

Registration of Free-Breathing 3D+t Abdominal Perfusion CT Images via Co-Segmentation

R. Prevost^{1,2}, B. Romain^{1,3}, R. Cuingnet¹, B. Mory¹,
L. Rouet¹, O. Lucidarme⁴, L.D. Cohen², and R. Ardon¹

¹ Philips Research Medisys, Suresnes, France

² CEREMADE UMR 7534, CNRS, Universite Paris Dauphine, Paris, France

⁴ MAS, Ecole Centrale Paris, Chatenay Malabry, France

³ Hopital La Pitie-Salpetriere, AP-HP, Paris, France

Abstract. Dynamic contrast-enhanced computed tomography (DCE-CT) is a valuable imaging modality to assess tissues properties, particularly in tumours, by estimating pharmacokinetic parameters from the evolution of pixels intensities in 3D+t acquisitions. However, this requires a registration of the whole sequence of volumes, which is challenging especially when the patient breathes freely. In this paper, we propose a generic, fast and automatic method to address this problem. As standard iconic registration methods are not robust to contrast intake, we rather rely on the segmentation of the organ of interest. This segmentation is performed jointly with the registration of the sequence within a novel co-segmentation framework. Our approach is based on implicit template deformation, that we extend to a co-segmentation algorithm which provides as outputs both a segmentation of the organ of interest in every image and stabilising transformations for the whole sequence. The proposed method is validated on 15 datasets acquired from patients with renal lesions and shows improvement in terms of registration and estimation of pharmacokinetic parameters over the state-of-the-art method.

1 Introduction

1.1 Clinical context

Dynamic contrast-enhanced (DCE) or perfusion imaging consists in acquiring a time sequence of images after a contrast agent injection. Parametric images are then generated by fitting at each voxel a pharmacokinetic model to its time-intensity curve. This technique is particularly used for oncologic applications, such as renal tumours follow-up, as the estimated parameters yield valuable information on healthy tissues and lesions [1].

Perfusion images can be acquired by MRI or CT systems. In this paper, we focus on DCE-CT sequences as it presents several advantages over DCE-MRI [1]. First, there is a linear relation between contrast agent concentration and image intensities (Hounsfield units), which simplifies pharmacokinetic models fitting. Second, CT is a cheaper and more widespread modality than MRI. However,

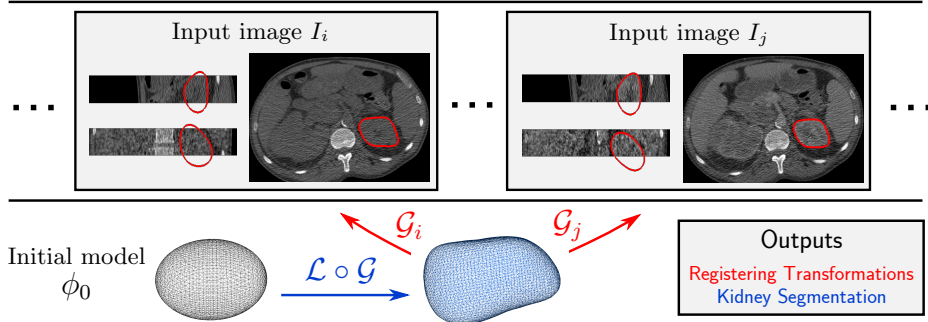


Fig. 1. Tracking via co-segmentation of the kidney in a sequence of volumes $(I_i)_i$. The segmentation is performed by template deformation ($\phi_0 \circ \mathcal{L} \circ \mathcal{G}$) using the whole sequence, while stabilising transformations $(\mathcal{G}_i)_i$ are simultaneously estimated.

this modality generates ionizing radiations that may harm the patient. In order to limit those risks, such acquisitions are performed in a very limited field of view with a reduced dose and a low framerate. This results in small, low-resolved volumes (both in spatially and temporally) that are noisier than static CT acquisitions (see Figure 1).

To capture the full dynamics of the contrast diffusion, a perfusion protocol lasts several minutes. Because of the patient’s breathing, voxels correspond to different anatomical locations across the sequence. The major challenge in parameter estimation is therefore to design a robust registration method, that cannot use temporal consistency because of the low framerate.

1.2 Related work and contributions

Several methods have already been proposed to register DCE sequences. Non-rigid iconic registration methods [2–4] are computationally demanding and rely on their similarity criterion. Standard choices such as mutual information are not effective in DCE-CT sequences [5]. In [6], Bhushan et al. used the pharmacokinetic model fitting error as registration criterion for DCE-MR, thus coupling the two tasks of sequence stabilisation and parameter estimation. However, the latter is a highly non-convex problem: including additional unknowns (namely the pharmacokinetic parameters) increases the risk of falling in local minima.

Some methods rather use a segmentation of the organ of interest to guide the registration [7–9]. Because of contrast diffusion, edge information is indeed more robust than region-based terms. In [7–9], segmentation and registration processes are performed sequentially. Yet they are inter-dependent (and equally challenging in DCE-CT images). In this paper, we propose a method to address both tasks simultaneously.

To do so, we extend the model-based segmentation algorithm proposed by Mory et al. [10] (that already proved effective kidney segmentation in CT images [11]) to a co-segmentation method that uses multiples images. This exten-

sion was inspired by a paper of Yezzi et al. [12], in which they segment a pair of CT/MR images with a single shape. However their work was based on active contours, while we adapt this approach to a more elaborate segmentation method and generalize it towards a novel model-based tracking method.

Concerning free-breathing DCE-CT registration, previous work is limited - to the best of our knowledge - to [5] in which Romain et al. proposed a registration by block-matching with a modified entropy-based similarity measure. This shall be considered as the baseline method.

Section 2 describes the segmentation framework and extends it to a tracking algorithm. In Section 3, we present results in terms of registration accuracy and parametric images and compare our approach to the baseline method [5]. Finally, Section 4 provides some discussion and concludes the paper.

2 Registration via kidney co-segmentation

As in [7–9], the proposed registration relies on a kidney segmentation. Our method is based on implicit template deformation [10], that we extend to a generic tracking method that simultaneously estimates the shape of the kidney and its pose in every frame of the sequence.

2.1 Template-based kidney segmentation

Let an initial shape \mathcal{S}_0 be implicitly represented by a function $\phi_0 : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}$, ie ϕ_0 is positive (resp. negative) inside (resp. outside) \mathcal{S}_0 and $\mathcal{S}_0 = \phi_0^{-1}(0)$. A space transformation $\psi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is determined such that the zero-level set of the deformed function $\phi = \phi_0 \circ \psi$ segments the target object in the image $I : \Omega \rightarrow \mathbb{R}$. The optimal transformation is then defined as the minimum of the energy

$$E(\psi) = \int_{\Omega} H(\phi_0 \circ \psi(\mathbf{x})) r(\mathbf{x}) d\mathbf{x} + \lambda d(Id, \psi) . \quad (1)$$

In the first term, H is the Heaviside function and r is a classification error function that depends on the image: for each point \mathbf{x} , $r(\mathbf{x}) < 0$ (resp. $r(\mathbf{x}) > 0$) if \mathbf{x} is likely to be an interior (resp. exterior) point. The choice of such a function will be detailed in Section 2.3. The second term d is a regularization penalty constraining the transformation ψ towards the identity and is weighted by the parameter λ . The transformation is decomposed as $\psi = \mathcal{L} \circ \mathcal{G}$ where

- \mathcal{G} is a global transformation, which accounts for the global pose and scale of the model (e.g. a similarity);
- \mathcal{L} is a non-rigid local deformation, expressed using a displacement field \mathbf{u} such that $\mathcal{L}(\mathbf{x}) = \mathbf{x} + (\mathbf{u} * K_{\sigma})(\mathbf{x})$. K_{σ} is a Gaussian kernel that provides built-in smoothness to such a transformation.

This decomposition allows d to be pose-invariant and to constrain only the non-rigid deformation : $d(Id, \psi) = d(Id, \mathcal{L}) = \frac{1}{2} \|\mathcal{L} - Id\|_2^2 = \frac{1}{2} \int_{\Omega} \|\mathbf{u} * K_{\sigma}\|_2^2$. This L_2 penalization on the amplitude of $\mathbf{u} * K_{\sigma}$ prevents large deviations from the original shape.

2.2 Tracking and stabilisation by co-segmentation

The previously described approach can be generalized to segment a given object in a collection of N images $(I_i)_{i=1..N}$. Following a similar approach to [12], a single shape shall segment all images. We introduce for every image I_i a transformation \mathcal{G}_i that we call the pose of the object in this image. The segmentation of the object in the i -th image is then the zero level-set of the function $\phi_0 \circ \mathcal{L} \circ \mathcal{G} \circ \mathcal{G}_i$. In the object's neighborhood, the transformations $(\mathcal{G}_i)_{i=1..N}$ act as stabilisation transformations from any image to a common reference (see Figure 1). Enforcing $\mathcal{G}_1 = Id$ sets this reference as the first image and resolves any possible ambiguity. Here we assume every \mathcal{G}_i to be a global rigid transformation. Indeed the kidney is a rigid organ whose motion across the sequence is mainly due to patient breathing. Note however that the proposed framework can be easily extended to any kind of stabilising transformations. We finally define the cosegmentation energy as a function that now depends on transformations $(\mathcal{G}_i)_i$:

$$E(\mathcal{L}, \mathcal{G}, (\mathcal{G}_i)_i) = \sum_{i=1}^N \int_{\Omega} H(\phi_0 \circ \mathcal{L} \circ \mathcal{G} \circ \mathcal{G}_i) \cdot r_i + \frac{\lambda}{2} \|\mathcal{L} - Id\|_2^2 \quad (2)$$

where r_i is the i -th image-based classification error function. This energy is minimized with a gradient descent, simultaneously performed on the deformation field \mathbf{u} and the parameters of the transformations \mathcal{G} , as well as the whole set $(\mathcal{G}_i)_i$. At the end of the process, we obtain both the kidney shape as the zero level-set of $\phi_0 \circ \mathcal{L} \circ \mathcal{G}$ and the transformations $(\mathcal{G}_i)_i$ that allow a global registration of the images.

2.3 Choice of the image-based term

The choice of the image-based term r is paramount for the segmentation. A common choice is $r(\mathbf{x}) = \log \frac{P_{ext}(I(\mathbf{x}))}{P_{int}(I(\mathbf{x}))}$ if intensities distributions are known inside (P_{int}) and outside (P_{ext}) the target object. However image intensities vary along the sequence because of contrast agent injection and kidney's appearance may change even between two successive acquisitions. It is therefore not robust to use such intensity models. We rather rely on edge information by only assuming that in every image I_i of the sequence, the kidney is brighter than its surrounding. This assumption is based on the fact that kidneys are highly vascularized organs whose contrast uptake is very early. The selected criterion to be minimized is the image gradient flux through the segmentation surface in image I_i denoted $\mathcal{S}_i = (\phi_0 \circ \mathcal{L} \circ \mathcal{G} \circ \mathcal{G}_i)^{-1}(0)$:

$$\int_{\mathcal{S}_i} - \langle \nabla I_i(\mathbf{x}), \vec{\mathbf{n}}(\mathbf{x}) \rangle dS(\mathbf{x}) = \int_{\text{inside } \mathcal{S}_i} -\Delta I_i(\mathbf{x}) dV(\mathbf{x}) \quad (3)$$

where Δ denotes the Laplacian operator. The right-hand term in this equation is obtained by application of divergence theorem. In practice, we only take into account relevant edges by applying a Gaussian filter on each image before computing its derivatives. The proposed image-based term falls into the described

framework by setting for each image, $r_i(\mathbf{x}) = -\Delta(K_{\sigma_{im}} * I_i)(\mathbf{x})$. The segmentations thus included tumors (as they do respond to the contrast) and excluded cysts (that do not). In both cases however, this does not hinder the registration since for a given sequence, the same structure is segmented in all frames.

3 Experiments and Results

3.1 Material

The experiments are based on 15 3D+t sequences coming from six different patients with renal tumours, enrolled in a longitudinal study. The data were acquired on a Brilliance iCT 256 Philips scanner. For each patient, a dynamic CT protocol of perfusion was used. 66 volumes were acquired per sequence (48 volumes every 2.5 seconds then 18 volumes every 10 seconds). The patients were asked to breathe normally during the whole exam. Typical image size was $512 \times 512 \times 22$ voxels with a spatial resolution of $0.68 \times 0.68 \times 2.5$ mm. The model ϕ_0 was set to an ellipsoid inside the kidney (with one click but we believe this could be easily automated) in the first frame and all transformations $(\mathcal{G}_i)_i$ were initialized to the identity. Our algorithm, implemented in C++, processes a whole 3D+t sequence in 30 seconds on a standard computer (Intel Core i5 2.67 GHz with 4GB RAM). Such a small computational time is possible because the registration is driven by the segmentation and thus only requires computations on the boundary of the segmented organ.

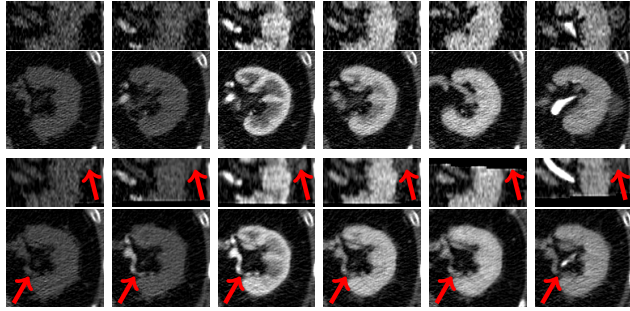


Fig. 2. Crops of coronal and axial slices between an original sequence (top) and the same sequence registered with our method (bottom) along the acquisition times. Note the stabilisation of the small blood vessel and the lesion (arrows).

3.2 Evaluation of registration

An example of a sequence before and after registration is given in Figure 2. We first assess the quality of the registration by computing error measures on landmarks. In our application, we are particularly interested in the region near the renal lesion: the selected landmark was therefore this lesion. For every sequence, the lesion has been manually segmented in each frame. If the registration were perfect, the segmentation (after compensation by the motion estimated via the

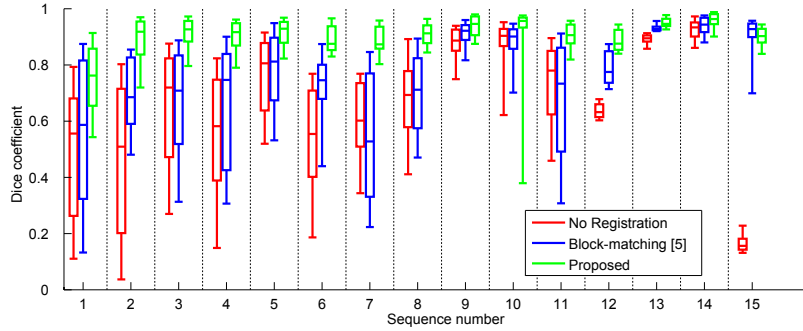


Fig. 3. Boxplots of lesions Dice coefficients from original sequences (red), sequences registered with entropy-based block matching [5] (blue) and the proposed method (green).

kidney) would be stable along the frames. We thus evaluated our registration by computing the Dice coefficient between the lesion in each frame and the lesion in the reference frame. Figures 3 shows this score, in comparison with the original sequence and the sequence registered by the block-matching method of [5]. Our method globally provides more precise and robust registration in the area of the tumour as the obtained Dice coefficients have both a higher mean and a lower variance. It outperforms [5] in every sequence but one, in which the lesion was extremely large (bigger than the kidney). Furthermore, the motion of center of the lesion, which was of 6.6mm (median over the datasets) in the original sequences, was reduced to 1.6mm after our stabilisation, which is below the resolution in z.

3.3 Parametric images

Our method was further evaluated by comparing parametric images estimated from the registered sequence. In each voxel, parameters (θ_1, θ_2) were obtained by fitting the time-intensity curve $I(t)$ - which is proportional to the contrast concentration $C_{tissue}(t)$ - to the solution of a Tofts model [13] :

$$\frac{dC_{tissue}(t)}{dt} = \theta_1 C_{aorta}(t) - \theta_2 C_{tissue}(t) \quad (4)$$

where C_{aorta} , which denotes the contrast concentration in the aorta, is modelled as the sum of two sigmoids and a Gaussian function. The sum of squared errors was minimized using a Levenberg-Marquardt method. In each voxel, we compute the residual fitting error at convergence, which quantifies how much the time-intensity curve deviates from the model. These are reported in Figure 4, which shows an improvement over baseline method in every tested sequence but one. This is illustrated by the example given in Figure 5. Our method provides much smoother curves, which improves the reliability of the subsequent parameter estimation. As the true parameters are unknown, we can only assess their value

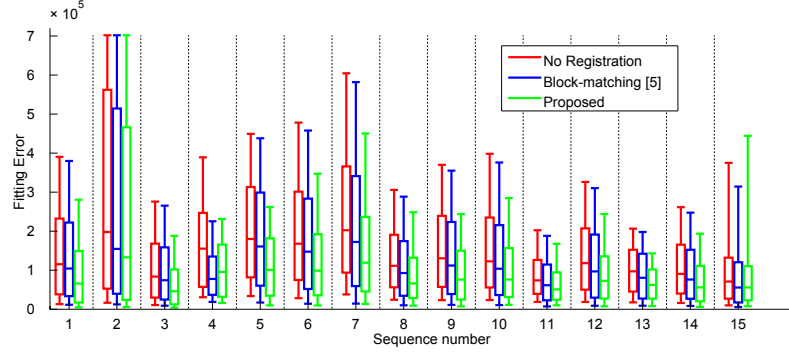


Fig. 4. Boxplots of pixelwise fitting errors (sum of squared differences with the model) for each sequence, estimated from original sequences (red), sequences registered with entropy-based block matching [5] (blue) and the proposed method (green).

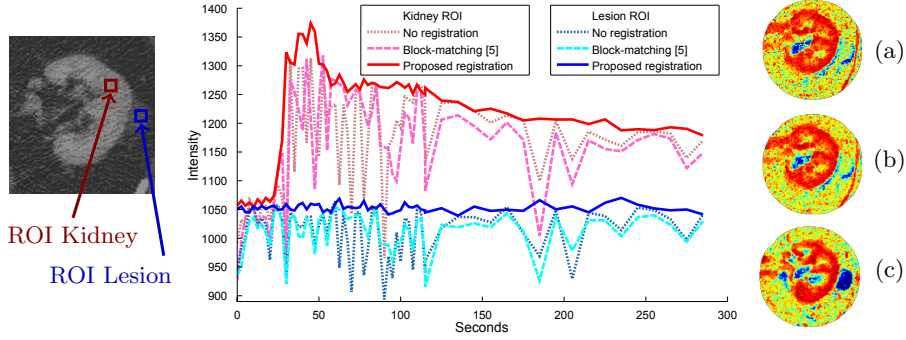


Fig. 5. (Left) Time-intensity curves in two different regions of interest of a sequence. (Right) Maps of parameter θ_2 estimated on (a) unregistered sequences, (b) sequences registered with entropy-based block matching [5] and (c) the proposed method. Note that the lesion is much more visible in our parametric image.

visually but one can clearly see that the lesion is much better distinguished in the parametric image obtained with our registration, which tends to prove the pertinence of our approach. The inner structures of the kidney are also more precisely delineated.

4 Discussion

In this paper, we proposed a fast, automatic and robust method to register 3D+t DCE-CT sequences. To be able to cope with contrast uptake, our approach relies on a segmentation of the organ of interest, rather than intensity-based similarity criteria. This segmentation is simultaneously estimated, within

a co-segmentation framework. Experiments showed that it provides better results than the state-of-the-art both quantitatively in terms of registration and qualitatively in terms of pharmacokinetic parameters estimation.

The proposed approach is generic and can be extended to other organs. For the kidney, a rigid transformation was enough to capture the movement of the region of interest. Other organs may undergo a different kind of movement, such as affine or even deformations. This can be taken into account within our framework by adapting $(\mathcal{G}_i)_i$ transformations. The co-segmentation could also have been directly applied to the tumour instead of the organ, but the definition of the image-based classification terms $(r_i)_i$ would have been challenging. In our experiments, the acquisitions frequency was so low (min 2.5 seconds) that no temporal coherence was enforced. For other applications, temporal consistency can however be useful. This could be achieved by adding extra terms in the energy to constrain the transformations $(\mathcal{G}_i)_i$ and is currently under investigation.

References

1. Kambadakone, A., Sahani, D.: Body perfusion CT: technique, clinical applications, and advances. *Radiologic clinics of North America* **47**(1) (2009) 161–178
2. Sance, R., et al.: Alignment of 3D DCE-MRI abdominal series for optimal quantification of kidney function. In: *Proceedings of IEEE ISPA 2007*. (2007) 413–17
3. Linguraru, M., et al.: Renal tumor quantification and classification in contrast-enhanced abdominal CT. *Pattern recognition* **42**(6) (2009) 1149–61
4. Zöllner, F., et al.: Assessment of 3D DCE-MRI of the kidneys using non-rigid image registration and segmentation of voxel time courses. *Computerized Medical Imaging and Graphics* **33**(3) (2009) 171–181
5. Romain, B., Letort, V., Lucidarme, O., d’Alche Buc, F., Rouet, L.: Registration of free-breathing abdominal 3D contrast-enhanced CT. In: *Abdominal Imaging. Computational and Clinical Applications*. Volume 7601. Springer (2012) 274–282
6. Bhushan, M., et al.: Motion correction and parameter estimation in DCE-MRI sequences: application to colorectal cancer. In: *MICCAI*. Volume 6891 of LNCS. Springer (2011) 476–83
7. Sun, Y., et al.: Contrast-invariant registration of cardiac and renal MR perfusion images. In: *MICCAI*. Volume 3216 of LNCS. Springer (2004) 903–10
8. Song, T., et al.: Integrated four dimensional registration and segmentation of dynamic renal MR images. In: *MICCAI*. Volume 4191 of LNCS. Springer (2006) 758–65
9. El-Baz, A., et al.: A new CAD system for the evaluation of kidney diseases using DCE-MRI. In: *MICCAI*. Volume 4191 of LNCS. Springer (2006) 446–53
10. Mory, B., et al.: Real-time 3D image segmentation by user-constrained template deformation. In: *MICCAI*. Volume 7510 of LNCS. Springer (2012) 561–8
11. Cuingnet, R., et al.: Automatic detection and segmentation of kidneys in 3D CT images using random forests. In: *MICCAI*. Volume 7512 of LNCS. Springer (2012) 66–74
12. Yezzi, A., et al.: A variational framework for integrating segmentation and registration through active contours. *MedIA* **7**(2) (2003) 171–85
13. Tofts, P., et al.: Estimating kinetic parameters from dynamic contrast-enhanced T1-weighted MRI of a diffusable tracer: standardized quantities and symbols. *Journal of Magnetic Resonance Imaging* **10**(3) (1999) 223–232