Abstract—In this paper, we extend the gradient vector flow field for robust variational segmentation of vector-valued images. Rather than using scalar edge information, we define a vectorial edge map derived from a weighted local structure tensor of the image that enables the diffusion of the gradient vectors in accurate directions through the 4D gradient vector flow equation. To reduce the contribution of noise in the structure tensor, image channels are weighted according to a blind estimator of contrast. The method is applied to biological volume delineation in dynamic PET imaging, and validated on realistic Monte Carlo simulations of numerical phantoms as well as on real images.

Index Terms—Deformable models, dynamic PET, gradient vector flow, structure tensor.

I. INTRODUCTION

Vector-valued images occur in contexts such as color images, images of the same field of view acquired at different wavelengths (i.e. multi- and hyper-spectral images), medical images acquired at different time intervals (e.g. dynamic PET or functional MR images), with different equipments (e.g. co-registered PET/CT images), different modes (e.g. multi-spin echo MR images) or textured images. The accurate segmentation of these images requires appropriate methods for exploiting additional information provided by the extra dimension [1]–[5].

Deformable models such as snakes, as originally proposed by Kass and Witkin [6] and active surfaces in 3D have become very popular in image segmentation, including for medical applications [7]. Active surface models attempt to recover the region of interest by conforming an evolving surface to the boundary of the object. The evolution of the surface is derived through a variational formulation of an energy functional which can be seen, when reaching equilibrium, as a force-balance relation between forces acting on the surface: internal forces, which control the smoothness of the model, and external forces, derived from image information. Such external forces can either use local information like edges [6], [8]–[10], global information based on region statistics [11], [12], hybrid approaches based on both local and global information [13]–[16] or based on edge pixels interactions [17].

Implicit representations of deformable models such as the geometric active contour model are able to handle topological changes through the level-set paradigm [18], [19]. This topological flexibility can however constitute a drawback when a single object has to be segmented, as it can be the case in medical image segmentation. In such cases, additional topology-preserving procedures must be implemented [20], [21]. Another drawback of implicit representations is the increased computational load, which can be prohibitive in 3D clinical imaging. On the other hand, parametric deformable models are particularly appropriate for single object delineation because of their inherent ability to preserve the topology of the initial model. In this study, we focus on parametric edge-based deformable models.

Efforts have been made to overcome the original limitations of parametric snakes, mainly through the derivation of new expressions for the external force field that guides the model toward the boundaries of the object [8]–[10], [22]–[24]. Among external force fields, the Generalized Gradient Vector Flow (GGVF) field [8], [9], and more recently the Vector Field Convolution (VFC) field [10] aroused great interest because of their reduced sensitivity to noise and their ability to progress into highly concave regions in the image, the latter property being due to their nonconservative nature. However, the efficiency of both GGVF and VFC force fields critically relies on the choice of an accurate scalar edge map, usually based on the spatial derivatives of the input image, e.g. a Canny filter [25]. Refinements of these methods have been proposed to tackle remaining issues such as sensitivity to initialization [26], [27], capture range and ability to move into long and thin concavities [28].

Aside from these improvements, deformable models tailored for vector-valued images either using edge-based or region-based approaches have been proposed [3], [14], [29]–[31]. In the pioneering work of Di Zenzo [32], the image is considered as a vector field whose dimension is the number of channels in the image. Edge detection is associated with the gradient of the vector-valued image, or...
vector gradient, derived from the norm of a local structure tensor (LST) that integrates the different gradient contributions to locate real edges, or vector edges more precisely. Structure tensors, also known as second moment matrices, can estimate magnitudes and directions of oriented structures like edges at a local scale in vector-valued images. They have been widely studied, especially in the field of image restoration [5], [33]–[35]. In particular, Tschumperlé and Deriche have devised a generic PDE (partial differential equations)-based formulation for the regularization of vector-valued images that exploits both amplitudes and directions of the vector gradient computed from local structure analysis. In image segmentation, LST have been studied first by Sapiro [2], [29] for geometric active contours, who set the edge-stopping term of the level set as a function of the norm of the LST. This approach has been also used by Xie and Mirmehdi [14]. Goldenberg et al. used an alternative metric tensor based on the Beltrami framework, where the color image is considered as a 2D surface living in a five-dimensional space [30]. Zhang et al. exploited the LST formalism for seeded segmentation based on anisotropic diffusion [36]. In the case of parametric active contours, the gradient magnitude in the Luo color space was incorporated in the GVF framework by Yang et al. [31] under the name Color GVF.

While the use of directions and magnitude derived from local geometry analysis is well established for image regularization using anisotropic diffusion, to our knowledge, there is no deformable model approach that also takes profit of the directional information carried by the LST for image segmentation. Current deformable models generally identify vector edges by a simple scalar value, i.e. the norm of the LST. Another drawback of existing methods is that the different channels may improve their localization. We introduce a new external force field. The validation setup used for the experiments is described in section IV. Results on synthetic images are presented and discussed in section V, followed by results on PET images in section VI. Finally, a conclusion is drawn in Section VII.

II. Background

In this section, we briefly describe the mathematical background of parametric active surfaces and of the GGVF of the literature.

A. Active Surface Model

In the continuous domain, a parametric active surface \( S^t \) at time \( t \) of its deformation is represented as a mapping of a bivariate parameter \((m,n)\) on a regular grid \( \Omega \), superimposed on the spatial image domain:

\[
S^t : \Omega = [0, 1] \times [0, 1] \rightarrow \mathbb{R}^3.
\]

More explicitly,

\[
(m, n) \mapsto S^t(m, n) = [x_1(m, n), x_2(m, n), x_3(m, n)]^T.
\]

\( S^t \) deforms under the influence of forces, iteratively minimizing the following energy functional:

\[
E(S^t) = \int_{m,n} \left[ E_{\text{int}}(S^t) + E_{\text{ext}}(S^t) \right] dmdn,
\]

where \( E_{\text{int}} \) is the internal energy and \( E_{\text{ext}} \), the external energy of the surface. The internal energy term imposes smoothness constraints on the surface, while the external energy term drives it toward the object of interest. Expanding the internal energy term, where the functionals of first and second order define smoothness measures, the energy can be expressed as:

\[
E(S^t) = \int_{m,n} \left[ \alpha \left( ||S^t_m||^2 + ||S^t_n||^2 \right) + \beta \left( ||S^t_{mn}||^2 + ||S^t_{nm}||^2 \right) + E_{\text{ext}}(S^t) \right] dmdn,
\]

where coefficients \( \alpha \) and \( \beta \) define the weighting of internal energies acting on the surface [37]. Elasticity terms are weighted by \( \alpha \), rigidity terms and resistance to twist by \( \beta \). Equation (3) can be seen as a force balance equation. At equilibrium, i.e. at minimum energy, one obtain the Euler-Lagrange equation:

\[
\alpha \Delta S^t - \beta \Delta^2 S^t - \nabla E_{\text{ext}} = 0,
\]
where $\nabla$ is the gradient operator and $\Delta$ is the Laplace operator. The first two terms define internal forces acting on the surface:

$$ F_{int}(S') = \alpha \Delta S' - \beta \Delta^2 S'. \quad (5) $$

The third term defines external forces derived from edge information:

$$ F_{ext}(S') = -\nabla E_{ext}, \quad (6) $$

so that an equilibrium is reached at the boundaries of the object.

### B. Generalized Gradient Vector Flow

The generalized gradient vector flow (GGVF) field [9] is the vector field $F_{ext}$ defined as the steady-state solution of the following vector partial differential equations:

$$ \frac{\partial F_{ext}}{\partial t} = g(|\nabla f|) \nabla^2 F_{ext} - h(|\nabla f|)(F_{ext} - \nabla f), \quad (7) $$

where each spatial component of $F_{ext}$ is solved independently. $\nabla^2$ is the vector Laplace operator, $f$ is an edge map derived from the image typically ranging from 0 to 1 and having strong values at the edges. $g = e^{-\frac{|\nabla f|^2}{\kappa}}$ and $h = 1 - g$ are two functions that control the trade-off between the first and second terms through parameter $\kappa$. The first term favors the isotropic diffusion of $F_{ext}$ where $|\nabla f|$ has low values, e.g. in homogeneous regions. The second term tends to conform the field to $-\nabla f$ in regions of strong gradients. The resulting vector field shares the desirable properties of providing a large capture range and of allowing the models to progress into narrow concavities.

### III. Method

In this section, we present the 4DGGVF external force field, a generalization of GGVF fields to vector-valued images.

#### A. Weighted Structure Tensor of a 3D Vector-Valued Image

In the continuous domain, we denote by $I$ a 3D vector-valued image consisting of $M$ channels:

$$ I(x, c) : (\Omega_x \otimes \Omega_c) \subset \mathbb{R}^3 \otimes \mathbb{N} \to \mathbb{R}, $$

where $\Omega_x$ is the 3D spatial domain of the image and $\Omega_c$ the channel dimension. $x = (x_1, x_2, x_3) \in \Omega_x$ is the spatial position of the voxel. We denote by $I_k = I(x, k)$ the $k^{th}$ channel of the image.

A natural generalization of the notion of gradient to vector-valued images consists in averaging the different gradient contributions in individual channels. However, this approach is generally not satisfying, for example when opposite contributions cancel out. To avoid this drawback, Di Zenzo proposed a geometrical approach in which a 2D color image is considered as a 2D $\rightarrow$ 3D vector field [32]. The gradient is then identified as the direction that maximizes the quadratic form of the total differential $dI$ of $I$. Further developments of Di Zenzo’s approach have expressed this maximization problem in tensor notation [38], by introducing the structure tensor of a vector-valued image. In the 3D case, $I$ is a 3D $\rightarrow$ MD vector field and its total differential is expressed as:

$$ dI = \frac{\partial I}{\partial x_1} dx_1 + \frac{\partial I}{\partial x_2} dx_2 + \frac{\partial I}{\partial x_3} dx_3. \quad (8) $$

A weighted quadratic form for $dI$, or first fundamental form, is expressed as:

$$ \|dI\|^2_{\omega} = dx^T G_\omega dx, \quad (9) $$

with $G_\omega$ a regularized, weighted structure tensor of the image:

$$ G_\omega = K_\sigma * \sum_{k=1}^{M} \omega_k (\nabla I_k \otimes \nabla I_k^T), \quad (10) $$

where $K_\sigma$ is a Gaussian kernel of scale $\sigma$, $*$ is the convolution operation, $\otimes$ is the tensor product, and $\omega_k$ is a weighting factor for channel $I_k$. The weighting factors $\omega_k$, $k \in \{1..M\}$ characterize the reliability of the different channels.

The contribution of this article is twofold. First, we establish a new weighting strategy tailored for deformable models and that is well adapted to vector-valued images in which the representativeness of the studied object is varying along the channels. Second, we exploit the above-mentioned geometrical framework to define a new gradient vector flow field through nonlinear diffusion of both directional and scalar information carried by the LST.

#### B. Weights Calculation

Without a priori knowledge, the extension of Di Zenzo’s approach to 3D consists in weighting all channels equally:

$$ \omega_1 = 1/M. \quad (11) $$

With such weights, contributions from noise in channels where the object is poorly represented might hamper edge detection. Different application-specific solutions have been proposed to weight the contribution of the LST, based for example on noise estimations in the diffusion framework [39], or local saliency for image fusion purposes [40].

In GVF-based approaches, it is desirable to maximize the contrast-to-noise ratio (CNR) of the gradient signal in order to perform accurate diffusion of the edge map gradient vectors throughout homogeneous regions of the image. In low contrast images, the gradient signal due to noise can be superior to the gradient signal due to true edges. This can bias the estimation of directions and magnitudes of vector edges and, consequently, the directions of the GVF force field. Here, we propose to exploit the active contours framework to define a new weighting strategy tailored for image segmentation, based on a blind estimation of the contrast of the object in each channel. Such a global weighted averaging scheme can maintain high contrast of the vector gradient signal while reducing its variance [41].

For each channel $I_k$, let $R_k^{in}$ be the set of voxels located inside $S'$, and let $R_k^{ext}$ be the set of voxels located outside $S'$ and inside $\lambda S'$, a morphological dilation of $S'$ of $\lambda$ units of length. The limitation of $R_k^{ext}$ to $\lambda S'$ prevents possible
influence from further regions. We define the weighting factor for channel $I_k$ as follows:

$$
ω_k := \left( \frac{|\bar{I}_k^{in} - \bar{I}_k^{out}|}{\sum_j |\bar{I}_j^{in} - \bar{I}_j^{out}|} \right)^\gamma,
$$

(12)

where summation is over the $M$ channels. $\bar{I}_k^{in}$ and $\bar{I}_k^{out}$ are the average intensities in $R_k^{in}$ and $R_k^{out}$ respectively. $\gamma$ is a parameter that controls the linearity of the influence of the channels.

Fig. 1 illustrates the weighting scheme on a conceptual 2D example. Fig. 1a displays a high contrast channel where average intensities in $R_k^{in}$ and $R_k^{out}$ in each channel.

![Fig. 1. Illustration of the weighting method on a 2D representation of the active surface $S^\tau$ around a region of interest (ROI). (a) High contrast channel. (b) Low contrast channel. A measure of contrast is established by comparing average intensities in $R_k^{in}$ and $R_k^{out}$ in each channel.](image)

The 4DGVF external force field is the result of nonlinear diffusion of the vectorial edge map $\vec{V}$ throughout the image. The eigenvectors of $G_\omega$ form an orthogonal set in the directions of maximal (gradient) and minimal (isophote) change. This illustrates the fact that the vectorial edge map $\vec{V}$ is orthogonal to vector edges, which is not necessarily the case for $\nabla N_\omega$ (and $a\text{ fortiori}$ for $\nabla N_1$).

D. The 4DGVF Equation

The 4DGVF external force field is the result of nonlinear diffusion of the vectorial edge map $\vec{V}$ throughout the image. In the vicinity of vector edges, as detected by $N_\omega$, the directions of the vectors are constrained by $\vec{V}$, while isotropic diffusion of $\vec{V}$ occurs in homogeneous regions. The 4DGVF field is defined as the steady-state solution of the following
where elasticity terms are weighted by $\alpha$ by $\beta$ (Fig. 3b) and the vectorial edge map vector partial differential equation:

$$\vec{J} = \vec{g} \nabla^2 \vec{F} - h \vec{F} + \vec{V},$$

where $g$ and $h$ are the functions used in eq. (7), replacing the gradient magnitude $|\nabla f|$ with $N_w$. At each iteration $\tau$ of the deformation, the surface $S^\tau$ undergoes locally the external force field $\vec{F}_{ext}$. To avoid convergence issues and ensure a smooth deformation, the deformation force field is projected on the normal direction to $S^\tau$. The surface is iteratively moved according to the following gradient descent flow:

$$\frac{\partial S^\tau}{\partial \tau} = \alpha \Delta S^\tau - \beta \Delta^2 S^\tau + \vec{F}_{ext} \cdot \vec{n},$$

where elasticity terms are weighted by $\alpha$ and rigidity terms by $\beta$. $\vec{n}$ denotes the normal direction to the local surface element $dS^\tau$. The LST is computed according to the proposed weighting scheme, and at each timestep $\tau$, weights are recomputed to construct a more accurate external force field for the next iteration.

Fig. 3 shows a comparison between the eigenvector field $\vec{\theta}_i$ (Fig. 3b) and the vectorial edge map $\vec{V}$ (Fig. 3c), superimposed on one frame of a dynamic PET image (Fig. 3a). The resulting 4DGVF field is shown in Fig. 3d. Fields are projected on a 2D slice. While the diagonalization of the LST does not uniquely specify the sign of the gradient vectors, equation (14) lifts the indeterminacy and orients the field toward vector edges, a desirable property for external force fields. The resulting 4DGVF field is consistent with the studied object.

E. Initialization

Minimizing equation (3) is equivalent to finding the ideal isosurface of minimal total energy $E$. In general, the energy landscape associated with the segmentation problem is not convex, requiring the initial model to be close to the desired optimum. To this end, we propose to initialize the 4DGVF model with an extension of the Poisson Inverse Gradient (PIG) approach, proposed by Li and Acton [27], to vector-valued images. The PIG approach approximates the potential energy $E_{ext}$, from which the external force field would derive. As fields such as GVF fields and VFC fields are nonconservative, this scalar potential does not exist and is estimated through a least-squares minimization problem. The initial model is identified as the isosurface of the reconstructed external energy $E_{ext}$ with lowest energy.

We adapt the PIG approach to vector-valued images and build an initialization field $\vec{F}^0_{ext}$ based upon the 4DGVF framework. We use equal weighting of all channels, following eq. (11), as finer weights such as proposed in section III-B can only be derived after an initial surface is defined. Once $\vec{F}^0_{ext}$ is computed, we estimate the scalar potential $E_{ext}$ by solving the Poisson equation:

$$\Delta E_{ext} = -\nabla \cdot \vec{F}^0_{ext},$$

This equation is solved numerically by matrix inversion for which Dirichlet boundary conditions are applied on the boundary $\partial \Omega$ of the image domain $\Omega$:

$$E_{ext}(\partial \Omega) = -N_w(\partial \Omega).$$

We scale $E_{ext}$ in the range $[0, -1]$, and perform $P$ triangulated isosurface reconstructions for different values $E_p = (E_1, E_2, ..., E_P)$, $E_p \in [0, -1]$, using a marching cubes algorithm [43]. In our experiments, we retain only closed surfaces as candidates models. We then select the surface model with minimal total external energy. This shape is then used for the computation of the initial weights prior to the deformation. To emphasize that vector-valued information is used in the initialization, we refer to this initialization in the following as Vector Poisson Inverse gradient (VPIG).

F. Numerical Implementation

We implemented our method using MATLAB®. The active surface $S^0$ was represented as a triangulated mesh and oriented such that the normals to the faces point inwards. Normals to each vertex were computed as the weighted average of the face normals incident to the vertex [44]. We solved eq. (16) with a standard finite difference approach expressed in a matrix form [6]. The Laplacian was linearly estimated at each vertex $v_i$ by the umbrella operator [45]:

$$\Delta(v_i) = \frac{1}{|i|} \sum_{j \in i^*} v_j - v_i,$$

where $i^*$ corresponds to the neighborhood of $v_i$ (vertices connected directly to $v_i$). As this number remained constant throughout the deformation and the internal forces only depended on the Laplacian, the neighborhood matrix corresponding to internal forces needed to be inverted only once. The deformable surface was considered to have converged when the maximum displacement of vertices between two iterations was less that 0.1 voxel side.
In our experiments, the amplitude maps $N_α$ were scaled in the range $[0, 1]$. In each image, the parameter $κ$ that controls the trade-off between field smoothness and gradient conformity was set so as to maximize the Jaccard similarity score between segmentation result and ground truth. The parameter $γ$ that controls the linearity of the weights in equation (12) was empirically set to 2, emphasizing the relevance of high contrast channels.

IV. Validation Setup

A. Comparison With Other Approaches

We compared the proposed 4DGVF approach with two single-channel and two vector-valued approaches of the literature.

- The Generalized Gradient Vector Flow (GGVF). For this single-channel approach, the diffusion of the gradient vectors is performed in each channel by finding the equilibrium solution of equation (7) [9], where the edge map of channel $I_k$ is defined as follows:

$$f_k = K_σ \ast |\nabla I_k|.$$  \hfill (20)

In our experiments, we scaled every $f_k$ in the range $[0, 1]$. The parameter $κ$ that controls the trade-off between field smoothness and gradient conformity was manually set so as to maximize the Jaccard similarity score between segmentation result and ground truth.

- The Vector Field Convolution (VFC) is also a single channel approach in which a convolution is performed between $f_k$ and a vector field kernel $C$ whose vectors point toward the kernel’s center [10]:

$$\hat{F}_{ext} = C \ast f_k.$$  \hfill (21)

In each image, the size of $C$ and the power parameter of the magnitude function were set so as to maximize the Jaccard similarity score.

In our experiments, GGVF and VFC approaches were performed in each channel of the tested images. For comparison with vector-valued approaches, we retained in each 4D image the channel that obtained the best Jaccard score.

- The Vector Geometric Snake (VGS) is an extension of the implicit geometric snake to 3D vector-valued images [2], where the edge-stopping term of the level-set function is based on the gradient magnitude of the vector-valued image from a local structure tensor that weighs all channels equally:

$$G_1 = K_σ \ast \sum_{k=1}^{N} \omega_l (\nabla I_k \otimes \nabla I_k^T),$$  \hfill (22)

The corresponding gradient magnitude $N_1$ is then derived from the eigenvalues of $G_1$. In our experiments, we scaled $N_1$ in the range $[0, 1]$.

- The Color Gradient Vector Flow (CGVF) [31] uses the gradient magnitude $N_1$ of the vector-valued image as the edge detector from which to perform gradient diffusion in the GVF equation (7).

Table I summarizes the edge detection terms used in the above-mentioned models.

B. Initialization

As results depend on the quality of the initialization, all comparative tests between models were performed using identical initialization. We generated results with the following initialization models:

- Ellipsoid initialization: an ellipsoidal shape centered around the object and fitting its shape.
- VPIG initialization: the approach described in section III-E, built upon the PIG approach and the initial 4DGVF field.

C. Validation Criterion

When a ground truth was available (synthetic 4D data, PET image simulations), the segmentation results were compared to the true volumes after convergence by using the Jaccard coefficient $J(A, B)$ that expresses volume similarity [46]. It is the ratio between the intersection and the union of the ground truth volume ($A$) and the segmented object ($B$). It ranges from 0 to 1, with 1 meaning a perfect match:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}.$$  \hfill (23)

For VGS, only visual results were analyzed, as the high noise levels of the studied images led to multiple spurious surface reconstruction in all the tested images, making the Jaccard score uninterpretable. This issue stressed the need for topology-preserving procedures for such applications, a known drawback of implicit representations [21].

V. Synthetic 4D Benchmark

We generated 5 synthetic images of dimensions $70 \times 50 \times 50 \times 10$ voxels featuring a spherical object of diameter 36 voxels inside a uniform background. In each image, the noise-free background intensity was set to 1 in every channel while the object intensity varied along the channels. By analogy with PET imaging studied hereafter in section VI, the curve that represents the different values of a voxel along the channels is referred to as its time-activity curve (TAC). We generated 5 4D images for which the object TACs are displayed in Fig. 4. From these noise-free images, a set of noisy 4D images was generated by adding white Gaussian noise in the channels ($σ = 0.2$). As a consequence, the number of channels in which the contrast between foreground and background was...
significantly superior to the noise-induced gradient amplitude varied along the different images of the dataset. These images presented two difficulties that can be found in 4D images, namely low SNR that makes edge detection challenging in individual channels, as well as variations in representativeness of the sphere along the channels.

For the model initialization, a spherical triangulated surface mesh of radius 10 voxels was placed at the center of mass of the object (initial Jaccard index of 0.23).

1) **Weighting Influence on Edge Detection:** The different channels in a 2D slice of one of the synthetic 4D images are displayed at the top of Fig. 5. The corresponding weights obtained with the proposed 4DGVF approach after initialization of the surface model are shown below. In this image, the weighting scheme was consistent with the observed variations of contrast. It was indeed desirable to lower the contribution of channels 4 to 7 in the calculation of the LST.

To study the influence of weighting the LST we compared the proposed weighted magnitude $N_{\omega}$ to the magnitude $N_1$ as well as the gradient amplitude $f_k$ obtained in the best contrasted channel. Fig. 6 displays representative amplitude edge maps of one slice, where the maximum value of each 3D map was set to 1. The edge signal was enhanced with $N_{\omega}$ compared to $f_k$ and $N_1$, and in homogeneous regions, spurious variations due to noise were kept at lower levels, leading to better edge detection with the $N_{\omega}$ amplitude edge map.

We studied quantitatively the quality of the vector gradient amplitude $N_{\omega}$ as a function of the number of high contrast channels used for its calculation. We computed $N_{\omega}$ using varying numbers of high contrast channels, ranging from 1 (best channel only) to 10 (all channels included: proposed method). The criterion for assessing the quality of the resulting amplitude maps was the CNR of $N_{\omega}$:

$$CNR = \frac{\mu_e - \mu_{bg}}{\sigma_{bg}},$$

where $\mu_e$ is the average intensity value of edge voxels, $\mu_{bg}$ is the average intensity of non-edge voxels (background), and $\sigma_{bg}$ is the standard deviation of the background. Fig. 7 displays CNR for $N_{\omega}$ maps averaged over the dataset as a function of the number of channels (ordered by decreasing contrast) used for their calculation. On average, using the $n = 7$ channels with the highest contrast values led to maximum CNR for $N_{\omega}$ (CNR = 2.36), with an increase of about 53% compared with the CNR obtained by only using the highest contrasted channel ($n = 1$, CNR = 1.53). Including all channels with the proposed weighted averaging did not lower significantly the CNR of the edge maps compared to the maximum value ($n = 10$, CNR = 2.34). This is due to the fact that low contrast channels have a reduced influence on the LST calculation with the proposed blind weighting scheme. In addition, this strategy does not require any prior selection of the number of channels that are considered valuable for edge localization.

2) **Orientations of Vector Edges:** We evaluated the accuracy of the force field around edges by comparing the 4DGVF field based on the diffusion of $\mathbf{V}$ to a CGVF field based on the diffusion of $\nabla N_{\omega}$ rather than $\nabla N_1$, in order to remove the influence of the weighting scheme.
Table II shows quantitative results averaged over the synthetic benchmark image dataset. For each image, a single result was obtained for vector-valued models (CGVF and 4DGVF) while we retained the channel corresponding to the best Jaccard similarity coefficient for single-channel approaches (GGVF and VFC). Overall, analysis of the Jaccard results suggests that the 4DGVF model improved segmentation results. CGVF provided the second best results after 4DGVF on average, but single channel approaches were able to perform better than CGVF in sufficiently well contrasted channels. The 4DGVF scheme was less hampered by the low contrast channels than CGVF due to the weighting scheme.

VI. VALIDATION ON DYNAMIC PET IMAGES

Dynamic PET imaging consists in the successive acquisition of different time frames of an identical field of view. A dynamic PET image can therefore be considered as a vector-valued image, where each time frame of the sequence corresponds to one channel. The resulting images reflect the dynamics of a radiotracer concentration in the body, but suffer from low resolution and low SNR. In these images, regions of interest, or kinetic regions have varying contrasts with respect to the surrounding regions over time. PET imaging is a functional imaging modality that can provide information unavailable in structural imaging modalities such as computed tomography. There has been a growing interest for its application to the early diagnosis of neurodegenerative pathologies such as Alzheimer’s disease [47], [48] or amyotrophic lateral sclerosis [49], and in the study of neuroinflammation [50], [51].

The validation of segmentation results using real clinical images is difficult due to lack of ground truth. While we show examples of application to real data in section VI-E, we assess the 4DGVF approach with quantitative results on realistic Monte Carlo simulations of dynamic PET images.

A. PET-SORTEO Benchmark Image Database

We used simulations of [11C]-Raclopride dynamic 3D+t PET images of the brain from the publicly available PET-SORTEO benchmark image database (http://sorteo.cermep.fr) [52]. We focused on the segmentation of the putamen in both cerebral hemispheres. These realistic images account for the inter-individual variability of anatomical structures by using different real MR images as numerical head models. Each dynamic PET volume has dimensions of 128 × 128 × 63 × 26 voxels, while each MR volume contains 181 × 217 × 181 voxels. We limited our experiments to 4 images of the benchmark: the Jacob, P02, P03 and P04 images. Each of the dynamic PET images were registered to their corresponding MR volumes with a rigid registration algorithm using the medical imaging software PMOD v.3.4. The provided labeled MR images were used as ground truth for the validation of the segmentation results. For a fair comparison, we used for this dataset identical
Fig. 10. Segmentation of the left putamen superimposed with a transaxial slice of image P02 of the PET-SORTEO benchmark. Ground truth (black wireframe), initial model (white wireframe) and 4DGVF result after convergence (solid yellow).

| JACCARD SEGMENTATION RESULTS FOR THE LEFT (L) AND RIGHT (R) PUTAMINA IN THE PET-SORTEO BENCHMARK IMAGE DATASET |
|--------------------------------------------------|----------------|----------------|----------------|----------------|
| Method                                          | GGVF (best ch.) | VFC (best ch.) | CGVF (best ch.) | 4DGVF (best ch.) |
| Image                                           | L  | R  | L  | R  | L  | R  | L  | R  |
| Jacob                                           | 0.65 | 0.68 | 0.67 | 0.68 | 0.62 | 0.63 | 0.77 | 0.77 |
| P02                                             | 0.61 | 0.58 | 0.62 | 0.61 | 0.60 | 0.61 | 0.68 | 0.70 |
| P03                                             | 0.61 | 0.58 | 0.62 | 0.61 | 0.60 | 0.54 | 0.65 | 0.66 |
| P04                                             | 0.57 | 0.58 | 0.59 | 0.61 | 0.48 | 0.58 | 0.60 | 0.64 |

Fig. 11. Simulations of dynamic PET images. (a) Zubal phantom (b) Zubal simulation (mid SNR) (c) Zubal simulation (low SNR).

initial ellipsoid models for every method. We centered a sphere at the center of mass of the considered putamen in the corresponding labeled MR image. The radius of the sphere was set to 10 voxels (initial Jaccard index of 0.46 on average).

Fig. 10 shows a cropped transaxial slice in the 20th channel of the P02 image around the putamina (slice 69/181). The 3D surfaces of the ground truth, of the initial model, and of the 4DGVF model after convergence are showed in black wireframe, white wireframe and solid yellow respectively. The 4DGVF model was able to capture the shape of the left putamen. For this image, the Jaccard score of the 4DGVF model after convergence was 0.68. According to this criteria, the 4DGVF model outperformed the other approaches for both left and right putamen segmentation on all tested images (table III).

B. Simulations of Realistic 4D PET Images With GATE

Additional dynamic PET images were simulated using GATE, a highly realistic medical image simulator based on the CERN’s GEANT4 particle interaction platform [53], [54]. We used the Zubal head phantom [55], describing the main brain structures as a voxelized source. Six regions were considered for the simulation: cerebellum, thalamus, parietal, frontal and occipital lobes, and the remaining parts of the brain were the background (Fig. 11a).

Time-activity curves, which represent the variations of each voxel intensity along the time frames, were generated according to a three-compartment model [56] that models the kinetics of the radiotracer in the body. The reconstruction of the PET image was performed using a fully 3D OP-OSEM (ordinary Poisson ordered-subset expectation-maximization) iterative method into $2.2 \times 2.2 \times 2.8$ mm$^3$ voxels. We performed two different reconstructions: one using 2 iterations and 16 subsets (Fig. 11b) and one using 10 iterations and 16 subsets (Fig. 11c) that resulted in different levels of SNR, called mid and low respectively. The simulation of these two images required 90 days of parallel computations on a 12 cores 48 GB RAM computer.

For each of the reconstructed images of the Zubal head phantom, we studied the segmentation of the cerebellum and of the thalamus, two structures showing different kinetics and volumes, colored in red in Fig. 11a.

C. Weighting Scheme

Fig. 12a displays 2D transverse slices of the 20 frames of the low SNR Zubal simulation around the thalamus. Thalamus can be distinguished as a hypersignal in frames of the second row and as a hyposignal in late frames. Estimated weights for the calculation of the corresponding LST are shown in Fig. 12b after convergence of the model. The 4DGVF weights were in good agreement with the subjective quality observed in each...
channel: the weight values followed the variations of contrast along the channels and thereby the representativeness of the object.

D. Segmentation Performances on 4D PET Simulations

Fig. 13 illustrates segmentation results in the mid SNR Zubal simulation around the cerebellum. 2D slices of the results are displayed for the sake of readability. The first column shows the ground truth and the ellipsoidal shape used as initialization. Columns b-f present segmentation results of the tested methods superimposed with the edge map used by the method and with the corresponding external force field. The intersection of the active surface with the slice is depicted as a red line and the ground truth as a black line. While GVF, VFC, CGVF and 4DGVF all led to consistent cerebellum segmentation, the best result was achieved with 4DGVF, followed by CGVF. 4DGVF was able to better capture the concavity of the cerebellum formed by the fourth ventricle (bottom part of Fig. 13f). Again, the implicit VGS model created numerous splitted reconstructed surfaces of various sizes and shapes and hence was not quantitatively evaluated.

Fig. 14 shows 3D representative segmentation results in the low SNR simulation for all tested methods, either using the VPIG initialization (top row, cerebellum) or using an ellipsoidal initialization (bottom row, thalamus). For both initialization methods, the overall shape of the two objects was best recovered by the 4DGVF approach. Despite low SNR conditions, the 4DGVF approach was able to capture the thalamus, a small region compared to the voxel size. The second best segmentation of the thalamus was obtained with CGVF, with segmentation results that were however visually less precise than the 4DGVF model (Fig. 14f and Fig. 14m).

Table IV shows the quantitative similarity criteria between the segmentation results and the ground truth after convergence for the two different simulations. For both images, segmentation results were improved by the 4DGVF approach. In the case of the segmentation of the cerebellum, the best VFC and GGVF results produced relatively high Jaccard. However, single-channels approaches require to select the channels of interest \textit{a priori}, which is not always feasible in practice. On the contrary, a single segmentation result was obtained for vector-valued models, with systematic improvement observed for 4DGVF.

In general, the 4DGVF external force field benefited from comparatively larger attraction range than other approaches, allowing initialization from farther distances, which is confirmed by the stability of the segmentation result against the initialization. The cumulative effect of exploiting accurate gradient directions drawn from the vectorial edge map and of weighting the image channels led to better performance of 4DGVF over CGVF, which obtained second best results on average. The improved robustness of 4DGVF under low
SNR is emphasized in the case of the noisier Zubal simulation, where the proposed method led to distinct improvements of figures of merit.

In the tested images, the weighting scheme led to systematic enhancement of the gradient magnitude map \( \nabla \omega \) over \( \nabla \omega \) to the benefit of the 4DGVF approach. The most computationally expensive aspect of the method lies in the re-calculation of the force field due to the re-evaluation of the weights at each iteration. However, in practice, recalculating weights at each iteration is not necessary because they converge to a steady result along with the surface model. For example, one can re-evaluate weights depending on the amount of global deformation of the model, as small deformations are likely to cause negligible change in the weights. While the proposed weighting scheme is convenient for numerous modalities and applications, the 4DGVF approach can be enriched by other type of weights, also based for example on noise estimation or available \textit{a priori} knowledge.

E. Illustration on Real Data

To illustrate the behavior of the 4DGVF approach in a preclinical context, we performed a dynamic PET acquisition of a rat using [18F]-DPAA714 injections, a radiotracer specific to the translocator protein (TSPO). This protein is over expressed under pathologic neuroinflammatory conditions and can therefore measure active diseases in the brain. The inflammation was produced by performing unilateral quinolinic acid lesions on the right striatum of the rat. Images were acquired on a microPET-CT GE Vista in list-mode and reconstructed using a 3D-OSEM iterative method with corrections for attenuation, in list-mode and reconstructed using [18F]-DPAA714 in the right striatum of the rat. Result is shown in red, initial shape in white (left), superimposed onto the 4DGVF field (middle, right).

We have proposed a novel external force field for the segmentation of vector-valued images using parametric active surfaces. The proposed 4DGVF field enables the segmentation of noisy 4D images where edge information cannot be deduced from a single channel, and where only the redundancy of edge information along the channels can help recover the feature of interest. In the 4DGVF approach, the gradient signal is weighted according to a blind estimation of contrast, favouring channels in which edges are better defined. While existing approaches based on deformable models applied to vector-valued images exploit local structure information in a scalar way to define vector edges, the 4DGVF field is produced through nonlinear diffusion of a vectorial edge map computed from the eigenvector analysis of the local structure tensor, improving robustness to noise.

Quantitative assessment on synthetic images and realistic simulations, and results on real dynamic PET images confirmed the potential impact of the method for segmentation of vector-based 2D or 3D imaging modality, such as dynamic PET, functional MRI or hyper-spectral imaging.

VII. CONCLUSION

We have proposed a novel external force field for the segmentation of vector-valued images using parametric active surfaces. The proposed 4DGVF field enables the segmentation of noisy 4D images where edge information cannot be deduced from a single channel, and where only the redundancy of edge information along the channels can help recover the feature of interest. In the 4DGVF approach, the gradient signal is weighted according to a blind estimation of contrast, favouring channels in which edges are better defined. While existing approaches based on deformable models applied to vector-valued images exploit local structure information in a scalar way to define vector edges, the 4DGVF field is produced through nonlinear diffusion of a vectorial edge map computed from the eigenvector analysis of the local structure tensor, improving robustness to noise.

Quantitative assessment on synthetic images and realistic simulations, and results on real dynamic PET images confirmed the potential impact of the method for segmentation of vector-based 2D or 3D imaging modality, such as dynamic PET, functional MRI or hyper-spectral imaging.

REFERENCES


Vincent Jasoen received the master’s degree in applied physics from the University of Rennes 1, Rennes, France, in 2012. He is currently pursuing the Ph.D. degree with the Imaging and Brain Unit, INSERM, François Rabelais University, Tours, France. He is also a Teaching Assistant of Signal Processing with the National Institute of Applied Sciences (INSA), School of Engineering, Blois, France. His main research interests are PDE-based segmentation and filtering approaches applied to medical image analysis. He is interested, in particular, in quantifying neuronalinflammatory disorders with PET imaging.

Simón Stute is currently a Researcher with the Service Hospitalier Frédéric Joliot for the Nuclear Energy Agency, Orsay, France. He received the Ph.D. degree regarding his work on the development and use of realistic Monte Carlo simulations for improved quantification in PET in 2010. He is also a part of the Biomedical Physics Team, and specialized in the development and application of new image reconstruction techniques in PET.
Denis Guilloteau is currently a Pharmacist, Biologist, and Head of the Department of Nuclear Medicine, CHU Bretonneau, Tours, France, where he leads research in the field of molecular imaging, and in particular, the development of new radiopharmaceuticals. He manages the Imaging and Brain Unit at INSERM, François Rabelais University, Tours, France, which includes radiochemists, mathematicians, biologists, radiologists, neurologists, geriatricians, and psychiatrists. One of the objectives of this unit is to develop imaging techniques applicable to the Alzheimer’s disease for the diagnosis and the follow-up of the efficiency of treatments. He has authored over 160 international articles.

Irène Buvat received the Ph.D. degree in particle and nuclear physics from Paris-Sud University, Orsay, France, in 1992, and then spent one year with University College London, London, U.K., and two years with the National Institutes of Health, Bethesda, MD, USA. She joined the Centre National de la Recherche Scientifique, Paris, France, in 1995, and is currently the Head of Research with the Service Hospitalier Frédéric Joliot, PET Center, Orsay. Her research activities focus on developing correction and tomographic reconstruction methods in PET and SPECT to improve the accuracy and to reduce the variability of measurements made from PET and SPECT images. She is currently a Spokesperson of the worldwide OpenGATE collaboration developing the GATE Monte Carlo simulation tool dedicated to Monte Carlo simulations in emission and transmission tomography and radiotherapy. She is also largely involved into making quantification in SPECT and PET a clinical reality and contributed to many studies demonstrating the clinical values of quantification to improve image interpretation. She has authored or co-authored over 110 peer-reviewed articles. She is an Associate Editor of the IEEE TRANSACTIONS ON MEDICAL IMAGING and the IEEE TRANSACTIONS ON NUCLEAR SCIENCE, and serves on the Editorial Board of the Journal of Nuclear Medicine and the International Advisory Board of Physics in Medicine and Biology.

Sylvie Chalon is currently a Neurobiologist and the Head of the Team Brain Molecular Imaging with the Imaging and Brain Unit, INSERM, University of Tours, Tours, France. She works on the development of new fluor-18 labeled PET tracers, and preclinical studies in animal models and clinical applications. Her current research focuses on PET, exploration of neurodegeneration and neuroinflammation in animal models of brain disorders, such as Parkinson and Alzheimer’s disease. She has authored over 120 peer-reviewed papers in this field.


He has been an Associate Professor with the University of Tours, Tours, France, since 2008, where he is currently with the UMRS INSERM U930 Imaging and Brain Unit. He is responsible for the Medical Image Processing Group with the Brain Molecular Imaging Team. His research interests are centered around medical image reconstruction, filtering, segmentation, registration, and quantification, with particular interest in exploiting the consistency of spatiotemporal information in dynamic image sequences.