

Abstract

Intelligent analysis and comparison of images is one of the most dynamic topic of research in both academia and industry. Describing and comparing images automatically is a critical issue for the full development of the «information society»

Search engines working on textual data have dramatically proved their value. However, there is currently no similar system for image-only data. One possible explanation is that we do not really have a language made for describing images, thus meaningful comparisons are much more difficult than in the case of text. Nevertheless, textual search engines have shown that it is not necessary that the machines understand what they analyse to return good results. Simple syntactic analysis methods, coupled with composition rules are enough to drive extremely efficient systems.

To enable machines to simulate the interpretation of images, it would be necessary to create descriptors playing the role of words in text and composition rules making it possible to compare images like search engines compare sentences. We already have at our disposal numerous methods to detected automatically simple objects or regions in images, by their common color, their identical motion, etc. Furthering the analogy, these objects could be seen as syllables. The difficulty lies in grouping them to form words, then sentences, and compare them while being robust to perturbations.

To achieve this, we use graphs to store these objects and their relationships. These can be either of a neighboring nature, or inclusion, which leads the graphs to be either planar graphs or trees. We will see several methods to construct either type as well as their pros and cons.

In a first step, we have used the graph-matching algorithms developed by Cristina Gomila at the end of her PhD thesis at the CMM (1998-2001). While working with the european project MASCOT studying the use of «metadata» to enhance video coding, we have studied in detail the algorithm and spotted its strengths and weaknesses. We first tested replacing the core of the matching algorithm by a better one. This resulted in slight improvements in both quality and computation time. Then we tried to reduce our sensitivity to variations in the segmentation process by using a spectral graph-matching algorithm. Despite good results on simple images, our tests on harder images have not succeeded. To improve our robustness with respect to the stability of the graphs, we then preferred working on the source material : images.

The second step of this work was the development of image-base techniques to reduce the sensitivity of our segmentation algorithms to noise and small variations. First, we developed a class of adaptive filtering operators, the «morphological amoebas», which proved extremely effective in reducing noise in image. Second, we created a robust color gradient operator that can detect contour lines in noisy images. These two operators have improved sometimes spectacularly the stability of our segmentations, hence that of our graphs and in the end the quality of the results.

The next step in this work has been the modeling of objects independently from the rest of the image. This approach was motivated by realizing that in some scenarii the content of the image out-

side some well-defined objects is not informative. We must thus analyse directly and as precisely as possible the objects themselves. We first supposed that the segmentation of the outline of the objects was a solved problem, and concentrated on creating a robust signature for each object. To get it, we modified a watershed algorithm in order to perform a top-down resegmentation of a morphological scalespace based on levelings. We used this resegmentation to build a robust tree of embedded regions, and we defined a distance between those trees. We tested the whole process on a commonly used database by the indexation community.

The last step was centered around applications. First by comparing the various approaches presented in this document, concentrating in particular on the speed versus robustness compromise.

Then we search for the best combination of techniques to build a videosurveillance application. In particular, we developed fast and robust segmentation techniques for the project PS26-27 «Intelligent Environment» in partnership with ST Microelectronics and the ORION group of the INRIA. This aim of this project is to build a technology demonstrator for videosurveillance applied to the detection of accidents in hospitals or at home. Our part of the work was the detection of the outline of people in video sequences.

Finally, by coupling these detectors to our tree-based objects descriptors, we were able to define robust signatures for people that could be use with great profit by automatic videosurveillance systems.

Keywords : segmentation, morphological scalespace, graphs, trees, noise reduction, morphological amoebas