

Indexing big colored image bank : Texture 3.0

**Etienne CAILLAUD, Thomas LE BRIS, Ibrahima GUEYE,
Gaetan ADIER**



XLIM-SIC Laboratory UMR CNRS 7252, Poitiers, France



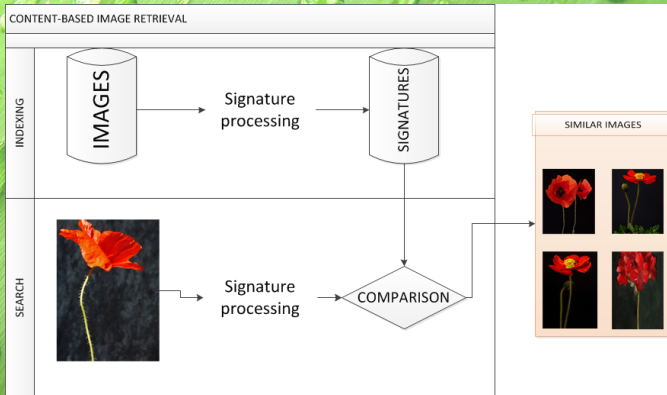
Outline

- 1 Introduction to the project context
- 2 Work and results
- 3 Project management
- 4 Conclusion

Project context (1/3)

Objective

Test a solution for content based image indexing flaw : standard descriptors (SIFT, SURF, etc) lacking real color and texture information.



Project context 2/3

What is a descriptor ?

Algorithm applied to an image which output is a short vector of numbers which is invariant to common image transformations and can be compared with other descriptors in a database.

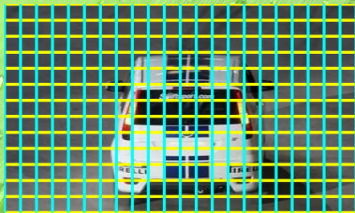


FIGURE: Densegrid

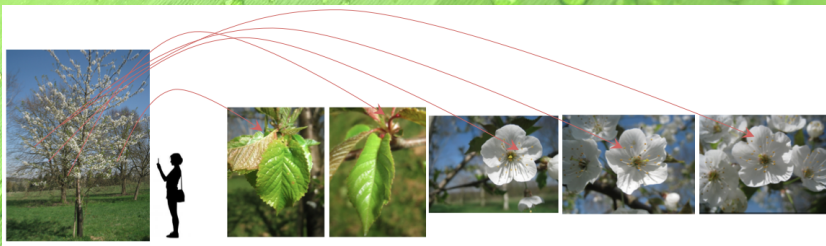


FIGURE: Interest points

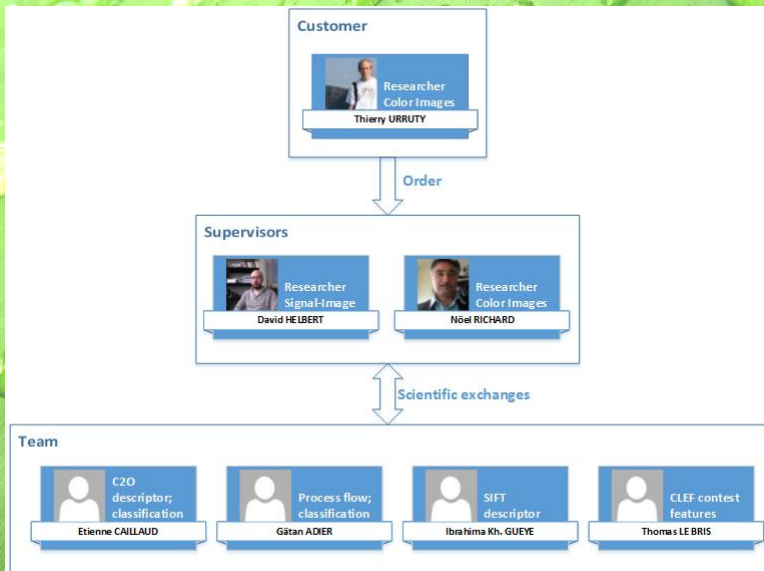
Project context 3/3

What is a CLEF ?

International contest organized every year since 2011 which purpose is to benchmark the progress in the area of plant identification from images.



Team presentation



User requirement

- Design software programs :
indexation of images database, calculate descriptor
according to nature images
- Adapt the last up to date designed color and texture
attributes to the current image classification
- Compare our results (using CLEF challenge metrics)
- Provide an abstract of the comparisons and a technical
report

Outline

- 1 Introduction to the project context
- 2 Work and results**
- 3 Project management
- 4 Conclusion

SIFT(1/2)

Key-points detection (x, y, σ)

- Scale-space extrema detection
- Key-point location
- Orientation assignment
- key-point descriptor

SIFT(2/2)

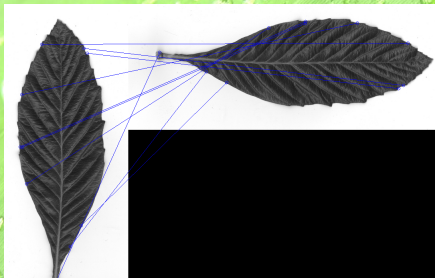


FIGURE: SIFT test1

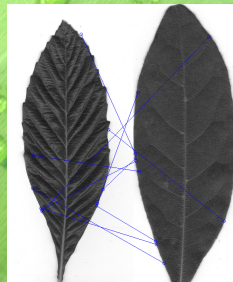


FIGURE: SIFT test2

C₂O (1/4)

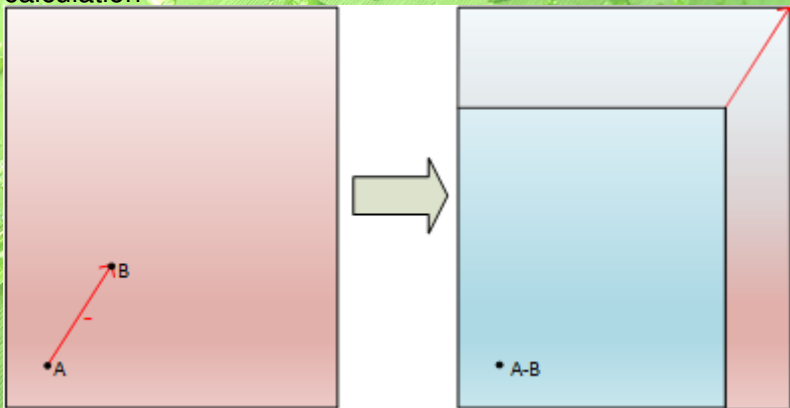
- Limitation of marginal approach
- Necessity to get a vectorial treatment
- Include better texture and color informations

C₂O (2/4)

- Conversion to a perceptual space (adapted to human perception).
- C₂O matrix calculation.
- C₂O signature extraction.

C₂O (3/4)

- Computation of the C₂O matrix by the color difference calculation



C_2O (3/4)

- The C_2O matrix for a poorly textured image :



FIGURE: Image to characterize

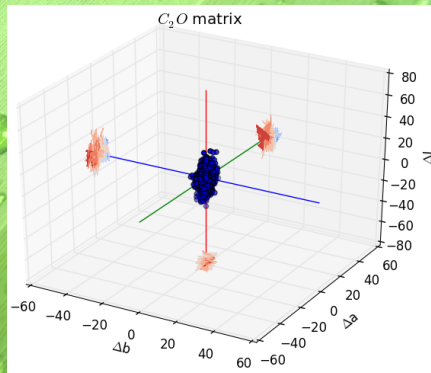


FIGURE: Signature

C_2O (3/4)

- The C_2O matrix for a poorly textured image :
- The C_2O matrix for a more textured image :



FIGURE: Image to characterize

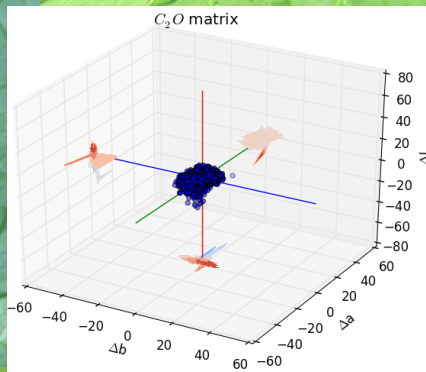


FIGURE: Signature

C₂O (3/4)

- The C₂O matrix for a poorly textured image :
- The C₂O matrix for a more textured image :
- The C₂O matrix for a more textured and colored image :



FIGURE: Image to characterize

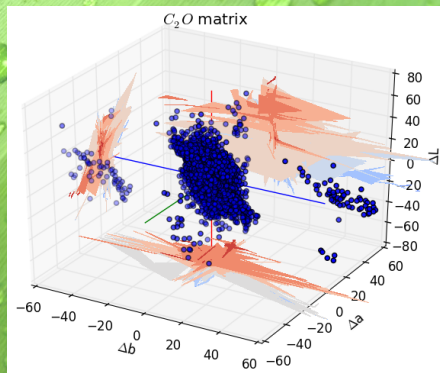
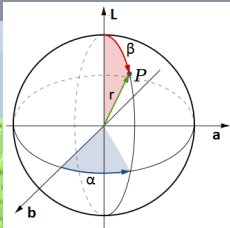
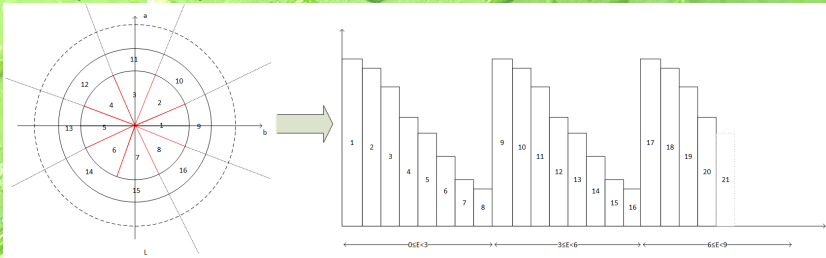


FIGURE: Signature

C₂O (4/4)

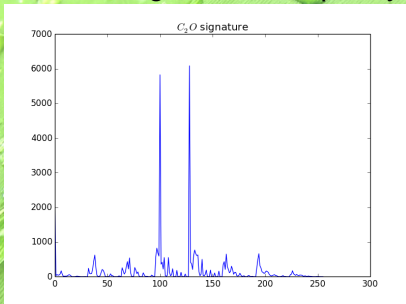
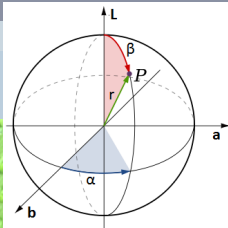


- The spherical quantization :



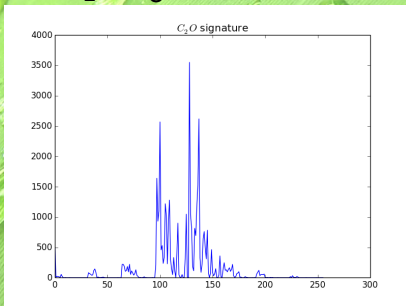
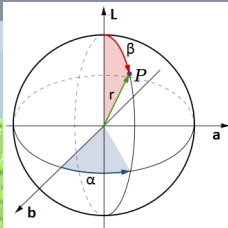
C₂O (4/4)

- The spherical quantization :
- The C₂O signature for a poorly textured image :

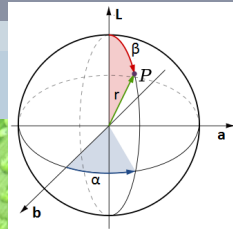


C₂O (4/4)

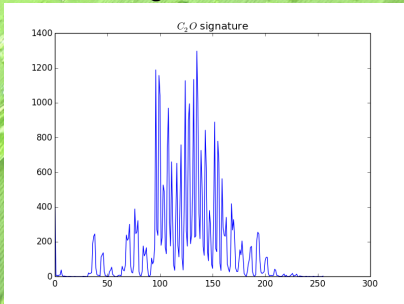
- The spherical quantization :
- The C₂O signature for a poorly textured image :
- The C₂O signature for a more textured image :



C₂O (4/4)



- The spherical quantization :
- The C₂O signature for a poorly textured image :
- The C₂O signature for a more textured image :
- The C₂O signature for a more textured and colored image :



Bag of word (1/2)

Reducing the number of points (100 in our case).

- K-means
 - Attribute the vectors to centroid vectors.

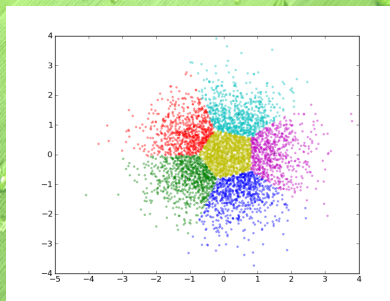


FIGURE: K-means

Bag of word (2/2)

- Signature
 - Design histogram in function of assignment of the vectors.

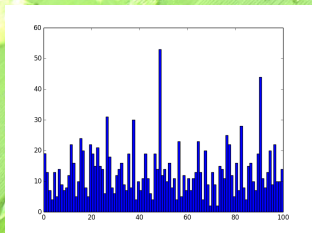


FIGURE: Signature 100 words -
1

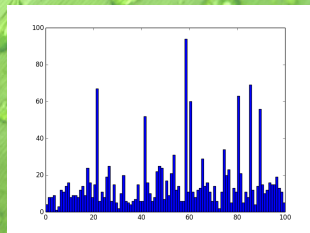
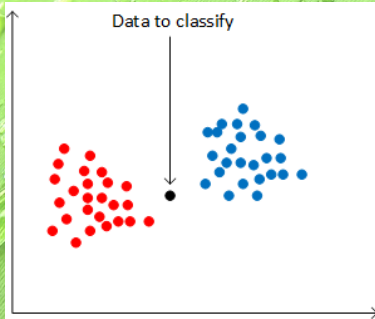


FIGURE: Signature 100 words -
2

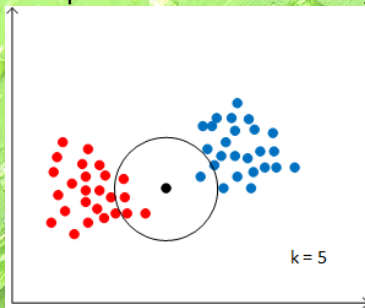
K-nn(1/2)

- The k nearest neighbor method
 - Comparison to the dictionary .



K-nn(1/2)

- The k nearest neighbor method
 - Comparison to the dictionary :



- 4 Occurrences of the 'red' class , - 1 occurrence of the 'blue' class
- The new point is attributed to the 'red' class

K-nn(2/2)

- Application for image classification
 - More complex data.
 - Distances on signature vectors extracted from the K-mean method.
 - One most adapted distance type for each descriptor .

Results (1/2)

- Reduce data-base of 100 images composed of only 4 species.



FIGURE: First specie



FIGURE: Second specie



FIGURE: Third specie



FIGURE: Fourth specie

Results (2/2)

- Compare the two descriptors SIFT and C₂O.

TABLE: SIFT result

ID	Training Base	Test Base	Correct	Accuracy
173	17	8	4	50%
1102	22	3	1	33%
1889	16	9	1	11%
2717	15	10	7	70%
Total	70	30	9	/

TABLE: C₂O result

ID	Training Base	Test Base	Correct	Accuracy
173	17	8	1	12.5%
1102	22	3	1	33%
1889	16	9	0	0%
2717	15	10	7	70%
Total	70	30	9	/

Discussion

- Classification

- To much reducing on the K-means (100 words).
- Euclidean distance not the most efficient or adapt.

- C_2O

- The concatenation way is not optimal.
- Parameters D, alpha, and beta has to be discussed regarding to the images.

Outline

- 1 Introduction to the project context
- 2 Work and results
- 3 Project management**
- 4 Conclusion

Scheduling (1/2)

- The previsionsal forecast Gantt chart :

ID	Task Name	Start	Finish	Duration	mai 2015							juin 2015	
					26/4	3/5	10/5	17/5	24/5	31/5	7/6		
1	Writing the state of the art	27/04/2015	01/05/2015	1w									
2	Preparing the database	01/05/2015	14/05/2015	2w									
3	Constitution of database structure	01/05/2015	04/05/2015	2d									
4	Calculation of SIFT descriptors	04/05/2015	08/05/2015	1w									
5	Calculation of SURF descriptors	08/05/2015	14/05/2015	1w									
6	Programming	14/05/2015	29/05/2015	2w 2d									
7	Programming C2O descriptor calculation	14/05/2015	20/05/2015	1w									
8	Programming the distance calculation function	20/05/2015	26/05/2015	1w									
9	Writing of the technical documentation	26/05/2015	29/05/2015	4d									
10	Writing of the report	29/05/2015	09/06/2015	1w 3d									
11	Preparation of the oral presentation	09/06/2015	15/06/2015	1w									

- All time affectation done before the beginning of the project
- Rarely respected in important project

Scheduling (2/2)

- The project backlog :

SPRINT 5 BACKLOG PRODUIT										
Sprint	Catégorie	Sous catégorie	Nom / Description	Importance	Estimation	Critères de Vérification	Acteur	Status	Notes	Bug
5	Dev Logiciel	Test	C2o - HULK	75	??			A faire		
5	Dev Logiciel	Test	SFT - calcul sur HULK	75	1,5		Ibrahima	A faire	Présent sprint 2	
5	Dev Logiciel	Recherche documentaire...	classification - doc	55	2			A faire		
5	Dev Logiciel	Redaction documentation	CLEF metrics - doc	65	0,5	présentation équipe scientifique	Thomas	A faire		
5	Dev Logiciel	Redaction documentation	documentation sur le processus "complet"	60	1	présentation à l'ensemble des acteurs du projet		A faire		
5	Dev Logiciel	Redaction documentation	SFT - doc	50	0,5	présentation équipe scientifique		A faire		
5	Présentation		Présentation - ecriture	40	2	présentation équipe scientifique		A faire		
5	Rapport		Ecriture du document final - synthèse des docs	40	2	présentation équipe scientifique		A faire		
5	Présentation		Présentation - préparation	30	2	présentation à l'équipe pédagogique		A faire		
5	Dev Logiciel		Procédure de validation	60	2					
5	Dev Logiciel	Redaction documentation	Analyse des résultats	75	2			A faire		
Total				15,5						

- Allow to change the affectation of a task
- Weekly time affectation : could be adapted to unforeseen

Our experience

- Minimal lack of time
- The possibility of changing task affectation is really useful
- An adaptation of the initial schedule has been realised

Outline

- 1 Introduction to the project context
- 2 Work and results
- 3 Project management
- 4 Conclusion

Sum-up of the situation

Starting objectives

- SIFT tests
- C2O programming
- classification programming
- Code optimizing for speed
- parallelization

Ending situation

- SIFT tests
- C2O programming
- classification programming

Issues

- C2O concatenation order
- distance calculation

Personal gains

- New way to organize teamwork
- Technical knowledge



Thanks for attention