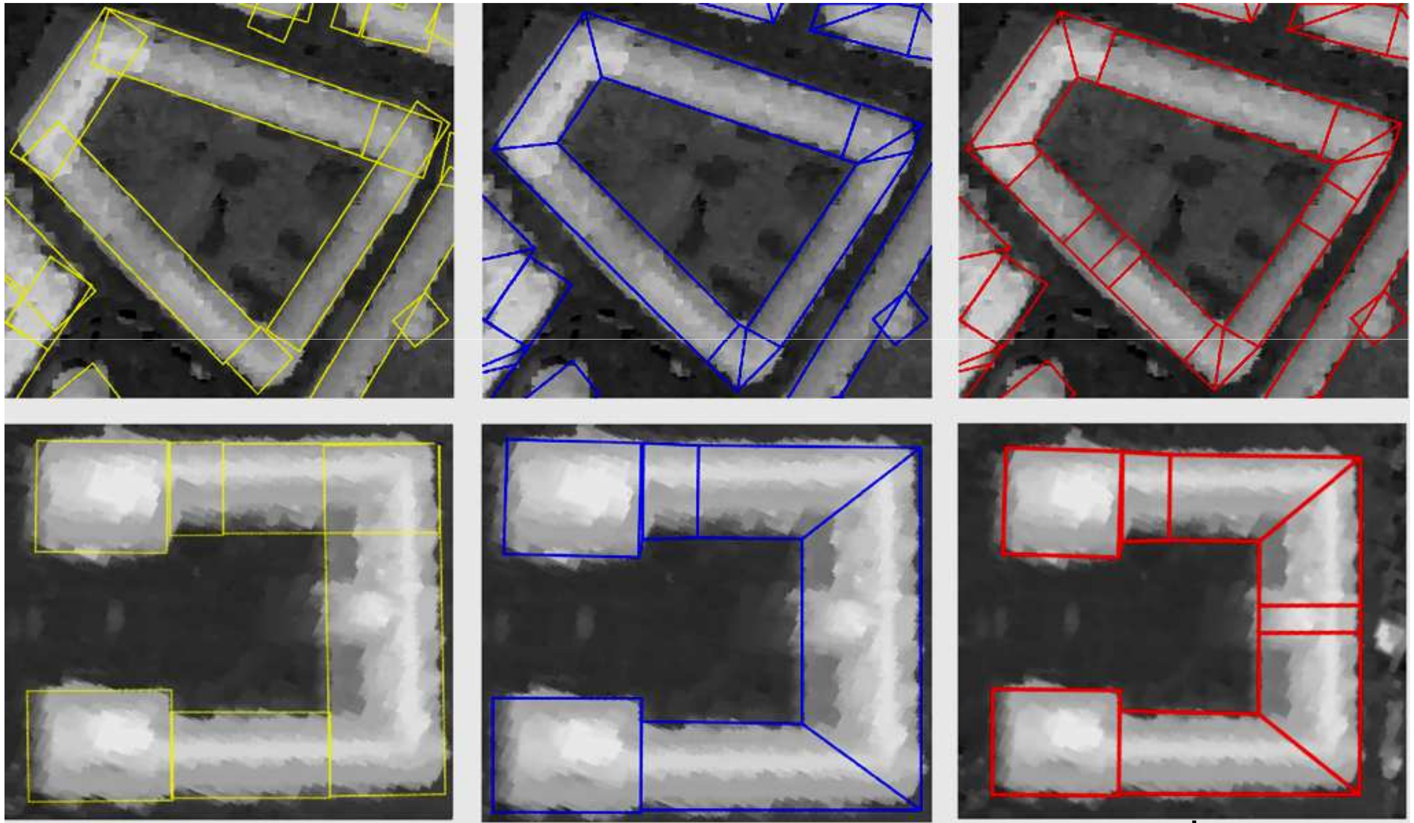


- partitioning of rectangles which represent different urban structures

PhDs: M. Ortner, F. Lafarge in collaboration with IGN/CNES/DGA

2D extraction of buildings

Examples © Ariana / INRIA



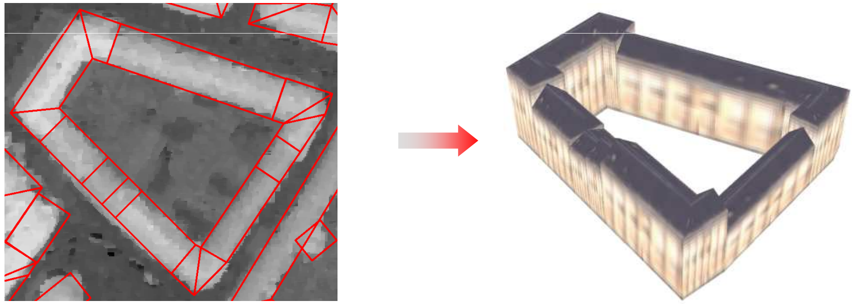
rectangular supports by [Ortner04]

“connected” supports by [Lafarge07]

structural supports by [Lafarge07]

PhDs: M. Ortner, F. Lafarge in collaboration with IGN/CNES/DGA

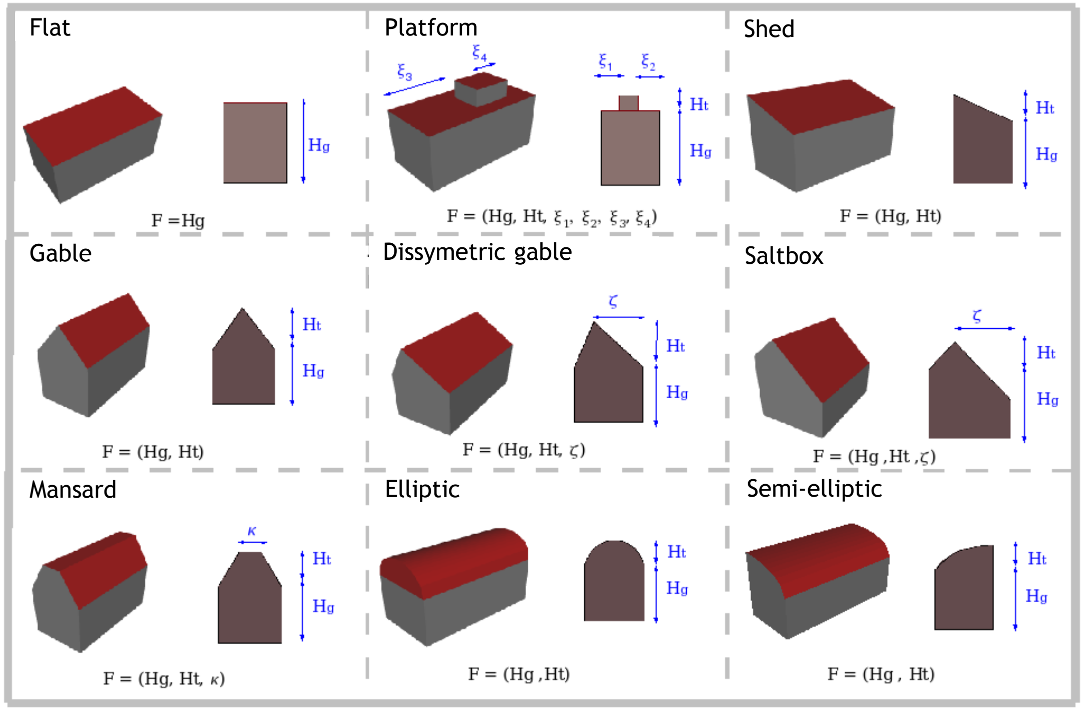
Stage 2:
3D reconstruction of buildings



PhDs: M. Ortner, F. Lafarge in collaboration with IGN/CNES/DGA

3D reconstruction of buildings

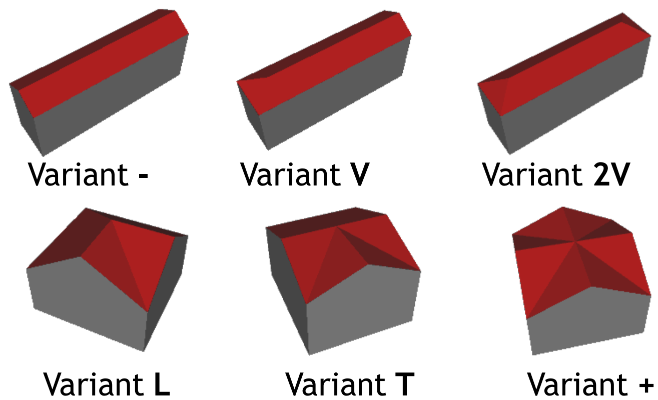
Library of 3D models



The roof shapes:

- 9 forms
- 1 to 6 parameters
- includes curved roofs

The variants:



- ends and junctions
- orientation of the object

PhDs: M. Ortner, F. Lafarge in collaboration with IGN/CNES/DGA

Inverse problem

Notations

- Q, a configuration of structural supports associated with the DEM Λ
- $\mathcal{Y}$, the data such that $\mathcal{Y} = (\mathcal{Y}_i)_{i \in Q}$ with $\mathcal{Y}_i = \{\Lambda(s) \in I / s \in S_i\}$
- x, a configuration of 3D objects $x = (x_i)_{i \in Q}$ where $x_i = (m_i, \theta_i)$ is an objet specified by a model m_i of the library and a parameter set θ_i
- $\mathcal{C}$, the set of 3D object configurations

Inverse problem

- to find the optimal configuration x from the observations $\mathcal{Y}$
- a posteriori density: $h(x) = h(x/\mathcal{Y}) \propto h_p(x)\mathcal{L}(\mathcal{Y}/x)$

Likelihood

Likelihood $\mathcal{L}(\mathcal{Y}/x)$

- to measure the coherence of the observations $\mathcal{Y}$ with an object configuration x

- hypothesis of conditional independence of data:

$$\mathcal{L}(\mathcal{Y}/x) = \prod_{i \in Q} \mathcal{L}(\mathcal{Y}_i/x_i)$$

- use of an altimetric distance between object and DEM:

$$\mathcal{L}(\mathcal{Y}_i/x_i) \propto \exp -\Gamma(\mathcal{S}_{x_i}, \mathcal{Y}_i)$$

where $\mathcal{S}_{x_i}$ corresponds to the roof altitude of object x_i
Γ is the distance (Lp norm)

A priori

A priori $h_p(x)$

- to introduce knowledge w.r.t. the assembling of the objects
  - to compensate for the lack of information contained in the DEM
  - to have realistic buildings
- must be simple (avoid too many tuning parameters)

➡ Solution: a unique type of binary interactions

- Neighboring relationship $\bowtie$ between 2 supports (common edge)
- assembling relation $\sim_a$ between 2 objects
- use of a Gibbs energy: $h_p(x) = \exp -U_p(x)$

A priori

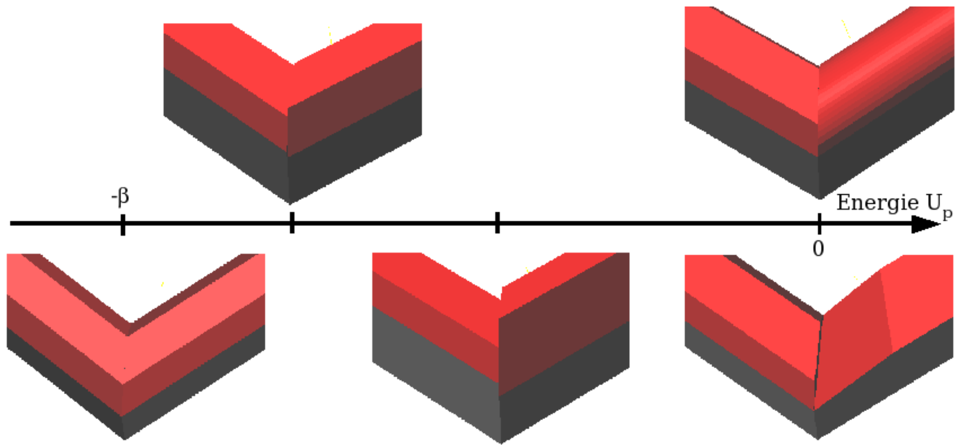
- the assembling relation $\sim_a$ between 2 objects is true if:

  - two objects have the same roof form
  - rooftop orientations are compatible
  - the common edge is not a roof height discontinuity

- A priori expression: $U_p(x) = \beta \sum_{i \bowtie j} \mathbb{1}_{\{x_i \sim_a x_j\}} g(x_i, x_j)$

where $\beta \in \mathbb{R}^+$ is a tuning parameter

g measures the distance between parameters of the objects



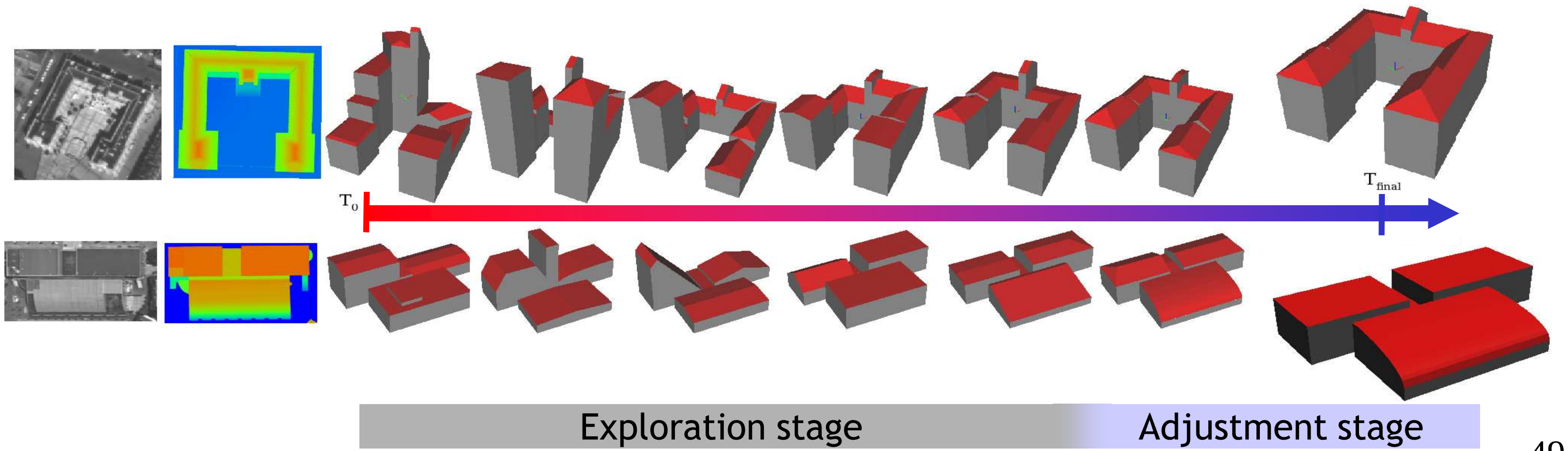
Optimization

MAP estimator: $x_{MAP} = \arg\max_{x \in \mathcal{C}} h(x)$

- non convex optimization problem in a large state space
- $\mathcal{C}$ is a union of spaces of different dimensions

RJMCMC sampler[Green95]

- consists in simulating a Markov chain $(X_t)_{t \in \mathbb{N}}$ on $\mathcal{C}$ which converges toward a target measure π specified by h



Results

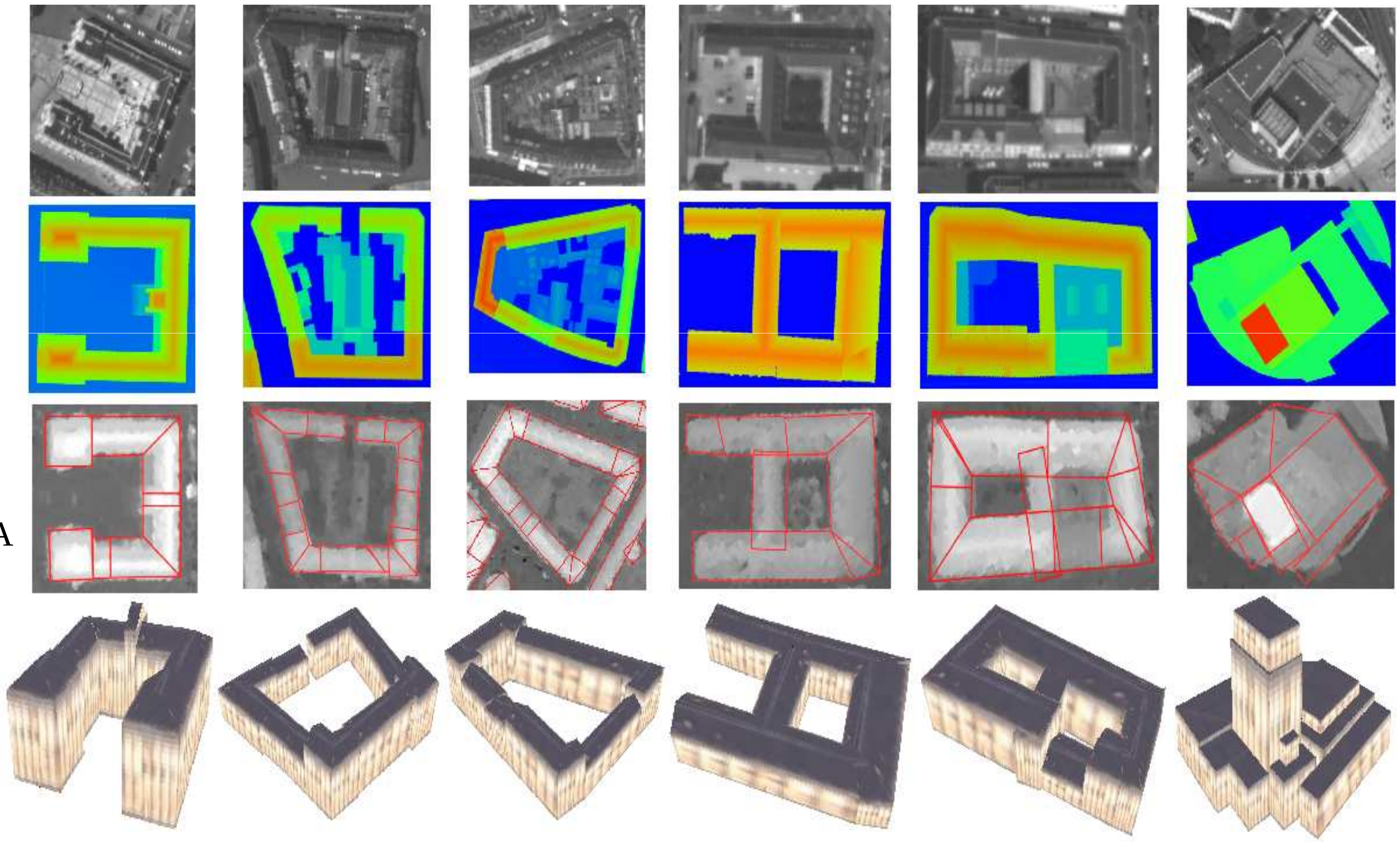
Reconstruction with automatic support extraction

PLEIADES simulations ©CNES

Ground truth ©IGN

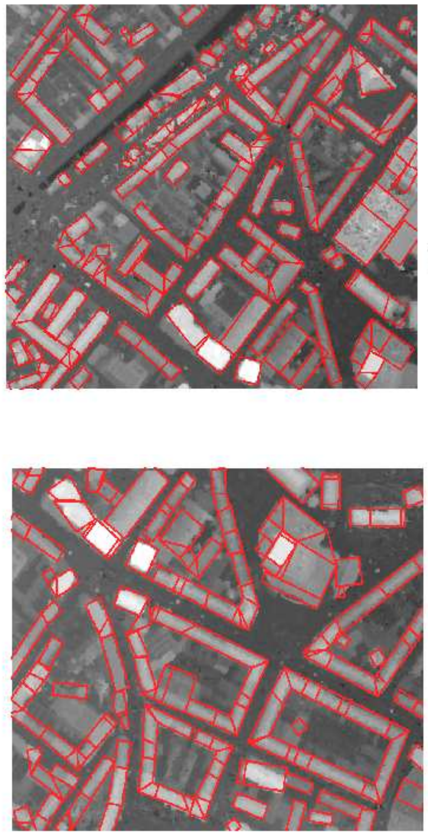
Buiding Extraction ©Ariana/INRIA

3D Reconstruction ©IGN/CNES

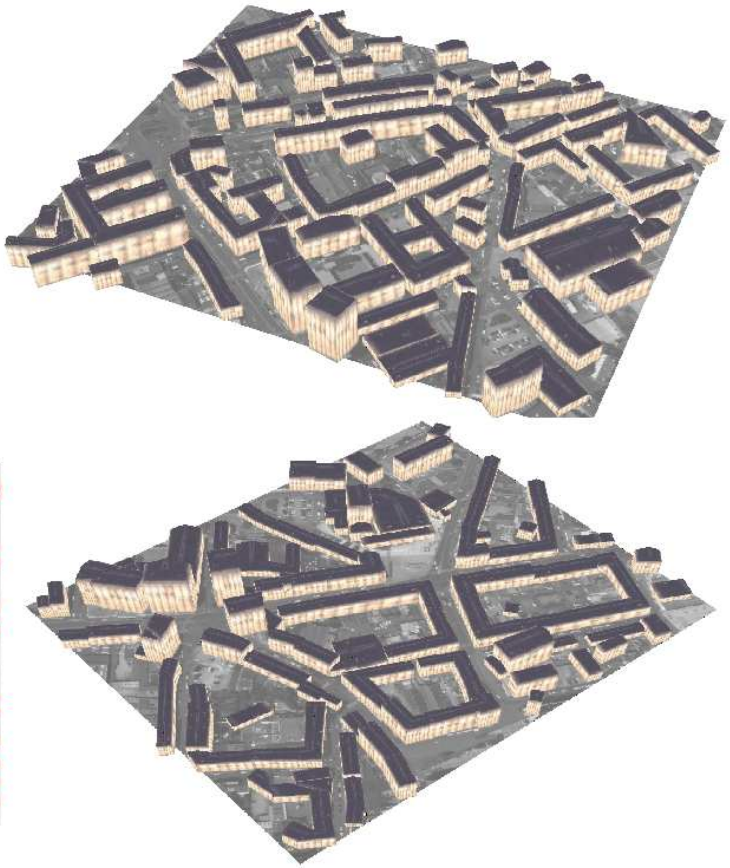


Results

PLEIADES simulations
© CNES



Building Extraction
© Ariana / INRIA

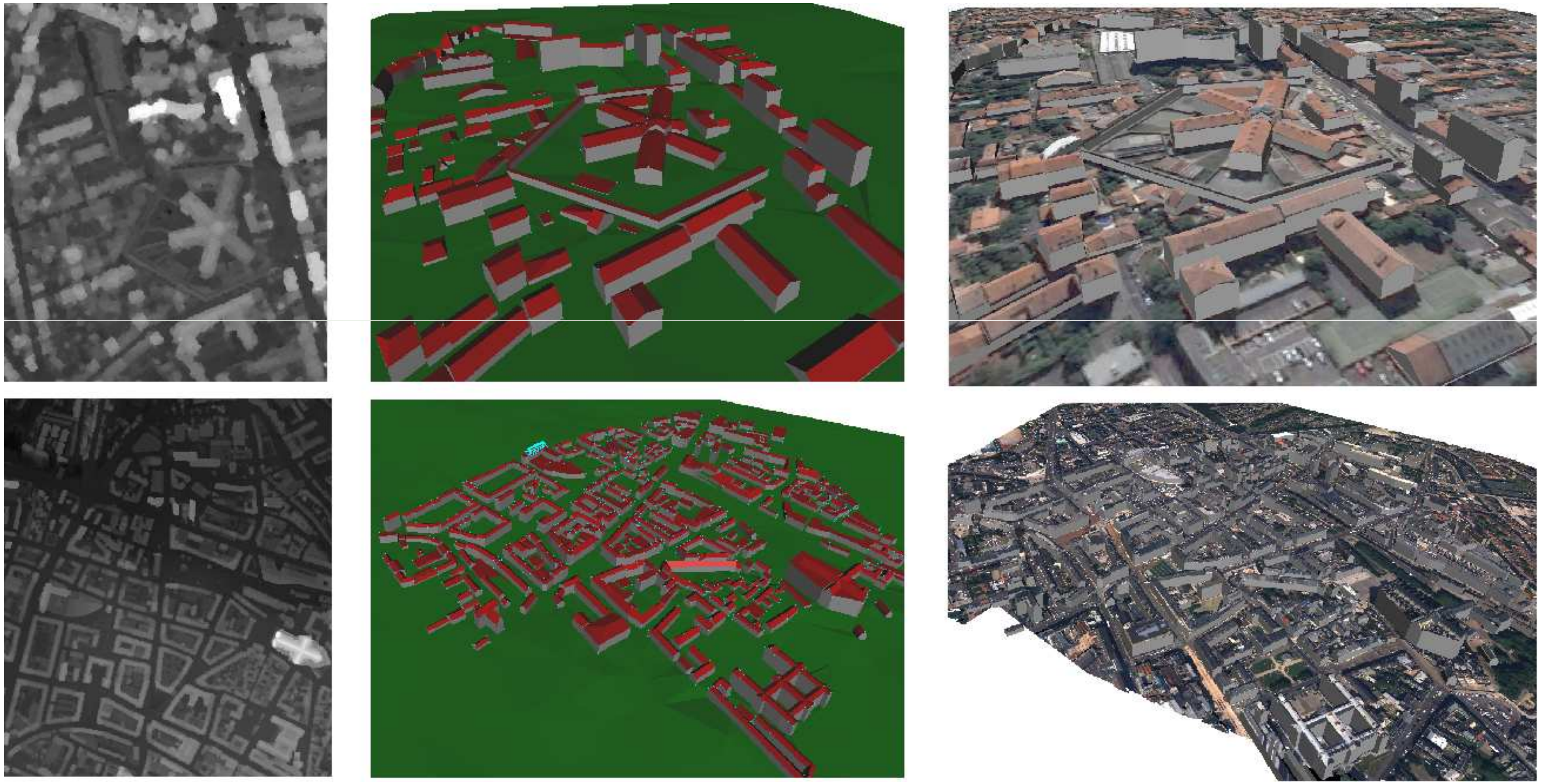


3D Reconstruction
© IGN / CNES

PhDs: M. Ortner, F. Lafarge in collaboration with IGN/CNES/DGA

Results

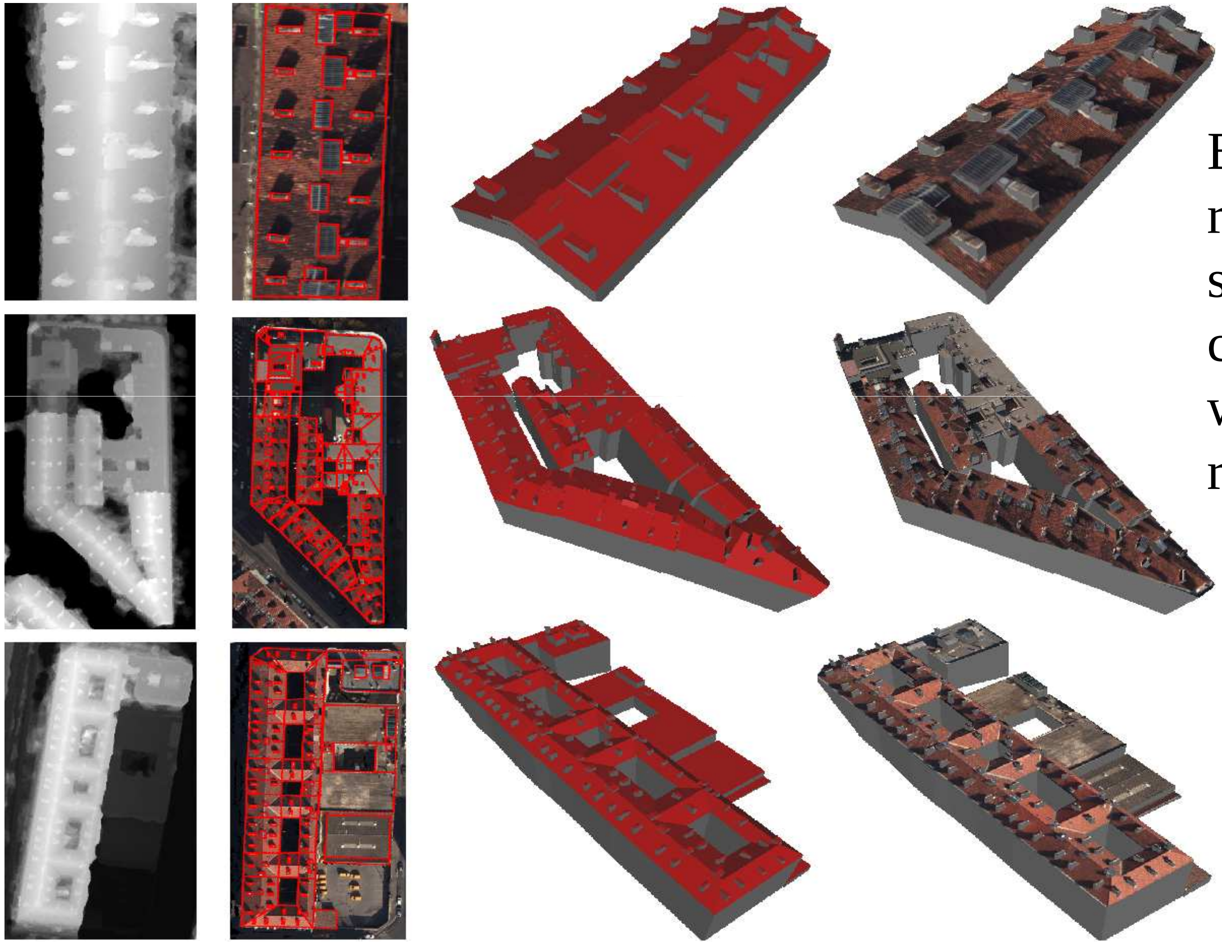
Reconstruction with interactive support extraction



Reconstruction of urban areas (Amiens downtown and St Michel prison in Toulouse)

Results

Reconstruction with interactive support extraction



Better reconstruction of superstructures: chimneys, dormer windows, glass roofs...

Reconstruction of buildings with superstructures (Marseille) from 0.1 meter resolution aerial DEM

Remarks

- Interesting characteristics
  - original and difficult context: satellite data – a single DEM – automatic without cadastral maps – dense urban areas
  - evolutive process (different roof models, various data resolutions)
  - possibility of using the extraction and reconstruction processes separately
- Limits
  - restricted use (possible problems if discontinuities in DEMs, vegetation, inner courtyards)
  - computing time
  - no 2D correction between extraction and reconstruction stages

General conclusion

- The marked point process framework extends the application domain of Markov Random Field approaches:
  - Data taken into account at the object level
  - Geometrical information taken into account
- Markov random fields are still an efficient tool (depending on the image resolution)

Future work

- Point process with marks living in a shape space:
  - More accurate definition of the geometry
  - Computing issues
- Multiple object detection [Lafarge09]
- New optimization dynamics
  - Diffusion processes (in progress at Ariana, in collaboration with IIPT, Moscow, RAS)
- Parameter estimation (in progress at Ariana, in collaboration with CNES)

Thanks to:

• CNES, IFN, IGN, Spot Image Corporation, Satellite Image Corporation, for providing the data.

• DGA, MAE, CNES, IGN, BRGM, ECP and INRIA, for partial financial support.

Bibliography

- **A Gibbs point process for road extraction in remotely sensed images**. R. Stoica et X Descombes et J. Zerubia. *International Journal of Computer Vision*, 57(2): pages 121-136, 2004.

- **Point Processes for Unsupervised Line Network Extraction in Remote Sensing**. C. Lacoste et X. Descombes et J. Zerubia. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 27(10): pages 1568-1579, 2005.

- **Adaptive Simulated Annealing for Energy Minimization Problem in a Marked Point Process Application**. G. Perrin et X. Descombes et J. Zerubia. In *Proc. Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR)*, St Augustine, Florida, USA, November 2005.

- **A comparative study of three methods for identifying individual tree crowns in aerial images covering different types of forests**. M. Eriksson et G. Perrin et X. Descombes et J. Zerubia. In *Proc. International Society for Photogrammetry and Remote Sensing (ISPRS)*, Marne La Vallee, France, July 2006.

Bibliography

- **Building Outline Extraction from Digital Elevation Models using Marked Point Processes**. M. Ortner et X. Descombes et J. Zerubia. *International Journal of Computer Vision*, 72(2): pages 107-132, 2007.

- **Automatic Building Extraction from DEMs using an Object Approach and Application to the 3D-city Modeling**. F. Lafarge et X. Descombes et J. Zerubia et M. Pierrot-Deseilligny. *Journal of Photogrammetry and Remote Sensing*, 63(3): pages 365-381, 2008.

- **Structural approach for building reconstruction from a single DSM**. F. Lafarge et X. Descombes et J. Zerubia et M. Pierrot-Deseilligny. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2009.

For more information:

http://www.inria.fr/ariana

Hammersley-Clifford Theorem

A MRF verifying a positivity constraint can be written as a Gibbs field:

$$P(X) = \frac{1}{Z} \exp - \left[\sum_{c \in C} V_c(x_s, s \in S) \right]$$

$S =$ all the pixels

$C =$ all the cliques associated to the neighborhood v

Markov process

- A point process density $\boldsymbol{f} : \boldsymbol{N}^{f} \rightarrow [0, \infty[$

  is Markovian under the neighborhood relation ~ if and only if there exists a measurable function $\phi : \boldsymbol{N}^{f} \rightarrow [0, \infty[$ such that:

$$\boldsymbol{f}(\boldsymbol{x}) = \alpha \prod_{\text{cliques}.\boldsymbol{y} \subseteq \boldsymbol{x}} \phi(\boldsymbol{y})$$

for all $\boldsymbol{x} \in \boldsymbol{N}^{f}$

Stability

• Condition required for proving the convergence of Markov Chain Monte Carlo sampling methods.

• A point process defined by its f(.) w.r.t. a reference measure π_ν(.) is locally stable if there exists a real number M such that:

$$\boldsymbol{f}\left(\boldsymbol{x} \cup \{\boldsymbol{u}\}\right) \leq \boldsymbol{M}\boldsymbol{f}\left(\boldsymbol{x}\right), \forall \boldsymbol{x} \in \boldsymbol{N}^{f}, \forall \boldsymbol{u} \in \chi$$