

WP2: General MI theory and tools

Last Updated Thursday, 03 March 2011

Development of general MI theory and tools is a major goal in MIPRCV. This is done in this WP through three research tasks on interaction, multimodality and adaptive learning. The basic, non-interactive technology viewed in WP 1 is used as a basis of development. The resulting theory is going to be applied in different application domains, in WPs 3, 4 and 5.

Some relevant publications:

S. Barrachina, O. Bender, F. Casacuberta, J. Civera, E. Cubel, S. Khadivi, A. Lagarda, H. Ney, J. Tom s, and E. Vidal. Statistical approaches to computer-assisted translation. *Computational and Linguistics*, 35(1):3-28, 2009.

A presentation of the theoretical framework and experimental results about interactive machine translation systems.

Alejandro Toselli, Verónica Romero, Moisés Pastor, and Enrique Vidal. Multimodal interactive o e transcription of text images. *Pattern Recognition*. 43(5), 1814-1825, 2009.

Grounded in the interactive-predictive general framework drawn in the WP2, in this article specific mathematical formulation and modelling are derived specifically for interactive transcription of handwritten text images. By adopting the so-called "passive left-to-right interaction protocol", an interactive approach for efficient transcription of handwritten text along with its more ergonomic and multimodal variants are presented. All these approaches, rather than full automation, aim at assisting the expert in the proper transcription process in an efficient way.

A. Cano, A. R. Masegosa and S. Moral. A method for integrating expert knowledge when learning Bayesian Networks from data. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*. Accepted for publication. 2011.

This work proposes a interactive methodology for learning Bayesian networks from data. This is a challenging task, particularly when the data is scarce and the problem domain contains a high number of random variables. In this work, we introduce a new methodology for the interactive integrating of expert knowledge, based on Monte Carlo simulations and which avoids the costly elicitation of these prior distributions. The great advantage of this integration approach is that only requests from the expert a very limited amount of information, to be more precise, about those direct probabilistic relationships between variables which can not be reliably discerned with the help of the data.

Â

L.M. de Campos, A.E. Romero, Bayesian network models for hierarchical text classification from a thesaurus, *International Journal of Approximate Reasoning* 50(7):932-944, 2009.

We propose a method which, given a document to be classified, automatically generates an ordered set of appropriate descriptors extracted from a thesaurus. The method creates a Bayesian network to model the thesaurus and uses probabilistic inference to select the set of descriptors having high posterior probability of being relevant given the available evidence (the document to be classified). Our model can be used without having preclassified training

documents, although it improves its performance as long as more training data become available. It is a multimodal system in the sense that it integrates three different sources of information: in addition to the textual content of the documents, it also manages structural information (the thesaurus' hierarchy of concepts), and semantic information (descriptors and non descriptors in the thesaurus).

Â

J. Oncina, Optimum Algorithm to Minimize Human Interactions in Sequential Computer Assisted Pattern Recognition, Pattern Recognition Letters, vol. 30, pp. 558-563 (2009)Â

Given a Pattern Recognition task, Computer Assisted Pattern Recognition can be viewed as a series of solution proposals made by a computer system, followed by corrections made by a user, until an acceptable solution is found. For this kind of systems, the appropriate measure of performance is the expected number of corrections the user has to make. In the present work we study the special case when the solution proposals have a sequential nature. Some examples of this type of tasks are: language translation, speech transcription and handwriting text transcription. In all these cases the output (the solution proposal) is a sequence of symbols. In this framework it is assumed that the user corrects always the first error found in the proposed solution. As a consequence, the prefix of the proposed solution before the last error correction can be assumed error free in the next iteration. Nowadays, all the techniques in the literature relies in proposing, at each step, the most probable suffix given that a prefix of the "correct" output is already known. In the present work we show that this strategy is not optimum when we are interested in minimizing the number of human interactions. Moreover we describe the optimum strategy that is simpler (and usually faster) to compute.