ENDOMICROSCOPIC IMAGE RETRIEVAL AND CLASSIFICATION USING INVARIANT VISUAL FEATURES

Barbara André, Tom Vercauteren, Aymeric Perchant, Anna M. Buchner, Michael B. Wallace, Nicholas Ayache

[1] Mauna Kea Technologies (MKT), Paris, France
[2] Asclepios Project-Team, INRIA Sophia-Antipolis, France
[3] Mayo Clinic, Jacksonville, Florida, USA

Abstract

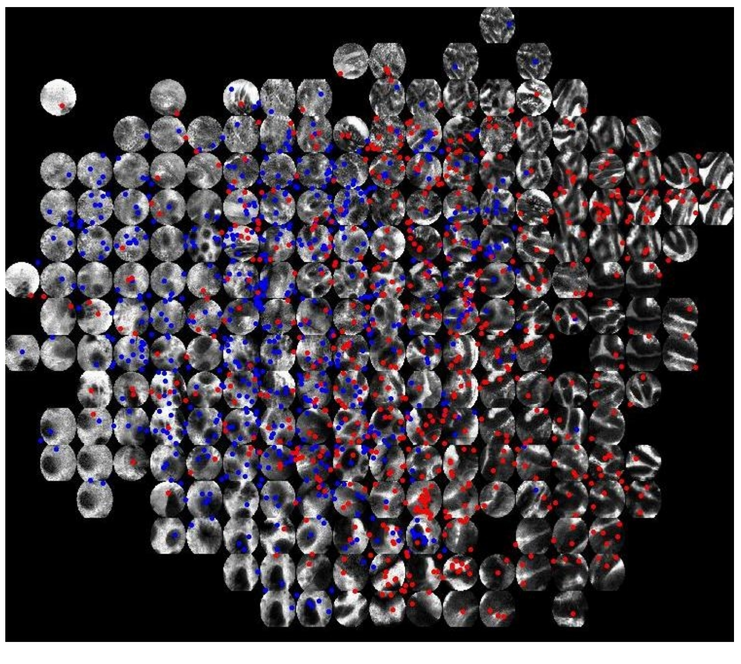
Probe-based Confocal Laser Endomicroscopy (pCLE) is a new technology that enables dynamic microscopic imaging of tissues in vivo, with a miniprobe, during ongoing endoscopy. Our study investigates the application of Content-Based Image Retrieval to pCLE images of colonic polyps, which would aid the physician in differentiating benign tissues and neoplastic (pathological) tissues.

We first explore the Bag of Visual Words method that maps an image into a visual feature signature which is a numerical vector invariant with respect to viewpoint changes, e.g., translations, rotations and scaling, and illumination changes, e.g., affine transformation of intensity.

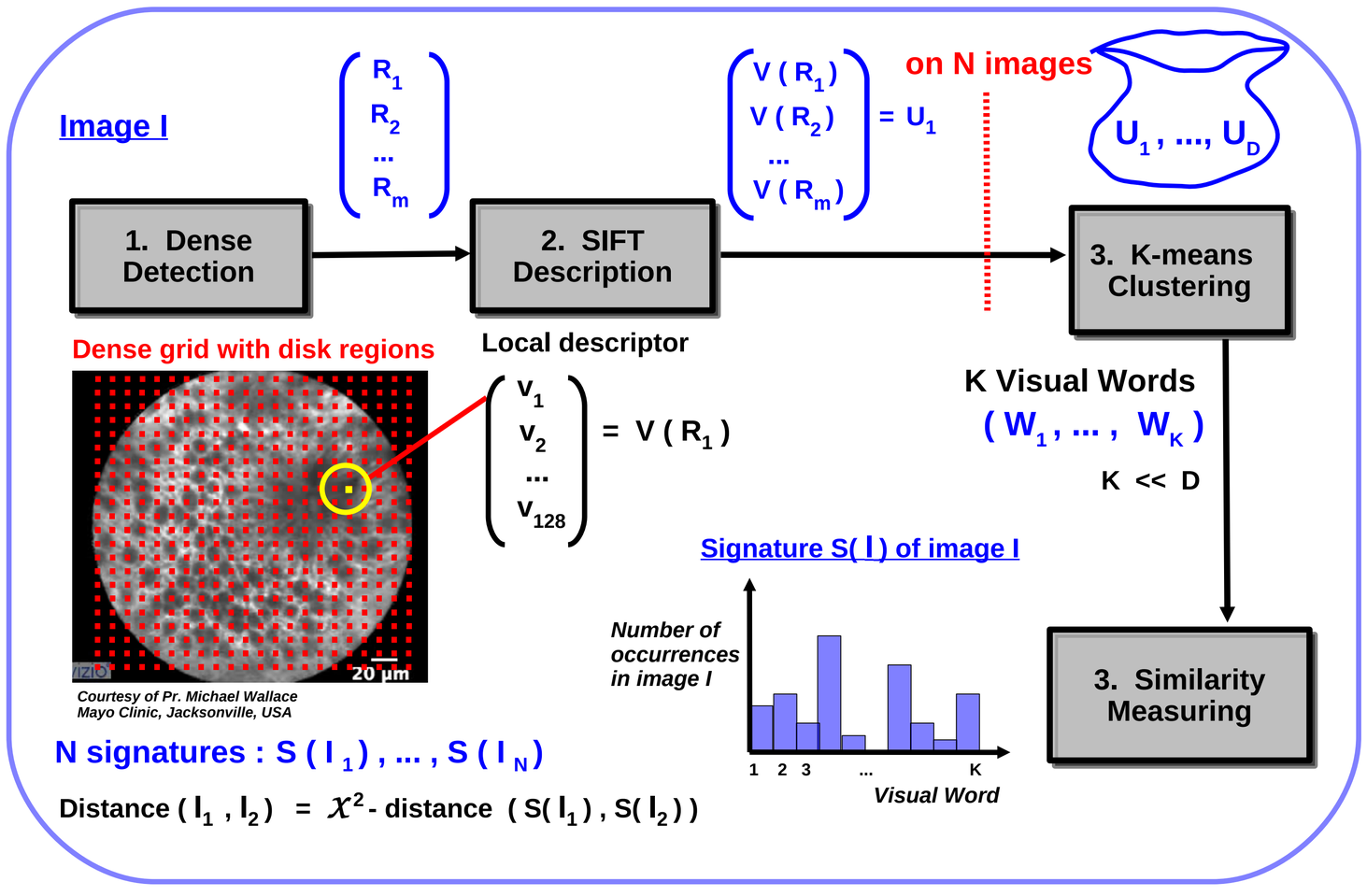
The relevance of the similarity results is then quantified by a k-nearest neighbors classification of the training database.

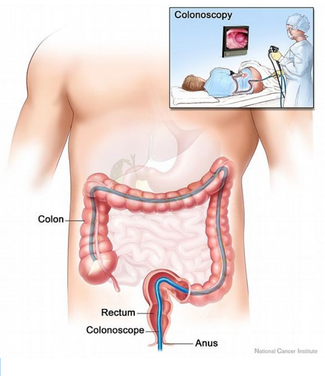
Compared with several previously published alternatives whose maximal accuracy rate is barely 67% on our pCLE database (while almost 99% on UIUCTex texture database), our approach yields an accuracy of 80% and offers promising perspectives.

Projection of the pCLE images from the K-dimensional signature space to a plane: Benign ● Neoplastic ●



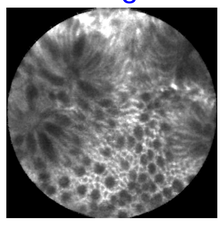
Bag of Visual Words Method

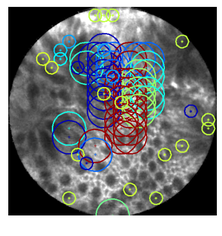




Imaging tissues of colonic polyps with Cellvizio®

Benign





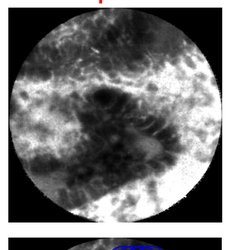
Learning the discriminative power of (group of) visual words

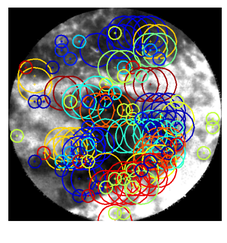
Same color ⇔ Same visual word

Microscopic regions: individual cells

Mesoscopic regions: groups of cells

Neoplastic



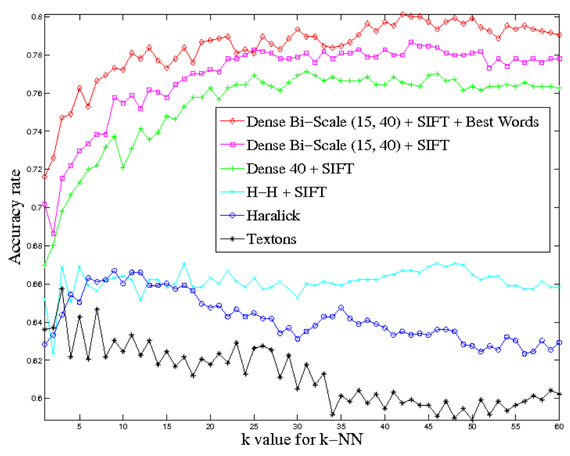


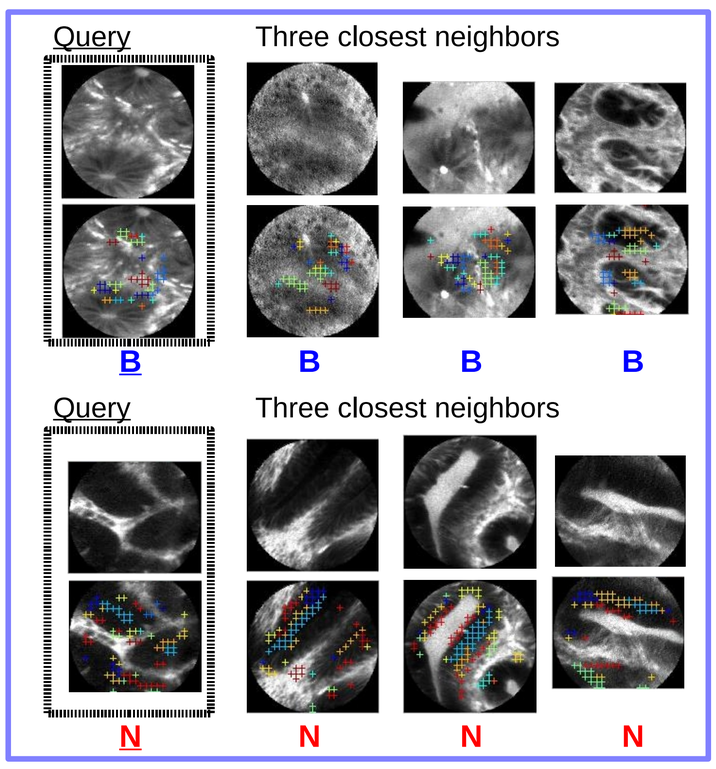
Retrieval and Classification Results

Training database: 1036 images
2 classes: Benign (B), Neoplastic (N)

k-NN leave-n-out classification

Accuracy of our method:
80 % correct retrievals for k=42 neighbors





Contributions and Conclusion

- Dense sampling of the image space
- Multi-scale description of the feature information, from cells to groups of cells
- Supervised selection of the most discriminative visual words

First attempt to classify endomiscroscopic images using content-based image retrieval. Generic framework for many image retrieval applications and for multi-class image classification.

Ongoing work

Find a robust validation framework with the aid of medical expertise.
Include spatial relationship between visual words and time relationship between video frames.

References

- J. Zhang, S. Lazebnik, and C. Schmid, “Local features and kernels for classification of texture and object categories: a comprehensive study,” IJCV 2007.
- J. Sivic and A. Zisserman, “Efficient visual search for objects in videos,” Proc. IEEE 2008.
- D.G. Lowe, “Distinctive image features from scale invariant keypoints,” IJCV 2004.
- B. Andre, T. Vercauteren, A. Perchant, A.M. Buchner, M.B. Wallace, and N. Ayache, “Endomicroscopic Image Retrieval and Classification using invariant visual features”, ISBI 2009, to appear.