Color Enhancement of Multispectral MR Images: Improving the Visualization of Subcortical Structures

Julianna Ward, Vincent Magnotta, Nancy C. Andreasen, Wendy Ooteman, Peg Nopoulous, and Ronald Pierson

Purpose: The current investigation was undertaken to evaluate a new method for creating MR multispectral color images, which we call "Superimages." They were developed to improve the delineation of small brain structures composed of mixed tissue types, such as the basal ganglia.

Method: To qualitatively validate the method, visual comparisons were made of six unimodal and multispectral images, including the Superimage. Quantitative validation was undertaken by comparing the reliability values for parcellation of the globus pallidus (GP) using either a gray scale (T1-weighted) image or the Superimage.

Results: Qualitative assessment of the Superimage revealed enhanced visualization of the GP, caudate, and putamen. Quantitative assessment resulted in good reliability for Superimage traces.

Conclusion: The Superimage significantly improves both the visualization and the parcellation of structures visualized by MRI.

Index Terms: Magnetic resonance imaging, multispectral—Parcellation—Basal ganglia—Globus pallidus.

MRI has quickly become a powerful research tool in the neurosciences (1–3). Research endeavors before MRI were limited to investigations of brain-behavior relationships with postmortem tissue. Such studies possessed the inherent limitations of age-, disease-, and mortality-related tissue changes that in turn limited the scope of investigation and generalizability of findings (4). MRI permits neuroscientists to conduct "antermortem autopsies" (4). These in vivo procedures have exponentially increased the research subject pool, allowing noninvasive investigation of the human brain across the life span and range of disease entities (4).

Advances in MR resolution and image enhancement technologies now permit visualization of soft tissues and circumscribed neuroanatomic structures (5–8). As a result, measurements of volumetric differences in brain structures such as the caudate and putamen have been undertaken between patient and control groups (9). In addition to hardware advances, software applications that aid in postacquisition image analysis entail continued development. The software package used in the current investigation was developed in-house and is entitled BRAINS (acronym for Brain Research: Analysis of Images, Networks, and Systems) (3,4). This package permits basic functions such as tracing, edge detection, and area and volume estimation. The most recent update of this software is entitled BRAINS2 and includes an option that allows simultaneous viewing of all MR sequences [T1, T2, and proton density (PD)] as one image enhanced with color.

Postacquisition processing techniques that synthesize numerous sources of information into one image (i.e., multispectral imaging) have been found to substantially enhance the visualization of small brain structures and result in greater accuracy on the part of the diagnostician and research scientist (6,8). Multispectral image display has been widely studied (10–16), as has the efficacy of color visualization for further enhancing medical imaging (7,8,17–20). Clarke and associates (21) stated that "gray scale segmentation methods may provide some useful information, but they are generally limited to relatively simple structures. For complex pathology, more information is often required, which is available in multispectral MRI data" (p. 347). As is true for "complex pathology," multispectral images have proven helpful in
delineating complex (i.e., hard-to-see, mixed tissue type) brain structures on MRI (6,8).

**MULTISPECTRAL IMAGING**

MRI is based on multiple tissue parameters, and the resultant image comprises an array of pixels (3,22). Three orthogonal parameters comprise each pixel (23): spin-lattice relaxation time (T1), spin-spin relaxation time (T2), and PD. Each parameter (pulse sequence) has its own unique tissue contrast pattern from which a "weighted image" can be produced, and the contrast patterns of these "weighted images" reflect the biophysical characteristics of the fluids and tissues scanned (17). Depending on the scan sequence used, MR images reveal varying aspects of tissue and structural composition (18) that have generally been viewed in gray scale. See Figs. 1, 2, and 3 for examples of T1-, T2-, and PD-weighted images of the basal ganglia at the temporal limb of the anterior commissure, viewed in the coronal plane. These three modalities each provide unique diagnostic information (22).

Early in the development of MRI, diagnosticians viewed T1-, T2-, and PD-weighted images separately, that is, unimodally (6,24). This approach was time consuming, requiring constant shifting among and readjustment to the anatomy and different contrast levels in the various images (6). Current MR technology permits the incorporation of all scan sequence data into one image, that is, a multispectral image (21). A technical requirement for deriving multispectral images involves processing MR sequences so that they are in the same orientation and at the same tissue slice (13). With data from these sequences combined, the amount of information available for interpretation is maximized and the likelihood of accurately identifying discrete brain structures, be they lesions or subcortical nuclei, enhanced (6,8).

Combination of MR sequencing data has been achieved through various techniques. Harris and associates (1) investigated the reliability and validity of "operator-dependent" and automated methods for generating tissue "training classes" that underwent discriminant analysis to segment tissue and generate discrete and continuous multispectral MR images. The discrete and con-
Continuous images created with this method were generated using raw data from the original MR sequences (T1, T2, and PD) wherein any given voxel comprised gray matter (GM), white matter (WM), CSF, or a combination of these (e.g., GM and WM or WM and CSF). Raw sequence data were classified as either categorical or continuous in the production of the discrete and continuous images, respectively. Discrete image production progressed from a voxel-by-voxel analysis, in which the percentage of GM, WM, and CSF in a given voxel was determined. Voxels were subsequently categorized as discrete tissue types based on these percentage values (e.g., a voxel containing 51% GM and 49% WM was categorized as solely GM using the discrete image technique). Therefore, the amount of information available for interpretation in each voxel of the discrete image was reduced by this forced choice method. The resulting data loss produced a somewhat degraded image (see Fig. 4).

The advantages of such data reduction methods lie in assessing structures that are grossly composed of one tissue type. Conversely, shortcomings of this technique are found in the image degradation caused by partial voluming that is worsened by data reduction. These problems are of particular difficulty when attempting to define the boundaries of structures containing a high proportion of mixed tissue types (e.g., GM and WM), such as the basal ganglia. In contrast to the discrete image, more raw data from the original MR sequences were preserved in the continuous image. As might be expected, continuous images were less vulnerable to partial voluming effects. Accordingly, the visualization of subcortical structural boundaries was enhanced (see Fig. 5). A more refined delineation of the caudate, putamen, and globus pallidus (GP) was possible with the continuous image. However, discrimination of the medial boundary of the GP from its medial neighbor (i.e., the internal capsule) remained quite difficult. Although Harris et al. (1) found their automated discrete image to be highly reliable and valid for distinguishing gross tissue classes, analysis of the special demands posed by brain structures that are harder to parcelate because of size and tissue composition constraints fell beyond the scope of their investigation. In response to such demands, the current study examined a methodology for producing multispectral color images that fully retain the raw data from the T1, T2, and PD sequences. The term “Superimage” has been coined to distinguish images obtained using this technique.

Although each MR sequence contains diagnostically useful information, the visual shifting among the respective images required to achieve a mental summation of the varying signal intensity data is time consuming and inefficient (6). The Superimage allows for a simultaneous presentation of all original data in the three MR sequences with the added benefit of color visualization (which will be discussed in more detail below). See Figure 6 for an example of the Superimage.

MR images, both unimodal and multispectral, have historically been viewed in gray scale (6). Gray scale presentation of MR images may reduce the contrast between structures in the brain, and this restriction is in part a result of the inadequacy of gray scale displays to capitalize on the full range of human visual perception (25). Hence, the application of color to medical images was undertaken (20).

COLOR IN MEDICAL IMAGING

The perception of color comprises three principal elements: hue, lightness, and saturation (26). A much
wider array of colors can be discriminated by the human eye than can levels of brightness (i.e., intensity) (25). As a result, color visualization possesses inherent advantages over gray scale presentation of certain data when they are to be interpreted by the human eye (25,26).

Color imaging has been demonstrated to adequately exploit human visual perception as well as scan sequence data (7,20,25,26), by augmenting the information available for visual inspection and thereby increasing the delineation of structural boundaries (6,8). Farrell (27) aptly described the benefits of using color in complex data interpretation by stating that “a color display has greater information capacity than a monochromatic display. Each display point has three attributes with a color display (hue, lightness, and saturation) and only one attribute with a monochromatic display (lightness)” (p. 197).

Most of the spectrum of colors visible to the human eye can be reproduced in an additive manner through varying combinations of red, green, and blue (25). Among the numerous methods for applying color to MR images is an older approach called “pseudocolor imaging,” by which one or several primary colors are arbitrarily assigned to the intensity values of an originally gray scale image (20,25). Using pseudocolor for medical imaging has become obsolete in light of sophisticated software applications that produce “color composite” and “classified” color images (7,8,13,18,20). However, debate exists regarding which of these newer techniques produces the more viable (i.e., most informative and anatomically accurate) image (7).

Color composite images can be produced using either two or three of the MR sequences (7,8). The three sequence technique has become much easier to perform since the transition from 8 bit to 24 bit color graphics displays (13). Brown et al. (7,8) and Phillips et al. (13) described these two methods in detail. As a first step in creating composite images, pixel intensities within each sequence were normalized so that 0 represented the least intensity and 255 the maximum possible intensity value. Subsequently, the intensity values of neighboring pixels determined the additive manner in which color was applied to each sequence. The contribution of each color to the images was chosen to maximize the different tissue characteristics of each MR sequence and could therefore be represented as a percentage value. In two sequence color composites, red was assigned to one image and both green and blue to another. When all three sequences were used, red, green, and blue were each assigned to a separate image, with the T1 sequence always being allocated to red (13). The additive nature of three sequence color compositing maximizes preservation of the original sequence data.

Statistical pattern recognition techniques have also been used to apply color to MR images (7). Brown and associates (7) applied a maximum likelihood (ML) algorithm to produce classified color images. In their method, regions of interest (ROIs) were first chosen from tissue types on images that had been verified as falling within a particular anatomic or pathologic classification. Numerical values representing different colors were then chosen from a color table and applied to each ROI. These ROIs were subsequently used in conjunction with an ML
algorithm to classify each pixel in a given image. Color values were applied to the rest of the image in relation to the signal intensity characteristics of a pixel and its similarity to a given ROI (i.e., tissue type) to which a color value had been previously applied. Inherent data reduction poses a limitation for this pattern recognition technique.

In a study comparing the viability of composite and classified color images, Brown and colleagues (7) found the overall quality of the classified images inferior to both the color composite and the original gray tone images. In addition to superior quality, color composite images were easier to create. The utility of pattern recognition techniques for producing classified color images was not dismissed, however, but appeared to lie in more restricted applications such as "the enhancement of specific tissues, particularly for the identification of probable sites of tumor metastases" (p. 154).

Alfano et al. (6) described a somewhat different method for rendering color composite images using the combined information from two MR sequences. These investigators termed their methodology "quantitative magnetic color imaging" and distinguished it from other approaches through their use of relaxation data rather than signal intensity values in generating color images.

The method for generating the Superimages investigated in the current study is best described as color compositing of a multispectral image that is derived by superimposing T1-, T2-, and PD-weighted images in a voxel-by-voxel alignment (see Fig. 6). As previously mentioned, production of this image does not involve data reduction and capitalizes on human visual perception through its use of color to enhance the delineation of brain structures.

PURPOSE OF THE CURRENT STUDY

The present investigation was designed to assess the efficacy of the Superimage for enhancing the visualization of subcortical nuclei (e.g., basal ganglia). As a quantitative test of the utility of the method, the reliability of manually tracing the GP using a Superimage and a gray scale image was also investigated. It was hypothesized that Superimage parcellation would result in higher reliability ratings than those previously attained at this site utilizing a gray scale T1-weighted image.

METHODS

Subjects

Two separate tracing studies were undertaken to assess the reliability of parceling the GP using the gray scale (T1-weighted) image and the Superimage. In the first study, two tracing technicians parcelled the GP using the gray scale (T1-weighted) image. Each technician in this study traced the GP on the scans of 30 subjects (22 control subjects and 8 patients). This sample comprised 22 men and 8 women with a mean age of 28.43 years (range 18–49 years, SD 8.14 years) that were randomly selected from our full MR database. The second study, which assessed the reliability of parceling the GP using the Superimage, was completed by two different tracing technicians. For this study, each technician traced the ROI on the scans of 10 subjects. This subject sample comprised five female and five male patients, all with schizophrenia, whose mean age was 29.70 years (range 19–51 years, SD 9.75 years).

The tracing technicians were blind to each other's work and to the diagnostic status (i.e., patient versus control) of the subjects' scans on which they worked. In accordance with institutional review board requirements, informed consent was obtained from all subjects after they were made aware of the procedures involved with this study.

MR Data Acquisition

Images were obtained on a 1.5 T GE Signa MR scanner. Three different sequences were acquired for each subject. T1-weighted images, using a spoiled GRASS sequence, were acquired with the following parameters: 1.5 mm coronal slices, 40° flip angle, 24 ms TR, 5 ms TE, 2 NEX, 26 cm FOV, and 256 × 192 matrix. The PD- and T2-weighted images were acquired with the following parameters: 3.0 or 4.0 mm coronal slices, 36 ms TE (for PD) or 96 ms TE (for T2), 3,000 ms TR, 1 NEX, 26 cm FOV, 256 × 192 matrix, and an echo train length of 8.

All images were rated for overall quality for movement artifacts by using a scale of 0–4 (4, excellent; 0, very poor). This scale included degradation due to motion artifacts, susceptibility artifacts, and loss of portions of the image due to the presence of metal, whole-brain coverage, and signal-to-noise ratio. Images with a quality rating of 4 were free of artifacts of any sort; 3, minor artifacts appeared in the image, such as slight motion artifacts or a slight signal intensity increase around the sinus area; 2, prominent motion artifacts or loss of small portions of the image existed owing to the presence of metal; and 1 or 0, deemed unusable in any further analysis owing to artifacts that dominated the image in general. Three separate ratings were acquired for each image for the anterior, middle, and posterior portions of the image sets. Only images with an average rating of 2.5 or greater were used in this study.

Postacquisition Processing

Immediately after acquisition, the data were transferred to our image-processing laboratory to undergo postacquisition processing on workstations (Silicon Graphics, Mountain View, CA, U.S.A.) by using locally
developed software (BRAINS; Iowa Mental Health Clinical Research Center, Iowa City, IA, U.S.A.) (3,4). The T1-weighted images were spatially normalized and resampled to 1.0 mm³ voxels so that the anteroposterior axis of the brain was realigned parallel to the anteroposterior commissure line and the interhemispheric fissure was aligned on the other two axes. The T2- and intermediate-weighted images were then aligned to the spatially normalized T1-weighted image by means of an automated image registration program (28,29). The data sets were then segmented by means of a Bayesian classifier based on discriminant analysis to reduce the variability in signal intensity across individual image sets and to correct for partial volume effects (30,31).

Superimages

As described above, T1-weighted images were used as the standard to which the T2- and PD-weighted images were spatially aligned. For the purposes of producing Superimages, a software program option was designed that aligned the three MR sequences voxel by voxel for presentation as a combined image on a Silicon Graphics workstation. The three MR sequences were assigned to the different components of red, green, and blue color spaces. Images of different parameters were fed into the different image memories that were displayed in red, green, or blue. The T1 image was assigned to red, the T2 image to green, and the PD image to blue. The color display unit enabled varying combinations of the three primary colors (red, green, blue), resulting in the capacity for a vast array of different color representations of the tissues and structures being viewed and traced. It was clearly important to select a color palette that enhanced the brain structure(s) of interest. Selecting the palette for a structure is a qualitative process, entailing modification of color combinations while viewing the entire image. As a feature of BRAINS2 software, new color palettes may be created for different brain structures of interest. An important technical note: It was found that to ensure the consistency of the color palette across different images, the slopes of the three primary colors could not be vertical. Color application as represented by a vertical slope was found to be artificially restricted within the image being viewed. This in turn increased the risk of imposing false structural boundaries dependent on signal intensity and tissue composition. A unique color palette was developed for tracing the GP (see Results for more detail).

GP Tracing

As part of the basal ganglia, the GP is an important subcortical structure that has been found to be morphologically abnormal in patients with schizophrenia (32,33). The GP consists of a mixture of WM and GM, making its boundaries hard to discern, particularly on a monochromatic display.

Tracing guidelines for the GP will be reviewed only briefly, given that the primary purpose of the current study was to investigate the utility of the Superimage methodology. The guidelines were as follows: Moving anteriorly to posteriorly (rostral to caudal), the lateral border was defined by the putamen throughout all slices on which the GP could be visualized. The superior border was defined by the joining of the internal capsule and putamen on more rostral slices and solely by the internal capsule further caudally. The medial boundary was defined by the internal capsule throughout. Finally, the inferior border was defined as the temporal limb of the anterior commissure (TLAC). Prior to tracing, slices were identified in the coronal plane on which the TLAC could be fully visualized. Cross-hairs (an available software option) were then set on the TLAC and left at these coordinates as a way of defining the inferior border of the GP on all coronal slices on which it was traced. As a point of interest: Although reliability and volumetric estimations were done using only traces laid down in the coronal plane, guide traces were made in the axial plane to aide the technicians in identifying the superior, medial, and lateral borders of the GP.

Qualitative Analyses

Qualitative assessment entailed visual inspection of the basal ganglia and surrounding structures on three unimodal gray scale images, two multispectral gray scale images, and the Superimage (see Figs. 1-6).

Statistical Analyses

Interrater reliability was estimated [overlap and intraclass correlation (ICC)] for right and left GP volumes traced in gray scale (T1-weighted image) and in multispectral color. Although overlap (i.e., the intersection of the two technicians' traces divided by the union of their traces) indicated the degree to which the two traces paralleled the same ROIs, the ICC value allowed for an estimation of the consistency and interchangeability of the two traces in applying the (GP tracing) methodology.

RESULTS

Prior to qualitative analysis and parcellation of the GP, a color palette was developed that enhanced visualization of this ROI.

Color Palette

The color palette that was selected for visualizing the GP was a mixture of pink, blue, and purple hues (see Fig. 6). This palette was chosen because it aided delineation of the lateral boundary, the putamen (visible in blue tones), and the medial and superior boundaries, the in-
ternal capsule (visible in pink tones). The anterior commissure used as a guide for the inferior boundary of the GP was distinctly visible in a deep, fluorescent pink. This particular palette facilitated the separation of the GP from surrounding white matter and other brain tissues. It was consequently used for the entire study for tracings done with the Superimage.

**Qualitative Analysis**

Qualitative analysis revealed an improvement in the visualization of subcortical structures on the Superimage as compared with all gray scale images. The Superimage particularly enhanced the entire structure of the GP, which is seen less distinctly on the gray scale images and as almost all WM on the discrete image. The distinction between the GP and its medial boundary, the internal capsule, was most remarkably improved. The desire to parcelate and estimate volumes for subcortical structures provided the genesis for this new technique. Anecdotally, the Superimage has provided the improvements in image enhancement for which it was conceived. Improved visualization of brain structures using this method has led to decrements in labor intensity and enhanced the accuracy of complex structural parcellation at our site.

**Quantitative Analysis**

Reliability estimates for manual tracings of the GP done using gray scale (T1-weighted) and multispectral color images were as follows (see Table 1): Overlap values for gray scale parcellation were 0.71 (left GP) and 0.68 (right GP). The consistency and interchangeability of the tracers as represented by the ICC ($r^2$) values were 0.64 and 0.30 for the left and right GP, respectively. Superimage parcellation resulted in overlap values of 0.80 (left GP) and 0.78 (right GP) and ICC ($r^2$) values of 0.72 (left GP) and 0.71 (right GP). Thus, the use of the Superimage improved reliability, as shown by both indexes. This indicates that the superior Superimage face validity (i.e., its visual superiority) is also supported by quantitative data, indicating that it enhances reliability. Given the complexity of anatomic boundary delineation of the GP, these reliability values are considered to be very favorable.

**TABLE 1. Reliability values for Superimage and gray scale parcellation of globus pallidus (GP)**

<table>
<thead>
<tr>
<th>Method</th>
<th>Left GP overlap</th>
<th>Right GP overlap</th>
<th>Left GP ICC</th>
<th>Right GP ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superimage</td>
<td>0.80</td>
<td>0.78</td>
<td>0.72</td>
<td>0.71</td>
</tr>
<tr>
<td>Gray scale (T1 weighted)</td>
<td>0.71</td>
<td>0.68</td>
<td>0.64</td>
<td>0.30</td>
</tr>
</tbody>
</table>

ICC, intraclass correlation.

**DISCUSSION**

The current study was undertaken to evaluate a new method for applying color to multispectral images. This technique was developed to enhance the visualization of hard-to-see brain structures such as subcortical nuclei. Structures of this type are often difficult to discern on gray scale display. Although numerous procedures exist for creating multispectral images (1,6,8,17), not all of these techniques are appropriate to every brain structure of interest. Segmentation of gross tissue classes using a discrete multispectral image is highly reliable (1). Unfortunately, the discrete method is vulnerable to the effects of partial voluming.

The GP, given its size and mixed tissue composition (i.e., WM and GM), presented an appropriately challenging structure with which to assess the efficacy of this new methodology. Qualitative assessment was undertaken using a variety of gray scale unimodal and multispectral images and the Superimage (Figs. 1–6). The new Superimage method, which afforded full preservation of the original MR sequence data, maximized visualization of the GP, caudate, and putamen. Enhancement of the medial boundary of the GP was particularly remarkable. Color visualization improved the delineation of structural boundaries above that seen in unimodal and multispectral gray scale.

Reliability values for gray scale and Superimage parcellation of the GP were then estimated and compared. This provided a quantitative assessment of the efficacy of the Superimage for enhancing visualization and reliability. Interrater reliability estimates of Superimage tracings were considerably improved beyond those obtained using a gray scale (T1-weighted) image. Tracers were also more consistent and interchangeable using the Superimage.

In addition to facilitating parcellation of the GP, other basal ganglia structures, and the whole thalamus, the Superimage is expediting measurements of even more exigent subcortical structures that are ongoing in our lab. In addition to manual estimations, ROIs that have been parcelled are “taught” to an artificial neural network (ANN) once tracing technicians attain adequate reliability. Labor intensity is substantially reduced using the ANN, with parcellation time being one half or less that required for manual tracing. Given that the ANN is only as reliable as the human traces from which it “learns” (34), any measure that improves validity and reliability also advances the research process. The Superimage appears to be a tool that will meet those demands. As ANNs are increasingly used for research, questions regarding the etiology, course, and treatment of disease entities will be addressed more rapidly.

**Acknowledgment:** This research was supported in part by NIMH grants MH19113, MH60990, MH40856, and MHCRC 43271.
REFERENCES


