Manifold-enhanced Segmentation through Random Walks on Linear Subspace Priors

Pierre Yves Baudin[1,2,4−6]
pierre-yves.baudin@ecp.fr

Noura Azzabou[5−7]
n.azzabou@institut-myologie.org

Pierre Carlier[5−7]
p.carlier@institut-myologie.org

Nikos Paragios[2−4]
nikos.paragios@ecp.fr

[1] SIEMENS Healthcare,
Saint Denis, FR

[2] Center for Visual Computing,
Ecole Centrale de Paris, FR

[3] Université Paris-Est,
LIGM (UMR CNRS),
Center for Visual Computing,
Ecole des Ponts ParisTech, FR

[4] Equipe Galen, INRIA Saclay,
Ile-de-France, FR

[5] Institute of Myology,Paris, FR

[6] CEA, I^{2}BM, MIRCen,
IdM NMR Laboratory, Paris, FR

[7] UPMC University Paris 06,
Paris, FR

Abstract

In this paper we propose a novel method for knowledge-based segmentation. Our contribution lies on the introduction of linear sub-spaces constraints within the random-walk segmentation framework. Prior knowledge is obtained through principal component analysis that is then combined with conventional boundary constraints for image segmentation. The approach is validated on a challenging clinical setting that is multi-component segmentation of the human upper leg skeletal muscle in Magnetic Resonance Imaging, where there is limited visual differentiation support between muscle classes.

1 Introduction

Although there is an abundant computer vision literature on the subject of automatic image segmentation, many concrete problems still wait to be solved, in particular in segmentation of organs in medical applications. Our particular interest lies in segmentation of skeletal muscles in 3D Magnetic Resonance Imaging, which poses some very specific issues: simultaneous multi-object segmentation, non-discriminative appearance of the muscles, partial contours between them, large inter-subject variations, spurious contours due to fat infiltrations.

Vastly different approaches have been used in the relatively small literature addressing automatic muscle segmentation. In [6, 7], segmentation is achieved with multi-object deformable models. Deformable models ([11, 12]) are surface models superimposed to the

target image in an initial state and deformed by minimizing an objective function composed of two terms: a data-term - attracting the surface to detected image contours - and a regularization term, ensuring the surface remains realistically smooth or similar enough to a reference surface (as in [2]). In [4], prior knowledge on the shape of one muscle is further enforced by imposing the model to reside in a hierarchical shape space built via diffusion wavelets decomposition on training examples. Relying on iterative local optimization procedures, deformable models are very sensitive to the proximity of the initialization state with the expected solution, and usually only yield non-global optima of the objective function. In [14], the muscle surface is represented with a set of landmarks in a graph framework, and a higher-order pose invariant shape prior is encoded by learning the distances within triplets of landmarks. Then, the surface is fit to the image though a discrete optimization procedure allowing to solve the objective function globally. Such method implies one can find consistent anatomical landmark points, an assumption which is hardly verified for muscles. Apart from surface models, a pixel-wise region-based approach was proposed in [1] for segmenting muscles with good results. In this framework, a segmentation is a labeling of the pixels of the target image. Shape prior is enforced by learning a low-dimensional space of allowed segmentations through Principal Component Analysis on a set of training segmentations projected in the Isometric Log-Ratio space. Optimization of the convex objective function is carried out through gradient descent. Like surface-based methods, this method explicitly relies on preliminary contour detection, a difficult and often unreliable process in muscle images.

Recent advances in numerical methods for computing globally large discrete regions-based segmentation methods - Graph-Cuts ([3]) and Random Walks ([10]) - have made them practical for large images. Due to being graph-based, these methods are intrinsically bottom-up - pixel-wise information has an influence on the whole segmentation. In particular, such property makes Random Walks notoriously robust to partial contours. On the other hand, few methods exist for enforcing prior knowledge on object shapes in such framework. Every method known to us proposes at best an affine deformation of a template shape[5, 13].

In this paper, we propose a segmentation method based on Random Walks, in which shape deformation is constrained to remain close to a Principal Component Analysis shape space built from training examples. Using the PCA allows us to model complex non-rigid shape variations relying on a few eigen-modes. Our method also benefits from the high performances of the RW optimization process, as it only requires a simple addition to the RW objective function.

This paper is organized as follow: in section 2, we review the RW segmentation framework, and describe how to include a PCA shape prior. Then in section 3, we demonstrate the potential of our method on a challenging data set of 3D muscle MR volumes. Section 4 sums up and concludes this paper.

2   Random walk segmentation with prior knowledge

2.1   Random walk formulation

In this section we briefly review the random walks approach for segmentation, based on the fully detailed presentation of [10]. A segmentation is formulated as a labeling problem of an undirected weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is a set of nodes and $\mathcal{E}$ is a set of edges. The i-th node v_i corresponds to the i-th pixel of I, the image to be segmented. Given a set of K
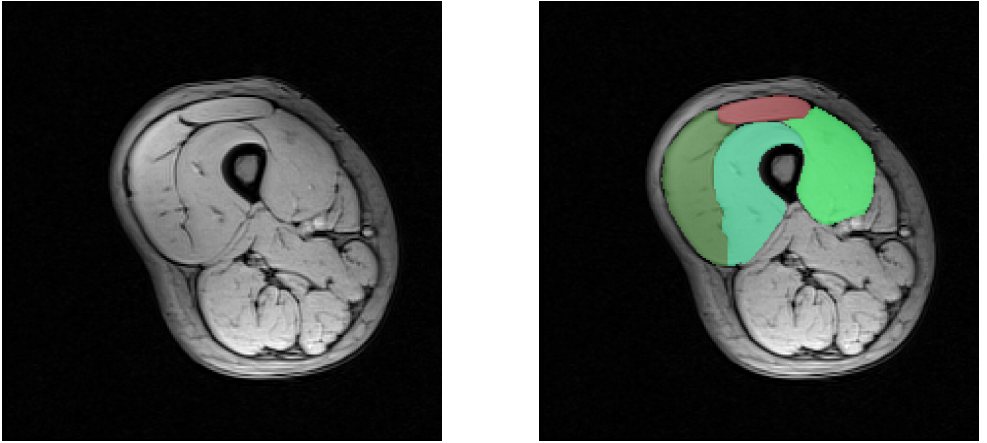
Figure 1: (left) Cross-section of an MR volume of the thigh. (right) Manual segmentation of four muscles.

labels $\mathcal{S}$ (e.g. the indices of the muscles), segmenting image I consists of assigning a label $s \in \mathcal{S}$ to each node $p \in \mathcal{V}$.

The RW method amounts to computing the probability x_i^s that the node $v_i \in \mathcal{V}$ is assigned to the label s. Assume we possess a set $\mathcal{V}_M$ of pre-labeled nodes for each label, also called *seeds* (typically, manually marked pixels). We denote the set of non pre-labeled nodes as $\mathcal{V}_U$. It was shown ([10]) that all *unknown* (non pre-labeled) entries of $x^s = [x_1^s, x_2^s, \ldots, x_N^s]^T$ - i.e. the probability that each node $v_i \in \mathcal{V}_U$ is assigned to label s - minimize the functional:

$$E_{\mathrm{RW}}^s (x^s) = x^{sT} L x^s \qquad (1)$$

where the *known* entries of x^s (the seeds) are set as follow:

$$\forall v_i \in \mathcal{V}_M, \; x_i^s = \begin{cases} 1 & \text{if pixel } i \text{ is marked with label } s \\ 0 & \text{if pixel } i \text{ is marked with another label} \end{cases} \qquad (2)$$

and where L is the combinatorial Laplacian matrix of the graph, defined as:

$$L_{i,j} = \begin{cases} \sum_k w_{kj} & \text{if } i = j \\ -w_{ij} & \text{if } (i,j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases} \qquad (3)$$

where w_{ij} is the weight assigned to the edge spanning vertices i and j. It is common to define w_{ij} as a Gaussian weighting function :

$$w_{ij} = \omega + \exp\left(-\beta \left(g_i - g_j\right)^2\right) \qquad (4)$$

where g_i is the gray-level of pixel i, β is a scaling parameter set accordingly to the image contrast, and ω is an regularization parameter which amounts to penalizing the gradient norm of x^s (no regularization if $\omega = 0$). Once computed x^s for each label s, the segmentation is obtained by retaining the label of maximum probability: $\hat{s}_i = \arg\max_s x_i^s$.

Since minimizing (1) is an independent process for each label, the whole RW process can be equivalently synthesized in one equation:

$$E_{\text{RW}}(x) = x^T \bar{L} x \tag{5}$$

where $x^T = \left[x^{1\ T} | x^{2\ T} | \dots | x^{K\ T}\right]$ ($x \in \mathbb{R}^{KN \times 1}$) is the concatenation of all the x^s in one vector, and

$$\bar{L} = \begin{bmatrix} L & \cdots & 0 \\ \vdots & L & \vdots \\ 0 & \cdots & L \end{bmatrix} \in \mathbb{R}^{KN \times KN} \tag{6}$$

is a block diagonal matrix with L repeated K times on the diagonal - i.e. as many times as there are labels. We keep this formulation for the remaining of the paper.

2.2 Prior shape model

Prior knowledge to the RW formulation was introduced in [9], through an estimate of the probability distribution of the gray-level intensity for each label. Such prior could be enforced by minimizing this cost function:

$$E^s_{\text{RWP}}(x^s) = x^{s\ T} L x^s + \gamma \left(x^{sT} \Lambda x^s - 2 x^{sT} \lambda^s\right) \tag{7}$$

where $\lambda^s(i)$ is the probability that the intensity at pixel i belongs to the intensity distribution for label s, and $\Lambda = \text{diag}\left(\sum_s \lambda^s\right)$ (we refer the reader to [9] for details).

In the context of muscle segmentation, since the intensity distributions for the labels (the muscles) are extremely similar, such prior would not be very efficient. As we could not find any discriminative features (textures, consistent anatomical points, etc.) to use within this framework, we designed a pixel-wise *shape* model, trained from a series of ground-truth segmentations.

Assume we possess such a shape space, in the form of an affine model (details in section 2.3) - i.e. any vector $\tilde{x}$ inside this space takes the form:

$$\tilde{x} = \Gamma \gamma + \bar{x}$$

where $\bar{x}$ is the mean shape, $\Gamma \in \mathbb{R}^{KN \times k}$ is the shape space matrix (e.g. obtained through PCA), $\gamma \in \mathbb{R}^k$ is the coordinate of $\tilde{x}$ within the shape space. Thus, vector x can be expressed as:

$$x = dx + \Gamma \gamma + \bar{x} \tag{8}$$

where $dx \in \mathbb{R}^{KN \times 1}$ is the deviation of x from the shape space. In order to obtain a segmentation which remains close to the shape space, we want to minimize the objective function 5 with respect to both dx and γ, while keeping dx small. This leads us to the following functional:

$$E_1(dx, \gamma) = (dx + \Gamma \gamma + \bar{x})^T \bar{L} (dx + \Gamma \gamma + \bar{x}) + \lambda \|dx\|^2 \tag{9}$$

where λ is a free parameter setting a constraint on the norm of dx. We reformulate 9 as:

$$E_2(y) = (Ay + \bar{x})^T \bar{L} (Ay + \bar{x}) + \lambda y^T B y \tag{10}$$

with

$$y = \begin{bmatrix} dx \\ \gamma \end{bmatrix}, \quad A = [I_{KN}\ \Gamma], \quad B = \begin{bmatrix} I_{KN} & 0 \\ 0 & 0 \end{bmatrix}. \tag{11}$$

where I_{KN} is the identity matrix of size $KN \times KN$.

The minimum of (10) verifies:

$$\left(A^T \bar{L} A + \lambda B \right) y = A^T \bar{L} \bar{x} \tag{12}$$

Although the system of equations (12) is invertible even when all entries of x are unknown (no seeds), as noted in [9], it is still possible - and useful - to use seeds in combination with a prior to obtain more robust segmentations.

2.3 Shape space

Any segmentation corresponding to a vector of probability x belongs to $\mathbb{R}^{KN \times 1}$. However, it is obvious that the entries of x are not independent from each other, but instead are highly tied to each other, at least for sets of close-by pixels. This implies that there exists an implicit lower dimensional space in which any true segmentations reside. The principle of a shape space is to design a low dimensional affine space approximating this implicit space. If we succeed, we expect any true segmentation to lie "not too far" from the shape space, and the projection of that segmentation into the shape space to be a good approximation of the segmentation itself. Principal Component Analysis notoriously provides a simple mean to compute such a shape space.

Assume we possess T co-registered segmented training volumes. We perform the PCA on vectors $\left\{ \hat{x}^i \right\}_{i=1...T}$, which are ground truth segmentations represented as probability vectors (actually binary vectors, since $x_i^s = 1$ if pixel i has label s, and 0 otherwise). Retaining the n eigen-modes of greatest variance, the projection of any segmentation in the shape space is represented as:

$$\tilde{x} = \bar{x} + \Gamma \gamma$$

where $\Gamma \in \mathbb{R}^{KN \times n}$ is a rectangular matrix whose columns are the n retained eigen-vectors, and $\gamma \in \mathbb{R}^{n \times 1}$ is the vector of coordinates of $\tilde{x}$ in the shape space.

3 Experimental validation

In this section, we present results obtained with our method on a set of 3D MR volumes of muscles. Our data set is comprised of 30 volumes of the right thigh of healthy subjects, covering a wide range of ages and morphologies. These volumes were acquired with a 3T Siemens scanner using a 3-point Dixon sequence ($TR = 10ms$, $TE_1 = 2.75ms$, $TE_2 = 3.95ms$, $TE_3 = 5.15ms$, rf flip angle $= 3°$) with resolution $1mm \times 1mm \times 5mm$ and initial size $224 \times 224 \times 128$. Four clinically relevant muscles were manually segmented - *rectus femoris*, *vastus lateralis*, *vastus intermedius*, *vastus medialis*, providing the ground truth segmentation against which the output of our method will be compared.

In order to compute the shape space (see section 2.3), we non-rigidly register all volumes and segmentation atlas to the same reference volume, using the method in [8] and the related software (Drop, at www.mrf-registration.net). Adopting a leave-one-out validation protocol, all the volumes in the training set, bar one, are registered to the remaining volume, on which we test our segmentation method.

Given the large size of the matrices, we solved the system of equations 12 with iterative methods like Bi-Conjugate Gradient. Computation time is approximately 15 min per
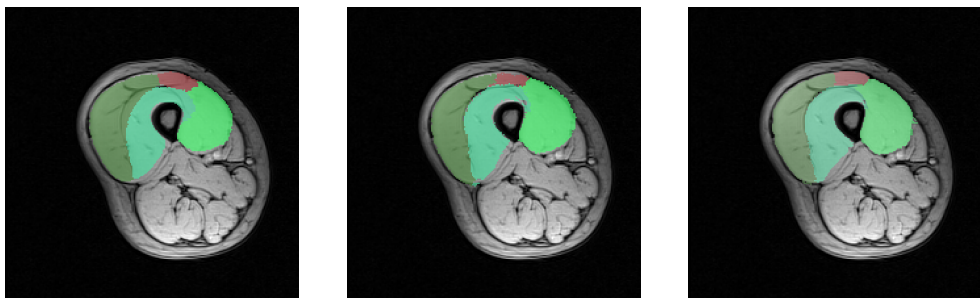
Figure 2: Effect of the PCA shape prior: (left) mean segmentation using $x = \bar{x}$, (middle) shape space segmentation using $x = \Gamma\gamma + \bar{x}$, (right) segmentation with shape prior using $x = dx + \Gamma\gamma + \bar{x}$. Notice that the shape space segmentation fits the boundaries better than the mean segmentation, but has fuzzy contours due to the approximation of projecting complex shapes into a linear subspace.
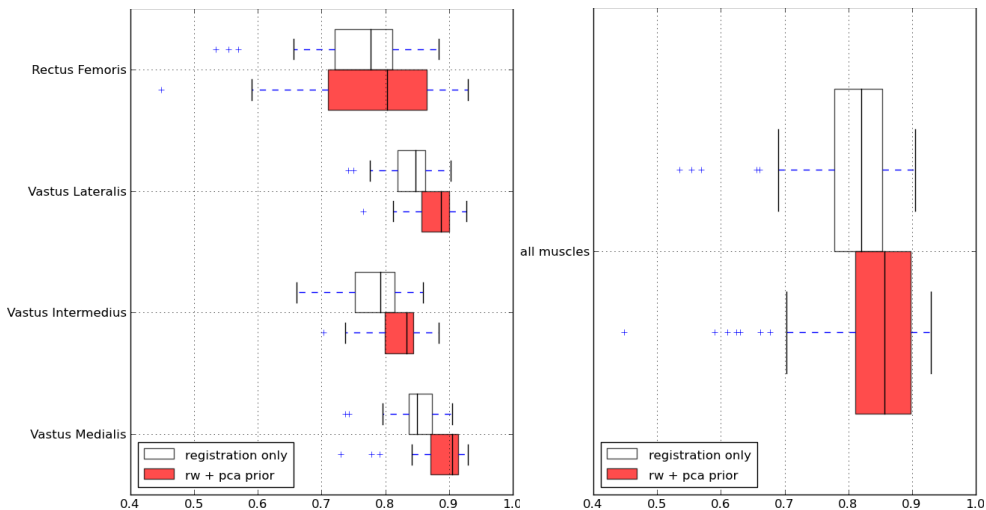

Figure 3: Segmentation results as Dice coefficients. This paper's method is represented in red, and the mean segmentation (using $x = \bar{x}$) is shown in white.
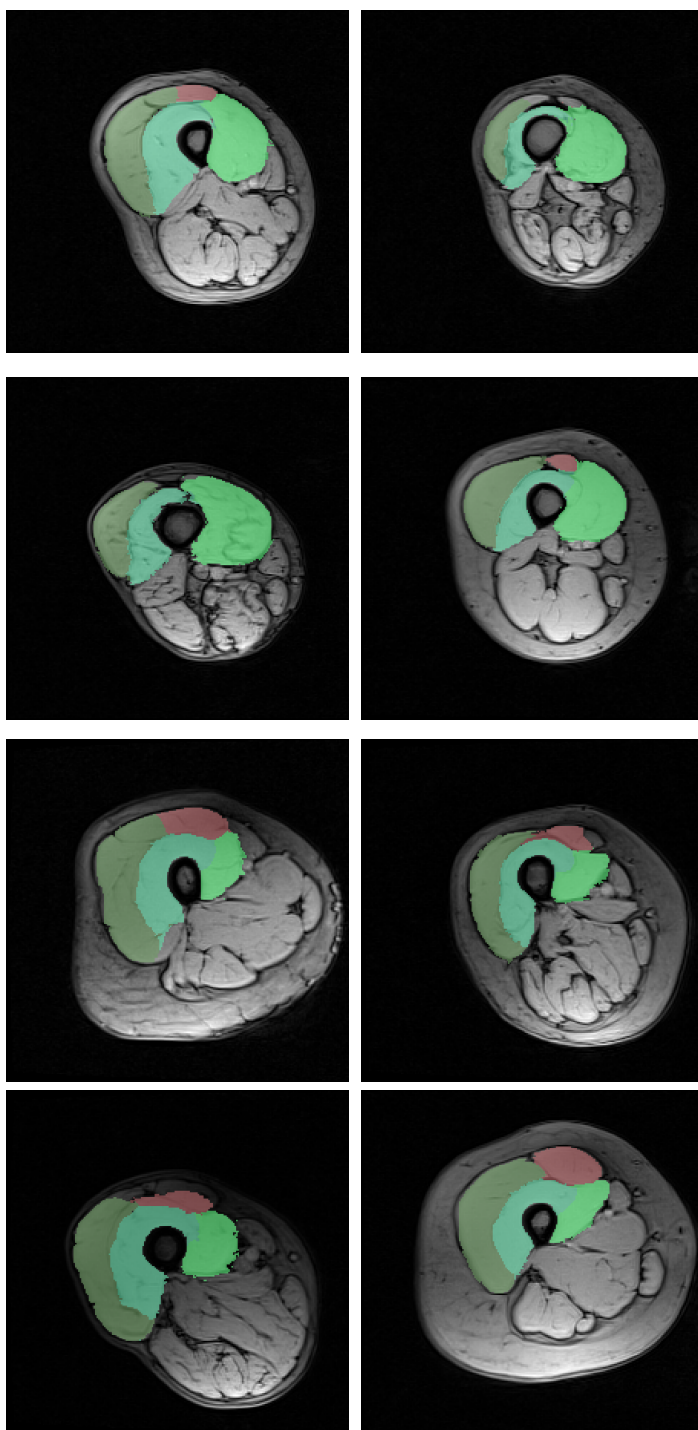
Figure 4: Examples of segmentation results

segmentation on a 2.8 GHz Intel process with 4GB of RAM. The parameter values, set empirically, were: $\beta = 100$, $\omega = 10^{-2}$, $\lambda = 10^{-2}$, $n = 30$.

On figure 2, we show the effect of the PCA shape prior on one example of our dataset. We see that the PCA model (formulated as $\Gamma \gamma + \bar{x}$) is deformed from its initial shape (i.e. $\bar{x}$) to fit the boundaries of the test image. Due to the rough approximation of modeling probabilistic atlases with a linear subspace, the muscles' boundaries are not well defined. However, when allowing a small deviation (i.e. dx) from the model, as does our method, the contours are more accurately defined. Notice there remains segmentation errors (the top red muscle in figure 2), as the PCA model is not capable of deforming enough to fit unusual shapes while retaining a realistic topology.

We assess the quality of the segmentation using Dice coefficients, whose expression is: $Dice = 2 |T \cap R| / (|T| + |R|)$, where T and R are the pixel sets for the same muscle in the computed segmentation and the ground truth segmentation respectively. We obtain an average Dice coefficient of 0.84 ± 0.08. On figure 3 we compare the results obtained with our method to those obtained when using $x = \bar{x}$ (We called it "registration only"). The latter method is equivalent to a segmentation through multi-atlas registration, with majority label voting. A previous method of ours [2] yielded an average Dice of 0.80 ± 0.19 for a different but similar dataset. In this method, we automatically generated appropriate seed positions with respect to the muscle classes. Unlabeled seeds were sampled across the test image, and were automatically labeled through a graph-based approach considering visual and topological properties. The resulting seeds were then fed to the standard RW algorithm [10].

In the presented method, the main segmentation errors are due to the muscles with unusual shape - e.g. more elongated, smaller, etc. Modeling probabilistic segmentations with PCA model does not allow representing shapes which differ too much from standard shapes. However, the amount of deformation that is allowed by such modeling is enough to capture a large number of common shapes and thus is valid in many cases. Some segmentation results are shown in figure 4.

4   Conclusion

In this paper, we presented a novel multi-object segmentation method with prior knowledge, based on the Random Walks segmentation algorithm. We designed our approach to be capable of handling muscle segmentation in 3D MRI, where partial contours, large inter-subject variations, absence of texture differences, etc. render typical segmentation methods inefficient. Building from the RW formulation, we add a shape space prior term to the objective function. Our shape space is built from PCA on a set of training segmentations. To reveal the potential of our method, we presented segmentation results on MR volumes of the thigh on 30 subjects.

Our model is capable of fitting most shapes encountered in our dataset. However, due to the approximation induced by the projection into a linear subspace, some more extreme shapes are not segmented well. In the future, we intend to design a model not relying on PCA and capable of handling more varied shapes.

References

[1] Shawn Andrews, Ghassan Hamarneh, Azadeh Yazdanpanah, Bahareh HajGhanbari, and W Darlene Reid. Probabilistic multi-shape segmentation of knee extensor and flexor muscles. In *MICCAI*, volume 14, pages 651–8. January 2011. URL http://www.ncbi.nlm.nih.gov/pubmed/22003755.

[2] Pierre-Yves Baudin, Noura Azzabou, Pierre G. Carlier, and Nikos Paragios. Automatic Skeletal Muscle Segmentation Through Random Walks and Graph-Based Seed Placement. In *ISBI*, pages 1036–1039, 2012.

[3] Yuri Boykov and M P Jolly. Interactive graph cuts for optimal boundary & region segmentation of objects in ND images. In *ICCV*, volume 1, pages 105–112. IEEE, 2002. ISBN 0769511430. URL http://iehttp//citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.129.8181&rep=rep1&type=pdfexplore.ieee.org/xpls/abs_all.jsp?arnumber=937505.

[4] Salma Essafi, Georg Langs, and Nikos Paragios. Hierarchical 3D diffusion wavelet shape priors. In *ICCV*, pages 1717–1724. IEEE, September 2009. ISBN 978-1-4244-4420-5. doi: 10.1109/ICCV.2009.5459385. URL http://dspace.mit.edu/handle/1721.1/59305http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5459385.

[5] D. Freedman. Interactive Graph Cut Based Segmentation with Shape Priors. In *CVPR*, volume 1, pages 755–762. Ieee, 2005. ISBN 0-7695-2372-2. doi: 10.1109/CVPR.2005.191. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1467344.

[6] Benjamin Gilles and Nadia Magnenat-Thalmann. Musculoskeletal MRI segmentation using multi-resolution simplex meshes with medial representations. *Medical image analysis*, 14(3):291–302, June 2010. ISSN 1361-8423. doi: 10.1016/j.media.2010.01.006. URL http://www.ncbi.nlm.nih.gov/pubmed/20303319.

[7] Benjamin Gilles and Dinesh K. Pai. Fast musculoskeletal registration based on shape matching. In *MICCAI*, volume 5242 of *Lecture Notes in Computer Science*, pages 822–829. January 2008. URL http://www.ncbi.nlm.nih.gov/pubmed/18982681.

[8] Ben Glocker, Nikos Komodakis, Georgios Tziritas, Nassir Navab, and Nikos Paragios. Dense image registration through MRFs and efficient linear programming. *Medical image analysis*, 12(6):731–41, December 2008. ISSN 1361-8423. doi: 10.1016/j.media.2008.03.006. URL http://www.ncbi.nlm.nih.gov/pubmed/18482858.

[9] L. Grady. Multilabel Random Walker Image Segmentation Using Prior Models. In *CVPR*, volume 1, pages 763–770. Ieee, 2005. ISBN 0-7695-2372-2. doi: 10.1109/CVPR.2005.239. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1467345.

[10] Leo Grady. Random walks for image segmentation. *Pattern Analysis and Machine Intelligence*, 28(11):1768–1783, 2006. URL http://www.ncbi.nlm.nih.gov/pubmed/17063682.

[11] T McInerney and D Terzopoulos. Deformable models in medical image analysis: a survey. *Medical image analysis*, 1(2):91–108, June 1996. ISSN 1361-8415. URL http://www.ncbi.nlm.nih.gov/pubmed/9873923http:// www.sciencedirect.com/science/article/pii/s1361-8415(96) 80007-7.

[12] Johan Montagnat, Hervé Delingette, and Nicolas Ayache. A review of deformable surfaces: topology, geometry and deformation. *Image and Vision Computing*, 19(2001): 1023 – 1040, 2004. URL http://www.sciencedirect.com/science/ article/pii/S0262885601000646.

[13] Nhat Vu. Shape prior segmentation of multiple objects with graph cuts. In *CVPR*, pages 1–8. Ieee, June 2008. ISBN 978-1-4244-2242-5. doi: 10.1109/CVPR.2008.4587450. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm? arnumber=4587450http://ieeexplore.ieee.org/xpls/abs_all. jsp?arnumber=4587450.

[14] Chaohui Wang, Olivier Teboul, Fabrice Michel, Salma Essafi, and Nikos Paragios. 3D knowledge-based segmentation using pose-invariant higher-order graphs. In *MIC-CAI*, volume 6363 of *Lecture Notes in Computer Science*, pages 189–196. 2010. URL http://www.ncbi.nlm.nih.gov/pubmed/20879399.