2013 IEEE 10th International Symposium on Biomedical Imaging:
From Nano to Macro
San Francisco, CA, USA, April 7-11, 2013

MAGNETIC RESONANCE IMAGE SYNTHESIS THROUGH PATCH REGRESSION

Amod Jog[1], *Snehashis Roy*[2], *Aaron Carass*[2], *Jerry L. Prince*[2]

[1]Dept. of Computer Science,
[2]Dept. of Electrical and Computer Engineering,
The Johns Hopkins University
{amodjog,snehashisr,aaron_carass,prince}@jhu.edu

ABSTRACT

Magnetic resonance imaging (MRI) is widely used for analyzing human brain structure and function. MRI is extremely versatile and can produce different tissue contrasts as required by the study design. For reasons such as patient comfort, cost, and improving technology, certain tissue contrasts for a cohort analysis may not have been acquired during the imaging session. This missing pulse sequence hampers consistent neuroanatomy research. One possible solution is to synthesize the missing sequence. This paper proposes a data-driven approach to image synthesis, which provides equal, if not superior synthesis compared to the state-of-the-art, in addition to being an order of magnitude faster. The synthesis transformation is done on image patches by a trained bagged ensemble of regression trees. Validation was done by synthesizing T_2-weighted contrasts from T_1-weighted scans, for phantoms and real data. We also synthesized 3 Tesla T_1-weighted magnetization prepared rapid gradient echo (MPRAGE) images from 1.5 Tesla MPRAGEs to demonstrate the generality of this approach.

Index Terms— Image synthesis, regression, brain

1. INTRODUCTION

MRI is the modality of choice for the study of brain structure and function. One of the many advantages of using MRI is the ability to image brain tissues in a variety of tissue contrasts via programmed pulse sequences. Certain pathologies are best understood when studied with a particular pulse sequence. For example, growth of multiple sclerosis (MS) lesions can be studied by measuring the lesions using T_2-weighted (T_2-w) images [1]. Infant brain surface reconstruction also relies on T_2-w images [2] as the T_1-w images lack sufficient tissue contrast. The left column of Fig. 1 shows a T_1-w MPRAGE and the corresponding T_2-w image (center column), which shows starkly different tissue contrasts.

In this paper, we address the problem of the missing pulse sequence (MPS). Due to the constant improvement of MRI scanners and the shifting focus from one particular tissue to another in a large research project, sometimes a desired tissue contrast (such as T_2-w), or a desired pulse sequence, remains unobserved thus necessitating the synthesis of the MPS. An intuitive and simple way to synthesize an image with a desired tissue contrast could be to perform histogram matching (HM). Given a subject image of a certain pulse sequence and an atlas image of the desired pulse sequence, we can map the intensity values by matching the two histograms. However, if the task involves synthesizing a T_2-w contrast from a given T_1-w contrast, HM matches the histograms, but does not yield actual intensities that are consistent with the desired tissue contrast. See Fig. 1 for an example of an undesirable result produced with histogram matching. We want to synthesize a T_2-w contrast for a given subject MPRAGE. Clearly, HM does not work for T_2-w image synthesis even though the histograms are closer.

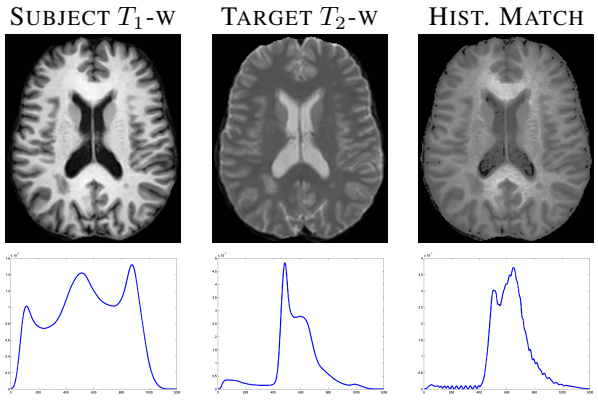

Fig. 1. Histogram matching result of a T_1-w MPRAGE to a T_2-w target is shown.

One of the classical approaches to synthesize a $\mathcal{C}_2$ contrast image from its $\mathcal{C}_1$ contrast acquisition is described in [3], wherein an atlas contains a pair of images $(\mathcal{A}_1, \mathcal{A}_2)$ having tissue contrasts $\mathcal{C}_1$ and $\mathcal{C}_2$, respectively, co-registered and sampled on the same voxel locations in space. Given a subject image $\mathcal{S}_1$ with contrast C_1, the atlas image $\mathcal{A}_1$ is registered deformably to $\mathcal{S}_1$ and the same transformation is applied to $\mathcal{A}_2$ to synthesize $\mathcal{S}_2$, which will thus have a $\mathcal{C}_2$ contrast. The synthesis result is heavily dependent on the

This work is funded by the NIH/NIBIB under Grant 1R21EB012765. Some data used in this paper are downloaded from Biomedical Informatics Research Network (BIRN) repository project accession number 2007-BDR-6UHZ1.

accuracy of the registration method used. However our approach is preferred in the case of abnormal anatomy, where registration is generally not successful. We show that our approach compares favorably with the registration approach, both in terms of quality of synthesis and speed. Our work is similar in spirit to the algorithm described in [4] where sparse priors were used to synthesize the different MR contrast images. We take a data-driven approach and model the synthesis transformation as a nonlinear regression of image patches. This regression is learned from data by a bagged ensemble of regression trees [5] and is significantly faster and better than the state-of-the-art. The paper is laid out as a description of the method and motivation, followed by a results section with experiments on real and phantom data, and we finish with a discussion on future directions for this work.

2. METHOD

An intuitive approach to image synthesis is through variations of histogram matching. Although this process is widely used and well-studied, it has certain deficiencies. Quantization artifacts are hard to avoid when matching histograms. Histogram matching assumes a one-to-one or a many-to-one relationship between the subject intensities and the target intensities. These relationships are often violated due to the spatial variation of the noise as well as the underlying MR tissue properties. A more general, one-to-many relationship, between the intensities is more appropriate considering the spatial information of intensities; which histogram matching does not offer. To improve upon the idea of histogram matching, we build a nonlinear regression learned from atlas image patches that generates transformations that are not one-to-one in the intensity space and also incorporates spatial smoothing by the use of image patches. Synthesis is done by applying the learned regression to the patches of a given subject image.

2.1. Training data

As stated earlier, an atlas is a pair of images $(\mathcal{A}_1, \mathcal{A}_2)$ having tissue contrasts $\mathcal{C}_1$ and $\mathcal{C}_2$, respectively, which are co-registered and sampled on the same voxel locations in space. As an example, we will consider $\mathcal{C}_1$ as the T_1-w MPRAGE and $\mathcal{C}_2$ as the T_2-w contrast (see Fig. 1). The contrast $\mathcal{C}_2$ is to be synthesized, thus addressing the *missing T_2-w sequence* problem. Our goal is to learn multiple local relationships between $\mathcal{A}_1$ and $\mathcal{A}_2$ and use these to synthesize the $\mathcal{C}_2$ contrast of a subject image $\mathcal{S}_1$, having $\mathcal{C}_1$ contrast. We do this by considering the patches of $\mathcal{A}_1$ and the intensities of $\mathcal{A}_2$ together. We extract the $p \times q \times r$ sized patches of $\mathcal{A}_1$ centered about the i^{th} voxel, having intensity x_i. The patches are stacked into $d \times 1$ sized vectors and are denoted by $\mathbf{x}_i \in \mathbb{R}^d$, $d = pqr$. In our experiments we have chosen $p = q = r = 3$ empirically, thus using $d = 27$ dimensional patch vectors. The corresponding i^{th} voxel in $\mathcal{A}_2$ is denoted by y_i. Here we note that HM constructs a one-to-one or many-to-one mapping between all x_i's and y_i's with the aim of matching their distributions, the x_i's being the independent variables, or attributes, and the y_i's are the dependent variables. We propose a nonlinear regression keeping the y_i's as the dependent variables, but with the attributes being $\mathbf{x}_i$'s, the image patch at the i^{th} voxel. Thus the training data consists of pairs of $\langle \mathbf{x}_i, y_i \rangle$.

2.2. Regression Trees

A regression tree ensemble is learned from this data by the algorithm described in [5]. A single regression tree learns a nonlinear regression by partitioning the d-dimensional space of $\mathbf{x}_i$ into regions based on a split at each node. At each split during training, one third of the attributes are randomly considered and the one that minimizes the least squares criterion is chosen as the attribute to split upon. We use a leaf size of five to prevent the tree from being too deep, thus avoiding overfitting of the training data. The value of the dependent variable assigned at a leaf is simply the average of the y_i's of the training data which end up accumulating at that leaf during training. The learned nonlinear regression is thus piecewise constant. A single regression tree is regarded as a weak learner and in general has higher error [6]. We use a bagged ensemble of regression trees, which reduces errors by bootstrap aggregation [6]. The ensemble consists of n trees, with $n = 100$ in our experiments, each learned from a bootstrapped dataset. To generate a bootstrapped dataset we randomly select a training sample, with replacement N times, where N is the size of the training dataset, which is $\sim 10^6$ in our case. Given a subject image, $\mathcal{S}_1$, with $\mathcal{C}_1$ contrast we generate patches for all voxel positions i and feed those to the trained ensemble. The outputs of each of the trees in the ensemble are aggregated by averaging to produce a final output intensity for the i^{th} voxel of the synthesized image $\mathcal{S}_2$. Training is computationally intensive but can be done beforehand. Image synthesis itself, i.e., applying the ensemble, is very fast. Training an ensemble takes about three hours with about 10 gigabytes of memory for atlas images of the size $256 \times 256 \times 173$. A trained ensemble can synthesize a new $256 \times 256 \times 173$ sized image in about 5–10 minutes, as compared to the sparse reconstruction method [4] ($\sim$ 2–3 hours) and the deformable registration-based method [3] ($\sim$ 1 hour) with the same computing resources.

3. RESULTS

3.1. Validation on Brainweb Phantom

In this experiment, the atlas $(\mathcal{A}_1, \mathcal{A}_2)$ consisted of a Brainweb T_1-w image (TR = 18ms, $\alpha = 30^\circ$, TE = 10ms) and a corresponding T_2-w image (TR = 3000 ms, TE_1 = 17 ms, $\text{TE}_2 = 80$ ms) [7]. The subject $\mathcal{S}_1$ was another T_1-w phantom with a different noise level or a different flip angle. The regression ensemble was trained on the atlas and then applied to the subject to produce a synthetic T_2-w image, $\mathcal{S}_2$. As the subject and the atlas have the same inherent biology, a perfect synthesis would result in an image that is exactly equal to the atlas $\mathcal{A}_2$. We used the mean squared error (MSE) between the synthesized image $\mathcal{S}_2$ and the atlas $\mathcal{A}_2$ as the compari-

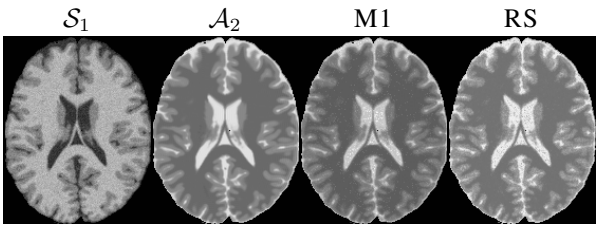

Fig. 2. For Brainweb subject $\mathcal{S}_1$, we synthesize using regression ensemble (RS) and sparse reconstruction (M1) based on the atlas pair $(\mathcal{A}_1, \mathcal{A}_2)$. $\mathcal{A}_2$ is the ground truth image against which RS and M1 results are compared.

son metric and compared the performance of our regression based synthesis (**RS**) with the sparse reconstruction method, denoted by **M1** [4]. We carried out two experiments: 1) we changed the noise level from $0-5\%$ keeping the flip angle constant at $30°$; 2) we varied the flip angle from $15-60°$ keeping the noise constant at 0%. In both cases the parameter is varied on the subject image $\mathcal{S}_1$ and we synthesized the corresponding $\mathcal{S}_2$ image, with results shown in Table 1. **RS** results in significantly lower MSE (p-value $= 0.002$) in both cases, indicating that **RS** is more robust than **M1** w.r.t. noise as well as imaging parameter change. Fig. 2 shows the atlas $\mathcal{A}_2$ in conjunction with our synthesis result and that produced by **M1**, when $\mathcal{S}_1$ is the 3% noise version of $\mathcal{A}_1$.

3.2. Synthesis of T_2-w contrasts

In this experiment, we synthesized T_2-w contrasts from T_1-w MPRAGE scans of real subjects. We compared the performance of **RS** with **M1**, as well as the deformable registration-based method [3], denoted by **D1**. The deformable registration algorithm used in **D1** is from [8]. We did not use **D1** in the phantom experiments as the atlas and the subject belonged to the same brain phantom sampled at the same locations in space, i.e., they were already registered. As the registration transformation is identity, **D1** would have resulted in a trivial, perfect synthesis. We used universal quality index (UQI) [9] as a metric in addition to MSE. UQI is an image quality metric that models the similarity that a human visual system would perceive between two images. If the images are identical, their UQI is 1, otherwise it lies between 0 and 1. We experimented on four subjects from the KIRBY dataset [10] where each subject has two MPRAGE acquisitions and two corresponding double spin echo T_2-w acquisitions. The resolution

Table 1. MSE ($\times 10^3$) values for synthesis of Brainweb phantoms with varying noise levels (%) and flip angles (°) are shown for **M1** and our regression ensemble (**RS**).

Noise (%)		0	1	3	5
MSE	**M1**	1.62	1.46	2.88	4.29
	RS	0.04	0.24	1.38	3.31
Flip Angle (°)		15	30	45	60
MSE	**M1**	2.24	1.62	1.57	1.96
	RS	2.57	0.03	1.19	2.04

Table 2. Average MSE ($\times 10^4$) and UQI values for synthesis of T_2-w images of two repeat scans of four subjects.

Subject #		1	2	3	4	**Mean**
MSE	**M1**	0.85	2.86	1.14	1.41	1.56
	D1	3.57	4.37	4.41	3.90	4.06
	RS	1.33	1.22	1.25	1.24	**1.26**
UQI	**M1**	0.77	0.62	0.69	0.67	0.69
	D1	0.78	0.76	0.77	0.75	0.76
	RS	0.86	0.83	0.84	0.84	**0.84**

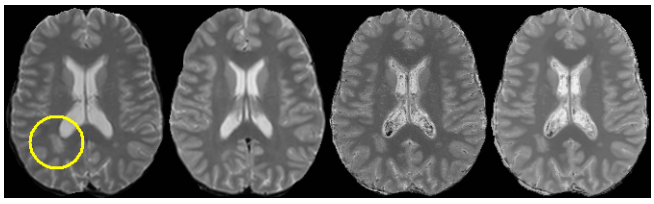

Fig. 3. Synthesis results of different algorithms vs true image on real data. Lesion circled in the true image is synthesized correctly by **M1** and **RS**, but not by **D1**.

of the MPRAGE acquisition is $1.2\times1.2\times1.2\text{mm}^3$, while that of the T_2-w acquisitions is $0.82\times0.82\times1.5\text{mm}^3$. All four images of a particular subject are co-registered. Although these images were acquired on the same scanner within a short duration, they differ slightly due to scanner parameter inconsistency and noise. As an atlas pair $(\mathcal{A}_1, \mathcal{A}_2)$, we selected one MPRAGE and T_2-w image from another subject in the dataset, and trained **RS** to transform MPRAGE intensities to the corresponding T_2-w intensities. We then used the trained **RS** to synthesize eight T_2-w images from the eight MPRAGE images. The quality of the synthesis of T_2-w images was measured by comparing with the corresponding true T_2-w images present in the data. MSE and UQI values after synthesis are reported in Table 2. Fig. 3 shows the results for all the algorithms in conjunction with the ground truth image. MSE-wise, **M1** and **RS** have a comparable performance (p-value $= 0.56$). UQI-wise, it is clear that **RS** outperforms both **M1** and **D1** where the images produced by **RS** are the closest to the ground truth, also seen visually. If the anatomy of the subject and the atlas differs, registration cannot truly reproduce it. This is evident in Fig. 3 (yellow circle) where a lesion is not synthesized by **D1** as it is not present in the atlas.

3.3. Synthesis of T_1-w images

In this experiment, we synthesized a 3T MPRAGE image from a 1.5T MPRAGE image. Generally, 3T MPRAGE images have a higher signal-to-noise ratio and better tissue contrast than 1.5T MPRAGE. We chose a normal subject from the BIRN dataset [11] as the atlas, with the $\mathcal{A}_1$ and $\mathcal{A}_2$ atlas images acquired on a 1.5T Siemens and a 3T GE scanner, respectively. We trained **RS** as before and applied it to a 1.5T MPRAGE scan acquired in the OASIS [12] dataset. If

Table 3. Dice (overlap with subject MPRAGE) and relative volumes after synthesis of one subject.

Tissues		CSF	GM	WM
Dice	**HM**	0.79	0.80	0.95
	RS	0.87	0.93	0.97
Rel. Vols	**HM**	0.22	0.42	0.36
	RS	0.35	0.33	0.32
	True	0.34	0.34	0.32

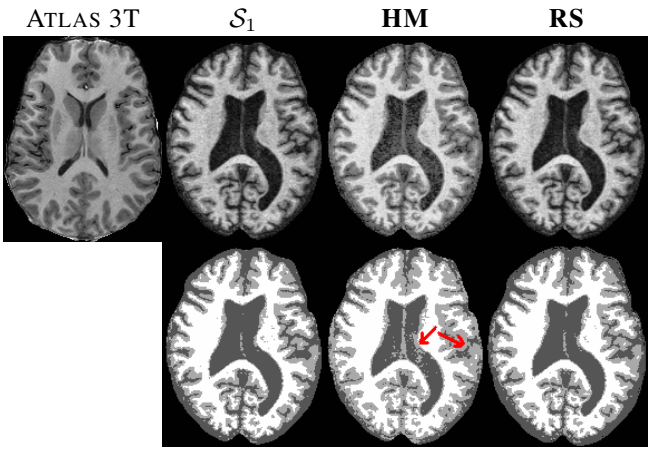

Fig. 4. Original and synthesized MPRAGE images along with their segmentations. The arrows show where CSF voxels are classified as GM.

the pulse sequences are similar, histogram matching has been shown to work reasonably well for synthesis. However when the anatomy and the intensity distributions of the subject and the atlas are quite different, histogram matching results in certain subject tissue voxels being forced to change their tissue class as a result of the matching of the intensity distributions. This phenomenon is evident in Fig. 4, where the subject has abnormally large ventricles (unlike that of the atlas) and thus has more cerebrospinal fluid (CSF) compared to the atlas. Hence, histogram matching results in CSF voxels receiving higher intensities and thereby being classified as gray matter (GM), as shown in Fig. 4, bottom row. In Table 3, we present results of performing segmentation [13] on the subject 1.5T MPRAGE ($\mathcal{S}_1$), and the two synthetic MPRAGEs and calculating Dice overlap coefficient for the three tissue classes. We also calculate the relative volumes (volume of tissue / total intra-cranial volume) of the tissues for all the three images. **RS** synthesis has the most overlap with the original subject image in terms of Dice and the tissue relative volumes are similar as well. It is noted from the relative volume numbers that a large amount of CSF voxels were classified as GM after histogram matching.

4. CONCLUSION AND FUTURE WORK

In this paper, we have proposed a simple, fast, and effective method to synthesize alternate tissue contrasts and normalize intensities for MR images through nonlinear regression on patches. A nonlinear regression is learned from data by a bagged ensemble of regression trees. The method is used to generate T_2-w contrasts from T_1-w images, where traditional histogram based methods fail. The quality of synthesis equals and in most cases, betters the current state-of-the-art and requires a fraction of the processing time. The method is also generic and can be used to synthesize other pulse sequences, for primarily image processing tasks. In future work, we intend to make this method faster and lighter by reducing the training time and memory required by appropriate atlas selection and tree pruning. We would also like to validate on larger datasets and include more pulse sequences for synthesis.

5. REFERENCES

[1] E. Geremia, O. Clatz, B. H. Menze, E. Konukoglu, A. Criminisi, and N. Ayache, "Spatial decision forests for MS lesion segmentation in multi-channel magnetic resonance images," *NeuroImage*, vol. 57, no. 2, pp. 378–390, 2011.

[2] F. Leroy, J.-F. Mangin, F. Rousseau, H. Glasel, L. Hertz-Pannier, J. Dubois, and G. Dehaene-Lambertz, "Atlas-Free Surface Reconstruction of the Cortical Grey-White Interface in Infants," *PLoS ONE*, vol. 6, no. 11, pp. e27128, 2011.

[3] Michael I. Miller, Gary E. Christensen, Yali Amit, and Ulf Grenander, "Mathematical textbook of deformable neuroanatomies," 1993.

[4] S. Roy, A. Carass, and J. L. Prince, "A Compressed Sensing Approach For MR Tissue Contrast Synthesis," in *22nd Conf. on Inf. Proc. in Medical Imaging (IPMI 2011)*, 2011, pp. 371–383.

[5] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*, Statistics/Probability Series. Wadsworth Publishing Company, U.S.A., 1984.

[6] L. Breiman, "Bagging Predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.

[7] R. K. S. Kwan, A. C. Evans, and G. B. Pike, "MRI simulation-based evaluation of image-processing and classification methods," *IEEE Trans. Med. Imag.*, vol. 18, no. 11, pp. 1085–1097, 1999.

[8] B.B. Avants, C.L. Epstein, M. Grossman, and J.C. Gee, "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain," *Medical Image Analysis*, vol. 12, no. 1, pp. 26 – 41, 2008.

[9] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Proc. Letters*, vol. 9, no. 3, pp. 81–84, 2002.

[10] B. A. Landman, A. J. Huang, A. Gifford, D. S. Vikram, I. A. L. Lim, J. A. D. Farrell, J. A. Bogovic, J. Hua, M. Chen, S. Jarso, S. A. Smith, S. Joel, S. Mori, J. J. Pekar, P. B. Barker, J. L. Prince, and P. van Zijl, "Multi-parametric neuroimaging reproducibility: A 3-T resource study," *NeuroImage*, vol. 54, no. 4, pp. 2854–2866, 2011.

[11] L. Friedman *et. al.*, "Test-Retest and Between-Site Reliability in a Multicenter fMRI Study," *Human Brain Mapping*, vol. 29, no. 8, pp. 958–972, August 2008.

[12] D. Marcus, T. Wang, J Parker, J. Csernansky, J. Morris, and R. Buckner, "Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults," *J. Cognitive Neuroscience*, vol. 19, no. 9, pp. 1498–1507, Sept. 2007.

[13] D. L. Pham, "Robust Fuzzy Segmentation of Magnetic Resonance Images," in *14th IEEE Symp. on Comp. Based Med. Systems (CBMS 2001)*, 2001, pp. 127–131.