Discrete Methods for Guide-wire Segmentation and Tracking

Nicolas Honnorat, Régis Vaillant, Nikos Paragios

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Discrete Methods for Guide-wire Segmentation and Tracking

Nicolas Honnorat and Régis Vaillant and Nikos Paragios

N° 7718

Août 2011

ISSN 0249-6399 ISRN INRIA/RR--7718--FR+ENG

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE

INRIA

centre de recherche
SACLAY - ÎLE-DE-FRANCE

Discrete Methods for Guide-wire Segmentation and Tracking

Nicolas Honnorat and Régis Vaillant and Nikos Paragios

Theme :
Équipes-Projets

Rapport de recherche n° 7718 — Août 2011 — 20 pages

Abstract: In this report, we propose a complete pipeline for guide-wire segmentation and tracking in the context of cardiac angioplasty. Segmentation is addressed through a bottom up approach which first detects interest points using appropriate features. These points are then clustered into local segments using geometric and proximity constraints. Ordering of these segments through a local permutation model is finally carried out towards complete delineation of the guide-wire. Tracking is addressed using a discrete iconic-geometric approach. Control points are used to represent the guide-wire. Their displacements in time aims at optimizing an image-driven likelihood (on the same space used for segmentation), preserving the geometric characteristics of the curve while establishing correspondences for meaningful landmark points automatically extracted along the guide-wire. The tracking and the landmark correspondence problem form a two layer inter-connected graphical model, that is solved using linear programming towards simultaneous recovery of the two set of variables. Large scale clinical validation and comparisons with state of the art methods demonstrate the potentials of our method.

Key-words: Markov Random Field, guide-wire, segmentation, tracking, clustering.

Centre de recherche INRIA Saclay – Île-de-France
Parc Orsay Université
4, rue Jacques Monod, 91893 ORSAY Cedex
Téléphone : +33 1 72 92 59 00

Méthodes discrètes pour la segmentation et le suivi de guides

Résumé : Dans ce rapport, nous proposons un système complet de segmentation et de suivi des guides utilisés en angioplastie cardiaque. La segmentation est réalisée selon une approche agglomérative qui consiste dans un premier temps à détecter des points intéressants en extrayant des caractéristiques appropriées des images. Ces points sont ensuite regroupés en prenant en compte des contraintes de géométrie et de proximité en paquets formant des segments de droites. L'ordonnancement de ces segments par permutations successives permet finalement de reconstruire le guide entièrement. Le suivi est réalisé selon une approche discrète couplant recalages iconique et géométrique. Des points de contrôle sont utilisés pour décrire le guide. Leur déplacement d'image en image est choisi de manière à maximiser une probabilité calculée à partir des images tout en préservant la forme du guide. Simultanément, des points d'intérêt extraits le long du guide sont mis en correspondance avec des candidats extraits dans l'image suivante. Le recalage iconique et la mise en correspondance de points d'intérêt forment un modèle graphique à deux couches inter-connectées qui est résolu en utilisant des techniques de programmation linéaire. Une validation clinique approfondie et une comparaison avec les techniques de l'état de l'art démontrent le potentiel de notre méthode.

Mots-clés : Champs de Markov, guides, segmentation, suivi, clustering.

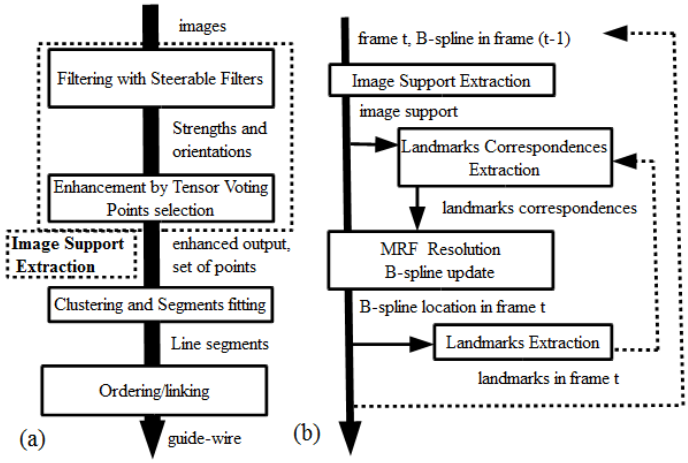

Figure 1: summary of our segmentation (a) and tracking method (b).

1 Introduction

Cardiac angioplasty interventions involve thin wires that are used to position balloons and stents inside the blood vessel, called guide-wires. In this context, fluoroscopic images are acquired on the fly and help the physicians to position these devices. Therefore automatic segmentation and tracking of these wires can be of great aide. However, addressing this task is a great challenge due to the fact that images suffer from low signal to noise ratio.

This problem was initially addressed in [1] and since then several works studied segmentation and tracking of these wires. Segmentation is challenging due to mis and false-detections. The low signal to noise ratio makes the detection of thin structures in a continuous manner almost infeasible while the presence of biological structures of similar appearance and geometry produces numerous "false"-detections. In terms of tracking, the visual support is also an issue while the observed motion is a combined one, that encompass cardiac, respiratory motion, and elongations controlled by the physician.

Therefore, issues mentioned earlier in terms of visual support are amplified in time (one can observe variations on the appearance of the guide-wire) while one has to face large, non-rigid and unforeseeable displacements.

Segmentation and tracking of guide-wires are tow distinct problems that share a common component referring to the extraction of visual support of the wire from the images. Many curvilinear structure detectors could and have been used like the Hessian [2], the Vesselness [3], steerable filters [4] and phase congruency [5]. Recent work combined dedicated interest points detectors along with concepts introduced in the machine learning community. Probabilistic Boosting Tree is an example that integrates Haar wavelet responses [6]. In [4], boosting acting on steerable filters and image statistics was also used. This image support has been further improved through curvilinear structure enhancement method, such as tensor voting [7, 8] and coherence enhancing diffusion [2].

Segmentation/guide-wire delineation was then addressed using three distinct methodologies. The first class of methods detects seeds belonging to the wire

and then expands/grows them towards complete recovery of the wire including its tip [9]. Multiple hypothesis testing (used recently for blood vessel extraction in [10]) and particle filtering [11] belong to this class as well. Bottom-up approaches are an alternative to the methods earlier presented. These methods extract primitives that are then grouped together through an optimization process. One can cite [6] among others, where edgels are linked according to the output of boosted classifiers and [4] related to the work presented here. The last group of methods rely on shortest path computations, such as [12] or fast marching methods [13]. Such approaches have been applied to blood vessel and microtubule segmentation [14] and recently adopted for segmenting guide-wires [12]. The common limitations of the aforementioned segmentation methods refer to the appropriate selection of the feature space, the fact that visual support of the wires varies locally and the presence of visually similar biological structures.

Tracking of filamentary structure in general and guide-wires in particular is also a problem well studied. B-splines are the most common representation, and tracking is achieved through the estimation of the optimal displacement of the control points. Such displacements are often driven from image-based likelihoods and are estimated using various optimization methods. In [5], a gradient descent is adopted where tracking is expressed within the formalism of B-spline snakes [15, 16]. Active surfaces have been used in [17, 18] to segment and track moving actin filaments. Discrete optimization is an alternative to descent-like method and was also considered either in the form of dynamic programming [19] or more recently Markov Random Fields in [9]. In order to cope with combined motion, [20] proposed a multi-resolution rigid registration step followed by a non rigid one. Despite important progress made, most of the aforementioned methods fail to cope with large displacements, are sensitive to the lack of visual support or computationally inefficient.

In this paper, we present a segmentation and a tracking method towards complete, robust and efficient delineation of guide-wires in fluoroscopic images. The segmentation is determined using a bottom-up approach: pixels likely to be part of the guide-wire are extracted and splitted into clusters assumed to correspond to line-segments. These segments are ordered using a local permutation model that aims at optimizing a global criterion involving the whole set of line segments and taking into account geometric and visual continuity. [Fig. (1.a)] summarizes our segmentation method.

Our tracking method unifies an iconic tracking part related to [9, 21] with a landmark matching. We adopt a B-spline representation that involves both interpolation/control and landmark points. Control points are used to determine the form of the guide-wire, while the landmark capture "geometrically" interesting features on it. The image likelihood refers to an integral along the spline, while the tracking components aim at establishing local correspondences between the parts of the guide-wire and the landmark points. This is achieved through the joint optimization of the position of the control and the landmark points while establishing automatic correspondences. [Fig. (1.b)] summarizes our tracking method.

The remainder of the paper is organized as follows: in section 2, we describe our approach for guide-wire segmentation. In section 3, we introduce a graph-based framework for parametric curve tracking unifying an iconic tracking approach with a geometric/landmark matching. Section 4 is dedicated to the clinical validation of the method while the last section concludes the paper.

2 Bottom-up Guide-wire Segmentation

Segmentation of guide-wires involves three components, (i) the selection of features and the definition of parameter space, (ii) an appropriate optimization problem that couples parameters with features, and (iii) the inference algorithm used to determine the optimal solution. We adopt a bottom approach as suggested in [4].

2.1 Guide-wire Detection

Numerous detectors exist in the literature for extraction of curvilinear structures. Optimal approaches are often observed with problem-specific trained detectors as suggested in [6, 4]. The main limitation of these methods refer to the generalization of their components learnt in training to unknown data, a problem that often arises in medical imaging either due the use of different parameter settings or different hardware devices. In order to present a more robust methodology we adopt a method based on steerable filters [22]. Such operators refer to a powerful feature extraction paradigm that can capture efficiently the strength and the orientations of edges. These orientations constitute in practice an extremely valuable information that will be exploited by the segmentation and the tracking component of our method.

We extract the visual support of the guide-wires using the most sensitive second-order steerable filter for edge detection. We obtain thus at each point $\mathbf{x}$ a filter response $\hat{f}(\mathbf{x})$ and an orientation $\hat{\omega}(\mathbf{x})$. Because guide-wires are dark structures, we consider the responses:

$$g(\mathbf{x}) = max(0, -\hat{f}(\mathbf{x}))$$

Such information space suffers from precision due to the low signal-to-noise ratio of fluoroscopic images, and therefore we consider the steerable filter output as a tensor field $\theta(\mathbf{x})$ by setting:

$$\theta(\mathbf{x}) = g(\mathbf{x}) \left[\mathbf{e}_{\hat{\omega}(\mathbf{x})} \otimes \mathbf{e}_{\hat{\omega}(\mathbf{x})}\right]$$

where $\mathbf{e}_\omega$ is the unit vector of orientation ω ; and we adopt a tensor voting scheme towards enhancing the wires. The second order voting field presented in [23] was used during our experiments. In cylindrical coordinates (r, ϕ) this field refers to:

$$V(r,\phi) = \frac{1}{16} e^{\left[-\frac{r^2}{2(\sigma_{TV})^2}\right]} cos^4(\phi) \begin{pmatrix} 1+cos(4\phi) & sin(4\phi) \\ sin(4\phi) & 1-cos(4\phi) \end{pmatrix}$$

In [23], the authors prove that applying this field leads to the new responses $f(\mathbf{x})$ and orientations $\omega(\mathbf{x})$:

$$\begin{aligned} f(\mathbf{x}) &= |A_{-2}(\mathbf{x})| \\ \omega(\mathbf{x}) &= \frac{1}{2} arg(A_{-2}(\mathbf{x})) \end{aligned}$$

where A_{-2} is given by five convolutions between complex-valued images and filters ($*$ denotes a convolution, $^-$ the complex conjugate and i the imaginary

unit):

$$\begin{aligned} A_{-2} &= (w_0 * \bar{c_2}) + 4(w_2 * c_0) + 6(w_4 * c_2) \\ &+ 4(w_6 * c_4) + (w_8 * c_6) \\ w_m(r,\phi) &= e^{\left[-\frac{r^2}{2(\sigma_{TV})^2}\right]} e^{im\phi} \qquad m \in \{0,2,4,6,8\} \\ c_m(\mathbf{x}) &= g(\mathbf{x}) e^{-im\hat{\omega}(\mathbf{x})} \qquad m \in \{0,2,4,6\} \end{aligned}$$

We compared the precision obtained with a standard Hessian filter and with the steerable filter, with or without tensor voting. [Fig. (2.b)] indicates that the tensor-voting clearly enhances the wires. On the contrary, the steerable filter demonstrates a negligeable improvement with respect to the standard Hessian.

For the sake of clarity we will denote the guide-wire support in the frame t with:

$$\begin{aligned} \mathcal{I}(\mathbf{x};t) &= (\mathcal{I}(\mathbf{x};t)_x, \mathcal{I}(\mathbf{x};t)_y) \\ &= (f(\mathbf{x};t)cos(\omega(\mathbf{x};t)), f(\mathbf{x};t)sin(\omega(\mathbf{x};t))) \end{aligned}$$

and abbreviate with $\mathcal{I}(\mathbf{x})$ the visual support of the first frame, that is used for segmentation. In order to reduce the computational complexity of the method only a small portion of the top responding pixels are retained. We will denote with X the set of detected points.

Bottom-up approaches assume that the structure of interest can be represented using a collection of local primitives. Our method assumes that guide-wires can be represented using a collection of local segments. Such a representation can be very rich (through the use of fine segments), can cope with outliers (individual detections cannot form a segment with reasonable support) and is compact. The segmentation of the wires will correspond to the ordering of these segments.

2.2 Primitives Extraction

In order to extract these segments, we adopt an unsupervised clustering formulation. To this end, we have to determine a cost of associating a detected point with a cluster (that is also represented by a point: its center). Local proximity constraints between points corresponding to the same segment is enforced through the use of the Euclidean distance, while geometric proximity can be enforced using the orientation information within the Elastica model [24]. Therefore, the cluster assignment cost between the cluster of center $k_i \in X$ and the point $x_j \in X$ can be expressed as follows:

$$d_\alpha(k_i, x_j) = d(k_i, x_j)^2 + \alpha e(\omega(k_i), \omega(x_j))$$

where $e(.,.)$ denotes the scale invariant Elastica criterion (and α is a positive parameter).

Given these definitions, the next step consists in determining the number of clusters, their centers and the optimal cluster assignments for each $x_j \in X$ with the cardinality of $\{X\}$ being N. This objective can be expressed as a discrete labeling problem involving two sets of binary variables: the $\delta(k_i, x_j)$ equal to one when the point x_j is assigned to the cluster of center k_i and the set of $\delta(k_i)$

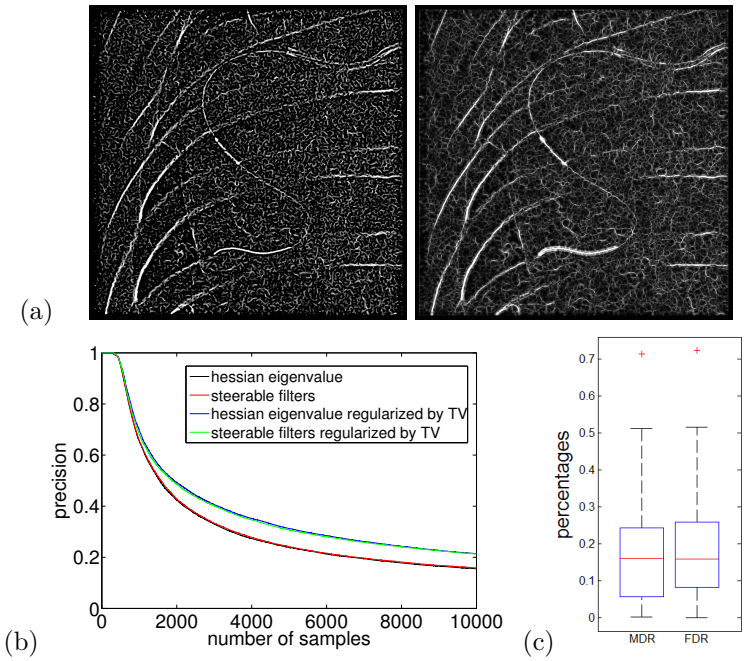

Figure 2: (a) Norm of the steerable filter response and norm of the tensor voting output. (b) Proportion of pixels extracted by the different filters belonging to the wires. These mean precisions were estimated on 3 images by ranking pixels by decreasing filter response. (c) Missed Detection Rate (MDR) and False Detection Rate (FDR) corresponding to the sequence of line segment provided by the segmentation.

equal to one when k_i is considered to be a center of a retained cluster. We assume that all the points have the same a priori likelihood of being a cluster center. In order to cope with outliers, we also introduce a virtual cluster center x_{N+1}. Considering once again that all the x_j have the same a priori likelihood of being outliers, we penalize all the outliers with the same cost d_{N+1} by setting:

$$d_\alpha(x_{N+1}, x_j) = d_{N+1} \qquad \forall x_j \in X$$

Given the above assumptions, the task of local segment extraction is equivalent with the optimization of the following functional:

$$E(\delta) = \sum_{\substack{i\in[1,N+1],\\ j\in[1,N],\ i\neq j}} \delta(x_i, x_j) d_\alpha(x_i, x_j) + K_c \sum_{i\in[1,N]} \delta(k_i)$$

where K_c is the cost of introducing a cluster, and under the constraint:

$$\delta(k_i, x_j) \leq \delta(k_i) \quad \forall k_i \in X, \forall x_j \in X$$

This problem is solved using LP-stability based clustering [25] that was designed to solve similar problems. This method can provide simultaneously the optimal number of clusters, their centers and the clusters/outliers attributions.

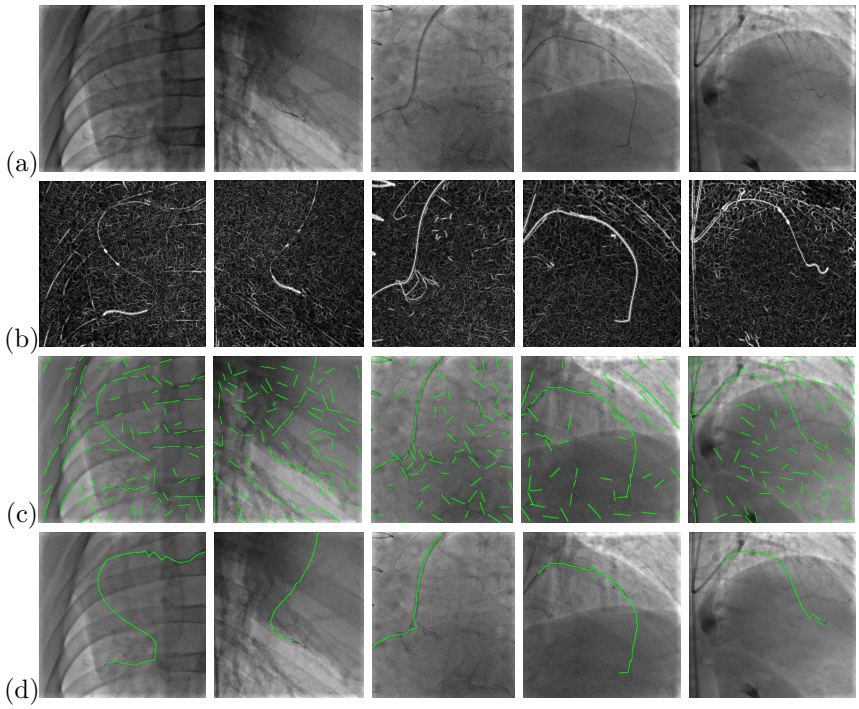

Figure 3: (a) input images (b) regularized responses (intensity has been windowed so as to increase the contrast) (c) extracted line segments (d) final segmentation

Once clusters and their attributes are determined, we approximate them with one line segment using the MSAC algorithm [26]. [Fig. (3.c)] presents some examples of extracted line segments. Let us denote with S the set of line segments found. Some of these primitives are part of the curvilinear structures that we want to segment and one now has to order them properly towards complete recovery of the guide-wire.

2.3 Guide-wire Delineation through Primitive Ordering

Given a set of primitives, guide-wire segmentation consists of selecting a subset and ordering it. We perform these tasks simultaneously thanks to the ordering method presented in [4].

Let us denote with $s_n \in S$ a segment primitive and $(\mathbf{a_n}, \mathbf{b_n})$ its extremities. We define a cost $C(\mathbf{x}, \mathbf{y})$ for each pair of primitive extremities $\{\mathbf{x}, \mathbf{y}\}$ and a cost for each segment $C(s_n)$ corresponding to the cost of removing of s_n from the list of inlier primitives.

Segmentation aims at selecting a set of inliers $\mathcal{J}$ and on ordering function $\sigma : \mathcal{J} \to \mathbb{N} \times \{0, 1\}$ associating a rank and an "orientation" to each inlier. By convention, an "orientation" equal to 1 will indicate that a primitive s_n has to be considered in the order $(\mathbf{b_n}, \mathbf{a_n})$ (and otherwise, the order $(\mathbf{a_n}, \mathbf{b_n})$ will be considered). This function allows to construct $\sigma^{-1}(i)$ indicating what is the i-th segment extremity being part of the wire. Given such notations, segmentation

of the guidewire can be determined through the minimization of the following functional:

$$\min_{\mathcal{J},\sigma^{-1}} \sum_{s_n \notin \mathcal{J}} C(s_n) + \sum_{i=1}^{|\mathcal{J}|-1} C(\sigma^{-1}(2i), \sigma^{-1}(2i+1))$$

In order to take into account the strength and the orientation of the image support $\mathcal{I}$ linking extremities $(\mathbf{x}, \mathbf{y})$, cost $C(\mathbf{x}, \mathbf{y})$ decreases when the strength of the image support increases and when the cosine of the angle between the orientation of the link and the local orientations increases. The cost $C(\mathbf{x})$ of removing a segment is defined as the complement (with respect to 1) of the image-based linking cost by taking as extremities the end points of the segment itself. We also penalize the euclidean distance and the curvature with the elastica criterioneading to the following functional ($\langle ., . \rangle$ denotes the standard scalar product):

$$\begin{aligned} C(\mathbf{x}, \mathbf{y}) &= \delta_{iconic} \int_{\mathbf{x}}^{\mathbf{y}} \psi \left(\frac{|\langle \mathcal{I}(\mathbf{s}), \mathbf{y} - \mathbf{x} \rangle|}{||\mathbf{y} - \mathbf{x}||} \right) ds \\ &+ \delta_{euclidean} ||\mathbf{y} - \mathbf{x}|| + \delta_{elastica} \left[e(\mathbf{x}, \mathbf{y}) \right] \\ C(s_n) &= \int_{\mathbf{a_n}}^{\mathbf{b_n}} 1 - \psi \left(\frac{|\langle \mathcal{I}(\mathbf{s}), \mathbf{b_n} - \mathbf{a_n} \rangle|}{||\mathbf{b_n} - \mathbf{a_n}||} \right) ds \end{aligned}$$

where δ_{iconic},$\delta_{euclidean}$ and $\delta_{elastica}$ are constant parameters fixing the relative influence of the support, the Euclidean distance and the elastica criterion and where (β being a positive parameter):

$$\psi(x) = e^{-\beta x}$$

2.4 Optimization

We consider a local search/permutation approach to solve the above mentioned problem [4]. To this end, we maintain two sets of primitives: a sequence of inliers, that once ordered will be the output of the algorithm, and a sequence of outliers. The line segments are considered as inliers by a first function that orders them greedily. Iteratively then, we order the inlier segments sequence, order the outlier segments sequence and then exchange subsequences of inliers with subsequences of outliers or flipped subsequences of outliers towards decreasing the objective energy.

The ordering of inliers and outliers sequences is addressed through by local search: segment permutations operations are applied while they reduce the sum of linking costs. Given that the local search methods suffer from local minima, we consider a large number of permutation operation: all the possible flips, displacements and displacements followed by a flip of any subsequence of line segments. These operations are illustrated in [Fig. (4)]. The convergence of our whole algorithm is guaranteed because all the operations applied reduce the objective function.

In order to cope with the presence of multiple guide-wires, we can add a semi-interactive component where the user determines the first and the last segment of the guide-wire. Such constraint is implemented by allowing permutations

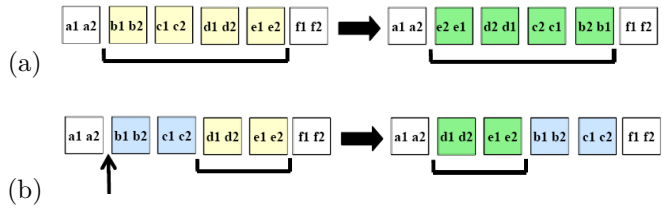

Figure 4: (a) flip of a subsequence. (b) displacement of a subsequence. The flip modify both the rank in the sequence and the orientation of line segments.

for all than the first and the last line segment, while introducing "suppression-flip" operator. This operation is simply a suppression of a subsequence at the beginning or at the end of the sequence followed by the flip of the first or the last element.

Last but not least additional user interactions constraints can be considered like for example indication of line segments being part of the guide [12] by modifying the singleton cost $C(s_n)$ of the segments s_n selected by the user. Contrary to [12], the user can also indicate line segments that are not part of the structure. In [Fig. (3)] segmentation results obtained with a maximum number of three user seeds are presented. Once the structure of interest has been determined, the next step consists of tracking the guide-wire in the whole sequence.

3 Geometric-iconic Curve Tracking

Landmark based methods are efficient object tracking techniques. These methods extract salient primitives that are matched between images under geometric constraints, like distances and angles preservation [27]. In this methodology, motion of rigid objects can be determined robustly with only a few correspondences. On the other hand, this approach is not suited for tracking structures undergoing large non-rigid deformations because the object support between primitives is not taken into account.

Iconic tracking methods [28] refer to a promising alternative since they consider dense image support. Coarse-to-fine strategies are often necessary to guide the inference of deformations parameters to a good minimum and deal with erroneous image support.

The motion of the guide-wires used during cardiac angioplasty interventions are large, unforeseeable and non-rigid. Therefore landmark or iconic methods will fail, the first due to the important internal geometric variation of the object of interest and the second due to the important motion and lack of strong visual support.

Let us describe the guide-wire as a B-spline of control points $\mathbf{c}(i)$ and denote with $\mathbf{p}(j)$ a set of landmarks associated to the guide-wire. The set of geometric landmarks is data-driven and can vary within the tracking sequence. We will denote with $\mathbf{c}(i;t)$ the control points describing the guide-wire found in the frame t, and with $\mathbf{p}(j;t)$ the landmarks associated with it. $\mathbf{p}^L(j;l_j;t+1)$ will denote a set of L candidates extracted in the frame $t+1$ considered as possible matches with the landmaks $\mathbf{p}(j;t)$. Our unified iconic-geometric tracking seeks

for the best $\mathbf{c}(i;t+1)$ given $\mathbf{c}(i;t)$, $\mathbf{p}(j;t)$, their possible matches $\mathbf{p}^L(j;l_j;t+1)$ and the image support $\mathcal{I}(\mathbf{x};t+1)$.

We adopt a unified MRF framework combining singleton and pairwise potentials involving the displacements of $\mathbf{c}(i;t)$, that are distinguished by the labels l_i and the correspondences of $\mathbf{p}(j;t)$, that are distinguished by the label l_j:

$$E_{MRF} = \sum_{x\in\mathcal{P}} \mu_x \left(\sum_{k\in\{i,j\}} \Theta_k^x(l_k) + \sum_{k,r\in\{i,j\}} \Theta_{k,r}^x(l_k,l_r) \right)$$

where $\mathcal{P}$ refers to the potentials determined according to the data, prior, landmarks and coupling terms:

$$\mathcal{P} = \{data,\ landmarks,\ template,\ coupling\}$$

3.1 Iconic Parametrized Curve Tracking

A real time MRF-based iconic tracking has been introduced in [9] and improved in [29]. The curve is described with a cubic B-spline and a MAP methodology is followed to find the most likely displacements of the control points $\mathbf{c}(i;t)$ from the last frame t given the new image support $\mathcal{I}(t+1)$. This framework considers the displacement of the control points as random variables and build a graph containing one node per control point.

We adopt a similar framework where the guide-wire is represented using a cubic B-spline of control points $\mathbf{c}(i;t)$ and their admissible displacements correspond to labels l_i. We introduce potentials involving pairs of succesive control points and we build a pairwise MRF. Let $s \in [0,1]$ be the curvilinear abscissa along the spline, $B_i(s)$ the basis function associated to the control point $\mathbf{c}(i;t)$. Then, the spline curve $\mathcal{C}(s)$ is described by:

$$\mathcal{C}(s) = \sum_i B_i(s)\mathbf{c}(i;t)$$

As suggested in [9] we adopt $B_i^{i+1}(.)$, an influence function that can approximate the fourth-order potentials with second order potentials:

$$B_i^{i+1}(s) = \frac{B_i(s)B_{i+1}(s)}{\sum_k B_k(s)B_{k+1}(s)}$$

Denoting with $\mathbf{v}(l_i)$ the displacement corresponding to the label l_i of the control point $\mathbf{c}(i;t)$, let $C_{l_i}^{l_{i+1}}(s)$ be the curve obtained when the control points $\mathbf{c}(i;t)$ and $\mathbf{c}(i+1;t)$ are displaced by $\mathbf{v}(l_i)$ and $\mathbf{v}(l_{i+1})$ respectively (and similarly ${C'}_{l_i}^{l_{i+1}}(s)$ its derivative).

We adopt two sums of pairwise potentials [9] involving successive control points. The first decreases when the curve passes through pixels likely to belong to the guide-wire and when the orientation of the derivative of the curve coïncides with the local orientation [2]:

$$\Theta_{i,i+1(l_i,l_{i+1})}^{data} = \int_{s(\mathbf{c}(i;t))}^{s(\mathbf{c}(i+1;t))} B_i^{i+1}(s)\psi\left(\frac{|\langle {C'}_{l_i}^{l_{i+1}}(s),\ \mathcal{I}(C_{l_i}^{l_{i+1}}(s);t+1)\rangle|}{||{C'}_{l_i}^{l_{i+1}}(s)||}\right) ds$$

where $s(\mathbf{c}(i;t))$ is the curvilinear abscissa of $\mathbf{c}(i;t)$. The second sum penalizes the local deformations of the derivatives of the spline with respect to a template $\mathcal{T}(s)$, that encodes the past curve location and is updated using exponential forgetting:

$$\mathcal{T}(s) \leftarrow \frac{memory-1}{memory}\mathcal{T}(s) + \frac{1}{memory}C(s)$$

$$\Theta^{template}_{i,i+1(l_i,l_{i+1})} = \int_{s(\mathbf{c}(i;t))}^{s(\mathbf{c}(i+1;t))} B_i^{i+1}(s)\frac{||\epsilon\ C'(s,l_i,l_{i+1}) - \mathcal{T}'(s)||^2}{||\mathcal{T}'(s)||^2}\,ds$$

This template imposes harder constraints than length preserving priors [5, 9].Contrary to [9], we neglect points of curvilinear abcissa out of the interval $[s(\mathbf{c}(i;t)), s(\mathbf{c}(i+1;t))]$ when computing pairwise potentials. As a result, the contribution of pixels far from the knots is higher than the contribution of pixels near the knots.

3.2 Geometric Tracking

We consider a MAP methodology to determine the optimal match $\mathbf{p}^L(j;l_j;t+1)$ of each landmark $\mathbf{p}(j;t)$;

Let us first discuss the extraction of landmarks and the choice of correspondences. We detect potential landmarks by considering the main eigenvalue of a matrix being related to the structure tensor [30]:

$$T(\mathbf{x};t) = \sum_{\mathbf{y}} e^{-\frac{||\mathbf{x}-\mathbf{y}||^2}{2{\sigma_L}^2}} \begin{pmatrix} \mathcal{I}(\mathbf{y};t)_x^2 & \mathcal{I}(\mathbf{y};t)_x\mathcal{I}(\mathbf{y};t)_y \\ \mathcal{I}(\mathbf{y};t)_x\mathcal{I}(\mathbf{y};t)_y & \mathcal{I}(\mathbf{y};t)_y^2 \end{pmatrix}$$

Such measure has a highest strength for pixels where strong and colinear steerable filter responses are accumulated. In order to take the orientation of the local structure into account as well when selecting the candidate correspondences $\mathbf{p}^L(j;l_j;t+1)$, we selected the L candidates $\mathbf{y}$ with the highest matching cost(where $\langle .,. \rangle$ denotes the inner product between matrices):

$$\mathscr{M}_j(\mathbf{y}) \quad = \quad \langle T(\mathbf{p}(j;t);t) | T(\mathbf{y};t+1) \rangle$$

Let us now describe the Markov Random Field built for the matching. We consider a fully connected pairwise MRF containing one node per landmark and we find the labels l_j minimizing a matching energy that penalizes the deformations of the cloud of points. In order to handle mis-detections and erroneous correspondences, we also introduce a label $(L+1)$ meaning that no correspondence for a landmark was found, being penalized by the following singleton potential:

$$\Theta_j^{landmarks}(l_j) = \begin{cases} C & if\ l_j = L+1 \\ 0 & otherwise \end{cases}$$

We used the following pairwise terms to penalize landmarks configurations changes:

$$\Theta^{landmarks}_{j,k(l_j,l_k)} \quad = \quad \begin{cases} 0 \ \ if\ l_j = L+1\ or\ l_k = L+1 \\ \min\left(||\mathbf{u}(j,l_j,k,l_k) - \mathbf{u}(j,k)||^2, \Gamma\right) \ \ else \end{cases}$$

where the configurations of pairs of landmarks are described by the following vectors:

$$\begin{aligned}\mathbf{u}(j,l_j,k,l_k) &= \mathbf{p}^L(j;l_j;t+1)-\mathbf{p}^L(k;l_k;t+1)\\ \mathbf{u}(j,k) &= \mathbf{p}(j;t)-\mathbf{p}(k;t)\end{aligned}$$

Such a penalization insures that the configurations matched are coherent in terms of Euclidean distance [31] as well as in term of orientation. In other words, translations are allowed with no cost but all the other geometric transformations of the set of landmarks are penalized. This penalization, however, mainly concerns pairs of neighboring landmarks, since the potential is quickly satured to the value Γ for other pairs. Matching results are presented in [Fig. (5)]. Based on the above methodology we can solve individually the iconic tracking and the landmark correspondence problem. Conventional methods to integrate them refer to expectation maximization that solves the problems in an alternating manner. We adopt a single shot optimization method that solves both problems at once through a coupling term.

3.3 Coupling Term

Let $\mathbf{w}_j(l_j)$ be the displacement of the interest points corresponding to the landmark matching:

$$\mathbf{w}_j(l_j)=\mathbf{p}^L(j;l_j;t+1)-\mathbf{p}(j;t)$$

Besides, let us note with $z_j \in [0,1]$ the curvilinear abscissa of $\mathbf{p}(j;t)$ and let recall that $\mathbf{v}_i(l_i)$ is the displacement of the control point $\mathbf{c}(i;t)$ with respect to the frame t and that $B_i(.)$ is the basis function of $\mathbf{c}(i;t)$. The landmark motion due to the displacement of the control points and the landmark motion corresponding to the landmark matching are equal if and only if:

$$\sum_i B_i(z_j)\mathbf{v}_i(l_i)=\mathbf{w}_j(l_j)$$

We consider the error-two norm to penalize deviations between the positions due to the iconic and landmark tracking:

$$\begin{aligned}E_j(\{l_i\},l_j) &= ||\sum_i B_i(z_j)\mathbf{v}_i(l_i)-\mathbf{w}_j(l_j)||^2\\ &= \sum_{i\neq k} B_i(z_j)B_k(z_j)\langle\mathbf{v}_i(l_i),\mathbf{v}_k(l_k)\rangle\\ &- 2\sum_i B_i(z_j)\langle\mathbf{v}_i(l_i),\mathbf{w}_j(l_j)\rangle\\ &+ \sum_i B_i(z_j)^2||\mathbf{v}_i(l_i)||^2+||\mathbf{w}_j(l_j)||^2\end{aligned}$$

The constraint in [27] can be considered as a relaxation of our coupling term. Unfortunately, such an approach does not insure that the iconic tracking will be guided by the landmark associations because many different sets of displacements of the $\mathbf{c}(i;t)$ correspond to the same displacement of $\mathbf{p}(j;t)$. In order to address this concern, we introduced pair-wise constraints between the control

points used for describing the position of each landmark $\mathbf{p}(j;t)$ with the aim of rigidifying the spline around the landmarks:

$$E_{r_j}(\{l_i\}) = \sum_{k \in v(j)} ||\mathbf{v}_k(l_k) - \mathbf{v}_{k+1}(l_{k+1})||^2$$

where $v(j)$ denotes the list of the indices of the control points related to $\mathbf{p}(j;t)$. This leads to the following coupling term:

$$\begin{aligned} \sum_k \Theta_k^{coupling}(l_k) \quad &+ \quad \sum_{k,r} \Theta_{k,r}^{coupling}(l_k, l_r) \\ &= \quad \sum_j E_j(\{l_i\}, l_j) + \eta \sum_j E_{r_j}(\{l_i\}) \end{aligned}$$

4 Experimental Validation

This section presents implementation details and the experiments carried out to validate our method.

4.1 Data Set

The validation was carried out on a database containing 20 fluoroscopic sequences of 200 frames of different sizes between 512×512 and 1000×1000 pixels, corresponding to 20 seconds of observations. These sequences were chosen in such a manner that they involve a variety of clinical settings as large as possible. Manual delineation from experts was used as ground truth.

4.2 Parameter Setting

The parameters were set using the first three sequences and the method was tested using the same set of parameters for the remaining 17 sequences.

For the segmentation, we tiled our images and considered only a small set of points from each tile. These points were the 120 points with strongest responses taken from a grid of step 4, and the size of the tiles was 200 x 200 points during our experiments. We used the steerable filters [32] at scale 1.8. The enhancement of their output was made at scalee $\sigma_{TV} = 4.5$.

The parameter α of the criterion used for pixel clustering was set to 100.0, we chose for the penalty of the cluster centers $K_c = 15000.0$ and the cost of the background pixels was: $d_{N+1} = 16000.0$. The MSAC was performed at scale 6.0. For the linking part parameters, we chose: $\delta_{iconic} = 6.5$, $\delta_{elastica} = 1.8$ and a small $\delta_{euclidean} = 0.1$.

For the tracking part, we set $\mu_{data} = \mu_{template} = 1.0$, $\mu_{landmarks} = 2.10^{-4}$, $\mu_{coupling} = 3.10^{-5}$, $\eta = 0.1$, $\epsilon = 0.9$, the cost $C = 45000$ for the landmarks not matched, displacements steps equal to $(25.0)2^{-\tau}$ for $\tau \in \{0 \dots 4\}$ and we sampled 25 displacements according to the sparse pattern in [9]. We also set $L = 24$, $\beta = 6.10^{-3}$,$\Gamma = 500$ and $memory = 7$. The landmarks were determined at scale $\sigma_L = 2.5$.

A coarse-to-fine approche was adopted, like in [20] towards accelerating the tracking. The positions of the control points were updated frame-by-frame towards their length-driven uniform distribution along the spline.We made landmark detections more homogeneous along the wire by defining intervals of fixed

length along the curve, separated by a fixed curvilinear abscissa and by choosing at most one landmark in each of them. Severe elongations of the guide-wire were managed by elongating the end of the B-spline by following its tangent $\mathbf{d}$ before the control-points update: pixels $\mathbf{q}$ were added to the end of the spline while: $\psi(< \mathbf{d}, \mathcal{I}(\mathbf{q}) >) < 0.4$. The solver FastPD [33, 34] known to be faster than other state-of-the art solvers was used for minimizing E_{MRF}. It already allowed to report real time results for guide-wire tracking [9].

4.3 Evaluation

The evaluation protocol was consisting of segmenting the guide-wire in the first frame using up to three user seeds and then automatically tracking it through the remaining sequence. The results are expressed in terms of missed detection rate (MDR) and false detection rate (FDR) for a distance of 5 pixels and for every 25 frames (what corresponds to 2.5 seconds or two heart beats). These rates correspond respectively to the proportion of pixels of the ground truth that lies too far from the spline and to the proportion of the spline that lies too far from the ground truth. [Fig. (2.c)] presents our segmentation results. Except for two sequences, where the image support was not sufficient to allow a precise segmentation, these results are of the same order as the state of the art [6] (mean FDR around 22% and mean MDR around 10%). Our tracking method was however robust enough to partially recover the missing parts of the wire in these two challenging sequences.

We have compared our method with [9][1]. The ground-truth was used as initial segmentation for both methods. We observed that the overall performance of our method is significantly better than the one of [9]. The gain of performance/added value for the geometric part of the method is demonstrated by turning off the corresponding component [Fig. (6)]. In that case [9] outperforms our method mostly due to the coarse to fine displacement strategy. The segmentation component is also a solid component of the pipeline since no loss of overall tracking performance is observed when the suggested method is used as initial guess instead of the ground truth. [Fig. (6)] also presents tip tracking error. Because the wire tip is easy to detect, this error is acceptable when smaller than 15 pixels and due to a slide of the tracker along the wire.

4.4 Computational Complexity

Our tracking method treats a frame in nearly 2 seconds (on a Intel Xeon 2.8 GHz processor) but the image support extraction could be easily accelerated. This computation time depends also dramatically on the size of the displacement space explored (quite over-sized with our current parameter choice) and on the time allocated to the enhancement of the visual support by tensor voting.

5 Conclusion

In this paper, a complete pipeline based on graphical models was proposed to delineate and track parametrized curves in the context of cardiac angioplasty.

[1] We have provided the test sequences, the ground truth and asked the authors to calibrate their method towards optimal performance.

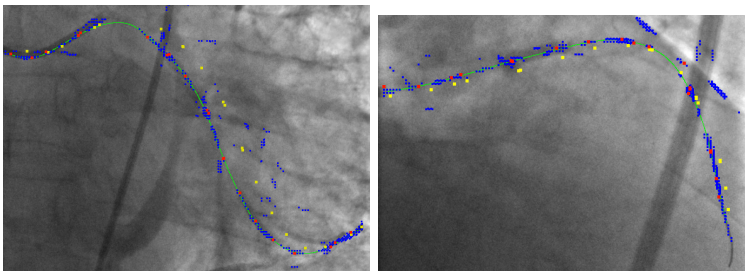

Figure 5: spline (green) landmarks detected along the previous spline (yellow), candidates for the matching (blue) and candidates chosen (red).

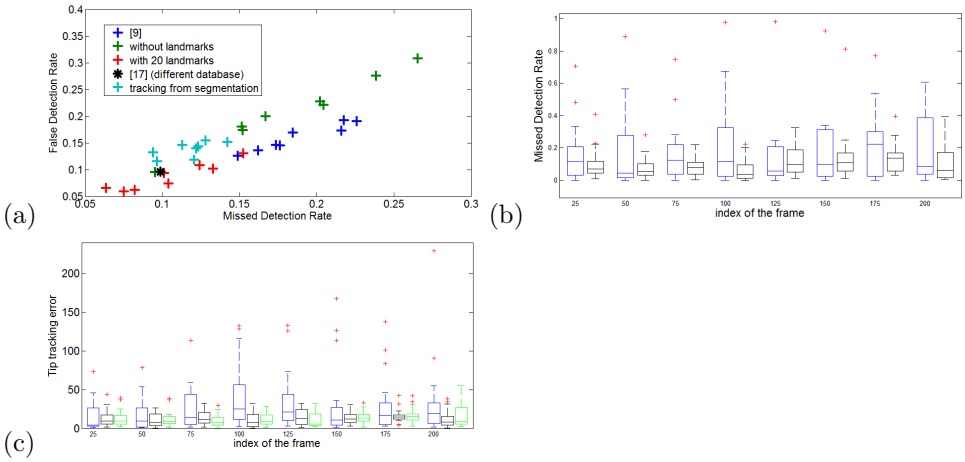

Figure 6: (a) Mean tracking results. [9] has been applied with a vesselness at scale 1.8 like our filters and $\lambda = 0.7$. The database for result [20] is different. The last result were obtained with our method by initializing our spline from the segmentation output and by considering 20 landmarks (b) comparison of MDR obtained with [9] (blue boxes) and with our method (black boxes) when ground truth is used as initialization. (c) tip tracking error in pixel, for method [9] (blue) and for our method when initialized with ground truth (black) or segmentation output (green)

The main strength of the method was the combination of visual support and geometric constraints through a couple segmentation/registration approach.

Future work will aim to reduce the user-interaction step from the segmentation component. This could be achieved through the fusion of fluoroscopic images and contrast-enhanced ones. Towards clinical use software/hardware acceleration of the method towards achieving real time performance is also feasible. The use of graphics processing units for the implementation of the image likelihoods could easily address this necessity. Because multiple guide-wires and confusing structures are present in fluoroscopic sequences, modifying our linking module toward performing semi-automatic simultaneous multiple structure delineation could increase the robustness of our approach for guide-wire segmen-

tation. Furthermore, the introduction of a motion model that could separate the observed motion between ego-motion of the wire and motion of the heart could improve robustness.

Acknowledgment

This work was supported by ANRT (grant 1062/2008) GE Healthcare and the European Research Council Starting Grant Diocles (ERC-STG-259112). The authors thank N. Komodakis for providing the clustering method and the MRF solver, and thank T. Hauke Heibel for performing validation experiments with the method [9].

References

[1] D. Palti-Wasserman, A. M. Bruckstein, and R. P. Beyar, "Identifying and tracking a guide wire in the coronary arteries during angioplasty from x-ray images," *IEEE Tr. on Bio. Eng.*, vol. 44, no. 2, pp. 152–164, feb. 1997.

[2] S. A. M. Baert, M. A. Viergever, and W. J. Niessen, "Guide-wire tracking during endovascular interventions," *IEEE TMI*, vol. 22, no. 8, pp. 965–972, aug. 2003.

[3] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *Wells,W.M., Colchester, A.C.F., Delp, S.L. (eds.) MICCAI 1998.*, ser. LNCS, vol. 1496. Springer, Heidelberg, 1998, pp. 130–137.

[4] N. Honnorat, R. Vaillant, and N. Paragios, "Guide-wire extraction through perceptual organization of local segments in fluoroscopic images," in *MICCAI 2010, Part III*, ser. LNCS, vol. 6363. Springer, 2010, pp. 440–448.

[5] G. Slabaugh, K. Kong, G. Unal, and T. Fang, "Variational guidewire tracking using phase congruency," in *MICCAI 2007, Part II*, ser. LNCS, vol. 4792. Springer, 2007, pp. 612–619.

[6] A. Barbu, V. Athitsos, B. Georgescu, S. Boehm, P. Durlak, and D. Comaniciu, "Hierarchical learning of curves application to guidewire localization in fluoroscopy," in *CVPR*, 2007.

[7] G. Guy and G. Medioni, "Inferring global perceptual contours from local features," *International Journal of Computer Vision*, vol. 20, pp. 113–133, 1996.

[8] E. Franken, P. Rongen, M. van Almsick, and B. M. ter Haar Romeny, "Detection of electrophysiology catheters in noisy fluoroscopy images," in *MICCAI*, 2006, pp. 25–32.

[9] T. Hauke Heibel, B. Glocker, M. Groher, N. Paragios, N. Komodakis, and N. Navab, "Discrete tracking of parametrized curves," in *CVPR*, 2009, pp. 1754–1761.

[10] O. Friman, M. Hindennach, C. Kühnel, and H.-O. Peitgen, "Multiple hypothesis template tracking of small 3d vessel structures," *Medical Image Analysis*, vol. 14, pp. 160–171, 2010.

[11] I. Smal, K. Draegestein, N. Galjart, W. Niessen, and E. Meijering, "Particle filtering for multiple object tracking in dynamic fluorescence microscopy images: Application to microtubule growth analysis." *IEEE TMI*, vol. 27, pp. 789–804, 2008.

[12] P. Wang, W.-s. Liao, T. Chen, S. K. Zhou, and D. Comaniciu, "Graph based interactive detection of curve structures in 2d fluoroscopy," in *miccai*, 2010.

[13] T. Deschamps and L. D. Cohen, "Fast extraction of minimal paths in 3d images and applications to virtual endoscopy (2001)," *Medical Image Analysis*, vol. 5, p. 281Ű299, 2001.

[14] S. Hadjidemetriou, D. Toomre, and J. Duncan, "Segmentation and 3d reconstruction of microtubules in total internal reflection fluorescence microscopy (tirfm)," in *MICCAI*, 2005.

[15] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models." *International Journal of Computer Vision*, vol. 4, p. 321Ű331, 1987.

[16] P. Brigger, J. Hoeg, and M. Unser, "B-spline snakes: A flexible tool for parametric contour detection." *IEEE Trans. on Image Processing 9*, vol. 9, p. 1484Ű1496, 2000.

[17] M. A. El Saban and B. S. Manjunath, "Tracking curvilinear structures using active contours and application to microtubule videos," in *ICIP*, 2005.

[18] H. Li, T. Shen, D. Vavylonis, and X. Huang, "Actin filament segmentation using spatiotemporal active-surface and active-contour models," in *MICCAI*, 2010, pp. 86–94.

[19] D. Geiger, A. Gupta, L. A. Costa, and J. Vlontzos, "Dynamic programming for detection, tracking, and matching deformable contours," *PAMI*, vol. 17, 1995.

[20] P. Wang, T. Chen, Y. Zhu, W. Zhang, S. K. Zhou, and D. Comaniciu, "Robust guidewire tracking in fluoroscopy," in *CVPR*, 2009.

[21] B. Glocker, A. Sotiras, N. Komodakis, and N. Paragios, "Deformable medical image registration: Setting the state of the art with discrete methods." *Annual Reviews of Biomedical Engineering*, vol. 13, pp. 219–244, 2011.

[22] E. H. A. William T. Freeman, "The design and use of steerable filters," *IEEE PAMI*, vol. 13, no. 9, pp. 891–906.

[23] E. Franken, M. van Almsick, P. Rongen, L. Florack, and B. ter Haar Romeny, "An efficient method for tensor voting using steerable filters," in *ECCV*, 2006.

[24] E. Sharon, A. Brandt, and R. Basri, "Completion energies and scale," *IEEE PAMI*, vol. 22, no. 10, pp. 1117–1131, 2000.

[25] N. Komodakis, N. Paragios, and G. Tziritas, "Clustering via lp-based stabilities," *NIPS*, 2008.

[26] P. Torr and A. Zisserman, "Robust computation and parametrization of multiple view relations," in *ICCV*, 1998, pp. 727–732.

[27] A. Sotiras, Y. Ou, B. Glocker, C. Davatzikos, and N. Paragios, "Simultaneous geometric - iconic registration," in *MICCAI 2010, Part II*, ser. LNCS, T. Jiang, N. Navab, J. P. Pluim, and M. A. Viergever, Eds., vol. 6362. Springer, Heidelberg, 2010, pp. 676–683.

[28] M. Isard and A. Blake, "Condensation conditional density propagation for visual tracking," *IJCV*, vol. 29, 1998.

[29] O. Pauly, T. Hauke Heibel, and N. Navab, "A machine learning approach for deformable guide-wire tracking in fluoroscopic sequences," in *MICCAI 2010, Part III*, ser. LNCS, vol. 6363. Springer, 2010, pp. 343–350.

[30] C. Harris and M. Stevens, "A combined corner and edge detector," in *4th Alvey Vision Conference*, 1988.

[31] Y. Ou, A. Besbes, M. Bilello, M. Mansour, C. Davatzikos, and N. Paragios, "Detecting mutually-salient landmark pairs with mrf regularization," in *ISBI*, 2010.

[32] M. Jacob and M. Unser, "Design of steerable filters for feature detection using canny-like criteria," *IEEE PAMI*, vol. 26, no. 8, pp. 1007–1019, 2004.

[33] N. Komodakis, G. Tziritas, and N. Paragios, "Performance vs computational efficiency for optimizing single and dynamic mrfs: Setting the state of the art with primal-dual strategies," *CVIU*, vol. 112, pp. 14–29, 2008.

[34] N. Komodakis and G. Tziritas, "Approximate labeling via graph-cuts based on linear programming," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007.

Contents

Centre de recherche INRIA Saclay – Île-de-France
Parc Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 Orsay Cedex (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Sophia Antipolis – Méditerranée : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399