Progress Report

Introduction

Color image processing has become a major issue since a few years, most of the colour texture discrimination having been explored using the marginal colours way. The issue is that we are now able to do colour image recognition on digital images but the results on nature pictures are rather mediocre.

The CLEF contest has been created as an answer to that problematic, making universities' and laboratories' own solutions compete against each other in order to find the best colour texture feature.

In this document we will introduce key-points and their use in the various descriptors. We will go first with the standard ones which are SIFT SURF and opponent SIFT. The last one being the descriptor used by FINKI, the laboratory from the last year contest we chose as reference to compare our results. We will then use a new descriptor offered by Noel Richard, the C_2O.

Activities and Results

On our first week we achieved all of what we had expected to do (mainly the state of the art and getting into the project). The second week has been less efficient in terms of tasks completed, we had to several several uncompleted tasks into the next week.

Work completed

- XML generation programmed : Since we decided on using the XML file format to stock the key-points calculated for each image we programmed a function doing that for each descriptor we have to test

- State of Art completed : we used the reports from the previous years CLEF challenge to define the most used descriptors and .

- Programming of SIFT : using the openCV library in python we proceeded to test the

- Data Base constitution : After registering for the contest we retrieved the image Data Base.

- Process Flow programming : we wrote what will be the skeleton of our final product

- Classification analysis :

Work in progress

- Bag of Words programming : this is part of the classification

- C_2O programming : new descriptor from Noel RICHARD

- Evaluation Metrics programming : we are programming the official method of the contest to

- SIFT Tests : See above

Work remaining

- Learning how to use HULK mesocenter : this is the calculating server we will use to process the descriptors on all the images of the database (which amount to more than 90k)

- Extend the automatic process to the whole CLEF database (once every descriptor is programmed correctly we will have to extend the computing to the whole database

- Programming of SIFT variants : mainly opponent SIFT descriptor

- Writing of the final report : will be done

- Documenting of the HULK usage : Can only be done once we get the access

- Various tasks yet to be defined

Work Schedule

The organisation method we use to schedule our work is the SCRUM method. Basically before every week of work we decide of a number of tasks we have to prioritize and then we allocate a a number of hour from the tasked member's pool. At the end of the week we then sum up what's done what remains to be done and follow up on the schedule of the next week. We call that a sprint (the period can be more than a week). We described all the expected tasks to be done in a table called back log.

Conclusion

As far as we are confident that we can deliver something at the defined date but we probably won't be able to meet the objectives at the limit date for the LifeCLEF challenge. Objectively we were a bit short on time to meet the contest requirements from the start. It was to be expected so we had already defined the contest participation as an optional objective.