# Retrieval of 4D Dual Energy CT for Pulmonary Embolism Diagnosis

Antonio Foncubierta–Rodríguez[1],
Alejandro Vargas[2], Alexandra Platon[2], Pierre–Alexandre Poletti[2],
Henning Müller[1,2], and Adrien Depeursinge[1,2]

[1]University of Applied Sciences Western Switzerland (HES–SO),
[2]University and University Hospitals of Geneva (HUG), Switzerland,
antonio.foncubierta@hevs.ch

**Abstract.** Pulmonary embolism is a common condition with high short–term morbidity. Pulmonary embolism can be treated successfully but diagnosis remains difficult due to the large variability of symptoms, which are often non–specific including breath shortness, chest pain and cough. Dual energy CT produces 4–dimensional data by acquiring variation of attenuation with respect to spatial coordinates and also with respect to the energy level. This additional information opens the possibility of discriminating tissue with specific material content, such as bone and adjacent contrast. Despite having already been available for clinical use for a while, there are few applications where Dual energy CT is currently showing a clear clinical advantage. In this article we propose to use the additional energy–level data in a 4D dataset to quantify texture changes in lung parenchyma as a way of finding parenchyma perfusion deficits characteristic of pulmonary embolism.

**Keywords:** 4D texture, pulmonary embolism, dual energy CT

## 1 Introduction

Pulmonary embolism (PE) consists of the obstruction of one or several arteries in the lungs as a complication of deep vein thrombosis (DVT). Studies have shown that acute pulmonary embolism mortality rates can reach 75% during initial hospital admission [1] and 30% after 3 years of hospital discharge [2]. Pulmonary embolism is an avoidable cause of death if treated immediately with anticoagulants. Delays in diagnosis of pulmonary embolism have shown to increase the risk of death [3, 4], making early diagnosis a key factor for successful treatment.

Schwickert et al. [5] showed that pulmonary embolism induced wedge–shaped pleura–based regions of heterogeneous increased attenuation in unenhanced computed tomography (CT) scans for 54 over 75 patients that were also visible on contrast–enhanced CT. Ganeshan et al. [6] observed that simple 3D texture attributes correlated well with ventilation and vascularization of the lung

parenchyma in contrast–enhanced CT scans. Texture information is therefore relevant to quantify pulmonary parenchyma ischemia in CT imaging.

Dual–energy computer tomography (DECT) contains 4D data: three spatial dimensions and the level of x–ray energy between 40 and 140 keV used for image acquisition. Iodine components from the contrast product have differing contrasts at varying energies and are related to the perfusion of the lung parenchyma. Several studies showed the value of DECT to quantify perfusion defects of the lung parenchyma [7–11] using iodine components, which can be derived from CT attenuation at two energy levels of 80 and 140 keV.

This work investigates the use of texture–based image retrieval of 4D DECT, where several energy levels add relevant information about the perfusion of the lung parenchyma. Texture is very hard to visualize for dimensions higher than 2D. Since DECT acquires not only spatially–sampled but also energy–sampled data, visual information is even more difficult to understand in an intuitive way. Computer–assisted analysis of DECT is therefore required. In this article, texture is described by using the wavelet transform together with a visual words approach. The use of texture analysis techniques and challenging 4D data allows to evaluate the possible impact of DECT in early diagnosis of pulmonary embolism.

## 2 Materials and methods

This section presents the data of 4D DECT images from PE patients and our approach to 4D texture quantification based on two main ideas: a 3D wavelet transform for multi–scale texture descriptors for each scale and visual words [12, 13], as a way of obtaining features based on patterns actually occurring in the data set.

### 2.1 Dataset

Pulmonary parenchyma ischemia in 4D dual energy CT (DECT) images were identified in collaboration with the emergency radiology of the University Hospitals of Geneva. A small set of 13 currently annotated patients was used to train and test the techniques.

For each patient, the five pulmonary lobes were manually segmented, and the Qanadli index [14] was computed as a measure of the obstruction on a lobe basis. The Qanadli index is computed by adding a score per artery in the lobe: 0 if there is no obstruction, 1 if there is partial obstruction and 2 if the artery is completely obstructed. The maximum value of the Qanadli index varies among lobes, depending on the number of arteries. The Qanadli index value was normalized using the maximum value per lobe, obtaining the percentage of the Qanadli index $Q(\%)$ as a measure of the pulmonary embolism severity.

The images in the dataset contain approximately 300 slices per patient and energy level. Energy levels are sampled from 40 to 140 keV in steps of 10 keV. The

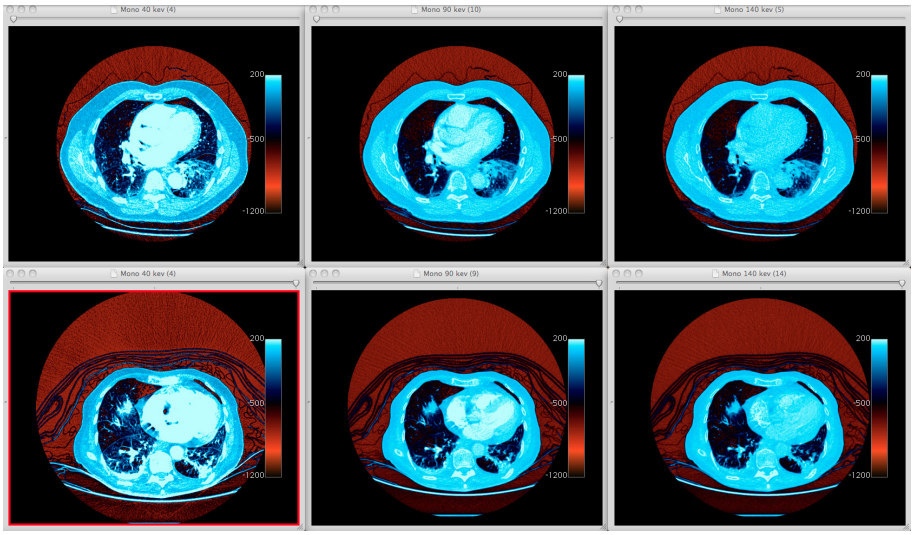

**Fig. 1.** DECT example with slices at three energy levels per row: 40, 90 and 140 keV.

total amount of data per patient is approximately $512 \times 512 \times 300 \times 11 = 865.08$ million voxels. Example images can be seen in Figure 1.

The image resolution is approximately isotropic in the spatial coordinates, with horizontal resolution of $0.83mm/voxel$ and vertical resolution of $1mm/voxel$. This allows for 3D analysis in the spatial domain, whereas the energy domain needs to be taken into account separately, given the different nature of the data. Therefore, experiments were conducted using a spatial, three–dimensional multi–resolution analysis of each energy level. Visual features were then aggregated through energy levels using a visual words approach.

### 2.2 4D texture analysis

**Energy of wavelet coefficients** The wavelet transform has been widely used to analyze images and videos at multiple resolutions [15–17]. In medical imaging, it showed to accurately characterize the lung parenchyma [18, 19, 12]. The mother wavelet chosen for texture description is the Difference of Gaussians (DoG). Since the image sampling is not fully isotropic, the multi–resolution analysis is based on the Gaussian function $g$ calculated in physical dimensions by scaling the variables $x$, $y$ and $z$ using the corresponding values for the voxel spacing in each direction ($\delta_x$,$\delta_y$,$\delta_z$):

$$g_{\boldsymbol{\sigma}}(\boldsymbol{x}) = \frac{1}{\sigma_x \sigma_y \sigma_z \sqrt{(2\pi)^3}} e^{-\left(\frac{(x\delta_x)^2}{2\sigma_x^2} + \frac{(y\delta_y)^2}{2\sigma_y^2} + \frac{(z\delta_z)^2}{2\sigma_z^2}\right)}. \quad (1)$$

The choice of using Gaussian functions is based on their good isotropic properties, which allow to analyze image texture without making prior choices of orientation, as opposed to co–ocurrence matrix based methods [20, 21].

The extracted coefficients are obtained by using the difference of Gaussians (DoG), which provides a good approximation to the Laplacian of Gaussians or Mexican Hat wavelets when the variance parameters $\sigma_{1,2}$ of two Gaussian functions $g_{1,2}$ satisfy $\sigma_2 \approx 1.6\sigma_1$. Since the functions are calculated in physical dimensions, there is no need for having anisotropic variance parameters of $g_{1,2}$. The resulting wavelets are shown in Equation 2. The number of scales $j$ used for the wavelet transform is five, with $j$ ranging from 1 to 5.

$$\psi_j(\boldsymbol{x}) = g_{\boldsymbol{\sigma_1}}(\boldsymbol{x}) - g_{\boldsymbol{\sigma_2}}(\boldsymbol{x}). \tag{2}$$

$$\boldsymbol{\sigma_2} = 1.6\boldsymbol{\sigma_1}. \tag{3}$$

$$\boldsymbol{\sigma_1} = \mathbf{2}^j. \tag{4}$$

The wavelet admissibility condition forces the mother function to have zero–mean. Therefore, the images filtered with $\psi_j(\boldsymbol{x})$ will be bandpass images (i.e., also zero–mean images). For this reason the raw wavelet coefficients are not used, since in the clustering phase all clusters would be located around $\mathbf{0}$. Instead, the energy of the wavelet coefficients averaged in a small neighborhood was chosen as a local feature for describing the images.

Given a voxel identified by its position $\boldsymbol{x}$, and a neighborhood $\mathcal{N}$ with $S_\mathcal{N}$ elements, the energy of the wavelet coefficients in the neighborhood is computed as the sum of the squared coefficients within the neighborhood. Since voxels near the boundaries of the image have fewer neighbors than those far from the boundaries, the mean energy over$\mathcal{N}$, $E_w(\boldsymbol{x})$, was chosen instead of the energy, as shown in equation 5. The neighborhood size was kept small ($6 \times 6 \times 6$) to maintain a sufficiently local operator on the images. Although the energy was averaged in neighborhoods, the wavelet transform was applied to the complete 3–dimensional image to avoid border effects.

$$E_w(\boldsymbol{x}) = \frac{1}{S_\mathcal{N}} \sum_{\boldsymbol{x} \in \mathcal{N}} \psi_j(\boldsymbol{x})^2. \tag{5}$$

**Visual words** The term *texture* often has a fuzzy definition and refers to the (sometimes regular or periodic) visual characteristics of the pixel values within a certain region and their relationships, which is not always explicit to human observers. Since the wavelet transform can describe the transient of the values in the voxel surroundings, a way of aggregating this information for all regions of interest is needed in order to describe the concept of texture which refers to a scale larger than voxels.

Visual words [22, 12] have been widely used in image retrieval and image classification for describing image content (or regions of interest). The approach is similar to the bag–of–words approach used for text retrieval or text similarity matching [23].

For each voxel, this technique maps a set of continuous low–level features in a given neighborhood, e.g. gray values or wavelet coefficients, into a compact discrete representation consisting of visual words.

Visual words are cluster centers in the feature space derived from the energy of wavelet coefficients across different scales and energy levels. This guarantees to have a set of visual features actually corresponding to discriminative patterns that do occur in the database. Every image or region is subsequently described by the histogram of the visual words within this region.

**Definition 1** *Let $\mathcal{F} = \{f_1, f_2, \ldots, f_m\}$ be the set of $m$ descriptors, $f_i \in \mathbb{R}^N$, describing visual characteristics of a given set of images or image regions. A visual vocabulary $W_{\mathcal{F},k} = \{w_1, w_2, \ldots, w_k\}$, with $w_i \in \mathbb{R}^N$ is constructed by grouping the elements of $\mathcal{F}$ into $k$ disjoint subsets or words, and selecting their $k$ centroids $w_j$ with $j \in \{1, \ldots, k\}$.*

*The bag–of–visual–words of an image or region $I$, described by $m_I$ visual descriptors $\{f_1, f_2, \ldots, f_{m_I}\}$, is defined as a vector $h_I = \{c_1, c_2, \ldots, c_k\}$ :*

$$c_j = \sum_{i=1}^{m_I} g_j(f_i) \quad \forall j \in \{1, \ldots, k\}$$

*where*

$$g_j(f) = \begin{cases} 1 \text{ if } d(f, w_j) \leq d(f, w_l) & \forall l \in \{1, \ldots, k\} \\ 0 \text{ otherwise} \end{cases}$$

*being $d(f, w)$ the distance between two vectors $f$ and $w$.*

### 2.3 Experimental setup

First, the images were analyzed using the wavelet transform at two to five scales; the energy of these wavelets in a $6 \times 6 \times 6$ neighborhood was computed. A mask of the lobes was used, storing only the values for the voxels contained in the lung lobes. This process was repeated for all patients and all energy levels.

Given the small number of patients available, the leave–one–patient–out cross–validation method was chosen. Leaving the features from one patient out at a time, k–means clustering was carried out on the rest of the feature space for all the concatenated features to find the visual words; five wavelet scales for each energy level. The number of clusters or visual words ranged from 50 to 150 in steps of 50 as this was computationally feasible.

For all images, the voxels were labeled with the identifier of the nearest cluster center of the vocabulary (e.g., visual word) built from all other patients. For each lobe, the histogram of visual words within the lobe was considered as an instance for retrieval using the $k$ nearest neighbors and a Euclidean distance. Figure 2 shows an overview of the complete experimental configuration with a focus on the relative dimensionality of the feature spaces before and after the visual word assignment.

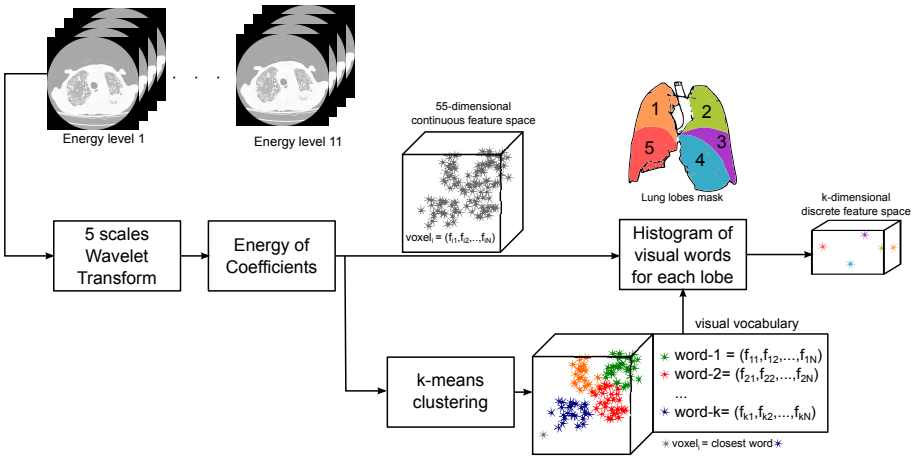


**Fig. 2.** System overview showing the data processing pipeline for each patient.

## 3 Results

In order to evaluate the retrieval precision, a relevance criterion needs to be defined. Since the degree of obstruction of obstruction of the lobe is quantified via the Qanadli index, and given the reduced size of the dataset, it is not possible to define relevance with a strictly equal condition based on the Qanadli index, only.

In this experiment, all lobes with Qanadli index larger than zero were considered not healthy, whereas only the lobes with Qanadli index equal to zero were considered healthy. Precision at one (P@1), five (P@5) and ten (P@10) were computed for a varying number of visual words, and wavelet scales. Table 1 shows the precision values for the DECT images (4D) whereas Table 1 shows the corresponding results for a single–energy 70KeV CT (3D). This energy level was chosen because it is the standard energy level used by radiologists to diagnose pulmonary embolism.

Table 3 shows the confusion matrices for the best performing configurations for the top retrieved lobes. Tables 4 and 5 show the corresponding matrices for precision of the first five and the first ten retrieved lobes.

| Visual Words | Scales | P@1(%) | P@5(%) | P@10(%) |
|---|---|---|---|---|
| 50 | 1 | 55 | 56 | 56 |
| 100 | 1 | 58 | 55 | 57 |
| 150 | 1 | 58 | 56 | 56 |
| 50 | 2 | 62 | 58 | 55 |
| 100 | 2 | 62 | **62** | **60** |
| 150 | 2 | **63** | **62** | **60** |
| 50 | 3 | 58 | 54 | 55 |
| 100 | 3 | 60 | 59 | 58 |
| 150 | 3 | 57 | **62** | 58 |
| 50 | 5 | 45 | 52 | 51 |
| 100 | 5 | 57 | 52 | 51 |
| 150 | 5 | 58 | 52 | 52 |

**Table 1.** Precision values for a varying number of visual words with 4D data, including all energy levels from 40KeV to 140KeV.

| Visual Words | Scales | P@1(%) | P@5(%) | P@10(%) |
|---|---|---|---|---|
| 50 | 1 | 60 | 56 | **56** |
| 100 | 1 | 57 | 57 | 55 |
| 150 | 1 | 58 | 56 | 54 |
| 50 | 2 | 55 | 57 | 55 |
| 100 | 2 | 58 | 57 | **56** |
| 150 | 2 | **63** | **60** | 54 |
| 50 | 3 | 51 | 55 | 53 |
| 100 | 3 | 49 | 53 | 55 |
| 150 | 3 | 55 | 55 | 55 |
| 50 | 5 | 45 | 50 | 52 |
| 100 | 5 | 51 | 49 | 52 |
| 150 | 5 | 51 | 52 | 53 |

**Table 2.** Precision values for a varying number of visual words with 3D data only (i.e. 70KeV).

**Table 3.** Confusion matrices for the best performing configurations for the first retrieved lobes.

(a) Dual energy CT (4D) with 2 wavelet scales per energy level and 150 visual words. P@1=63%

| | Healthy (%) | Not healthy (%) |
|---|---|---|
| Healthy | 47.6 | 52.4 |
| Not healthy | 29.5 | 70.5 |

(b) Single energy CT (3D) with 2 wavelet scales per energy level and 150 visual words. P@1=63%

| | Healthy (%) | Not healthy (%) |
|---|---|---|
| Healthy | 52.4 | 47.6 |
| Not healthy | 31.8 | 68.2 |

**Table 4.** Confusion matrices for the best performing configurations for the first five retrieved lobes.

(a) Dual energy CT (4D) with 2 scales per energy level and 100 visual words. P@5=62%

| | Healthy (%) | Not healthy (%) |
|---|---|---|
| Healthy | 43.8 | 56.2 |
| Not healthy | 29.5 | 70.5 |

(b) Dual energy CT (4D) with 2 wavelet scales per energy level and 150 visual words. P@5=62%

| | Healthy (%) | Not healthy (%) |
|---|---|---|
| Healthy | 43.8 | 56.2 |
| Not healthy | 29.9 | 70.1 |

(c) Dual energy CT (4D) with 3 wavelet scales per energy level and 150 visual words. P@5=62%

| Healthy (%) | Not healthy (%) | |
|---|---|---|
| Healthy | 44.8 | 55.2 |
| Not healthy | 29.1 | 70.9 |

**Table 5.** Confusion matrices for the best performing configurations for the first ten retrieved lobes.

(a) Dual energy CT (4D) with 2 wavelet scales per energy level and 100 visual words. P@10=60%

| | Healthy (%) | Not healthy (%) |
|---|---|---|
| Healthy | 42.4 | 57.6 |
| Not healthy | 31.1 | 68.9 |

(b) Dual energy CT (4D) with 2 wavelet scales per energy level and 150 visual words. P@10=60%

| | Healthy (%) | Not healthy (%) |
|---|---|---|
| Healthy | 40 | 60 |
| Not healthy | 31.1 | 68.9 |

## 4 Discussion and conclusions

This article presents an approach for solid texture analysis on 4D dual energy CT data to better detect and quantify pulmonary embolisms for a more efficient and faster treatment in emergency radiology by retrieving visually similar cases to support clinical decisions. To the best of the author's knowledge such an apprach has not yet been described in the literature.

Results in Section 3 show that four dimensional texture contains patterns related to the pulmonary embolism severity. Six out of the seven best performing configurations are obtained with dual energy data. The number of visual words and wavelet scales also have an impact on the accuracy. The optimal configuration for the multiscale framework was found for 2 scales, and the number of visual words that performed best was 150 visual words, which agrees with [12].

The confusion matrices from Tables 3, 4 and 5 show that the accuracy for PE lobes is much better than for healthy lobes. There are two possible explanations for these results. First, the dataset contains only patients with PE, which can affect the lung parenchyma of the lobes without embolisms which can be overloaded by redirected blood flows. Second, healthy lobes can contain a very rich set of patterns possibly linked to other diseases.

In this paper, a novel use of dual energy CT data is proposed. The limitations of visual inspection for higher dimensional data are underlined and an automatic analysis is suggested as an aid for clinicians for detection and quantization of pulmonary embolisms. Despite the small size of the dataset, the relationship between four dimensional texture and pulmonary embolism severity was shown, proposing a new diagnosis tool that may help particularly emergency radiologists to diagnose and quantify pulmonary embolism and potentially reduce mortality through quicker and more accurate treatment.

A small scale evaluation was performed as there are currently no large data sets with corresponding ground truth available and we only started acquiring new cases with a DECT protocol. This only allows a few basic conclusions on the techniques employed. However, the results state a baseline for future work and point at the use of medical image computing as a means to overcome the limitations of human perception and understanding of high dimensional image data. Nevertheless, computational complexity is very high: the offline feature extraction took approximately 4 hours per patient on a 24–core machine with 96 gigabytes of main memory.

We currently work on creating a larger database to allow for a better evaluation and also for more training data for our system. The future dataset is foreseen to contain control cases that will allow for a better characterization of the healthy lobes.

## 5 Acknowledgments

This work was partially supported by the Swiss National Science Foundation (FNS) in the MANY project (grant 205321–130046), the EU $7^{th}$ Framework Program in the context of the Khresmoi project (FP7–257528), and the Center for Biomedical Imaging (CIBM).

## References

1. Goldhaber, S.Z., Visani, L., Rosa, M.D.: Acute pulmonary embolism: clinical outcomes in the international cooperative pulmonary embolism registry (icoper). The Lancet **353**(9162) (1999) 1386 – 1389
2. Anderson, Frederick A., J., Wheeler, H.B., Goldberg, R.J., Hosmer, D.W., Patwardhan, N.A., Jovanovic, B., Forcier, A., Dalen, J.E.: A population-based perspective of the hospital incidence and case-fatality rates of deep vein thrombosis and pulmonary embolism: The worcester dvt study. Archives of Internal Medicine **151**(5) (1991) 933–938
3. Alonso-Martínez, J., Sánchez, F.A., Echezarreta, M.U.: Delay and misdiagnosis in sub-massive and non-massive acute pulmonary embolism. European Journal of Internal Medicine **21**(4) (2010) 278 – 282
4. Ozsu, S., Oztuna, F., Bulbul, Y., Topbas, M., Ozlu, T., Kosucu, P., Ozsu, A.: The role of risk factors in delayed diagnosis of pulmonary embolism. The American Journal of Emergency Medicine **29**(1) (2011) 26 – 32
5. Schwickert, H.C., Schweden, F., Schild, H.H., Piepenburg, R., Düber, C., Kauczor, H.U., Renner, C., Iversen, S., Thelen, M.: Pulmonary arteries and lung parenchyma in chronic pulmonary embolism: preoperative and postoperative CT findings. Radiology **191**(2) (1994) 351–357
6. Ganeshan, B., Miles, K.A., Young, R.C.D., Chatwin, C.R.: Three–dimensional selective–scale texture analysis of computed tomography pulmonary angiograms. Investigative Radiology **43**(6) (June 2008) 382–394
7. Chae, E.J., Seo, J.B., Jang, Y.M., Krauß, B., Lee, C.W., Lee, H.J., Song, K.S.: Dual–energy CT for assessment of the severity of acute pulmonary embolism: Pulmonary perfusion defect score compared with CT angiographic obstruction

score and right ventricular/left ventricular diameter ratio. American Journal of Roentgenology **194**(3) (2010) 604–610
8. Lee, C., Seo, J., Song, J.W., Kim, M.Y., Lee, H., Park, Y., Chae, E., Jang, Y., Kim, N., Krauß, B.: Evaluation of computer–aided detection and dual energy software in detection of peripheral pulmonary embolism on dual-energy pulmonary CT angiography. European Radiology **21**(1) (2011) 54–62
9. Nakazawa, T., Watanabe, Y., Hori, Y., Kiso, K., Higashi, M., Itoh, T., Naito, H.: Lung perfused blood volume images with dual–energy computed tomography for chronic thromboembolic pulmonary hypertension: Correlation to scintigraphy with single–photon emission computed tomography. Journal of Computer Assisted Tomography **35**(5) (2011) 590–595
10. Thieme, S.F., Becker, C.R., Hacker, M., Nikolaou, K., Reiser, M.F., Johnson, T.R.C.: Dual energy CT for the assessment of lung perfusion—correlation to scintigraphy. European Journal of Radiology **68**(3) (2008) 369–374
11. Thieme, S.F., Johnson, T.R.C., Lee, C., McWilliams, J., Becker, C.R., Reiser, M.F., Nikolaou, K.: Dual–energy CT for the assessment of contrast material distribution in the pulmonary parenchyma. American Journal of Roentgenology **193**(1) (2009) 144–149
12. Foncubierta-Rodríguez, A., Depeursinge, A., Müller, H.: Using multiscale visual words for lung texture classification and retrieval. In Greenspan, H., Müller, H., Syeda Mahmood, T., eds.: Medical Content–based Retrieval for Clinical Decision Support. Volume 7075 of MCBR–CDS 2011., Lecture Notes in Computer Sciences (LNCS) (September 2012) 69–79
13. Avni, U., Greenspan, H., Konen, E., Sharon, M., Goldberger, J.: X–ray categorization and retrieval on the organ and pathology level, using patch–based visual words. IEEE Transactions on Medical Imaging **30**(3) (2011) 733–746
14. Qanadli, S.D., El Hajjam, M., Vieillard-Baron, A., Joseph, T., Mesurolle, B., Oliva, V.L., Barré, O., Bruckert, F., Dubourg, O., Lacombe, P.: New CT index to quantify arterial obstruction in pulmonary embolism. American Journal of Roentgenology **176**(6) (2001) 1415–1420
15. Nevel, A.V.: Texture classification using wavelet frame decompositions. In: Conference Record of the Thirty–First Asilomar Conference on Signals, Systems & Computers. Volume 1. (November 1997) 311–314
16. Smith, J.R., Lin, C.Y., Naphade, M.: Video texture indexing using spatio–temporal wavelets. Proceedings of the International Conference on Image Processing **2** (September 2002) 437–440
17. Chenouard, N., Unser, M.: 3D steerable wavelets and monogenic analysis for bioimaging. In: 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. (April 2011) 2132–2135
18. Korfiatis, P., Skiadopoulos, S., Sakellaropoulos, P., Kalogeropoulou, C., Costaridou, L.: Automated 3D segmentation of lung fields in thin slice CT exploiting wavelet preprocessing. Computer Analysis of Images and Patterns **4673** (August 2007) 237–244
19. Korfiatis, P., Kalogeropoulou, C., Karahaliou, A., Kazantzi, A., Skiadopoulos, S., Costaridou, L.: Texture classification–based segmentation of lung affected by interstitial pneumonia in high–resolution CT. Medical Physics **35**(12) (2008) 5290–5302
20. Korfiatis, P.D., Kalogeropoulou, C., Karahaliou, A.N., Kazantzi, A.D., Costaridou, L.I.: Vessel tree segmentation in presence of interstitial lung disease in MDCT. IEEE Transactions on Information Technology in Biomedicine **15**(2) (March 2011) 214–220

21. Kovalev, V.A., Kruggel, F.: Texture anisotropy of the brain's white matter as revealed by anatomical MRI. IEEE Transactions on Medical Imaging **26**(5) (May 2007) 678–685
22. Sivic, J., Zisserman, A.: Video google: A text retrieval approach to object matching in videos. In: Proceedings of the Ninth IEEE International Conference on Computer Vision - Volume 2. ICCV '03, Washington, DC, USA, IEEE Computer Society (2003) 1470–1477
23. van Rijsbergen, C.J.: Information Retrieval. Prentice Hall, Englewood Cliffs, New Jersey, USA (1979)