

**Year of CfP: 2007**

**Project No: 07004 Completed**

**Project title:** Automated identification of weed species in Camargue rice fields using techniques for the recognition of visual content

**Unit managing the project:** AMAP (botany and computational plant architecture) (CIRAD, CNRS, INRA, IRD, UMII)

**Project leader:** Daniel Barthélémy (barthelemy(a)cirad.fr)

**Countries involved in the project:** Philippines, India, Colombia, West Africa, Madagascar

**Other research units from the Foundation's scientific network involved:** Innovation

**Subthematic axes:** IPB-1 (Integrative Plant Biology 1: *Genetics and genomics, plant breeding, ecophysiology*), IPB-2 (Integrative Plant Biology 2: *Plant pests and diseases, integrated crop protection, population ecology*), STDI-1 (Socio-Technical Dynamics of Innovation 1: *Agri-environmental innovations, agri-ecosystems, resources management*)

**Objectives:**

The difficulties encountered in taxonomic identification are known for being one of the major obstacles to the application of the Convention on Biological Diversity. Since this obstacle was shown at the Conference of the Parties, much effort has been devoted to disseminating taxonomic information and identification tools that are faster, more relevant and more effective.

Numerous text and photographic taxonomic databases have thus been generated and disseminated using New Information and Communication Technology (NICT). However, they are still difficult to use for reasons of their size and complexity.

This project is aimed at achieving the automatic identification of weeds in rice fields in the Camargue using appropriate algorithms and an image recognition system linked with a botanical taxonomic database. The work will be conducted on a limited flora in which we already possess great expertise and that comprises several thousand images. The research will be focused on the species shown in the book *Les plantes des rizières de Camargue*, published jointly by CIRAD, the Parc naturel régional de Camargue and the Centre français du riz.

It must make it possible to analyse the feasibility of an approach in a real case and develop generic methods for the automatic identification of plant species by means of a computer system of recognition of visual content. The results will form the basis for subsequent, more ambitious projects on the automated identification of weeds of tropical crops, invasive plants or the plant species that form various natural ecosystems, especially in the tropics.

A photographic database will be assembled using standardised illustrations of 178 weed species in Camargue rice fields. This database, together with the existing database for this flora, will be used by image recognition tools for the optimisation of the latter. Present tools will be appraised and improved using series of tests conducted jointly by data processing specialists, botanists and weed scientists. The results will be reported in publications drafted jointly by the various partners and a report on the databases and the tools developed will be published on the websites of the various participating bodies.

**Action carried-out and results obtained:**

A preliminary phase of this project consisted in setting up an acquisition protocol of weeds pictures, adapted to the evaluation of the visual search engine IKONA. The realization of this protocol as well as the computer processing of the results was made on a picture database of Orchids of Laos. A second phase of the project, consisted to adapt this protocol to the specific pictures of rice weeds,

that were made during regular missions conducted from June to October 2008. All images realized were then annotated and transferred to Imedia project-team through the visual search engine Ikona.

These results showed the high performance of global images descriptors on well standardized pictures, and less good results on non-standardized ones. Tests led with local descriptors (such as points of interest) showed in some cases good results, due to the fact that they allow to be free of the background' constraints, however there are until now no mathematical tools adapted to their evaluation for taxonomic identification.

All the results of the project were put on-line in the form of tables and graphics, to show the pertinence of the various descriptors methods on the various images types. Two different applications permit us to use on-line the visual search engine Ikona on picture archives produced within the framework of this project. The specific interface of the research engine allows furthermore to access on each image to the species descriptive page on the website of the weeds of the rice fields of the Camargue

#### **Publication:**

Le Bourgeois T., Bonnet P., Edelin C., Grard P., Prospero J., Théveny F., Barthélémy D., 2008. L'identification des adventices assistée par ordinateur avec le système IDAO. Innovations Agronomiques, 3 : 167-175

#### **Prospects for the future:**

The results supplied within the framework of this project, allow to orientate our future researches for the adaptation of the visual research system Ikona for a more efficient taxonomic identification.

Activities are in progress for evaluation of a specific shape descriptor of leaves forms and to compare it with generic visual descriptors. The requests made with local descriptors of images seem to offer interesting perspectives for the extension of this approach, they must however be improved in order (i) to allow their exploitation and evaluation on a wider scale, (ii) to run in a simultaneous way with shape and colors descriptions, (iii) to allow the description of forms with local descriptors according to their respective geometrical position.

Our works also revealed the interest to be able to make combined request by texts and images, as well as to be able to make multi-class requests. The multi-class request (the possibility to search a species by several characters types, as leaves scans, flowers pictures, etc.) will be thus deeply studied within the framework of the project PI@ntNet. The learning techniques were not able to be estimated during the AF 07 004 project, they seem to be sharply able to improve the answers of the system, that's why exploitation of relevance feedback with points of interests is the object of a thesis subject financed by INRA and INRIA. These encouraging results on weeds species require additional researches that will be implemented within the framework of PI@ntNet. They have to allow the realization of a multiapproaches identification tool (visual, graphical and textual). The approaches recommended during these researches want to be the most generic as possible. That's why results of these works will also serve as base for the extension of this approach for the identification of plants and/or plants diseases, by several images types and of several identification characters.

**Total Agropolis Fondation funding:** € 15,547 (travel expenses, 3 months contract for an engineer, photographic equipment)

**Funding categorie(s):** Agropolis Fondation small grants (support for small exploratory, risky and innovative projects ("proof of concept", new frontier research, etc.)

**Project duration:** January 7 2008 – January 7 2009

**Keywords:** taxonomy – photography – database – weeds – rice field