

Doctoral Thesis Defense

Speaker:	Wumo Pan
Supervisor:	Drs. T. D. Bui, C. Y. Suen
Examining Committee:	Drs. Szabo, Cheriet, Krzyzak, Lam
Title:	Pattern Detection and Recognition using Over-Complete and Sparse Representations
Date:	March 30, 2009
Time:	14:00
Place:	EV 1.162

ABSTRACT

Recent research in harmonic analysis and mammalian vision systems has revealed that over-complete and sparse representations play an important role in visual information processing. The research on applying such representations to pattern recognition and problems has become an interesting field of study. The main contribution of this thesis is to propose two feature extraction strategies - the global strategy and the local strategy - to make use of these representations.

In the global strategy, over-complete and sparse transformations are applied to the input pattern as a whole and features are extracted in the transformed domain. This strategy has been applied to the problems of rotation invariant texture classification and script identification, using the Ridgelet transform. Experimental results have shown that better performance has been achieved when compared with Gabor multi-channel filtering method and Wavelet based methods.

The local strategy is divided into two stages. The first one is to analyze the local over-complete and sparse structure, where the input 2-D patterns are divided into patches and the local over-complete and sparse structure is learned from these patches using sparse approximation techniques.

The second stage concerns the application of the local over-complete and sparse structure. For an object detection problem, we propose a sparsity testing technique, where a local over-complete and sparse structure is built to give sparse representations to the text patterns and non-sparse representations to other patterns. Object detection is achieved by identifying patterns that can be sparsely represented by the learned structure. This technique has been applied to detect texts in scene images with a recall rate of 75.23% (about 6% improvement compared with other works) and a precision rate of 67.64% (about 12% improvement).

For applications like character or shape recognition, the learned over-complete and sparse structure is combined with a Convolutional Neural Network (CNN). A second text detection method is proposed based on such a combination to further improve (about 11% higher compared with our first method based on sparsity testing) the accuracy of text detection in scene images. Finally, this method has been applied to handwritten Farsi numeral recognition, which has obtained a 99.22% recognition rate on the CENPARMI Database and a 99.5% recognition rate on the HODA Database. Meanwhile, the SVM with gradient features achieves recognition rates of 98.98% and 99.22% on these databases respectively.