Functional semi-automated segmentation of renal DCE-MRI sequences using a Growing Neural Gas algorithm

Béatrice Chevaillier, Damien Mandry, Yannick Ponvianne, Jean-Luc Collette, Michel Claudon, Olivier Pietquin

FUNCTIONAL SEMI-AUTOMATED SEGMENTATION OF RENAL DCE-MRI SEQUENCES USING A GROWING NEURAL GAS ALGORITHM

Chevaillier B.[(1)]*, Mandry D.*[(2)]*, Ponvianne Y.*[(2)] *, Collette J.L.*[(1)]*, Claudon M.*[(2)] *and Pietquin O.*[(1,2)]

[(1)] SUPELEC-Metz campus, IMS Research Group, 2 rue E. Belin, 57070 Metz (France)
[(2)] IADI, INSERM, ERI 13, Tour Drouet CHU Nancy Brabois 54500 Vandoeuvre-les-Nancy - France
Nancy University Nancy - Tour Drouet CHU Nancy Brabois 54500 Vandoeuvre-les-Nancy - France
phone: + (33) 387 764 711, fax: + (33) 387 764 700, email: Beatrice.Chevaillier@supelec.fr
web: www.supelec.fr

ABSTRACT

In dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI), segmentation of internal kidney structures like cortex, medulla and pelvo-caliceal cavities is necessary for functional assessment. Manual segmentation by a radiologist is fairly delicate because images are blurred and highly noisy. Moreover the different compartments cannot be delineated on a single image because they are not visible during the same perfusion phase for physiological reasons. Nevertheless the differences between temporal evolution of contrast in each anatomical region can be used to perform functional segmentation. We propose to test a semi-automated split and merge method based on time-intensity curves of renal pixels. Its first step requires a variant of the classical Growing Neural Gas algorithm. In the absence of ground truth for results assessment, a manual anatomical segmentation by a radiologist is considered as a reference. Some discrepancy criteria are computed between this segmentation and the functional one. As a comparison, the same criteria are evaluated between the reference and another manual segmentation.

1. INTRODUCTION

Among the different MRI techniques aiming at renal function study, DCE-MRI with gadolinium chelates injection is the most widely used [1]. Several parameters can be non-invasively computed from perfusion curves of different regions of interest (ROI). So segmentation of internal anatomical kidney structures like cortex, medulla and pelvo-caliceal cavities is crucial for functional assessment. Manual segmentation by a radiologist is fairly delicate because images are blurred and highly noisy. Moreover the different compartments are not visible during the same perfusion phase because of contrast changes; consequently they cannot be delineated on a single image. Radiologists have to examine the whole sequence in order to choose the two most suitable frames: the operation is time-consuming and functional analysis can vary greatly in case of misregistration or through-plane motion. Some semi-automated methods using thresholding, boundary or region-based techniques are often used in the medical field but few of them have been tested on renal DCE-MRI sequences (see [2] for a review). As the contrast temporal evolution is different in every compartment for physiological reasons, pixels can be classified according to their time-intensity curves. Nevertheless a functional segmentation using some unsupervised classification method and resulting in only three ROIs corresponding to cortex, medulla and cavities can hardly be obtained directly. This is mainly due to considerable contrast dissimilarities between pixels in a same compartiment despite some common characteristics [3].

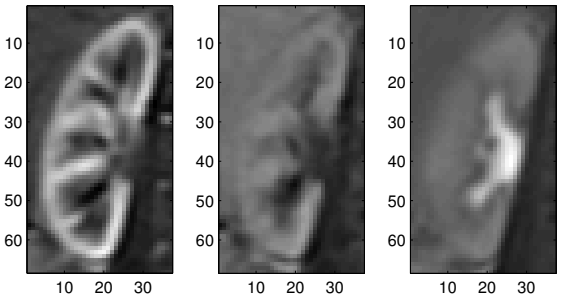


Figure 1: Examples of frames from a DCE-MRI sequence during arterial peak (left), filtration (middle) and late phase (right)

We propose to test a semi-automated split (2.1) and merge method (2.2) for renal functional segmentation. The kidney pixels are first classified in several clusters according to their contrast evolution using a vector quantization algorithm. These clusters are then merged thanks to some characteristic criteria of their prototype functional curves to get the three final anatomical compartments. Operator intervention consists only in a coarse tuning of two independent thresholds for merging, and is thus easy and quick to perform. The method is also relatively robust because the whole sequence is used instead of only two frames for manual segmentation. In the absence of ground truth for results assessment, a manual anatomical segmentation by a radiologist is considered as a reference. Some discrepancy criteria are computed between this segmentation and functional ones. As a comparison, the same criteria are evaluated between the reference and another manual segmentation.

2. METHOD FOR FUNCTIONAL SEGMENTATION

2.1 Vector quantization of time-intensity curves

The temporal evolution of contrast for each of the N pixels of a kidney results from a DCE-MRI registered sequence: examples of three frames for different perfusion phases can be seen in figure 1.

In order to have similar dynamic for any kidney, all intensities I are normalized, being replaced by $\frac{I-I_B}{I_L-I_B}$, where I_B

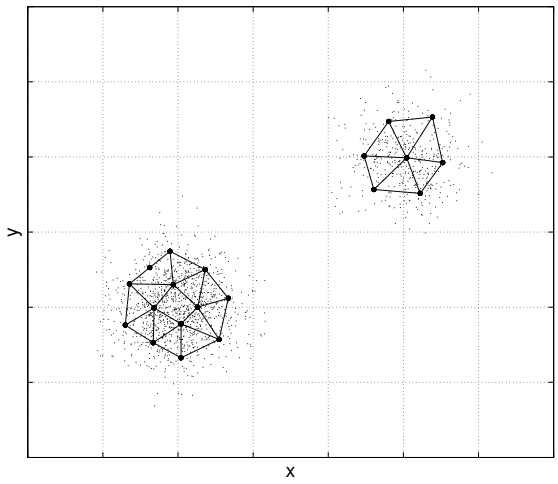


Figure 2: Results of a GNG procedure: small dots represent samples of the distribution, large dots are the resulting nodes linked with edges.

is the mean value for baseline and I_L the mean value during late phase for the time-intensity curve of entire kidney. Let $\xi_i = (\xi_{i1}, \ldots, \xi_{iN_T})$ be the N_T-components normalized vector associated with each pixel, where ξ_{ip} is the contrast at time p for the pixel x_i.

The N vectors ξ_i are considered as samples of an unknown probability distribution over a N_T dimensional space X with a density of probability $p(\xi)$. The aim is to find a set $\{w_j\}_{1\leq j\leq K} \subset X$ of prototypes (or nodes) that maps the distribution with a given distorsion. Let be $w(\xi) = argmin_{w_j}\{\|\xi - w_j\|^2\}$. The Growing Neural Gas with targeting (GNG-T) [4], which is a variant of the classical Growing Neural Gas algorithm [5], minimizes the cost function

$$E = \int_X \|w(\xi) - \xi\|^2 p(\xi) d\xi = \sum_{j=1}^{K} E_j \qquad (1)$$

where

$$E_j = \int_{V_j} \|w_j - \xi\|^2 p(\xi) d\xi \text{ and } V_j = \{\xi \in X : w(\xi) = w_j\} \qquad (2)$$

The results consist in both:

- a set $\{w_j\}_{1\leq j\leq K}$ of prototypes,
- a graph structure defining a topological neighborhood relation in the parameter space, so that two prototypes with similar behaviour are linked with an edge.

The number K of prototypes is iteratively determined to reach a given average node distortion T. Let us note that the aim of the algorithm is not to classify pixels but to perform vector quantization, that is why it tends to give a fairly large K value. As an example, the quantization results for a two-dimensional Gaussian mixture distribution is given in figure 2. For this ideal density every class corresponds to a connected nodes set but this is seldom true for real cases because of noise and because the distributions are not obviously separable. Most of the time a single connected network is obtained. So a merging step is indispensable to obtain the final segmentation in three anatomical compartments.

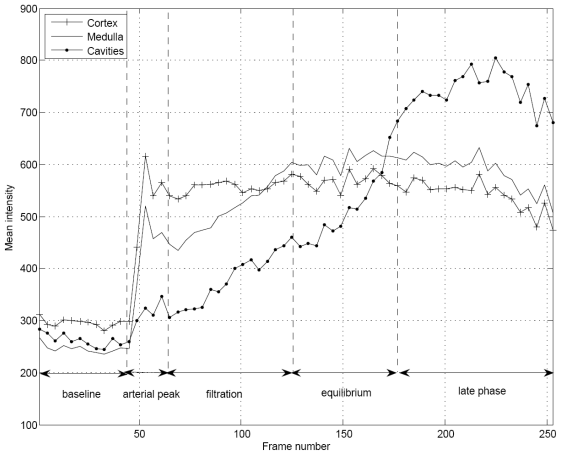


Figure 3: Typical time-intensity curves for cortex, medulla and cavities

2.2 Merging

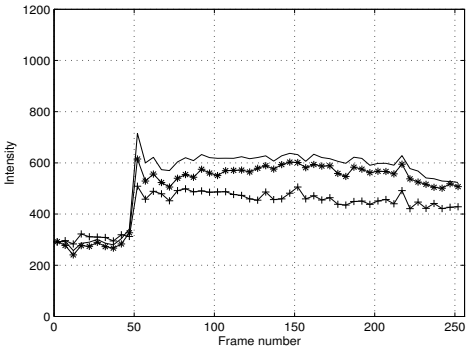
Each node has then to be attributed to one of the three anatomical compartments. Typical time-intensity curves with the main perfusion phases (baseline, arterial peak, filtration, equilibrium and late phase) are shown in figure 3.

Nevertheless, for a given kidney, noticeable differences can be observed in figure 4 inside each compartment. The Euclidean distance between curves is not a criterion significant and robust enough to regroup nodes. Indeed distance between two prototypes of two distinct ROIs may often be smaller than disparity within a given compartment. This is true even if distance is evaluated only for points of filtration, during which contrast evolutions should be the most different. That is why some physiology related characteristics of the contrast evolution have to be used to carry out the merging step.

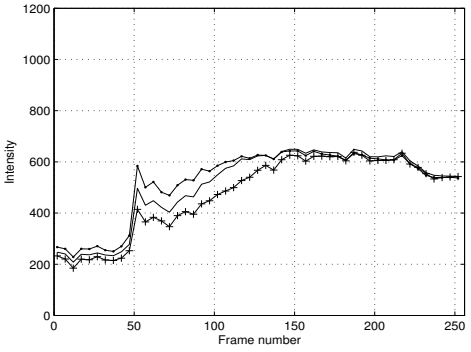
We proceed as follows:

- First, as cavities should be the brighter structure in the late phase, nodes whose average intensity during this stage is greater than a given threshold t_1 and that are directly connected to each other in the GNG graph are considered as cavities.
- In a second step, filtration phase is used to separate cortex and medulla. Filtration rate depends on the tissu nature. So the slope of time-intensity curves during filtration is evaluated using standard linear regression for all remaining prototypes: a node is attributed to the cortex if the corresponding slope is less than a given threshold t_2, else it is assigned to the medulla.

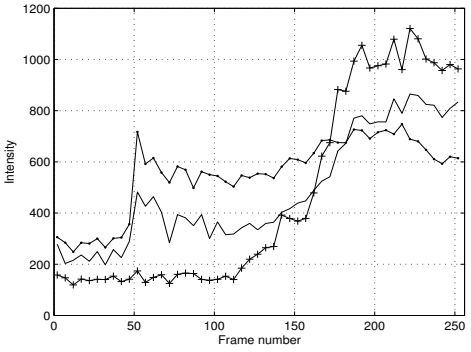
The two thresholds t_1 and t_2 are initialized so that cortex represents approximately 50% and cavities about 20% of kidney area and are adjusted by an observer. This is the only manual intervention of the whole operation. As the algorithm is very fast, the tuning step can be done in real time. Let us stress that the second criterion would not be sufficient to distinguish cavities from cortex and medulla because of a theoretically unexpected but fairly high arterial peak that can be observed in figure 4(c): this peak is induced by the great vascularization of the whole kidney and may appear in all ROIs. Furthermore the use of topological edges for cavities determination avoids to classify in this compartment some nodes



(a) Cortex prototypes



(b) Medulla prototypes



(c) Cavities prototypes

Figure 4: Some examples of time-intensity curves for prototypes attributed to cortex, medulla and cavities of a given kidney

that have a similar contrast in late phase but whose behaviour differs sufficiently in other filtration phases.

3. EXPERIMENT

3.1 Materials

Eight two-dimensional low resolution DCE-MRI sequences of normal kidney perfusion with 256 images were used (acquisition duration: about 12 minutes, temporal resolution: about 3 s). The examinations were performed on a whole-body 1.5T MR-scanner (General Electric Healthcare). A 3D ultrafast gradient echo LAVA sequence was used with the following parameters: 15° flip angle, TR/TE 2.3 ms/1.1 ms. The slice that contained the largest surface of renal tissue was then selected. The initial matrix size was 256×256 with pixel size between 1.172 mm and 1.875 mm (slice thickness: 10 mm). A rectangular area containing kidney was delineated (size between 47×35 and 84×59). In-plane movements due to respiration were corrected by a rigid registration algorithm including translations and rotation. Because of rapid and high contrast changes during perfusion mutual information was chosen as a similarity criterion [6]. Anyway through-plane motions remained and frames were highly noisy. An example of frames for three different perfusion phases is shown in figure 1.

3.2 Manual segmentations by radiologists

The different sequences were presented to two experienced radiologists (OP1 and OP2) after automatic registration. They had to delineate three ROIs, namely the cortex, the medulla and the pelvo-caliceal cavities as well as a global kidney mask. To do so, the following procedure was set up:

1. visualization of the complete sequence,
2. selection of a late phase frame were cavities contrast is maximum and manual segmentation of the cavities,
3. identification of the frame corresponding to the cortical enhancement peak and manual segmentation of the cortex
4. segmentation of medulla by difference with cortex and cavities already segmented.

A global mask was then extracted as the common area of the two manual segmentations, including the three ROIs delineated by the two radiologists. This mask was subsequently used for functional segmentation. Two examples of manual segmentations can be seen in figure 5.

3.3 Discrepancy criteria for segmentation comparison

For each of the eight cases a manual segmentation is considered as a reference. The functional segmentation obtained thanks to the proposed method or another manual one will both be compared to this reference.

Each segmentation can be considered as a binary map, with label 1 inside the ROI and label 0 outside. Let be R the reference segmentation and T the tested one. Four types of pixels can then be defined, according to their labels in R and T:

Pixel type	Label in R	Label in T
True Positive (TP)	1	1
False Negative (FN)	1	0
False Positive (FP)	0	1
True Negative (TN)	0	0

Four discrepancy measures between R and T are evaluated for each ROI:

- percentage overlap $PO = 100 \times TP/(TP+FN)$, i.e. percentage of pixels of the reference ROI that are in the test ROI too,
- percentage extra $PE = 100 \times FP/(TP+FN)$, i.e. the number of pixels that are in the test ROI while they are out of reference ROI, divided by the number of pixels in the reference ROI,
- similarity index $SI = (2 \times TP)/(TP+FN+FP)$. SI is sensitive to both differences in size and location [7]. For

Segmentation by	OP2	GNG-T
Well classified pixels (%)	69.8	83.7
Overlap (%)	71.8	83.2
Extra pixels(%)	9.2	21.9
Similarity index	0.79	0.81
Mean distance to reference contour	0.6	0.8

(a) Cortex

Segmentation by	OP2	GNG-T
Well classified pixels (%)	84.8	73.0
Overlap (%)	84.0	73.0
Extra pixels(%)	56.6	33.7
Similarity index	0.70	0.71
Mean distance to reference contour	1.0	0.8

(b) Medulla

Segmentation by	OP2	GNG-T
Well classified pixels (%)	74.9	68.7
Overlap (%)	73.9	69.8
Extra pixels(%)	16.1	11.1
Similarity index	0.77	0.77
Mean distance to reference contour	0.8	0.7

(c) Cavities

Segmentation by	OP2	GNG-T
Well classified pixels (%)	74.9	77.6

(d) Global kidney

Table 1: Discrepancy measures for segmentations of the three ROIs, where OP1 is considered as a reference

Segmentation by	OP1	GNG-T
Well classified pixels (%)	89.7	85.8
Overlap (%)	89.0	85.7
Extra pixels(%)	36.1	34.8
Similarity index	0.79	0.78
Mean distance to reference contour	0.7	0.9

(a) Cortex

Segmentation by	OP1	GNG-T
Well classified pixels (%)	60.3	69.7
Overlap (%)	60.5	69.9
Extra pixels(%)	11.8	16.3
Similarity index	0.70	0.75
Mean distance to reference contour	1.0	0.9

(b) Medulla

Segmentation by	OP1	GNG-T
Well classified pixels (%)	81.8	74.9
Overlap (%)	82.2	76.4
Extra pixels(%)	32.4	12.6
Similarity index	0.77	0.80
Mean distance to reference contour	0.9	0.5

(c) Cavities

Segmentation by	OP1	GNG-T
Well classified pixels (%)	74.9	76.7

(d) Global kidney

Table 2: Discrepancy measures for segmentations of the three ROIs, where OP2 is considered as a reference

instance two equally sized ROIs that share half of their pixels would yield $SI = 1/2$. A ROI covering another that is twice as little would give $SI = 2/3$. For a perfect segmentation the SI value would be 1.

- mean distance (in pixel) between contours of test and reference segmentation.

SI is the only selected criterion that is independent of the chosen reference, however its values appear twice in the tables in order to facilitate comparisons.

4. RESULTS

Examples of two manual segmentations and of a functional semi-automated one can be seen in figure 5. For this case, size of ROIs varies between 532 and 700 pixels for cortex, 375 and 559 for medulla, 161 and 217 for cavities. The delineated contours are superimposed on MR images of the renal pixels (region out of the global kidney mask is black). Frames correspond to perfusion phases during which each compartment is visible at best:

- arterial peak for cortex and medulla
- late phase for cavities.

For almost all the tested kidneys, a very good visual qualitative consistency between these frames and functional segmentation is obtained.

In table 2, on the other hand, the reference is OP2. The percentage of well classified pixels for a given compartment is the sum of TP pixels over the eight cases divided by the total number of pixels for this type of ROI. For global kidney it is the sum of TP pixels for all ROIs over the eight cases divided by the total number of pixels of all kidney global masks. Results for small kidneys have less influence on this percentage than on mean overlap.

Similarity measures between functional segmentation and any manual segmentation are very similar to those computed between the two manual ones. A better score for overlap is always compensated by an increase of extra pixels. The percentage of globally well classified pixels is even higher for functional segmentation, and the similarity index and the mean distance between contours are most of the time better. This was not the case for k-means clustering of the time-intensity curves, where scores were almost less good for functional segmentations [3]: for instance an increase of 3 to 6% in the percentage of globally well classified pixels can be noted for the new method. Furthermore results do not depend significantly on the type of ROI. The quality of functional segmentation does not change with the region size: cavities that are much smaller than the two other ROIs are recovered as well as them. Moreover this technique is fast and user-friendly. The size of the prototype set stemming from GNG-T varies between 10 and 30 nodes: it depends essentially on temporal comportment complexity induced in particular by vascularization artifacts and highly noisy acquisition but little on kidney size. Nevertheless operator has only to adjust two thresholds: the first allows to extract cavities by adding or taking off the most relevant nodes, whereas the second is used in the same way to set relative areas of cortex and medulla. The splitting step with GNG-T allows to consider first the global time-intensity evolution and to reduce noise effect; the threshold adjustment is then easier be-

cause at each tuning level a relatively numerous set of pixels with homogenous temporal evolution is added.

5. CONCLUSIONS AND PERSPECTIVES

A semi-automated method for functional segmentation of internal kidney structures using DCE-MRI sequences was tested and compared with manual segmentations by radiologists. Good qualitative consistency between the two types of segmentation is observed. Similarity measures between a manual segmentation and a functional one are similar and often better than the same criteria evaluated between two manual segmentations. Let us note that the derived time-intensity curves of each compartment are almost identical for functional or manual segmentation. Thus the method is suitable for renal segmentation from DCE-MRI. Moreover this technique is user friendly because the only manual intervention during the whole segmentation process consists in the coarse real-time tuning of two independent thresholds. It is also greatly faster than manual segmentation: the latter requires 12 to 15 minutes for one sequence, versus about 30 seconds for the former, including threshold adjustment.

REFERENCES

[1] N. Grenier, F. Basseau, M. Ries, B. Tyndal, R. Jones, and C. Moonen, "Functional mri of the kidney," *Abdominal Imaging*, vol. 28, no. 2, pp. 164–175, March 2003.

[2] N. Michoux, J.P. Vallee, A. Pechere-Bertschi, X. Montet, L. Buehler, and B. Beers, "Analysis of contrast-enhanced mr images to assess renal function," *MAGMA Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 19, no. 4, pp. 167 – 179, 2006.

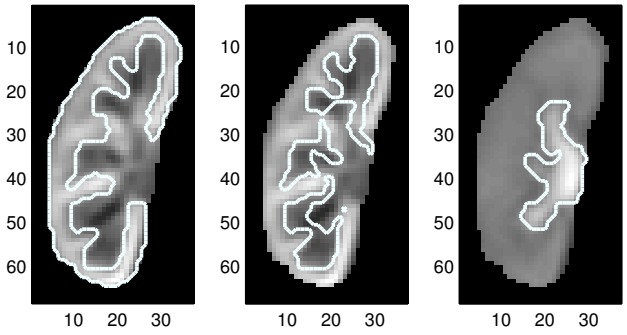
[3] B. Chevaillier, Y. Ponvianne, J.L. Collette, D. Mandry, M. Claudon, and O. Pietquin, "Functional semi-automated segmentation of renal dce-mri sequences," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008)*, Las Vegas (NV, USA), April 2008, 4 pages.

[4] H. Frezza-Buet, "Following non-stationary distributions by controlling the vector quantization accuracy of a growing neural gas network," *Neurocomputing*, vol. 71, no. 7-9, pp. 1191–1202, March 2008.

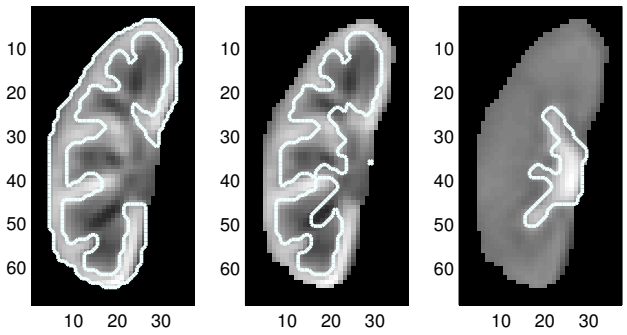
[5] B. Fritzke, "A growing neural gas network learns topologies," in *Advances in Neural Information Processing Systems 7*, G. Tesauro, D. S. Touretzky, and T. K. Leen, Eds., pp. 625–632. MIT Press, Cambridge MA, 1995.

[6] J.P.W. Pluim, J.B.A. Maintz, and M.A. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Trans. Med. Imaging*, vol. 22, no. 8, pp. 986 – 1004, 2003.

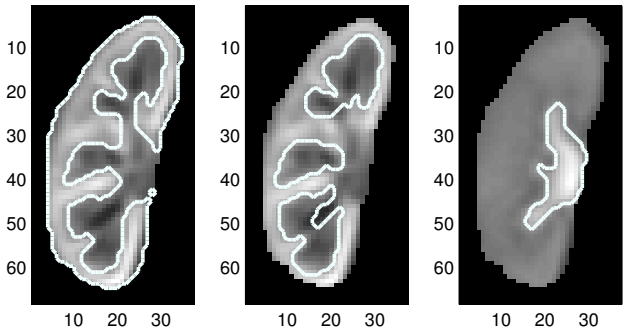
[7] A.P. Zijdenbos, B.M. Dawant, R.A. Margolin, and A.C. Palmer, "Morphometric analysis of white matter lesions in mr images: method and validation," *IEEE Trans. Med. Imaging*, vol. 13, no. 4, pp. 716 – 24, 1994.



(a) Anatomical manual segmentation (OP1)



(b) Anatomical manual segmentation (OP2)



(c) Functional semi-automated segmentation

Figure 5: Example of cortex (left), medulla (middle) and cavities (right) segmentations