Kidney Detection and Segmentation in Contrast-Enhanced Ultrasound 3D Images

Raphael Prevost, Benoit Mory, Remi Cuingnet, Jean-Michel Correas, Laurent D. Cohen, and Roberto Ardon

Abstract Contrast-enhanced ultrasound (CEUS) imaging has lately benefited of an increasing interest for diagnosis and intervention planning, as it allows to visualize blood flow in real-time harmlessly for the patient. It complements thus the anatomical information provided by conventional ultrasound (US). This chapter is dedicated to kidney segmentation methods in 3D CEUS images. First we present a generic and fast two-step approach to locate (via a robust ellipsoid estimation algorithm) and segment (using a template deformation framework) the kidney automatically. Then we show how user interactions can be integrated within the algorithm to guide or correct the segmentation in real-time. Finally, we develop a co-segmentation framework that generalizes the aforementioned method and allows the simultaneous use of multiple images (here the CEUS and the US images) to improve the segmentation result. The different approaches are evaluated on a clinical database of 64 volumes.

Introduction

Ultrasound imaging (US) is a widely used modality due to its versatility, low cost, and real-time capabilities. Such acquisitions have been for a long time limited to 2D images but the recent development of 3D US allowed to consider new problems

R. Prevost (✉) • B. Mory • R. Cuingnet • R. Ardon
Philips Research Medisys, Suresnes, France
e-mail: prevost.raphael@gmail.com; benoit.mory@philips.com; remi.cuingnet@philips.com; roberto.ardon@philips.com

J.-M. Correas
Adult Radiology Department, Necker Hospital, Paris, France
e-mail: jean-michel.correas@nck.aphp.fr

L.D. Cohen
Paris Dauphine University, CEREMADE UMR 7534, Paris, France
e-mail: cohen@ceremade.dauphine.fr

such as volumetric assessments of organs or image registration. In addition to conventional US, three-dimensional real-time visualization of vascularization can be achieved with contrast-enhanced ultrasound (CEUS) imaging. This rather new modality provides very useful information for lesions diagnosis or large vessels monitoring [1]. Gas-filled microbubbles, acting as amplifiers of the blood back-scattering signal, are used as a contrast agent. Because the bubbles are naturally eliminated by metabolism processes, this modality is considered as completely safe for the patients even with renal or liver failure (unlike contrast-enhanced CT, for example).

However the usually poor quality of CEUS images makes any computer-based analysis challenging: in addition to having powerful speckle noise, the image is very grainy and almost binary as a result of ultrasound interactions with individual bubbles. Unlike in conventional US [2], very few segmentation methods of 3D CEUS images have been reported. Among them, Gasnier et al. [3] introduced an interactive approach to segment and analyze tumors in this modality. However, their framework was specific to lesion segmentation, just as the automatic methods proposed in [4, 5]. In [6], Ma et al. developed an automatic algorithm to segment the heart left ventricle. This method, although applicable to other organs, does not provide any natural way to refine or correct the result interactively. Besides, it has been designed for images acquired with a particular transducer, producing sparse rotated slices instead of a whole 3D volume.

In this chapter, we address the problem of kidney segmentation in 3D CEUS images. This challenging issue is of great importance to assess quantitatively the volume of renal tissues. First, we present a generic and fast approach to automatically segment a kidney in CEUS volumes. Our method consists in detecting it in the image as an ellipsoid, and then deforming this ellipsoid to match precisely its boundary. Second, we extend this framework in order to take into account other kinds of information :

- *user interactions:* Because of the poor image quality or pathologies, image information may be sometimes unreliable and even misleading. In such cases, the clinician user should be able to guide or correct the segmentation easily and with a real-time feedback.
- *simultaneous use of another image:* Because of shadowing effects, pathologies, and limited field of view, parts of the kidney may be hardly visible in the image. In such cases even expert users may have difficulty delineating the true boundary of the organ by solely relying on one CEUS image. In clinical routine every CEUS acquisition is preceded by a conventional US acquisition to locate the kidney. Hence, the latter would be useful to complement the CEUS image and thus cope with missing and corrupted information.

Prior work on kidney segmentation in CEUS is limited to two of our conference papers [7, 8], of which this chapter is an extended version.

The remainder of the chapter is organized as follows. First of all, section "Material" is dedicated to the description of the material used throughout the chapter in validation experiments. In section "Kidney Detection via Robust Ellipsoid
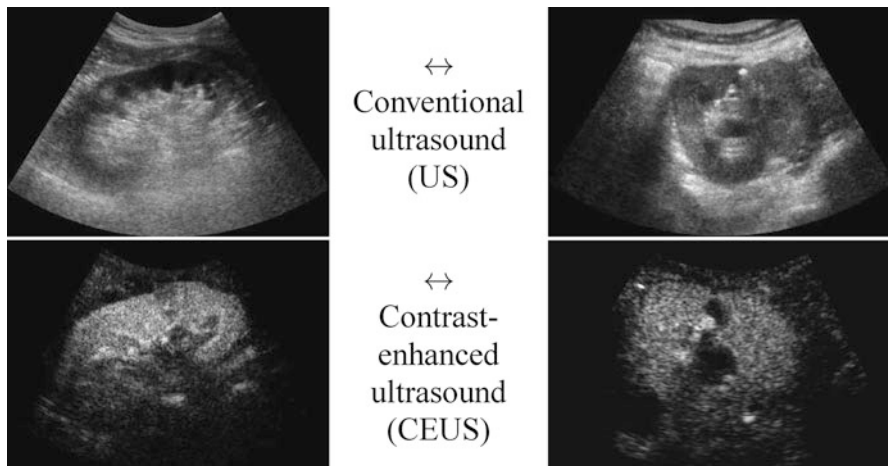
Fig. 1 Slices of conventional and contrast-enhanced ultrasound 3D images of the kidney for two different patients (*left* and *right*)

Estimation", we introduce a fast and robust method to estimate roughly the center, orientation, and sizes of the kidney. This is a done via an original variational framework for ellipsoid detection. The outcome of this step is then used as the prior model of a segmentation algorithm, based on template deformation, described in section "Kidney Segmentation via Implicit Template Deformation". Because of the inherent ambiguities in CEUS images, the obtained segmentation may be improved by using additional information. In section "Segmentation with User Interactions", we show how user interactions can be used inherently in our framework to correct the result in real-time. Then we extend our approach to multiple images, namely the CEUS and the US volumes (section "Joint Co-segmentation and Registration") which are not aligned. Thus a generic framework for joint co-segmentation and registration is introduced and applied to both the kidney detection and segmentation. We show that by taking additional information into account, the automatic kidney segmentation is more robust. Finally, we conclude the chapter by discussing potential improvements.

Material

This section describes the material used throughout the chapter. Our database is composed of 64 pairs of CEUS and US volumes acquired from 35 different patients, via an iU22 ultrasound system (Philips, The Netherlands). In order to have a clinically representative database, both healthy and diseased kidneys were considered. Images were acquired using different probes, namely V6-2 and X6-1 (Philips, The Netherlands) US probes, with various fields of view. The volumes size was

$512 \times 510 \times 256$ voxels with different spatial resolutions ($0.25 \times 0.25 \times 0.55$ mm in average). The acquisition protocol was as follows: first, the clinician scouted for the patient's kidney using conventional US and acquired a US volume. Then, 2.4 mL of Sonovue (Bracco, Italy) contrast agent was injected to the patient and a CEUS acquisition was performed after a few seconds. Indeed, dynamic CEUS images of a kidney show a cortical enhancement shortly followed by a medullary enhancement. Better visualization of kidney tissue is then available when the contrast agent has diffused as it is completely hyperechoic whereas its fatty surrounding produces no signal. Figure 1 shows a comparison of US and CEUS images for two patients of our database. Note that the US and CEUS images are not aligned as the clinician may have slightly moved the probe between the two acquisitions.

For each image, an expert was asked to segment the kidney with a semiautomatic tool. This segmentation was considered as the ground truth. The different approaches described in the chapter will be evaluated by computing the Dice coefficient between the segmentation result S and the ground truth GT, defined as

$$\mathrm{Dice}(S, GT) = 2 \; \frac{\mathrm{Vol}(S \cap GT)}{\mathrm{Vol}(S) + \mathrm{Vol}(GT)} \; , \tag{1}$$

where $\mathrm{Vol}(X)$ denotes the volume of a region X. Thus the higher the Dice coefficient, the better the segmentation is. In particular, this score is equal to 1 for a perfect segmentation and 0 for a completely non-overlapping segmentation.

All proposed methods were implemented in a C++ prototype and the computational times will be given for a standard computer (Intel Core i5 2.67 Ghz, 4GB RAM).

Kidney Detection via Robust Ellipsoid Estimation

Since kidney shape can be roughly approximated by an ellipsoid, the kidney automatic detection problem in CEUS images can be initially reduced to finding the smallest ellipsoid encompassing most of the hyperechoic voxels. A large number of methods (e.g., Hough transforms [9, 10]) have already been proposed to detect ellipses in images [11]. However their extension to 3D, though possible, is usually computationally expensive mainly because of the number of parameters to estimate (9 for a 3D ellipsoid). Furthermore, they do not explicitly use the fact that only one ellipsoid is present in the image. On the other hand, statistical approaches like robust Minimum Volume Ellipsoid (MVE) estimators [12] are better suited but require prior knowledge on the proportion of outliers (here the noise, artifacts, or neighboring structures), which may vary from one image to another and is thus not available. We therefore propose an original variational framework, which is robust and fast, to estimate the best ellipsoid in an image $I : \Omega \subset \mathbb{R}^3 \to \mathbb{R}^+$.

A Variational Framework for Robust Ellipsoid Estimation

In the considered framework, an ellipsoid is implicitly represented using an implicit function $\phi : \Omega \to \mathbb{R}$ that is positive inside the ellipsoid and negative elsewhere. ϕ can be parametrized by the center of the ellipsoid $\mathbf{c} \in \mathbb{R}^3$ and its sizes and orientations encoded by a 3×3 positive-definite matrix $\mathbf{M}$. We therefore define the implicit equation of an ellipsoid as

$$\phi_{\mathbf{c},\mathbf{M}}(\mathbf{x}) = 1 - (\mathbf{x} - \mathbf{c})^T \mathbf{M}(\mathbf{x} - \mathbf{c}) = 0 \ . \tag{2}$$

The detection method should be robust to outliers, i.e. bright voxels coming from noise, artifacts, or other neighboring structures. Excluding those outliers is done by estimating a weighting function w (defined over the image domain Ω into [0,1]) that provides a confidence score for any point $\mathbf{x}$ to be an inlier. The ellipsoid estimation is then formulated as an energy minimization problem with respect to $\mathbf{c}$, $\mathbf{M}$, and w:

$$\min_{\mathbf{c},\mathbf{M},w} \left\{ E_{\det}(\mathbf{c}, \mathbf{M}, w) = - \int_\Omega \phi_{\mathbf{c},\mathbf{M}}(\mathbf{x}) \, I(\mathbf{x}) \, w(\mathbf{x}) \, d\mathbf{x} \right.$$

$$\left. + \mu . \log\left(\frac{\mathrm{Vol}(\mathbf{M})}{|\Omega|}\right) . \left(\int_\Omega I(\mathbf{x}) \, w(\mathbf{x}) \, d\mathbf{x}\right) \right\} \tag{3}$$

$$\text{with} \quad \phi_{\mathbf{c},\mathbf{M}}(\mathbf{x}) = 1 - (\mathbf{x} - \mathbf{c})^T \mathbf{M}(\mathbf{x} - \mathbf{c})$$

$$\text{and} \quad \mathrm{Vol}(\mathbf{M}) = \frac{4\pi}{3}\sqrt{\det \mathbf{M}^{-1}} \quad \text{the ellipsoid volume.}$$

The ellipsoid detection energy $E_{\det}$ is composed of two terms:

- a *data-fidelity term*: The first term is an integral over the whole image domain Ω of the product $\phi_{\mathbf{c},\mathbf{M}}$ by $w\,I$. Note that $w\,I$ is highly positive at voxels that have a high intensity but are not outliers. To minimize the energy, such voxels must therefore be included inside the ellipsoid, i.e. where ϕ is positive.
- a *regularization term*: The second term penalizes the volume of the ellipsoid Vol $(\mathbf{M})$ with respect to the domain volume $|\Omega|$. The logarithm provides a statistical interpretation of the problem and eases the minimization of the energy, as will be seen in the next subsection. It is normalized by $\int w\,I$ and weighted by a trade-off parameter $\mu > 0$.

Numerical Optimization

This ellipsoid estimation process can be viewed as fitting a Gaussian distribution to the bright pixels of the image by minimizing its negative log-likelihood. Therefore $E_{\det}$ has a statistical meaning and when w is fixed, the minimizers $(\mathbf{c}^*,\mathbf{M}^*)$ of $E_{\det}$

$(\cdot, \cdot, w)$ have a closed form. Indeed, $\mathbf{c}^*$ is the barycenter of all voxels $\mathbf{x}$ weighted by $I(\mathbf{x})w(\mathbf{x})$, while $\mathbf{M}^*$ is the inverse of the covariance matrix[1] of the same data. Besides, E_{det} is linear with respect to w which is by definition restricted to [0,1]. Therefore, at every voxel $\mathbf{x}$ the minimizer $w^*(\mathbf{x})$ is equal to 0 or 1, depending only on the sign of $\phi_{\mathbf{c},\mathbf{M}} - \mu \log\left(\frac{\mathrm{Vol}(\mathbf{M})}{|\Omega|}\right)$. w^* is then the indicator of the current ellipsoid estimation which has been dilated proportionately to μ. Its purpose is to remove the contribution of the points which are far away from the current ellipsoid and may hinder its refinement.

The weighting function w is initialized to 1 everywhere. Minimization of E_{det} is then performed with an alternate iterative scheme that successively updates the variables $\mathbf{c}$, $\mathbf{M}$, and w, as summarized in Algorithm 1. As the energy E_{det} decreases at each step, the algorithm is guaranteed to converge. In practice, few iterations are required for convergence and total computational time is less than a second for a 3D image.

Algorithm 1: Robust ellipsoid detection algorithm

---

initialization $\forall\, \mathbf{x} \in \Omega,\, w(\mathbf{x}) \leftarrow 1$
repeat
  // Estimation of center $\mathbf{c}$ and matrix $\mathbf{M}$
  $\mathbf{c} \leftarrow \frac{1}{\int_\Omega Iw} \int_\Omega I(\mathbf{x})\, w(\mathbf{x})\, \mathbf{x}\, d\mathbf{x}$
  $\mathbf{M}^{-1} \leftarrow \frac{2}{\mu \int_\Omega Iw} \int_\Omega I(\mathbf{x})\, w(\mathbf{x})\, (\mathbf{x} - \mathbf{c})\, (\mathbf{x} - \mathbf{c})^T\, d\mathbf{x}$
  // Update of the weighting function w for each $\mathbf{x} \in \Omega$
  if $(\mathbf{x} - \mathbf{c})^T \mathbf{M}\, (\mathbf{x} - \mathbf{c}) \leq 1 - \mu \log\left(\frac{Vol(\mathbf{M})}{|\Omega|}\right)$ **then**
    $w(\mathbf{x}) \leftarrow 1$
  else
    $w(\mathbf{x}) \leftarrow 0$
until *convergence*;

The choice of μ is paramount as it controls the number of points that are taken into account for the ellipsoid matrix estimation. It should be set to values close to $\frac{2}{5}$ in 3D and $\frac{1}{2}$ in 2D (the proof is deferred in the appendix).

Figure 2 shows such a process for a synthetic 2D image. The first ellipse estimate is too large as all voxels are considered but far points are progressively eliminated via the weighting function w until the algorithm converges towards the good solution. We also present results on real CEUS data in Fig. 3. The estimated ellipsoids are not perfectly accurate but robust and close enough to be used as an initialization for a segmentation algorithm.
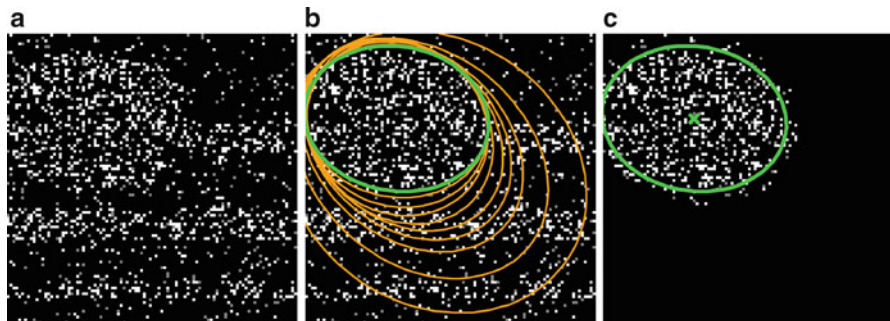
---

[1] Up to a constant multiplier.

Fig. 2 (**a**) Original 2D synthetic image, corrupted by salt-and-pepper noise. (**b**) Evolution of the ellipse along the iterations (*orange*) and final result (*green*). (**c**) Ellipse contour and center superimposed on the product $w\,I$ at convergence
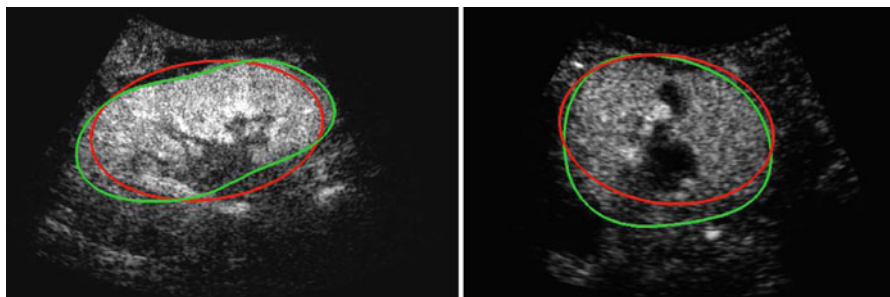

Fig. 3 Results of the ellipsoid detection (*red*) compared to the ground truth (*green*), on slices of the volumes shown in Fig. 1

Kidney Segmentation via Implicit Template Deformation

The previously detected ellipsoid will now be deformed to segment the kidney more precisely. We follow the template deformation framework described in [13, 14] and extended in [15], as it is a very efficient model-based algorithm and it has already been applied successfully to kidney segmentation in CT images [16].

Implicit Template Deformation Framework

Implicit template deformation is a framework where an implicit shape defined by a function $\phi_0 : \Omega \to \mathbb{R}$, called the *template*, is deformed so that its zero level set segments a given image $I : \Omega \to \mathbb{R}^+$. The segmenting implicit shape is the zero level set of a function $\phi : \Omega \to \mathbb{R}$, therefore defined with respect to this template

and a transformation of the space $\psi : \Omega \rightarrow \Omega$ that becomes the unknown of the problem : $\phi = \phi_0 \circ \psi$. In our application, the template is the implicit function of the previously estimated ellipsoid $\phi_0 = \phi_{\mathbf{c}}{}^{*}, \mathbf{M}^{*}$ and ψ is sought such that the image gradient flux across the surface of the deformed ellipsoid $(\phi_0 \circ \psi)^{-1}(0)$ is maximum. The segmentation energy is then

$$E_{\text{seg}}(\psi) = \int_{\{\phi_0 \circ \psi = 0\}} -\left\langle \vec{\nabla} I(\mathbf{x}) \, , \, \vec{\mathbf{n}}(\mathbf{x}) \right\rangle \, dS(\mathbf{x}) + \lambda \mathcal{R}(\psi), \qquad (4)$$

where $\vec{\mathbf{n}}(\mathbf{x})$ denotes the vector normal to the surface of the segmentation at point $\mathbf{x}$. $\mathcal{R}(\psi)$ is a regularization term which prevents large deviations from the original ellipsoid. Its choice will be detailed in section "Transformation Model" hereafter. λ is a positive scalar parameter that controls the strength of this shape constraint.

Using the divergence theorem, the first data-fidelity term can be rewritten as

$$\int_{\{\phi_0 \circ \psi = 0\}} -\left\langle \vec{\nabla} I(\mathbf{x}), \, \vec{\mathbf{n}}(\mathbf{x}) \right\rangle \, dS(\mathbf{x}) = - \int_{\{\phi_0 \circ \psi \geq 0\}} div(\nabla I(\mathbf{x})) \, d\mathbf{x} = - \int_{\{\phi_0 \circ \psi \geq 0\}} \Delta I(\mathbf{x}) \, d\mathbf{x} \quad (5)$$

where Δ denotes the Laplacian operator. Introducing H the Heaviside function ($H(a) = 1$ if a is positive, 0 otherwise) yields a more convenient formulation of the segmentation energy :

$$E_{\text{seg}}(\psi) = - \int_{\Omega} H(\phi_0 \circ \psi(\mathbf{x})) \, \Delta I(\mathbf{x}) \, d\mathbf{x} + \lambda \mathcal{R}(\psi), \qquad (6)$$

Transformation Model

The choice of the space of possible solutions ψ to Problem (6) is, in our case, intrinsically linked to the notion of *shape*. A *shape* can be considered as a set of objects sharing the same visual aspect. It should be invariant to geometric transforms such as translation, rotation, scaling, or shearing. We will refer to such a global transformation as the *pose*. To set up a clear distinction between the pose and the subsequent shape *deformation*, similarly to [17], we design our template transformation model ψ as a functional composition of a global transformation $\mathcal{G}$ and a nonrigid local transformation $\mathcal{L}$ (see Fig. 4):

$$\psi = \mathcal{L} \circ \mathcal{G} \qquad (7)$$

Pose. $\mathcal{G} : \Omega \rightarrow \Omega$ is chosen as a parametric transform that coarsely aligns the template with the target surface in the image. It will basically correct or adjust the global position and scaling of the ellipsoid and can be chosen as a similarity. $\mathcal{G}$ is thus represented by a matrix in homogeneous coordinates defined by 7 parameters $\mathbf{p} = \{p_i\}_{i=1\ldots7}$ and noted $\mathcal{G}_{\mathbf{p}}$.
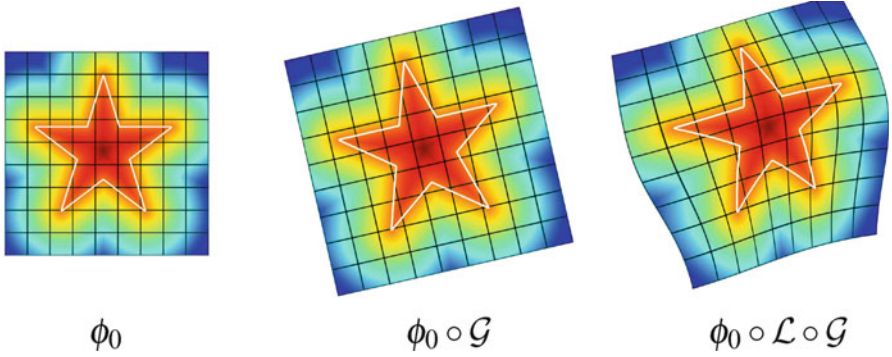
Fig. 4 Decomposition of the transformation ψ. The implicit template ϕ_0 undergoes a global transformation $\mathcal{G}$ and a local deformation $\mathcal{L}$

Deformation. $\mathcal{L} : \Omega \rightarrow \Omega$ is expressed using a displacement field $\mathbf{u}$ in the template referential $\mathcal{L} = Id + \mathbf{u}$. Similarly to problems in image registration and optical flow algorithms [18], $\mathbf{u}$ should be smoothly varying in space. While adding penalizations on differential terms of $\mathbf{u}$ to $\mathcal{R}(\psi)$ is a valid approach, efficient implementations are difficult to derive. Taking advantage of efficient linear filtering, smoothness of the displacement $\mathbf{u}$ is set as a built-in property defining it as a filtered version of an integrable unknown displacement field $\mathbf{v}$

$$\mathbf{u}(\mathbf{x}) = [K_\sigma * \mathbf{v}](\mathbf{x}) = \int_\Omega K_\sigma(\mathbf{x} - \mathbf{y}) \, \mathbf{v}(\mathbf{y}) \, d\mathbf{y} \tag{8}$$

where K_σ is a Gaussian kernel of scale σ. The overall transformation that can therefore be parametrized by $\mathbf{p}$ and $\mathbf{v}$ will be noted $\psi_{\mathbf{p},\mathbf{v}}$.

The proposed decomposition allows to define the shape prior term independently from the pose: $\mathcal{R}(\psi) = \mathcal{R}(\mathcal{L})$. $\mathcal{R}$ thus quantifies how much the segmenting implicit function ϕ deviates from the prior *shape* ϕ_0. Using the L_2 norm we choose to constraint $\mathcal{L}$ towards the identity :

$$\mathcal{R}(\mathcal{L}) = \frac{1}{2} \|\mathcal{L} - Id\|_2^2 = \frac{1}{2} \int_\Omega \|\mathbf{u}(\mathbf{x})\|^2 \, d\mathbf{x} \tag{9}$$

The optimization problem to solve finally reads:

$$\min_{\mathbf{p},\mathbf{v}} \left\{ E_{\text{seg}}(\psi_{\mathbf{p},\mathbf{v}}) = - \int_\Omega H(\phi_0 \circ \psi_{\mathbf{p},\mathbf{v}}(\mathbf{x})) \, \Delta I(\mathbf{x}) \, d\mathbf{x} + \frac{\lambda}{2} \int_\Omega \|K_\sigma * \mathbf{v}\|^2 \right\} \tag{10}$$

with $\psi_{\mathbf{p},\mathbf{v}} = (Id + \mathbf{u}) \circ \mathcal{G}_{\mathbf{p}}$ and $\mathbf{u} = K_\sigma * \mathbf{v}$

surface/pointwise forces   coarse-to-fine $\phi_0 \circ \mathcal{L}$ convolved deformation
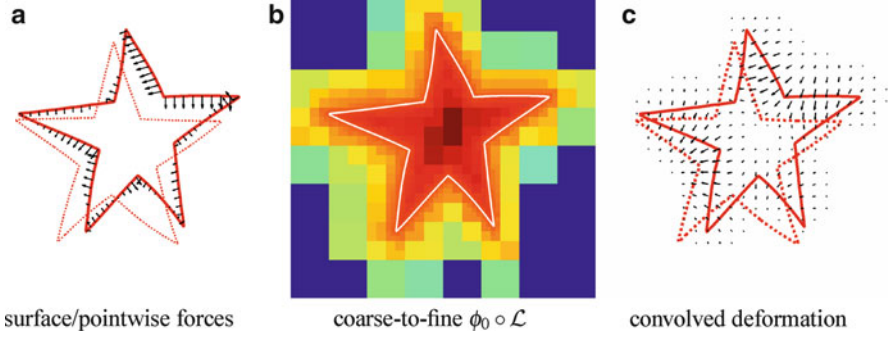
Fig. 5 Fast template deformation with coarse-to-fine distance warp and convolutions. (**a**) Surface/pointwise forces. (**b**) Coarse-to-fine $\phi_0 \circ \mathcal{L}$. (**c**) Convolved deformation

Numerical Implementation

Problem (10) is minimized via a standard gradient descent simultaneously on the parameters of the pose $\mathcal{G}_p$ and the deformation field $\mathbf{v}$. The descent evolution equations are obtained by applying calculus of variations to E_{seg}. We omit the tedious details but the final equations, after a variable substitution, read

$$
\begin{cases}
\dfrac{\partial \mathbf{p}}{\partial t} = -\displaystyle\int_{\Omega} \delta(\phi_0 \circ \mathcal{L}) \, . \, \left\langle \nabla \phi_0 \circ \mathcal{L}, \ (Id + J_{\mathbf{u}}) \dfrac{\partial \mathcal{G}}{\partial \mathbf{p}} \mathcal{G}^{-1} \right\rangle \, . \, \Delta I \circ \mathcal{G}^{-1} \\[2mm]
\dfrac{\partial \mathbf{v}}{\partial t} = - \left[\delta(\phi_0 \circ \mathcal{L}) \, . \, \nabla \phi_0 \circ \mathcal{L} \, . \, \Delta I \circ \mathcal{G}^{-1} \ \ + \lambda \mathbf{v} \right] * K_\sigma
\end{cases}
\tag{11}
$$

where δ denotes the Dirac distribution and $J_{\mathbf{u}}$ is the Jacobian matrix of the displacement field $\mathbf{u}$.

A quick analysis of Eq. (11) reveals several key aspects for an efficient implementation. Interpolating $\phi_0 \circ \mathcal{L}$ and $\nabla \phi_0 \circ \mathcal{L}$ over the whole domain Ω would be extremely time-consuming. Nevertheless, since it is multiplied by $\delta(\phi_0 \circ \mathcal{L})$, the warped gradient field $\nabla \phi_0 \circ \mathcal{L}$ is only needed on the set $\{\phi_0 \circ \mathcal{L} = 0\}$ (Fig. 5a) which highly reduces the computational burden. Moreover, precise knowledge of the warped template $\phi_0 \circ \mathcal{L}$ is only necessary near its zero level set. We use a coarse-to-fine approach using octrees. At each level a decision is made to further refine the cell depending on the distance measure (Fig. 5b) drastically dropping complexity. Finally, stemming from the displacement model, extrapolating image and pointwise forces to the whole space boils down to a convolution with K_σ (Fig. 5c). In practice, our current 3D implementation supports up to 100 time steps per second for a discretization of the implicit function on a $64 \times 64 \times 64$ lattice.
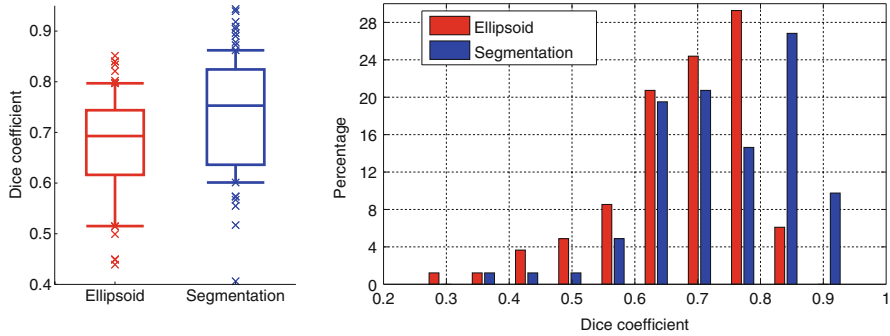
Fig. 6 Kidney detection (*red*) and segmentation (*blue*) results in terms of Dice coefficients shown as boxplots (*left*) and histograms (*right*). Boxplots show, respectively, the first decile, the first quartile, the median, the third quartile, and the ninth decile. Extreme points are shown separately

Results for Automatic Segmentation in CEUS Images

This validation has been performed on the CEUS images of the dataset presented in section "Material". The completely automatic pipeline had a computational time of around 5 s.

Quantitative results are reported in Fig. 6. The overall median Dice coefficient is 0.69 for the detection and 0.76 for the segmentation and 25 % of the database have a very satisfying segmentation (Dice coefficient higher than 0.85), given the very poor image quality and the presence of pathologies.

Figure 7 shows the obtained result for the two cases introduced in Fig. 1. The segmentations are very similar to the ground truth and can be considered as satisfying. Some cases are, however, more difficult (e.g., Fig. 10 in the next section) and will require additional information.

Segmentation with User Interactions

The previously described approach is fast and automatic, but fails in some difficult cases. Indeed ultrasound shadows or kidney pathologies make the image information unreliable and thus hinder the segmentation algorithm. It is therefore important to provide the clinician a way to guide or correct the segmentation easily and with a real-time feedback. As proposed in [15], this can be done easily within the implicit template deformation framework that was presented in section "Kidney Segmentation via Implicit Template Deformation".
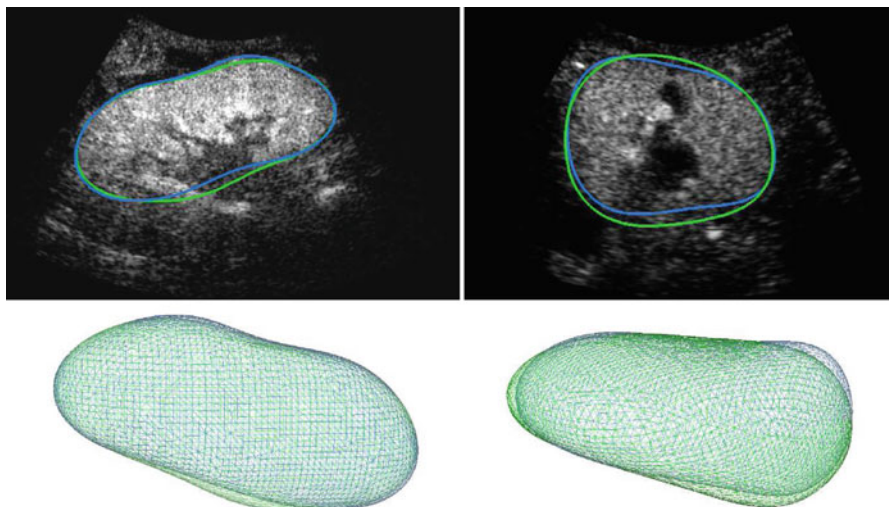
Fig. 7 Result of the automatic segmentation (*blue*) compared to the ground truth (*green*) on a particular slice (*top*) and in 3D (*bottom*)
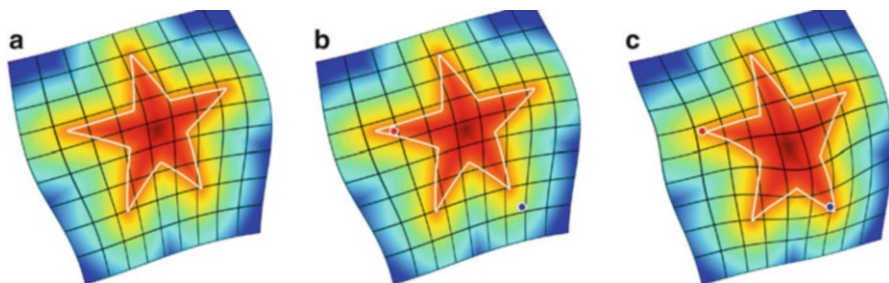

Fig. 8 User interactions as inside/outside points. (**a**) Template deformed without constraints. (**b**) User indicates points that should be inside (*blue*) and outside (*red*) the segmentation. (**c**) New segmentation that satisfies these constraints

User Interactions as Constraints

In this section, we show how the user can guide the segmentation by indicating points that should be inside or outside the segmentation (see Fig. 8).

Consider that the user provides N points $\{\mathbf{x}_k\}_k \subset \Omega^N$ in the image domain labeling each one as inside or outside of the surface to extract (which can be done via simple interactions such as a left click on an inside point, and a right click on an outside point). The implicit formulation allows to express this

information merely as inequality constraints on the deformed implicit function, at points $\{\mathbf{x}_k\}_k$:

$$\forall k \in [|1,N|], \quad \gamma_k \, . \, \phi_0 \circ \psi(\mathbf{x}_k) \geq 0 \tag{12}$$

where $\gamma_k = 1$ (resp. -1) for inside (resp. outside) points. Note that it is also possible to specify a point that should be exactly on the segmentation surface by labeling it as both inside and outside: the two inequality constraints are equivalent to an equality constraint. Then, putting together the initial formulation in Eq. (6) and the constraints of Eq. (12) yields a general formulation of implicit template deformation with user interactions, as the following minimization problem:

$$\min_{\psi} \quad \left\{ E_{\text{seg}}(\psi) = - \int_{\Omega} H(\phi_0 \circ \psi(\mathbf{x})) \, \Delta I(\mathbf{x}) \, d\mathbf{x} + \lambda \mathcal{R}(\psi) \right\}$$
$$\text{subject to } \forall k \in [1,N], \quad \gamma_k \, . \, \phi_0 \circ \psi(\mathbf{x}_k) \geq 0 \tag{13}$$

In the next subsection we propose a method to solve this problem efficiently. For the sake of genericity, no assumption is made on the representation of the deformation ψ and the model $\psi = \mathcal{L} \circ \mathcal{G}$ will be just a particular implementation of the approach described hereafter.

Optimization Scheme

Since $E_{\text{seg}}(\psi)$ is a non-convex functional and has to be minimized under a set of nonlinear constraints, no specifically tailored algorithms are available. For this matter, we follow a general augmented Lagrangian methodology [19] and define an equivalent unconstrained problem that can be locally minimized by gradient descent. The constrained problem (13) can equivalently be written as an unconstrained minimization problem of the form

$$\min_{\psi} \left\{ \tilde{E}_{\text{seg}}(\psi) = \max_{\alpha \geq 0} \left\{ E_{\text{seg}}(\psi) - \sum_{k=1}^{N} \alpha_k c_k(\psi) \right\} \right\}$$
$$\text{with } c_k(\psi) = \gamma_k \, . \, \phi_0 \circ \psi(\mathbf{x}_k) \tag{14}$$

where α_k is the Lagrange multiplier associated with the kth constraint. Equation (14) has the same set of solutions as the original problem in Eq. (13): if ψ satisfies all constraints c_k, then $\tilde{E}_{\text{seg}}(\psi) = E_{\text{seg}}(\psi)$; otherwise, $\tilde{E}_{\text{seg}}(\psi)$ is infinite. Since $\tilde{E}_{\text{seg}}$ jumps from finite to infinite values at the boundary of the feasible set, it is difficult to minimize it as such. A more practical approach is to introduce a smooth

approximation $\tilde{E}^{\nu}_{\text{seg}}$ that depends on a quadratic penalty parameter ν. Parameter ν will be used to constrain the maximizers $(\alpha_k)_k$ to finite values. These multipliers are estimated iteratively and we introduce $(\alpha_k^j)_k$ the multipliers estimates at the jth iteration, in order to define the energy approximation

$$\tilde{E}^{\nu}_{\text{seg}}(\psi, \alpha^j) = \max_{\alpha \geq 0} \left\{ E_{\text{seg}}(\psi) - \sum_{k=1}^{N} \alpha_k c_k(\psi) - \frac{1}{2\nu} \sum_{k=1}^{N} \left(\alpha_k - \alpha_k^j \right)^2 \right\} \qquad (15)$$

The maximizing Lagrange multipliers associated with each constraint $c_k(\psi)$ have a closed-form solution :

$$\alpha_k^{j+1} = \begin{cases} 0 & \text{if } \alpha_k^j - \nu c_k(\psi) \leq 0 \\ \alpha_k^j - \nu c_k(\psi) & \text{otherwise.} \end{cases} \qquad (16)$$

Substituting (16) into (15) yields the following expression of the smooth approximation $\tilde{E}^{\nu}_{\text{seg}}$:

$$\tilde{E}^{\nu}_{\text{seg}}(\psi, \alpha^j) = E_{\text{seg}}(\psi) + \sum_{k=1}^{N} F_{\nu}\left(c_k(\psi), \alpha_k^j \right)$$

$$\text{with} \quad F_{\nu}(a, b) = \begin{cases} -ab + \dfrac{\nu}{2} a^2 & \text{if } \nu a \leq b \\ -\dfrac{1}{2\nu} b^2 & \text{otherwise.} \end{cases} \qquad (17)$$

Finally, the alternate scheme described in Algorithm 2, in which the penalty parameter ν is gradually increased, will provide a local minimizer of E_{seg} that eventually satisfies the user constraints. Within this process, Step (1) is straightforward and Step (2) is very similar to the gradient descent proposed in section "Numerical Implementation":

$$\begin{cases} \dfrac{\partial \mathbf{p}}{\partial t} \leftarrow \dfrac{\partial \mathbf{p}}{\partial t} - \sum_{k=1}^{K} \gamma_k F(\alpha_k) \left\langle \nabla \phi_0 \circ \mathcal{L} \circ \mathcal{G}(\mathbf{x}_k), \ (Id + J_{\mathbf{u}}) \dfrac{\partial \mathcal{G}}{\partial \mathbf{p}}(\mathbf{x}_k) \right\rangle \\ \\ \dfrac{\partial \mathbf{v}}{\partial t} \leftarrow \dfrac{\partial \mathbf{v}}{\partial t} - \left[\sum_{k=1}^{K} \gamma_k \delta_{\mathcal{G}(\mathbf{x}_k)} \ F_{\nu}(\alpha_k) \nabla \phi_0 \circ \mathcal{L} \right] * K_{\sigma} \end{cases} \qquad (18)$$

Note that the additional terms in Eq. (18) are just pointwise contributions that do not influence the overall computational time.

Algorithm 2: Augmented Lagrangian Scheme For Inequality Constraints

---

initialization choose $v^0 > 0$ and set $\forall k,\ \alpha_k^0 \leftarrow 0$,

repeat

    choose $v^t > v^{t-1}$,

    set $j \leftarrow 0$,

    repeat

        (1) ψ being fixed, update the Lagrange multipliers α^j using Eq (16)

        (2) α^j being fixed, update ψ by minimizing $\tilde{E}_{seg}^{v^t}(\psi, \alpha^j)$ with gradient descent on Eq (17)

        (3) increment $j \leftarrow j + 1$

    until *convergence*;

    increment $t \leftarrow t + 1$

until *a local minimum of $E_{seg}(\psi)$ satisfying $\forall k,\ c_k(\psi) \geq 0$ is found*;

Influence of User Interactions on Kidney Segmentation in CEUS

Validation of the user interactions has been performed on a subset of 21 CEUS volumes from 21 different patients of our database. For each case, the automatic segmentation has been run and its result was refined with user interactions from an expert. Figure 9 reports the Dice coefficients obtained as a function of the number of clicks. The score gradually increases as the user interacts with the algorithm but rapidly converges: most of the time, less than 3 clicks are needed for a fairly precise result (Dice $\geq$ 0.9).[2] The results also show that even when the initialization produces a low score, very few interactions can improve a lot the segmentation. The influence of user interactions is illustrated in Fig. 10, where we show results on a difficult case. The patient has a lot of renal cysts that are anechogenic and hinders the automatic segmentations. With 3 clicks, the segmentation is much closer to the ground truth.

Nevertheless, in some applications user interactions are not possible and the segmentation has to be automatic. In the next section, we propose to improve the kidney segmentation by using simultaneously and automatically the conventional US image

---

[2] The ground truth may not exactly be reached because of the high intra-operator variability.
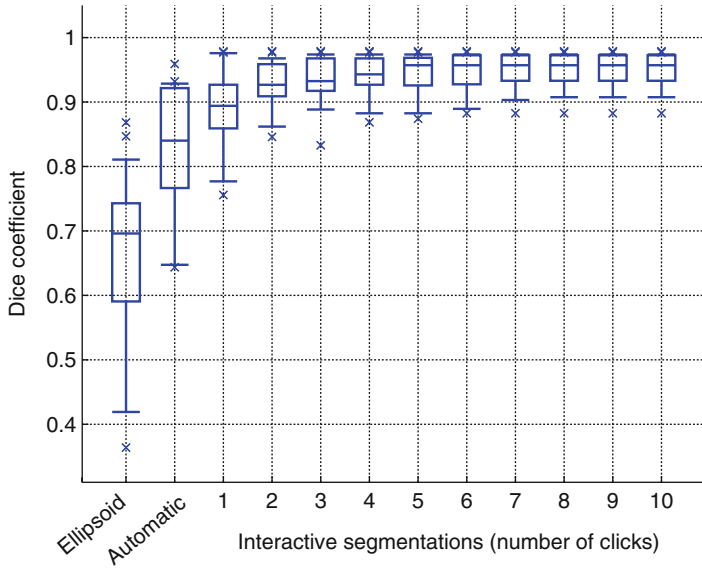
Fig. 9 Boxplots of the Dice coefficient between the ground truth and the segmentation at different steps of the proposed algorithm. Boxplots show, respectively, the first decile, the first quartile, the median, the third quartile, and the ninth decile. Extreme points are shown separately
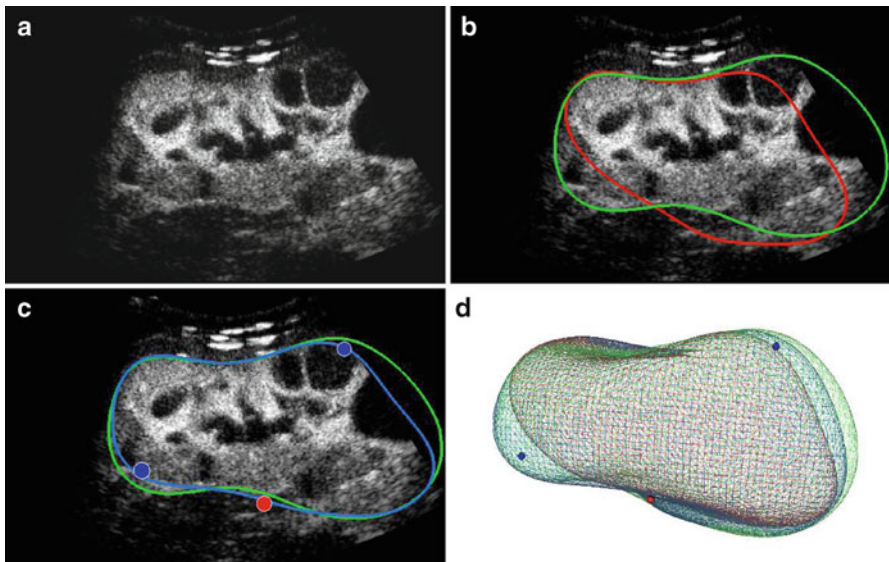

Fig. 10 Example of a segmentation with user interactions. (**a**) Slice of the original CEUS volume. (**b**) Comparison of the ground truth (*green*) and automatic segmentation (*red*). (**c**) Corrected segmentation (*blue*) with 2 inside points (*blue dots*) and one outside point (*red dot*). (**d**) 3D visualization of the ground truth (*green*), the automatic (*red*), and corrected (*blue*) segmentation, with constraint points

Joint Co-segmentation and Registration

Co-segmentation often denotes the task of finding an object in each image that shares the same appearance but not necessarily the same shape [20]. Here we look for the exactly same organ in two images but with a different appearance. As simultaneous acquisition of US and CEUS is not possible on current 3D imaging systems, the two images are in arbitrary referentials and need to be aligned. However, standard iconic registration methods are not adapted since visible structures, apart from the kidney itself, are completely different in US and CEUS. Co-segmentation shall therefore help registration, just as registration helps co-segmentation. This calls for a method that jointly performs these two tasks (see Fig. 11).

Although segmentation and registration are often seen as two separate problems, several approaches have already been proposed to perform them simultaneously. Most of them rely on an iconic registration guiding the segmentation (e.g., [21–23]). Yet they assume that the segmentation is known in one of the images, which is not the case in our application of co-segmentation. Moreover, as stated before, CEUS/US intensity-based registration is bound to fail since visible structures do not correspond to each other. Instead of registering the images themselves, Wyatt et al. [24] developed a MAP formulation to perform registration on label maps resulting from a segmentation step. However no shape model is enforced and noise can degrade the results. In [25], Yezzi et al. introduced a variational framework that consists in a feature-based registration in which the features are actually the segmenting active contours.

In this section, we aim at extending both the previously described kidney detection and segmentation in a 3D CEUS image to a pair of 3D CEUS and US images. To that end, we develop a generic joint co-segmentation and registration framework inspired by [25]. This results in a fully automated pipeline to obtain both an improved kidney segmentation in CEUS and US images and a registration of them. But first of all, in order to use conventional US, we need to learn how the kidney looks like in such images.
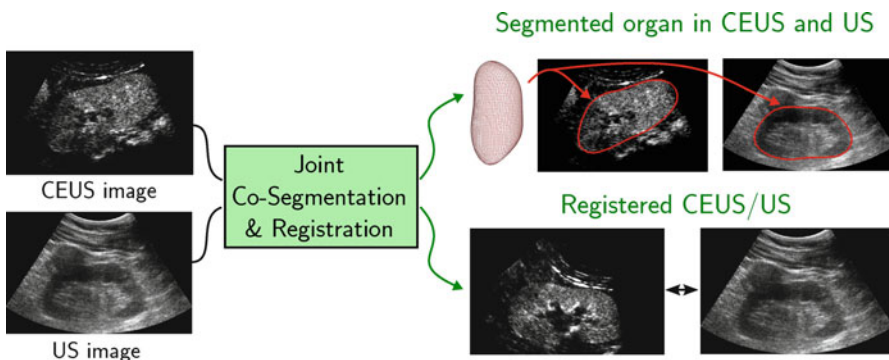

Fig. 11 Joint co-segmentation and registration. Given two different non-aligned images of the same object, the proposed method aims at segmenting this object in both images as well as estimating a rigid transformation between them
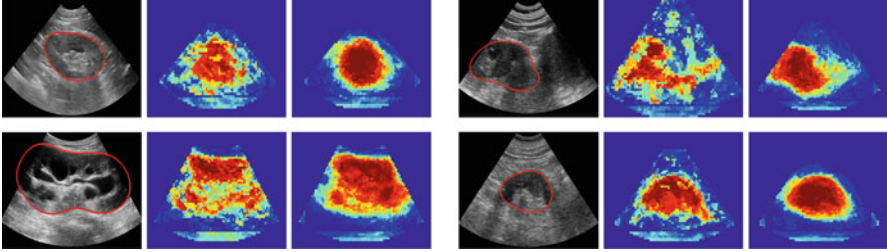
Fig. 12 Kidney appearance in US images (the kidney boundary is denoted in *red*). (*Left*) Original images showing the high variability of the database. (*Middle*) Kidney probability given by the first classifier. (*Right*) Final kidney probability p_{US}

Learning Appearance in Conventional Ultrasound

In CEUS images, bright areas indicate the presence of contrast agent which is mainly localized in the kidney. This is why we directly used the image intensity as a voxel probabilities to be inside the kidney. However in conventional US images, this does not hold and we need to transform the image into a more elaborate kidney probability map.

The kidney appearance has a much higher variability in US images, although their structure is consistent: kidneys are always composed of a bright sinus surrounded by a darker parenchyma (see Fig. 12). As intensity itself is not reliable enough, we chose to combine multiple image features using decision forests [26] to obtain a class posterior map p_{US}.

Recent work [27–31] demonstrated that adding contextual information allows to improve spatial consistency and thus classification performance. Here we propose to exploit the kidney structure in a simple yet efficient way. Similarly to the auto-context framework introduced by Tu et al. [32], contextual information is included by using two classifiers in cascade. A first classification (kidney vs background) is performed in each voxel using a decision forest. Then we use these class posterior probabilities as additional input of a second random forest that will give the final kidney probability p_{US}. In the remainder of the chapter, we will work on this map instead of the original US image.

The features used for the first decision forest were the intensity of the image and its Laplacian at the considered voxel as well as at its neighbors' within a $7 \times 7 \times 7$ local patch, at three different scales ($\sigma = 2,4,6$ mm). Intensities were normalized in each patch. For the second forest, we added the estimated class posterior as additional channels. Each forest was composed of 10 trees with maximum depth 15.

To validate this probability estimation, the patient database was split into two groups. Results on the whole dataset were then obtained using a two-fold cross-validation. Figure 13 shows the ROC and Precision-Recall curves computed (1) by
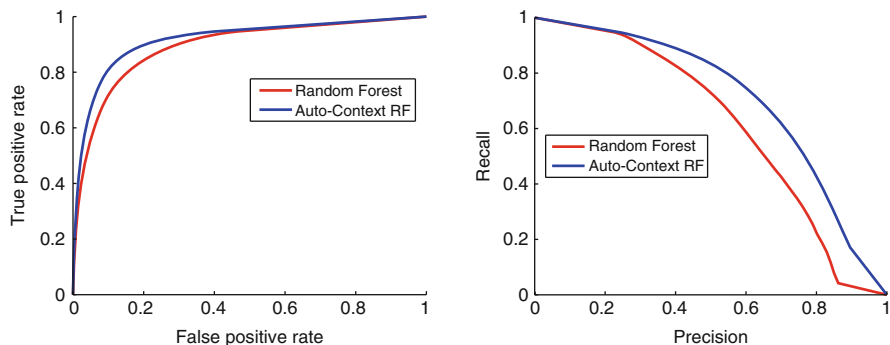
Fig. 13 Comparison of classification results for the single decision forest and the auto-context approach. (*Left*) ROC curve. (*Right*) Precision-Recall curve

the first decision forest and (2) using the auto-context approach with another forest in cascade. The latter provides better kidney probabilities with respect to all reported statistics. Indeed, taking into account structural information helps, for example, in distinguishing the kidney sinus from the background or the parenchyma from shadows and allows a more spatially coherent classification (see Fig. 12).

Generic Framework for Co-segmentation and Registration

In sections "Kidney Detection via Robust Ellipsoid Estimation" and "Kidney Segmentation via Implicit Template Deformation", we presented two variational methods to, respectively, detect and segment the kidney. They both consist in seeking ϕ as the minimizer of a functional of the following generic form

$$E_I(\phi) = \int_\Omega f(\phi(\mathbf{x})) \, r_I(\mathbf{x}) \, d\mathbf{x} + \mathcal{R}(\phi) \tag{19}$$

where f is a real-valued function and $r_I(\mathbf{x})$ denotes a pointwise score on whether $\mathbf{x}$ looks like an interior or exterior voxel in the image I. This is a standard setting in which the optimal implicit function ϕ must achieve a trade-off between an image-based term and a regularization term $\mathcal{R}$.[3]

---

[3] For example, the seminal method of Chan and Vese [33] falls in this framework with $f = H$ the Heaviside function and $r_I(\mathbf{x}) = (I(\mathbf{x}) - c_{\text{int}})^2 - (I(\mathbf{x}) - c_{\text{ext}})^2$ with c_{int} and c_{ext} denoting mean intensities inside and outside the target object.

We are interested in the case where a pair of images $I_1 : \Omega_1 \to \mathbb{R}$ and $I_2 : \Omega_2 \to \mathbb{R}$ of the same object are available. If those images were perfectly aligned, the energy in Eq. (19) can be straightforwardly generalized to perform co-segmentation :

$$E_{I_1,I_2}(\phi) = \int_{\Omega_1} f(\phi(\mathbf{x})) \ (r_{I_1}(\mathbf{x}) + r_{I_2}(\mathbf{x})) \ d\mathbf{x} + \mathcal{R}(\phi) \ . \tag{20}$$

Unfortunately, such an assumption rarely holds in medical applications unless the two images are acquired simultaneously. A more realistic hypothesis is to assume that the target object, segmented by ϕ, is not deformed between the two acquisitions, but only undergoes an unknown rigid transformation $\mathcal{G}_r$. The co-segmentation energy thus reads

$$E_{I_1,I_2}(\phi, \mathcal{G}_r) = \int_{\Omega_1} f(\phi(\mathbf{x})) \ r_{I_1}(\mathbf{x}) \ d\mathbf{x} + \int_{\Omega_2} f(\phi \circ \mathcal{G}_r(\mathbf{x})) \ r_{I_2}(\mathbf{x}) \ d\mathbf{x} + \mathcal{R}(\phi) \ . \tag{21}$$

Note that, after a variable substitution, it can be equivalently written

$$E_{I_1,I_2}(\phi, \mathcal{G}_r) = \int_{\Omega_1} f(\phi(\mathbf{x})) \ (r_{I_1}(\mathbf{x}) + r_{I_2} \circ \mathcal{G}_r^{-1}(\mathbf{x})) \ d\mathbf{x} + \mathcal{R}(\phi) \ . \tag{22}$$

Minimizing E_{I_1,I_2} with respect to ϕ and $\mathcal{G}_r$ simultaneously can be therefore interpreted as performing jointly segmentation (via ϕ) and rigid registration (via $\mathcal{G}_r$). This generalizes a more common co-segmentation approach (e.g., [34]) where the images are first aligned in a preprocessing step.

In the following, we apply this framework to the robust ellipsoid detection (section "Kidney Detection via Robust Ellipsoid Estimation") and implicit template deformation (section "Kidney Segmentation via Implicit Template Deformation") to build a completely automated workflow for kidney segmentation in CEUS and US images. Note that the kidney, which is surrounded by a tough fibrous renal capsule, is a rigid organ. The hypothesis of non-deformation is therefore justified.

Application to Kidney Detection

The robust ellipsoid detection setting of Eq. (3) falls into the framework described in Eq. (19) with :

- $f = I d$ and $r_I = -wI$;
- $\mathcal{R}(\phi_{\mathbf{c},\mathbf{M}}) = \mathcal{R}(\mathbf{M}) = \mu . \int_{\Omega} Iw. \log\left(\frac{\mathrm{Vol}(\mathbf{M})}{|\Omega|}\right).$

Expanding this algorithm to another image I_2 requires the introduction of another weighting function w_2. Following Eq. (21), we can now define the co-detection energy as

$$
\begin{aligned}
E_{\text{co−det}}(\mathbf{c}, \mathbf{M}, w_1, w_2, \mathcal{G}_r) = &-\int_{\Omega} \phi_{\mathbf{c},\mathbf{M}}(\mathbf{x})\ w_1(\mathbf{x})\ I_1(\mathbf{x})\ d\mathbf{x} \\
&-\int_{\Omega} \phi_{\mathbf{c},\mathbf{M}} \circ \mathcal{G}_r(\mathbf{x})\ w_2(\mathbf{x})\ I_2(\mathbf{x})\ d\mathbf{x} \\
&+\mu \left(\int_{\Omega} w_1 I_1 + w_2 I_2\right)\ \log\left(\frac{\text{Vol}(\mathbf{M})}{|\Omega|}\right)
\end{aligned}
\tag{23}
$$

$$
\text{with}\quad \text{Vol}(\mathbf{M}) = \frac{4\pi}{3}\sqrt{\det \mathbf{M}^{-1}}\quad \text{the ellipsoid volume.}
$$

To facilitate the resolution of such a problem, $\mathcal{G}_r$—as a rigid transformation—can be decomposed into a rotation and a translation. We can therefore equivalently write the energy as a function of the ellipsoid center $\mathbf{c}_2$ in the second image and the rotation matrix $\mathbf{R}$:

$$
\begin{aligned}
E_{\text{co−det}}(\mathbf{c}_i, w_i, \mathbf{R}, \mathbf{M}) = &-\int_{\Omega} \phi_{\mathbf{c}_1,\mathbf{M}}(\mathbf{x})\ w_1(\mathbf{x})\ I_1(\mathbf{x})\ d\mathbf{x} \\
&-\int_{\Omega} \phi_{\mathbf{c}_2,\mathbf{R}^T\mathbf{M}\mathbf{R}}(\mathbf{x})\ w_2(\mathbf{x})\ I_2(\mathbf{x})\ d\mathbf{x} \\
&+\mu \left(\int_{\Omega} w_1 I_1 + w_2 I_2\right)\ \log\left(\frac{\text{Vol}(\mathbf{M})}{|\Omega|}\right)
\end{aligned}
\tag{24}
$$

Minimization of such functional is done in an alternate three-step process:

1. The statistical interpretation still holds for the ellipsoid centers and matrix: minimizers $c_1{}^*$ and $c_2{}^*$ are weighted centroids while minimizer $\mathbf{M}^*$ is related to the weighted covariance matrix of pixels coming from both images.
2. The unknown matrix $\mathbf{R}$ accounts for the possible rotation between the two images and can be parametrized by a vector of angles $\Theta \in \mathbb{R}^3$. A gradient descent is performed at each iteration to minimize the energy with respect to Θ.
3. The weights w_1 and w_2 are finally updated as indicator functions (up to a slight dilation) of the current ellipsoid estimates.

The complete minimization strategy is summarized in Algorithm 2. This algorithm is computationally efficient: closed-form solutions are available (except for $\mathbb{R}$) and the process, though iterative, usually converges in very few iterations.

Algorithm 3: Robust ellipsoid co-detection algorithm

---

initialization $\forall\, \mathbf{x} \in \Omega,\ w_1(\mathbf{x}) \leftarrow 1,\ w_2(\mathbf{x}) \leftarrow 1$

repeat

   // Estimation of centers $\mathbf{c}_1$ and $\mathbf{c}_2$ and matrix $\mathbf{M}$

      $\mathbf{c}_1 \leftarrow \dfrac{1}{\int_\Omega w_1 I_1}\ \int_\Omega w_1(\mathbf{x})\, I_1(\mathbf{x})\, \mathbf{x}\, d\mathbf{x}$

      $\mathbf{c}_2 \leftarrow \dfrac{1}{\int_\Omega w_2 i_2}\ \int_\Omega w_2(\mathbf{x})\, I_2(\mathbf{x})\, \mathbf{x}\, d\mathbf{x}$

    $\mathbf{M}^{-1} \leftarrow \dfrac{2}{\mu \int_\Omega w_1 I_1 + w_2 I_2}\Big(\int_\Omega w_1(\mathbf{x})\, I_1(\mathbf{x})\, (\mathbf{x}-\mathbf{c}_1)\,(\mathbf{x}-\mathbf{c}_1)^T\, dx$

                         $+\int_\Omega w_2(\mathbf{x})\, I_2(\mathbf{x})\, \mathbf{R}\,(\mathbf{x}-\mathbf{c}_2)\,(\mathbf{x}-\mathbf{c}_2)^T\, \mathbf{R}^T dx\Big)$

   // Update of the rotation matrix by gradient descent with step Δt

   repeat

    | $\mathbf{R}(\Theta) \leftarrow \mathbf{R}\left(\Theta - \Delta t\ \nabla_\Theta E_{co-det}(\Theta)\right)$

   until *convergence*;

   // Update of the weighting functions w_1 and w_2 for each $\mathbf{x} \in \Omega$

   if $(\mathbf{x}-\mathbf{c})^T\, \mathbf{M}\,(\mathbf{x}-\mathbf{c}) \le 1 - \mu \log\left(\dfrac{Vol(\mathbf{M})}{|\Omega|}\right)$ **then**

    |__ $w_1(\mathbf{x}) \leftarrow 1$ **else** $w_1(\mathbf{x}) \leftarrow 0$

   if $(\mathbf{x}-\mathbf{c}_2)^T\, \mathbf{R}^T \mathbf{M} \mathbf{R}\,(\mathbf{x}-\mathbf{c}_2) \le 1 - \mu \log\left(\dfrac{Vol(\mathbf{M})}{|\Omega|}\right)$ **then**

    |__ $w_2(\mathbf{x}) \leftarrow 1$ **else** $w_2(\mathbf{x}) \leftarrow 0$

until *convergence*;

Figure 14 shows an example of ellipse co-detection in synthetic images, where the probability of belonging to the target object is the image intensity. Despite the noise, the simulated shadow, and the reduced field-of-view effect, the co-detection algorithm provides a good estimate on the ellipse position, size, and orientation in both images.

Application to Kidney Segmentation

Implicit template deformation, as previously described in section "Kidney Segmentation via Implicit Template Deformation", is part of the framework defined in Eq. (19) with :

- $f = H$ and $r_I = -\Delta I$;
- $\mathcal{R}(\phi_0 \circ \psi) = \mathcal{R}(\mathcal{L}) = \frac{\lambda}{2}\, \|\mathcal{L} - Id\|_2^2.$

We can therefore extend it to co-segmentation using Eq. (21) by considering the following functional

$$
\begin{aligned}
E_{\text{co-seg}}(\phi_0 \circ \mathcal{L} \circ \mathcal{G}, \mathcal{G}_r) =\ & E_{\text{co-seg}}(\mathcal{L}, \mathcal{G}, \mathcal{G}_r) \\
=\ & -\int_\Omega H(\phi_0 \circ \mathcal{L} \circ \mathcal{G})\, \Delta I_1(\mathbf{x})\, d\mathbf{x} \\
& -\int_\Omega H(\phi_0 \circ \mathcal{L} \circ \mathcal{G} \circ \mathcal{G}_r)\, \Delta I_2(\mathbf{x})\, d\mathbf{x} \\
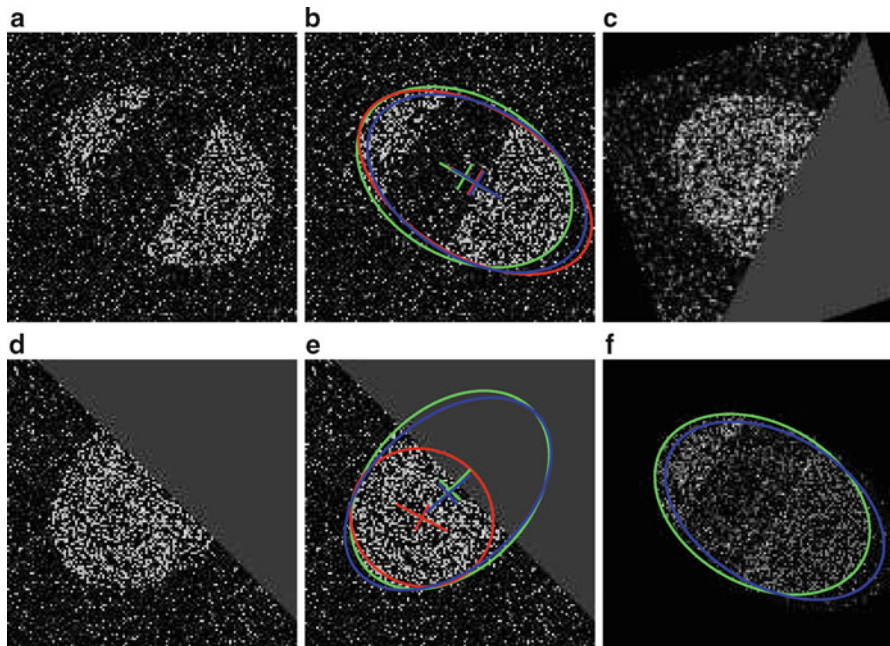& +\frac{\lambda}{2}\|\mathcal{L} - Id\|_2^2.
\end{aligned}
\tag{25}
$$

Fig. 14 Ellipse detection on two synthetic images I_1 (**a**) and I_2 (**d**). Detected ellipses with their center and main axes are shown in (**b**) and (**e**) for independent ellipse detection (*red*) and proposed method for co-detection (*blue*) compared to the ground truth (*green*). (**c**) Second image registered with the estimated transform $\mathcal{G}_r^{-1}$. (**f**) Combination of image terms $w_1 I_1 + (w_2 I_2) \circ \mathcal{G}_r^{-1}$ used for ellipse estimation at convergence

The energy $E_{\mathrm{co-seg}}$ is then minimized with respect to the parameters of $\mathcal{G}$, $\mathcal{G}_r$ and each component of the vector field **u**, through a gradient descent similar to section "Numerical Implementation".

Results for Kidney Co-segmentation in CEUS and US

The average overall computational time for kidney probability estimation in US, ellipsoid co-detection, and kidney co-segmentation was around 20 s with our implementation.

Validation was performed by comparing the co-segmentation approach to a segmentation from a single image (in both CEUS an US cases). Dice coefficients and relative error on the measured kidney volume are reported in Fig. 15. Using simultaneously the complementary information from US and CEUS images significantly improves the segmentation accuracy in both modalities. More specifically, the median Dice coefficient is increased from 0.74 to 0.81 in CEUS (p-value < 10^{-4}) and 0.73 to 0.78 in US (p-value < 10^{-4}). Furthermore, the proposed approach
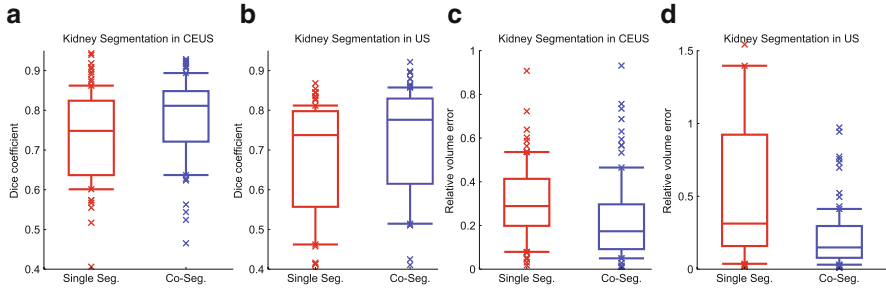
Fig. 15 Boxplots of segmentation results for kidney segmentation in US and CEUS images, in terms of Dice coefficients (**a**)–(**b**) and relative volume error (**c**)–(**d**). The proposed co-segmentation compares favorably to independent segmentation with a p-value $< 10^{-4}$. Boxplots show, respectively, the first decile, the first quartile, the median, the third quartile, and the ninth decile. Extreme points are shown separately
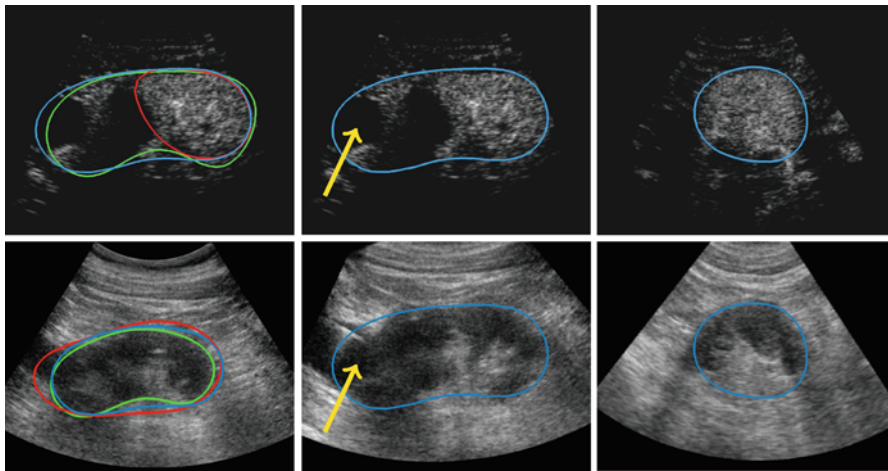

Fig. 16 Example of joint co-segmentation and registration for a CEUS (*top*) and a US (*bottom*) images. (*Left*) Comparison of independent segmentations (*red*) and the proposed co-segmentation (*blue*) with respect to the ground truths (*green*). (*Middle, Right*) Two views of the registered volumes that can be assessed by considering the position of the lesion (*yellow arrow*)

provides more reliable clinical information as the median error on the kidney volume is almost divided by two in CEUS (29 % versus 15 %) and in US (25 % versus 13 %). Figure 16 shows the joint co-segmentation and registration results for one case. Independent segmentation fails in both US and CEUS images because of the kidney lesion (indicated by the yellow arrow) that looks like the background in CEUS but like the kidney in US. Conversely, the proposed co-segmentation manages to overcome this difficulty by combining information from the two modalities. Furthermore, for this example, one can assess the estimated registration by comparing the location of the lesion in the two modalities. Results on another case were also displayed in Fig. 11.

Conclusion

This chapter addressed the problem of kidney segmentation in 3D CEUS images. Such a task is challenging because of the noise, the artifacts, and the partial occultation of the organ (due to the limited field of view).

A robust ellipsoid detector has been introduced to coarsely locate the kidney. The ellipsoid is then deformed to segment the kidney more precisely, by maximizing the image gradient flux through the segmentation boundary, using the template deformation framework. This method yields a fully automatic pipeline that provides a satisfying segmentation in a large number of cases but may fail when the image information is too ambiguous (shadows, pathologies, etc).

To overcome such difficulties, two extensions of this approach have been proposed to take into account additional information. First, we showed how user interactions can be exploited to guide the segmentation in real-time, by letting the user indicate points that should be inside/outside/on the segmentation. Then, we introduced a generic co-segmentation framework that generalizes any segmentation method to allow the simultaneous use of multiple images (here the CEUS and the US images). This results in both a better estimate of the organ shape and a registration of the images. The two aforementioned extensions are compatible and including user interactions in multiple images would be straightforward.

The kidney detection can still be improved by including more anatomical prior knowledge. A possible solution would be to constrain the ellipsoid's axis lengths or volume to be close to clinically meaningful values. Another way is the use of CT images of the same patient to extract a tailored model of the kidney and help both the CEUS detection and segmentation.

Appendix: Choice of the Parameter μ for Ellipsoid Detection

The choice of μ in Eq. (3) is paramount as it controls the number of points that are taken into account for the ellipsoid matrix estimation. To find a suitable value, let us consider an ideal case of an image I_0 in which there is one white ellipsoid ($I_0 = 1$) on a black background ($I_0 = 0$), whose implicit function is $\phi_{\mathbf{c}_0, \mathbf{M}_0}$. We also assume that the confidence weight is $w \equiv 1$ everywhere. Then the matrix estimated by our approach would be

$$\mathbf{M}^* = \mathrm{argmin}_{\mathbf{M}} \; E_{\mathrm{det}}(\mathbf{c}_0, \mathbf{M}, \mathbf{1})$$
$$= \left[\frac{2}{\mu} \frac{1}{\int_{\Omega} I_0} \int_{\Omega} I_0(\mathbf{x}) \; (\mathbf{x} - \mathbf{c}_0)(\mathbf{x} - \mathbf{c}_0)^T d\mathbf{x} \right]^{-1} \tag{26}$$

Using the fact that $I_0 = \mathbf{I}_{\{1 - (\mathbf{x} - \mathbf{c}_0)^T \mathbf{M}_0 (\mathbf{x} - \mathbf{c}_0) \geq 0\}}$ is the indicator of the ellipsoid yields

$$\mathbf{M}^* = \left[\frac{2}{\mu} \frac{1}{\text{Vol}(\mathbf{M}_0)} \int_{\{1-(\mathbf{x}-\mathbf{c}_0)^T \mathbf{M}_0 (\mathbf{x}-\mathbf{c}_0) \geq 0\}} (\mathbf{x} - \mathbf{c}_0)(\mathbf{x} - \mathbf{c}_0)^T d\mathbf{x} \right]^{-1} \qquad (27)$$

After a variable substitution $\mathbf{x} \leftarrow \mathbf{M}_0^{1/2}(\mathbf{x} - \mathbf{c}_0)$, this expression becomes

$$\mathbf{M}^* = \left[\frac{2}{\mu} \frac{\det(\mathbf{M}_0^{-1/2})}{\text{Vol}(\mathbf{M}_0)} \mathbf{M}_0^{-1/2} \int_{\{\|\mathbf{x}\| \leq 1\}} \mathbf{x}\mathbf{x}^T d\mathbf{x} \ \mathbf{M}_0^{-1/2} \right]^{-1} \qquad (28)$$

With $\text{Vol}(\mathbf{M}_0) = \frac{4\pi}{3} \sqrt{\det(\mathbf{M}_0^{-1})} = \frac{4\pi}{3} \det(\mathbf{M}_0^{-1/2})$, we then obtain

$$\mathbf{M}^* = \left[\frac{2}{\mu} \frac{3}{4\pi} \mathbf{M}_0^{-1/2} \int_{\{\|\mathbf{x}\| \leq 1\}} \mathbf{x}\mathbf{x}^T d\mathbf{x} \ \mathbf{M}_0^{-1/2} \right]^{-1} \qquad (29)$$

Note that the integral $\int_{\{\|\mathbf{x}\| \leq 1\}} \mathbf{x}\mathbf{x}^T d\mathbf{x}$ denotes the covariance matrix of a 3D unit ball, which is actually a scalar matrix that can be easily computed

$$\int_{\{\|\mathbf{x}\| \leq 1\}} \mathbf{x}\mathbf{x}^T d\mathbf{x} = \begin{pmatrix} 2\pi\frac{2}{3}\frac{1}{5} & 0 & 0 \\ 0 & 2\pi\frac{2}{3}\frac{1}{5} & 0 \\ 0 & 0 & 2\pi\frac{2}{3}\frac{1}{5} \end{pmatrix} = \frac{4\pi}{15} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad (30)$$

Combining Eqs. (29) and (30) leads to

$$\mathbf{M}^* = \left[\frac{2}{\mu} \left(\frac{1}{5} \mathbf{M}_0^{-1} \right) \right]^{-1} \qquad (31)$$

which yields the following relationship between $\mathbf{M}^*$ and $\mathbf{M}_0$:

$$\mathbf{M}^* = \frac{5}{2}\mu\mathbf{M}_0 \qquad (32)$$

This shows that the exact solution $\mathbf{M}_0$ is retrieved for $\mu = \frac{2}{5}$. This value actually depends on the dimension of Ω. Here we assumed $\Omega \subset \mathbb{R}^3$ but for 2D images, the optimal value would rather be $\mu = \frac{1}{2}$.

References

1. Albrecht T et al (2004) Guidelines for the use of contrast agents in ultrasound. Ultraschall Med 25(4):249–256
2. Noble JA, Boukerroui D (2006) Ultrasound image segmentation: a survey. IEEE Trans Med Imag 25(8):987–1010
3. Gasnier A, Ardon R, Ciofolo-Veit C, Leen E, Correas J (2010) Assessing tumour vascularity with 3D contrast-enhanced ultrasound: a new semi-automated segmentation framework. In: Proceedings of IEEE ISBI 2010, pp 300–303. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5490351
4. Kissi A, Cormier S, Pourcelot L, Bleuzen A, Tranquart E (2004) Contrast enhanced ultrasound image segmentation based on fuzzy competitive clustering and anisotropic diffusion. In: IEEE IEMBS 2004, vol 1, pp 1613–1615. http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=1403489
5. Prevost R, Cohen L, Correas J, Ardon R (2012) Automatic detection and segmentation of renal lesions in 3D contrast-enhanced ultrasound images. Proc. SPIE 8314:83141D–83141D1
6. Ma M, Stralen M, Reiber J, Bosch J, Lelieveldt B (2009) Left ventricle segmentation from contrast enhanced fast rotating ultrasound images using three dimensional active shape models. In: Proceedings of FIMH 2009, pp 295–302. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6235871
7. Prevost R, Mory B, Correas JM, Cohen LD, Ardon R (2012) Kidney detection and real-time segmentation in 3D contrast-enhanced ultrasound images. In: Proceedings of IEEE ISBI, pp 1559–62
8. Prevost R, Cuingnet R, Mory B, Correas JM, Cohen LD, Ardon R (2013) Joint co-segmentation and registration of 3D ultrasound images. In: To appear in Proceedings of IPMI, pp 268–279. Springer, Berlin, Heidelberg
9. Guil N, Zapata E (1997) Lower order circle and ellipse Hough transform. Pattern Recogn 30 (10):1729–1744
10. McLaughlin RA (1998) Randomized Hough transform: improved ellipse detection with comparison. Pattern Recogn Lett 19(3):299–305
11. Wong C, Lin S, Ren T, Kwok N (2012) A survey on ellipse detection methods. In: 2012 I.E. international symposium on industrial electronics (ISIE), pp 1105–1110. IEEE, New York
12. Van Aelst S, Rousseeuw P (2009) Minimum volume ellipsoid. Wiley Interdisciplin Rev: Comput Statist 1(1):71–82
13. Saddi K, Chefd'Hotel C, Rousson M, Cheriet F (2007) Region-based segmentation via non-rigid template matching. In: Proceedings of ICCV, pp 1–7. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4409152
14. Somphone O, Mory B, Makram-Ebeid S, Cohen LD (2008) Prior-based piecewise-smooth segmentation by template competitive deformation using partitions of unity. In: Proceedings of ECCV, pp 628–641. http://link.springer.com/chapter/10.1007/978-3-540-88690-7_47
15. Mory B, Somphone O, Prevost R, Ardon R (2012) Real-time 3D image segmentation by user-constrained template deformation. In: Proceedings of MICCAI. LNCS, vol 7510, pp 561–568. Springer, Berlin
16. Cuingnet R, Prevost R, Lesage D, Cohen LD, Mory B, Ardon R (2012) Automatic detection and segmentation of kidneys in 3D CT images using random forests. In: Proceedings of MICCAI. LNCS, vol 7512, pp 66–74. Springer, Berlin
17. Yezzi A, Soatto S (2003) Deformotion: deforming motion, shape average and the joint registration and approximation of structures in images. IJCV 53(2):153–167
18. Schunck B, Horn B (1981) Determining optical flow. In: Image understanding workshop, pp 144–156. http://www.sciencedirect.com/science/article/pii/0004370281900242
19. Nocedal J, Wright SJ (1999) Numerical optimization. Springer, Berlin
20. Vicente S, Kolmogorov V, Rother C (2010) Cosegmentation revisited: models and optimization. In: Proceedings of ECCV, vol 6312, pp 465–479

21. Wang F, Vemuri B (2005) Simultaneous registration and segmentation of anatomical structures from brain MRI. In: Proceedings of MICCAI. LNCS, vol 3749, pp 17–25. Springer, Berlin
22. Pohl K, Fisher J, Grimson W, Kikinis R, Wells W (2006) A Bayesian model for joint segmentation and registration. NeuroImage 31(1):228–239
23. Lu C, Duncan J (2012) A coupled segmentation and registration framework for medical image analysis using robust point matching and active shape model. In: IEEE Workshop on MMBIA, pp 129–136
24. Wyatt P, Noble JA (2002) MAP MRF joint segmentation and registration. In: Proceedings of MICCAI. LNCS, vol 2488, pp 580–587. Springer, Berlin
25. Yezzi A, Zöllei L, Kapur T (2003) A variational framework for integrating segmentation and registration through active contours. MedIA 7(2):171–185
26. Breiman L (2001) Random forests. Mach Learn 45(1):5–32
27. Payet N, Todorovic S (2010) 2-Random forest random field. In: Proceedings of NIPS 2010
28. Montillo A, Shotton J, Winn J, Iglesias J, Metaxas D, Criminisi A (2011) Entangled decision forests and their application for semantic segmentation of CT images. In: Proceedings of IPMI. LNCS, vol 6801, pp 184–196. Springer, Berlin
29. Kontschieder P, Bulò S, Criminisi A, Kohli P, Pelillo M, Bischof H (2012) Context-sensitive decision forests for object detection. In: Proceedings of NIPS, pp 440–448
30. Glocker B, Pauly O, Konukoglu E, Criminisi A (2012) Joint classification-regression forests for spatially structured multi-object segmentation. In: Proceedings of ECCV. LNCS, vol 7575, pp 870–881. Springer, Berlin
31. Zikic D, Glocker B, Konukoglu E, Criminisi A, Demiralp C, Shotton J, Thomas O, Das T, Jena R, Price S (2012) Decision Forests for Tissue-specific Segmentation of High-grade Gliomas in Multi-channel MR. In: Proceedings of MICCAI. LNCS, vol 7512, pp 369–376. Springer, Berlin
32. Tu Z, Bai X (2010) Auto-context and its application to high-level vision tasks and 3D brain image segmentation. IEEE TPAMI 32(10):1744–1757
33. Chan T, Vese L (2001) Active contours without edges. IEEE TIP 10(2):266–277
34. Han D, Bayouth J, Song Q, Taurani A, Sonka M, Buatti J, Wu X (2011) Globally optimal tumor segmentation in PET-CT images: A graph-based co-segmentation method. In: Proceedings of IPMI, pp 245–256. http://link.springer.com/chapter/10.1007/978-3-642-22092-0_21

Biography

Raphael Prevost is currently working as a Ph.D. candidate at Philips Research

Medisys (Suresnes, France) in collaboration with Universite Paris Dauphine (Paris, France). Born in France in 1988, he received in 2010 the M.Sc. degree in Applied Mathematics from the engineering school ENSTA ParisTech (Paris, France), and the M.Sc. degree in Image Processing and Machine Learning from Ecole Normale Superieure (Cachan, France). His research interests lie in various fields of medical image processing such as segmentation, registration, and pattern recognition.

Benoit Mory born in Paris in 1976, received an engineer degree in signal and image processing in 1998 from Ecole Supérieure d'Electricité, France. After 10+ years of experience as a research scientist with Philips Research France in the area of multimedia indexing (MPEG-7) and medical image processing, he received a Ph.D. in Electrical Engineering in 2011 from Ecole Polytechnique Fédérale de Lausanne, Switzerland. He is currently a senior researcher at Medisys, Philips Research lab located in Suresnes, France. His latest research interests include 3D image segmentation with implicit surfaces and multi-modal image registration.

Rémi Cuingnet received the B.Sc. from Ecole Polytechnique in 2006, the M.Sc. in Electrical Engineering from Telecom ParisTech in 2007, and the Ph.D. in Physics from University Paris-Sud in 2011. From 2007 to 2011, he was a research assistant within the Research Center of the Brain and Spine Institute (CR-ICM) in Paris. He is now a research scientist at Philips Research (Suresnes, France). His research interests are in the area of statistical learning and medical image computing.

Jean-Michel Correas is a Professor of Radiology at the Paris-Descartes University and is a vice chairman of the Department of Adult Radiology in Necker University Hospital, Paris, France. He got an M.D. Ph.D. degree from the University of Tours under the supervision of Pr. Leandre Pourcelot. He spent almost 2 years in the USA as part of his Ph.D. participating in the preclinical and clinical research programs in a small company developing contrast agents for ultrasound. He was Visiting Professor of Radiology at the University of Toronto department of Imaging Research, where he was in charge of several research projects on Ultrasound Contrast Agents under Pr. PN Burns supervision, as part of his doctorate of science on Ultrasound Contrast Agents. He was deeply involved in the development of interventional uroradiology. After ablation of more than 80 hepatocellular carcinomas, he developed together with the Adult Urology department the minimally invasive treatment of renal tumors, starting in 2004. He performed more than 400 ablative procedures of renal masses.

Pr. Correas is a member of the French Society of Radiology, the French Society of Uroradiology, and the European Society of Radiology. He was elected as President of the French Society of Ultrasound in 2008.

Laurent David Cohen was born in 1962. He was student at the Ecole Normale Superieure, rue d'Ulm in Paris, France from 1981 to 1985. He received the Master's and Ph.D. degrees in Applied Mathematics from University of Paris 6, France, in 1983 and 1986, respectively. He got the Habilitation à diriger des Recherches from University Paris 9 Dauphine in 1995.

From 1985 to 1987, he was member at the Computer Graphics and Image Processing group at Schlumberger Palo Alto Research, Palo Alto, California and Schlumberger Montrouge Research, Montrouge, France and remained consultant with Schlumberger afterwards. He began working with INRIA, France in 1988, mainly with the medical image understanding group EPIDAURE.

He obtained in 1990 a position of Research Scholar (Charge then Directeur de Recherche 1st class) with the French National Center for Scientific Research (CNRS) in the Applied Mathematics and Image Processing group at CEREMADE, Universite Paris Dauphine, Paris, France. His research interests and teaching at university are applications of Partial Differential Equations and variational methods to Image Processing and Computer Vision, like deformable models, minimal paths, geodesic curves, surface reconstruction, Image segmentation, registration, and restoration.

He is currently an editorial member of the *Journal of Mathematical Imaging and Vision*, the *International Journal for Computational Vision and Biomechanics*, and the journal *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*. He was also member of the program committee for about 40 international conferences. He authored about 200 publications in international Journals and conferences or book chapters, and 5 patents.

In 2002, he got the CS 2002 prize for Signal and Image Processing. In 2006, he got the Taylor & Francis Prize: "2006 prize for Outstanding innovation in computer methods in biomechanics & biomedical engineering."

He was 2009 laureate of Grand Prix EADS de l'Academie des Sciences.

He was promoted as IEEE Fellow 2010 for contributions to computer vision technology for medical imaging.

Roberto Ardon was born in Guatemala in 1977. He received an engineer degree in Applied Mathematics in 2001 from Ecole Centrale Paris, a master's degree in computer vision in the same year from Ecole Normale Supérieure de Cachan, and a Ph.D. with honors in Applied Mathematics from the Universite Paris Dauphine in 2005 within a collaboration with Philips Research Lab in Paris, Medisys. He is currently a senior researcher in Medisys, a team in which he joined as a permanent member in 2005. His research interests are in medical image analysis and computer vision.