Automatic quality assessment for multi-slice 2D FLAIR MR imaging

B. Mortamet[1], M. A. Bernstein[2], C. R. Jack[2], J. L. Gunter[2], M. Shiung[2], R. Meuli[3], J-P. Thiran[4], and G. Krueger[1]
[1]Advanced Clinical Imaging Technology, Siemens Healthcare Sector IM&WS S - CIBM, Lausanne, Switzerland, [2]Mayo Clinic, Rochester, MI, United States, [3]CHUV, Radiology, Lausanne, Switzerland, [4]Signal Processing Laboratory (LTS5) Ecole Polytechnique Fédérale de Lausanne

Introduction: Many clinical MRI protocols use the fluid attenuated inversion recovery (FLAIR) contrast to better delineate tissue abnormalities such as white matter lesions. Most FLAIR protocols acquire data in a 2D fashion. FLAIR images are often degraded by patient motion, especially when scanning uncooperative patients. Typical motion patterns induce inter-slice misalignment, ghosting and blurring artifacts which can obscure the pathology or mislead automated image analysis algorithms. In this work, we propose the use of an automated quality assessment algorithm for clinical T2-weighted 2D-FLAIR data.

Material & Methods: In this work data from 99 patient scans are analyzed. All images had been acquired on a Signa 3T scanner (GE Healthcare, Milwaukee, WI, USA) using a T2-weighted 2D FLAIR pulse-sequence (TR-TE=11000-147ms, 0.8594x0.8594x3mm^3 resolution, 256x256x42-48 matrix size, 3-4 concatenations and acquisition time 4.5-6 minutes). Qualitative ratings of image quality were assigned on all data by an experienced reader. Quantitative ratings are performed with an approach that extends recent work (1). Here, the 2D origin of the data is specifically addressed.

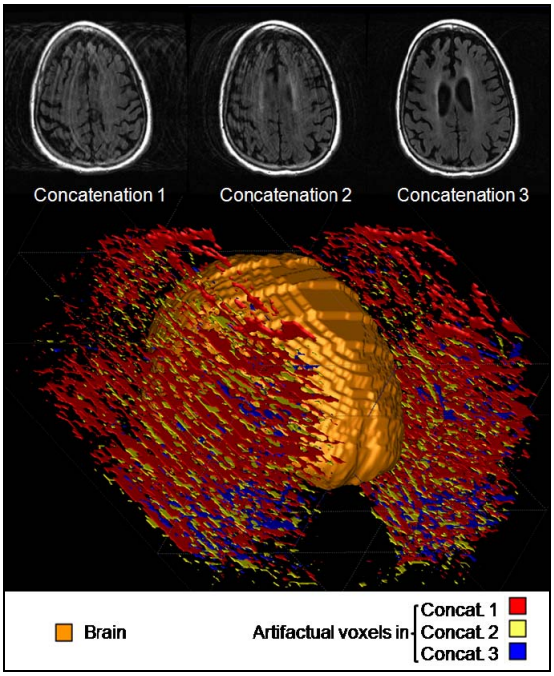
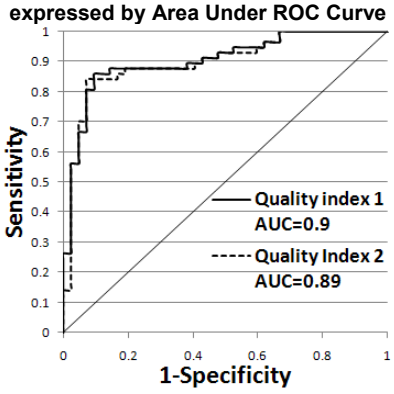


Fig1. Magnitude axial slices from a 2D-FLAIR low-quality dataset with severe/moderate/mild motion in 1st/2nd/3rd concatenations as reflected by detected artifactual voxels in 3D rendering

The proposed algorithm for image quality assessment is based on the assumption that motion leads to projection of artifactual signal intensities into the air background over which a careful analysis is performed (1) in a three-step process. First, all axial slices are combined to create a 3D volume in order to extract the relevant background regions using an edge detection and atlas-based refinement of the volume of interest. The remainder of the analysis considers each concatenation to produce independent image volumes. Second, a model-free quality index QI1 is computed as the proportion of artifactual voxels relative to the background size for each concatenation. Artifactual voxels are isolated by means of anisotropic morphological operations (see the 3D-rendering, Fig1). This first quality measure is sensitive to artifacts exhibiting clustered property (i.e. motion patterns) and is strengthened by computing an additional model-based quality index QI2 that detects more subtle artifacts (i.e. blurring). More specifically, QI2 examines the noise intensity distribution by fitting the histogram of background (without artifactual voxels) with chi distribution (2) using maximum likelihood estimation. The goodness-of-fit (absolute error) added up to QI1 is the second quality index QI2 as described previously (1). Finally, overall quality scores are computed as the summation of each QI over concatenations (i.e. if motion severely corrupts one concatenation, the entire scan will be classified as low-quality). Sensitivity and specificity are considered as the true positives (high-quality prediction) rate and the true negatives (low-quality prediction) rate, respectively. Receiver Operating Characteristics (ROC) curve (Fig2) represents the range of combinations of sensitivity and specificity achievable over the range of possible cutoff points for our quality indices (QIs). ROC is used (a) to evaluate the performance of each QI measured by area under the curve (Fig2) and (b) to compare the discriminative abilities of QIs in order to identify the preferred one. QIs cutoff values are determined by equalizing sensitivity and specificity. Analysis of variance is performed to detect significant mean differences among quality groups (low-/high-quality) resulting from QI1-/QI2-based classification.

Fig2. Performance of quality tests expressed by Area Under ROC Curve



Results: Both quality indices exhibit excellent prediction performances (AUC~0.9) and perform equally well (no conclusion can be drawn concerning which test performs best). When comparing to the expert quality ratings and trading off sensitivity and specificity at approximately equal rates (88.22% and 88.53% on average for QI1 and QI2, respectively) provides cutoff points of 0.3% and 0.8%, respectively. ANOVA reveals significant differences among the two quality groups for each QI ($p<0.001$). Overall, the model fits the data well and our quality indices appear to be both accurate and consistent.

Discussion: In this work, a method for automatic quality assessment, recently validated on 750 3D T1-weighted datasets, is adapted to measure image quality in multi-slice 2D MRIs with FLAIR contrast. To verify this extension, we evaluate 99 brain scans with FLAIR contrasts and achieve sensitivity and specificity of 88% to classify high- or low-quality scans. The results exhibit very similar performances as observed in the 3D-approach and prove to efficiently and robustly predict image quality in a 2D-data collection. This suggests that the technique could be used with other contrast types as long as the method assumptions are not violated (i.e. sufficient background 3D volume and MR signal strength is given). The method provides unbiased exclusion criteria for assessing overall image quality as well as identifying concatenations (i.e. slices) that are likely to be misaligned. It could have a great potential in both routine clinical practice and multicenter studies. In particular, if integrated with the online image reconstruction, it could provide immediate feedback to the MR technologist to repeat low-quality scans within the same session.

References: [1] Mortamet B et al. *Automatic quality assessment in structural brain magnetic resonance imaging*. MRM 2009;62:365–372. [2] Constantinides CD et al. *Signal-to-noise measurements in magnitude images from NMR phased arrays*. MRM 1997;38(5):852-857.
Acknowledgements: NIH grant AG11378 for data collection. Supported by CIBM (UNIL/UNIGE/HUG/CHUV/EPFL) and Leenaards and Jeantet Foundations