Modality Propagation: Coherent Synthesis of Subject-Specific Scans with Data-Driven Regularization

D. H. Ye[1], D. Zikic[2], B. Glocker[2], A. Criminisi[2], E. Konukoglu[3]

[1] Department of Radiology, University of Pennsylvania, Philadelphia, PA, USA
[2] Microsoft Research, Cambridge, UK
[3] Martinos Center for Biomedical Imaging, MGH, Harvard Medical School, MA, USA

Abstract. We propose a general database-driven framework for coherent synthesis of subject-specific scans of desired modality, which adopts and generalizes the patch-based label propagation (LP) strategy. While modality synthesis has received increased attention lately, current methods are mainly tailored to specific applications. On the other hand, the LP framework has been extremely successful for certain segmentation tasks, however, so far it has not been used for estimation of entities other than categorical segmentation labels. We approach the synthesis task as a *modality propagation*, and demonstrate that with certain modifications the LP framework can be generalized to continuous settings providing coherent synthesis of different modalities, beyond segmentation labels. To achieve high-quality estimates we introduce a new data-driven regularization scheme, in which we integrate intermediate estimates within an iterative search-and-synthesis strategy. To efficiently leverage population data and ensure coherent synthesis, we employ a spatio-population search space restriction. In experiments, we demonstrate the quality of synthesis of different MRI signals (T2 and DTI-FA) from a T1 input, and show a novel application of modality synthesis for abnormality detection in multi-channel MRI of brain tumor patients.

1 Introduction

Medical imaging enjoys a multitude of image modalities, each locally quantifying and mapping different characteristics of the underlying anatomy. For instance, while CT images display local tissue densities, diffusion weighted images quantify the tissue directionality. Even derived quantities such as local fractional anisotropy (FA) can be seen as modalities, as they quantify certain characteristics of the anatomy. There is increased interest in methods which perform subject-specific synthesis of a certain *target modality*, from a given *source modality*. The ability to automatically generate different appearances of the same anatomy, without an actual acquisition, enables various applications such as creating virtual models [1], multi-modal registration [2–4], super-resolution [5], atlas construction [6] and virtual enhancement [7]. Most of the current methods are tailored to specific applications. Some approaches perform synthesis using explicit models, for example for simulating US from CT [2] or modality conversion between T1 and T2 [3]. Explicit modeling is application-dependent and does not generalize easily. Other approaches use existing databases for synthesis. In [8] the FLAIR channel is created from T1 and T2 for brain MRI, the MR inhomogeneity field is estimated in [9], synthesis of an alternative modality to facilitate the registration in correlative microscopy is considered in [4], and in [10, 11] synthesis of high-resolution images is tackled. Despite being database-driven, the above works focus on specific problems, and propose case-specific approaches that differ significantly from one another. Admittedly, the general synthesis problem poses additional difficulties over the specific versions, due to the multitude of possible scenarios. Different modalities characterize different physical properties, and the target modality might contain richer information than the source.

We propose a general framework for modality synthesis which avoids explicit modeling and operates by utilizing a database of images with arbitrary target and source modalities. For each point in the target image, we perform a local patch-based search in the database and nearest neighbor information is used for estimating the target modality value for this point. The intuition behind our method comes from the observation that local and contextual similarities observed in one modality often extend to other modalities. Our method can be seen as a generalization of so-called label propagation (LP) strategies (cf. [12–14]), especially of the recent patch-based approaches [15, 16]. LP has been very successful for certain segmentation tasks, being the de facto standard for brain anatomy segmentation. To our best knowledge, LP has so far been exclusively used in the setting of segmentation, with discrete, categorical labels. We show in this work that with certain generalizing modifications, the LP framework can be adapted to continuous non-categorical tasks with high-quality synthesis of arbitrary modalities. We refer to our approach as *modality propagation* (MP).

Label propagation operates in two steps. First, for each location in the target image a set of label candidates is determined from the database. Second, these candidates are fused into a single label based on one of the numerous fusion strategies. Label fusion has been shown to be crucial for achieving high-quality results for segmentation. However, currently available fusion strategies are specific to discrete categorical labels. For our general setting, dealing with continuous values, we replace the fusion step with a data-driven regularization approach. We show that this improves the quality – similar to fusion strategies – while being applicable to general problems, beyond segmentation. Our data-driven regularization is inspired by Image Analogies [17], using both source *and* target modalities within an iterative search-and-synthesis strategy.

After describing our MP framework in the next section, we demonstrate its properties on synthesis of MR-T2 and DTI-FA maps from MRI-T1 in (Sec 3.1). In Sec 3.2, we present the potential of our framework by proposing a novel approach for abnormality detection in multi-channel brain MRI where we synthesize patient-specific pseudo-healthy T2 images from T1, in order to highlight differences to the actual patients' T2 scans.

2 Coherent Synthesis via Modality Propagation

In the context of modality propagation, the process of synthesis refers to the following task: Given a source image I with modality-specific information of the underlying anatomy, we seek to generate a corresponding target image S of the same anatomy but from another modality. The new image S is constructed using both subject-specific image I and a population database. This database contains N exemplar image pairs $\mathcal{T} = \{(I_n, S_n)\}_{n=1}^{N}$ where images I_n and S_n are assumed to be spatially aligned. The idea behind MP is that local similarities between structures both visible in I and in the database images $\{I_n\}$ should indicate similarities between $\{S_n\}$ and S, the image to-be-synthesized. By finding correspondences between input I and database $\{I_n\}$, information can be transferred from the set $\{S_n\}$ in order to synthesize S. This approach can be applied to arbitrary pairs of source and target modality.

For each image point $\mathbf{x}$ the estimate $S(\mathbf{x})$ is determined through patch-based nearest neighbor search within the population database $\mathcal{T}$. Previous works perform this search based on information extracted from images I and $\{I_n\}$ only. In this context, we develop an iterative search-and-synthesis strategy which is inspired by Image Analogies [17]. The key idea is to incorporate the partly synthesized image into the nearest neighbor search.

Our algorithm is based on two main components which aim for improved coherency: 1) patch-based search in a restricted space, 2) iterative synthesis using intermediate results yielding a data-driven regularization.

Search Space Restriction: For each image point $\mathbf{x}$ we perform a patch-based nearest neighbor search within the database images $\mathcal{T}$. The restriction of the search space is achieved on two different layers, a *spatial* and *population* layer. We enforce spatial restriction by assuming affine alignment of all subjects both within the population database and between the database and the input image I. Depending on the expected accuracy of the registration, this allows us to restrict the search to a small search window $W_{\mathbf{x}}$ centered at the point $\mathbf{x}$. A search window placed in a particular database image I_n is denoted by $W_{\mathbf{x}}^{I_n}$.

Restriction at the population layer is achieved by obtaining a subset of database images that are most "similar" to I. The input image is divided into equally sized cells, and for each cell, the k nearest neighbors in terms of sub-image dissimilarity are determined within the population database. The size of the sub-images is equal to the size of the cells, and the k neighbors can be different for different cells. The patch-based search for all image points within a particular cell is restricted to the same set of k database images. For notational convenience, we define a set of the kNN image indices as $\mathcal{N}_{\mathbf{x}}^{k}$ linked to an image point $\mathbf{x}$, instead of a cell.

The search space restriction is important for two reasons. First, a smaller search space yields lower computation times, in particular when dealing with larger databases. Second, and more importantly, the spatial restriction increases correlation between high patch similarity and correct anatomical correspondence.

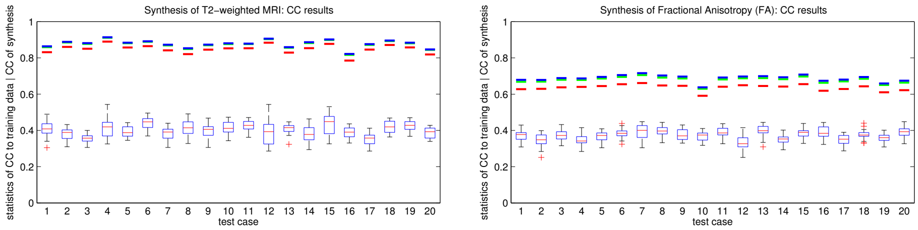

Fig. 1. Correlation coefficients between synthesis and ground truth after 1st, 2nd, and 3rd iteration in red, green, and blue for T2 and FA. The box plots summarize the CC distribution between ground truth and all database images.

Iterative Synthesis: The actual task of synthesizing the output image S is performed in an iterative manner. The first iteration is similar to well-known label propagation methods. A patch $P^I_{\mathbf{x}}$ centered at the image point $\mathbf{x}$ is extracted from the input image I. Based on the restriction strategy mentioned above, we perform a nearest neighbor search by evaluating patch dissimilarities between $P^I_{\mathbf{x}}$ and all patches in the spatio-population search space as

$$(\hat{n}, \hat{\mathbf{y}}) = \underset{n \in \mathcal{N}^k_{\mathbf{x}};\, \mathbf{y} \in W^{I_n}_{\mathbf{x}}}{\arg\min}\; d(P^I_{\mathbf{x}}, P^{I_n}_{\mathbf{y}}) \quad . \tag{1}$$

For d we consider sum of squared differences (SSD), though other definitions, e.g. based on correlation, are possible. Based on Equation (1), we determine the best matching patch to be the one centered at $\hat{\mathbf{y}}$ in the database image $I_{\hat{n}}$. The synthesis image is then updated by setting $S(\mathbf{x}) := S_{\hat{n}}(\hat{\mathbf{y}})$. Once this is performed for all image points, we obtain a fully synthesized image S.

A limitation of this process is that the rich information in $\{S_n\}$ remains unused during search. This might yield inaccurate anatomical coherency and noisier outputs. Coherency of S can be greatly improved when intermediate results are considered in a subsequent refinement. After obtaining an initial estimate of S, for subsequent iterations the patch-based nearest neighbor search is performed using a modified version of Equation (1) where we incorporate information from the synthesized modality:

$$(\hat{n}, \hat{\mathbf{y}}) = \underset{n \in \mathcal{N}^k_{\mathbf{x}};\, \mathbf{y} \in W^{I_n}_{\mathbf{x}}}{\arg\min}\; (1-\alpha)\, d(P^I_{\mathbf{x}}, P^{I_n}_{\mathbf{y}}) + \alpha\, f(P^S_{\mathbf{x}}, P^{S_n}_{\mathbf{y}}) \quad . \tag{2}$$

Note, this definition covers the one for the first iteration when $\alpha = 0$. The metric f depends on the nature of the target modality. In case of scalar-valued modalities, it can be the same as d. Integrating the synthesized data into the search yields a significant improvement on the spatial coherency. One can think of the function $f(P^S_{\mathbf{x}}, P^{S_n}_{\mathbf{y}})$ as a data driven regularization term. The effect of this term will be demonstrated in our experiments.

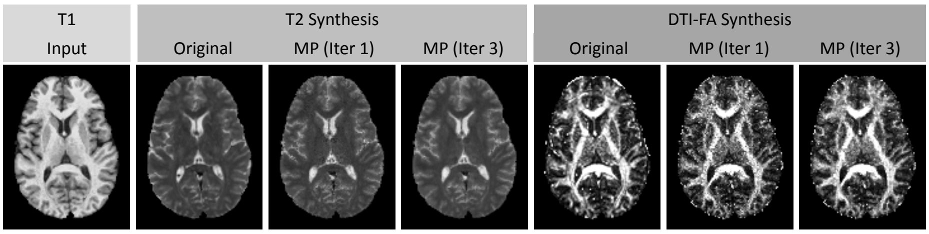

Fig. 2. Visual results for synthesis of T2 and FA maps from T1 MRI input. Synthesis is shown for the 1st and 3rd iteration. Note the reduction of noise and improved consistency of structural appearance both in T2 and FA. In particular, the white matter tracts in FA and the ventricle structures in T2 are well synthesized.

3 Experiments

Modality Propagation (MP) is a general tool that can be applied to possibly any source-target modality pair. Here we present two different applications of the algorithm: synthesis of T2 and FA MRI signals from T1, and abnormality detection via modality synthesis. We also analyze the impact of individual components, such as the iterative search-and-synthesis strategy.

We use the same parameter settings throughout all experiments. The size of patches P is $3 \times 3 \times 3$ voxels, and the local search window W is set to $9 \times 9 \times 9$. The database sub-image indices $\mathcal{N}^k$ are determined for cells with the same size as the search windows. We determine $k = 5$ nearest neighbors from the database for each cell. Synthesis starts with a weighting factor $\alpha = 0$, as no synthesized information is available initially. In subsequent iterations, the weighting is increased to reach $\alpha = 1$ yielding increased importance of the intermediate synthesis. This has the effect of data driven regularization. We found that 3 iterations are in general sufficient for MP to converge.

3.1 Synthesis of T2 and DTI-FA from T1 MRI

The first application is synthesis of T2 and FA from corresponding T1-weighted MR brain images. The goal of this experiment is to demonstrate synthesis of different structures mapping different physical properties, from the same input source. A possible use case for this is modality translation, e.g. for multi-modal registration [3]. We use the NAMIC database which has T1, T2 and Diffusion Tensor images (DTI) for 20 individuals, 10 normal controls and 10 schizophrenia patients (http://hdl.handle.net/1926/1687). We compute FA maps with the 3D Slicer software (http://www.slicer.org). Images are linearly registered, skull-stripped, inhomogeneity corrected, histogram-matched within each modality, and resampled to 2 mm resolution.

We apply a leave-one-out cross-validation, so each synthesis result is based on 19 subjects. This allows us to compute the similarity between the synthesized and the real images. Graphs in Figure 1 show the quantification of this image

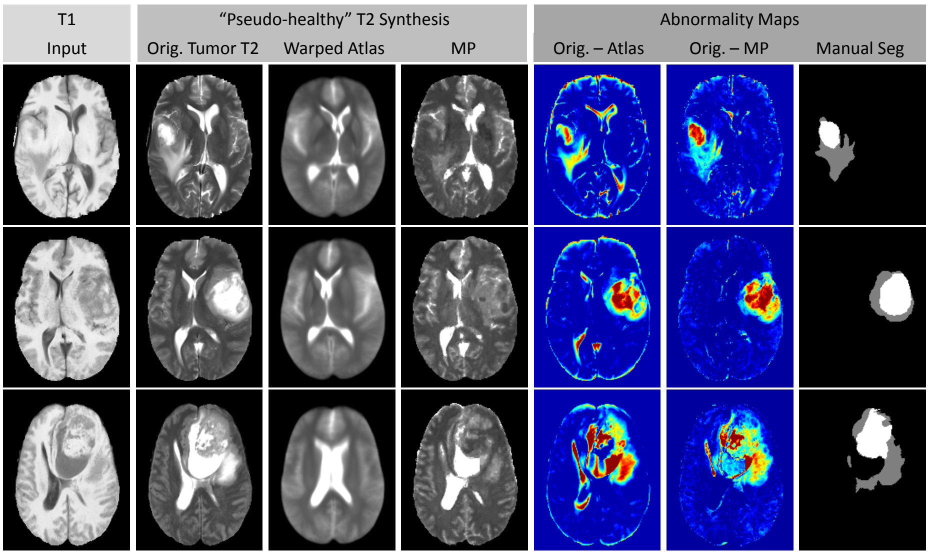

Fig. 3. Visual results for the abnormality detection using synthesis of "pseudo-healthy" images. We compare our results with a deformable atlas approach. The abnormality maps are computed by subtracting the warped atlas and the synthesized T2 from the original T2. For comparison, we also show manual tumor segmentations.

similarity. We use correlation coefficient (CC) as its normalized values allow us to compare the different experiments. The horizontal lines shown in red, green and blue indicate the CC after the 1st, 2nd and 3rd iteration of MP, respectively. To provide a comparative context, we also compute the CC between the ground truth and all database images. These values are summarized as box plots in the same figure. We make several observations: 1) The synthesized images are significantly closer to the real one than any other database image. 2) Synthesis for T2 gets closer to ground truth than for FA. We believe this is due to the fact that physical properties quantified by FA are very different from the ones quantified by T1 and T2. T1 and T2 capture more similar structures. Furthermore, FA contains more local information in terms of anisotropy and directionality of the tissue, and synthesizing FA from T1 is less accurate. 3) The improvement obtained by the 2nd and 3rd iteration is higher for FA than for T2. As the type of local information in FA is different, the anatomical coherency is not entirely captured using T1 as input. However, using the synthesized channel in the nearest neighbor search greatly improves the structural coherency.

Visual examples for T2 and FA synthesis after the 1st and 3rd iteration are shown in Figure 2. These images demonstrate the impact of the iterative algorithm. We observe that the prominent structures, such as the ventricles or the large white matter bundles, are accurately constructed. Smaller structures, such as small fiber tracts in the FA maps or thin sulci in T2, are less accurate. A larger training database could yield higher accuracies for smaller structures.

3.2 Abnormality Detection in Multi-Channel MRI

In our second application we build a system for abnormality detection which is based on the concept of comparing a suspicious image to an image of a healthy subject. Automatic techniques often make use of an atlas constructed from a healthy population. The atlas is spatially aligned with the test image, and subtraction of the two reveals differences (or abnormalities) in the test image. The main difficulty is the nonlinear registration in the presence of pathologies.

We take a synthesis approach making use of the properties of different MR signals commonly acquired for tumor patients. While for this particular tumor, pathological tissue enhances in T2 yielding a hyper-intense appearance, it does not substantially alter the intensity profile in (non-contrasted) T1. The patient's T1 image serves as input for synthesis of a "pseudo-healthy" T2 image by employing a database of 100 healthy subjects for which both T1 and T2 images are available (IXI database http://biomedic.doc.ic.ac.uk/brain-development/). Abnormality maps are computed by subtracting the synthesized T2 from the original T2 image. We use images of 20 tumor patients from the BRATS dataset (http://www2.imm.dtu.dk/projects/BRATS2012/)[1].

For comparison, we use an atlas constructed from the healthy-subjects and non-linearly registered onto the patients' T1 images. Figure 3 presents the visual results. The first four columns display patients' T1 and T2 images, the aligned atlas and the pseudo-healthy image synthesized using MP. The fifth and sixth column display the abnormality maps obtained by computing the differences between the patient's T2 image and the registered atlas, and the synthesized image. The last column displays manual tumor delineations. The abnormality maps obtained using MP are much cleaner, especially in areas not related to tumor, and abnormal areas are more prominent. Our approach provides a simple alternative which avoids the challenging nonlinear registration problem in presence of pathologies. These maps could be used for further segmentation steps.

4 Summary

We propose Modality Propagation, a general algorithm for patient-specific synthesis of arbitrary modalities, which employs population data, and generalizes the patch-based label propagation scheme. As main contributions of our work, we see: 1) Generalization of the LP strategy to arbitrary modalities, and showing that this general framework can obtain high-quality coherent results for applications beyond segmentation; 2) We propose an efficient data-driven regularization scheme for improvement of the result quality, as a general alternative to LP fusion strategies; 3) We propose a novel scheme for abnormality detection in multi-channel brain MRI, which utilizes the proposed MP framework. With increasing availability of population databases we believe our work can be an important component for developing subject-specific analysis tools.

[1] BRATS is organized by B. Menze, A. Jakab, S. Bauer, M. Reyes, M. Prastawa, and K. Van Leemput. The database contains fully anonymized images from following institutions: ETH Zurich, Univ. of Bern, Univ. of Debrecen, and Univ. of Utah.

References

1. Prakosa, A., Sermesant, M., Delingette, H., Marchesseau, S., Saloux, E., Allain, P., Villain, N., Ayache, N.: Generation of Synthetic but Visually Realistic Time Series of Cardiac Images Combining a Biophysical Model and Clinical Images. IEEE TMI **32**(1) (2013) 99–109
2. Wein, W., Brunke, S., Khamene, A., Callstrom, M., Navab, N.: Automatic CT-ultrasound registration for diagnostic imaging and image-guided intervention. MedIA **12**(5) (2008) 577–585
3. Kroon, D., Slump, C.: MRI modalitiy transformation in demon registration. In: IEEE ISBI. (2009) 963–966
4. Cao, T., Zach, C., Modla, S., Powell, D., Czymmek, K., Niethammer, M.: Registration for correlative microscopy using image analogies. In: WBIR. LNCS, Springer, Heidelberg (2012) 296–306
5. Zhang, Y., Wu, G., Yap, P.T., Feng, Q., Lian, J., Chen, W., Shen, D.: Hierarchical Patch-Based Sparse Representation - A New Approach for Resolution Enhancement of 4D-CT Lung Data. IEEE TMI **31**(11) (2012) 1993–2005
6. Commowick, O., Warfield, S., Malandain, G.: Using Frankensteins creature paradigm to build a patient specific atlas. In: MICCAI. LNCS, Springer, Heidelberg (2009) 993–1000
7. Nuyts, J., Bal, G., Kehren, F., Fenchel, M., Michel, C., Watson, C.: Completion of a Truncated Attenuation Image from the Attenuated PET Emission Data. IEEE TMI **32**(2) (2013) 237–246
8. Roy, S., Carass, A., Shiee, N., Pham, D., Prince, J.: MR contrast synthesis for lesion segmentation. In: IEEE ISBI. (2010) 932–935
9. Roy, S., Carass, A., Bazin, P., Prince, J.: Intensity inhomogeneity correction of magnetic resonance images using patches. In: SPIE Med. Imaging. (2011)
10. Rousseau, F.: Brain Hallucination. In: ECCV. LNCS, Springer, Heidelberg (2008) 497–508
11. Rueda, A., Malpica, N., Romero, E.: Single-image super-resolution of brain MR images using overcomplete dictionaries. MedIA **17**(1) (2013) 113–132
12. Rohlfing, T., Russakoff, D., Maurer, C.: Expectation maximization strategies for multi-atlas multi-label segmentation. In: IPMI. LNCS, Springer, Heidelberg (2003) 210–221
13. Wolz, R., Chu, C., Misawa, K., Mori, K., Rueckert, D.: Multi-organ Abdominal CT Segmentation Using Hierarchically Weighted Subject-Specific Atlases. In: MICCAI. LNCS, Springer, Heidelberg (2012) 10–17
14. Artaechevarria, X., Munoz-Barrutia, A., Ortiz-de Solorzano, C.: Combination strategies in multi-atlas image segmentation: Application to brain MR data. IEEE TMI **28**(8) (2009) 1266–1277
15. Coupé, P., Manjón, J., Fonov, V., Pruessner, J., Robles, M., Collins, D.: Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation. NeuroImage **54**(2) (2011) 940–954
16. Rousseau, F., Habas, P., Studholme, C.: A supervised patch-based approach for human brain labeling. IEEE TMI **30**(10) (2011) 1852–1862
17. Hertzmann, A., Jacobs, C., Oliver, N., Curless, B., Salesin, D.: Image Analogies. In: ACM Conference on Computer Graphics and Interactive Techniques. (2001)