ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 30 (2006) 436 – 443

Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data

Jorge Jovicich,[a,*] Silvester Czanner,[a] Douglas Greve,[a] Elizabeth Haley,[a] Andre van der Kouwe,[a] Randy Gollub,[a] David Kennedy,[a] Franz Schmitt,[b] Gregory Brown,[c] James MacFall,[d] Bruce Fischl,[a,e] and Anders Dale [f]

[a]*MGH/MIT/HMS Athinoula A. Martinos Center for Biomedical Imaging, Building 149, 13th Street, Radiology/CNY149-Room 2301, Charlestown, MA 02129, USA*
[b]*Siemens Medical Solutions, MRPF, Manufacturing Field Generating Units, Allee am Roethelheimpark 2, 91052 Erlangen, Germany*
[c]*Laboratory of Cognitive Imaging, Department of Psychiatry, University of California, San Diego, La Jolla, CA 92092, USA*
[d]*Neuropsychiatric Imaging Research Laboratory (NIRL), Duke University, Durham, NC 27708, USA*
[e]*Computer Science and Artificial Intelligence Lab, Massachusetts Institute of Technology, Cambridge, MA 02142, USA*
[f]*Departments of Neurosciences and Radiology, University of California San Diego, San Diego, CA 92092, USA*

Received 28 April 2005; revised 5 September 2005; accepted 19 September 2005
Available online 21 November 2005

Longitudinal and multi-site clinical studies create the imperative to characterize and correct technological sources of variance that limit image reproducibility in high-resolution structural MRI studies, thus facilitating precise, quantitative, platform-independent, multi-site evaluation. In this work, we investigated the effects that imaging gradient non-linearity have on reproducibility of multi-site human MRI. We applied an image distortion correction method based on spherical harmonics description of the gradients and verified the accuracy of the method using phantom data. The correction method was then applied to the brain image data from a group of subjects scanned twice at multiple sites having different 1.5 T platforms. Within-site and across-site variability of the image data was assessed by evaluating voxel-based image intensity reproducibility. The image intensity reproducibility of the human brain data was significantly improved with distortion correction, suggesting that this method may offer improved reproducibility in morphometry studies. We provide the source code for the gradient distortion algorithm together with the phantom data.

Keywords: Multi-site calibration; Gradient non-linearity distortions; Human structural MRI

Introduction

Multi-site and longitudinal neuroimaging studies are increasingly becoming a standard element of clinical neuropsychiatric research for diagnosing and evaluating neurological impairments (Ashburner et al., 2003; Grundman et al., 2002; Fox and Schott, 2004). One of the challenges of both longitudinal and multi-site studies is to minimize image variability caused by technological factors (e.g., hardware differences, hardware imperfections), as such variability may be confounded with specific disease-related changes in the images thereby limiting the power to detect and follow the progression of disease biomarkers. Optimization of image reproducibility motivates the calibration of acquisition protocols and the characterization and correction of scanner-specific image variability effects. This is particularly important when data from multiple sites and MRI vendors are to be combined.

An important task in this effort is to correct for site-specific image distortions in order to allow accurate cross-site comparisons of quantitative morphometry results. Image distortions can potentially affect the accuracy of volume (Fischl et al., 2002), shape (Miller, 2004) and boundary (Barnes et al., 2004) measurements. Although distortions in MRI can arise from several factors, one of the most prominent in structural MRI is imaging gradient non-linearity, which degrades both geometric and image intensity accuracy. While in principle gradient distortions may be addressable using manufacturer-supplied software, the currently available correction algorithms work only in two-dimension (2-D) providing an incomplete solution to the problem (Wang et al., 2004a). Three-dimensional (3-D) algorithms to correct gradient non-linearity distortions have been investigated using phantoms. To summarize, two main correction methods have been developed: (a) 3-D measurement of the geometric displacements due to distortions using specially designed phantoms followed by an image transformation to perform the correction (Wang et al., 2004b,c; Langlois et al., 1999) and (b) 3-D calculation of the geometric displacements from the spherical harmonic expansion for the representation of the

* Corresponding author. Fax: +1 617 726 7422.
E-mail address: jovicich@nmr.mgh.harvard.edu (J. Jovicich).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2005.09.046

magnetic fields generated by the gradient coils (Schmitt, 1985; Janke et al., 2004; Wald et al., 2001). The second method was used in this work. As yet, there is no quantitative study that systematically compares these correction methods. More importantly, no study investigates the effects of these distortion correction methods on test–retest reproducibility of multi-site human structural MRI data.

The purpose of this work was (i) to quantitatively characterize and correct site-specific image distortions caused by gradient non-linearity in a phantom study, and (ii) to assess if gradient non-linearity distortion correction improves image reproducibility when the same subjects are scanned at multiple sites in multiple sessions. To keep our results independent of brain morphometry analysis tools, here, we focus only on the reproducibility of image intensity for the human data. Parts of these results have been presented at recent meetings (Jovicich et al., 2004; Jovicich et al., 2003).

Materials and methods

Human and phantom image data acquisitions

Four sites with clinical 1.5 T whole body scanners used in regular functional and structural MRI studies participated in this study. These systems included: (a) GE Medical Systems with Cardiac Resonator Module (CRM) gradient coils (maximum strength = 40 mT/m, slew rate = 150 T/m/s) at Duke University Medical Center (Duke) and at Brigham and Women's Hospital (BWH); (b) GE Medical Systems with Brain Resonator Module (BRM) gradient coils at the University of California San Diego (UCSD, 22 mT/m, 120 mT/m/ms); and (c) Siemens Medical Systems with Sonata gradient coils at the Massachusetts General Hospital (MGH, Sonata gradients, 40 mT/m, 200 T/m/s). The GE scanners allowed a default 2-D in-plane distortion correction, which was used for the acquisitions on the GE platforms tested. The version of the Siemens Sonata system used did not enable this, so that data were acquired with no distortion correction at all. Therefore, the 3-D gradient distortion correction effects on image reproducibility can be evaluated against acquisitions with no correction at all (our Siemens data), or acquisitions that had the vendors' 2-D corrections (our GE data).

Phantom data (from the 4 sites) and test–retest human data (from MGH, UCSD and Duke) were collected using an acquisition protocol that included a 3-D-spoiled gradient echo volume (TR = 20 ms, TE = 6 ms, flip angle 30°, 256 × 192, 1.3 mm thick 124 sagittal slabs, FOV 25 cm, 8 min 12 s acquisition) and used the vendors' standard head RF coil. Five healthy volunteers gave written informed consent to participate in this multi-site study, which was approved by the institutional review boards at each participating site. Each subject (1 female, 4 males, and average age 39) was scanned twice on each site, in different sessions using this acquisition protocol. For within-site repetitions the average time interval for rest–retest scans was 19 days (minimum time 1 day, maximum time 8 months), and for across-site repetitions the average time interval was 8 months (minimum, 2 months; maximum, 15 months).

A special cylindrical phantom (250 mm length × 220 mm diameter) consisting of 25 plastic plates, each 10 mm thick was specifically built for assessing geometric distortions from gradient non-linearities (Franz Schmitt, personal communication). Each plate had a pattern of holes 3 mm in diameter, going through the thickness of the plate and perpendicular to the faces of the plate. These holes formed a 2-D rectangular grid with 10 ± 0.05 mm spacing. In addition, on the two sides of each plate, each hole was enlarged with a half spherical depression of 7 mm in diameter. In this way, with the plates glued to each other and the 3 mm holes aligned, the phantom formed a 3-D 10 ± 0.05 mm grid of 7 mm diameter spheres that could be filled with a fluid (water doped with salt) through the passing 3 mm holes. The phantom acquisition protocol scan was as for the humans, but with 2 mm slices and a 45-cm FOV. In addition to the 3-D geometric phantom, which was scanned at all participating sites and is not suitable for measuring image intensity uniformity, a standard saline filled uniform cylindrical phantom (250 mm, length; 150 mm, diameter) was scanned at one of the sites (MGH) to evaluate image intensity uniformity improvements from gradient distortion correction. This uniform phantom was scanned three times within the same session, at slightly different positions within the field of view, using the standard 3-D T1-weighted sequence described above.

Gradient distortion correction

The goal of the distortion correction is to transform (interpolate) the original distorted image into a corrected one by displacing each voxel into an estimate of its correct 3-D location and by scaling each voxel's intensity to account for voxel-size distortions. A laboratory-based coordinate system was used to calculate the displacements in Euclidean coordinates (x, y, z). This coordinate axis system had its origin at the iso-center of the scanner, with the z direction aligned with the main field B_0, the x direction the left/right and the y direction superior/inferior. The 3-D displacements along each of the three directions can be calculated from the non-linear terms of the magnetic field generated by each of the gradient coils (Janke et al., 2004). These fields are usually provided by the vendor in the form of a truncated series of spherical harmonic coefficients and instructions for how these coefficients are normalized to calculate field magnitudes (5 terms were available for the GE CRM and BRM gradients and 11 for Siemens Sonata gradients). The intensity correction is the Jacobian determinant calculated from the non-linear magnetic field terms. The displacements and intensity correction tables were pre computed (using trilinear 3-D interpolation) in a 3-D cylindrical grid (300 mm in diameter, 300 mm long) that is large enough to contain any imaging volume acquired within the head RF coil.

Evaluation of unwarping effects on test–retest reproducibility

The quantitative characterization of the distortion corrections (unwarping) on the phantom data was evaluated in two ways: (a) measurement of the phantom diameter at several positions along the z axis in the raw and corrected images and comparison of these values with the true phantom diameter (220 mm) to quantify how the % errors were reduced with distortion correction, and (b) estimation of the range of image intensity errors and magnitude displacements within a spherical region of interest centered at the magnet's iso-center and large enough to encompass the full brain volume in a conventional scan.

The unwarping effects on the human data were evaluated by assessing if voxel-based image intensity variability was reduced with distortion correction, both within and across sites. For each subject, the T1-weighted volume was skull stripped, co-registered with the other scans of the same subject, and intensity normalized (brain mean 100). The images were then grouped in the following ways to compute variability: within-site test–retest (for each site