Belief functions for image analysis and processing

S. Le Hégarat-Mascle

University Paris-Sud (France)

Introduction to image processing

Image definition :

- Bidimentional signal with finite support and bounded values $\{y(i,j), i \in [1,N], j \in [1,M]\}$ values may be
  - binary $\Rightarrow y(i,j) \in \{0,1\}$,
  - grey level $\Rightarrow y(i,j) \in [y_{min}, y_{max}]$ (generally [0,255]),
  - RGB $\Rightarrow y(i,j) = (y_{red}, y_{green}, y_{blue})$ with $y_{red} \in [y_{min}, y_{max}]$ etc.,
  - multispectral $y(i,j) = (y_{\lambda_1}, y_{\lambda_2}, \ldots, y_{\lambda_n})$ with λ_i the i^{th} wavelength and $y_{\lambda_i} \in [y_{min}, y_{max}]$,
  - multitemporal $y(i,j) = (y_{t_1}, y_{t_2}, \ldots, y_{t_n})$ with t_i the i^{th} time sample and $x_{t_i} \in [y_{min}, y_{max}]$.
- Stochastic process $\{y(s), s \in [1, N \times M]\}$, Random vector $(y(1), \ldots y(N \times M))$
- Surface $(i,j,x(i,j)) \in [1,N] \times [1,M] \times \Re$

Some problems of image processing

'Cartography' → image should be interpreted as a map of …
depth values (based on stereo-image pair),
soil surface in remote sensing observation (Earth…),
organs and tissues in medical imaging (given single-slice),
... label in classification...

⇒ to decide what is the unknown value in every pixel, this estimation is for an unknown parameter or for a label

'Detection/identification' → image should be interpreted to derive…
the objects of interest,
the current features of the objects of interest,
…
(Numerous processing provide indicators for this high level problem: edges, interest points, local image features (color, histogram…))

⇒ to detect and characterize the objects present in the image or video sequence, e.g. classification at object level, object tracking.

What can provide the belief functions to these problems? (I)

TBF main availabilities:

- able to model both uncertainty and imprecision,
- able to model ignorance,
- able to deal with the source correlation ($\rightarrow$ idempotent combination),
- able to measure the conflict between sources ($m(\emptyset)$, etc.),
- …

$\Rightarrow$ TBF is interesting for classification problems:

- When a source gives information about some classes but not about other ones, in particular when a source does not distinguish some classes, TBF allows dealing with class ambiguity,
- When the global source reliability depends on the source, TBF allows to discount the less reliable sources,
- When the class set or the class parameters have to be validated a posteriori, TBF provides several conflict measures,
- When combining different classifiers that are not independent, TBF provides idempotent rules.

What can provide the belief functions for these problems? (II)

Now there are two main features of image data

- There are a lot of pixels
- Pixels are not spatially independent

  ⇒ TBF is interesting:

  - When mage processing use several sources (data images or outputs of image processing algorithms) → to deal with the source imprecision, source combination, etc.
    - Sources are complementary (partially) in terms of class/object detection/identification: e.g. images acquired in different modalities (wavelength, polarisation…) → to deal with each source contextual ambiguities/imprecision, or local ignorance,
  - To model pixel spatial relationships → to deal with spatial imprecision & to introduce imprecise spatial information
  - For unsupervised classification → to deal with discernment frame dynamic estimation and/or to check class validity a posteriori

Outline of the presentation

Belief function for multisource classification

→ Supervised case, pixel level

From spatial imprecision to spatial information

→ Spatial imprecision introduces ambiguities at class border

→ Spatial information viewed as an independent information source

Automatic estimation of the discernment frame

→ Case of unsupervised classification

→ Case of sequential estimation at image level

Video sequences and object tracking problem

→ Data association sub-problem

Belief function for multisource classification

Supervised case, pixel level

Fusion may be performed in each pixel to remove some classification errors

Image classification problem

A data image is a realisation y of a random field $Y = \{Y_s, s \in \mathcal{S}\}$, with $\mathcal{S}$ the set of pixels (location), $|\mathcal{S}|$ is the number of pixels, and $Y_s \in \Re$ (or $\Re^d$ etc.)

$\exists$ another random field $X = \{X_s, s \in \mathcal{S}\}$, whose realisation x is hidden, $X_s \in \Omega$, $|\Omega|$ is the number of labels or classes ;

The aim of classification is to retrieve x the label field knowing the observation one y.

Different criteria: distance, ML (maximum of likelihood), MAP, MPM, etc.
Different constraints: supervised /unsupervised approach, etc.

- Blind classification:

For every $s \in \mathcal{S}$, estimation of x_s knowing y_s : $\forall s \in \mathcal{S}, x_s = \arg\max_{\omega \in \Omega} P(x_s = \omega / y_s) = \arg\max_{\omega \in \Omega} P(y_s / x_s = \omega).P(\omega)$

- Markovian models:

For every $s \in \mathcal{S}$, estimation de x_s knowing $\{y_s, s \in V_S\}$

Multisource classification

Case of *n* observation sources:

$Y = \{Y_s, s \in \mathcal{S}\}$ is now $Y^{(1)} = \{Y_s^{(1)}, s \in \mathcal{S}\}$, $Y^{(2)} = \{Y_s^{(2)}, s \in \mathcal{S}\}$, ..., $Y^{(n)} = \{Y_s^{(n)}, s \in \mathcal{S}\}$

Classical assumptions:

- The random variables $(Y_s^{(1)},...., Y_s^{(n)})_{s \in \mathcal{S}}$ are independent conditionally to X,
  $\rightarrow$ X contains all the dependencies between pixels
- For given pixel s, the distribution of $(Y_s^{(1)},...., Y_s^{(n)})$ conditional to X is equal to the distribution of $(Y_s^{(1)},...., Y_s^{(n)})$ conditional to X_s,
  $\rightarrow$ X_s gives the distribution of the observation $(Y_s^{(1)},...., Y_s^{(n)})$ in s
- For given s, the random variables $Y_s^{(1)},...., Y_s^{(n)}$ are independent conditionally to X_s,
  $\rightarrow$ can be relaxed taking into account the source correlation (e.g. using cautious conjunctive rule)

Case of **homogeneous** multi-sources : 1 label field

- data fusion interest $\rightarrow$ take advantage of source redundancy to remove classification noise
- TBF interest $\rightarrow$ model the ambiguities between some classes, of the different sources

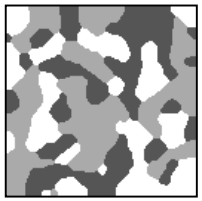
Multisource classification

supervised case, pixel level

Blind classification : dependencies between pixels are disregarded!

Assume several sources = radiometric images $Y^{(1)}$, $Y^{(2)}$, ..., $Y^{(n)}$

→ Each one has its own ability to distinguish classes :



3 classes

$\Omega=\{C_1,C_2,C_3\}$



S_1 →2 classes

$\{C_1, C_2 \equiv C_3\}$



S_2 → 2 classes

$\{C_2, C_1 \equiv C_3\}$

e.g. S_1=radar band L , S_2=radar band C, C_1=dense vegetation, C_2=bare soil, C_3=sparse vegetation

⇒ Ω is a refinement monosource discernment frames

⇒ bbas are defined in every pixel

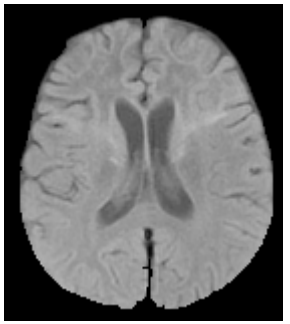
e.g. (in the ex.)

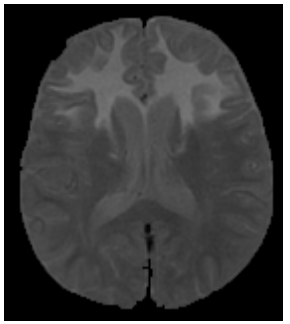
$$\begin{cases} m_1(\Theta)=1-\alpha_1 \\ m_1(C_2\cup C_3)=\alpha_1 t \\ m_1(C_1)=\alpha_1(1-t) \end{cases} \quad \begin{cases} m_2(\Theta)=1-\alpha_2 \\ m_2(C_1\cup C_3)=\alpha_2 u \\ m_2(C_2)=\alpha_2(1-u) \end{cases}$$

with $(t,u)\in[0;1]^2$

Bloch I., 1996, "Some aspects of Dempster-Shafer evidence theory for classification of multi-modality medical images taking partial volume effect into account", *Pattern Recognition Letters*, **17**(8): 905-919.

Objective: classifying brain tissues of patients suffering from adrenoleukodystrophy (ADL)
Sources = 2 dual echo MRI images (one slide)





$\Omega=\{C_1,C_2,C_3\}$ with C_1=Ventricles (V) and cerebro-spinal fluid (CSF), C_2=White Matter (WM) and Grey Matter (GM), C_3=ALD

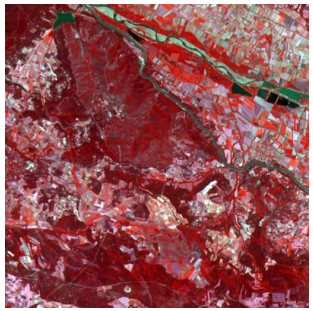
First modelling of the sources →
-for S_1, $C_2 \cup C_3$ is focal element whereas $m(C_1 \cup C_2) = m(C_1 \cup C_3) = m(\Omega)$, and
-for S_2, $C_1 \cup C_2$ is focal element but not $C_1 \cup C_3$, $C_2 \cup C_3$, Ω.

Second modelling of the sources: taking into account partial volume effect between ADL and WM →
-for S_2, $C_2 \cup C_3$ also is focal element.

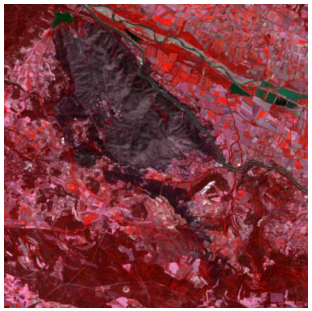
Le Hégarat-Mascle S. and Seltz R., 2004, "Automatic change detection by evidential fusion of change indices", *Remote Sensing of Environment*, **91** (3-4):390-404.

Objective: detection of the change affecting continental surfaces, e.g. forest fire damage evaluation.

July 15, 1989



Sept. 4, 1989



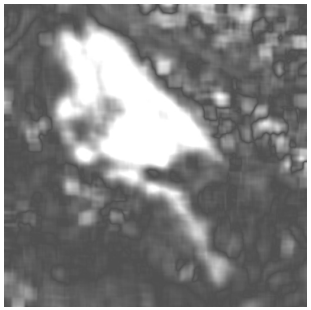
Logical sources
- the SAVI normalized (i.e. histogram matching) difference values,
- the texture (cluster shade) difference,
- the mutual information, locally to a window.

Discernment frame
Ω={Change=C, NoChange=NC}

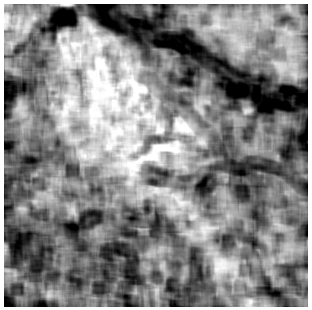
Preprocessing of the sources
→ 1 class (NC) or 2 classes (C,NC)



'normalised' SAVI difference



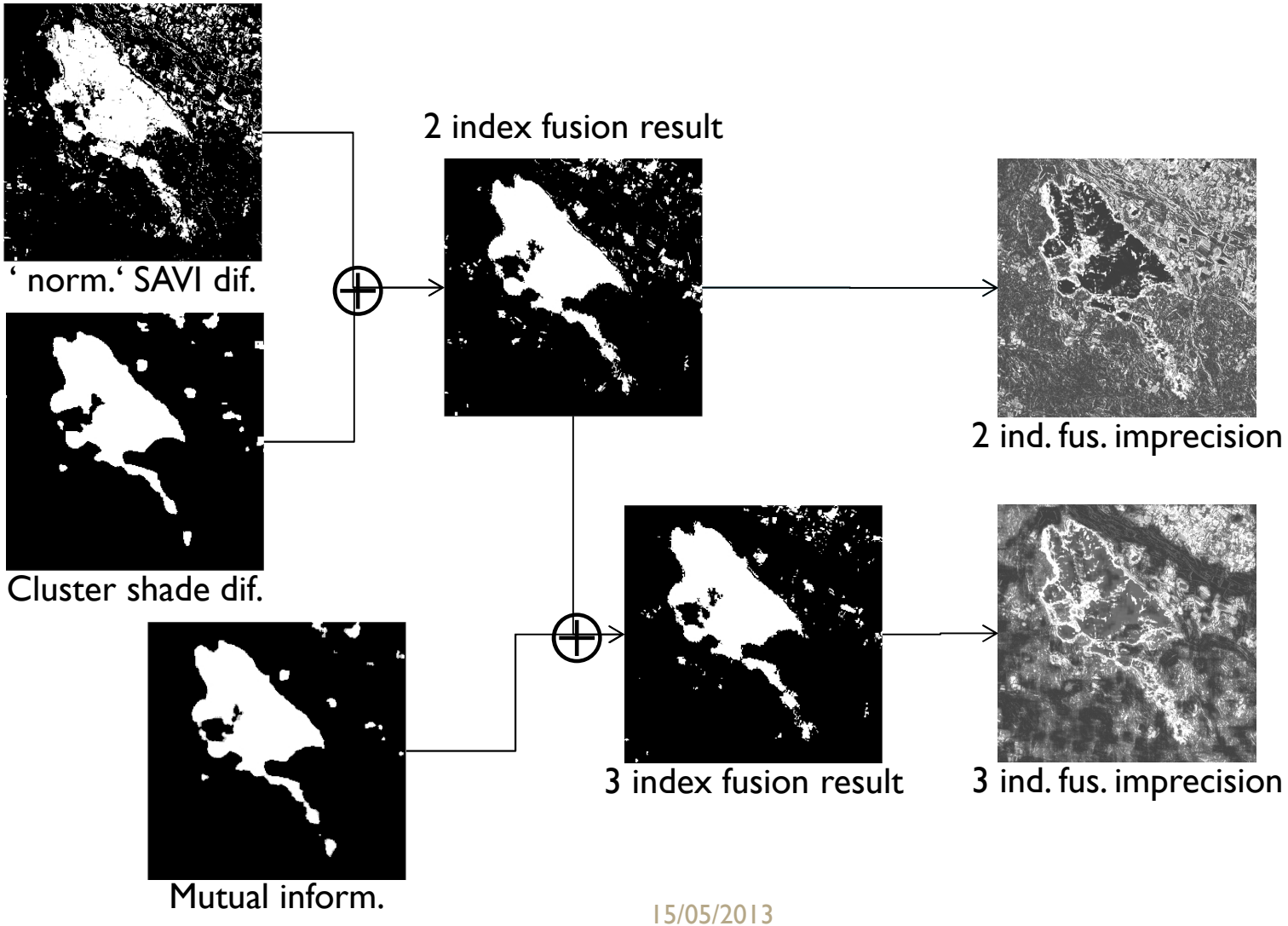
Cluster shade difference



Local mutual information

Discernment frame: $\Omega=\{C, NC\} \Rightarrow$ closed world and $\forall A \in \{C,NC\}$, $Pl(A)$-$Bel(A)=m(\Omega)$

Source modelling : $m(\Omega)$ maximum at class border when 2 classes are detected



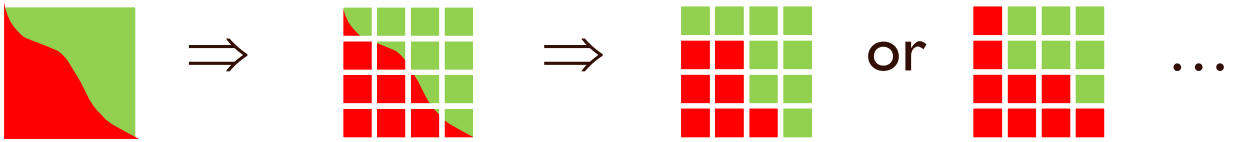
From spatial imprecision to spatial information

Spatial imprecision introduces ambiguities at class border

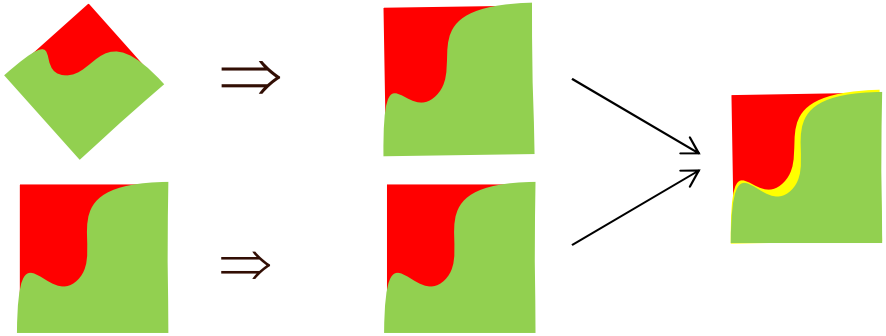
Spatial imprecision

Two main causes of spatial inaccuracy:

- Intrinsic to an image: mixing pixels

ex.:  or …

- Between two images: imprecision in spatial registration

ex. S1 

S2

Bloch I., 2008, "Defining belief functions using mathematical morphology – Application to image fusion under imprecision", *IJAR*, **48**: 437-465.

Basic idea: use the mathematical morphology operators to learn the determine the disjunction beliefs

Use a structuring element to evaluate the mass function imprecision:
Consider that bba are also as a function on the image domain $\mathcal{S}$: $\forall s \in \mathcal{S}, \forall A \in \Omega, m_s(A)(s)$

Case of a bba m_0 'without' imprecision and 'at most' 2 disjoint focal elements A, B:
$\forall s \in \mathcal{S}, m_0(B)(s) = 1 - m_0(A)(s)$
→ Imprecision can be added to m_0 as follows:

$$\forall s \in \mathcal{S}, \forall H \in \{A,B\}, Pl(H)(s) = \delta_V(m_0(H)(s)), Bel(H)(s) = \varepsilon_V(m_0(H)(s))$$
$$\Rightarrow \forall s \in \mathcal{S}, \forall H \in \{A,B\}, m(H)(s) = \varepsilon_V(m_0(H)(s)), m(A \cup B)(s) = \delta_V(m_0(H)(s)) - \varepsilon_V(m_0(H)(s))$$

Using duality: $\forall s \in \mathcal{S}, Bel(B)(s) = \varepsilon_V(m_0(B)(s)) = \varepsilon_V(1 - m_0(A)(s)) = 1 - \delta_V(m_0(A)(s)) = 1 - Pl(A)(s)$

$\Rightarrow$ m is a bba well defined (values in [0,1], sum=1)

$\Rightarrow$ $A \cup B$ is now a focal element of m, with value defined accordingly to the context observed through the structuring element

Case of closing and opening instead of dilatation and erosion MM operators:

$$\forall s \in \mathcal{S}, \forall H \in \{A,B\}, Pl(H)(s) = \varphi_V(m_0(H)(s)), Bel(H)(s) = \gamma_V(m_0(H)(s))$$
$$\Rightarrow \forall s \in \mathcal{S}, \forall H \in \{A,B\}, m(H)(s) = \gamma_V(m_0(H)(s)), m(A \cup B)(s) = \varphi_V(m_0(H)(s)) - \gamma_V(m_0(H)(s))$$

Math. morphology operators: background notions (I)

Math. Morphology (MM) operators defined on **complete lattice** $\mathcal{L}$ (partially ordered set w.r.t. $\leq$, with a supremum and an infimum for every non-void subpart) $\rightarrow$ For grey level images, $\mathcal{L}$ = the lattice of functions on the image domain $\mathcal{S}$, $\leq$ = the partial order on numerical functions.

MM operators are defined through a structuring element g (invariance by translation)
For g flat ($\mathcal{S}$ subset), denoting E_s the support of g translated in $s \in \mathcal{S}$:

$$\forall s \in \mathcal{S}, \quad \delta_g(y)(s) = \sup_{s' \in E_s} y_{s'} \qquad \varepsilon_g(y)(s) = \inf_{s' \in E_s} y_{s'}$$

Duality: $\forall y = \{y_s, s \in \mathcal{S}, y_s \in [0,1]\}, \ \delta_g(y) = 1 - \varepsilon_g(1-y)$

Properties:

- Monotonically increasing relatively to y : $y \leq y' \Rightarrow \delta_g(y) \leq \delta_g(y')$ and $\varepsilon_g(y) \leq \varepsilon_g(y')$
- (Anti-)Extensivity if $0 \in E_x$ the support of g : $\varepsilon_g(y) \leq y \leq \delta_g(y)$
- Monotonically (De)Increasing relatively to g : $g \leq g' \Rightarrow \delta_g(y) \leq \delta_{g'}(y)$ and $\varepsilon_g(y) \geq \varepsilon_{g'}(y)$
- Adjunction : $y \leq \varepsilon_g(y') \Leftrightarrow \delta_{\check{g}}(y) \leq y'$
- Commutations: e.g. $\delta_{g \vee g'}(y) = \delta_g(y) \vee \delta_{g'}(y)$, $\varepsilon_{g \vee g'}(y) = \varepsilon_g(y) \wedge \varepsilon_{g'}(y)$ etc.

Math. morphology operators: background notions (II)

For g flat ($\mathcal{S}$ subset), denoting $\breve{g}$ the transpose of g:

$$\gamma_g(y)(s)=\delta_{\breve{g}}\left(\varepsilon_g(y)(s)\right)=\sup_{s'\in\breve{E}_s}\left\{\inf_{s''\in E_{s'}}\{y_{s''}\}\right\} \qquad \varphi_g(y)(s)=\varepsilon_{\breve{g}}\left(\delta_g(y)(s)\right)=\inf_{s'\in\breve{E}_s}\left\{\sup_{s''\in E_{s'}}\{y_{s''}\}\right\}$$

Duality: $\forall y=\{y_s, s\in\mathcal{S}, y_s\in[0,1]\},\ \varphi_g(y)=1-\gamma_g(1-y)$

Properties:

Monotonic increasing relatively to y: $y\le y' \Rightarrow \begin{cases}\gamma_g(y)\le\gamma_g(y')\\ \varphi_g(y)\le\varphi_g(y')\end{cases}$

(Anti-)Extensivity: $\gamma_g(y)\le y\le\varphi_g(y)$

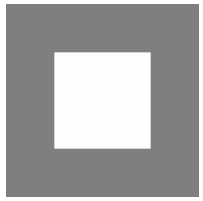
(De)Increase relatively to g: $g\le g'\Rightarrow\begin{cases}\gamma_g(y)\ge\gamma_{g'}(y)\\ \varphi_g(y)\le\varphi_{g'}(y)\end{cases}$

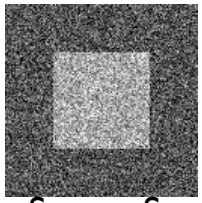
$$\begin{cases}\gamma_{g^m}\left(\gamma_{g^n}(y)\right)=\gamma_{g^{max(m,n)}}(y)\\ \varphi_{g^m}\left(\varphi_{g^n}(y)\right)=\varphi_{g^{max(m,n)}}(y)\end{cases}\Rightarrow\text{Idempotence}$$

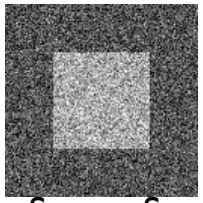
→ Morphological filters

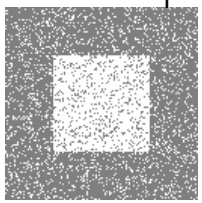
$$\varepsilon_g(y)\le\gamma_g(y)\le y\le\varphi_g(y)\le\delta_g(y)$$

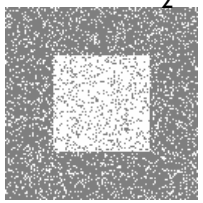
Ex. on simulated data




Source S_1

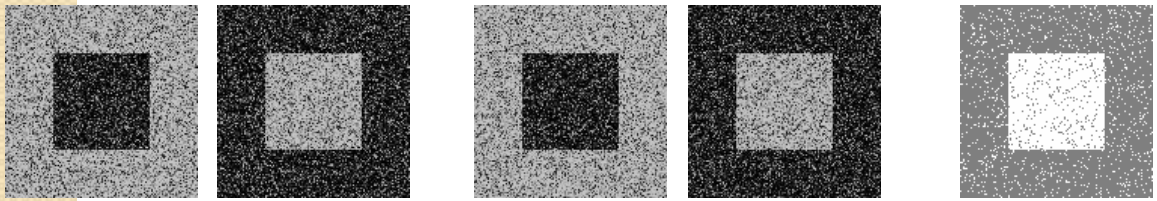

Source S_2


S_1 classif.

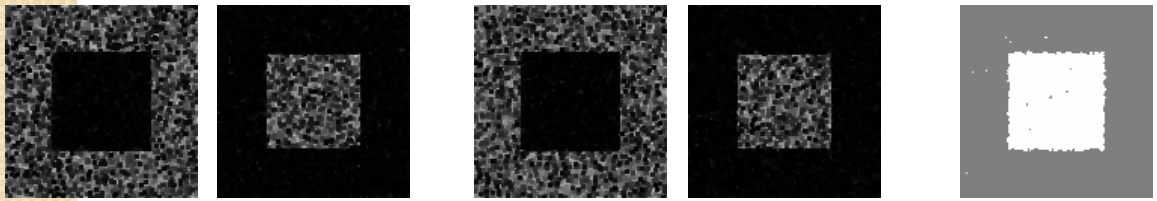

S_2 classif.

$B = \overline{A}$

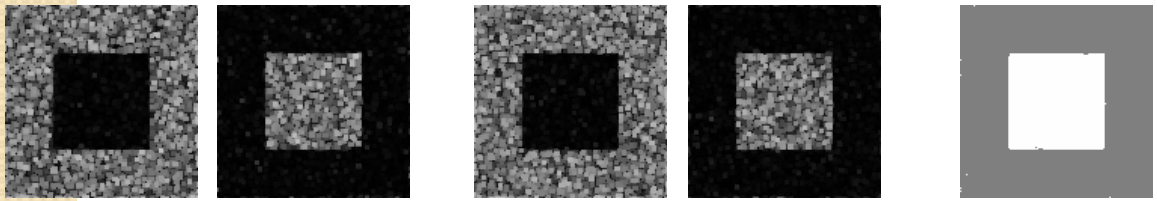
$m_1(A)$ $m_1(B)$ $m_2(A)$ $m_2(B)$ $S_{1\&2}$ classif.



Erosion $\rightarrow \varepsilon(m_i)$

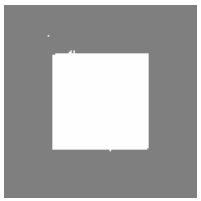


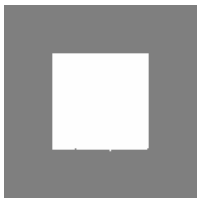
Opening $\rightarrow \gamma(m_i)$



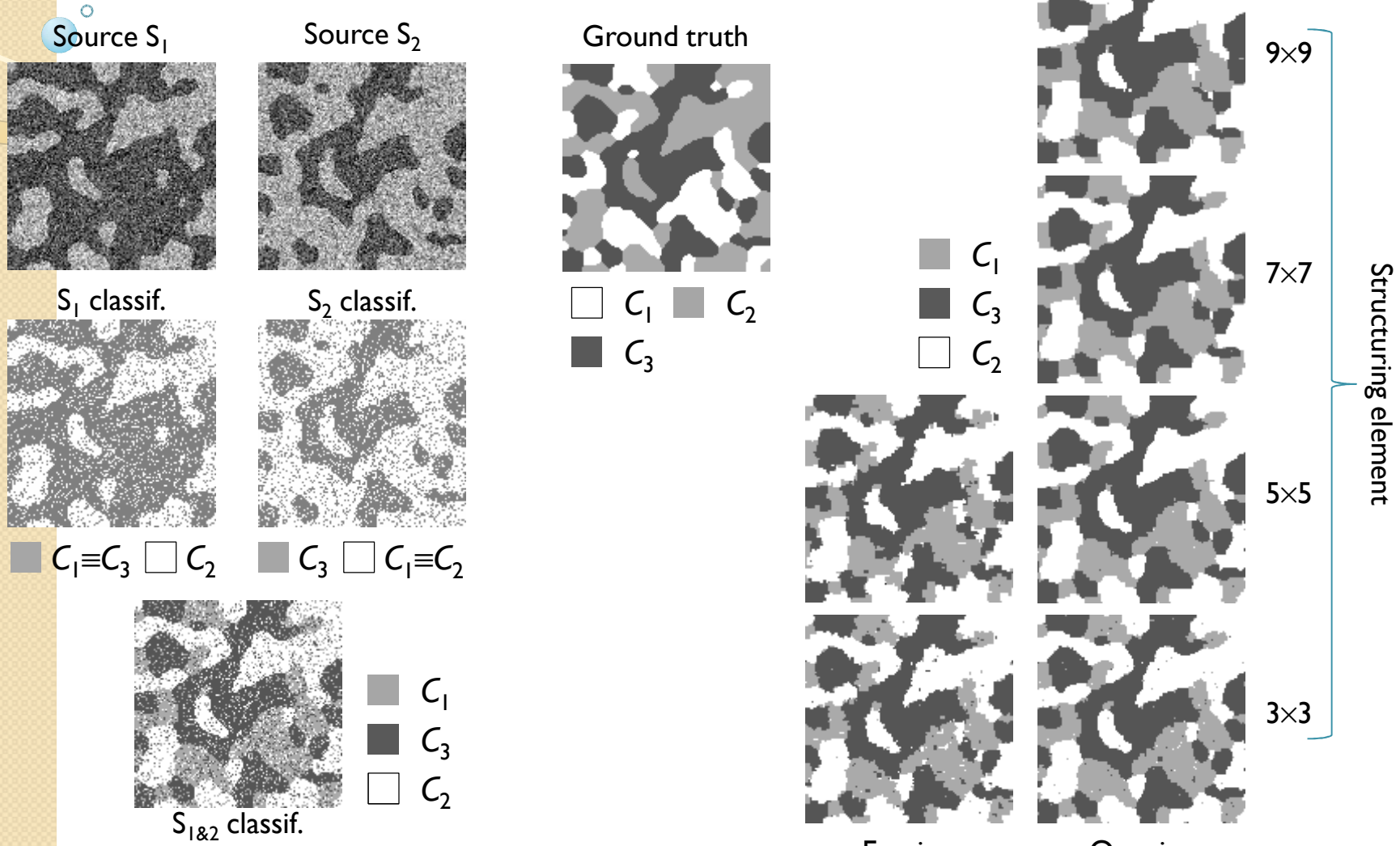
Structuring element 3×3


Structuring element 5×5




Structuring element 7×7

Ex. on simulated data



Spatial information viewed as an independent information source

Link with A Priori Probability and Markov Random Fields models

Contextual classification

Basic idea: a priori on the probability of a configuration of *X*

Classical models: regular configurations more probable than irregular ones

Configurations defined on a neighbourhood, e.g. 4-connectivity, 8-connectivity, ...



Basic idea: neighbour label configuration is an information source

⇒ Modelling it through a bba $\boldsymbol{m_N}$ and combining it with the other bbas

1st example:

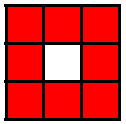
forbidden configurations → conditioning on the possible (allowed) configurations

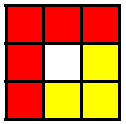
e.g. a priori about isolated pixels ⇒ not possible:

Assume a 3 label/color discernment frame ({Red,Green,Yellow}).

Given the neighbourhood N_1, there is only one focal element: $\boldsymbol{m_N}(\{Red\})=1$, and

Given the neighbourhood N_2, $\boldsymbol{m_N}(\{Green\})=0$ and {Red,Yellow} is focal element

N_1 

N_2 

2nd example:

More 'credible' configurations → ad-hoc bba

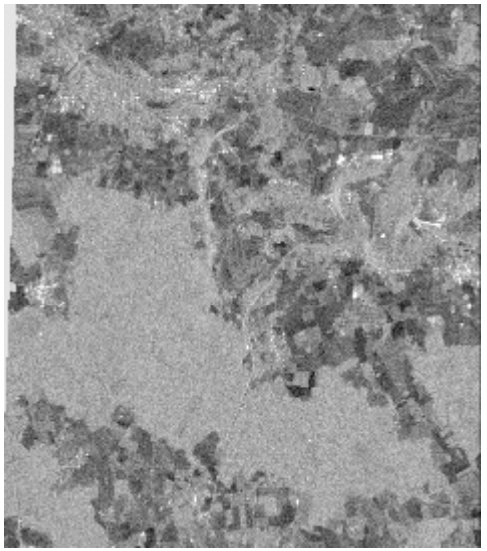
e.g. Contextual a priori probabilities => bba allocation from these probabilities

e.g. Given neighbourhood N_2, $\boldsymbol{m_N}(\{Red\})=2/9$, $\boldsymbol{m_N}(\{Red,Yellow\})=2(3/9)$, $\boldsymbol{m_N}(\{Red,Green,Yellow\})=1/9$

Le Hégarat-Mascle S. et al., 1998, "Introduction of neighborhood information in evidence theory and application to data fusion between radar and optical images with partial cloud cover", *Pattern Recognition*, **31**(11):1811-1823.

Objective: forest area detection
The sources: radar and optical sensor
→ cloud ⇒ ignorance,
→ cloud boundaries ⇒ imprecision,
→ speckle ⇒ imprecision.





Discernment frame Ω= {Forest (F), unforested area (NF)}

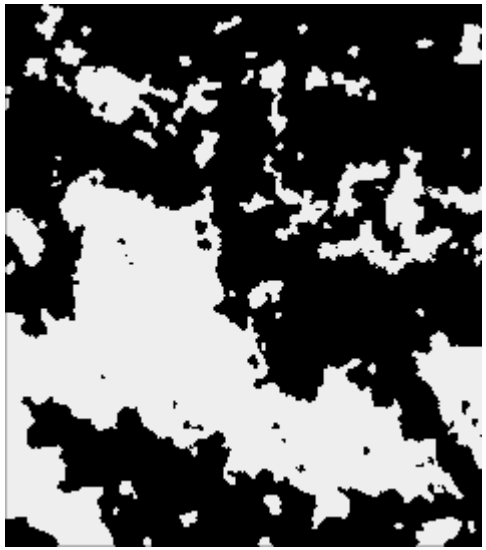
Modelling the sources:
Ad-hoc mass allocation
- Radar image → global discounting based on monosource performance (learning step)
- Optical image → $m_0(\Omega)=1$ on the cloud mask, $m_0(\Omega)$ ↗ around the cloud mask

⇒ Ability to give more importance to radar image and neighbourhood under and around the clouds

Bba representing neighbourhood information:

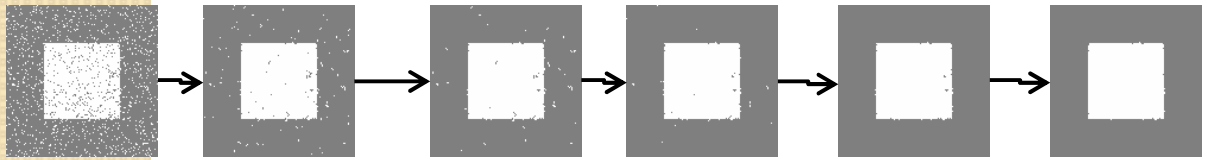
→ based on empiric frequencies of F and NF in the pixel neighbourhood

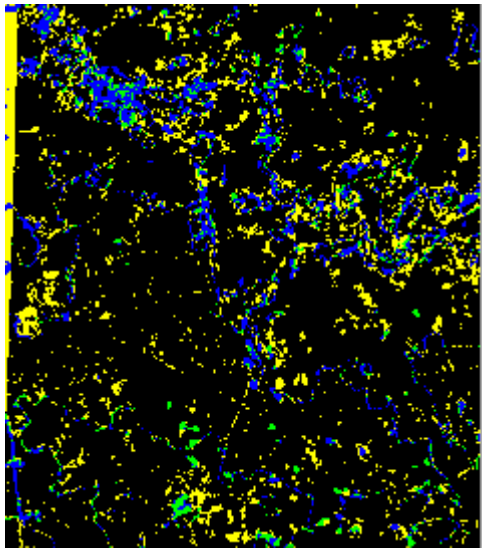
→ discounted to weight it relatively to remote sensing data bba ($m_R \oplus m_O$)



Data fusion ≡ Iterative process

1. compute remote sensing bba $m_D = m_R \oplus m_O$ and spatial bba m_N s.t. $m_N(\Omega)=1$
2. perform multisource classification $m = m_D \oplus m_N$
3. update m_N
4. if stopping criterion is not verified goto step 2.





Comparison with SAR errors:
Yellow=corrected pixels, blue=uncorrected errors, green=new errors

Link with Markov random fields

Bendjebbour A. et al., 2001, "Multisensor image segmentation using Dempster-Shafer fusion in Markov fields context", *IEEE Trans. Geoscience & Remote Sensing*, **39**(8):1789-1798.

Formal extension of probabilities defined on $\Omega=\{\omega_1,...,\omega_k\}$ to 2^Ω

→ $P(X=x)$ replaced by M_0 et $P(Y=y/X=x)$ replaced by M_y

→ If M_0 is a Markov Random Field with $M_0[x]=\gamma.e^{-U(x)}$, where $U(x)$ is a sum potential over the clique set, and if M_y is such that $\forall A=(A_s)_{s\in S}, A_s\subseteq\Omega, M_y(A)=\prod_{s\in S}\left(m_s^1\oplus...\oplus m_s^n\right)(A_s)$ where n is the number of sources
then $M_0\oplus M_y$ is a Markov field

The main difference with previous model is the spatial bba definition. Indeed:

$$M_0[x]=\gamma.e^{-U(x)}=\gamma.e^{-\sum_{c\in C}U_c(x_s,s\in c)}=\gamma.\prod_{s\in S}e^{-\frac{1}{|c|}\sum_{c\in C,s\in c}U_c(x_s,s\in c)}\Rightarrow\forall x_s\in\Omega, m_0(x_s)\propto e^{-\frac{1}{|c|}\sum_{c\in C,s\in c}U_c(x_s,s\in c)}$$

Alternative def. : Richards J. and Jia C., 2007, "A Dempster-Shafer relaxation approach to context classification", *IEEE Trans. Geoscience & Remote Sensing*, **45**(5):1422-1431.



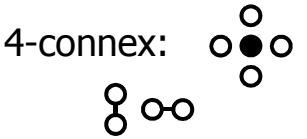
Markov random fields: background notions

X is a MRF , i.e. $P(X_s = x_s / x_{-s}) = P(X_s = x_s / x_t, t \in V_s)$ where $x_{-s} = \{x_t, t \in S, t \neq s\}$ and $P(X)>0 \ \forall X$

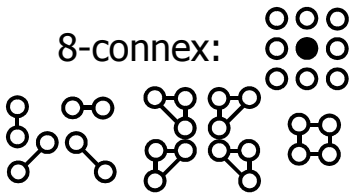
$\Leftrightarrow X$ is a Gibbs field, i.e. $P(X = x) = \frac{1}{Z} exp(-U(x))$ with $U(x) = \sum_{c \in C} U_c(x_s, s \in c)$

where C is the set of cliques (a clique being a subset of S so that it is a set of pixels mutually neighbours w.r.t. the considered neighbourhood)

4-connex:



8-connex:



Then: $P(X_s = x_s / X_{-s} = x_{-s}) = \frac{exp(-U_s(x_s, x_t, t \in V_s))}{\sum_{l_s \in \Omega} exp(-U_s(x_s = l_s, x_t, t \in V_s))}$

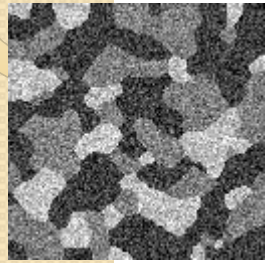
Maximum A Posteriori criterion $\Rightarrow x^*(y) = \underset{x' \in \Omega^{|S|}}{arg\,max} P(y / x').P(x')$

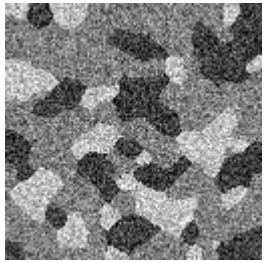
Using assumptions:

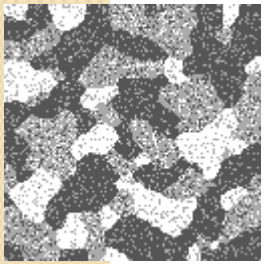
- ✓ $P(Y=y / X=x) = \prod_{s \in S} P(Y_s=y_s / X_s=x_s)$ (conditional indep. w.r.t. X)
- ✓ $\forall s \in S, P(Y_s=y_s / X_s=x_s) > 0 \Rightarrow U_s^0(x_s,y_s) = -\ln(P(y_s / x_s))$

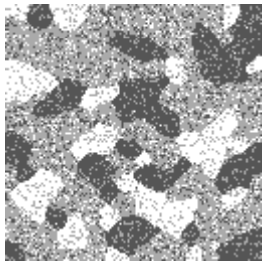
$\Rightarrow$ The couple (X,Y) is MRF: $P(X = x, Y = y) = \frac{1}{Z} \exp\left\{ - \sum_{c \in C} U_c(x_S, s \in c) - \sum_{s \in S} U_s^0(x_s, y_s) \right\}$

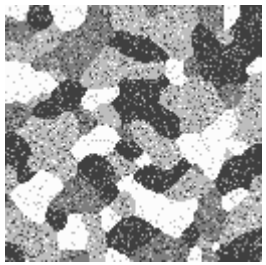
Comparison on simulated data

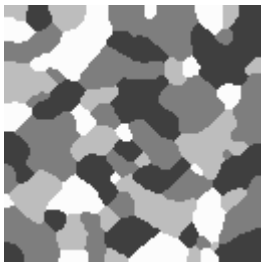

Source S_1


Source S_2


S_1 classif.


S_2 classif.

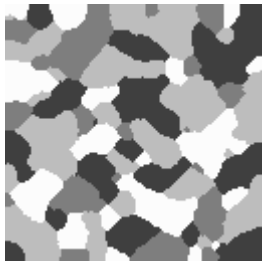

$S_{1\&2}$ classif.

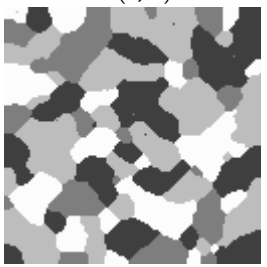

Ground truth

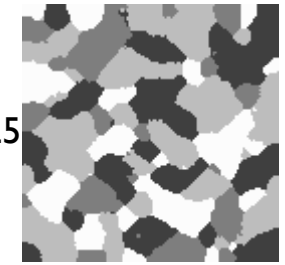
α: discounting parameter applied to $m_D=m_1 \textcircled{\cap} m_2$

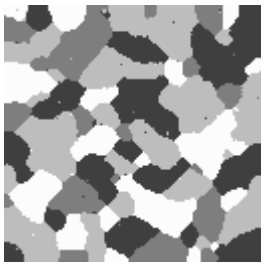
β: Potts model parameter

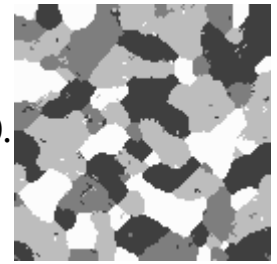
$$U_{C=(s,s')}(x_s,x_{s'})=\begin{cases}-\beta \text{ if } x_s=x_{s'}\\ +\beta \text{ if } x_s\neq x_{s'}\end{cases}$$

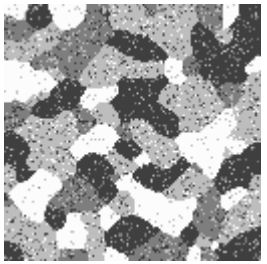

α=0.5


β=25.


α=0.25


β=10.


α=0.


β=1.

Ad-hoc neighbour mass

Markov model-Potts mass

Automatic estimation of the discernment frame

Case of unsupervised classification

Minimum multisource discernment frame

Case of unsupervised classification:
Discernment frame (DF) = set of distinguishable classes → How learning this set?

Basic idea: monosource classification are information sources for multisource DF estimation
→ Consider the non-empty intersections of every pairs (2 source case) of monosource classes (i.e. Singleton hyp. in the monosource DF)
→ Such a DF is the minimum (in terms of number of elements) DF that is a common refinement of monosource DF

Ex. $S_1 \rightarrow \{A_1,A_2\}$, $S_2 \rightarrow \{B_1,B_2\}$, the min. common DF is $\{A_i \cap B_j \text{ s.t } A_i \cap B_j \neq \varnothing, (i,j) \in \{1,2\}^2\}$

Now some classification errors may produce fictitious class (intersections)
→ remove them based on a criterion of minimum number of pixels after *robust* classification (e.g. multisource) → iterative algorithm

Iterative DF estimation

Basic algorithm (2 sources Y^1, Y^2, defined over lattice S)

- Compute the monosource classifications X^1, X^2: $\forall s \in \mathcal{S}, x_s^1 \in \{A_i, 1 \le i \le c_1\}, x_s^2 \in \{B_i, 1 \le i \le c_2\}$
- Compute the bbas of the sources $m_1^{\Omega_1}$ and $m_2^{\Omega_2}$ in their respective discernment frames Ω_1 and Ω_2: $\Omega_1 = \{A_i, 1 \le i \le c_1\}, \Omega_2 = \{B_i, 1 \le i \le c_2\}$
- Initialize $\Omega_{1,2}$ (DF) to the set of non empty intersections between monosource classes

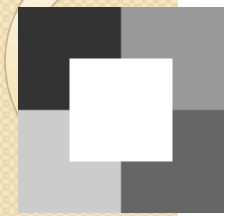
$$\Omega_{1,2} = \left\{ H_i = A_j \cap B_k \,/\, \left| \left\{ x_s^1 = A_j \right\} \cap \left\{ x_s^2 = B_k \right\} \right| > 0 \right\}$$

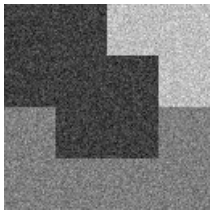
- Repeat until stop ($\Omega_{1,2}$ does not change)
  - Extent the bbas $m_1^{\Omega_1}$ and $m_2^{\Omega_2}$ to $\Omega_{1,2}$: $m_1^{\Omega_{1,2}}$ and $m_2^{\Omega_{1,2}}$
  - Compute the combination: $m_{1,2}^{\Omega_{1,2}} = m_1^{\Omega_{1,2}} \textcircled{\cap} m_2^{\Omega_{1,2}}$
  - Compute the multisource classification $X^{1,2}$
  - Remove the hypotheses of $\Omega_{1,2}$ not enough supported:

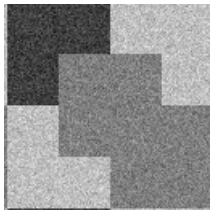
$$\Omega_{1,2} \leftarrow \Omega_{1,2} \setminus \left\{ H_i \,/\, \left| \left\{ x_s^{1,2} = H_i \right\} \right| < n_{min} \right\}$$

- Perform final multisource classification (if different decision criterion and/or introduction of supplementary information)

Simulated example (step by step)

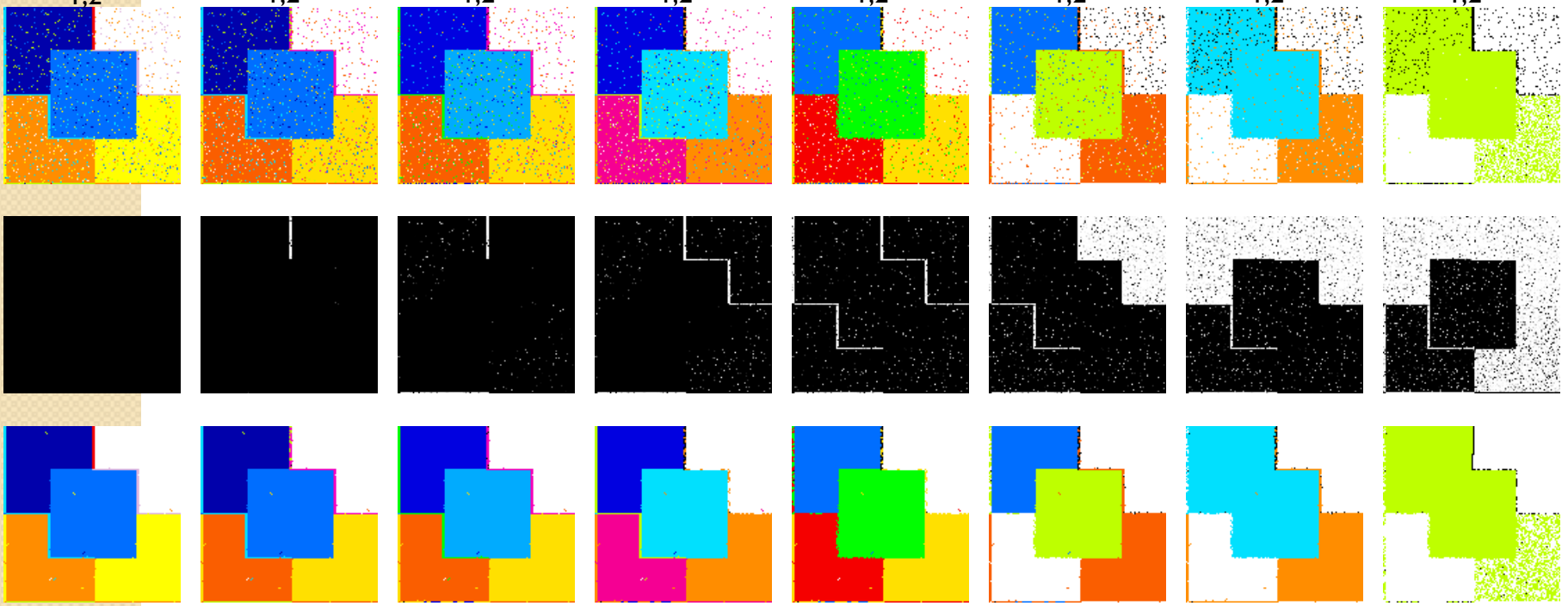

Ground truth


Source S_1

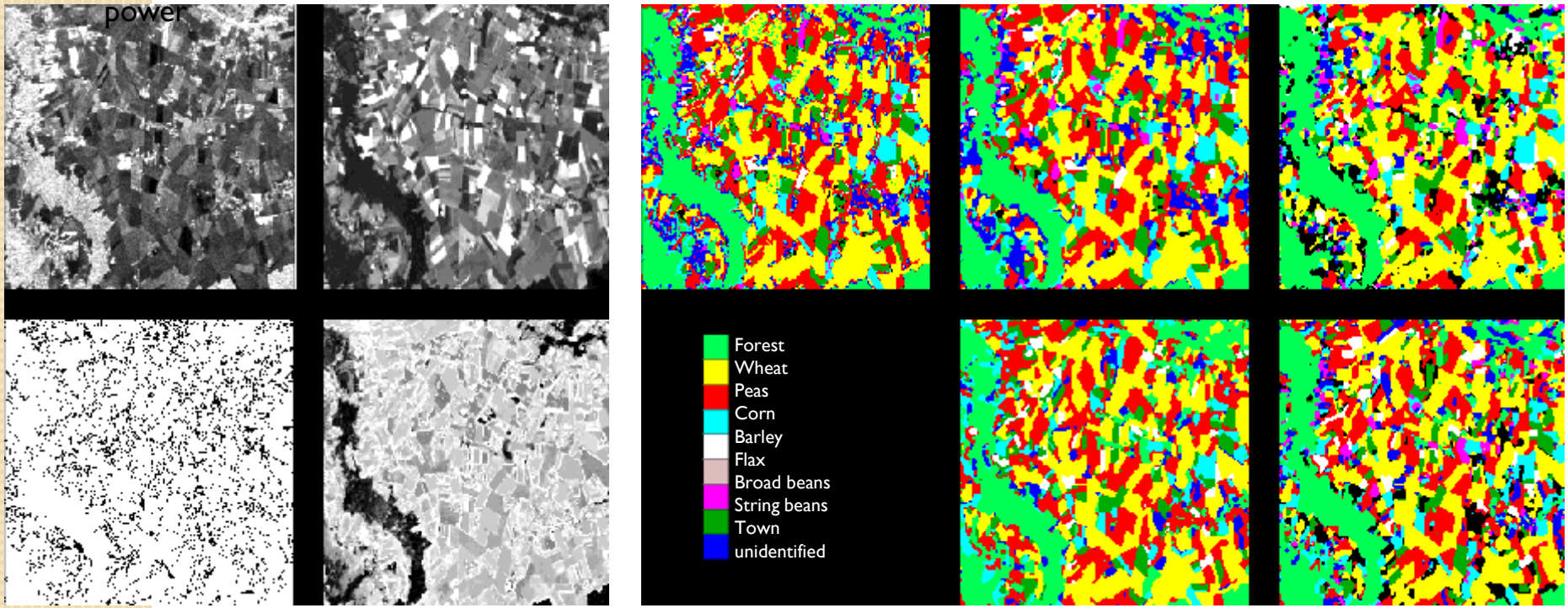

Source S_2

2 kinds of imprecision:
-Class ambiguities
-Superimposition error (1-2 pixels)

$\Omega_{1,2}=9$ $\Omega_{1,2}=8$ $\Omega_{1,2}=7$ $\Omega_{1,2}=6$ $\Omega_{1,2}=5$ $\Omega_{1,2}=4$ $\Omega_{1,2}=3$ $\Omega_{1,2}=2$



Le Hégarat-Mascle S. et al., 1997, "Application of Dempster-Shafer evidence theory to unsupervised classification in multisource remote sensing", *IEEE Transactions on Geosciences and Remote Sensing*, **35**(4):1018-1031.

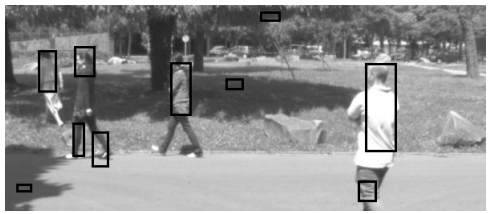


Sequential estimation of the discernment frame

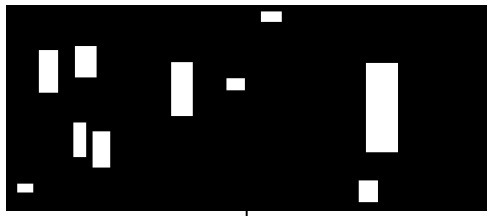
Case of classification at image level

Classification problem at object level

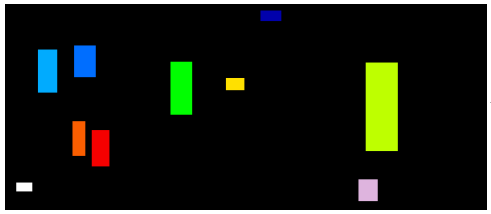
Example in videosurveillance application:



Different kinds of methods : Segmentations methods, bounding box of areas of interest, etc. → binary image at 'window' resolution



Connected component (CC) labeling step →



⇒ Problem is rewritten as a problem of multilabel classification at image level:
Assuming N CCs noted $\{O_i\}_{i\in\{1,\ldots,N\}}$ decide which ones correspond to actual objects:

Discernment frame about the object actual existence ≡ cross-product space

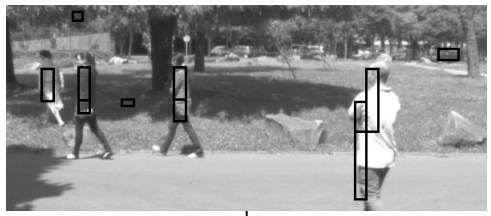
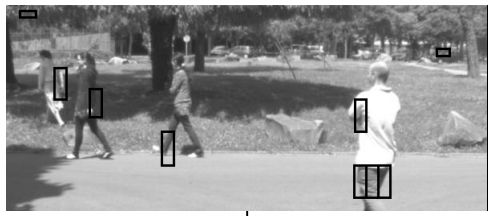
$$\{O_1,\overline{O_1}\}\times\ldots\times\{O_i,\overline{O_i}\}\times\ldots\times\{O_N,\overline{O_N}\}$$

Ex. $\Omega=\{O_1,\overline{O_1}\}\times\{O_2,\overline{O_2}\}\times\{O_3,\overline{O_3}\}=\{(\overline{O_1},\overline{O_2},\overline{O_3}),(\overline{O_1},\overline{O_2},O_3),(\overline{O_1},O_2,\overline{O_3}),(\overline{O_1},O_2,O_3),\ldots,(O_1,O_2,O_3)\}$

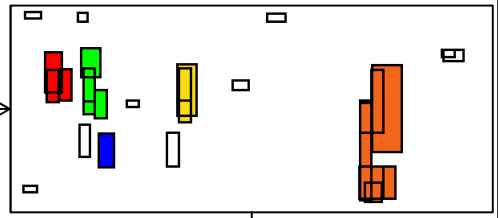
⇒ Unsupervised classification → determine object features (size, location, etc.) simultaneously to classification

Subproblem: dynamic updating of the set of the objects

Sources are assumed either simultaneous (time static) or compensated for object motion
→ Only cope with 2 source imperfections : false alarm existence and object fragmentation

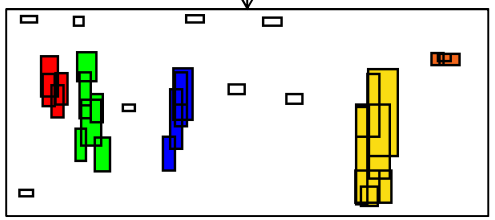
⇒ CC association is simply based on spatial relationship (e.g. spatial overlapping) →



Fusion is performed sequentially

⇒ updating of the set of potential objects → Ω

⇒ Updating the beliefs on Ω




Rekik, W. et al., 2013, "Dynamic estimation of the discernment frame in belief function theory", *Int. Conf. on Information Fusion*.

Assume several sources → As many cross-product spaces to represent the beliefs about the object actual existence $\{O_1,\overline{O}_1\}\times...\times\{O_N,\overline{O}_N\}$ $\{O'_1,\overline{O}'_1\}\times...\times\{O'_{N'},\overline{O}'_{N'}\}$

✓ Case where Ω (updated DF) is simply an extension of Ω_O:

$$\Omega=\Omega_O\times\{O'_{i_1},\overline{O}'_{i_1}\}\times...\times\{O'_{i_I},\overline{O}'_{i_I}\}$$

- For each BBA m_O defined on its own cross-product space $\Omega_O=\underset{i=1}{\overset{N}{\times}}\{O_i,\overline{O}_i\}$
  - for each element $\{O_i,\overline{O}_i\}$ of Ω_O,
    - marginalize m_A to $\{O_i,\overline{O}_i\}$ $\rightarrow m_O^{\Omega_O\downarrow\{O_i,\overline{O}_i\}}$
  - extend the obtained BBA on common discernment frame Ω → $\left(m_O^{\Omega_O\downarrow\{O_i,\overline{O}_i\}}\right)^{\uparrow\Omega}$
- Combine the BBAs extended on Ω → $m^{\Omega}=\left(m_O^{\Omega_O\downarrow\{O_i,\overline{O}_i\}}\right)^{\uparrow\Omega}\left(m_{O'}^{\Omega_{O'}\downarrow\{O'_i,\overline{O}'_i\}}\right)^{\uparrow\Omega}$

✓ Case where some O_i have been found to represent the same object:

After marginalization $m_O^{\Omega_O\downarrow\{O_i,\overline{O}_i\}}$, fusion/redefinition the marginalized BBAs that represent the same object

⇒Decision:

Maximum of pignistic probability → choose, for every potential object, the hypothesis maximizing the marginal mass $m^{\Omega\downarrow\{O_i,\overline{O}_i\}}$ among $\{O_i,\overline{O}_i\}$

Example: $\Omega_O=\{o_1,\bar{o}_1\}$ $\Omega_{O'}=\{o'_1,\bar{o}'_1\}\times\{o'_2,\bar{o}'_2\}$, $o_1\equiv o'_1 \Rightarrow \Omega=\{o_1,\bar{o}_1\}\times\{o_2,\bar{o}_2\}$

Hyp.	$(\bar{o}_1,\bar{o}_2)$	$(\bar{o}_1,o_2)$	$(\bar{o}_1,\{\bar{o}_2,o_2\})$	$(o_1,\bar{o}_2)$	(o_1,o_2)	$(o_1,\{\bar{o}_2,o_2\})$	$(\{\bar{o}_1,o_1\},\bar{o}_2)$	$(\{\bar{o}_1,o_1\},o_2)$	Ω
$m_1^{\Omega\downarrow\{o_1,\bar{o}_1\}}$	0	0	0.4	0	0	0.5	0	0	0.1
$m_2^{\Omega\downarrow\{o_1,\bar{o}_1\}}$	0	0	0.4	0	0	0.5	0	0	0.1
$m_2^{\Omega\downarrow\{o_2,\bar{o}_2\}}$	0	0	0	0	0	0	0.3	0.6	0.1
m^{Ω}	0.072	0.144	0.024	0.105	0.21	0.035	0.003	0.006	0.001
$BetP^{\Omega}$	0.1429	0.2654	/	0.2071	0.3846	/	/	/	/
$m^{\Omega}[H](\varnothing)$	0.5	0.425	/	0.456	0.348	/	/	/	/

$m^{\Omega}(\varnothing)=0.4$

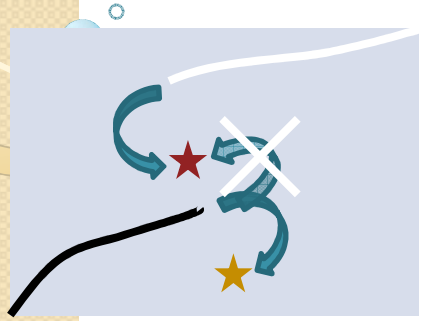
Hyp.	$(\bar{o}_1,\bar{o}_2)$	$(\bar{o}_1,o_2)$	$(\bar{o}_1,\{\bar{o}_2,o_2\})$	$(o_1,\bar{o}_2)$	(o_1,o_2)	$(o_1,\{\bar{o}_2,o_2\})$	$(\{\bar{o}_1,o_1\},\bar{o}_2)$	$(\{\bar{o}_1,o_1\},o_2)$	Ω
$m_1^{\Omega\downarrow\{o_1,\bar{o}_1\}}$	0	0	0.4	0	0	0.5	0	0	0.1
$m_2^{\Omega\downarrow\{o_1,\bar{o}_1\}}$	0	0	0.4	0	0	0.5	0	0	0.1
$m_2^{\Omega\downarrow\{o_2,\bar{o}_2\}}$	0	0	0	0	0	0	0.15	0.35	0.5
m^{Ω}	0.036	0.084	0.12	0.0525	0.1225	0.175	0.0015	0.0035	0.005
$BetP^{\Omega}$	0.1633	0.245	/	0.2367	0.355	/	/	/	/
$m^{\Omega}[H](\varnothing)$	0.4375	0.3875	/	0.366	0.294	/	/	/	/

Video sequences and object tracking problem

Def.: 1 *target* = 1 object detected at time *t*

1 *track* = 1 object detected at different times

Data association sub-problem



If 1-1 associations AND additive cost function → Matrix of the association costs

	★	★
⌒	2	4
⌒	1	2

Solution ⇒ Hungarian (Munkres) assignment algorithm

Matrix of the association costs based on

- Criteria of feature similarity:
  - Radiometric, colour, texture,
  - Pattern: surface, height/width ratio,
  - Others : speed…
- Criteria of distance:
  - in the image domain,
  - 3D…

⇒ Predict tracks at *t* (from *t*-1) to assess their similarity / distance with the targets at *t*:
- Pattern features (colour etc.) → generally assumed to be constant in time …
- Spatial locations (image, 3D) → generally predicted assuming a regular motion

⇒ Extension to the case of 0-1 or 1-0 associations: *N* targets, *M* tracks
- Extent the cost matrix to $(N+M)\times(N+M)$ (or to $\max(N,M)$)
- Define costs of non association (of a track and of a target)

Mercier D. et al., 2011, "Object association with belief functions, an application with vehicles", *Information Sciences*, **181**(24):5485-5500

- BF represent the information regarding the association of pairs (target, track)
- 3 discernment frames: $\Omega_{i,j}=\{0,1\}$, $\Omega_{Oi}=\{T_1,..,T_M,*\}$, $\Omega_{Tj}=\{O_1,..,O_N,*\}$
- Basic knowledge deals with the potential association of target O_i with track T_j $\Rightarrow$ expressed on the discernment frame $\Omega_{i,j}=\{0,1\} \rightarrow m^{\Omega i,j}$

  $\Rightarrow$ interest of TBF = possibility to model (partial) ignorance relative to the relevance of this association.
- Refined discernment frames $\Omega_{Oi}=\{T_1,..,T_M,*\}$ or $\Omega_{Tj}=\{O_1,..,O_N,*\}$ (* = non association)

$$\Rightarrow\ m^{\Omega_{i,j}\uparrow\Omega_{O_i}}(T_j)=m^{\Omega_{i,j}}(1)\ ,\quad m^{\Omega_{i,j}\uparrow\Omega_{O_i}}(\overline{T}_j)=m^{\Omega_{i,j}}(0)\ ,\ m^{\Omega_{i,j}\uparrow\Omega_{O_i}}(\Omega_{O_i})=m^{\Omega_{i,j}}(\Omega_{i,j})$$

- Combination $\Rightarrow$ $m^{\Omega_{O_i}}=\bigcirc_{j=1}^{M} m^{\Omega_{i,j}\uparrow\Omega_{O_i}}$ or/and $m^{\Omega_{T_j}}=\bigcirc_{i=1}^{N} m^{\Omega_{i,j}\uparrow\Omega_{T_j}}$
- Decision: association function $\hat{a}(.)=\underset{a(.)}{arg\,max}\left[\prod_{i=1}^{N} BetP^{\Omega_{O_i}}(T_{a(i)})\right]$

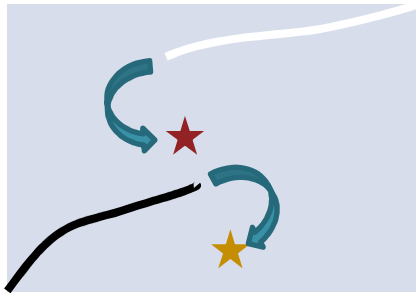
Relatively to the 'classic' approach:

$\Rightarrow$ replace the costs of the potential association of O_i with the different T_j by $BetP^{\Omega_{O_i}}(T_j)$

$\rightarrow$Derive the cost from beliefs taking into account the partial ignorance

$\Rightarrow$replace the sum (of the costs) by the product $\rightarrow$ take the log to use Hungarian algorithm

Toy example: [M11] solution



- Basic belief about association between track and target

Hyp.	yes	no	?
★ ≡ ⌒	0.7	0.1	0.2
★ ≡ ⌒	0.5	0.1	0.4
★ ≡ ⌒	0.5	0.1	0.4
★ ≡ ⌒	0.1	0.7	0.2

Hyp.	⌒	⌒	⌒
★ ≡ ⌒ ↑★	0.7	0.1	0.2
★ ≡ ⌒ ↑★	0.1	0.5	0.4
↑★	0.37	0.19	0.08
★ ≡ ⌒ ↑★	0.5	0.1	0.4
★ ≡ ⌒ ↑★	0.7	0.1	0.2
↑★	0.73	0.07	0.08

BetP	⌒	⌒
★	0.6406	0.3594
★	0.875	0.125

$\Rightarrow 0.875 \times 0.3594 > 0.6406 \times 0.125$

$\Rightarrow$ ☆ ≡ ⌒ and ☆ ≡ ⌒

Ristic B. and Smets P., 2006, "The TBM global distance measure for the association of uncertain combat ID declarations", *Information Fusion*, **7**:276-284

- Use the same frame of discernment for both problems of class estimation and data association problem
  - → 1 discernment frame: $\Theta=\{\theta_1,..., \theta_N\}$
  - → Basic knowledge deals with the potential class of an object $O_i \rightarrow m^{\Theta}\{O_i\}$
- The objects are unlabeled → 2 sets of objects with unknown correspondences
- for n researched pairs of objects observed at t and at $t+1$, discernment frame for multiple objects assignment is Θ^{2n}
- Maximising the plausibility of the hypothesis $H \subseteq \Theta^{2n}$ forcing the same class for associated objects ⇒
- The data association is specified by the association function $\hat{a}(.)$, that is a permutation of the indexes $i \in \{1...n\}$, so that:

$$\hat{a} = \underset{a(.)}{arg\,max}\left[pl^{\Theta\uparrow\Theta^{2n}}\left(\theta(O_i)=\theta\left(O_{a(i)}\right), i=1...n\right)\right] = \underset{a(.)}{arg\,max}\left[\prod_{i=1}^{n}\left(1-m^{\Theta}_{\{O_i\}\cap\{O_{a(i)}\}}(\varnothing)\right)\right]$$

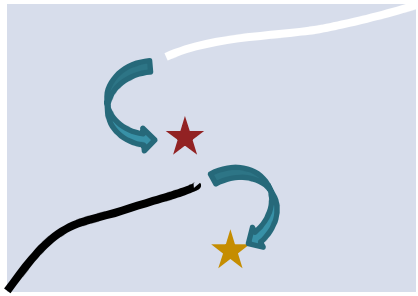
Transposition to the data association problem between tracks and targets:

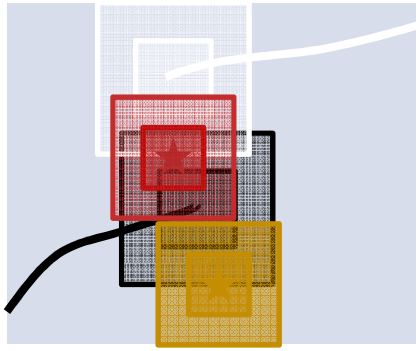
⇒ generalize Θ to the features on with tracks and targets have to agree

→ Localisation, speed, etc.

⇒ replace the 2 sets of objects by the set of targets O_i and tracks T_j

Toy example: [RS06] solution





- Θ=set of elementary blocks forming a partition of the region

Hyp.								
★	0.5	0.5	0	0	0	0	0	0
★	0	0	0.5	0	0	0	0	0
	0	0	0	0	0.5	0.5	0	0
	0	0	0	0	0	0	0.5	0.5

$\cap = \cap(\cup\cup) = \cap(\cup) = \varnothing$

$\Rightarrow$

m(∅)	★	★
	0.25	1
	0	0.25

$\Rightarrow (1-0.25)(1-0.25) > (1-0)(1-1) \Rightarrow$ ★ ≡ and ★ ≡

