

# Indexing big colored image bank : Texture 3.0

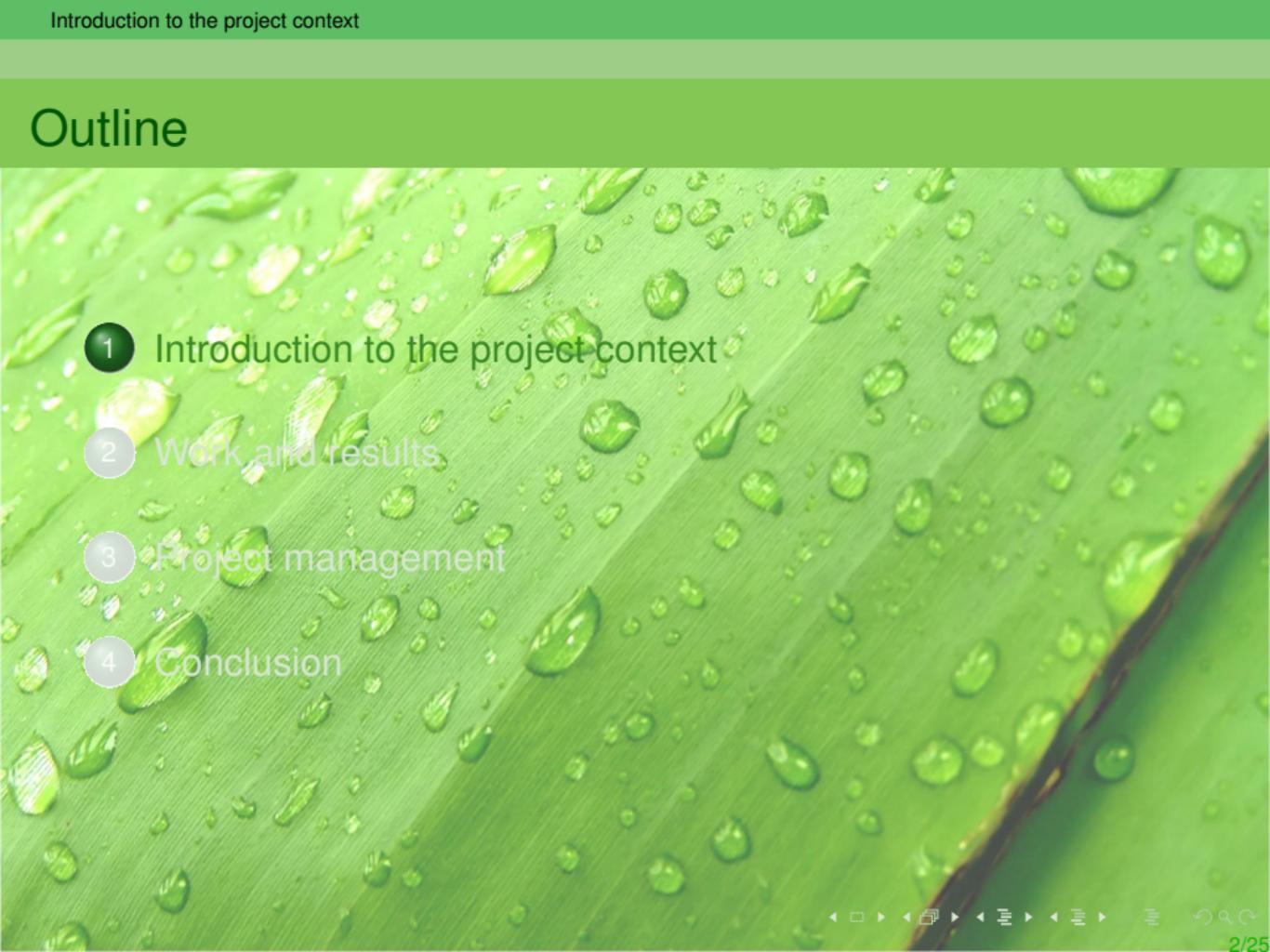
**Etienne CAILLAUD, Thomas LE BRIS, Ibrahima GUEYE,  
Gaetan ADIER**



XLIM-SIC Laboratory UMR CNRS 7252, Poitiers, France



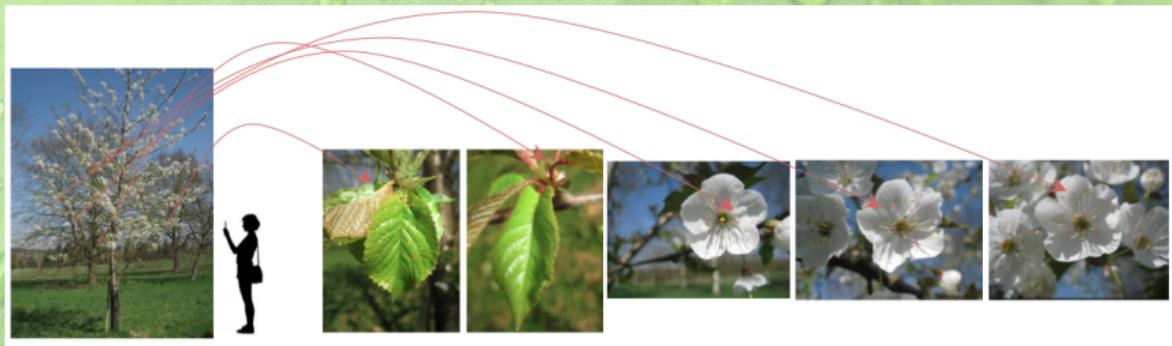
# Outline

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

# Project context (1/2)

## What is a CLEF ?

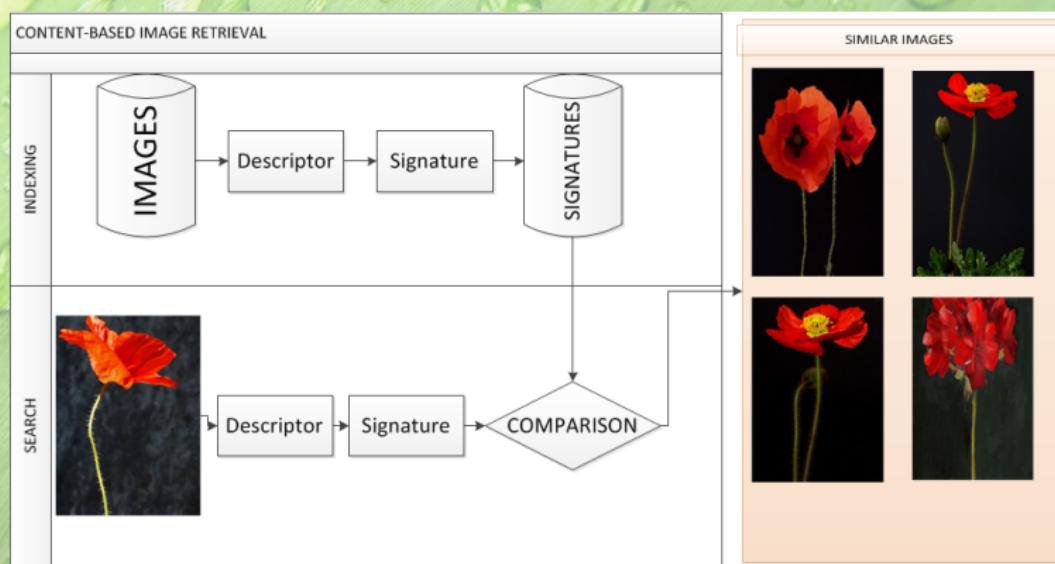
International contest organized every year since 2011 which purpose is to benchmark the progress in the area of plant identification from images.



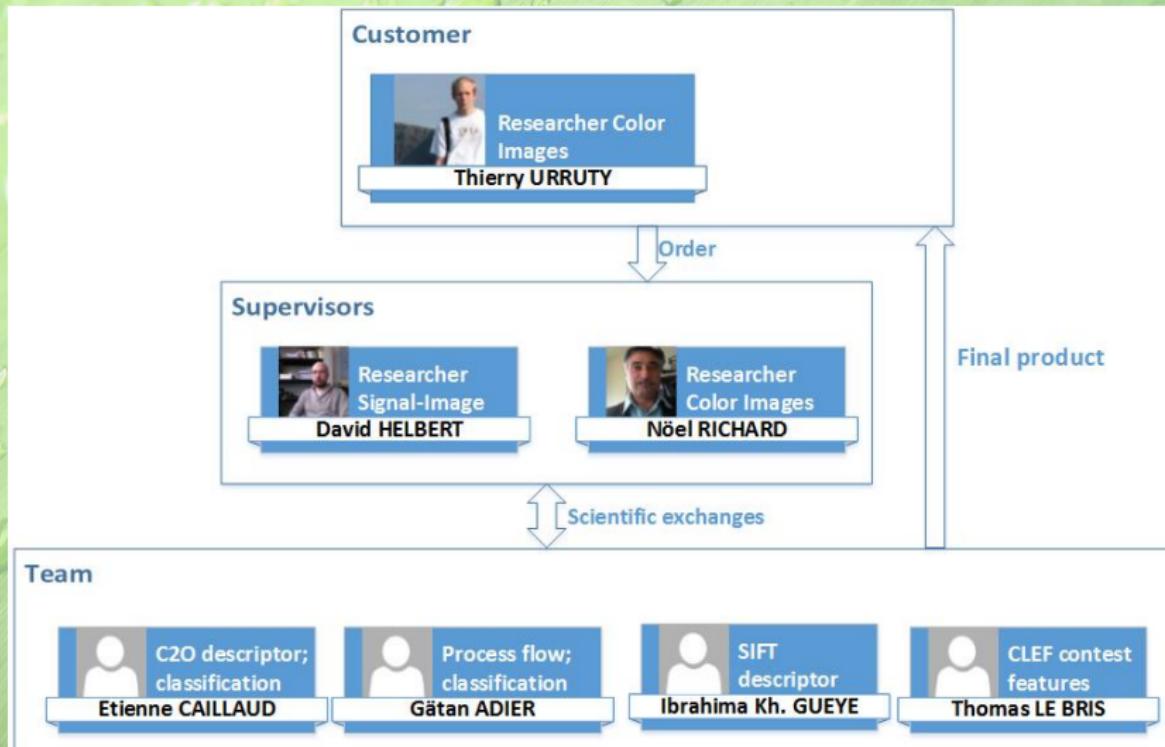
# Project context (2/2)

## Objective

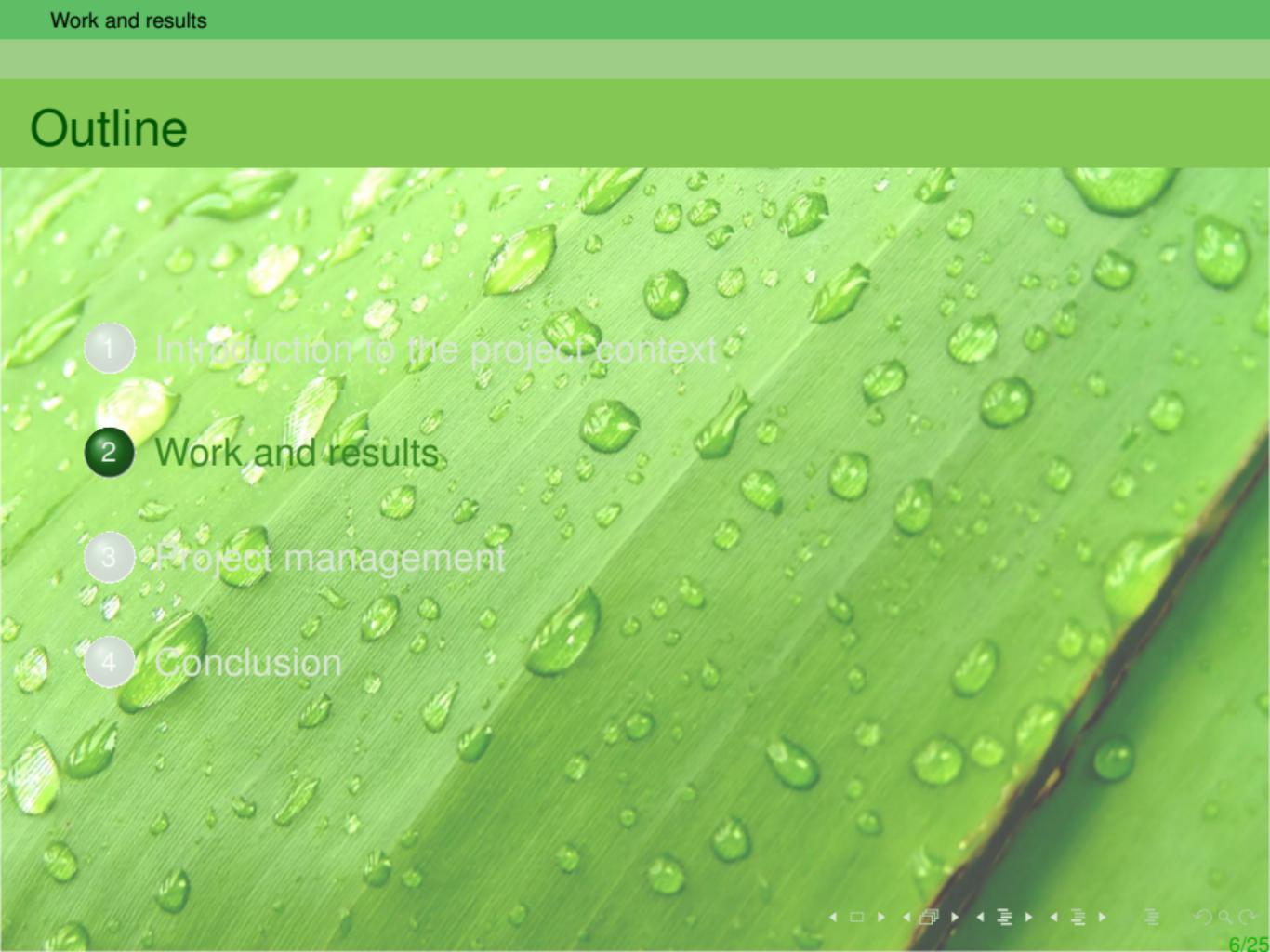
Test a solution for content based image indexing flaw : standard descriptors (SIFT, SURF, etc) lacking real color and texture information.



# Team presentation



# Outline

- 
- A close-up photograph of a bright green leaf with numerous small, clear water droplets scattered across its surface. The leaf has distinct veins running through it.
- 1 Introduction to the project context
  - 2 Work and results
  - 3 Project management
  - 4 Conclusion

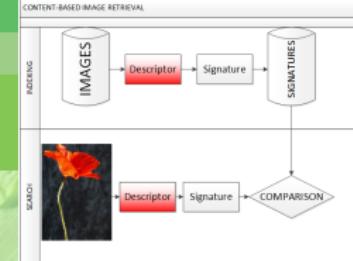
# SIFT(1/2)

Key-points detection ( $x, y, \sigma$ )

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



FIGURE: SIFT test2



## SIFT(2/2)

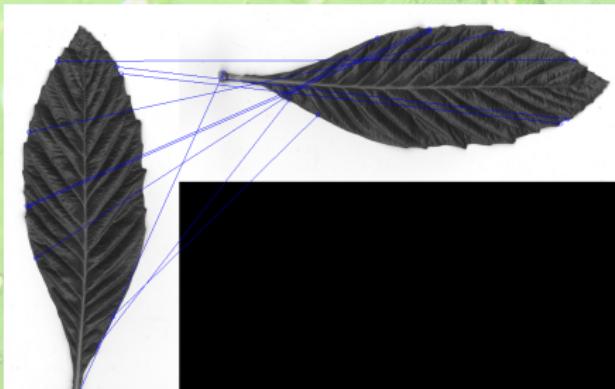
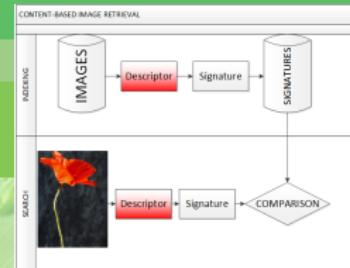


FIGURE: SIFT test1

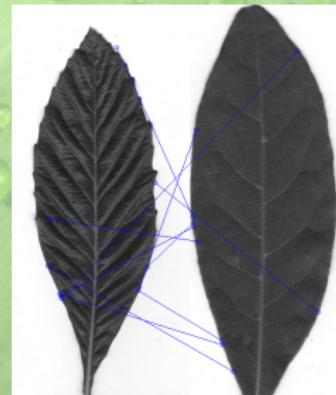
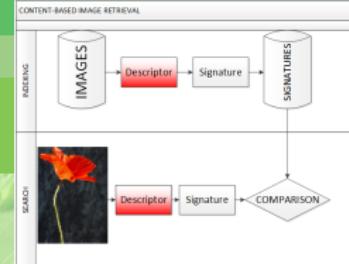


FIGURE: SIFT test2

# What about nature images ?



## SIFT

- Description using orientation of shapes
- Natively used on grayscale images
- Only marginal methods for color images
- Unable to get the texture information from image

## C<sub>2</sub>O

- Description based on color difference
- Natively conceived for color images
- Take account of the texture information

# C<sub>2</sub>O (1/2)

- The C<sub>2</sub>O matrix for a poorly textured image :



FIGURE: Image to characterize

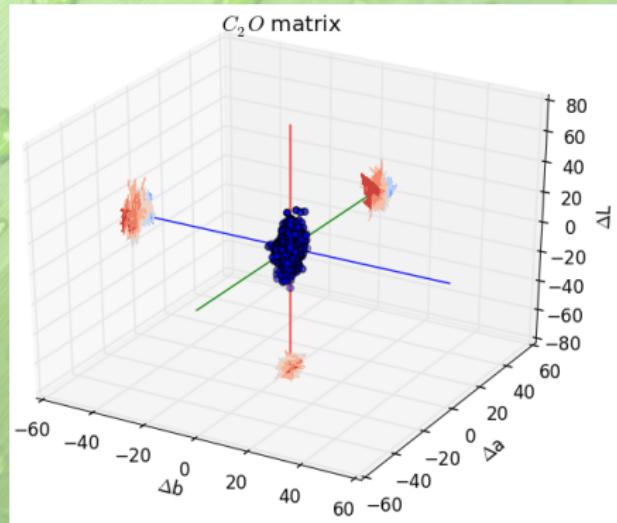


FIGURE: Signature

# C<sub>2</sub>O (1/2)

- The C<sub>2</sub>O matrix for a poorly textured image :
- The C<sub>2</sub>O matrix for a more textured image :



FIGURE: Image to characterize

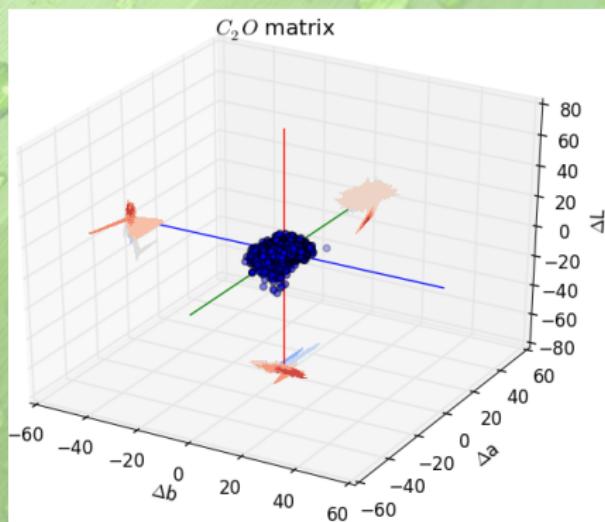
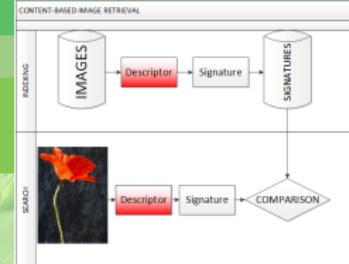


FIGURE: Signature



# C<sub>2</sub>O (1/2)

- The C<sub>2</sub>O matrix for a poorly textured image :
- The C<sub>2</sub>O matrix for a more textured image :
- The C<sub>2</sub>O matrix for a more textured and colored image :



FIGURE: Image to characterize

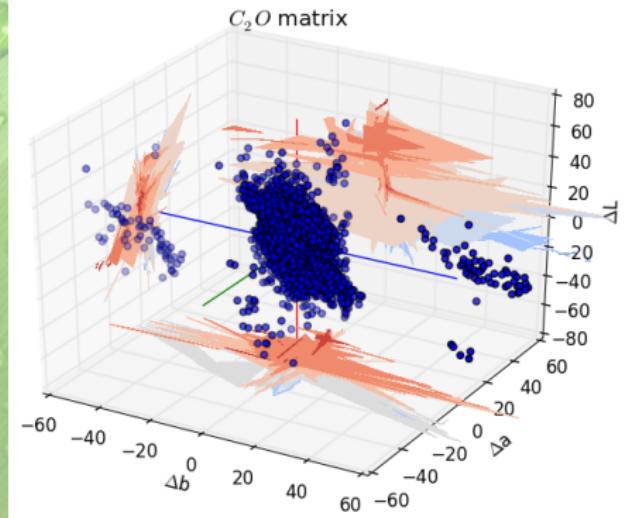
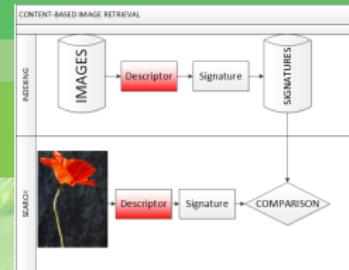
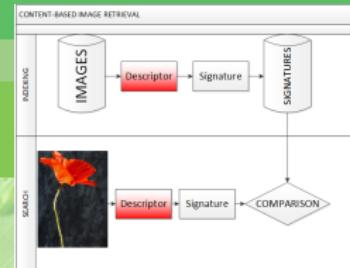


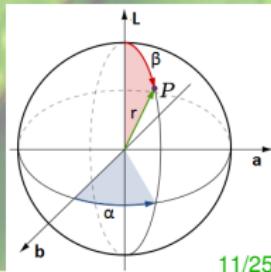
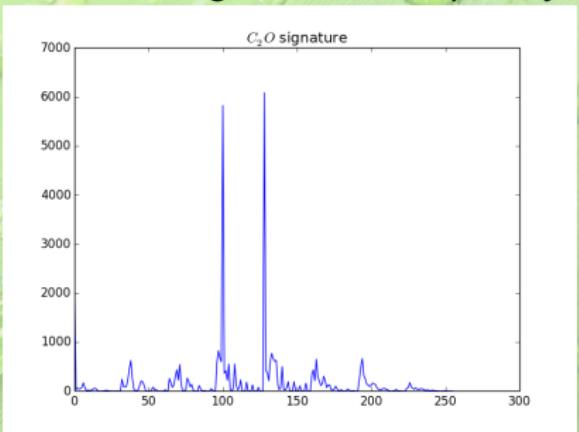
FIGURE: Signature



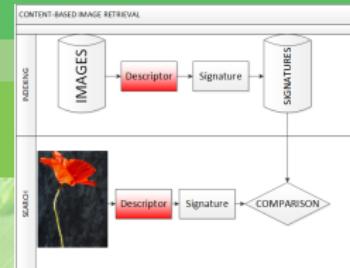
# C<sub>2</sub>O (2/2)



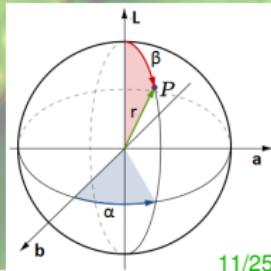
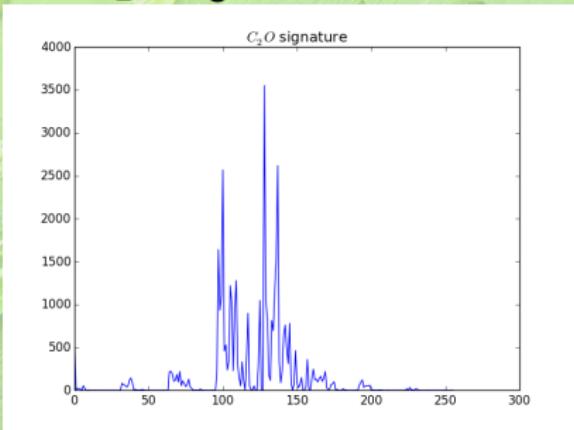
- The C<sub>2</sub>O signature for a poorly textured image :



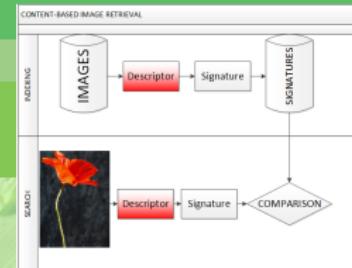
# C<sub>2</sub>O (2/2)



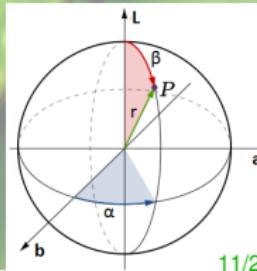
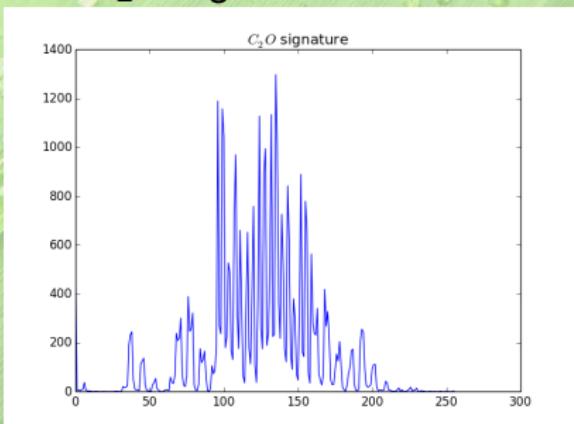
- The C<sub>2</sub>O signature for a poorly textured image :
- The C<sub>2</sub>O signature for a more textured image :



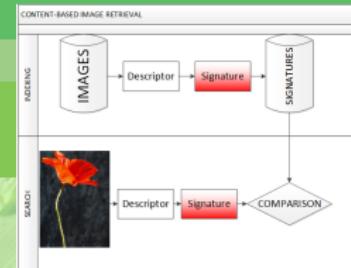
# C<sub>2</sub>O (2/2)



- The C<sub>2</sub>O signature for a poorly textured image :
- The C<sub>2</sub>O signature for a more textured image :
- The C<sub>2</sub>O 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.

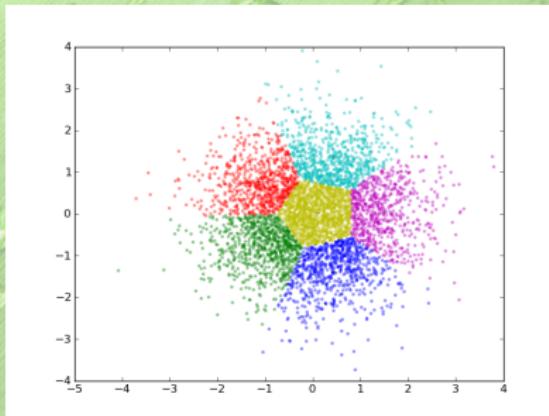
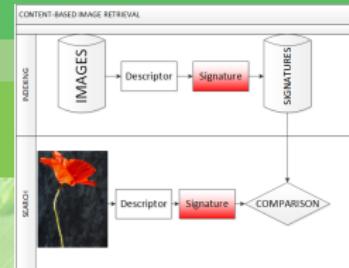


FIGURE: K-means

# Bag of word (2/2)



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

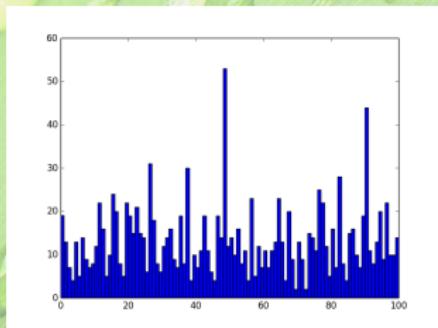


FIGURE: Signature 100 words -  
1

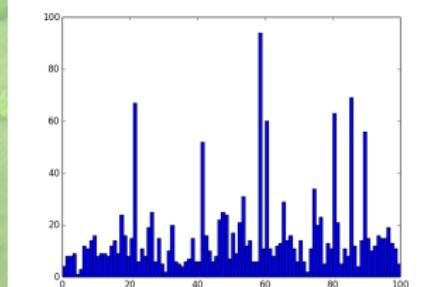
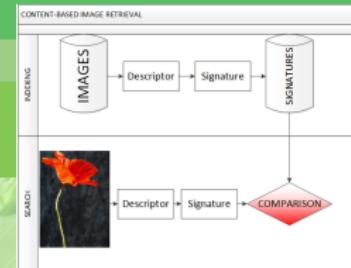
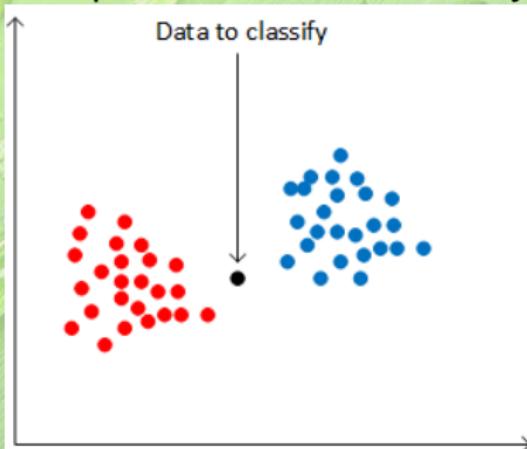


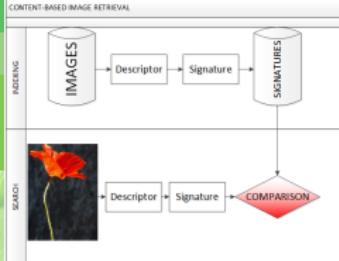
FIGURE: Signature 100 words -  
2

# K-nn(1/2)



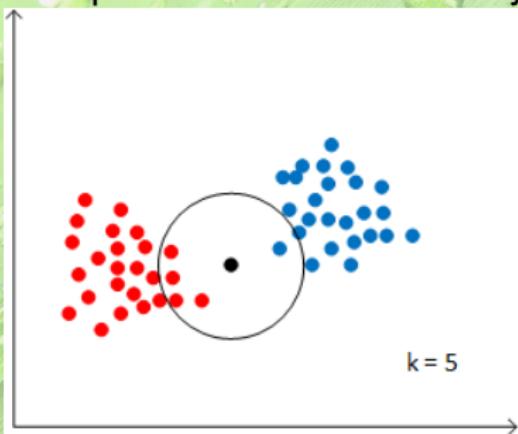
- The k nearest neighbor method
  - Comparison to the dictionary .





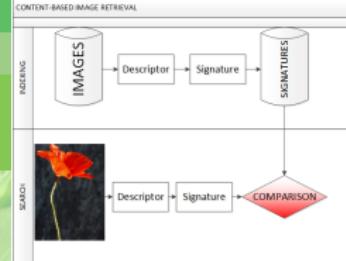
# K-nn(1/2)

- 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 species

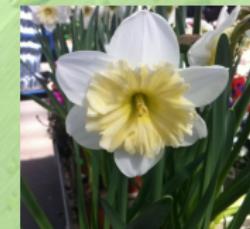


FIGURE: Second species



FIGURE: Third species



FIGURE: Fourth species

## Results (2/2)

- Compare the two descriptors SIFT and C<sub>2</sub>O.

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	/

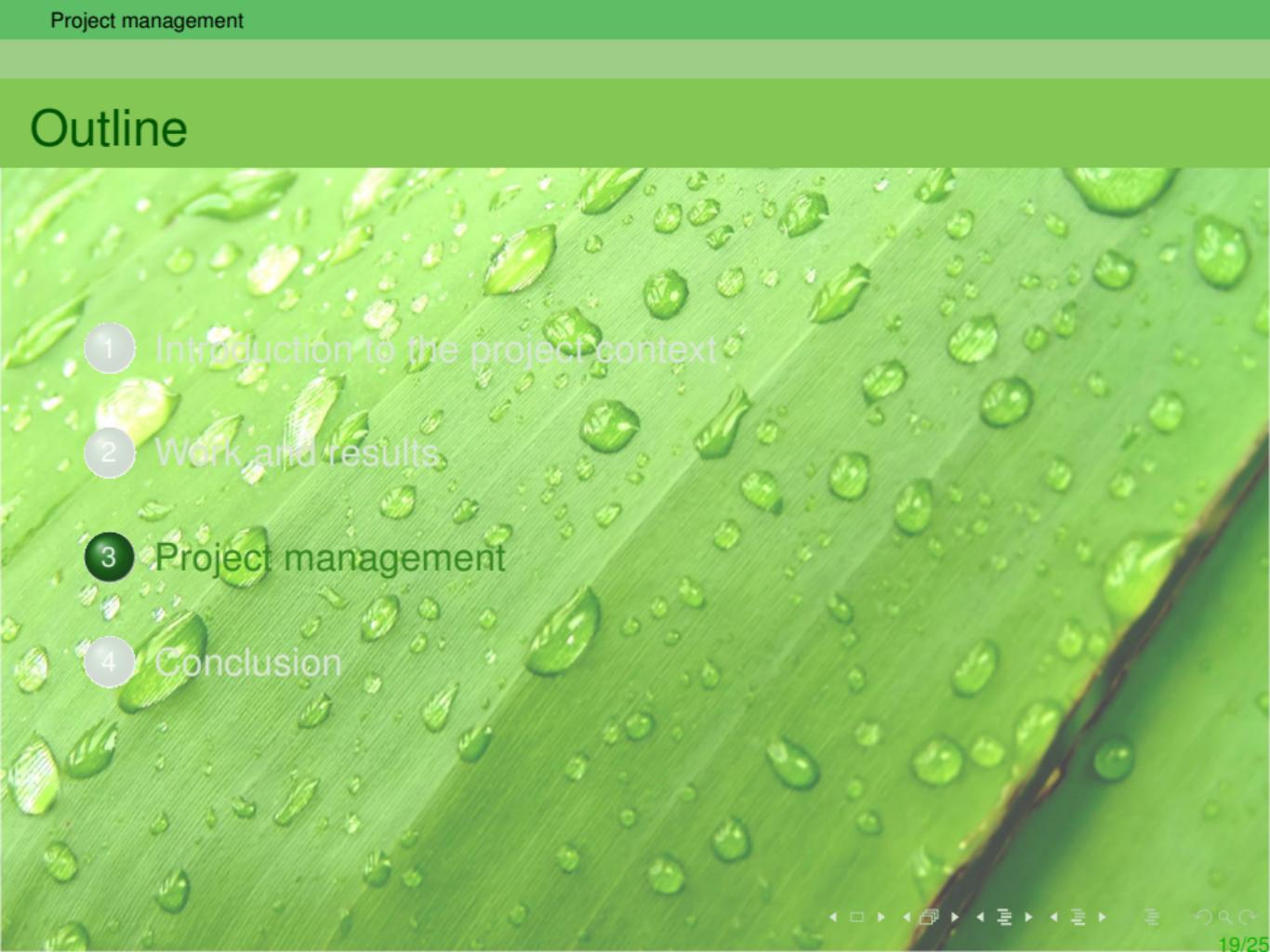
TABLE: C<sub>2</sub>O result

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

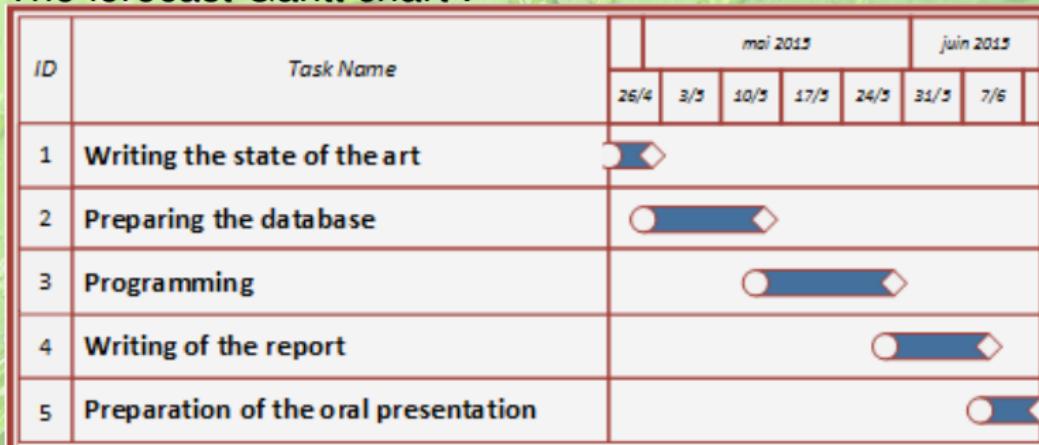
- Classification
  - To much reducing on the K-means (100 words).
  - Euclidean distance not the most efficient or adapt.
- C<sub>2</sub>O
  - The concatenation way is not optimal.
  - Parameters D, alpha, and beta has to be discussed regarding to the images.

# Outline

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

# Scheduling (1/2)

- The forecast Gantt chart :



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

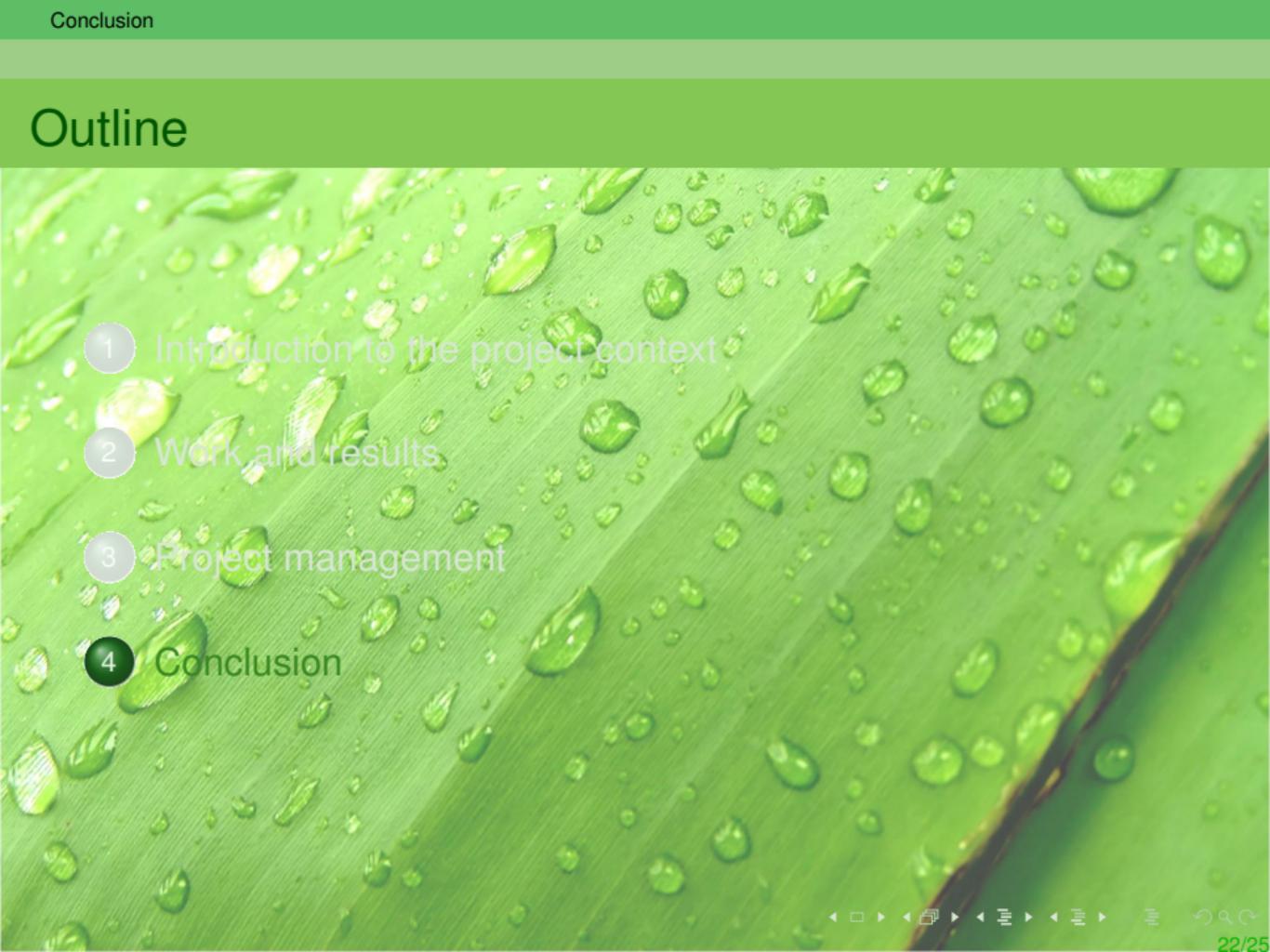
## Scheduling (2/2)

- The project backlog :

Spécificité	Catégorie	Sous catégorie	Nom / Description	Importance	Estimation	Critères de Vérification	Acteur	Status
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

- Division of each main task in subtasks
- Time attribution for each subtask
- Tasks sorted by priority
- Each subtask attributed to team member
- Allow to change the affectation of a task
- Weekly time affectation : could be adapted to unforeseen

# 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 conclusion

## Personal gains

- New way to organize teamwork
- Technical knowledge
- Contest participation context
- Code management on a project scale

## Perspectives

- Fixing technical issues
- Test on the whole database
- Classification programming



Thank you for attention