SOLO-MEDICINE IN OPTICAL BIOPSIES
A Way to Practice Telemedicine

Olga Ferrer, Francesco Ettorre, Xiomara Santos
University of La Laguna, La Laguna, Canary Islands, Spain

Thomas Zinkl, Ruben Tous, Jaime Delgado
Department of Computer Architecture, Universitat Politecnica de Catalunya (UPC-BARCELONATECH), Barcelona, Spain

Keywords: Solo-medicine, Optical biopsy, Confocal laser endomicroscopy, CLE, CBIR, Medical image retrieval, Query by image, ISO-15938-12, MPEG query format, MPQF, ISO 24800-3, JPSearch, JPEG query format, JPQF, Artificial intelligence, Multimedia standard.

Abstract: A way to practice Telemedicine is to access a data-base capable to assist you in medical procedures (diagnosis, treatment and prognosis), similarly to consult a book or to ask a college. In many countries the lack of specialists and training capabilities demand to practice solo-medicine, that in the case of surgery require robots capable to induce anesthesia or help in vision or handling instruments. A relevant case is the diagnostic self-training requirements for optical biopsies (OBs) obtained with confocal laser endomicroscopy (CLE) or the assistance in the diagnosis of pathology slides. In both cases it is required a training set of digital images against which to compare the question case by means of image-queryformat. The present paper present a content-based image retrieval system (CBIR) based on the MPEG Query Format Standard in order to provide a set of similar pictures and the corresponding diagnosis to help on diagnosis or just to train the doctor. The paper defined the Image Solo-Medicine Paradigm (ISMP) architecture merging medical image standards and MPEG and JPEG standards. It tested the solution with normal, and benign colon OBs with 90% congruency. The ISMP is of particular interest viewed the proliferation of iPhone medical applications aiming to train doctors and support medical decisions.

1 INTRODUCTION

Nowadays the number of medical applications for iPhone proliferate (CATAI, 2010), attracting the interest of relevant medical Journals such as the British Medical Journal (BMJ Group, n.d.) to build applications that help doctor to make their decisions and auto-train themselves. The time of training books, with periodical updates for new diagnosis or treatments, is arriving to an end; doctors will have on-line and on mobile phones that information, and will use mobile phones for a variety of medical applications (Ferrer-Roca and Marcano, 2009; Ferrer-Roca and Marcano, 2010; Ferrer-Roca, 2010).

We have been working in a diagnostic medical application based on images and on which "gold-standards" are still on the way (Hersch et al., 2005; Kiesslich et al., 2008; Ferrer-Roca et al., 2010). This is the so-called optical biopsy (OB) (Wang and VanDam 2004). A non-intrusive optic diagnostic method, capable to analyze the tissue in surface and in deepness with one of the following techniques: laser, OCT, infrared, fluorescence, spectroscopy etc. This means, that it is not necessary to extract the tissue from the body. Tissue is accessed through the surface of the body through the skin or by endoscopy.

In OBs images are obtained in real time together with complementary information that allows evaluating the disease in vivo, but "gold-standards" are still lacking while in surgical pathology standards lay on the histology of the normal fixed tissue (Ferrer-Roca, 2009). To provide training and self-confidence on OB diagnosis, two possibilities are open: (a) Tele-consult to a pathologist or (b) Train themselves with a non-supervised search for a "similar image" on the Net using multimedia query and image mining techniques (Chen et al., 2006).

Ferrer O., Ettorre F., Santos X., Zinkl T., Tous R. and Delgado J..
SOLO-MEDICINE IN OPTICAL BIOPSIES - A Way to Practice Telemedicine.
DOI: 10.5220/0003090804410445
In *Proceedings of the International Conference on Health Informatics* (HEALTHINF-2011), pages 441-445
ISBN: 978-989-8425-34-8

The present proposes a standardized ISMP (Image Solo-Medicine Paradigm) architecture based in the usage of two novel standards, the MPEG Query Format (MPQF) and the JPEG's JPSearch project (Tous, 2006). While MPQF provides a uniform language for querying multimedia databases, JPSearch provides an interoperable architecture for images' metadata management. Preliminary results on Internet image search and discovery system for diagnostic medical purpose are showed. Results were based on a training-set of CLE-OB images annotated with specific CLE semantics.

2 MATERIAL AND METHODS

The proposed system to allow users to navigate searching similar images considered to be golden standards (due to the pathology confirmation and availability of pathology image) in a pair image database integrated by OB-CLE images together with the histological counterpart.

Images used in the present paper were provided by one of the authors (OFR) or taken from data published in Internet. All were JPEG images.

2.1 ISMP (Image Solo Medicine Paradigm) Architecture

The ISMP system provided tools to annotate an unknown OB-CLE image with key-words and image structural information for content based image retrieval (CBIR). Figure 1 summarizes the overall architecture.

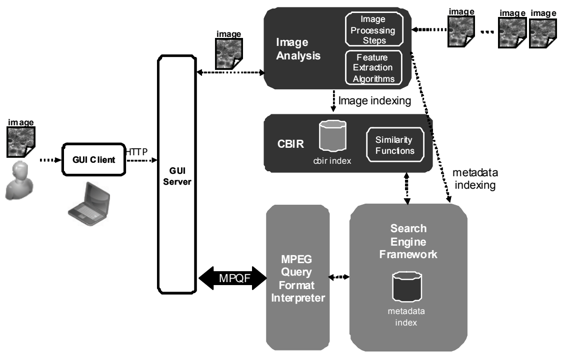

Figure 1: Overall architecture of Image Solo-medicine Paradigm(ISMP).

Image Solo-Medicine Paradigm (ISMP) architecture was integrated by four main modules:

1) Image processing: Offline extraction of medium-level and high-level metadata from the images in the database, and also to the on-the-fly extraction of the same metadata from an example image submitted by a user as a query. We used the ImageJ (ImageJ, n.d.) Java library to implement an adhoc algorithm.

2) CBIR index construction: We generated an index for query-by-example search by means of selection of a feature vector and a similarity function.

3) Search Engine Framework: We built a query processor capable of solving text-based queries, CBIR queries and combinations of both.

4) MPEG Query Format Interpreter: In order to effectively ensure interoperability with potential third-party applications we built a standard interface based on ISO/IEC 15938-12:2008 (MPEG Query Format, MPQF).

3 RESULTS

3.1 ISMP Training Set

It is composed by 25 OB-CLE images obtained with a PENTAX CLE with their re-sulting histological images (50 images in total).

3.1.1 ISMP Preprocessing

ISMP preprocessing was done in two steps: 1- Normalization (to minimize light in homogeneities caused by laser light source) that included several image processing steps (enhanced contrast, equalization, etc.). 2- Grey level reduction using *pixel value range reduction and region merging* algorithms as seen in Figure 2.

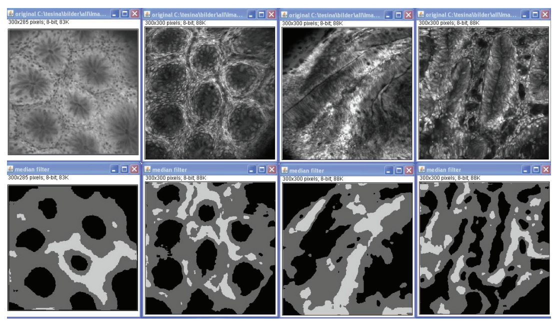

Figure 2: ISMP pre-processing of original images (top). Results in the bottom line.

3.1.2 ISMP Feature Extraction

ISMP feature extraction was done in two steps:

1) *The Local Binary Pattern (LBP)* (Pietikainen et

al., 2000) operator (A gray-scale invariant texture measure derived from a general definition of texture in a local neighbor-hood). The process included (a) Integration: On each pixel, we calculated an array of bits of 0 and 1 comparing the original pixel value and its neighbors in a certain radius. (b) Decision maker: The array values are summed up. The higher *lbpSum* for a pixel indicated more likely to be the center of one of the big black areas (Figure 2).

2) The modified density-based DBSCAN algorithm, highlighted the various crypts and their boundaries.

With the *lbpSum* value for every pixel, we apply a clustering algorithm to cluster to a certain crypt. In the clustering process we used a *modified density-based DBSCAN* algorithm originally proposed in (Sander et al., 2004). See Figure 3.

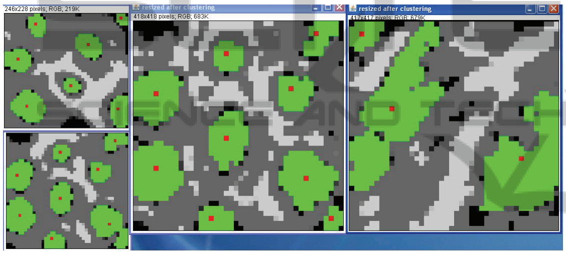

Figure 3: Clustering for gland identification. Normal (left) and hiperplastic glands (right).

3.1.3 ISMP Feature Measurement

The results of this process allow us to extract and measure certain features such as the silhouette coefficient, the crypt compactness, the crypt roundness or the inter-crypt distance.

3.1.4 ISMP-indexing and ISMP-retrieval in Two Steps

1) We defined a feature vector and normalized it applying *linear scaling unit range normalization*.

2) We retrieved similar images to a given one, using the similarity function. The selected function operated over the vector of selected features, whose composition determines which is the nature of the similarity being considered (similarity is relative in a multidimensional space).

The test set demonstrated that the *manhattan* and the *euclidian* distance, in combination with the *linear scaling unit range normalization*, provide better performance.

3.1.5 ISMP Data-base Query

MPQF queries were evaluated against one or more multimedia databases which were unordered set of Multimedia Contents-MC (combination of multimedia data and its associated metadata).

1) Data-base: It was a dual database model (Figure 4) by content and by metadata. The MPQF operated over sequences of *evaluation items*.

2) Condition Tree: It was dual condition tree since it (a) Combined filtering ele-ments (conditions) from the *BooleanExpressionType* and (b) Interconnected them with Boolean operators (AND, OR, NOT and XOR).

The (image retrieve) IR-like condition used *QueryByExample* query type and in-cluded the Base64 encoding of the binary contents of a JPEG image. The (data retrieve) DR-like condition specified, in the present case, that the metadata field *FileSize* must be less than 1000 bytes. Each condition acted over a sequence of *evaluation items* and, for each one, re-turned a value. For IR-like conditions, returned any value in the range of [0.1]. For DR-like conditions returned 1 or 0 (true /false).

A threshold value within a condition was used to indicate the minimum value the score to be processed in the training set.

3.1.6 ISMP Image Retrieval

The ISMP retrieval system over the web interface present de problem image for query and retrieve a list of similar images (as many as possible) from the data base.

3.2 Test Set

The web user interface used both 1) the query-by-image in combination with 2) classic XML metadata-based criteria.

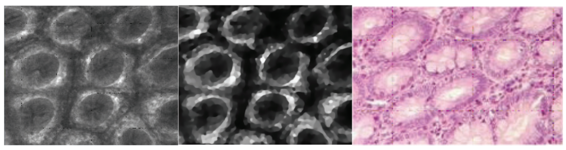

Figure 4: ISMP retrieval in the Test set. Original OB image (left), preprocessing (middle) and retrieved image (right), in this case a histological image.

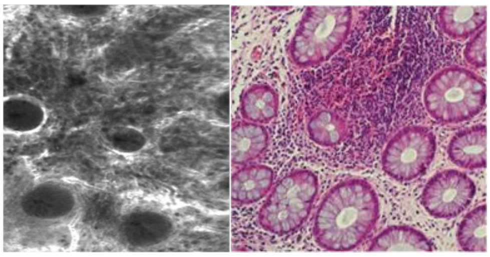

Figure 5: ISMP retrieval in the Test set . Original OB image (left) and retrieved image (right) only histological images were retrieved.

The rate of adequate image retrieval from normal, benign and hyperplasic images using the threshold values indicated in the Section 3.1.5 was 90%.

4 DISCUSSION

The use of solo-medicine (Ferrer-Roca and Marcano, 2009) will be a common practice in a near future, and there-fore professional will require support by any media, including mobile phone. Two are the characteristics of this support: 1) is going to be on line (books will be soon obsolete) and 2) it will be required during patient interventions, therefore the possibility to get access to mobile phones is of paramount importance. The latter is even more relevant considering that the majority of this solo-medicine will be carried out in remote, isolated or developing countries where satellite mobile phones will be, probably, the only available technology.

The solution brought in this paper merge the standardization process of mass-used multimedia standards with the medical image standardization process, and built a ISMP (Image solo medicine paradigm) architecture, that in the present paper was applied to colon OBs (Optical Biopsies).

One of the main functionalities of the ISMP system architecture is the ability to combine conventional search criteria (keywords, metadata ranges) with the direct usage of an example image (query-by-example paradigm) to retrieve similar precedent cases. In the data retrieve DR-like conditions, the MPQF standard acts as a conventional Boolean-based filtering language, while with respect to (Information retrieve) IR-like conditions MPQF acts preserving scores as a fuzzy-logic system. The standard specifies the behaviour of the provided Boolean operators in presence of non-Boolean values.

The result showed that automatic feature-extraction by image analysis on Black & White images coming from the CLE, as well as colour images from surgical specimens, reached the 90% congruency. Thus indicating that the image-query solution proposed in the ISMP architecture is an adequate one to give professional support in medicine at least in the normal and benign cases. Nevertheless we have to test the borderline and malignant ones to detect the sensitivity and specificity of the proposed solution.

ACKNOWLEDGEMENTS

This work has been partly supported by the Spanish government (TEC2008-06692-C02-01) and the CATAI association.

REFERENCES

BMJ Group. *Differential Diagnosis iPhone application*. <http://bestpractice.bmj.com/best-practice/marketing/differentials.html>

CATAI 2010, *En India: El hospital mas eficiente del mundo*. < http://catai.net/blog//2010/06/en-india-el-hospital-mas-eficiente-del-mundo/>.

CATAI 2010, *Anatomia Patologica en iPhone*. <http://catai.net/blog//2010/06/anatomia-patologica-en-iphone/>.

Chen, N., Shatka, H., Blostein, D., 2006. *Use of Figures in Literature Mining for Biomedical Digital Libraries.* Dial, pp.180-197, Second International Conference on Document Image Analysis for Libraries (DIAL'06).

Ferrer-Roca, O., 2009. *Telepathology and optical biopsy.* Int J Telemed Appl Vol: 2009 Pages: 740712 doi:10.1155/2009/740712.

Ferrer-Roca, O., Marcano F., 2009. *Anatomia Patologica Digital. Control de calidad y patoinformatica.*, Rev.Esp.Patol 42(2): 85-85.

Ferrer-Roca, O., Marcano F., 2010. *Computed assisted microscopy. The era of the Small Size Virtual Slides and the 4M microscopes*. HEALTHINF 2010 Vol: 2010 Pp: 517-522. Portugal, ISBN: 978-989-674-018-4.

Ferrer Roca, O., Duval V., Delgado J., Rolim C., Tous R, 2010 *Query by image medical training. Optical Biopsy with confocal endoscopy (OB-CEM).* HEALTHINF 2010 Vol: 2010 Pp: 166-172. Portugal, ISBN: 978-989-674-016-0.

Ferrer-Roca, O., 2010. *Mobile phones in pathology.* J.Telemed & Telecare 16(3): 165.

Hersch W. R., Bhuptiraju R. T., Ross L., Johnson P., Cohen A. M., Kraemer D. F., 2005. *TREC 2004 Genomics Track overview.* Proc of TREC 2004 NIST Special Publication <http://ir.ohsu.edu/genomics>.

ImageJ. <http://rsbweb.nih.gov/ij/>.

Ishii, N., Koike, A., Yamamoto, Y., Takagi T., 2008. *Figure Classification in Biomedical Literature towards Figure Mining*. bibm, pp.263-269, IEEE InternationalConference on Bioinformatics and Biomedicine, 2008.

Kiesslich R., Galle P. R., Neurath M. F., 2008. *Atlas of endomicroscopy*. Springer-Verlag Heidelberg. ISBN 978-3-540-34757-6.

Pietikainen M., Ojala T. and Maenpaa T., 2000. *Gray Scale and Rotation Invariant Texture Classification with Local Binary Patterns*. In Proceedings of the Sixth European Conference on Computer Vision (ECCV2000, pages 404–420.

Sander, et al., 2000. *Density-based Clustering In Spatial Databases: The Algorithm Gdbscan and its Applications*. Data Min. Knowl. Discov., 2(2):169–194.

Tous, R., 2008. *Query formats for multimedia applications ISO/IEC 15938-12 (MPEG Query Format) & ISO/IEC 24800 (JPSearch)*. In CATAI 2009: Super-resolution and optical Biopsy. CATAI editions. Tenerife. Pp25-32.

Wang, T. D., VanDam, J., 2004. *Optical Biopsy: A New Frontier in Endoscopic Detection and Diagnosis.*. Clin Gastroenterol Hepatol 2(9): 744–753.