Spectral Edge Image Fusion: Theory and Applications

David Connah[1], Mark Samuel Drew[2], and Graham David Finlayson[3]

[1] University of Bradford, UK
[2] Simon Fraser University, Vancouver, Canada
[3] University of East Anglia, Norwich, UK

Abstract. This paper describes a novel approach to the fusion of multidimensional images for colour displays. The goal of the method is to generate an output image whose gradient matches that of the input as closely as possible. It achieves this using a constrained contrast mapping paradigm in the gradient domain, where the structure tensor of a high-dimensional gradient representation is mapped *exactly* to that of a low-dimensional gradient field which is subsequently reintegrated to generate an output. Constraints on the output colours are provided by an initial RGB rendering to produce 'naturalistic' colours: we provide a theorem for projecting higher-D contrast onto the initial colour gradients such that they remain close to the original gradients whilst maintaining exact high-D contrast. The solution to this constrained optimisation is closed-form, allowing for a very simple and hence fast and efficient algorithm. Our approach is generic in that it can map any N-D image data to any M-D output, and can be used in a variety of applications using the same basic algorithm. In this paper we focus on the problem of mapping N-D inputs to 3-D colour outputs. We present results in three applications: hyperspectral remote sensing, fusion of colour and near-infrared images, and colour visualisation of MRI Diffusion-Tensor imaging.

Keywords: Image fusion, gradient-based, contrast, dimensional reduction, colour, colour display.

1 Introduction

As imaging technology has developed to solve a variety of problems, so the richness of imaging systems data has increased. Hyperspectral imaging systems used in remote sensing, for example, routinely capture > 200 channels of spectral data [5], while medical imaging systems capture multi-dimensional, and multimodal image sets [19]. Ultimately these images are often interpreted by human observers for analysis or diagnosis, and it is therefore crucial that dimensionality is reduced such that the image can be displayed on an output device such as a colour monitor. This process is termed *image fusion*.

Thus, in the image fusion problem, there can be 10, or 20, or hundreds of values per pixel, and we are interested in reducing the number to 1 for a representative greyscale output or 3 for colour visualization. The simplest way to visualise

D. Fleet et al. (Eds.): ECCV 2014, Part V, LNCS 8693, pp. 65–80, 2014.

the information is to simply average the values to produce a greyscale. This approach preserves basic scene structure but suffers from *metamerism*, where different multi-valued inputs are assigned the same output value.

Where the input values correspond to radiances at different wavelengths, a colour output can be generated by mapping the visible part of the spectrum to display RGB via projection onto a set of colour matching functions, which represent human sensitivity to wavelength [20]. At least such an approach produces a 'naturalistic' RGB image, where we define 'natural' as the colours that would be seen by a human observer, but it begs the question of how to take into account the influence of spectral values beyond the human visual system's sensitivity. One idea is to simply stretch the colour matching functions over the full wavelength range of the data [20]; in this case the displayed output produces a false-colour RGB visualisation of the entire spectral range. In general false-colour visualisations can be hard to interpret when object colours are very different from their natural appearance. Furthermore, these spectral projection methods do not say how to fuse non-spectral multi-valued data, e.g. multi-modal medical data.

In order to incorporate non-spectral data a more general approach is required. Generic dimensionality techniques such as PCA [37] or ISOMAP [8] can be applied to map multi-valued data to a 3-D space that is then interpreted as colour values. These approaches maximise the separation of colours in the output image, i.e. minimise the incidence of metamerism, but again produce false colourings. Also, while the incidence of metamerism may be minimised relative to some global objective function, there often aren't enough degrees of freedom to remove it completely.

To get closer to the preservation of all the multi-valued information in the output image, spatial information must be taken into account [30]. This can be done, for example, by transforming images into a multiscale representation, merging information at each spatial scale, and then inverting the multiscale transformation to produce an output image [4,24]. Practically, while this has the potential to preserve more information, artefacts such as haloing and ghost images are common. Also, the outputs are rendered in greyscale, which is a disadvantage. One way around this is to retain the RGB colour information whilst swapping in the new greyscale to take the place of the original luminance (i.e., intensity) information [35,31]. However, while such an approach does produce colour output in the fused image, the 3-D nature of colour is not fully harnessed.

An alternative approach to incorporating spatial information is to work in the gradient domain, where edge information is represented. Gradient domain processing has attracted significant interest due to the importance of edges in human perception [7], and has been applied in a range of fields such as HDR processing [12], image editing [28], and computational photography [3] among others. In the area of image fusion a key paper is the contrast preserving variational algorithm of Socolinsky and Wolff [33] who generate a greyscale image such that its gradient matches that of a multi-channel image as closely as

possible. This approach preserves key image information in the output, but still generates a greyscale output.

In this paper we present a gradient domain approach to image fusion that: generates colour outputs, incorporates constraints to allow a more 'natural' colour labelling, and can be applied to both spectral and non-spectral data. The approach is motivated by the work of Socolinsky and Wolff and the colorisation work of Drew and Finlayson [11], who use a gradient domain decomposition to apply the gradient from a greyscale image to a colour image, which they use to regulate the output of a colorization algorithm. The key contribution in the present paper is a theorem similarly yielding a gradient decomposition, but one which can be applied to the more general N-D to M-D mapping. This result allows us generalise Socolinsky and Wolff's work [33] to map N-D images to a colour, rather than just greyscale, output while also exactly matching contrast[1].

Our Spectral Edge (SpE) method is applicable to any domains where a) a transformation is required from an N-D space to an M-D space, b) the images in the individual channels are registered, and c) there is a a putative M-D image available with a viable colour scheme; this image may be captured by a separate device, or generated from the image data. The generality of the method makes it applicable to a wide range of problems, including: mapping multispectral / hyperspectral images to RGB; fusing RGB and NIR images; colour to greyscale; mapping 3D colour images to 2D to enhance images for colour-deficient observers; pan-sharpening; multi-exposure; dark flash; and visualisation of high-D medical image data such as MRI or time-activity curve data, to name a few. In this paper we report results for the applications of remote sensing, RGB / NIR fusion, and medical DTMRI data, with the output a colour image ($M = 3$) and $N > M$. Clearly for visualising medical data there is no concept of a 'natural' colour image; in these cases we can constrain the output colours using a putative false-colour labelling that is appropriate for the task.

The paper is organised as follows: in the next section we review related work in the application areas that we tackle in this paper; in §3 we describe the underlying mathematical formulation, and algorithmic details, of the method; in §4 we show the results of the method for three representative applications; and we conclude the paper in §5.

2 Related Work

The image fusion literature encompasses a wide range of applications and techniques. Different channels are typically treated as independent greyscale images and mapped to a single greyscale output, e.g. by averaging them. A popular framework is to decompose each channel into a multi-scale representation, fuse the images at each scale – e.g. by choosing the maximum wavelet coefficient over all images for that pixel / region – and inverting the decomposition step to recover a greyscale output. This approach has been followed using Laplacian

[1] U.S. patent granted March 2014 [6].

pyramids [4] and their variants [36], wavelets [24], complex wavelets [23], perceptual transforms using centre-surround filters [38], bilateral filtering [21], or multi-scale representations of the first fundamental form [32]. These methods are often complex and intensive to compute, as well as being prone to generating artefacts when conflicting information appears in different image channels, making them more suited to fusing pairs of images rather than multiple channels. Finally, the base layer of the pyramid, or wavelet decomposition, is often a low-pass average image, which can lead to poor colour separation at edges for low spatial scales.

Socolinsky and Wolff [33] cast image fusion as a variational problem, where the goal is to find a greyscale output with gradient information as similar as possible to the input image set. This approach solves the problem of greyscale separation at low spatial scales, but can also be prone to warping artefacts close to edges. These are exacerbated by the ambiguity of gradient ordering at each pixel [15]. Piella [29] uses a variational approach to generate an output that simultaneously preserves the underlying geometry of the multivalued image, similarly to Socolinsky and Wolff, and performs an edge enhancement to improve greyscale separation at object boundaries. The integration of gamut constraints means that potential for artefacts is greatly reduced using this method, but necessitates that the objective function is minimised using an iterative gradient descent scheme, which restricts the speed of the method. As with the wavelet-based approaches, the outputs are in greyscale only.

Several strategies exist for mapping high-dimensional images to RGB, rather than just greyscale. Jacobson *et al.* [20] investigate different fixed projections; these have an advantage over adaptive methods that colours remain fixed across different visualisations, but the disadvantage that they preserve less information. Adaptive approaches using standard decompositions such as PCA and ICA have also proved popular. Tyo *et al.* [37] use PCA to extract a 3-D subspace from the spectral data, and then rotate the basis of this space so that the final 3D co-ordinates form a plausible RGB image. While this approach is information preserving, the false coloured output can deviate from the 'natural' representation, and the global nature of the transform means that localised metamerism may still be common.

In particular applications greyscale fusion schemes can also be applied to generate colour outputs. Schaul *et al.* [31] employ fusion of near-infrared (NIR) and RGB images as part of a de-hazing scheme. They firstly decompose the RGB image into an opponent-based representation and then use an edge-aware multiscale representation to fuse the NIR and luminance channels into a single greyscale. This greyscale is then swapped into the original image as the luminance component. Our approach differs in that it maps the contrast of each of the R, G and B, channels as well as the NIR image, rather than just luminance and NIR. Fay *et al.* [13] use dual-band RGB / long-wave infrared (LWIR) to improve night-vision in low-light settings. This work, which results in fused colour imagery, is specifically focused on a low-light-sensitive visible-light CCD imager.

The approach we outline here is generic, in that it can be applied to a range of input and output dimensions. In this respect our work is closely related to that of Lau *et al.* [22] who proposed an optimisation based approach to colour mapping. They firstly cluster input colours into groups, and then maximise separation of those groups in a target colour space. They also include constraints on how far resulting colours can deviate from the colours in the target colour space such that the output remains 'naturalistic'. Although the goal and application of our technique is similar, our approach is markedly different in that we work on gradients, thus focusing the colour separation on spatial boundaries between objects or segments. The speed and low complexity of our method also makes it more suitable for visualising hyperspectral images.

In medical imaging, high-D information such as Time-Activity Curve multi-dimensional data is routinely reduced to RGB output, using various strategies such as false-colour renderings of the final sample or of the integral under the curve (cf. [19]). Our approach can be adapted to any application where a viable 3D colour output is available, whether one that is natural or colour obtained as a pseudocolour rendering. Our gradient domain approach focuses on separating colours at object boundaries, and can be used to improve colour visualisations derived from global mappings such as ICA or PCA.

3 Spectral Edge Image Fusion (SpE)

3.1 Definition of Gradient and Contrast

The goal of our method is to preserve the gradient of a high-dimensional image in a low-dimensional representation. The gradient of a multi-channel image $\boldsymbol{C}$ at a single pixel is given by the gradient matrix:

$$\boldsymbol{\nabla C} = \begin{bmatrix} C^1_{,x} & C^1_{,y} \\ \vdots & \vdots \\ C^N_{,x} & C^N_{,y} \end{bmatrix}, \tag{1}$$

where the function C^i is the ith channel of an N-channel image $\boldsymbol{C}$ and subscripts x and y denote derivatives in the x- and y-directions. The gradient matrix $\boldsymbol{\nabla C}$ contains the partial derivatives of $\boldsymbol{C}$ in the x and y directions; the gradient in direction $\boldsymbol{d} = [\cos\theta, \sin\theta]^T$ is $\boldsymbol{\nabla C}\,\boldsymbol{d}$. Assuming a Euclidean metric, the squared magnitude of the gradient in direction $\boldsymbol{d}$ is given by:

$$m^2 = \boldsymbol{d}^T(\boldsymbol{\nabla C})^T\boldsymbol{\nabla C}\,\boldsymbol{d}. \tag{2}$$

The 2×2 matrix $\boldsymbol{Z}_C = (\boldsymbol{\nabla C})^T\boldsymbol{\nabla C}$ is known in differential geometry as the First Fundamental Form, and was introduced to the image processing literature by Di Zenzo [10] as the structure tensor.

The structure tensor representation is powerful because it encodes magnitude information for the N-dimensional matrix in 2 dimensions: given $\boldsymbol{Z}_C$ we can compute the gradient magnitude in any direction $\boldsymbol{d}$.

A fundamental idea behind our method, therefore, is: *in order for a low-dimensional image (low-D) to have an identical contrast to a high-dimensional image (high-D), the structure tensor for both must be identical.*

3.2 Exact Contrast Mapping

In Socolinsky and Wolff [33], the authors have a similar goal in mapping high-D contrast, defined by the the structure tensor, to a scalar image, approximately. In the first stage of their algorithm they define a scalar gradient field ∇I by multiplying the first eigenvector of $\boldsymbol{Z}_C$ by the first eigenvalue of $\boldsymbol{Z}_C$; the resulting gradient field has the closest possible possible structure tensor – $\boldsymbol{Z}_I$ – that a scalar field can have to $\boldsymbol{Z}_C$ in the least squares sense.

In the novel approach presented here, instead of creating a scalar gradient-field we create M gradient fields, where M is the number of channels in our output image; we refer to this set of gradient fields as an M-D gradient-field. By doing this we can now generate an $M-$D gradient field whose structure tensor matches the original structure tensor $\boldsymbol{Z}_C$ *exactly*.

In order to ensure that the output is coloured naturally, we suppose that we have access to a putative low-D version $\widetilde{\boldsymbol{R}}$ of the high-D image data which has naturalistic colours: this image may either be captured by a specific device (e.g. an RGB camera), or generated from the high-D using some algorithm (e.g. a true colour rendering of remote sensing data). We then use the contrast information from the high-D image, and the colour information from the putative low-D image, to generate a new low-D gradient field, which we finally reintegrate to generate a colour output. This idea motivates the following theorem:

Spectral Edge (SpE) Projection Theorem:
Given a multidimensional image $\boldsymbol{C}$ and a putative RGB "guiding" image $\widetilde{\boldsymbol{R}}$, we can generate a new RGB gradient matrix $\boldsymbol{\nabla R}$ that is as close as possible to the gradient of the RGB image, and whose contrast matches that of $\boldsymbol{C}$ exactly.

For the most common application of our method, we start with an N-D higher-dimensional input image $\boldsymbol{H}$, with the *goal* of generating a 3-band colour image $\boldsymbol{R} = (R, G, B)$. We denote the desired colour gradient at each pixel by $\boldsymbol{\nabla R}$, which is a 3×2 gradient matrix:

$$\nabla \boldsymbol{R} \;=\; \begin{pmatrix} R_{,x} & R_{,y} \\ G_{,x} & G_{,y} \\ B_{,x} & B_{,y} \end{pmatrix} \tag{3}$$

This is the output of our algorithm, which is subsequently to be reintegrated.

We also have a putative RGB colour image, generated by some initial algorithm or captured by a colour camera, which we denote $\widetilde{\boldsymbol{R}}$. We denote the gradient matrices of these images as $\boldsymbol{\nabla H}$, $\boldsymbol{\widetilde{\nabla R}}$, and $\boldsymbol{\nabla R}$ for respectively the high-D image, putative RGB image, and output RGB image. We notate $\boldsymbol{\widetilde{\nabla R}}$ carefully since it is in fact the putative colour gradient we wish to alter to create an output RGB gradient.

For our 3 gradient fields, the Di Zenzo matrices are defined as:

$$\boldsymbol{Z}_H = (\boldsymbol{\nabla H})^T(\boldsymbol{\nabla H}),\ \boldsymbol{Z}_R = (\boldsymbol{\nabla R})^T(\boldsymbol{\nabla R}),\ \widetilde{\boldsymbol{Z}}_R = (\widetilde{\boldsymbol{\nabla R}})^T(\widetilde{\boldsymbol{\nabla R}}) \quad (4)$$

Now we aim to satisfy two conditions: **(1)** For a generated $\boldsymbol{\nabla R}$, i.e. the result of the theorem, we wish $\boldsymbol{Z}_R$ to equal $\boldsymbol{Z}_H$, the structure tensor for the higher-D image, so that contrast is mapped exactly from high-D to low-D; and **(2)** the output gradient $\boldsymbol{\nabla R}$ should approximate as closely as possible the putative gradient $\widetilde{\boldsymbol{\nabla R}}$, so that no large colour shifts are obtained. That is, we desire an altered colour gradient $\boldsymbol{\nabla R} \simeq \widetilde{\boldsymbol{\nabla R}}$, subject to **(1)** and **(2)**.

A solution obeying **(1)** can be found easily if we keep only within the span of colour gradient $\widetilde{\boldsymbol{\nabla R}}$, and seek a 2×2 linear matrix transform $\boldsymbol{A}$ such that

$$\boldsymbol{\nabla R} = \widetilde{\boldsymbol{\nabla R}}\ \boldsymbol{A} \quad (5)$$

so that the colour gradient will not differ greatly from the approximation. In that case the desired relation between Di Zenzo matrices is as follows:

$$\begin{aligned} & \boldsymbol{Z}_R \equiv \boldsymbol{Z}_H \\ \Rightarrow\ & \boldsymbol{Z}_R = \boldsymbol{\nabla R}^T\boldsymbol{\nabla R} = \boldsymbol{A}^T\widetilde{\boldsymbol{\nabla R}}^T\widetilde{\boldsymbol{\nabla R}}\ \boldsymbol{A} \equiv \boldsymbol{Z}_H \\ \Rightarrow\ & \boldsymbol{A}^T\widetilde{\boldsymbol{Z}}_R\boldsymbol{A} \equiv \boldsymbol{Z}_H \end{aligned} \quad (6)$$

Given this relation, we satisfy **(1)** above provided matrix $\boldsymbol{A}$ is any solution of (6). For example, one solution is given by:

$$\boldsymbol{A} = \left(\sqrt{\widetilde{\boldsymbol{Z}}_R}\right)^+ \sqrt{\boldsymbol{Z}_H} \quad (7)$$

where the matrix square root is the unique symmetric root [18] of the real positive semi-definite symmetric matrices $\widetilde{\boldsymbol{Z}}_R$ and $\boldsymbol{Z}_H$, and $^+$ indicates the Moore-Penrose pseudoinverse (even though $\sqrt{\widetilde{\boldsymbol{Z}}_R}$ is square, nonetheless we guard against instability by using the pseudoinverse rather than the inverse).

To show that $\boldsymbol{A}$ is indeed a valid solution we can see that:

$$\boldsymbol{A}^T\widetilde{\boldsymbol{Z}}_R\boldsymbol{A} = (\sqrt{\boldsymbol{Z}_H}\sqrt{\widetilde{\boldsymbol{Z}}_R}^+)\widetilde{\boldsymbol{Z}}_R(\sqrt{\widetilde{\boldsymbol{Z}}_R}^+\sqrt{\boldsymbol{Z}_H}) = \boldsymbol{Z}_H \quad (8)$$

since $\sqrt{\boldsymbol{Z}_H}$ and $\sqrt{\widetilde{\boldsymbol{Z}}_R}^+$ are symmetric.

The *complete* set of solutions solving (6) then consists of all matrices $\boldsymbol{A}$ that are any 2×2 orthogonal transform $\boldsymbol{O}$ away from (7):

$$\boldsymbol{A} = \left(\sqrt{\widetilde{\boldsymbol{Z}}_R}\right)^+ \boldsymbol{O}\ \sqrt{\boldsymbol{Z}_H},\quad \boldsymbol{O}^T\boldsymbol{O} = \boldsymbol{I}_2 \quad (9)$$

since any such solution satisfies (6):

$$\boldsymbol{A}^T\widetilde{\boldsymbol{Z}}_R\boldsymbol{A} = (\sqrt{\boldsymbol{Z}_H}\boldsymbol{O}^T\sqrt{\widetilde{\boldsymbol{Z}}_R}^+)\widetilde{\boldsymbol{Z}}_R(\sqrt{\widetilde{\boldsymbol{Z}}_R}^+\boldsymbol{O}\sqrt{\boldsymbol{Z}_H}) = \boldsymbol{Z}_H \quad (10)$$

To produce realistic colours we also wish to fulfil constraint **(2)** , that the adjusted gradient $\boldsymbol{\nabla R}$ approximates as closely as possible the putative colour gradient $\widetilde{\boldsymbol{\nabla R}}$. From (5), this implies a constraint on rotation $\boldsymbol{O}$ as follows:

$$\begin{aligned} & \boldsymbol{\nabla R} \simeq \widetilde{\boldsymbol{\nabla R}} \\ \Rightarrow\ & \widetilde{\boldsymbol{\nabla R}}\ \boldsymbol{A} \simeq \widetilde{\boldsymbol{\nabla R}} \\ \Rightarrow\ & \boldsymbol{A} \simeq \boldsymbol{I}_2 \\ \Rightarrow\ & \sqrt{\widetilde{\boldsymbol{Z}}_R}^{+}\ \boldsymbol{O}\ \sqrt{\boldsymbol{Z}_H} \simeq \boldsymbol{I}_2 \\ \Rightarrow\ & \boldsymbol{O}\ \sqrt{\boldsymbol{Z}_H} \simeq \sqrt{\widetilde{\boldsymbol{Z}}_R} \end{aligned} \tag{11}$$

with $\boldsymbol{I}_2$ the 2×2 identity matrix. The last line of (11) says that $\boldsymbol{O}$ should be chosen to rotate $\sqrt{\boldsymbol{Z}_H}$ such that it is as close as possible to $\sqrt{\widetilde{\boldsymbol{Z}}_R}$. This problem is known as the Orthogonal Procrustes Problem [18]; the solution in the least-squares sense is to firstly use a singular value decomposition to express the product of square roots of $\widetilde{\boldsymbol{Z}}_R$ and $\boldsymbol{Z}_H$:

$$\sqrt{\widetilde{\boldsymbol{Z}}_R}\ \left(\sqrt{\boldsymbol{Z}_H}\right)^T = \boldsymbol{D}\ \boldsymbol{\Gamma}\ \boldsymbol{E}^T \tag{12}$$

with $\boldsymbol{\Gamma}$ diagonal (the transpose on the second term above is actually unnecessary since $\sqrt{\widetilde{\boldsymbol{Z}}_H}$ is symmetric but we include it to agree with the formulation in [18]). Then the solution $\boldsymbol{O}$ that minimises the last line of (11) in terms of Least Squares is given by:

$$\boldsymbol{O} = \boldsymbol{D}\ \boldsymbol{E}^T \tag{13}$$

We can now obtain $\boldsymbol{A}$ by substituting this solution for $\boldsymbol{O}$ into equation (9), and then directly derive a modified colour gradient $\boldsymbol{\nabla R}$ using (5). ■

Importantly, we note that in this theorem $\boldsymbol{Z}_H$ is not in fact restricted to being derived from a higher-dimensional image — it can be any Di Zenzo matrix from an image of any dimension, e.g. that for a greyscale Near-Infra-Red (NIR) image, or alternatively that for a 4-D image generated by appending the NIR image to RGB. Similarly $\boldsymbol{R}$ could refer to an output of any dimension, provided a putative gradient $\widetilde{\boldsymbol{\nabla R}}$ can be specified.

In summary, starting from a lower-D image containing a naturalistic rendering of the scene $\widetilde{\boldsymbol{R}}$, at *each pixel* we find a transform $\boldsymbol{A}$ of the $M \times 2$ gradient matrix of the lower-D image such that (i) the altered gradient has an identical *contrast* as that for the higher-D image – i.e. we transfer the higher-D contrast to the lower-D image; and (ii) the altered lower-D gradient $\boldsymbol{\nabla R}$ *remains in the span* of the unaltered gradient, at each pixel; i.e. the new $M \times 2$ gradient is a 2×2 linear transform away from the putative gradient.

3.3 Reintegration

The contrast mapping process results in an M-D gradient matrix $\boldsymbol{\nabla R}$ at each pixel location. We would like to treat $\boldsymbol{\nabla R}$ as a set of M gradient fields, one

for each output channel, defined by the rows of $\boldsymbol{\nabla R}$. The final phase of the algorithm is to reintegrate each gradient field in turn to generate M new output channels. However, in general each of the approximate gradient fields will be *non-integrable*, i.e. will not in fact be the gradient for a scalar image. An output image must therefore be reconstructed by computing an image whose gradient matches that of the target field as closely as possible, by minimising some error function. Interestingly, however, we have more information available here than in the traditional reintegration problem of forming a greyscale image I from a gradient-approximation – we have the actual, N-D image dataset itself.

If we denote the approximate gradient field from the i-th channel of $\boldsymbol{\nabla R}$ as $P^i = \left(R^i_{,x} R^i_{,y}\right)$, then we seek a scalar image I such that:

$$R^i = \arg\min_I \|P^i - \nabla I\|_n \tag{14}$$

where n defines the norm used in the error function. For $n = 2$ the solution could be given by the solution to Poisson's equation, and a number of approaches have been applied to do this, e.g. [16,1]. However since here we also have the N-D data $\boldsymbol{H}$, we can use the look-up-table approach of Finlayson *et al.* in [15,14], which minimises the error function in (14) for $n = 2$ using a LUT mapping from the high-D image $\boldsymbol{H}$ to each R^i. This constraint means that the final image is guaranteed to be free of artefacts, and facilitates the operation of the algorithm in real time. Importantly, in [15] it was shown that if a multi scale gradient is approximately integrable across multiple scales then a LUT mapping is the correct reintegrating function.

3.4 Implementation Details

To compute the gradient matrices $\widetilde{\boldsymbol{\nabla R}}$ and $\boldsymbol{\nabla H}$ we use local finite differencing, i.e. for an image $\boldsymbol{C}$ at pixel (x, y) and channel i, $C^i_{,x}(x, y) = C^p(x-1, y) - C^i(x, y)$ and $C^p_{,y}(x, y) = C^p(x, y-1) - C^p(x, y)$, although other gradient operators, e.g. Sobel operators, would serve the purpose just as well. Furthermore, given the global nature of the reintegration approach in [15], the gradient operator could also be applied at different spatial scales, and reintegrated simultaneously. For other reintegration techniques the finest spatial scale is advised to reduce blurring in the output. There is a potentially large discrepancy in image dimensionalities between input and output, i.e. $N \gg M$ for input dimensionality N and output M, and as a result the total high-D contrast may not be displayable within the low-D gamut. Here, we mitigate this with a simple contrast scaling approach whereby 99% of pixel values are mapped within the image gamut, although more complex gamut mapping strategies could also be employed [25] as post-processing after applying the algorithm.

The complexity of the contrast projection algorithm is $O(P)$, where P is the number of pixels. The complexity of the reintegration is also $O(P)$ [15], although using other approaches, such as iterative Poisson solvers, can increase the complexity. Memory requirements are low, since most of the calculations are performed on 2×2 structure tensor matrices. In our case the chosen reintegration

[15] increases memory requirements since the high-dimensional image needs to be stored and used in the reintegration. But the chief advantage of this method is its ability to remove artefacts.

The method is general in the choice of output colour space. We represent images in sRGB for the applications here, but the putative low-D image could be represented in a different space, e.g. a perceptually uniform space such as CIELAB, and then mapped to sRGB for display. This would be a good approach in applications where Euclidean distances in sensor-space should correlate with the magnitude of perceived differences in the output.

4 Experiments

4.1 Experiment Paradigms

In this paper we show results of our method in three application areas: i) hyperspectral / multispectral remote sensing, ii) fusion of NIR / LWIR (thermal imaging) with RGB; and iii) medical MRI diffusion-tensor imaging. Each of the applications falls naturally within the same computational framework; we explain below how to adapt this framework for each application.

Remote Sensing Applications. Images captured for remote sensing applications, e.g. from satellite or airborne imaging systems, typically span the visible, near infra-red and far-infra red wavelength spectrum. Here we use data from two publicly available datasets: a) Landsat 7 [27], and b) AVIRIS [26]. The Landsat 7 satellite captures 8 separate images; 3 in the visible range, 4 IR images (including one thermal image) and a panchromatic detail image; these images are captured using a scanning radiometer. The three visible images are captured from 450-515nm (blue), 525-605 (green), and 630-690 nm (red), and we use these as the B, G and R channels respectively of $\widetilde{\boldsymbol{R}}$; $\boldsymbol{H}$ then consists of the three RGB channels, and three IR images captured at: 750-900nm (NIR); 1550-1750nm (SWIR); and 2090-2350nm (SWIR). We omit the thermal and panchromatic channels as they have different spatial resolutions than the other images.

The AVIRIS data is captured from an airborne imaging system, and uses a "sweep-broom" hyperspectral camera with 224 adjacent spectral channels, which span a spectral range 380-2500 nm and are sampled at approximately 10nm intervals. To generate $\widetilde{\boldsymbol{R}}$ in this case we project the visible wavelengths, 380-730nm, onto the sRGB colour matching functions [34], to generate a true-colour sRGB rendering; $\boldsymbol{H}$ is composed of all 224 channels.

Visualising NIR / LWIR Images. Pairs of RGB and NIR (or thermal) images can be captured using different methods, e.g. using a beamsplitter and two CCD arrays to capture registered NIR and RGB, or taking successive photographs with an IR filter ("hot mirror") present and absent.

To apply our technique to this problem we construct a 4-D image $\boldsymbol{H}$ by appending the NIR channel to the colour image. This 4D image is used to calculate

the high-D gradient $\boldsymbol{\nabla H}$ while the original RGB image is used to calculate the putative gradient $\widetilde{\boldsymbol{\nabla R}}$.

We compare our technique with: (a) "alpha blending", where the RGB outputs, R_{out} G_{out}, and B_{out} are constructed as convex combinations of the RGB and NIR input images, e.g. $R_{out} = \alpha R + (1 - \alpha)NIR$ for $0 \geq \alpha \leq 1$; (b) "luminance replacement", where the RGB image is firstly mapped to YIQ space, and the luminance component, Y, is then replaced by the NIR image; (c) the colour-cluster optimisation method of Lau *et al.* [22].

Medical Applications. In some fusion applications there is no "true-colour" rendering of the input image available, but labelling the input data using colour still has value for interpreting the data. In medical imaging, for example, multimodal and multidimensional imaging devices such as PET, MRI and diffusion tensor imaging (DTI) systems are used to gather physiological data that is displayed as an image, and used by clinicians to aid diagnosis.

Here we apply our algorithm the problem of visualising MRI Diffusion-Tensor data. In this application the data consists of 3×3 symmetric positive semi-definite matrices at each spatial location, and is hence 6-D. To preserve its character, 6-D vectors are formed respecting a Log-Euclidean metric [2]. The most common method for visualising such data is to display loadings on the first 3 principal component vectors [37]. A more perceptually meaningful approach than PCA is to carry out multi-dimensional scaling (MDS) on the 6-vectors, descending to 3-D [19]; then the result is conceived as approximately perceptually uniform CIELAB colour and then mapped to standard gamma-corrected sRGB display space. To apply our algorithm we use both PCA and MDS approaches to generate different putative RGB outputs $\widetilde{\boldsymbol{\nabla R}}$, and the 6-D tensor output to calculate $\boldsymbol{\nabla H}$.

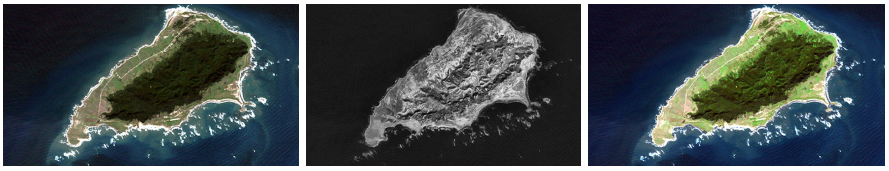
4.2 Results

Results from the multispectral Landsat data are shown in Figs. 1 and 2, and a result from the hyperspectral AVIRIS data is shown in Fig. 3. Each example includes the RGB rendering, an example IR image, and the output of our method. For Fig. 3 we also show the result of using a stretched colour-matching function approach (cf. [20]). In each case the content of the output SpE image shares the same colour scheme as the putative true-colour RGB output, and as well integrates the information from the additional channels. In particular, because of the inclusion of IR data, the presence of bodies of water becomes more pronounced than in the original.

Figures 4 and 5 show results for the problem of merging RGB and NIR images. In Figs. 4(c,d) alpha-blending and luminance replacement outputs significantly alter the natural colouring. The method of Lau *et al.* Fig. 4(e) attempts to incorporate detail from the NIR image and does keep natural colours. Our approach

(a) RGB (Putative) (b) SWIR (c) SpE output

Fig. 1. A remote sensing example using images from Landsat [27] (see text for details)



(a) RGB (putative) (b) MWIR (c) SpE output

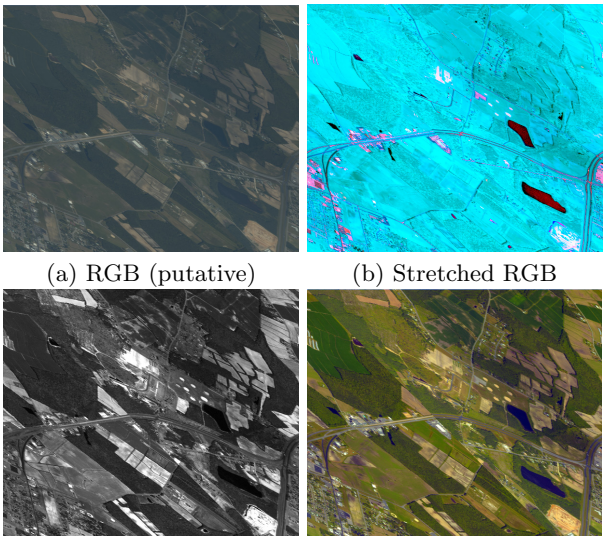
Fig. 2. A second image set taken from the Landsat database

focuses on preserving the information content in all four input image planes; as a result the presence of the NIR image is much more noticeable in regions of low contrast in the original RGB, e.g. around the trees. In Fig. 5 we succeed in keeping colour information intact while displaying NIR information more visibly. As in the non-gradient approach [17], age-spots are removed, along with freckles; but as well, more of the NIR content is displayed using the SpE method.

We also demonstrate that our method can be used to fuse RGB with longer-wave, even thermal, IR (wavelengths $> 10\mu$m). Figure 6 shows fusion results for an image from the OTCBVS dataset [9], which contains registered RGB and thermal images. The fusion is successful, with hidden structures made visible.

In Fig. 7 we show results for the medical, DTI, application. This data consists of 55 axial images of brain slices, each representing a different depth plane. In Fig. 7(a), we use PCA weightings to generate a putative RGB for a single axial slice, mapped to RGB and with each colour channel mapped to [0..1]. In Fig. 7(b) we show results of the SpE method for the same slice; the image clearly better incorporates full 6-D contrast information.

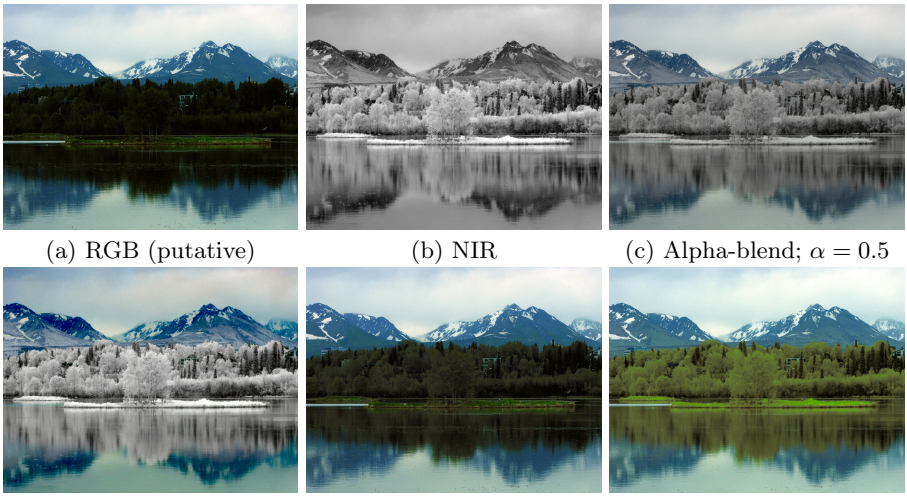
In Fig. 7(c) we show the same slice, where the putative RGB is generated from an MDS scaling; this image is taken from [19]. In Fig. 7(c) the output is already optimising information content as global data, but our SpE projection in Fig. 7(d) shows the substantive effect of SpE in including more of the higher-D information, and focusing colour separation on boundaries between regions.



(a) RGB (putative) (b) Stretched RGB

(c) SWIR (d) SpE output

Fig. 3. Example of hyperspectral image fusion; images taken from AVIRIS dataset [26]. In (b), the largely blue output is due to most of the energy measured in each pixel spectrum residing in the visible band, which is on the small-wavelength end in the full measured spectrum extending from 370.5nm to 2507.6nm.



(a) RGB (putative) (b) NIR (c) Alpha-blend; $\alpha = 0.5$

(d) Luminance replacement (e) Lau *et al* (f) SpE approach

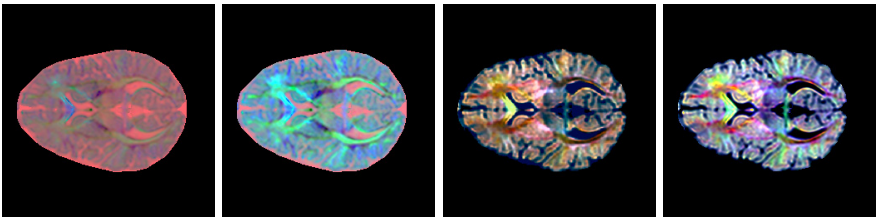
Fig. 4. Comparison of SpE with other methods for an RGB + NIR fusion application

(a) RGB (putative) (b) Result from [17] (c) SpE approach

Fig. 5. RGB + NIR fusion application

(a) Thermal (7-14μm) (b) RGB (putative) (c) SpE output

Fig. 6. Example of thermal + RGB fusion; images taken from OTCBVS dataset [9]



(a) PCA (putative) (b) PCA+ SpE (c) MDS (putative) (d) MDS+SpE

Fig. 7. Visualization of 6-D DTMRI data: (a,b) PCA approach, (c,d) MDS method

5 Conclusion

In this work we have presented a novel, gradient-domain, approach for mapping images of any dimension to images of any other dimension. The method is based on mapping contrast, defined by the structure tensor matrix, exactly onto a low-dimensional gradient field, and incorporates constraints on the naturalness of output colours borrowed from a putative RGB rendering. The approach is formulated as a constrained optimisation with a closed-form solution, making the method both fast and efficient. We have demonstrated applications in mapping high-dimensional images to RGB outputs for display, and will expand the applicability to new areas in future work.

References

1. Agrawal, A., Chellappa, R., Raskar, R.: An algebraic approach to surface reconstruction from gradient fields. In: Int. Conf. on Comp. Vision, pp. 174–181 (2005)
2. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Log-euclidean metrics for fast and simple calculus. Mag. Res. in Medicine 56, 411–421 (2006)
3. Bhat, P., Zitnick, C.L., Cohen, M., Curless, B.: Gradientshop: A gradient-domain optimization framework for image and video filtering. ACM Trans. Graph. 10, 10:1–10:14 (2010)
4. Burt, P.J., Adelson, E.H.: Merging images through pattern decomposition. In: Proc. SPIE 0575, Applications of Digital Image Processing VIII. pp. 173–181 (1985)
5. Campbell, J.B., Wynne, H.: Introduction to Remote Sensing, 5th edn. Guilford Press (2011)
6. Connah, D., Drew, M., Finlayson, G.: Method and system for generating accented image data. In: U.S. patent No. 8682093 and UK patent GB0914982.4 (March 25, 2014)
7. Cornsweet, T.: Visual Perception. Academic Press, New York (1970)
8. Cui, M., Hu, J., Razdan, A., Wonka, P.: Color to gray conversion using ISOMAP. Vis. Comput. 26, 1349–1360 (2010)
9. Davis, J., Sharma, V.: Background-subtraction using contour-based fusion of thermal and visible imagery. Comp. Vis. and Im. Und.; IEEE OTCBVS WS Series Bench 106 (2-3), 162–182 (2007)
10. Di Zenzo, S.: A note on the gradient of a multi-image. Comp. Vision, Graphics, and Image Proc. 33, 116–125 (1986)
11. Drew, M., Finlayson, G.: Improvement of colorization realism via the structure tensor. Int. J. Image and Graphics 11(4), 589–609 (2011)
12. Fattal, R., Lischinski, D., Werman, M.: Gradient domain high dynamic range compression. ACM Trans. on Graphics 21, 249–256 (2002)
13. Fay, D., Waxman, A.M., Aguilar, M., Ireland, D., Racamato, J., Ross, W., Streilein, W.W., Braun, M.I.: Fusion of multi-sensor imagery for night vision: Color visualization, target learning and search. In: 3rd Int. Conf. Information Fusion, pp. 215–219 (2000)
14. Finlayson, G., Connah, D., Drew, M.: Image reconstruction method and system, U.S. and U.K. filing, British Patent Office Application Number GB0914603.6 (August 20, 2009)
15. Finlayson, G.D., Connah, D., Drew, M.S.: Lookup-table-based gradient field reconstruction. IEEE Trans. Im. Proc. 20(10), 2827–2836 (2011)
16. Frankot, R.T., Chellappa, R.: A method for enforcing integrability in shape from shading algorithms. IEEE Trans. on Patt. Anal. and Mach. Intell. 10, 439–451 (1988)
17. Fredembach, C., Barbuscia, N., Süsstrunk, S.: Combining visible and near-infrared images for realistic skin smoothing. In: Color Imaging Conf. (2009)
18. Golub, G., van Loan, C.: Matrix Computations. John Hopkins U. Press (1983)
19. Hamarneh, G., McIntosh, C., Drew, M.S.: Perception-based visualization of manifold-valued medical images using distance-preserving dimensionality reduction. IEEE Trans. on Medical Imaging 30(7), 1314–1327 (2011)
20. Jacobson, N., Gupta, M., Cole, J.: Linear fusion of image sets for display. IEEE Trans. on Geosciences and Remote Sensing 45, 3277–3288 (2007)
21. Kotwal, K., Chaudhuri, S.: Visualization of hyperspectral images using bilateral filtering. IEEE Trans. Geosci. and Remote Sen. 48(5), 2308–2316 (2010)

22. Lau, C., Heidrich, W., Mantiuk, R.: Cluster-based color space optimizations. In: Int. Conf. on Comp. Vision, pp. 1172–1179 (2011)
23. Lewis, J., O'Callaghan, R., Nikolov, S., Bull, D., Canagarajah, C.: Region-based image fusion using complex wavelets. In: 7th Int. Conf. on Information Fusion, vol. 1, pp. 555–562 (2004)
24. Li, H., Manjunath, B., Mitra, S.: Multisensor image fusion using the wavelet transform. Graphical Models and Im. Proc. 57(3), 235–245 (1995)
25. Morovic, J.: Color Gamut Mapping. John Wiley & Sons (2008)
26. NASA: Aviris: Airborne visible / infrared imaging spectrphotometer (2013), http://aviris.jpl.nasa.gov/
27. NASA: Landsat imagery (2013), http://glcf.umd.edu/data/gls/
28. Pérez, P., Gangnet, M., Blake, A.: Poisson image editing. ACM Trans. Graph. 22(3), 313–318 (2005)
29. Piella, G.: Image fusion for enhanced visualization: a variational approach. Int. J. Comput. Vision 83(1), 1–11 (2009)
30. Pohl, C., Genderen, J.L.V.: Multisensor image fusion in remote sensing: concepts, methods and applications. Int. J. of Remote Sensing 19(5), 823–854 (1998)
31. Schaul, L., Fredembach, C., Süsstrunk, S.: Color image dehazing using the near-infrared. In: Int. Conf. on Im. Proc. (2009)
32. Scheunders, P.: A multivalued image wavelet representation based on multiscale fundamental forms. IEEE Trans. Im. Proc. 11(5), 568–575 (2002)
33. Socolinsky, D., Wolff, L.: Multispectral image visualization through first-order fusion. IEEE Trans. Im. Proc. 11, 923–931 (2002)
34. Stokes, M., Anderson, M., Chandrasekar, S., Motta, R.: A standard default color space for the internet – sRGB (1996), http://www.w3.org/Graphics/Color/sRGB
35. Toet, A.: Natural colour mapping for multiband nightvision imagery. Infor. Fusion 4, 155–166 (2003)
36. Toet, A., Ruyven, J.J.V., Valeton, J.M.: Merging thermal and visual images by a contrast pyramid. Optical Eng. 28(7), 789–792 (1989)
37. Tyo, J., Konsolakis, A., Diersen, D., Olsen, R.: Principal-components-based display strategy for spectral imagery. IEEE Trans. on Geosciences and Remote Sensing 41, 708–718 (2003)
38. Waxman, A., Gove, A., Fay, D., Racamoto, J., Carrick, J., Seibert, M., Savoye, E.: Color night vision: Opponent processing in the fusion of visible and ir imagery. Neural Networks 10, 1–6 (1997)