

No-Reference Quality Measure in Brain MRI Images using Binary Operations, Texture and Set Analysis

Michael Osadebey^{a,*}, Marius Pedersen^b, Douglas Arnold^c, Katrina Wendel-Mitoraj^d

^a MRI Reader Group, NeuroRx Research Inc, 3575 Parc Avenue, Suite # 5322, Montreal, QC, H2X 3P9, Canada
e-mail: mosadebey@neurorx.com

^bDepartment of Computer Science, Norwegian University of Science and Technology, Teknologivegen 22, N-2815 Gjøvik,
Norway email: marius.pedersen@ntnu.no.

^c NeuroRx Research Inc, 3575 Parc Avenue, Suite # 5322, Montreal, QC, H2X 3P9, Canada

^dBrainCare Oy, Finn-Medi 1, PL 2000, 33521 Tampere, Finland

Abstract

We propose a new application-specific post-acquisition quality evaluation method for brain MRI images. The domain of a MRI slice is regarded as the universal set. Four feature images; grayscale, local entropy, local contrast and local standard deviation are extracted from the slice and transformed into the binary domain. Each feature image is regarded as a set enclosed by the universal set. Four qualities attribute; lightness, contrast, sharpness and texture details are described by four different combinations of the feature sets. In an ideal MRI slice the four feature sets are identically equal. The degree of distortion in real MRI slice is quantified by the fidelity between the sets that describe a quality attribute. Noise is the fifth quality attribute and it is described by the slice Euler number region property. The total quality score is the weighted sum of the five quality scores. Our proposed method addresses the current challenges in image quality evaluation. It is simple, easy-to-use and easy-to-understand. Incorporation of binary transformation in the proposed method reduces computational as well as operational complexity of the algorithm. We provide experimental results that demonstrate the efficacy of our proposed method on good quality images and on common distortions in MRI images of the brain.

Keywords: Magnetic Resonance Imaging, Image Quality, Texture, Contrast, Standard Deviation, Entropy and Euler Number.

1. Introduction

Magnetic resonance imaging (MRI) system is a popular tool in the field of medicine for the study of human anatomy [42], [13], segmentation to aid pathology [7], [28] and the clinical trials of drugs for the treatment and monitoring of neurological diseases [3], [2]. Quality evaluation is required at the acquisition and post-acquisition stages of an imaging workflow. Quality evaluation assess the integrity of information contained in images to ensure that they are of acceptable quality before post-acquisition processing and analysis. Image quality evaluation is one of the basic criteria for the performance evaluation of MRI system devices at the acquisition stage and the performance evaluation of automated image analysis systems at the post-acquisition stage. An example of post-acquisition utility of brain MRI images is in clinical research organizations (CRO) that manage the clinical trials of drugs for the treatment and monitoring of multiple sclerosis and Alzheimer's diseases. Daily clinical research organizations receive large volumes of MRI data acquired from several clinical trial sites around the globe.

*Corresponding author at: MRI Reader Group, NeuroRx Research Inc, 3575 Parc Avenue, Suite # 5322, Montreal, QC, H2X 3P9, Canada: E-mail address: mosadebey@neurorx.com

Brain measurements derived from MRI images are susceptible to differences in MRI system sequence parameters [4]. Thus quality evaluation is important to evaluate the variations in the quality attributes of MRI images from the different MRI system manufacturers and the different clinical trial sites to ensure that they conform with the acquisition protocols set by the sponsoring pharmaceutical organizations.

The human visual system is the gold standard in image quality evaluation. However subjective evaluation by human observers will be grossly inefficient in real-world scenarios such as the CRO where large volumes of brain MRI images are evaluated before processing by automated image analysis systems. Routine quality evaluation by a trained MRI reader involves negotiation around several different types and levels of distortions [32]. Some quality evaluation task require the consent of other trained readers because the human eyes does not possess clearly defined quality index threshold for distinguishing a good quality image from a poor quality image. Often it is difficult for trained MRI readers to arrive at a consensus for images that are perceived to lie on the borderline between acceptable and unacceptable quality. The consequence is high intra-reader and inter-reader variability, and the task of quality evaluation becomes cumbersome.

In recent years objective image quality evaluation have been an active research area due to technological advancements which encourage the use, storage and transmission of images. Objective quality assessment methods are categorized according to the availability of a reference image. Full reference image quality assessment predicts image quality by comparing the image under investigation to another image whose image quality has already been determined [38], [48], [52], [22]. Reduced reference image quality assessment uses partial information from the reference image [55], [29], [43], [47]. No-reference method predict the quality of an image without a reference image and use only the information about the image under investigation [33], [12], [45], [8], [44], [1]. It is noteworthy that most contributions in the literature such as [52], [58], [24], [35] are driven by the need to solve computer vision problems such as storage, compression and transmission. The popular image quality evaluation techniques are peak signal-to-noise ratio (PSNR), signal-to-noise ratio (SNR) and structural similarity index [52]. These techniques are strongly related as they are both generic and full-reference methods [21].

Most brain MRI quality assessment method focus on the acquisition stage of the imaging workflow. Quality evaluation methods at the acquisition stage involves modifying the design of system hardware, optimizing acquisition parameters and the implementation of undersampling and constrained reconstruction schemes to reduce acquisition time and generate high quality images. Proposed methods in this category include [20], [6], [37], [32], [53], [18]. There are very few contributions on post-acquisition quality evaluation for brain MRI images. One of the few contributions in the literature [15] proposed signal-to-noise ratio as a quality metric within an automated quality control system. The authors in [34] adopt artifactual voxels and noise level as the two quality attributes to measure image quality. The report in [54] apply analysis of variance (ANOVA) algorithm to assess the variation of several quality measures with different levels of distortions.

The contribution by [15] which adopt the popular signal-to-noise ratio (SNR) is diagnostically misleading [32] because it cannot discriminate the quality of two images that are perceptually dissimilar [41]. There are many definitions of SNR which makes it difficult to compare quality measures from different imaging system, modalities and researchers [25]. Quality indices derived from these popular methods does not always correlate with the performance of observers using the imaging system on the task for which they are intended [25]. The adoption of only artifacts and noise in [34] are too few attributes to account for the several types of distortions which combine with ideal features to manifest as image quality [33]. The report in [54] use the analysis of variance (ANOVA) to demonstrate the variations of several quality measures with different types and levels of distortions. However the proposed technique cannot transform the different levels of distortion into a quality index. ANOVA based measures does not discriminate quality measures which successive values are decreasing, increasing or random. Thus, there is the risk of ambiguity in quality measures that are based on the use of ANOVA models [31].

We acknowledge that, as at when they were proposed, current post-acquisition quality evaluation

techniques demonstrate state-of-the-art performance in the industry. Rapid technological advancements in the design of MRI systems coupled with increasing clinical interest in brain MRI images demands better technique which overcomes their limitations, meet real-world challenges and keep pace with advances in technology. The challenges include design of objective algorithms with reduced complexity and to make them easy-to-use and easy-to-understand [5], [51]. In real-world scenarios it is not always possible to have prior knowledge of all the possible distortions that can degrade an image [57]. Therefore there is need for new quality models that extends beyond the single most relevant quality attribute in an image [26]. A robust objective quality assessment must be based on efficiently managed quality attributes that accounts for all possible distortion in an image [39].

We propose a new application-specific no-reference image quality assessment for MRI images of the brain. The proposed method addresses the challenges outlined in [51] to reduce the complexity of image quality models and to make them easy-to-use and easy-to-understand. It is based on the piecewise constant model of ideal MRI image, binary operations, local texture analysis and set theory. The system operates on individual slices in a MRI volume. Four feature images are derived from each slice. They are the grayscale image of the test image, local entropy image, local contrast image and local standard deviation image. The four feature images are binarized using their first moments as global thresholds. In the design we show that, for an ideal MRI slice, the four binary images are identically equal. We apply binary morphological operation and set theory analysis on different combinations of the feature images to define five quality factors; noise, lightness, contrast, sharpness and texture details. Total quality factor is the weighted sum of the five quality factors. The quality score of an ideal MRI slice for each quality factor is unity. The quality score of a real MRI slice is its fidelity to an ideal MRI slice. Performance evaluation results showed that our proposed method is simple, efficient in automated environment and correlates with different levels of distortion.

This report is organized as follows. The next section explains how the problem of quality evaluation is formulated. The setup of the experiment to evaluate the quality of MRI data is in section 3. Section 4 describe the operation of the algorithm. In section 5 the results of the experiment are displayed. The experimental results are discussed in section 6. Section 7 gives an insight into future research work. Section 8 concludes this report.

2. Problem Formulation

This section begins with description of a model MRI slice. Quality attributes of the model are derived. The problem of image quality evaluation is quantified by the fidelity between a specific quality attribute of an ideal MRI image and corresponding quality attribute of a real MRI image.

2.1. Attributes of an Ideal MRI image

An ideal MRI slice $I_c(x, y) \in \mathbf{R}^2$ where (x, y) are pixel locations is modeled as a piecewise constant image [59]. Each piecewise constant region belongs to one of two intensity classes v_r where $r = \{1, 2\}$. The spacing between adjacent pixels is ι and the constant regions have clearly defined boundary $b(x, y)$. The ideal image is the union of the two tissues classes and the boundary pixels:

$$I_c(x, y) = I_c(x, y)_{v_1} \cup I_c(x, y)_{v_2} \cup b(x, y) \quad (1)$$

Pixels $I_c(x, y)_{v_r}$ in each tissue class are assigned one of the two possible pixel intensity values, $k = \{0, 1\}$ in a binary image. A local region within the image is denoted $i_c(x, y)_{v_r}$ and has constant dimension $m \times m$ with constant number n_{v_r} of pixels:

$$n_{v_r} = m^2 \quad (2)$$

2.1.1. Local Regions Within Constant Regions

In local regions $i_c(x, y)_{v_r}$ that lie entirely within constant regions:

The pixels intensity level k_{v_r} and the mean of pixel intensity levels:

$$\begin{aligned} i_c(x, y)_{v_r} &= k_{v_r} \\ \mu_{i_c(x, y)_{v_r}} &= k_{v_r} \end{aligned} \quad (3)$$

are constants. The variance $\sigma_{i_c(x, y)_{v_r}}$ of pixel intensity levels:

$$\sigma_{i_c(x, y)_{v_r}} = 0 \quad (4)$$

is zero. The gradient $\nabla_{i_{v_r x}}, \nabla_{i_{v_r y}}$ in the x and y directions:

$$\begin{aligned} \nabla_{i_{v_r x}} &= \|i_c(x + \iota, y)_{v_r} - i_c(x, y)_{v_r}\| = 0 \\ \nabla_{i_{v_r y}} &= \|i_c(x, y + \iota)_{v_r} - i_c(x, y)_{v_r}\| = 0 \end{aligned} \quad (5)$$

The probability $P(k_{v_r})$ of the pixel intensity level:

$$P(k_{v_r}) = \left(\frac{n_{k_{v_r}}}{m^2} \right) = 1 \quad (6)$$

Therefore,

The Shannon entropy s_{v_r} [17]:

$$s_{v_r} = - \sum P(k_{v_r}) \log P(k_{v_r}) = 0 \quad (7)$$

The contrast c_{v_r} [39]:

$$c_{v_r} = \frac{i_{v_r max} - i_{v_r min}}{i_{v_r max} + i_{v_r min}} = 0 \quad (8)$$

The sharpness d_{v_r} [19]:

$$d_{v_r} = \frac{1}{m^2} \sum \left| \frac{\nabla_{i_{v_r}}}{i_{v_r}} \right|^2 i_{v_r} = 0 \quad (9)$$

2.1.2. Local Regions Containing Boundary pixels

At local regions that includes the boundaries between different piecewise constant regions:

There is diversity of pixel intensity levels,

$$i_c(x, y)_b = 0 \parallel 1 \quad (10)$$

The variance $\sigma_{i_c(x, y)_b}$ and the magnitude of the gradient are positive

$$\begin{aligned} \sigma_{i_c(x, y)_b} &> 0 \\ |\nabla_{i_{bx}}| &> 0 \\ |\nabla_{i_{by}}| &> 0 \end{aligned} \quad (11)$$

The mean of pixel intensity levels is a new constant k_b determined by the number n of pixels belonging to the different binary classes, 0 or 1;

$$k_b = \frac{1}{m^2} \sum (n_0 + n_1) \quad (12)$$

The probability $P(i(x, y))$ of a specific pixel intensity level lies between 0 and 1,

$$\begin{aligned} 0 &\leq P(i(x, y))_{i(x,y)=0} \leq 1 \\ 0 &\leq P(i(x, y))_{i(x,y)=1} \leq 1 \end{aligned} \quad (13)$$

Therefore,

Entropy, contrast and sharpness are positive semi-definite

$$\begin{aligned} s_{vb} &\geq 0 \\ c_{vb} &\geq 0 \\ d_{vb} &\geq 0 \end{aligned} \quad (14)$$

2.1.3. Quality Attributes

We define the universal set Ω as the domain \mathbf{R}^2 of an ideal MRI slice. Given that four binary feature images; grayscale I_c , local entropy S_c , local contrast C_c and local standard deviation D_c are extracted from the original image.

The followings hold:

Each feature image is a subset of the universal set:

$$I_c \in \Omega, \quad S_c \in \Omega, \quad C_c \in \Omega, \quad D_c \in \Omega \quad (15)$$

The feature images are identically equal;

$$I_c = S_c = C_c = D_c \quad (16)$$

2.2. Attributes of Real MRI image

In real MRI image $I_t(x, y)$ the piecewise constant regions v_r of ideal MRI image are replaced by homogeneous regions h . There is intra-class pixel intensity variations within each homogeneous region. Local regions that includes the boundaries between the homogeneous regions are denoted e .

2.2.1. Local Regions Within Homogeneous Regions

In local regions within the homogeneous regions the entropy s_{h_r} , contrast c_{h_r} and sharpness d_{h_r} is not necessarily equal to zero value of the corresponding ideal image:

$$s_{h_r}(x, y) \geq s_{v_r}(x, y) \quad (17)$$

$$c_{h_r}(x, y) \geq c_{v_r}(x, y) \quad (18)$$

$$d_{h_r}(x, y) \geq d_{v_r}(x, y) \quad (19)$$

2.2.2. Local Regions Containing Boundary pixels

In the local regions containing boundaries between the two homogeneous regions the entropy $s_t(x, y)_e$, contrast $c_t(x, y)_e$ and sharpness $d_t(x, y)_e$ are determined by the proportion $P(i(x, y))_{i(x,y)=0}$ of the dark pixels and the proportion $P(i(x, y))_{i(x,y)=1}$ of the bright pixels:

$$\begin{aligned} s_t(x, y)_e &\propto P(i(x, y))_{e_r} \\ c_t(x, y)_e &\propto P(i(x, y))_{e_r} \\ d_t(x, y)_e &\propto P(i(x, y))_{e_r} \end{aligned} \quad (20)$$

Thus at local regions containing boundaries the relationship between a real MRI slice and its corresponding ideal slice can be expressed as follows;

$$s_t(x, y)_e \geq s_c(x, y)_b \quad (21)$$

$$c_t(x, y)_e \geq c_c(x, y)_b \quad (22)$$

$$d_t(x, y)_e \geq d_c(x, y)_b \quad (23)$$

2.3. Quality Measure in Real MRI images

Let I_t , I_{t_s} , I_{t_c} and I_{t_d} denote grayscale, local entropy, local contrast and local standard deviation feature images extracted from the real MRI slice I_t . The notations, BW_t , BW_s , BW_c , BW_d denote the threshold versions of the grayscale, entropy, contrast and standard deviation feature images at global thresholds μ_t , μ_s , μ_c and μ_d , respectively. Let $f(x, y)$ denote the foreground image extracted from the grayscale feature image and n_f the total number of foreground pixels. The total number of bright pixels and the Euler number in BW_t are denoted n_t and EUL , respectively.

We define five quality scores for a real MRI slice based on its fidelity to corresponding quality attribute of an ideal MRI image. They are noise $q1$, lightness $q2$, contrast $q3$, sharpness $q4$ and texture details $q5$.

2.3.1. Noise Quality Factor

The noise quality factor $q1$ measure how the homogeneous regions in real MRI slice deviates from the piecewise constant regions in ideal MRI slice:

$$q1 = 1 - \left(\frac{\log(EUL)}{\log(n_f - n_t)} \right) \quad (24)$$

The reasoning here is that isolated single pixels in the homogeneous regions within the binary image of a real MRI slice are mainly derived from the presence of noise. Thus the Euler number as a region property of the binary image is a suitable metric for noise measure.

2.3.2. Lightness Quality Factor

The lightness quality factor $q2$ is a measure of distortion in the pixel intensity levels in the observed image. It is quantified by the proportion of the intersection of two sets; the binary image of the grayscale image and the binary image of the entropy image:

$$q2 = \left(\frac{BW_t \cap BW_s}{n_t} \right) \quad (25)$$

2.3.3. Contrast Quality Factor

The contrast quality factor $q3$ is the proportion of the intersection of three sets; the binary image of the grayscale image, binary image of the entropy image and the binary image of the contrast image:

$$q3 = \left(\frac{BW_t \cap BW_s \cap BW_c}{n_t} \right) \quad (26)$$

2.3.4. Sharpness Quality Factor

The sharpness quality factor $q4$ is the proportion of the intersection of three sets; the binary image of the grayscale image, binary image of the entropy image and the binary image of the standard deviation image:

$$q4 = \left(\frac{BW_t \cap BW_s \cap BW_d}{n_t} \right) \quad (27)$$

2.3.5. Texture Details Quality Factor

The texture details quality factor $q5$ is the proportion of the average of the bright pixels n_s, n_c, n_d in the binary images of the entropy image, the contrast image and the standard deviation image:

$$q5 = \left(\frac{n_s + n_c + n_d}{3} \right) \left(\frac{1}{n_t} \right) \quad (28)$$

2.3.6. Perceptual Weights

Perceptual weight assigned to each quality factor is derived from the results of the experiment reported in [36]. The experiment models a MRI slice as a Markov random field (MRF) [16], [59], [27] but without reference to a prior model image. A clean image, assumed to have zero noise variance, is at equilibrium energy state. Increasing levels of noise degradation correspond to energy states that are higher than the equilibrium energy. On the other hand increasing levels of blur degradation corresponds to energy states that are lower than the equilibrium energy. The energy at each level of noise is referred to as the MRF or total clique potential (TCP) energy. It is computed from the sum of local interactions of pixels, referred to as local clique potential energy. Figure 1 and Fig. 2 shows how a simple phantom obtained from McGill University BrainWeb Simulated Database [9] can be used to demonstrate the experiment in [36]. The original image of a MRI slice, assumed to be free of noise, is shown in Fig. 1a. It can be perceived by the human eye as having pleasant appearance because of the strong cluster of similar pixels and high contrast between homogeneous regions. Increase in Rician noise (Fig. 1a - Fig. 1c) and blur (Fig. 2a - Fig. 2c) from level 0 through level 3, level 7 to level 10, results in gradual weakening of the clusters that characterize the image. The human eye perceives the gradual weakening of clusters as variations in contrast, sharpness and lightness quality attributes.

The experiment in [36] propose a power model that express the relationship between the normalized total clique potential E_t and the standard deviation $\hat{\sigma}$ of Gaussian noise for brain MRI images:

$$\begin{aligned} E_t &= a\hat{\sigma}^b + c, \quad \{E_t : E_t \leq 1\} \\ a_{b_g} &= -1.67, a_{f_g} = -0.6863 \\ b_{b_g} &= -0.6764, b_{f_g} = -0.3663 \\ c_{b_g} &= 1.053, c_{f_g} = 1.105 \end{aligned} \quad (29)$$

where $a_{b_g}, a_{f_g}, b_{b_g}, b_{f_g}, c_{b_g}, c_{f_g}$ are the model parameters for the foreground (f_g) and background (b_g) modes. The foreground mode of the power model shown in Fig. 2d is of interest to our research as our proposed method evaluates the quality of a MRI slice by considering only the foreground pixels. Cursory view of Fig. 2d shows that the plot does not intercept with the y-axis. The plot is asymptotic to the y axis, $\min_{E_t} \approx 0.5$ because real MRI images were utilized for the experiment. Only ideal MRI slice has zero variance. There is no real MRI slice with exact zero variance. We adopt the horizontal asymptote $y = 0.5$ as threshold T_h to classify the MRF energies into two energy bands, noise energy band \mathcal{E}_N and blur energy band \mathcal{E}_B :

$$\mathcal{E} = \begin{cases} E_t \geq T_h & \text{Noise Energy Band} \\ E_t < T_h & \text{Blur Energy Band} \end{cases} \quad (30)$$

The lower figure of Fig. 2d indicates increasing dominance of noise and blur above and below the threshold, respectively and conforms with the suggestion in [14] that noise can be regarded as a unique type of information. Both degradation processes results in loss of sharpness. Quality factors that complements each other are assigned to the same energy band. They are also assigned equal weights within the same energy band since they provide complementary information.

The noise energy band consists of the noise quality factor $q1$ and the lightness quality factor $q2$ with perceptual weights:

$$w1 = w2 = \left(\frac{E}{2}\right) \quad (31)$$

where $w1, w2$ are the weights assigned to the noise and lightness quality factors, respectively.

The blur category consists of the contrast quality factor $q3$, sharpness quality factor $q4$ and texture details quality factor $q5$ with perceptual weights:

$$w3 = w4 = w5 = \left(\frac{1-E}{3}\right) \quad (32)$$

where w_3 , w_4 and w_5 are the weights assigned to contrast, sharpness and texture details quality factors.

2.3.7. Total Quality Score

For a m percent Rician noise level the standard deviation of the equivalent Gaussian noise is given by

$$\sigma \approx \mathcal{N}\left(0, \frac{\tau m}{100}\right) \quad (33)$$

where τ is the maximum pixel intensity [10]. Given an estimate of the noise level in a MRI slice the total quality score Q is the weighted sum of the five quality factors:

$$Q = q_1 \left(\frac{E}{2}\right) + q_2 \left(\frac{E}{2}\right) + q_3 \left(\frac{1-E}{3}\right) + q_4 \left(\frac{1-E}{3}\right) + q_5 \left(\frac{1-E}{3}\right) \quad (34)$$

3. Experimental setup

Data used for the performance evaluation of our proposed method were obtained from a clinical research organization, NeuroRx research Inc (<https://www.neurorx.com>) and a medical device development and research organization, BrainCare Oy (<http://braincare.fi/>).

The experiment was performed on more than 100 hundred T2, T1 and fluid attenuation inverse recovery (FLAIR) magnitude MRI volume data. To comply with the required page limits for this report we provide results on only four MRI volume data. Three of the four dataset are a 60-slice, 2.4mm thickness, 256×256 data from NeuroRx Research Inc. The first is a T1-weighted brain MRI data originally acquired with various levels of degradation by bias fields. The second data is a T2-weighted data which can be perceived as having acceptable quality. The third data is a FLAIR having quality that can be considered as borderline between good and bad quality. The data from BrainCare Oy is T2-weighted and is considered a good quality image. It consists of 24 slices. Each slice has dimension 448×390 and thickness of 7.4mm.

Three types of degradation; blur, motion blur and Rician noise at different levels were artificially induced on the test data. Blur was simulated by convolving a slice with circular averaging filter of radius $\{r : 1 \leq r \leq 10\}$ pixels. 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 pixels and the angular distance in degree. Both parameters were scaled from 1 to 10. Two identical Gaussian noise levels were generated and each was separately added to the test slice and its identical slice. This simulates the real and imaginary components in the complex plane of MRI acquisition process. Rician noise is added to the test image by computing the magnitude of the complex data. The Rician noise is scaled from 1 to 10 percent.

4. System Operation

The algorithm was implemented on Matlab computing environment. The flow chart in Fig. 3 and the images in Fig. 4 explains the four stages of the quality evaluation experiment. The algorithm regard individual slice as a separate image and thus operates slice-by-slice on a MRI volume data.

4.1. Extraction of Foreground

The initial task in the experiment is the extraction **FRX** of the foreground **FRG** from the FLAIR image **TIM** shown in Fig. 4a. The indices of pixels in the foreground is used to determine the mean pixel intensity level and the number of pixels n_f in the foreground of the test image.

4.2. Texture Filtering

The test image is convolved separately with three texture filters **TEX**; local entropy filter, local contrast filter and local standard deviation filter to obtain local entropy (Fig. 4b), local contrast (Fig. 4c) and local standard deviation (Fig. 4d) feature images, respectively. The algorithm is sensitive to the size of filter. We recommended the use of a 3×3 filter as the finest filter for an image where the row and column dimensions are ≤ 256 and a 5×5 filter for an image where the row and the column dimensions are > 256 .

4.3. Hard Thresholds

First moments of four feature images; the grayscale test image and the three local texture images are computed with reference to the foreground pixels. These moments are the global thresholds used to binarize **BIT** the four feature images. Binarization classifies the intensity levels in the image into regions of similar attributes [11] and reduce the computational as well as the operational complexity of the algorithm [46]. Furthermore binarization conforms with the two-class piecewise constant ideal MRI slice explained in section IIA and follows the same reasoning in [49] which regard the observed image as a blurred version of an ideal binary image.

4.4. Quality Evaluation

Since the four feature images are directly derived from the same slice there is no need to incorporate additional resources such as image registration into the algorithm. This design feature will potentially reduce the operational complexity of the system. Binary morphological operation **MPH** is executed on the foreground and the three binary local texture feature images to generate four region properties. They are Euler number **EuL** associated with the white matter region, area of the entropy image **aE**, area of the contrast image **aC** and the area of the standard deviation image **aS**. The Euler number region property and the number of foreground pixels are used to determine the noise quality factor **q1** according to Eq. 24. The intersection of the binarized test image and the entropy image gives the lightness quality factor **q2** according to Eq. 25. The intersection of three binary images; test image, entropy and contrast determines the contrast quality factor **q3** according to Eq. 26. The intersection of another set of three binary images; test image, entropy image and standard deviation image gives the sharpness quality score **q4** according to Eq. 27. Texture details **q5** computed according to Eq. 28 is determined from the average of the areas of the entropy, contrast and standard deviation binary images. The total quality score is the weighted sum of the five quality scores according to Eq. 34.

5. Experimental Results

In this results section we display quality evaluation on one T2-weighted MRI volume data each from NeuroRx and BrainCare. There is also quality evaluation results for a T1-weighted volume data from NeuroRx which the slices were originally degraded by different levels of bias fields. A slice was selected from each of the T2-weighted data to demonstrate the efficacy of our proposed method on images with different levels of noise, blur and motion blur. In the plots, the five quality criteria; noise, lightness, contrast, sharpness and texture details are colored deep blue, light blue, green, pink and brown, respectively. The total quality score is in a separate plot with deep blue color.

5.1. Acceptable Quality

Two slices in a MRI volume data from BrainCare and NeuroRx considered as having acceptable image quality are shown in Fig. 5a - Fig. 5b and Fig. 6a - Fig. 6b, respectively. Figure 5c - Fig. 5d and Fig. 6c - Fig. 6d are the five quality scores and the total quality scores of successive slices in the MRI volume data, respectively.

5.2. Bias Fields

Two slices in a MRI volume data from NeuroRx degraded by bias fields are shown in Fig. 7a - Fig. 7b. Figure 7c - Fig. 7d are the five quality scores and the total quality scores of successive slices in the MRI volume data.

5.3. Blur

The image in Fig. 8a is a slice in the volume data from NeuroRx. Its blurred version at blur level of 10 is shown in Fig. 8b. The plot of increasing levels of blur versus the five quality criteria is displayed in Fig. 8c. The total quality score of the slice at each level of blur is in Fig. 8d.

5.4. Motion blur

The image in Fig. 9a is a slice in the volume data from BrainCare. Its motion blurred version at motion blur level of 10 is shown in Fig. 9b. The plot of the variation of increasing levels of motion blur with the five quality criteria is displayed in Fig. 9c. Figure 9d shows how the total quality score of the slice varies with the different levels of motion blur.

5.5. Noise

The image in Fig. 10a is the same slice in Fig. 9a but degraded by Rician noise level of 10 as shown in Fig. 10b. The plot of increasing level of noise with the five quality criteria is displayed in Fig. 10c. The variation of the total quality score with noise levels is shown in Fig. 10d.

5.6. Comparative Performance Evaluation

Figure 11 and Fig. 12 are displayed to compare our proposed quality measure method to the existing methods that adopt global sharpness and noise level quality measures to predict the quality of brain MRI images. The global sharpness measure was formulated in [19] and implemented as a quality measure in [54]. Noise level was estimated using the fast noise variance estimation algorithm proposed by [23]. The images for comparative performance evaluation had their pixel intensity levels normalized to lie between 0 and 1 [40].

The image in Fig. 11a is a slice in T1 volume data originally acquired with bias fields. The index number of the slice in the MRI volume data is 19. Global sharpness measure for 21 successive slices in the T1 MRI volume data is displayed in Fig. 11c. The image in Fig. 11b is a slice in a T2 MRI volume data degraded by Rician noise level of 10 percent. The plot of variation of global sharpness with Rician noise levels increasing from 0 to 10 percent is shown in Fig. 11d.

The T1 MRI slice in Fig. 12a is the same slice in Fig. 11a. The estimate of noise levels in 21 successive slices in the T1 MRI volume data is shown in Fig. 12c. The image in Fig. 12b is the same slice in Fig. 11b but degraded by motion blur level of 10. The variation of global sharpness with motion blur levels increasing from 1 to 10 is displayed in Fig. 12d.

6. Discussion

6.1. Classification Across Perceived Good Quality Images

The algorithm demonstrated very good classification across perceived good quality images. Two examples are the good quality MRI volume data from BrainCare and NeuroRx displayed in Fig. 5 and Fig. 6, respectively. The average quality scores of the slices in the BrainCare volume data shown in Fig. 5c are 0.55 for noise, 0.65 for lightness, 0.5 for contrast, 0.5 for sharpness and 0.75 for texture details. The average of the total quality score shown in Fig 5d is 0.55. The data from NeuroRx in Fig. 6c had average quality scores of 0.6 for noise, 0.8 for lightness, 0.6 for contrast, 0.6 for sharpness and 0.9 for details. The average of the total quality score shown in Fig. 6d is 0.75.

6.2. Classification Across Perceived Poor Quality Images

The T1-weighted images in Fig. 7a - Fig. 7b and the plots of quality scores shown in Fig. 7c - Fig. 7d for successive MRI slices degraded by bias fields demonstrates that our proposed method can give good classification across poor quality images. The average of the five quality scores are 0.5 for noise and 0.55 for lightness, 0.1 for contrast, 0.1 for sharpness and 0.7 for texture. The average of the total quality score is 0.35 as shown in Fig. 7d.

6.3. Classification Across Different Quality Factors

The plots in Fig. 7c, Fig. 8c, Fig. 9c and Fig. 10c demonstrates that our proposed method can give good classification across different image quality factors. An example is the plot in Fig. 7c that shows the five quality criteria for the image degraded by bias fields. The image can be perceived as free of noise, clearly visible grayscale pixels with texture and low contrast between the different anatomic structures. These quality perception is consistent with the average quality score of 0.5 for noise, 0.55 for lightness. The average quality score of 0.1 for contrast and 0.1 for sharpness is also acceptable because bias fields results in perceived low contrast and sharpness between the white matter and cortical grey matter regions.

6.4. Relationship Between Different Quality Factors

The plots in Fig. 7d, Fig. 8d, Fig. 9d and Fig. 10d demonstrate how the different quality factors relate with each other for the different levels of degradation. All the plots shows a general decrease in the total quality factor for different levels of degradation because our proposed method gave a fairly accurate estimate of the various quality criteria. The plot in Fig. 10d indicate gradual loss of lightness, contrast, sharpness and texture details as a result of increasing levels of Rician noise.

6.5. Consistent Quality Prediction Across Different Levels of Distortion

In Fig. 8 - Fig. 10 there is general reduction in the total quality scores for increasing levels of the dominant distortion for T2-weighted images degraded by blur, motion blur and noise. It is obvious that the T2-weighted images in Fig. 8a, Fig. 9a and Fig 10a are better in quality than their corresponding distorted versions in Fig. 8b, Fig. 9b and Fig. 10b, respectively. Even state-of-the-art automated image analysis system will fail to accurately classify the images degraded by bias fields shown in Fig. 7 into two-tissue class consisting of white matter and gray matter regions. The objective quality score recorded by our proposed method demands that these images be corrected for bias field before analysis. The performance evaluation results is an indication that our proposed method has the potential to demonstrate good correlation with human observers in a subjective performance evaluation experiment.

6.6. Suitability for Automated Environment

Quality control in fully automated image analysis environments require physicians or trained readers to assess the performance of image analysis systems on large volumes of MRI data. State-of-the-art image analysis systems are generally efficient in the analysis of good quality images. There is limit to their performance and their performance can be said to be related to the quality of the image. Quality control task becomes difficult and cumbersome when the quality of the output of the system is perceived to lie on the borderline between acceptable and unacceptable quality. Our proposed method assess the quality of an image based on five different quality factors. In automated environments our proposed method can be used in conjunction with human observers in a subjective evaluation experiment. This can help determine generally acceptable objective quality scores for borderline bad and borderline good images.

6.7. Limitations of the Proposed Algorithm: Modeling Error

The performance of our proposed algorithm is limited by modeling error. In the ideal model the cortical gray matter characterized by high density of edges and the smoothly varying ventricle are two distinct tissue classes. The modeling error result from the transformation of the test image into a binary image. In the binary image the cortical gray matter and the ventricular system are regarded as belonging to the same tissue class because of the narrow distance in their mean pixel intensity levels. The effect of the modeling error is evident in the total quality score displayed in Fig. 5d and Fig. 6d. The increase in quality score with slice number correlates with the reduction in the percentage volume of the ventricular system with increasing slice number. This phenomenon is clearly evident in Fig. 6d. The ventricular system is about 17 percent of the brain volume [30] hence the modeling error does not have significant impact on our proposed quality measure technique.

6.8. Quality Scores Cut-off Mark

A slice in a MRI volume is classified as good or bad with reference to the asymptotic behavior of the power model shown in Fig. 2d. The cut-off for each quality score is set at the horizontal asymptote, 0.5. The asymptote is a natural threshold which demarcates the energy band into two separate regions dominated by noise and blur, respectively. The slice monotonically loses its sharpness quality as the distance on either side of the threshold increases. To account for variations in the perceptual weights of the different quality scores the cut-off for the total quality score is set at 0.45, about 10 percent lower than the cut-off for individual quality factors.

6.9. Comparative Performance Evaluation

There are several problems that hinders direct objective comparative performance evaluation of the few currently available post-acquisition quality evaluation methods for brain MRI images. Existing methods adopt different distortion models. The SNR quality measure proposed in [15] is a full-reference quality assessment method. The reports in [54] and [34] adopt ANOVA as quality index. Evaluation results reported on existing methods can be said to be data dependent because there is currently no ground truth data available for clinical brain MRI images. The use of phantom is not an effective approach for comparative performance evaluation because it lacks the natural anatomical variability and image acquisition artifacts that are usually encountered in real images [50]. It is for these reasons that our proposed method is compared to existing methods such as [15], [54] and [34] based on their design characteristics.

The plot in Fig. 11d demonstrate that global sharpness, one of the several global quality measures in [54], can efficiently discriminate between different levels of distortion. Evaluation parameters derived from ANOVA does not transform to a quality index and the parameters can have ambiguous meanings [31]. ANOVA does not indicate if the quality measure variable decrease, increase or is random with increasing levels of distortion. Both proposals can be regarded as distortion detection techniques because it does not transform the distortion levels into an index of image quality. Quality evaluation is the transformation of the different levels of degradation into a quality index. The prediction accuracy and hence the perceptual quality of an image quality model can only be assessed based on a quality index [56]. The absence of a quality index makes it impossible to compare the two separate sharpness-distortion levels variations shown in Fig. 11c and Fig. 11d with other algorithms.

We will now consider a scenario where we choose to disregard the absence of a quality index. The sharpness quality attribute which effectively discriminate different levels of motion blur in Fig. 11d can be said to be an erroneous quality indicator for images degraded by bias fields. The average global sharpness of 2×10^4 recorded for the slices in Fig. 11c is greater than the global sharpness of 7×10^2 recorded for the T2 slice in the absence of distortion. There is insignificant level of noise $\sigma < 1$ recorded for the slices in Fig. 12c. The global sharpness of 4×10^2 recorded for the T2 slice degraded by blur in Fig. 12d is comparable to the sharpness attribute of 7×10^2 recorded for the T2 slice in the absence

of distortion in Fig. 11d. Thus sharpness and noise attribute can be said to be an erroneous quality indicator for an image degraded by bias fields and motion blur. The implication of this demonstration is that single or few relevant quality attributes are not sufficient to capture all the possible distortions that can be present in an image [26]. The contribution by [15] which adopt signal-noise-ratio attribute will be limited in their performance because it relies only on gauging the intensity attributes of distortions [26]. The quality evaluation method proposed by [34] will also be limited in performance because the detection of noise and artifacts are too few attributes to capture all the possible distortions in an image.

7. Future Work

Image quality evaluation is the focus of collaborative research by our group. In the near future we hope to develop new quality evaluation techniques that addresses the different challenges in the use of brain MRI images for clinical research. New quality evaluation methods will incorporate three new features. First is a region segmentation technique for the three major anatomical structures of the brain; white matter, gray matter and the ventricles. This approach will extend the performance of our proposed method to include region-of-interest quality measure alongside current global quality measure. Second feature is a new approach to automatically compute the perceptual weight associated with each quality attribute. This can be achieved through the use of a full phantom experiment to explore the various simulation frameworks for the generation and quantification of artifacts. The third feature is a validation study to compare how a proposed method compares with subjective evaluation by human observers.

8. Conclusion

Medical images exhibit characteristics which distinguish them from each other and the other classes of natural images. It will be futile to design a generic image quality metric because of intra-class and inter-class variability in the characteristics of images. We hereby propose a new application-specific objective post-acquisition quality assessment for MRI images of the brain which takes into account the major anatomical structures of the brain. Processing of the quality metrics is on binary images, hence the proposed method is simple and efficient for both automated and manual environments. Experimental results demonstrates that our proposed method can objectively classify perceived good and poor quality images based on different quality criteria. Thus our proposed method will meet current challenges and ensure accurate prediction of image quality, improve the discernment of the trained MRI reader and encourage time-efficient consensus in subjective quality evaluation.

References

- [1] Abdalmajeed, S., Shuhong, J., 2014. No-reference image quality assessment algorithm based on weibull statistics of log-derivatives of natural scenes. *Electronics Letters* 50 (8), 595–596.
- [2] Arnold, D., You, X., Shang, S., Sperling, B., Evilevitch, V., 2015. Long-term efficacy in mri and no evidence of disease activity outcomes in patients with relapsing-remitting multiple sclerosis treated with peginterferon beta-1a. *Neurology* 84 (14 Supplement), P7–266.
- [3] Arnold, D. L., 2005. Changes observed in multiple sclerosis using magnetic resonance imaging reflect a focal pathology distributed along axonal pathways. *Journal of neurology* 252 (5), v25–v29.
- [4] Chalavi, S., Simmons, A., Dijkstra, H., Barker, G. J., Reinders, A. S., 2012. Quantitative and qualitative assessment of structural magnetic resonance imaging data in a two-center study. *BMC medical imaging* 12 (1), 27.

- [5] Chandler, D. M., 2013. Seven challenges in image quality assessment: past, present, and future research. *ISRN Signal Processing* 2013.
- [6] Chen, H., Tao, J., Sun, Y., Ye, Z., Qiu, B., 2015. Magnetic resonance image reconstruction via lo-norm minimization. In: 2015 IET International Conference on Biomedical Image and Signal Processing (ICBISP 2015). IET, pp. 1–6.
- [7] Chen, Y., Li, J., Zhang, H., Zheng, Y., Jeon, B., Wu, Q. J., 2016. Non-local-based spatially constrained hierarchical fuzzy c-means method for brain magnetic resonance imaging segmentation. *IET Image Processing* 10 (11), 865–876.
- [8] Ciancio, A., Da Costa, A. T., Da Silva, E., Said, A., Samadani, R., Obrador, P., 2009. Objective no-reference image blur metric based on local phase coherence. *Electronics Letters* 45 (23), 1162–1163.
- [9] Collins, D. L., Zijdenbos, A. P., Kollokian, V., Sled, J. G., Kabani, N. J., Holmes, C. J., Evans, A. C., 1998. Design and construction of a realistic digital brain phantom. *IEEE transactions on medical imaging* 17 (3), 463–468.
- [10] Coupé, P., Manjón, J. V., Gedamu, E., Arnold, D., Robles, M., Collins, D. L., 2010. Robust rician noise estimation for mr images. *Medical image analysis* 14 (4), 483–493.
- [11] Deravi, F., Pal, S. K., 1983. Grey level thresholding using second-order statistics. *Pattern Recognition Letters* 1 (5-6), 417–422.
- [12] Fang, Y., Ma, K., Wang, Z., Lin, W., Fang, Z., Zhai, G., July 2015. No-reference quality assessment of contrast-distorted images based on natural scene statistics. *Signal Processing Letters, IEEE* 22 (7), 838–842.
- [13] Frisoni, G. B., Fox, N. C., Jack, C. R., Scheltens, P., Thompson, P. M., 2010. The clinical use of structural mri in alzheimer disease. *Nature Reviews Neurology* 6 (2), 67–77.
- [14] Gabarda, S., Cristóbal, G., 2007. Blind image quality assessment through anisotropy. *JOSA A* 24 (12), B42–B51.
- [15] Gedamu, E. L., Collins, D., Arnold, D. L., 2008. Automated quality control of brain mr images. *Journal of Magnetic Resonance Imaging* 28 (2), 308–319.
- [16] Geman, S., Geman, D., 1984. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence* (6), 721–741.
- [17] Gray, R. M., 1990. Entropy and information. In: *Entropy and Information Theory*. Springer, pp. 21–55.
- [18] Guo, Y., Lebel, R. M., Zhu, Y., Lingala, S. G., Shiroishi, M. S., Law, M., Nayak, K., 2016. High-resolution whole-brain dce-mri using constrained reconstruction: Prospective clinical evaluation in brain tumor patients. *Medical physics* 43 (5), 2013–2023.
- [19] Hadjidemetriou, E., Grossberg, M. D., Nayar, S. K., July 2004. Multiresolution histograms and their use for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26 (7), 831–847.
- [20] Han, Y., Du, H., Gao, X., Mei, W., 2017. Mr image reconstruction using cosupport constraints and group sparsity regularisation. *IET Image Processing* 11 (3), 155–163.

- [21] Horé, A., Ziou, D., 2013. Is there a relationship between peak-signal-to-noise ratio and structural similarity index measure? *IET Image Processing* 7 (1), 12–24.
- [22] Huynh-Thu, Q., Ghanbari, M., 2008. Scope of validity of psnr in image/video quality assessment. *Electronics letters* 44 (13), 800–801.
- [23] Immerkaer, J., 1996. Fast noise variance estimation. *Computer vision and image understanding* 64 (2), 300–302.
- [24] Kim, D., Park, R., 2012. New image quality metric using random projection. *IET Image Processing* 6 (9), 1246–1255.
- [25] Kupinski, M., Barrett, H., 2007. *Small-Animal SPECT Imaging*. Springer US.
URL `isbn:9780387252940`
- [26] Larson, E. C., Chandler, D. M., 2010. Most apparent distortion: full-reference image quality assessment and the role of strategy. *Journal of Electronic Imaging* 19 (1), 011006–011006.
- [27] Li, S. Z., 2009. *Markov random field modeling in image analysis*. Springer Science & Business Media.
- [28] Li, W., Ogunbona, P., Desilva, C., Attikiouzel, Y., 2011. Semi-supervised maximum a posteriori probability segmentation of brain tissues from dual-echo magnetic resonance scans using incomplete training data. *IET image processing* 5 (3), 222–232.
- [29] Li, X., Tao, D., Gao, X., Lu, W., 2009. A natural image quality evaluation metric. *Signal Processing* 89 (4), 548–555.
- [30] Lüders, E., Steinmetz, H., Jäncke, L., 2002. Brain size and grey matter volume in the healthy human brain. *Neuroreport* 13 (17), 2371–2374.
- [31] McCarthy, G., Wood, C. C., 1985. Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section* 62 (3), 203–208.
- [32] Miao, J., Huang, F., Narayan, S., Wilson, D. L., 2013. A new perceptual difference model for diagnostically relevant quantitative image quality evaluation: A preliminary study. *Magnetic resonance imaging* 31 (4), 596–603.
- [33] Mittal, A., Moorthy, A., Bovik, A., Dec 2012. No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing* 21 (12), 4695–4708.
- [34] Mortamet, B., Bernstein, M. A., Jack, C. R., Gunter, J. L., Ward, C., Britson, P. J., Meuli, R., Thiran, J.-P., Krueger, G., 2009. Automatic quality assessment in structural brain magnetic resonance imaging. *Magnetic Resonance in Medicine* 62 (2), 365–372.
- [35] Nirmala, S., Dandapat, S., Bora, P., 2007. Image quality assessment in retinal image compression systems. In: *IET-UK International Conference on Information and Communication Technology in Electrical Sciences (ICTES 2007)*, 2007. ICTES. IET, pp. 737–742.
- [36] Osadebey, M., Bouguila, N., Arnold, D., 2013. The clique potential of markov random field in a random experiment for estimation of noise levels in 2d brain mri. *International Journal of Imaging Systems and Technology* 23 (4), 304–313.

- [37] Parikh, P., Sandhu, G., Blackham, K. A., Coffey, M., Hsu, D., Liu, K., Jesberger, J., Griswold, M., Sunshine, J. L., 2011. Evaluation of image quality of a 32-channel versus a 12-channel head coil at 1.5 t for mr imaging of the brain. *American Journal of Neuroradiology* 32 (2), 365–373.
- [38] Pedersen, M., 2015. Evaluation of 60 full-reference image quality metrics on the CID: IQ. In: 2015 IEEE International Conference on Image Processing, ICIP 2015, Quebec City, QC, Canada, September 27-30, 2015. pp. 1588–1592.
- [39] Pedersen, M., Bonnier, N., Hardeberg, J. Y., Albregtsen, F., 2010. Attributes of image quality for color prints. *Journal of Electronic Imaging* 19 (1), 011016–011016–13.
- [40] Peli, E., 1990. Contrast in complex images. *JOSA A* 7 (10), 2032–2040.
- [41] Prieto, F., Guarini, M., Tejos, C., Irarrazaval, P., 2009. Metrics for quantifying the quality of mr images. In: Proceedings of the 17th Annual Meeting of ISMRM. p. 4696.
- [42] Ramasamy, U., Arulprakash, G., 2016. Mid-sagittal plane detection in brain magnetic resonance image based on multifractal techniques. *IET Image Processing* 10 (10), 751–762.
- [43] Rehman, A., Wang, Z., 2012. Reduced-reference image quality assessment by structural similarity estimation. *IEEE Transactions on Image Processing* 21 (8), 3378–3389.
- [44] Ryu, S., Sohn, K., 2011. No-reference sharpness metric based on inherent sharpness. *Electronics letters* 47 (21), 1178–1180.
- [45] Saad, M., Bovik, A., Charrier, C., 2012. Blind image quality assessment: A natural scene statistics approach in the dct domain. *IEEE Transactions on Image Processing* 21 (8), 3339–3352.
- [46] Shi, Y.-Q., Xia, X., 1997. A thresholding multiresolution block matching algorithm. *IEEE Transactions on Circuits and Systems for Video Technology* 7 (2), 437–440.
- [47] Soundararajan, R., Bovik, A. C., 2012. Rred indices: Reduced reference entropic differencing for image quality assessment. *IEEE Transactions on Image Processing* 21 (2), 517–526.
- [48] Tan, H. L., Li, Z., Tan, Y. H., Rahardja, S., Yeo, C., 2013. A perceptually relevant mse-based image quality metric. *IEEE Transactions on Image Processing* 22 (11), 4447–4459.
- [49] Tsai, W.-H., 1985. Moment-preserving thresholding: A new approach. *Computer Vision, Graphics, and Image Processing* 29 (3), 377–393.
- [50] Vovk, U., Pernus, F., Likar, B., 2007. A review of methods for correction of intensity inhomogeneity in mri. *IEEE transactions on medical imaging* 26 (3), 405–421.
- [51] Wang, Z., 2016. Objective image quality assessment: Facing the real-world challenges. In: Society for Imaging Science and Technology, International Symposium on Electronic Imaging, Image Quality and System Performance XII. pp. 1–6.
- [52] Wang, Z., Bovik, A. C., Sheikh, H. R., Simoncelli, E. P., 2004. Image quality assessment: From error visibility to structural similarity. *Trans. Img. Proc.* 13 (4), 600–612.
- [53] Weizman, L., Eldar, Y. C., Bashat, D. B., 2015. Compressed sensing for longitudinal mri: An adaptive-weighted approach. *Medical physics* 42 (9), 5195–5208.
- [54] Woodard, J., Carley-Spencer, M., 2006. No-reference image quality metrics for structural mri. *Neuroinformatics* 4 (3), 243–262.

- [55] Wu, J., Lin, W., Shi, G., Liu, A., 2013. Reduced-reference image quality assessment with visual information fidelity. *IEEE Transactions on Multimedia* 15 (7), 1700–1705.
- [56] Xue, W., Zhang, L., Mou, X., Bovik, A. C., 2014. Gradient magnitude similarity deviation: a highly efficient perceptual image quality index. *IEEE Transactions on Image Processing* 23 (2), 684–695.
- [57] Ye, P., Doermann, D., 2012. No-reference image quality assessment using visual codebooks. *IEEE Transactions on Image Processing* 21 (7), 3129–3138.
- [58] Yong, W., Yuqing, W., Xiaohui, Z., 2016. Complex number-based image quality assessment using singular value decomposition. *IET Image Processing* 10 (2), 113–120.
- [59] Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging* 20 (1), 45–57.

Acknowledgment

Marius Pedersen have been supported by the Research Council of Norway, project no. 247689 IQMED: Image Quality enhancement in MEDical diagnosis, monitoring and treatment.

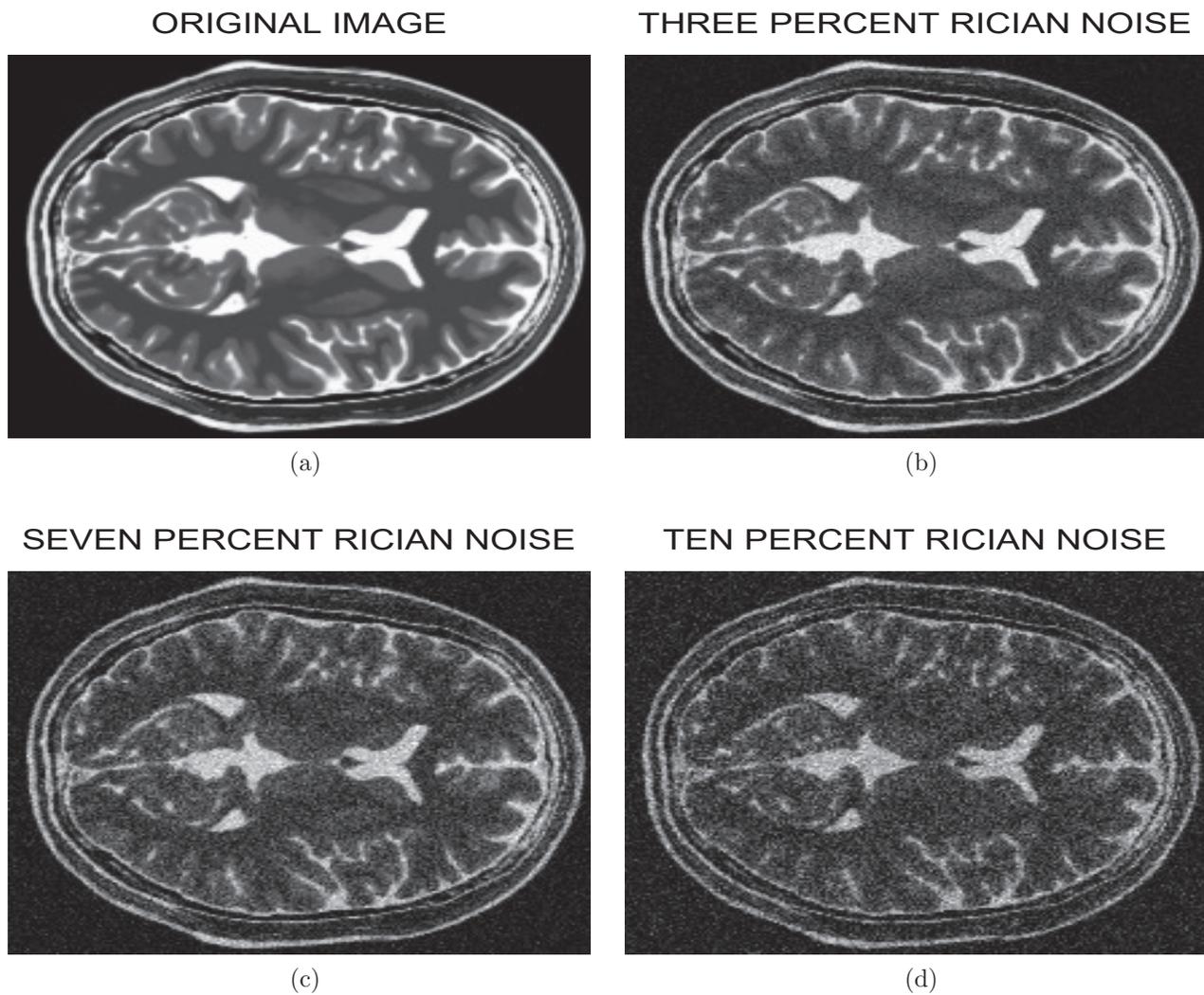


Figure 1: An MRI slice image at various levels of degradation by Rician noise (a) 0 percent (b) 3 percent (c) 7 percent (d) 10 percent.

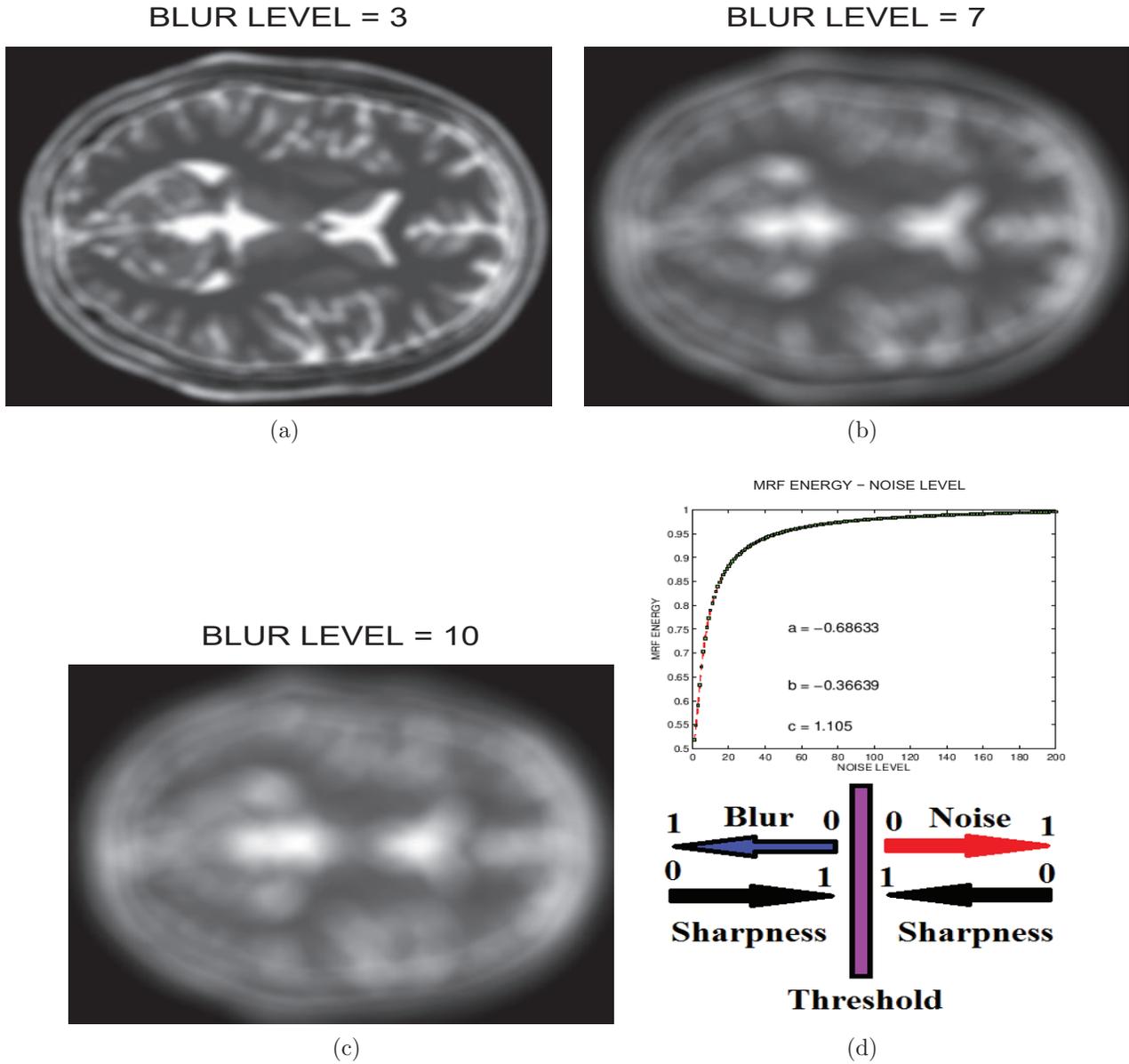


Figure 2: An MRI slice image at various levels of degradation by blurring with circular disk of radius (a) 0 pixels (b) 3 pixels (c) 7 pixels (d) The upper figure is the plot of the mathematical model for describing the relationship between MRF energy and noise level in the foreground of a MRI slice. The lower figure explains how the threshold set at the horizontal asymptote demarcates the MRF energies into the noise and blur energy bands. Sharpness monotonically decreases with increasing distance from the threshold.

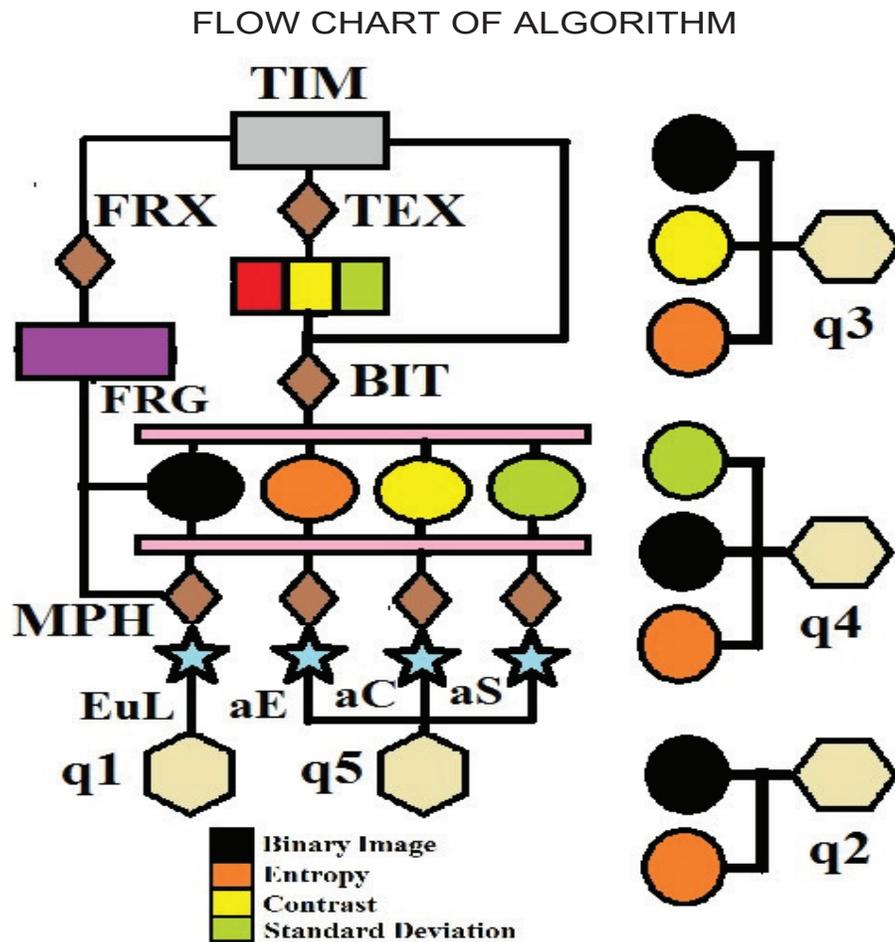


Figure 3: The flow chart of the method for quality evaluation in a brain MRI slice (**TIM**). The three initial steps are transformation to the binary domain (**BIT**), foreground extraction (**FRX**) and local texture analysis (**TEX**). These initial steps generate a binary image (**black colored circle**), foreground image (**FRG**) and three local texture images; local entropy (**orange colored circle**), local contrast (**yellow colored circle**) and local standard deviation (**green colored circle**). Morphological operation on the binary image and the binary images of the three texture images produces Euler number (**EuL**) and area properties (**aE**, **aC**, **aS**) for the computation of the noise quality score (**q1**) and texture details quality score (**q5**). The binary image of the test image is combined with the entropy image to compute the lightness quality score (**q2**). The contrast quality score (**q3**) is computed by the combination of the binary image, entropy image and contrast image. The combination of the binary image, the entropy image and the standard deviation image gives the sharpness quality score (**q5**)

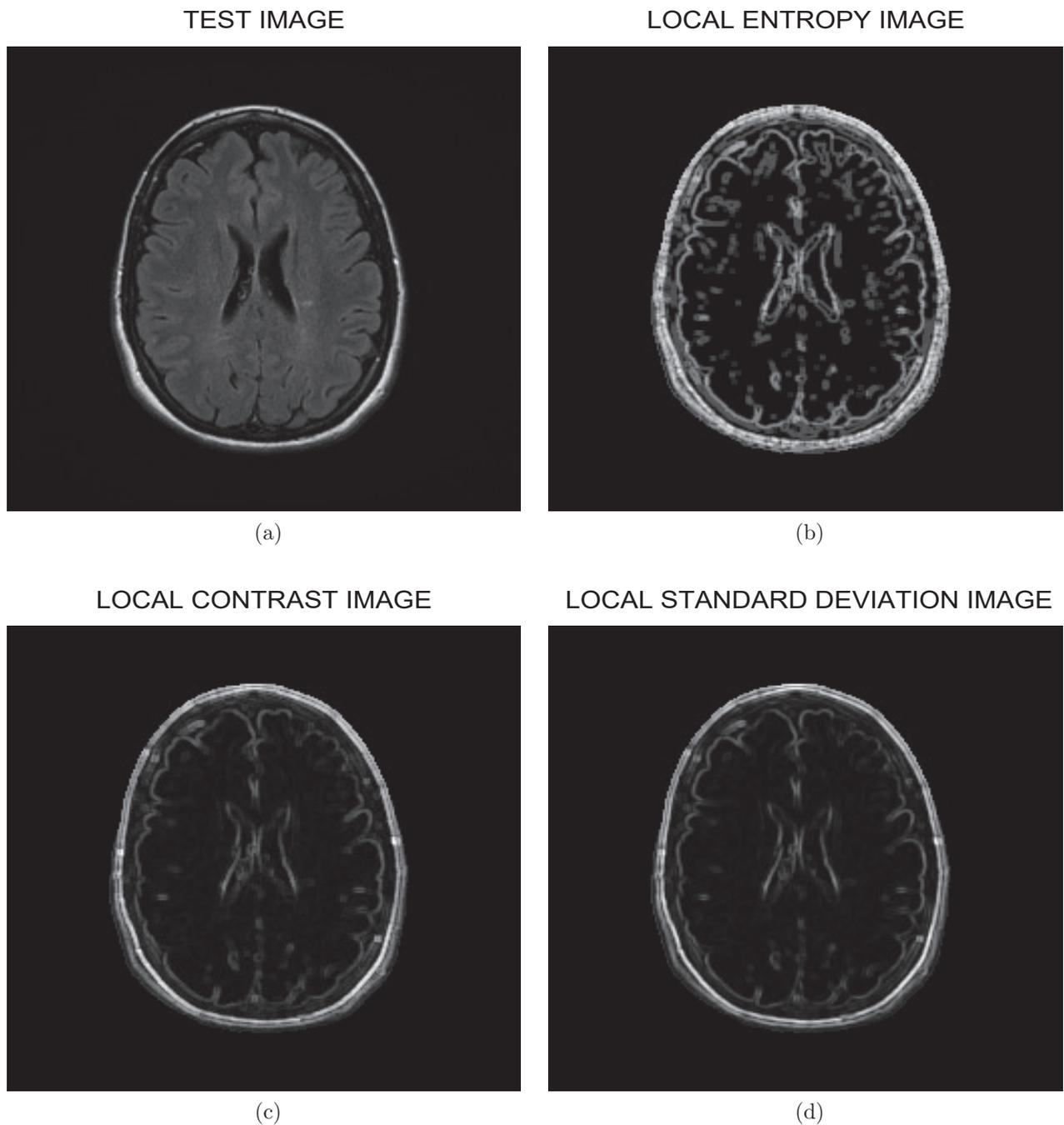


Figure 4: A FLAIR MRI slice (a) in a volume data. Texture filtering is applied to derive local entropy (b), local contrast (c) and local standard deviation (d) images.

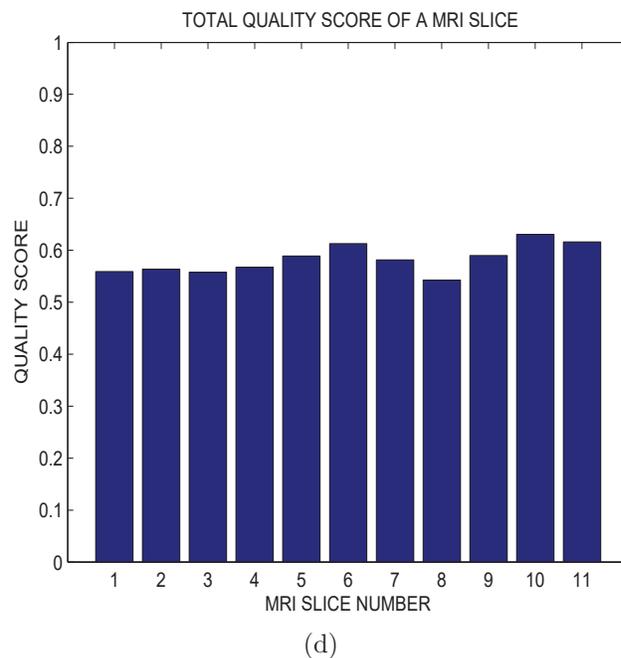
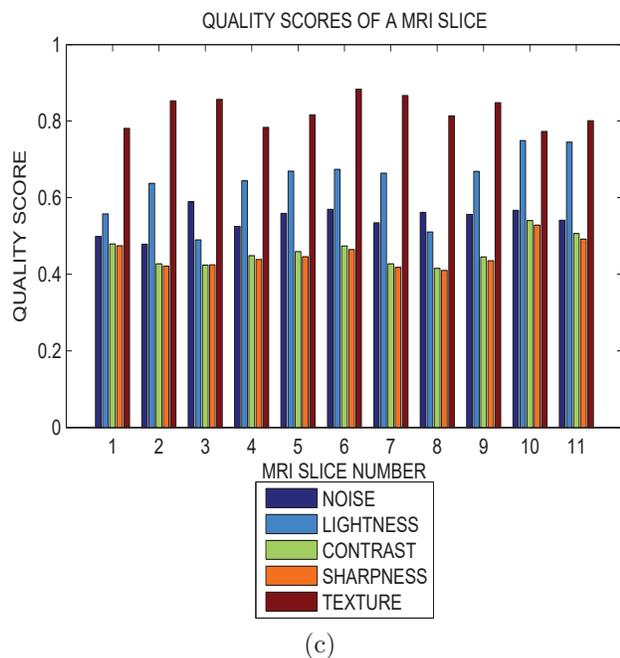
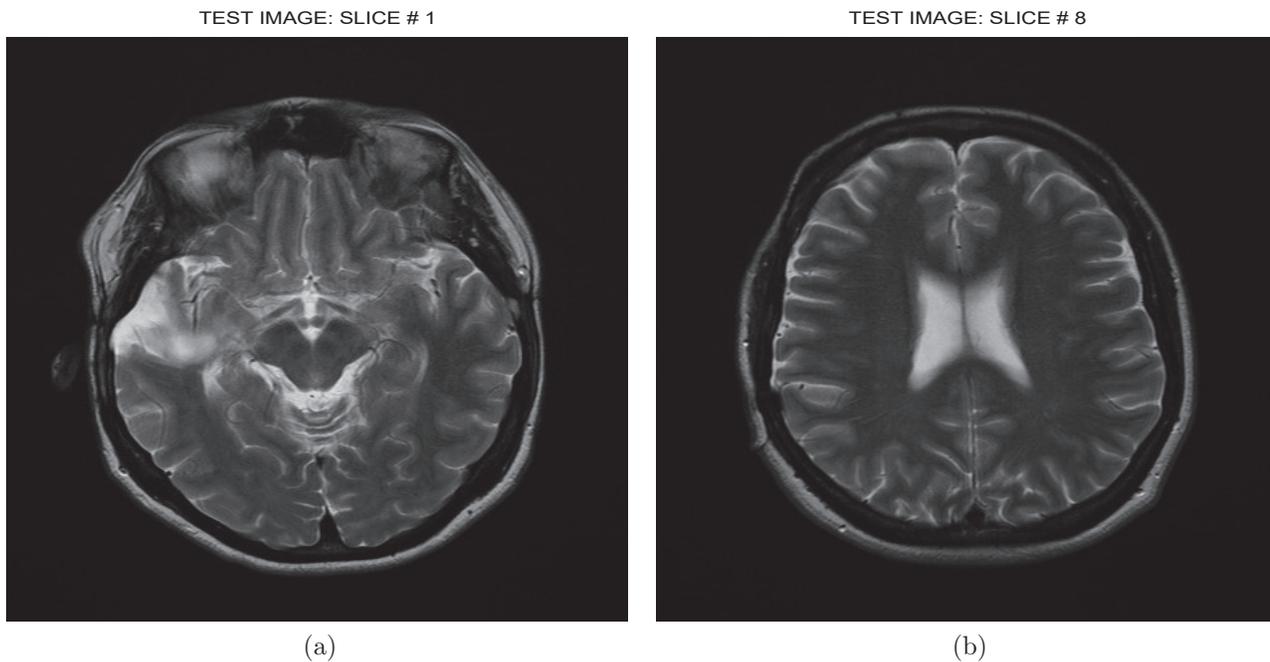


Figure 5: The images in (a) and (b) are slices with indices 1 and 8 in a T2-weighted MRI volume data from BrainCare. The quality scores of 11 successive slices in the volume data based on five quality criteria are in (c). The total quality score for each slice is in (d)

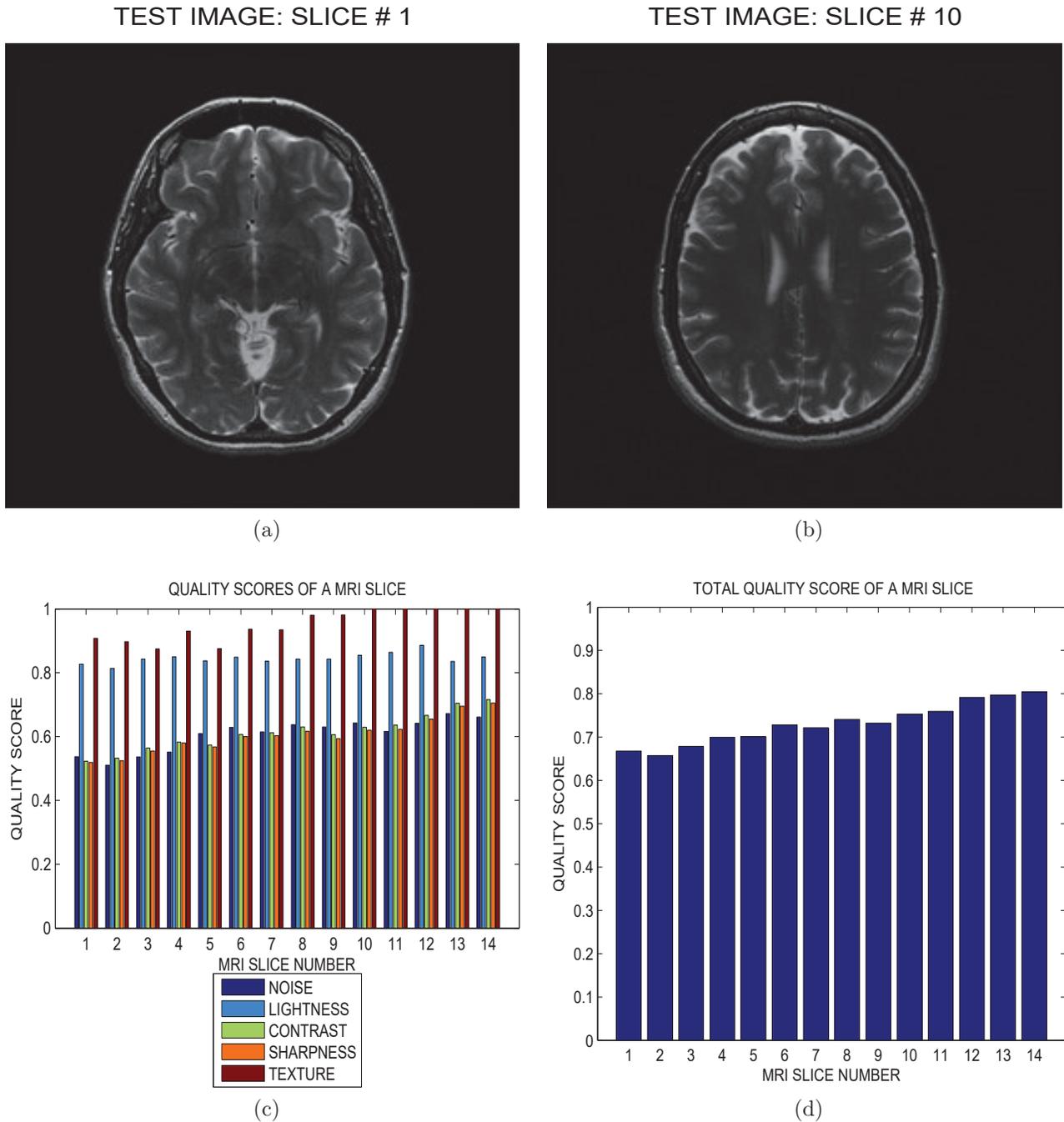
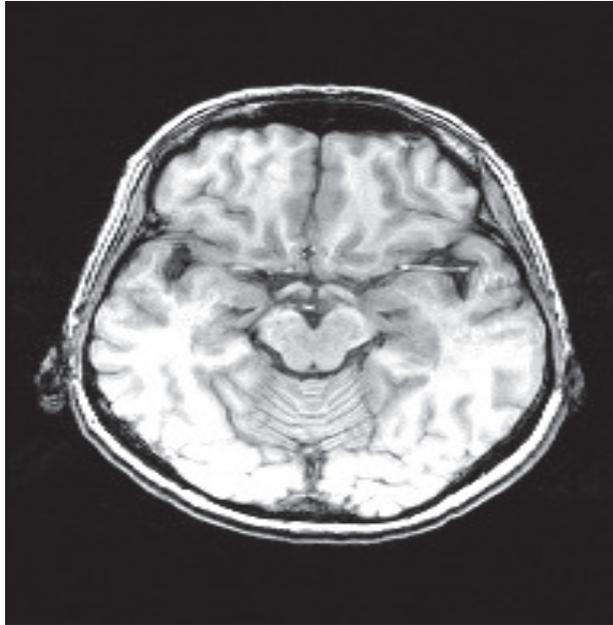


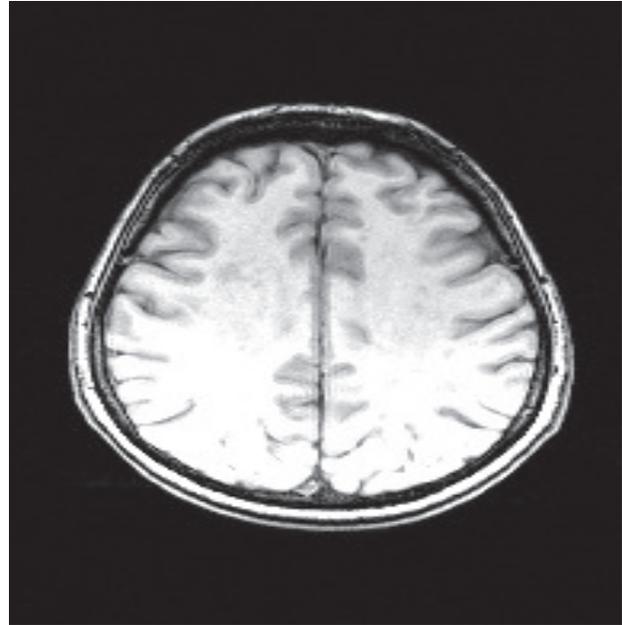
Figure 6: The images in (a) and (b) are slices with indices 1 and 10 in a T2-weighted MRI volume data from NeuroRx. The quality scores of 14 successive slices in the volume data based on five quality criteria are in (c). The total quality score for each slice is in (d)

TEST IMAGE: SLICE # 1

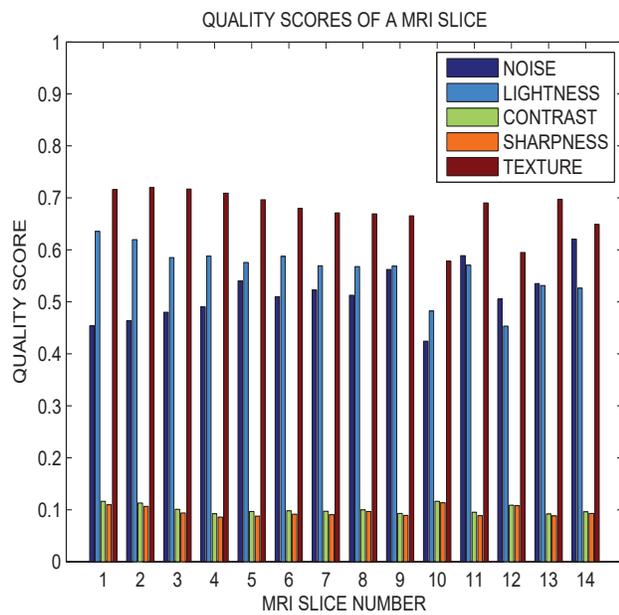


(a)

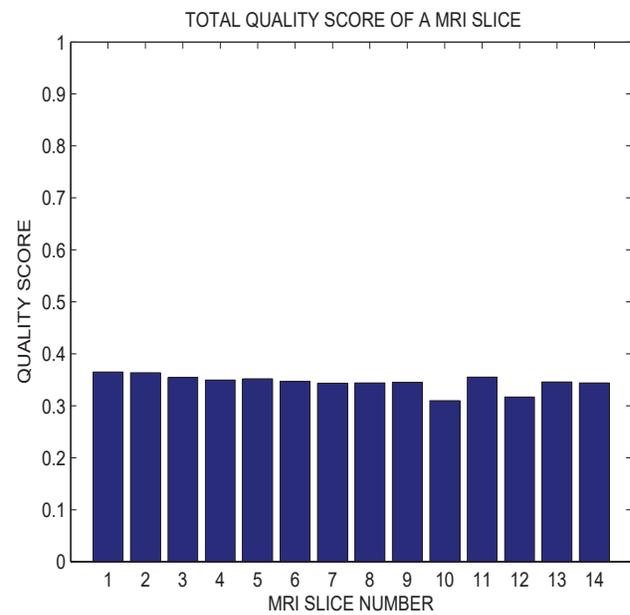
TEST IMAGE: SLICE # 14



(b)



(c)



(d)

Figure 7: The images in (a) and (b) are slices with indices 1 and 14 in a T1-weighted MRI volume data from NeuroRx. The MRI data is degraded by bias fields. The quality scores of 14 successive slices in the volume data based on five quality criteria are in (c). The total quality score for each slice is in (d)

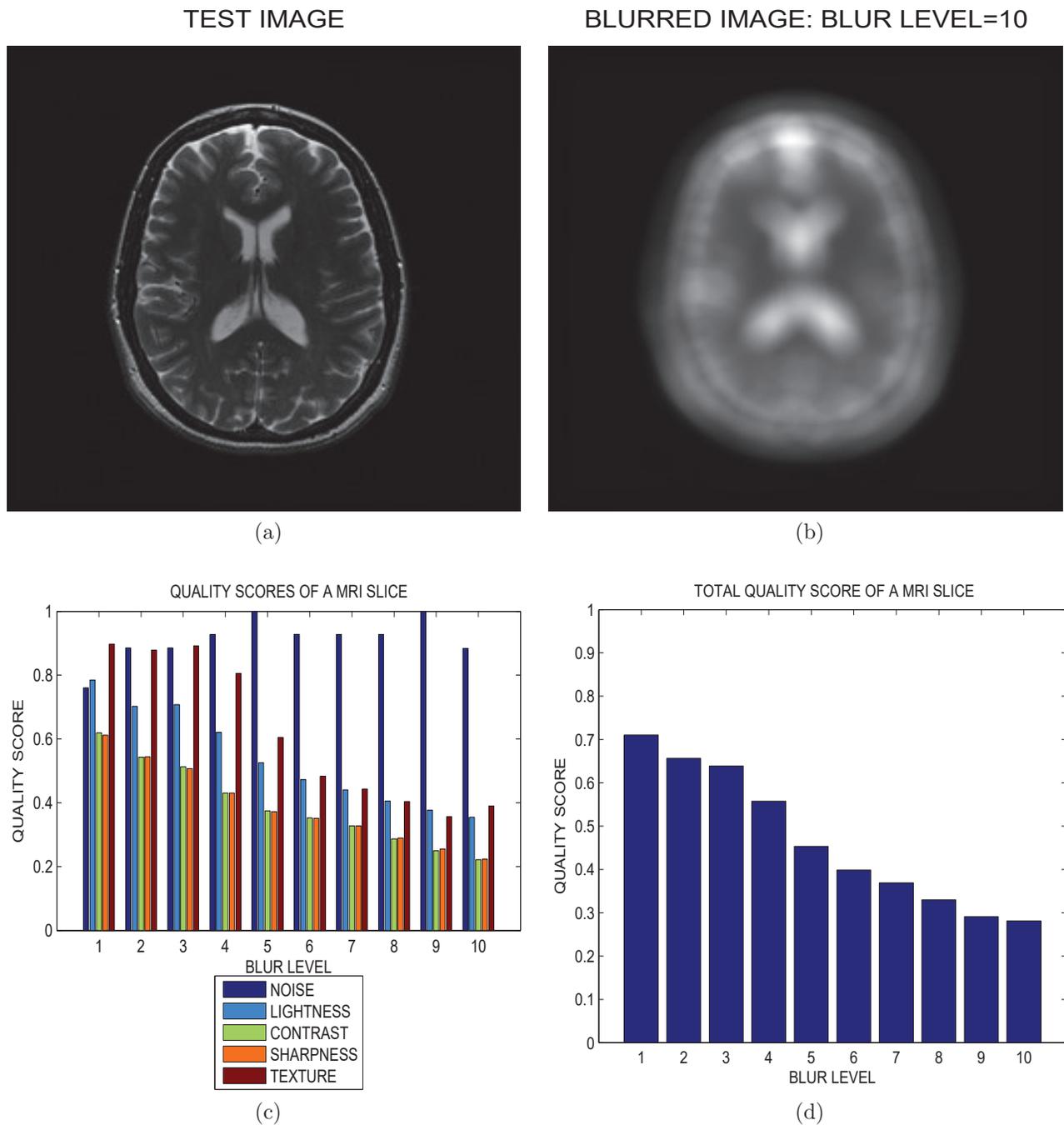


Figure 8: The image in (a) is a slice in a MRI volume data from NeuroRx shown in Fig. 3. It is degraded with increasing levels of blur using a circular averaging filter. The image in (b) is the image in (a) but degraded by blur level of 10. The quality scores of the image in (a), for blur levels increasing from 0 to 10, based on the five quality criteria are in (c). The total quality scores for each blur level are in (d)

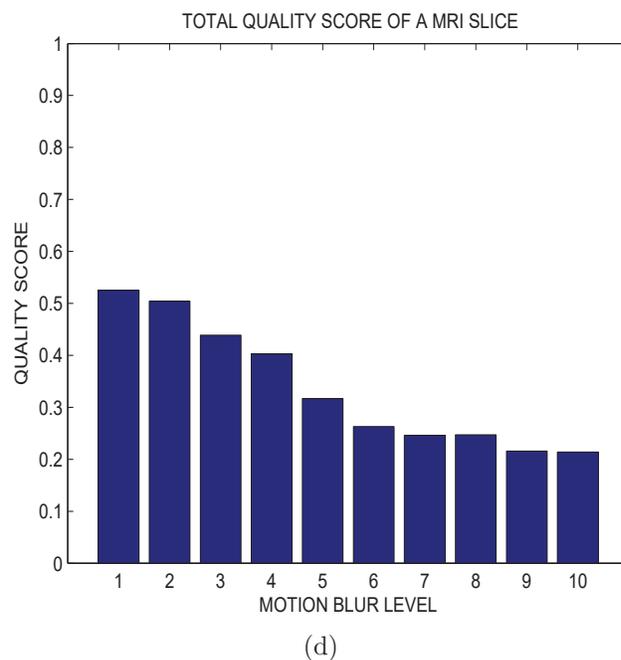
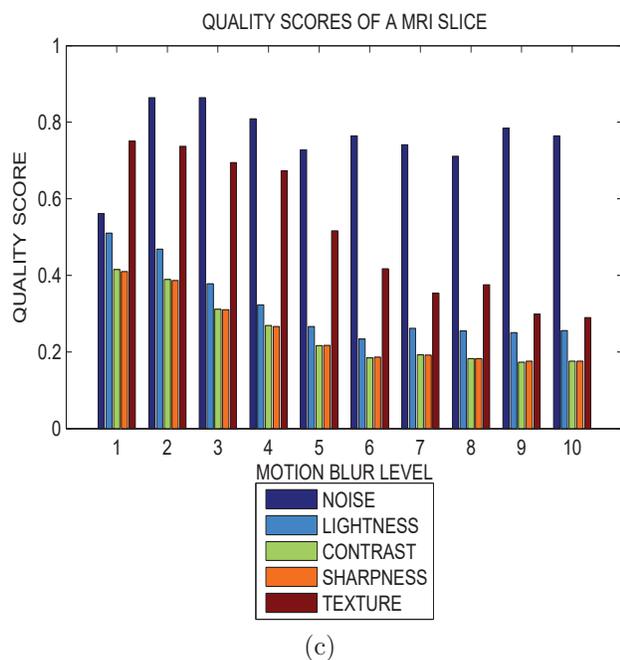
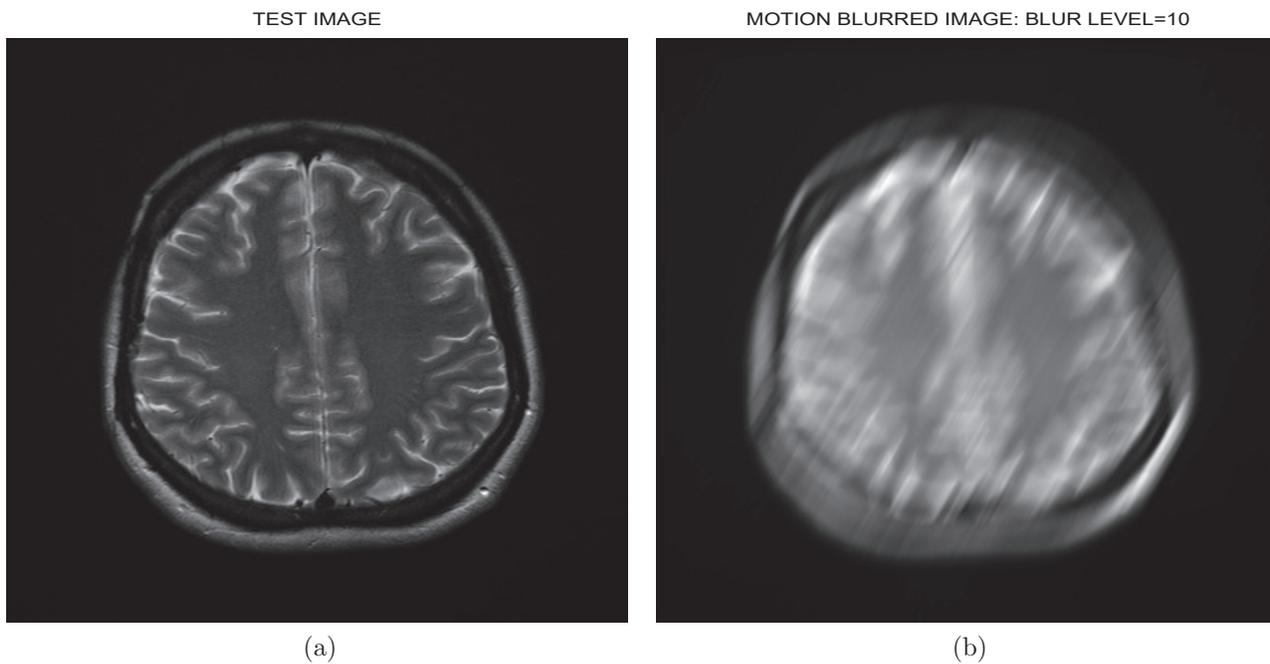


Figure 9: The image in (a) is a slice in a MRI volume data from BrainCare shown in Fig. 5. It is degraded with increasing levels of motion blur modeled as linear motion of a camera. The image in (b) is the image in (a) but degraded by motion blur level of 10. The quality scores of the image in (a), for motion blur levels increasing from 0 to 10, based on the five quality criteria are in (e). The total quality scores for each motion blur level are in (d)

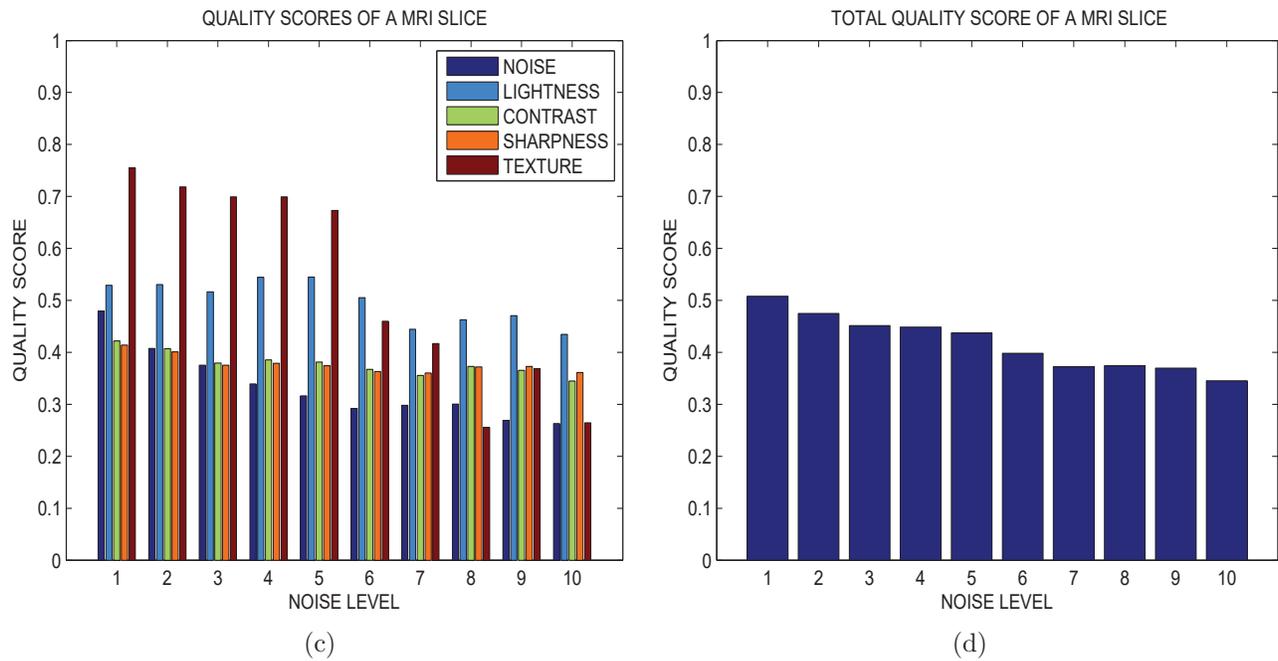
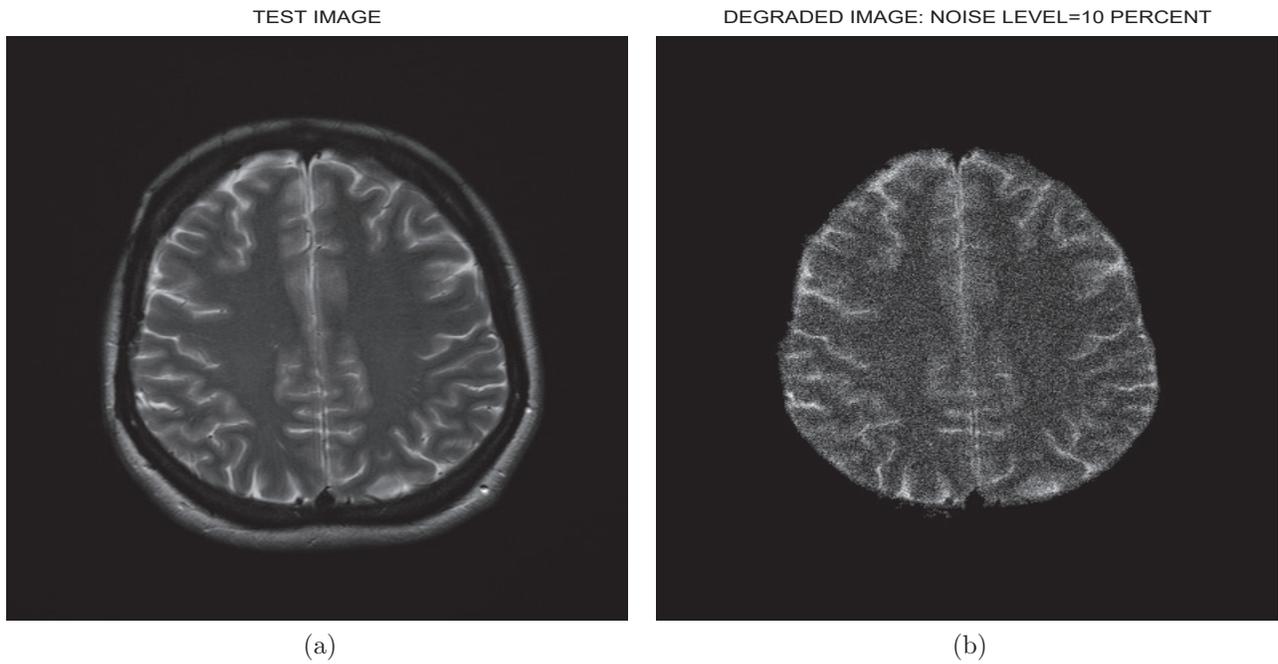
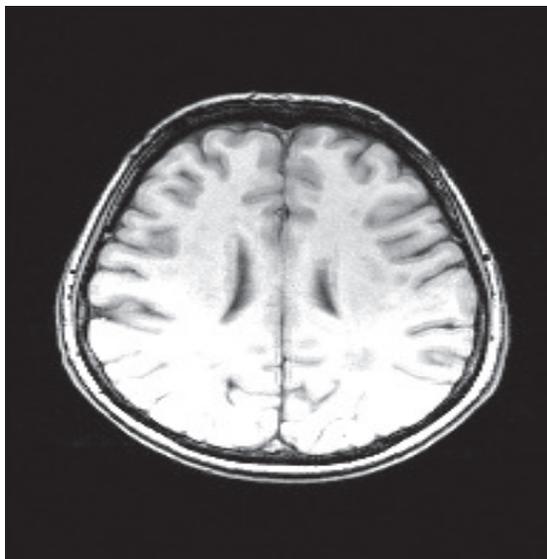


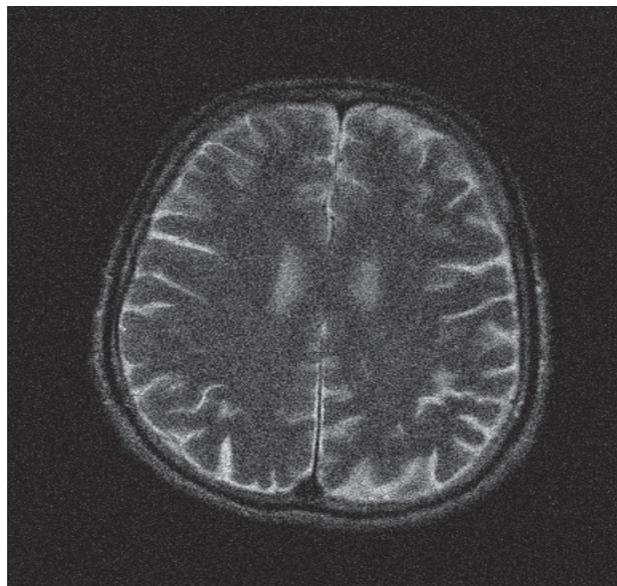
Figure 10: The image in (a) is a slice in a MRI volume data from BrainCare shown in Fig. 5. It is degraded with increasing levels of Rician noise. The image in (b) is the image in (a) but degraded by noise level of 10. The quality scores of the image in (a), for noise levels increasing from 0 to 10, based on the five quality criteria are in (e). The total quality scores for each noise level are in (d)

TEST IMAGE: SLICE # 19

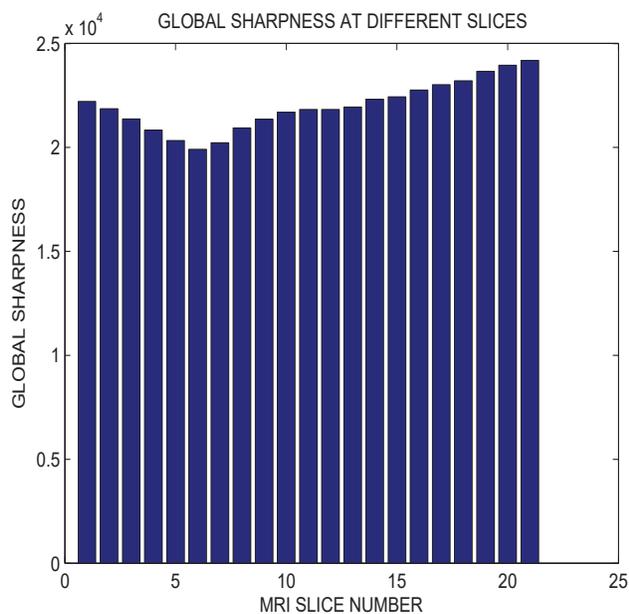


(a)

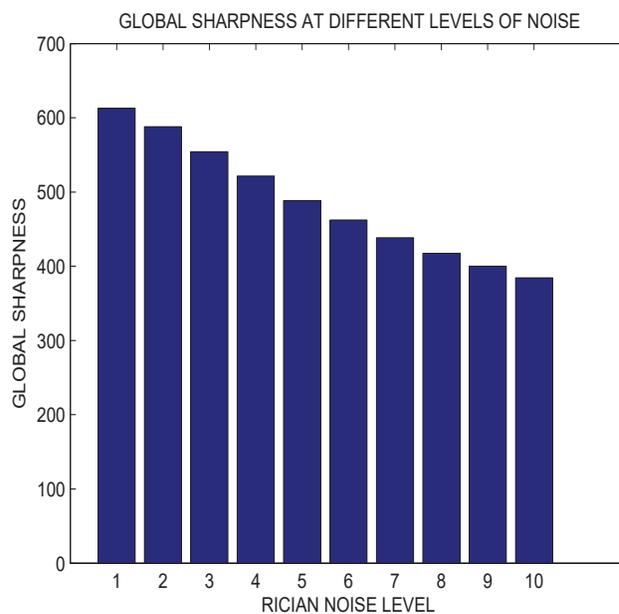
IMAGE DEGRADED BY RICIAN NOISE



(b)



(c)



(d)

Figure 11: The image in (a) is a slice in a T1 MRI volume data originally acquired with bias fields. The image in (b) is a slice in a T2 MRI volume data degraded by Rician noise level of 10 percent. (c) The global sharpness measures of 21 successive slices in the T1 MRI volume data shown in (a). (d) The global sharpness measure for Rician noise levels increasing levels from 1 to 10 in the image shown in (b).

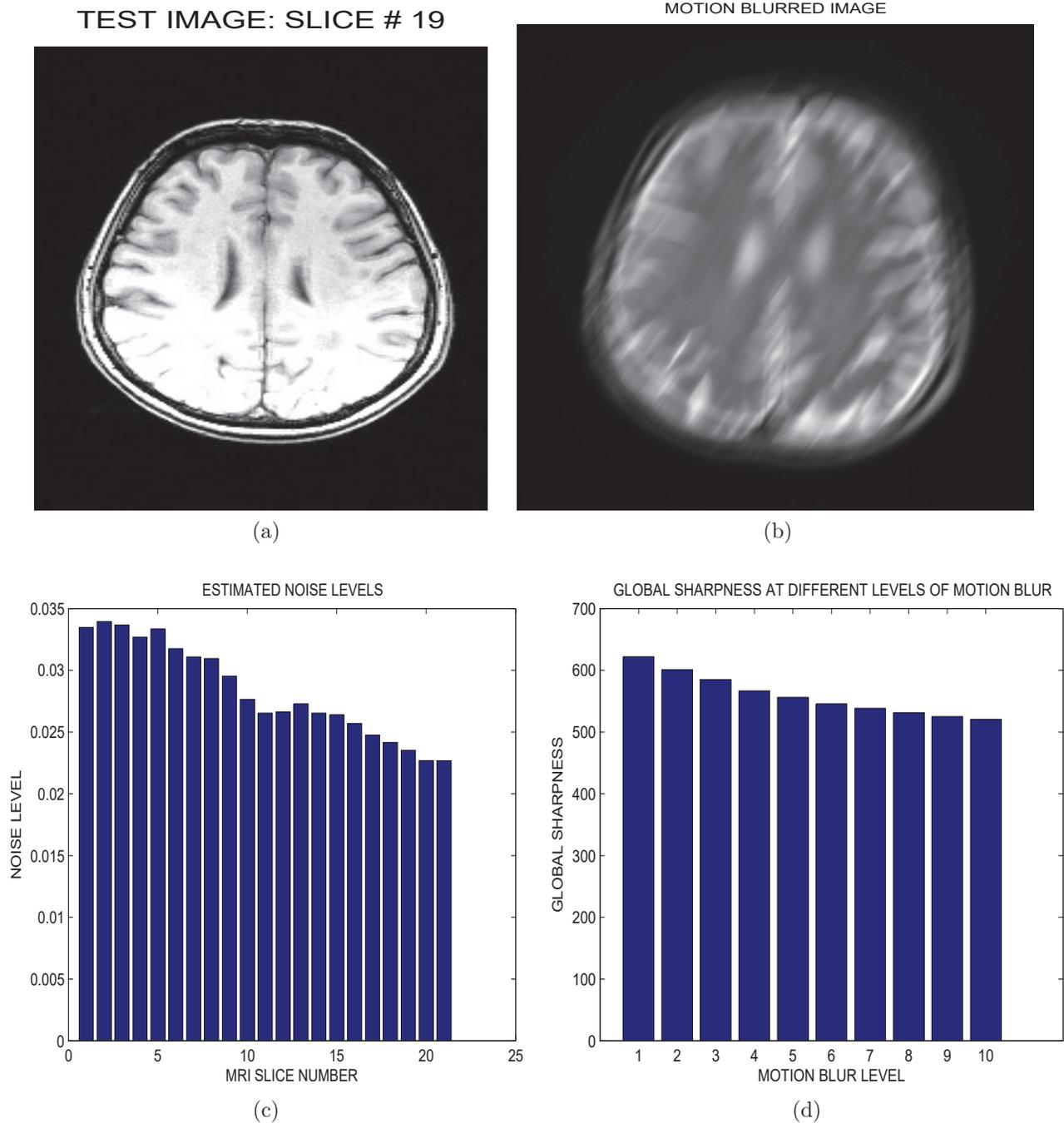


Figure 12: The image in (a) is the same slice in a T1 MRI volume data originally acquired with bias fields shown in Fig. 11a. The image in (b) is the same slice in a T2 MRI volume data shown in Fig. 11b but degraded by motion blur level of 10 . (c) The estimated noise levels of 21 successive slices in the T1 MRI volume data shown in (a). (d) The global sharpness measure for motion blur levels increasing from 1 to 10 in the image shown in (b).