



HAL
open science

Characterization of endomicroscopic images of the distal lung for computer-aided diagnosis

Aurélien Saint-Réquier, Benoît Lelandais, Caroline Petitjean, Chesner Désir,
Laurent Heutte, Mathieu Salaün, Luc Thiberville

► **To cite this version:**

Aurélien Saint-Réquier, Benoît Lelandais, Caroline Petitjean, Chesner Désir, Laurent Heutte, et al.. Characterization of endomicroscopic images of the distal lung for computer-aided diagnosis. International Conference on Intelligent Computing, Sep 2009, Ulsan, South Korea. pp.994-1003. hal-00439110

HAL Id: hal-00439110

<https://hal.science/hal-00439110>

Submitted on 6 Dec 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Characterization of endomicroscopic images of the distal lung for computer-aided diagnosis

Aurélien Saint-Réquier¹, Benoît Lelandais¹, Caroline Petitjean¹,
Chesner Désir¹, Laurent Heutte¹, Mathieu Salaün², and Luc Thiberville²

¹ Université de Rouen, LITIS EA 4108

BP 12, 76801 Saint-Etienne-du-Rouvray, France,

² CHU de Rouen, LITIS EA 4108, 76031 Rouen, France

{caroline.petitjean,chesner.desir,laurent.heutte,
mathieu.salaun,luc.thiberville}@univ-rouen.fr

Abstract. This paper presents a new approach for the classification of pathological vs. healthy endomicroscopic images of the alveoli. These images, never seen before, require an adequate description. We investigate two types of feature vector for discrimination: a high-level feature vector based on visual analysis of the images, and a pixel-based, generic feature vector, based on Local Binary Patterns (LBP). Both feature sets are evaluated on state-of-the-art classifiers and an intensive study of the LBP parameters is conducted. Indeed best results are obtained with the LBP-based approach, with correct classification rates reaching up to 91.73% and 97.85% for non-smoking and smoking groups, respectively. Even though tests on extended databases are needed, first results are very encouraging for this difficult task of classifying endomicroscopic images of the distal lung.

1 Introduction

The lungs are the essential respiration organ. They are divided into two anatomic and functional regions: (i) the air conduction system, that includes the trachea, bronchi, and bronchioles, and (ii) the gas-exchange region, or lung parenchyma, made of alveolar sacs. These sacs are made up of clusters of alveoli, tightly wrapped in blood vessels, that allow for gas exchange. Whereas the conduction airways can be explored in vivo during bronchoscopy, the alveolar region was until recently unreachable for in vivo morphological investigation. Therefore, the pathology of the distal lung is currently assessed only in vitro, using invasive techniques such as open lung biopsies. No real time imaging was available.

A new endoscopic technique, called Fibered Confocal Fluorescence Microscopy (FCFM), has recently been developed that enables the visualisation of the more distal regions of the lungs in-vivo [7]. The technique is based on the principle of fluorescence confocal microscopy, where the microscope objective is replaced by a fiberoptic miniprobe, made of thousands of fiber cores. The miniprobe can be introduced into the 2mm working channel of a flexible bronchoscope to produce

in-vivo endomicroscopic imaging of the human respiratory tract in real-time. Real-time alveolar images are continuously recorded during the procedure and stored for further analysis. This very promising technique could replace lung biopsy in the future and might prove to be helpful in a large variety of diseases, including interstitial lung diseases [8].

A clinical trial is currently being conducted that collects FCFM images in several pathological conditions of the distal lungs. This trial also includes a control group of smoker and non smoker healthy volunteers. This strategy provides a dataset of normal images, that can be compared to pathologic ones.

The images recorded within the alveolar regions of the lungs have not been very well described so far. These images represent the alveolar structure, made of elastin fiber, with an approximate resolution of $1\mu\text{m}$ per pixel. This structure appears as a network of (almost) continuous lines. This elastic fiber framework can be altered by distal lung pathologies and as one can see on Figure 1, images acquired on pathological subjects differ from the ones acquired on healthy subjects. The great complexity of these new images justifies the development of reproducible software tools for computer aided diagnosis, that enables automatic image description for diagnosis and follow up of pathological situations.

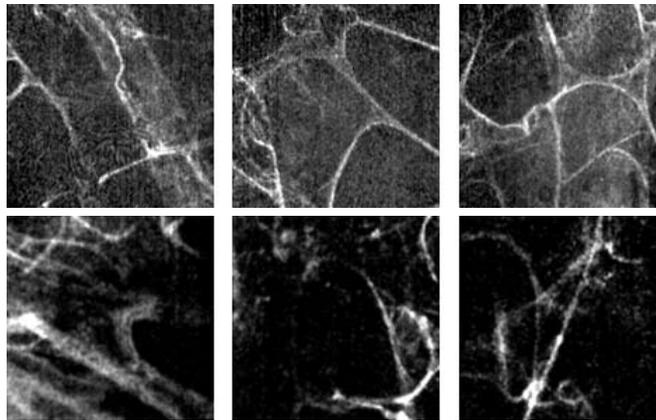


Fig. 1. FCFM images of healthy cases (above) and pathological cases (below)

We wish to conceive and develop methods for the automatic analysis of FCFM images, so as to classify them as healthy cases or pathological cases. As usual when designing such a pattern recognition system, the most crucial step is the choice of a suitable feature space in which healthy and pathological subjects can be discriminated. However, as it is shown in Figure 1, some images present strong similarities where pathological cases can be visually misclassified

for healthy ones and vice versa. The choice of suitable features is therefore critical. Furthermore, to the best of our knowledge, there is no reference work upon which we can rely to find the best way of characterizing these totally new images. Therefore, as it is commonly done when designing the feature extraction stage, two opposite strategies can be investigated to build the feature vector that will describe an image. One can first rely on the visual analysis of the images, searching for general properties that enable to discriminate classes, transforming them numerically into a feature vector. In this case, it boils down to characterizing images by knowledge-based features that one can a priori assume that they will be adapted to the problem. On the other hand, one can design a set of totally unspecific, low-level, pixel-based features that have been proved to be efficient for characterizing a large variety of images. Because we are currently looking for the best way to describe our FCFM images, we propose in this paper to investigate these two opposite strategies for image characterization: problem-dependent, knowledge-based features vs general-purpose, low level features.

Note that a knowledge-based feature vector has already been designed and evaluated on a preliminary database of FCFM images [5]. Our alternative to design a general-purpose feature vector for characterizing our FCFM images has been to investigate the LBP operator, which is a generic tool to describe an image texture [3]. The aim of this work is thus twofold: (i) to study the influence on classification performance of the parameters of an LBP-based characterization, and (ii) to compare performance obtained with generic features and knowledge-based features.

The remaining of this paper is organized as follows: our classification method, including the description of the two feature sets, is described in Section 2, and results and discussion are provided in Section 3. Section 4 concludes and draws some perspectives for this work.

2 Image classification method

2.1 Feature extraction

LBP-based feature set The LBP operator is a powerful means of texture description [3]. It consists in locally identifying and coding some binary patterns and in using their distribution to characterize the image. The original invariant against gray-scale LBP analyzes an eight-pixel neighborhood, using the value of the center pixel as a threshold [3]. The LBP code for a neighborhood is obtained by combining the thresholded neighborhood with powers of two and summing up the result (Figure 2). Histogram of the LBP codes can then be used as a texture descriptor.

The method is extended in [4] by defining a circularly symmetric neighbor set (Figure 2). The number of neighbors and the radius are denoted P and R , respectively. The couple (P, R) is denoted scale factor.

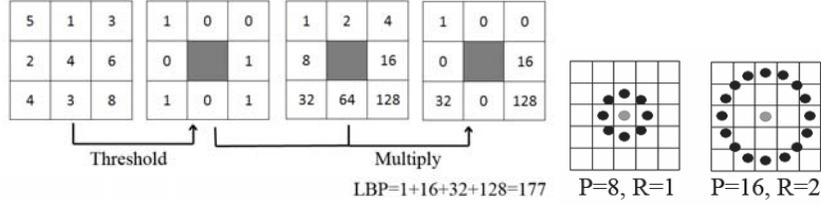


Fig. 2. Circularly symmetric neighbor set for different values of (P, R) (left), the original LBP operator (right)

The $LBP_{P,R}$ operator can be defined as:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) \cdot 2^p$$

where g_c is the gray value of the center pixel, g_p defines the gray value of neighbor p and $s(x)$ is the thresholding function defined as:

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

In order for the LBP code to be rotation invariant, each code is rotated back to a reference position having the smallest LBP code. The rotation-invariant LBP operator, denoted as $LBP_{P,R}^i$, is then defined as :

$$LBP_{P,R}^i = \min\{ROR(LBP_{P,R}, i) \mid i = 0, 1, \dots, P-1\}$$

where $ROR(x, i)$ circularly shifts the P -bit binary number x i times to the right and $\min(\cdot)$ is the minimum function.

Let us now denote by U the number of spatial transitions in a pattern. For example, patterns "00000000" and "11111111" have no spatial transition ($U = 0$), while pattern "00001111" has two transitions in the pattern ($U = 2$). Patterns with U values less than or equal to 2 are considered as "uniform". It has been observed that a vast majority of patterns falls into this category [4]. This allows to define a new operator denoted as $LBP_{P,R}^{riu2}$ that outputs only $P+2$ distinct values:

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & \text{if } U(LBP_{P,R}) \leq 2 \\ P+1 & \text{otherwise} \end{cases}$$

with :

$$U(LBP_{P,R}) = |s(g_{P-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)|$$

This $LBP_{P,R}^{riu2}$ operator is a gray-scale and rotation-invariant measure of the spatial structure of the image texture. For sake of simplicity, the $LBP_{P,R}^{riu2}$ operator is denoted by $LBP_{P,R}$ in the remaining of the paper.

Texture can also be characterized by its contrast, that can be measured with local variance :

$$VAR_{P,R} = \frac{1}{P} \sum_{p=0}^{P-1} (g_p - \mu), \text{ where } \mu = \frac{1}{P} \sum_{p=0}^{P-1} g_p$$

$VAR_{P,R}$ is by definition invariant against rotation. Note that this variance measure has a continuous-valued output, which requires quantization. Quantization intervals are defined by assigning, between each cut value, the same amount of values from the total distribution of variance. As suggested in [4], the number of intervals is empirically determined. Number of intervals between 10 and 22 have been tested on our training set of images and a number of 15 intervals for quantization of the variance values has finally been chosen, as it provided the best recognition rates on our training set of images.

As shown above, the characterization of images through LBP and VAR operators finally requires to choose only the value of the scale factor, i.e. the number P of neighbors and the radius R . However, to the best of our knowledge, there is no commonly agreed rule upon which one can rely to fix its value according to the characteristics of the images we want to classify. Therefore, in our experiments, we investigated several ways to characterize our images in order to determine how these two parameters act on classification performance. For example, one can first test each operator LBP and VAR independently, with varying values for the scale factor (P, R) : (8,1), (16,2), (24,3). Note that a multiresolution analysis can also be performed by concatenating values obtained with operators having varying values of the scale factor. For example, a 2-level LBP operator can be obtained by concatenating $LBP_{8,1}$ values and $LBP_{16,2}$ values. Doing so, three levels of multiresolution analysis can be investigated. Moreover, as $LBP_{P,R}$ and $VAR_{P,R}$ operators allow to respectively characterize the spatial structure and the contrast, one can note that they are of complementary nature. Their combination is thus expected to be a more complete measure of local texture. A new operator denoted by $LV_{P,R}$ can then be created and tested by concatenating the two single operators. Finally, Table 1 lists all the feature vectors that have been tested for comparison and gives their respective feature numbers.

Knowledge-based feature set Several general characteristics can be observed from the visual analysis of the images. As shown in Figure 1, the alveolar structure in healthy subjects can be described as contrasted continuous lines and curves. On the opposite, in the pathological subset, the disorganization of the meshing is illustrated by the numerous irregularities and the tangle of the fibered structures (see Figure 1). Differences are mostly visible for the structure shape, image texture and contrast. A total of 148 features are computed for

Table 1. LBP-based feature vectors used to characterize FCFM images

Resolution levels	P, R	Feature number		
		$LBP_{P,R}$	$VAR_{P,R}$	$LV_{P,R}$
1	8,1	10	15	25
	16,2	18	15	33
	24,3	26	15	41
2	8,1+16,2	28	30	58
	8,1+24,3	36	30	66
	16,2+24,3	44	30	74
3	8,1+16,2+24,3	54	45	99

this knowledge-based feature vector. Number of features are given in parenthesis.

The **structure contrast** seems to be an important property of the images we deal with. For example, we could suppose that pathological images will have higher values of densities than healthy ones because of an emphasized disorganization of the meshing in pathological images. Therefore the structure contrast can be characterized by studying first order pixel gray level distribution and computing pixel densities. The features retained for characterizing the image contrast are: (i) first order statistics on the image histogram (5), (ii) pixel densities obtained on binarized images using Otsu thresholding (1), (iii) the sum of the image gradient values, obtained using Prewitt operator (1).

The **complexity of the structure shape** can be characterized by studying the image skeleton. After skeletonization [1] obtained on the binary image, the number of junction points is computed. One can suppose that on clearly organized, healthy images, this number will be small, contrary to pathological images where the meshing mess will induce a higher number of points (1).

The **image texture** can be characterized by Haralick parameters computed from co-occurrence matrix [2]. Co-occurrence matrix provides the joint distribution of gray-level intensities between two image points. These two points are located according to several configurations, that represent different distances and rotation angles. We chose the following classical 10 translation vectors: [0 1], [-1 1], [-1 0], [-1 -1], [0 2], [-1 2], [-1 -2], [-2 1], [-2 0], [-2 -1]. From the features originally proposed by Haralick, we retain the following ones: energy, contrast, correlation, variance, inverse different moment, entropy, sum average, sum entropy, sum variance, difference entropy, difference variance, and two information measures of correlation. The only discarded feature is the maximum correlation coefficient, which is too computationally expensive. To these 13 parameters we added dissimilarity, a measure of homogeneity [6]. All these 14 parameters are computed over the 10 co-occurrence matrices (140).

2.2 Classifier

On the previously cited feature sets two standard classifiers have been implemented. First a 1-Nearest Neighbour (1-NN) classifier is used to assess the discriminating power of the features. Due to the high computational cost of the 1-NN classifier, we have also implemented a Support Vector Machine (SVM) classifier on our features [9]. SVM is one of the most performing and most used classification algorithm. The support vector machine classifier is a binary classifier algorithm that looks for an optimal hyperplane as a decision function in a high-dimensional space. A classical choice for the kernel is the cubic polynomial kernel.

2.3 Experimental protocol

Because of the large difference between non-smoker and smoker images, experiments have been conducted separately on those two groups. In non-smoker, FCFM images exclusively represent the elastin framework of the alveolar ducts. In smokers, tobacco-tar induced fluorescence allows to observe the alveolar walls and the presence of macrophages (cells which digest cellular debris), as shown in Figure 3. Note that a histogram equalization has been applied to the images, initially quite dark.

The non-smoker database includes 133 images, 31 originating from healthy subjects, and 102 from pathological ones. The smoker database includes 93 images, 60 originating from healthy subjects, and 33 from pathological ones. Because of the relatively small number of images in the non-smoker and the smoker bases, a leave-one-out cross validation process is used, which ensures unbiased generalization error estimation. It consists in extracting one sample image from the image base for validation, the rest of the base being used for learning. Classification rate is computed over all the samples.

3 Results

The influence of the LBP method parameters, i.e. scale factor (P, R), operator combination and number of resolution levels, is first assessed with a 1-NN classifier. Then, LBP-based features and knowledge-based (K) features are compared.

3.1 Influence of LBP parameters

Correct classification rates are provided in Table 2, for LBP, VAR and combination of both operators, on the non-smoker and smoker groups using a 1-NN classifier. Analyzing the results allows to underline the following points:

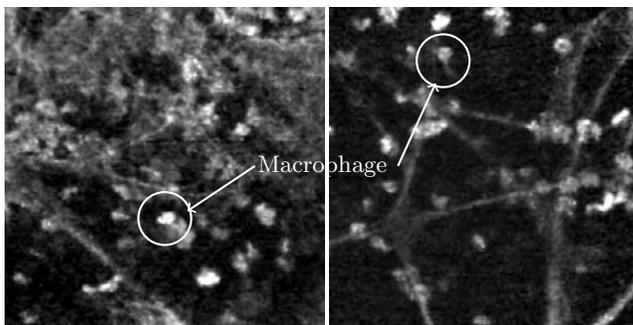


Fig. 3. FCFM images of smoker, healthy (left) and pathological (right) cases. Notice the presence of macrophages.

- The LBP operator is better than the VAR operator, in all cases, when comparing single operators. Recognition rate reaches up to 95.7% for the smoking group and 91.73% for the non-smoking group, whereas the VAR operator only reaches 89.25% and 85.71% on the two groups respectively.
- Combining LBP and VAR operators yields better or equal results than the LBP alone, in 10 out of 14 cases (71%). A more precise analysis of the results shows that the non-smoking group especially takes benefit of the concatenation of both complementary type of vectors. This highlights the difference of content between non smoking and smoking images. Presence of macrophages could explain the perturbation of local contrast.
- Increasing the number of resolution levels does not improve the performance. In 20 out of 24 cases (83%) where resolution is increased, recognition rates are decreased or remain unchanged. This suggests that information contained in the image is not sensitive to the scale factor.

Note that the 3-level $LV_{P,R}$ operator obtains the better compromise of performance for both smoking and non-smoking groups. This has led us to choose it for comparison with the knowledge-based feature vector in the next section.

Table 2. Classification rates (%) on non-smoker and smoker databases

P,R	Non-smoker database			Smoker database		
	$LBP_{P,R}$	$VAR_{P,R}$	$LV_{P,R}$	$LBP_{P,R}$	$VAR_{P,R}$	$LV_{P,R}$
8,1	91,73	85,71	90,98	93,55	86,02	92,47
16,2	87,97	82,71	89,47	94,62	83,87	94,62
24,3	84,21	81,20	88,72	94,62	75,27	93,55
8,1+16,2	90,98	84,21	90,98	95,7	89,25	94,62
8,1+24,3	83,46	84,21	90,98	94,62	87,1	94,62
16,2+24,3	89,47	82,71	90,98	94,62	83,87	94,62
8,1+16,2+24,3	87,21	83,46	91,73	94,62	86,02	94,62

3.2 Comparative study between LBP-based features and knowledge-based features

In this section, performance obtained with the knowledge-based feature set and the $LV_{P,R}$, 3-level operator, are compared. Results provided in Table 3 include recognition and error rates, as well as number of false positive (FP) and false negative (FN). False negative is the number of healthy instances that are erroneously reported as pathological and false positive is the number of pathological cases considered healthy.

LBP-based features yield better results on both databases, and whatever the classifier used. Note that, as observed in the previous study, results on the smoker group are better than those obtained on the non smoker group. For the smoker group, the 2 recognition rates provided by the 2 feature sets are pretty close for a given classifier. Performance are increased by about 1.6%, vs 5.7% for the non-smoker group. This might suggest that smoker images could benefit from a higher level description in addition to the low-level, LBP-based feature vector. Nonetheless, the image database needs to be extended, and FP and FN are still quite high, which is encouraging us to continue our work.

Table 3. Comparison between results obtained by 1-NN and best results obtained by the cascade for the KFV and LBPFV

	Non-smoker database				Smoker database			
	LBP (99)		K (148)		LBP (99)		K (148)	
	1-NN	SVM	1-NN	SVM	1-NN	SVM	1-NN	SVM
Recognition rate (%)	91,73	91,73	86,47	85,71	94,62	97,85	92,47	96,77
Error rate (%)	8,27	8,27	13,53	14,29	5,38	2,15	7,53	3,23
FP (number of images)	5	6	7	10	1	1	3	2
FN (number of images)	6	5	11	9	4	1	4	1

4 Conclusions

We presented in this paper an original system for the automatic classification of normal versus abnormal endomicroscopic images of the respiratory alveolar system with an emphasis on feature extraction. The first evaluated feature vector does not require any a priori knowledge and is pixel-based. It includes a local description of spatial structure (LBP operator) and contrast (VAR operator) of the image texture. Several configurations are possible with these two operators, which have been considered for the experiments (operator combination, number of resolution levels, different values for the scale factor). The second feature set is a combination of descriptors (histogram, density, gradient, skeleton, cooccurrence matrices) deduced from a visual inspection of healthy and pathological images.

Recognition rates obtained after SVM based classification reach 91.73% and 97.85% with the LBP approach on non-smoker and smoker databases respectively, whereas they are 85.71% and 96.77% with the knowledge-based feature vector. These results show that the LBP-based feature vector is more adapted to the description of FCFM images, especially the non-smoker images. In particular, the study of the different configurations shows that combining LBP and VAR operators increases correct classification rates. This could suggest to further investigate combination of LBP and VAR operators, by computing for instance their joint distribution, or combining LBP and VAR with different values of P and R . The behavior difference of the two image databases (smoking and non-smoking) suggests that different feature sets could be also required for these two groups.

Future work will concern rendering the process real-time, so as to aid the clinician during examination in real time. Classification methods could also give information about which part of the image is the most discriminant or which part of the structure might be more altered by pathologies. A future goal will also be to discriminate between different pathologies : interstitial lung diseases (abestosis, systemic sclerosis, fibrosis, sarcoidosis), carcinomatous lesions etc.

References

1. G.S. Dibajaa and E. Thiel: Skeletonization algorithm running on path-based distance maps. *Image and Vision Computing*, vol.14, p.47-57 (1996)
2. R.M. Haralick, K. Shanmugam and I. Dinstein: Textural Features for Image Classification. *Systems, Man and Cybernetics*, vol.3, no.6, p.610-621 (1973)
3. T. Ojala, M. Pietikäinen and D. Harwood: A comparative study of texture measures with classification based on featured distribution. *Pattern Recognition*, 29(1):51-59 (1996)
4. T. Ojala, M. Pietikäinen and T. Mäenpää: Multiresolution gray-scale and rotation invariant texture classification with Local Binary Patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971-987 (2002)
5. C. Petitjean, J. Benoist, L. Thiberville, M. Salaün and L. Heutte: Classification of In-vivo Endomicroscopic Images of the Alveolar Respiratory System. In *IAPR Conference on Machine Vision Applications*, Yokohama, Japan (2009)
6. W.K. Pratt: *Digital Image Processing*, 2nd Edition, John Wiley & Sons (1991)
7. L. Thiberville, S. Moreno-Swirc, T. Vercauteren, E. Peltier, C. Cave and G. Bourg-Heckly: In vivo imaging of the bronchial wall microstructure using fibered confocal fluorescence microscopy. *American Journal of Respiratory and Critical Care Medicine*, vol.175, no.1, p.22-31 (2007)
8. L. Thiberville, G.Bourg-Heckly, M. Salaün, S. Dominique and S. Moreno-Swirc: Human in-vivo confocal microscopic imaging of the distal bronchioles and alveoli. *Chest Journal*, vol.132, no.4, p.426 (2007)
9. V. Vapnik: *The nature of statistical learning theory*. Springer (1995)