

Slice-to-volume deformable registration: efficient one-shot consensus between plane selection and in-plane deformation

Enzo Ferrante¹  · Vivien Fecamp¹ · Nikos Paragios¹

Received: 23 February 2015 / Accepted: 3 April 2015 / Published online: 23 April 2015
© CARS 2015

Abstract

Purpose This paper introduces a novel decomposed graphical model to deal with slice-to-volume registration in the context of medical images and image-guided surgeries.

Methods We present a new non-rigid slice-to-volume registration method whose main contribution is the ability to decouple the plane selection and the in-plane deformation parts of the transformation—through two distinct graphs—toward reducing the complexity of the model while being able to obtain simultaneously the solution for both of them. To this end, the plane selection process is expressed as a local graph-labeling problem endowed with planarity satisfaction constraints, which is then directly linked with the deformable part through the data registration likelihoods. The resulting model is modular with respect to the image metric, can cope with arbitrary in-plane regularization terms and inherits excellent properties in terms of computational efficiency.

Results The proof of concept for the proposed formulation is done using cardiac MR sequences of a beating heart (an artificially generated 2D temporal sequence is extracted using real data with known ground truth) as well as multimodal brain images involving ultrasound and computed tomography images. We achieve state-of-the-art results while decreasing the computational time when we compare with another method based on similar techniques.

Conclusions We confirm that graphical models and discrete optimization techniques are suitable to solve non-rigid slice-to-volume registration problems. Moreover, we show that decoupling the graphical model and labeling it using two

lower-dimensional label spaces, we can achieve state-of-the-art results while substantially reducing the complexity of our method and moving the approach close to real clinical applications once considered in the context of modern parallel architectures.

Keywords Slice-to-volume registration · 2D–3D registration · Discrete optimization · Graphical models · Markov random fields

Introduction

The problem of slice-to-volume deformable image registration consists in aligning a sliced 2D image (e.g., ultrasound or US) to its corresponding plane from a 3D volume (e.g., computer tomography or CT). We call it deformable registration because the 2D image can be deformed during the registration process.

This problem finds applications in many medical image-related contexts such as computer-aided biopsy [19], motion correction for image reconstruction [5], tumor ablation [22] and image-guided surgery (IGS) [23]. In the case of image-guided procedures, a preoperative 3D image and several intra-operative 2D acquisitions are to be fused toward providing position and navigation information to the surgeons. Nowadays, this fusion is mainly performed using two different tracking technologies: optical (OTS) and electromagnetic (EMTS) tracking systems. In the first case, OTS requires a line of sight to be maintained between the tracking device and the instrument to be tracked; this fact can disturb doctors during their work and is not always convenient. In the second case, EMTS does not have line-of-sight requirements, but it is very susceptible to distortion from nearby metal sources and

✉ Enzo Ferrante
enzo.ferrante@ecp.fr

¹ Center for Visual Computing (CVN), CentraleSupélec – Galen Team, INRIA, 92295 Chatenay-Malabry, France

presents limited accuracy compared to optical tracking [4]. Moreover, neither OTS nor EMTS can deal with deformations between intra- and preoperative images. In this work, we propose to use 2D–3D slice-to-volume registration algorithms which are purely image based to solve this challenging problem and overcome the limitations presented by current technologies.

The problem of deformable image registration has been a pillar of computer vision (optical flow) and medical imaging (image fusion), and therefore, one can cite numerous methods to perform 2D–2D and 3D–3D registration [1, 11]. However, the problem of 2D–3D registration, and particularly the problem of slice-to-volume registration, deserves separate investigation and specific methods development. While a single 2D slice contains less information than a 3D volume, the solution remains a 3D mapping function (a deformation field in case of non-rigid registration or a transformation matrix in case of rigid registration) as in the case of 3D–3D registration. This fact converts 2D to 3D slice-to-volume registration in a really challenging problem. The other case of 2D–3D registration problems, where projective 2D images such as X-ray images are registered with volumetric images (CT, for example), has received more attention in the last years [15, 18] and is not covered in this paper.

A variety of methods has been proposed to deal with slice-to-volume registration. In [3], standard optimization approaches and heuristics (as simplex and simulated annealing algorithms) are applied on FluroCT to CT registration, testing with different intensity-based similarity measures. Dalvi and Abugharbieh [6] present a feature-based method that performs slice-to-volume registration, using several slices in order to improve the quality of the results. Gill et al. [10] track intra-operative MRI slices of prostate images with a preoperative MRI volume. This monomodal registration (MRI intra-operative slices to MRI preoperative volume) is designed to provide patient tracking information for prostate biopsy performed under MR guidance. A similar problem is tackled by [25] where a two-step algorithm (rigid registration in the first step and deformable registration in the second one) is applied to register three orthogonal intra-operative MR slices with a preoperative volume. San José Estépar et al. [23] propose a method to register endoscopic and laparoscopic US images with preoperative CT volumes. It is based on a new phase correlation technique called LEPART, and it manages only rigid registration in quasi-real time. Osechinskiy and Kruggel [21] present a flexible framework for intensity-based slice-to-volume non-rigid registration algorithms that was used to register histological sections images to MRI of the human brain.

The main limitations of the aforementioned methods are their specificity to the clinical context (they are derived and can be used for specific clinical applications), the requirement of anatomical segmentations in some of them that

increases their complexity and often their sequential nature where first plane is selected and then in-plane deformation is determined. Graphical models are powerful formalisms that could be amended to overcome these limitations. Casting computer vision problems as labeling ones through the use of Markov random field (MRF) theory has gained attention since [9]. It has been widely used to solve non-rigid image registration in the last years [11, 16, 17], mainly for 2D–2D or 3D–3D. In [26], a method based on MRFs to perform 2D–3D registration is presented, but it estimates just rigid transformations and works with projective images. Regarding slice-to-volume registration using MRF, our previous work [7] presents a MRF framework based on a high-dimensional label space to solve this problem; we will refer to it as the overparameterized method.

In this work, our aim is to introduce a low-rank graphical model that is able to simultaneously perform plane selection and estimate the in-plane deformation between the 2D source image and the corresponding slice from the 3D volume. We decouple a physical control point of a regular grid in two nodes of the MRF graph, one taking labels from the plane selection label space and the other one from the in-plane deformations label space. In that way, the complexity of the model reduces to the square of the cardinality of the biggest label space (instead of being quadratic in the product of the cardinalities of the two spaces), with a slight increase in the graphical model connectivity. This technique has been previously applied in 2D–2D registration [24]. The main advantage is related to the fact that, while the number of nodes augments linearly, the number of labels is decreased in a quadratic order.

The main contributions of our paper with respect to our previous work [7] are therefore two-fold. Firstly, we propose a new way of decoupling the plane selection and the in-plane deformation label spaces toward a novel low-rank model of order 3 (instead of a model of order 5 as in [7]); it results in a more tractable problem in terms of getting the optimal solution. Secondly, we obtain a substantial decrease in the search space size (order of 10), allowing much richer sampling of the label space, thus in theory more precise solutions. Moreover, by decoupling the label spaces, it is possible to explore both of them with different sparseness levels.

The framework is intensity based and independent of the similarity measure, so it can be adapted to different image modalities or new measures. We tested our approach on two different datasets: a monomodal dataset where 2D MRI images of the heart are registered with MRI volumes and another multimodal dataset where 2D US images are fused with CT volumes [20]. Both datasets were also used in [7].

The paper is organized as follows: In Sect. 2, we present the decoupled MRF formulation together with a complete explanation about the label spaces and the energy terms. In Sect. 3, the validation tests and results are presented and

discussed. Finally, Sect. 4 concludes our paper and provides some ideas on relevant future directions.

Method description

Non-rigid slice-to-volume registration can be seen as an optimization problem. We aim at optimizing an energy function by choosing the optimal plane (slice) $\hat{\pi}[J]$ from target volume J and the optimal deformation field \hat{T}_D as indicates the following equation:

$$\hat{T}_D, \hat{\pi} = \underset{T_D, \pi}{\operatorname{argmin}} \mathcal{D}(I \circ T_D(x), \pi[J](x)) + \mathcal{R}(T_D, \pi), \quad (1)$$

where I is the source 2D image, \mathcal{D} represents the data term, and \mathcal{R} is the regularization term. Given the 2D source image I and the 3D target volume J , we seek the slice $\hat{\pi}[J]$ from volume J that best matches the image I . We call it non-rigid registration because image I can be deformed by the deformation field \hat{T}_D . The data term \mathcal{D} measures the similarity between the source and the target, while the regularization term imposes smoothness constraints on the solution.

From this general optimization problem, we can derive different formulations. In [7], we proposed a high-dimensional label space-based approach considering local labels of dimension five (plane + in-plane deformations). One of the main problems related to this high dimensionality is its consequently high computational cost. In this work, we try to avoid this problem by decoupling the label space in two different ones and reforming the structure of the graph to still capture rigid plane displacements and in-plane deformation.

Our formulation consists in an undirected pairwise graph $G_D = \langle V, E \rangle$ with a set of nodes $V = V_I \cup V_P$ and a set of edges $E = E_I \cup E_P \cup E_D$. V_I and V_P have a four-neighbor grid structure and the same cardinality. Nodes in V_I are labeled with in-plane deformation labels, while labels

used in V_P represent the plane position. Edges from E_I and E_P correspond to a conventional pairwise neighborhood connection system for nodes in V_I and V_P , respectively; they are associated with regularization terms (E_I corresponds to in-plane deformation regularizers, and E_P to the plane selection regularizers). Edges in E_D link every node from V_I to its corresponding node from V_P , creating a graph with a sort of three-dimensional structure (see Fig. 1); those terms associated to E_D encode the data terms (i.e., the similarity measure).

In order to get a better understanding of the model, we can think of a single hypothetical grid similar to the one defined in [7], where every control point p_k from this grid is associated with two nodes from our approach, i.e., $v_k^I \in V_I$ and $v_k^P \in V_P$. This idea is depicted in Fig. 1, and it will be useful to understand the energy terms.

Label space

We define two different label spaces: One associated with nodes in V_I (called L_I) and the other one associated with nodes in V_P (called L_P).

The first label space, L_I , is a bidimensional space that models in-plane deformation using displacement vectors $l^I \in E_I = (d_x, d_y)$.

The second label space, L_P , indicates the plane in which the corresponding control point is located. It consists of labels l^P associated with different planes. In order to specify the plane and the orientation of the grid on it, we store an orthonormal basis of this plane together with the position of a reference point in this plane. Using this information, we can reconstruct the position of the rest of the control points in the grid. This way of storing the planes allow us to implement different plane space sampling methods. In this work, we chose a simple uniform sampling around the current plane position, varying rotation and translation parameters in a given range.

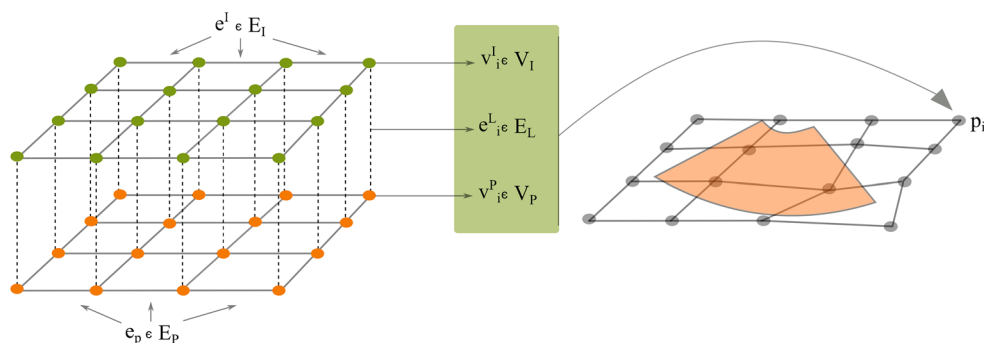


Fig. 1 Structure of the decoupled graph. The green nodes (top grid) are included in V_I and orange ones (bottom grid) in V_P modeling in-plane deformations and plane position, respectively. Edges connecting V_I nodes are part of E_I and those connecting V_P nodes are part of E_P ; they are associated with regularization terms. Dotted lines represent

cliques in E_D that encode the matching similarity measure. Using this information, we can reconstruct a deformed grid that is interpreted as a free-form deformation model. In the image, we can appreciate how we associate two nodes of the graph with one control point of the grid

This is an important advantage of our method: We could use prior knowledge to improve the way we explore the plane space, just by changing the plane space sampling method.

To compute the final position of a control point, we use both labels. First, the corresponding label in L_P defines a 3D point belonging to a plane space with a given basis. Then, we use the corresponding label in L_I to move the point in the 2D plane thanks to its basis.

Objective function

The energy that guides the optimization process is defined on the pairwise terms. Two types of edges represent regularization terms, while the last one represents the data terms; the energy is thus defined as:

$$\mathcal{E}(I, P, D) = \min \left\{ \gamma \sum_{(i,j) \in E_I} e_{i,j}^I(l_i^I, l_j^I) + \alpha \sum_{(i,j) \in E_P} e_{i,j}^P(l_i^P, l_j^P) + \beta \sum_{(i,j) \in E_D} e_{i,j}^D(l_i^I, l_j^P) \right\}, \quad (2)$$

where γ, α and β are positive weighting factors, $e_{i,j}^I \in I$ are the in-plane regularizers (associated with edges in E^I), $e_{i,j}^P \in P$ are the plane regularizers (associated with edges in E^P), and $e_{i,j}^D \in D$ are the data terms (associated with edges in E^D). l_i^I and l_j^P are labels from both label spaces L_I and L_P , respectively. Data and regularization terms are detailed in the following sections.

Data likelihood

The data term is defined for interconnected pairs of nodes (i, j) between the two graphs (where $i \in V^I, j \in V^P$) and their corresponding labels $l^I \in L_I, l^P \in L_P$. It is encoded in the pairwise terms $e^D \in E_D$. As we described before, a plane and an in-plane deformation 2D vector are associated with every control point. Combining both labels, we calculate the final position of the control point p_k and extract an oriented patch Ω_k over the plane π_k (centered in p_k) from the volume J , so that the similarity measure δ can be calculated between that patch and the corresponding area over the 2D source image:

$$e_{i,j}^D(l_i^I, l_j^P) = \int_{\Omega_k} \delta(I(x), \pi_k[J](x)) dx. \quad (3)$$

The patch-based similarity measure δ (defined on the sub-domain Ω_k) can encompass a wide choice of intensity-based

measures. One of the simplest and most used similarity measures is the sum of absolute differences (SAD). It is useful in the monomodal scenario, where two images of the same modality are compared. Its formulation is as follows:

$$e_{SAD,i,j}^D(l_i^I, l_j^P) = \int_{\Omega_k} |I(x) - \pi_k[J](x)| dx. \quad (4)$$

In multimodal scenarios, where different modalities are compared (e.g., CT with US images), statistical similarity measures such as mutual information (MI) are generally used since we cannot assume that corresponding objects have the same intensities in the two images. MI is defined using the joint intensity distribution $p(i, j)$ and the marginal intensity distribution $p(i)$ and $p(j)$ of the images as:

$$e_{MI,i,j}^D(l_i^I, l_j^P) = - \int_{\Omega_k} \log \frac{p(I(x), \pi_k[J](x))}{p(I(x))p(\pi_k[J](x))} dx. \quad (5)$$

As we could see in the previous examples, our framework can be endowed with any similarity measure defined on two bidimensional images. In this work, we use SAD for the monomodal heart dataset and MI for the multimodal brain dataset.

Regularization terms

We define two different regularization terms, one regularizing the plane selection and the other one the in-plane deformation. The first regularization term penalizes the average distance between the nodes $i, j \in V^P$ and the plane corresponding to the neighboring one. If $D_{\pi}(p)$ indicates the point-to-plane distance between the point p and the plane π , we define the regularization term e^P as the average of these distances for two neighboring points i, j and their corresponding planes:

$$e_{i,j}^P(l_i^P, l_j^P) = \frac{1}{2} (D_{\pi_j}(p'_i) + D_{\pi_i}(p'_j)), \quad (6)$$

where p'_i and p'_j are the positions after applying label l_i^P, l_j^P to p_i, p_j , respectively. This value is 0 when both points lie in the same plane.

The second regularization term controls the in-plane deformation and is defined between nodes i and j included in V_I . We use a distance-preserving approach which is symmetric, based on the ratio between the current position of the control points p_i, p_j and their original position $p_{o,i}, p_{o,j}$:

$$\psi_{i,j}(l_i^I, l_j^I) = \frac{\| (p_i + l_i^I) - (p_j + l_j^I) \|}{\| (p_{o,i}) - (p_{o,j}) \|}. \quad (7)$$

Once defined ψ_{ij} , we need our regularizer to fulfill two conditions: First, we want it to be symmetric with respect

to the displacement of the points, i.e., to penalize with the same cost whenever the control points are closer or more distant; second, we need the energy to be zero when the points are preserving distances and bigger than zero otherwise. The following regularization term fulfills both conditions for a couple of nodes $i, j \in V^I$ labeled with labels I_i^I, I_j^I :

$$e_{i,j}^I(I_i^I, I_j^I) = \left(1 - \psi_{i,j}(I_i^I, I_j^I)\right)^2 + \left(1 - \psi_{i,j}(I_i^I, I_j^I)^{-1}\right)^2. \tag{8}$$

Note that both types of pairwise terms are not sub-modular since we include the current position of the points (which can be arbitrary) in their formulation and therefore sub-modularity constraint is not fulfilled.

Implementation details

We adopt a pyramidal approach, using different grid resolution levels, from coarse to fine spacing between the control points. For each grid resolution, some iterations of the registration algorithm are performed, choosing the best possible set for each one and updating the control point positions with this information. During the inner iterations of one grid level, the size of the displacement vectors that form the deformation label space as well as the parameter variation of the plane label space is reduced in order to improve the search space sampling. A pseudocode of the algorithm is shown in Algorithm 1.

Algorithm 1 Pseudocode corresponding to the pyramidal approach based slice-to-volume registration

```

1: procedure 2D3DREGISTRATION( $I$ : Source,  $J$ : Target,  $T_0$ : Initial guess)
2:    $G \leftarrow \text{initializeGraph}(T_0)$   $\triangleright$  Initialize the graph in the position indicated by  $T_0$ 
3:    $\text{bestEnergy} \leftarrow \infty$ 
4:   for  $i=1$  to  $\text{gridLevels}$  do
5:      $L \leftarrow \text{updateLabelSpace}(L, i)$   $\triangleright$  Update the label space for the given level
6:     for  $j = 1$  to  $\text{iterationSteps}$  do
7:        $\text{newEnergy}, \text{newLabeling} \leftarrow \text{optimizeGraphicalModel}(G, L)$ 
8:       if  $\text{newEnergy} < \text{bestEnergy}$  then
9:          $\text{applyLabeling}(G, \text{newLabeling})$ 
10:         $\text{bestEnergy} = \text{newEnergy}$ 
11:       end if
12:        $\text{refineLabelSpace}()$ 
13:     end for
14:   end for
15:   return  $\text{bestEnergy}, G$ 
16: end procedure

```

The pairwise graphical model is optimized using the loopy belief propagation algorithm (other discrete optimiza-

tion algorithms can be used as well) implemented in the OpenGM2 library [12]. In [7], we used FastPD [14] instead of loopy belief propagation for optimizing our pairwise model, which is among the most efficient optimization algorithms. However, due to its construction (lifting of the duality gap minimization), FastPD requires in general (toward optimizing complexity) an equal number of labels for all nodes, which is an issue in our setting given the different dimensionality of the graph spaces (3D and 2D). Furthermore, while it can converge to a minimum even for non-submodular graphs, it is known that the quality of the linear programming (LP) relaxation is far from being satisfied and therefore the solution itself might be a very bad local minimum. Message passing methods like loopy belief propagation do not inherit the computational constraints of FastPD while it is known (at least experimentally) that they do good job as well even with highly non-submodular pairwise functions.

Validation and results discussion

We validate our method in two different scenarios, and we compare the results with our previous method [7]. The first one corresponds to a monomodal sequence of 2D MRI images randomly extracted from a 3D MRI temporal series of a beating heart. The second one is a multimodal brain dataset formed by 2D US images and 3D CT extracted from [20].

In order to compare both methods in a fair way, we exhaustively tested different parameter configurations (empirically for every dataset) on a grid of discretized values, and we took the best combination for each method.

Heart dataset

The MRI heart dataset consists of ten sequences of twenty bidimensional MRI slices each, which are registered with a MRI volume, giving a total of 200 registration cases. In order to generate them, as it was described in [7], we took a temporal series of 20 MRI volumes of a beating heart, and we extracted ten random trajectories of twenty slices I_i each (one slice for every volume M_i). Starting from a random initial rotation $R_0 = (R_{x_0}, R_{y_0}, R_{z_0})$ and translation $T_0 = (T_{x_0}, T_{y_0}, T_{z_0})$, we extracted a 2D slice I_0 from the initial volume M_0 . In every sequence, the position of slice I_i was generated adding Gaussian noise to the position of slice I_{i-1} with $\sigma_r = 3^\circ$ and $\sigma_t = 5$ mm to every translation (T_x, T_y, T_z) and rotation (R_x, R_y, R_z) parameters, respectively. It gives maximum distances of about 25 mm between the current and its succeeding slice. The MRI resolution was $192 \times 192 \times 11$, and the voxel size was $1.25 \times 1.25 \times 8$ mm³.

For every sequence, we initialize the registration adding the same noise (with the same parameters than before) to the ground truth. During the registration process, given two consecutive slices of the same sequence, the estimated trans-

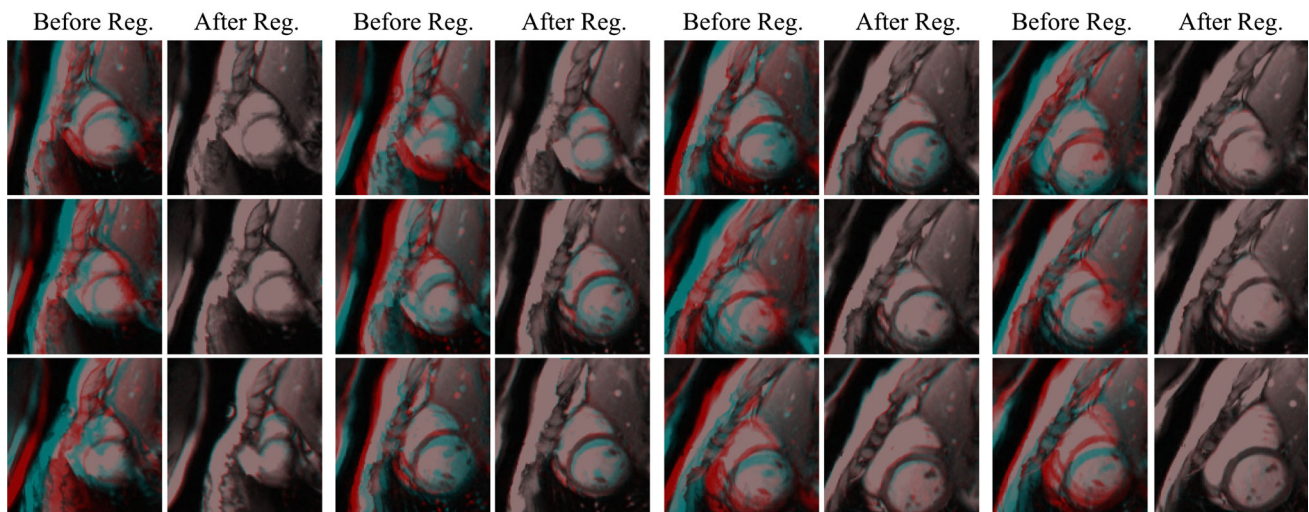


Fig. 2 12 registration cases of the same sequence, before and after registration. The overlapping images (in light blue, we show the source image, and in red, the target) showed before registration correspond to the source image and a slice taken from the volume at the initial position.

The overlapping images after registration correspond to the deformed source image and the slice taken from the volume at the estimated plane position

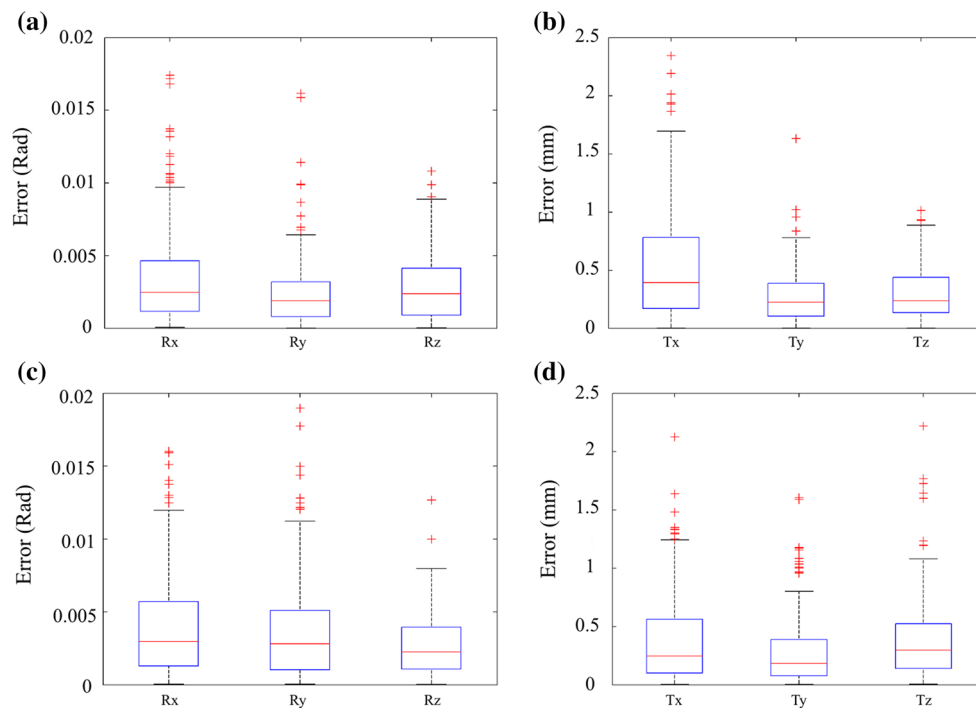


Fig. 3 Comparison of the error estimation for plane parameters (R_x , R_y , R_z) and (T_x , T_y , T_z) for our decoupled method (figures (a) and (b)) and the overparameterized approach presented by [7] (figures (c) and

(d)). For presentation clarity, three outliers between 0.02 and 0.05 rad as well as one at 4 mm have been removed at figures (c) and (b), respectively

formation for slice I_i was used as initialization for the registration of slice I_{i+1} .

Figure 2 shows the overlapping between the source image and the corresponding target plane, before and after registration, for 12 cases of one sequence. As we can observe in a qualitative way, the overlapping increases after registration.

Figure 3 compares our results in a quantitative way with the ones obtained using our previous method. We measure the error between the estimated transformation parameters and the ground truth. The mean error was (0.0036, 0.0024, 0.0029) rad for rotation and (0.5403, 0.2713, 0.2966) mm for translation parameters, with a standard

deviation of (0.0034, 0.0024, 0.0024)rad and (0.4914, 0.2296, 0.2236) mm, respectively. The average running time was around 60s for every registration case. Using the method presented in [7], we obtained (0.0051, 0.0051, 0.0031)rad and (0.4164, 0.2874, 0.4847) mm for rotation and translation parameters error, and standard deviation equal to (0.0122, 0.0134, 0.0051)rad and (0.4720, 0.2976, 1.1546)mm. Results are presented in Table 1. Every registration case took around 220s (almost 3.5 times more than our method). As we can see, the quality of the results was preserved (and improved in some cases), while the computational time was reduced approximately 3.5 times (keeping equivalent grid and label space sizes, sampling patch size and number of algorithm iterations).

Validation of in-plane deformation was performed over 20 registration cases, deforming an initial segmentation of the left endocardium using the estimated deformation field

T_{D_i} . We measure the average DICE coefficient between the segmentations, before and after deforming the initial one, to measure the impact of the deformation on the registration process. The average DICE before deformation was 0.858 and after registration was 0.907, showing that our method can capture in-plane deformations and select the correct plane at the same time.

Common parameters used for both methods were 3 grid levels, 5 iterations per level, initial control point distance of 40 mm and minimum sampling patch size of 20 mm. In case of the decoupled model, we use $\gamma = 1$, $\beta = 0.2$, $\alpha = 0.8$, 41 labels in the plane label space and 91 labels in the deformations label space. In case of the overparameterized model, we use 13,122 labels and $\alpha = 0.9$ (for a complete understanding of these parameters, refer to [7]). We run the experiments on an Intel Xeon W3670 with 6 Cores, 64bits and 16GB of RAM.

Table 1 Error estimation for plane parameters (R_x, R_y, R_z) and (T_x, T_y, T_z) for our decoupled method and the previous overparameterized approach presented in [7]

	R_x	R_y	R_z	T_x	T_y	T_z
Decoupled method						
Mean	0.0036	0.0024	0.0029	0.5403	0.2713	0.2966
SD	0.0034	0.0024	0.0024	0.4914	0.2296	0.2236
Overparameterized method [7]						
Mean	0.0051	0.0051	0.0031	0.4164	0.2874	0.4847
SD	0.0122	0.0134	0.0051	0.4720	0.2976	1.1546

Brain dataset

The brain dataset consists of a preoperative brain MRI volume (voxel size of $0.5 \times 0.5 \times 0.5 \text{ mm}^3$ and resolution of $394 \times 466 \times 378$ voxels) and 6 series of 9 US images extracted from the patient 01 of the database MNI BITE presented in [20]. The size of the US images was $48 \times 38 \text{ mm}$, and the pixel resolution $0.3 \times 0.3 \text{ mm}$. The ventricles were manually segmented by specialists in both modalities and used to calculate DICE coefficient and contour mean distance (CMD) to evaluate and compare the quality of the results. Initializations were done following the

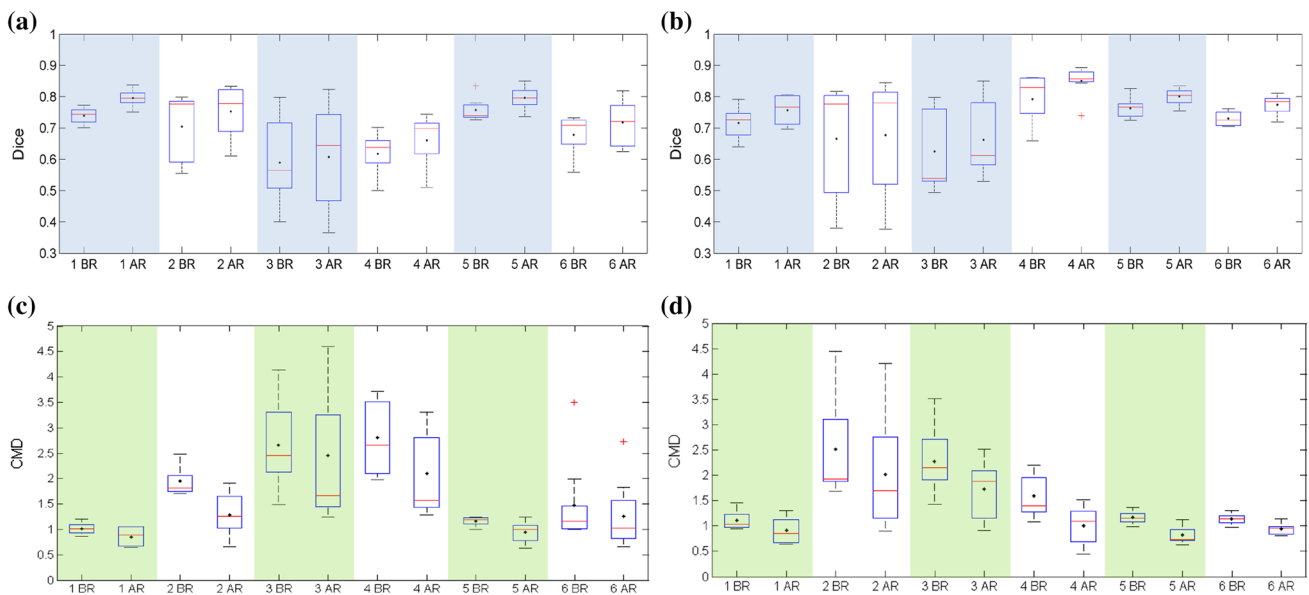


Fig. 4 Figures show a quantitative comparison of the two methods, before (BR) and after (AR) registration for the six sequences of brain data. Figures (a) and (c) show results for our decoupled method (DICE

and CMD, respectively), while figures (b) and (d) show results for the overparameterized approach presented in [7] (DICE and CMD, respectively)

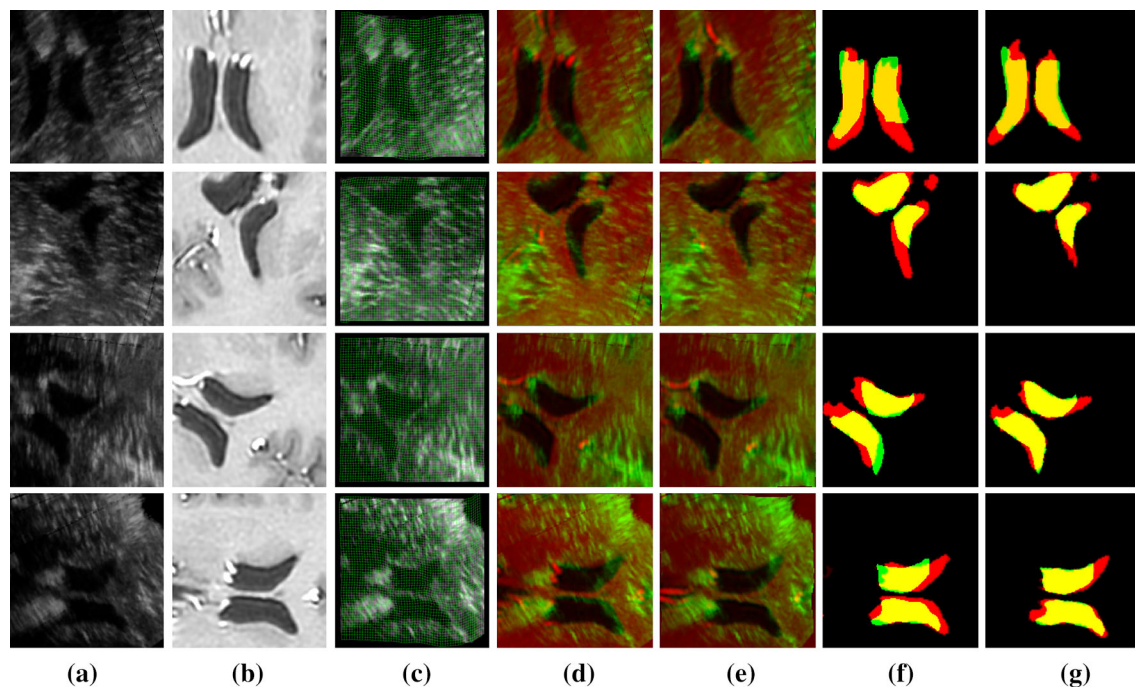


Fig. 5 Results for one slice from four of the six brain sequences (each row correspond to a different sequence). **a** Source 2D ultrasound image. **b** Slice extracted from the MRI corresponding to the initial position of the plane. **c** Deformed source image overlapped with the estimated

deformation field. **d** Blending between initial images (US and corresponding MRI slice). **e** Blending between final images (deformed US image and estimated MRI slice). **f** Overlapping between initial segmentations. **g** Overlapping between segmentations after registration

same methodology that we described for the heart dataset (Section 3.1).

Figure 4 summarizes the average DICE and CMD coefficients for each series while Fig. 5 shows some qualitative results. It shows that, using our decoupled method, the mean DICE increases after the registration process an average of 0.0405, a little bit more than the 0.0380 obtained with [7] method. Regarding the CMD, the average decrement for our method is 0.3654 mm while for the other one is 0.3943 mm. Even if our new method performs better in average, we can observe that results are almost equivalent in terms of DICE and CMD. However, there is a big difference in terms of computing time: while our method is taking around 3 min per registration case, the overparameterized method takes around 10 min running in the same computer using the same configuration. To perform the experiments with both methods, we used the same configuration given by 3 grid levels, initial control point distance of 8 mm, 4 iterations per level and minimum sampling patch size of 13 mm. In case of the decoupled model, we set $\gamma = 1$, $\beta = 0.05$, $\alpha = 0.2$, 41 labels in the plane label space and 91 labels in the deformations label space. For the overparameterized method, we set $\alpha = 0.8$ and 6174 labels. We run the experiments in the same Intel Xeon W3670 with 6 Cores, 64bits and 16GB of RAM used for the heart dataset (Fig. 5).

Discussion and comparison with other methods

As we have shown, our method is able to achieve state-of-the-art results while decreasing the computational time when we compare to another MRF-based method (namely [7]). In the monomodal case, we reduce it from around 3.5 to 1 min, while in the multimodal one, we go from 10 to 3 min, giving a time factor reduction of about three times.

The main strength of the proposed formulation is the linear complexity of the inference process with respect to the product of the label spaces. This allows to go even further for challenging cases (brain tumor removal) where precision is required to substantially increase the label space. This is not the case for the approach presented in [7] due to the complexity of the label space.

An interesting point to discuss about is the five-fold improvement in the standard deviation error of parameter T_z that we obtain with the new method. In [7], the justification for the poor performance of the method when estimating T_z was told to be that image resolution in z axis was lower than in x and y. We think that the new algorithm is less sensitive to image resolution anisotropy mainly because of the different way we explore the plane selection label space by allowing a deeper exploration when decoupling it without exponentially increasing the amount of labels.

It is important to remark that both, the decoupled and overparameterized methods, are highly dependent on the initialization given for the first slice of the sequence. Since these algorithms optimize the energy based on a limited search space (determined by the label space), if the solution is not reachable from the initial position using the current label space, the algorithm will fail. Another factor that is crucial for the success of the algorithm is the similarity measure used to decide whether or not two patches coming from different images correspond to the same anatomical structure. The study of different similarity measures is outside the scope of this paper; however, note that in order to use the method in other image modalities, it will be necessary to choose an accurate similarity measure and calibrate the parameters accordingly.

Comparison with other methods in the field of slice-to-volume registration is a complicated task, mainly because of the lack of public datasets. Here, we include some of the results reported by other state-of-the-art methods for their own datasets, in terms of accuracy and/or performance. In [10], for example, authors report a mean target registration error (TRE) lower than 1 mm when estimating rigid transformations in a monomodal MRI dataset of prostate images (for a pixel size of $1.5 \times 1.5 \times 3$ mm). Random initializations were generated by modifying the ground truth position with displacements of 10 mm and rotations of 10° maximum. The MATLAB implementation of their algorithm took between 36 and 107 s depending on the algorithm configuration. In [23], authors tested on a multimodal dataset formed by 2D ultrasound and CT volumes of the heart. They report errors around 1.56 ± 0.78 mm when estimating rigid transformations on CT images with 0.6 mm isotropic resolution, using initializations with uniformly random shifts in the range -5 to 5 mm. They achieve quasi-real-time performance with execution times around 4 s. Another interesting example to compare with is the multi-slice-to-volume registration case tackled [25], who applies it to MRI-guided transperineal prostate biopsy. Authors report that deformable registrations were accurate to within 2 mm in images with a slice spacing of 3.6 mm. The execution time for the complete deformable registration algorithm is about 30 s. Even if it is not possible to do a fair comparison mainly because of the lack of standard benchmarks, by observing these examples, we can clearly remark that our results are in the state-of-the-art level. Moreover, visual assessment on the obtained results seems to confirm that these are satisfactory in the context of a clinical setting.

In terms of complexity, it is interesting to remark the difference with respect to our previous method. The optimization complexity/difficulty heavily depends on the maximum number of label combinations that the pairwise cliques can take (this is the bottleneck for most optimization algorithms). In this perspective, the complexity of the overparameter-

ized model is given by $O(|L|^2)$, where $|L|$ is the cardinality (number of labels) of the label space. In our new approach, we introduce two label spaces L_1 and L_2 that decouple the previous one. To give an idea about the reduction in the complexity of our new model, let us say that $|L| = |L_1 \cdot L_2|$. Because of the way in which we construct our decoupled graph (as it is indicated in Fig. 1), it is straightforward to show that the complexity of the new model reduces now to $O(\max(|L_1|, |L_2|)^2)$. Therefore, because of the decoupling strategy, the complexity of the model reduces to the square of the cardinality of the biggest label space (instead of being quadratic in the cardinalities of the joint space), with a slight increase in the graphical model connectivity. Consequently, while the number of nodes augments linearly, the number of labels is decreased in a quadratic order.

Conclusions

We presented a new method to perform slice-to-volume registration based on a decoupled model that associates two local graphs to the plane selection and the in-plane deformations while imposing consistency through direct connections between the corresponding nodes. In order to solve this problem, we seek the plane and the in-plane deformation that best matches our energy function. It is important to remark that we just look for the in-plane deformations given the nature of the problems we are trying to solve (mainly image fusion for IGS), where it is not useful to find out-of-the-plane deformations at least for visualization purposes, even if they can exist.

As we have shown in the previous section, our method achieves state-of-the-art results while decreasing substantially the time of computation when it is compared to our previous MRF-based method that uses a unique high-dimensional label space [7]. It confirms our initial hypothesis, meaning that decoupling the graphical model and labeling it using two lower-dimensional label spaces, we can achieve the same results while reducing the complexity of our method.

We have also shown that the method is robust with respect to the type of images we are registering. Since slice-to-volume registration has multiple applications, other problems are under investigation (it should be noted that such a task is complex due to the complete absence of public ground truth). To this end, two clinical scenarios are currently under investigation, the first refers to liver tumor resection guidance, while the second to US guidance during prostate biopsy through fusion of intra-operative ultrasound and preoperative CT/MR.

In order to improve the quality of the results, specially in multimodal cases, feature engineering must be considered. Future work includes adapting and using features specifically designed for multimodal registration such as

the LC^2 presented in [8] and the MIND descriptor presented in [2]. Furthermore, energy regularizers inspired on precise biophysical modeling and tissue properties could lead to accuracy improvements as well. The underlying idea is to adapt the “smoothness” constraint of the deformation model by explicitly taking into account organ-specific motion/deformation constraints, for example, in the context of liver biopsies or brain tumor ablation.

Finally, we are investigating new methods to improve the parameters estimation procedure. Energy parameters estimation based on machine learning techniques [13] have to be considered as a future work if we want to exploit at the maximum level the potential of the proposed method.

Conflict of interest The authors declare that they have no conflict of interest.

Funding This research was partially supported by European Research Council Starting Grant Diocles (ERC-STG-259112).

Ethical approval This article does not contain any studies with animals performed by any of the authors. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

References

- Baker S, Scharstein D, Lewis J, Roth S, Black MJ, Szeliski R (2011) A database and evaluation methodology for optical flow. *Int J Computer Vision* 92(1):1–31
- Bardera A, Feixas M, Boada I, Sbert M (2006) High-dimensional normalized mutual information for image registration using random lines. In: Pluim J, Likar B, Gerritsen F (eds) *Biomedical Image Registration, Lecture Notes in Computer Science*, vol 4057, Springer, Berlin, Heidelberg, pp 264–271
- Birkfellner W, Figl M, Kettenbach J, Hummel J, Homolka P, Scherthaner R, Nau T, Bergmann H (2007) Rigid 2D/3D slice-to-volume registration and its application on fluoroscopic CT images. *Med Phys* 34(1):246. doi:10.1118/1.2401661
- Birkfellner W, Hummel J, Wilson E, Cleary K (2008) Tracking devices. In: *Image-guided interventions*, Springer, pp 23–44
- Chandler AG, Pinder RJ, Netsch T, Schnabel JA, Hawkes DJ, Hill DL, Razavi R (2008) Correction of misaligned slices in multi-slice MR cardiac examinations by using slice-to-volume registration. *J Cardiovas Magn Reson* 10:13
- Dalvi R, Abugharbich R (2008) Fast feature based multi slice to volume registration using phase congruency. In: *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pp 5390–5393
- Ferrante E, Paragios N (2013) Non-rigid 2d–3d medical image registration using Markov random fields. In: *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2013*, Springer, pp 163–170
- Fuerst B, Wein W, Muller M, Navab N (2014) Automatic ultrasound–MRI registration for neurosurgery using the 2d and 3d lc^2 metric. *Med Image Anal* 18(8):1312–1319. Special Issue on the 2013 Conference on Medical Image Computing and Computer Assisted Intervention
- Geman S, Geman D (1984) Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *Pattern Anal Mach Intell IEEE Trans* 6:721–741
- Gill S, Abolmaesumi P, Vikal S, Mousavi P, Fichtinger G (2008) Intraoperative prostate tracking with slice-to-volume registration in MRI. In: *Proceedings of the 20th International Conference of the Society for Medical Innovation and Technology*, pp 154–158
- Glocker B, Sotiras A, Komodakis N, Paragios N (2011) Deformable medical image registration: setting the state of the art with discrete methods. *Annu Rev Biomed Eng* 13:219–244. doi:10.1146/annurev-bioeng-071910-124649
- Kappes JH, Andres B, Hamprecht FA, Schnörr C, Nowozin S, Batra D, Kim S, Kausler BX, Lellmann J, Komodakis N, Rother C (2013) A comparative study of modern inference techniques for discrete energy minimization problem In: *CVPR 2013*
- Komodakis N (2011) Efficient training for pairwise or higher order crfs via dual decomposition. In: *CVPR*, pp 1841–1848
- Komodakis N, Tziritas G, Paragios N (2007) Fast, approximately optimal solutions for single and dynamic mrf. In: *Computer vision and pattern recognition, 2007. CVPR'07. IEEE Conference on*, pp 1–8
- Kotsas P, Dodd T (2011) A review of methods for 2d/3d registration. *WASET Conference Paris*, pp 14–16
- Lee K, Kwon D, Yun I, Lee S (2008) Deformable 3d volume registration using efficient mrf model with decomposed nodes. In: *British Machine Vision Conference*, pp 1–10
- Mahapatra D, Sun Y (2008) Nonrigid registration of dynamic renal mr images using a saliency based mrf model. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2008*, pp 771–779
- Markelj P, Tomaževič D, Likar B, Pernuš F (2012) A review of 3d/2d registration methods for image-guided interventions. *Med Image Anal* 16(3):642–661
- Marks L, Young S, Natarajan S (2013) Mri-ultrasound fusion for guidance of targeted prostate biopsy. *Curr Opin Urol* 23(1):43
- Mercier L, Del Maestro RF, Petrecca K, Araujo D, Haegelen C, Collins DL (2012) Online database of clinical mr and ultrasound images of brain tumors. *Med Phys* 39:3253
- Osechinskiy S, Kruggel F (2010) Slice-to-volume nonrigid registration of histological sections to Mr images of the human brain. *Anatomy Research International* 2011. doi:10.1155/2011/287860
- Penney G, Blackall J, Hayashi D, Sabharwal T, Adam A, Hawkes D (2001) Overview of an ultrasound to ct or mr registration system for use in thermal ablation of liver metastases. In: *Proceedings of Medical Image Understanding and Analysis, Citeseer*, vol 1, p 6568
- San José Estépar R, Westin C, Vosburgh K (2009) Towards real time 2d to 3d registration for ultrasound-guided endoscopic and laparoscopic procedures. *Int J Computer Assist Radiol Surg* 4(6):549–560
- Shekhovtsov A, Kovtun I, Hlaváč V (2008) Efficient mrf deformation model for non-rigid image matching. *Computer Vision Image Underst* 112(1):91–99. doi:10.1016/j.cviu.2008.06.006
- Xu H, Lasso A, Fedorov A, Tuncali K, Tempny C, Fichtinger G (2014) Multi-slice-to-volume registration for mri-guided transperineal prostate biopsy. *Int J Computer Assist Radiol Surg*, pp 1–10 CARS. doi:10.1007/s11548-014-1108-7
- Zikic D, Glocker B, Kutter O, Groher M, Komodakis N, Kamen A, Paragios N, Navab N (2010) Linear intensity-based image registration by Markov random fields and discrete optimization. *Med Image Anal* 14(4):550–562. doi:10.1016/j.media.2010.04.003