Multispectral Image Registration Based on Local Canonical Correlation Analysis

Mattias P. Heinrich[1], Bartłomiej W. Papież[2], Julia A. Schnabel[2], and Heinz Handels[1]

[1] Institute of Medical Informatics, University of Lübeck, Germany
[2] Institute of Biomedical Engineering, Department of Engineering, University of Oxford, UK
heinrich@imi.uni-luebeck.de
http://www.mpheinrich.de

Abstract. Medical scans are today routinely acquired using multiple sequences or contrast settings, resulting in multispectral data. For the automatic analysis of this data, the evaluation of multispectral similarity is essential. So far, few concepts have been proposed to deal in a principled way with images containing multiple channels. Here, we present a new approach based on a well known statistical technique: canonical correlation analysis (CCA). CCA finds a mapping of two multidimensional variables into two new bases, which best represent the true underlying relations of the signals. In contrast to previously used metrics, it is therefore able to find new correlations based on linear combinations of multiple channels. We extend this concept to efficiently model local canonical correlation (LCCA) between image patches. This novel, more general similarity metric can be applied to images with an arbitrary number of channels. The most important property of LCCA is its invariance to affine transformations of variables. When used on local histograms, LCCA can also deal with multimodal similarity. We demonstrate the performance of our concept on challenging clinical multispectral datasets.

Keywords: multichannel, canonical correlation, multimodal, MRI.

1 Introduction

Multispectral imaging, in particular multi-sequence magnetic resonance imaging (MRI), is increasingly becoming available in clinical practice. Image registration forms an integral part in the analysis pipeline for computer aided diagnosis and interventions based on medical imaging. However, to this date, few algorithms have been proposed to explicitly handle multichannel image data [1,2]. A further difficulty arises when not all exact same sequences are available for all patients in a study. In order to establish correspondences across multichannel scans of different patients (or with respect to an atlas), image registration relies on a robust similarity metric. For different MRI sequences, such as T1-weighted, T2-weighted or fluid attenuated inversion recovery (FLAIR), a certain degree of

P. Golland et al. (Eds.): MICCAI 2014, Part I, LNCS 8673, pp. 202–209, 2014.

correlation across channels can be expected, since they are all based on the same physical principle of the magnetisation of water protons. It is however difficult to establish a priori which channels correlate best with each other and represent a true correspondence based on the underlying physiology.

In order to deal with similarity of multichannel images one approach is to average the cross-correlations [3] calculated for each channel individually. This, however, disregards all cross-channel correlation. Generalised correlation coefficients for diffusion tensor images have been proposed in [4], which are used to calculate a scalar correlation value (but also ignore cross-channel correlations). This concept was extended to multi-tensor images in [5] using the matrix logarithm of diffusion matrices. The linear correlation of linear combination (LC^2) metric [6], captures the similarity of a multichannel and a scalar image. Image synthesis based on aligned training data [7], or general polynomial models [8] have been used to predict the appearance of a different modality or contrast. Multichannel demons [1], use squared intensity differences as force fields derived independently from multiple channels. Multi-variate mutual information has been proposed in e.g. [2], but certain approximations have to be made to overcome the large complexity. In [9], image registration is performed by choosing the optimal subset of Gabor wavelet features of multichannel images, using independent component analysis and a 'choose max' fusion strategy.

In this work, we overcome the challenges of multispectral image similarity using a established statistical technique: canonical correlation analysis (CCA) [10]. CCA has been previously used in the context of medical image analysis, e.g. to detect neural activity in functional MRI [11]. It has also been used as a metric in [12] for log tensor images, calculated over disjunct blocks, and was called generalised correlation coefficient in that work. CCA measures the linear relationships between two multi-dimensional random variables (see Sec. 2.1). We apply CCA to define similarity across multichannel images, because it allows us to find a mapping of the images into a new space where they maximally correlate. The canonical correlation is invariant to affine transformations or permutations of the input variables. Since linear relations between channels in medical images do not hold globally, we propose an efficient scheme to evaluate the **local** canonical correlation (LCCA). For this purpose, we extend the recently proposed guided filter [13], which uses a (multichannel) image as a guidance to filter a second scalar input image (in our work both images have multiple channels). The filter can be implemented using box filters, with a computational complexity independent of the local neighbourhood size. We make a further contribution by applying LCCA to local histogram images enabling multimodal registration. In Sec. 3, we demonstrate the state-of-the-art performance of our approach for two challenging clinical multispectral datasets, evaluated with manual segmentations.

2 Methods

Let us consider two general 3D multichannel images $\mathbf{I}$ and $\mathbf{J}$, where each location x is represented by an intensity vector of length m and n respectively. Given a

neighbourhood radius r, image patches $\mathbf{X}(x)$ (in $\mathbf{I}$) and $\mathbf{Y}(x)$ (in $\mathbf{J}$) are defined at every x within a spatial window Ω_x with a size of $(r \times 2 + 1)^3$. The individual channels of each patch are denoted by a subscript: $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$. The similarity at location x between two patches can be defined for the most trivial case of two scalar images as the normalised correlation coefficient (NCC):

$$\rho^2(X_1, Y_1, x) = \frac{(\sum_{\Omega_x}(X_1 - \bar{X_1})(Y_1 - \bar{Y_1}))^2}{\sum_{\Omega_x}(X_1 - \bar{X_1})^2 \sum_{\Omega_x}(Y_1 - \bar{Y_1})^2} \tag{1}$$

where $\bar{X_i} = \frac{1}{|\Omega_x|}\sum_{\Omega_x} X_i$ represents the mean intensity of an image patch. The scalar NCC is invariant to any affine intensity transform $X' = aX + b$ of the image patches. It can be extended to multichannel images (MC-NCC), defining a vectorial correlation coefficient $\boldsymbol{\rho}$:

$$\boldsymbol{\rho}^2(\mathbf{X}, \mathbf{Y}, x) = \{\rho^2(X_1, Y_1, x), \ldots, \rho^2(X_{d_{\min}}, Y_{d_{\min}}, x)\} \tag{2}$$

where $d_{\min} = \min(m, n)$ is the minimal channel dimensionality. We note that $\boldsymbol{\rho}$ is not well adapted to images with differing numbers of channels $m \neq n$, since a choice of which extra channel to drop has to be made. Furthermore, it is not invariant to arbitrary linear transformations of the input variables.

2.1 Canonical Correlation Analysis (CCA)

CCA is able to overcome these limitations by finding the two bases $\mathbf{w}_X$ and $\mathbf{w}_Y$, which maximise the sum over $\boldsymbol{\rho}$ for the transformed variables $\mathbf{X} \cdot \mathbf{w}_X$ and $\mathbf{Y} \cdot \mathbf{w}_Y$. In order to find the canonical basis vectors, we first need to construct a full correlation matrix D. Let us define a $m \times n$ cross-covariance matrix $\boldsymbol{\Sigma}_{\mathbf{XY}}$, where the entry in the kth row and lth column is defined to be:

$$\Sigma_{X_k Y_l} = \frac{1}{|\Omega_x|}\sum_{\Omega_x}(X_k - \bar{X_k})(Y_l - \bar{Y_l}). \tag{3}$$

The variance matrices $\boldsymbol{\Sigma}_{\mathbf{XX}}$ and $\boldsymbol{\Sigma}_{\mathbf{YY}}$ are defined analogously. The full correlation matrix D then becomes:

$$D(\mathbf{X}, \mathbf{Y}) = \boldsymbol{\Sigma}_{\mathbf{XX}}^{-1} \boldsymbol{\Sigma}_{\mathbf{XY}} \boldsymbol{\Sigma}_{\mathbf{YY}}^{-1} \boldsymbol{\Sigma}_{\mathbf{XY}}^{T} \tag{4}$$

The basis vectors can be obtained by performing an eigendecomposition, where $\mathbf{w}_X$ is build up from the eigenvectors of $D(\mathbf{X}, \mathbf{Y})$ and $\mathbf{w}_Y$ from the eigenvectors of $D(\mathbf{Y}, \mathbf{X})$. The canonical correlations are the eigenvalues $\boldsymbol{\lambda}_{\mathbf{XY}} = \{\lambda_1, \ldots, \lambda_{d_{\min}}\}$ of D. A scalar similarity metric $\mathcal{S}(\mathbf{X}, \mathbf{Y})$ of two patches can now be defined based on the sum over $\boldsymbol{\lambda}_{\mathbf{XY}}$ divided by $d_{\min}$. Note, that even though the number of non-zero eigenvalues is limited by $\min(rank(m), rank(n))$ every image channel has the same influence weight on calculating the eigenvalues. Another simplification can be obtained, based on the fact that for any matrix the sum of eigenvalues equals the trace of that matrix. So finally we get:

$$\mathcal{S}(\mathbf{X}, \mathbf{Y}, x) = 1 - \frac{1}{d_{\min}} trace(D(\mathbf{X}, \mathbf{Y}, x)) \tag{5}$$

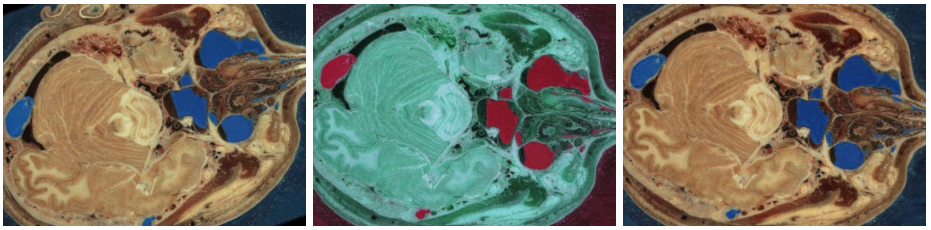

Fig. 1. Visual example of (global) canonical correlation analysis applied to colour images. Left: **I** slice of the visible human dataset with its original colour, but rotated by 20°. Centre: **J** same slice without rotation, but with cyclically shifted hue. Right: $\mathbf{J}'$ reconstructed true slice estimated using CCA (of **I** and **J**) and $\mathbf{J}' = \mathbf{w}_Y \mathbf{w}_Y^{-1} \mathbf{J}$.

In contrast, to Eq. 2 the canonical correlation is invariant to affine intensity mappings of multichannel images. Figure 1 demonstrates the capabilities of finding an intensity mapping, which affects all image channels simultaneously even when one image is geometrically transformed. $\mathcal{S}$ is symmetric and also independent of the order of $\mathbf{X}$ and $\mathbf{Y}$. When one of the images is scalar $n = 1$, the correlation matrix D becomes a scalar, and it can be shown that LCCA is equivalent to the linear correlation of linear combination (LC^2) metric [6] for this special case.

Local Canonical Correlation (LCCA) Computation: Linear correlations rarely hold for the whole image, yet to get a point-wise local evaluation of the similarity metric, we have to calculate D for every voxel, using all $(2r+1)^3$ voxels within the patch Ω_x. Fortunately, the calculation can be simplified without loss of accuracy using box filters, following the ideas of guided image filtering [13]. The similarity has to be evaluated for every voxel, so the complexity is greatly reduced and independent of the patch-size by first rearranging Eq. 3:

$$\sum_{\Omega_x} (X - \bar{X})(Y - \bar{Y}) = \sum_{\Omega_x} XY - \bar{X}\bar{Y} \tag{6}$$

and then replacing the summation in Eq. 6 by a convolution kernel K_r. The moving average kernel has a constant complexity independent of kernel size and the local means $\bar{X}$ and $\bar{Y}$ can be precomputed (again by convolution). For the purpose of image registration, where multiple displacements have to be evaluated, we can furthermore pre-calculate the inverse variance matrices $\boldsymbol{\Sigma}_{\mathbf{XX}}^{-1}$ and $\boldsymbol{\Sigma}_{\mathbf{YY}}^{-1}$ once (per iteration). Thus only mn box filter convolutions (for the whole image) and two matrix multiplications (for every voxel) are required. For a typical dual channel image with one million voxels this enables us to process the patch-wise similarity for every voxel and up to 40 displacements per second.[1]

Non-functional Intensity Mappings Using Local Histograms: A limitation of NCC for medical image registration is that it is usually not applicable to multimodal images. We present an interesting second application of LCCA to address this shortcoming. Given a pair of scalar images I and J from different

[1] Source code for LCCA will be made public at `http://www.mpheinrich.de/software.html`

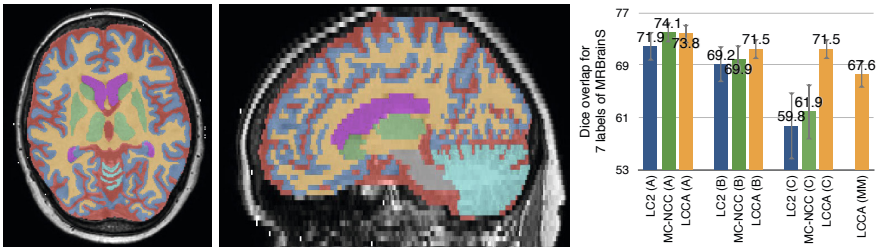

Fig. 2. Left-centre: MRBrainS dataset with manual segmentations: ■ Cortical gray matter, ■ Basal ganglia, ■ White matter and White matter lesions, ■ Cerebrospinal fluid in the extracerebral space, ■ Ventricles, ■ Cerebellum, ■ Brainstem. Right: Quantitative results of 4 multi-spectral registration experiments on training set of 5 patients. MC-NCC and LCCA both outperform LC^2 for experiment A, while LCCA yields higher accuracy for all other tests (descriptions see text) than MC-NCC or LC^2.

modalities, they can be transformed into multichannel images $\mathbf{I} = \{I_1, \ldots, I_m\}$ and $\mathbf{J} = \{J_1, \ldots, J_m\}$, where each channel represents the distance from a range of m quantised intensity value $\mathbf{v} = \{v_1, \ldots, v_m\}$, with $I_k = \exp(-(I - v_k)^2/\sigma^2)$. Here, σ represents the uncertainty in image intensity similar to the Parzen window in mutual information calculation. This representation enables LCCA to model even non-functional intensity mappings. In practice a very low number m of histogram channels is sufficient. For this application the invariance of LCCA to permutations of the channels is of great importance.

Regularisation and Optimisation: LCCA is employed in a discrete non-rigid registration framework, which we have recently published in [14]. An explicit search is performed over a discrete displacement space $\mathbf{d} \in \mathcal{L} = \{0, \pm q, \ldots, \pm q_{\max}\}^3$ (with quantisation step q) using the patch-wise formulation of CCA as similarity metric. Starting from an initial estimate using the argmin $\mathbf{v}$ of the discrete search, a globally smooth displacement field $\mathbf{u}$ is iteratively estimated by two alternating steps. First, a Gaussian smoothing $\mathbf{u} \leftarrow K_\sigma \mathbf{v}$ of the current field is performed. Second, an auxiliary term $\frac{1}{\theta}(\mathbf{u} - \mathbf{v})^2$ is added to the local similarity distributions and a new argmin $\mathbf{v}$ selected. When the parameter θ is subsequently reduced, a convergence ($\mathbf{u} = \mathbf{v}$) is reached after few iterations.

3 Experiments

We compare the performance of the presented canonical correlation analysis based similarity metric (LCCA) to multichannel NCC (MC-NCC) and the LC^2 metric on two different multispectral image datasets.

Multispectral Brain Image Registration: First, we employ data from the multispectral brain segmentation challenge (MRBrainS) held at MICCAI 2013 [15]. The organisers provided training datasets of five patients with T1-weighted, T1 with inversion recovery (IR) and FLAIR MRI modalities. The scans have a resolution of 1x1x3 mm, are rigidly aligned (to the FLAIR sequence) and manually segmented into 8 labels (since white matter lesions cannot be detected

using atlas-based segmentation, we merge labels 3 and 4). Pair-wise registrations between all five patients are performed using a three-level approach of our registration framework and the following parameters: down-sampling factors of $\{3, 2, 1\}$, maximal displacement range of $q_{\max} = \{6, 2, 1\}$ and patch-radius $r = 2$ voxels. The Dice overlap before registration is on average 45.5%. We perform four different experiments using all of the three metrics, whenever they are applicable on dual-channel MRI scans (within the same optimisation framework):

- **A**: both channels are of the same modality for both moving and target image and in the right order (e.g. $\mathbf{I} = \{T1, T2\}, \mathbf{J} = \{T1, T2\}$)
- **B**: one channel is of the same modality for both images the other channel is from two different modalities (e.g. $\mathbf{I} = \{T1(IR), T1\}, \mathbf{J} = \{T1(IR), T2\}$)
- **C**: analogous to **B**, but the order of the channels is swapped for one image (e.g. $\mathbf{I} = \{T2, T1\}, \mathbf{J} = \{T1(IR), T2\}$)
- **MM** (multimodal registration): only one MR sequence is used for the target image and a different one for the moving image. Three histogram channels are constructed for both single channel MR images.

MC-NCC performs best when the exact same two channels are available, closely followed by LCCA. Both methods improve significantly over LC^2 ($p < 0.01$ for Wilcoxon rank sum test), which can only utilise one multichannel image. When only one channel is of common modality across scans, LCCA achieves significantly better registration accuracy than MC-NCC or LC^2 ($p < 0.01$), which demonstrate that it correctly finds new relationships between different multispectral channels. When using local histograms for multimodal registration (T1→FLAIR) LCCA nearly reaches its performance from experiments **B** and **C**. However, further experiments and comparisons are necessary to confirm its suitable for other multimodal registration tasks. We also tested, if a combination of global and local CCA would be beneficial, but found no improvements.

Canine Muscle Segmentation: To facilitate further comparison, we perform registration experiments on the canine dataset (22 training subjects) from the MICCAI SATA challenge [16] for which SyN [3] was used to provide standardised registrations using both MRI channels (T2 and T2 with fat suppression). Manual segmentations of seven proximal leg muscles have been provided (see Fig. 3) to study muscular dystrophy. The significant differences in size and appearance of the studied dogs render this a very challenging registration task. We include an affine alignment step (using block matching with LCCA and trimmed least squares [17]) and increase the patch-radius r to 3 voxels. Our registration approach significantly outperforms SyN (both use affine+deformable transformations) with an improvement in overlap of 15% ($p < 0.01$). When initialising SyN using the affine transforms obtained for LCCA, the results improve to D=52.4±15%, which is still significantly ($p = 0.008$) inferior to our results of D=55.5±13%. These results for single-atlas segmentation can be further improved using label fusion [16]. Already a simple majority voting results in a Dice overlap of D=73.0±14% for our approach (compared to 55.1% for SyN in [16]). Our algorithm is with a runtime of ≈40 sec. several times faster than SyN.

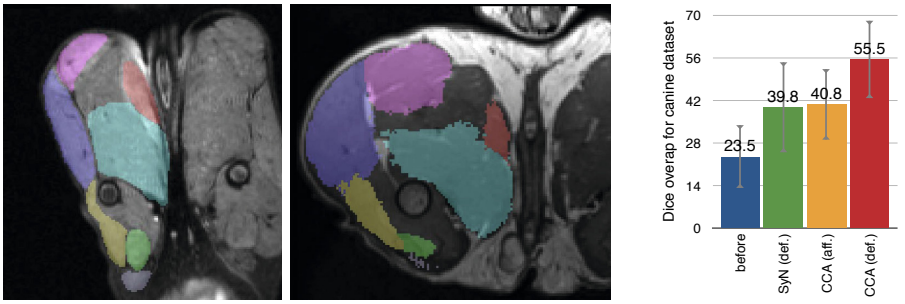

Fig. 3. Results of registration of canine dataset with LCCA. Left: Slice of atlas scan with manual segmentation. Centre: Target scan with automatically transferred labels. Right: Quantitative evaluation of Dice overlap shows improvement using our approach with deformable (def.) registration (D=55.5 %) compared to ANTs SyN (D=39.8 %). When applying majority voting, LCCA achieves a segmentation accuracy of D=73.0%.

4 Conclusion

We have presented a novel similarity metric for registration of multispectral images. Local canonical correlation analysis (LCCA) is based on an established statistical concept for multivariate variables. An efficient computational scheme inspired by guided image filtering is used to calculate dense patch-wise similarities. The main benefit of LCCA is its ability to find new linear relationships **across** channels. Our new metric generalises established techniques, such as NCC [3] and LC^2 [6], and works for images with an arbitrary number of channels. When applied to local histograms it is able to deal with multimodal data. The validation results demonstrate its advantages over current state-of-the-art methods especially for challenging multispectral data.

LCCA can be very useful in clinical practice where not always the same sequences are acquired for every patient and scanner parameters (e.g. repetition time, flip angle) may vary. In future work, we plan to apply this concept within an ongoing clinical study to improve the detection and segmentation of stroke lesions in multispectral MRI. A further application is its use to analyse sequences, where time-points are not necessarily temporally aligned [1]. LCCA can also be used to achieve rotational invariance for multidimensional orientated image descriptors [18], removing the need for recalculation/reorientation of them. It could also be used to find optimal correlations across feature-based image representations. The metric can easily be applied to multi-atlas based label fusion and extend it to multichannel images. Future work may investigate whether selecting only the strongest correlations (akin to dimensionality reduction) and/or the calculation of the eigenvector mapping on a coarser scale are beneficial.

Acknowledgements. B.W.P. and J.A.S. would like to acknowledge funding from the CRUK/ EPSRC Cancer Imaging Centre, Oxford. M.P.H. thanks Adrian Dalca and Ramesh Sridharan for fruitful discussions, which initiated this article.

References

1. Peyrat, J.M., Delingette, H., Sermesant, M., Xu, C., Ayache, N.: Registration of 4D cardiac CT sequences under trajectory constraints with multichannel diffeomorphic demons. IEEE Transactions on Medical Imaging 29(7), 1351–1368 (2010)
2. Rohde, G., Pajevic, S., Pierpaoli, C., Basser, P.: A comprehensive approach for multi-channel image registration. In: Gee, J.C., Maintz, J.B.A., Vannier, M.W. (eds.) WBIR 2003. LNCS, vol. 2717, pp. 214–223. Springer, Heidelberg (2003)
3. Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C.: Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med. Imag. Anal. 12(1), 26–41 (2008)
4. Ruiz-Alzola, J., Westin, C., Warfield, S., Alberola, C., Maier, S., Kikinis, R.: Nonrigid registration of 3D tensor medical data. Med. Imag. Anal. 6(2), 143–161 (2002)
5. Taquet, M., Macq, B., Warfield, S.K.: A generalized correlation coefficient: Application to DTI and multi-fiber DTI. In: MMBIA 2012, pp. 9–14. IEEE (2012)
6. Wein, W., Brunke, S., Khamene, A., Callstrom, M.R., Navab, N.: Automatic CT-ultrasound registration for diagnostic imaging and image-guided intervention. Med. Imag. Anal. 12(5), 577–585 (2008)
7. Iglesias, J.E., Konukoglu, E., Zikic, D., Glocker, B., Van Leemput, K., Fischl, B.: Is synthesizing MRI contrast useful for inter-modality analysis? In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013, Part I. LNCS, vol. 8149, pp. 631–638. Springer, Heidelberg (2013)
8. Guimond, A., Roche, A., Ayache, N., Meunier, J.: Three-dimensional multimodal brain warping using the demons algorithm and adaptive intensity corrections. IEEE Transactions on Medical Imaging 20(1), 58–69 (2001)
9. Li, Y., Verma, R.: Multichannel image registration by feature-based information fusion. IEEE Transactions on Medical Imaging 30(3), 707–720 (2011)
10. Hotelling, H.: Relations between two sets of variates. Biometrika 28(3-4), 321–377 (1936)
11. Friman, O., Cedefamn, J., Lundberg, P., Borga, M., Knutsson, H.: Detection of neural activity in functional MRI using canonical correlation analysis. Magnetic Resonance in Medicine 45(2), 323–330 (2001)
12. Suarez, R.O., Commowick, O., Prabhu, S.P., Warfield, S.K.: Automated delineation of white matter fiber tracts with a multiple region-of-interest approach. NeuroImage 59(4), 3690–3700 (2012)
13. He, K., Sun, J., Tang, X.: Guided image filtering. IEEE Transactions on Pattern Analysis and Machine Intelligence 35(6), 1397–1409 (2013)
14. Heinrich, M.P., Papież, B.W., Schnabel, J.A., Handels, H.: Non-parametric discrete registration with convex optimisation. In: Ourselin, S., Modat, M. (eds.) WBIR 2014. LNCS, vol. 8545, pp. 51–61. Springer, Heidelberg (2014)
15. Mendrik, A.: Evaluation framework for MR brain image segmentation. In: MICCAI Grand Challenge (2013), `http://mrbrains13.isi.uu.nl`
16. Landman, B., Warfield, S.: Segmentation: Algorithms, theory and applications. In: MICCAI SATA (2013), `https://masi.vuse.vanderbilt.edu/workshop2013`
17. Ourselin, S., Roche, A., Prima, S., Ayache, N.: Block matching: A general framework to improve robustness of rigid registration of medical images. In: Delp, S.L., DiGoia, A.M., Jaramaz, B. (eds.) MICCAI 2000. LNCS, vol. 1935, pp. 557–566. Springer, Heidelberg (2000)
18. Heinrich, M.P., Jenkinson, M., Papież, B.W., Brady, S.M., Schnabel, J.A.: Towards realtime multimodal fusion for image-guided interventions using self-similarities. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013, Part I. LNCS, vol. 8149, pp. 187–194. Springer, Heidelberg (2013)