REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-01-

0840

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
		Final - 01 September 1998 to 31 August 1999

4. TITLE AND SUBTITLE
Differential Geometric Tools for Image Sensor Fusion

5. FUNDING NUMBERS
F49620-98-C-0047

6. AUTHOR(S)
Dr. Lawrence B. Wolff

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Equinox Corporation
9 West 57th Street
New York, NY 10019

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NL
801 North Randolph Street, Room 732
Arlington, VA 22203-1977

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This report describes work performed under AFOSR STTR Phase I contract ~ F49620-98 C-0047 from September 1, 1998 through August 31, 1999. We present a new formalism for the treatment and understanding of multispectral images and multisensor imagery based on first order contrast information. Although little attention has been paid to the utility of multispectral contrast, we develop a theory for multispectral contrast that enables us to produce an optimal grayscale visualization of the first order contrast of an image with an arbitrary number of bands. We demonstrate how our technique can reveal significantly more interpretive information to an image analyst, who can use it in a number of image understanding algorithms. Existing grayscale visualization strategies are reviewed. A variety of experimental results are presented to support the performance of the new method.

20010227 111

14. SUBJECT TERMS	15. NUMBER OF PAGES
Imagery, multisensory, multispectral	20
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclas	Unclas	Unclas	

Differential Geometric Tools for Image Sensor Fusion

Final Report
AFOSR STTR Phase I contract # F49620-98-C-0047

Diego A. Socolinsky, Lawrence B. Wolff

Equinox Corporation
9 West 57th Street
New York, New York 10019

Abstract

This report describes work performed under AFOSR STTR Phase I contract # F49620-98-C-0047 from September 1, 1998 through August 31, 1999. We present a new formalism for the treatment and understanding of multispectral images and multisensor imagery based on first order contrast information. Although little attention has been paid to the utility of multispectral contrast, we develop a theory for multispectral contrast that enables us to produce an optimal grayscale visualization of the first order contrast of an image with an arbitrary number of bands. We demonstrate how our technique can reveal significantly more interpretive information to an image analyst, who can use it in a number of image understanding algorithms. Existing grayscale visualization strategies are reviewed. A variety of experimental results are presented to support the performance of the new method.

1 Introduction

The advent of new remote sensing and imaging technologies provides us with ever increasing volumes of multispectral data. Faced with this information explosion, it has become necessary to develop methods for analysis of such high dimensional datasets. One key aspect of this process is the visualization of multispectral data, to be used for photointerpretation. This allows an image analyst to determine regions of interest and important features in the image for further analysis or segmentation. In order to take full advantage of the human visual system, a Red-Green-Blue composite image is usually generated from the data by one of a number of statistical methods which we review in section 2. We aim to produce a one-band, grayscale visualization image from a given multispectral dataset. We propose to do this in such a way as to preserve as much local image contrast 'feature information' as possible.

Computation of contrast, which includes computation of gradient and zero-crossings, has been used in computer vision as one of the primary methods for extracting grayscale and color features

[16, 2, 4, 10, 6]. It seems plausible, therefore, that the correct way to compare versions of the same image in terms of feature information is through their contrast content. Contrast, however, is defined *a priori* only for grayscale images, so we cannot readily compare multiband images amongst themselves or multiband images with grayscale images. The first step is thus to define contrast for a multiband image. This is achieved in section 3.1 through the introduction of a differential form on the image, computed in terms of the spectral map and a metric defined on photometric space. This reduces to the standard notion of contrast in grayscale images.

Once contrast has been defined for an arbitrary image, it is natural to ask which grayscale image most closely matches the contrast information of a given multiband image. Or, how should we convert a multiband image to grayscale while preserving as much contrast information as possible? This problem is taken up in section 4, where we show the mathematical formulation of the problem, together with its solution and experimental results.

It should be noted that the solution to this problem has multiple applications. For the remote sensing community, this algorithm provides a visualization tool for realizing the full edge information content in multispectral images, such as those obtained through satellite imaging. Such high-dimensional photometric data is not easily tractable by traditional methods; in this context our reduction method yields a useful data analysis tool. In medical imaging, image fusion can be used to simultaneously visualize multiple data modalities such as CT, PET and MRI [19].

2 Review of existing techniques

Perhaps the simplest possible transformation from a multispectral image to a grayscale image is averaging of the spectral bands. This produces a visualizable image which contains information from all the bands in a unified way. However, as is easily seen, this method fails to take into account any measure of the information content in the dataset. A minor modification can be obtained by considering a weighted average, where different bands will contribute differently to the final result, depending on some pre-assigned assesment of their relative relevance in the overall image. Since it may be difficult or even impossible to determine *a priori* which bands should be emphasized over others, this method suffers from similar problems as unweighted averaging. Furthermore, relative grayvalues in different bands may be such that features are completely obliterated by this process, for example consider averaging two black and white checkerboard patterns with grayvalues reversed.

In order to overcome the shortcomings of averaging methods, we can take into account statistical information about the multispectral image. Principal Component Analysis (PCA) achieves this by considering an n-band image as a set of vectors in an n-dimensional vector space. A grayscale visualization is obtained from a multipectral image by projecting the entire distribution of spectral values onto the line spanned by the eigenspace of the covariance matrix with largest eigenvalue, and then perhaps re-scaling the result to fit the dynamic range of the output device (printer,

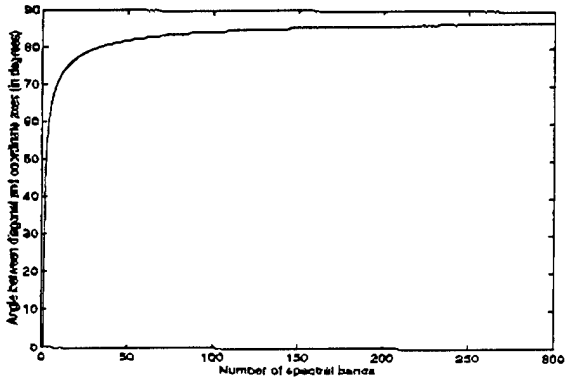


Figure 1: Angle between diagonals and coordinate axes as a function of spectral dimension.

monitor, etc). This technique has been extensively used for visualization of remote sensing imagery [13], as well as multimodal medical imaging [18]. The difference, between PCA and weighted averaging is that the line onto which we project is chosen ahead of time in the latter, whereas in the former it is determined by the global statistics of the particular image at hand. However since both methods have a common geometric foundation, they share a common problem. To see this clearly we resort to the following argument [12]. It is easy to see (Figure 1) that the cosine of the angle θ between any diagonal vector in an n-dimensional vector space and any one of the coordinate axis is given by $\cos(\theta) = \frac{1}{\sqrt{n}}$. Hence as the dimension increases, diagonals tend to become orthogonal to the coordinate axes. It follows that upon projecting the spectral measurements onto a fixed axis or a principal axis in photometric space, the contrast between adjacent pixels is always foreshortened, and it follows from the previous remark that this phenomenon becomes more severe as the dimensionality of the data increases. Any method based on linear projections will be adversely affected by this situation.

Multiresolution methods based on pyramidal decompositions [22, 3, 21] and wavelet transforms [9] have been reported. A common feature to all these methods is a selection rule which determines which band of the multiband image is 'most relevant' in a neighborhood of a given pixel; the features of the selected band are then incorporated into the fused image through various procedures. By using such a selection rule, these methods implicitly assume that there is only one dominant band at each pixel. It is well-known that multispectral imagery often exhibits large inter-band correlation, so this assumption is often violated. The consequence is that such methods do not allow for small but consistent contrast features across different bands to compound and form more salient features in the fused image. The loss of potential contrast resulting from such a selection rule is easily seen to be proportional to the square root of the number of bands.

Recently a number of other visualization methods have been proposed, most notably those based on self-organizing maps [1] and optimal projection maps [11]. While very different in their

ideas and implementation, both methods rely on global image statistics to produce a linear or non-linear projection of multispectral space onto a grayscale axis. The use of global statistics in any visualization technique has undesireable side-effects, which we exemplify in Section 4.3.

3 Contrast in a multiband image

A standard reference for all geometric notions used in this section is [20].

3.1 The contrast form

By a multiband image, we mean a rectangle $\Omega \subset \mathbb{R}^2$ together with a spectral map $s : \Omega \to \mathbb{P}^n$, where $\mathbb{P}^n$ denotes n-dimensional photometric space. We will assume that $\mathbb{P}^n$ is given an 'extended' *RGB* coordinate system, in which each band takes on values between 0 and $M < \infty$. The extension to other coordinate systems is straightforward. In this context, a grayscale image is a map $s : \Omega \to \mathbb{P}^1$. We let $\mathbb{P}^n$ have an arbitrary Riemannian metric g, which can be used to address the issue of sensor noise as in [23, 7], to introduce psychophysically correct metrics [8], or simply to manually influence the fusion (see example in Section 4.3). We expand on this in section 4.2. Finally, if the different bands in the image are the result of measurements taken with different sensors and/or at different times, then we assume that these images have been properly registered with respect to each other and brought to a common spatial resolution.

For the remainder, let $s : \Omega \to \mathbb{P}^n$ be a multiband image, let p be a point in Ω, and v an arbitrary unit vector in $\mathbb{R}^2$. In analogy with the grayscale case, we seek to define the contrast in s at p in the direction of v as the speed of spectral variation within the image in that direction. To do so, consider the following construction. Let $\gamma : [-\epsilon, \epsilon] \to \Omega$ be a curve defined on a small interval, such that $\gamma(0) = p$ and $\gamma'(0) = v$. The speed of spectral variation at p in the direction of v is given by the magnitude of the vector $s_*(v) \equiv \frac{d}{dt}(s \circ \gamma)(t)|_{t=0}$, as evaluated by the metric on $\mathbb{P}^n$. Note that since the metric of $\mathbb{P}^n$ is not assumed to be constant, this magnitude must be evaluated with respect to the metric at $s(p)$. In the language of differential geometry, the vector $s_*(v)$ is called the *pushforward* of v by s, and its expression in local coordinates is given by

$$s_*(v) = J_p v, \tag{1}$$

where J_p is the Jacobian matrix of s at the point p. Let g_q denote the matrix for g at $q \in \mathbb{P}^n$ with respect to a coordinate system. Then the contrast of s at p in the direction of v is given by the quantity

$$(J_p v)^t g_{s(p)} (J_p v) = v^t (J_p^t g_{s(p)} J_p) v. \tag{2}$$

Let $g_p^s = (J_p^t g_{s(p)} J_p)$. It follows from Equation (1) that in a coordinate system the components of g_p^s are given by

$$(g_p^s)_{ij} = \sum_{k,l=1}^{n} (g_{s(p)})_{kl} \frac{\partial s_k}{\partial x_i} \frac{\partial s_l}{\partial x_j} \quad 0 \leq i,j \leq 2. \tag{3}$$

Finally, define $\chi^2(p) = g_p^s$ to be the image *contrast form* of s. Thus the contrast of s at $p \in \Omega$ in the direction of v is given by $\chi^2(p)(v)$. Once again, in the language of differential geometry, the differential form χ^2 is called the *pullback* of g by s. In a coordinate system, the image contrast form can be expressed as

$$\chi^2(p) = \sum_{i,j=1}^{2} (g_p^s)_{ij} dx_i dx_j, \tag{4}$$

which upon evaluation on a vector $v = (v^1, v^2)$ yields the non-negative real number

$$\chi^2(p)(v) = \sum_{i,j=1}^{2} \sum_{k,l=1}^{n} (g_{s(p)})_{kl} \frac{\partial s_k}{\partial x_i} \frac{\partial s_l}{\partial x_j} v^i v^j. \tag{5}$$

This differential form encodes all first order contrast information about a multispectral image. Any false color image visualization whose corresponding image contrast differential form is identical to that for a multispectral image, replicates its first order contrast information.

Note that if all bands are brought to a common resolution prior to the computation of the contrast form (by re-sampling, interpolation or any other suitable method), then the choice of g is independent of the original resolution since g is a metric on photometric space, not on the image plane. However, if one chooses to carry the computations in the original resolutions, then g can be used to compensate for the disparity, as long as the difference in resolution is uniform in both spatial coordinates. Thus, for example if one band has half the resolution another, then we can use a metric that assigns the lower resolution band half as much weight as to the other. It is more straightforward and efficient to bring all bands to a common resolution prior to computing the contrast form, and so we use this method in general.

From Equation (3) we see that $\chi^2(p)$ is a symmetric matrix with real entries, therefore its eigenvalues are both real and non-negative. Let λ_p denote the largest eigenvalue of $\chi^2(p)$. We define the *absolute contrast* of s at $p \in \Omega$ to be equal to $\sqrt{\lambda_p}$, and we say that the eigenspace of $\chi^2(p)$ corresponding to λ_p is the *direction of maximal contrast* at p. Note that this direction does not have a preferred orientation *a priori*, i.e. the eigenspace is a line without a preassigned orientation.

Let us pause for a moment and see how these definitions simplify in the case where Ω is a grayscale image and $\mathbb{P}^1$ is given the standard Euclidean metric. Using coordinates x, y on Ω and

letting subscripts denote partial differentiation, we see that the Jacobian matrix of s is $\begin{pmatrix} s_x & s_y \end{pmatrix}$, and therefore

$$\chi^2 = \begin{pmatrix} s_x^2 & s_x s_y \\ s_x s_y & s_y^2 \end{pmatrix}. \tag{6}$$

We can readily compute the eigenvalues of this matrix to be 0 and $|\nabla s|^2$. Thus $\lambda_p = |\nabla s(p)|$, and the direction of maximal contrast is the line spanned by $\nabla s(p)$. Hence we recover the standard notion of grayscale contrast, modulo orientation. In this case, note that the direction of maximal contrast can be given the orientation determined by ∇s. For a general multiband image, there is no canonical orientation for the direction of maximal contrast; this constitutes one of the main differences between multiband and grayscale images in terms of contrast.

It follows from the previous considerations that two images carry the same contrast information if and only if they have the same contrast form. Note that the contrast form of a multiband image will, in general, have two non-zero eigenvalues. Since the contrast form of a grayscale image, shown above, always has one null eigenvalue, we have that the contrast information in a multiband image cannot, in general, be exactly reproduced by a grayscale image. This observation will be necessary in section 4.1.

3.2 The contrast vector field

The definition of contrast given in the previous section is not quite sufficient to tackle the optimal visualization problem. This is so precisely because, as we noted before, there is no canonical orientation for the direction of maximal contrast at a given point in a multiband image. We must remedy this by introducing such an orientation in a well-defined manner.

Let $\mathcal{I}(p) = dist(0, s(p))$, where 0 is the point in $\mathbb{P}^n$ all of whose coordinates are zero, and *dist* is the geodesic distance function for the metric of $\mathbb{P}^n$. The function $\mathcal{I} : \Omega \to \mathbb{R}$ is the *spectral intensity* function. Note that in the case of Euclidean photometric space we simply have $\mathcal{I}(p) = \sqrt{\sum s_i(p)^2}$. In general, $\mathcal{I}$ induces an ordering on $\mathbb{P}^n$, given by $x \geq y$ if $\mathcal{I}(x) \geq \mathcal{I}(y)$, for $x, y \in \mathbb{P}$. The function *dist* could in principle be replaced here by any non-negative function $\phi : \mathbb{P}^n \to \mathbb{R}$ whose level sets are hypersurfaces foliating $\mathbb{P}^n$, satisfying the relation $\phi^{-1}([0,a]) \subseteq \phi^{-1}([0,b])$, for all $0 \leq a \leq b$. The spectral intensity defined using ϕ instead of *dist* would thus induce a different ordering of the set of colors.

Now, we construct the *contrast vector field* V as follows. Let V_p be the unique vector of length $\sqrt{\lambda_p}$ along the direction of maximal contrast at p, for which $\nabla_{V_p}\mathcal{I} > 0$. If $\mathbb{P}^n$ is given an arbitrary Riemannian metric, this procedure can be achieved by considering the sign of the Riemannian inner product between V_p and the outward unit normal to the geodesic sphere of radius $\mathcal{I}(p)$ at p instead of the sign of $(\nabla_w \mathcal{I})(p)$. If a different function ϕ is chosen, then 'geodesic sphere'

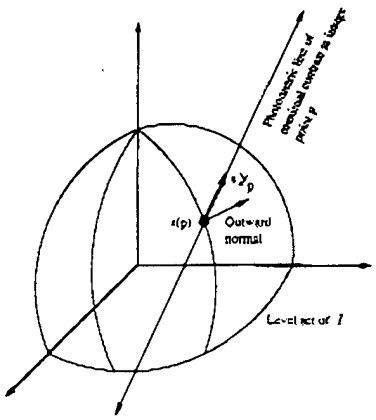


Figure 2: The direction of the contrast vector at p, is the image direction at the point p along the largest eigenvector of χ^2, that produces a positive variation in $\mathcal{I}$.

should be replaced by 'level set of ϕ' in the previous sentence. Essentially, we orient V to point in the direction of increasing spectral intensity.

It is quite important to allow for different choices of ϕ, as can clearly be seen by considering an example. If we take a 2-band image whose bands are very negatively correlated, the orientation of V_p will be inconsistent along image edges. An artificial situation exhibiting this behavior can be constructed by taking the two bands to be $f+\eta$ and $-f+\eta'$, where f is any intensity function, and η and η' are random gaussian fields representing image noise. In this case, the orientation of V_p using the ordering induced by the Euclidean metric is entirely dependent on the noise, and completely unrelated to the image features. One valid choice of ϕ in this case is any of the projections constructed in [11] (but note that these projections may not preserve the 'natural' relation of dark and bright in the image).

The contrast vector field V constructed above, encodes the first order spectral contrast information of Ω, together with intensity information. It constitutes a bridge between the multispectral and grayscale realms.

4 Grayscale visualization of local contrast

For the analytic aspects of this section, the reader may refer to [5]. The numerical methods can be found in [14].

4.1 Mathematical formulation

In light of the definitions of section 3.2, the statement of the problem is now the following: Given a multiband image $s : \Omega \to \mathbb{P}^n$ with contrast vector field V, find the grayscale image whose contrast vector field is closest to V. Equivalently, we seek the function $f : \Omega \to \mathbb{R}$ whose gradient is closest to V.

To find such a function, we would like to solve the equation $\nabla f = V$. However, this equation will in general have no solution, since it follows from the remarks at the end of section 3.1 that V need not be integrable. This means that in general it will not be possible for a grayscale image to exactly reproduce the contrast information of a multispectral image. We propose instead to find the function f for which the following functional is minimized

$$\int_{\Omega}\int |\nabla f - V|^2 \mathrm{d}x\mathrm{d}y\,. \tag{7}$$

The Euler-Lagrange equation for this functional can be easily found to be

$$\begin{cases} \Delta f = \mathrm{div} V, & \text{on } \Omega, \\ \nabla f \cdot \vec{n} = V \cdot \vec{n}, & \text{on } \partial\Omega, \end{cases} \tag{8}$$

where $\vec{n}$ is the outward unit normal to Ω. There is no natural way to assign Dirichlet boundary conditions, so we choose Neumann conditions, consistent with the idea that V should approximate the gradient of the solution image intensity function.

We would like to point out that this formulation represents a new paradigm in multispectral image visualization. The standard approach to the problem has been to seek a projection (usually linear) from photometric space onto a one-dimensional grayscale axis. Our formulation omits the projection and instead seeks the best grayscale image itself. The main consequence of this is that we have much more freedom to reproduce contrast variations. On the other hand, a side effect is that pixels with the same photometric values in the multispectral image may not have the same grayscale value in our visualization. We have not found this to be a problem, since it simply reflects the fact that in our formulation, contrast is a local quantity. In the conclusion we outline a solution that produces a unique mapping of grayvalues at the expense of reduced contrast fidelity. Also, in section 4.3 we show how one can combine the utility of the new method with some desireable properties of statistical projections, to obtain high-contrast color composites whose color scheme provides a unique correspondence between raw photometric values and image colors.

There are a number of methods for the solution of Poisson's Equation (8) with Neumann boundary conditions on a rectangle. A simple iterative scheme based on the standard five-point approximation for the Laplacian can be written for a discrete pixel grid of dimensions $[0, I] \times [0, J]$.

Letting f^0 be any initial guess, we write

$$f_{i,j}^{t+1} = f_{i,j}^t + \frac{1}{4}\left[\Delta f_{i,j}^t - (div\ V)_{i,j}\right], \tag{9}$$

for $(i,j) \in [0,I] \times [0,J]$. For interior pixels $(i,j) \in [1,I-1] \times [1,J-1]$, we use

$$\Delta f_{i,j}^t = f_{i+1,j}^t + f_{i-1,j}^t + f_{i,j+1}^t + f_{i,j-1}^t - 4f_{i,j}^t. \tag{10}$$

The divergence term must be discretized appropriately to avoid artifacts in the solution image. For the Laplacian operator in (10), we construct the contrast vector field using forward difference derivatives in the computation of χ^2. The divergence is then computed using backward difference derivatives of the components of V. In this way we insure that the correct discrete solution is found in the case where V is itself the gradient of some function.

At the boundary pixels, Equation (10) is undefined. Here, the Neumann condition in (8) is used to modify (10). For example, if (i,J), $1 \le i \le I-1$, is a pixel on the lower horizontal boundary segment of the image, the boundary condition becomes $f_{i,J+1} - f_{i,J} = V_{i,J}^2$, which implies the Laplacian operator

$$\Delta f_{i,J}^t = f_{i+1,J}^t + f_{i-1,J}^t + f_{i,J-1}^t + V_{i,J}^2 - 3f_{i,J} \tag{11}$$

The boundary condition can similarly be used to modify the computation of V and $div\ V$ at boundary pixels. Note that convergence is guaranteed regardless of the initial guess, and all solutions agree upto an overall additive constant.

A common modification of (9) in the interest of computational speed is the addition of over- or under-relaxation to speed convergence [14, 15]. In this case (9) becomes

$$\begin{cases} f_{i,j}^{t+\frac{1}{2}} &= f_{i,j}^t + \frac{1}{4}\left(\Delta f^t - div\ V\right)_{i,j} \\ f_{i,j}^{t+1} &= (1-\omega) f_{i,j}^t + \omega f_{i,j}^{t+\frac{1}{2}}, \end{cases} \tag{12}$$

for some choice of relaxation parameter $0 \le \omega \le 2$. It can be shown [14] that for the discretization above on the pixel grid $[0,I] \times [0,J]$, the ideal relaxation parameter is given by

$$\omega = \frac{2}{1+\sqrt{1-\rho^2}}, \quad \text{for} \quad \rho = \frac{\cos(\pi/I) + \cos(\pi/J)}{2}$$

Since we normally work on large pixels grids with $I, J > 100$, the relaxation parameter is very close to 2, and large speed improvement results from over-relaxation.

4.2 Noisy data

It may be the case that all bands of a multispectral image have comparable resolutions, but still each band has different noise characteristics. This may be due to a number of reasons: different technologies are used for sampling at different wavelengths, penetration of the atmosphere at different wavelengths is not uniform, cloud cover in a satellite image may introduce noise only in the visible range, but not in the infrared, etc. In some of these cases, noise characteristics of the sensor at each band may be known *a priori*. The approach described in [23] allows us to temper the effect of noise in this situation: the $n \times n$ covariance matrix associated to the sensor is computed (or interpolated) at each point in n-dimensional photometric space, yielding a positive definite bilinear form which can be thought of as a Riemannian metric on photometric space. This metric encodes the noise characteristics of the sensor in such a way that unit (in the Euclidean norm) tangent vectors in the direction of noisier bands will have shorter lengths than those in the direction of bands of higher fidelity. We use this as the metric g in Equation (2) and proceed with the rest of the algorithm with no further changes. The overall effect is to have the higher fidelity bands contribute more to the grayscale composite than those which are known to be noisier.

If we have no *a priori* knowledge of the sensor noise characteristics, but we assume independent identically distributed noise in all bands, with zero mean and distribution symmetric about the mean, then we can use a hybrid of the method above, and band averaging. Let G_σ be a 2-dimensional Gaussian distribution with zero mean and standard deviation σ. Compute the low- and high-frequency components of the multiband image s as

$$l_\sigma = G_\sigma * s, \quad \textit{and} \quad h_\sigma = s - G_\sigma * s. \tag{13}$$

Let χ^2_σ be the contrast form of the low-pass filtered image l_σ, and f_σ be the grayscale fused visualization resulting from χ^2_σ, via the method introduced above. Lastly, construct the final visualization as

$$\mathsf{f} = \frac{1}{n}\sum_{i=1}^{n}(h_\sigma)_i + f_\sigma. \tag{14}$$

The assumptions on the noise distribution are such that high frequency noise in the first term of (14) will tend to cancel out, while lower frequency features from each band will combine within the second term to produce a fusion of the low-pass filtered bands. The parameter σ clearly controls the frequency threshold past which we consider the data to be too noisy, and in the limit as $\sigma \to 0$ we recover the original procedure, with no noise attenuation.

A more involved procedure based on an adaptive metric is developed in [7]. A complete description of the method is outside the scope of the current article, but we include an example image here to demonstrate the effect of the metric in the context of noise suppression. The images in Figure 3 were obtained through the application of the technique introduced in this paper, to a 4-band aerial

image with rather severe noise corruption. The image in 3(a) was created using a Euclidean metric whereas the one in 3(b) exploits the adaptive construction in [7].

(a) (b)

Figure 3: Fusion of 4-band aerial image corrupted by noise: (a) with Euclidean metric, (b) with adaptive metric from [7]

4.3 Experimental results

We performed a series of experiments to validate our theoretical approach, using 7-dimensional data from the EOSAT thematic mapper and 12-dimensional remote sensing imagery, and 210-band data from the HYDICE project, as well as a variety of multimodal medical images. All the fused images shown in this section were computed using a Euclidean metric, unless otherwise noted.

Figures 4(a) and 4(b) show the result of applying PCA and our algorithm, respectively, to a 12-band image. A number of corresponding regions have been selected in both images to highlight some of the differences between them. Note how new features are visible in our visualization in the areas labeled 1, 2. The continuation of a road which gets lost in the PCA image is visible in area 4. More contrast detail is present in area 3, to the extent that it is possible to identify a distinctly rectangular region almost invisible in the PCA visualization. Area 5 shows higher resolution of objects in the image. The clear advantage of our visualization technique would allow an image analyst to more accurately and reliably select regions of interest in the image for further processing by image understanding algorithms. Since these algorithms become significantly more time consuming as the number of bands grows, the ability to quickly reduce the areas to which they must be applied results in increased processing efficiency.

Figures 5 and 6 show more detailed views of our methodology at work versus principal component analysis. We can see that not only does our technique highlight more detail, but it even uncovers features that were obliterated by the statistical method. It is hard to overestimate the advantage this affords the image analyst. In Figure 5 we see how a number of small islands along the coastline of Lake Worth, Florida, dissapear from the statistical visualization, while they are easily visible in ours. Likewise Figure 6 shows how we recover more of the original man-made (probably airport runways) and natural structures from a 9 band EOSAT image than PCA.

Figures 7(a) and 7(b) show the results of applying PCA and our algorithm to an artificial movie sequence produced by sliding a window over area 2 of the 12-band image in Figure 4, respectively. First note the greatly enhanced level of feature information conveyed by the images obtained by our method. In fact the roads and divisions between land plots are hardly visible in 7(a). We also see that the same physical feature looks different in different frames of the sequence in Figure 7(a). The grayscale values for a given feature do not remain constant through the sequence, as they depend on the global statistics of the particular image for which the principal components were computed. We should mention that this shortcoming is not unique to PCA; it affects any visualization method which exploits the global statistics of the image. In sharp contrast to this phenomenon, the movie sequence produced by our visualization algorithm is consistent in terms of grayscale values. This property makes it possible for us to create consistent optimal visualizations of multispectral video. We are not aware of any other methods which are currently capable of this.

Grayscale consistency on overlapping regions is an asset for multispectral image registration as well. Many manual and semi-automatic registration methods rely on an operator selecting a number of matching features from both images. These features are then used as tie-points for a discrete optimization algorithm. When distinctive features occur across different spectral bands, it is advantageous for the operator to use band composites for feature selection. However, if statistical methods are used to create the composites, then the overlapping region between the images may appear quite different in each image. This renders the tie-point selection process more difficult and less reliable. By using our visualization algorithm it is possible to obtain images with both consistent shading in the overlapping region and rich feature content.

Figure 8 shows the relative performance of our method and the multiresolution wavelet method in [9] for the fusion of chest CT. We used two contrast windows of body CT to produce these unified visualizations. Note how the wavelet method yields a lower contrast image, in which many of the details appear washed out, while our method produces crisper detail throughout. No post-processing was done on either image.

Let us see by example how particular choices of metric in photometric space can be used to achieve various desired effects. Figure 9 shows three modalities of brain MRI of the same patient. In Figure 10(a) we have the result of using a Euclidean metric in Equation (5) for the fusion of these three modalities. Note among other details, how the tumor from 9(b) and the skull from 9(a,c) are both clearly visible. To create the image in 10(b), we used a metric which emphasizes

contrast in the T2 weighted MRI for dark pixels (10–50 gray counts out of 256), which has the effect of increasing contrast between gray and white matter in the fused image without affecting, for example, the visibility of the tumorous lesion. The opposite effect is displayed in the fused image 10(c), where the metric ignores the contribution of contrast in the T2 weighted band between gray and white matter. As a result, the gray to white matter contrast in 10(c) is solely dependent on that in 9(a,c), whereas the visibility of the tumor is due largely to 9(b).

The contrast fusion technique in this paper can also be used to create color visualizations whose utility surpasses that of current methods. Perhaps the most widely used and acknowledged procedure for producing color composites which reveal the structure of multi-band remotely sensed imagery is to combine the top three principal component images. Two standard techniques are to use the three top principal components as the red, green and blue channels of a color image, or alternativelly as the intensity, hue and saturation components (in HSV coordinates), in that order [13]. Figures 11(b) 11(c) show examples of these techniques for a 4-band aerial image (red, green, blue and near-infrared), while 11(a) shows the three bands in the visible spectrum. We can use the contrast fusion method introduced in this article to create the intensity band of a color composite in the HSV coordinate system. Figure 11(d) shows the result of this procedure using the hue and saturation components from 11(a), and 11(e) shows the results of using the hue and saturation from 11(b). In both cases we see that the quality of the resulting color composite is improved by the addition of the contrast fused image as an intensity component. Most notably, in both cases it becomes easier to discriminate trees from their shadows and their surrounding background. In Figure 11(e), the image aquires a more 'natural' look, and the road markings become visible.

In terms of computational time, the algorithm is rather efficient. For example, for a 4-band image measuring 500×500 pixels, the time to complete 30 iterations of (12) is just 2.7 seconds on a Pentium III processor at 700Mhz. While our theoretical formulation requires (12) to reach a steady state, it is usually the case that visually acceptable results are reached quite quickly, and further iteration adds imperceptible changes to the fused image. Note that the computational scheme proposed is of quadratic order in the number of bands and of linear order in the number of image pixels.

5 Conclusion

The need for image fusion techniques is now greater than ever. Proliferation of imaging modalities and sensor technologies continues to flood image analysts with increasing amounts of digital information for processing. New and effective visualization methods that reduce the workload and increase the efficiency of analysts constitute therefore a crucial area for algorithm development.

In this paper, we presented an analytic definition of contrast applicable to a general multiband image and showed how our definition agrees with the standard one in the case of a grayscale

image. The problem of grayscale visualization with optimal contrast was stated then in terms of an image contrast vector field, which allowed us to obtain a clean mathematical formulation. The instantiation of this formulation yields an algorithm capable of producing consistently high quality visualizations of multiband images. The enhanced feature content of these visualizations is a powerful aid for the image analyst, who can more effectively select regions of interest or distinctive image features for further processing by image understanding algorithms. The range of applicability extends from automatic target recognition and aerial mapping to detection of brain lesions in medical imaging. A number of examples and comparisons support our performance claim, and attest to the utility of our formalism.

Two directions for further research appear necessary. First, it would be interesting to solve the reduction problem not for an optimal grayscale image, as we have done, but for an optimal projection $\pi : \mathbb{P}^n \to [0, M]$. It follows from Equation (8) that the Euler-Lagrange equation for this problem is $\Delta(\pi \circ c) = \mathrm{div} V$. Unfortunately this problem is rather ill-posed, so more constraints should be introduced in order to solve it in a meaningful way. Of course by solving this problem we give up the local consistency property of our method. Regardless, such a solution is relevant in cases where relative brightness is important in a global sense and there is little or no need for local consistency on overlapping sub-images. This may be the case for current exploitation of multimodal medical imaging of the head, where tissue attenuation is correlated with brightness. Of course, the color composites shown in Section 4.3 can also be used in this situation to convey both contrast and attenuation.

Secondly, it would be desirable to create a hybrid method exhibiting the best qualities of the multiresolution techniques [22, 3, 21, 9], and the method developed above. A challenging series of images are the Ishihara pseudo-isochromatic plates used for testing color blindness [17] (see Figure 12). We see that in these plates, most colored circles are immediately adjacent to white background pixels, and not to other colored circles. Thus the short-scale contrast in the image is not dependent on the respective colors. The eye perceives the larger scale contrast, and so should any algorithm which seeks to extend the performance of the human eye beyond the visible spectrum. Of course, by blurring the plates we can create short-scale contrast between the colored areas, and this is essentially what wavelet-based multiresolution algorithms do. Thus we see that it would be beneficial to incorporate the best aspects of multiscale techniques with those of the algorithm introduced in this article. Such a programme is currently underway, and we hope to report on the results in the near future.

References

[1] A. Manduca. Multispectral Image Visualization with Nonlinear Projections. *IEEE Transactions on Image Processing*, 5(10):1486–1490, 1996.

[2] D. Ballard and C. Brown. *Computer Vision.* Prentice Hall, 1982.

[3] P. Burt and R. Lolczynski. Enhanced image capture through fusion. In *Proceedings of IEEE 4th International Conference on Computer Vision*, volume 4, pages 173–182, 1993.

[4] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:679–698, 1986.

[5] R. Courant and D. Hilbert. *Methods of Mathematical Physics.* Interscience, 1953.

[6] A. Cumani. Edge detection in multispectral images. *CVGIP: Graphical Models and Image Processing*, (53):40–51, 1991.

[7] D. A. Socolinsky. A Locally-Adaptive Metric for Contrast-Fusion of Noisy Multimodal Imagery. In *Proceedings of Signal and Image Processing 2000*, Las Vegas, 2000. IASTED.

[8] G. Wyszecki and W. S. Styles. *Color Science: Concepts and Methods, Quantitative Data and Formulae.* John Wiley & Sons, New York, 1982.

[9] S. K. Mitra H. Li, B. S. Manjunath. Multisensor image fusion using the wavelet transform. *Computer Vision, Graphics, and Image Processing: Graphical Models and Image Processing*, 57(3):627–640, 1995.

[10] R. M. Haralick. Digital step edges from zero crossings of second directional derivatives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:58–68, 1984.

[11] G. Harikumar and Y. Bresler. Feature extraction for exploratory visualization of vector valued imagery. *IEEE Trans. on Image Processing*, 5(9):1324–1334, 1996.

[12] D. Landgrebe. On information extraction principles for hyperspectral data. Whitepaper, 1997.

[13] M. Lillesand and R. Kiefer. *Remote Sensing and Image Interpretation.* John Wiley & Sons, 3rd edition, 1994.

[14] W. Press. *Numerical Recipes in C.* Cambridge University Press, 2nd edition, 1992.

[15] J.-L. Lions R. Glowinski and R. Trémolières. *Numerical ananlysis of variational inequalities.* North-Holland, Amsterdam, 1981.

[16] A. Rosenfeld and A. Kak. *Digital Picture Processing.* Academic Press, 2nd edition, 1982.

[17] S. Ishihara. *Tests for colour-blindness.* Kanehara Shuppan, Tokyo, 1954.

[18] U. Schmiedl, D. A. Orthendahl, A. S. Mark, I. Berry, and L. Kaufman. The utility of principal component analysis for the image display of brain lesions: A preliminary, comparative study. *Magnetic Resonance in Medicine*, (4):471–486, 1987.

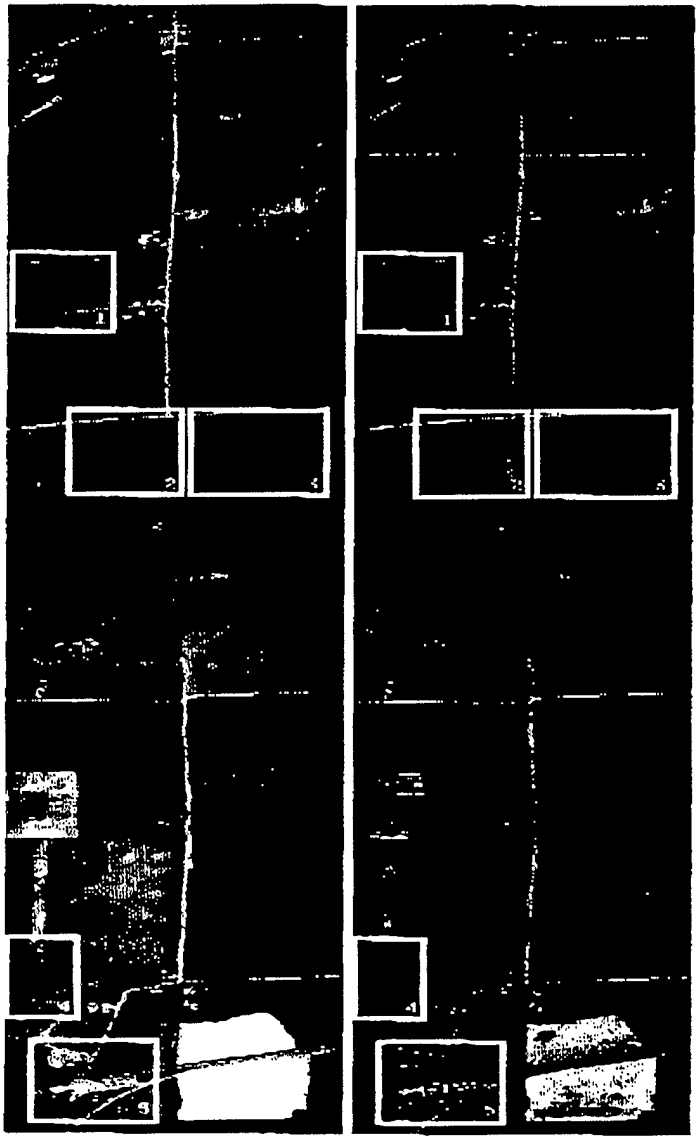
[19] D. A. Socolinsky and L. B. Wolff. Image fusion for enhanced visualization of brain imaging. In *Proceedings: SPIE Medical Imaging '99*, San Diego, February 1999.

[20] M. Spivak. *A comprehensive introduction to differential geometry*. Publish or Perish, Houston, 1979.

[21] S. K. Rogers T. A. Wilson and M. Kabrisky. Perceptual-based image fusion for hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 35(4):1007–1017, July 1997.

[22] A. Toet. Hierarchical image fusion. *Machine Vision and Applications*, pages 1–11, March 1990.

[23] L. Wolff and D. Socolinsky. Theory and analysis of color discrimination for the computation of color edges using camera sensors in machine vision. In *Proceedings: 8th Congress of the International Colour Association*, pages 515–518, Kyoto, May 1997.



(a) (b)

Figure 4: (a) Grayscale version of 12-band image computed through PCA. (b) Grayscale version of the same image computed through our algorithm.

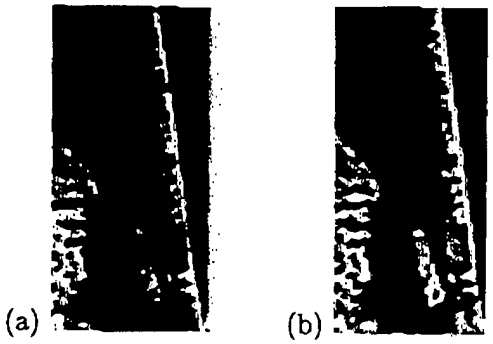


Figure 5: (a) Grayscale version of 9-band image computed through PCA. (b) Grayscale version of the same image computed through our algorithm.

Figure 6: (a) Grayscale version of 9-band image computed through PCA. (b) Grayscale version of the same image computed through our algorithm.

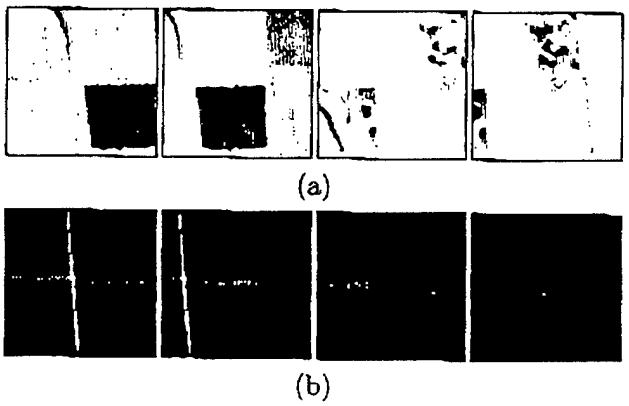
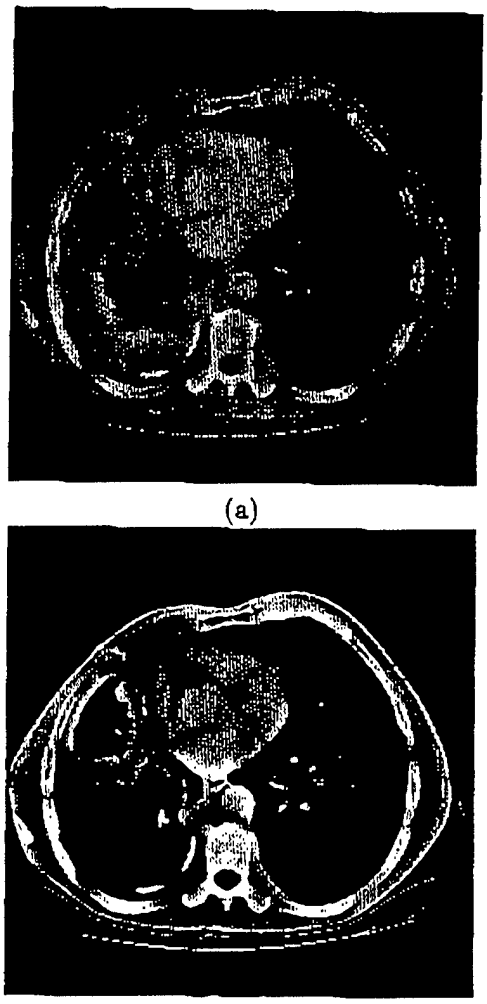


Figure 7: Visualizations of successive frames of a 12-band image sequence: (a) Through PCA. (b) Through our algorithm.



(a)

(b)

Figure 8: Comparison between (a) wavelet fusion and (b) proposed method, for the fusion of two contrast windows of chest CT.