Hotelling Observer Efficiency Image Quality Metric for Compressed Sensing MRI

Christian G. Graff[1]

[1]*Division of Imaging and Applied Mathematics, U.S. Food and Drug Administration, Silver Spring, MD, United States*

Target Audience: This work should be of interest to developers and users of advanced MR image reconstruction algorithms such as compressed sensing who wish to quantitatively evaluate the performance of these algorithms with image quality metrics that reflect clinical utility.

Purpose: Traditional Cartesian MR data reconstruction has well-understood noise properties. K-space data contaminated by Gaussian electronic noise are reconstructed via the linear FFT algorithm to obtain magnitude images containing Rician noise, often well approximated as Gaussian. Modern dynamic and quantitative imaging however benefits from faster imaging, achieved by acquiring less data and compensating through parallel imaging or constrained iterative reconstructions such as compressed sensing (CS). CS algorithms are non-linear, invalidating the original linear systems-based analysis of noise in MR images and may produce images of high apparent quality measured by metrics such as pixel SNR ($\mathrm{SNR_p}$) through de-nosing or other non-linear effects, while incongruently suppressing or modifying clinically-relevant image features. The purpose of this work is to develop an image quality metric based on task-specific Hotelling SNR ($\mathrm{SNR_h}$) that reflects clinical utility, overcoming limitations of $\mathrm{SNR_p}$ and related metrics to allow for more meaningful quantitative comparisons and optimization of accelerated acquisitions and CS reconstructions.

Methods: Consider an imaged object (human or phantom) represented by a spatial distribution of MR parameters, $\boldsymbol{f}(\boldsymbol{r}) = \{T_1(\boldsymbol{r}), T_2(\boldsymbol{r}), \rho(\boldsymbol{r}) \ldots\}$, and a detection task, i.e., determine the presence (hypothesis $\mathrm{H_1}$) or absence (hypothesis $\mathrm{H_0}$) of a clinically important object feature. The object $\boldsymbol{f}$ passes through the MR imaging system $\mathcal{H}$ to generate k-space data $\boldsymbol{k}$ corrupted by Gaussian noise $\boldsymbol{n}$: $\boldsymbol{k} = \mathcal{H}\boldsymbol{f}(\boldsymbol{r}) + \boldsymbol{n}$. K-space data $\boldsymbol{k}$ is then passed to a CS reconstruction algorithm $\mathcal{R}$ to generate image $\boldsymbol{g}$, thus $\boldsymbol{g} = \mathcal{R}\{\boldsymbol{k}\} = \mathcal{R}\{\mathcal{H}\boldsymbol{f}(\boldsymbol{r}) + \boldsymbol{n}\}$. The Hotelling observer[1] is the linear observer function that maximizes separability of object classes by de-correlating the noise in the image $\boldsymbol{g}$ and subsequently multiplying by a template, the expected difference in the image between hypotheses $\mathrm{H_0}$ and $\mathrm{H_1}$. The corresponding Hotelling observer SNR is defined by the equation

$$\mathrm{SNR_h^2}=[\langle \boldsymbol{g}|\mathrm{H_1}\rangle - \langle \boldsymbol{g}|\mathrm{H_0}\rangle]^{\dagger} K_{\boldsymbol{g}}^{-1}[\langle \boldsymbol{g}|\mathrm{H_1}\rangle - \langle \boldsymbol{g}|\mathrm{H_0}\rangle],$$

where $\langle \boldsymbol{g}|\mathrm{H_j}\rangle$ represents the expectation of the image under hypotheses $\mathrm{H_j}$ and $K_{\boldsymbol{g}}$ is the image noise covariance matrix. The feature of interest is assumed to be small enough to not affect the noise correlation structure. As a measure of image quality we propose the *Hotelling efficiency* ε_h, defined as the ratio of $\mathrm{SNR_h}$ for a fully-sampled Cartesian acquisition and $\mathrm{SNR_h}$ for an under-sampled CS acquisition, thus representing the loss in task performance by resorting to an under-sampled acquisition,

$$\varepsilon_h = \frac{\mathrm{SNR_h}(\mathcal{H} = \text{ under-sampled}, \mathcal{R} = \text{ CS reconstruction})}{\mathrm{SNR_h}(\mathcal{H} = \text{Cartesian fully-sampled}, \mathcal{R} = \text{ FFT reconstruction})}.$$

To test this metric a simulated imaging system and high-resolution anthropomorphic digital brain phantom[2] were utilized to generate image recons for a 2D Cartesian spin-echo sequence (TE=30ms, TR=1500ms, matrix=256x256) and under-sampled 2D pseudo-radial spin-echo sequence with the same timings (64 or 128 views). The CS reconstruction was an implementation of the Split-Bregman technique[3] with total variation and wavelet sparsity constraints. The task was to detect a localized difference in proton density ρ of diameter 4 mm at a given location within the brain. The change in ρ was calibrated for each location such that the Hotelling observer had a 95% chance of correctly discriminating between a random pair of $\mathrm{H_0}$ and $\mathrm{H_1}$ images (corresponding to $\mathrm{SNR_h} \simeq 2.34$) for the fully-sampled acquisition with the variance of $\boldsymbol{n}$ set to achieve $\mathrm{SNR_p} \simeq 30$ in gray matter.

Results: Figure 1 compares ε_h for 64 and 128-view acquisitions using the same CS recon. Mean ε_h was 0.689 and 0.948 respectively with low variation across the FOV. Figure 2 compares ε_h and SNRp for the 128-view data and a 64-view dataset reconstructed with increased regularization. $\mathrm{SNR_p}$ was higher in 74.8% of ROIs for the highly regularized 64-view data, while ε_h indicates greater task performance for the 128-view dataset.

Discussion: Figure 2 illustrates the dangers of using $\mathrm{SNR_p}$ in the presence of non-linear regularization, which can have arbitrary amounts of smoothing or de-noising depending on the choice of regularization parameters. At some level de-noising will begin to negatively impact clinical utility, which is not captured by $\mathrm{SNR_p}$, while ε_h that measures task performance more effectively captures the impact of these types of regularization.

Conclusion: We have developed and tested an image quality metric that is able to compare and optimize CS regularized reconstructions. This method better reflects clinical utility and is not susceptible to the problems encountered when using $\mathrm{SNR_p}$ for non-linear reconstructions.

References: 1. Barrett, HH, Myers KJ, *Foundations of Image Science*, Hoboken NJ: Wiley 2004. 2. Graff, CG, *Framework for Task-Based Assessment of MR Image Quality*, Proc. ISMRM 2013 p701. 3. Goldstein T, Osher S, *The Split Bregman Method for L1-Regularized Problems*, SIAM Imag. Sci. v**2**:2, p323-343.

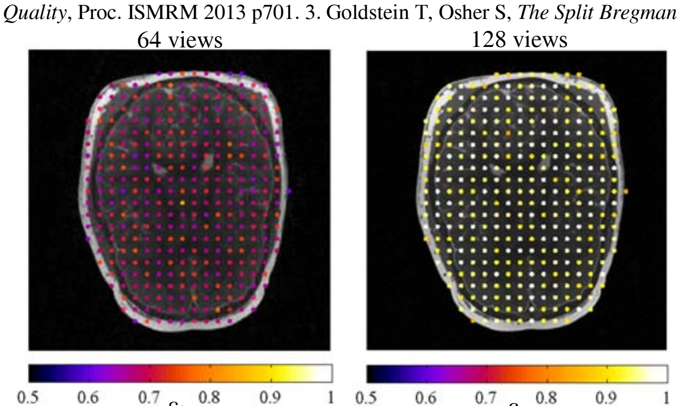

Fig 1. Calculated efficiency ε_h for locations throughout the brain for 64 radial views (l) and 128 radial views (r).

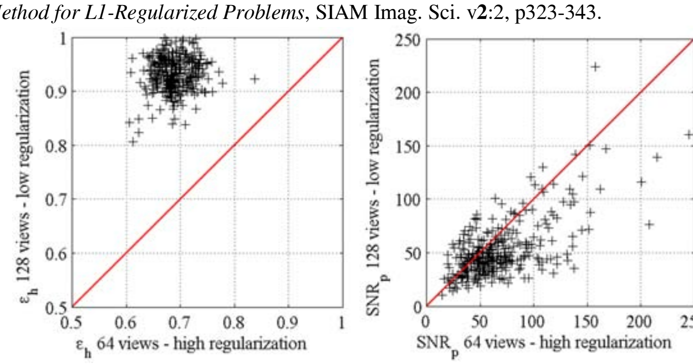

Fig 2. The ε_h (l) and $\mathrm{SNR_p}$ (r) metrics for 64-view recon with high regularization vs. 128-view recon with lower regularization for each ROI shown in Fig. 1. $\mathrm{SNR_p}$ indicates better performance with 64-view data for 74.8% of ROIs while ε_h indicates 128-view data are superior for all ROIs.