

IMAGE QUALITY EVALUATION IN CLINICAL RESEARCH: A CASE STUDY ON BRAIN AND CARDIAC MRI IMAGES IN MULTI-CENTER CLINICAL TRIALS

¹V. NIKHILA, ²T.MAHABOOB RASOOL

¹PG Student, Dept of ECE(DECS), BCETFW, Kadapa, AP, India.

²Associate Professor, Dept of ECE, BCETFW, Kadapa, AP, India.

ABSTRACT- MRI system images are important components in the development of drugs because it can reveal the underlying pathology in diseases. Unfortunately the processes of image acquisition, storage, transmission, processing and analysis can influence image quality with the risk of compromising the reliability of MRI-based data. Therefore, it is necessary to monitor image quality throughout the different stages of the imaging workflow. This report describes a new approach to evaluate the quality of an MRI slice in multi-center clinical trials. The design philosophy assumes that an MRI slice, like all natural images, possess statistical properties that can describe different levels of contrast degradation. A unique set of pixel configuration is assigned to each possible level of contrast-distorted MRI slice. Invocation of the central limit theorem results in two separate Gaussian distributions. The central limit theorem says that the mean and standard deviation of pixel configuration assigned to each possible level of contrast degradation will follow a normal distribution. The mean of each normal distribution corresponds to the mean and standard deviation of the underlying ideal image. Quality prediction processes for a test image can be summarized into four steps. The first step extract local contrast feature image from the test image. The second step computes the mean and standard deviation of the feature image. The third step separately standardizes each normal distribution using the mean and standard deviation computed from the feature image. This gives two separate z-scores. The fourth step predicts the lightness contrast quality score and the texture contrast quality score from cumulative distribution function of the appropriate normal distribution. The proposed method was evaluated objectively on brain and cardiac MRI volume data using four different types and levels of degradation. The four types of degradation are Rician noise, circular blur, motion blur and intensity nonuniformity also known as bias fields. Objective evaluation was validated using a proposed variation of difference of mean opinion scores. Results from performance evaluation show that the proposed method will be suitable to monitor and standardize image quality throughout the different stages of imaging workflow in large clinical trials.

I. INTRODUCTION

Cardiac MRI is used for the assessment of cardiac structure and function such as the characterisation of myocardial tissue, blood volume and blood flow measurements. Other applications include the diagnosis of cardiac amyloidosis, identification of regions with left ventricle hypertrophy. evaluation of pathology in congenital heart disease, cardiac masses, cardiomyopathies and valvular heart diseases. A typical setup of large clinical trial consist of multiple locations across the globe. Each location is referred to as clinical trial site. Multiple clinical trial sites interact with a clinical research organization (CRO). The CRO manages the clinical trial of drugs for the sponsoring pharmaceutical organization. Three major activities are carried out at the clinical trial sites. They are enrolment of subjects, administration of the drug under trial to subjects and the acquisition of images from MRI system. Daily, several thousands of slices contained in hundreds of MRI volume data are routed from the clinical trial sites to a CRO. The quality of a medical image can be assessed either in terms of its measurable physical properties or the medical goal of the image. Measurable physical properties include visual attributes such as texture, contrast, sharpness and noise. The medical goal of the image is the task of the imaging system from which the image was acquired. The following characteristics and requirements in a large clinical trial justify the need for a no-reference objective quality evaluation of MRI images.

1) Quality Monitoring Through Stages of Imaging Workflow: The task of acquisition, storage, transmission, processing and analysis can have adverse effect on image quality. The potential of an MRI system to generate high contrast image can be reduced by improper system parameter settings. Concern for patient comfort may require trade-off between signal-to-noise ratio, image resolution and length of scan time. MRI signal is sensitive to motion. It is extremely difficult for a subject to maintain ideal pose during every visit for image acquisition. Patient motion and the manifestations of physiological functions such as breathing and heart beat introduces blur and artifacts during acquisition. The processes of storage and

transmission of images can introduce blur and blocking which reduces image details and sharpness. Wavelet and total variation approaches to the removal of noise, deblurring and enhancement introduce ringing artifacts and blurred edges resulting in loss of diagnostic information in the images.

2) Limitations of Subjective Quality Evaluation: Subjective evaluation by human observers is regarded as the gold standard for quality evaluation. However several factors limits its application in large clinical trials. Trained experts cannot cope with the large volume of data that are processed in clinical trials. Human emotions, environmental and lighting conditions influence subjective evaluation by radiologists and trained MRI readers resulting in intra and inter expert variability. Efficient processing and the management of MRI data demand real-time operation offered by objective quality evaluation. There is little tolerance for the cumbersomeness and the variability of the outcomes of subjective image quality evaluation.

3) Inter-Site Variations in MRI System Parameters: Cost-saving measures by the pharmaceutical companies requires that only MRI systems available at the clinical trial sites are utilized for acquisition. The consequences are variations in the quality of images from the different scanner manufacturers. It is impractical to construct a single image model to act as reference image for the evaluation of images from the different trial sites. In the real world there is no image having ideal qualities that can be regarded as a reference image. Thus no-reference method based on image quality attributes is a more practical approach to evaluate image quality. Good clinical practice demands high level of integrity from clinical data. The reliability of metrics derived from MRI-based images acquired from the different clinical trials sites, to a large extent, is dependent on the reevaluation and standardization of image quality before data analysis.

4) Intra-Subject Variations in Acquisition Parameters: Images of a subject acquired at different time points requires registration before analysis. There is also the possibility of scanner change at the clinical trial sites for clinical trials that span over a period of time. Intensity mismatch is common occurrence between images acquired at different time points. Processing tasks such as intensity normalization and image registration demands quality re-evaluation to assess the integrity of information contained in the images.

5) Conformity with Acquisition Protocols: Brain measurements derived from MRI systems are susceptible to differences in imaging sequence parameters. In clinical trials the sponsoring pharmaceutical organization outline acquisition protocols which include requirements on image

quality to ensure optimal utility of the images and avoid inaccurate diagnosis. Post-acquisition image quality evaluation at the CRO is one of key steps towards conformity with the acquisition protocols. Signal-to-noise ratio (SNR), mean square error (MSE) and peak signal-to-noise ratio (PSNR) are the popular full reference quality evaluation methods at the acquisition stage of MRI images. Several post-acquisition evaluation methods have been proposed. The report in apply analysis of variance (ANOVA) algorithm to assess the variation of several quality measures with different levels of distortions. The authors in combine the detection of artifacts and estimation of noise level to measure image quality. Recently the report in propose a no-reference method which predict brain MRI quality based on five quality attributes. The attributes are lightness, contrast, sharpness, texture details and noise. The report in predict image quality by casting the relationship between entropy and classical image quality attributes on Bayesian framework. Another report computes image quality by using three separate geo-spatial feature vectors extracted from a test image to standardize corresponding Gaussian distributed quality models. Other recent reports assess image quality based on how subject motion during acquisition bias structural information and metrics derived from the image. This report provide only a brief review of image quality evaluation. Detailed review of quality evaluation methods for medical images are available.

Current quality evaluation methods for MRI images are designed using different quality evaluation models for specific stages of the imaging workflow. There is no specific quality evaluation method that can effectively evaluate the quality of an image from acquisition through the different stages of the imaging workflow. Current methods such as which assume that background noise voxels contain information pertinent to quality of images have several shortcomings. First, they cannot be applied to parallel imaging technique which the noise level is variable across the image field of view. Second, background-based noise estimation methods are only suitable for images in which the field of view allows the MRI system to capture air-tissue boundary and generate images with background that describe the surrounding air. For this reason background-based noise estimation methods will be suitable for brain MRI images but useless for cardiac and lung MRI images with small field of view because background voxels are not available. Even for brain MRI images, the performance of these algorithms can be significantly limited by underestimation of noise level when the number of background voxels are limited or corrupted by artifacts. The need for a large population to extract relevant features for the construction of quality model can be regarded as a

drawback for the reports. This drawbacks makes it difficult to achieve the much desired consistent quality evaluation required in good clinical practice. Thus, it can be said that current algorithms are unsuitable for large clinical trials. This paper describes a new objective, no-reference attribute based quality evaluation method for MRI images. It is based on the application of moments-preserving property of additive linear degradation model, labeling problem and the central limit theorem to the pixel configurations that describe each possible level of contrast degradation in an MRI slice. Labeling problem is used to classify the different levels of degradation in an image.

II. EXISTING SYSTEM

Image segmentation is one of the most important and active research area in the medical imaging domain. It can be defined as the delineation of one or several structures of interest within the image. Automated methods are sought in order to avoid the time consuming burden of manually contouring the structures. The problem is particularly difficult in the context of brain tumors. Indeed, most tumors have heterogeneous appearances and their intensity range overlap with the healthy tissues'. The presence of a necrotic core is frequent (especially for glioblastomas, but it also occurs for DLGGs) resulting on a strong contrast with the "active" tumor. Prior information regarding the shape of the tumor cannot be used as they have variable sizes and shapes. DLGGs in particular, have very fuzzy and irregular boundaries due to their infiltrative nature. Edema (swelling of brain tissue around the tumor) and mass effect (tissue displacement induced by the tumor) are quite uncommon due to the slow-growing nature of the DLGGs [Sanai 2011]. In this context, the simplest segmentation methods such as thresholding or region growing are insufficient [Gibbs 1996]. Despite extensive and promising work in the tumor segmentation field, obtaining accurate and reliable segmentations of brain tumors remains a difficult task. Segmentation methods can be grouped in two categories: surface and region-based approaches. The objective of surface based methods is to find the organ or tumor's boundary by propagation a curve/surface with a flow that is determined according to curvature and image constraints (generally the image gradient). Snakes and level sets are typically used in this context. The former defines the object's boundary explicitly as a parametric curve, while the latter defines the contour via an implicit function allowing for more complex geometries and topological changes. Region based methods consider the segmentation problem from a different angle. Here, the goal is to identify all voxels belonging to the object and separate them from the rest of the image. Early

work relied on fuzzy clustering to regroup similar voxels but quickly showed limitations. Supervised statistical pattern classification techniques have been the basis of the majority of recent region based tumor segmentation methods. The voxels are separated by a classification score or probability, the ultimate goal being to label each pixel of the image as tumor or background. The tumor is frequently separated into active tumor, necrosis and an additional class is introduced in the context of fast growing tumors for the induced edema around the tumor. The obtained segmentations are refined by the use of graph based methods that model spatial dependencies and prior anatomical knowledge.

A. BRAIN IMAGING

There exists a variety of different imaging modalities that enable the study of the brain. This section presents a brief overview of the different imaging methods, then focus on Magnetic Resonance Imaging (MRI) which is the most common modality for brain tumor observation.

1. IMAGING MODALITIES

Brain Imaging modalities can be grouped in two categories according to the information being captured by them: structural and functional imaging. The structural or anatomical modalities aim at visualizing the different structures and tissues of the brain. Among them, the most popular for Neuroimaging studies are the Computed Tomography (CT) and MRI. CT imaging relies on X-ray technology that is based on the absorption of X-rays beams as they pass through the different tissues of a patient's body. CT scans are constructed by using a series of X-ray beams that rotate around the patient's head. Each beam yields a 2D image at a specific angle that are used to construct a 3D volume by tomographic reconstruction. CT scans are well contrasted and high resolution images. The main drawback of this technique is the fact that it uses ionizing radiation. MRI relies on the magnetic property of the hydrogen nuclei present in large quantity in the human body. It is the modality of choice for brain studies due to its high tissue contrast and details and non ionizing property. In the recent years, new techniques have developed such as Diffusion Tensor Imaging (DTI). This modality enables the reconstruction of the white matter tracts connecting the different parts of the neural networks in the brain by measuring the anisotropic diffusion of water inside the tissues. The presence of a tumor could have a direct impact on the fibers structures by causing disruption, displacement or infiltration of the fibers [Wei 2007]. Modalities such as MR spectroscopy (measurement of major metabolites in tumor tissue) and Perfusion MRI (measurement of the relative cerebral blood volume using a contrast agent) provide additional informations for the

diagnosis and study of brain tumors [Soffietti 2010]. The goal of functional imaging is to study the human brain's function based on physiological changes caused by the brain's activity. Electroencephalography (EEG) and Magneto encephalography (MEG) are techniques that offer means of directly measuring the brain activity. EEG detects the electrical impulses in the brain due to the neuronal activity via electrodes placed on the scalp. MEG measures the magnetic flux changes using sensors positioned closed to the scalp. The methods are popular due to their simplicity, the fact that they are non invasive and their very high temporal resolution. Determining precisely the spatial origin of the observed signal is however difficult. Position Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT) are nuclear imaging techniques that measure changes in the cerebral blood flow and tissue metabolism and how it is altered by brain disorders. A biological molecule is marked with a radioactive isotope and injected in the bloodstream and accumulated in areas where the molecule has affinity. The advantage of this method is that the radioactive tracer can be designed specifically to target specific organs or processes related to an illness. For instance, a radioactive isotope typically used is the Fluorodeoxyglucose (FDG) that behaves like glucose molecules and therefore is able to trace the brain's metabolic activity. However, the technique is invasive and potentially harmful due to the use of radioactive tracers while at the same time producing the isotopes is costly and difficult. Nowadays, CT and PET technologies are being combined in one device, allowing to combine the anatomical information recovered from the CT scans with the metabolic information given by the PET scans. Functional MRI (fMRI) exploits the Blood Oxygen Level Dependent (BOLD) contrast [Ogawa 1990] to detect changes in the neuronal activity induced by sensorimotor or cognitive tasks. Neuronal activation causes an increase of the blood flow to compensate the oxygen consumption and therefore reduces the amount of deoxygenated hemoglobin molecules. The detection of changes in hemoglobin oxygenation relies on the paramagnetic properties of the deoxygenated hemoglobin which impacts the measured NMR signal. fMRI is often used for tumor surgery planning in order to identify the spatial relationship between the lesion and the functional area and evaluate the corresponding risks. Time constraints in the acquisition of the signal cause the fMRI images to be of lower quality than the structural MRI. Functional imaging is particularly important for brain tumor surgery planning. It enables to establish a relationship between the lesion and the functional area and plan the surgery accordingly (extent of resection, awake craniotomy...).

2. MAGNETIC RESONANCE

Imaging Magnetic Resonance Imaging is one of the most popular medical imaging modalities. It has the advantage of being a non invasive and non ionizing technique that produces images with excellent tissue contrast, making it the modality of choice for brain (and more specifically brain tumors) study. MRI is based on the principle of Nuclear Magnetic Resonance (NMR), that is commonly used in spectroscopy to study the physical and chemical properties of molecules. Consider a set of random variables $Y = y_1, y_2, \dots, y_n$, and a set of labels $X = x_1, x_2, \dots, x_k$. The objective of segmentation problem is to assign a label from the set X to each variable of set Y . Random fields provides a probabilistic framework for labeling problem, in which we can incorporate neighborhood information easily. Two popular type of random field models are markov random fields (MRF) and conditional random fields (CRF). MRF calculate the joint probability of $p(x,y)=p(y/x)p(x)$, where $P(y/x)$ and $P(x)$ are likelihood and prior probabilities. Likelihood can be factorized under markovian property of restrictive independence, such as $p(y/x)=\prod_i p(y_i/x_i)$. This assumption does not allow to model complex dependencies between data and labels, so we use another class of random field models called CRF. In CRF we directly model the posterior probability of labels given data $p(x/y)$. The posterior distribution $p(x/y)$ over the labeling in CRF given as : $p(x/y) = \frac{1}{Z} \exp(-E(x, y))$ (using Hammersley-Clifford theorem), where Z is a normalization constant and $E(x,y)$ is an energy function. The best labeling of a given random field is describe as maximum posterior probability.

$$X_{max} = \operatorname{argmin}(E(x, y))$$

In our problem random variables(y) correspond to intensity value at each pixel in an image, and the label set(x) correspond to $\{0, 1\}$. Label-0 corresponds to tumor class and label-1 corresponds to non tumor class. For modeling the energy function in CRF requires initial labeling. For initial labeling we use Fuzzy C-means algorithm (FCM). After that define the energy function using this initial labeling and optimize the energy function using graph cut approach.

I. Initial Segmentation

First remove the background of FLAIR image using the T1-weighted map(in ADC map background have zero value). For initial segmentation(labeling) used FCM algorithm on intensity features. FCM segments image into pre-specified number of clusters (K). FCM gives a fuzzy membership(U) to describe the degree of similarity of one pixel to each cluster. Fuzzy membership U is a $C \times N$ matrix, where C is the number of clusters and N is number of pixels. Applied threshold on membership function to classify the result in pre-define classes. Segmenting

the FLAIR image into tumor and non-tumor class applied FCM with K=2. Applied threshold on data matrix

results probable tumor regions (with false positive) and nontumor regions. Compare the tumor region with T1-weighted and remove the false regions. If slice does not contain any infarct region then we stop the algorithm and declare slice is normal, otherwise slice goto next step for fine segmentation. into infarct and non infarct classes

II. Energy Function

Based on initial labeling on FLAIR and T1-weighted(have same lables as FLAIR) define an energy function. Energy function $E(x,y)$ can be expressed as a sum of clique potential $E(x, y) = \sum_{c \in C} \phi_c$, where C is the set of cliques formed by the neighborhood system Ω and ϕ_c is the potential function of the clique. This potential function is describe as local interaction in neighborhood system.

$$E(x, y) = \sum_i \phi_i(x_i) + \sum_{i,j \in \Omega} \phi_{i,j}(x_i, x_j, y)$$

where $\phi_i(x_i)$ and $\phi_{i,j}(x_i, x_j, y)$ are called as unary potential and pairwise potential respectively, and Ω is neighborhood of pixel i . Unary potential, model as association of each pixel i with class x . We use the sigmoid function to model the unary potential, such as

$$\phi_i(x_i \in \{0,1\}) = \frac{1}{1 + \exp(-(\ln P(y/x_i) + w_0))}$$

where $p(y/x)$ and w_0 are likelihood probabilities in FLAIR image and scaling parameter. For calculate the likelihood probabilities $p(y/x)$ of tumor class and non-tumor class separately fit the distribution on Initial labeling. w_0 is data dependent term which depends on variance of the class. pairwise potential model the interaction of each pixel i with its neighborhood Ω (use 8 connected neighborhood). CRF uses observation and their labels to model pairwise potential. We model the pairwise potential as a inverse of similarity between the neighboring pixel, when the pixels are most similar potential value is highest.

$$\phi_{i,j}(x_i, x_j, y) \propto \frac{1}{\text{dist}(x_i, x_j, y)}$$

where dist is the euclidean distance. $\text{dist}(x_i, x_j, y)$ is define as

$$\text{dist}(x_i, x_j, y) = \text{dist}(P_{FLAIR}, P_{T1})$$

where PFLAIR is define in FLAIR image as $P_{FLAIR} = (p(y/x_i = 1), p(y/x_j = 1))$ and similarly PT1 in T1 for tumor class. If neighborhood of pixel i have same labels in FLAIR and T1-weighted both then this distance is less, otherwise viceversa. For the non-stroke class we used the same pairwise potential function, where

PFLAIR is define in the FLAIR image as $P_{FLAIR} = (p(y/x_i = 0), p(y/x_j = 0))$ and PT1 is defined in the corresponding T1-weighted. Now pairwise potential is define using sigmoid function

$$\phi_{i,j}(x_i, x_j, y) = \frac{1}{1 + \exp(-(\ln(\text{dist}(P_{FLAIR}, P_{T1}))) + w_0)}$$

Construct the graph based on this above defined energy function. The nodes in graph corresponds to the pixels in the image, and edge weights of the graph are define according to the unary and pairwise potential functions. For finding the global optimal solution use Graph-cut approach. Graph cut minimize energy function using min-cut/max-flow algorithm in polynomial time. For finding the minimum value of energy function we cut a graph in two parts in such a way that cost of cut(sum of edge weights) is minimum.

III. PROPOSED SYSTEM

A. Materials

1) Sources and Description of Test Data: The test data were retrospectively acquired from different models of General Electric (GE) and Siemens 1.5 and 3T scanners that use different coils, and were obtained from four different sources. The sources of data are NeuroRx research, the Alzheimer's disease neuroimaging initiative (ADNI) and the Department of Diagnostic Imaging of the Hospital for Sick Children in Toronto, Canada. There are thirty nine brain MRI volume data. They consist of fifteen T2 weighted, ten T1 Magnetization-Prepared Rapid Gradient Echo (MPRAGE) pulse sequence and fourteen conventional T1 weighted images. All the T2 volume data were without perceived degradation. There are five, seven and three T2 volume data from NeuroRx, ADNI and BrainCare, respectively. Each T2 volume data from NeuroRx and ADNI contain 60 slices. Each slice has dimension 256x256 and 2.4 mm thickness. There are 24 slices in the T2 volume data from BrainCare, each with dimension 448x390 voxels and 2.6 mm thickness. The MPRAGE pulse sequence images from ADNI were without perceived degradation. Each data has 150 slice with dimension 190x160 voxels and 1.2 mm thickness. All the conventional T1 MRI volume data from NeuroRx were originally acquired with various configurations of bias fields.

2) Cardiac MRI Data: There are 16 cardiac MRI volume data from the Department of Diagnostic Imaging of the Hospital for Sick Children in Toronto, Canada. The data were acquired as short axis MRI data. The images were acquired using the Fast Imaging Employing Steady State Acquisition (FIESTA) sequence protocol. The images reveal the endocardial and epicardial structures of the ventricle. The data were among the experimental data in the report which describe the framework for the analysis of short axis cardiac MRI using

statistical models of shape and appearance. Each volume data contain 20 frames. The number of slices in each frame varies from 8 to 15. The dimension of each slice is 256x256 along the long axis.

3) Artificial Degradation: Three different types of degradation; circular blur, motion blur and Rician noise at different levels were artificially induced on the foreground and background voxels of the test data. Circular blur was simulated by convolving a slice in a volume data with circular averaging filter of radius r , $f_r : 0 < r < 7g$ voxels. The range of the radius of the circular averaging filter was scaled from level 1 to level 15 in unit step. The motion blur was induced on a slice by convolving it with a special filter which approximates the linear motion of a camera. The linear motion is described by two parameters, the linear distance in voxels and the angular distance in degree. Both parameters were scaled from 1 to 15 in unit step. Two separate and identical Gaussian noise levels were generated to simulate the real and imaginary components in the complex plane of MRI acquisition process. Rician noise was added to the data by computing the magnitude of the complex data. The noise level was scaled from 1 to 15 in unit step.

B. Problem Formulation

1) Structural and Acquisition Models of MRI: Following on the contribution in we model an ideal MRI slice as statistically simple and structurally piecewise constant. A slice is regarded as a two-tissue class binary image. The two tissue class MRI slice follows the same reasoning that regard the observed grayscale image as a blurred version of an underlying binary image. With reference to a T2 weighted MRI slice the bright voxels describe the high density of edges that describe the cortical gray matter, ventricular system and the boundaries between the different anatomical structures. The white matter and other anatomical structures are described by the dark voxels. MRI image acquisition follows the mathematical model of a 2D linear shift-invariant imaging system, expressed by:

$$I_d = \mathcal{H}I_f + n$$

where I_d is the observed grayscale image, I_f , the underlying ideal image, \mathcal{H} can be either space-invariant point spread function or multiplicative spatially varying factor and n is random noise. The random noise is independent of the image spatial coordinates and modeled as a Gaussian distribution with mean and variance. An MRI slice is formed on a rectangular lattice. The lattice consist of sites S corresponding to the location of image voxels in Euclidean space:

$$S = \{1, \dots, m\}$$

where $f_l; \dots; m_g$ are the indices of the sites. Label L is the set of pixel intensity levels that can be assigned to a site:

$$\mathcal{L} = \{1, \dots, M\}$$

where $f_l; \dots; M_g$ are the indices of the labels. Image labeling problem is the assignment of a label from the set L to each of the site in S .

2) Ideal MRI Acquisition: In the absence of any degradation, there is no random noise, the multiplying spatial varying factor and the space-invariant point spread function are identity matrix:

$$n = 0$$

$$\mathcal{H} = I$$

Under this condition the observed MRI slice possess its full natural properties and is considered the exact replica of the underlying ideal image:

$$I_d = I_f$$

Let I_c denote the local contrast feature LCF image derived through the use of appropriate filter to extract local information from the observed image. Under an ideal acquisition condition, based on the two-tissue class model, the LCF image is a replica of the observed image as well as a replica of the underlying ideal image;

$$I_c = I_d$$

$$I_c = I_f$$

Therefore, the first and second moments of the LCF image and the observed image are equal

$$\mu_c = \mu_d$$

$$\sigma_c^2 = \sigma_d^2$$

3) Real MRI Acquisition: The mathematical model of image acquisition expressed in Eq. 1 indicates that all the different types of degradation which are present during real MRI acquisition process are derived from three sources; random noise, the multiplying spatial varying factor and the space invariant point spread function. In this report we generalize invariant features proposed for blur degradation by to include all the different types of degradation. Specifically, in the presence of any degradation, the first and second moments of the ideal image are preserved in the observed image:

$$\mu_f = \mu_d$$

$$\sigma_f^2 = \sigma_d^2$$

The severity of any type of degradation is denoted by integer numbers $f_l : 0; 1; 2; \dots; L_g$ where $l = 0$ implies absence of degradation, that is, image acquisition under ideal condition. At each level of degradation, a unique set of label L_l referred to as image pixel configuration is assigned to each site on the grid of the observed image I_d . With

reference to 8-bit grayscale image, each image pixel configuration is a sample of size 256, obtained at random, with replacement, from the population of 256 formed by 8-bit grayscale voxels. The total number of possible degradation levels is all the possible random samples $L = 256256$. The mean of the LCF image extracted from the observed image I_d at each level l of degradation is a random variable X . According to the central limit theorem, if the number of possible degradation levels L tends to infinity and is finite, the distribution of X approaches a normal distribution with mean X and variance $2X$::

$$\lim_{L \rightarrow \infty} P(X) = \frac{1}{\sigma_X \sqrt{2\pi}} \exp - \left(\frac{(x - \mu_X)^2}{2\sigma_X^2} \right)$$

where f is the variance of the underlying ideal image I_f . Using the same hypothesis, the variance $2c_l$ of the LCF image extracted from the observed image I_d at each level l of degradation is also a random variable Y . The central limit theorem says that the distribution of Y approaches a normal distribution with mean Y and variance $2Y$::

$$\lim_{L \rightarrow \infty} P(Y) = \frac{1}{\sigma_Y \sqrt{2\pi}} \exp - \left(\frac{(y - \mu_Y)^2}{2\sigma_Y^2} \right)$$

where f is the variance of the underlying ideal image I_f .

4) Quality Prediction: Quality prediction is based on making analogy between the Gaussian distributions expressed in Eq. 12 and Eq. 13 and the power spectral density of an image. The pixel configurations assigned to an observed image at each possible level of degradation is the equivalent of all the possible frequencies contained in the image. The power of the observed image at a specific frequency is the variance of the LCF image extracted from the observed image. The maximum possible total power in the spectrum is the area under the curve that describe each probability distribution. The total power corresponds to the maximum possible image contrast. Two quality scores, lightness contrast quality score and the texture contrast quality score of the test image It can be predicted from the appropriate Gaussian distribution. Given the combined effect of all the possible distortions, the lightness contrast quality score q_1 is the magnitude of perceived visual differences of local structures within the image. It is expressed by the normal cumulative distribution function of Y

$$q_1 = P(z_c \leq y_t)$$

where y_t the standard deviation of the LCF image extracted from the test image It standardize the normal distribution Y to obtain z_c , the z-score:

$$z_c = \frac{\|y_t - \mu_Y\|}{\sigma_Y}$$

C. Objective Evaluation

In this section we use the flow chart of Fig. 1 and the images displayed in Fig. 2 to describe the eight steps to implement our proposed no-reference quality evaluation for MRI images. The MRI slice is from a MRI volume data provided by BrainCare.

1) Step 1 - Intensity Rescaling: The intensity level of the test image TIM shown in Fig. 2a is rescaled RES to lie between 0 and 1 so that the rescaled test image RES in Fig.2b can be regarded as a blurred version of a binary image.

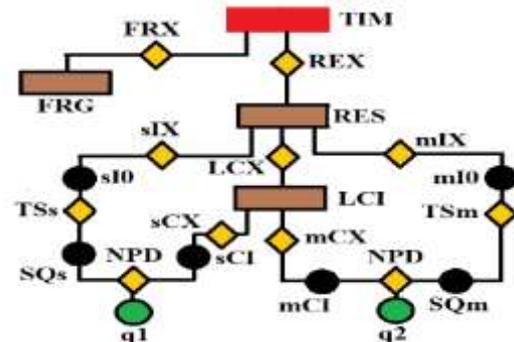


Fig. 1. The flow chart of our proposed no-reference quality evaluation for MRI images. The first step rescale RES the intensity level of the test image TIM to lie between 0 and 1, followed by the extraction FRG of foreground FRG . The third step computes (mIX , sIX) the first $mI0$ and second $sI0$ moments of the test image. The local contrast feature image LCI is extracted LCX in the fourth step. The fifth step computes (mCX , sCX) the first mCX and second sCX moments of the local contrast feature image. The lightness contrast quality score q_1 and the texture contrast quality score q_2 are computed in the sixth and seventh steps from the cumulative normal distribution function NPD of the random variables evaluated at ($X = sCX$, $Y = mCX$). In the last step, the total quality score is computed from the weighted sum of the lightness and contrast quality scores.

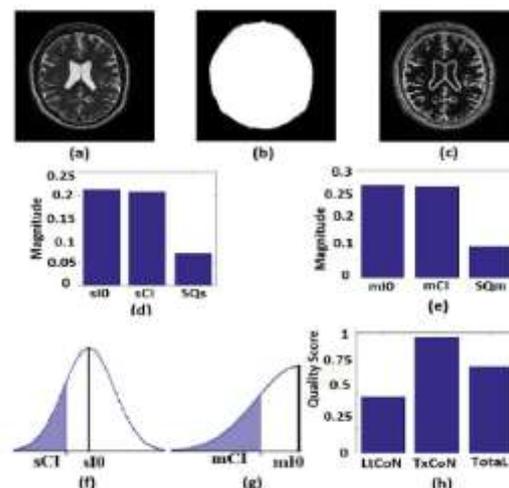


Fig. 2. Description of the proposed no-reference quality evaluation for MRI images. (a) The test image has its pixel intensity level rescaled to lie between 0 and 1. (b) Foreground of the test image in (a) is extracted. (c) Local contrast feature image is extracted from the test image. (d) The second moments ($sI0$, sCI) is computed from the test image and the local contrast feature image. The variance SQs of the normal distribution is also computed. (e) The first moments ($mI0$, mCI) computed from the test image and the local contrast feature image as well as the variance SQm of the normal distribution. (f) Two-tail cumulative distribution function for the computation of lightness contrast quality score. (g) One-tail cumulative distribution function for the computation of texture contrast quality score. (h) Bar chart of the lightness contrast quality score, texture contrast quality score and the total quality score.

2) Step 2 - Foreground Extraction: Foreground extraction FRX extracts the foreground voxels FRG shown in Fig. 2c. Foreground extraction excludes the background voxels so that quality evaluation is computed from only the foreground voxels which contains the anatomical structures in the test image.

3) Step 3 - Compute Image Moments of the Test Image: Two actions, mIX and sIX refer to the foreground voxels in step 2 to extract the first moment $mI0$ and the second moment $sI0$ of the test image

4) Step 4 - Contrast Feature Image Extraction: Local contrast feature image LCI shown in Fig. 2d is extracted LCX from the test image by convolving the test image with a local range filter of appropriate size. We hereby emphasize the need for the use of local range filter of appropriate size because the algorithm is sensitive to the size of filter. Larger filter size causes loss of fine details while smaller filter size will result in loss of spatial coherence in the filtered image. A 7×7 filter is recommended for images with either row r or column c dimensions $f200$ (r ; c) $512g$. Standardization of image quality across different clinical trial sites is attained through the combination of intensity rescaling in step 1, foreground extraction in step 2 and the use of fixed size filter for feature extraction in step 3.

5) Step 5 - Compute Image Moments of the Local Contrast Feature Image: The first moment mCI and the second moment sCI of the local contrast feature image are computed mCX , sCX with reference to the foreground voxels.

6) Step 6 - Lightness Contrast Quality Score: The lightness contrast quality score $q1$ is the cumulative normal distribution function NPD of Y evaluated at $Y = sCI$ according to Eq. 14 and Eq. 15. The variance SQs of the normal distribution shown in

Fig. 2g is computed by the three sigma rule TSs according to Eq. 18 using inputs from the second moment sI (see Fig. 2e) of the test image and the second moment sCX (see Fig. 2e) of the local contrast feature image.

7) Step 7 - Texture Contrast Quality Score: The texture contrast quality score $q2$ is the cumulative normal distribution function NPD of X evaluated at $X = mCI$ according to Eq. 16 and Eq. 17. The first moment $mI0$ (see Fig. 2f) of the test image and the first moment mCX (see Fig. 2f) of the local contrast feature image are the inputs for the computation of the variance SQm of the normal distribution shown in Fig. 2h. The variance is computed using the three sigma rule TSm according to Eq. 18.

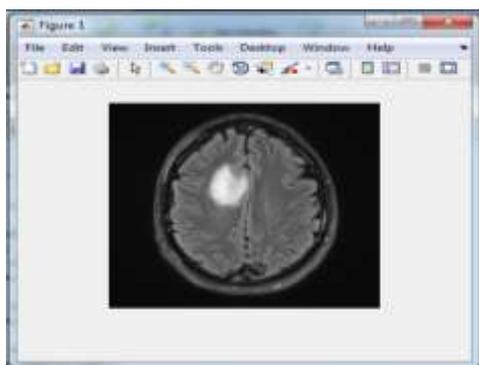
8) Step 8 - Total Quality Score: The total quality score is computed according to Eq. 19.

D. Subjective Evaluation

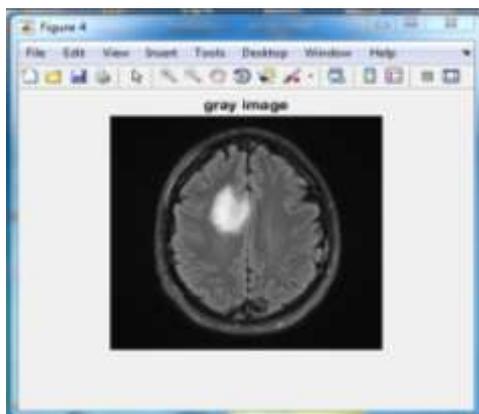
The objective experiment was validated using QuickEval, a web-based tool for psychometric image evaluation provided by the Norwegian Colour and Visual Computing Laboratory (www.colourlab.no/quickeval) at the Norwegian University of Science and Technology, Gjøvik, Norway. The observers are one radiologist and one MRI reader. MRI reader is a trained professional with experience working on MRI images that are affected by pathology.

There are ten categories of the subjective experiment. The ten categories can be split into two major categories; MRI volume data without perceived degradation and MRI volume data degraded by different types of degradation. The category of MRI volume data without perceived degradation can be further classified into three categories. They are cardiac MRI without perceived degradation, T2 brain MRI without perceived degradation and T1 MPRAGE brain MRI without perceived degradation. There are seven categories under the main category of degraded MRI volume data. Each different levels of degradation by Rician noise, circular blur and motion blur has two categories from brain and cardiac MRI volume data to form a total of six categories. The seventh category is T1 MRI volume data originally acquired with bias fields. Three hundred and sixty slices from different MRI volume data are utilized for each category of the experiment. The observer assigns a score between 0 and 100, in unit steps, to each slice. Each score assigned by the observer is divided by 100 to ensure that the subjective and objective scales are in the same range. In the category of MRI volume data with artificially induced degradation, each observer was first presented with an undistorted version of an MRI slice, followed by increasing degradation levels of the original slice. The distorted levels are 5, 10 and 15.

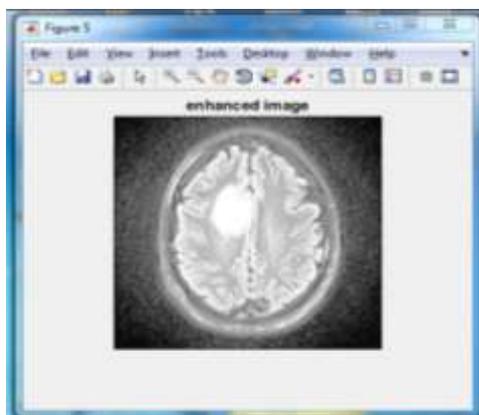
IV. RESULTS



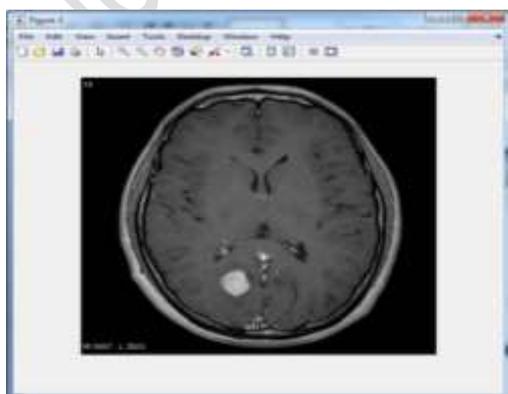
Input image 1



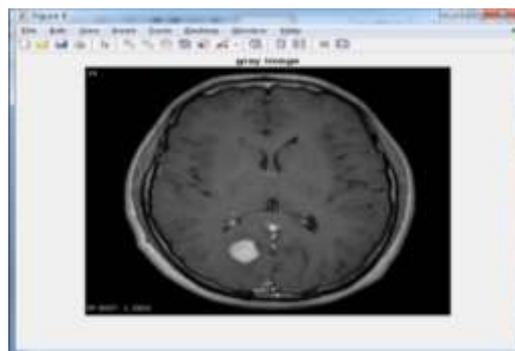
Grey scale of Image 1



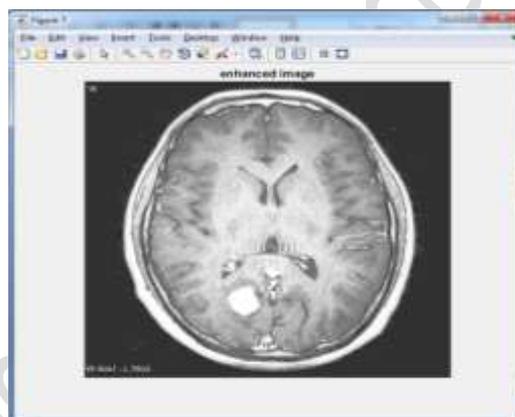
Enhanced image of image 1



Input image 2



Grey scale of Image 2



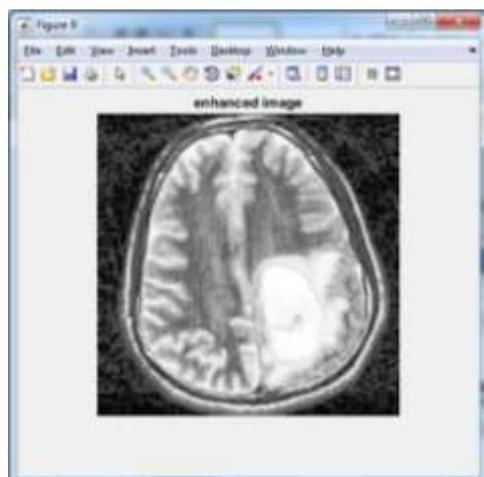
Enhanced image of image 2



Input image 3



Grey scale of Image 3



Enhanced image of image 3

V. CONCLUSION

There is increasing clinical interest in the use of MRI images for the study of human anatomy, treatment and diagnosis of diseases. Currently MRI images are being considered the primary endpoints in large clinical trials of drugs for the treatment of neurological and cardiovascular diseases. In large clinical trials large volumes of MRI data are processed. Thus no-reference objective quality assessment is highly desired. The reliability of metric derived from quantitative analysis of MRI images is strongly dependent on rigorous monitoring throughout the various stages of the imaging workflow. We hereby propose a new method to evaluate the quality of brain MRI images from acquisition through processing to the analysis stages of the imaging workflow. Our proposed quality evaluation method re-evaluates and standardizes the quality of MRI images acquired from different clinical trial sites across the globe and through all the stages of the imaging workflow. Experimental results demonstrate that our proposed method had good correlation with human visual judgement and gives fairly accurate quality evaluation within and across good quality images and different levels of degradation.

REFERENCES

- [1] G. Liney, MRI in clinical practice. Springer Science & Business Media, 2007.
- [2] T. E. Yankeelov and J. C. Gore, "Dynamic contrast enhanced magnetic resonance imaging in oncology: theory, data acquisition, analysis, and examples," *Current medical imaging reviews*, vol. 3, no. 2, pp. 91–107, 2007.
- [3] F. Fazekas, F. Barkhof, M. Filippi, R. Grossman, D. Li, W. McDonald, H. McFarland, D. Paty, J. Simon, J. Wolinsky et al., "The contribution of magnetic resonance imaging to the

diagnosis of multiple sclerosis," *Neurology*, vol. 53, no. 3, pp. 448–448, 1999.

[4] S. J. Teipel, M. Wegrzyn, T. Meindl, G. Frisoni, A. L. Bokde, A. Fellgiebel, M. Filippi, H. Hampel, S. Klöppel, K. Hauenstein et al., "Anatomical MRI and DTI in the diagnosis of Alzheimer's disease: a European multicenter study," *Journal of Alzheimer's Disease*, vol. 31, no. s3, pp. S33–S47, 2012.

[5] S. J. Khoury, J. Rochon, L. Ding, M. Byron, K. Ryker, P. Tosta, W. Gao, M. S. Freedman, D. L. Arnold, P. H. Sayre et al., "Acclaim: A randomized trial of abatacept (CTLA4-Ig) for relapsing-remitting multiple sclerosis," *Multiple Sclerosis Journal*, p. 1352458516662727, 2016.

[6] R. Gold, D. L. Arnold, A. Bar-Or, M. Hutchinson, L. Kappos, E. Havrdova, D. G. MacManus, T. A. Yousry, C. Pozzilli, K. Selmaï et al., "Long-term effects of delayed-release dimethyl fumarate in multiple sclerosis: Interim analysis of endorse, a randomized extension study," *Multiple Sclerosis Journal*, p. 1352458516649037, 2016.

[7] J. Maranzano, D. Rudko, D. Arnold, and S. Narayanan, "Manual segmentation of MS cortical lesions using MRI: A comparison of 3 MRI reading protocols," *American Journal of Neuroradiology*, vol. 37, no. 9, pp. 1623–1628, 2016.

[8] H. Lu, L. M. Nagae-Poetscher, X. Golay, D. Lin, M. Pomper, and P. van Zijl, "Routine clinical brain MRI sequences for use at 3.0 Tesla," *Journal of Magnetic Resonance Imaging*, vol. 22, no. 1, pp. 13–22, 2005.

[9] D. L. Arnold, E. Fisher, V. V. Brinar, J. A. Cohen, A. J. Coles, G. Giovannoni, H.-P. Hartung, E. Havrdova, K. W. Selmaï, M. Stojanovic et al., "Superior MRI outcomes with alemtuzumab compared with subcutaneous interferon β -1a in MS," *Neurology*, vol. 87, no. 14, pp. 1464–1472, 2016.

[10] L. Kappos, H. Wiendl, K. Selmaï, D. L. Arnold, E. Havrdova, A. Boyko, M. Kaufman, J. Rose, S. Greenberg, M. Sweetser et al., "Daclizumab hyp versus interferon β -1a in relapsing multiple sclerosis," *New England Journal of Medicine*, vol. 373, no. 15, pp. 1418–1428, 2015.