3D Riesz–wavelet based Covariance descriptors for texture classification of lung nodule tissue in CT

Pol Cirujeda[1], Henning Müller[2], Daniel Rubin[3], Todd A. Aguilera[4], Billy W. Loo Jr.[4], Maximilian Diehn[4], Xavier Binefa[1], Adrien Depeursinge[2]

***Abstract*— In this paper we present a novel technique for characterizing and classifying 3D textured volumes belonging to different lung tissue types in 3D CT images. We build a volume-based 3D descriptor, robust to changes of size, rigid spatial transformations and texture variability, thanks to the integration of Riesz-wavelet features within a Covariance-based descriptor formulation. 3D Riesz features characterize the morphology of tissue density due to their response to changes in intensity in CT images. These features are encoded in a Covariance-based descriptor formulation: this provides a compact and flexible representation thanks to the use of feature variations rather than dense features themselves and adds robustness to spatial changes. Furthermore, the particular symmetric definite positive matrix form of these descriptors causes them to lay in a Riemannian manifold. Thus, descriptors can be compared with analytical measures, and accurate techniques from machine learning and clustering can be adapted to their spatial domain. Additionally we present a classification model following a "Bag of Covariance Descriptors" paradigm in order to distinguish three different nodule tissue types in CT: solid, ground-glass opacity, and healthy lung. The method is evaluated on top of an acquired dataset of 95 patients with manually delineated ground truth by radiation oncology specialists in 3D, and quantitative sensitivity and specificity values are presented.**

I. INTRODUCTION

Clinical research has identified morphological tissue properties as indicators of cancer aggressiveness [1]. Texture and size of the solid and ground-glass opacity (GGO) components of a nodule, as observed from CT images, can provide reliable cues in order to assess medical examination criteria [2], but region texture delineation and classification is still an open and time-demanding problem. According to the clinical knowledge about the typology of ground-glass opacity (GGO) and solid tissue, it is established that the compactness, size, density and homogeneity of a nodule is differentiated from healthy lung regions, despite the large variability of normal lung tissue. We believe that the discriminative capabilities of these visual cues can be tackled from a pattern recognition approach, which motivates and settles the basis of the work presented in this paper.

[1] Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain,

[2] Institute of Information Systems, University of Applied Sciences Western Switzerland Sierre (HES-SO),

[3] Department of Radiology and Medicine (Biomedical Informatics), Stanford University, Stanford, CA, USA,

[4] Department of Radiation Oncology and Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA, USA

This work was supported by the Swiss National Science Foundation (Grant PZ00P2_154891) and the Spanish Government Ministry of Economy and Competitivity (Grants TIN2012-39203 and IPT-2012-0630-020000)

In computer vision research, several descriptors for 3D object classification have appeared ([3]–[6]). Nevertheless, these descriptors are usually targeted to 3D surfaces instead of 3D dense volumes as is the case in CT images. In the medical imaging domain, the survey conducted in [7] points out relevant techniques in applied 3D solid texture analysis and highlights the importance of multi-scale directional convolutional approaches that are non-separable to characterize subtle and discriminative properties of 3D biomedical textures. In this area, Riesz-wavelet features have demonstrated great representative capabilities: they characterize the morphology of tissue density thanks to their response to changes in CT intensities. These features are expressed by the response magnitudes to a set of 3D multiscale filters applied to the CT volume. We use this theoretically solid texture definition and propose its integration into a covariance-based descriptor, with the goal of establishing a paradigm for 3D region definition and classification. The main benefits of Covariance descriptors include the robustness to spatial transformations such as rotations, as well as the tolerance to changes in shape, size and resolution in the 3D domain. This is due to the fact that feature variation observations inside a region are used, instead of absolute feature values, and any structural information about feature location is discarded. Furthermore, as Covariance descriptors are embodied by covariance matrices, they lie in a meaningful and geometrically coherent descriptor space: similarly textured regions appear clustered in a low dimensional and analytically operable space. We then propose a part-based model which is able to represent the entire possible space of lung tissue types in this particular manifold, and model the underlying three classes of interest (GGO, solid and healthy tissue).

Our final goal is to provide a classification mechanism in order to assist clinical diagnosis and to aim towards unsupervised lung nodule segmentation. In this paper we provide the basis of the so-called Riesz-covariance descriptor and the "Bag-of-covariances" classification method, and present our results in the form of accuracy performance estimated on a dataset containing 100 patients with 3D regions delineated by experienced radiologists.

II. METHODOLOGY

A. Patients and dataset

100 patients from Stanford Hospital and Clinics with biopsy–proven early stage non-small cell lung carcinoma were used to estimate the performance of our approach. The nodule region present in each patient lungs was delineated

978-1-4244-9270-1/15/$31.00 ©2015 IEEE

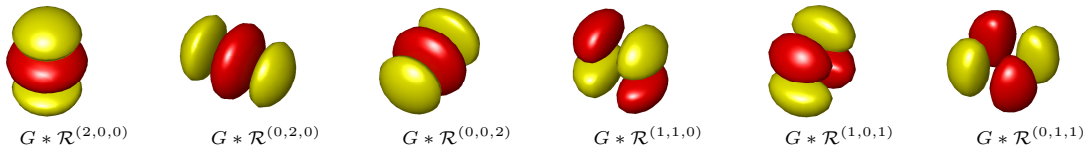

Fig. 1. Second–order Riesz kernels $\mathcal{R}^{(n_1,n_2,n_3)}$ convolved with isotropic Gaussian kernels $G(\boldsymbol{x})$. Responses to a linear combination of the filterbank represented by these kernels are used as discriminative representations of the underlying 3D tissue texture.

in 3D by the treating radiation oncologist, then the GGO and solid components were contoured separately using lung and mediastinal windows. The MATLAB software was used for post-processing of the available CT images and data, including region ground-truth preparation and resampling of volumes in order to have isotropic voxels of $0.8 \times 0.8 \times 0.8$ mm^3 using cubic spline interpolation. 5 patients of the dataset were discarded due to annotation artifacts, therefore the final dataset contains 95 patients.

B. 3D Riesz-wavelet features

3D Riesz filterbanks are used to characterize the texture of the lung parenchyma in 3D CT. The N–th order Riesz transform $\boldsymbol{\mathcal{R}}^{(N)}$ of a three–dimensional signal $f(\boldsymbol{x})$ is defined in the Fourier domain as:

$$\widehat{\mathcal{R}^{(n_1,n_2,n_3)}} f(\boldsymbol{\omega}) = \sqrt{\frac{n_1+n_2+n_3}{n_1!n_2!n_3!}} \frac{(-j\omega_1)^{n_1}(-j\omega_2)^{n_2}(-j\omega_3)^{n_3}}{||\boldsymbol{\omega}||^{n_1+n_2+n_3}} \hat{f}(\boldsymbol{\omega}), \tag{1}$$

for all combinations of (n_1, n_2, n_3) with $n_1 + n_2 + n_3 = N$ and $n_{1,2,3} \in \mathbb{N}$. Eq. (1) yields $\binom{N+2}{2}$ templates $\mathcal{R}^{(n_1,n_2,n_3)}$. The filterbanks are obtained by coupling the Riesz transform with isotropic band–limited wavelets [8]: Fig. 1 depicts the second–order Riesz filterbank when convolved with anisotropic Gaussian kernels. Rotation–covariance is obtained by locally aligning all Riesz components $\mathcal{R}^{(n_1,n_2,n_3)}$ of all scales based on the locally prevailing orientation [9]. The steerability property of the filterbanks allows synthesizing the responses of the Riesz components at any orientation from a linear combination of themselves, and does not require additional convolutions with oriented components. Locally prevailing orientations are estimated using a regularized version of the structure tensor computed directly from the Riesz components [10].

C. 3D spatial Covariance-based descriptors

As defined in Eq. (1), 2^{nd} order 3D Riesz features yield to a 6-dimensional response to a filterbank according to the texture of the tissue volume. For a 3D CT image of size $W \times H \times S$, we can obtain a new volume with the responses to each one of the second–order Riesz kernels, of size $6 \times W \times H \times S$. Nevertheless, for the task of tissue classification, a more compact and accurate representation is desirable, where feature characteristics can be encoded to a specific common format regardless of the size of any given volumetric region.

Covariance matrices were first used as descriptors in the computer vision domain by Tuzel *et al.* [11] for the recognition of objects and faces in 2D color images, as a way of relating visual cues such as edges, curvature or color values inside a region of interest of arbitrary size. This framework has been extended to other domains such as unstructured point clouds [12], 3D color texture + surface description [13], or to spatio-temporal gesture recognition in 3D depth image sequences [14], [15].

Due to their construction, covariance-based descriptors are robust to noisy inputs and lose structural information about the observed features. Therefore, they are suitable for unstructured, abstract texture characterization inside a region, regardless of spatial rigid transformations such as rotation, scale or translations. This representation is based on the statistical notion of covariance as a measure of how several random variables change together – 3D Riesz responses in our case. In order to define a given texture, it captures the intrinsic correlations between the n-dimensional distribution of these features, and uses this as a descriptive signature.

In order to formally define the 3D Riesz-Covariance descriptors, we denote a feature selection function $\Phi(ct, v)$ for a given 3D CT image ct and a selected subvolume region v inside the boundaries of ct as:

$$\Phi(ct, v) = \{\phi_{x,y,z}, \forall x, y, z \in v\}, \tag{2}$$

where $\phi_{x,y,z}$ is the vector of variables obtained at the set of the coordinates $\{x, y, z\}$ contained in the volume v, and is defined as:

$$\phi_{x,y,z} = \left(\mathcal{R}^{(n_1,n_2,n_3)}_{x,y,z}, \|\mathcal{R}\|_{x,y,z}, ct_{x,y,z}\right). \tag{3}$$

These features include the 6 Riesz features at each one of the coordinates in the set, as well as their norm and the CT intensity values in Hounsfield Units. We have empirically observed that this feature selection is capable of encoding the texture and the tissue nature.

Then, for a given region v of the CT image, the associated Covariance descriptor can be obtained as:

$$Cov\left(\Phi(ct, v)\right) = \frac{1}{N-1} \sum_{i=1}^{N} \left(\phi_{x,y,z} - \mu\right)\left(\phi_{x,y,z} - \mu\right)^T, \tag{4}$$

where μ is the vector mean of the set of vectors $\{\phi_{x,y,z}\}$ within the volumetric neighbourhood made of N samples.

The resulting 8×8 matrix Cov is a symmetric matrix where the diagonal entries will represent the variance of each feature channel, and the non-diagonal elements represent their pairwise covariance, as seen in Fig. 2.

D. Riemannian manifold topology

Riesz-Covariance descriptors do not only provide an interesting representation of data for their compactness and flexibility: being covariance matrices they lie in the Riemannian

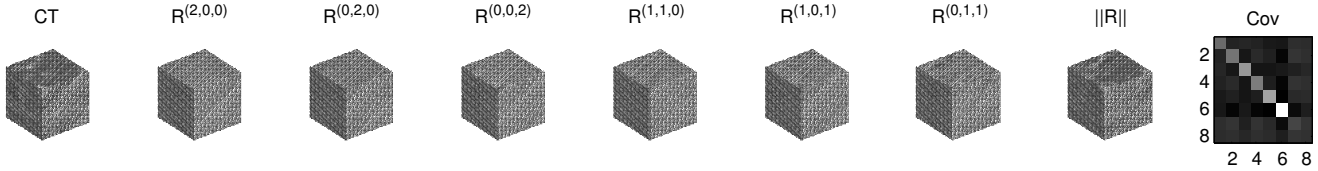

Fig. 2. Cues involved in the descriptor calculation for a given CT cubic region. The 8 first cubes depict the values within a $40 \times 40 \times 40$ pixel volume, with the CT intensities, 3D-Riesz wavelet responses (for one fixed scale) and Riesz norm features. The 8×8 matrix in the right subfigure depicts the resulting Covariance descriptor, encoding the different correlations between the distributions of the observed cues.

manifold of symmetric definite positive matrices Sym_d^+. This spatial variety is geometrically meaningful as 3D regions sharing similar texture characteristics remain under close areas in the descriptor space. Nevertheless, descriptors have to be treated carefully in order to analytically exploit their non Euclidean spatial distribution via their projection into the tangent space at a particular point of the manifold.

According to [16], the Riemannian metric can be approximated in close neighborhoods of the Sym_d^+ manifold by the Euclidean metric in its tangent space, T_Y, where the symmetric matrix Y is a reference projection point in the manifold. T_Y is formed by a vector space of $d \times d$ symmetric matrices. The tangent mapping of a manifold element X to $x \in T_Y$ is made by the point-dependent $\log_Y$ operation:

$$x = \log_Y(X) = Y^{\frac{1}{2}} \log\left(Y^{-\frac{1}{2}} X Y^{-\frac{1}{2}}\right) Y^{\frac{1}{2}}. \quad (5)$$

The geodesic distance between two points X_1 and X_2 on Sym_d^+ is defined as

$$\delta(X_1, X_2) = \sqrt{Trace\left(\log\left(X_1^{-\frac{1}{2}} X_2 X_1^{-\frac{1}{2}}\right)^2\right)}, \quad (6)$$

or more simply $\delta(X_1, X_2) = \sqrt{\sum_{i=1}^{d} \log(\lambda_i)^2}$, where λ_i are the positive eigenvalues of $X_1^{-\frac{1}{2}} X_2 X_1^{-\frac{1}{2}}$.

After observing the spatial distribution of the different lung tissue classes, via multidimensional scaling of the space of descriptors obtained for the GGO, solid and healthy lung areas, we map all the data in a unique tangent space with respect to a unique point in Sym_d^+, as depicted in Fig. 3. In order to simplify the computation, we choose this projection point as the identity matrix I_d, and the tangent mapping is computed as a regular matrix logarithm as

$$x = \log_{I_d}(X) = log(X) = U \log(D) U', \quad (7)$$

with U and D being the elements of the single value decomposition (SVD) of $X \in Sym_d^+$.

The resulting symmetric matrices in the tangent space T_{I_d} contain only $d(d+1)/2$ independent coefficients, in their upper or lower triangular parts. Therefore it is possible to apply the vectorization operation in order to obtain a linear orthonormal space for the independent coefficients:

$$\hat{x} = vect(x) = (x_{1,1}, x_{1,2}, ..., x_{1,d}, x_{2,2}, x_{2,3}, ..., x_{d,d}), \quad (8)$$

where x is the mapping of $X \in Sym_d^+$ to the tangent space, resulting from Eq. (7). The obtained vector $\hat{x}$ will lie in the Euclidean space $\mathbb{R}^m$, where $m = d(d+1)/2$.

E. Texture classification via Bag-of-Covariances

The "Bag of features" paradigm is an established classification technique in the machine learning domain [17]. It is conceived as an implicit part-based modeling from a learning set of instances, where collections of different parts of objects can be gathered in order to cover the intra-class variability. Later on, this set of part representations, often referred to as *dictionary*, is used to encode a learning set of instances in terms of frequency histograms of the part repetitions found on these instances. The same representation is done for classification samples and the final decision criteria are made in terms of histogram similarities. This directly suits our classification problem: due to the small number of samples and low resolution of features, we can model a vast dictionary of all the possibilities in tissue types – inner texture and margin of solid and GGO nodule components, vessels, air, blood or even fiducial markers in the healthy lung tissue.

We define our so-called "Bag of Covariances" in three stages: dictionary learning, modeling of tissue classes by word frequencies, and classification of test regions. In order to build the so-called dictionary, we denote by $P = \{\text{CT}\ {}^{c}_{1:p}\}$ as the set of CT images for all patients p, and their delineated regions for each class c (solid, GGO and healthy lung). From this data, we can obtain the set of vectorized 3D Riesz-Covariance descriptors as defined in the previous section: $\hat{x}\ {}^{c}_{v,p} = vect(\log_{I_d}(Cov^C_{V,P}))$, for a set of learning patients p and classes c. v denotes the set of 3D subvolumes inside the region class for which the descriptors are computed, and it is obtained randomly inside the manually annotated class regions. See Fig. 4 for a clarification of this learning. All these elements, so-called words, can be stored as a matrix, and data clustering algorithms such as K-means can be applied in order to reduce dimensionality of those over-represented samples.

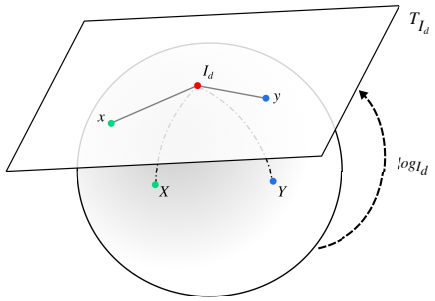

Fig. 3. Mapping of points in Sym_d^+ manifold to the tangent space T_{I_d}.

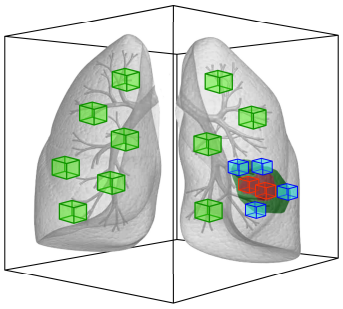

Fig. 4. Representation of the descriptors at a patient level: cube colors denote the 3D patches used for the construction of the dictionary according to the three tissue classes.

In order to model the different classes, a new set of 3D Riesz-Covariance descriptors for different parts of all three classes is again obtained, from a second set of learning patients. The descriptors are mapped to the nearest words in the dictionary via their Euclidean distance, as the parts are projected to the tangent space T_{I_d}. This gives a set of histogram representations in which each one of the tissue instances for all the patients are defined as the frequency of part appearances present in the dictionary.

For the classification of a new sample, a new set of 3D Riesz-Covariance descriptors $\hat{y}_v = vect(\log_{I_d}(Cov_{ct,v}))$ is obtained, where v indicates different patches inside the CT image. Again, the descriptors can be quantized in terms of dictionary frequencies, and the final classification criteria are made according to the closest histogram representation in the available model:

$$class(ct) = \underset{i}{\operatorname{argmin}}\, D(h_{ct}, h_i), \tag{9}$$

where h_{ct} denotes the histogram representation of the CT test sample, h_i denotes the learned model of dictionary frequency representations, and D is the χ^2 distance used for the comparison of histograms.

III. EXPERIMENTAL EVALUATION

In order to estimate the classification accuracy achieved by the proposed method, we performed cross-validation over the presented dataset, keeping 35 patients for learning the model and the remaining 60 patients for testing, for 10 iterations. The "Bag-of-covariances" method was trained by modelling 60 parts for each class, for each patient, therefore creating a dictionary size of 6300 words at each iteration. The classification performance accuracy for the three modelled classes is reported in terms of sensitivity $(TP/TP+FN)$ and specificity $(TN/TN+FP)$, with average values of **82.2**% (σ = 2.55%) and **86.2**% (σ = 5.85%) respectively, according to the available ground-truth annotations defined in Section II.a.

Recent methods as [18], [19] reported similar accuracy, which settles our presented approach amongst state of the art performance levels. Even if these methods are focused on interstitial lung diseases rather than nodule separation, their definition of GGO and solid areas is consistent with our approach so comparing against their outcomes is coherent.

IV. DISCUSSION AND CONCLUSIONS

We proposed an integrated approach for the characterization and classification of different lung tissue types in 3D, particularly focused to the separation of GGO and solid nodule areas. Despite the high intra-class variability, our method obtained reliable classification results, thanks to a theoretically solid descriptor for encoding feature variations.

The contribution of this work sets the basis for further CT image classification, not only in terms of different tissue classes but also at an intra-class level in order to model temporal stages of a nodule or to learn specific models for characterizing the response to treatment of specific diseases. These techniques may be clinically useful for identifying and characterizing suspicious tissue regions in lesions.

REFERENCES

[1] T. Yokose, K. Suzuki, K. Nagai, Y. Nishiwaki, S. Sasaki, and A. Ochiai, "Favorable and unfavorable morphological prognostic factors in peripheral adenocarcinoma of the lung 3 cm or less in diameter," *Lung Cancer*, vol. 29, no. 3, pp. 179–188, 2000.

[2] A. Depeursinge, M. Yanagawa, A. N. Leung, and D. L. Rubin, "Predicting adenocarcinoma recurrence using computational texture models of nodule components in lung CT," *Medical Physics*, 2015.

[3] A. Zaharescu, E. Boyer, K. Varanasi, and R. Horaud, "Surface feature detection and description with applications to mesh matching," in *CVPR*, pp. 373–380, 2009.

[4] F. Tombari, S. Salti, and L. Stefano, "Unique signatures of histograms for local surface description," in *ECCV*, vol. 6313, pp. 356–369, 2010.

[5] R. Rusu, N. Blodow, and M. Beetz, "Fast point feature histograms (fpfh) for 3d registration," in *ICRA*, pp. 3212–3217, 2009.

[6] A. Flint, A. R. Dick, and A. Van Den Hengel, "Thrift: Local 3d structure recognition.," in *Digital Image Computing Techniques and Applications*, vol. 7, pp. 182–188, 2007.

[7] A. Depeursinge, A. Foncubierta-Rodríguez, D. Van De Ville, and H. Müller, "Three–dimensional solid texture analysis and retrieval in biomedical imaging: review and opportunities," *Medical Image Analysis*, vol. 18, no. 1, pp. 176–196, 2014.

[8] P. Pad, V. Uhlmann, and M. Unser, "VOW: Variance optimal wavelets for the steerable pyramid," in *IEEE International Conference on Image Processing*, ICIP 2014, pp. 2973–2977, October 2014.

[9] A. Depeursinge, P. Pad, A. C. Chin, A. N. Leung, D. L. Rubin, H. Müller, and M. Unser, "Optimized steerable wavelets for texture analysis of lung tissue in 3–D CT: classification of usual interstitial pneumonia," in *ISBI*, 2015.

[10] N. Chenouard and M. Unser, "3D steerable wavelets and monogenic analysis for bioimaging," in *2011 IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro*, pp. 2132–2135, April 2011.

[11] O. Tuzel, F. Porikli, and P. Meer, "Region covariance: A fast descriptor for detection and classification," *ECCV*, pp. 589–600, 2006.

[12] D. Fehr, A. Cherian, R. Sivalingam, S. Nickolay, V. Morellas, and N. Papanikolopoulos, "Compact covariance descriptors in 3d point clouds for object recognition," in *ICRA*, pp. 1793–1798, 2012.

[13] P. Cirujeda, X. Mateo, Y. Dicente, and X. Binefa, "MCOV: A covariance descriptor for fusion of texture and shape features in 3d point clouds," in *International Conference on 3D vision (3DV)*, 2014.

[14] P. Cirujeda and X. Binefa, "4DCov: A nested covariance descriptor of spatio-temporal features for gesture recognition in depth sequences," in *International Conference on 3D vision (3DV)*, 2014.

[15] M. Hussein, M. Torki, M. Gowayyed, and M. El-Saban, "Human action recognition using a temporal hierarchy of covariance descriptors on 3d joint locations," in *Int. Conf. on Art. Intel.*, pp. 2466–2472, 2013.

[16] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Log-euclidean metrics for fast and simple calculus on diffusion tensors," *Magnetic Resonance in Medicine*, vol. 56, no. 2, pp. 411–421, 2006.

[17] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," in *CVPR*, pp. 2169–2178, 2006.

[18] Y. Song, W. Cai, Y. Zhou, and D. D. Feng, "Feature–based image patch approximation for lung tissue classification," *IEEE Transactions on Medical Imaging*, vol. PP, no. 99, 2013.

[19] A. Depeursinge, D. Van De Ville, A. Platon, A. Geissbuhler, P.-A. Poletti, and H. Müller, "Near–affine–invariant texture learning for lung tissue analysis using isotropic wavelet frames," *IEEE Trans. on Inf. Tech. in BioMedicine*, vol. 16, no. 4, pp. 665–675, 2012.