The Relation between Color Spaces and Compositional Data Analysis Demonstrated with Magnetic Resonance Image Processing Applications

Omer Faruk Gulban

Maastricht University

Abstract

This paper presents a novel application of compositional data analysis methods in the context of color image processing. A vector decomposition method is proposed to reveal compositional components of any vector with positive components followed by compositional data analysis to demonstrate the relation between color space concepts such as hue and saturation to their compositional counterparts. The proposed methods are applied to a magnetic resonance imaging dataset acquired from a living human brain and a digital color photograph to perform image fusion. Potential future applications in magnetic resonance imaging are mentioned and the benefits/disadvantages of the proposed methods are discussed in terms of color image processing.

Keywords: compositional data analysis, color, magnetic resonance imaging, image fusion.

1. Introduction

Compositional data analysis can be applied to color images in the context of image processing and analysis. Color images are stored as triplets of non-negative integers representing the additive primary colors red, green and blue (referred as RGB channels). Once the color image is formed, the transformation from RGB to hue, saturation and intensity (HSI) coordinate system is an often used first step to improve the image visualization (Smith 1978; Ledley, Buas, and Golab 1990; Grasso 1993) (for the mathematical expression of this transformation see Appendix A.2). HSI color space is commonly used in computer applications because of the commonly accepted property of its components relating to human color perception (Joblove and Greenberg 1978; Levkowitz and Herman 1993; Pohl and van Genderen 2016). Hue relates to the dominant component of the primary colors (red, green, blue) in the mixture, saturation relates to the distance from an equal mixture (gray), and intensity relates to distance from total darkness (black to white). In this article, inspiration is drawn from RGB to HSI color space transformation to propose an analogous method based on compositional data analysis. During this process, a simple vector decomposition is outlined to justify the use of compositional data analysis (Aitchison 1982; Pawlowsky-Glahn, Egozcue, and Tolosana-Delgado 2015) and the simplex space is leveraged to manipulate vector fields to perform image enhancement. The proposed method is demonstrated using a magnetic resonance imaging

(MRI) dataset. The results are also visualized using a digital color photograph to provide a general intuition. Due to the interdisciplinary nature of this work a glossary is provided in Appendix A.1 to accompany readers from different fields.

2. Materials and methods

2.1. MRI data acquisition and preprocessing

Whole head T1 weighted (T1w), proton density weighted (PDw) and T2* weighted (T2*w) images at 0.7 mm isotropic volumetric cubic element (voxel) resolution were acquired in one male participant using a three dimensional magnetization prepared rapid acquisition gradient echo sequence with a 32-channel head coil (Nova Medical) on a 7 Tesla whole-body scanner (Siemens) in Maastricht Brain Imaging Center. These measurements reflect different intrinsic properties of the tissues, for instance T1w image shows the optimal contrast between white matter and gray matter, PDw image shows the density of the hydrogen atoms, T2*w image shows the iron content. For the detailed report of the data acquisition parameters see Appendix A.3. For a review on ultra high field MRI (($\geq$ 7 Tesla)) see Ugurbil (2014).

As a standard preprocessing step, a masking operation was performed based on the PDw image using FSL-BET software (version 2.1; Smith (2002); Smith, Jenkinson, and Woolrich (2004)). The resulting volumetric mask was applied to all images in order to discard parts containing non-brain tissues (e.g. bones, muscles, skin, air). This masking step was necessary to reduce the overall processing time in the following analyses by reducing the total number of voxels from 26214400 ($320 \times 320 \times 256$) to $\sim$ 4.5 million (4543582 to be exact). Measurements (T1w, PDw and T2*w) stored at each one of the $\sim$ 4.5 million voxels are considered as components of three separate scalar fields constituting a vector field when combined.

2.2. Barycentric decomposition

The image analysis illustrated in this paper starts from the following vector decomposition, where the real space of n dimensions is indicated with $\mathbb{R}^n$ and the simplex space is indicated with $\mathbb{S}^n$ symbols:

$$\vec{v} = [v_1, v_2, \ldots, v_{n-1}, v_n], \text{ where } v_1, v_2, \ldots, v_{n-1}, v_n \in \mathbb{R}^n_{>0},$$

$$\vec{v} = \frac{1}{k}C(\vec{v})s, \text{ where } C(\vec{v}) \in \mathbb{S}^n, \text{ and } s \in \mathbb{R}^1_{>0}. \tag{1}$$

where $\mathbb{R}^n_{>0}$ indicates positive real numbers, k is an arbitrary scalar and the letter C stands for the closure operation used in compositional data analysis (Aitchison 2002; Pawlowsky-Glahn *et al.* 2015):

$$C(\vec{v}) = k\frac{\vec{v}}{s}, \tag{2}$$

The letter s is another scalar that is the sum of the vector components:

$$s = \sum_{i=1}^{n} v_i. \tag{3}$$

Note that k disappears from the expression in Eq. 1 and 2 when selected as one. The decomposition (Eq. 1) of the vector $\vec{v}$ into its barycentric coordinates ($C(\vec{v})$) and a scalar (s) allows the application of the compositional data analysis to the barycentric coordinates. Historically, the closure operation was used by August Ferdinand Mobius (Fauvel, Flood, and Wilson 1993) and the resulting vector was interpreted as relating to the barycenter (center of mass) of a simplex, hence giving the name to the vector decomposition demonstrated here.

2.3. Compositional image analysis

In this section, operations performed on the the barycentric coordinates are laid out. These operations are mostly adapted from Pawlowsky-Glahn *et al.* (2015) to fit the context of the analyzed vector field. Tensor notation is used for explicit formulations:

Let A be a tensor with three dimensions, $A_{x,y,z}$, where the coordinates x, y, z of an element represents the spatial location. As mentioned the in Section 2.1, there are ~ 4.5 million tensors (i.e. voxels) of interest in total.

Let the superscript of tensor A indicate different MRI measurements acquired at every voxel; $A_{x,y,z}^{T1w}$, $A_{x,y,z}^{PDw}$, $A_{x,y,z}^{T2^*w}$. As the first step, all three tensors are vectorized (flattened):

$$\mathsf{vec}(A^t) = [a_{1,1,1}, \ \dots, \ a_{x,1,1}, a_{1,2,1}, \ \dots, \ a_{1,y,1}, a_{1,2,2}, \ \dots, \ a_{1,2,z}]^T. \tag{4}$$

T stands for the transpose operator and $t \in T1w, PDw, T2^*w$.

Concatenation denoted by $\|$ is used to matricize the vectorized tensors as the next step:

$$V = \mathsf{vec}(A^{T1w}) \ \| \ \mathsf{vec}(A^{PDw}) \ \| \ \mathsf{vec}(A^{T2w}), \tag{5}$$

where number of rows of the matrix V is $x \times y \times z$ ($n = \sim 4.5$ million), and the number of columns is 3. These operations are done for convenient indexing in the following equations.

Let V indicate the set of voxels where the vector v_i is the composition of T1w, PDw and T2*w measurements:

$$V = [v_i, \ \dots, \ v_n] \text{ where } i \in [1, 2, \ \dots, \ n] \text{ and } v_i = [v_i^{T1w}, v_i^{PDw}, v_i^{T2^*w}]. \tag{6}$$

At this stage, $V \in \mathbb{R}^3$. Voxel-wise (i.e. row-wise) closure is applied to V to acquire the barycentric coordinates of every composition:

$$X = C(V) = [C(v_i), \ \dots, \ C(v_n)], \ X \in \mathbb{S}^3. \tag{7}$$

The set of compositions X was centered by finding the *sample center* and *perturbing* each composition with the inverse of the sample center:

$$\hat{X} = X \oplus \mathsf{cen}(\mathbf{X})^{-1}, \tag{8}$$

where $\oplus$ denotes the perturbation operator (analogous to addition in real space):

$$x \oplus y = C[x_{T1w}\mathbf{y}_1, x_{PDw}\mathbf{y}_2, x_{T2^*w}\mathbf{y}_3]. \tag{9}$$

Multipliers of the components are indicated with $\mathbf{y}_1$, $\mathbf{y}_2$ and $\mathbf{y}_3$. The term $\mathsf{cen}(X)$ is a vector that stands for the component-wise (i.e. column-wise) geometric mean across all voxels:

$$\mathsf{cen}(X) = [\mathbf{g}_{T1w}, \mathbf{g}_{PDw}, \mathbf{g}_{T2^*w}] = \left(\prod_{i=1}^{n} x_{ij}\right)^{1/n}, j = [T1w, PDw, T2^*w]. \tag{10}$$

After centering, the data is standardized:

$$\hat{\hat{X}} = \hat{X} \odot \mathsf{totvar}[X]^{-1/2}, \tag{11}$$

where $\odot$ symbol stands for the power operator (analogous to scaling in real space). The exponent p is applied to every component:

$$x \odot p = C[x_{T1w}^p, x_{PDw}^p, x_{T2*w}^p], \tag{12}$$

and the total variance is computed by:

$$\mathsf{totvar}[X] = \frac{1}{n}\sum_{i=1}^{n} d_a^2(x_i, \ \mathsf{cen}(X)), \tag{13}$$

where d_a^2 indicates squared Aitchison distance:

$$d_a(x, y) = \sqrt{\frac{1}{2D} \sum_{j=1}^{D} \sum_{k=1}^{D} \left(\ln \frac{x_j}{x_k} - \ln \frac{y_j}{y_k} \right)^2}. \tag{14}$$

In the current example $D = 3$ because of operating on a set of three part compositions X (see Equation 7).

At this stage some visual intuition could be gained with regards to the vector field (MR images) that is decomposed and processed via compositional data analysis methods. For illustration purposes, this is done by generating virtual image contrasts computed through metrics of simplex space (see Figure 2).

The norm in S^3 image in Figure 2 is generated by computing Aitchison norm voxel-wise:

$$\| X \|_a = [\| x_i \|_a, \ \ldots, \ \| x_n \|_a]. \tag{15}$$

The subscript a stands for the vector norm defined in simplex space (analogous to Euclidean norm in real space). Following Pawlowsky-Glahn *et al.* (2015), Aitchison norm is defined as:

$$\| x \|_a = \sqrt{\frac{1}{2D} \sum_{j=1}^{D} \sum_{k=1}^{D} \left(\ln \frac{x_j}{x_k} \right)^2}, \tag{16}$$

where $D = 3$ considering three different MRI measurements (T1w, PDw, T2*w).

It should be noted that this norm image is analogous to saturation dimension in the aforementioned RGB to HSI color space transformation (see Appendix A.2).

The angular difference in S^3 image in Figure 2 is the voxel-wise angular difference ($\angle$) between the set of compositional vectors (X) and a reference vector (r):

$$X_{\angle} = [\angle x_i r, \ \ldots, \ \angle x_n r], \tag{17}$$

where

$$\angle x_i r = \arccos \left(\frac{\langle x_i, r \rangle_a}{\| x_i \|_a \| r \|_a} \right). \tag{18}$$

$\| x \|_a$ stands for Aitchison norm of the vector x as defined in Equation 16 and $\langle x, r \rangle_a$ stands for inner product in simplex space:

$$\langle x, r \rangle_a = \frac{1}{2D} \sum_{j=1}^{D} \sum_{k=1}^{D} \ln \frac{x_j}{x_k} \ln \frac{r_j}{r_k}. \tag{19}$$

Here, $D = 3$ again and it should be noted that the choice of the reference vector r is arbitrary. In this example the reference vector is selected as $r = [0.05, 0.9, 0.05]$. This angle difference image is similar to hue dimension in RGB to HSI color space transformation (see Appendix A.2).

Recognizing the similarity of norm and angular difference in simplex space to saturation and hue dimensions of HSI color space at this stage leads to the proposal of a color balance algorithm which is used together with centering (Eq. 8) and standardization (Eq. 11) to generate the color image in Figure 3 lower row:

$$\mathsf{truncate}(x) = \begin{cases} x \odot \left(\dfrac{\| x \|_a}{\lambda} \right) & , \| x \|_a > \lambda \\ x & , \| x \|_a \leq \lambda \end{cases} \tag{20}$$

This method is inspired from other simple color balance methods that makes use of the dynamic range of RGB channels of rendered color images (Limare, Lisani, Morel, Petro, and

Sbert 2011). Truncate function is useful when when there are a small amount of outliers compositions. In essence this functions pulls the compositions that are too far away from the center of the simplex space towards the center. In the current work, distance threshold was arbitrarily chosen $\lambda = 3$ based on visual inspection of the resulting color enhancement. It is conceivable that other operations based on percentiles can also be used to the make the decision less arbitrary (e.g. 99th percentile of Aitchison norm distribution consisting of all vectors of X).

The ilr transformation (Egozcue, Pawlowsky-Glahn, Mateu-Figueras, and Barceló-Vidal 2003; Pawlowsky-Glahn *et al.* 2015) was performed voxel-wise to acquire real space coordinates (ilr coordinates) of the set compositions to generate Figure 3 right column:

$$X_{\mathsf{ilr}} = [\mathsf{ilr}(x_i), \ \ldots, \ \mathsf{ilr}(x_n)], \tag{21}$$

where the ilr transformation is defined as:

$$\mathsf{ilr}(x_i) = \ln(x_i) \cdot \mathbf{H}. \tag{22}$$

$\mathbf{H}$ indicates the Helmert sub-matrix, chosen by following Tsagris, Preston, and Wood (2011); Lancaster (1965). In this case $\mathbf{H}$ consists of 3 rows and 2 columns and selected as the following:

$$\mathbf{H} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ 0 & -\sqrt{\frac{2}{3}} \end{bmatrix}. \tag{23}$$

To explore the usefulness of the ilr coordinates and probe physically interpretable compositional characteristics, a real-time interactive visualization method is used with joint ilr-coordinates and image space representations of the data. This method is similar to using pre-defined modulation transfer functions for mapping the bins of a 2D histogram data representation to 3D image space used in volume rendering software (Kniss, Kindlmann, and Hansen 2005). Major brain tissues such as gray matter, white matter, cerebrospinal fluid arteries and sinuses are identified manually by the author and delineated in both ilr-coordinates and brain images using neuro-anatomical expertise.

All analysis steps demonstrated in this work are implemented in a free and open source Python package (available at https://github.com/ofgulban/compoda; Gulban (2018)) using Numpy (Van der Walt, Colbert, and Varoquaux 2011), Scipy (Jones, Oliphant, Peterson *et al.* 2001–), Matplotlib (Hunter 2007), Nibabel (Brett, Hanke, Côté, Markiewicz, Ghosh, Wassermann, Gerhard, Larson, Lee, Halchenko, Kastman, M, Morency, Maloney, Rokem, Cottaar, Millman, jaeilepp, Gramfort, Vincent, McCarthy, van den Bosch, Subramaniam, Nichols, embaker, markhymers, chaselgrove, Basile, Oosterhof, and Nimmo-Smith 2017) and scikit-image (Van der Walt, Schönberger, Nunez-Iglesias, Boulogne, Warner, Yager, Gouillart, Yu, and the scikit-image contributors 2014) as auxiliary scientific libraries. The joint exploration of tissues in ilr-coordinates and image space is performed by using a specialized software developed by the author (Gulban and Schneider 2017).

3. Results

Figure 1 shows the similarity between HSI and the analogous metrics demonstrated in methods section. When comparing Fig. 1 rows 2-3 column 1 to column 3, it can be seen that hue and angular difference in simplex space are both invariant to intensity gradient visible in the sky (brightness change from upper right corner to lower left). The same observation can also be made for the saturation and norm in simplex space images (compare Fig. 1 rows 2-3 column 2 to column 3). These observations can be related to the scale invariance principle of compositional data analysis. By applying the barycentric decomposition (Eq.1) to color triplets in each pixel, scale information is separated, revealing compositional vectors invariant
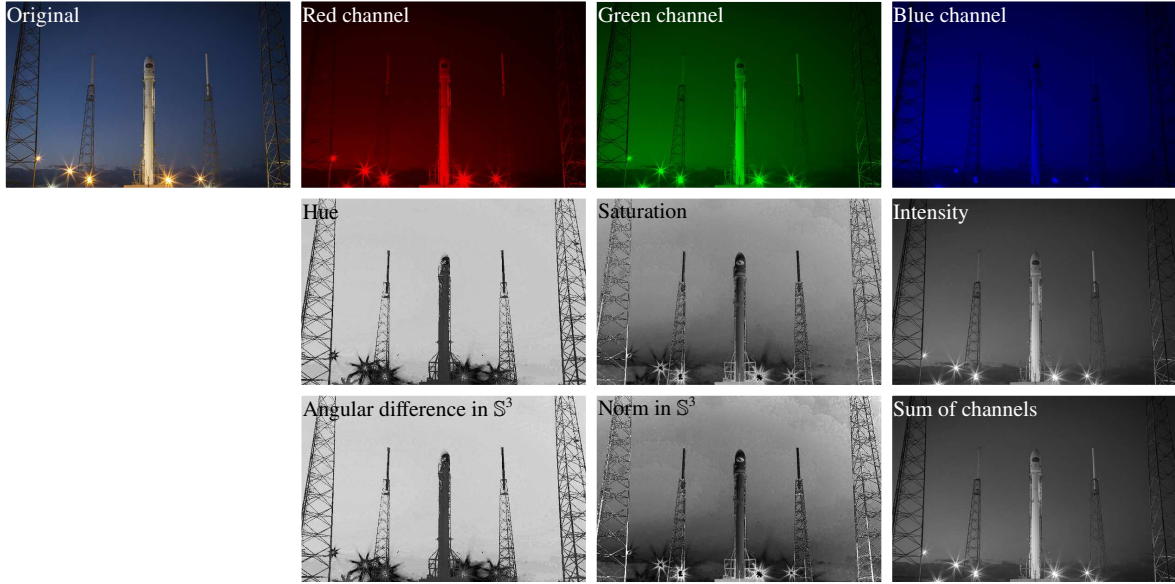
Figure 1: Visual demonstration of the similarity between RGB to HSI transformation and the corresponding analogous concepts laid out in Sections 2.2 and 2.3. In the first row the original image consisting of three channels and each of its channels can be seen. The second row shows hue, saturation and intensity components after RGB-HSI transformation [A.2]. The third row shows angular difference in simplex space (analogous to hue), norm in simplex space (analogous to saturation) and derived by using compositional data methods (Equations 4-18 and the scalar component in real space (Equation 3). Color maps of all scalar images range between the minimum and the maximum values of the corresponding image (darker shades are assigned to low values). The image is acquired from official SpaceX photo gallery (https://www.flickr.com/photos/spacex) licensed under CC0 1.0 public domain.

to the underlying multiplicative scalar field (i.e. brightness). These image contrasts based on compositional metrics are useful in object recognition tasks, see how the clouds are invisible in hue and angular difference images or the well-defined tip of the rocket in saturation and norm images. Figure 2 depicts two different slices of the 3D brain images visualized in gray scale depicting different image contrasts. It can be seen that the smooth, artefactual multiplicative scalar field (referred as bias field in MRI literature, see arrows with asterisks in Fig. 2) is separated from the compositional components (compare Fig. 2 rows 1-2 column 2 with row 3 column 2 in both panels). This is similar to the separation of the brightness gradient in the sky in Fig. 1 as mentioned in the previous paragraph. The compositional image contrasts can be considered as virtual contrasts which enhance the visual appearance by being invariant to certain artefactual features. It should be mentioned that other image contrasts can be generated by changing the reference vector in angular difference images (see Eq. 18) to highlight specific tissues. Similarly, the compositions can also be perturbed (see Eq.8) differently to generate different contrasts in norm images. The generation of task-specific virtual contrasts could potentially be useful in medical imaging applications since the input channels and the virtual contrasts are physically interpretable.

For the physical interpretation of the compositional vectors, a joint representation of color images and ilr coordinates of compositional vectors are presented in Figure 3. The difference between the rows in left column demonstrates the effect of color balance (see Equations 8 and 11). It can be seen that the image contrast is enhanced to the level of easily recognizing major brain tissues by associating them to different colors. This is a more convenient visualization compared to inspecting T1w, PDw, T2*w image contrasts individually considering that the color image contains information from all three inputs at the same time. Right column in Fig. 3 shows the 2D histograms of the ilr coordinates (see Eq. 22) of the vector compositions.
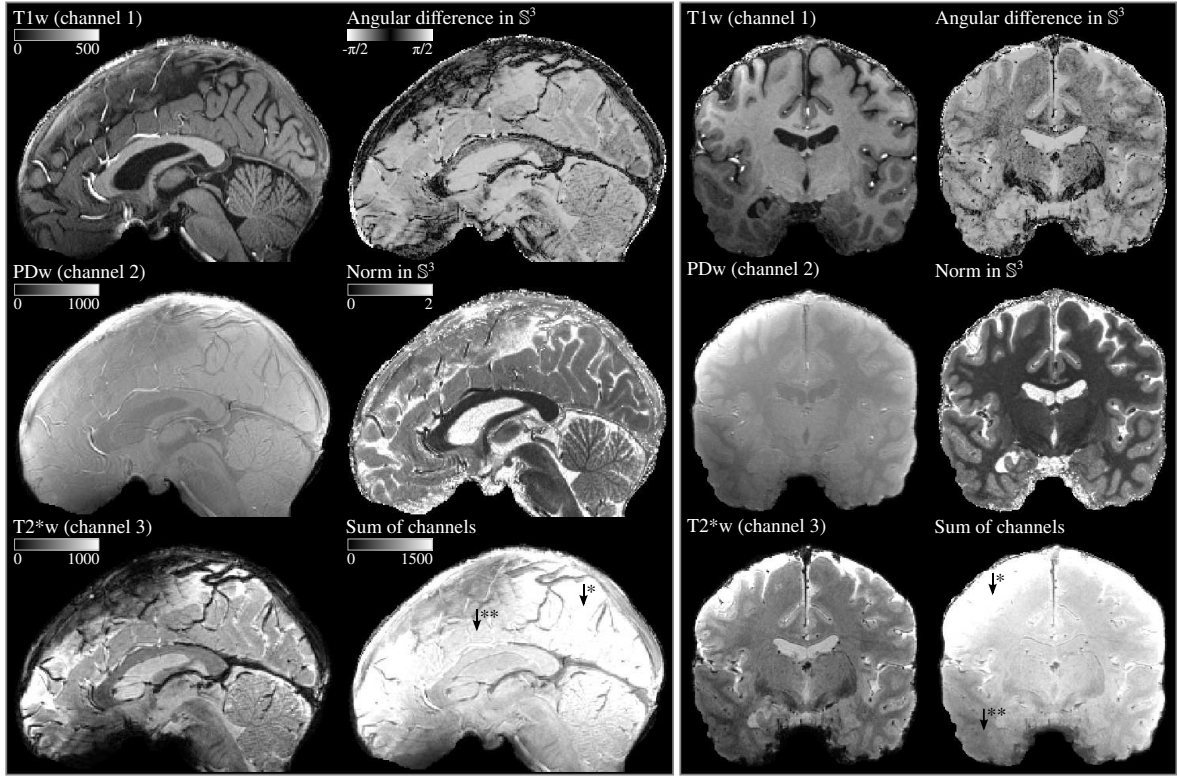
Figure 2: T1w, PDw and T2*w MR images rendered together with angular difference in simplex space (Eq. 18), norm in simplex space (Eq. 16) and scalar channel sum component (Eq. 3). Left panel shows the sagittal slices (y-z plane) and the right panel shows a coronal slices (x-z plane) of the 3D images. Arrows with one asterisk (*) indicates the parts of the images that are too bright and arrows with two asterisks (**) indicates the parts that are too dim as a result of data acquisition imperfections. The legends on the upper left corner in left panel shows the colors bars which are also valid for the images in right panel.

The effect of color balance becomes apparent when the compositions are interpreted with regards to the embedded RGB color cube primary axes (real space axes). For instance most of the clusters are on the left hand side of the center, close to the green arrow (PDw). This indicates that PDw measurements were initially dominating the compositions causing the green heavy color image. After the color balance, compositions spread more equally along the primary color axes, indicating that color contrast in the image will be richer (i.e. less dominated by a single component). The color balanced visualization is more useful for human observers in terms of intuitively recognizing the tissues. For instance the arteries have mostly reddish-white colors and the sinuses appear in green, which corresponds to the areas delineated for these tissues in ilr coordinates when the positions of clusters are considered relative to the embedded primary axes of RGB color cube. Although both of these are blood vessels, the difference between arteries and sinuses is meaningful because sinuses contain mostly deoxygenated hemoglobin, leading to a rapid decay of the MRI signal. In contrast, arteries contain oxygenated blood with slower MR signal decay. The difference in signal decay times of blood vessels effects T2*w and PDw measurements separating the compositional properties in relation of T1w images. Similarly the compositional change from white matter to gray matter to cerebrospinal fluid can be seen as an approximately straight line which corresponds to the change from red-heavy color of white matter to cyan of cerebrospinal fluid. It can be seen that these tissues are not dispersed towards PDw or T2*w axes, which can be intepreted as PDw and T2*w measurements not revealing different compositions in relation to T1w measurements when white matter, gray matter and cerebrospinal fluid is considered.
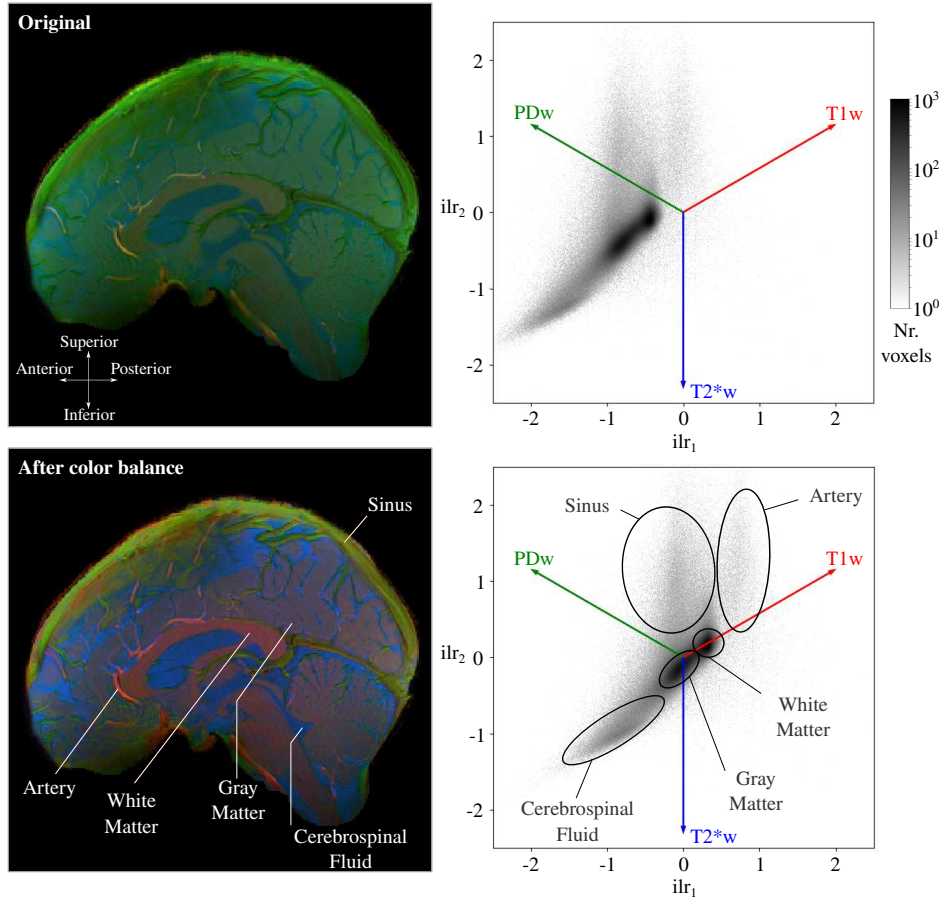
Figure 3: MRI measurements rendered as a color image. Red channel is assigned to T1w, green to PDw and blue to T2*w measurements. Left column shows the 2D histograms of the corresponding ilr coordinates (Eq. 22). The projection of primary axes of RGB color cube (in R^3) to ilr coordinates are embedded to provide an intuitive reference for the characteristics of the compositions. The effect of centering and standardization inside the simplex (Eq. 8, 11) and the compositional truncation (Eq. 20) is visible as color balance improvement in the rendered brain slice. The labels pointing to the tissues in brain image and circles in 2D ilr coordinate histograms shows the relation of the compositional characteristics with the coloration. The circles in 2D histograms are the edges of the 2D transfer functions used to manually probe the tissue-ilr coordinate relationship.

4. Discussion

In this work, compositional data analysis methods are used to reformulate RGB-HSI color space transformation. It is visually demonstrated that the compositional metrics such as angular difference and norm in simplex space relate to hue and saturation concepts in color space literature. This reformulation of hue and saturation would be advantageous for having a well-principled framework when operating on images with n-dimensions. For instance, a potential future application would be to analyze MRI datasets with more than three types of measurements by adding cerebral blood volume measurements (Uludağ and Blinder 2018) or multi-echo echo planar imaging (Poser and Norris 2009) to T1, T2 and PD weighted measurements at ultra high fields. Another application area would be processing of multi-spectral images for image fusion purposes (Pohl and van Genderen 2016). In cases where more than three measurements are acquired for each element in an image, dimensionality reduction methods such as principal component analysis in simplex space (Wang, Shangguan, Guan, and Billard 2015) would become relevant to maximize the visualized information content of color images. This is the disadvantage of being limited to three primary additive colors

for color image rendering. However the virtual image contrasts (scalar images) could still be explored without dimensionality reduction. For instance the norm in simplex space can still be straightforwardly computed for n-dimensional compositions or the angular difference images can be explored can be explored by selecting different reference vectors to track the positions of n-dimensional compositions on an n-sphere.

The disadvantages of compositional metrics presented here should also be mentioned. If the signal to noise ratio of the acquired images are low, compositional image contrasts would become less useful. For instance, the parts of images where only the measurement noise is recorded (e.g. thermal noise) the compositional metrics would carry no meaning. In such cases the angular difference and norm in simplex space would return enhanced noise patterns therefore the local brightness (i.e. intensity) should be taken into account before the compositional image analysis. The domain-specific exploration of the parameter spaces are also needed. For instance what is the extent of the usefulness of the compositional image processing when the noise properties are dissimilar between different types of measurements or when the measurements have different spatial scales (e.g. images with different resolutions). Further investigation is necessary to establish the relevance of the proposed application of compositional data analysis specifically to MRI data and generally to image processing and image fusion.

5. Acknowledgements

A. Appendix

A.1. Glossary

Barycentric coordinates: Center of mass-centric coordinates. Used in relation to the center of mass of an n-simplex.

Color balance: Used to indicate centering and standardizing compositional vectors in this work.

HSI: A three dimensional space with named axes relating to human color perception: hue, saturation, intensity.

MRI: Magnetic resonance imaging.

PDw: Proton density weighted MRI signal acquisition. Indicates density of the hydrogen atoms in different tissues.

RGB: A three dimensional space with named axes relating to color cube: red, green, blue

Simplex: Generalized geometrical notion of triangles. For example, 0-simplex is a point, 1-simplex is a line segment, 2-simplex is a triangle, 3-simplex is a tetrahedron etc.

T1w: T1 weighted MRI signal acquisition. Optimized to give the highest image contrast between white matter and gray matter brain tissues.

T2*w: T2* weighted MRI signal acquisition. Optimized to give the highest image contrast related to iron concentration between brain tissues.

Voxel: Volumetric cubic element, in other words a three dimensional pixel.

A.2. RGB to HSI transformation

$$Intensity = I = \left(\frac{R}{\max(R,\ G,\ B)} + \frac{G}{\max(R,\ G,\ B)} + \frac{B}{\max(R,\ G,\ B)} \right) \div 3.$$

$$Saturation = S = \begin{cases} 0 & ,I = 0 \\ 1 - \dfrac{\min(R,\ G,\ B)}{I} & ,I > 0. \end{cases}$$

$$Hue = H = \begin{cases} 0° & ,\Delta = 0 \\ 60° \times \left(\dfrac{G' - B'}{\Delta} \mod 6 \right) & ,C_{max} = R' \\ 60° \times \left(\dfrac{B' - R'}{\Delta} + 2 \right) & ,C_{max} = G' \\ 60° \times \left(\dfrac{R' - G'}{\Delta} + 4 \right) & ,C_{max} = B' \end{cases}$$

where $\Delta = \max(R,\ G,\ B) - \min(R,\ G,\ B)$

and $R' = \dfrac{R}{\max(R,\ G,\ B)}$, $G' = \dfrac{G}{\max(R,\ G,\ B)}$, $B' = \dfrac{B}{\max(R,\ G,\ B)}$.

A.3. MRI data acquisition parameters and ethics statement

Whole head images were acquired using a three dimensional magnetization prepared rapid acquisition gradient echo (MPRAGE) sequence. The data consisted of a T1w image (repetition time [TR] = 3100 ms; time to inversion [TI] = 1500 ms [adiabatic non-selective inversion pulse]; time echo [TE] = 2.42 ms; flip angle = 5°; generalized auto-calibrating partially parallel acquisitions [GRAPPA] = 3 (Griswold, Jakob, Heidemann, Nittka, Jellus, Wang, Kiefer, and Haase 2002); field of view [FOV] = $224 \times 224\ mm^2$; matrix size = 320×320; 256 slices; 0.7 mm isotropic voxels; pixel bandwidth = 182 Hz/pixel; first phase encode direction anterior to posterior; second phase encode direction left to right), a PDw image (0.7 mm isotropic) with the same 3D-MPRAGE sequence but without the inversion pulse (TR = 1380 ms; TE = 2.42 ms; flip angle = 5°; GRAPPA = 3; FOV = $224 \times 224\ mm$; matrix size = 320×320; 256 slices; 0.7 mm isotropic voxels; pixel bandwidth = 182 Hz/pixel; first phase encode direction anterior to posterior; second phase encode direction left to right), and a T2*w anatomical image using a modified MPRAGE sequence TR (= 4910 ms; TE = 16 ms; flip angle = 5°; GRAPPA = 3; FOV = $224 \times 224 mm$; matrix size = 320×320; 256 slices; 0.7 mm isotropic voxels; pixel bandwidth = 473 Hz/pixel; first phase encode direction anterior to posterior; second phase encode direction left to right). Only magnitude images are stored after reconstruction.

The experimental procedures were approved by the ethics committee of the Faculty for Psychology and Neuroscience at Maastricht University, and were performed in accordance with the approved guidelines and the Declaration of Helsinki. Informed consent was obtained from the participant before conducting the data acquisition.

References

Aitchison J (1982). "The Statistical Analysis of Compositional Data." *Journal of the Royal Statistical Society*, **44**(2), 139–177.

Aitchison J (2002). "A Concise Guide to Compositional Data Analysis." *CDA Workshop Girona*, **24**, 73–81.

Brett M, Hanke M, Côté MA, Markiewicz C, Ghosh S, Wassermann D, Gerhard S, Larson E, Lee GR, Halchenko Y, Kastman E, M C, Morency FC, Maloney B, Rokem A, Cottaar M, Millman J, jaeilepp, Gramfort A, Vincent RD, McCarthy P, van den Bosch JJ, Subramaniam K, Nichols N, embaker, markhymers, chaselgrove, Basile, Oosterhof NN, Nimmo-Smith I (2017). "nipy/nibabel: 2.2.0." `doi:10.5281/zenodo.1011207`. URL `https://doi.org/10.5281/zenodo.1011207`.

Egozcue JJ, Pawlowsky-Glahn V, Mateu-Figueras G, Barceló-Vidal C (2003). "Isometric Logratio Transformations for Compositional Data Analysis." *Mathematical Geology*, **35**(3), 279–300. ISSN 08828121. `doi:10.1023/A:1023818214614`. URL `http://link.springer.com/10.1023/A:1023818214614`.

Fauvel J, Flood R, Wilson R (1993). *Mobius and His Band: Mathematics and Astronomy in Nineteenth-century Germany*. Oxford University Press. URL `https://books.google.de/books?id=z-XuAAAAMAAJ`.

Grasso DN (1993). "Applications of the IHS Color Transformation for 1:24,000-Scale Geologic Mapping: A Low Cost SPOT Alternative." *Photogrammetric Engineering & Remote sensing*, **59**(1), 73–80.

Griswold MA, Jakob PM, Heidemann RM, Nittka M, Jellus V, Wang J, Kiefer B, Haase A (2002). "Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA)." *Magnetic Resonance in Medicine*, **47**(6), 1202–1210. ISSN 07403194. `doi:10.1002/mrm.10171`.

Gulban OF (2018). "Compoda v0.3.1." `doi:10.5281/zenodo.1207293`. URL `https://doi.org/10.5281/zenodo.1207293`.

Gulban OF, Schneider M (2017). "Segmentator v1.3.0." `doi:10.5281/zenodo.999487`. URL `https://doi.org/10.5281/zenodo.999487`.

Hunter JD (2007). "Matplotlib: A 2D graphics environment." *Computing In Science & Engineering*, **9**(3), 90–95. `doi:10.1109/MCSE.2007.55`.

Joblove GH, Greenberg D (1978). "Color Spaces for Computer Graphics." *ACM SIGGRAPH Computer Graphics*, **12**(3), 20–25. ISSN 00978930. `doi:10.1145/965139.807362`. URL `http://www.sciencedirect.com/science/article/pii/S1049965283710199http://portal.acm.org/citation.cfm?doid=965139.807362`.

Jones E, Oliphant T, Peterson P, *et al.* (2001–). "SciPy: Open Source Scientific Tools for Python." [Online; accessed 2015-07-13], URL `http://www.scipy.org/`.

Kniss J, Kindlmann G, Hansen CD (2005). "Multidimensional Transfer Functions for Volume Rendering." *Visualization Handbook*, pp. 189–209. ISSN 1077-2626. `doi:10.1016/B978-012387582-2/50011-3`.

Lancaster HO (1965). "The Helmert Matrices." *The American Mathematical Monthly*, **72**(1), 4. ISSN 00029890. `doi:10.2307/2312989`. URL `http://www.jstor.org/stable/2312989?origin=crossref`.

Ledley R, Buas M, Golab T (1990). "Fundamentals of True-color Image Processing." In *[1990] Proceedings. 10th International Conference on Pattern Recognition*, volume i, pp. 791–795. IEEE Comput. Soc. Press. ISBN 0-8186-2062-5. `doi:10.1109/ICPR.1990.118218`. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=118218http://ieeexplore.ieee.org/document/118218/`.

Levkowitz H, Herman G (1993). "GLHS: A Generalized Lightness, Hue, and Saturation Color Model." *CVGIP: Graphical Models and Image Processing*, **55**(4), 271–285. ISSN 10499652. `doi:10.1006/cgip.1993.1019`.

Limare N, Lisani Jl, Morel Jm, Petro AB, Sbert C (2011). "Simplest Color Balance." *Image Processing On Line*, **1**(1), 125–133. ISSN 2105-1232. doi:10.5201/ipol.2011.llmps-scb. URL http://www.ncbi.nlm.nih.gov/pubmed/24608318http://doi.wiley.com/10.1111/mono.12083http://www.ipol.im/pub/art/2011/llmps-scb/?utm{_}source=doi.

Pawlowsky-Glahn V, Egozcue JJ, Tolosana-Delgado R (2015). *Modelling and Analysis of Compositional Data*. John Wiley & Sons, Ltd, Chichester, UK. ISBN 9781119003144. doi:10.1002/9781119003144. URL http://doi.wiley.com/10.1002/9781119003144.

Pohl C, van Genderen J (2016). *Remote Sensing Image Fusion*. CRC Press, Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742. ISBN 978-1-4987-3002-0. doi:10.1201/9781315370101. URL http://www.crcnetbase.com/doi/book/10.1201/9781315370101.

Poser BA, Norris DG (2009). "Investigating the Benefits of Multi-echo EPI for fMRI at 7T." *NeuroImage*, **45**(4), 1162–1172. ISSN 10538119. doi:10.1016/j.neuroimage.2009.01.007.

Smith AR (1978). "Color Gamut Transform Pairs." *ACM SIGGRAPH Computer Graphics*, **12**(3), 12–19. ISSN 00978930. doi:10.1145/965139.807361. URL http://portal.acm.org/citation.cfm?doid=965139.807361.

Smith S, Jenkinson M, Woolrich M (2004). "Advances in functional and structural MR image analysis and implementation as FSL." *Neuroimage*, **23**(November), S208–S219. ISSN 1053-8119. doi:10.1016/j.neuroimage.2004.07.051.

Smith SM (2002). "Fast robust automated brain extraction." *Human Brain Mapping*, **17**(3), 143–155. ISSN 10659471. doi:10.1002/hbm.10062. NIHMS150003.

Tsagris MT, Preston S, Wood ATA (2011). "A Data-based Power Transformation for Compositional Data." *arXiv preprint*, **1**, 1–9. 1106.1451, URL http://arxiv.org/abs/1106.1451.

Ugurbil K (2014). "Magnetic Resonance Imaging at Ultrahigh Fields." *IEEE Transactions on Biomedical Engineering*, **61**(5), 1364–1379. ISSN 0018-9294. doi:10.1109/TBME.2014.2313619. 3, URL http://ieeexplore.ieee.org/document/6778087/.

Uludağ K, Blinder P (2018). "Linking Brain Vascular Physiology to Hemodynamic Response in Ultra-high Field MRI." *NeuroImage*, **168**, 279–295. ISSN 1095-9572. doi:10.1016/j.neuroimage.2017.02.063. URL http://www.ncbi.nlm.nih.gov/pubmed/28254456.

Van der Walt S, Colbert SC, Varoquaux G (2011). "The NumPy array: A Structure for Efficient Numerical Computation." *Computing in Science & Engineering*, **13**(2), 22–30.

Van der Walt S, Schönberger JL, Nunez-Iglesias J, Boulogne F, Warner JD, Yager N, Gouillart E, Yu T, the scikit-image contributors (2014). "scikit-image: Image Processing in Python." *PeerJ*, **2**, e453. ISSN 2167-8359. doi:10.7717/peerj.453. URL http://dx.doi.org/10.7717/peerj.453.

Wang H, Shangguan L, Guan R, Billard L (2015). "Principal Component Analysis for Compositional Data Vectors." *Computational Statistics*, **30**(4), 1079–1096. ISSN 16139658. doi:10.1007/s00180-015-0570-1.

Affiliation:

Omer Faruk Gulban
Cognitive Neuroscience Department, Faculty of Psychology and Neuroscience
Maastricht University
6224 EV Maastricht, Netherlands
E-mail: faruk.gulban@maastrichtuniversity.nl
URL: https://orcid.org/0000-0001-7761-3727