A Simple Measure for Acuity in Medical Images

Frithjof Kruggel

Abstract—**An automatic and objective assessment of image quality is important in an era, where large-scale processing of imaging data from multi-center studies becomes commonplace. Based on a comprehensive statistical image model that includes noise and blur, a measure for image acuity is derived here as the ratio of the maximal gradient magnitude and the intensity difference at a boundary. Acuity may be affected by the object under study, the image acquisition, reconstruction processes, and any post-processing steps. The acuity measure presented here is post-hoc, intuitive to understand, simple to compute, and easily integrates with other standard measures of image quality. Three applications in medical imaging are included where our acuity measure is useful in the objective and automatic assessment of image quality.**

Index Terms—**Image quality, image artifacts, motion, medical imaging.**

I. INTRODUCTION

THE assessment of image quality is important in the development of new imaging technology, image acquisition and processing methods [11], [19], [26]. Before medium- and large-sized image data bases (e.g., anatomical and functional group studies in neuro-biology) are subjected to an image processing chain, it is important to ensure that pre-defined quality criteria are met, because later stages may fail or produce erroneous results. It is evident that image quality is best optimized during image acquisition, by applying the best imaging technology and protocol available for a specific task. At any level of technology, quantitative metrics are useful to rate and compare any effort to ensure or improve image quality.

Noise is a well-known influence that degrades image quality, and there is an abundance of literature on the origin and estimation of noise in medical images [2], [17], [18], [20], [36]. Most typically, the signal-to-noise (SNR) and contrast-to-noise ratio (CNR) are widely employed as quality measures [24]. While the SNR quantifies the amount of random influences on the intensity statistics of a specific region, CNR quantifies the contrast of two neighboring regions in the presence of noise. Here, we introduce acuity as another basic parameter of image quality, which is understood as the ability of an imaging process to retain edges. If an edge is sharp in the imaged object with respect to the image resolution, any smoothing is understood as a loss of acuity due to the imaging processing. A complementary term for

acuity is image blur [37]. Acuity in medical images is affected by (a) parameters and mode of image acquisition, (b) image reconstruction method, (c) spatial, temporal, and intensity resolution, (d) physiological noise, (e) object motion, (e) image processing (e.g., filtering and interpolation). As we will demonstrate below, our measure is largely independent of the SNR and CNR, so can be used as a complementary measure.

Early approaches of analytically modeling edges blurred by Gaussian kernels are described by Elder and Zucker [12] and Kayargadde and Martens [21]. However, their intent was to determine the precise location of an edge, not to estimate acuity. In a different context, Panda and Rosenfeld [33] demonstrated the usefulness of intensity × gradient magnitude histograms to distinguish regions of pure materials from mixtures at edges. In this two-dimensional histogram, voxels with pure materials map to low-gradient regions at a specific intensity, while voxels with two-material mixtures map to "arcs" between both materials (see Fig. 4). This idea was picked up by Kindlemann and Durkin [23] and combined with the mathematical edge model of Elder and Zucker [12] for the semi-automatic generation of transfer functions in volume rendering. Serlie *et al.* [39] provided an analytical form that models the shape of the arcs. Finally, Chiverton [10] developed an image segmentation method that rigorously incorporated a noise and gradient model in a Bayesian framework for estimating partial volume fractions from these histograms.

In the following section, we establish an image model with noisy and blurred edges to derive a simple acuity measure and discuss its properties in experiments on simulated and phantom images. To illustrate the use of this measure, we include three application examples in real medical images:

A. Detection of Motion-Corrupted Images in Large-Scale Data Bases

Gross body and physiologic motion is a well-known source of image degradation [31]. Unwanted motion is best minimized during acquisition, e.g., by using restraining mechanisms such as cushions, inflatable masks or mouth guards [34]. However, the application of such mechanisms may be limited in specific subject groups (e.g., newborn children [30]) or find limited tolerance (e.g., children with autism spectrum disorders, elderly demented subjects). As an alternative, within-scan correction schemes were developed that use information from external trackers [15], [29] or image-based measures to estimate and correct for motion [8], [15], [27], [41]. Such approaches require modification of the imaging protocol and/or upgrades to the hardware inventory that may not (yet) be available in a specific setting. Post-hoc motion
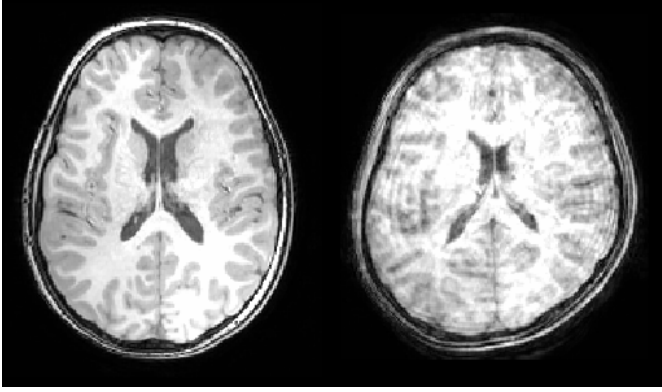
Fig. 1.   Axial slices of a MRI data set of the human head with acceptable quality (left, Acuity $= 1.441$, SNR $= 47.8$, CNR $= 10.2$), and a motion-corrupted data set (right, Acuity $= 1.057$, SNR $= 40.3$, CNR $= 9.31$). While SNR and CNR are similar, the much lower acuity indicates the presence of motion blur.

correction schemes are widely applied in situations where several short, repeated scans are acquired (e.g., in functional MRI), such that within-scan motion is neglected and between-scan motion corrected by image registration [25], [43]. In a clinical setting, trained personnel will typically identify a degraded scan immediately, and repeat or reorder an exam, eventually while sedating the patient during the exam. In an academic research environment, sedation is not feasible or viable (e.g., in functional imaging), scan time scheduling often does not allow for an immediate repeat, and re-invitations may not be possible due to a limited budget. Due to the huge amount of data acquired in specific protocols (e.g., in diffusion-weighted MRI), and their relatively low image quality, motion may visually be difficult to detect. Given an image data base, we consider as due diligence to check that all data meet or exceed pre-defined quality criteria before any further image processing is applied. Our application example is based on a medium-scale data base of images acquired in children with Autism Spectrum Disorder (ASD, see also Fig. 1).

B. Differentiation of Moving and Stationary Boundaries in an Image

We use a time-series of ECG-gated cardiac images as an example and distinguish motion on the heart boundary from a stationary region at the spine. Imaging technology to detect and quantify heart motion is well developed these days, and elaborate schemes exist for tracking motion [42], e.g., during angiography [14]. Most easily, information about moving boundaries could be sampled from difference images in the time domain. However, we estimate regional motion within a scan based on our acuity measure, as a post-hoc measure, at a single time point. We include this example to demonstrate that within-scan motion can be measured, where the presence of moving and stationary regions is obvious.

C. Sampling of a Common Anatomical Template

The problem of generating a high-quality study-specific anatomical template has been under considerable study,

given the huge popularity and wide application of voxel- and tensor-based morphometry these days. The sampling process of such a template is computationally demanding, due to the high-resolution non-linear registration process underneath. While early approaches used the full image set for template generation [32], later publications demonstrated that sampling from a subset is advantageous [13]. This strategy does not only reduce the computational demand, but also yields a sharper template, which improves the registration process for the full set. Other authors emphasized the importance of including the anatomical variability of the sample, and developed sampling schemes to select specific cases. Our example demonstrates an application of the acuity measure in a sampling scheme that leads to a high-quality, sharp template. We note that our process can be integrated with other objectives, such as the population-based sampling above.

II. METHODS

Before deriving the acuity measure, statistical properties of intensities and gradient magnitudes at edges in the presence of noise and image blur are revised.

A. Gradient Magnitudes in Uniform Noisy Images

Consider a $D > 1$ dimensional image I discretized on a regular lattice at sites s. Let the image consist of a single region R with intensity μ, corrupted by independent Gaussian-distributed noise with variance σ_i^2. The probability distribution function (PDF) of the intensities x is given as:

$$x \sim \mathcal{N}(\mu, \sigma_i^2) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left[-\frac{(x - \mu)^2}{2\sigma_i^2}\right]. \qquad (1)$$

Convolution operators are employed to estimate the gradient magnitude that uses a linear combination of local intensity differences. e.g., for central differences:

$$g_s = \frac{1}{2}\left[\sum_{d=1}^{D}(x(s_l) - x(s_r))^2\right]^{1/2}, \qquad (2)$$

where $x(s_l), x(s_r)$ denote the intensities of both the N_6-connected neighbors (for $D = 3$) of site s in direction d. By definition, the gradient magnitude g follows a *central* χ-distribution with $k > 1$ degrees-of-freedom (DOF):

$$g \sim \chi_k(0, \sigma_g^2) = \frac{g^{k-1}}{\sigma_g^k 2^{k/2-1}\Gamma(k/2)} \exp\left[-\frac{g^2}{2\sigma_g^2}\right]. \qquad (3)$$

The mean of this distribution is $E\{g\} = \sqrt{2\sigma_g^2}\,\Gamma(k/2 + 1/2)/\Gamma(k/2)$, where $\Gamma(\cdot)$ denotes the Gamma function [1]. Let us make the following remarks:

(a) Note that the DOF k of this distribution is related to the number of summation terms of the gradient filter (Eq. 2), not the dimensionality of the image.

(b) If $x_1 \sim \mathcal{N}(\mu_1, \sigma_i^2)$ and $x_2 \sim \mathcal{N}(\mu_2, \sigma_i^2)$, then $d_{1,2} = x_2 - x_1 \sim \mathcal{N}(\mu_2 - \mu_1, 2\sigma_i^2)$. Thus, $\sigma_g^2 = 2\sigma_i^2$.

(c) Any (directional) weighting factor applied in the convolution filter applies to σ_g. For the central difference filter above: $\sigma_g^2 = \frac{1}{2}\sigma_i^2$. The noise level σ_i in an image region defined above can be estimated as: $\sigma_i = E\{g\}\sqrt{\pi}/2$.
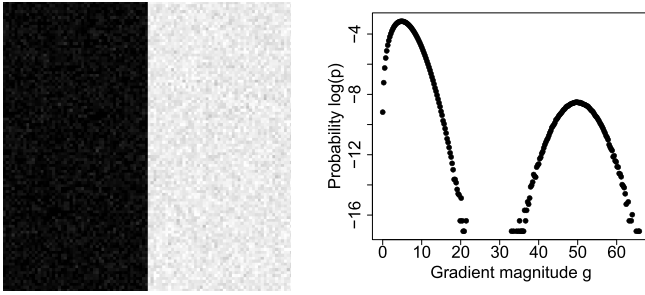
Fig. 2. Left: Two-class synthetic image with equally-sized regions of intensity $\mu_1 = 50$, $\mu_2 = 150$, with added normally distributed noise $\sigma_i = 5$. Right: Histogram of gradient magnitudes, computed by central differencing. A central $\chi_3(0, \sigma_g^2)$ distribution with $E\{g\} = 2\sigma_i/\sqrt{\pi} = 5.6$ corresponds to homogeneous regions. The non-central $\chi_1(\lambda, \sigma_g^2)$ distribution with $\lambda = \frac{1}{2}(\mu_2 - \mu_1) \approx 50$ arises from voxels at the one-dimensional boundary.

(d) If filter weights w_d depend on the direction, the resulting distribution of the gradient magnitudes corresponds to a linear combination of central χ-distributions $\chi_D(0, 2(w_d\sigma_i)^2)$.

B. Gradient Magnitudes at Noisy Boundaries

Next, consider an image with two regions R_1, R_2 with mean intensities μ_1, μ_2; $\mu_1 < \mu_2$, and the same noise level σ_i, separated by a boundary parallel to one image dimension (Fig. 2). Intensities in this image follow a mixture of two Gaussian distributions (Eq. 1), weighted by the region size. Gradient magnitudes determined within a region are χ_k-distributed, as above. However, gradient magnitudes at the boundary follow a *non-central* χ-distribution with k DOF and $\lambda = \mu_2 - \mu_1$, denoted as $\chi_k(\lambda, \sigma_g^2)$. The PDF for $k > 1$ is given by [3]:

$$g \sim \chi_k(\lambda, \sigma_g^2) = \frac{\lambda}{\sigma_g^2}\left(\frac{g}{\lambda}\right)^{k/2}\exp\left(-\frac{g^2 + \lambda^2}{2\sigma_g^2}\right)I_{\frac{k}{2}-1}\left(\frac{\lambda g}{\sigma_g^2}\right),$$
(4)

where $I_v(\cdot)$ denotes the modified Bessel function [1]. For computational stability, it is advantageous to use the equivalent form:

$$g \sim \frac{\lambda}{\sigma_g^2}\left(\frac{g}{\lambda}\right)^{k/2}\exp\left(-\frac{(g-\lambda)^2}{2\sigma_g^2}\right)I_{\frac{k}{2}-1}^*\left(\frac{\lambda g}{\sigma_g^2}\right),$$
(5)

where $I_v^*(\cdot)$ denotes the exponentially scaled modified Bessel function [1]. The mean of the non-central χ distribution is given as (Fig. 3):

$$E\{g\} = \sqrt{2\sigma_g^2}\,\frac{\Gamma(k/2 + 1/2)}{\Gamma(k/2)}\,M\left(-\frac{1}{2}, \frac{k}{2}, -\frac{k\lambda^2}{2\sigma_g^2}\right),$$
(6)

where $M(a, b, x)$ is the confluent hyper-geometric (Kummer's) function of the first kind. We note:

(a) If the non-centrality parameter λ is large (i.e., $\lambda \gg \sigma_g$), then (i) the mean is independent of the noise: $E\{g\} \approx \lambda\sqrt{k}$ (see Appendix), and (ii) $I_{\frac{k}{2}-1}^*\left(\frac{\lambda g}{\sigma_g}\right) \to 1$, i.e., the distribution $\chi_k(\lambda, \sigma_g^2)$ is approximated by a Gaussian distribution $\mathcal{N}(\lambda\sqrt{k}, \sigma_g^2)$.
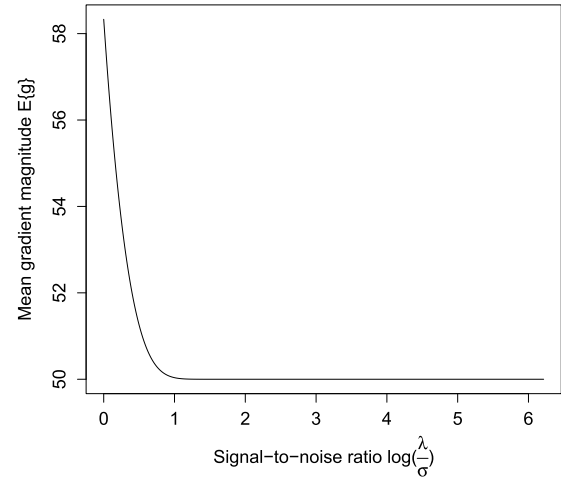

Fig. 3. Mean value of the gradient magnitude of a non-central $\chi_1(\lambda, \sigma_g^2)$ distribution for $\lambda = 50$ and various signal-to-noise ratios $\frac{\lambda}{\sigma_g}$. Note that $E\{g\} \approx \lambda\sqrt{k}$ for $\frac{\lambda}{\sigma_g} > 5$.
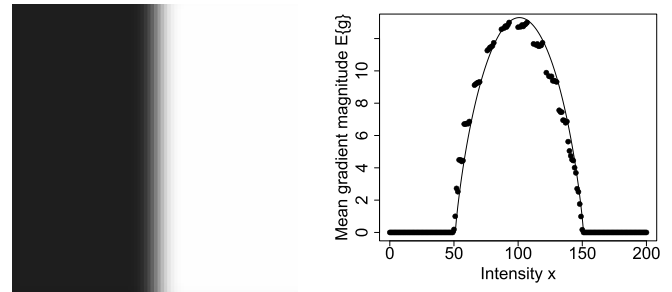

Fig. 4. Left: Two-class synthetic image with equally-sized regions of intensity $\mu_1 = 50$, $\mu_2 = 150$, convolved with a one-dimensional Gaussian filter of width $\tau = 3$. Right: Gradient magnitude g vs. intensity x. Dots correspond to the averaged gradient magnitude $E\{g\}$, computed by central differencing and sampled at a given intensity x in the image. Due to the spatial discretization, not all intensities are present in the image, so some intensity bins are empty. The continuous line shows the gradient magnitudes predicted by the $\text{arch}(\cdot)$ function (see Eq. 9). The maximum gradient magnitude is found at the mean of $(\mu_1, \mu_2) : g(\mu_1, \mu_2, \tau) \approx \frac{100}{\tau\sqrt{2\pi}} = 13.3$.

(b) Suppose we construct an image with a complex boundary by randomizing voxels in Fig. 2. The gradient filter yields equal proportions of directional components centered at $\Delta x = 0$ and $\Delta x = \frac{1}{2}(\mu_2 - \mu_1)$. The resulting distribution of gradient magnitudes follows a linear combination of χ_3 distributions with proportions $\frac{1}{8}\chi_3(0, \sigma_g^2)$, $\frac{3}{8}\chi_3(\frac{1}{2}(\mu_2 - \mu_1), \sigma_g^2)$, $\frac{3}{8}\chi_3(\frac{\sqrt{2}}{2}(\mu_2 - \mu_1), \sigma_g^2)$, and $\frac{1}{8}\chi_3(\frac{\sqrt{3}}{2}(\mu_2 - \mu_1), \sigma_g^2)$.

(c) Applying a more complex gradient magnitude filter (e.g., a Zucker-Hummel filter) will also result in a linear combination of χ_k distributions, where k corresponds to the number of directional terms of the gradient filter.

(d) It is possible to derive analytical solutions for linear combinations of non-central χ-distributions [22]. These solutions are complex, and coefficients depend on the interaction of the boundary and the gradient filter. As we will demonstrate later, it is not necessary to explicitly use an analytical solution for our intended application.

C. Gradient Magnitudes at Blurred Boundaries

Now, we study the influence on image blur, first for the noise-free case of the two-region image as shown in Fig. 4. We model image blur by a convolving the image I along
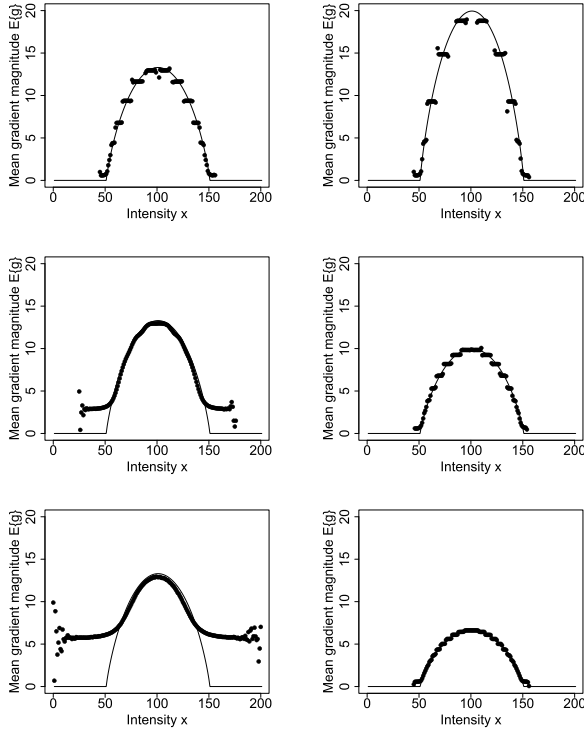
Fig. 5. Gradient magnitude g vs. intensity x computed in the image shown in Fig. 4. Dots correspond to the averaged gradient magnitude $E\{g\}$, computed by central differencing and sampled at a given intensity x in the image, and the continuous line to the prediction. Left column: Varying the noise level $\sigma_i = 1$ (top), $\sigma_i = 5$ (middle), and $\sigma_i = 10$ (below), at constant blur $\tau = 3$. Right column: Varying the blur $\tau_i = 2$ (top), $\tau_i = 4$ (middle), and $\tau_i = 6$ (below), at constant noise level $\sigma = 1$.

coordinate d and boundary at d_b with a one-dimensional Gaussian filter $f(\tau)$ [12]:

$$I_f(d_b, \mu_1, \mu_2, \tau) = I(d_b, \mu_1, \mu_2) \times f(\tau) \qquad (7)$$

$$= \mu_1 + (\mu_2 - \mu_1)\mathrm{erf}\left(\frac{d - d_b}{\tau\sqrt{2}}\right), \quad (8)$$

where $\mathrm{erf}(\cdot)$ corresponds to the error function [1]. The gradient magnitude g in relation to the intensity x along the transition from μ_1 to μ_2 is described by:

$$g(x|\mu_1, \mu_2, \tau) = \frac{\mu_2 - \mu_1}{\tau\sqrt{2\pi}} \mathrm{arch}\left(\frac{x - \mu_1}{\mu_2 - \mu_1}\right) \qquad (9)$$

$$\text{with } \mathrm{arch}(x) = \exp\left[-\{\mathrm{erf}^{-1}(2x - 1)\}^2\right]. \qquad (10)$$

For a fast approximation to $\mathrm{erf}^{-1}(\cdot)$, refer to [40]. Note that the gradient magnitude is inversely proportional to the width of the blurring filter: $g \propto \frac{1}{\tau}$. Adding Gaussian-distributed noise to the image in Fig. 4, we recognize that gradient magnitudes at a given intensity x are now distributed according to a non-central $\chi(\lambda, \sigma_g^2)$ distribution, with means $E\{g\}$ according to Eq. 6 (see Fig. 5). Note that the prediction by Eq. 9 does not depend on the noise level, but on the blur τ only. For a reasonably high SNR, the approximation in Eq. 13 (Appendix) still holds, i.e., $E\{g\} \approx \lambda\sqrt{k}$. Adding noise followed by blurring introduces spatial correlations in the noise, invalidating the image model from Section II-A. However, we find that Eq. 13 is still valid in the presence of correlated noise.
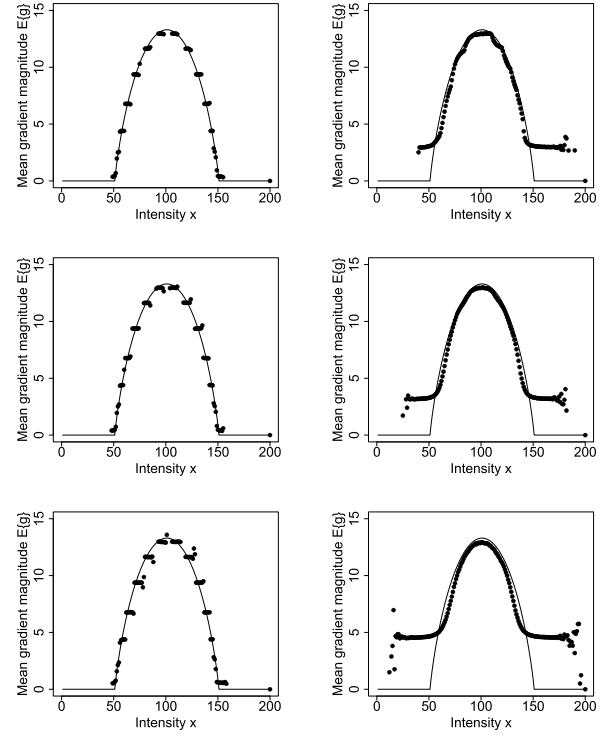

Fig. 6. Gradient magnitude g vs. intensity x computed in the image shown in Fig. 4. Dots correspond to the averaged gradient magnitude $E\{g\}$, computed by central differencing and sampled at a given intensity x in the image, and the continuous line to the prediction. Images were corrupted by χ_2-distributed (Rician) noise (top row), χ_{16}-distributed noise (middle row), and Poisson noise (below), with var $= 1$ (left column) and var $= 64$ (right column).

D. Non-Gaussian Noise Distributions

Noise in reconstructed MR images follows a $\chi_n(0, \sigma^2)$ distribution, where n corresponds to twice the number of receiver coils. Common examples include the case of a single receiver coil ($n = 2$, Rician noise), and an 8-channel head coil ($n = 16$). A similar, albeit more involved theoretical derivation can be made for images corrupted by this noise distribution. We provide a simulation experiment to compare the prediction according to Eq. 9 in images corrupted by χ_2-, χ_{16}-, and Poisson-distributed noise. In Fig. 6, we compiled simulation results for these three distributions, for a low noise (var $= 1$, left column) and high noise (var $= 64$, right column) scenario. We conclude that the prediction according to Eq. 9 still holds. Because these distributions have a non-zero mean, μ_1 and μ_2 above correspond to the sample means in these regions, not the un-corrupted intensities.

E. Image Acuity Measure

Now, the theoretical framework is assembled to define our image acuity measure. From the simulation above, remember that the gradient magnitude at a boundary is proportional to the width of the blurring filter, i.e., $g \propto \frac{1}{\tau}$. Thus, we propose to use the ratio of the peak gradient magnitude and the intensity difference at a boundary as the dimensionless acuity measure $A(\cdot)$:

$$A(\mu_1, \mu_2) = \frac{E\{g\left[\frac{1}{2}(\mu_2 + \mu_1)\right]\}}{\mu_2 - \mu_1}, \quad \mu_1 < \mu_2. \qquad (11)$$

Again, let us make some remarks:

(a) Note that $A \to \infty$ if the image blur $\tau \to 0$, reflective of infinite acuity. Along such a boundary in a discretized image, only voxels with pure intensities μ_1 or μ_2 are found. Let us denote such a boundary as "sharp".

(b) The acuity measure is (largely) independent of the noise level, the absolute image intensity and contrast. Thus, acuity can be used as a complementary measure of image quality, independent of SNR and CNR measures.

(c) Note that image noise is often estimated from a gradient magnitude image, as indicated in Section II-B. For the estimation of SNR and CNR, the intensity of homogeneous regions must be determined (e.g., from intensities of local regions with low gradient magnitudes). If this information is available, acuity can readily be computed according to Eq. 11.

(d) For one-directional boundaries and Gaussian blur, the width τ of the Gaussian point spread function can be estimated from the proportionality factor (see Eq. 9). We study this relationship for more complex images, gradient filters, and other blur in the next section.

(e) Noise in reconstructed MR images follows a $\chi_n(0, \sigma^2)$ distribution, where n corresponds to twice the number of receiver coils. We demonstrated in the experiment of Section II-D that differences in the acuity measure between images corrupted by different noise distributions at the same variance are negligible.

(f) For other types of image blur, e.g. due to directed motion, the Gaussian convolution model introduced in the previous section does not hold. For real images such as studied in the next section, intensity and gradient magnitudes still follow an arch-like function with a maximum gradient magnitude at the mean of both pure intensities.

(g) The "modulation transfer function" (MTF) is used to quantify properties of an imaging process, and is defined as the quotient (frequency-dependent) image contrast and object contrast [6], similar to this definition of acuity. The "subjective quality factor" is obtained by integrating the MTF over a specified frequency range, and is used as a quantitative property of imaging systems that correlates with human visual perception [16]. Both MTF and acuity aim at quantifying image degradation using image properties at edges. However, acuity is defined here as a post-hoc measure for which the true object contrast is not available, at least in the example use cases in medical imaging discussed here.

III. Experiments and Results

The analytic framework developed above is first studied in simulated data and then applied to real images to demonstrate the use of the proposed acuity measure.

A. Relationship of Acuity and Blur

First, we used an image with a one-directional boundary (see Fig. 4), simulated blur by convolution with a Gaussian filter of width τ, and added Gaussian-distributed noise of level σ_i. As shown in Fig. 7 left, the proportionality factor $\sqrt{2\pi}$ is approximately found for a considerable range of image blur and noise. The deviance for low image blur is due to the approximation error of the gradient magnitude on the
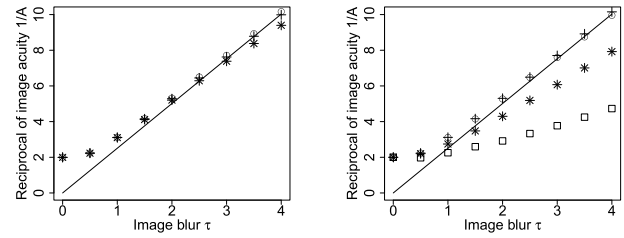

Fig. 7. Left: Relationship of image blur τ and the reciprocal of image acuity $1/A$ for different noise levels ($\sigma_i = 0$: circle; 2: plus, 4: star). Right: Relationship of image blur τ and the reciprocal of image acuity $1/A$ for different gradient magnitude filters (central gradient magnitude: circle; Zucker-Hummel: plus, Gaussian derivative, $\sigma = 1.0$: star, Gaussian derivative, $\sigma = 2.0$: square).
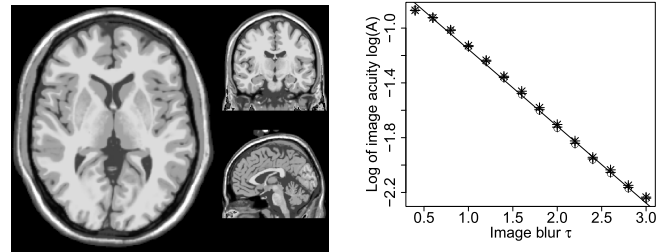

Fig. 8. Left: Orthogonal sections of the brain web phantom [7]. The gray and white matter compartments were selected for the following simulation. Right: Relationship of image blur τ and the logarithm of image acuity $\log A$ for a Zucker-Hummel filter at noise levels of $\sigma_i = \{0, 2, 4, 6, 8, 10\}$, for Gaussian noise: circle, χ_2-distributed noise: plus, and χ_{18}-distributed noise: star.

discretized spatial grid, the deviance for high blur/high noise due to the approximation error on the mean of the gradient magnitude for low SNR. Next, we tested the influence of the type of gradient magnitude filter on the proportionality in Eq. 9 (Fig. 7, right) at a constant noise level $\sigma_i = 2.0$. We conclude that the amount of (Gaussian) image blur can be estimated if simple gradient magnitude filters are used. The application of more complex edge filters (with a higher stability against noise) is not required.

Next, we used the brain web phantom [7] as an example for a realistic image with complex boundaries. We assume that the provided image is free of noise and "sharp" as defined above. Image acuity was determined for regions corresponding to the gray and white matter of the human brain, with mean intensities $\mu_{GM} = 110$, $\mu_{WM} = 150$, and an isotropic spatial resolution of 1 mm. As examples of image deterioration, the provided image was convolved by a Gaussian filter of different widths, simulating an increase in image blur. Gaussian-, χ_2-, χ_{18}- and Poission-distributed noise was added, corresponding to a SNR between 13 and 65. Image acuity was determined by the procedure described in Section II-E.

Empirically, we found a log-linear relationship of the form $\log(A) \approx -0.568\tau - 0.563$ in the interval of $\tau \in [0.6, 3.0]$. The difference between data points corresponding to different noise levels were too small for visualization, so the average is shown in Fig. 8. We understand this relationship as arising from a linear combination of arch-like functions (Eq. 9) with higher-order terms τ, τ^2, τ^3 in the denominator, due to blurring at one-, two-, and three-directional boundaries.
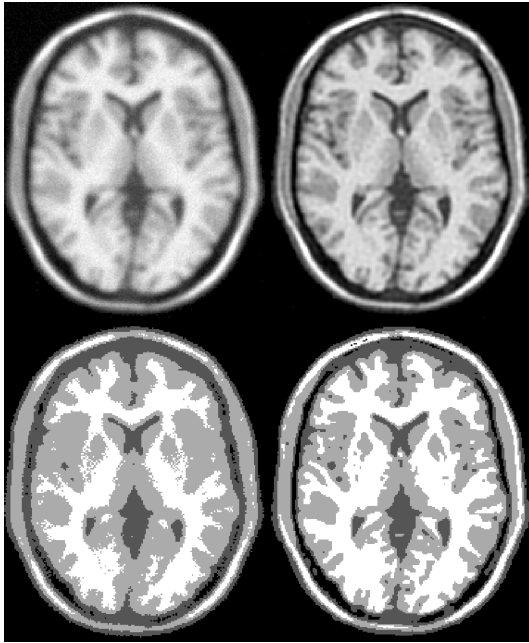
Fig. 9. Top left: Axial slice of the brain web phantom [7], corrupted by image blur $\tau = 1.5$ and Gaussian noise $\sigma = 4$. Top right: Image restored by deconvolution. Below left: Segmentation of the corrupted image into three classes (NMI = 0.548). Below right: Segmentation of the restored image (NMI = 0.732).

Fig. 10. Left: Acuity of the phantom, corrupted by image blur τ and Gaussian noise (low noise: square, high noise: plus) and restored by deconvolution (low noise: circle, high noise: star). Right: Similarity of a three-class segmentation with reference segmentation for the corrupted images (low noise: square, high noise: plus) and restored images (low noise: circle, high noise: star).

Although coefficients of these models depend on the configuration of the boundaries, i.e., the imaging content, we have deliberately chosen the simple two-parameter model above. The relationship between blur and acuity is (almost) independent of the noise level. Given an example image that is sufficiently close in image content (e.g., the similar MR image of the human head), the equation above can be used to estimate the amount of blur in a real image of the same spatial resolution.

Finally, we wanted to assess the use of acuity in image restoration by deconvolution [9]. The brain web phantom was corrupted by image blur $\tau_i = \{0, 0.5, 1, 1.5, 2, 2.5, 3\}$ and Gaussian noise $\sigma_i = \{2, 4\}$ (Fig. 9, top left), then restored (Fig. 9, top right). Corrupted and restored images were segmented into three classes based on a Gaussian mixture model, roughly corresponding to 1: Cerebro-spinal fluid and skull; 2: Gray matter and connective tissue; 3: White matter and fat. Example results are shown in Fig. 9, for a corrupted image (bottom left) and its restored version (bottom right). Label confusion matrices were computed w.r.t. the reference segmentation of the un-corrupted phantom. To rate the restoration quality, we used the normalized mutual information (NMI) of the confusion matrix as a similarity metric, which ranges from 0 to 1 (perfect correspondence). The improvement in acuity and similarity are shown in Fig. 10 for different levels of noise and image blur. At each iteration of the restoration process, measures acuity, SNR, CNR, and the segmentation similarity were determined. Most interestingly, the quotient of acuity and similarity reached a constant level after 15-20 iterations, indicating that acuity can be used as a convergence criterion of the restoration process if the ground truth is unknown. The computation of these quality measures adds a computational
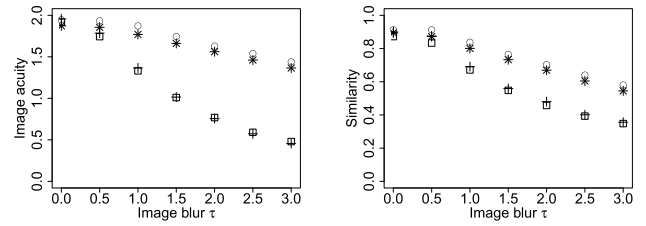
overhead of about 20% per iteration, and is largely due to the fact that a segmentation must be computed in order to determine the class-wise intensities and gradient magnitudes to estimate the noise level. The additional time required to compute acuity is negligible. It must be added that the restoration quality (according to the similarity metric) depends much on the setting of the damping factor. Thus, acuity cannot be used to optimize this parameter.

B. Application Examples

In the following, we demonstrate the use of acuity in three application examples that were described in the introduction.

1) Detection of Motion-Corrupted Images in Large-Scale Data Bases: A straightforward application of the acuity measure is to detect images corrupted by motion blur. A typical scenario involves the analysis of a pre-clinical neuro-imaging study where individual anatomical head data sets were acquired to be used as an anatomical reference. As discussed in the introduction, subject motion cannot always be avoided (refer to Fig. 1 for an example). We consider an initial quality check as helpful before later stages of an image processing chain fail or produce erroneous results.

We selected a data base of head images selected from a study that aims at finding possible differences in brain macro-anatomy in young subjects with autism spectrum disorder (ASD). We received 20 structural MRI examinations of subjects with a behavioral diagnosis of ASD (5 female, 15 male, age 12–20), and an age-matched control group of 10 normally developing children (1 female, 9 male, age 12–18). T_1-weighted MR images were acquired on a Philips Achieva 3T scanner, equipped with an 8-channel phased array coil, using a TFE sequence with TR 11 ms, TE 3.7 ms, flip angle 18 degrees, 150 sagittal slices with a matrix of 240×240 voxels, corresponding to an isotropic resolution of 1.0 mm. Images were analyzed by the procedure described above (Section II-E) to obtain measures of image acuity, SNR, and CNR, based on the white and gray matter compartments of the brain (Fig. 11). These quality measures are helpful to identify data sets that may be considered for exclusion (Fig. 1). In the whole sample, SNR (resp. CNR) are found to be statistically independent of acuity, conforming our initial notion of a complementary measure. A principal component analysis of the measures can be used to map measures onto a single dimension that can
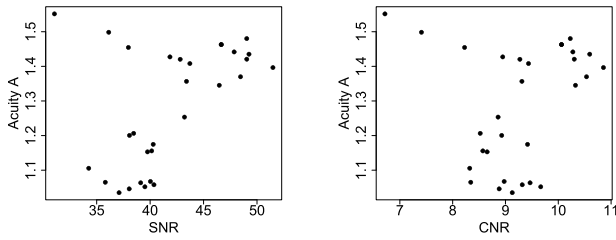
Fig. 11. Plots of acuity vs signal-to-noise ratio (SNR, left) and contrast-to-noise ratio (CNR, right) for all data sets of the study. Acuity is independent of SNR (Pearson correlation, $p = 0.11$) and CNR ($p = 0.43$).

Fig. 13. Left: Axial section of the optimal template. Right: Relationship of acuity A and the number of images n in the template set T.
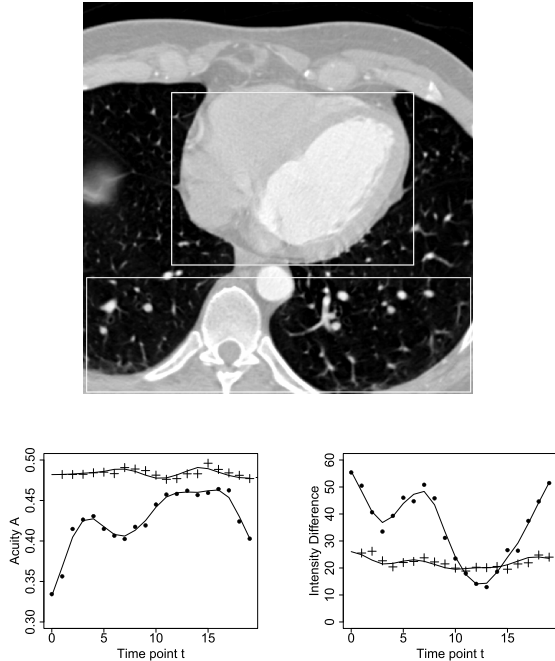

Fig. 12. Top: Example slice from the dynamic CT heart study. Boxes mark the heart and chest region. Below left: Acuity vs. acquisition time point for the heart region (dots) and the spine region (crosses). Note that the heart region has a lower acuity and a considerable dependency on the heart cycle. Below right: Intensity difference between consecutive time points vs. acquisition time point for the heart region (dots) and the spine region (crosses). The intensity difference is used as a simple detector for motion between consecutive time points. Similar to the panel on the left, the heart region shows a considerable motion during the cycle, while the spine regions remains mostly stationary.

loosely be denoted as "image quality". We suggest to visually inspect data sets in one tail to decide if a specific case should be excluded from further evaluation.

2) Differentiation of Moving and Stationary Boundaries in an Image: We received a dynamic CT study of the human heart, obtained from a publicly available data base [35]. ECG-gated CT images were acquired in a healthy volunteer at 20 time steps throughout the R-R cycle on a Siemens Syngo scanner (matrix 512×512 voxels, 66 slices, 0.43×0.43 mm in-plane resolution, 3 mm slice thickness). We selected a region enclosing the heart (which is presumably moving throughout the cycle) and a region enclosing the spine and posterior chest (presumably stationary; Fig. 12, top). We measured the acuity at the border between soft tissue and air in both regions and at all time points (Fig. 12, below left).
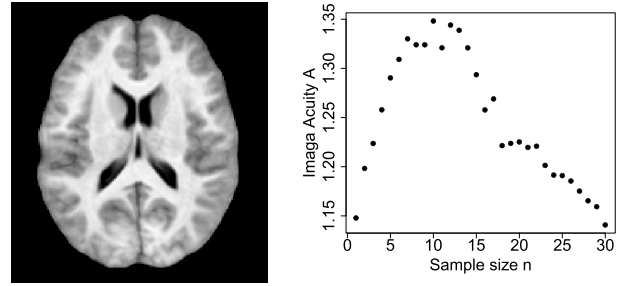
In contrast to the spine region, the acuity in the heart region is shows a dependency on the heart cycle. We interpret this lower acuity as due to the higher motion-related within-scan blur at the tissue-air boundary. As a simple verification, we computed difference images of consecutive time points, and determined the variance within the ROIs, as a simple surrogate of motion. We found that the variance in the heart ROI shows a considerable dependency on the cycle while the variance in the spine ROI is mostly independent of the cycle, but not always lower than the heart (Fig. 12, below right). From the difference images, we understand that lung nodules in the spine ROI contribute to the detected motion. Thus, the acuity measure can be used to differentiate stationary and moving regions in an image. Note that our motion estimation is based on the assessment of motion-related blur at single time points.

3) Sampling of a Common Anatomical Template: In the context of voxel or tensor-based morphometry (VBM, TBM), it is common practice to develop a study-specific anatomical template to which all data sets are registered before statistical analysis. There has been some discussion about the optimal number and selection of data sets to be included in the formation of the template [5], [38]. We propose to use the set of images that yields the highest acuity of the template. A simple greedy approach was used to accomplish this task. We formed two sets of images, the set of candidate images C (initially, $n = 30$), and the set of template images T (initially, $n = 0$). The image with the highest acuity was selected from C and moved to T. Next, we selected the image from C that yielded the template with the highest acuity when combined with T. This procedure was repeated until C was empty. The set of images in T with the highest overall acuity was used to generate the template.

We used the children's data base described above, and registered images via the nonlinear diffeomorphic demon approach, with cross-correlation as image similarity metric [4]. We found a maximum around $n = 10$ data sets with a maximal acuity of 1.35 (Fig. 13). Note that the combined set had a considerably higher acuity than any single data set, while the acuity of the maximal set $n = 30$ was much lower and similar to the best image alone.

IV. DISCUSSION

The statistical properties of intensities and gradient magnitudes at region boundaries in the presence of noise and image

blur were examined and discussed in this work. A simple measure of image acuity (Eq. 11), based on the ratio of the maximal gradient magnitude and intensity difference at a boundary, was established and applied to both simulated and non-simulated data. Assuming that image blur can be modeled by convolution with a Gaussian filter of width τ, the filter width can be estimated from the acuity measure by a log-linear relationship (Fig. 8), within a range of image noise and blur typically found in medical images. Three use cases demonstrated applications of this measure, the detection of motion corrupted data sets, the detection of moving regions in an image, and the selection of the best set of images in the generation of a common anatomical template. Further application areas for this acuity measure are conceivable but were not tested here:

- Optimizing the efficiency of procedures for within-scan motion correction (e.g., prospective motion correction, cardiac gating).
- Assessing and optimizing the spatial resolution of statistical image reconstruction algorithms (e.g., in PET imaging).
- Assessing and comparing the spatial resolution of non-linear image registration procedures.
- Employing the statistical context developed here for segmentation to better resolve mixtures in partial-volume voxels.

We emphasize that our acuity measure can be determined *post-hoc*, i.e., a phantom with known properties is not needed. As discussed above, we assume that the imaged object has a sharp boundary (with respect to the spatial and temporal image resolution) between two regions of (approximately) homogeneous intensity. This condition is often fulfilled in medical images, at least in sub-sections of a scene. We are aware of a few caveats in the application of the acuity measure that are discussed in the following.

A. Multiple Neighboring Regions

Suppose we have three regions R_1, R_2, R_3 with mean intensities $\mu_1 < \mu_2 < \mu_3$ in an image with boundaries $B_{1,2}, B_{1,3}, B_{2,3}$. The gradient magnitude distributions from voxels at boundaries $B_{1,2}$ and $B_{2,3}$ do not overlap, however, the distributions at $B_{1,2}, B_{1,3}$ and $B_{2,3}, B_{1,3}$ may overlap if the intensity differences are small and/or the image blur and noise are large. Thus, we suggest to select a sub-image with just two regions. As an alternative, the tracing procedure described for accumulating LH-histograms [39] can be used to discriminate voxels due to different boundaries.

B. Non-Uniform Regions

The presence of (large) intensity inhomogeneities in the image invalidates assumptions made above. Thus, we suggest to either correct inhomogeneities first or select an (approximately) homogeneous sub-image. Likewise, the presence of texture in a region is not in accordance with our simple image model. For small-scale texture with small intensity variations (e.g., as in anatomical MR images of the brain's white matter), texture will be accounted for as an increase in noise in our

model. Regions with strong texture violate the assumption of a uniform region intensity, so that acuity cannot be estimated in this image area based on our approach.

C. Blur Model

We assumed that image blur can be modeled by a convolution with a Gaussian filter. This assumption approximately holds for a variety of causes that influence image acuity, e.g., filtering, interpolation, (non-linear) registration, and small-scale motion. Images corrupted by large, directed motion or ghosting may not have a lower acuity as defined here.

D. Correlated Noise

Our statistical image model considers the case of additive noise that is uncorrelated in space. While this simplification is often made in medical image analysis, it is important to note that either properties of the imaged object (e.g., texture), the image acquisition and reconstruction method (e.g., CT tomographic reconstruction), or any post-processing (e.g., spatial filtering) violate this assumption. Considering the case that two neighboring region may have different spatial correlation statistics, it may not be feasible to derive closed-form description for the distribution of intensity and gradients in the vicinity of a boundary. Instead, Monte-Carlo simulations could be used to study the influence of correlated noise on acuity [28]. In a cursory repetition of the experiment in Section II-C, we added noise first, then blurred the image, thus, introducing spatial correlation in the noise. Here, we found that the arch-shaped relationship is still retained.

Our image acuity measure is simple to understand and straightforward to implement. We demonstrated applications in medical imaging where our acuity measure is useful in the objective and automatic assessment of image quality.

APPENDIX
MEAN OF GRADIENT MAGNITUDE

We were interested in approximating the mean $E\{g\}$ of the gradient magnitude for $\lambda \gg \sigma$, e.g., if the intensity difference at an edge is much larger than the noise. Recall that:

$$E\{g\} = \sqrt{2\sigma^2} \frac{\Gamma(k/2 + 1/2)}{\Gamma(k/2)} M\left(-\frac{1}{2}, \frac{k}{2}, -\frac{k\lambda^2}{2\sigma^2}\right). \quad (12)$$

We used Kummer's transformation $M(a, b, z) = e^z M(b - a, b, -z)$ [1] and identify $a = k/2 + 1/2$, $b = k/2$, and $z = (k\lambda^2)/(2\sigma^2)$. For $a > b$ and $z \gg 0$, Kummer's function is approximated by $M(a, b, z) \approx (\Gamma(b)/\Gamma(a)) e^z z^{a-b}$ [1]. Insertion yields:

$$E\{g\} \approx \sqrt{2\sigma^2} \frac{\Gamma(a)}{\Gamma(b)} e^{-z} \frac{\Gamma(b)}{\Gamma(a)} e^z z^{a-b} \quad (13)$$

$$= \sqrt{2\sigma^2} \left(\frac{k\lambda^2}{2\sigma^2}\right)^{1/2} = \lambda\sqrt{k}. \quad (14)$$

Thus, the mean is independent of the noise level, and proportional to the intensity difference at the edge, given the dimensionality of the image.

REFERENCES

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York, NY, USA: Dover, 1972.

[2] S. Aja-Fernández and G. Vegas-Sánchez-Ferrero, *Statistical Analysis of Noise in MRI*. New York, NY, USA: Springer, 2016.

[3] D. A. Anderson, "Maximum likelihood estimation in the non-central chi distribution with unknown scale parameter," *Indian J. Stat. B*, vol. 43, no. 1, pp. 58–67, 1981.

[4] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain," *Med. Image Anal.*, vol. 12, no. 1, pp. 26–41, 2008.

[5] P. Avants *et al.*, "The optimal template effect in hippocampus studies of diseased populations," *NeuroImage*, vol. 49, pp. 2457–2466, Feb. 2010.

[6] G. D. Boreman, *Modulation Transfer Function in Optical and Electro-Optical Systems*. Bellingham, WA, USA: SPIE, 2001.

[7] *Brain Web Phantom*. Accessed: Jul. 1, 2018. [Online]. Available: http://www.bic.mni.mcgill.ca/brainweb/

[8] E. Caruyer, I. Aganj, C. Lenglet, G. Sapiro, and R. Deriche, "Motion detection in diffusion MRI via online ODF estimation," *Int. J. Biomed. Imag.*, vol. 2013, Jan. 2013, Art. no. 849363, doi: 10.1155/2013/849363.

[9] S. H. Chan, R. Khoshabeh, K. B. Gibson, P. E. Gill, and T. Q. Nguyen, "An augmented Lagrangian method for total variation video restoration," *IEEE Trans. Image Process.*, vol. 20, no. 11, pp. 3097–3111, Nov. 2011.

[10] J. P. Chiverton, "Probabilistic partial volume modelling of biomedical tomographic image data," Ph.D. dissertation, School Electron. Phys. Sci., Univ. Surrey, Guildford, U.K., 2006.

[11] L. S. Chow and R. Paramesran, "Review of medical image quality assessment," *Biomed. Signal Process. Control*, vol. 27, no. 1, pp. 145–154, 2016.

[12] J. H. Elder and S. W. Zucker, "Local scale control for edge detection and blur estimation," in *Computer Vision—ECCV*, 1996, pp. 55–69.

[13] A. C. Evans, A. L. Janke, D. L. Collins, and S. Baillet, "Brain templates and atlases," *NeuroImage*, vol. 62, pp. 911–922, Aug. 2012.

[14] T. A. Fuchs *et al.*, "Impact of a new motion-correction algorithm on image quality of low-dose coronary CT angiography in patients with insufficient heart rate control," *Acad. Radiol.*, vol. 21, pp. 312–317, Mar. 2014.

[15] S. Fürst *et al.*, "Motion correction strategies for integrated PET/MR," *J. Nucl. Med.*, vol. 56, pp. 261–269, Feb. 2015.

[16] E. M. Granger and K. N. Cupery, "An optical merit function which correlates with subjective image judgments," *Photogr. Sci. Eng.*, vol. 16, no. 3, p. 221–230, 1972.

[17] H. Gudbjartsson and S. Patz, "The Rician distribution of noisy MRI data," *Mag. Reson. Med.*, vol. 34, no. 6, pp. 910–914, 1995.

[18] N. Gupta, A. P. Shukla, and S. Agarwal, "Despeckling of medical ultrasound images: A technical review," *Int. J. Inf. Eng. Electron. Bus.*, vol. 8, pp. 11–19, May 2016.

[19] T. Ihalainen, D. Sipilä, and S. Savolainen, "MRI quality control: Six imagers studied using eleven unified image quality parameters," *Eur. Radiol.*, vol. 14, no. 10, pp. 1859–1865, 2004.

[20] P. Irrera, I. Bloch, and M. Delplanque, "A flexible patch based approach for combined denoising and contrast enhancement of digital X-ray images," *Med. Image Anal.*, vol. 28, pp. 33–45, Feb. 2016.

[21] V. Kayargadde and J. B. Martens, "Estimation of edge parameters and image blur using polynomial transforms," *Graph. Models Image Process.*, vol. 56, pp. 442–461, Nov. 1994.

[22] G. K. Karagiannidis and S. A. Kotsopoulos, "On the distribution of the weighted sum of L independent Rician and Nakagami envelopes in the presence of AWGN," *J. Commun. Netw.*, vol. 3, no. 2, pp. 1–8, Jun. 2001.

[23] G. Kindlmann and J. W. Durkin, "Semi-automatic generation of transfer functions for direct volume rendering," in *Proc. IEEE Symp. Vol. Vis.*, Oct. 1998, pp. 79–86.

[24] C. G. Koay, E. Özarslan, and C. Pierpaoli, "Probabilistic identification and estimation of noise (PIESNO): A self-consistent approach and its applications in MRI," *J. Magn. Reson.*, vol. 199, no. 1, pp. 94–103, 2009.

[25] C. Kolbitsch, C. Prieto, C. Tsoumpas, and T. Schaeffter, "A 3D MR-acquisition scheme for nonrigid bulk motion correction in simultaneous PET-MR," *Med. Phys.*, vol. 41, no. 8, p. 082304, 2014, doi: 10.1118/1.4890095.

[26] F. Kruggel, J. Turner, and L. T. Muftuler, "Impact of scanner hardware and imaging protocol on image quality and compartment volume precision in the ADNI cohort," *NeuroImage,* vol. 49, no. 3, pp. 2123–2133, 2010.

[27] T. Kästner *et al.*, "Self-navigated 4D Cartesian imaging of periodic motion in the body trunk using partial k-space compressed sensing," *Magn. Reson. Med.*, vol. 78, no. 2, pp. 632–644, 2017.

[28] B. A. Landman, P.-L. Bazin, and J.-L. Prince, "Estimation and application of spatially variable noise fields in diffusion tensor imaging," *Magn. Reson. Imag.*, vol. 27, no. 6, pp. 741–751, 2009.

[29] J. Maclaren, M. Herbst, O. Speck, and M. Zaitsev, "Prospective motion correction in brain imaging: A review," *Magn. Reson. Med.*, vol. 69, no. 3, pp. 621–636, 2013.

[30] C. Malamateniou *et al.*, "Motion-compensation techniques in neonatal and fetal MR imaging," *Amer. J. Neuroradiol.*, vol. 34, no. 6, pp. 1124–1136, 2013.

[31] J. R. McClelland, D. J. Hawkes, T. Schaeffter, and A. P. King, "Respiratory motion models: A review," *Med. Imag. Anal.*, vol. 17, no. 1, pp. 19–42, 2013.

[32] A. Mechelli, C. J. Price, K. J. Friston, and J. Ashburner, "Voxel-based morphometry of the human brain: Methods and applications," *Current Med. Imag. Rev.*, vol. 1, no. 2, pp. 105–113, 2005.

[33] D. P. Panda and A. Rosenfeld, "Image segmentation by pixel classification in (grey level, edge value) space," *IEEE Trans. Comput.*, vol. C-27, no. 9, pp. 875–879, Sep. 1978.

[34] N. Pannetier *et al.*, "Quantitative framework for prospective motion correction evaluation," *Magn. Reson. Med.*, vol. 75, no. 2, pp. 810–816, 2016.

[35] Pixmeo SARL. *Cardiac-CT With Systolic and Diastolic Images—Diastolix*. Accessed: Jul. 1, 2018. [Online]. Available: http://www.osirix-viewer.com/datasets

[36] A. Rahmim and J. Tang, "Noise propagation in resolution modeled PET imaging and its impact on detectability," *Phys. Med. Biol.*, vol. 58, no. 19, pp. 6945–6954, 2013.

[37] J. C. Russ, *The Image Processing Handbook*, 6th ed. Baltimore, MD, USA: CRC Press, 2011.

[38] M. L. Senjem, J. L. Gunter, M. M. Shiung, R. C. Petersen, and C. Jack, Jr., "Comparison of different methodological implementations of voxel-based morphometry in neurodegenerative disease," *NeuroImage*, vol. 26, no. 2, pp. 600–608, 2005.

[39] I. W. O. Serlie, F. M. Vos, R. Truyen, F. H. Post, and L. J. V. Vliet, "Classifying CT image data into material fractions by a scale and rotation invariant edge model," *IEEE Trans. Image Process.*, vol. 16, no. 12, pp. 2891–2904, Dec. 2007.

[40] A. Soranzo and E. Epure, "Very simply explicitly invertible approximations of normal cumulative and normal quantile function," *App. Math. Sci.*, vol. 8, no. 87, pp. 4323–4341, 2014.

[41] M. K. Stam *et al.*, "Navigators for motion detection during real-time MRI-guided radiotherapy," *Phys. Med. Biol.*, vol. 57, no. 21, pp. 6797–6805, 2012.

[42] V. Tavakoli and A. A. Amini, "A survey of shaped-based registration and segmentation techniques for cardiac images," *Comput. Vis. Image Understand.*, vol. 117, no. 9, pp. 966–989, 2013.

[43] M. A. Viergever, J. B. A. Maintz, S. Klein, K. Murphy, M. Staring, and J. P. W. Pluim, "A survey of medical image registration—Under review," *Med. Imag. Anal.*, vol. 33, pp. 140–144, Oct. 2016.

Frithjof Kruggel received the Diploma degree in chemistry and the M.D. degree in medicine from Ruhr University, Bochum, Germany, in 1897 and 1983, respectively, and the Ph.D. degree from Ludwig Maximilian University, Munich, Germany, in 1987. He received his clinical training in neurology at the Bogenhausen City Hospital and Klinikum Rechts der Isar, Munich, until 1995. He headed the Workgroup on Signal and Image Processing, Max Planck Institute on Cognitive Neuroscience, Leipzig, Germany, from 1995 to 2005. Since 2005, he has been a Full Professor with the Department of Biomedical Engineering, University of California, Irvine, USA. He has authored over 150 journal and conference papers. His research interests include medical image and signal processing with applications in neurobiology.