PhD Position 1 - *Patch-based Image Representation*

Supervisor: Prof M Cheriet
Location: Synchromedia Lab, ETS, Montreal, Quebec, Canada
Starting Date: September 1st, 2014
Keywords: *Data-driven, Data representation, Patch-based image representation.*

Context

Ancient manuscripts constitute a primary carrier of cultural heritage globally, and they are currently being intensively digitized all over the world to ensure their preservation, and, ultimately, the wide accessibility of their content. Critical to this research process are the legibility of the documents in image form, access to live texts, and similarity between document image within huge collections (big data). Several state-of-the-art methods and approaches have been proposed and developed to address the challenges associated with processing these manuscripts. The key issue to be tackled to face these challenges is the learning of prior representation of the document images that can be used in many computer vision and image processing algorithms.

Beyond pixel-based representation, patch-based image models enforced by non-parametric sampling shown increasingly more vitality and they could be very interesting representation in many successful applications, i.e. exemplar-based image inpainting, image denoising, such as non-local mean, patch-based image denoising, BM3D, exemplar-based super-resolution, sparsity-based, object recognition (i.e. SIFT, LBP, SURF, and SOR), image retrieval and document image binarization. All these could be seen somehow related to representation learning.

Patch-based representations shown to be more reliable to smoothly preserve implicit information carried by an image, such as its spatial and frequential (for instance phase information) relations. The two aspect of information carried by document images, i.e., spatial relations, and frequential relations, are all of equivalent importance to the thesis position/proposal, and the candidate is required to contribute to all of them by introducing separate or hybrid representations.

Patch-based representation could be extended smoothly to deal with vector-valued (color and multispectral) document images. The latter contain additional information in a widespread electromagnetic spectrum (visible and invisible) more suitable for the design of more discriminative features. In addition to the spatial and frequential relations that can be learned from similar patches, spectral relations-along the different bands-is an important feature that may play a key role in the design of reliable processing models of enhancement, restoration, among others.

Objective

The goal of this research is to develop image "representations", which in a smooth go beyond the traditional pixel-wise representations, and preserves implicit information carried by an image, such as its spatial, spectral, and frequential relations. These new representations should be applicable and transformable. Not

limited to that but also to be more generic in order to be used in many application of computer vision, image processing, and machine learning. Therefore, developing the associated models, algorithms, and methods to enhance, restore, or do any other process on images is part of this PhD research/proposal.

Research Questions

In addition to the following research questions, the student is requested to open new possible questions and provide with innovative hypothesis, in order to supply his thesis. Three issues could be raised.

- What are the attribute or the features that we need to capture -from the patches- for providing accurate datasets allowing to design robust and reliable systems, for learning as well as for testing; and also to respond to specific questions posed by end-users (scholar, librarian, ..) about their documents.
- What is the best way to model the prior models of an image: using visual data directly from the image or using explicit intermediate representation (i.e. transformation to features)
- What is the best metric and what is the appropriate metric space (Cartesian, Riemannian,…) are suitable for dealing with big data.
- Is it possible to develop generic methods to solve more than one type of degradations of documents images? Otherwise, several questions arise: how many types of degradations are there and how many methods should be developed?
- What would be a representation that can absorb spatial, spectral, frequency, among other information carried by images with minimal information loss and maximum transformability? What would be it in the case of document images, in the case of a manuscript image, and in the case of a collection of manuscript images?

Requirements

The candidate must hold a M.Sc. in Computer Science (or equivalent). He/she must very good skills in image processing, statistics, pattern recognition, algebra, optimization, and matlab (and C/C++) programming. The candidate should have good writing skills in English. He/she must be highly motivated, independent, with a real ability to organize and follow a schedule.

References

[1]   A. Criminisi and K. Toyama, ''Object removal by exemplar-based inpainting'', in
        IEEE CVPR, 2003.
[2]   A. Criminisi, P. Perez, and K. Toyama, ''Region filling and object removal by examplar-based inpainting'', IEEE Trans. On Image Processing, vol. 13, pp. 1200-1212, Sept. 2004.
[3]   A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," CVPR, vol. 2, pp. 60–65, 2005.
[4]   C. Kervrann and J. Boulanger, "Unsupervised patch-based image regularization and representation," in ECCV06, pp. IV: 555–567, 2006.
[5]   J. Lee, "Digital image enhancement and noise filtering by use of local statistics," IEEE Transactions on Pattern Analysis and Machine Intelligence, no. 2, pp. 165–168, 1980.
[6]   W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," IEEE Computer Graphics and Applications, vol. 22, pp. 56–65, 2002.
[7]   J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution as sparse representation of raw

image patches," IEEE Conference on Computer Vision and pattern Recognition, pp. 1–8, 2008.

[8]   D. G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. of Comp. Vis., vol. 60, pp. 91–110, 2004.

[9]   T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Transactions on pattern analysis and machine intelligence, pp. 971–987, 2002.

[10]   H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," ECCV 2006, pp. 404–417, 2006.

[11]   Sauvola, J. and M. Pietikainen. February 2000. " Adaptive document image binarization". Pattern Recognition, vol. 33, n◦ 2, p. 225–236.

[12]   Rachid Hedjam, Reza Farrahi Moghaddam, and Mohamed Cheriet. 2011. A spatially adaptive statistical method for the binarization of historical manuscripts and degraded document images. Elsevier Pattern Recognition (PR). 44(9), Sept. 2011, 2184-2196.

[13]   Farrahi Moghaddam, Reza and Mohamed Cheriet. 2012. " AdOtsu: An adaptive and parameterless generalization of Otsu's method for document image binarization ". Pattern Recognition, vol. 45, n◦ 6, p. 2419–2431.

[14]   Rachid Hedjam and M. Mignotte, "A hierarchical graph-based markovian clustering approach for the unsupervised segmentation of textured color images," in ICIP'09, 2009, pp. 1365–1368.

[15]   Rachid Hedjam, ''Visual Image Processing in Various Representation Spaces for Documentary Heritage Preservation'', Ph.D. thesis, ETS, University of Quebec, Montreal, Quebec, Canada, April, 30 2013.

[16]   Rachid Hedjam and Mohamed Cheriet, "Historical document image restoration using multispectral imaging system," Pattern Recognition, vol. 46, no. 8, pp. 2297–2312, Aug. 2013.

[17]   Reza Farrahi Moghaddam and Mohamed Cheriet, "Real-time knowledge-based processing of images: Application of the online NLPM method to perceptual visual analysis," IEEE Transactions on Image Processing, vol. 21, no. 8, pp. 3390–3404, 2012

[18]   Reza Farrahi Moghaddam, Fereydoun Farrahi Moghaddam, and Mohamed Cheriet, "A new framework based on signature patches, microregistration, and sparse epresentation for optical text recognition," in ISSPA'12, 2012, pp. 1259–1265

[19]   M Cheriet, R Farrahi Moghaddam, R Hedjam, '' Visual language processing (VLP) of ancient manuscripts: Converting collections to windows on the past'', GCC Conference and Exhibition (GCC), 2013 7th IEEE, 407-412

[20]   Hossein Ziaei Nafchi, Hamidreza Rashidy Kanan: Rectangular based binary image representation: Theory, applications, and dataset introduction. ICPR 2012: 190-193

[21]   Cheriet, Mohamed, Reza Farrahi Moghaddam, and Rachid Hedjam. 2012. "A learning framework for the optimization and automation of document binarization methods ", Computer Vision and Image Understanding, 2013.

[22]   Stefan Roth and Michael J. Black, ''Fields of experts'', International Journal of Computer Vision, (2008).

[23]   Bengio, Y.; Courville, A. & Vincent, P. Representation Learning: A Review and New Perspectives IEEE Transactions on Pattern Analysis and Machine Intelligence, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2013, 35, 1798-1828.