Edge- and Detail-Preserving Sparse Image Representations for Deformable Registration of Chest MRI and CT Volumes

Mattias P. Heinrich[1,2,⋆], Mark Jenkinson[2], Bartlomiej W. Papież[1], Fergus V. Glesson[3], Sir Michael Brady[4], and Julia A. Schnabel[1]

[1] Institute of Biomedical Engineering, University of Oxford, UK
[2] Oxford University Centre for Functional MRI of the Brain, UK
[3] Department of Radiology Churchill Hospital, Oxford, UK
[4] Department of Oncology, University of Oxford, UK
mattias.heinrich@eng.ox.ac.uk
http://users.ox.ac.uk/~shil3388

Abstract. Deformable medical image registration requires the optimisation of a function with a large number of degrees of freedom. Commonly-used approaches to reduce the computational complexity, such as uniform B-splines and Gaussian image pyramids, introduce translation-invariant homogeneous smoothing, and may lead to less accurate registration in particular for motion fields with discontinuities. This paper introduces the concept of sparse image representation based on supervoxels, which are edge-preserving and therefore enable accurate modelling of sliding organ motions frequently seen in respiratory and cardiac scans. Previous shortcomings of using supervoxels in motion estimation, in particular inconsistent clustering in ambiguous regions, are overcome by employing multiple layers of supervoxels. Furthermore, we propose a new similarity criterion based on a binary shape representation of supervoxels, which improves the accuracy of single-modal registration and enables multi-modal registration. We validate our findings based on the registration of two challenging clinical applications of volumetric deformable registration: motion estimation between inhale and exhale phase of CT scans for radiotherapy planning, and deformable multi-modal registration of diagnostic MRI and CT chest scans. The experiments demonstrate state-of-the-art registration accuracy, and require no additional anatomical knowledge with greatly reduced computational complexity.

Keywords: supervoxels, sliding motion, multi-modal fusion, pulmonary.

1 Introduction

Registering medical images of three (or more) dimensions with high spatial image resolution results in a highly complex optimisation problem, which is usually

⋆ We thank EPSRC and Cancer Research UK for funding this work in the Oxford Cancer Imaging Centre.

J.C. Gee et al. (Eds.): IPMI 2013, LNCS 7917, pp. 463–474, 2013.

impractical to solve directly. A common approach to address this is the use of coarse-to-fine optimisation schemes. Representing the images by a Gaussian or spline pyramid [13], in which the spatial resolution is decreased by a certain factor between pyramid levels, is a popular strategy to tackle the challenges involved with high resolution data. It involves first solving a relaxed problem using a low-dimensional data representation and then refining the solution at a finer scale. A disadvantage of such multi-resolution approaches is the loss of detail at lower resolutions, which can and does lead to errors that cannot be compensated for at a finer scale. In addition, most approaches are limited by the use of a Cartesian grid representation and restricted spatially uniform subsampling. Furthermore, multi-resolution schemes do not preserve image boundaries and edges, which has a negative influence when estimating complex motion (with discontinuities).

In this work, we discuss the use of an alternative, low parametric, image representation, namely supervoxels. The motivation for the use of supervoxels is their ability to group voxels, which are close both **spatially and visually** into perceptually meaningful clusters. The locations of cluster centres do not have to lie on a regular grid, making them more versatile than traditional parametric image representations such as B-splines (with the exception of the recent work on sparse free-form deformations [12]). Supervoxels are not restricted to be uniformly shaped, enabling a better preservation of image edges. Figure 1 shows an example over-segmentation of an axial slice of a CT scan using 250 superpixels and the image reconstruction, obtained by assigning the mean intensity value of each cluster to each pixel within a superpixel. Note, that although only 2D superpixels are shown all steps in our implementation use 3D supervoxels. The advantages of preserving image edges and small anatomical detail is shown in comparison to a low-parametric B-spline representation. The disadvantage is that because superpixel segmentation relies on a piecewise constant image model, it shows poor performance in homogeneous or gradually changing image regions, resulting in inconsistent clustering. This reason has largely restricted the use of superpixels to image segmentation or classification tasks. These shortcoming will be addressed in this paper.

2 Related Work Using Supervoxels

Motion estimation or tracking using supervoxels relies on the assumption that all pixels contained in a cluster have the same (translational) motion. This precondition improves the robustness against image noise, reduces the ambiguity in homogeneous regions, and substantially reduces the complexity of the optimisation problem, in turn simplifying global regularisation of the motion field. Another advantage of this piecewise constant motion model is that it enables the preservation of motion discontinuities, so long as they coincide with intensity steps. In [9], stereo depth estimation is performed based on an over-segmentation of one of the two views. Employing a segmentation in only one view does not reduce the space of the potential motion vectors, hence extending this approach to higher dimensional problems substantially increases the complexity. An alternative approach, which is used in this work, is to perform supervoxel clustering

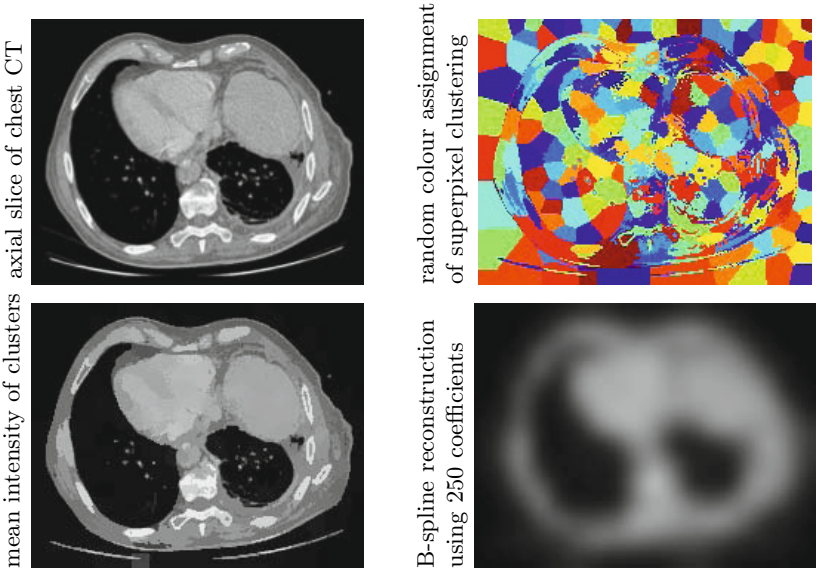

Fig. 1. Examples of sparse image representations. Image representation using 250 superpixels is able to preserve image edges and small-scale structures. While a reconstruction using 250 equally spaced coefficients [13] loses most fine details.

in both images and restrict the motion to corresponding cluster centres, thus directly matching segments across images. The main challenge here is the inconsistency of supervoxel clustering across images. [17] attempts to address this challenge by alternating steps of segmentation and matching to obtain a final segmentation, which is consistent across images.

The concept we introduce in this paper overcomes the disadvantages of previous approaches by using multiple layers of supervoxels. Sections 3.1 and 3.2 describe how two volumes (to be registered) are clustered into layers of supervoxels, which enable a better image representation in ambiguous (homogeneous or gradually changing intensity) regions, while at the same time preserving important details and edges. Sec. 3.2 shows that these complementary (not hierarchical) 3D layers of supervoxels form a low-parametric image representation with similar qualities as the popular joint bilateral filter [8] with very low computational complexity. We make further use of the clustering by introducing a new similarity criterion between single- and multi-modal scans, using a binary representation of the shape of the supervoxel (see Sec. 3.3). Section 3.4 describes the graph based optimisation, which uses a discretised, sparse displacement space. For each 3D layer, a separate graph connects supervoxels of the reference image, which are close both spatially and based on intensity. A combined energy term (Eq. 7) consisting of a diffusion regularisation term and similarity criterion is optimised using belief propagation [3] to obtain piecewise smooth transformations. The optimal transformations for all layers are then finally combined voxelwise in the original image domain (see Sec. 3.5).

3 Methods

3.1 Supervoxel Clustering

Supervoxel clustering performs an over-segmentation of an image that respects image boundaries. Supervoxels remove the redundant intensity information of voxels within homogeneous areas, which are likely to belong to the same object. Supervoxels enable a more compact image representation with little loss of detail. Due to their flexibility they are also more adaptive to the local shape of the images. Because they significantly reduce the computationally complexity, they have attracted a lot of attention in a range of image analysis tasks like stereo matching [9], optical flow [17], and segmentation [10].

We adapt a very recent algorithm, "simple linear iterative clustering" (SLIC) [1] for supervoxel clustering. Its complexity is linear w.r.t. the number of pixels N and therefore easily applicable to large datasets. It clusters voxels based on their grayscale similarity and spatial Euclidean distance. The algorithm is designed to create approximately K equally-sized supervoxels $\mathbf{S} = \{S_1, S_2, \ldots, S_k\}$. It starts from a set of equally spaced seed cluster centres with distance $s \approx \sqrt[3]{N/K}$. The distance d_{ik} between a voxel $p_i = [l_i, x_i, y_i, z_i]^T$ (where l is intensity value) and a cluster centre $S_k = [l_k, x_k, y_k, z_k]^T$ is given by Eq. 1.

$$d_{ik}^v = |l_k - l_i| \ , \ d_{ik}^{xyz} = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2 + (z_k - z_i)^2}$$
$$d_{ik} = d_{ik}^v + \frac{m}{s} d_{ik}^{xyz} \tag{1}$$

The weighting m determines the compactness of the clusters, where higher values result in more regularly shaped supervoxels. It is assumed that the spatial extent of a supervoxel lies within a compact search region R of spatial extent $3s$ x $3s$ x $3s$. Therefore each voxel p_i has to be compared only to all centres which are within R and is afterwards assigned to the closest cluster (see Eq. 3.1):

$$S(p_i) = \underset{S_j \in R}{\arg\min}\, d_{ij} \tag{2}$$

After one pass over all pixels, the cluster centres are recomputed. This process is iterated until the clusters no longer change, or a fixed iteration number is reached. The algorithm does not guarantee connectedness within clusters, thus a simple connectivity enforcing step could be performed afterwards to eliminate stray labels. This method achieves substantial improvements in computation time compared to state-of-the-art methods, while yielding similar or better segmentation accuracy [1]. An example output of the method is displayed in Fig. 1, in which we compare the representation of an axial CT slice with the same number (250) of free parameters, using the supervoxels clustering and then a cubic B-spline basis function. It can be seen that the supervoxels better preserve image details and edges.

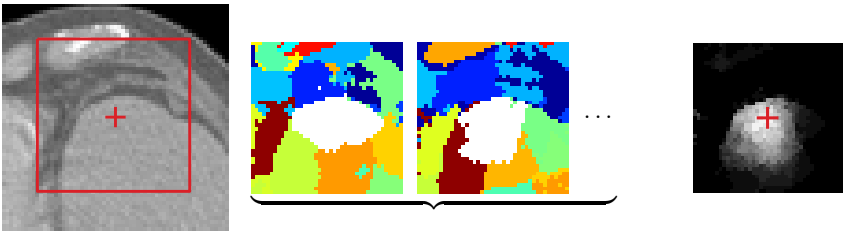

Fig. 2. Concept of layers of supervoxels for particular region (anterior part of liver) in axial CT plane, with central voxel p_c marked with red cross. Multiple clusterings $\mathbf{S}$ are obtained using different seeds. Shape of supervoxels is used as similarity criterion (see Sec. 3.3). Linear combination of all clusters of which p_c is part of, results in a spatial weighting similar to the bilateral filter.

3.2 Multiple Layers of Supervoxels

A disadvantage of supervoxel clustering is its inconsistent clustering in homogeneous or gradually changing image regions. In the context of motion estimation this is a major limitation, because it is important which correspondence within a homogeneous region is selected. One of the main contributions of this work is to use multiple layers of supervoxels to obtain a piecewise smooth motion model for accurate deformable registration.

To create multiple layers of supervoxels, the clustering algorithm of Sec. 3.1 is run several times with slightly different initialisation (random offset of seed locations with maximum magnitude $s/2$). Image regions with sufficient structural content (e.g. edges) are not affected by this disturbance and the clustering is therefore very similar for all layers. Homogeneous or gradually changing areas do not provide sufficient guidance for the supervoxel, resulting in arbitrarily different clusters for each layer. This means that when a separate optimisation is performed for each layer of supervoxels, the combination of these transformation is a smooth average in homogeneous regions, but adheres to discontinuities at image boundaries. Figure 2 demonstrates this concept. Two different layers of supervoxels of an axial CT slice are shown. The linear combination of the spatial layout of supervoxels of different layers for one particular voxel is shown. The close relation to the joint bilateral filter [8], and the ability to model smooth regions and at the same time preserve edges can be clearly seen. Furthermore, the supervoxel based approach enables a low parametric representation for each layer, which is beneficial for the optimisation, and the complexity of the clustering is independent of the number or size of the clusters.

3.3 Shape of Supervoxel as Similarity Criterion

The spatial context of a voxel's intensity has been used in many applications as it provides a good descriptor of local geometric and anatomical structure [5].

The census transform [16] uses a binary representation $\mathbf{B}$ of local shape. For each voxel p_i in an image, $\mathbf{B}$ is obtained by comparing its intensity l_i to the intensities l_n of all voxels in a certain spatial neighbourhood $\mathcal{I}$ and setting $B(p_i, p_n) = 1$ if $l_n < l_i$ and $B(p_i, p_n) = 0$ if $l_n \geq l_i$ for all $p_n \in \mathcal{I}$. It has been shown that these representations are well suited to define similarity across images, because they are sensitive to edges, local orientation and invariant to monotonic grayscale transformations. However, it cannot be employed for multi-modal images where gradients of corresponding anatomies might be reversed. We therefore propose to use the shape of supervoxels as a multi-modal similarity criterion. Given a supervoxel clustering $\mathbf{S}$, the binary vector $\mathbf{B}_i$ for a voxel p_i is defined as:

$$B(p_i, p_n) = \begin{cases} 1, & \text{if } S(p_i) = S(p_n) \\ 0, & \text{if } S(p_i) \neq S(p_n) \end{cases} \text{ for all } p_n \in \mathcal{I} \tag{3}$$

We show examples of this binary shape representation in Fig. 2. We define the neighbourhood $\mathcal{I}$ to be the subset of the 320 closest points p_n (multiples of 64 are well suited for computation) based on the distance $|d_{in}^{xyz} - r|$ of p_n to the surface of a sphere centred at p_i with radius of $r = \frac{1}{2} s \sqrt[3]{6/\pi}$. The binary representation has the advantage that the L_1 distance between two vectors $|\mathbf{B}_i - \mathbf{B}_j|$ can be efficiently evaluated by the Hamming weight (bit count of all pair-wise unequal binary values [16]). The computation for a vector of size 64 takes less time than calculating a single absolute difference of two intensities.

3.4 MRF-Based Optimisation

Discrete optimisation methods have become very popular for deformable registration, due to their flexibility, computational efficiency and guarantees of optimality [4]. Our aim is to find the best correspondence for each supervoxel $S_p \in \mathbf{S}$ of the reference scan, in terms of similarity and subject to a global regularisation. This optimisation problem is solved independently for each 3D layer in the reference scan. The combination of several optimal transformations and the reverse mapping into the voxelwise image domain is detailed in Sec. 3.5. The set of potential displacement labels $\mathbf{f}_p = \{f_p^1, f_p^2, \ldots\}$ is defined for a supervoxel S_i as the sparse set of all supervoxel locations within a rectangular search window W (centred around $\mathbf{x}_p$) in any layer of the moving image (see Fig. 3 for a visual example). Each label f_p^j corresponds to a spatial displacement $\Delta\mathbf{x}(f_p^j) = \mathbf{x}_p - \mathbf{x}_j$ between reference and moving scan. The similarity cost of a label f_p based on the intensities I and I' (normalised to [0,1]) of reference and moving scan respectively is then defined as:

$$D(f_p) = |I(\mathbf{x}_i) - I'(\mathbf{x}_i + \Delta\mathbf{x}(f_p))| \tag{4}$$

Likewise, the similarity criterion based on the local shape of supervoxels, which was introduced in Sec. 3.3, can now be defined for a certain label f_p:

$$\bar{D}(f_p) = \sum_{n \in \mathcal{I}} |B(\mathbf{x}_p, n) - B'(\mathbf{x}_p + \Delta\mathbf{x}(f_p), n)| \tag{5}$$

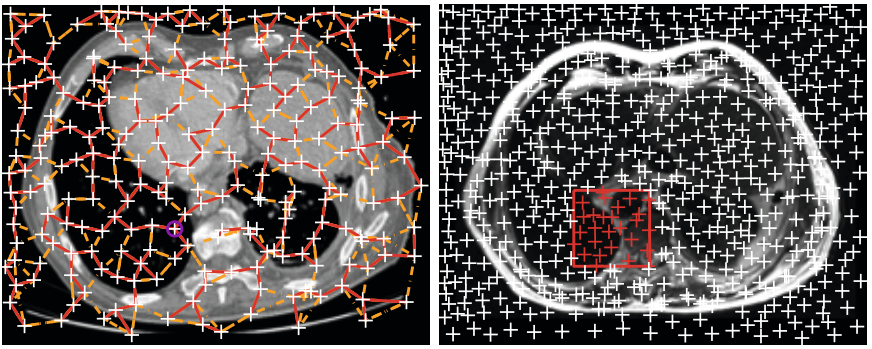

Fig. 3. Concept of supervoxel matching across scans. Cluster centres are denoted by white crosses. Left: Edges of kNN spanning graph in the reference image are shown with yellow dashed lines, edges which are also part of the Euclidean minimum spanning tree (EMST) with red lines. The possible displacements for a voxel of interest (purple circle) within a rectangular window in moving image are shown in red. Note that all 3D layers of the the moving image are considered simultaneously, resulting in more cluster centres.

In order to enforce spatial regularity of the displacements of all supervoxels, a graph is introduced, which connects all clusters that are close both spatially and in intensity. A diffusion regularisation term $R(f_p, f_q)$ is defined between two displacements f_p and f_q of two clusters, which penalises the squared Euclidean distance of displacements divided by the distance of the supervoxels (see Eq. 1):

$$R(f_p, f_q) = \frac{||\Delta\mathbf{x}(f_p) - \Delta\mathbf{x}(f_q)||^2}{\frac{s}{m}|l_p - l_q| + ||\mathbf{x}_p - \mathbf{x}_q||} \tag{6}$$

Notice that the regularisation between clusters with different appearances is reduced. This enables the preservation of discontinuities, which are likely to coincide with image boundaries. Since the cluster centres are non-uniformly distributed across the image grid, conventional neighbourhood connections do not apply. We employ two steps to obtain a suitable graph representation. First, a k-nearest neighbour (kNN) graph with edges $\mathcal{E}$ is extracted based on the supervoxel cluster distances (see Eq. 1). The value of k is incrementally increased until a spanning graph (connecting all supervoxels in the reference image) is found. For large numbers of supervoxels K the naïve nearest neighbour search with K^2 complexity would be impractical, so we employ an accelerated (∼50x) search using the vantage-point tree (see [15] for details).

The global energy term to be minimised is a weighted combination of intensity similarity D, shape similarity $\bar{D}$ and a diffusion regularisation term:

$$E(f) = \underbrace{(1-\alpha)\sum_{p\in\mathbf{S}} D(f_p)}_{\text{intensity similarity}} + \underbrace{\alpha\sum_{p\in\mathbf{S}} \bar{D}(f_p)}_{\text{shape similarity}} + \underbrace{\lambda\sum_{(p,q)\in\mathcal{E}} R(f_p, f_q)}_{\text{diffusion regularisation}} \tag{7}$$

Using the kNN graph, this energy can be minimised with a number of discrete optimisation methods (see [7] for an overview). Similar to [6], we make a further approximation to employ belief propagation on a tree (BP-T), which guarantees a global optimum of the energy on this relaxed graph. By iteratively removing edges with highest distance between clusters (and therefore smallest influence on the regularisation) we find a Euclidean minimum-spanning-tree, which is shown with red lines in Fig. 3.

3.5 Linear Combination of Optimal Transformations

Once the optimal transformation is found for each 3D layer in the reference image, the final deformation field can be obtained as a linear combination of all transformations. For each voxel p_i the resulting displacement is found by averaging over all displacement vectors of the particular supervoxel of which p_i is part in the respective 3D layer. This results in a spatial weighting of displacement vectors, which is similar to the bilateral filter, as displayed in Fig. 2. The advantage is that this linear combination can both preserve discontinuities in the motion field and model gradually changing motion magnitudes.

4 Experiments

We perform experiments on two clinical 3D datasets to demonstrate the suitability of our contributions. First, the benefits of using multiple layers of supervoxels to model piecewise smooth, edge-preserving deformations is evaluated on 4D-CT scans. Second, the similarity criterion based on the shape of supervoxels is applied successfully to the multi-modal fusion of chest CT and MRI scans.

4.1 Inhale-Exhale Volumetric CT Registration

Sliding motion is a particular challenge when estimating motion of lung CT scans acquired at different phases of the breathing cycle, which is clinically useful for radiotherapy of lung cancer and the assessment of breathing disorders. Most approaches to this problem have used a manual or automatically detected segmentation of the thoracic cage [11], [14]. Our approach relies solely on the over-segmentation from the supervoxel clustering to preserve motion boundaries and requires no further manual interaction. We perform our experiments on the extreme phases (inhale and exhale) of 4D respiratory CT scans of 5 lung cancer patients. The dataset has been made publicly available by the DIR-lab, University of Texas, manually annotated with 300 landmark per scan pair [2]. We use the most challenging cases #6-10, which have a very large diaphragm displacement and particularly strong sliding motion at the lung/rib cage interface.

With a slice thickness of 2.5 mm, the volumes have strongly anisotropic voxel-sizes, so the axial resolution was resampled to 1.5 mm. The scans were slightly cropped to exclude regions outside the body. We found that around 15000 supervoxels are sufficient to represent the details of these images, which yields

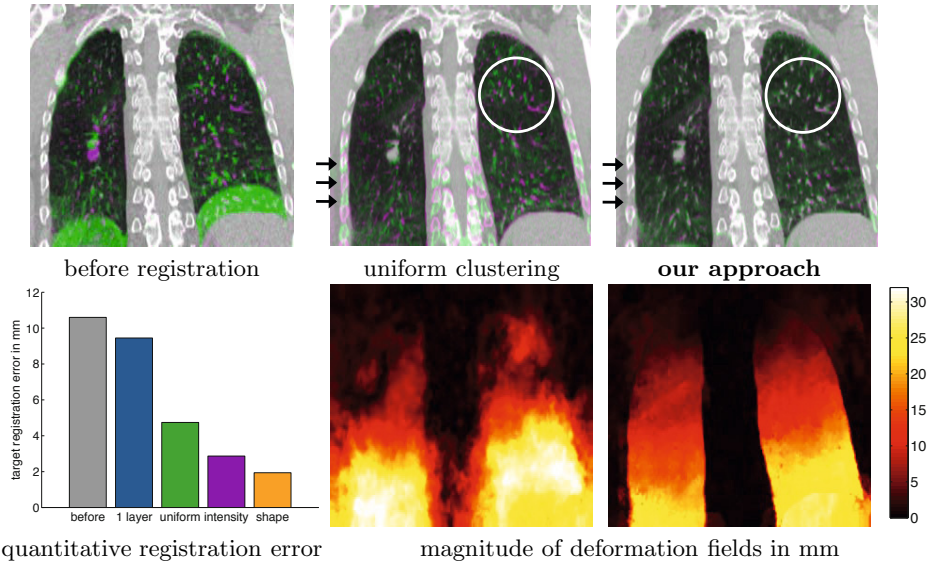

Fig. 4. Example of deformable registration of an inhale-exhale CT scan pair. Overlay before and after registration is shown in green (inhale phase) and magenta (exhale phase). Uniform clustering approximates the traditional coarse-scale representation. Our approach using layers of supervoxels outperforms this for the matching of small vessels (see white circle) and the preservation of sliding motion (see black arrows). The quantitative evaluation shows a significantly lower registration error of our method using absolute intensity differences and a further improvement when employing the binary shape similarity criterion.

an average distance of 5 voxels between cluster centres. The supervoxel compactness parameter m is chosen empirically as 20 (for an intensity range of [0, 255] and similar to [1]). We restrict the spatial extent of the (sparsely sampled) displacement space to ± 9 voxels in the axial plane and ± 15 voxels in proximal-distal direction (to cover the expected respiratory motion). For each 3D layer of supervoxels the computation takes roughly 5 sec. for the clustering, 1 sec. for similarity evaluation and graph extraction and 4 sec. for the BP-T optimisation using a C++ implementation on a quad-core CPU. We used 20 layers of supervoxels, tuned all parameters on a single scan pair and found their setting to be relatively insensitive. For the regularisation parameter $\lambda = \frac{1}{8}$ is chosen, but doubling or halving the value leads to a deviation of less than 10% in accuracy. The registration quality was examined visually and evaluated quantitatively using 300 manually annotated landmarks per scan pair.

Four different variations of our proposed method were tested to assess the individual influence of our contributions. The initial average landmark displacement was **10.6±7.5 mm**. First, our method was applied using only a **single layer** of supervoxels, similar to previous work on supervoxel matching. It yielded unsatisfactory results, as only the outer lung boundaries are aligned. Second, multiple layers were used and the image-adaptivity of the supervoxels was substantially

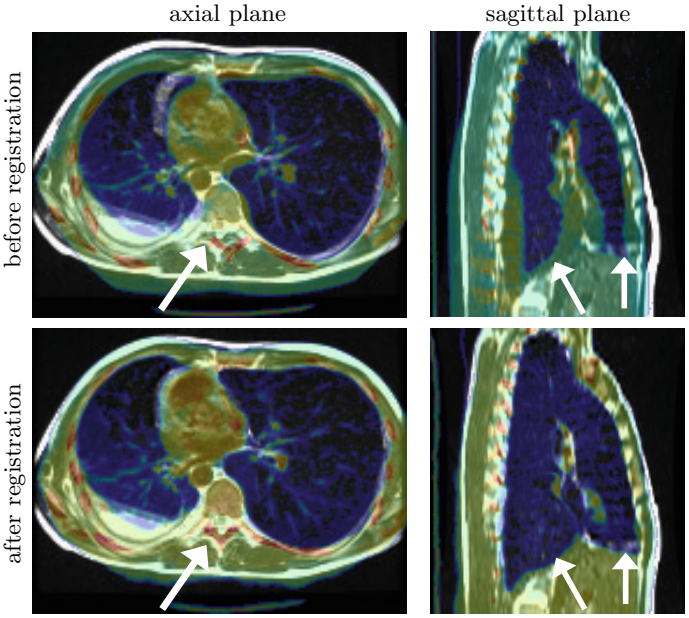

Fig. 5. Example of multi-modal fusion using the presented method. The sagittal and axial planes of a CT/MRI scan pair are shown. The MRI scan is displayed with greyscale intensities, while the CT scan is shown in pseudo-colours (red indicates high intensities and blue low intensities). A clearly improved alignment can be seen after deforming the CT scan with the estimated deformation field.

reduced ($m = 2000$) resulting in a very **uniform clustering** and a target registration error (TRE) of **4.72±3.5 mm**. Figure 4 demonstrates the main problem of this approach, which is similar to a traditional coarse-scale image representation. The motion field is smooth across the interface at which discontinuous sliding motion occurs (see black arrows). Additionally small details (lung vessels) are lost due to uniform smoothing, resulting in an inaccurate alignment of them (see white circle). Third, our approach was tested using only **intensity-based similarity** ($\alpha = 0$ in Eq. 7), achieving a significant improvement and a TRE of **2.87±1.9 mm**. Finally, the **shape similarity** $\bar{D}$ was included with a weighting of $\alpha = \frac{4}{5}$. Adding this structural image information, with negligible computational complexity, has clear advantages to match fine image details and further reduces the registration error to **1.94±1.3 mm**. Figure 4 demonstrates the accurate alignment of our approach and the well-preserved sliding motion at the thoracic cage.

4.2 Multi-modal Fusion of MRI and CT Chest Scans

Additionally, we applied our method to five pairs of longitudinal diagnostic MRI and CT scans from patients with lung diseases. The additional challenges here

are lower scan quality in the MRI, large slice thicknesses of up to 8 mm, and pathological changes. The volumes have been manually cropped to a similar field of view (compensating global translation) and resampled to form isotropic voxels of 2x2x2 mm^3. The parameter settings were kept the same as before, except now only the shape similarity criterion was used $\alpha = 1$, since there is no direct relation of intensities between CT and MRI. The number of supervoxels was set to 15000 as before, which resulted in a spacing of 7 voxels. The search space is defined to be ± 10 voxels in all dimensions. Figure 5 shows a successful multi-modal fusion of MRI in greyscale with the deformed CT as pseudo-colour overlay. A small number of anatomical landmarks (12 per scan) has been manually annotated by a radiologist, this tasks, however, is very difficult and resulted in a large intra-observer error ($\geq$ 5mm). The landmark distance could be reduced significantly (p =0.028) from **10.43±7.1 mm** to **7.33±4.3 mm** for this challenging fusion.

5 Conclusion

We have presented a novel concept for sparse image representation with application to deformable registration and fusion. Using multiple layers of supervoxels, each representing the grouped voxels with one constant intensity and motion vector, enables a very efficient piecewise smooth transformation model with very low computational complexity. We have demonstrated that our approach is able to deal well with complex lung motion containing both smooth deformations and discontinuities. Our resulting TRE of **1.94 mm** for the registration of inhale-exhale CT compares favourable with the published results of [11] and [14], which specifically address the challenge of sliding motion using a segmentation mask, and achieve a TRE of **2.68 mm** and **2.09 mm** for these particularly challenging scan pairs (#6-10 of [2]). It is worth noting that they found similar results (TRE 4.27 and 4.23 mm) when not using a lung segmentation compared to our uniform clustering variant, which confirms the advantages of the presented image-adaptive supervoxel clustering. Our second contribution, the formulation of a binary shape representation based on the clusters as a similarity criterion, improves accuracy of single-modal registration and enables multi-modal fusion.

While our results already demonstrate state-of-the-art performance for the given tasks, further improvements are possible (e.g. by employing twice as many layers of supervoxels, the TRE of the 4DCT registration is reduced to 1.75 mm – at the cost of higher computational complexity). Making use of the symmetry of the registration problem by enforcing inverse consistency would further improve the robustness. The use of an optimisation strategy, which can include a larger set of edges (e.g. loopy BP [3]) is likely to perform better. Rather than resampling the scans to avoid large anisotropy of voxel-sizes, isotropic supervoxels could be calculated directly using world instead of voxel coordinates. Finally, this concept has potential use in many other image analysis tasks, such as probabilistic segmentation and parameter map estimation.

References

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S.: SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans. Pattern Anal. Mach. Intell. 34(11), 2274–2282 (2012)
2. Castillo, E., Castillo, R., Martinez, J., Shenoy, M., Guerrero, T.: Four-dimensional deformable image registration using trajectory modeling. Phys. Med. Biol. 55(1), 305 (2009)
3. Felzenszwalb, P., Huttenlocher, D.: Efficient Belief Propagation for Early Vision. Int. J. Comp. Vis. 70(1), 41–54 (2006)
4. Glocker, B., Komodakis, N., Tziritas, G., Navab, N., Paragios, N.: Dense image registrations through MRFs and efficient linear programming. Med. Imag. Anal. 12(6), 731–741 (2008)
5. Heinrich, M.P., Jenkinson, M., Bhushan, M., Matin, T., Gleeson, F.V., Brady, M., Schnabel, J.A.: MIND: Modality independent neighbourhood descriptor for multimodal deformable registration. Med. Imag. Anal. 16(7), 1423–1435 (2012)
6. Heinrich, M., Jenkinson, M., Brady, M., Schnabel, J.: MRF-based deformable registration and ventilation estimation of lung CT. IEEE Trans. Med. Imag. (2013)
7. Kolmogorov, V., Rother, C.: Minimizing nonsubmodular functions with graph cuts - a review. IEEE Trans. Pattern Anal. Mach. Intell. 29(7), 1274–1279 (2007)
8. Kopf, J., Cohen, M.F., Lischinski, D., Uyttendaele, M.: Joint bilateral upsampling. ACM Trans. Graph. 26(3), 96 (2007)
9. Lei, C., Selzer, J., Yang, Y.H.: Region-tree based stereo using dynamic programming optimization. In: CVPR, pp. 2378–2385. IEEE (2006)
10. Lucchi, A., Smith, K., Achanta, R., Knott, G., Fua, P.: Supervoxel-based segmentation of mitochondria in EM image stacks with learned shape features. IEEE Trans. Med. Imag. 31(2), 474–486 (2012)
11. Schmidt-Richberg, A., Werner, R., Handels, H., Ehrhardt, J.: Estimation of slipping organ motion by registration with direction-dependent regularization. Med. Imag. Anal. 16(1), 150–159 (2012)
12. Shi, W., Zhuang, X., Pizarro, L., Bai, W., Wang, H., Tung, K.P., Edwards, P., Rueckert, D.: Registration using sparse free-form deformations. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part II. LNCS, vol. 7511, pp. 659–666. Springer, Heidelberg (2012)
13. Unser, M.A., Aldroubi, A., Gerfen, C.R.: Multiresolution image registration procedure using spline pyramids. In: Int. Symp. on Optics, Imaging, and Instrumentation, SPIE, pp. 160–170 (1993)
14. Vandemeulebroucke, J., Bernard, O., Rit, S., Kybic, J., Clarysse, P., Sarrut, D.: Automated segmentation of a motion mask to preserve sliding motion in deformable registration of thoracic CT. Med. Phys. 39, 1006 (2012)
15. Yianilos, P.N.: Data structures and algorithms for nearest neighbor search in general metric spaces. In: ACM-SIAM Symp. on Discrete Algorithms, pp. 311–321 (1993)
16. Zabih, R., Woodfill, J.: Non-parametric local transforms for computing visual correspondence. In: Eklundh, J.-O. (ed.) ECCV 1994. LNCS, vol. 801, pp. 151–158. Springer, Heidelberg (1994)
17. Zitnick, C.W., Jojic, N., Kang, S.B.: Consistent segmentation for optical flow estimation. In: ICCV, pp. 1308–1315. IEEE (2005)