# Growing time-homogeneous neighborhoods for denoising and clustering Dynamic Contrast Enhanced-CT sequences

Yves Rozenholc[a,b,*], Markus Reiß[d], Daniel Balvay[c], Charles A. Cuenod[a,c]

[a] *Université Paris Descartes*
[b] *Dept of Appl. Math., MAP5, UMR CNRS 8145, 45 rue des Saints-Pères, 75006 Paris - France*
[c] *LRI, INSERM U970 and Assistance Publique - Hôpitaux de Paris (AP-HP) - European Georges Pompidou Hospital (HEGP)*
[d] *Humboldt University, Institute of Mathematics, Berlin - Germany*

## 1. Introduction

Dynamic Contrast Enhanced Imaging (DCE-imaging) is beginning to be widely used in medical imaging of brain or cancer [1][2][3][4][5][6]. Following the evolution of a bolus of contrast agent injected during a sequential imaging acquisition with Computed Tomography, Magnetic Resonance Imaging or Ultrasound imaging (DCE-CT, DCE-MRI or DCE-US) allows the assessment of such microcirculation parameters as tissue-blood perfusion, permeability, blood and interstitial volume *in vivo* [4][7][8][9]. Taking into account the Arterial Input Function (AIF), such estimations explain the local tissue characteristics [10][11]. Thus DCE-imaging has great potential for cancer detection and characterization, as well as for following and monitoring *in vivo* the effects of treatments [12][13][14][15]. Clinicians are now expecting such information to be given at a high resolution up to pixel-level and to be displayed as parametric maps [16]. This is of main interest in cancerous tumors as they are known to be heterogeneous with areas going from non-perfused and necrotic to hypervascular "hot spots" [17].

Recently, Cao *et al.* [18] showed the ability of DCE-CT to assess intra-tumor physiological heterogeneity in tumors. This offers an *in vivo* tool for the evaluation and optimization of new therapeutic strategies. At pixel-level, however, high-frequency acquisition is achieved with a poor signal to noise ratio. Consequently, to improve the signal to noise ratio either large Regions Of Interests (ROI) are used or the sequences are denoised by spatial averaging or filtering

*Corresponding author
*Email addresses:* yves.rozenholc@parisdescartes.fr (Yves Rozenholc ), mreiss@mathematik.hu-berlin.de (Markus Reiß ), daniel.balvay@yahoo.fr (Daniel Balvay ), charles-andre.cuenod@egp.aphp.fr (Charles A. Cuenod)

techniques [19]. Both procedures have the disadvantage of mixing dynamics which may not be homogeneous, leading to inaccurate parameter estimation. Using parametric modeling (usually derived from pharmacokinetic models) or smoothing techniques, it is possible to regularize dynamic enhancements [20][21]. Unfortunately, such *time*-regularization techniques lead to information loss due either to smoothing out the dynamics or to the inaccuracies of a specific model [22]. This is an important concern in DCE-CT where, in order to obtain a dynamic sequence, the unitary X-ray dose per image should be as low as possible, which leads to a very poor signal to noise ratio.

Thanks to the linear relationship between the concentration of the contrast agent and the attenuation, using DCE-CT instead of DCE-MRI for example, one can access the contrast agent concentration directly [31]. Let us stress that even if an improvement of the acquisition techniques in CT may be expected, the need to control X-ray doses will remain for ethical reasons. In the last decade, many techniques have been introduced to address the problem of the tomographic Radon reconstruction of one (2D- or 3D-) image when undersampled measurements (potentially noisy), which strongly violate the Nyquist theorem, are used. We can distinguish between techniques A/ based on iterative reconstruction which rely on the accurate modeling of the distribution of noise in the acquired data [23]; B/ using sparsity inside the tomographic image and related to compressed sensing [24, 25] such as Prior Image Constrained Compressed Sensing (PICCS) [26] or Vastly undersampled imaging with projections (VIPR); C/ which benefit from the correlation between voxels in time-dependant images like Highly Constrained Back Projection (HYPR), see the review [27] and references within; D/ aimed at smoothing, estimating or interpolating the sinogram before reconstruction [28, 29]. Allowing undersampled measurements in 2D or 3D tomographic acquisitions, all these techniques offer the possibility to reduce the X-ray dose in CT and have generalized the use of so-called "low dose" CT. Unfortunately, these techniques, which are for the practitioner implemented inside the scanner by the constructors as black-boxes, produce as output one image (2D-slice or 3D-volume) by acquisition time and do not take advantage of the long total acquisition time used in DCE-imaging which is of order 100 seconds. Even multi-band filtration [30], which uses multiple acquisitions from the same image in order to increase the signal to noise ratio by using a Maximum A Posteriori (MAP) approach, relies on the assumption that images do not change during acquisition. A hypothesis which is clearly violated in DCE-CT due to the variations of the contrast agent concentrations.

We aim to improve the signal to noise ratio of DCE-CT *after* the Radon reconstruction of each slice obtained at each time acquisition. Because the DCE-CT sequences, we focus on, are made from about 100 CT-slices, a grail would be to reduce the X-ray dose of each slice by the same 100 factor. It should be clear that our aim is not to replace the technics used to achieved "low dose" CT but to complete such technics by taking advantage of the dynamical acquisition in order to go further in the reduction of the X-ray dose. Here, the increase of signal to noise ratio is achieved in two steps using a global procedure, based on statistical multiple testing to compare the dynamics observed at each voxel of

the image. Using a test of zero mean for the difference between enhancements or estimated enhancements, no specific modeling assumption of the dynamics is either used or made. Being adaptive to the unknown smoothness of the curves, multiple tests have enough power to properly differentiate between dynamics.

The first step is a voxel-wise denoising step : at each voxel, the dynamic is compared to those from the voxels around the selected location. Homogeneous dynamics are aggregated iteratively in order to build a denoised version of the dynamic coming from the selected voxel. At each iteration, using multiple tests, the dynamics from the next closest voxels around the selected voxel, are compared to each previous denoised versions in order to check for homogeneity. As a result, around the selected voxel, a neighborhood of locations is build which is aimed to satisfy the bias-variance paradigm of non-parametric statistics : 1/ being large enough, to reduce the variance and improve the signal to noise ratio; 2/ being not too large to ensure a good homogeneity in between the selected voxels. This step provides high resolution information with an improvement in signal to noise ratio which allows, for example, to build parametric maps.

The second step automatically builds a clusterization of all the dynamics, preserving dynamic homogeneity in each cluster. The clustering procedure is built from the result of the denoising step: it uses as input the neighborhoods grown in the denoising procedure and their associated denoised enhancements. This clusterization depends neither on an *a priori* knowledge of the number of classes like $k$-means [38] nor on any assumption of behavior inside a class such as a Gaussian mixture classification with EM algorithm [39]. It is based on the same multiple tests as those previously used in the denoising step. The clustering procedure constructs classes made from a set of voxels with enhancements statistically close to an estimated enhancement (built as a class centroid).

The clusters can be viewed as an automatic ROI-partition selection with respect to the typical behavior of the dynamics. This automatic ROI-partition provides a summary of the full dynamic sequence into few typical denoised dynamics which preserve the heterogeneity of the tissues in general and of the tumor in particular. It can be compared to a piecewise constant representation of a function having functional values.

From a clinical point of view, this denoising and clustering procedure is a necessary step to allow a relevant evaluation of the microcirculation *in vivo* by using pharmacokinetic models on a pixel by pixel basis.

Only one hyper-parameter is used in these two stages, namely the level of the multiple test. In the denoising step, it controls how easily the dynamics are aggregated. In the clustering procedure, it plays a role equivalent to a penalization ensuring an adaptive control on the number of classes.

The article is organized as follows: we first introduce the statistical framework and comment on the assumptions in Section 2. In Section 3 we outline our denoising method for DCE-CT. The automatic and unsupervised classification of tissues is tackled in Section 4. Finally in Section 5, we show an application to DCE-CT data on liver metastases and some simulations in order to validate our methods. The two main statistical tools used to construct our denoising procedure are described in the Appendix : multiple testing in 7.1 and neighbor-

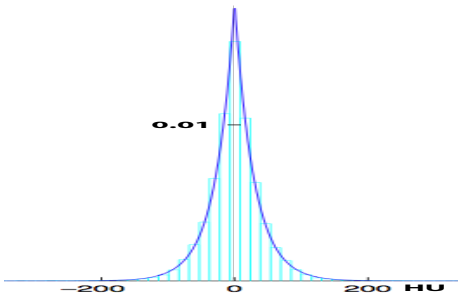


Figure 1: Histogram of the noise in DCE-CT baseline images with fitted Laplace distribution with parameter $\lambda = 29$ ($\sigma = 41HU$)

hood/ring growth in 7.2.

## 2. Statistical framework for DCE-CT

In the rest of this paper, we consider a DCE-CT sequence as a finite sequence of noisy images indexed by both time and space:

$$I = \{I_x(t), x \in \mathcal{X}, t \in \{t_1, t_2, \ldots, t_K\}\},$$

where $I_x(t)$ denotes the noisy enhancement at time $t$ and voxel location $x$. We will denote by $I_x$ the vector

$$(I_x(t_1), \ldots, I_x(t_K)),$$

of the discretely observed dynamics at location $x$ at all observation times $t_1$,..., $t_K$. Here $\mathcal{X}$ denotes the finite voxel grid and the $t_j$ are the acquisition times.

We assume that the observable gray level $I_x(t)$ may be written as

$$I_x(t) = i_x(t) + \sigma\varepsilon_x(t), \tag{1}$$

where $i_x(t)$ denotes an unobservable true gray level, $\varepsilon_x(t)$ denotes a standardized noise and $\sigma$ the noise level. We assume that the noises $\varepsilon_x(t_j)$ are independent with respect to both space location $x$ and time location $t_j$. While time independence may be easily justifiable, the assumption of independence between locations is known not to hold in CT. Especially, it is well known that CT creates radial artifacts. We use this assumption as a simplified model. A more complete modeling should certainly be written as:

$$I_x(t) = i_x(t) + \eta_x(t) + \sigma\varepsilon_x(t), \tag{2}$$

where $\eta$ is a time-independent, but space-correlated noise which explains spatial artifacts.

Even if these spatial artifacts are clearly visible by the human eye, it seems that the noise $\eta$ may be considered as negligible from a statistical point of view.

In the framework of DCE-CT where, because of low X-ray dose, $\sigma$ is clearly large, our results (see Section 5) indicate that this simplified model (1) may be applied.

To help the presentation, we will assume in the following that the parameter $\sigma$ is known even if this is not necessary. The knowledge of $\sigma$ may be ensured by either a proper calibration of the scanner or an independent estimation. We have assumed for the presentation that the distribution of the errors $\varepsilon_x(t)$ is Gaussian and used estimates based on the mean of enhancement vectors. A simple histogram of the DCE-CT noise, see Figure 1, shows that this noise distribution is certainly closer to a Laplace distribution. In this Laplace setting, median based estimates are more advisable (see [36]) and we have used this setting in practice (see Section 5.2). From an industrial point of view, it is possible to avoid extra calibration or noise level estimation, using the result in [34]. In such case the only assumption needed is that, for any couple of locations $x$ and $y$, the distribution of $\varepsilon_x(t) - \varepsilon_y(t)$ is symmetrical without mass in 0 for all time $t$.

## 3. Denoising DCE-CT

We outline here the statistical procedure used to denoise DCE-CT which is summarized by the flowchart in Figure 2 and the Algorithm 1 in Appendix 7.3. It is based on two statistical tools introduced in Appendix 7.1 and 7.2.

Our method is based on the statistical comparison between two observed enhancement vectors $I_x$ and $I_y$ at spatial locations $x$ and $y$ as presented in Appendix 7.1. Two enhancement vectors are considered indistinguishable if their difference does not deviate significantly from the zero vector. This is controlled by a statistical multiple test on whether the difference vector $I_x - I_y$ has mean zero vector or not. In the positive, we will call $I_x$ and $I_y$ to be *statistically close* or *time homogeneous* and we write $I_y \equiv^{\alpha}_{\sigma^2} I_x$ where $\alpha$ refers to the test level and $\sigma^2$ to the noise variance.

Such comparisons of vectors using multiple test procedures have been developed in the Gaussian framework by [32, 33] and for heterogenous symmetric noise by [34]. The power of these test procedures are known to be adaptive to the regularity of the underlying signal. Hence, we do not need to specify the behavior of the enhancements in order to be able to deal adaptively and automatically with the particularities of the enhancements. We only need the enhancements not to be too wild in the sense that their differences should have a certain minimal (Hölder ) regularity [37]. The latter being clearly satisfied in the context of contrast enhancements in DCE-CT.

The use of differences ensures that noise can be assumed symmetrically distributed, thus avoiding typical problems that spring from existing structures in tomographic sequences as described in [35].

At each spatial location $x \in \mathcal{X}$, we aim to construct a spatial $x$-neighborhood $\mathcal{V}_x$ made of voxels $y \in \mathcal{X}$ such that $I_y$ is statistically close to $I_x$ and such that the statistical error may be controlled. The estimated enhancement vector at

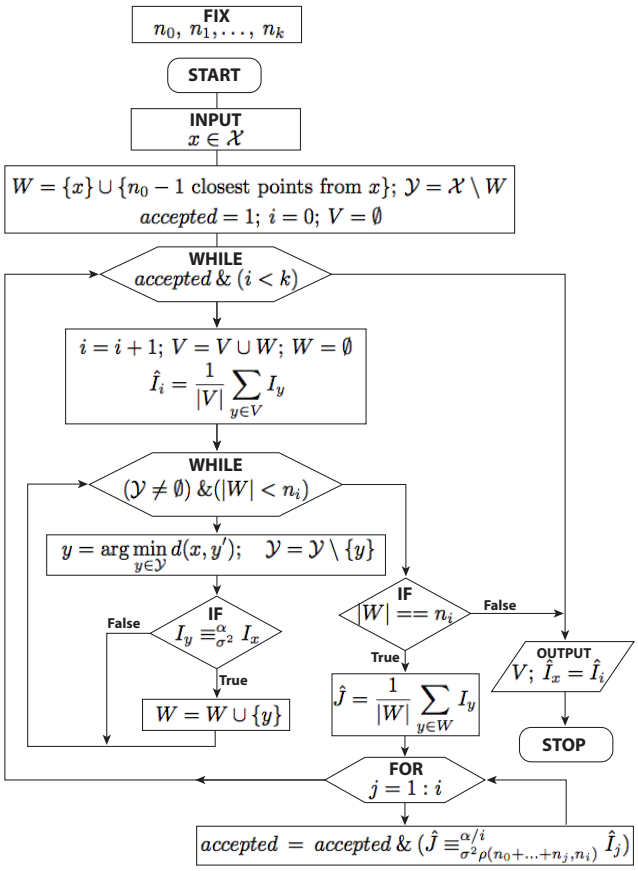


Figure 2: Flowchart of the spatially pointwise denoising procedure. See Appendix 7.1 for precisions on the test denoted $\equiv^{\alpha}_{\sigma^2}$ and Appendix 7.2 for precisions on $\rho$. Here, the centroids $\hat{I}_i$ and $\hat{J}$ are the means of the enhancements in their corresponding neighborhoods; other possibilities such as generalized medians are discussed in Apendix 7.2.

location $x$ is then derived from this neighborhood as a centroid given for example by the empirical mean or by the generalized median (see Eq. (3) and (4) in Appendix 7.2).

Controlling the statistical error means that we aim at including as many voxels as possible in $\mathcal{V}_x$ to reduce the variance while the bias due to over-smoothing is kept small.

To that end, as presented in detail in Appendix 7.2, at each spatial location (called center), an increasing sequence of "time homogeneous" neighborhoods of voxels is grown according to the following steps: (i) using the difference between their enhancements, compare the center to the voxels which are spatially close and select the voxels with statistically close enhancement; (ii) construct a sequence of estimates – each built on one neighborhood – having decreasing

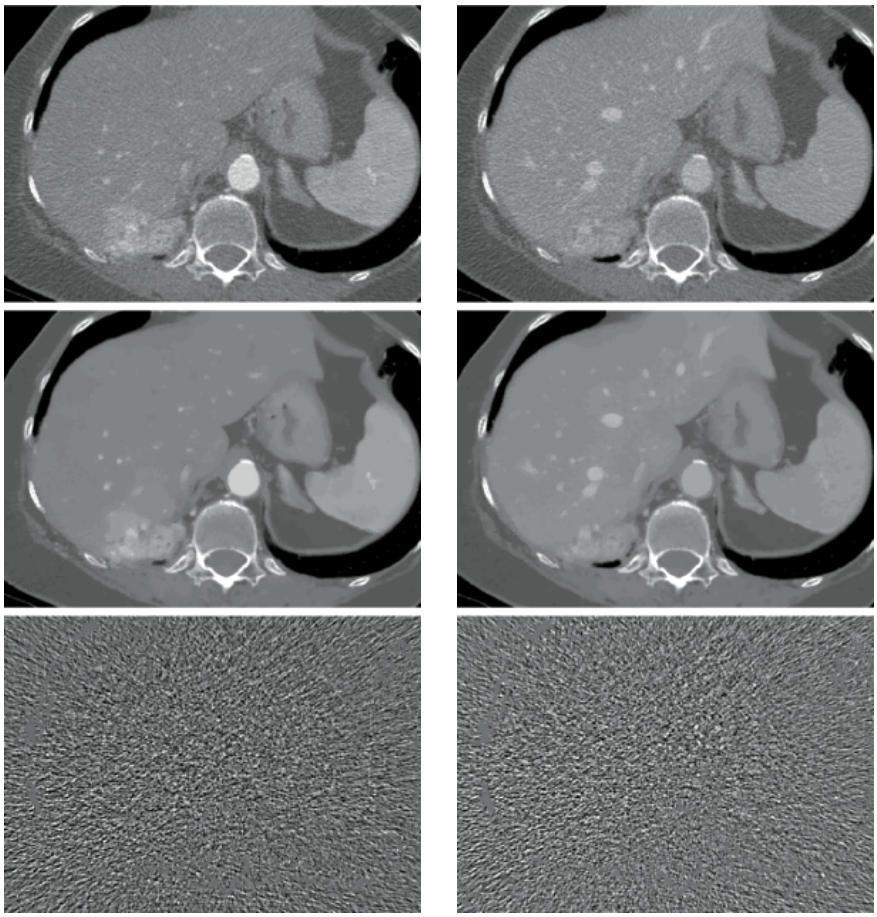

Figure 3: Axial CT images of the abdomen centered on a liver metastasis. These images belong to a dynamic series, they have been acquired 25 seconds (left column) after the beginning of the acquisition, during the arterial phase and respectively 45 seconds (right colum) after the beginning of the acquisition, after the arterial phase. For each (column) time : Original (top) - Denoised (middle) - Residuals (bottom).

variance; (iii) from this sequence select the largest statistically acceptable neighborhood which is expected to realize the statistical paradigm of the bias-variance tradeoff.

Step (i), see Eq. (5), is a pre-selection step which allows the denoising construction to obtain "neighborhoods" built from different objects of the same type over long distances. This is useful in medical images where the same kind of tissue (e.g. small arteries, see Figure 5-a) may reappear in different areas. At step (ii), taking a centroid (mean or median) of the enhancements associated to the voxels included in the spatial neighborhood provides an enhancement estimate. The last step (iii) involves a generalization of the above multiple test

in order to decide whether or not two estimated enhancements $\hat{I}_V$ and $\hat{I}_W$, constructed on two disjoint sets $V$ and $W$, are statistically indistinguishable.

Classically, because of the increasing property of the neighborhoods, the known variances of these estimates decrease and can be used in order to select an estimate that realizes a good statistical trade-off between bias and variance in step (iii) [36]. The precise description of the neighborhood growth is described in Appendix 7.2.

## 4. Clustering

Using the test procedure introduced in Appendix 7.1 and denoted by $\equiv^{\alpha}_{\sigma^2}$, we develop an automatic clustering procedure. It does not rely on the knowledge of the number of classes like in the $k$-means algorithms [46, 38]. It also does not need any Gaussian assumption, often made to describe the behavior inside a class. The classes are built in order to keep statistical homogeneity. Let us denote by $\mathcal{C}$ the clusterization *i.e.* the set of all classes. The estimated enhancement associated to a class $c \in \mathcal{C}$, called *center* of $c$, will be denoted $\hat{\mathbf{I}}_{\mathbf{c}}$. The clustering algorithm is a recursive algorithm with a main loop which can be decomposed in three main steps : a/ define a class from a single voxel ; b/ test if a new class center is statistically close or not from any existing class center ; c/ gather the classes with statistically close centers. The algorithm stops when all single voxels have been assigned to a class and when classes cannot be merged anymore.

Using the construction described in Appendix 7.2, Step a/ can be done in an efficient way, leading to a feasible algorithm. Suppose that, using the denoising procedure described in Section 3, we have grown at each voxel location $x$ of $\mathcal{X}$ a neighborhood $\mathcal{V}_x$. Given a list $\mathcal{L} \subset \mathcal{X}$ of voxels of interest (a Region Of Interest (ROI) or the full image), we define the child/ancestor relation $\preceq$ by "$z \preceq y$ if $z \in \mathcal{V}_y$" and call "children of $x$ in $\mathcal{L}$" the set

$$N_x^{\mathcal{L}} = \{z \in \mathcal{L}, \exists y_1, ..., y_k \in \mathcal{X} \text{ s.t.} \\ z \preceq y_k \preceq \ldots \preceq y_1 \preceq x\}.$$

The list $N_x^{\mathcal{L}}$ of all children of $x$ is constructed in the function `Children`, see Algorithm 3 in the Appendix.

**While the list $\mathcal{L}$ is non empty, the following four steps are processed :**

**Step 0 - Next** : Consider one voxel $x$ in the list $\mathcal{L}$ with the largest neighborhood size $|\mathcal{V}_x|$. Compute its child list $N_x^{\mathcal{L}}$ and set $c = N_x^{\mathcal{L}}$.

**Step a - Class construction** : From $c$ build a *robust center* $\hat{\mathbf{J}}$ (see hereafter for the construction of this robust center). A new (possible) class $c_{\hat{\mathbf{J}}}$ is defined by its center $\hat{\mathbf{J}}$ and the list $c_{\hat{\mathbf{J}}}$ of voxels $y$ in $c$ such that $\hat{I}_y \equiv^{\alpha}_{\sigma^2 \rho(1, \min(|\mathcal{V}_y|, |c|))} \hat{\mathbf{J}}$

is accepted. Set $\mathcal{L} := (\mathcal{L} \cup c) \setminus c_{\hat{\mathbf{J}}}$ to remove selected points from the list.

**Step b - Class checking** : The new class $c_{\hat{\mathbf{J}}}$ with center $\hat{\mathbf{J}}$ made at the previous step is checked for closeness with all previously constructed classes. If a class $c$ is found to satisfy $\hat{\mathbf{I}}_c \equiv^{\alpha}_{\sigma^2 \rho(1,\min(|c|,|c_{\hat{\mathbf{J}}}|))} \hat{\mathbf{J}}$, then $c_{\hat{\mathbf{J}}}$ and $c$ are merged together to form a new class $c$ and the algorithm returns to Step a/ with this new $c$.

**Step c - Update Cluster List** : The class $c_{\hat{\mathbf{J}}}$ is added as a new class to the cluster list.

The clustering procedure is synthetically described in Algorithm 2 in the Appendix.

We now give the details of the *robust center* construction. Given a set $c$ of voxels, we iteratively define $\hat{\mathbf{J}} = \hat{\mathbf{I}}_c$ and update $c$ as the set of the $y$ in $N_x^{\mathcal{L}}$ such that the test $\hat{I}_y \equiv^{\alpha}_{\sigma^2 \rho(1,\min(|\mathcal{V}_y|,|c|))} \hat{\mathbf{J}}$ is "strongly" accepted. This procedure, which can be viewed as a robust $k$-means procedure with $k = 1$, converges quickly and will stop after a few iterations.

*Remark 1:* In this setting, the larger class attracts the smaller class in its statistical surrounding defined by a "noise level radius" of order $\sigma^2$ around its center. Moreover the center of the smaller class is located up to a statistical distance of order $\sigma^2 / \min(|c|, |c'|)$. Hence, the two classes are merged if the "statistical distance" between their centers is less than $\rho(1, \min(|\mathcal{V}_y|, |c|))$ (see Appendix 7.2).
*Remark 2:* Thanks to the decreasing order in the neighborhood sizes, the last unclusterized voxels are mainly due to exceptional behaviors like body movements and are far from all previously constructed centers. Because of this observation one can stop the clustering algorithm when the neighborhood size is less than a prescribed $v_0$. The remaining voxels then define an extra class of unselected voxels. This extra class of movements can be used as a prior input for a registration algorithm.

## 5. Results

### *5.1. Data material*

We use a DCE-CT sequence of 53 images obtained, in 90 seconds, at the same level of the upper abdomen on a patient who wass asked to hold his breath. Acquisition parameters were fixed at 80 Kv and 50 mAs. These images were obtained with an in-plane resolution of 512×512 pixels. The sequence is split into three periods of 30-second breath-hold separated by 8-second pauses to allow for free breathing. Each period is characterized by the time delay between two images: one second between images for the first 30 images; two seconds for the next 15 images and three seconds for the last 8 images. After the beginning of the acquisition, at time 3 seconds, an intravenous bolus of

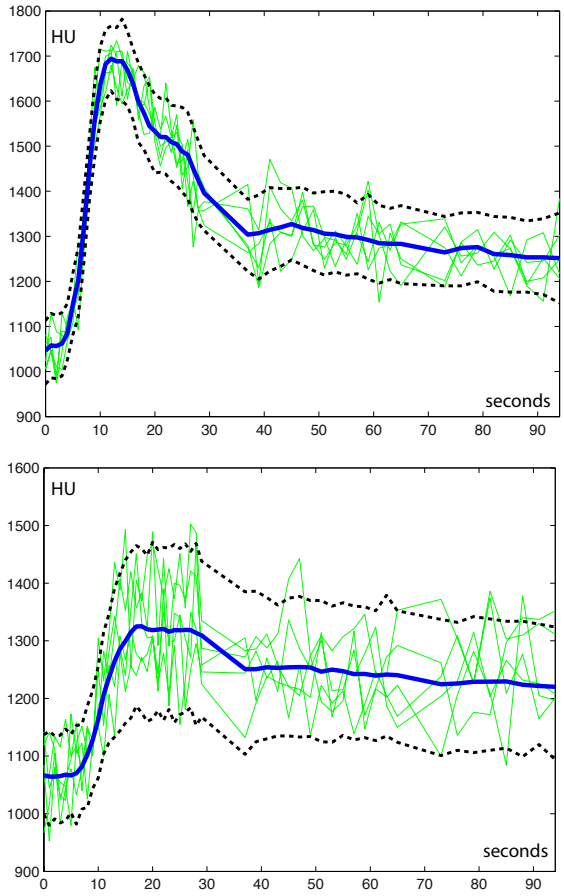


Figure 4: Enhancement curves obtained from manually drawn ROI containing more than 100 voxels within the Aorta (top) and the tumor (bottom). For each sub-figure, the mean curve inside the ROI is the thick plain line, the 5% and 95% quantiles of the curves of the selected points are in thick dotted lines and finally the curves of 5 randomly selected points inside the ROI are drawn in fine lines.

80 ml of Iobitridol (Xenetix, Guerbet, France), an iodinated contrast medium, is injected at the rate of 4-5 ml/s and followed inside the tissues through their enhancements. At time $t$, the contrast agent concentration is proportional to the difference of gray levels between time $t$ and time 0. On CT images, gray-level ranges of the different tissues, which are measured in Hounsfied Units (HU), are related to their chemical composition (fat, air, bone, water,...) and their content in contrast media. On the images, the liver, the spleen, the aorta, the stomach, a vertebra and some blood vessels are visible. However, the images suffer from a poor signal to noise ratio, due to the limited irradiation dose used for the sequential acquisition. Figure 4 shows typical enhancements obtained from voxels in manually selected ROI (of size larger than 100 voxels) within the

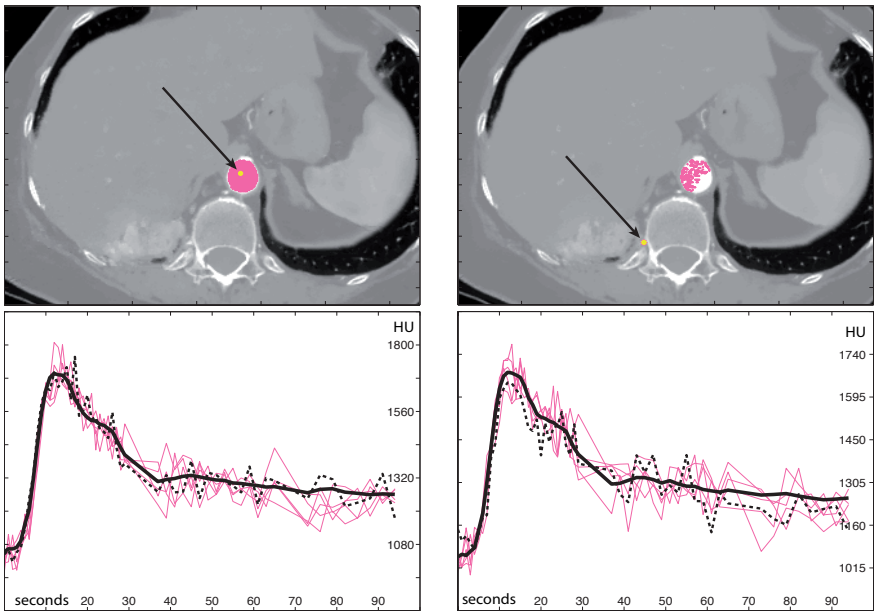


Figure 5-a: Left column : Voxel within the aorta - Right column : Voxel within a small dorsal artery on the right side of the spine. For each column, (Top) the arrow points to the selected voxel $x$ on one axial slice of the sequence; the pink dots show the constructed neighborhood $\mathcal{V}_x$. (Bottom) the thick dotted black line shows the tissue enhancement at voxel $x$; the thick black line is the denoised dynamic, centroid of the $\mathcal{V}_x$ dynamics. In addition, individual curves of 5 randomly selected voxels in $\mathcal{V}_x$ are shown in the background in fine pink line.

Aorta (top), known to be homogeneous, and within the tumor (bottom), where heterogeneity is expected. In order to provide a visual idea of the noise level and the variability inside a manual ROI, for each selected ROI, we have constructed a mean curve, and the 5% and 95% quantile curves by considering each time separately and we have drawn 5 curves of randomly selected individual voxels.

Using the first four images without enhancement (baseline images), we have derived a sample of the noise distribution as follows: 1/ compute a (voxel-by-voxel) baseline median image from the first four images; 2/ compute the residual images of the differences between each baseline image and the baseline median image; 3/ use the values inside the residual images as a sample (of size $4\times512\times512$) of the noise distribution. We expect in step 1/, if there are only a few movements, that, in each image, most of the voxels represent the same tissue and the variation in their gray levels are just due to noise. The histogram of the noise distribution obtained from this sample is represented in Figure 1 together with the fitted Laplace distribution.

### *5.2. Results for the denoising procedure*

In Appendix 7.2, we fix the increasing sizes $n_i$ of the neighborhoods to be successively 1, 5, 15, 38, 91, 211, 476. As our procedure involves several multi-

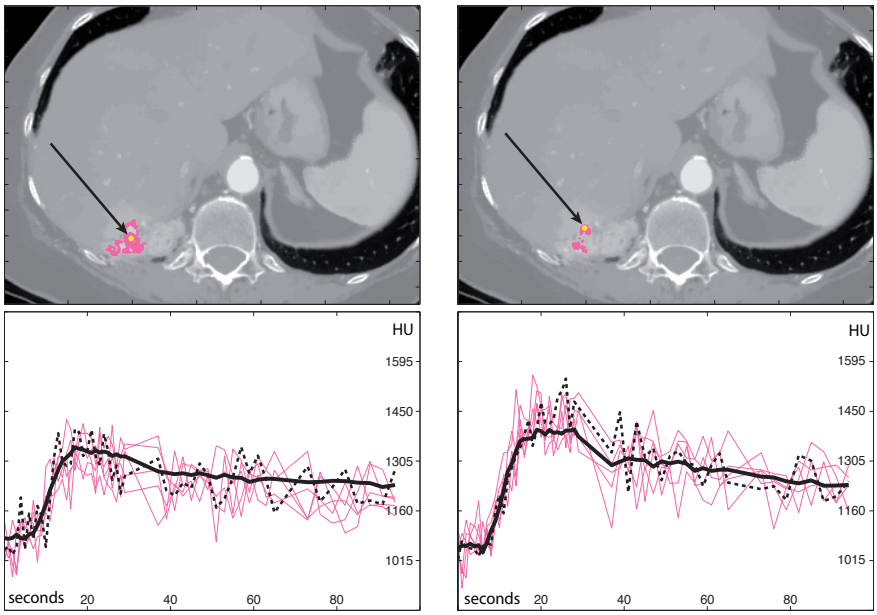


Figure 5-b: Left column : Voxel within the tumor - Right column : Voxel within a tumor "hot spot" : The estimated curve shows a quicker and larger enhancement providing a stronger initial slope that shows a perfusion and a more important blood volume than for the voxel within the tumor itself (left column). See Figure 5-a for more information.

tests, we use a False Discovery Rate (FDR) approach [40] calibrated to take into account the multiplicity due to time (see Appendix 7.1) and due to neighborhood growth (see Appendix 7.2). We apply our algorithm using three different values for $\sigma$ : 41, 50 and $60HU$. These values have been calibrated for this specific sequence as follows: using the noise distribution sample made from the baseline images (see Section 5.1), fit a Laplace distribution as shown on Figure 1; the parameter of this Laplace distribution is $1/29$, which corresponds to a $29\sqrt{2} \approx 41HU$ standard deviation. An estimation of the standard deviation on the same sample leads to a $60HU$ value. The difference between these values may be explained by the heavy tails due to tomographic artifacts and movements. The value $50HU$ has been tuned manually as a mean value between these two estimations. Due to the maximal size of the neighborhood, we may expect for the minimal residual noise to be of level $\sigma/\sqrt{476} \approx 2.75$. Hence, in the worst case considering that $\sigma = 60HU$, our maximal gain in the signal to noise ratio is about $60/2.75 \approx 22$.

Because our procedures rely on tests, a good trade-off between the choice of $\sigma$ and those of $\alpha$ should be made. The value $\sigma = 41HU$, which ideally fits our model, does not denoise sufficiently for any $\alpha$ as it leads to too small neighborhoods for which growth has been stopped artificially by tomographic artifacts. For the value $\sigma = 50HU$, this phenomenon remains at least in the lower left corner of the sequence. With $\sigma = 60HU$ and values of $\alpha$ in a range

0.2 to 0.001 our algorithm has a stable behavior, leading to similar denoised sequences and similar clusterizations showing more details as $\alpha$ gets larger. With $\sigma = 60HU$ the sequence is properly denoised and most of the tomographic artifacts are removed.

The clustering procedure using $\sigma = 50HU$ or $\sigma = 60HU$ with the corresponding $\alpha$ value for the tests involved in the clustering procedure leads to similar and stable segmentations. Parameter $\alpha$ is a good tuning parameter which may be used by doctors to tune the clusterization to let, or not, some details appear. It is remarkable that, when using this statistical tuning parameter $\alpha$, the geometry of the clusters is stable and varies slowly while preserving the morphological details.

In what follows, we use the setting $\sigma = 60HU$ and $\alpha = 0.05$ to illustrate both the denoising and the clustering procedures.

First in Figures 3 and 3, we present the result of our method on two images of our DCE-CT sequence at time 25s and 45s, respectively : (top) the original image; (middle) the denoised image and (bottom) the residuals, difference between original and denoised images. At time 25s, the image shows a clear enhancement in the aorta and in the tumor. At time 45s, the details of the return to the veinous system are visible. On both images, the same quality of details can be found. Morphological information such as shapes and borders of the organs are clearly visible on the denoised image (middle), while most of the noise has been removed. One can notice two types of structures on the residual image. The first structure comes from movements (see for example on the ribs in upper-left corner or around the stomach) as we apply our method on the sequence without using any kind of (pre-) registration or motion correction algorithm. Our algorithm suffers from these movements on the borders of moving tissues and should be associated with a registration procedure in case of strong movements. The second structure comes from CT radial artifacts. In the model given by Eq. (2), the noise $\eta$ is also removed when using our procedure with $\sigma = 60HU$. Again this is a surprise as nothing, from a theoretical point of view, has been used to catch this noise known to have spatial structures. Tomographic artifacts are organized in one image along directions which are not necessarily the same at different time instances. Thanks to the use of the whole temporal structure in the testing, the directions are mixed and the tomographic artifacts do not have a significant effect on the denoising result. The benefit of using an approach which compares not only voxels in one image locally but directly the dynamics of these voxels in the full sequence is clear : one can apply a stronger denoising procedure without losing any details in the 2D structure.

Figures 5-a to 5-b present the result of our method applied to four voxels $x$ in (respectively): (a) the aorta; (b) a small dorsal artery that is located near the spine and crosses our CT-plan on only one or two voxels; (c) the tumor; (d) a tumor hot-spot.

Figures 5-a to 5-b each show the result of our procedure within a voxel and are divided into two sub-figures. Top: the slice at a specific time, with the selected voxel $x$ (yellow circle pointed to by an arrow) and the voxels from its neighborhood $\mathcal{V}_x$ (pink dots). Bottom: the original enhancement vector

$I_x$ (black thick dotted curve) ; five randomly selected associated enhancement vectors $I_y$ for $y \in \mathcal{V}_x$ (pink background curves) and the estimated enhancement (black thick curve) using a generalized median as centroid of the selected neighborhood (see Appendix 7.2).

In Figure 5-a, the neighborhood $\mathcal{V}_x$ of the voxel $x$ is not connected. This voxel $x$ has been selected in a small artery covering only a few voxels and most of its neighborhood is included in the aorta and made from voxels disconnected with the original location. This is not a surprise, as physiologically, the flow in the artery system is such that, at our time resolution, the enhancements in different arteries cannot be distinguished.

These figures show the second benefit of our approach : the sequence is not denoised in a "slice-by-slice " approach as a movie but as a single image showing dynamics. Hence, it is not the gray levels of images which are denoised but the full dynamics, over time, at each voxel. This approach clearly improves the signal to noise ratio by reducing the noise without changing the signal and one can then proceed to a proper inspection of kinetic enhancement curves to derive the characteristics of the underlying physiological processes.

Let us emphasize that the maximum of the estimated enhancements in the hot-spot inside the tumor (Figure 5-b) is larger than those associated to the surroundings of the tumor (Figure 5-b). This is in accordance with clinical knowledge as the flow in such voxels is larger.

*5.3. Results for the clusterization procedure*

We present here the classifications obtained by our automatic clusterization procedure described in Section 4 applied to the denoised sequence obtained with the setting $\sigma = 60HU$ and $\alpha = 0.05$. See Section 5.2 for discussion about this choice. Like in Section 4, the voxels with neighborhoods of size smaller than 60 voxels, which correspond to movements, are not classified. They appear in Figures 6 and 7 in dark blue. The final clusterization of this series contains more than 600 clusters. Most of these clusters are small and correspond to movements.

Figure 6-(top) shows (with artificial colors) the result of our procedure applied to the enhancements in order to focus on functional phenomena. The enhancements are obtained after removing, at each voxel, a baseline intensity estimated on the first 5 times prior to the contrast agent injection. Only 30 clusters have a size larger than 60 voxels. In this figure, the 30 clusters are shown and nine typical clusters are localized. Air and tissues without enhancement are localized in Cluster 4. Cluster 1 mainly consists of the liver. The aorta is well distinguished (Cluster 3). Hepatic veins correspond to Cluster 5 and smaller vessels to Cluster 6. The tumor is split into four types of behavior (7, 2, 8 and 9) extending from tissues at the periphery, which are compressed by the tumor, to hot spots. Due to similar enhancement patterns in reaction to the specific contrast agent used to obtain the DCE-CT sequence, the tissues of the spleen and the stomach are not properly distinguished. While Cluster 5 corresponds to the center of hepatic veins, Cluster 6, which contains smaller vessels, also contains the borders of the same hepatic veins. This may be due

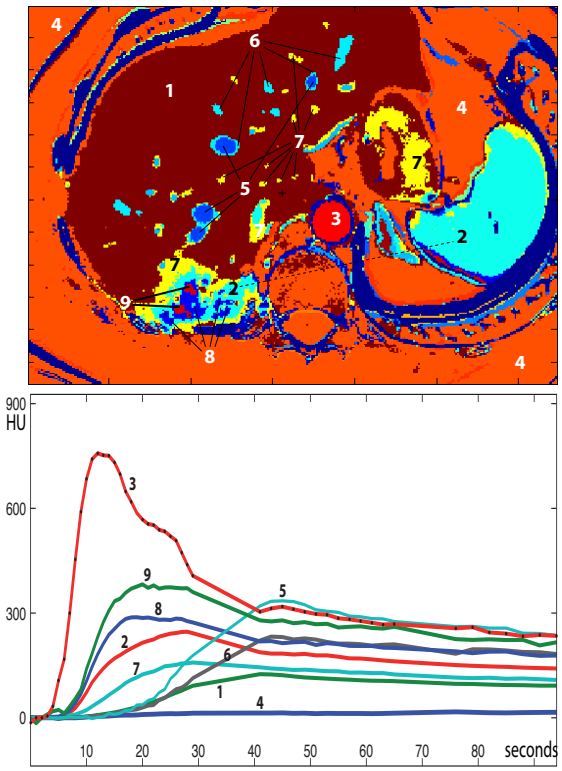


Figure 6: Clusterization result based on the enhancement curves (variation from the baseline) - (top) clusters in artificial colors ; (bottom) estimated enhancements, "centroids" of the clusters. The numbers link the estimated enhancements to the clusters.

to a partial volume effect or due to surface tension in large vessels: the flow on their borders being not as fast as in their centers.

The centers (estimated enhancement) of these nine relevant clusters are presented in Figure 6-(bottom) with their associated label number. These curves show very high signal to noise ratio without time regularization. These curves appear as a proper summary of the dynamic information that exists in the sequence. From this summary, one can easily derive the characteristics of the underlying physiological processes.

Figure 7 shows (with artificial colors) the result of our procedure directly applied to the original image, without baseline removal. This strategy leads to a differentiation of tissues with a different baseline even if they have same enhancement. For tissues showing enough enhancement, the clusterization is similar to what is obtained in Fig.6 (see tumor tissues (2,7,8,9) for example). The other tissues (4) are here distinguished through their baseline offering a different information to the physiologist. In this clusterization, 45 clusters have more than 60 voxels. Only these 45 clusters are displayed.

*5.4. Simulation results*

In order to offer a validation for evaluating our method, we have run artificial experiments on simulated data in the ideal case when no movements occur during the sequence. We have built a synthetic dynamical image ($120\times60\times64$) without baseline with 64 observation times using the synthetic enhancements shown at the bottom of Figure 8-a. We have chosen time 30 to present our synthetic dynamical image and to illustrate our simulation results. At the top of Figure 8-a, the homogenous regions are surrounded by artificial fine white lines in order to help the visualization of the pattern. These homogenous regions aim to represent air (0), normal tissue (1), two types of veinous systems (2, 3) and the aorta (8). The large ball in the upper-right corner of the Figure 8-a represents a tumor with behaviors that vary continuously in a piecewise linear way from the center to the border. Inside this synthetic tumor, enhancements may be computed from the radius using the profile given by the thick line in Figure 9 and enhancements 4 to 7 in Figure 8-a. At radius 0.7 for example, the profile yields the value 5.2 which corresponds to a mixture of enhancements 5 and 6 in proportion 20% and 80%, respectively.

From this synthetic dynamical sequence, we have constructed a simulated noisy version by adding (time and space) independent Laplace noise with standard deviation 41HU as shown, at time 30, in Figure 8-b.

Our denoising algorithm (see Section 3) has been applied to this sequence with $\sigma = 60HU$ and $\alpha = 0.05$ using the same values as in Section 5.2. The result of this denoising step is presented (at time 30) with the same gray-scale as the observations in Figures 8-c and (to show the gain in contrast) with its full gray-scale range in Figure 8-d. To help the comparison, this last gray-scale is also used in the top Figure 8-a which shows the original data.

Considering the amount of noise in Figure 8-b, any *slice-by-slice* denoising procedure will fail to show the weak enhancements in the synthetic venous

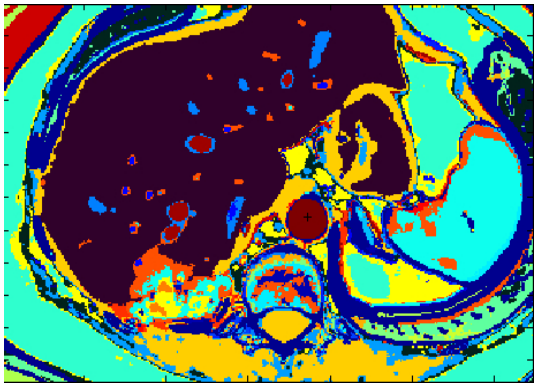

Figure 7: Clusterization result using the entire signal temporal dynamics including the specific baseline level of each tissue – colors are artificial. For tissues showing enough enhancement, the clusterization is similar to what is obtained in Fig.6 (see tumor tissues (2,7,8,9) for example). The other tissues (4) are here distinguished through their baseline.

systems (2 and 3) already visible in Figure 8-c and even clearer in Figure 8-d.

Using the same gray-scale, the result of our clusterization algorithm (see Section 4) applied to this sequence is presented at the top of Figure 8-e with the associated estimated enhancements inside each cluster shown at the bottom of the same figure. The estimated clusters and their associated estimated enhancements (centroids of each cluster) are shown and linked by capital letters. The enhancements are well recovered as shown by the comparison of the bottom part of Figures 8-a and 8-e. The clusterization of the synthetic tumor respects the circular geometry used to build the synthetic tumor in this sequence. This clusterization has to be understood as the construction of an adaptive piecewise constant map with unknown number of steps. The selected number of steps depends on the underlying enhancement, on the standard deviation of the noise and on the choice of $\alpha$, the level of the test procedure, and on the unknown localizations. In Figure 9, the fine dotted line, shows a possible stepwise profile which could be associated to this clusterization: each step represents the 4 estimated enhancements D to G along a diameter of the synthetic tumor.

The choice of a continuous synthetic tumor may seem surprising but it is driven by two ideas: on one hand continuous changes exist inside tumors, and

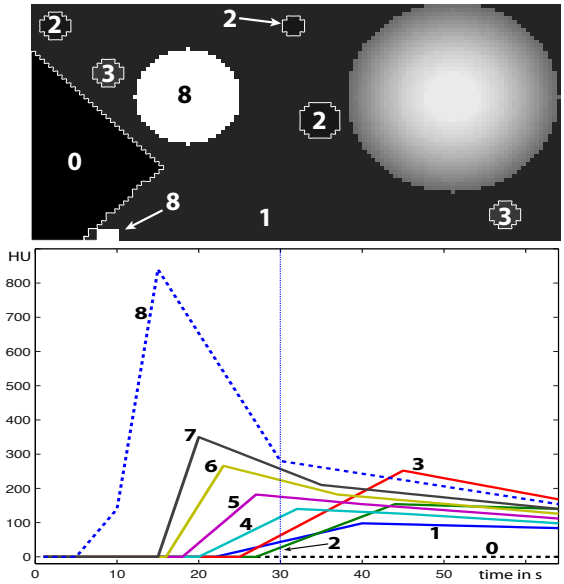


Figure 8-a: Artificial Simulated Dynamical Image : (bottom) The nine artificial enhancements used in our simulations. The dotted vertical line shows time 30: used for the presentation of the results of denoising and clusterization on these artificial data. The numbers link the used enhancements with their locations. (top) The dynamical image shown at time-slice 30. In the upper-right of the sub-figure, the large ball simulates a tumor with behaviors changing continuously from behavior 7 (in the middle) to 4 (on the border). The exact profile along a diameter of this tumor is given in Figure 9. The dark homogenous areas are surrounded by a fine white line to emphasize the borders. The gray scale comes from Figure 8-c

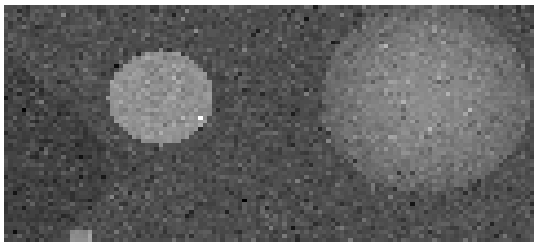

Figure 8-b: Simulated noisy dynamical image obtained by adding Laplace noise with standard deviation 41HU to the Artificial Dynamical Image see Figure 8-a. Dynamical image shown at time-slice 30

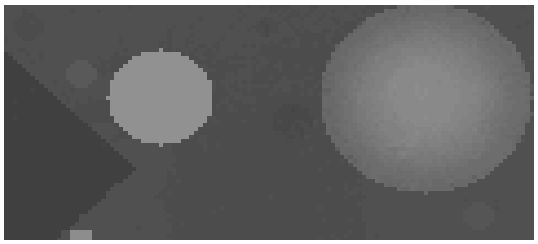

Figure 8-c: Denoised version of the artificial data using the same grayscale as the observed noisy data. Dynamical image shown at time-slice 30

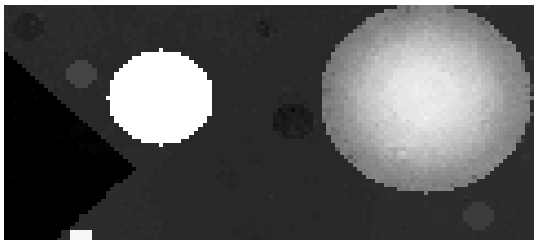

Figure 8-d: Denoised version of the artificial data using its full grayscale. Dynamical image shown at time-slice 30

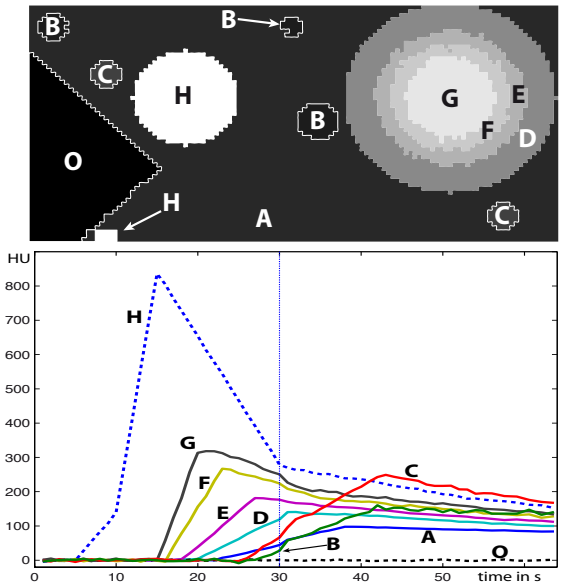


Figure 8-e: (top) Clusterization of the artificial data. Scale comes from Figure 8-c. The letters refer to the estimated enhancements (centers) of the clusters (see Figure 8-e). Dynamical image shown at time-slice 30. (bottom) Estimated enhancements (*center* of the clusters). The dotted vertical line shows time 30: used to present the results of denoising and clusterization on these artificial data.

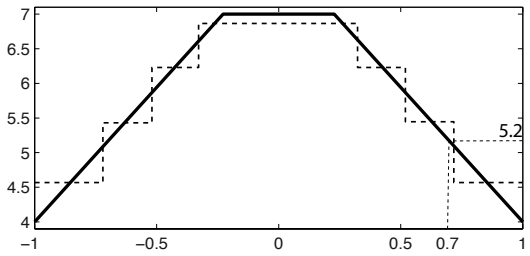


Figure 9: Representation of the profile inside the simulated tumor (thick line). The abscissa is the radius and the ordinate defines the enhancement mixture: a value 5.2 obtained at radius 0.7 represents a mixture of 20% of enhancement 5 combined with 80% of enhancement 6. The fine dotted line represents a feasible piecewise constant representation of this profile.

on the other hand this condition is more challenging for the algorithm. In addition, we have performed simulations (not presented here) with a synthetic tumor showing a piecewise constant profile or split into one continuous part and one piecewise constant part. In each case, the estimation and the clusterization showed the same quality of behavior. The estimation of each enhancement clearly improved as soon as the synthetic tumor exhibited some constant part in its profile.

## 6. Discussion

The goal of DCE-imaging is to create functional images (pictures) at voxel-level and not to deal with ROI which can mix several functional behaviors resulting in loss of localized information. For example, the heterogeneity of tumors is believed to provide important information which may be properly evaluated using DCE-CT. Let us point out that for ethical reasons, even if acquisition techniques in CT are expected to improve, it will remain necessary to control X-ray doses. Therefore images with poor signal to noise ratio, hardly usable to properly evaluate micro-circulation parameters, will still be produced. This is why we propose a two step procedure, based on the same multiple tests to compare random vectors, in order to improve the signal to noise ratio of the DCE-CT sequences. The first step enables to denoise the dynamical sequence voxel-by-voxel. The second step builds a spatial segmentation of the tissues based on the differentiation of the full dynamics.

The denoising procedure constructs around each voxel a neighborhood of (dynamically) homogeneous voxels with a size related to 1/ the size of the homogenous tissue it belongs to; 2/ the maximal gain in signal to noise ratio provided by the user. Through this step, at each voxel location, an individual denoised enhancement is obtained providing complete denoised information on the heterogeneity of the dynamics. The size of each neighborhood, which plays the role of a window or bandwidth, is chosen adaptively thus preventing under- or over-smoothing.

In addition to this local approach, the clustering step pulls together the homogeneous voxels (dynamics) in order to create a synthetic map (segmentation) of dynamical behaviors that sums up in a few dynamics all the information contained in the DCE-CT sequence. Although an over-reduction of the information in this clustering step could be feared, the amount of details found in the shape of the organs and the good level of heterogeneity in the tumor show the proper behavior of this technique. The clusterization provides an adaptive piecewise constant representation of the dynamical behaviors for which clear steps at the border of the organs and more continuous changes inside the tumor may be imagined. This is illustrated by the dynamics 2, 7, 8 and 9 of Figure 6, which sum up in a piecewise way the vascularity of the tumor from low vascularization to hot-spot.

In an image-by-image approach, an adaptive "bandwidth" selection aims at choosing the largest subset of pixels with statistically similar intensities in order to improve the signal to noise ratio. For DCE-CT, since each individual

voxel is too noisy, such a time-by-time pulling approach fails. Given a voxel $x$, we propose to find voxels with statistically similar enhancements: pulling those enhancements together provides a complete estimated enhancement at voxel $x$. Clearly, we avoid a (time-by-time) maximal approach and build potentially smaller (at some observation time) dynamical homogenous neighborhoods. Hence, around a voxel $x$, the selected voxels form a large subset with respect to the size of the homogenous tissue they belong to, and not necessarily large with respect to the gray-level obtained at a specific time.

Going further, it would be of interest to take advantage of the dynamical homogeneity of the tissues before the Radon reconstructions (of all acquisition times) by, for example, either comparing statistically partial sinograms or by looking for a sparse representation of the full dynamics. We do not have addressed this issue because usually scanners automatically provide one reconstructed image per acquisition time without giving access to the Radon reconstruction which works as a black-box inside. Our expectation is that the (black-box) Radon reconstruction has already benefited from larger local homogeneity inside each image. Applying a dynamical approach like ours directly during the Radon reconstruction in order to benefit from dynamical homogeneity remains a challenge.

During a long acquisition period as needed in DCE imaging, it is hard to avoid patient's movements. In Figures 6 and 7, we clearly see navy bands at the periphery of the organs (interface regions). They are characteristic, easy to identify and remarkably fine. Most movements should be removed by prospective or retrospective techniques like respiratory gating or registration. These different techniques are not conflicting and can even benefit one from the other: Registration techniques can be improved by the knowledge of these navy bands and a denoised dynamics, or part of it, could be the target of a registration algorithm. Denoising and clustering based on the comparison of entire dynamics are not strongly affected by small movements. This is promising for further developments using such a dynamical approach combined with techniques that deal with movements.

In order to make things clearer, we have used a homoscedastic setting with a known constant noise level $\sigma$ assuming independence in time and in space. This modelization is certainly not perfect: 1/ the noise level may depend on the intensity which varies not only among the tissues but also with time as the contrast agent goes through the tissue leading to contrast beam hardening artifacts; 2/ CT-artifacts are spatially not independent. Surprisingly, although our construction ignores CT-artifacts, they are properly removed without affecting the overall quality, when using a larger $\sigma$ than those prescribed directly by the estimation. This is an advantage of our approach that combines spatial locations, preserving time structures thanks to the use of statistically powerful multiple tests. The first issue is addressed by [34] which proposes a test of zero mean in random vector even when noise distribution presents strong departures from the simpler Gaussian homoscedastic case: noise distribution needs only to be symmetric with an unknown noise level which may even vary in time. As our construction is based on differences, this symmetric assumption is fulfilled

and ensures that the hardening artifacts will be taken into account. Even if the results we obtain are already satisfying, an industrial application will benefit from such an implementation.

In the simplified setting we have used, $\alpha$ and $\sigma$ play the role of tuning parameters. Using the baseline image, parameter $\sigma$ could be evaluated as we did or by taking a ROI outside the patient. Hence, the true remaining tuning parameter of our algorithm is $\alpha$. In the denoising step, at each voxel, it controls the level of the test used to choose the proper neighborhood and controls how smooth spacially the denoised dynamical image (meaning the full sequence) is. In the clustering step, it controls the number of clusters used to describe all the tissues. It plays a similar role as a penalty in the description of a function by a piecewise constant function. Decreasing $\alpha$ will produce fewer clusters preserving the easiest locations where the behaviors of the tissues are constant or slowly varying. In the case of DCE-CT, it will preserve large organs and strong details like the veinous system, aorta, bones, etc. Inside a heterogenous area like the tumor where the profile can vary slowly, the clusterization will show fewer behaviors slowly when $\alpha$ decreases but will keep a realistic piecewise constant description of the profile and hence of the tumor (see Figure 9). Changing the value of $\alpha$ in the clusterization will mostly affect only the description of the navy bands.

Our clustering procedure does not require the knowledge of the number of classes and a proper initialization like the $k$-means algorithms does [46, 38]. It also does not rely on assumptions to describe the behavior inside the classes like the Expectation Maximization algorithms does [39]. Neither does it create a synthetical representation of the enhancements based on linear combinations of a few artificial enhancements, eigenvectors of a functional PCA [47]. In our clusterization, the classes are built in order to and only to achieve a statistical homogeneity that offers a good and exhaustive description involving a few realistic enhancements.

We have implemented the denoising algorithm (see Section 3) in Matlab©. The processing of a 512×512×53 dynamical image as presented in Section 5 takes three hours on a bi-3 Ghz Quad-core Intel Xeon PowerMac as it uses a loop on the 512×512 voxels. As the processing at one location does not involve processing at other locations, this loop may be highly distributed with a cheap parallelized implementation in open-CL on a Graphic Processing Unit (GPU): after a transfer in the shared memory of the GPU of the global information of the DCE-CT (here the 512×512×53 numbers encoded on 8 bits), an industrial implementation will send to each chip a number of the 512×512 voxels inversely proportional to the number of chips on the GPU.

It is worth noting that even if we have presented our technic on 2D DCE-CT, it could be straightforwardly generalized to 3D DCE-CT sequences using 3D neighborhoods instead of 2D neighborhoods. In such case, the result will be a clusterization of the dynamical behaviors existing in the tissues of the 3-dimensionnal image.

## 7. Conclusion

Using a two step procedure, each step based on a statistical multiple hypothesis testing, we introduce a novel algorithm to denoise dynamical images, where each point of the picture is coding a complete time series. Based on the comparison of the dynamics, this algorithm preserves their full structure. It does not rely on any a priori knowledge of the features in the image and runs unsupervised up to one tuning parameter, which has a clear statistical interpretation as a significance level. The efficiency of our algorithm is shown on DCE-CT data used to follow the vascular and tissular distribution of a contrast agent. The quality of the denoised dynamical image is shown by the details which can be found by clinicians especially in small structures and allows a clear recognition of the tumor heterogeneity. As a byproduct, the tomographic artifacts are removed. Based on the denoising procedure, an efficient clustering algorithm is proposed: it relies neither on the knowledge of the number of classes nor on the distribution inside a class nor on the modelization of typical enhancements. The clustering algorithm runs with a statistical tuning parameter which acts like a penalization of the number of clusters. The result of the clusterization provides Regions Of Interest that are meaningful at a physiological level. They automatically sum up the typical dynamics of the tissue behavior for further analysis of the microcirculation. Artificial experiments validate these results on simulated data.

## Appendix

### *7.1. "Multiple testing for the comparison of random vectors"*

Given two spatial locations $x$ and $y$, we present the statistical test used to compare to the zero vector the enhancement difference vector with components

$$Z_k = I_y(t_k) - I_x(t_k), \quad k = 1 \dots K.$$

For the sake of simplicity, we write $Z = f + \sigma\varepsilon$ and we introduce these tests in their simplified version namely the Gaussian case with $\sigma$ known. Such a test of comparison to a zero vector is derived from the theoretical work of [32, 33] and [34] who consider general frameworks where $\sigma$ can be unknown and where $\varepsilon$ need not necessarily be Gaussian but at least symmetrical, which is ensured by the use of differences. We aim to test whether the mean vector $f$ is zero or not and hence consider the hypotheses

$$\mathcal{H}_0 : "f = 0" \text{ against } \mathcal{H}_1 : "f \neq 0".$$

Let us emphasize that for a prescribed level, due to their good properties of adaptation, these tests allow a good control of the second kind error which is important in our setting.

To simplify the presentation, let us suppose that $K$, the time number of the DCE-CT sequence, is of the form $K = 2^d$. We consider the regular dyadic

decomposition of the observation times $t_1 \dots t_{2^d}$. For $j = 0 \dots d-1$, we denote by $T_1^j \dots T_{2^j}^j$ the $2^j$ intervals with $2^{d-j}$ time indices

$$T_l^j = \{t_k, k = 2^{d-j}(l-1) \dots 2^{d-j} l\}, \quad l = 1, \dots, 2^j.$$

Given $j$ in $0 \dots d-1$, let us denote the projection of $Z$ onto the space generated by the vectors with same components on each time index $T_l^j$:

$$\Pi_j Z = (\underbrace{m_1^j, \dots, m_1^j}_{2^{d-j} \text{ times}}, \dots, \underbrace{m_{2^j}^j, \dots, m_{2^j}^j}_{2^{d-j} \text{ times}}),$$

where

$$m_l^j = \frac{1}{2^{d-j}} \sum_{t \in T_l^j} Z_t.$$

The test is based on the comparison to zero of the squared Euclidean norm $\|\Pi_j Z\|_K^2$ equal to

$$2^{d-j} \sum_{l=1}^{2^j} \left(m_l^j\right)^2 = \frac{1}{2^{d-j}} \sum_{l=1}^{2^j} \left( \sum_{t \in T_l^j} Z_t \right)^2.$$

Under $\mathcal{H}_0$, the difference vector $Z = I_y - I_x$ is a centered Gaussian vector with covariance matrix $2\sigma^2 Id_K$ where $Id_K$ denotes the identity matrix in $\mathbb{R}^K$. Hence, under $\mathcal{H}_0$, $\|\Pi_j Z\|_K^2 / \sigma^2$ follows a $\chi^2$-distribution with $2^j$ degrees of freedom. Our test procedure works as follows:

$$\text{Reject } \mathcal{H}_0 \text{ at level } \alpha \text{ if for any } j = 0 \dots d-1,$$

$$\|\Pi_j Z\|_K^2 / 2\sigma^2 > \Psi_{2^j}^{-1}(\alpha/d),$$

where $\Psi_D^{-1}$ denotes the quantile function of a $\chi^2(D)$-distribution.

To simplify the presentation, we have used a Bonferroni correction to ensure that this multiple testing procedure is of level $\alpha$. Some cleverer corrections could be applied as proposed in [33] and [34] or, as we do in practice (see Section 5), by using an FDR approach [40].

From a clinical point of view, it is clear that the unobservable true enhancements $i_x$ and $i_y$ are functions of time. So does the function $F$ defined by $F(t) := i_x(t) - i_y(t)$. Moreover, from a clinical point of view, $F$ is smooth (regular). We have $f_k = F(t_k)$ and in this setting, these multi-tests are adaptive with respect to the unknown Hölder regularity $s$ (see [37]) of the function $F$ and, for a given fixed power, this test automatically achieves the best rate of testing $\rho_s(K)$ for all regularities $s > 1/4$ (see [33, Thm 1], [34, Thm 4 and 5] and [41]).

We now have the tool to compare enhancement sequences at two voxel locations $x$ and $y$ with respect to the known noise level $\sigma$.

We write "$I_y \equiv_{\sigma^2}^{\alpha} I_x$" when the enhancement difference average vector $i_x - i_y$

is accepted to be the zero vector at level $\alpha$ with respect to a noise level $\sigma^2$ following the above construction. This defines precisely the "statistical closeness" involved in Section 3.

*7.2. Growing Time Homogeneous Neighborhood*

We now present the construction of the neighborhood $\mathcal{V}_x$ for a fixed spatial location $x$.

Given an estimation (or denoising) procedure of the enhancement using a set of locations $V$, for example the empirical mean defined by

$$\hat{I}_V = \frac{1}{|V|} \sum_{y \in V} I_y \tag{3}$$

or a generalized median defined, for example, by

$$\hat{I}_V = \arg \min_{J \in \mathbb{R}^K} \sum_{y \in V} \sum_{i=1}^{K} |I_y(t_i) - J(t_i)|. \tag{4}$$

Given a maximal number of iterations $K$ and a increasing sequence of integers $n_1, \ldots, n_K$ with a geometrical growth.

We consider in a first step the set

$$\mathcal{W}_x = \{y \in \mathcal{X} \text{ such that } y \neq x \text{ and } I_y \equiv^{\alpha}_{\sigma^2} I_x\} \tag{5}$$

of the spatial locations for which enhancements are statistically similar to those of $x$ with respect to the multi-test introduced in Appendix 7.1 and we set $V_0 = \emptyset$ and $i = 0$.

Then, setting $i = i + 1$, we sequentially grow rings, denoted $W_i$, around $x$ and set the neighborhood $V_i = V_{i-1} \cup W_i$. The ring $W_i$ is made of the $n_i$ closest points in $\mathcal{W}_x$ not in $V_{i-1}$. At each step, $W_i$ is tested for statistical coherence, defined below, with previously built neighborhoods $V_j$, $j < i$. When the statistical coherence of $W_{i+1}$ is refused or when $i = K$, the algorithm stops and returns $\mathcal{V}_x = V_i$ as the selected neighborhood and $\hat{I}_x = \hat{I}_{V_i}$ as the denoised enhancement.

The skeleton of our construction is given by the flowchart in Figure 2 and Algorithm 1 in Appendix 7.3.

This method is illustrated in Figure 10. It shows a zoom on the axial upper abdominal section focused on the right posterior part of the sequence presented in Section 5: a vertebra (1), the aorta (2), the liver (3) and a tumor (4). This figure shows for a specific voxel $x$ (white dot designed by the white arrow) inside the tumor the first four rings $W_i$, $i = 1 \ldots 4$ (in order : red, yellow, green and blue). Due to the pre-selection of voxels in $\mathcal{W}_x$, the rings are neither convex nor connected and follow the heterogeneity of the tumor.

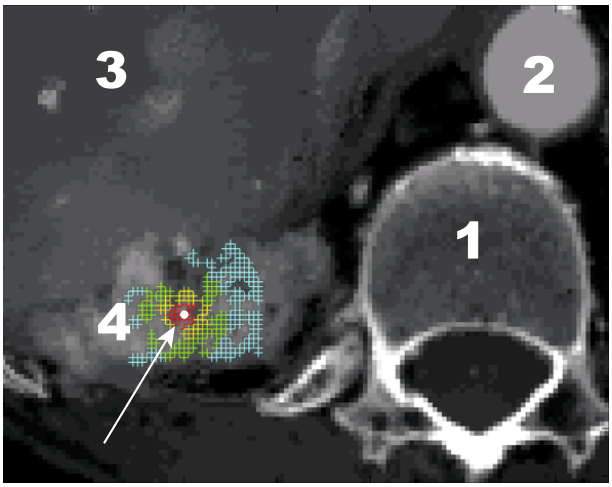


Figure 10: First four rings $W_i$ (in order : red, yellow, green and blue) selected around the selected voxel $x$ specified by the white dot and the white arrow.

During the iterations, the statistical coherence between $W_i$ and the previously build neighborhood $V_j$, $j < i$, is ensured by the test defined by

$$\hat{J}_i \equiv^{\alpha/i}_{\sigma^2\rho(|V_j|,|W_i|)} \hat{I}_j, \quad j = 1 \dots i-1,$$

where $\hat{J}_i = \hat{I}_{W_i}$ and $\hat{I}_j = \hat{I}_{V_j}$. This is a generalization of the multi-test introduced in Appendix 7.1 which compares $\hat{J}_i$, the estimate on the ring $W_i$, with all estimates $\hat{I}_1, \dots, \hat{I}_i$ built on the previous nested neighborhoods $V_1$, ..., $V_i$. This multiple testing procedure compares the statistical hypothesis

$$\mathcal{H}_0 : \text{"}\mathbb{E}(\hat{J}_i - \hat{I}_j) = 0, \text{ for all } j = 1 \dots i\text{"},$$

against

$$\mathcal{H}_1 : \text{"}\mathbb{E}(\hat{J}_i - \hat{I}_j) \neq 0, \text{ for at least one } j\text{"},$$

where $\mathbb{E}Z$ denotes the expectation of a vector $Z$.

The use of $\alpha/i$ is a classical Bonferroni correction to ensure that the level of this test is $\alpha$. Cleverer corrections are possible, see Appendix 7.1. The correction in the noise level, defined by $\rho(|V_j|, |W_i|)$, aims at taking into account the fact that the estimates come from independent samples with respective sizes $|V_j|$ and $|W_i|$. This correction depends on the choice of the estimates: for example, if the estimate on a set of locations derives from an empirical mean defined in (3) we set

$$\rho(|V_j|, |W_i|) = |V_j|^{-1} + |W_i|^{-1},$$

which derives from the mean equality test for two Gaussian samples with same known variance (see Wald test or Likelihood test in [42]). The sizes $n_i$ of the

rings $W_i$ are chosen to ensure that the sizes of the $V_i$ grow at least geometrically and that the number of tests is of logarithmic order.

*Remark 1:* The pre-selection step specified by the set $\mathcal{W}_x$, which is neither necessarily convex nor connected, allows the denoising construction to obtain "neighborhoods" built from different objects of the same type over long distances. This is useful in medical images where the same kind of tissue may reappear in different areas.
*Remark 2:* This method is closely related to Lepski's method (see e.g. [43] and [44]). This method relies on comparing directly $\hat{I}_{i+1}$ with the previously built estimates $\hat{I}_0 \dots \hat{I}_i$. Because they are built on nested subsets, these estimates are not independent. This lack of independence is the drawback of this method from a practical point of view : the growth of neighborhoods often stops too late, leading to over-smoothing [36] and in the DCE-CT case to mixing dynamics. This drawback is corrected by the use of rings.
*Remark 3:* The neighborhood comparison protects from extra bias which could appear in a one step procedure using a subset of $\mathcal{W}_x$ without more control (see [36]). The sizes grow geometrically to ensure that the number of tests remains logarithmic with respect to the number of used voxels, allowing a good control of the power of this multi-test procedure.
*Remark 4:* In the second step of the algorithm, if the number of locations in $\mathcal{W}_x \setminus V_i$ (set of voxels in $\mathcal{W}_x$, but not in $V_i$) is too small, the algorithm ends and returns the last estimate.
*Remark 5:* The maximal size of $V_i$ which controls the gain in the signal to noise ratio is fixed for computational time reasons.
*Remark 6:* In the case of a generalized median defined in (4), to take into account that (i) for Laplace-distributed noise the asymptotic for the median differs from that of the mean by a factor 2 in variance (see [45] or [42, cor. 21.5]); (ii) median and mean are the same when sets contain only one element, we propose the quickly computable approximation

$$\rho(|V_j|, |W_i|) = \frac{1}{4}\left(\frac{1 + 7/|V_j|}{|V_j|} + \frac{1 + 7/|W_i|}{|W_i|}\right).$$

*7.3. Algorithms*

In this section, we describe synthetically our denoising and clustering algorithms. Given two lists $A$ and $B$, the notations $[A; B]$, $A \cup B$ and $A \setminus B$ denote the concatenation without deletion, the union and the complement of B relative to A, respectively.

*Denoising algorithm*

**Input**: $x \in \mathcal{X}$
**Output**: $\mathcal{V}_x$ and $\hat{I}_x$
**Initialization :** `// define first neighborhood`
$i := 0$ ; $V_0 := \emptyset$ ; $W_0 := \{x\}$ ; $accepted := 1$
`// Main loop : find new neighbors and test statistical closeness`
**while** *accepted* **do**
  `// build new neighborhood`
  $i := i + 1$ ; $V_i := V_{i-1} \cup W_{i-1}$
  `// estimate build on the new neighborhood`
  Compute $\hat{I}_i$ using locations in $V_i$
  `// find a "ring" around new neighborhood`
  Find $W_i$ the subset of the $n_i$ closest points to $V_i$ in $\mathcal{W}_x \setminus V_i$.
  `// estimate build on the ring`
  Compute $\hat{J}_i$ the estimated enhancement using locations in the ring $W_i$.
  `// test closeness of ring estimate with previous estimates`
  $level := \alpha / i$
  **for** $j := 1 \dots i$ **do**
    $var := \sigma^2 \rho(|V_j|, |W_i|)$
    **if** $not(\hat{J}_i \equiv_{var}^{level} \hat{I}_j)$ **then** accepted := 0
    *(See Section 7.2 for precisions on $\rho$ and these tests.)*
**return** *( $\mathcal{V}_x := V_i$ ; $\hat{I}_x := \hat{I}_i$ )*

**Algorithm 1**: Spatially pointwise denoising algorithm

The Algorithm 1 is summarized by the flowchart given by Figure 2.

*Clustering algorithm*

**Input**: $\{(\hat{I}_x; \mathcal{V}_x), x \in \mathcal{L}\}$ ; $v_0$
**Output**: $\mathcal{C}$ and $\mathcal{I}$ `// cluster list and cluster centers`
**Initialization:** set $\mathcal{C} := \emptyset$ and $\mathcal{J} := \emptyset$ ; lastchange:=0;
**while** $\mathcal{L} \neq \emptyset$ **do**
  `// build a new cluster`
  **if** lastchange=0 **then**
    $x := \arg\max_{y \in \mathcal{L}} |\mathcal{V}_y|$
    $c :=$ Children($x$)
    lastchange:=length(C)+1
  **else**
    $c := \mathcal{C}$(lastchange)
  $(\hat{\mathbf{J}}, c_{\hat{\mathbf{J}}}) :=$ RobustKmeans($c$)
  $\mathcal{L} := (\mathcal{L} \cup c) \setminus c_{\hat{\mathbf{J}}}$
  $\mathcal{C}$(lastchange) $:= c_{\hat{\mathbf{J}}}$
  $\mathcal{J}$(lastchange) $:= \hat{\mathbf{J}}$
  lastchange:=CheckClusterList(lastchange)

**Algorithm 2**: Clustering algorithm

**function CheckClusterList($j$)**
  **Input**: $j$ cluster to check for merging
  **Output**: 0 if not merged else merged cluster number
  `// check if` $\mathcal{C}(j)$ `need to be merged`
  **for** $i := [1..j-1] \cup [j+1..\text{length}(\mathcal{C})]$ **do**
    **if** $\hat{\mathbf{I}}_{\mathcal{C}(i)} \equiv^{\alpha}_{\sigma^2 \rho(1, \min(|\mathcal{C}(i)|, |\mathcal{C}(j)|))} \hat{\mathbf{I}}_{\mathcal{C}(i)}$ **then**
      $\mathcal{C}(i) := \mathcal{C}(i) \cup \mathcal{C}(j)$;
      $\mathcal{C} := \mathcal{C}([1 : j-1, j+1 : \text{length}(\mathcal{C})])$;
      break and return(i);
    **end**
  **end**
  return(0);

**function Children($x$)**
  **Input**: $x$ a voxel
  **Output**: $N$ the children of $x$ for the relation $\preceq$
  **Initialization:** $N := \mathcal{V}_x$ ; $i := 0$;
  **while** $i <$ length($N$) **do**
    { $i := i + 1$ ; $N := [N; (\mathcal{V}_{N(i)} \setminus N) \cap \mathcal{L}]$ };

**Algorithm 3**: Functions CheckClusterList and Children

**Acknowledgment**

This work was supported in 2008 by a grant "Bonus Qualité Recherche" from the University Paris Descartes for the project "Cancer Angiogénèse et Outils

Mathématiques".

## References


[1] Goh, V., A. R. Padhani, et al., *Functional imaging of colorectal cancer angiogenesis*, Lancet Oncol., **8**-3, pp. 245-55, 2007.

[2] Goh V., Halligan S., Hugill J.A., Gartner L., Bartram C.I., *Quantitative colorectal cancer perfusion measurement using dynamic contrast-enhanced multidetector-row computed tomography: effect of acquisition time and implications for protocols*, J. Comput. Assist. Tomogr., **29**, pp. 59-63, 2005.

[3] Cuenod C.A., Fournier L., Balvay D., Guinebretière J.M., *Tumor angiogenesis: pathophysiology and implications for contrast-enhanced MRI and CT assessment*, Abdom Imaging., **31**-2, pp. 188-93, 2006.

[4] Miles, K. A., *Functional CT imaging in oncology*, Eur. Radiol., **13**-suppl. 5, M134-8, 2003.

[5] Padhani, A. R., C. J. Harvey, *et al.*, *Angiogenesis imaging in the management of prostate cancer*, Nat. Clin. Pract. Urol., **2**-12, pp. 596-607, 2005.

[6] Bisdas S., Konstantinou G.-N., Lee P.-S., Thng C.-H., Wagenblast J., Baghi M. and Koh T.-S., *Dynamic contrast-enhanced CT of head and neck tumors: perfusion measurements using a distributed-parameter tracer kinetic model. Initial results and comparison with deconvolution-based analysis.*, Physics in medicine and biology, **52**-20, pp. 6181-96, 2007.

[7] Tofts P.-S., *Modelling tracer kinetics in dynamic Gd-DTPA MR imaging*, J Magn. Reson. Imaging, **7**-1, pp. 91-101, 1997.

[8] Brix G, Kiessling F, Lucht R, Darai S, Wasser K, Delorme S, Griebel J., *Microcirculation and microvasculature in breast tumors: pharmacokinetic analysis of dynamic MR image series*, Magn Reson Med., **52**-2, pp. 420-9, 2004.

[9] de Bazelaire C., Siauve N., Fournier L., Frouin F., Robert P., Clement O., de Kerviler E., Cuenod C.A., *Comprehensive model for simultaneous MRI determination of perfusion and permeability using a blood-pool agent in rats rhabdomyosarcoma*, Eur Radiol, **15**, pp. 2497-505, 2005.

[10] Brochot C., Bessoud B., Balvay D., Cuénod C.-A., Siauve N., Bois F.-Y., *Evaluation of antiangionenic treatment effects on tumors' microcirculation by Bayesian physiological pharmacokinetic modelling and magnetic resonance imaging.*, J Magn. Reson. Imaging, **24**, pp. 1059-67, 2006.

[11] Krishnamurthi G., Stantz K.-M., Steinmetz R., Gattone V.-H., Minsong Cao ; Hutchins G.-D. ; Yun Liang, *Functional imaging in small animals using X-ray computed tomography- : study of physiologic measurement reproducibility*, IEEE Trans. on Medical Imaging, **24**-7, pp. 832-43, 2005.

[12] Jain R., Scarpace L., Ellika S., Schultz L.R., Rock J.P., Rosenblum M.L., Patel S.C., Lee T.Y., Mikkelsen T., *First-pass perfusion computed tomography: initial experience in differentiating recurrent brain tumors from radiation effects and radiation necrosis*, Neurosurgery, **61**, pp. 778-86; with discussion, 2007.

[13] Fournier L., Thiam R., Cuénod C.-A., Medioni J., Trinquart L., Balvay D., Banu E., Balcaceres J., Frija G. and Oudard S., *Dynamic contrast-enhanced CT (DCE-CT) as an early biomarker of response in metastatic renal cell carcinoma (mRCC) under anti-angiogenic treatment.*, J. of Clinical Oncology - ASCO Annual Meeting Proceedings (Post-Meeting Edition), **25**, 2007.

[14] Rosen M.A., Schnall M.D., *Dynamic contrast-enhanced magnetic resonance imaging for assessing tumor vascularity and vascular effects of targeted therapies in renal cell carcinoma*, Clin Cancer Res., **13-2**, pp. 770-6, 2007.

[15] Zhu A.X., Holalkere N.S., Muzikansky A., Horgan K., Sahani D.V., *Early antiangiogenic activity of bevacizumab evaluated by computed tomography perfusion scan in patients with advanced hepatocellular carcinoma*, Oncologist, **13**, pp. 120-5, 2008.

[16] Sorensen AG, Tievsky AL, Ostergaard L, Weisskoff RM, Rosen BR., *Contrast agents in functional MR imaging*, J Magn Reson Imaging, **7**-1, pp. 47-55, 1997.

[17] Kiessling F., Greschus S., Lichy M.P., Bock M., Fink C., Vosseler S., Moll J., Mueller M.M., Fusenig N.E., Traupe H., Semmler W., *Volumetric computed tomography (VCT): a new technology for noninvasive, high-resolution monitoring of tumor angiogenesis*, Nat. Med., **10**, pp. 1133-8, 2004.

[18] Cao M.,Liang Y., Shen C., Miller K.D., Stantz K.M., *Developing DCE-CT to Quantify Intra-Tumor Heterogeneity in Breast Tumors With Differing Angiogenic Phenotype*, IEEE Trans. on Medical Imaging, to appear in 2009.

[19] Buades A., Coll B., Morel J.-M., *Nonlocal Image and Movie Denoising* Int. J. of Computer Vision, **76**-2, pp. 123-39, 2007.

[20] Barbier EL, Lamalle L, Décorps M, *Methodology of brain perfusion imaging*, J Magn Reson Imaging, **13**-4, pp. 496-520, 2001.

[21] Collins, D.J.; Padhani, A.R., *Dynamic magnetic resonance imaging of tumor perfusion*, IEEE Engineering in Medicine and Biology Magazine, **23**-5, pp. 65 - 83, 2004.

[22] Balvay D, Troprès I, Billet R, Joubert A, Péoc'h M, Cuenod CA, Le Duc G., *Mapping the zonal organization of tumor perfusion and permeability in a rat glioma model by using dynamic contrast-enhanced synchrotron radiation CT.*, Radiology, **250**-3, pp. 692-702, 2009.

[23] Thibault J.-B., Bouman C.A., Sauer K.D., and Hsieh J., *A Recursive Filter for Noise Reduction in Statistical Iterative Tomographic Imaging*, in Computational Imaging IV, edited by Charles A. Bouman, Eric L. Miller, Ilya Pollak, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 6065, 2006.

[24] Donoho, David L., *Compressed sensing*, IEEE Trans. Inform. Theory, **52**-4, pp. 1289–306, 2006.

[25] Candes E.J., Tao T., Near-optimal signal recovery from random projections: universal encoding strategies?, IEEE Trans. Inform. Theory, **52**-12, pp. 5406–5425, 2006.

[26] Chen G.H., Tang J., Leng S., *Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets*, Med Phys., **35**-2, pp. 660-3, 2008 Feb.

[27] Mistretta C.A., *Undersampled Radial MR Acquisition and Highly Constrained Back Projection (HYPR) Reconstruction: Potential Medical Imaging Applications in the Post-Nyquist Era*, Journal of Magnetic Resonance Imaging, **29**-3, pp. 501-16, 2009.

[28] Happonen A.P., Ruotsalainen U., *A Comparative Study of Angular Extrapolation in Sinogram and Stackgram Domains for Limited Angle Tomography*, in Lecture Notes in Computer Science **3540**, Springer Berlin / Heidelberg, pp. 1047-56, 2005.

[29] Wang J. , Li T., Lu H. and Liang Z., *Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose X-ray computed tomography*, IEEE Trans. on Medical Imaging, **25**, pp. 1272-83, 2006.

[30] Chen Y., Cheng L., Fang T., Raupach R., *A Multi-Image Restoration Method for Image Reconstruction from Projections*, proc. of IEEE 11th International Conference on Computer Vision, pp. 1-8, 2007.

[31] Axel L., *Cerebral blood flow determination by rapid-sequence computed tomography: Theoretical analysis*, Radiology, **137**-3, pp. 679-86, 1980.

[32] Spokoiny V. G., *Adaptive hypothesis testing using wavelets*, Ann. Statist., **24**-6, pp. 2477-98, 1996.

[33] Baraud, Y., Huet, S. and Laurent, B., *Adaptive tests of linear hypotheses by model selection*, Ann. Statist., **31**-1, pp. 225–51, 2003.

[34] Durot C. and Rozenholc Y., *An adaptive test for zero mean*, Math. Methods Statist., **15**-1, pp. 26-60, 2006.

[35] Gravel P., Beaudoin G. and de Guise J. A., *A method for modeling noise in medical images*, IEEE Trans. Med. Imaging, **23**-10, pp. 1221-32, 2004.

[36] Reiß, M., Rozenholc, Y. and Cuenod, C.-A., *Pointwise adaptive estimation for robust and quantile regression*, Math arXiv:0904.0543, `http://arxiv.org/abs/0904.0543`

[37] Wasserman, L., *All of nonparametric statistics*, Springer Texts in Statistics, Springer, New York, 2006.

[38] Jain A.K. and Dubes R.C, *Algorithms for Clustering Data*, Prentice Hall, Englewood Cliffs, NJ, 1988.

[39] Dempster A., Laird N., and Rubin D., *Maximum likelihood from incomplete data via the EM algorithm*, J. of the Royal Statistical Society, Series B, **39**-1, pp. 1–38, 1977.

[40] Benjamini Y., Hochberg Y., *Controlling the False Discovery Rate: a practical and powerful approach to multiple testing*, J. Royal Stat. Soc. Series B, **57**-1, pp. 289-300, 1995.

[41] Gayraud G. and Pouet C., *Adaptive minimax testing in the discrete regression scheme*, Probab. Theory Rel. Fields., **133**-1, pp. 531-58, 2005.

[42] van der Vaart, A., *Asymptotic Statistics*, Cambridge University Press, 1998.

[43] Lepskii 0.V., *Asymptotically minimax adaptive estimation I: Upper bounds. Optimally adaptive estimates.*, Theory Probab. Appl., **36**, pp. 682-97, 1991.

[44] Lepskii 0.V. and Spokoiny V.G., *Optimal pointwise adaptive methods in nonparametric estimation*, Ann. Statist., **25**-6, pp. 2512-46, 1997.

[45] Ferguson T., *Asymptotic Joint Distribution of Sample Mean and a Sample Quantile*, unpublished: http://www.math.ucla.edu/∼tom/papers/unpublished/meanmed.pdf, 5 pages,

[46] Hartigan J. A., *Clustering algorithms*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, New York-London-Sydney, 1975.

[47] Ramsay J. O. and Silverman B. W., *Functional data analysis*, Springer Series in Statistics, Springer, New York, 2005.