
BeaulieuJM.ca/publi/Bea2019a

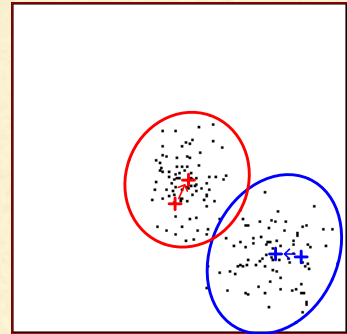
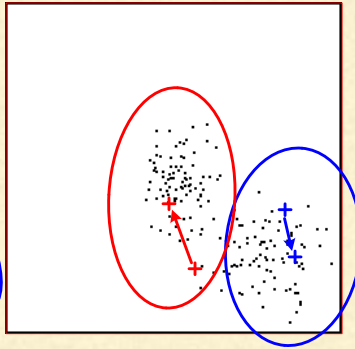
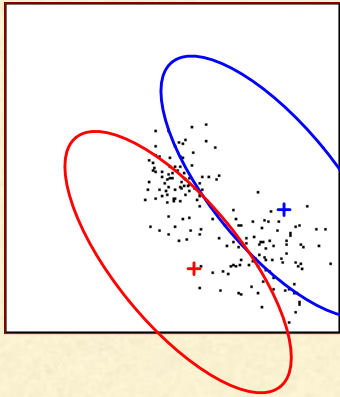
CONTRÔLE DU VOISINAGE POUR UN REGROUPEMENT HIÉRARCHIQUE EFFICACE

Jean-Marie Beaulieu

**Le regroupement hiérarchique est difficile
long temps de calcul**

**Peut être un outils intéressant
avec une implantation rapide**

- **Regroupement Itératif**
- **Algorithme des K-Centres — Déplace les centres**
- **Le nombre de groupes est fixe**



K-Centres: Regroupement itératif

Calcule des distances entre les pixels et les centres

M centres

N pixels



+



- - -

Nombre d'opérations

$M \times N$ (par itérations) \times L itérations \approx L M N

10 centres

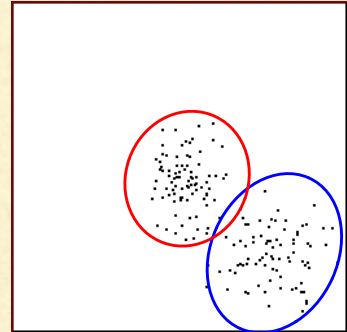
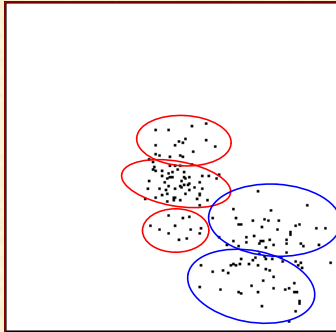
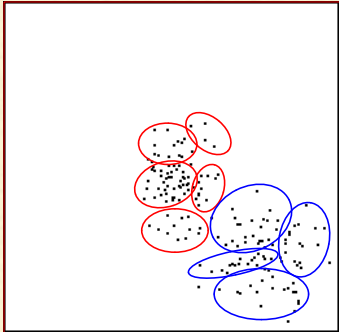
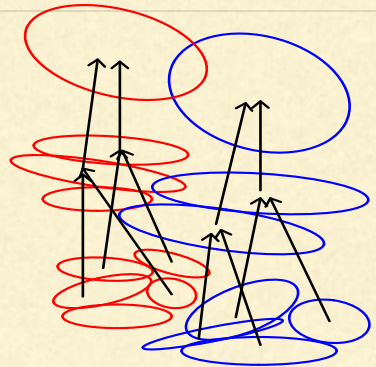
1,000,000 pixels

10 itérations

$\rightarrow 10^8$

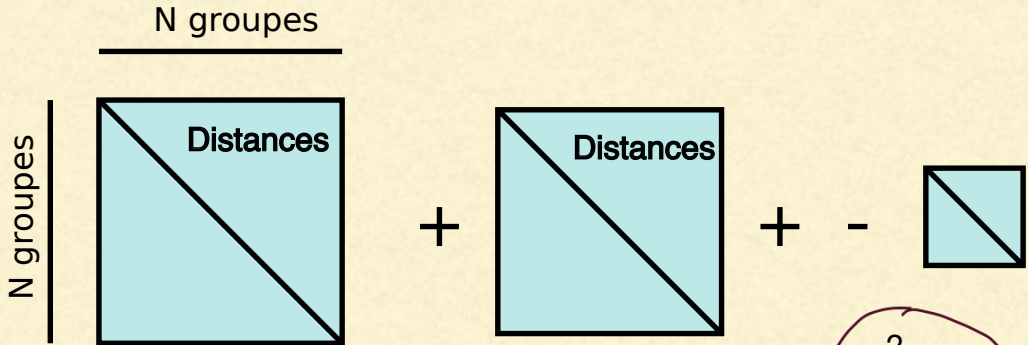
\rightarrow 100 M op

- **Regroupement Hiérarchique**
- **Fusion Séquentielle des groupes**
- **Sélection de la meilleure paire**
- **Construis un arbre des fusions**



Regroupement Hiérarchique

Calcule des distances entre groupes



Nombre d'opérations

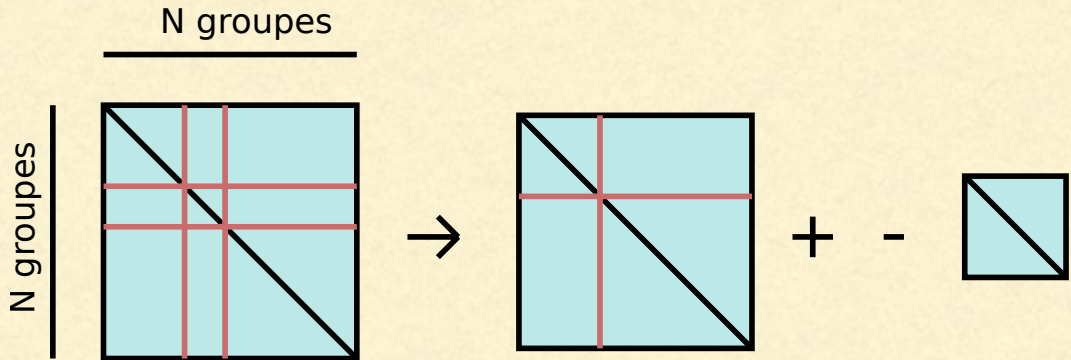
$N \times N$ (par itérations) \times (N-1) itérations $\approx \alpha N^3$

$10^6 \times 10^6 \times 10^6 \rightarrow 10^{18} \rightarrow 1,000,000$ Tera op

$$\sum_{n=N}^2 n^2$$

Regroupement Hiérarchique

Mise à jour des distances



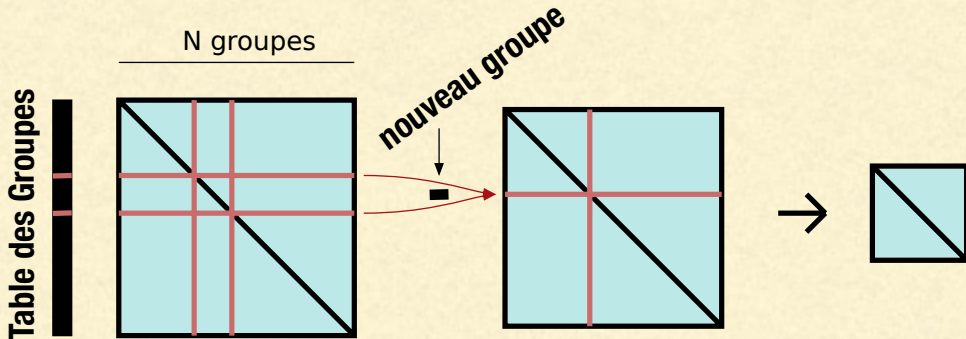
Nombre d'opérations

$$N \times N \text{ (initialisation)} + \sum_{n=N}^2 n \approx \alpha N^2$$

$$10^6 \times 10^6 \rightarrow 10^{12} \rightarrow 1 \text{ Tera op}$$

Regroupement Hiérarchique

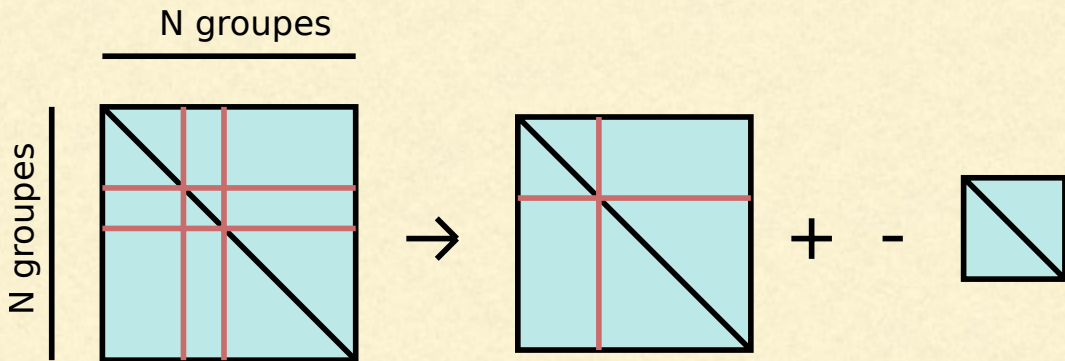
Mise à jour des distances



Nombre d'opérations

$$N \times N \text{ (initialisation)} + \sum_{n=N}^2 n \approx \alpha N^2$$

$$10^6 \times 10^6 \rightarrow 10^{12} \rightarrow 1 \text{ Tera op}$$



Nombre d'opérations

$$N \times N \text{ (initialisation)} + \sum_{n=N}^2 n \approx \alpha N^2$$

Espace mémoire

$$N \times N \rightarrow 10^6 \times 10^6 \rightarrow 10^{12} \rightarrow 1 \text{ Tera valeurs}$$

-
- **Image Polsar multi-vue**
 - Z_k suit une distribution de Wishart complexe

$$p(Z_k | \Sigma) = \frac{L^{3L} |Z_k|^{L-3} \exp \left\{ -L \operatorname{tr} \left(\Sigma^{-1} Z_k \right) \right\}}{\pi^3 \Gamma(L) \Gamma(L-1) \Gamma(L-2) |\Sigma|^L}$$

- **Distance entre groupes D(G_i, G_j)**
- **Log du Rapport de Vraisemblance**

$$D(G_i, G_j) = (n_i + n_j) \ln \left| \hat{\Sigma}_{G_i \cup G_j} \right| - n_i \ln \left| \hat{\Sigma}_{G_i} \right| - n_j \ln \left| \hat{\Sigma}_{G_j} \right|$$

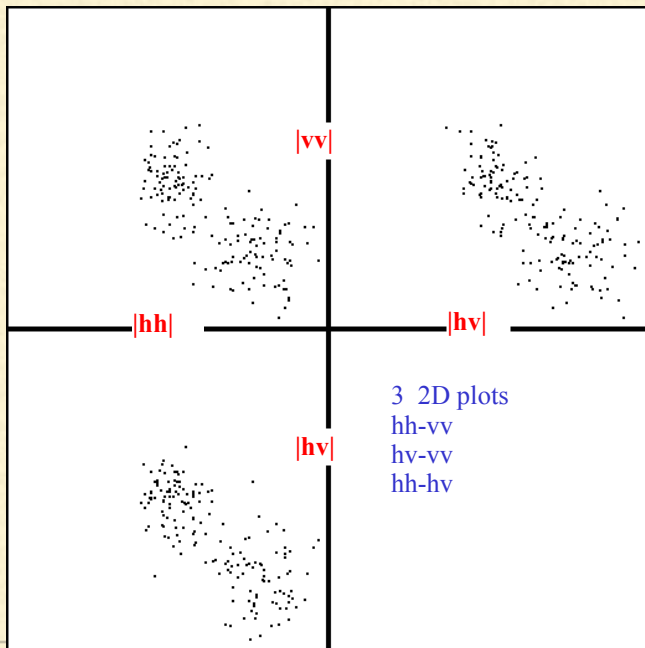
- **Attributes or feature space** (many dimensions)
- **Radiometric information** (or color/spectral)

Radar 1-look

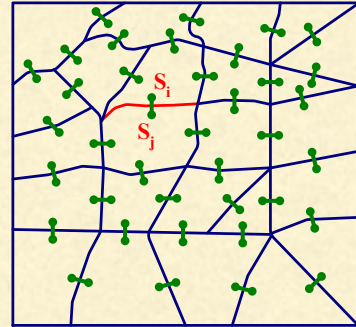
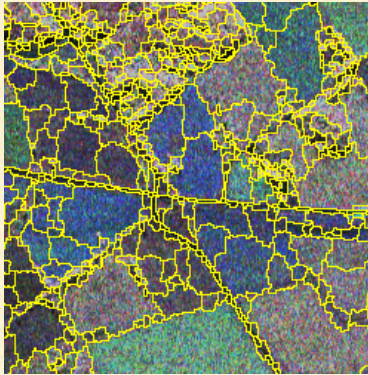
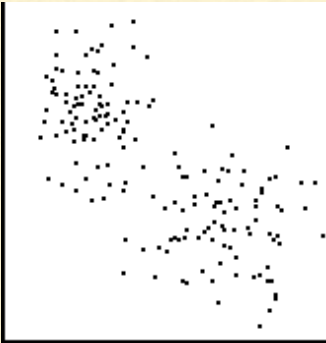
$$x = \begin{bmatrix} hh \\ hv \\ vv \end{bmatrix}$$

Radar multi-look

$$Z = \begin{bmatrix} \overline{hh \ hh^*} & \overline{hh \ hv^*} & \overline{hh \ vv^*} \\ \overline{hv \ hh^*} & \overline{hv \ hv^*} & \overline{hv \ vv^*} \\ \overline{vv \ hh^*} & \overline{vv \ hv^*} & \overline{vv \ vv^*} \end{bmatrix}$$

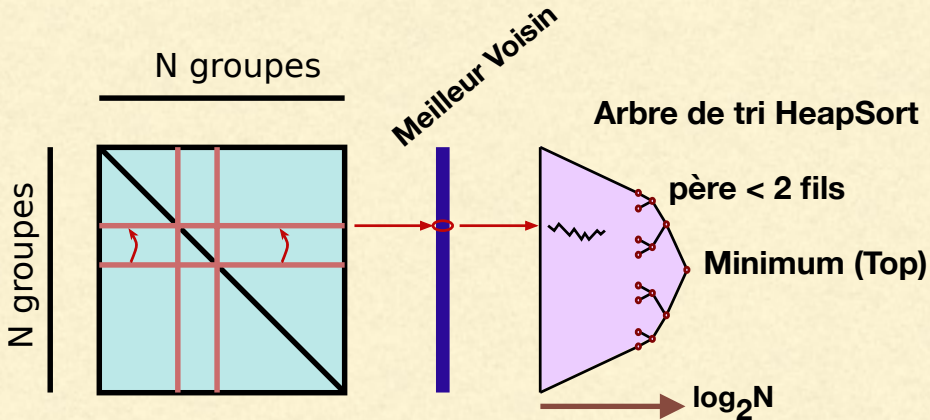


- **Spatial information - position in the image**
- **Clustering -- distance between points $D(G_i, G_j)$**
- **Segmentation -- only adjacent regions**



Recherche du Minimum

Arbre de tri HeapSort — $\propto N \log_2 N$



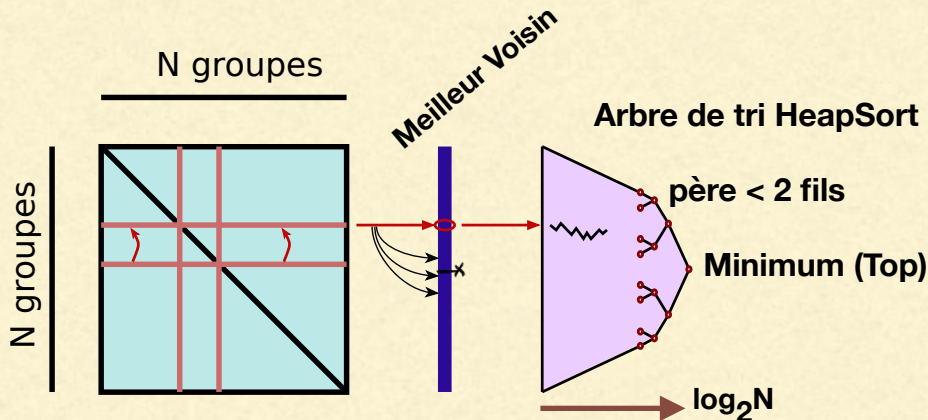
Le minimum indique quel groupe fusionner et son meilleur voisin indique avec qui il est fusionné

Pas besoin de conserver la matrice des distances

Mise à jour Retardée pour les voisins

quand le groupe est choisi comme minimum

- vérifie si déjà fusionné
- élimine ou met à jour (facteur 2 à 5)



Pré-Sélection Rapide

Test rapide → élimine 90% à 95% des calculs de distances

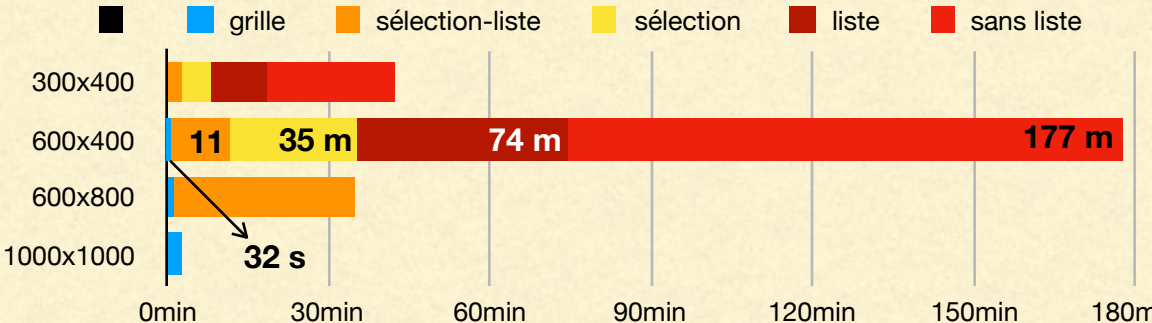
Seuillage → rapport $|hh|_i/|hh|_j$, $|hv|_i/|hv|_j$ et $|vv|_i/|vv|_j$

Listes des Plus Proches Voisins

pour les fusions seulement

Grille des Groupes

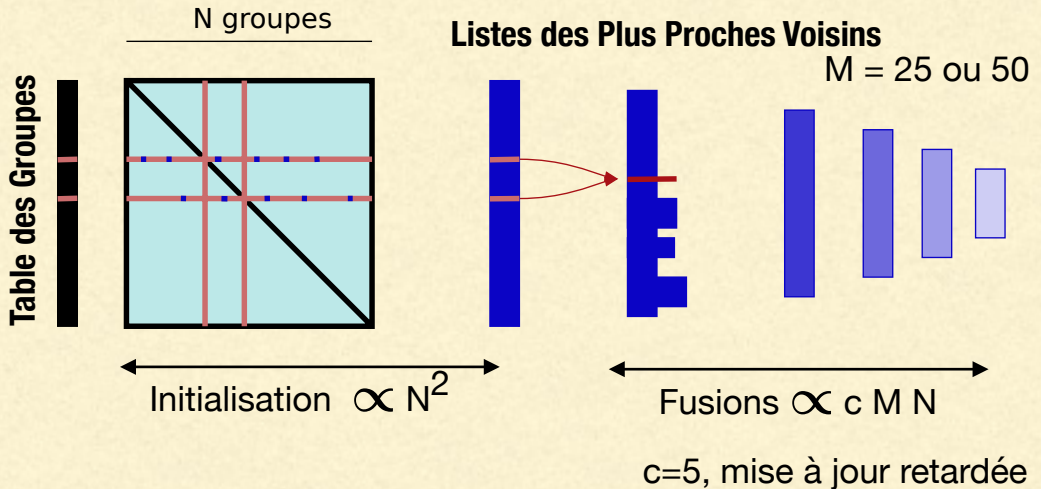
pour l'initialisation des listes



Listes des Plus Proches Voisins

Calcul les distances seulement pour les plus proches voisins
pour les fusions seulement $\propto M N$

temps d'initialisation long $\propto N^2$



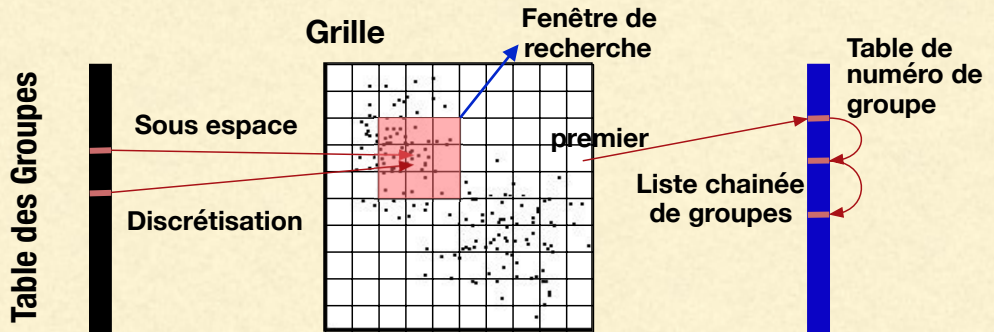
Grille des Groupes

Sous espace des attributs divisé en cellules ($5D, 8 \times 8 \times 8 \times 8 \times 8$)

- Discrétisation → numéro de la cellule
- Cellule → liste chaîné de groupes (un pointer)

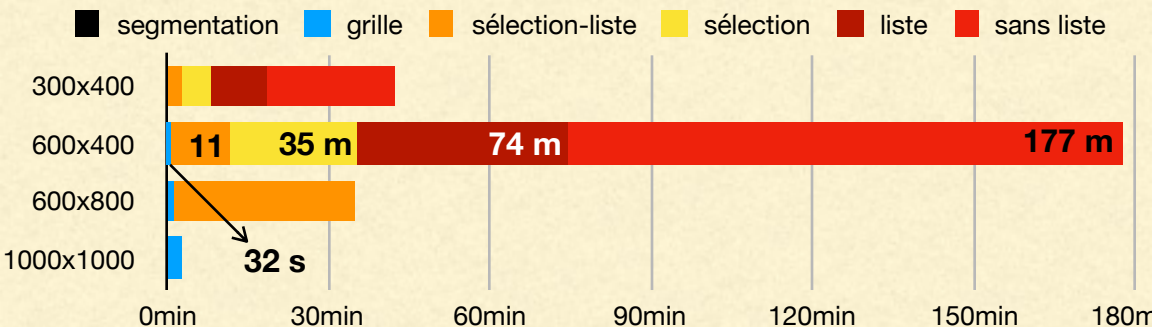
Recherche des plus proches voisins

- Examine les cellules dans une fenêtre ($3 \times 3 \times 3 \times 3 \times 3$)
- Examiner moins de groupes mais les plus pertinents



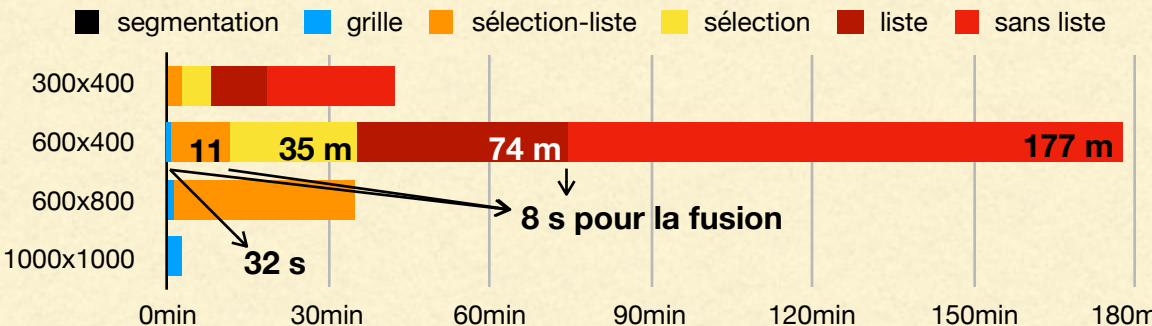
Temps de Calcul (CPU)

Taille de l'image	segmen- tation	grille	sélection + liste	sélection	liste	sans liste
300x400	0s 400ms	15s 200ms	2min 50s	8min 21s	18min 36s	42min 20s
600x400	0s 830ms	31s 500ms	11min 31s	35min 18s	74min 42s	177min 29s
600x800	1s 780ms	1min 7s 400ms	35min 3s			
1000x1000	3s 860ms	2min 31s 300ms				



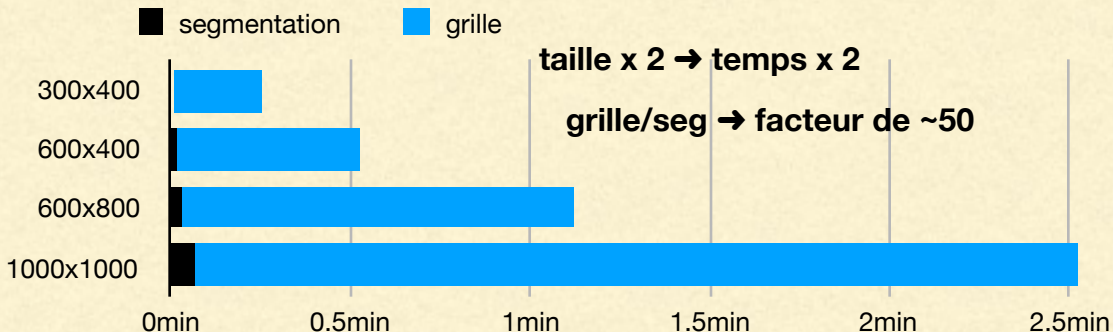
Temps de Calcul (CPU)

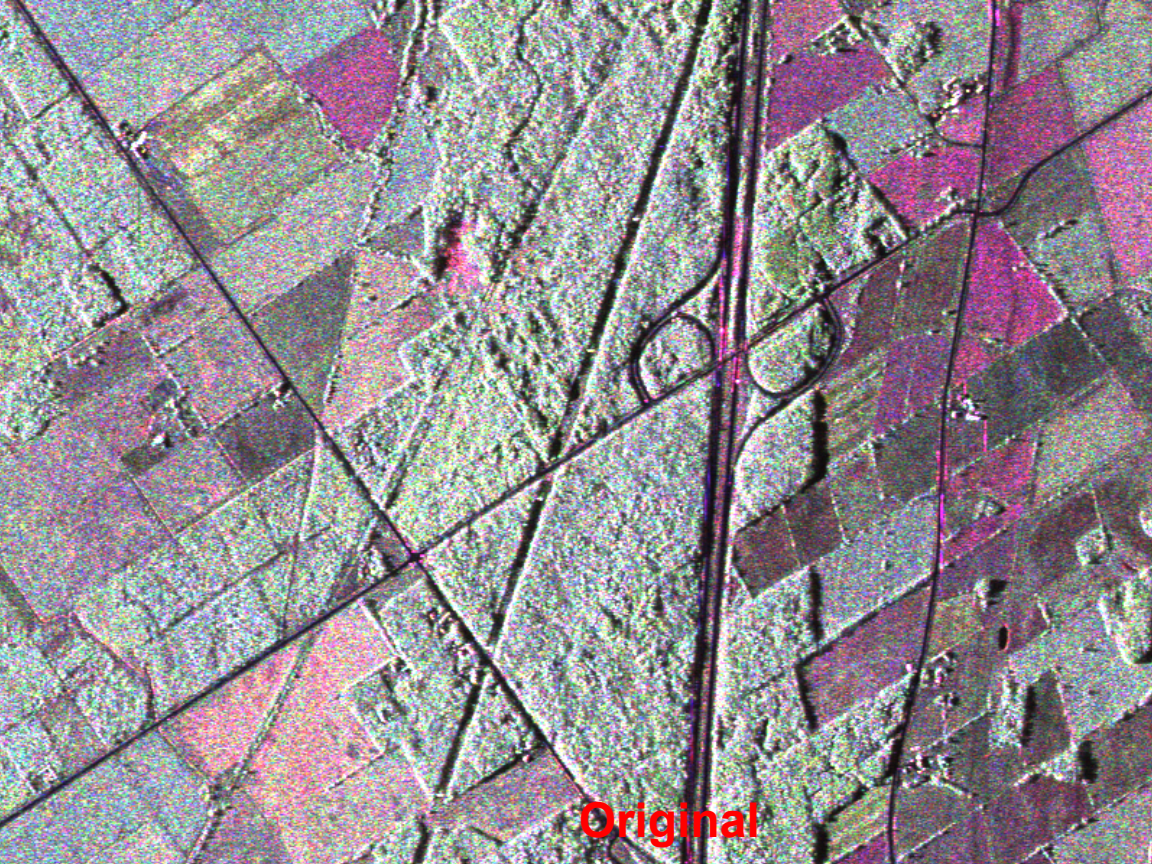
Taille de l'image	segmentation	grille	sélection + liste	sélection	liste	sans liste
300x400	0s 400ms	15s 200ms	2min 50s	8min 21s	18min 36s	42min 20s
600x400	0s 830ms	31s <u>500ms</u>	11min <u>31s</u>	35min 18s	<u>74min 42s</u>	177min 29s
600x800	1s 780ms	1min 7s 400ms	35min 3s			
1000x1000	3s 860ms	2min 31s 300ms	taille x 2 → temps x 4			



Temps de Calcul (CPU)

Taille de l'image	segmentation	grille	sélection + liste	sélection	liste	sans liste
300x400	0s 400ms	15s 200ms	2min 50s	8min 21s	18min 36s	42min 20s
600x400	0s 830ms	31s 500ms	11min 31s	35min 18s	74min 42s	177min 29s
600x800	1s 780ms	1min 7s 400ms	35min 3s			
1000x1000	3s 860ms	2min 31s 300ms				

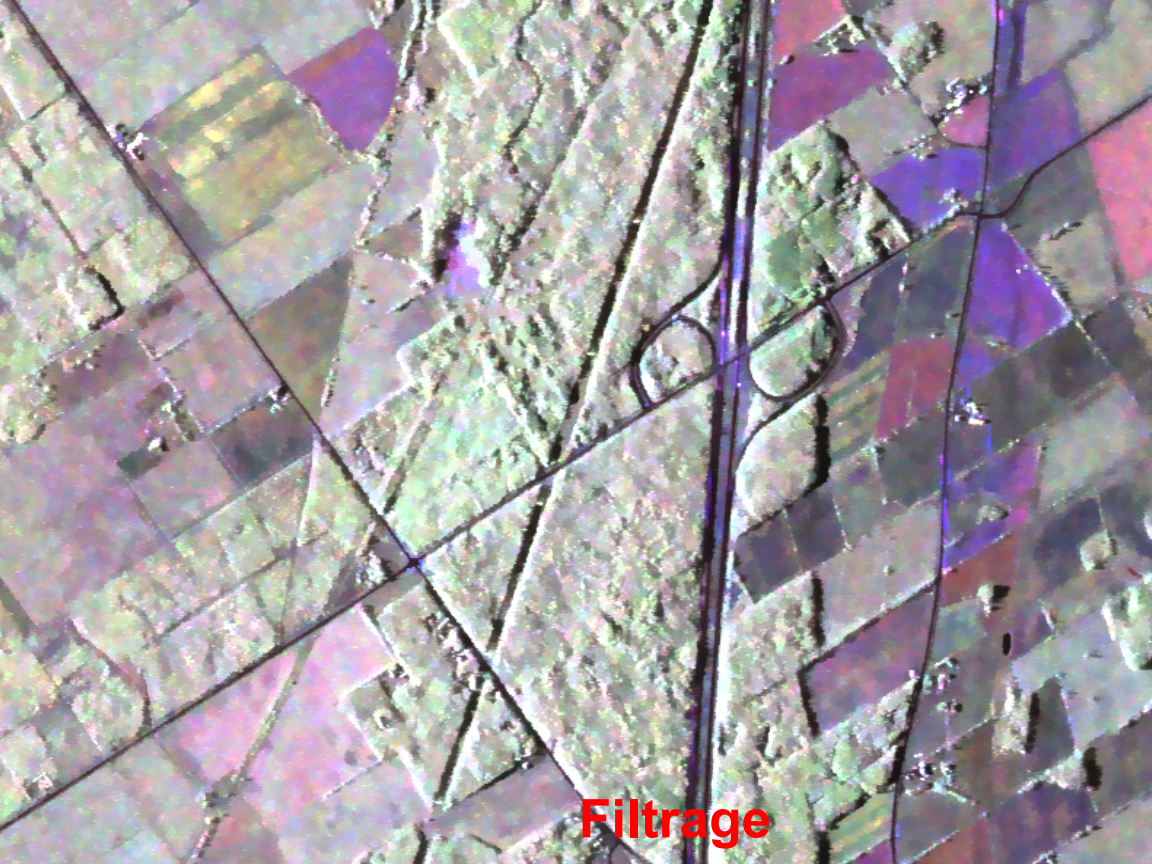




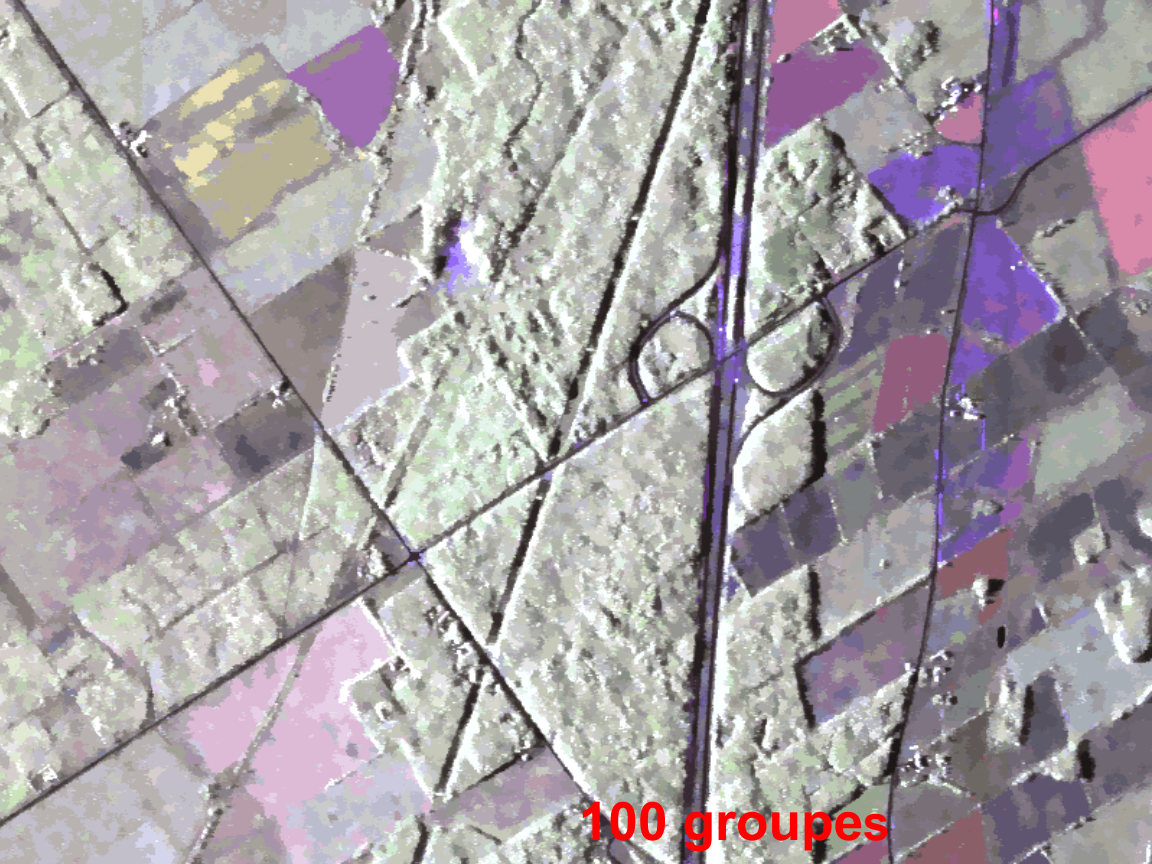
Original



100 groupes



Filtrage



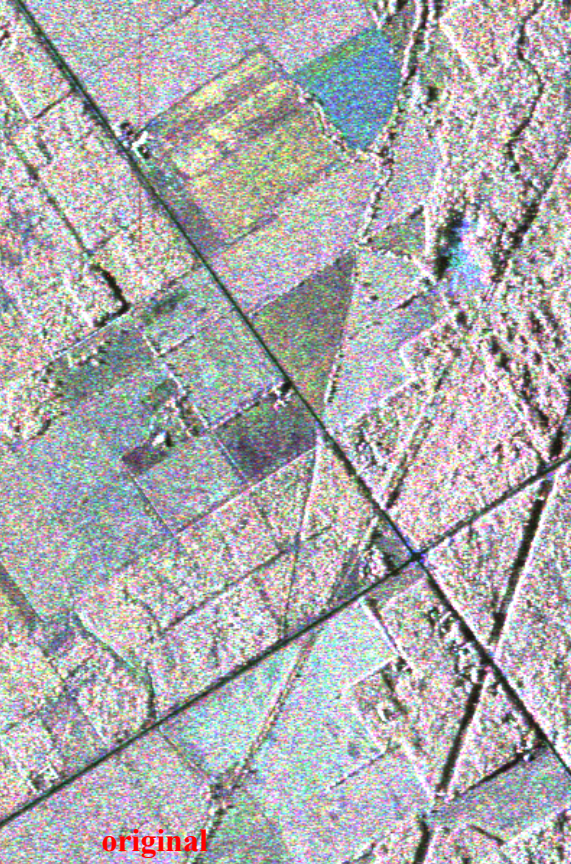
100 groupes

Conclusion

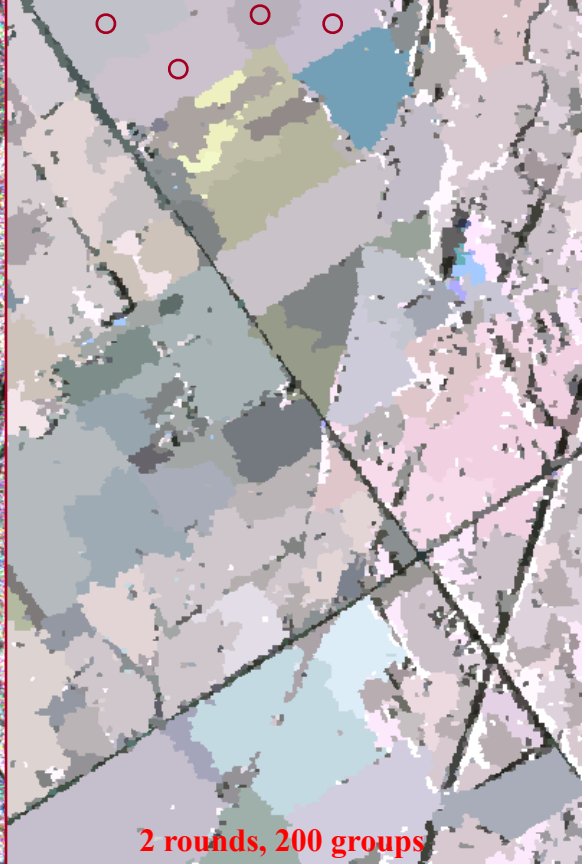
**Le regroupement hiérarchique était difficile
long temps de calcul**

**Devient un outils intéressant
avec une implantation rapide**

Déterminer quand et comment l'utiliser

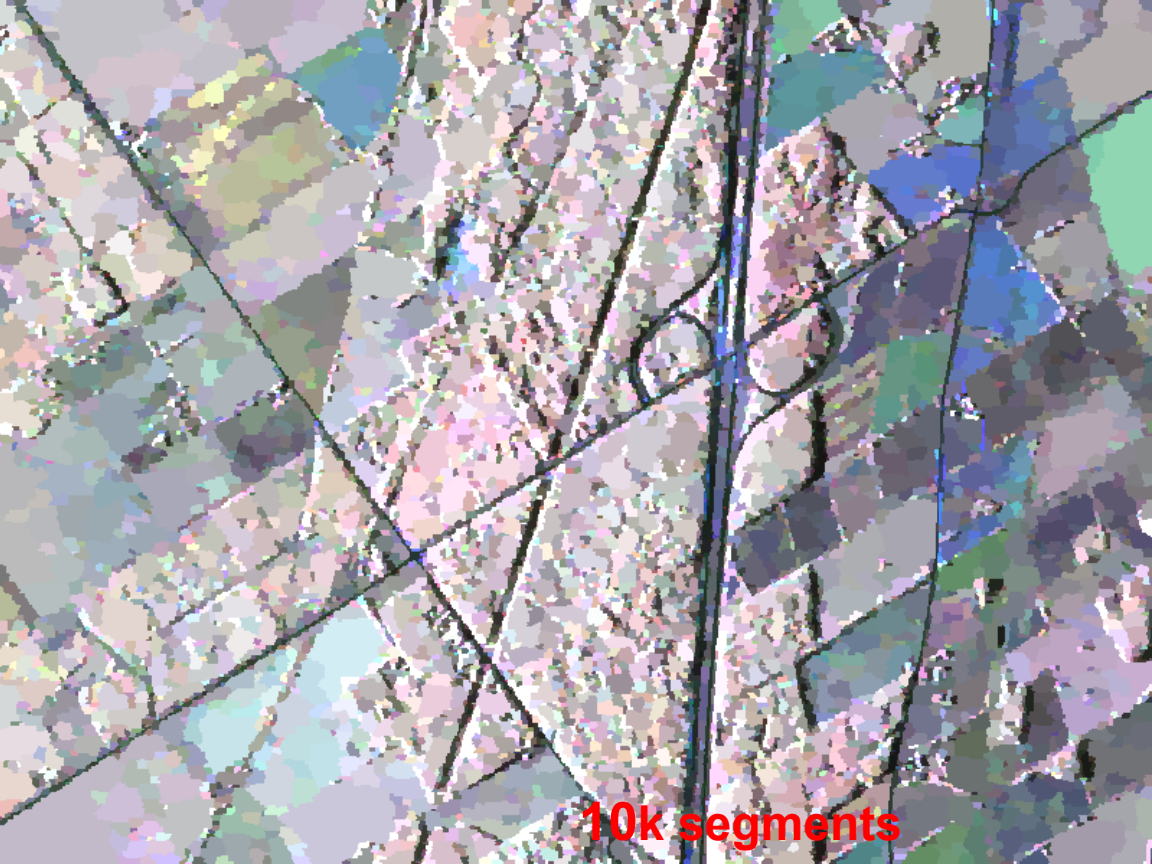


original

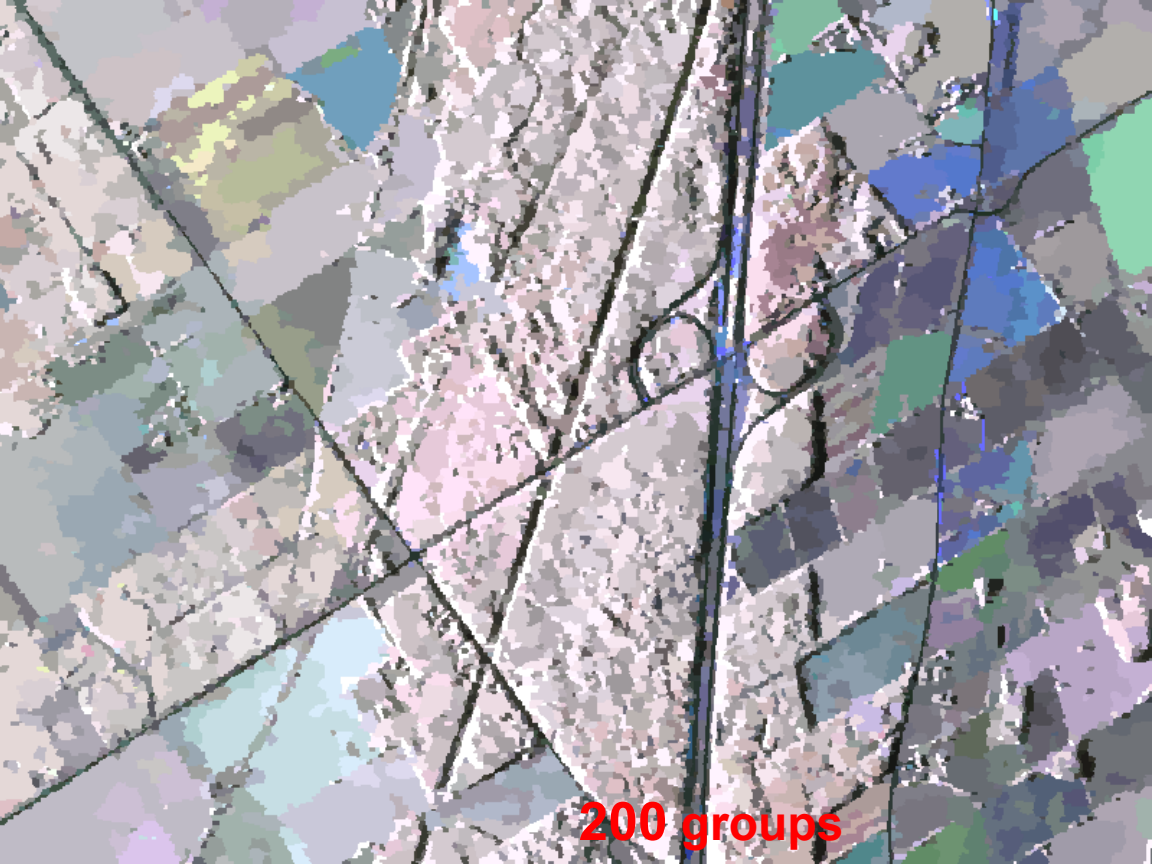


2 rounds, 200 groups

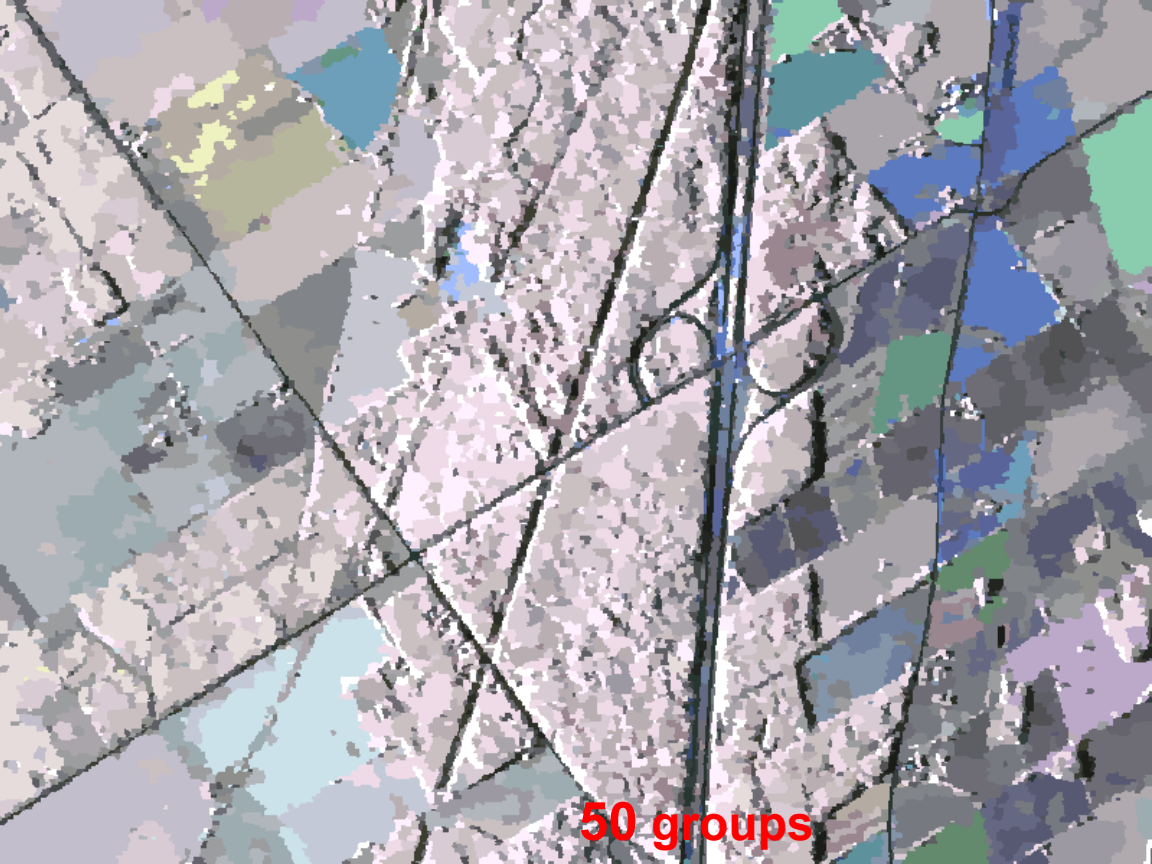




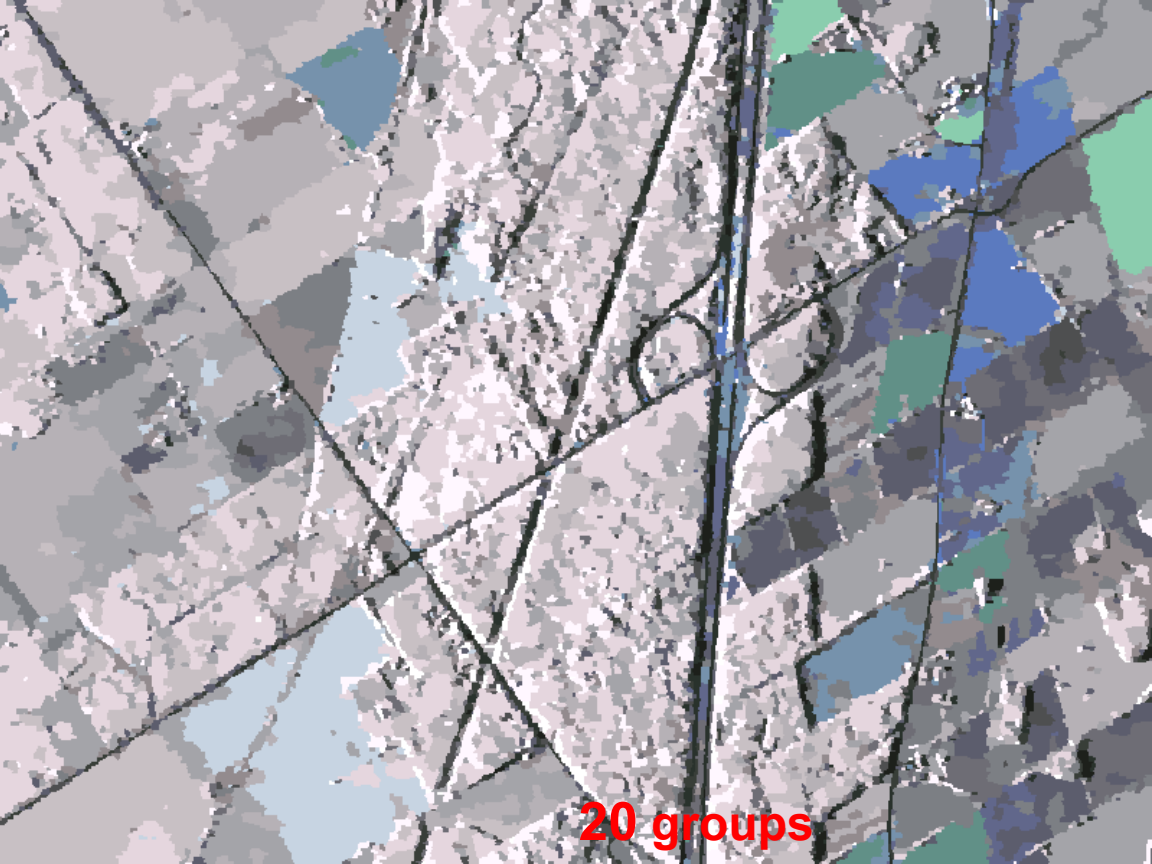
10k segments



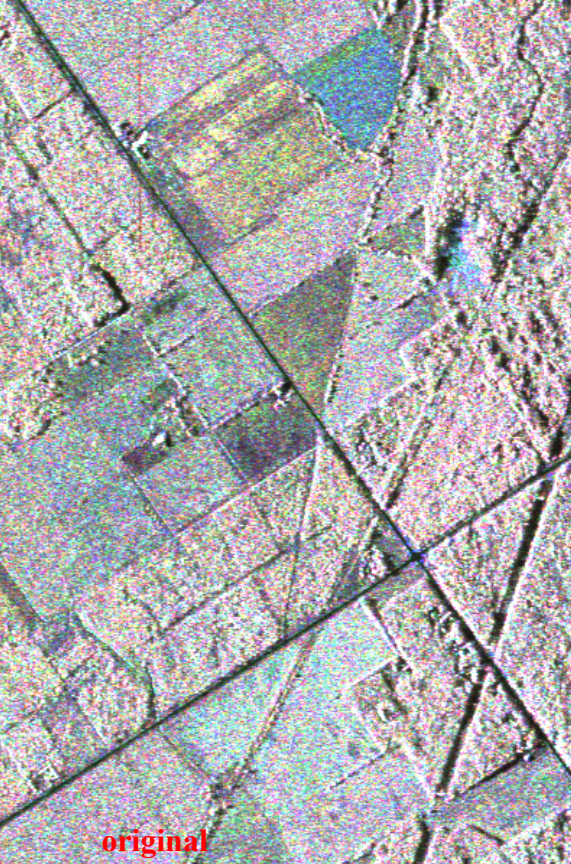
200 groups



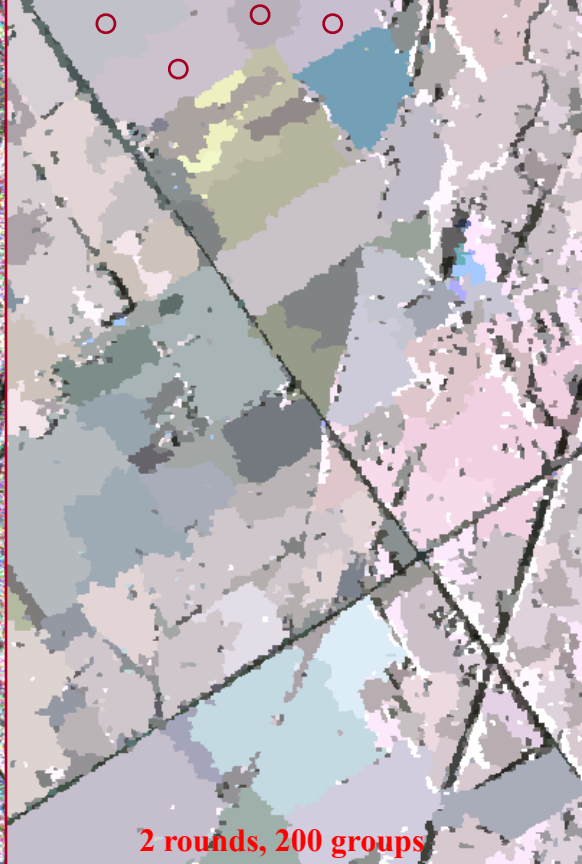
50 groups



20 groups



original



2 rounds, 200 groups

Conclusion

**Le regroupement hiérarchique était difficile
long temps de calcul**

**Deviens un outil intéressant
avec une implantation rapide**

Déterminer quand et comment l'utiliser
