

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

J. Zerubia

Joint work with S. Ben Hadj, F. Chatelain, P. Craciun, S. Descamps, X. Descombes, P. Gernez, C. Lacoste, F. Lafarge M. Ortner, G. Perrin and R. Stoica

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



# Bayesian Approach

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

$P(Y | X)$ : likelihood

$P(X)$  : prior

$P(X | Y)$ : posterior

# Example : Classification

Prior : Markov Random Field

Likelihood : conditional independence assumption

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

No contextual information in the likelihood:

1 - uncorrelated noise

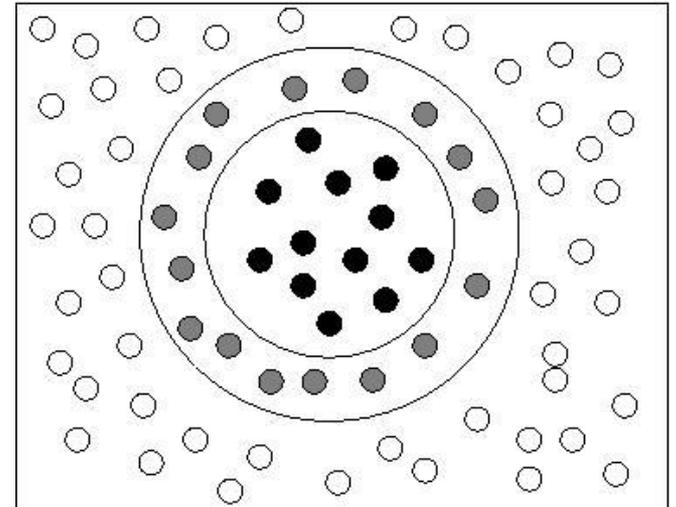
2 - no texture

# Markov Random Field Approach

Markov Random Field :

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

$v_s$  being the neighborhood of  $s$



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

# 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  $\nu$

# Potts Model

$$C = \{((i, j), (i, j + 1)); ((i, j), (i + 1, j)), s = (i, j) \in S\}$$

$$V_c(x_s, x_t) = \begin{cases} 0 & \text{if } x_s = x_t \\ \beta > 0 & \text{if } x_s \neq x_t \end{cases}$$

$$P(X) = \frac{1}{Z} \exp -\#_c$$

$\#_c$  : Number of heterogeneous cliques

# Simple modeling

Prior : Potts model

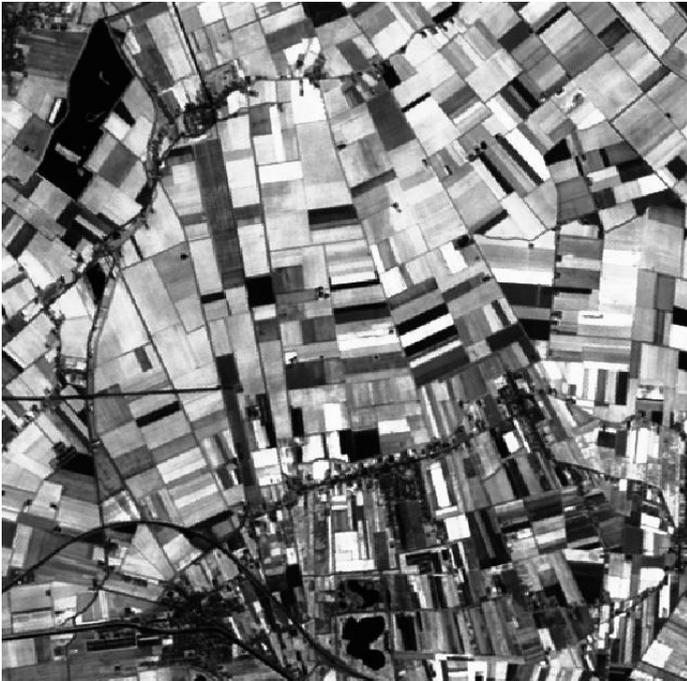
Likelihood : Gaussian model

$$P(X | Y) \propto \exp - [U(X) + U(Y | X)]$$

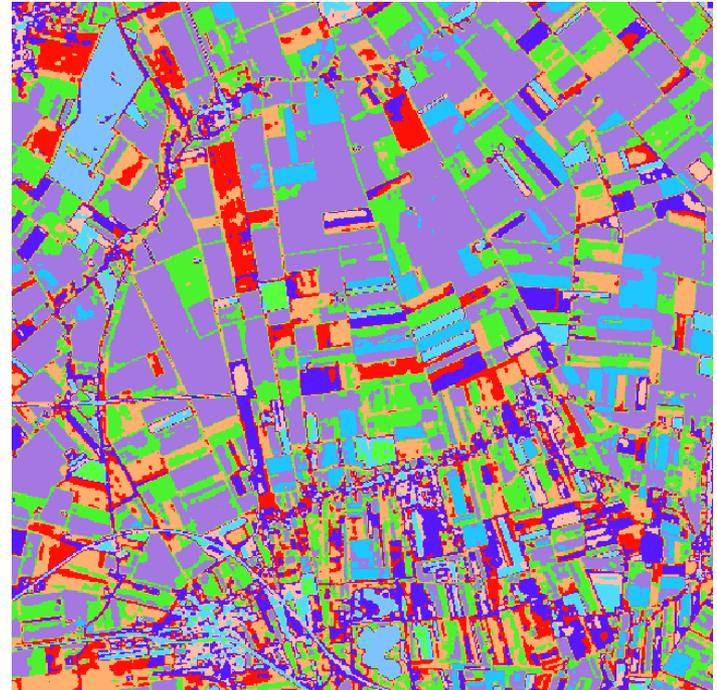
$$U(X) = \beta \sum_{c=\{s,t\} \in C} \delta_{x_s \neq x_t}$$

$$U(Y | X) = \sum_{s \in S} \sum_i \left( (y_s - \mu_i)^2 + \frac{1}{2} \log(\sigma_i^2) \right) \delta_{x_s=i}$$

# Example : Classification



SPOT image © CNES /  
Airbus D & S



Classification result  
© Ariana / INRIA

# From context to geometry



SPOT image © CNES



IKONOS image © Satellite imaging Corporation



IKONOS image © Satellite image Corporation<sup>9</sup>

# From context to geometry



SPOT image © CNES



aerial image © IGN

# From pixels to objects

- Goals :
  - 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).
- But we do not know the number of objects (Markov random fields on graphs are excluded).



Solution : Marked point processes

# Marked Point process defined by a density w.r.t. the Poisson process

- A **marked point process**  $X$  on  $\chi = \mathcal{P} \times \mathcal{M}$  is a point process on  $\chi$  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  $\pi_\nu(\cdot)$  of a Poisson process known as the reference process ( $\nu(\cdot)$  is the intensity measure) :

$$f : N^f \rightarrow [0, \infty[ : \int_{N^f} f(\mathbf{x}) d\pi_\nu(\mathbf{x}) = 1$$

# Markov process

- A point process density  $f : N^f \rightarrow [0, \infty[$  is Markovian under the neighborhood relation  $\sim$  if and only if there exists a measurable function  $\phi : N^f \rightarrow [0, \infty[$  such that :

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

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

# Stability

- Condition required for proving the convergence of MCMC sampling methods.
- A point process defined by its  $f(\cdot)$  w.r.t. a reference measure  $\pi_\nu(\cdot)$  is locally stable if there exists a real number  $M$  such that :

$$f(\mathbf{x} \cup \{\mathbf{u}\}) \leq Mf(\mathbf{x}), \forall \mathbf{x} \in N^f, \forall \mathbf{u} \in \chi$$

# Sampling : Birth and death algorithm (Geyer/Moller-94)

- Birth : with probability  $1/2$ , propose to add a new point  $u$  in  $\chi$  following some density  $\mu(\cdot)$  to the current configuration  $x$ .

Let  $y = x \cup \{u\}$ , compute the ratio :

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

- Death : with probability  $1/2$ , propose to remove a point  $v$  uniformly chosen in  $x$ . Let  $y = x / \{v\}$ , compute the ratio :

$$R_2(x, y) = \frac{f(y) n(x)}{f(x) \nu(\chi)}$$

- With probability  $\alpha_i = \min(1, R_i)$   $i = 1, 2$  accept the proposition  $x_{t+1} = y$ , otherwise accept the proposition  $x_{t+1} = x$ .

# Sampling : RJMCMC (Green-95)

- Mixture of several proposition kernels :

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

- Convergence condition exists.

# RJMCMC

- **Algorithm:**

*At time  $t$ :*

1) *Select randomly a kernel  $\mathbf{q}_m$  using the discrete law  $(\mathbf{p}_m(\mathbf{x}))$*

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

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

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

5) *With probability*      •  $\alpha$       *set:*       $\mathbf{X}_{t+1} = \mathbf{y}$

                                 •  $(1-\alpha)$       *set:*       $\mathbf{X}_{t+1} = \mathbf{x}$

# Optimization

- **Goal** : Estimate a configuration maximizing  $f(\cdot)$
- **Simulated annealing** :

*Successive simulations of  $f_t(\mathbf{x})$   $n(\mathbf{dx})$  using an RJMCMC algorithm with :*

$$f_t(\mathbf{x}) = f(\mathbf{x})^{\frac{1}{T_t}}$$

*where  $(T_t)$  ( $\equiv$  temperature) decreases toward zero.*

- Logarithmic decrease  $\Rightarrow$  global maximum.
- **In practice** : geometric decrease.  
At each step,  $T_{t+1} = T_t \times c$ , where  $c$  is a constant close to 1.  
( $c=0.99999$  or  $c=0.999999$  depending on the difficulty of the detection)

# Summary

## Goal :

To model the observed scene as a configuration of objects (roads, rivers, buildings, trees, flamingos).

- **Stochastic modeling:**

Set of objects in the scene  $\equiv$  realization of a marked point process,  $\mathbf{X}$ .

- **Density:**

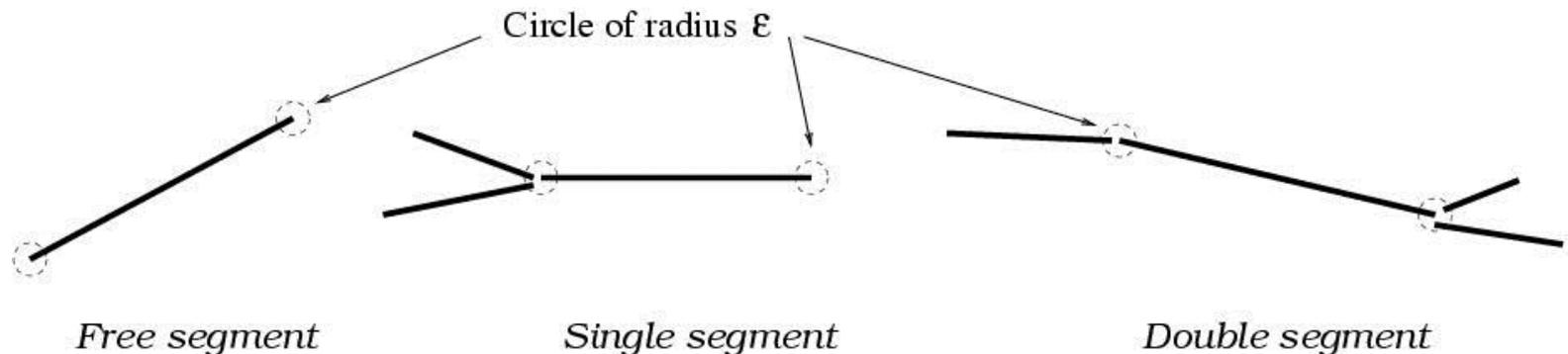
$$f(\mathbf{x}) = \frac{1}{Z} h_d(\mathbf{x}) h_p(\mathbf{x})$$

*data term* ↙ ↘ *prior*

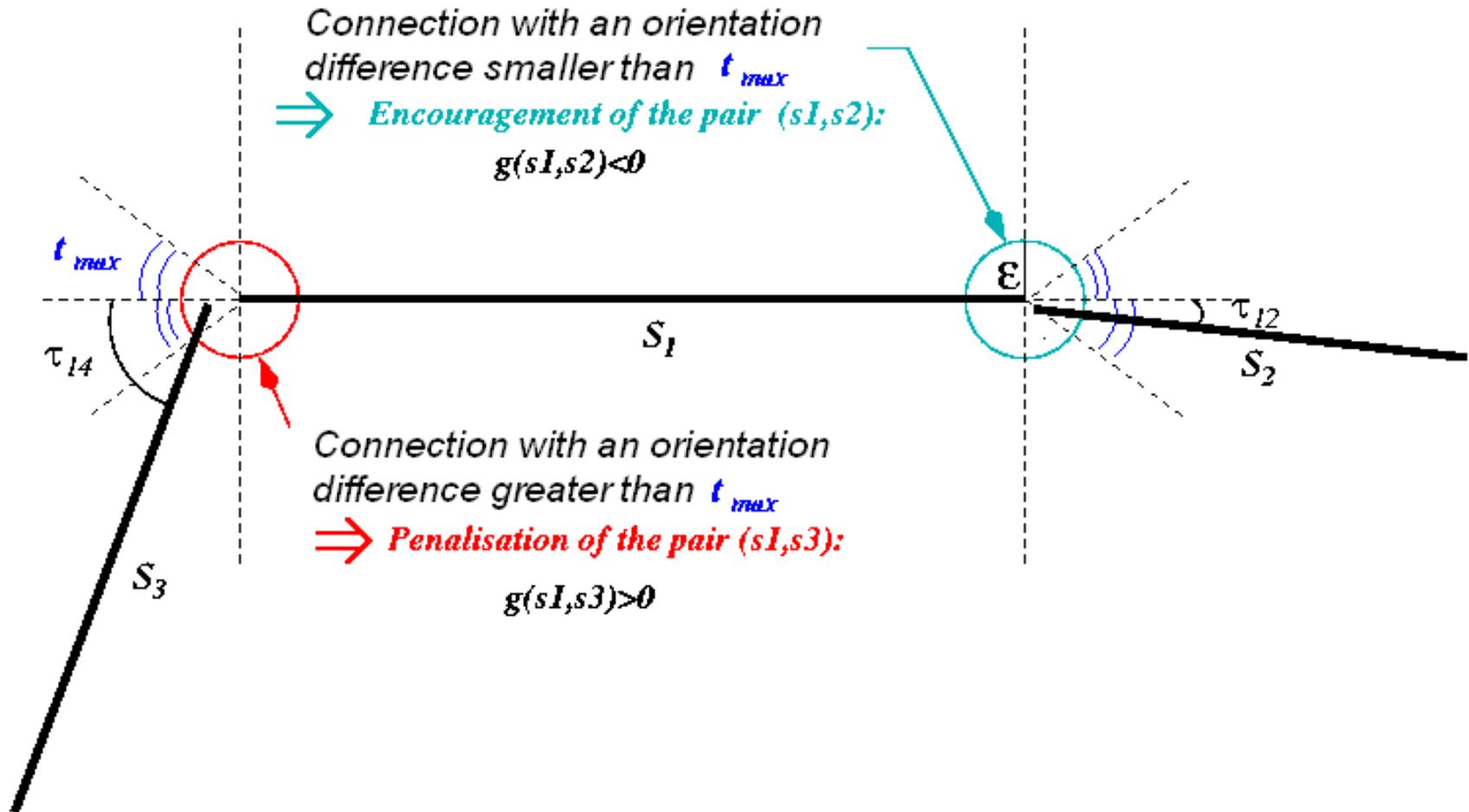
- **Algorithm :** Monte Carlo sampler (e.g. RJMCMC) + simulated annealing

# First example : Quality Candy Model for road network extraction

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

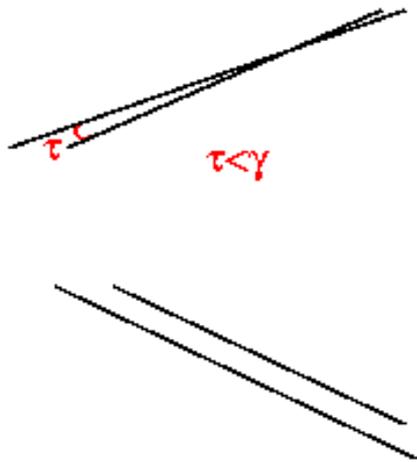


# First example : Quality Candy Model for road network extraction



# First example : Quality Candy Model for road network extraction

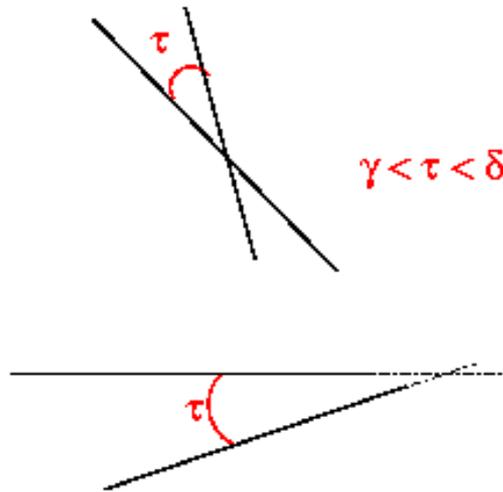
*Very slight difference  
of orientation*



⇒ **Clique forbidden**

$$g(u,v)=\infty$$

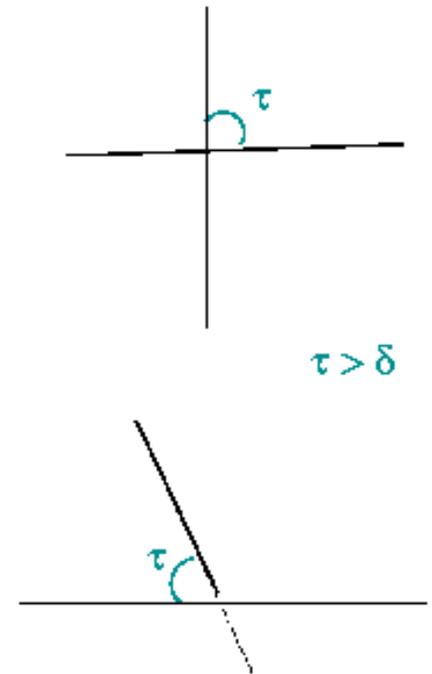
*Slight difference of orientation*



⇒ **Clique penalised**

$$g(u,v)>0$$

*Difference of orientation  
close to a right angle*

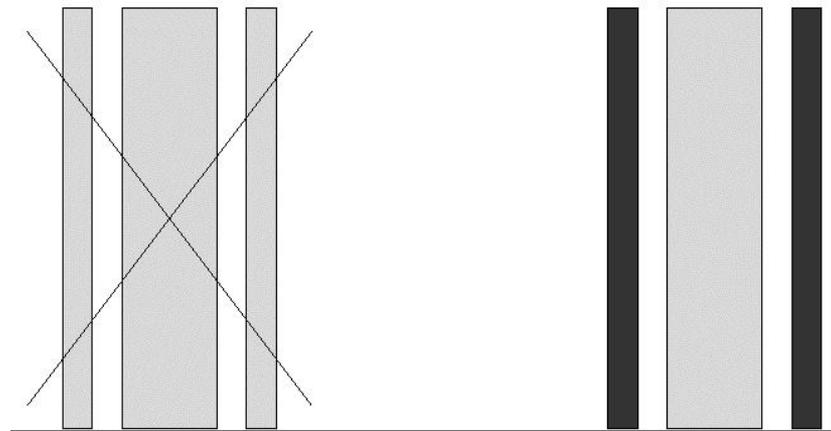


⇒ **Clique not penalised**

$$g(u,v)=0$$

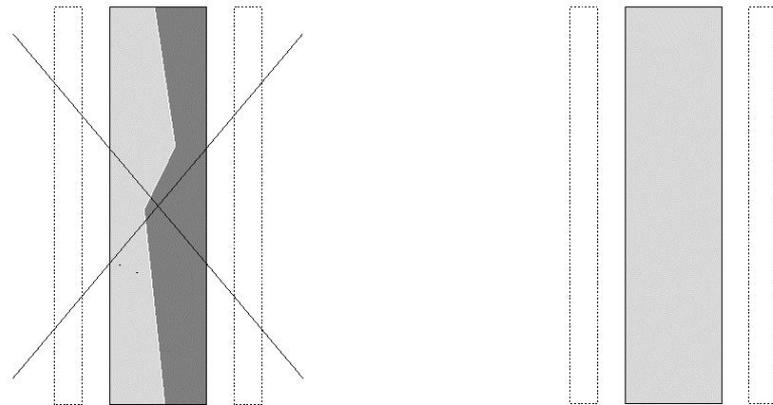
# First example : Quality Candy Model for road network extraction

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



# First example : Quality Candy Model for road network extraction

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



# 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

# Results



# Results



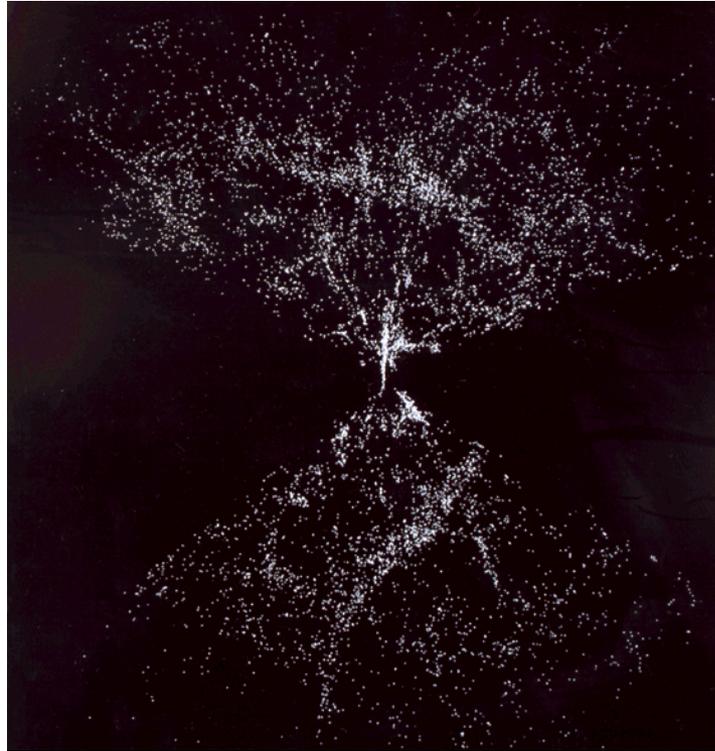
# Results



# Results



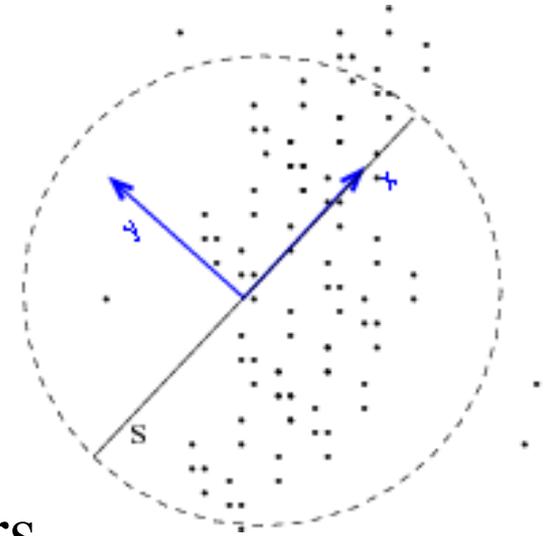
# Galaxy Filament Detection



© Center for Astrophysics (Harvard University)

# Galaxy Filament Detection

- Prior : Quality Candy model
- Assumptions for the data term :
  - Segments live in dense areas
  - Segments live in the center of clouds
  - Segments live in elongated clusters





© Center for Astrophysics (Harvard University)



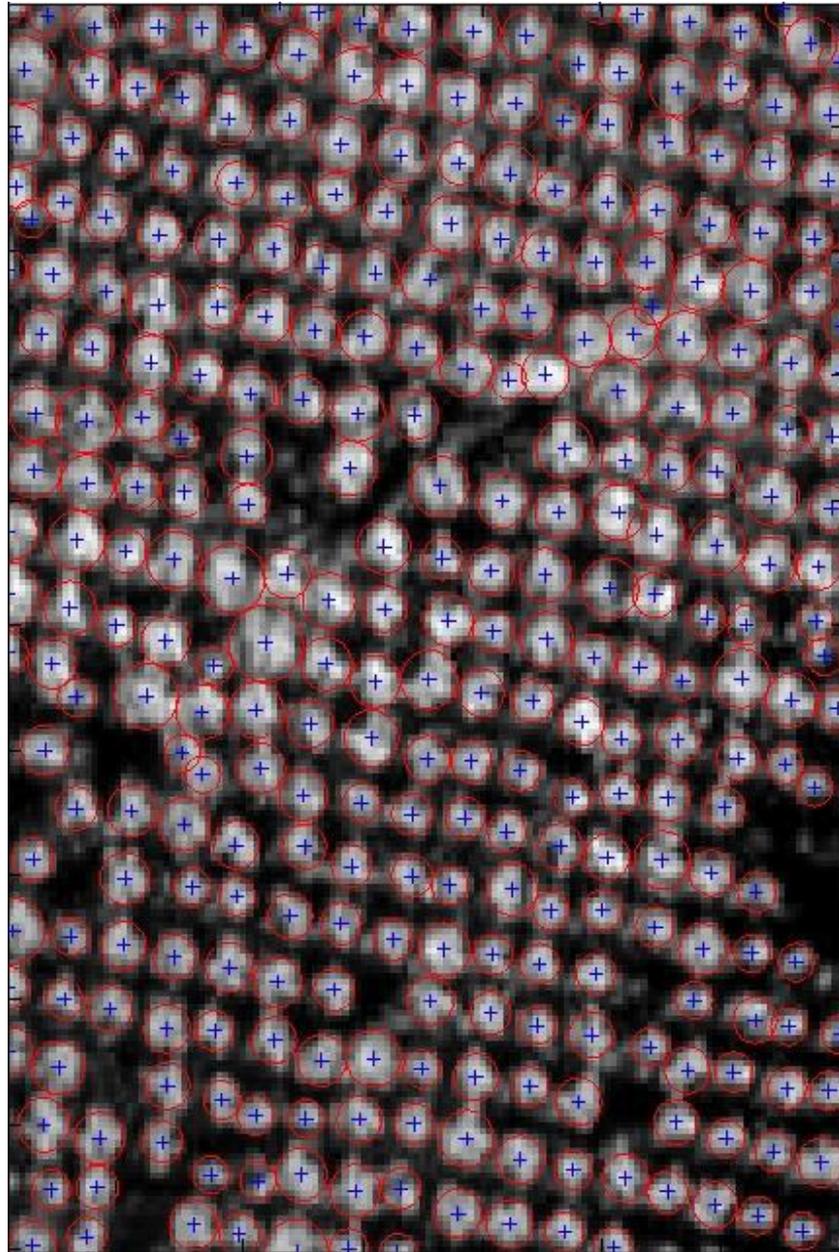
Results



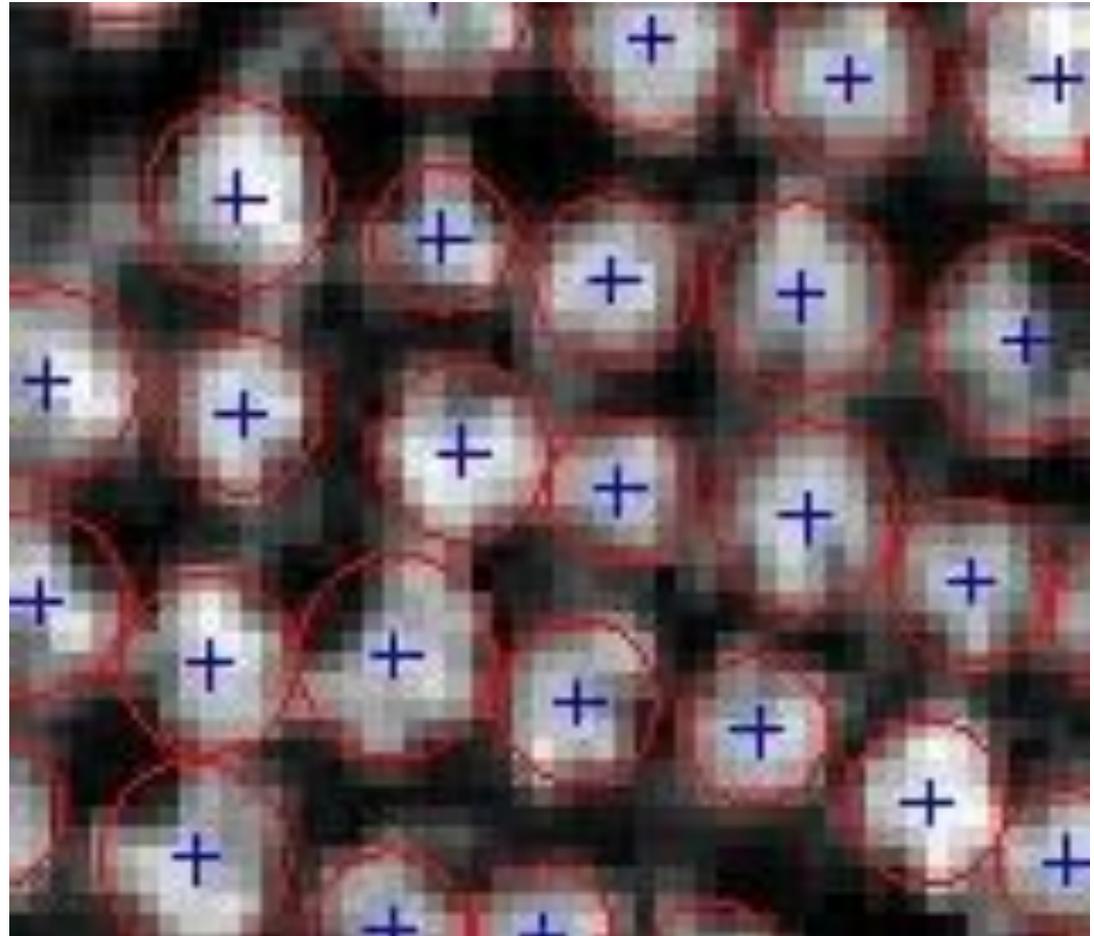
# Second example : tree crown extraction

- Object : disk  $r \in [r_{\min}, r_{\max}]$
- Prior : non-overlapping  $d_i \propto_{\rho} d_j \Leftrightarrow \|c_i c_j\| < r_i + r_j$   
 $N_{\propto_{\rho}}(d_i) = \{d_j \neq d_i : d_i \propto_{\rho} d_j\}$   
 $t_{\rho}(d_i) = \gamma^{A(S(d_i) \cap S(N_{\propto_{\rho}}(d_i)))}$  A : area
- Data : Gaussian likelihood  
 $A_y(S(x)) = \prod_{p \in S(x)} P_{tree}(y_p) \prod_{p \notin S(x)} P_{notree}(y_p)$

# Result

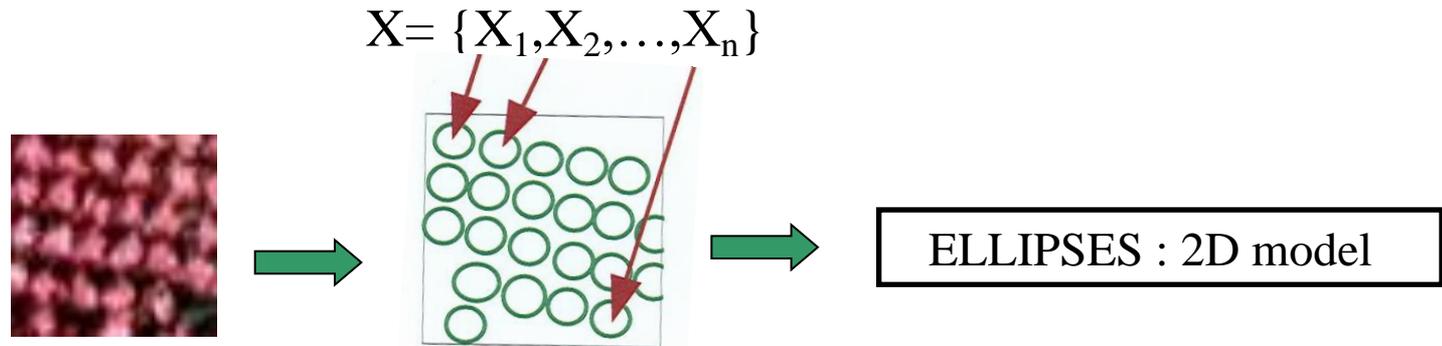


# Result

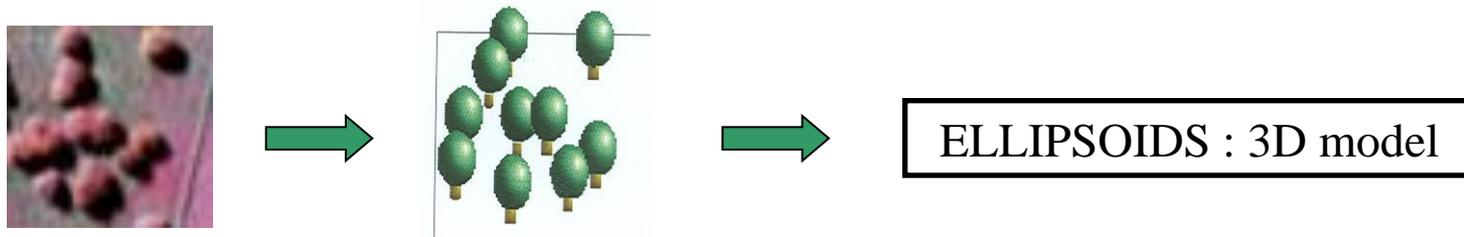


# Proposed method

- Marked point processes : find an unknown number of geometric objects (ellipses or ellipsoids) whose positions and sizes are unknown.
- Find the best configuration of objects :



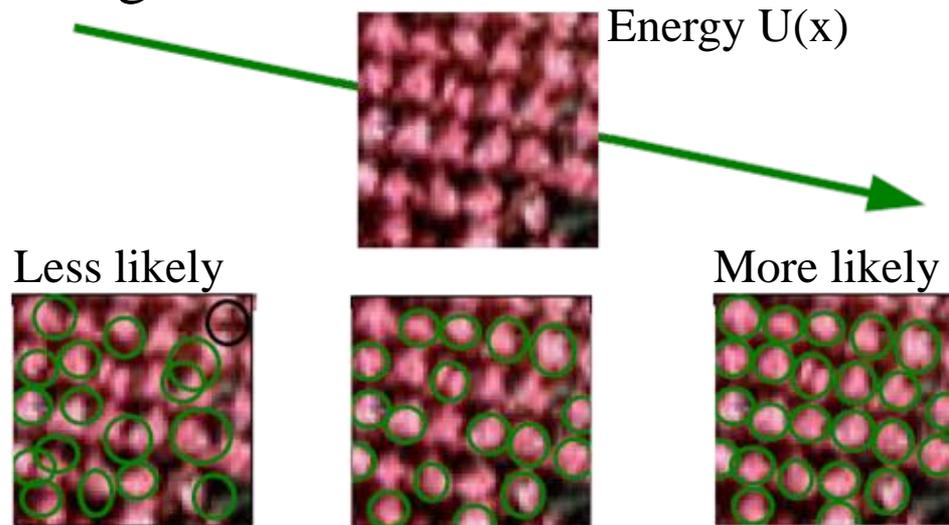
Dense area : plantation (merged shadows)



Sparse vegetation (drop shadows)

# Density of the process

- Goal : design the density of the mpp in order to make tree configurations be the most likely configurations.
- Minimise the energy :  $U(x) : f(x) = \frac{1}{Z} \exp(-U(x))$
- Mathematical tools : Markov Chain Monte Carlo algorithms + simulated annealing.



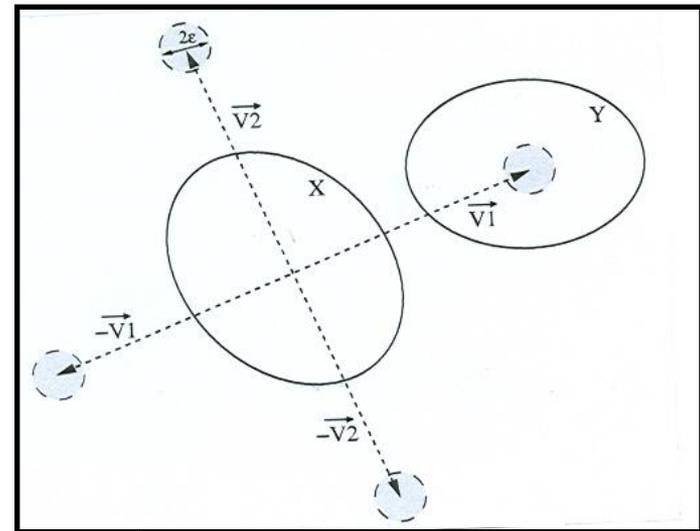
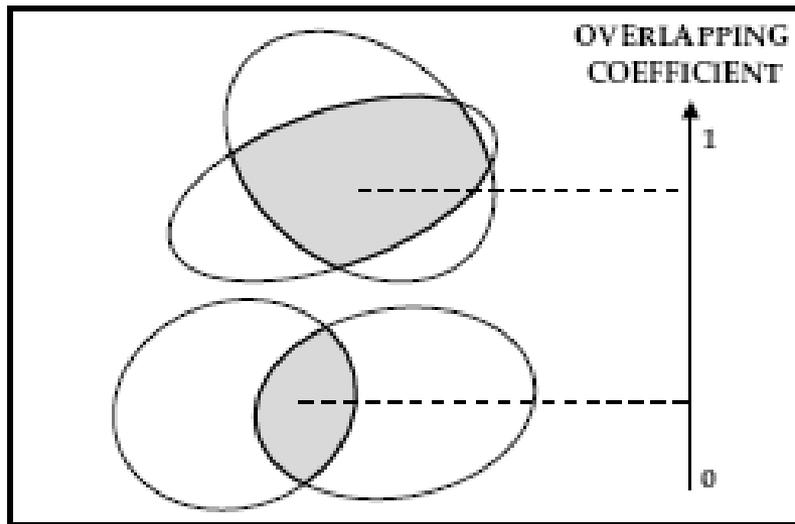
Poplars to be extracted with ellipses

# Energy of the model

- Regularizing term + Data term :

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

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

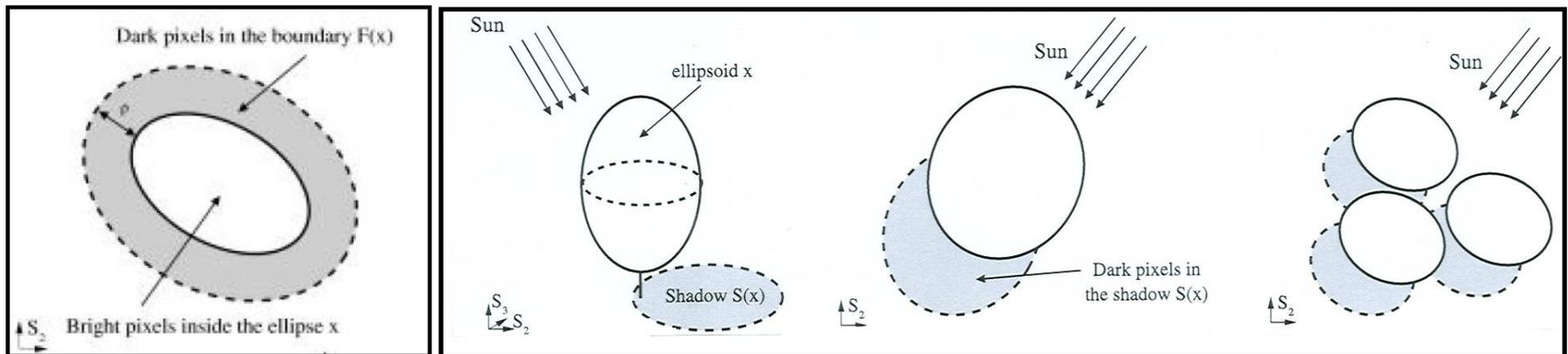


- $U_d(x)$  : data term = fitting the object into the image.

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

# Data energy term $U_d(x)$

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



# Results with the 2D model (1)

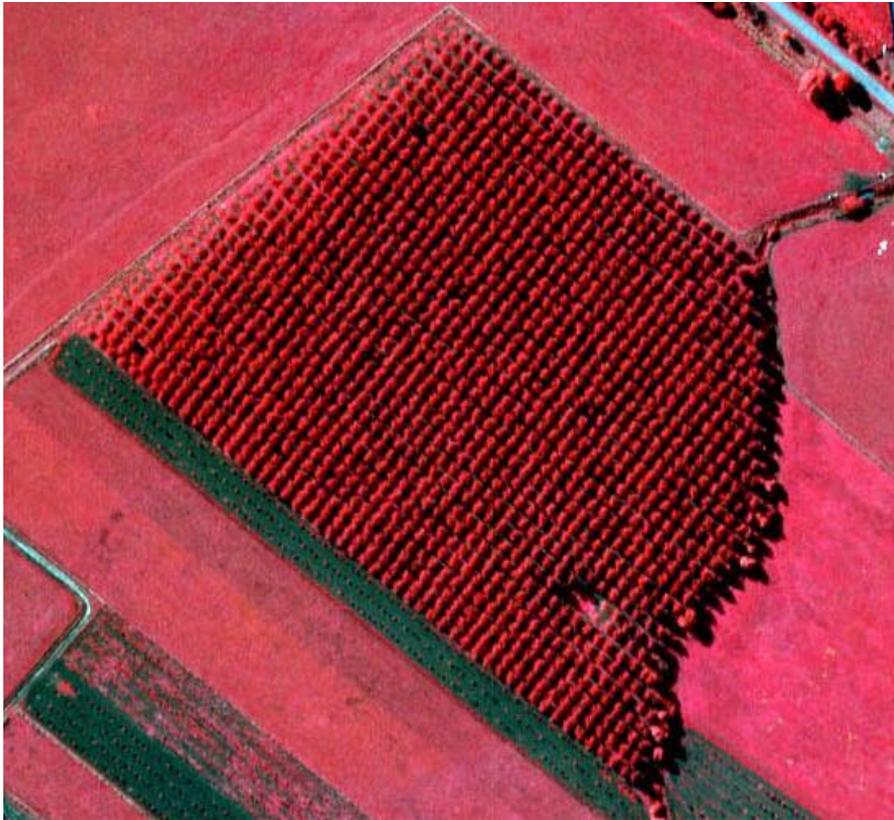


Poplar plantation. 1 ha ©IFN (now IGN).

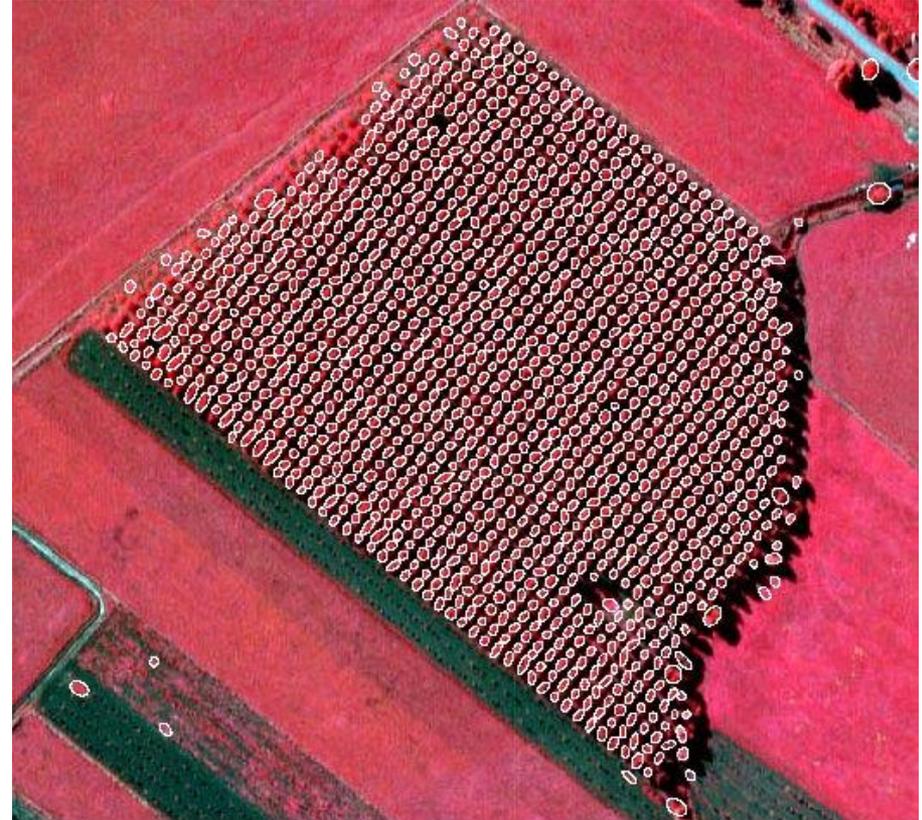


2D model extraction. © Ariana / INRIA

# Results with the 2D model (2)



Poplar plantation. 7 ha ©IFN (now IGN)



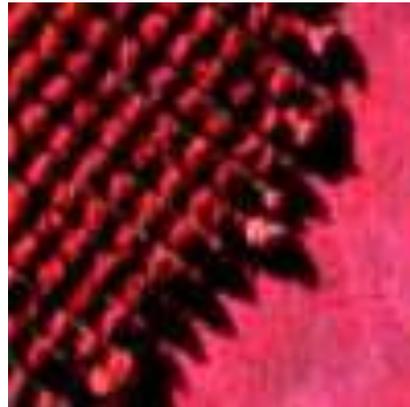
2D model : more than 1300 objects. © Ariana / INRIA

# 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 (now IGN)



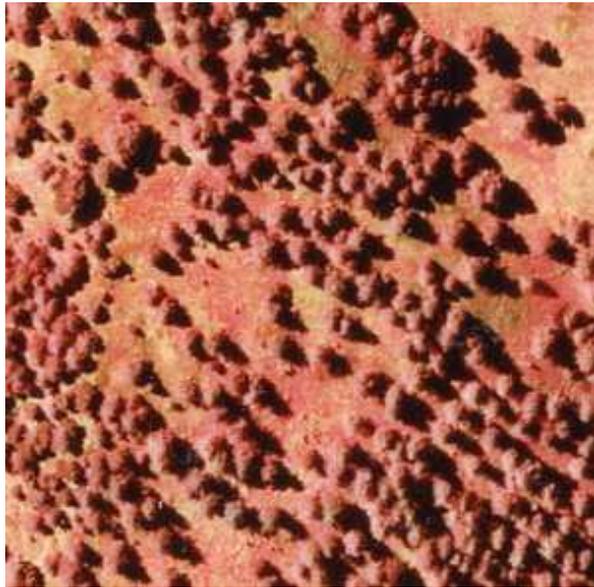
© IFN (now IGN)



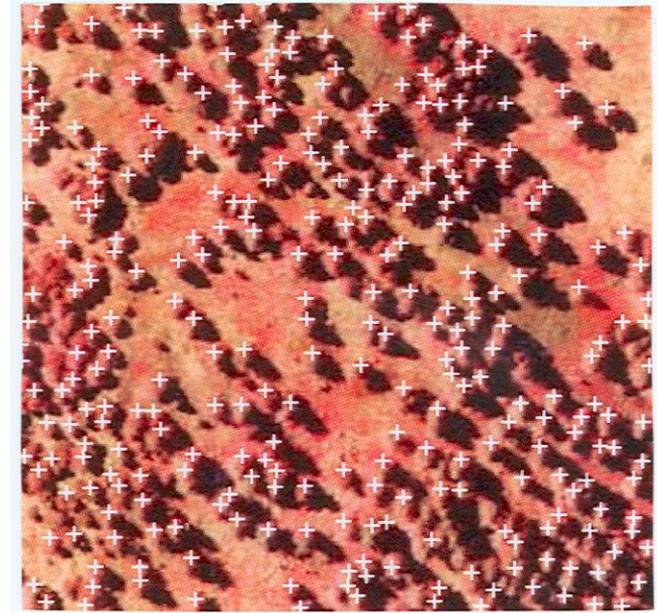
© IFN (now IGN)

# Results with the 3D model (2)

- 3D model extraction in **sparse vegetation**.



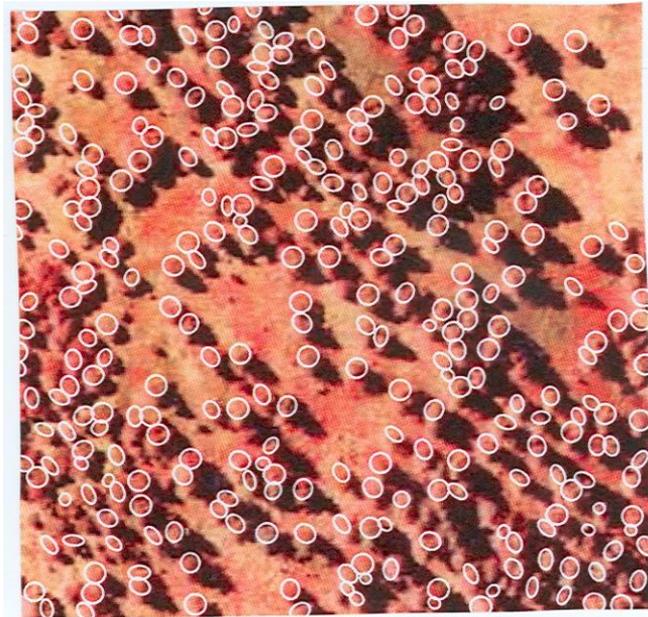
2.5 ha (Alpes Maritimes)  
© IFN (now IGN).



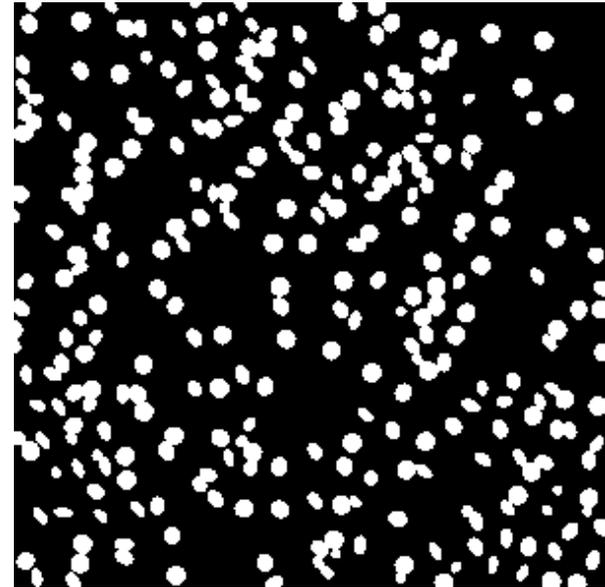
3D model extraction © Ariana / INRIA

# Results with the 3D model (3)

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



3D model extraction. © Ariana / INRIA



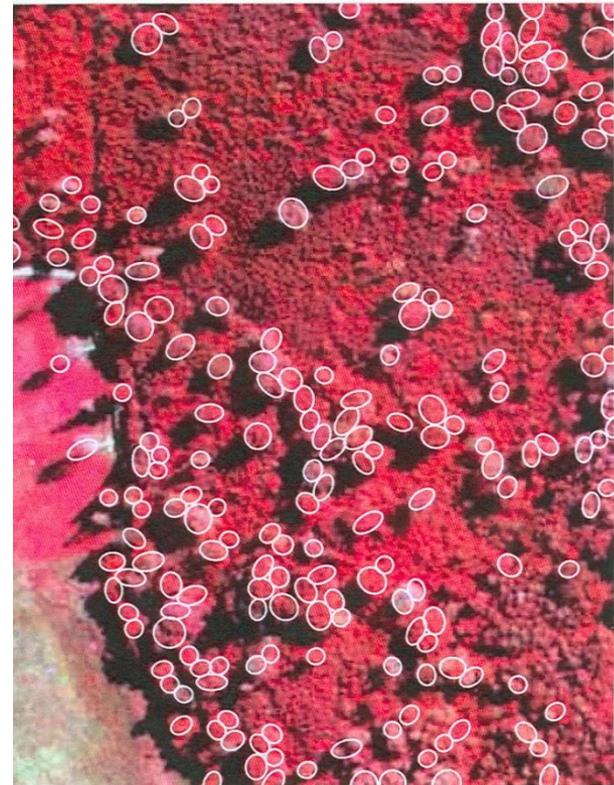
Binary image of the vegetation.

# Results with the 3D model (4)

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



Mixed height stand (3 ha) © IFN (now IGN).



3D model extraction © Ariana / INRIA

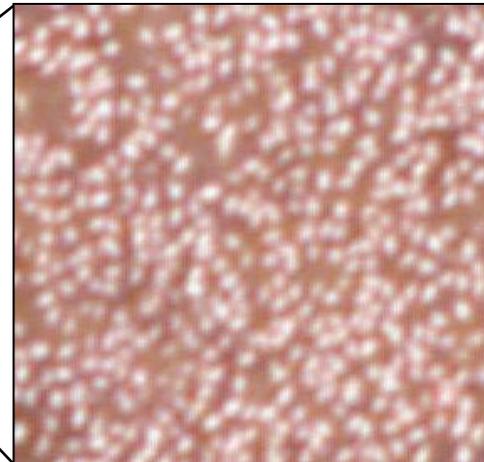
# Third example : Counting Flamingos

## Previous techniques for counting flamingos :

- ✓ Manually, by sampling, counting then extrapolating
- ✓ Tricky because of the low quality of the aerial images
- ✓ Time consuming and in the end not accurate



© Tour du Valat



© Tour du Valat

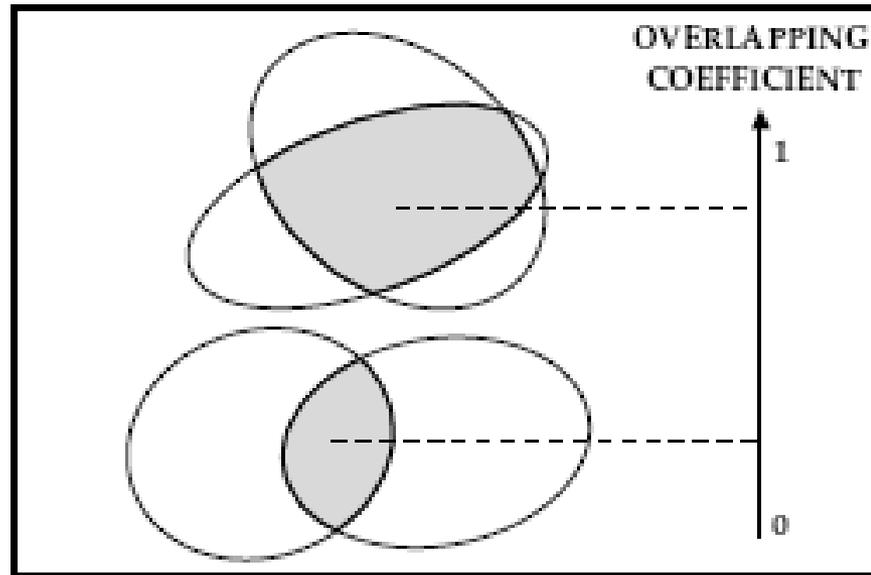
➡ **Need to develop a method to count flamingos automatically**

(free software available with CeCILL C licence at [http://www.flamingoatlas.org/dwld\\_flamingo.php](http://www.flamingoatlas.org/dwld_flamingo.php))

# Model for the extraction of flamingos

A priori model : Interaction between objects [G.Perrin et al., 06]

➔ Penalization of overlapped objects



*Top: high energy, Bottom: low energy*

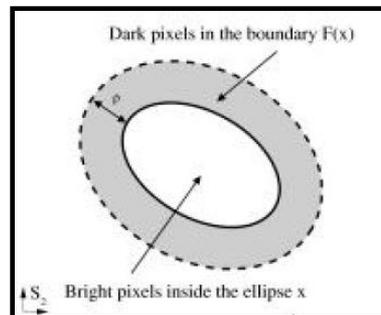
# Model for the extraction of flamingos

## Data model : to adapt objects to flamingos

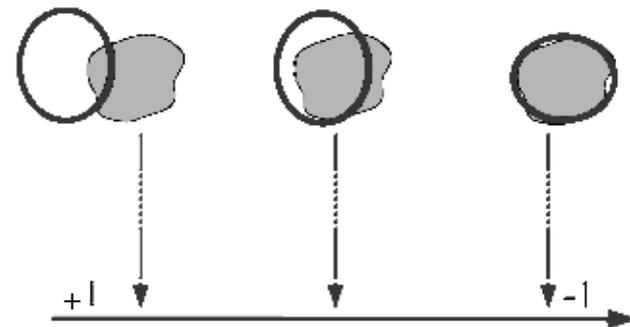
Flamingos considered as bright ellipses making a contrast with their crowns.

- ✓ Bhattacharya distance computation from the pixel distributions in the ellipse and in its crown [G.Perrin et al., 06].
- ✓ Comparison of the center of the ellipse with the mean value of a flamingo in the image. [S. Descamps et al., 08]

We favor good objects, we penalize bad ones (automatic computation of the limit  $L$ ):



*Ellipse and its crown*



*Different levels of energy*

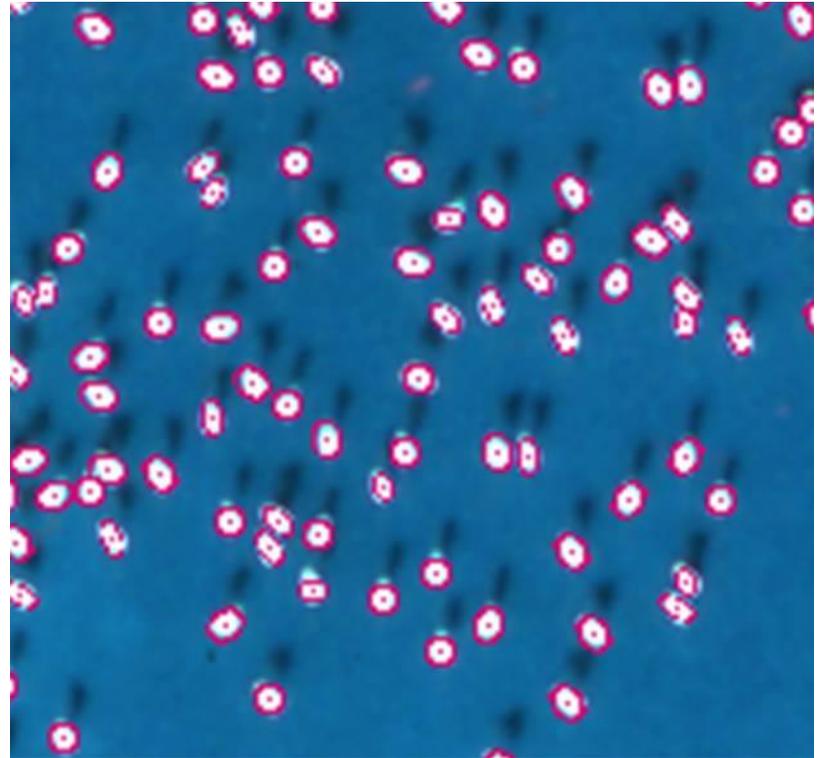
# Results

## Estimation of size of a French colony:

✓ 560 flamingos detected



© IFN (now IGN)

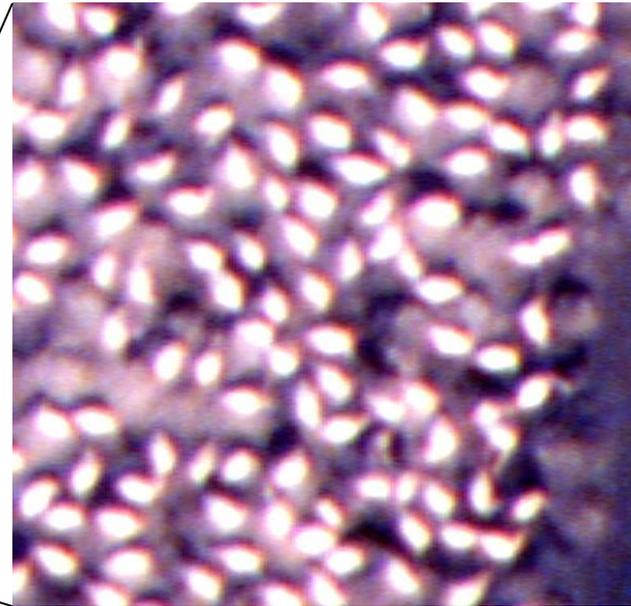
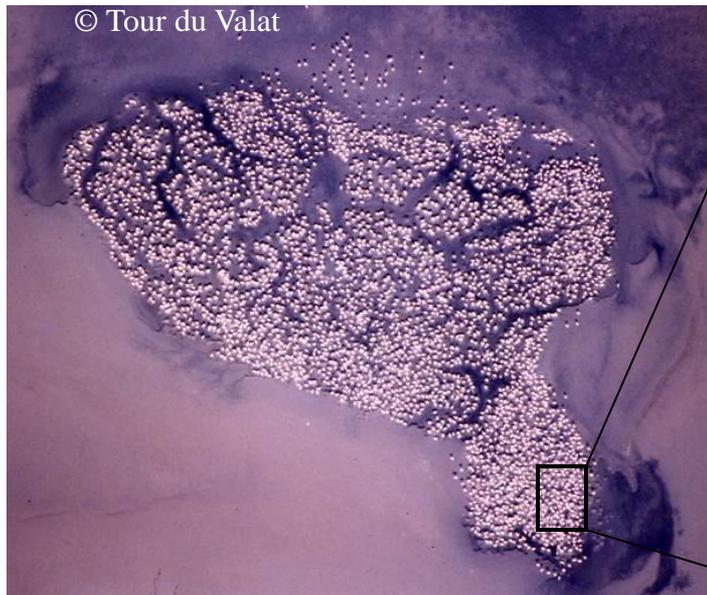


© Ariana / INRIA

# Results

## Estimation of size of a Turkish colony (2005):

✓ Low density

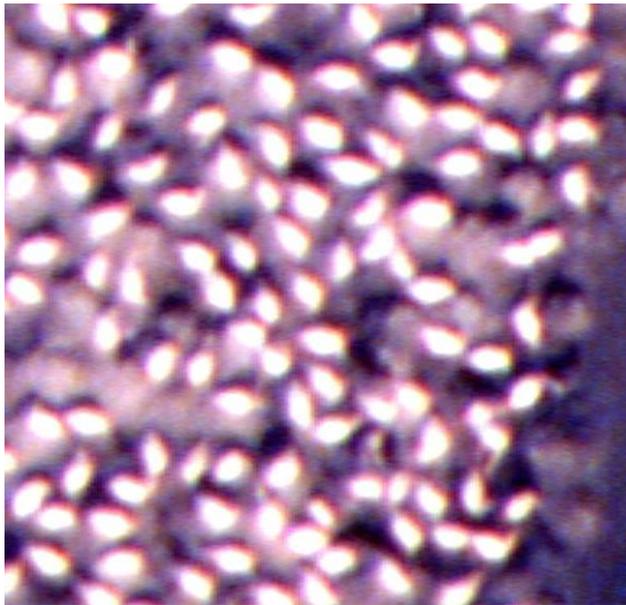


© Tour du Valat

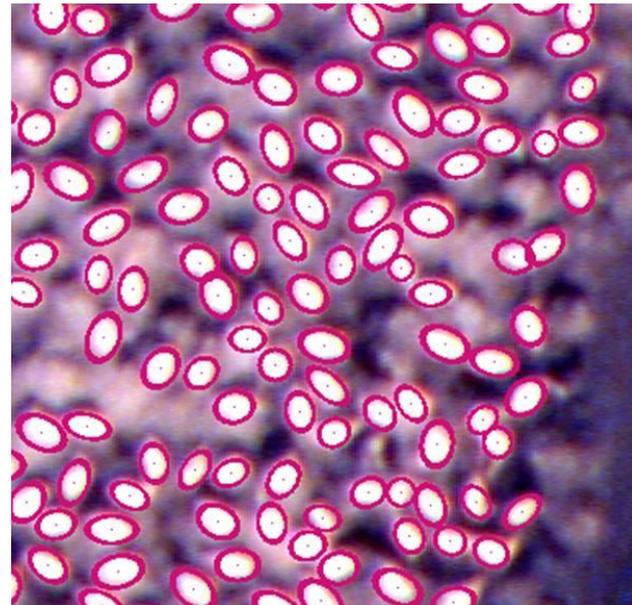
# Results

## Estimation of size of a Turkish colony (2005):

- ✓ Automatic: 3680 flamingos;                      Manually: 3682 flamingos
- ✓ Computation time: 80 minutes
- ✓ Image size: 6080 x 4128



© Tour du Valat

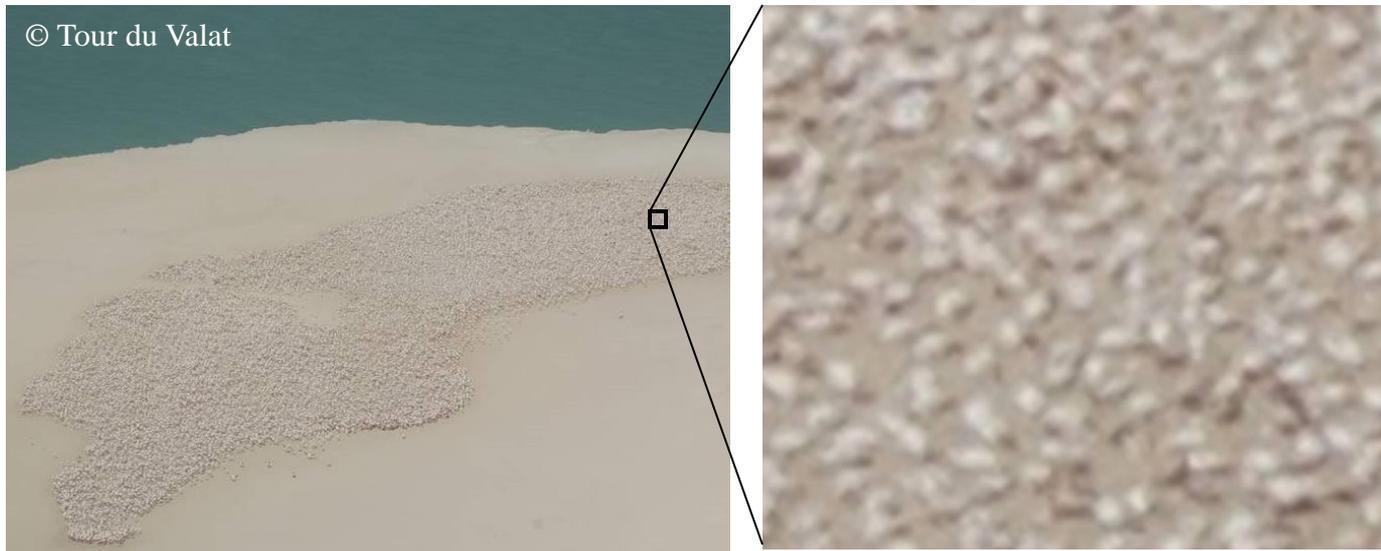


© Ariana / INRIA

# Results

## Estimation of size of an African colony (2004):

- ✓ High density



© Tour du Valat

# Results

## Estimation of size of an African colony (2004):

- ✓ Automatic: 14300 flamingos
- ✓ Manually: 13650 flamingos
- ✓ Computation time: 15 minutes
- ✓ Image size: 3008 x 2000



© Ariana / INRIA

# Fourth example : boat detection

Harbor management is a difficult problem due to :

- ✓ Big increase in the recreational fleet
- ✓ Reduced anchoring space



© CNES



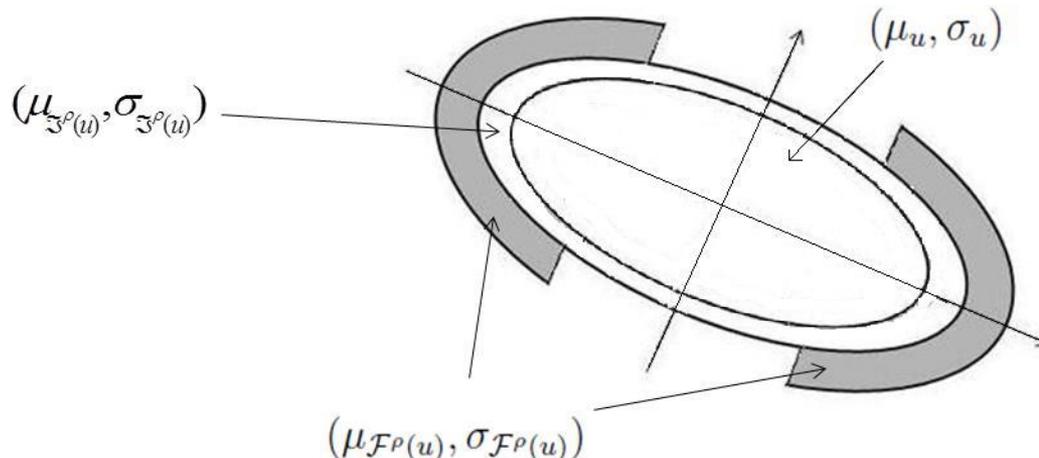
© CNES

# Model for the detection of boats

A new data term : to adapt the objects to boats

Boats tend to have a small dark area in the middle

- ✓ Distance similar to Bhattacharyya distance between interior of the object and the surrounding border [*G. Perrin et al. 06*]
- ✓ Distance similar to Bhattacharyya distance between inner border of the object and the surrounding border [*P. Craciun et al. 13*]

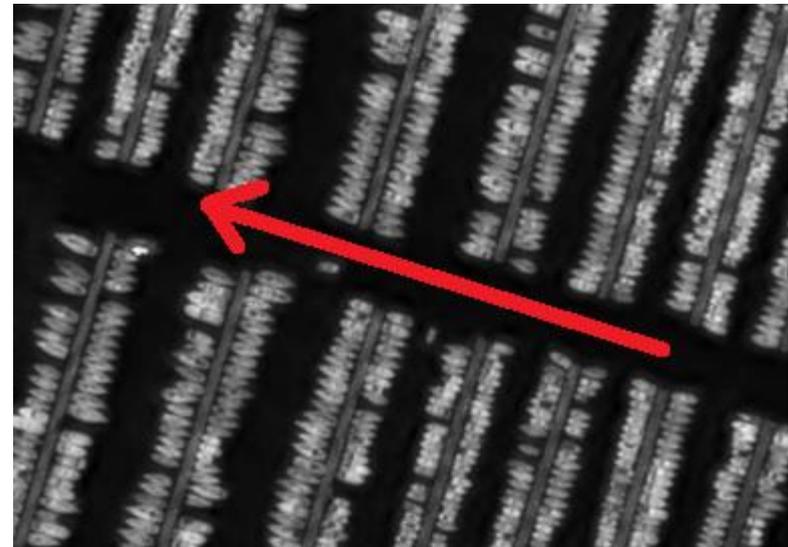
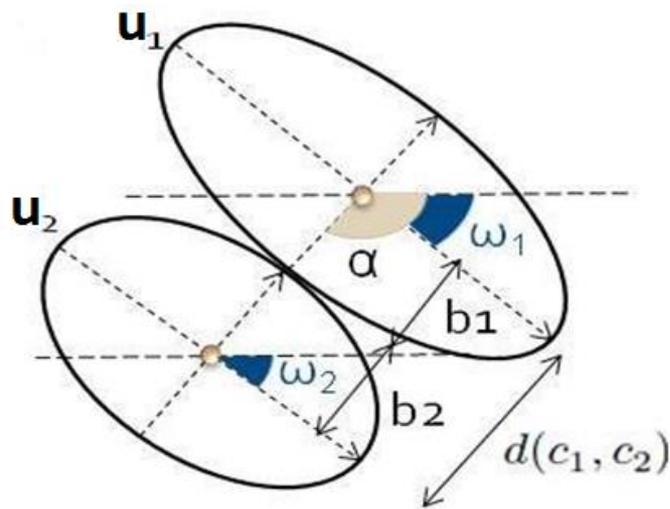


# Model for the detection of boats

A first prior model : to adapt the interactions between objects

Boats in harbors are aligned and close to each other

- ✓ Favor close and aligned objects which have the same global orientation [S. Ben Hadj et al. 10]



# Model for the detection of boats

A second prior model : to handle multiple orientations of objects

Docks in harbor can have multiple orientations and thus, boats too

- ✓ Favor close and aligned objects
- ✓ Pre-compute orientation of water area [*P. Craciun et al. 13*]

Boats are perpendicular to the local orientation of the water

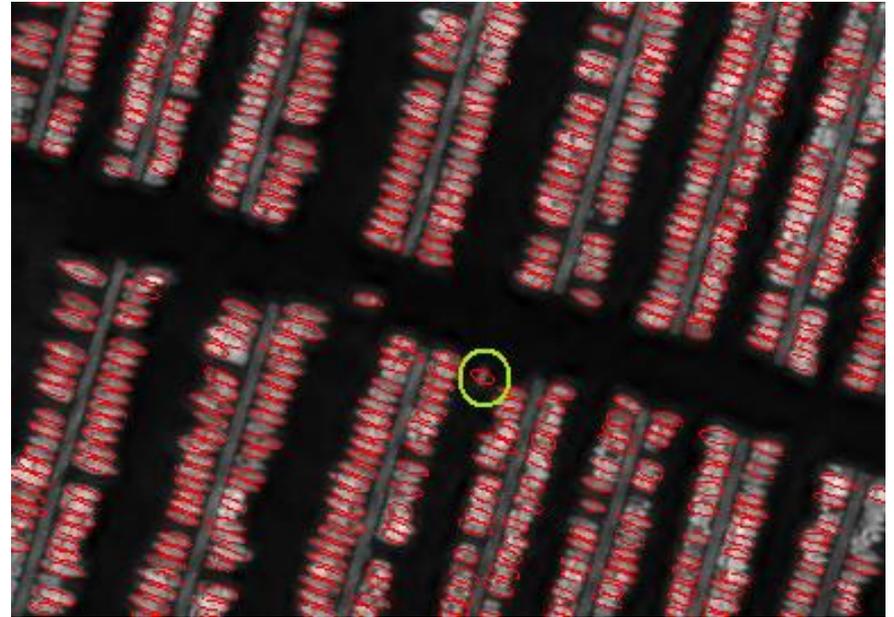


Local orientation of the water © Ayin / INRIA

# Results

## Estimation of recreational fleet in harbor (Ben Hadj's model):

- ✓ Automatic: 518 boats
- ✓ Manually: 523 boats
- ✓ Computation time: 55 minutes
- ✓ Image size: 375 x 285

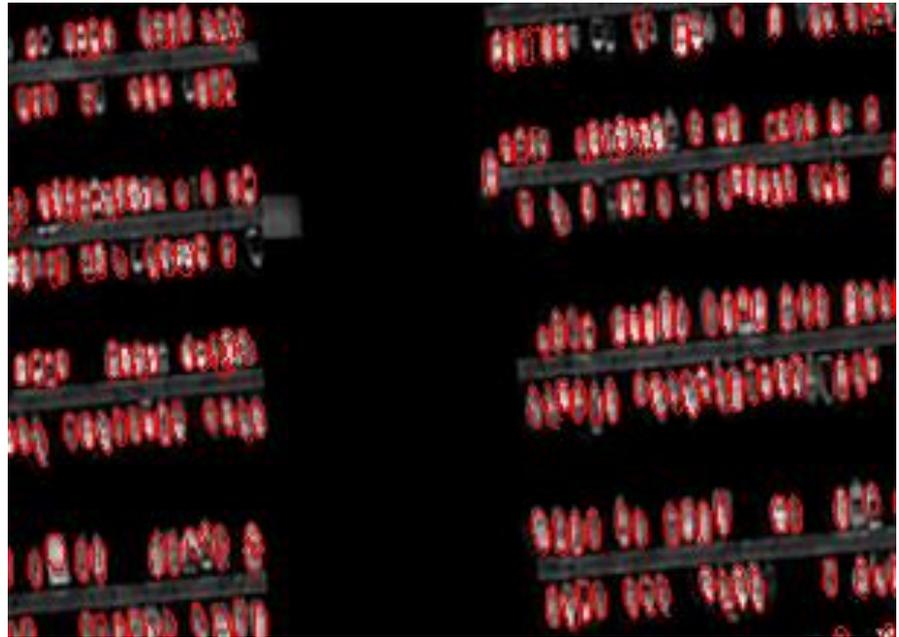


© Ayin / INRIA

# Results

## Estimation of recreational fleet in harbor (Ben Hadj's model):

- ✓ Automatic: 233 boats
- ✓ Manually: 234 boats
- ✓ Computation time: 32 minutes
- ✓ Image size: 304 x 220



© Ayin / INRIA

# Results

## Estimation of recreational fleet in harbor (Craciun's model):

- ✓ Automatic: 168 boats
- ✓ Manually: 190 boats
- ✓ Computation time: 36 minutes
- ✓ Image size: 326 x 226



© Ayin / INRIA

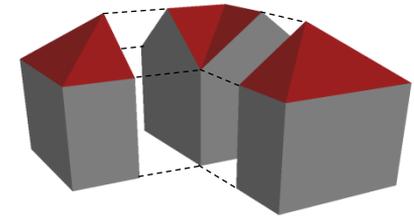
# Fifth example : building extraction

## Context

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

## Toward structural modeling

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

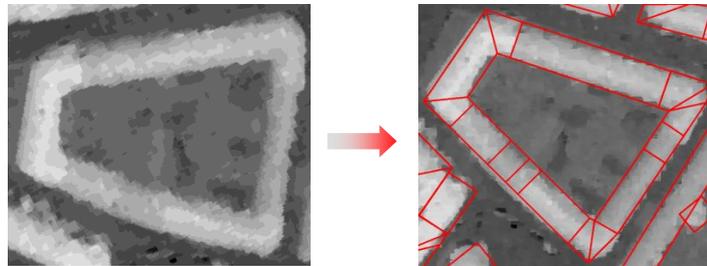


A building = an assembly of simple urban structures

## 2 stages : 2D extraction, then 3D reconstruction

- computation is greatly reduced

# Stage 1: 2D extraction of buildings



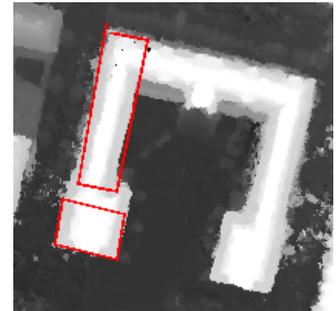
# 2D extraction of buildings

Outlines of buildings by marked point processes [Ortner04]

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

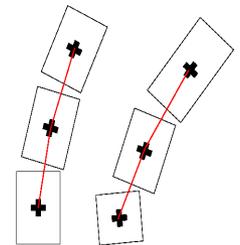
▶  $U_{ext}$  data term

⊕ coherence between the location of a rectangle and discontinuities in the DEM



▶  $U_{int}$  regularizing term

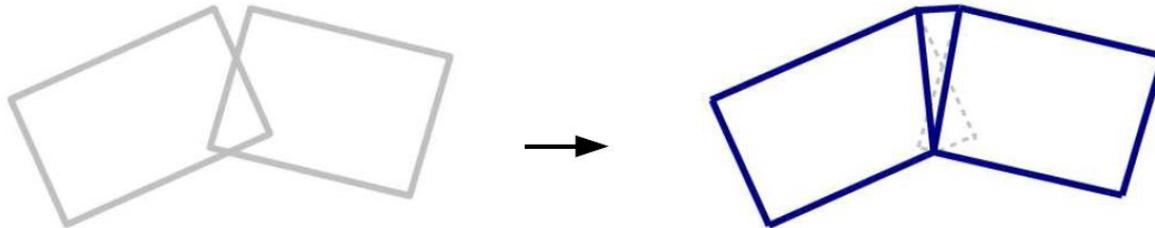
⊕ introduction of prior knowledge about the object layout (alignment, paving, completion)



# 2D extraction of buildings

Transformation of rectangles into structural supports

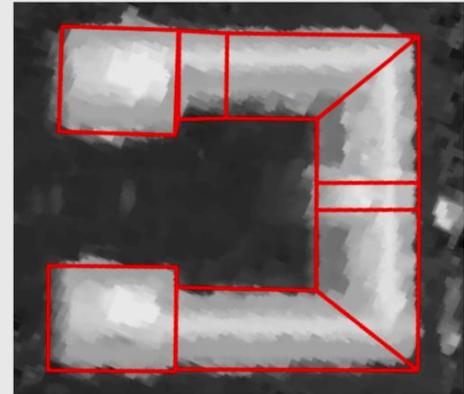
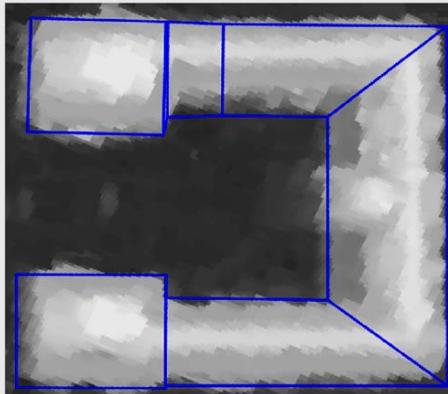
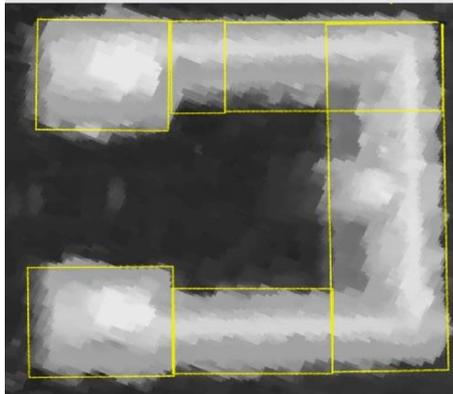
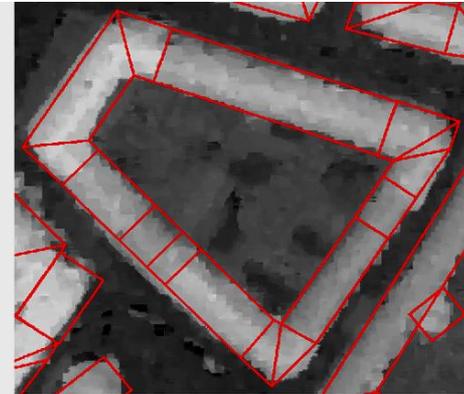
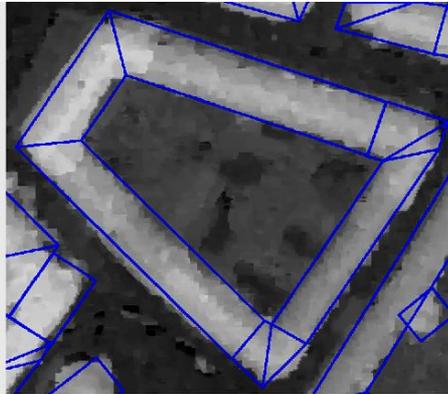
- transformation of rectangles into **unspecified quadrilaterals which are ideally connected** (without overlapping, with a common edge)



- partitioning of rectangles that represent different urban structures

# 2D extraction of buildings

Examples © Ariana / INRIA

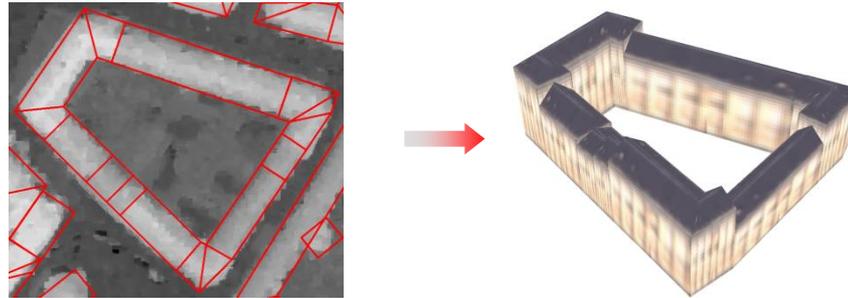


rectangular supports  
by [Ortner04]

"connected" supports  
by [Lafarge07]

structural supports  
by [Lafarge07]

## Stage 2: 3D reconstruction of buildings

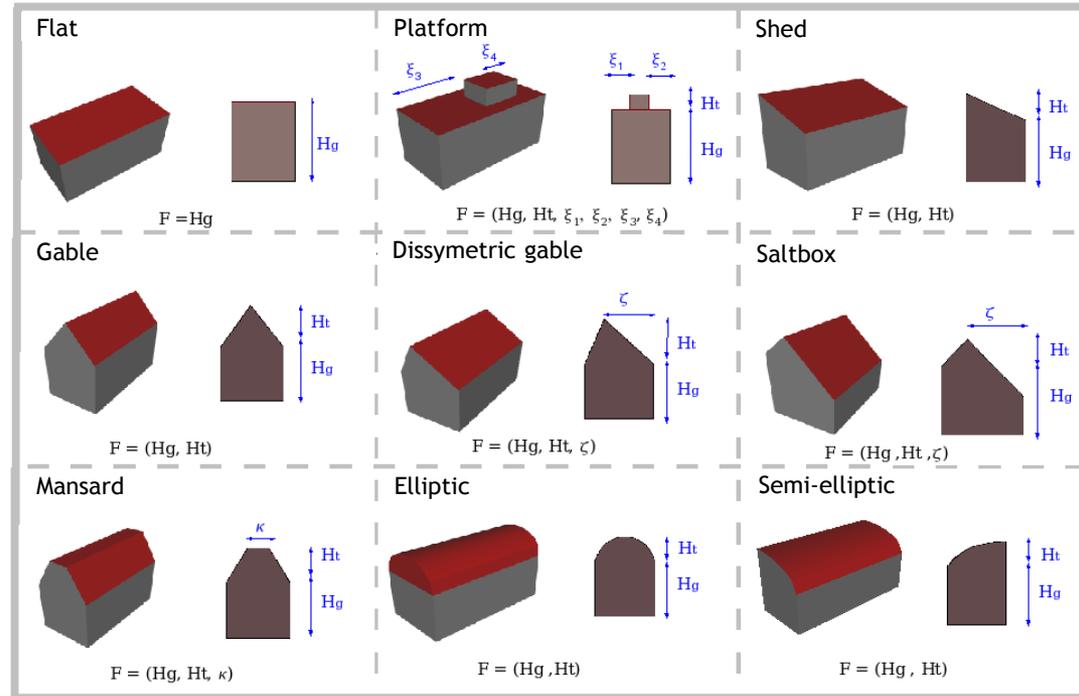


# 3D reconstruction of buildings

## Library of 3D models

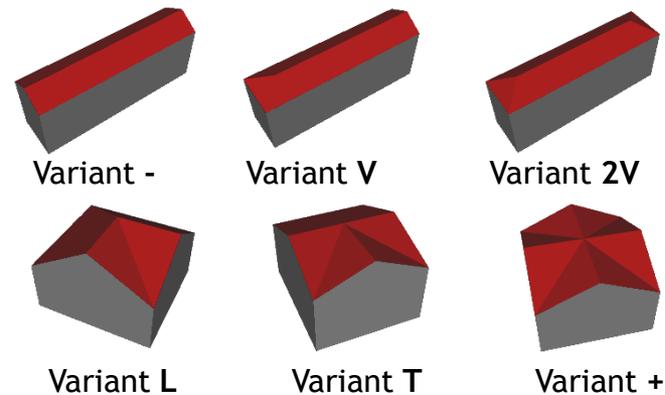
### The roof shapes :

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

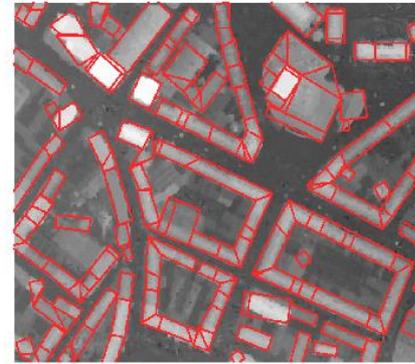
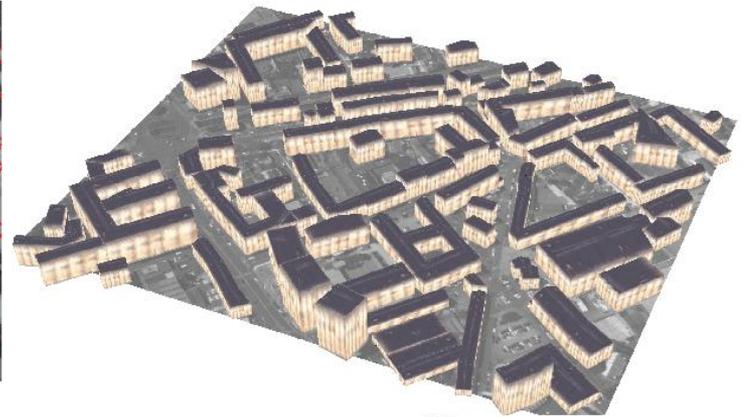


### The variants :

- ends and junctions
- orientation of the object



# Results



PLEIADES simulations  
©CNES

Building Extraction  
© Ariana / INRIA

3D Reconstruction  
© INRIA / IGN / CNES

# Conclusion

- The marked point process framework extends the application domain of Markov 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)

# Current extensions

- Point process with marks living in a shape space
  - Multiple object detection [F. Lafarge et al.]
  - Computational issues [Y. Verdie et al., P. Craciun et al.]
- New optimization dynamics
  - Diffusion processes [X. Descombes et al.]
  - Multiple Birth and Cut [A. Gamal et al.]
- Different applications
  - Vascular network detection in the brain in 3D [X. Descombes et al.]
  - Cell counting [X. Descombes et al.]
  - Pimple detection in dermatology [J. Zerubia et al.]

# Future work

- Point process with marks living in a shape space
  - More accurate definition of the geometry
- Parameter estimation
  - Quasi-MCMC techniques
  - Genetic algorithms
- New applications
  - Object tracking in image sequences

# Thanks to :

- CNES, IGN, Airbus D&S, Satellite Image Corporation, Tour du Valat and Harvard University, for providing the data.
- Airbus D&S, DGA, MAE, CNES, IGN, BRGM, ECP and INRIA, for partial financial support.

# For more information :

- Previous team:

<http://www.inria.fr/ariana>

- Current teams:

<http://team.inria.fr/ayin>

<http://www-sop.inria.fr/morpheme/>

<http://team.inria.fr/titane>

# References

- P. Green.** Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. **In *Biometrika* (82), 1995.**
- R. Stoica, X. Descombes, J. Zerubia.** A Gibbs point process for road extraction in remotely sensed images. **In *Int. J. Comput. Vis.* (57), 2004.**
- C. Lacoste, X. Descombes, J. Zerubia.** Point processes for unsupervised line network extraction in remote sensing. **In *IEEE Trans. PAMI* (27), 2005.**
- G. Perrin, X. Descombes, J. Zerubia.** 2D and 3D Vegetation Resource Parameters Assessment using Marked Point Processes. **In *Proc. ICPR*, 2006.**
- M. Ortner, X. Descombes, J. Zerubia.** Building outline extraction from digital elevation models using marked point processes. **In *Int. J. Comput. Vis.* (72), 2007.**
- S. Descamps, X. Descombes, A. Bèchet, J. Zerubia.** Automatic flamingo detection using multiple birth and death process. **In *Proc. ICASSP*, 2008.**
- S. Ben Hadj, F. Chatelain, X. Descombes, J. Zerubia.** Parameter estimation for marked point process within the framework of multidimensional shape extraction from remote sensing images. **In *Proc. ISPRS XXXVIII*, 2010.**
- F. Lafarge, G. Gimel'farb, X. Descombes.** Geometric feature extraction by a multi-marked point process.. **IEEE Trans. PAMI (32), 2010**
- A. Gamal-Eldin, X. Descombes, G. Charpiat, J. Zerubia.** Multiple birth and cut algorithm for multiple object detection. **Journal of Multimedia Processing and Technologies, 2012.**
- Y. Verdie, F. Lafarge.** Efficient Monte Carlo sampler for detecting parametric objects in large scenes. **In *Proc. ECCV*, 2012.**
- P. Craciun, J. Zerubia.** Unsupervised marked point process model for boat extraction in harbors from high resolution optical remotely sensed images. **In *Proc. ICIP*, 2013.**