

Small Bowel Tumors Detection in Capsule Endoscopy by Gaussian Modeling of Color Curvelet Covariance Coefficients

Maria M. Martins, Daniel J. Barbosa, Jaime Ramos and Carlos S. Lima, *Members IEEE*

Abstract— This paper is concerned with the classification of tumoral tissue in the small bowel by using capsule endoscopic images. The followed approach is based on texture classification. Texture descriptors are derived from selected scales of the Discrete Curvelet Transform (DCT). The goal is to take advantage of the high directional sensitivity of the DCT (16 directions) when compared with the Discrete Wavelet Transform (DWT) (3 directions). Second order statistics are then computed in the HSV color space and named Color Curvelet Covariance (3C) coefficients. Finally, these coefficients are modeled by a Gaussian Mixture Model (GMM). Sensitivity of 99% and specificity of 95.19% are obtained in the testing set.

I. INTRODUCTION

UNLIKE conventional endoscopy the endoscopic capsule allows to examine the entire gastrointestinal tract (GI) with a minimal invasive procedure, not requiring highly trained personal for navigating the equipment and diminishing the injury risks for the patients [1]. The patient ingests the CE and the images are captured as the capsule moves through the GI, propelled by peristalsis. More than 50.000 images are acquired per exam at a rate of two per second for about 8 hours [2]. Usually the physician is required to analyze this huge amount of data, selecting the frames that he considers important regarding diagnosis purposes. This task is prone to subjective errors, time consuming (40-60 min) and increases costs. So, it's important to develop a computer assisted diagnosis tool to assist in this task. The texture alteration of the small intestine mucosa has been previously reported in the work of Karkanis *et al.* [3] where the use of statistic textural descriptors has been successfully applied in classification schemes for identification of abnormalities in colonoscopy videos. Therefore this is a promising classification technique for automatic detection of abnormalities. The extraction of a relevant feature set is usually done using wavelet methods.

During the past two decades, wavelets theory has been widely used because wavelets provide a powerful tool for multi-resolution analysis of images [4].

In two-dimensional case the commonly used separable wavelets obtained by a tensor-product of 1D wavelets are only good at capturing discontinuities at edge points, but do not see the smoothness along edges. As a matter of fact,

wavelets transform extracts directional details that only capture horizontal, vertical and diagonal activities in an image, and these three directions cannot in general provide enough directional information in images. In an attempt to overcome the weakness of traditional multiscale representations using wavelets and to capture more directional features in an image, Discrete Curvelet Transform (DCT) has recently emerged as a new multi-resolution analysis tool. This approach is proposed in this paper to extract the textural descriptors from CE video frames.

Texture classification techniques have been based on several classifiers such as Multi layer Perceptron (MLP) networks, Radial Basis Functions (RBF's) and Support Vector Machines (SVM). These techniques have been shown to have a great ability to differentiate the normal regions of abnormal regions, presented by the images of the CE. Unfortunately, these techniques have unavoidable drawbacks when dealing with the modeling of statistical dependence between the various features [5]. In fact statistical dependence is not explicitly modeled in neural based pattern recognition schemes. So, in order to cover this drawback, we discuss in this paper, the use of Gaussian Mixture Model (GMM), which allows extending the approach to time series modeling by extending the GMM to Hidden Markov Model (HMM). In this framework the diagnosis will be based on frame sequence analysis, which means that contextual information can be captured improving the knowledge that can be used in the classification process. The Expectation-Maximization (EM) Algorithm was used for finding maximum likelihood estimates of the GMM parameters and can be easily extended to the HMM framework.

The classification system described in this article uses statistical texture descriptors extracted from Curvelet Discrete Transform, and a Gaussian mixture model as a pattern recognition module.

II. DISCRETE CURVELET TRANSFORM

In an attempt to overcome the weakness of traditional multiscale representations and to capture more directional characteristics in an image, curvelet transform has recently emerged as a new multi-resolution analysis tool. In 2000, Candes and Donoho introduced the Curvelet Transform [6]. The basic idea of curvelet transform is to represent a curve as a superposition of functions of various lengths and widths obeying a specific scaling law. Regarding 2D images, it can be done first by decomposing an image into wavelet sub-

Manuscript received April 1, 2010.

C. S. Lima and Maria M. Martins are with the Industrial Electronics Department, Minho University, Portugal.

Daniel J. Barbosa is with the Catholic University of Leuven, Belgium.

Jaime Ramos is with the Capucho's Hospital, Lisbon, Portugal.

bands, i.e., separating the object into a series of disjoint scales. Curvelet exhibits optimally sparse representation of otherwise smooth objects, and as such it has generated a great deal of interest in recent years [7].

So the Continuous Curvelet Transform (CCT) can be defined by a pair of windows $W(r)$, a radial window, and $V(t)$, an angular window. These are smooth, nonnegative, and real-valued, with W taking positive real arguments and supported on $r \in (1/2, 2)$ and V taking real arguments and supported on $t \in [-1, 1]$. These windows will always obey the admissibility conditions:

$$\sum_{j=-\infty}^{\infty} W^2(2^j r) = 1, \quad r \in \left(\frac{3}{4}, \frac{3}{2}\right) \quad (1)$$

$$\sum_{j=-\infty}^{\infty} V^2(t - 1) = 1, \quad t \in \left(-\frac{1}{2}, 1/2\right) \quad (2)$$

Now, for each $j \geq j_0$, it is introduced the frequency window U_j defined in the frequency domain by:

$$U_j(r, \theta) = 2^{-\frac{3j}{4}} W(2^{-j} r) V\left(\frac{2^{\lfloor j/2 \rfloor} \theta}{2\pi}\right) \quad (3)$$

where $\lfloor j/2 \rfloor$ is the integer part of $j/2$, r and θ are polar coordinates in frequency domain. The support of this window is called polar wedge which depends on the support of W and V .

Assuming U_j is the Fourier transform of a function $\varphi_j(x)$, we can consider $\varphi_j(x)$ as the “mother” curvelet function since all the curvelets at scale 2^j may be obtained by rotations and translations of $\varphi_j(x)$. We further introduce equispaced sequence rotation angles $\theta_l = 2\pi l \cdot 2^{-\frac{j}{2}}, 0 \leq l \leq 2^{\frac{j}{2}} - 1$, and translation parameters $k = (k_1, k_2)$. With the above rotation, a curvelet can be defined as a function of $x = (x_1, x_2)$ at scale 2^j , orientation θ_l and position $\chi_k^{(jL)}$ by

$$\varphi_{j,l,k}(x) = \varphi\left(F_{\theta_l}\left(x - \chi_k^{(jL)}\right)\right) \quad (4)$$

where $\chi_k^{(jL)} = F_{\theta_l}^{-1}\left(k_1 \cdot 2^{-j}, k_2 \cdot 2^{-j}\right)$ and F_{θ} is the rotation matrix in radians:

$$F_{\theta} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \quad (5)$$

A curvelet coefficient is then defined as the inner product of ir denent $f \in L^2(\mathbb{R}^2)$ and a curvelet $\varphi_{j,l,k}$, i.e.,

$$c(j, l, k) = \langle f, \varphi_{j,l,k} \rangle = \int_{\mathbb{R}^2} f(x) \overline{\varphi_{j,l,k}(x)} dx \quad (6)$$

Fig. 1a) shows the polar wedge represented by U_j in the frequency domain. Hence, in Fourier space curvelets are

supported by “parabolic” wedges, and the shaded area represents such a generic wedge.

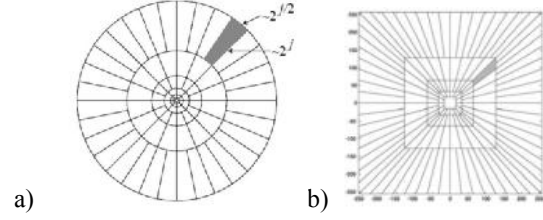


Fig. 1 a) Curvelet wedge in frequency domain; b) Basic wedge of the digital coronization process

Reference [8] proposes two different schemes for the discretization of the CCT, namely the USFFT algorithm and the wrapping algorithm. Both rely in the transformation of the frequency coronae of the CCT of Fig. 1b) in a “Cartesian coronae”, which are based on concentric squares (instead of circles) and shears, in a process designated by digital coronization. The motivation for this digital coronization is the fact that coronae and rotations are not especially adapted to Cartesian arrays, which difficult their computation. Since it is stated that the wrapping algorithm may be simpler to understand and implement, this approach was chosen to calculate the Discrete Curvelet Transform (DCT) in the present work. Further details about the CCT and its discretization can be found in [8]. In the proposed approach, the CE frames were processed with the wrapping algorithm for three scales, with sixteen angles.

III. GAUSSIAN MIXTURE MODEL

Mixture Models are a composite density model which comprises a number of component functions, usually Gaussians. These component functions are combined to provide a multimodal density [9]. They can be employed to model the texture descriptors of an image captured by CE in order to perform tasks such as the classification of the state of the intestine mucosa. Mixture models are a semi-parametric alternative to non-parametric histograms [10] (which can also be used as densities) and provide greater flexibility and precision in modelling the underlying statistics of sampled data.

Once a model is generated, conditional probabilities can be computed for textual descriptors [9]. Let the probability density function $p(\bar{x})$ of vector observations \bar{x} (corresponds to the curvelets coefficients) be a mixture with M component densities:

$$p(\bar{x}) = \sum_{j=1}^M p(\bar{x}|j) c_j \quad (7)$$

where a mixing parameter c_j corresponds to the prior probability of observation \bar{x} has been generated by the density j and where $\sum_{j=1}^M c_j = 1$. Each mixture component is a Gaussian with mean vector $\bar{\mu}$ and covariance matrix Σ :

$$p(\bar{x}|j) = \frac{1}{2\pi|\Sigma_j|^2} \exp^{-\frac{1}{2}(\bar{x}-j)^T \Sigma_j^{-1}(\bar{x}-j)} \quad (8)$$

The Expectation-Maximisation (EM) algorithm [11] is a two step based procedure that iteratively maximizes the likelihood of the training data.

IV. STATISTICAL TEXTURE DESCRIPTORS

Various statistical features can be extracted from the curvelet domain as texture descriptors, and the most common are the mean, standard deviation, energy and entropy of each DCT sub-band [12].

Although the images captured by the capsule are square, the information is restricted to a circular area in the center of the image, by restrictions in the acquisition process. Given the scale anisotropy of the DCT, the calculation of textural descriptors will be made only for the coefficients corresponding to the circular region of CE frames, N .

The proposed texture descriptors on this paper can be calculated as

$$\mu = E\{P(i,j)\} = \frac{1}{N} \sum_i \sum_j P(i,j) \quad (9)$$

$$\sigma = \sqrt{E\{(P(i,j) - \mu)^2\}} = \sqrt{\frac{1}{N} \sum_i \sum_j (P(i,j) - \mu)^2} \quad (10)$$

where $P(i,j)$ is the pixel at position (i,j) .

The most important scales of the DCT, regarding texture encoding information are those in which high frequencies are present. As suggested in [13] the most relevant information for classification purposes are encoded as high frequency content in the scale corresponding to the highest level of detail. Thus, for the scales whose coefficients correspond to low and medium frequency content (coarsest scale coefficients), the texture descriptors weren't computed. Besides, the coarsest scale coefficients of the DCT do not possess directional sensitivity as a consequence for not being directional.

It is expected that textural descriptors correlate among the different color channels as supposed in [3] and these correlations allow distinguishing normal from abnormal texture patterns. Thus, in the present work it is proposed an analogous framework, named as Color Curvelet Covariance (3C) coefficients, where the classification feature is the covariance of textural descriptors in the different color channels. In this framework, the high directional sensitivity of the DCT will be expected to lead to more vigorous descriptors than a similar scheme used in textural descriptors taken from Discrete Wavelet Transform coefficients. The Color Curvelet Covariance of a texture descriptor can be calculated as:

$$CCC(a,b,s,m) = \sum_c (F_m(a,s,\alpha) - E\{F_m(a,s,\alpha)\}) \cdot (F_m(b,s,\alpha) - E\{F_m(b,s,\alpha)\}) \quad (11)$$

where \cdot stands for the multiplication operation, a, b correspond to the different color channels in the covariance computation, F_m is the statistic textural descriptor, α is the angle of the DCT coefficients, s is the DCT scale and $E\{F_m(a,s,\alpha)\}$ the average of the statistical textural descriptor F_m over the different angles α , in the color channel a or b . See that if $i=j$, the previous equation is nothing more than the variance of the textural descriptor F_m with the DCT angle α . In opposition, $CCC(i,j,s,m)$, for $i \neq j$, gives a measure of the likeness of the variation of the textural descriptor F_m taken from the DCT coefficients at different angles, between two color channels.

As the value of the statistical descriptor F_m already has useful information for the texture classification process, and as the proposed $CCC(i,j,s,m)$ features only report for the covariance of F_m between color channels, it is proposed to also include in the feature set the average value $E\{F_m\}$. This portion is the mathematical expectancy for the value of F_m and so does not have as much information as the sequence of the different F_m values taken from the DCT coefficients at different angles, but the insertion of the full sequence would lead to feature vectors with very high dimensionality, which would compromise the training of the classifier and the speed of the classification step.

V. IMPLEMENTATION AND RESULTS

The set of experimental data was built with frames of video capsule endoscopy segments from different patients.

The data set consists of 400 normal frames, which were divided into two sets, one for the Gaussian training and another for the test; and 200 abnormal frames, which were also equally divided into two sets. An example of each type of frame from CE that was used for the experiment is shown in Fig. 2.

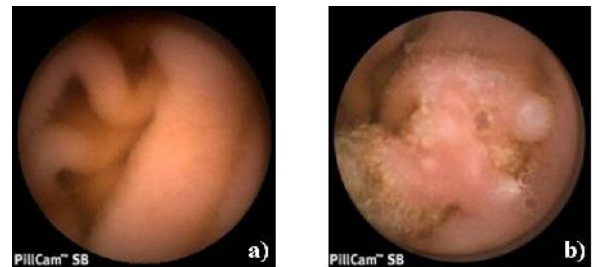


Fig. 2. Example of a normal (a) and an intestinal tumor (b) CE frames

A 2.4 GHz Pentium Core Duo processor-based, with 3 GB of RAM, was used with MATLAB to run the proposed algorithm and the average processing time is approximately 1.5 seconds per frame. As the proposed algorithm has not been optimized for speed, the processing can still be improved. The DCT calculation was done with the routines implemented in the tool CurveLab (available for research purposes at www.curvelet.org). In the proposed approach, the CE frames were processed with the wrapping algorithm with sixteen angles, leading to finest detail coefficients in the DCT domain for each color channel. The selected color

space was the HSV, since is more similar to the physiological perception of human eye [14], and therefore more satisfactory than the standard RGB color space. To evaluate the performance of GMM as a classification system in this approach, we used the values of sensitivity and specificity from previous work [15] that calculated the coefficients with the DCT, as well, but used the Multi layer Perceptron, as a classification system. The feature set was constructed with the Color Curvelet Covariance values, since it was shown in [15], that the Color Curvelet Covariance features can be successfully used for CE frames classification purposes and that the best statistical textural descriptors in the present framework are the mean and the standard deviation of the DCT coefficients.

Instead of measuring the rate of successful recognized patterns, more reliable measures for the evaluation of the classification performance can be achieved by using the sensitivity (true positive rate) and the specificity (100-false positive rate) measures.

Table 1 shows the results for the different classification systems. Note that the proposed system has a significant better sensitivity than the MLP, but in terms of specificity it is slightly poorer.

TABLE I
CLASSIFICATION PERFORMANCE OF DIFFERENT SYSTEMS

Feature Set	GMM	MLP
Specificity (%)	95.19	95.9
Sensitivity (%)	99	95.9

VI. CONCLUSION AND FUTURE WORK

This paper proposes modeling the Color Curvelet Correlation (3C) coefficients by a Gaussian Mixture Model. 3C coefficients are obtained by processing endoscopic capsule imaging in the HSV color space. Although the results can be highly dependent on the used database GMM classification seems to be superior to MLP classification. The advantage of using GMM is that statistical dependence can be explicitly modeled if full covariance matrices are used. Additionally context information can be easily included if GMM is changed to HMM modeling. Inclusion of context information will be considered as near future developments and is frequently used by physicians for improving the diagnosis.

REFERENCES

[1] G. Idden, G. Meron, A. Glukhovsky and P. Swain. "Wireless capsule endoscopy", *Nature*. 2000, pp. pp. 415-417.
[2] Pennazio, M. "Capsule endoscopy: Where are we after 6 years of clinical use?", 2006, Vols. 38, pp. 867-878.
[3] S. Karkanis, D. Iakovidis, D. Maroulis, D. Karras and M. Tzivras. "Computer-Aided Tumor Detection in Endoscopic Video Using Color Wavelet Features". *IEE Trans. On Information Technology in Biomedicine*, Vols. 7, no. 3, Sep. 2003.
[4] N.D. Minh, Martin Vetterli. The finite ridgelet transform for image representation. *IEEE Transactions on Image Processing* 12 (1) pp. 16-28, 2003.

[5] Yanfang Han, Pengfei Shi. An adaptive level-selecting wavelet transform for texture defect detection. *Image and Vision Computing* 25 (2007) pp.1239-1248.
[6] Donoho, E. Candes and D. "Curvelets, multi-resolution representation, and scaling laws". *Wavelet Applications in Signal and Image Processing VIII*, SPIE vol. 4119, pp 1-12, 2000. .
[7] J.L Starck, D.L. Donoho, E.J. Candes. Astronomical image representation by the curvelet transform. *Astronomy and Astrophysics* (398) (2003) pp. 785-800.
[8] Candes E, Demanet L, Donoho D and Ying L. Fast Discrete curvelet transforms. *SIAM Multiscale Modeling Simul*, 2006.
[9] Y. Raja, S. McKenna, and S. Gong. "Segmentation and tracking using colour mixture models," . in *Asian Conference on Computer Vision*, Hong Kong, January 1998.
[10] Bishop, C. *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.
[11] A. P. Dempster, N. M. Laird and D. B. Rubin. "Maximum Likelihood from Incomplete Data via the EM Algorithm," . *Journal of the Royal Statistical Society, B*, vol. 39, no. 1, pp. 1-38, 1977.
[12] Semler, L. Dettori and L. "A comparison of wavelet, ridgelet, and curvelet-based texture classification algorithms in computed tomography," *Computers in Biology and Medicine*, vol. 37, no. 2, pp 486-498, 2007.
[13] D. Barbosa, J. Ramos and C. Lima. "Detection of Small Bowel Tumors in Capsule Endoscopy Frames Using Texture Analysis based on the Discrete Wavelet Transform," in *Proceedings of the IEEE-EMBC2008*, Vancouver, Canada, 2008, pp 3012-3015.
[14] L. Baopu and M. Meng, "Computer-based detection of bleeding and ulcer in wireless capsule endoscopy images by chromaticity moments," *Computers in Biology and Medicine*, vol. 37, no. 4, pp 141-147, 2009
[15] D. Barbosa, J. H. Correia, J. Ramos and C. S. Lima, Texture Classification of Images from Endoscopic Capsule by Using MLP and SVM – A Comparative Approach" *Proceedings of the World Congress on Medical Physics and Biomedical Engineering 2009; Munich 7-12 September 2009*, Vol. 5 pp. 271-274.