Iterative Spectral Edge Image Fusion

Graham D. Finlayson, Alex E. Hayes

Abstract

The Spectral Edge method of image fusion fuses input image details, while maintaining natural color. It is a derivative-based technique, based on the structure tensor, and lookup-table-based gradient field reconstruction. It has many applications, including RGB-NIR image fusion and remote sensing.

In this paper, we propose adding an iterative step to the method. We use the output Spectral Edge image as the putative color image for another fusion step, and repeat this for several iterations. We show that this creates an output image with a structure tensor field closer to that of the high-dimensional input than the output of the original method. We perform a psychophysical experiment using the iterative Spectral Edge method for RGB-NIR image fusion, which shows that the result of multiple iterations is preferred.

Introduction

Image gradients are a natural and intuitive way of representing image information. An important set of image fusion methods are based on the color structure tensor, a way of representing gradient information across multiple image channels.

In these methods, a set of output gradients is calculated which represents the maximum variation within the input images. These gradients are then reintegrated to produce an output image. Unfortunately, implemented naively (solving a Poisson equation [9]), the reintegration process produces artifacts.

To remedy this problem, Finlayson *et al.* introduced lookup-table-based gradient field reconstruction, which avoids artifacts by posing gradient field reintegration as a lookup-table (LUT) problem. Specifically, that the output image - and the reintegrated gradient field, should be a look-up table mapping from the input image - This LUT constraint prevents new structure (not present in any of the image bands being fused) from appearing in the final image. An important detail of the LUT approach is that the form of the mapping is a second order polynomial function. The restricted form of the underlying function is chosen both to ensure the determination of model parameters to be robust (e.g. to image noise) but also that the function should be smooth[5].

In this paper we extend the “Spectral Edge” Image fusion method developed by Connah *et al.*. The input RGB image gradients are transformed in a way which captures the high-dimensional image gradient information, while remaining faithful to the color of a putative RGB image (often the input RGB image).

While Spectral Edge image fusion produces impressive and preferred [7] results, the penalty of using the LUT reintegration method is that the recovered image only approximates the desired gradient. The hypothesis we examine in this paper is that by running Spectral Edge iteratively we will be able to further improve the image fusion result. Pairwise preference experiments validate our technical work. We find that the results produced after 2 or 3 iterations are preferred over the single application of Spectral Edge (which previous work demonstrated was already preferred over several competing contemporary algorithms).

The background to our work is presented in the second section of this paper. Then, the proposed iterative method is presented. Experiments demonstrate 2 or 3 iterations provide a positive increment on the images produced (they are preferred). The paper finishes with a short conclusion.

Background

The color structure tensor

Di Zenzo introduced the color structure tensor[3]. The motivation behind this tensor was to find a way to represent the relation between the gradients in the x and y directions across multiple image channels, and find a direction of maximal contrast at a particular image pixel. This is accomplished by representing the gradient in each channel as a vector of length 2. These vectors are stacked to form the image Jacobian:

$$J = \begin{bmatrix} \frac{\partial I_1}{\partial x} & \frac{\partial I_1}{\partial y} \\ \frac{\partial I_2}{\partial x} & \frac{\partial I_2}{\partial y} \\ .. & .. \\ \frac{\partial I_N}{\partial x} & \frac{\partial I_N}{\partial y} \end{bmatrix} \tag{1}$$

Where I_n is the nth input channel, out of a total of N input channels.

The next step is to use this Jacobian matrix to create what Di Zenzo defined as the structure tensor[3], which is a symmetric tensor field of rank 2, and in differential geometry is known as the First Fundamental Form:

$$Z = J^T J \tag{2}$$

If $\vec{V}$ is a 2 by 1 unit length vector specifying a direction in the image plane, we can calculate the squared magnitude of the change across all input channels in this direction using the structure tensor Z.

Socolinsky and Wolff (SW) made the important observation that using the structure tensor alone we can solve for the $\vec{V}$ where there is a maximum change in the underlying image and the magnitude of this change. Their idea then was to compute the direction and magnitude of maximum change at each point of the image, in effect calculating a single gradient field (from the N gradient fields in (1)). And, then for image fusion this ‘equivalent’ gradient field is reintegrated. In fact we can’t quite do this directly as the calculated $\vec{V}$ is only unique up to a sign. SW fixed this sign to be the same as the gradient in this direction of the mean image (over the N input channels).

Gradient field reintegration

Let us denote the gradient field derived via Socolinsky and Wolff[9] as ∇J. It could be that there is no image that has derivatives equal to the ones we seek. After all, for every pixel we have an x and y derivative yet the reintegrated image has a single pixel value. Thus, the typical away to solve this reintegration problem is to solve the Poisson equation

$$\arg\min_{I} ||\nabla I - \nabla J|| \tag{3}$$

In finding the image I it is often the case that the reintegrated image has details not in any of the original N-image planes. Indeed I will typically have haloes and/or bending artifacts. The gradient field is not integrable and in solving for I (in a least-squares sense) the error manifests itself in these visible artifacts.

One way to remove artifacts from the reintegrated image is to place a constraint on I. Let us denote all images that are a a linear combination of the N-image planes as

$$I \in P_1(I) \tag{4}$$

Or, if we also allow second order polynomial terms (for an RGB image this would be R^2, G^2, B^2 and RG, GB and BG) we write

$$I \in P_2(i) \tag{5}$$

where P_n denotes the order of the polynomial expansion. Finlayson *et al.* proposed solving for I as

$$\arg\min_{I \in P_2(I)} ||\nabla I - \nabla J|| \tag{6}$$

The advantage of ensuring that the output image is a function of the input is bending and halo artifacts cannot occur (a unique N-pixel in the input maps to a unique greyscale). Further, adopting a low order polynomial ensures the function is smooth and that the computational process is rapid.

Spectral Edge image fusion

The Spectral Edge (SE) method is a derivative domain image fusion technique. It is based on the color structure tensor, but instead of solving for a greyscale output image with gradient as close as possible to the ideal SW gradient, it finds an output color image whose structure tensor is simultaneously as close as possible to the Socolinsky and Wolff structure tensor, meaning it contains the most important details, while also remaining as close as possible to that of the input RGB image, meaning its color remains the same[2].

So if we define J as the image Jacobian at a pixel, and the inner product of the Jacobian, the structure tensor, as Z, the aim of the SE method is to find an image whose structure tensor exactly matches that of the high-dimensional input image, termed Z_H, while simultaneously keeping the output gradient as close as possible to the input RGB gradient (this constraining gradient can also be a false colour mapping or other supplied image). The input color gradient is defined as $\tilde{\nabla} R$, and the final output image gradient as ∇R.

Mathematically, we find a new Jacobian ∇R (per pixel) such that

$$\arg\min_{\nabla R} ||\nabla R - \tilde{\nabla} R|| \text{ s.t. } K^T K = J^T J \tag{7}$$

Where J is the Jacobian of the input high-dimensional image H, and K the Jacobian of the output RGB image.

Details of how to solve this minimization can be found in Connah *et al.* [2]. Typically ∇R is a 3-dimensional gradient field (for a color image). The individual color planes are again found by look-up-table reintegration[5].

Iterative Image Fusion

Spectral Edge image fusion finds an RGB image R from an n-dimensional image H for which a guide RGB image is known

$$R = SE(RGB, H) \tag{8}$$

The output image R has a gradient structure similar to H but has colors similar to RGB. The main idea of this paper is apply the spectral edge algorithm in iteration

$$R_i = \mathrm{SE}(R_{i-1}, H) \tag{9}$$

where

$$R_0 = \mathrm{SE}(\mathrm{RGB}, H) \tag{10}$$

By construction the Spectral Edge algorithm forces the integrated edge to be a function of the polynomial expansion of the input hyperspectral image. This constraint is applied to avoid the reintegrated gradient field having artifacts and it also makes the whole gradient field reintegration very rapid. However, the gradient field of R_0 (∇R_0) may be quite far from the ones sought (esp. compared to the Poisson reintegration eq. 4). But, if we run the algorithm in iteration we should move to a better approximation, because the need to be close to the original RGB image is relaxed.

Fig. 1 shows the structure tensor error, measured as the L2 norm of the difference between the structure tensor of the high-dimensional input and the structure tensor of the output image, summed across the image,

$$\sum_{x \in X} \sum_{y \in Y} ||Z_H(x,y) - Z_{R_i}(x,y)||_2 \tag{11}$$

Where X and Y are the sets of possible x and y coordinates in the image.

Fig. 2 shows the RGB error, measured as the L2 norm of the difference between each RGB pixel value and that of the output image,

$$\sum_{x \in X} \sum_{y \in Y} ||RGB(x,y) - R_i(x,y)||_2 \tag{12}$$

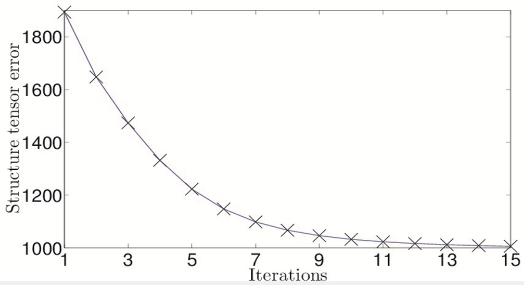

Figure 1: Structure tensor error by iteration

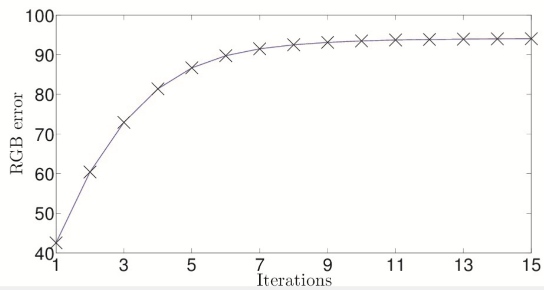

Figure 2: RGB error by iteration

Where $RGB(x,y)$ and $R_i(x,y)$ are length 3 vectors of the *R*, *G* and *B* pixel values at that *x* and *y* location.

These values have been averaged over the 16 RGB-NIR image pairs used in our psychophysical experiment, from the EPFL RGB-NIR data set[1].

As the number of iterations increases, the result has a structure tensor closer and closer to that of the high-dimensional input, meaning more of the gradient information is present in the output image. However, the result also differs more and more from the original RGB image, which may at some point lead to a less natural and pleasing output image. Both graphs appear to be approaching an asymptote, at which the output image will reach its minimum structure tensor error, and maximum RGB error. While this final output image is mathematically optimal, it may not be the preferred output for human observers.

The images appear more colorful with more iterations and there should be a point after which subjective preference decreases with increasing colorfulness[4] and detail.

Results

We performed a psychophysical experiment, comparing the standard SE method to the result after 2, 4 and 8 iterations, for RGB-NIR color image fusion. 16 test images from the EPFL RGB-NIR data set were used[1]. The result after 8 iterations is close to the final result after an infinite number of iterations.

In previous work, Hayes *et al.* compared the standard SE method to the original RGB image and the RGB-NIR color image fusion methods of Schaul *et al.*[8] and Fredembach and Süsstrunk[6]. That work demonstrated that the single iteration Spectral Edge method was preferred, so here we simply compare against that method.

Psychophysical results with 8 observers indicate a preference ranking, as shown in fig. 3. The SE method with 2 iterations is the most preferred, followed by 4 iterations, then the original method, then 8 iterations. Results were not obtained for 3 iterations, but from the ranking it looks like either 2 or 3 iterations is the peak for preference.

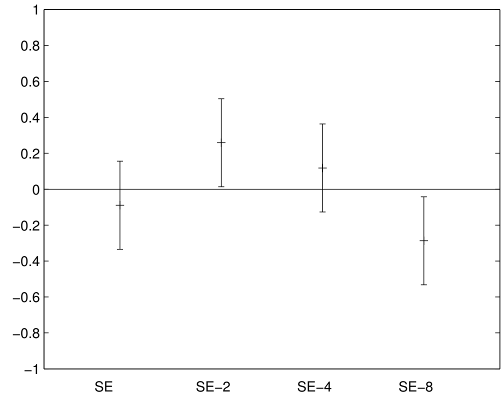

Figure 3: Psychophysical experiment results

It seems that the first few iterations produce extra detail and color which is beneficial for subjective preference, but then with successive iterations, the output image becomes too extreme and unnatural. Figs. 4, 5 and 6 demonstrate this tendency. In fig. 4, the vegetation becomes increasingly green, and the water increasingly blue with each iteration. At first this is an enhancement, as the color vividness and contrast is increased, but eventually it becomes too much and unnatural. Figs. 5 and 6 follow a similar pattern with regards to vegetation, and also in fig. 6 the road's contrast is at first improved, but then it becomes unnaturally purple after 4-8 iterations.

Conclusion

In this paper, we have proposed adding iteration to the Spectral Edge image fusion method. We have found, with a psychophysical experiment, that the result after 2-3 iterations is the most preferred.

We have demonstrated a way of improving the Spectral Edge image fusion method, by creating an output image with a structure tensor closer to that of the high-dimensional input, therefore containing more of the input detail. The color of the output image is also heightened.

However, it appears that there is a limit of detail and color intensity beyond which observer preference diminishes. Beyond this point, the image appears unnaturally detailed and colorful.

The Spectral Edge image fusion method is a fast and efficient technique, which produces vivid and detailed results. It can be used for a wide variety of applications, such as RGB-NIR image fusion, remote sensing, and medical imaging.

References

[1] Matthew Brown and Sabine Susstrunk. Multi-spectral sift for scene category recognition. *Computer Vision and Pattern Recognition, IEEE Conference on*, pages 177–184, 2011.

[2] David Connah, Mark S. Drew, and Graham D. Finlayson. Spectral edge image fusion: Theory and applications. *European Conference on Computer Vision, IEEE Conference on*, pages 65–80, 2014.

[3] Silvano Di Zenzo. A note on the gradient of a multi-image. *Computer Vision, Graphics, and Image Processing*, 33(1):116–125, 1986.

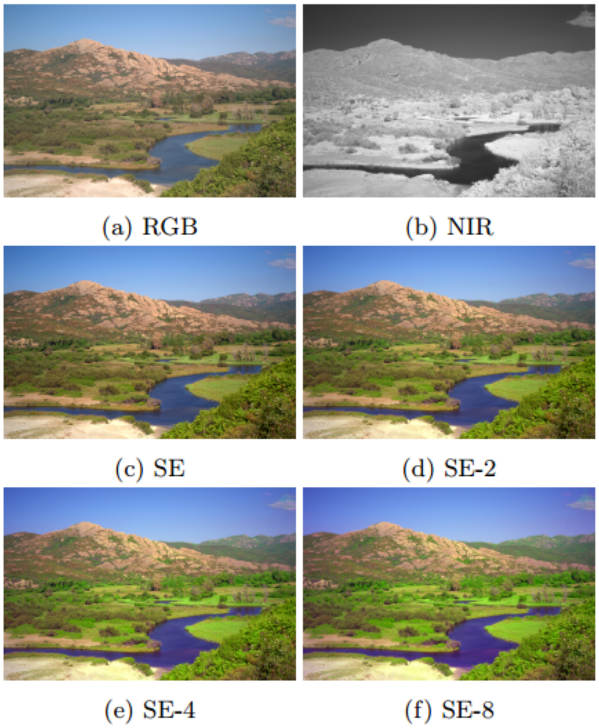

Figure 4: RGB-NIR Image Fusion: ‘Country04’ Comparison

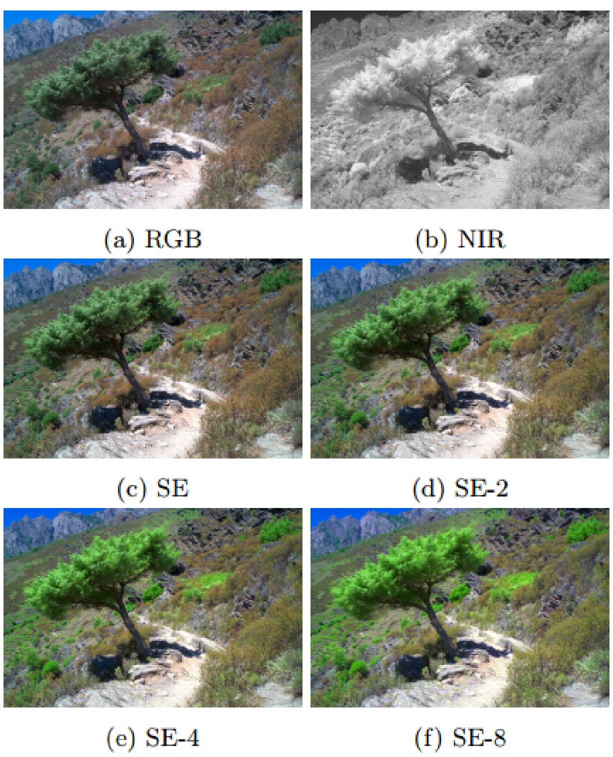

Figure 5: RGB-NIR Image Fusion: ‘Country08’ Comparison

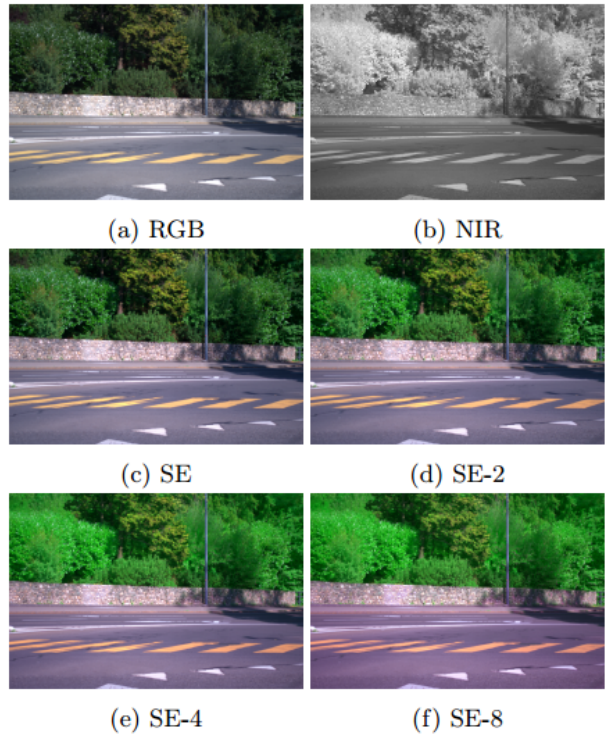

Figure 6: RGB-NIR Image Fusion: ‘street42’ Comparison

[4] Elena A Fedorovskaya, Huib de Ridder, and Frans JJ Blommaert. Chroma variations and perceived quality of color images of natural scenes. *Color Research & Application*, 22(2):96–110, 1997.

[5] Graham D Finlayson, David Connah, and Mark S Drew. Lookup-table-based gradient field reconstruction. *Image Processing, IEEE Transactions on*, 20(10):2827–2836, 2011.

[6] Clément Fredembach and Sabine Süsstrunk. Colouring the near-infrared. *Color and Imaging Conference*, pages 176–182, 2008.

[7] Alex E Hayes, Graham D Finlayson, and Roberto Montagna. RGB-NIR color image fusion: metric and psychophysical experiments. *IS&T/SPIE Electronic Imaging*, 2015.

[8] Lex Schaul, Clément Fredembach, and Sabine Süsstrunk. Color image dehazing using the near-infrared. *International Conference on Image Processing*, pages 1629–1632, 2009.

[9] Diego A Socolinsky and Lawrence B Wolff. Multispectral image visualization through first-order fusion. *Image Processing, IEEE Transactions on*, 11(8):923–931, 2002.

Author Biography

Graham D. Finlayson is a Professor of Computer Science at the University of East Anglia. He joined UEA in 1999 when he was awarded a full professorship at the age of 30. He was and remains the youngest ever professorial appointment at that institution. Graham trained in computer science first at the University of Strathclyde and then for his masters and doctoral degrees at Simon Fraser University where he was awarded a ‘Dean’s medal’ for the best PhD dissertation in his faculty. Prior to joining UEA, Graham was a lecturer at the University of York and then a founder and Reader at the Colour and Imaging Institute *at the University of Derby. Professor Finlayson is interested in 'Computing how we see' and his research spans computer science (algorithms), engineering (embedded systems) and psychophysics (visual perception).*

Alex E. Hayes received his MA in English from the University of St Andrews in 2009, and his MSc in Games Development from the University of East Anglia in 2013. He is currently pursuing his PhD studies at the University of East Anglia. His research is focused on image fusion, particularly methods based on derivatives, and applied to RGB-NIR image fusion.