# Algorithme de calcul de l'arbre des composantes avec applications à la reconnaissance des formes en imagerie satellitaire

Anthony Baillard<sup>1</sup> Christophe Berger<sup>2</sup> Emmanuel Bertin<sup>1</sup> Thierry Géraud<sup>2</sup> Roland Levillain<sup>2</sup> Nicolas Widynski<sup>2</sup>

<sup>1</sup>Institut d'Astrophysique de Paris (IAP)

<sup>2</sup>Laboratoire de Recherche et de Développement de l'EPITA (LRDE)

Conférence GRETSI, 12 septembre 2007



# Algorithme de calcul de l'arbre des composantes avec applications à la reconnaissance des formes en imagerie satellitaire

- Motivation
  - Filtres connectés
  - Le cas des images d'astronomie
- 2 Un nouvel algorithme de calcul de l'arbre des composantes
  - Calcul de l'arbre
  - Calcul d'attributs et marquage
  - Résultats et applications
- Conclusions et perspectives



- Finalité : application de filtres connectés de morphologie mathématique sur des images d'astronomie
- Caractéristiques des images d'astronomie traitées

- Besoin de nouveaux outils pour écrire ces filtres connectés, et notamment, l'arbre des composantes.
- Travail en collaboration avec l'IAP et le LRDE dans le cadre du projet EFIGI (Extraction de Formes Idéalisées de Galaxies en Imagerie)



- Finalité : application de filtres connectés de morphologie mathématique sur des images d'astronomie
- Caractéristiques des images d'astronomie traitées
  - volumineuses (ordre de grandeur : 100 Mo − 1,5 Go
  - codées sur des valeurs à virgule flottante
  - intervalle de dynamique élevé
- Besoin de nouveaux outils pour écrire ces filtres connectés, et notamment, l'arbre des composantes
- Travail en collaboration avec l'IAP et le LRDE dans le cadre du projet EFIGI (Extraction de Formes Idéalisées de Galaxies en Imagerie)



- Finalité : application de filtres connectés de morphologie mathématique sur des images d'astronomie
- Caractéristiques des images d'astronomie traitées
  - volumineuses (ordre de grandeur : 100 Mo 1,5 Go)
  - codées sur des valeurs à virgule flottante
  - intervalle de dynamique élevé
- Besoin de nouveaux outils pour écrire ces filtres connectés, et notamment, l'arbre des composantes
- Travail en collaboration avec l'IAP et le LRDE dans le cadre du projet EFIGI (Extraction de Formes Idéalisées de Galaxies en Imagerie)



- Finalité : application de filtres connectés de morphologie mathématique sur des images d'astronomie
- Caractéristiques des images d'astronomie traitées
  - volumineuses (ordre de grandeur : 100 Mo 1,5 Go)
    - codées sur des valeurs à virgule flottante
  - intervalle de dynamique élevé
- ⇒ Besoin de nouveaux outils pour écrire ces filtres connectés, et notamment, l'arbre des composantes.
- Travail en collaboration avec l'IAP et le LRDE dans le cadre du projet EFIGI (Extraction de Formes Idéalisées de Galaxies en Imagerie)



- Finalité : application de filtres connectés de morphologie mathématique sur des images d'astronomie
- Caractéristiques des images d'astronomie traitées
  - volumineuses (ordre de grandeur : 100 Mo − 1,5 Go)
  - codées sur des valeurs à virgule flottante
  - intervalle de dynamique élevé
- ⇒ Besoin de nouveaux outils pour écrire ces filtres connectés, et notamment, l'arbre des composantes.
- Travail en collaboration avec l'IAP et le LRDE dans le cadre du projet EFIGI (Extraction de Formes Idéalisées de Galaxies en Imagerie)



- Finalité : application de filtres connectés de morphologie mathématique sur des images d'astronomie
- Caractéristiques des images d'astronomie traitées
  - volumineuses (ordre de grandeur : 100 Mo − 1,5 Go)
  - codées sur des valeurs à virgule flottante
  - intervalle de dynamique élevé
- ⇒ Besoin de nouveaux outils pour écrire ces filtres connectés, et notamment, l'arbre des composantes.
  - Travail en collaboration avec l'IAP et le LRDE dans le cadre du projet EFIGI (Extraction de Formes Idéalisées de Galaxies en Imagerie)



- Finalité : application de filtres connectés de morphologie mathématique sur des images d'astronomie
- Caractéristiques des images d'astronomie traitées
  - volumineuses (ordre de grandeur : 100 Mo 1,5 Go)
  - codées sur des valeurs à virgule flottante
  - intervalle de dynamique élevé
- ⇒ Besoin de nouveaux outils pour écrire ces filtres connectés, et notamment, l'arbre des composantes.
  - Travail en collaboration avec l'IAP et le LRDE dans le cadre du projet EFIGI (Extraction de Formes Idéalisées de Galaxies en Imagerie)



# Arbre des composantes

- Représentation pratique et polyvalente d'une image



# Arbre des composantes

- Représentation pratique et polyvalente d'une image
- L'inclusion spatiale se traduit par une relation de parenté dans l'arbre



# Arbre des composantes

- Représentation pratique et polyvalente d'une image
- L'inclusion spatiale se traduit par une relation de parenté dans l'arbre
- Applications
  - Classification
  - Filtrage d'images
  - Segmentation
  - Recalage
  - Compression



GRETSI 2007

- Propriétés
  - Simplification de l'image
  - Préservation des contours
- Relations avec l'arbre des composantes



- Propriétés
  - Simplification de l'image
  - Préservation des contours
- Relations avec l'arbre des composantes



- Propriétés
  - Simplification de l'image
  - Préservation des contours
- Relations avec l'arbre des composantes
- Un filtre connecté peut s'exprimer comme une
  - transformation sur l'arbre des composantes n'ajoutant pass
  - de branche



- Propriétés
  - Simplification de l'image
  - Préservation des contours
- Relations avec l'arbre des composantes
  - Un filtre connecté peut s'exprimer comme une transformation sur l'arbre des composantes n'ajoutant pas de branche



- Propriétés
  - Simplification de l'image
  - Préservation des contours
- Relations avec l'arbre des composantes
  - Un filtre connecté peut s'exprimer comme une transformation sur l'arbre des composantes n'ajoutant pas de branche





- Image résultant d'une convolution par une fonction instrumentale (Point-Spread Function)
- Valeurs flottantes sur 32 bits
- Majorité des pixels : ciel (bruité en arrière-plan (sombres)
- Pixels clairs: objets (étoiles, galaxies) et artefacts visuels (halos, etc.)





- Image résultant d'une convolution par une fonction instrumentale (Point-Spread Function)
- Valeurs flottantes sur 32 bits
- Majorité des pixels : ciel (bruité en arrière-plan (sombres)
- Pixels clairs: objets (étoiles, galaxies) et artefacts visuels (halos, etc.)





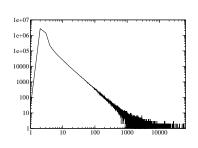
- Image résultant d'une convolution par une fonction instrumentale (Point-Spread Function)
- Valeurs flottantes sur 32 bits
- Majorité des pixels : ciel (bruité) en arrière-plan (sombres)
- Pixels clairs: objets (étoiles, galaxies) et artefacts visuels (halos, etc.)





- Image résultant d'une convolution par une fonction instrumentale (Point-Spread Function)
- Valeurs flottantes sur 32 bits
- Majorité des pixels : ciel (bruité) en arrière-plan (sombres)
- Pixels clairs: objets (étoiles, galaxies) et artefacts visuels (halos, etc.)

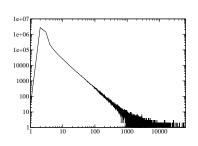
#### Histogramme (échelle log-log)



#### Quantification linéaire sur 16 bits



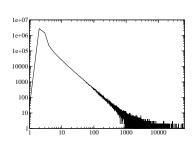
# Histogramme (échelle *log-log*)



- Quantification linéaire sur 16 bits
- Majorité des pixels entre 0 et 255
- Pixels entre 256 et 65 535 dus à la présence de flou
- ⇒ Nécessité d'une quantification optimale ; ou
- traitement direct sans quantification sur les valeur flottantes.



# Histogramme (échelle *log-log*)

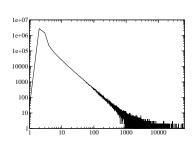


- Quantification linéaire sur 16 bits
- Majorité des pixels entre 0 et 255
- Pixels entre 256 et 65 535 dus à la présence de flou
- ⇒ Nécessité d'une quantification optimale ; ou
- ⇒ traitement direct sans quantification sur les valeur flottantes.



**GRETSI 2007** 

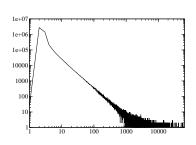
#### Histogramme (échelle log-log)



- Quantification linéaire sur 16 bits
- Majorité des pixels entre 0 et 255
- Pixels entre 256 et 65 535 dus à la présence de flou
- ⇒ Nécessité d'une quantification optimale; ou



# Histogramme (échelle *log-log*)



- Quantification linéaire sur 16 bits
- Majorité des pixels entre 0 et 255
- Pixels entre 256 et 65 535 dus à la présence de flou
- ⇒ Nécessité d'une quantification optimale ; ou
- ⇒ traitement direct sans quantification sur les valeur flottantes.

Calcul de l'arbre

# Aperçu

- S'appuie sur une variante de l'Union-Find [Tarjan, 1975]





- S'appuie sur une variante de l'Union-Find [Tarjan, 1975]
- Stratégie en trois étapes
  - Occupation d'un parcours des points basé sur une relation d'ordre  $\mathcal{R}$ , tel que

$$p \mathcal{R} q \Leftrightarrow \left\{ egin{array}{ll} f(p) &> f(q), \text{ ou} \\ f(p) &= f(q) \text{ et } p \text{ est situé avant } q \\ & \text{dans l'ordre de balayage usuel} \end{array} \right.$$

- Calcul effectif d'un arbre de composantes
- Ganonisation de l'arbre obtenu (compression)



GRETSI 2007

- S'appuie sur une variante de l'Union-Find [Tarjan, 1975]
- Stratégie en trois étapes
  - Oconstruction d'un parcours des points basé sur une relation d'ordre  $\mathcal{R}$ , tel que

$$p \mathcal{R} q \Leftrightarrow \left\{ egin{array}{ll} f(p) &>& f(q), \, ext{ou} \\ f(p) &=& f(q) \, ext{et } p \, ext{est situ\'e avant } q \\ & ext{dans l'ordre de balayage usuel} \end{array} 
ight.$$

- Calcul effectif d'un arbre de composantes
- Canonisation de l'arbre obtenu (compression)



- S'appuie sur une variante de l'Union-Find [Tarjan, 1975]
- Stratégie en trois étapes
  - Construction d'un parcours des points basé sur une relation d'ordre  $\mathcal{R}$ , tel que

$$p \mathcal{R} q \Leftrightarrow \left\{ egin{array}{ll} f(p) &>& f(q), \, ext{ou} \\ f(p) &=& f(q) \, ext{et } p \, ext{est situ\'e avant } q \\ & ext{dans l'ordre de balayage usuel} \end{array} 
ight.$$

- Calcul effectif d'un arbre de composantes





- S'appuie sur une variante de l'Union-Find [Tarjan, 1975]
- Stratégie en trois étapes
  - Occupance Construction d'un parcours des points basé sur une relation d'ordre  $\mathcal{R}$ , tel que

$$p \mathcal{R} q \Leftrightarrow \left\{ egin{array}{ll} f(p) &>& f(q), \, ext{ou} \\ f(p) &=& f(q) \, ext{et } p \, ext{est situ\'e avant } q \\ & ext{dans l'ordre de balayage usuel} \end{array} 
ight.$$

- Calcul effectif d'un arbre de composantes
- Canonisation de l'arbre obtenu (compression)

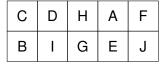


GRETSI 2007

# Calcul de l'arbre

# Arbre des composantes et *max-tree*

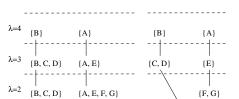




 $\mathcal{R}$ 

{F, G}

{H, I, J}





 $\lambda=1$  {A, B, C, D, E, F, G, H, I, J}

# Calcul de l'arbre

# Compacité et canonisation du max-tree

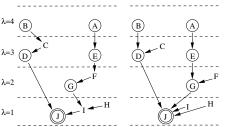


λ=2

 $\lambda = 1$ 



 $\mathcal{R}$ 







# Ensembles de niveaux de l'image

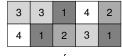




GRETSI 2007

# Calcul de l'arbre Calcul d'attributs et marquage Résultats et applications

# Ensembles de niveaux de l'image



ì

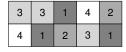
С	D	Н	Α	F
В	I	G	Е	J

 $\mathcal{R}$ 



# Calcul de l'arbre Calcul d'attributs et marquage Résultats et applications

# Ensembles de niveaux de l'image





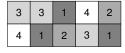
Niveau 4 ( $\lambda = 4$ )

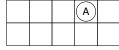
С	D	Н	Α	F
В	ı	G	Е	J

 $\mathcal{R}$ 



# Ensembles de niveaux de l'image





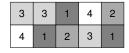
f

Niveau 4 ( $\lambda = 4$ )

С	D	Н	Α	F
В	ı	G	Е	J

 $\mathcal{R}$ 







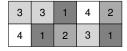
Niveau 4 ( $\lambda = 4$ )

С	D	Н	Α	F
В	Ι	G	Е	J

R



# Ensembles de niveaux de l'image



B Nivery 4 (1) 4)



1

Niveau 4 ( $\lambda = 4$ )

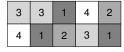
Niveau 3 ( $\lambda = 3$ )



R



## Ensembles de niveaux de l'image



(A)



1

Niveau 4 ( $\lambda = 4$ )

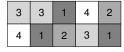
Niveau 3 ( $\lambda = 3$ )

С	D	Н	Α	F
В	1	G	Е	J

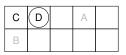
 $\mathcal{R}$ 



## Ensembles de niveaux de l'image



(A)



f

Niveau 4 ( $\lambda = 4$ )

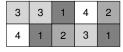
Niveau 3 ( $\lambda = 3$ )

С	D	Н	Α	F
В	ı	G	Е	J

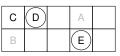
 $\mathcal{R}$ 



## Ensembles de niveaux de l'image



B



1

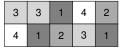
Niveau 4 ( $\lambda = 4$ )

Niveau 3 ( $\lambda = 3$ )



R

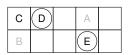




Niv

		A	
В			

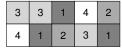
Niveau 4 ( $\lambda = 4$ )



Niveau 3 ( $\lambda = 3$ )

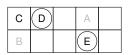
 $\mathcal{R}$ 





		A	
В			

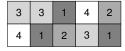
Niveau 4 ( $\lambda = 4$ )



Niveau 3 ( $\lambda = 3$ )

 $\mathcal{R}$ 

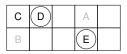




f

		A	
В			

Niveau 4 ( $\lambda = 4$ )

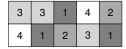


Niveau 3 ( $\lambda = 3$ )

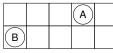
R



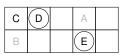
# Ensembles de niveaux de l'image



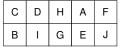
t



Niveau 4 ( $\lambda = 4$ )



Niveau 3 ( $\lambda = 3$ )



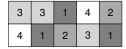
R



Niveau 2 ( $\lambda = 2$ )



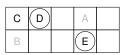
# Ensembles de niveaux de l'image



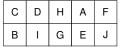
1

		(A)	
В			

Niveau 4 ( $\lambda = 4$ )



Niveau 3 ( $\lambda = 3$ )

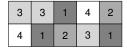


R



Niveau 2 ( $\lambda = 2$ )

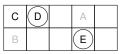




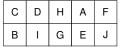
t

		$\bigcirc$	
В			

Niveau 4 ( $\lambda = 4$ )



Niveau 3 ( $\lambda = 3$ )

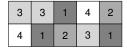


R



Niveau 2 ( $\lambda = 2$ )

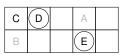




t

		(A)	
B			

Niveau 4 ( $\lambda = 4$ )



Niveau 3 ( $\lambda = 3$ )



R



Niveau 2 ( $\lambda = 2$ )



#### Construction

#### COMPUTE-TREE

```
FIND-ROOT(X)
    if zpar(x) = x then return x
2 else { zpar(x) \leftarrow FIND-ROOT(zpar(x)) ; return zpar(x) }
COMPUTE-TREE(f)
    for each p, zpar(p) \leftarrow undef
   R \leftarrow \text{REVERSE-SORT}(f) // traduit R comme un tableau
    for each p \in R in direct order
       parent(p) \leftarrow p; zpar(p) \leftarrow p
5
       for each n \in \mathcal{N}(p) such as zpar(n) \neq undef
6
         r \leftarrow \text{FIND-ROOT}(n)
         if r \neq p then { parent(r) \leftarrow p ; zpar(r) \leftarrow p }
    DEALLOCATE(zpar)
    return pair(R, parent) // une fonction « correcte »
9
```

#### Canonisation

**CANONIZE-TREE** 

```
CANONIZE-TREE(parent, f)

1 for each p \in R in reverse order

2 q \leftarrow parent(p)

3 if f(parent(q)) = f(q) then parent(p) \leftarrow parent(q)

4 return parent // une fonction « canonisée »
```



#### Exemple

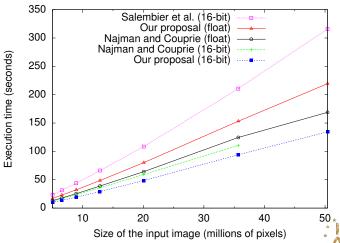
Calcul de l'aire des composantes.

```
COMPUTE-AREA(f, R, parent)
```

- 1 **for** each  $p \in R$ ,  $area(p) \leftarrow 1$  // initialisation
- 2 **for** each  $p \in R$  in direct order
- 3  $area(parent(p)) \leftarrow area(parent(p)) + area(p) // MAJ$ 
  - Processus simple
  - Calcul effectué de façon itérative
  - Complexité linéaire

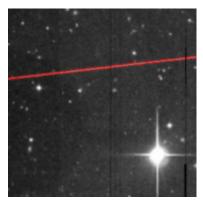


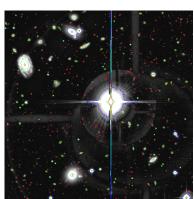
## Comparaison en temps



(Processeur Bi-Xeon à 3 Ghz, avec 2 × 1 Mo de mémoire cache et

# **Applications**









#### Conclusions

- Performant sur les images à valeurs hautement quantifiées ou sans quantification
- Comparable à l'algorithme connu le plus rapide, et Nécessite deux fois moins de mémoire
- Permet de construire des filtres connectés efficaces sur les images d'astronomie à traiter



# Algorithme de calcul de l'arbre des composantes avec applications à la reconnaissance des formes en imagerie satellitaire

- Motivation
  - Filtres connectés
  - Le cas des images d'astronomie
- Un nouvel algorithme de calcul de l'arbre des composantes
  - Calcul de l'arbre
  - Calcul d'attributs et marquage
  - Résultats et applications
- Conclusions et perspectives



### Bibliography I



Géraud, T. (2005).

Ruminations on Tarjan's Union-Find algorithm and connected operators.

In Mathematical Morphology: Proc. of the Intl. Symp. (ISMM), pages 105–116. Springer.



Jones, R. (1997).

Component trees for image filtering and segmentation.

In Coyle, E., editor, *IEEE Workshop on Nonlinear Signal and Image Processing*, Mackinac Island.



#### Bibliography II



Meijster, A. (2004).

Efficient sequential and parallel algorithms for morphological image processing.

PhD thesis, University of Groningen (NL).



Meijster, A. and Wilkinson, M. H. F. (2002).

A comparison of algorithms for connected set openings and closings.

IEEE Trans. Pattern Anal. Machine Intell., 24(4):484-494.

Calcul de l'arbre des composantes et imagerie satellitaire



Najman, L. and Couprie, M. (2006).

Building the component tree in quasi-linear time.

IEEE Trans. Image Processing, 15(11):3531-3539.



20 / 21

#### Bibliography III



Antiextensive connected operators for image and sequence processing.

IEEE Trans. Image Processing, 7(4):555–570.

Tarjan, R. E. (1975).

Efficiency of a good but not linear set union algorithm. *Journal of the ACM*, 22(2):215–225.

Vincent, L. (1993).

Grayscale area openings and closings: their applications and efficient implementation.

In Intl. Symp. on Mathematical Morphology, pages 22-27,

