Generalized Ball-Scale: Theory, Algorithms, and Application to Image Inhomogeneity Correction

Anant Madabhushi[a], Jayaram K. Udupa[a], and Andre Souza[a]

[a]Medical Image Processing Group, Department of Radiology
University of Pennsylvania, Philadelphia, PA 19104.

ABSTRACT

The concept of generalized scale (g-scale) was introduced previously to overcome the shape, topological, and anisotropic constraints imposed by previous local morphometric scale models. Roughly speaking, the g-scale of a voxel in a scene was defined as the largest set of all voxels associated with it, that satisfy some homogeneity criterion. g-scale was shown to have interesting theoretical properties, and its superiority to an existing image background inhomogeneity correction method was demonstrated. In this paper, we present a variant of g-scale that we refer to as g_B-scale. The difference between g- and g_B-scale is that, while for g-scale, individual voxels are included into the g-scale set one at a time, the g_B-scale set is grown by including hyperballs, the hyperball corresponding to the local ball scale at every voxel (which, briefly, is the radius of the largest hyperball of homogeneous intensity centered at the voxel). The g_B-scale model was found to be more resistant to severe levels of inhomogeneity and noise compared to g-scale. A methodology to perform image background inhomogeneity correction based on the idea of g_B-scale was qualitatively and quantitatively compared on nearly 250 clinical and phantom datasets, with a ball-scale- and a g-scale-based correction methodology. For scenes containing inhomogeneity but no noise, the g-, and g_B-scale methods performed comparably and were superior to the ball-scale method. For scenes containing both noise and inhomogeneity, the g_B-scale-based method outperformed both the g- and ball-scale correction methods.

Keywords: Scale, generalized scale, noise, inhomogeneity correction, MRI, Image analysis, Image processing.

1. INTRODUCTION

Scale is a fundamental concept useful in almost all image processing and analysis tasks including segmentation, filtering, interpolation, registration, visualization, and quantitative analysis. A big challenge in all of these image processing tasks is to get a handle on the spatially varying level of image detail. The notion of scale emerged in image processing to tailor the processing of the image to local object size. Broadly speaking, scale concepts can be categorized into (i) *multi-scale* or *scale-space* representations, (ii) *local scale* models, and (iii) *locally adaptive scale* models.

Motivated by the multi-scale nature of real-world images, early attempts at handling scale focused on developing a *multi-scale representation* of images which included scale-space[1] and pyramidal representations.[2] Traditional scale-space theory does not, however, directly address the problem of selecting appropriate scales and image structures from the multi-scale representation for further analysis.[3] Typically, an object contains structures at many different scales, and it is not usually clear which scale is appropriate for each location in the image. An important problem in computer vision and image processing has been to find such locally stable scales, or rather to generate hypotheses about interesting local structures for further processing, without prior information about what the image can be expected to contain.

Consequently, there have been several different *local scale* models proposed, geared toward specific image processing applications.[3] The core idea behind *local scale* has been to extract meaningful scales from a multi-scale representation of the image. Most local scale estimation methods have however relied on some image smoothing, which could cause blurring of the image boundaries. Further, local scale models do not explicitly deal with image structure.

Correspondence to Jayaram K. Udupa: E-mail: {anantm, jay, andre}@mipg.upenn.edu, Telephone: (215) 662-6780, Fax:(215) 898-9145. This work was supported by DHHS grant NS37172.

Medical Imaging 2005: Image Processing, edited by J. Michael Fitzpatrick, Joseph M. Reinhardt, Proc. of SPIE Vol. 5747 (SPIE, Bellingham, WA, 2005)
1605-7422/05/$15 · doi: 10.1117/12.595360

Owing to some of the limitations of local scale models, recently, there has been interest in the concept of *locally adaptive scale*.[4,5] In this paradigm, the size of the scale is allowed to change in different parts of the image. By definition, locally adaptive scale in regions with fine details or in the vicinity of boundaries is small, while it is large in the interior of a large homogeneous object. This notion of scale differs from the spirit in which it is used in scale-space literature, where it typically refers to the size of the σ parameter in the Gaussian kernel and its derivatives. It is also different from the local scale idea in that a homogeneity-based region continuity criterion is used in its definition, and it is not affected by locally disconnected nearby structures irrespective of whether or not they originate from the same object. Further, an explicit morphometric interpretation is built into the definition of locally adaptive scale models.

In,[4] a spherical model of scale, called *ball scale* (from here on abbreviated as *b-scale*), was proposed to determine the size of local structures at every space element (abbreviated from now on as *spel*) in the image. The b-scale value at every spel was defined as the radius of the largest ball centered at the spel such that all spels within the ball satisfy a pre-defined homogeneity criterion. Similar to the b-scale idea, a *tensor scale* model (or *t-scale* for short) was proposed by Saha,[5] which, at every spel, was defined as the parametric representation of the largest ellipse (in 2D) or ellipsoid (in 3D) centered at that spel and contained in the same homogeneous region under a pre-defined homogeneity criterion. While the shape of the b- and t-scale models is constrained by the choice of model (ball or ellipse), the largest allowable scale size at any image location for both these scale models[4,5] gets determined by the chosen shape. In,[6] we presented the generalized scale model (from here on abbreviated as g-scale), which like existing locally adaptive scale models[4,5] uses a homogeneity-based region continuity criterion for inclusion of image elements. It, however, differs from them in that it imposes no shape, size, or anisotropic constraints. This *semi-locally* adaptive behavior of g-scale was shown to be useful in an image processing application.[6]

Because of various sources of noise and artifacts and their ability to alter scene intensities in intractable ways, it is very unlikely that a purely theoretical definition of *local homogeneity* can be arrived at which can effectively address these artifacts. Consequently, all definitions of local homogeneity may have to be *ad hoc* in nature. The effectiveness of these definitions with respect to the artifacts is difficult to establish theoretically and can be established only experimentally. In this spirit, from the point of view of tolerance to noise, we believe that the process used in b-scale definition[4] of iteratively expanding the ball radius by 1 and comparing the mean intensity in the peripheral shell to the central spel's intensity is a better strategy than that employed in the g-scale definition of determining local homogeneity at the level of individual spels. It seems implausible to devise a similar strategy for g-scale. Consequently, we propose using the b-scale formulation to arrive at an alternative definition of g-scale, which we refer to as generalized ball-scale (from here on abbreviated as g_B-*scale*), which we anticipate will outperform g-scale for noisy images. In this paper we compare the properties of these two generalized scale models and their utility in an important MR image processing application, namely, in correcting for slowly varying background spatial intensity in the image.

2. THEORY

Before introducing the concept of g_B-scale, we first present the required preliminaries from fuzzy connectedness,[7] and concepts associated with the g- and b-scale notions.[4,6]

R^n, Z^n, spels

Let n-dimensional Euclidean space R^n be subdivided into hypercuboids by n mutually orthogonal families of orthogonal hyperplanes. We shall assume, with no loss of generality for our purposes, that the hyperplanes in each family have equal unit spacing so that the hypercuboids are unit hypercubes, and we shall choose coordinates so that the center of each hypercube has integer coordinates. The hypercubes will be called *spels* (an abbreviation for space elements). When n=2 and n=3, the spels are called *pixels* and *voxels*, respectively.

Fuzzy Adjacency, Hard Adjacency

A fuzzy relation α in Z^n is said to be a *fuzzy adjacency* if it is reflexive and symmetric. We denote by μ_α the membership function of α in Z^n. The basic idea is that the strength $\mu_\alpha(c,d)$ of adjacency between spels c and d in Z^n is a monotonically non-increasing function of the distance between c and d. The hard adjacency relationships commonly used in digital topology are special cases of fuzzy spel adjacencies. Any hard adjacency

relation α_b is a binary relation in Z^n that indicates the adjacency relationship between every two spels in Z^n.

Digital Space, Scene, Binary Scene
The pair (Z^n, α_b) is called a *digital space*. A *scene over* Z^n is a pair $\mathcal{C} = (C, f)$ where $C = \{c | -h_j \leq c_j \leq h_j$ for some $h \in Z^n{}_+\}$ (where $Z^n{}_+$ is the set of n-tuples of positive integers) and f is a function whose domain is C and whose range is a set of numbers. We call $\mathcal{C}$ *a binary scene over* Z^n if $\mathcal{C}$ is a scene over Z^n in which the range of f is {0,1}.

Fuzzy Affinity
A fuzzy relation κ in $\mathcal{C}$ is said to be a *fuzzy spel affinity* in $\mathcal{C}$ if it is reflexive and symmetric. The strength of the fuzzy affinity κ between any spels c and d in C is denoted by $\mu_\kappa(c, d)$. One functional form of $\mu_\kappa(c, d)$ that captures the degree of local hanging togetherness of c and d based on intensity homogeneity is given below.

$$\mu_\kappa(c, d) = \mu_\alpha(c, d)\mu_\psi(c, d), \tag{2.1}$$

where μ_ψ represents the *homogeneity-based component* of affinity and indicates how similar the intensities of c and d are. We assume that the homogeneity between two spels c and d in a scene $\mathcal{C}=(C, f)$ can be characterized by $|f(c) - f(d)|$ and express $\mu_\psi(c, d)$ as some function (W_ψ) of $|f(c) - f(d)|$.

$$\mu_\psi(c, d) = W_\psi(|f(c) - f(d)|). \tag{2.2}$$

g-scale
The formal definition of g-scale was presented in.[6] Roughly speaking, the g-scale $G(c)$ of spel c in a scene $\mathcal{C}$ represents the largest, homogeneous, connected region (of any shape) associated with spel c in $\mathcal{C}$. The g-scale set essentially represents a partitioning of C into fuzzy connectedness (by homogeneity only) regions of strength greater than a pre-defined threshold (θ).

Before formally defining g_B-scale, we first reproduce some ideas related to b-scale.[4]

Fraction of object, Ball Scale, Ball scale scene
A *hyperball* $B_k(c)$ of radius $k \geq 0$ and with center at $c \in C$ in a scene $\mathcal{C}=(C, f)$ over (Z^n, α) is given by,

$$B_k(c) = \left\{ e \in C \;\middle|\; \sqrt{\sum_{i=1}^{n} (c_i - e_i)^2} \leq k \right\}. \tag{2.3}$$

Note that $B_k(c)=\{c\}$ for $k = 0$. In,[4] for a hyperball $B_k(c)$ of radius $k \geq 0$ and centered at c, a fraction $FO_k(c)$ ("fraction of object") that indicates the fraction of the ball boundary occupied by a region which is sufficiently homogeneous with c, was defined by

$$FO_k(c) = \frac{\sum_{e \in [B_k(c) - B_{k-1}(c)]} W_\psi(|f(c) - f(e)|)}{|B_k(c) - B_{k-1}(c)|}, \tag{2.4}$$

where $|B_k(c) - B_{k-1}(c)|$ is the number of spels in $B_k(c) - B_{k-1}(c)$, and W_ψ is the homogeneity function. The *b-scale at* any *spel c in* any scene $\mathcal{C}=(C, f)$ over (Z^n, α) is defined to be the largest integer k-1 such that $FO_k(C) \geq \tau_b$, where τ_b is a fixed parameter in [0, 1]. The *b-scale scene* of $\mathcal{C}$ is a scene $\mathcal{C}_B=(C, f_B)$ over (Z^n, α) such that, for any spel $c \in C$, $f_B(c)$ is the b-scale at c in $\mathcal{C}$. We will refer to the hyperball of radius $f_B(c)$ centered at c as the *b-scale region of c in $\mathcal{C}$* and denote it by $B(c)$.

Definition 1. For any scene $\mathcal{C} = (C, f)$ over (Z^n, α), let $\mathcal{C}_B=(C, f_B)$ be its b-scale scene, and let ${\mathcal{C}_B}^b=(C, f_B^b)$ be the binary scene resulting from thresholding $\mathcal{C}_B$ at a b-scale value $\tau_g > 0$; i.e., for any $c \in C$, $f_B^b(c)=1$ if $f_B(c) \geq \tau_g$, and $f_B^b(c)=0$, otherwise. Let, for any $c \in C$, $[c]_{\alpha_b}$ denote the α_b-connected component, that contains c, of the spels with $f_B^b(c) = 1$ in $\mathcal{C}_B^b$. Then the *g_B-scale $G_B(c)$ of c in $\mathcal{C}$* is defined as the subset of C such that,

$$G_B(c) = \begin{cases} \{c\}, & \text{if } f_B(c) < \tau_g \\ [c]_{\alpha_b}, & \text{if } f_B(c) \geq \tau_g. \end{cases}$$

Definition 2. For any scene $\mathcal{C}=(C,f)$ over (Z^n,α) and for any positive integer τ_g, the *g_B-scale set* $\mathcal{G}_B(\mathcal{C})$ *of* $\mathcal{C}$ is the set $\{G_B(c) \mid c \in C\}$.

A consequence of definitions 1 and 2 is that not only the g_B-scale regions collectively cover the entire scene domain but also they do not overlap - that is, any two g_B-scale regions are either the same region in C or they are completely disjoint. This implies that (as for g-scale[6]), for any spel $d \in G_B(c)$, there is no need to compute $G_B(d)$.

2.1. Algorithms

Before describing the algorithm for computing g_B-scale, for completeness, we first reproduce algorithm *OSE* for b-scale estimation from.[4]

Algorithm *OSE*
Input: $\mathcal{C}$, $c \in C$, W_ψ, a fixed threshold τ_b.
Output: $f_B(c)$.
begin
0. set $k = 1$;
1. *while* $FO_k(c) \geq \tau_b$ *do*
2. set k to $k+1$;
3. *endwhile*;
4. set $f_B(c)$ to k-1;
5. output $f_B(c)$;

end

The algorithm for computing g_B-scale essentially calls a (hard) connected component labeling algorithm to track and output each connected component of spels with value 1 in the binary scene ${\mathcal{C}_\mathcal{B}}^b$ resulting from thresholding $\mathcal{C}_\mathcal{B}$ at τ_g.

Algorithm g_BSE
Input: $\mathcal{C}_\mathcal{B}$, a threshold τ_g.
Output: $\mathcal{G}_B(\mathcal{C})$.
begin
0. threshold $\mathcal{C}_\mathcal{B}$ at τ_g and create binary scene $\mathcal{C}_B^b$;
1. track and output each α_b-connected component of the 1-valued spels of $\mathcal{C}_B^b$;
2. output each 0-valued spel c of $\mathcal{C}_B^b$ as a separate g_B-scale region {c};

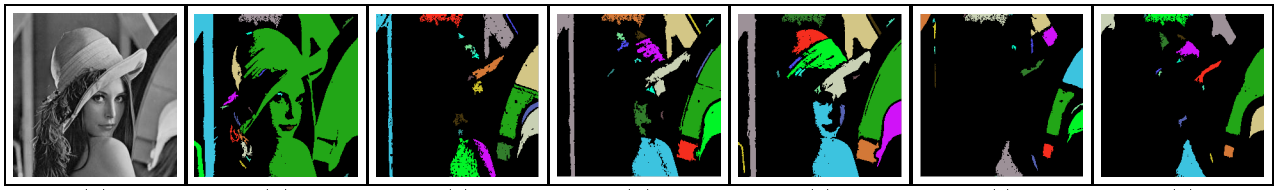
end

3. PROPERTIES OF GENERALIZED SCALE

3.1. Parameters for g_B-scale

The definition of g-scale involves two parameters, ζ, which controls the magnitude of the homogeneity parameter σ_ψ and an inclusion parameter θ. The definition of g_B-scale also involves two parameters: ζ and τ_g. (The definition of b-scale involves two parameters σ_ψ and τ_b. The value of τ_b is set to 0.83 as recommended in.[4]) In a manner similar to g-scale computation,[6] increasing the value of ζ results in larger b-scale values for individual spels. The parameter τ_g determines the criterion for including spels into the g_B-scale. Higher values of τ_g make the criterion for a spel to be determined as noise more stringent, which results in a larger number of smaller g_B-scale regions. Conversely, for lower values of τ_g, a smaller number of larger g_B-scale regions result (assuming ζ is fixed).

Figures 1(b)-(g) show the g_B-scale sets resulting from different combinations of values of τ_g and ζ for the original scene (Figure 1(a)). All spels in the same g_B-scale region in $\mathcal{G}_B(\mathcal{C})$ were assigned the same color. Also we have shown only the 20 largest g_B-scale regions in each $\mathcal{G}_\mathcal{B}(\mathcal{C})$ in Figure 1, and smaller g_B-scale regions were all assigned the black color. Higher values of τ_g and lower values of ζ cause more severe partitioning, while lower values of τ_g and higher values of ζ result in a smaller number of larger g_B-scale regions (Figure 1(b)). The values of ζ and τ_g used for computing $\mathcal{G}_\mathcal{B}(\mathcal{C})$ in Figures 1(d) and (e) seem to yield visually meaningful partitions of the scene.



(a) (b) (c) (d) (e) (f) (g)

Fig. 1: (a) Original scene $\mathcal{C}$. $\mathcal{G}_{\mathcal{B}}(\mathcal{C})$ corresponding to (b) τ_g=1, ζ=0.8, (c) τ_g=2, ζ=0.6, (d) τ_g=2, ζ=0.75, (e) τ_g=2, ζ=0.9, (f) τ_g=5, ζ=0.75, (g) τ_g=9, ζ=0.95. Only the largest 20 g-scale regions are shown in different colors. Smaller g_B-scale regions are all assigned the same color, namely black.

3.2. Selecting optimal values for g_B-scale parameters

As in,[6] our guiding factors for the selection of optimal values for ζ and τ_g are: (1) obtaining scale regions which represent the same object region, and (2) obtaining a meaningful partitioning of the scene domain. Our approach for selecting optimal values for the g_B-scale parameters is similar to our approach in[6] for g-scale. We consider a set E of MRI scenes of the brain for which accurately segmented white matter (WM) masks were available. Let $T(\mathcal{C})$ denote the subset of C constituting this segmented result for any scene $\mathcal{C}$ in E. Let $\mathcal{G}_{B,\tau_g,\zeta}(\mathcal{C})$ denote the g_B-scale set of scene $\mathcal{C}$ in E for a given τ_g and ζ. For each pair of (ζ, τ_g) values, let the largest g_B-scale regions in $\mathcal{G}_{B,\tau_g,\zeta}(\mathcal{C})$ be denoted by $G^{\max}_{B,\tau_g,\zeta}(\mathcal{C})$. Let $Y = G^{\max}_{B,\tau_g,\zeta}(\mathcal{C})$. The trade-off referred to previously manifests itself in the form of the balance between *false negative volume fraction*, $FNVF(\mathcal{C})=\frac{|T(\mathcal{C})-Y|}{|T(\mathcal{C})|}$ and *false positive volume fraction*, $FPVF(\mathcal{C})=\frac{|Y-T(\mathcal{C})|}{|\mathcal{C}-T(\mathcal{C})|}$ (see[8]), where $\mid X \mid$ denotes the cardinality of set X. $FPVF(\mathcal{C})$ indicates the extent of the object region falsely included in Y as compared to the true mask $T(\mathcal{C})$. $FNVF(\mathcal{C})$ similarly indicates the extent of true object region that is missed by Y. As described in,[8] this classical dilemma surfacing in the context of image segmentation can be summarily quantified via the notion of *delineation operating characteristic* (DOC) curve, which is akin to the ROC curve used in decision theory.[9] The DOC curve is a plot of the variation of *delineation sensitivity* $SN(\mathcal{C})$ with respect to *delineation specificity* $SP(\mathcal{C})$,

$$SN(\mathcal{C}) = 1 - \mathcal{FNVF}(\mathcal{C}), \quad SP(\mathcal{C}) = 1 - \mathcal{FPVF}(\mathcal{C}), \tag{3.5}$$

as the values of the parameters of the segmentation method (in our case, of the scale methods) are varied, and it represents the inter-dependence (trade-off) between $FNVF(\mathcal{C})$ and $FPVF(\mathcal{C})$. For each of the two parameters, we select a set of values X_1 and X_2 of the parameters. Then, for all possible combinations of values of the two parameters in $X_1 \times X_2$, for each scene $\mathcal{C}$ in E, we compute $SN(\mathcal{C})$ and $SP(\mathcal{C})$, which yields one point on the DOC curve. The set of tuples in $X_1 \times X_2$ thus yields a set of points on the DOC curve for $\mathcal{C}$. A smooth curve is then fit to these points. The curves resulting from the scenes in E are then averaged to yield one DOC curve for E. The parameter values corresponding to the computed point on the DOC curve which is closest to the ideal point for which SN and SP values are both 100% is then chosen to be the optimal parameter setting (Figure 2). The operating points for g_B-scale for scenes with no noise or inhomogeneity (Figure 2(a)), for scenes with no noise but containing inhomogeneity (Figure 2(b)), and for scenes containing both noise and inhomogeneity (Figure 2(c)), correspond roughly to τ_g=1, 2, and 3, respectively, suggesting that higher values of τ_g are more appropriate for noisy scenes. The optimal values for ζ and τ_g were found to be 0.7 and 2 respectively.

3.3. Effect of background inhomogeneity on the estimation of g- and g_B-scale

We introduced varying levels of background inhomogeneity (zero-mean Gaussian) into a mathematical phantom scene (Figures 3(b)-(d)) and computed the resulting g- and g_B-scale sets (Figures 3(e)-(l)). The phantom (Figure 3(a)) comprises of two homogeneous regions of different intensities against a uniform dark background. Figures 3(e)-(h) show the corresponding g-scale sets, all computed for the same value of θ and ζ, and Figures 3(i)-(l) show the corresponding g_B-scale sets, all computed for the same value of τ_g and ζ. For no, low level of, and

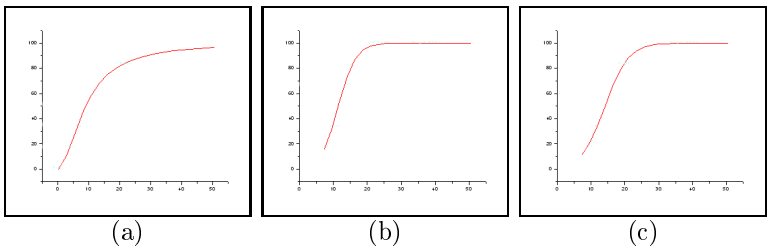


Fig. 2: DOC curves for $\mathcal{G}_{B,\tau_g,\zeta}(\mathcal{C})$ averaged over (a) 5 phantom PD-weighted scenes with no noise or inhomogeneity, (b) 5 phantom PD-weighted scenes containing inhomogeneity (40%), but no noise, and (c) 5 phantom T2-weighted scenes containing both noise (3%) and inhomogeneity (40%).

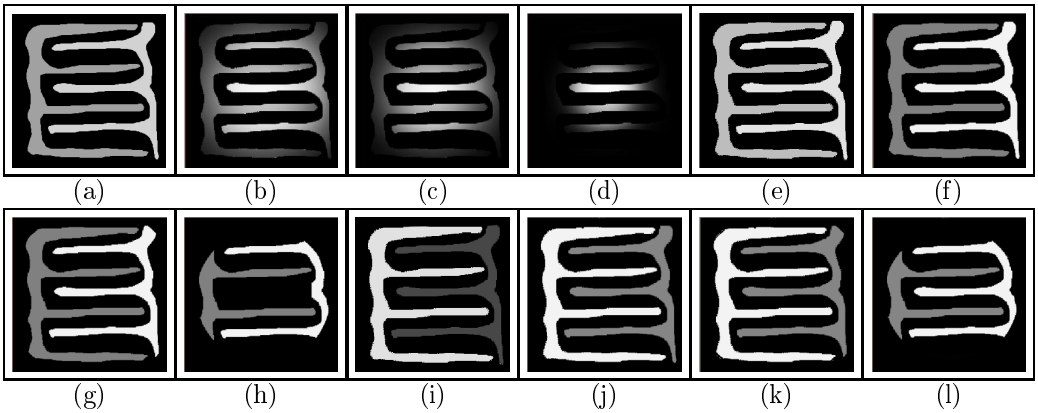


Fig. 3: Phantom with (a) no, (b) low, (c) medium, and (d) high level of inhomogeneity, (e)-(h) g-scale set corresponding to (a)-(d), (i)-(l) g_B-scale set corresponding to (a)-(d).

medium level of inhomogeneity, both g- and g_B-scale are able to accurately estimate the two foreground regions (Figures 3(e)-(g) and (i)-(k)). For the case of severe inhomogeneity, g_B-scale appears to perform better than g-scale (Figures 3(h), (l)). The experiment in Figure 3 ascertains that the process of estimating generalized scale regions is resistant to slowly varying background intensity.

3.4. Effect of noise on the estimation of g- and g_B-scale

In order to study the noise-resistance properties of the g- and g_B-scale models, we introduced different levels of uncorrelated noise into the phantom of Figure 3(a). For the scenes resulting from adding three different levels

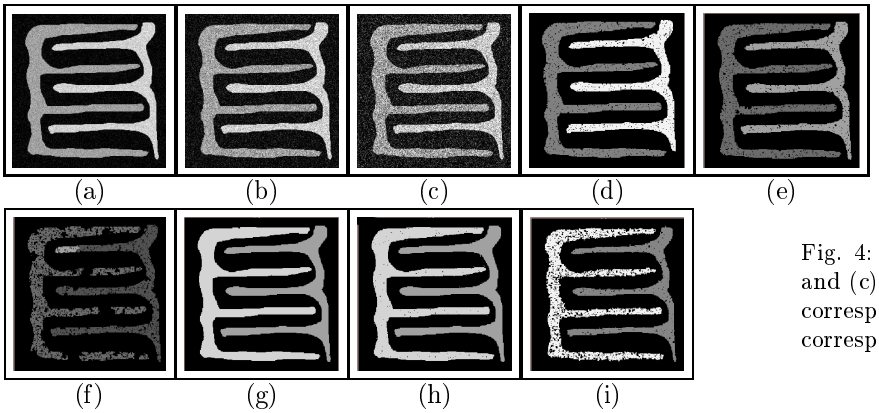


Fig. 4: Phantom with (a) low, (b) medium, and (c) high level of noise; (d)-(f) g-scale sets corresponding to (a)-(c); (g)-(i) g_B-scale sets corresponding to (a)-(c).

of noise (Figures 4(a)-(c)) to the scene in Figure 3(a), the resulting g_B-scale sets (Figures 4(g)-(i)) are clearly superior to the resulting g-scale sets (Figures 4(d)-(f)) at all noise levels. The g_B-scale definition builds on the b-scale idea which has been shown to be noise-resistant.[4] The computation of g-scale on the other hand is based on comparison of pair-wise spel intensities, making it more susceptible to noise.

4. INHOMOGENEITY CORRECTION

Intensity non-uniformity is a serious problem in many image processing areas. In MRI, this may be caused by a number of different factors, including poor RF coil uniformity, static field inhomogeneity, RF penetration, gradient-driven eddy currents, and patient anatomy both inside and outside the field of view. Several methods of inhomogeneity correction have been proposed since the first paper on the subject.[10] Most of these methods make certain assumptions about the datasets[11–13] or require some amount of manual intervention.[14,15] In,[16] a fully automated b-scale-based correction methodology was proposed. The method assumes that thresholding the scene using intensity statistics obtained from the largest b-scale region corresponds to the largest tissue region. One limitation of using b-scale to estimate scene inhomogeneity is that since it ignores local anisotropy and orientation, it is not very effective along elongated structures and edges. More importantly, however, thresholding by itself cannot guarantee a sufficiently large region of the same tissue. A g-scale based approach was found to significantly outperform the b-scale correction method on clinical and phantom studies.[6] In this paper, we propose a g_B-scale-based correction method that is very similar to the g-scale method.[6] We believe that both generalized scale-based approaches are superior to existing methods on account of (i) both g- and g_B-scales having the inherent property of not being affected by slow background variations, (ii) the computation of both g- and g_B-scales being completely automatic requiring no *a-priori* knowledge about the MRI protocol used, body region imaged, or the scanner, and (iii) both g- and g_B-scales resulting in a partitioning of the image into a number of homogeneous regions, each homogeneous region corresponding to a single object/tissue region. Hence, without explicit segmentation, it is possible to sample the same object/tissue region, and hence, to accurately estimate true intensities.

4.1. Methodology

We will use the following notation for the g- and g_B-scale correction methods.

$\mathcal{C}$=(C,f): a given scene corrupted by background variation.
$\mathcal{C}_{bt}$=(C,f_{bt}): the true background variation component in $\mathcal{C}$.
$\mathcal{C}_u$=(C,f_u): $\mathcal{C}$ without the background variation component $\mathcal{C}_{bt}$.
$\mathcal{C}_c$=(C,f_c): the scene resulting from applying the b-scale correction method to $\mathcal{C}$.
$\mathcal{C}_{c_g}$=(C,f_{c_g}): the scene resulting from applying the g-scale correction method to $\mathcal{C}$.
$\mathcal{C}_{c_{g_B}}$=(C,$f_{c_{g_B}}$): the scene resulting from applying the g_B-scale correction method to $\mathcal{C}$.
$\mathcal{C}_{bb}$=(C,f_{bb}): the background variation component in $\mathcal{C}$ estimated by the b-scale correction method.
$\mathcal{C}_{bg}$=(C,f_{bg}): the background variation component in $\mathcal{C}$ estimated by the g-scale correction method.
$\mathcal{C}_{bg_B}$=(C,f_{bg_B}): the background variation component in $\mathcal{C}$ estimated by the g_B-scale correction method.

Image intensity inhomogeneity in MR images is approximately a multiplicative phenomenon.[10] For an MRI scene $\mathcal{C}$, the inhomogeneity at any spel $c \in C$ can be expressed as,

$$f_{bt}(c) = \frac{f(c)}{f_u(c)}. \tag{4.6}$$

The central idea of most correction methodologies is thus to estimate $f_{bt}(c)$ by obtaining a good estimate of the true scene intensity $f_u(c)$. In our methodology, we use generalized scales to obtain an estimate of $f_u(c)$. Once the g_B-scale set is determined, the correction methodology follows the same steps as the g-scale method.[6]

4.2. Results

For the purpose of qualitative evaluation, we tested the b-, g-, and g_B-scale correction methods on several clinical MR images corresponding to different MRI protocols and body regions. The scenes resulting from the

application of the three methods were then visually inspected for residual inhomogeneity. The three methods were also quantitatively compared on a total of 210 phantom scenes from the MNI database *.

4.3. Qualitative

Slices from three clinical MRI scenes corresponding to different regions of the human body and the results obtained from these scenes for the b-, g-, and g_B-scale-based correction methods are shown in Figure 5. Figures 5(a), (e) show slices from clinical MRI scenes of the abdomen and shoulder joint, respectively; both scenes are corrupted by severe inhomogeneity. The results of g- and g_B-scale correction in Figures 5(c), (g) and (d), (h), respectively, show that not only most of the inhomogeneity in Figures 5(a), (e) is corrected, but also significantly more anatomic detail is revealed in these scenes than in the results obtained by b-scale correction (Figures 5(b), (f)). Even for scenes with mild inhomogeneity, as in the MRI scene of the foot (Figure 5(i)), both g- and g_B-scale correction (Figures 5(k), (l)) outperform b-scale correction (Figure 5(j)).

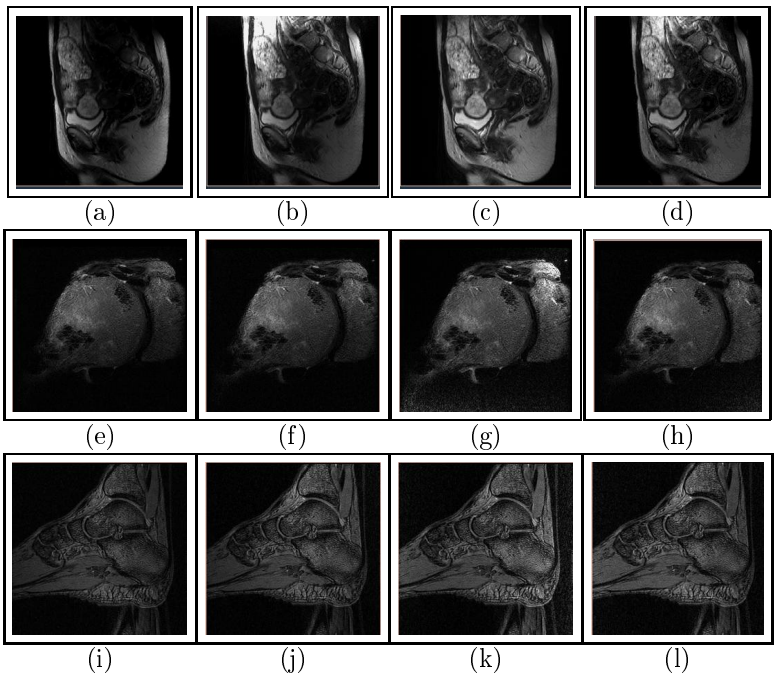


Fig. 5: (a), (e), (i) Slices from original MRI scenes; corresponding slices from (b), (f), (j) b-scale, (c), (g), (k) g-scale, and (d), (h), (l) g_B-scale corrected scenes.

4.4. Quantitative

For quantitative evaluation of the generalized scale correction methods, we considered a total of 210 scenes corresponding to (i) no inhomogeneity and no noise; (ii) inhomogeneity, but no noise; (iii) inhomogeneity and noise. Three levels of inhomogeneity corresponding to 0%, 20%, 40% and three levels of noise corresponding to 0%, 1%, 3% were used in our experiments. Since, for phantom scenes, the true signal intensity is known, we decided to use the RMS difference between the original and corrected scene intensities as the figure of merit to evaluate the effectiveness of the correction method. All scenes were standardized by using the method in[17] prior to evaluation but following inhomogeneity correction as recommended in.[18]

4.4.1. No inhomogeneity, no noise

Ideally a correction algorithm should be expected to leave a scene containing no inhomogeneity unchanged. Hence, the RMS intensity difference between the original scenes and the scenes obtained by the application of a correction method would be an indicator of the amount of artificial inhomogeneity introduced by the method.

*http://www.bic.mni.mcgill.ca/brainweb

In Table 1, the mean and standard deviation of the RMS intensity differences between corrected and original scenes are listed for the PD studies for different methods. In all 30 scenes, the g- and g_B-scale correction methods outperformed b-scale, there being no statistically significant difference in performance between the g- and g_B-scale methods. Overall, while g_B-scale marginally outperformed g-scale for the normal studies, the trend reversed for the MS studies (Table 1). Similar results were observed with T1 and T2 scenes. Table 2 shows the p-values of the paired t-tests conducted to compare methods pairwise utilizing all 30 data sets. Figure 6

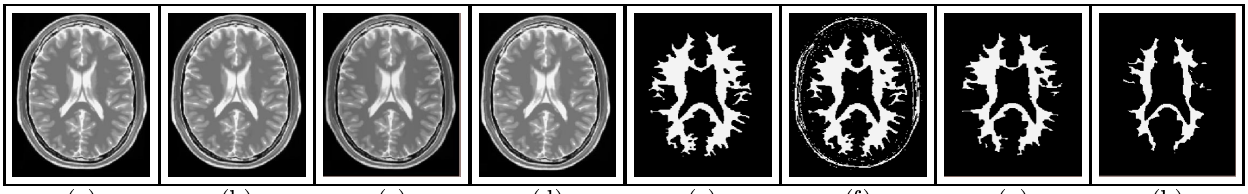
μ, σ of	Normal			MS		
RMS diff.	b	g	g_B	b	g	g_B
μ	0.281	0.074	0.073	0.189	0.083	0.130
σ	0.122	0.024	0.081	0.059	0.037	0.110

Tab. 1: The mean and standard deviation of RMS of intensity differences over 5 scenes for each of the PD scenes with 0% noise and 0% inhomogeneity, and corrected with the three scale methods.

Normal			MS		
b-g	b-g_B	g-g_B	b-g	b-g_B	g-g_B
$3.714e^{-4}$	$3.907e^{-4}$	0.200	$1.951e^{-4}$	$1.087e^{-4}$	0.138

Tab. 2: p-values of paired t-tests for comparing RMS values for different pairs of methods.

provides a qualitative explanation for the results in Tables 1 and 2. While the correction results (Figure 6(b)-(d)) on a T2-weighted scene (with 0% noise and 0% inhomogeneity) all appear visually similar, the largest scale regions estimated by the g- and g_B-scale methods (Figures 6(g), (h)) are more similar to the true WM region (obtained from the MNI database) in Figure 6(e) compared to the largest region determined via b-scale (Figure 6(f)). The largest region obtained by the b-scale method includes both WM and some cranial tissues, while the largest g-scale (Figure 6(g)) and g_B-scale regions (Figure 6(h)) are completely within the WM region.



(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 6: Displays of a slice from a normal T2 scene for various situations. (a) $\mathcal{C}$ with no noise and no inhomogeneity, (b) $\mathcal{C}_c$, (c) $\mathcal{C}_{c_g}$, (d) $\mathcal{C}_{c_{g_B}}$, (e) the true WM region, and largest regions determined via (f) b-scale, (g) g-scale, and (h) g_B-scale.

4.4.2. With inhomogeneity but no noise

The mean and standard deviation of the RMS intensity differences between corrected scenes and original scenes (with no inhomogeneity and no noise) are listed in Table 3 for PD scenes obtained by using different methods. In all 60 scenes, the g- and g_B-scale correction methods outperformed b-scale, there being no statistically significant difference in performance between the g- and g_B-scale methods. Overall, however, g-scale marginally outperformed g_B-scale at both inhomogeneity levels (Table 3, bottom). Similar results were observed with T1 and T2 scenes. Table 4 shows the p-values of the paired t-tests conducted to compare methods pairwise for the two levels of inhomogeneity. Figure 7 shows the result of the three methods on a slice of a PD scene with 40% inhomogeneity and 0% noise (Figure 7(a)). Figures 7(b)-(d) show the same slices from scenes $\mathcal{C}_c$, $\mathcal{C}_{c_g}$, and $\mathcal{C}_{c_{g_B}}$. Figures 7(f)-(h) represent the largest scale regions extracted by using the b-, g-, and g_B-scale

Category	Protocol	μ and σ of RMS difference	20%			40%		
			b	g	g_B	b	g	g_B
Normal	PD	μ	3.092	1.538	1.433	5.106	3.282	3.125
		σ	0.720	0.141	0.272	0.472	0.180	0.595
MS	PD	μ	2.950	1.571	1.455	5.617	3.171	3.160
		σ	0.840	0.119	0.423	0.417	0.445	0.383

Tab. 3: The mean (μ) and standard deviation (σ) of RMS of intensity differences for PD scenes with 0% noise, and with 20% and 40% inhomogeneity, and corrected by using the three scale methods.

	Normal			MS		
Inhomogeneity	b-g	b-g_B	g-g_B	b-g	b-g_B	g-g_B
20%	$1.242e^{-4}$	$9.997e^{-5}$	0.071	$9.561e^{-5}$	$9.991e^{-5}$	0.094
40%	$1.322e^{-5}$	$1.490e^{-5}$	0.085	$3.307e^{-8}$	$3.096e^{-6}$	0.099

Tab. 4: p-values of paired t-tests for comparing RMS values for different pairs of methods.

methods. The images displayed in Figures 7(e)-(h) correspond to the slices in $\mathcal{C}$, $\mathcal{C}_c$, $\mathcal{C}_{c_g}$, and $\mathcal{C}_{c_{g_B}}$ displayed in Figures 7(i)-(l). The reason for showing the images in Figures 7(a)-(d) is that the differences in inhomogeneity for the methods are more apparent in this slice of the scene near the bottom of the brain than in a slice near the middle of the brain, while the WM regions are bigger near the middle of the brain (Figures 7(i)-(l)). The largest g-scale (Figure 7(g)) and g_B-scale regions (Figure 7(h)) constitute more accurate samplings of the true WM region (Figure 7(e)) than the region obtained via b-scale (Figure 7(f)). Since the largest region is estimated by thresholding the original scene $\mathcal{C}$ based on thresholds derived from the largest b-scale region, its estimation is affected by the presence of inhomogeneity. These differences are reflected in the estimated bias fields shown in Figures 7(n)-(p). The similarity between $\mathcal{C}_{bg}$, $\mathcal{C}_{bg_B}$ (Figures 7(o), (p)) and the true inhomogeneity $\mathcal{C}_{bt}$ (Figure 7(m)) explains why $\mathcal{C}_{c_g}$ and $\mathcal{C}_{c_{g_B}}$ (Figures 7(c), (d)) contain significantly less residual inhomogeneity than $\mathcal{C}$ and $\mathcal{C}_c$.

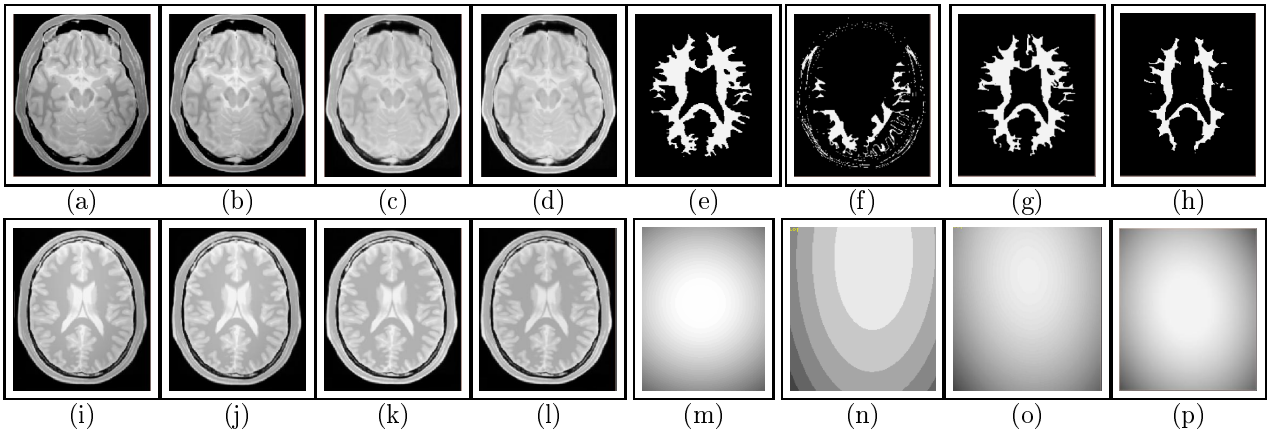


Fig. 7: Displays of a slice from a normal PD scene for various situations. (a) $\mathcal{C}$ with 40% inhomogeneity and 0% noise, (b) $\mathcal{C}_c$, (c) $\mathcal{C}_{c_g}$, (d) $\mathcal{C}_{c_{g_B}}$. Displays of another slice (i)-(l) from the same scene for the situations in (a)-(d). (e) True WM mask, largest regions estimated via (f) b-scale, (g) g-scale, and (h) g_B-scale, (m) $\mathcal{C}_{bt}$, (n) $\mathcal{C}_{bb}$, (o) $\mathcal{C}_{bg}$, and (p) $\mathcal{C}_{bg_B}$.

4.4.3. With inhomogeneity and noise

The mean and standard deviation of the RMS intensity differences between corrected scenes and original scenes are listed in Table 5 for PD scenes for different methods. In all 120 scenes, the g- and g_B-scale correction methods out-performed the b-scale method, all differences being statistically significant. While no statistically significant difference was observed between g- and g_B-scale for lower noise levels (1%), the g_B-scale method was significantly superior at higher noise levels (3%). Overall, however, g_B-scale outperformed g-scale at both levels of noise and inhomogeneity. Only results for scenes with high levels of noise (3%) have been shown (Table 5). Similar results were observed with T1 and T2 scenes. Table 6 shows the p-values of the paired t-tests conducted to compare methods pairwise for the two levels of noise and inhomogeneity. Figure 8 shows the result of the

Category	Protocol	μ and σ of RMS difference	20%			40%		
			b	g	g_B	b	g	g_B
Normal	PD	μ	4.885	4.708	4.288	6.349	5.705	4.735
		σ	0.129	0.197	0.453	0.613	0.694	0.028
MS	PD	μ	4.689	4.343	3.824	5.906	5.109	4.371
		σ	0.058	0.181	0.294	0.387	0.490	0.085

Tab. 5: The mean and standard deviation of RMS intensity differences for PD scenes with noise (3%) and inhomogeneity (20%, 40%), and corrected by using the three scale methods.

		Normal			MS		
Noise	Inhomo	b-g	b-g_B	g-g_B	b-g	b-g_B	g-g_B
3%	20%	$6.734e^{-3}$	$4.951e^{-4}$	0.011	$7.653e^{-3}$	$6.485e^{-4}$	$4.087e^{-3}$
3%	40%	0.023	$2.376e^{-3}$	$1.896e^{-3}$	0.011	$1.193e^{-3}$	$3.539e^{-3}$

Tab. 6: p-values of paired t-tests for comparing RMS values for different pairs of methods.

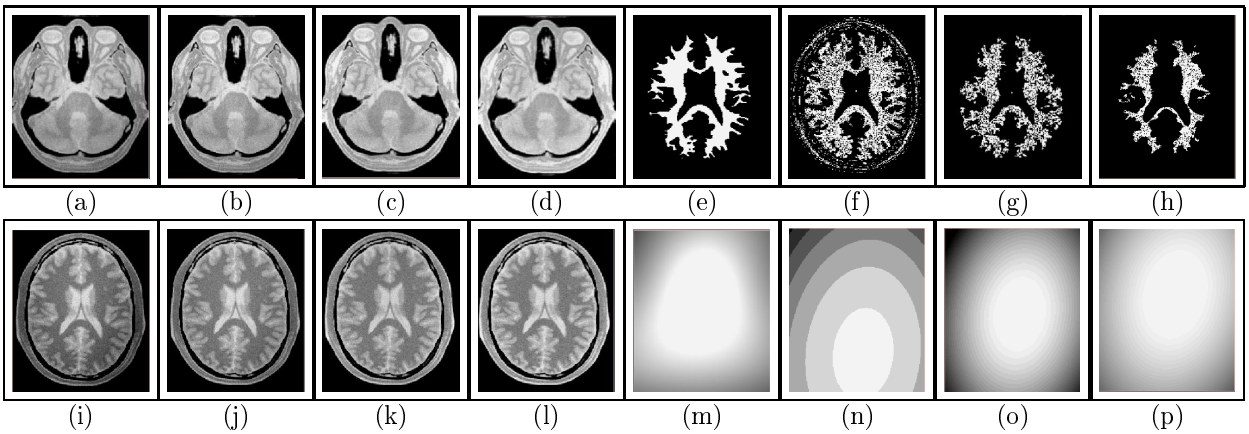


Fig. 8: Displays of a slice from a normal T2 scene for various situations. (a) $\mathcal{C}$ with 40% inhomogeneity and 3% noise, (b) $\mathcal{C}_c$, (c) $\mathcal{C}_{c_g}$, (d) $\mathcal{C}_{c_{g_B}}$. Displays of another slice (i)-(l) from the same scene for the situations in (a)-(d). (e) True WM mask, largest regions estimated via (f) b-scale, (g) g-scale, and (h) g_B-scale, (m) $\mathcal{C}_{bt}$, (n) $\mathcal{C}_{bb}$, (o) $\mathcal{C}_{bg}$, and (p) $\mathcal{C}_{bg_B}$.

three methods on a slice of a T2 scene with 40% inhomogeneity and 3% noise. The arrangement for this figure is as for Figure 7. The largest g_B-scale region (Figure 8(h)) constitutes a more accurate sampling of the true

WM region (Figure 8(e)) than the corresponding b- and g-scale regions (Figures 8(f), (g)). These differences are reflected in the estimated bias fields shown in Figures 8(n)-(p). The similarity between $\mathcal{C}_{b_{g_B}}$ (Figure 8(p)) and the true inhomogeneity $\mathcal{C}_{bt}$ (Figure 8(m)) explains why $\mathcal{C}_{c_{g_B}}$ (Figure 8(d)) contains significantly less residual inhomogeneity than $\mathcal{C}$, $\mathcal{C}_c$, and $\mathcal{C}_{c_g}$.

5. CONCLUDING REMARKS

In this paper, we have presented a variant of generalized scale called generalized ball scale (g_B-scale), which in addition to possessing the advantages of g-scale, also possesses the noise resistance properties of b-scale. g_B-scale requires only two parameters for its specification. We have described a method to automatically estimate these parameters in an optimal manner. Further, g_B-scale, like g-scale has the inherent property of being resistant to the presence of inhomogeneity in a scene. In addition, since g_B-scale inherits the properties of b-scale, it is resistant to the presence of noise. Both g- and g_B-scale correction methods were found to be significantly superior to the b-scale correction method on clinical images and phantoms, across MRI protocols, and at different levels of noise and inhomogeneity. On phantom scenes containing no noise and no inhomogeneity, these methods were found to alter the original scenes significantly less than b-scale correction. g_B-scale performed significantly better than g-scale for scenes with high noise levels and severe inhomogeneity. At low levels of noise, the behavior of the generalized scale models is comparable. The computational costs are slightly higher for g_B-scale than for g-scale (typical total processing time of 24 s for g_B-scale versus 22 s for g-scale method on a 2.3 GHz Pentium IV PC). Consequently, for most clinical scenes with low levels of noise, the g-scale method should be adequate. The g_B-scale method should be preferred when noise levels are moderate or high.

REFERENCES

1. A.P. Witkin, "Scale-space filtering", *Proc. of 8th Intl. Joint Conf. Artificial Intelligence*, pp. 1019-22, 1983.
2. P. Burt, "Fast Filter Transform for Image Processing", *Computer Graphics and Image Processing*, pp. 20-51, 1981.
3. Tony Lindeberg, "Scale-Space Theory in Computer Vision", *Kluwer Academic Publishers*, 1993.
4. P. Saha, J. Udupa, D. Odhner "Scale-based Fuzzy Connected Image Segmentation: Theory, Algorithms, and Validation", *CVIU*, vol. 77, pp. 145-174, 2000.
5. P. Saha, "Novel Theory and Methods for Tensor Scale: A Local Morphometric Parameter", *Proc. of SPIE: Medical Imag.*, vol. 2, pp. 314-24, 2003.
6. A. Madabhushi, J. Udupa, "Generalized Scale: Theory, Algorithms, and Application to Inhomogeneity Correction", *Proc. of SPIE: Medical Imag.*, vol. 5370, pp. 765-776.
7. J. K. Udupa, S. Samarasekera, "Fuzzy Connectedness and Object Definition: Theory, Algorithms, and Applications in Image Segmentation", *CVGIP*, vol. 58(3), pp. 246-61, 1996.
8. J. Udupa, Y. Zhuge, "Delineation operating characteristic (DOC) curve for assessing the accuracy behavior of image segmentation algorithms", *Proc. of SPIE: Medical Imag.*, 2004, vol. 5[24], pp. 640-47.
9. J. Hanley, B. McNeil "The meaning and use of the area under the ROC curve", *Radiol.*, vol. 143, pp. 29-36, 1982.
10. L. Axel, J. Constantini, J. Listerud, "Intensity Correction in Surface-Coil MR Imaging", *America. J. of Radiol.*, vol. 148, pp. 418-420, 1987.
11. W. Wells, W. Grimson, R. Kiknis, F. A. Jolez, "Adaptive Segmentation of MRI data", *IEEE Trans. Med. Imaging*, vol. 15, pp. 429-443, 1996.
12. M. Styner, C. Brechbuhler, G. Szekely, G. Gerig, "Parametric Estimate of Intensity Inhomogeneities Applied to MRI", *IEEE Trans. Med. Imaging*, vol. 19, pp. 153-165, 2000.
13. J. Sled, A. Zijdenbos and A. Evans, "A Non-parametric Method for Automatic Correction of Intensity Nonuniformity in MRI Data", *IEEE Trans. Med. Imaging*, vol. 17, pp. 87-97, 1998.
14. B. M. Dawant, A. P. Zjidenbos, R. A. Margolin, "Correction of intensity variations in MR images for computer-aided tissue classification", *IEEE Trans. Med. Imaging*, vol. 12, pp. 770-81, 1993.
15. C. R. Meyer, P. H. B. Land, J. Pipe, "Retrospective Correction of intensity inhomogeneities in MRI", *IEEE Trans. Med. Imaging*, vol. 14, pp. 36-41, 1995.
16. Y. Zhuge, J. Udupa, J. Liu, P. Saha "A Scale-Based method for correcting background intensity variation in acquired images", *Proc. of SPIE: Medical Imag.*, vol. 4684, pp. 1103-1111, 2002.
17. L. G. Nyul, J. Udupa, "On standardizing the MR Image Intensity Scale", *Mag. Reson. in Med.*, vol. 42, pp. 1072-1081, 1999.
18. A. Madabhushi, J. Udupa, "Interplay of Intensity Standardization and Inhomogeneity Correction in MR Image Analysis", *IEEE Trans. Med. Imaging*, In Press.