







MRI Denoising and Artefact Removal Using Self-Organizing Maps for Fast Global Block-Matching

Lee B. Reid¹(✉) , Ashley Gillman^{1,2}, Alex M. Pagnozzi¹ ,
José V. Manjón³ , and Jurgen Fripp¹ 

¹ The Australian e-Health Research Centre, Commonwealth Scientific and Industrial
Research Organisation, Brisbane, QLD, Australia

lee.reid@csiro.au

² Faculty of Medicine, The University of Queensland, St Lucia, QLD, Australia

³ Universitat Politècnica de València, ITACA, 46022 Valencia, Spain

Abstract. Image noise and motion degrade the quality of MR images. Block-matching methods are a well-demonstrated means of improving signal-to-noise ratios in such images. Ideally, block-matching methods would search within the entire image for matching patches to a target, leveraging an image's full informational redundancy, but this carries impractical computational costs. A well-known workaround, implemented in the traditional Non-Local Means (NLM) filter, is to search for matching patches only within a local neighborhood. Here, we detail a Global Approximate Block-matching (GAB) method that, via a self-organizing map, rapidly searches an entire image for patches similar to a target. Four sets of five T1 + five FLAIR images were acquired. GAB and NLM both denoised the T1s; the results were compared to subject-wise mean images with very low noise. GAB reliably produced images that were more similar to these 'templates' than NLM. This was repeated for the same images with motion-like artefacts artificially added. GAB, again, outperformed NLM. For this task, GAB further improved with multichannel inputs, even if the FLAIR image contained artefacts. GAB's competitive performance appeared to be due to a better balance between preserving image features and removing noise/artefacts. The performance of GAB and NLM variants hinted that GAB's advantage was not brute-force processing, but its ability to effectively search the whole image.

Keywords: Block-matching · Self-organizing map · Image noise
MRI artefacts · Non-local means

1 Introduction

Images with motion artefacts and/or high levels of noise are an unfortunate reality for many magnetic resonance projects and clinical facilities. Block-matching

(BM) techniques are typically designed to remove such noise. These methods typically accept a single input image, match similar cubes of voxels (patches) to one another, and compute a weighted average according to the patch similarity. This approach is known as ‘non-local means’ [2]. BM leverages the small patterns that exist throughout an image, such as the repeated folding of brain tissue. BM techniques are unable to take full advantage of this redundancy, however, because exhaustively comparing patches to one another is extremely computationally expensive. To circumvent this, many patch-based techniques such as the traditional non-local means method [3, 4], herein referred to as ‘NLM’, search only a selected local area for patches similar to a target, assuming that most of the pattern redundancy is local.

An alternative, or complementary, way to reduce this computational cost is through dimensionality reduction. The self-organizing map (SOM) is a flexible non-linear dimensionality reduction technique [5]. Briefly, SOMs are implemented as collection of nodes which each have local connectivity, a fixed position in low dimensional space (e.g. forming a 2D grid), and a trainable position in the high-dimensional space. SOMs train quickly through competitive learning, in a manner that results in a smooth projection between the two spaces. Intuitively, a trained 1D SOM can be thought of as optimally ‘snaking’ through high-dimensional space, much as principal components analysis (PCA) draws a straight line through such a space.

We have developed a ‘Global Approximate Block-matching’ algorithm (GAB) which performs a whole-image search for patches matching each target. To reduce this operation’s computational burden, GAB collapses each patch into a singular value (SV) through a dimensionality reduction method such as the SOM, which allows binary search lookups. Here, we describe GAB and demonstrate its use for removal of noise and motion artefacts in T1 images, utilizing both single- (T1) and multi-channel (T1 + FLAIR) input data. Performances of four variants of GAB are compared to that of NLM.

2 Methods

2.1 Overview

We tested the ability of GAB and NLM to perform two related tasks: (1) reduction of noise, and (2) removal of motion artefacts. Images used were acquired from a modern scanner, with simulated motion artefacts added for the second task. Performance was measured by the quantitative similarity between the processed T1 image and a low-noise template that was an average of five scans of the same subject.

2.2 Global Approximate Block Matching and Comparator Methods

GAB can run in single- or multi-channel mode, and is summarized in Fig. 1. To conserve memory, input image(s) are linearly scaled and voxel intensities stored

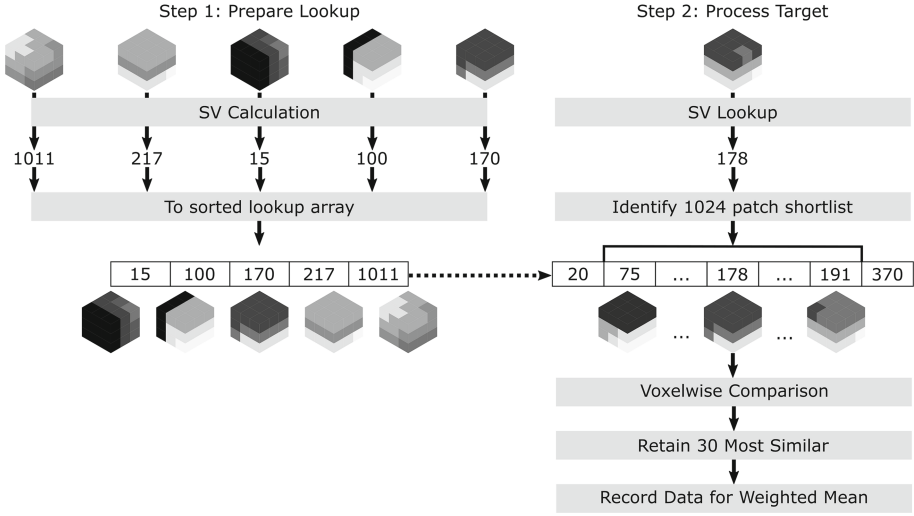


Fig. 1. In Step 1 (top left), the image was split into overlapping $3 \times 3 \times 3$ voxel patches, or $3 \times 3 \times 3 \times 2$ voxel patches in the case of dual-channel processing. For each patch, a singular value (SV) was calculated using one of four methods. Patches were then sorted by these SVs. In Step 2, for each target patch, the 1024 atlas patches with the most similar SVs were selected (the ‘shortlist’). The voxelwise sum of square differences (SSD) was calculated for these patches versus the target, and the 30 patches with the lowest SSDs selected to contribute toward final image reconstruction. See the text for details on final image reconstruction.

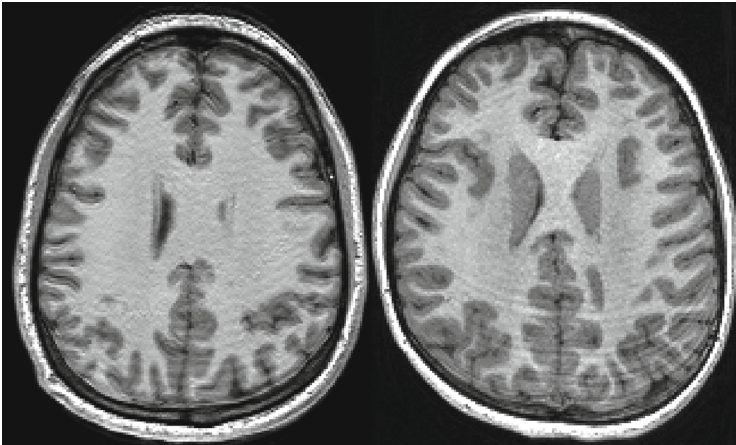


Fig. 2. T1 MPRAGE images with simulated (left) and real motion artefacts (right), acquired with the same sequence and scanner.

as 8-bit integers. The image within the brain mask is split into overlapping $3 \times 3 \times 3$ voxel patches ($3 \times 3 \times 3 \times n$ -channels for multi-channel mode). A SV is calculated for each patch (see below), by which patches are then sorted. During denoising, the location of a target patch can be found through a simple SV-indexed binary search. A 1024 patch shortlist is then identified as those items from 512 positions before to 511 indices after this index. The voxelwise sum of square differences (SSD) is calculated between the target patch and shortlist, excluding the original patch itself, identifying the 30 most similar patches to the target. These patches are each multiplied by their weight ($1/(SSD + 10^{-6})$), filtered by a Gaussian of $\sigma = 1$ voxel, and (all 27 voxels per weighted patch) are added to a ‘sum’ image. These weights, also multiplied by this Gaussian are added to a ‘weights’ image. Upon completion of all block matching, the sum image is divided by the weights image to generate a final reconstruction. This non-intuitive unweighting is required because each voxel in the ‘sum’ image is contributed to by up to 27 block-matching operations, each operation in turn averaging 30 weighted patches from the shortlist.

SVs were calculated solely from voxel intensities. Here, four SV methods were compared: PCA (ϵ_0), mean intensity, random (SV randomly generated), and SOM. PCA and SOM training sets were up to 10^7 patches randomly selected from the input image. Each SOM was arranged as a 1D array of 4096 equally-spaced nodes. Training took 10–40s for a single-channel input. SV calculation using the SOM was performed by locating a patch’s continuous position in this array (i.e. between the best matched node and its most similar neighbor) based on voxelwise SSD.

For a comparator method, we utilized NLM with a Rician noise model [3, 4], as implemented in DIPY. This is single-channel and utilizes $3 \times 3 \times 3$ voxel patches. Two search radii were tested: 5 voxels (728 patches; the default setting; NLM-5) and 6 voxels (1330 patches; NLM-6). Other settings were left as default.

2.3 Dataset

Four adults (28–32y) were scanned on a 3T Siemens Prisma. MPRAGE and FLAIR images (both 1 mm^3 resolution) were acquired in an alternating fashion. Five of each were acquired per subject. Noise and intensity biases can vary within an image, due in part to voxels’ differing locations within the B0 field and proximity to the head coil. We wished to reduce the likelihood of images displaying local intensity biases, or higher noise levels, at consistent anatomical locations, so as to generate more representative ‘ground truth’ (template) images (see below). As such, subjects reoriented their head before each T1 acquisition. The field of view was coarsely adjusted each time only to ensure that the full head was contained within the image. Subjects remained still during acquisition and gave written informed consent. Ethical approval was granted through the Herston Imaging Research Facility. All acquired images were N4 bias corrected [6] and linearly intensity scaled such that 95% of non-zero voxels were within the range of 1–100, prior to all processing and analysis.

Low Noise Template. For each subject, a low noise template T1 was generated as follows: (1) the aforementioned bias-corrected and intensity normalized images were upsampled to 0.5 mm^3 with trilinear interpolation; (2) an initial mean image was calculated, and brainmask generated by affine transforming the FSL MNI 152 T1 brainmask into this space using ANTS [1]; (3) each interpolated image was rigid-registered to the mean (from its original orientation) using ANTS; (4) another mean was then calculated. Steps 3 and 4 were repeated four times, the final mean constituting the ‘template’ for that subject. These templates exhibited a very low level of noise (Fig. 3). Regarding Step 1, upsampling was performed to compensate for the interpolation effects later during resampling to target image space. A final brainmask was generated for each by affine transforming the FSL MNI 152 T1 brainmask into this space using ANTS.

Motion Corruption. In this work, motion corruption was simulated as head rotation around a central point. To do so, we manipulated images in Fourier space, to represent partial filling of k-space at two head orientations. Although many factors contribute to image reconstruction, this method provided sufficiently-realistic artefacts for the purpose of testing artefact removal. An example of artificial motion corruption is provided in Fig. 2.

We generated a motion-corrupted copy of each acquired bias-corrected image (I.E. five T1 and five FLAIR images per subject). This center of rotation was defined in voxel space as the position between the two thalami on the mid (i.e. third) image acquired for each subject. As subjects were reoriented between scans, this meant that the anatomical center of rotation differed in each image whilst still being at an anatomically-plausible location. Motion corrupted images were constructed by combining Fourier space data from unaltered images with that from images that had been rotated in image space. T1 and FLAIR images were rotated by $(5^\circ, 1^\circ, \text{ and } 1^\circ)$ and $(1^\circ, 2^\circ, \text{ and } 2^\circ)$, respectively, to give non-correlated motion. After an FFT, these rotated images were multiplied by two to ensure a moderate artefact effect. Real and imaginary T1 images were constructed by combining coronal-planes 0–80 of the rotated data with 81–175 of the original image. Real and imaginary FLAIR images, which were acquired a different plane to our T1 images, were constructed from sagittal planes 0–70 of the rotated data and 71–191 of the original image. An FFT then reconstructed these data back into image space.

2.4 Performance Metric

Each BM method performed two sets of tasks: (1) reduction of noise in T1 images and (2) removal of the aforementioned motion artefacts from T1 images. For *noise reduction*, NLM accepted a single input image (T1). GAB was run with both a single-channel input (T1) and with multi-channel (T1+FLAIR) input. For *motion artefact* removal, NLM again accepted a single input image (corrupted T1), whilst GAB was run with single-channel (corrupted T1), multi-channel (corrupted T1 + FLAIR) and double-corrupted multichannel (corrupted T1 +

corrupted FLAIR) inputs. In all multi-channel conditions, analyses were run for the 25 possible combinations of T1 + FLAIR image pairs for each subject. For these conditions, FLAIR images were rigid-registered to the T1 using ANTS [1].

To calculate error, the high-resolution template for the appropriate subject was rigid registered to the target ('noise' or 'artefact' containing) T1 using ANTS [1] and resampled to 1 mm, providing voxelwise correspondence between images. BM-processed T1s were linearly scaled to match the histogram of this template image. The mean squared error (MSE) of image intensities was calculated between these processed images and their template within the template's brain-mask. Identical templates, registrations, and target images were used for each method. All registrations were visually checked.

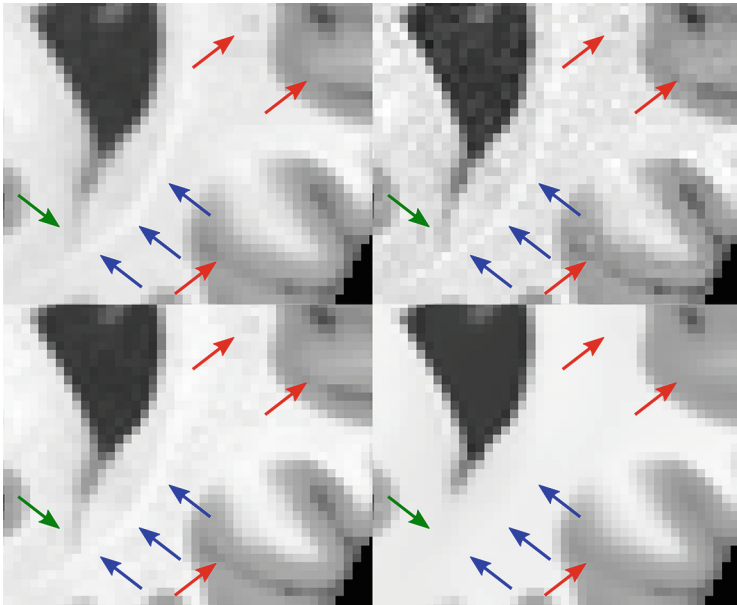


Fig. 3. Representative images of a template (top left); target bias-corrected T1 (top right); denoised result from single-channel GAB-SOM (bottom left); and denoised result from NLM performed with default settings (bottom right). Images are cropped to the brain mask. Note how GAB preserves image features such as sulci, the ventricular shape (green arrow) and intensities of structures including the optic radiation (blue arrows). (Color figure online)

3 Results

Tests were run on a single Tesla P100 GPU, 10 cores of a Xeon E5-2690 v4 compute node, and 64GB of RAM. GAB was run on Mono 4.6.2 and OpenCL 1.2. NLM ran through a Python 2.7.13 wrapper for compiled code. The mean

Table 1. Mean MSE \pm SD across all subjects, for each of the tested methods and conditions. The left two columns reflect denoising; the right three reflect artefact removal. The best denoising and artefact removal scores are highlighted. Abbreviations: MC, Motion Corrupted; NLM-5, traditional Non-Local Means (search radius = 5); NLM-6, traditional Non-Local Means (search radius = 6); PCA, Principal Components Analysis; GAB, Global Approximate Block Matching; SOM, Self-Organizing Map.

Method	Noise reduction		Motion artefact removal		
	T1	T1+FLAIR	MC T1	MC T1 + FLAIR	MC T1 + MC FLAIR
GAB-SOM	4.3 \pm 1.7	6.0 \pm 1.2	14.2 \pm 1.1	12.3 \pm 1.0	12.9 \pm 1.0
GAB-Mean	5.8 \pm 1.5	11.0 \pm 1.0	14.4 \pm 1.2	15.8 \pm 1.3	16.4 \pm 1.3
GAB-PCA	6.4 \pm 1.4	13.5 \pm 1.3	15.2 \pm 1.4	18.1 \pm 1.7	18.9 \pm 1.7
NLM-5	7.3 \pm 1.7	N/A	16.5 \pm 1.4	N/A	N/A
NLM-6	7.7 \pm 1.6	N/A	16.4 \pm 1.4	N/A	N/A
GAB-Random	16.4 \pm 6.7	28.8 \pm 6.8	21.1 \pm 3.9	30.8 \pm 8.1	30.0 \pm 7.3
No Correction	12.1 \pm 2.0	12.1 \pm 2.0	26.8 \pm 1.8	26.8 \pm 1.8	26.8 \pm 1.8

run time (of all variants) of GAB was 247 s for single-channel processing; GAB-SOM was \sim 30 s slower than other variants. Multi-channel processing increased run times by \sim 100% for GAB-SOM and \sim 30% for other SV methods. NLM ran in 40 s and 70 s on average for radii of 5 and 6 voxels, respectively.

3.1 Noise Reduction

Twenty single-channel, and 100 multi-channel, tests were conducted and are summarized in the left two columns of Table 1. When interpreting MSE values, note that images were intensity normalized to an approximate range of 0–100. Non-denoised T1 images had a MSE of 12.14 ± 2.00 (Mean \pm SD), compared to their respective templates. The most effective method was single-channel GAB-SOM (4.24 ± 1.65), which appeared to better balance noise removal with image feature preservation than NLM (7.29 ± 1.66 ; Fig. 3). NLM performance was not bolstered by a greater search radius. Multi-channel GAB provided poorer performance than single-channel GAB.

3.2 Artefact Removal

One motion-corrupted T1 image was unable to be registered with its template, and was discarded, resulting in 19 T1s contributing to 19 single-channel, and 95 multi-channel, tests. Results are summarized in the right three columns of Table 1. Corrupted T1 images showed a MSE of 26.83 ± 1.75 (Mean \pm SD). Single channel GAB-SOM (14.20 ± 1.07), GAB-Mean, and GAB-PCA, outperformed NLM (16.39 ± 1.38). Qualitatively, NLM often appeared to better reduce this artefact, but again this came at the expense of modification or deletion of features apparent in the template image. Including FLAIR data substantially improved GAB-SOM effectiveness (12.34 ± 1.00), even if the FLAIR image contained artefacts (12.86 ± 0.98).

3.3 Discussion

We proposed a BM method that initially performs a rapid global search for a shortlist of approximately similar patches that are to be compared voxelwise to a target. Ultimately, GAB, performed more accurately and reliably than NLM in removing image artefacts and noise, whilst preserving image features (Table 1; Fig. 3). NLM executed faster than GAB, but GAB's run time of $\sim 3\text{--}8$ min is still well within acceptable bounds for image preprocessing. Although we relied on artificial motion-like artefacts, these artefacts were a fair approximation of the genuine artefacts we see in our facility (Fig. 2). It is unlikely that any shortcomings in our simulation biased results as BM methods were not optimized for motion artefacts, nor particular sequences or scanners.

In single-channel mode, GAB's advantage appeared to be primarily due to its better leveraging of image redundancies, rather than brute force computation. This is indicated by its poor performance when patch shortlists were random, and the lesser performance of NLM even when performing 30% more voxelwise comparisons than GAB (NLM-6). GAB performed more poorly when denoising a good-quality T1 in multichannel mode than when that T1 was the only input. Results for GAB-Random implied that this may be due to inappropriate scaling of patch weighting factors ($1/\text{SSD}$), given such a high-dimensional space, and/or the FLAIR image being interpolated during reslicing to T1 space. When processing images with artefacts, however, multichannel information improved the performance of GAB-SOM, even with mild artefacts in the FLAIR image. Here, the SOM's non-linearity presumably enabled a meaningful 54-to-1 dimensionality reduction, given the substantially poorer results of GAB-PCA.

In conclusion, we proposed a global approximate BM method that uses a SOM for dimensionality reduction. In real images this outperformed traditional NLM when removing both image noise and simulated motion artefacts.

References

1. 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. Image Anal.* **12**(1), 26–41 (2008)
2. Buades, A., Coll, B., Morel, J.M.: A non-local algorithm for image denoising. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 60–65. IEEE (2005)
3. Coupé, P., Manjón, J., Robles, M., Collins, D.: Adaptive multiresolution non-local means filter for three-dimensional magnetic resonance image denoising. *IET Image Process.* **6**(5), 558 (2012)
4. Coupé, P., Yger, P., Prima, S., Hellier, P., Kervrann, C., Barillot, C.: An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. *IEEE Trans. Med. Imaging* **27**(4), 425–441 (2008)
5. Kohonen, T.: Self-organized formation of topologically correct feature maps. *Biol. Cybern.* **43**(1), 59–69 (1982)
6. Tustison, N.J., et al.: N4ITK: improved N3 bias correction. *IEEE Trans. Med. Imaging* **29**(6), 1310–20 (2010)