Decomposition of 3D Medical Images into Visual Patterns

Raquel Dosil, Xosé M. Pardo and Xosé R. Fdez-Vidal

Abstract— In this paper we present a method for the decomposition of a volumetric image into its [1]most relevant *visual patterns*, which we define as features associated to local energy maxima of the image. The method involves the clustering of a set of predefined band-pass energy filters according to their ability to segregate the different features in the image, thus generating a set of composite-feature detectors tuned to the specific visual patterns present in the data. Clustering is based on a measure of statistical dependence between pairs of frequency features. We will illustrate the applicability of the method to the initialization of a 3D geodesic active model.

Index Terms—Low Level Representation, Multiresolution Analysis, Phase Congruence, Active Model Initialization.

1 Objectives

In this paper we present a method for low level representation of 3D images consisting on the identification its most relevant low level features, which we call *visual patterns*. Some authors agree in that relevant features correspond to points with locally maximal Phase Congruence (PC), which measures the local degree of matching among the phase of Fourier components [1]. PC maxima are also maxima of the local energy of the image [2]. Local energy can be estimated using energy filters [1], [2]. They are scalable operators that measure local energy as the sum of the squared responses of a pair of filters in quadrature. To study PC, many approaches use multiscale analysis. The image is first decomposed into a set of subband images and then the information is integrated again. Usually, integration leads to unique "primal sketch" of the image [3]. However, in the RGFF model [4] one response map is generated for each visual pattern. Integration of bands correspondent to each composite-feature is accomplished by clustering of subband images according to a measure of *statistical congruence* among them.

Based on these ideas, we have developed our representation model. It involves hierarchical clustering of a

Manuscript received July 19, 2004. This work was supported in part by the Xunta de Galicia, under Grant PGIDIT04TIC206005PR.

R. Dosil and X. M. Pardo are with the Dept. of Electronics and Computer Science of the Univ. de Santiago de Compostela. Campus Univ. Sur, s/n, 15782, Santiago de Compostal, Spain. Phone: +34 981563100, Fax: +34 981528012, e-mail: rdosil@usc.es.

X. R. Fdez-Vidal is with Dept. of Applied Physics, of the Univ. of Santiago de Compostela. Escola Politécnica Superior, Campus Univ., s/n, 27002, Lugo, Spain.

set of frequency features, obtained by means of a 3D bank [5] of filters with rotational symmetry [6], according to their ability to discriminate a given visual pattern. Thus, the method generates a set of composite-feature detectors tuned to the specific visual patterns present in the data. Here, the clustering criterion is PC, estimated using a global distance based on the Normalized Mutual Information (NMI) [7] of the energy of the subband images. It reflects their degree of statistical dependence and, hence, the degree of alignment of their energy maxima. The partitioning method is integrated with a geodesic active model as an initialization step. Isolating visual patterns from each other allows to focus local deformation on a given target object, avoiding interference of other features. We will show that the isolated visual patterns can place the initial model near the target.

The method for visual pattern extraction is described in section 2. Integration with a geodesic active model is described in section 3. In section 4, we illustrate the behavior of the method along with a geodesic active model. In section 5 we present the conclusions and future work.

2 Visual pattern detection

Our representation model is illustrated in Fig. 1. It involves subband decomposition of the image using *active* channels in a 3D log Gabor filter bank, with high information content, followed by clustering of frequency features. The 3D log Gabor transfer function T is designed as the product of a 1D log Gabor function in the radial frequency [8] and a Gaussian in the angular distance, which has rotational symmetry in spherical coordinates [6]

$$T(\rho,\phi,\theta;\rho_i,\phi_i,\theta_i) = \exp\left(-\frac{\left(\log\left(\rho/\rho_i\right)\right)^2}{2\left(\log\left(\sigma_\rho/\rho_i\right)\right)^2}\right)\exp\left(-\frac{\alpha(\phi_i,\theta_i)^2}{2\sigma_\alpha^2}\right)$$

where (ρ_i,ϕ_i,θ_i) is the filter's central frequency, σ_ρ and σ_α are the standard deviations, $\alpha(\phi_i,\theta_i)=\text{acos}(\mathbf{f}\cdot\mathbf{v}/\|\mathbf{f}\|)$, $\mathbf{v}=(\cos\phi_i\cdot\cos\theta_i, \cos\phi_i\cdot\sin\theta_i, \sin\phi_i)$ and $\mathbf{f}$ is a point in the frequency space expressed in Cartesians.

To generate the bank of filters, the radial frequency dimension is tiled by octave bands, centered at wavelengths $\lambda_i = \lambda_{\min}\cdot 2^{(i-1)}$, with $i = \{1, \ldots, N\}$ and N being the number of bands. Elevation is sampled uniformly. The number of azimuth values decreases with elevation, instead of taking uniform angular distances, in order to keep the constant filter "density", by maintaining equal arc-length between adjacent azimuth values over the unit radius sphere.

To decrease computational cost, bands with low energy contribution, or non *active*, can be discarded, measured as $E = \log(|F| + 1)$, where F is the Fourier transform of the image. We assume high frequency energy caused mainly by noise and flat spectral noise energy [3]. Hence, a band is *active* if it comprises any energy value over the estimated maximum level of the spectral noise $m + 3\sigma$, where m and σ are estimated as the average and the standard deviation of E in a band of high frequencies. Subsequently, radial median filtering is applied to eliminate spurious noise peaks. It considers only neighbors of a pixel previous and following in the radial direction, eliminating spurious peaks but preserving the continuity of structures along scales.

The measure of dissimilarity between pairs of energy maps is a function of their NMI, a measure of statistical dependence between two signals. Let H represent entropy and X and Y the energies of two active frequency features. Then, the actual distance employed is:

$$D_{NMI}(X,Y) = \left(1 - \sqrt{\mathrm{NMI}(X,Y)}\right)^2 = \left(1 - \sqrt{2 \cdot \left(1 - \frac{\mathrm{H}(X,Y)}{\mathrm{H}(X) + \mathrm{H}(Y)}\right)}\right)^2 .$$

D_{NMI} is much simpler compared to the measure used in the RGFF, which is highly parameterized, and is less computationally expensive, with time $O(F^2 \cdot N)$, compared to $O(F^2 \cdot N^2)$ in the RGFF, where N is the volume size and the filter bank consists of F filters.

To group filter's responses, a hierarchical clustering method has been chosen. It works directly with the dissimilarity matrix, not involving data coordinates. Cluster distances are evaluated using a ward's link strategy. To determine the optimal number of clusters N_c, we run the algorithm for each possible N_c and evaluate the quality of each resulting configuration according to the modified Davies-Bouldin validity index [9]. It is a graph-theory based measure for the compactness of the clusters in relation to their separation.

3 Geodesic Model Initialization

A geodesic active model represents objects using an implicit function defined in the whole image domain, where the surface of the objects corresponds to its zero level set. The specific implementation used here is the one described in [10], with the stopping criterion being $g=1/(1+(\|G_\sigma * I\|/\lambda)^2)$, where G_σ is a Gaussian operator with scale σ, I is the original image, and λ is a real constant. Neither balloon forces nor *a priori* knowledge on the domain are incorporated. Usually, the model is initialized defining a window function

enclosing all objects or by user interaction. In the present application, the initial model is defined manually by choosing one among the generated visual patterns that best represents the target object. The selected visual pattern response is taken as the linear summation of the real part or half-wave rectified real part of the responses of its filters, instead of the amplitude, which would produce a less precise initialization. The initial zero level set results form rescaling the selected visual pattern to the interval [–1, 1].

4…Results

In this section we illustrate both the ability of the system to identify and extract visual patterns of different types and its usefulness in medical image applications. These results have been obtained using a bank of 92 filters, with wavelengths 4, 8, 16 and 32 pixels, 2 octave bandwidth and σ_α= 25º. The orientation space has been sampled taking 3 elevations and 6 azimuths in the semi-equator, with $\theta = 0$ and $\phi = [0, \pi)$, producing a total of 23 orientations under the condition of equal arc-length.

The first example, in Fig. 2, corresponds to the segmentation of real 3D medical imaging data. It is a volumetric data set of size 32×64×20 from Single Photon Emission Computerized Tomography (SPECT) of the left ventricle of the heart. The method detects two visual patterns, one gathering the background artifacts and the other one mainly representing the ventricle wall, which is the one selected to initialize the active model. The segmentation result shows that both the outer and inner side of the ventricle wall have been correctly captured, since the initial surface was already very close to them. Thresholding produces artifacts due to noise. If the threshold level is lowered the ventricle wall presents holes. Segmentation using a square box as initialization is not able to reach the ventricle cavity without overcoming the ventricle wall.

The second medical image case, in Fig. 3.a and b, is 3D data set of size 71×107×35 of a T2-weighted Magnetic Resonance (MR) phantom image of the lateral ventricles of the brain. It can be seen that the first visual pattern enhances the corpus callosum, while the second visual pattern accounts for the lateral ventricles. If each of these two visual patterns is used to initialize one geodesic model, both structures can be independently segmented. The use of a square box as initialization leads to a segmentation that gathers different parts of the ventricles, the corpus callosum and the caudate nucleus all together. Thresholding does not succeed either, as the corpus callosum has the same color than the lateral lobes and the ventricles.

5 Conclusions

In this paper we have presented a method for the isolation of visual patterns in volumetric images by subband decomposition followed by clustering based on PC. The proposed dissimilarity measure, based on the NMI, reduces the computational burden of the process, is less parameterized and less complex that other approaches. Image segmentation results show that the method is suitable for its combination with active models. Model initialization by manual selection of visual patterns enhances results and fastens convergence. The learned visual patterns can be reutilized when dealing with large data sets within the same domain.

The defined initialization methodology mainly deals with low level information, since the active model evolves locally. To take advantage of high level information, an intermediate step of matching prior shape models to the visual patterns can be incorporated before deformation [11]. Future work will point to this direction and also to automate visual pattern selection.

References

[1] M.C. Morrone, R. Owens. Feature Detection from Local Energy. PRL, 6(5), 303-313, 1987.

[2] S. Venkatesh, R. Owens. On the Classification of Image Features. PRL, 11, 339-349, 1990.

[3] P. D. Kovesi. Invariant Measures of Image Features from Phase Information, PhD. Thesis, The University or Western Australia, 1996.

[4] R. Rodríguez-Sánchez et al.. The RGFF Representational Model: A System for the Automatically Learned Partition of "Visual Patterns" in Digital Images, PAMI, Vol. 21(10), pp. 1044-1073, 1999.

[5] J. Bigun. Speed, Frequency, and Orientation Tuned 3-G Gabor Filter Banks and their Design. In: ICPR94, Jerusalem, pp. C-184-187, 1994.

[6] F. G. A. Faas, L. J. van Vliet. 3D-Orientation Space; Filters and Sampling. In: J. Bigun, T. Gustavsson (eds.), Scandinavian Conference on Image Analysis, Springer, 36-42, 2003.

[7] C. Studholme, D. L. G. Hill, D. J. Hawkes. An Overlap Invariant Entropy Measure of 3D Medical Image Alignment. PR, 32, 71-86, 1999.

[8] D. J. Field. Relations between the Statistics of Natural Images and the Response Properties of Cortical Cells. JOSA A, 4, 2379-2394, 1987.

[9] N. Pal, J. Biswas. Cluster Validation Using Graph Theoretic Concepts. Pattern Recognition, 30(6), pp. 847-857, 1997.

[10] J. Weickert, G. Kühne. Fast Methods for Implicit Active Contour Models. In: S. Osher, N. Paragios (eds.): Geometric Level Set Methods in Imaging, Vision and Graphics, Springer, pp. 43-58, 2003.

[11] R. Dosil and X. Pardo. Generalized Ellipsoids and Anisotropic Filtering for Segmentation Improvement in 3D Medical Imaging. Imag. and Vis. Comp., 21(4), pp. 325-343, 2003.

Figure Captions

Fig 1. Scheme of the visual pattern partitioning process

Fig 2. (a) Cross sections of a SPECT image of the left ventricle of the heart. (b, c) Cross sections of the two visual patterns identified. (d) Segmentation using visual pattern in (c) as initial model. (e). Isosurface of level 100. (f) Segmentation using a square box as initial model for three stages of evolution. (g, h, i) 3D reconstruction of the surfaces in (d), (e) and (f) for 60 iterations.

Fig 3. *Top Row*: Sample slices of a phantom MR image of the brain ventricles and the corpus callosum. *Rows 2 and 3*: The two visual patterns isolated –the first one is inverted. *Rows 4 and 5*: Segmentations using visual patterns in 2nd and 3rd rows respectively as initialization. *Row 6*: Segmentation using a square box as initial model. *Row 7*: 3D reconstructions of *Left*: isosurface of level 200 –corpus callosum– and *Right*: isosurface of level 240 –ventricles–. *Bottom Row:* ·3D reconstructions of segmentations *Left*: on 6th row and *Right*: segmentations on 4th and 5th rows, in the same plot, 4th in black and 5th in white.

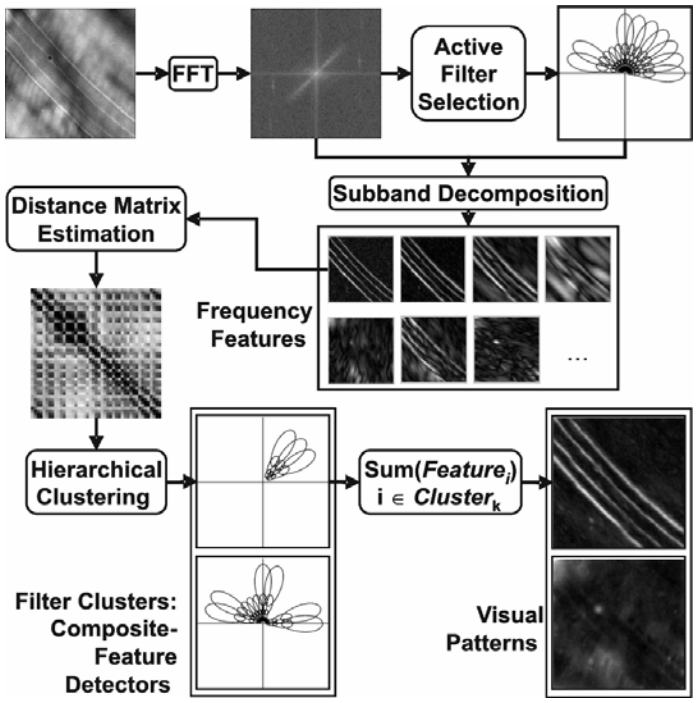
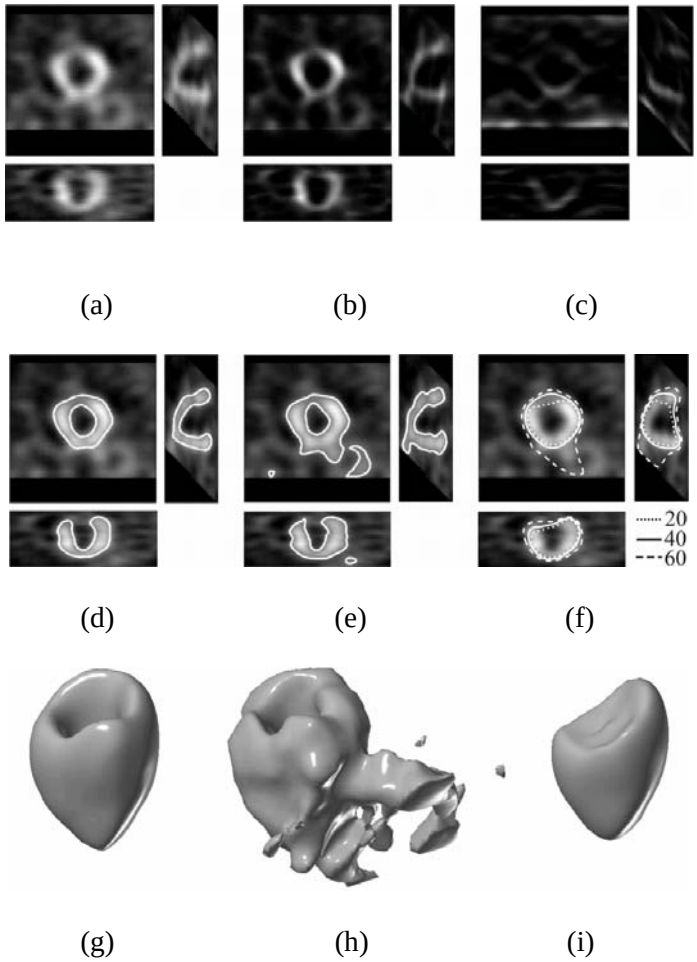

FFT
Active Filter Selection
Subband Decomposition
Distance Matrix Estimation
Frequency Features
...
Hierarchical Clustering
Sum(Feature_i) i ∈ Cluster_k
Filter Clusters: Composite-Feature Detectors
Visual Patterns

Fig. 1.



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Fig. 2.

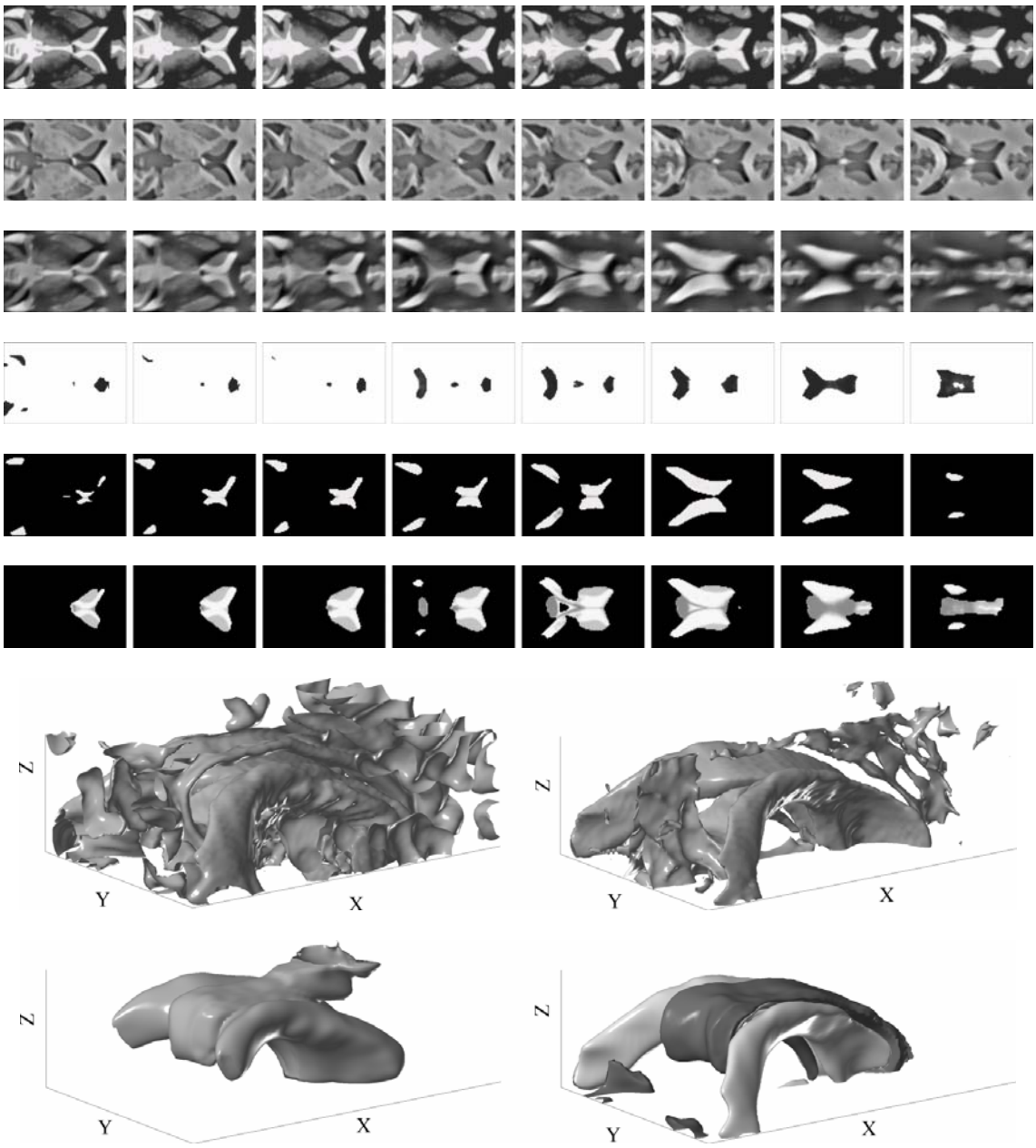


Fig. 3.