

Journal of Medical Imaging

MedicalImaging.SPIEDigitalLibrary.org

Formulation of image fusion as a constrained least squares optimization problem

Nicholas Dwork
Eric M. Lasry
John M. Pauly
Jorge Balbás

SPIE.

Nicholas Dwork, Eric M. Lasry, John M. Pauly, Jorge Balbás, "Formulation of image fusion as a constrained least squares optimization problem," *J. Med. Imag.* 4(1), 014003 (2017), doi: 10.1117/1.JMI.4.1.014003.

Formulation of image fusion as a constrained least squares optimization problem

Nicholas Dwork,^{a,*} Eric M. Lasry,^b John M. Pauly,^a and Jorge Balbás^c

^aStanford University, Department of Electrical Engineering, Stanford, California, United States

^bStanford University, Pre-Collegiate Summer Institutes, Stanford, California, United States

^cCalifornia State University in Northridge, Department of Mathematics, Northridge, California, United States

Abstract. Fusing a lower resolution color image with a higher resolution monochrome image is a common practice in medical imaging. By incorporating spatial context and/or improving the signal-to-noise ratio, it provides clinicians with a single frame of the most complete information for diagnosis. In this paper, image fusion is formulated as a convex optimization problem that avoids image decomposition and permits operations at the pixel level. This results in a highly efficient and embarrassingly parallelizable algorithm based on widely available robust and simple numerical methods that realizes the fused image as the global minimizer of the convex optimization problem. © 2017 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMI.4.1.014003](https://doi.org/10.1117/1.JMI.4.1.014003)]

Keywords: image fusion; convex optimization; medical imaging.

Paper 16230RR received Oct. 11, 2016; accepted for publication Feb. 8, 2017; published online Feb. 28, 2017.

1 Introduction

Image fusion is the process of combining multiple images into a single one so as to provide the most relevant and complete information to the viewer in a single frame. In medical imaging, for instance, it is a common practice to fuse a higher resolution monochromatic image with a 3-channel color image of lower resolution that provides relevant diagnostic information to incorporate spatial context. If implemented naively, however, fusion may lead to the loss of relevant information from the combined images. So it is crucial for a robust fusion algorithm to be guided by rigorous principles that ensures its capability to preserve the right amount of information from all images.

1.1 Applications

A good example of the benefits of fusing a lower resolution color image with a higher resolution monochromatic one is the combination of positron emission tomography (PET) with either magnetic resonance (MR) or computed tomography (CT) imagery for diagnosing cancer.^{1,2} In this application, the monochrome PET image is converted into a 3-channel false color image through the use of a colormap (examples are shown in Fig. 1). The lower resolution PET (3 to 8 mm³) presents metabolic information and permits *in vivo* oncology studies. However, the PET signal is only strong where the PET tracer is located. This limits the amount of anatomical information, which can be provided by the higher resolution of MR (~1 mm³) or CT (~0.3 mm³). To best align the two modalities, whole-body PET/CT and whole-body PET/MR systems are used to capture the data.⁴

Image fusion can also restore the loss of information that results from scattering of fluorescence in whole-body imaging with epi-illumination, transillumination, or hybrid systems.⁵ In these systems, the light scatters as it travels through the body,

reducing the resolution and the signal-to-noise ratio (SNR) of the color image. Fusing the fluorescence image with a higher resolution visible band monochromatic image helps regain resolution and incorporates spatial context (see, for example, Fig. 7 in Sec. 3).^{7,8}

Image fusion may also prove beneficial in the diagnosis of bladder cancer where blue-light cystoscopy (BLC) can be combined with white-light cystoscopy (WLC). The BLC has been shown to increase the sensitivity and specificity of lesion detection over WLC.^{9,10} Fusing these two images may provide clinicians with a single image that combines higher sensitivity of the color BLC image with the increased SNR of the monochromatic WLC, avoiding possible mistakes resulting from alternate viewing of the two separate images.

In all these applications, it is crucial that the information incorporated into the fused image from the monochromatic image add valuable spatial context without distorting the important information provided by the 3-channel color image. The algorithm proposed in this paper works particularly well in this respect as the fused image is the result of a convex optimization problem free of *a priori* bias favoring one type of information (color or monochromatic) over the other.

1.2 Brief Overview of Existing Image Fusion Techniques

To assess the performance of the proposed algorithm, results are compared with results from four types of image fusion algorithms commonly employed in medical imaging: pixel averaging,¹¹ color-altering,¹² multiresolution methods,¹¹ and hybrid methods. Throughout this document, $X = (R, G, B)$ represents the color image, where $R, G, B \in \mathbb{R}^{M \times N}$ are the red, green, and blue channels, respectively. The variable $Y \in \mathbb{R}^{M \times N}$ represents the monochrome image and $F = (F_R, F_G, F_B)$ is the fused color

*Address all correspondence to: Nicholas Dwork, E-mail: ndwork@stanford.edu

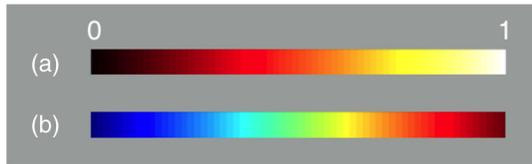


Fig. 1 Colormaps created by Mathworks:³ (a) hot colormap and (b) jet colormap.

image, where $F_R, F_G, F_B \in \mathbb{R}^{M \times N}$. The values in all images are assumed to be bounded between 0 and 1.

Unlike the images fused in surveillance applications or remote sensing, in medical applications the sensing machine is often constructed so that the gathered data are well aligned and no preprocessing is needed prior to fusing the images.¹¹ This is certainly the case in whole-body transillumination, PET/CR, and PET/MR systems; each system aligns the imagery so that fusion can operate at the pixel level.¹³

Several classes of image fusion algorithms are now presented along with an example of each.

1.2.1 Pixel averaging

A common and fast approach to image fusion is to perform a convex combination of the images to be fused: $F = w_1 X_1 + w_2 X_2 + \dots + w_k X_k$, where $\sum_i w_i = 1$ and $w_i \geq 0$ for all i . When averaging a multispectral image with a monochrome image, one can assume that the monochrome image has a uniform spectral intensity: $X_Y = (Y, Y, \dots, Y)$.

An example of a pixel averaging algorithm is alpha-blending; it is applicable when fusing two images. The alpha-blending algorithm to fuse a 3-color channel image with a monochrome image is summarized below:^{14,15}

$$F_R = \alpha R + (1 - \alpha) Y,$$

where $0 \leq \alpha \leq 1$. The procedure is repeated for the green and blue color channels.

Alpha-blending is computationally simple and very fast. However, the best value of α is image dependent and cannot be known *a priori*. Additionally, for any α value, incorporating information from one image comes at the cost of reducing the information from the other image without a clear or uniform criterion to favor one over the other.

1.2.2 Color-altering

With color-altering algorithms, the fusion is performed in an alternate color space.¹² For example, rather than fusing the red–green–blue (RGB) data, the fusion can occur in the intensity–hue–saturation (IHS), hue–saturation–value, or luma–in-phase–quadrature color space. The main steps of color-altering fusion algorithms are (1) transform the color image from RGB into another color space, (2) combine the monochrome image with a single channel in the transformed color space, and (3) inverse transform back to RGB. The effect of color-altering algorithms is to sharpen the color imagery with the structure present in the monochrome image.

An example of a color-altering fusion algorithm is IHS fusion;¹⁶ the fusion is performed between the monochrome image and the intensity channel, which is defined to be $L = R/3 + G/3 + B/3$.¹⁷ The IHS fusion result can be obtained



Fig. 2 Image pyramid with four levels and a reduction factor of 2.

solely with addition operations using the fast IHS-fusion method:¹⁸

$$\begin{bmatrix} F_R \\ F_G \\ F_B \end{bmatrix} = \begin{bmatrix} R + (Y - L) \\ G + (Y - L) \\ B + (Y - L) \end{bmatrix}. \quad (1)$$

With the fast algorithm, IHS fusion is computationally efficient and embarrassingly parallelizable, so it can be computed very quickly. IHS fusion spectrally degrades the color image,¹⁷ which is problematic for medical imaging applications where the relevant information is largely present in the hue channel.

1.2.3 Multiresolution methods

Rather than fusing two images at their native resolutions, multi-resolution methods fuse the data at several different scales.¹¹ The final fused image is a function of the results from all scales. The set of scaled images is called a pyramid; an example is shown in Fig. 2. For this figure, the next level down is created by convolving the image of the current level with a Gaussian kernel (i.e., low-pass filtering the image) and then downsampling by 2.

Discrete wavelet transforms (DWTs) are often used to compute the pyramid.¹⁹ An example of a multiresolution fusion method is wavelet fusion (WF), described in Algorithm 2 (in Sec. 3).

1.2.4 Hybrid methods

Several fusion algorithms combine aspects of the fusion algorithm classes previously described.¹⁹ An example of a hybrid algorithm, referred to as the Wav-IHS fusion algorithm,²⁰ is shown in Algorithm 4 (in Sec. 3).

Hybrid methods are able to incorporate the benefits of many of the previous fusion algorithms (along with the computational cost of each technique). The Wav-IHS fusion algorithm results in fused images of especially high quality. The results presented in Sec. 3, however, demonstrate that the proposed constrained least squares (CLS) fusion algorithm more robustly preserves the salient information from both images. Additionally, CLS is embarrassingly parallelizable and lends itself well to a graphics processing unit (GPU) implementation.

1.2.5 Pansharpening

This paper presents an algorithm for the branch of image fusion that combines multiple images into a single image. A related branch of image fusion is called pansharpening. The inputs

to a pansharpening algorithm are a set of low-resolution multispectral images $\{M_k : k = 1, \dots, K\}$ and a high-resolution panchromatic image P .²¹ The goal of pansharpening is to output high-resolution multispectral images; this is a special case of the image deblurring problem,²² where the panchromatic image informs the result. Two assumptions are generally made:²¹ (1) the low-resolution images are related to a high-resolution image through a known degradation modeled by a convolution and additive noise [as shown in Eq. (2)] and (2) the panchromatic image is an affine function of the low-resolution images [as shown in Eq. (3)]

$$M_k = D_k(H_k * P) + n^k, \tag{2}$$

$$P = \sum_{k=1}^K w_k M_k + \Delta + n^p, \tag{3}$$

where D_k is a downsampling operator of the k 'th image, $*$ represents convolution, and n^k represents the noise in the k 'th image. The weight w_k is a function of the absorption spectrums of the k 'th multispectral camera and the panchromatic camera. With these assumptions, pansharpening can be accomplished by solving a regularized least squares problem²³

$$\begin{aligned} & \text{minimize}_{\{M_k \text{ for all } k\}} \left\| P - \sum_{k=1}^K w_k M_k \right\|_2^2 \\ & + \lambda \sum_{k=1}^K \|M_k - D_k(H_k * P)\|_2^2 + \mathcal{R}(M), \end{aligned} \tag{4}$$

where $\|\cdot\|_2$ represents the $L2$ norm, $\lambda \in \mathbb{R}$ is a parameter, $M = (M_1, \dots, M_K)$, and \mathcal{R} is an optional regularization function. Several different regularization functions and dictionary-based methods have been investigated.^{21,23-25}

Though pansharpening is fulfilling a different purpose, we draw inspiration from the literature and formulate image fusion as a (constrained) least squares problem. Our optimization problem will be similar to (but substantially different from) that of Eq. (4); the difference permits a geometric interpretation of the problem that will lead to a computationally efficient embarrassingly parallelizable algorithm.

This paper is organized as follows. Section 2 describes the proposed algorithm and provides a geometric interpretation of it. Results are presented in Sec. 3, where the algorithm is tested on images of the types mentioned in Sec. 1.1 and compared to the techniques described in Sec. 1.2. Section 4 provides a summary on the performance of the algorithm and suggests future lines of work.

2 Constrained Least Squares Image Fusion

The CLS algorithm formulates image fusion as a convex optimization problem, a common approach in image processing.²⁶⁻²⁸ This approach avoids image decomposition, which permits the problem to be separated at the pixel level and facilitates the parallelization of the algorithm. Convexity also makes the problem more robust as convex problems are characterized by having a global minimizer; thus, their solution is optimal and can be found with simple and efficient numerical methods. A fused image realized as the global minimizer of a convex functional incorporates information from both images while staying as close to each of them as possible, always incorporating both

sources and avoiding bias toward color or monochromatic information.

2.1 Proposed Algorithm

Recall that $X = (R, G, B)$ denotes the color image, where $R, G, B \in \mathbb{R}^{M \times N}$ correspond to the red, green, and blue channels, respectively, and $Y \in \mathbb{R}^{M \times N}$ denotes the monochrome image. All values in X and Y satisfy $0 \leq X, Y \leq 1$. The four channels will be combined into a single-color image $F = (F_R, F_G, F_B)$. Intuitively, a fused image of high quality would be one that is close to the color image X and simultaneously close to the monochrome image Y . This goal could be accomplished by finding an F that minimizes the following two quantities:

$$\|X - F\|_F = \left\| \begin{bmatrix} R \\ G \\ B \end{bmatrix} - \begin{bmatrix} F_R \\ F_G \\ F_B \end{bmatrix} \right\|_F, \quad \|f(F_R, F_G, F_B) - Y\|_F, \tag{5}$$

where $\|\cdot\|_F$ denotes the Frobenius norm and $f(F_R, F_G, F_B) = w_R F_R + w_G F_G + w_B F_B$ is a weighted average of the three color channels. This is a multicriterion optimization problem.²⁹ It can be converted into a convex optimization problem as follows:

$$\begin{aligned} & \text{minimize} \quad \|X - F\|_F^2 + \gamma \|f(F_R, F_G, F_B) - Y\|_F^2 \\ & \text{subject to} \quad 0 \leq F \leq 1, \end{aligned} \tag{6}$$

where $0 \leq F \leq 1$ means that each component of F is in $[0, 1]$ and γ is a parameter set by the user. A block diagram depicting the scalarization of the objective function is shown in Fig. 3.

Problem [Eq. (6)] is equivalent to the following CLS problem:

$$\begin{aligned} & \text{minimize} \quad \left\| \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \\ w_R \sqrt{\gamma} & w_G \sqrt{\gamma} & w_B \sqrt{\gamma} \end{bmatrix} \begin{bmatrix} \hat{F}_R \\ \hat{F}_G \\ \hat{F}_B \end{bmatrix} - \begin{bmatrix} \hat{R} \\ \hat{G} \\ \hat{B} \\ \hat{Y} \sqrt{\gamma} \end{bmatrix} \right\|_2 \\ & \text{subject to} \quad 0 \leq \hat{F} \leq 1, \end{aligned} \tag{7}$$

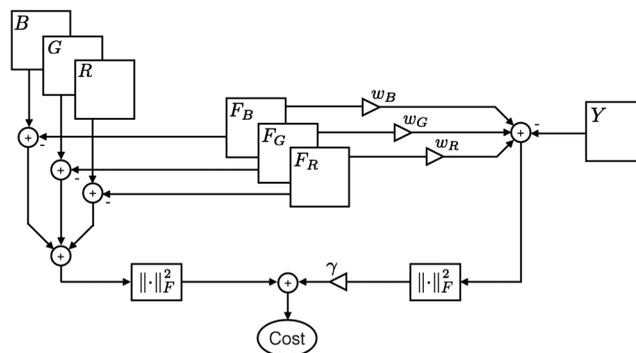


Fig. 3 Block diagram depicting the calculation of the cost that gets minimized.

where \hat{F}_R denotes the column-major extension of the matrix F_R , $\hat{F} = (\hat{F}_R, \hat{F}_G, \hat{F}_B)$, and I is the identity matrix.

Note that problem of [Eq. (7)] is completely separable across pixels. That is, the result of each pixel is independent of any other pixel. Thus, the problem is embarrassingly parallelizable, where the value of the (i, j) 'th pixel in the fused image can be found by solving the following problem:

$$\begin{aligned} & \text{minimize} \left\| \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \underbrace{w_R\sqrt{\gamma} \quad w_G\sqrt{\gamma} \quad w_B\sqrt{\gamma}}_A & \underbrace{\begin{bmatrix} (F_R)_{ij} \\ (F_G)_{ij} \\ (F_B)_{ij} \end{bmatrix}}_{x_{ij}} - \underbrace{\begin{bmatrix} R_{ij} \\ G_{ij} \\ B_{ij} \\ Y_{ij}\sqrt{\gamma} \end{bmatrix}}_{b_{ij}} \end{bmatrix} \right\|_2^2 \\ & \text{subject to } 0 \leq F_{ij} \leq 1. \end{aligned} \quad (8)$$

Optimization problems of this form can be solved using the fast iterative shrinkage threshold algorithm (FISTA).³⁰ However, FISTA is computationally expensive. Instead, a much more efficient algorithm can be used to solve this problem; this method is presented in Sec. 2.2.

2.2 Geometric Interpretation: Convex Optimization and Algorithm Robustness

Considering each term of Eq. (5) geometrically leads to an efficient algorithm for solving problem [Eq. (8)]. Since the problem is separable across pixels, without loss of generality, only the (i, j) 'th pixel will be considered in this analysis.

The geometric interpretation of minimizing $\|X_{ij} - F_{ij}\|_2$ is straight forward; for any pixel, the corresponding pixel of the fused image should be close (in a Euclidean distance sense) to the original color image pixel's value (R_{ij}, G_{ij}, B_{ij}) .

A geometric interpretation of the $\|f[(F_R)_{ij}, (F_G)_{ij}, (F_B)_{ij}] - Y_{ij}\|_2$ term will require more analysis. Consider the function f ; any level set of this function is a plane perpendicular to the vector $w = (w_R, w_G, w_B)$. Minimizing this term of the objective function places the point F_{ij} on the Y_{ij} level set of the function f .

A depiction of the geometric interpretation is shown in Fig. 4. The point (R_{ij}, G_{ij}, B_{ij}) is the location of the (i, j) 'th color pixel in a three-dimensional space. The vector w represents the three element weight vector. The planar segment represents a

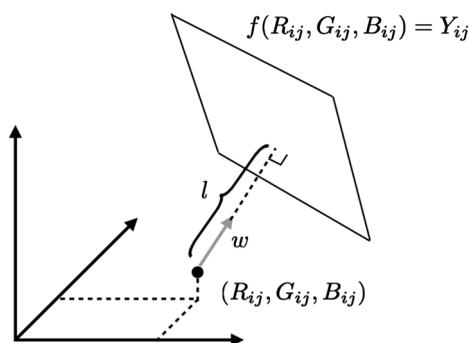


Fig. 4 The point that minimizes the objective function lies on the line segment l . If $\gamma = 0$ then (R_{ij}, G_{ij}, B_{ij}) is the optimal point. As γ is increased, the optimal point moves along the line segment in the direction w toward the Y_{ij} level set of the function f .

Algorithm 1 CLS fusion.

Prestep: Calculate \mathbf{A}^\dagger (the pseudoinverse of \mathbf{A}).

for Every pixel (i, j) **do**

Step 1: Find $\hat{\mathbf{x}}_{ij} = \mathbf{A}^\dagger \mathbf{b}_{ij}$.

Step 2: Project $\hat{\mathbf{x}}$ onto $[0, 1]^3$ along l .

end

portion of the Y_{ij} level set of f . If γ is 0, then the optimal point is the coordinate (R_{ij}, G_{ij}, B_{ij}) , which makes the objective function evaluate to 0. If γ has a very large value, then any point on the Y_{ij} level set minimizes the second term of the objective function. The line segment l consists of all possible optimal points of problem [Eq. (8)]. The parameter γ determines where on that line segment the optimal point lies. If $\gamma = 0$, then the optimal point is (R_{ij}, G_{ij}, B_{ij}) ; as γ is increased, the optimal point moves along l in the direction w toward the Y_{ij} level set of f . This interpretation leads to the key attribute of this fusion algorithm; even when γ is high, the algorithm retains the color information.

The constraint of problem [Eq. (8)] requires that the optimal point satisfy $0 \leq F_{ij} \leq 1$. If the point that minimizes the objective lies outside the feasible range then the optimal point is the intersection of the line segment l with the boundary of the $[0, 1]^3$ cube (where the exponent here means the Cartesian cross product).

With this interpretation, it becomes evident that the solution to problem [Eq. (9)] can be found by minimizing $\|\mathbf{A}\mathbf{x}_{ij} - \mathbf{b}_{ij}\|_2$ and then performing a projection of the result onto the cube $[0, 1]^3$ along the line l parameterized by λ in the following equation:

$$l(\lambda) = \begin{cases} (R_{ij}, G_{ij}, B_{ij}) + \lambda w & \text{if } F_{ij} > 1 \\ (R_{ij}, G_{ij}, B_{ij}) - \lambda w & \text{otherwise} \end{cases} \quad (9)$$

The overdetermined linear system $\mathbf{A}\mathbf{x}_{ij} = \mathbf{b}_{ij}$ only has three unknowns, so the entire system can easily be stored in a computer's memory. The $\|\mathbf{A}\mathbf{x}_{ij} - \mathbf{b}_{ij}\|_2$ term can be minimized by computing the pseudoinverse of \mathbf{A} and right multiplying by \mathbf{b}_{ij} .^{31,32} Algorithmic efficiency can be gained by computing the pseudoinverse \mathbf{A}^\dagger only once for all vectors \mathbf{b}_{ij} .

This algorithm, minimizing $\|\mathbf{A}\mathbf{x}_{ij} - \mathbf{b}_{ij}\|_2$ for each pixel and projecting the result onto $[0, 1]^3$, is called the CLS fusion algorithm. CLS fusion is summarized in Algorithm 1.

By formulating image fusion as a convex optimization problem, existing numerical methods to determine the optimal solution are immediately applicable;²⁹ e.g., FISTA can be used to solve Eq. (8). The geometric interpretation has led to the simpler and more efficient CLS algorithm; its computational complexity is $\mathcal{O}(MN)$.

An approximation to the CLS fusion algorithm is to perform a Euclidean projection of the result of step 1 onto $[0, 1]^3$. This approximation requires fewer operations than a projection along l and does not significantly change the results (data not shown).

3 Results

This section presents results of the CLS fusion algorithm for several different applications and compares them to the techniques presented in Sec. 1.2. The weight vector was chosen

to be $w = (1/3, 1/3, 1/3)$ so that $f(X)$ yields the intensity channel of the color image.¹⁷ Results were generated using the MATLAB[®] programming language version R2016a by Mathworks.

Figures 5–9 only show results on images that were previously published; the images are reprinted with permission and the necessary language stating so is included in the caption of each relevant figure. The MR data of Fig. 10 were gathered with Institutional Review Board (IRB) approval, Health Insurance Portability and Accountability Act (HIPAA) compliance, and patient informed assent/consent. The WLC and BLC images of bladder tissue for Fig. 11 were obtained from Dr. Joseph C. Liao as part of standard clinical care at the Veterans Affairs Palo Alto Health Care System using a clinical system (D-Light, Karl Storz Endoscopy, El Segundo, California). These images were de-identified and no protected health information was provided to the engineering team.

3.1 Positron Emission Tomography/Magnetic Resonance

Figure 5 shows data of a 3'-[¹⁸F] fluoro-3'-deoxythymidine PET image and a T1 weighted MR image of a living BALB/c mouse bearing a CT26 colon carcinoma.³³ A hot (or fire) color map was used to convert the PET image into a false color image.

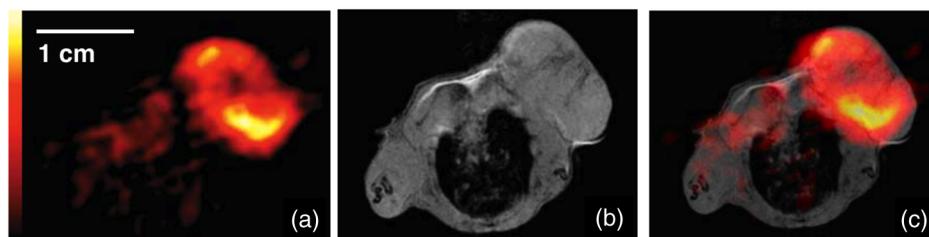


Fig. 5 (a) PET image of a BALB/c mouse bearing a CT26 carcinoma, (b) T1 weighted MR image, and (c) the CLS fused image with $\gamma = 3$. Figures (a) and (b) were adapted by permission from Macmillan Publishers Ltd.,³³ Copyright © 2008.

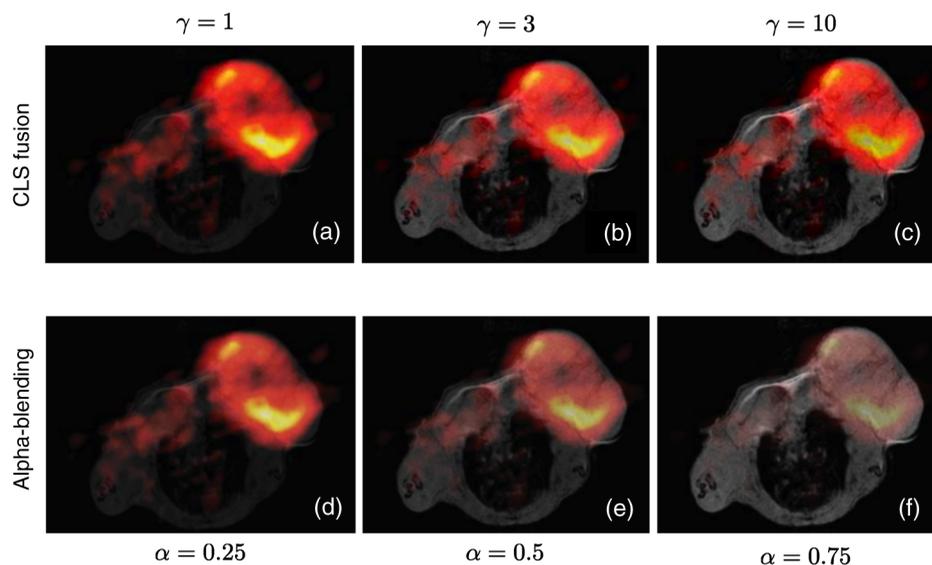


Fig. 6 (top) CLS fusion results and (bottom) alpha-blending results. (a) $\gamma = 1$, (b) $\gamma = 3$, (c) $\gamma = 10$, (d) $\alpha = 0.25$, (e) $\alpha = 0.5$, and (f) $\alpha = 0.75$.

The top row of Fig. 6 shows results for the CLS fusion algorithm for various values of γ . As is observed, increasing γ increases the intensity of the monochrome image in the fused result. These results are compared to that of the alpha-blending algorithm. For comparable levels of intensity from the monochrome image, the CLS fusion algorithm is able to retain much more of the information from the color imagery.

3.2 Transillumination with Fluorescence

As discussed in Sec. 1.1, fusion provides spatial context to whole-body fluorescence imaging. This section shows results of the CLS algorithm applied to fluorescent false color imagery fused with monochrome imagery captured by a visible light camera. These results are presented alongside results obtained with WF, which is summarized in Algorithm 2.

Figure 7 shows images captured with a small animal whole-body transillumination fluorescence imaging system.⁶ With the colormap presented, unlike the fire colormap, the information is largely encoded into the hue of the image (and not the saturation or intensity). Thus, it is important that the fusion algorithm retains the appropriate hue so that the clinician can properly interpret the amount of fluorescence present at each location. In Fig. 7(c), one can observe that the WF alters the saturation of much of the image until the hue is no longer discernible. This

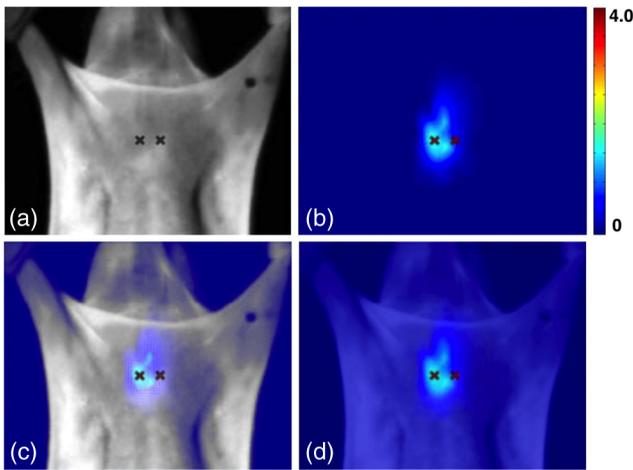


Fig. 7 (a) Monochrome reflectance image, (b) Net tdTomato fluorescence (counts/s/nm/ 10^6 cells), (c) WF result, and (d) CLS result with $\gamma = 2$. The crosses indicate the positions of the illumination sources. Figures (a) and (b) were originally published in Ref. 6. Copyright © 2008 by SPIE and Photonics. They are reprinted with permission.

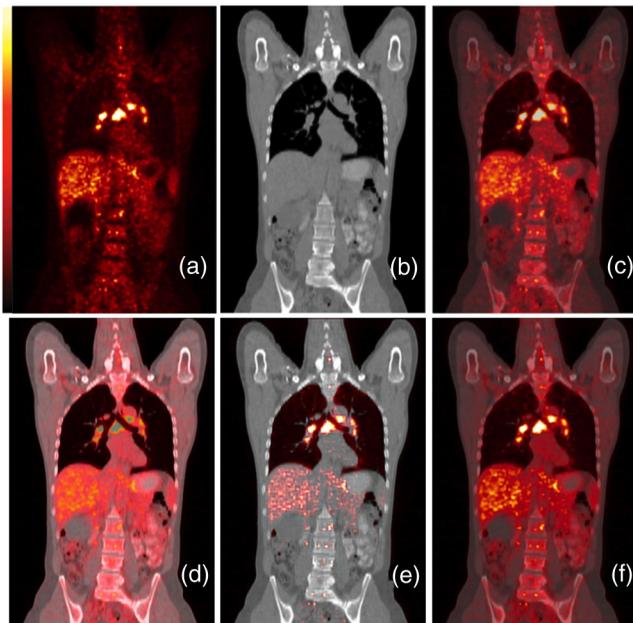


Fig. 8 (a) PET image, (b) CT attenuation map, (c) CLS fusion result with $\gamma = 2$, (d) IHS fusion result, (e) WF fusion, and (f) Wav-IHS fusion result. Figures (a) and (b) were originally published in Ref. 34. Copyright © by the Society of Nuclear Medicine and Molecular Imaging, Inc. They are reprinted with permission.

makes WF an inappropriate technique to fuse this type of imagery. CLS is able to incorporate the structure of the monochrome image without significantly altering the hue of the result, thus a clinician can discern how much fluorescence is present in the image and where the fluorescence is located.

3.3 Positron Emission Tomography/Computed Tomography

This section presents the paper's most rigorous comparison of the different fusion algorithms. PET and CT images are fused

together with CLS, IHS, WF, and Wav-IHS. IHS fusion is summarized in Algorithm 3, and Wav-IHS fusion is summarized in Algorithm 4.

Figure 8 shows results from fusing a fluorodeoxyglucose ^{18}F -FDG PET image and a CT attenuation map of a patient with multiple avid mediastinal and hilar lymph node metastases.³⁴ The data were collected using a combination PET/CT system. The IHS fusion result shown in Fig. 8(d) retains much of the spatial information from the CT image and shows metabolic activity from the PET image. However, much of the color has been significantly altered and the brightest spots of metabolic activity in the heart have been drastically reduced. These changes do not present an accurate representation of metabolic information to the clinician. The WF result shown in Fig. 8(e) only retains the brightest spots of the PET image; much of the more subtle information is lost. The CLS and Wav-IHS results are able to successfully provide context with the monochrome data and retain the information present in the color image.

In some cases, CLS significantly outperforms Wav-IHS. In Fig. 9, holes were artificially inserted into the CT attenuation coefficient map using the mask shown in Fig. 9(a); the two holes in the mask are labeled "A" and "B." Figure 9(b) shows the masked CT image. Figures 9(c) and 9(d) show the Wav-IHS fusion and CLS fusion results, respectively. The Wav-IHS result fills in hole A almost completely with data from the color image; only the outline of that structure remains in the fused image; the structure of hole B is retained intact. However, the CLS result properly retains the structure from both holes of the CT image in the fused image. This suggests that while the fusion scheme of the Wav-IHS algorithm may neglect information from one of the images in some regions, CLS blends together whatever salient information may be present in both images.

3.4 Magnetic Resonance Anatomy and Flow

Figure 10 shows another case where CLS outperforms Wav-IHS fusion. MR can be used to simultaneously collect velocity and anatomical data with a four-dimensional flow collection sequence. The data presented in Fig. 10 were collected using the variable-density sampling and radial view-ordering sequence with image navigation and soft gating to suppress motion artifacts.³⁵ Since there are three dimensions of velocity flow, the velocity in each dimension can be represented as a different color. Typically, the velocity field is overlaid on top of an anatomical image; for example, individual velocity arrows might be drawn, or the velocity image may be made opaque and only those pixels with negligible velocity show anatomy.³⁵ Figure 10(a) displays the maximum intensity projection (MIP) MR image, which shows the anatomy. Figure 10(b) displays the MIP velocity flow field. The right-left (RL)/anterior-posterior (AP)/superior-inferior (SI) motion are represented by red/green/blue, respectively. In this figure, larger velocities are represented with larger color values. Figures 10(c) and 10(d) show the result of the Wav-IHS and CLS fusion algorithms, respectively. The Wav-IHS result shows artifacts resulting from the structure of the wavelet kernels; these artifacts are absent in the CLS result. Additionally, CLS incorporates more anatomical information without sacrificing the information in the velocity flow field.

3.5 Blue-Light Cystoscopy in the Bladder

In this section, we fuse together a monochrome image with a true color fluorescent image. We will show that the CLS fusion

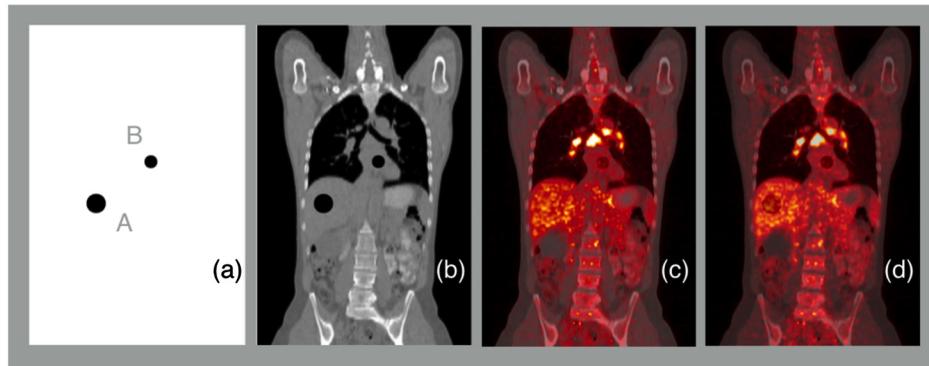


Fig. 9 (a) CT mask, (b) masked CT attenuation map, (c) Wav-IHS fusion result, and (d) CLS fusion result with $\gamma = 2$.

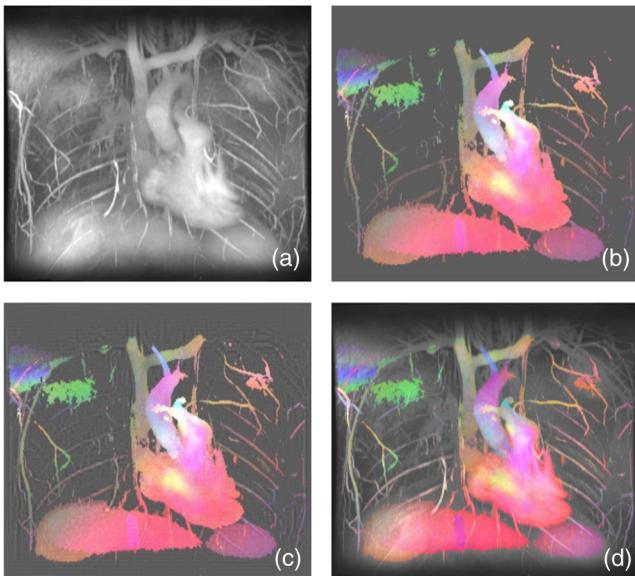


Fig. 10 (a) MIP anatomical MR image, (b) MIP MR flow image where red/green/blue represent RL/AP/SI motion, respectively, (c) Wav-IHS fusion result, and (d) CLS fusion result with $\gamma = 9$. The data presented here were gathered with IRB approval, HIPAA compliance, and patient informed assent/consent.

algorithm can improve the SNR of a color image by fusing it with a higher SNR monochrome image.

WLC has been used to detect bladder cancer by searching for anomalies in the bladder tissue surface. By labeling the bladder

tissue with hexaminolevulinate and illuminating the tissue with blue light, cancerous lesions fluoresce a pink color, which makes detection easier; this technique is known as BLC.^{9,10} The BLC image has a lower SNR due to the reduced signal intensity and the increased scattering of blue light in water, and so the WLC image is more detailed than the BLC image. Figure 11 shows results of fusing together a BLC color image with a monochrome WLC image. The fusion retains the color information of the BLC and shows the tumors well. Moreover, the increased signal level and contrast of the WLC aids the spatial understanding of the clinician; the vasculature that is easily visible in the WLC remains visible in the fused image.

Unlike the PET/MR or PET/CT systems, this data were not collected simultaneously, and so it was not perfectly aligned prior to fusion. A projective transformation was used to align the data. Harris corners were identified in the luminance channel of the color image.³⁶ Each corner was tracked into the WLC image using template matching with normalized cross correlation as the matching metric.³⁷ These points were used to determine the homography that aligned the WLC image with the BLC image.³⁸ To make this determination resilient to points that were erroneously matched, the random sample consensus algorithm was used with a threshold of 1.5 pixels.³⁸

3.6 Processing Time

The processing time for CLS ranges between 0.02 and 0.2 s with a MATLAB[®] implementation for the results shown in this paper (image sizes range from 320×350 to 1096×1336). Since CLS

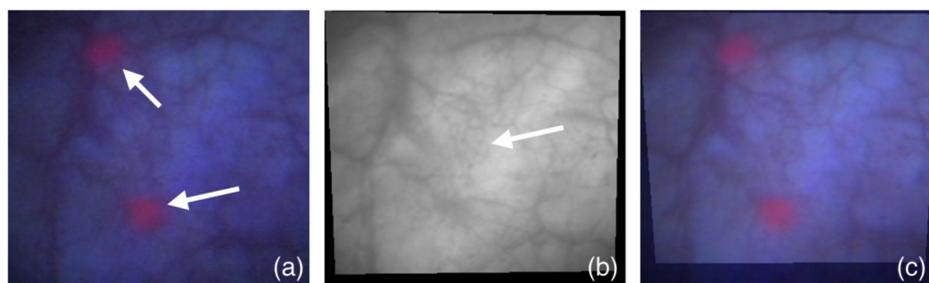


Fig. 11 (a) BLC image of a bladder, arrows point to cancerous lesions highlighted with pink fluorescence, (b) WLC image of a bladder aligned to BLC image, arrow points to vasculature that is not clearly seen in the BLC image, and (c) the fused image with $\gamma = 2$, cancerous lesions and vasculature are both clearly seen.

Algorithm 2 WF¹⁹

Step 1: Perform a Haar DWT on the monochrome image and on each channel of the color image.

Step 2: Create a new set of data representing the DWT of the fused image. Set the value of each pixel according to the following expression:

$$(\tilde{F}_R)_{ij} = \begin{cases} \tilde{R}_{ij} & \text{if } |\tilde{R}_{ij}| > |\tilde{Y}_{ij}|, \\ \tilde{Y}_{ij} & \text{otherwise} \end{cases},$$

where i and j are the row and column indices of the pixel and \tilde{F}_R is the DWT of the red channel of the fused image. Repeat this for the blue and green channels of the fused image.

Step 3: Perform an inverse DWT on each channel of the fused wavelet data.

Algorithm 3 IHS fusion^{16,18}

Step 1: Convert the RGB image to an IHS image.

Step 2: Replace the intensity channel with the monochromatic image.

Step 3: Convert the new IHS image back to an RGB image.

Algorithm 4 Wav-IHS hybrid fusion²⁰

Step 1: Convert the color image to the IHS color space.

Step 2: Perform the WF of Algorithm 2 on the intensity channel and the monochrome image.

Step 4: Set the intensity channel of the fused image to the result of step 2. Set the hue and saturation channels to the corresponding channels of the original color image.

Step 5: Convert the fused image to the RGB color space.

is a pixel-based fusion algorithm, it lends itself well to a GPU implementation, which would make CLS a real-time fusion algorithm.

3.7 Quality Metrics

To quantitatively compare the performance of the different fusion algorithms, we chose two metrics: the ERGAS metric and mutual information (MI). For the alpha-blending algorithm, a value of 0.5 was used for α .

The ERGAS metric was intended to describe an overall image quality and is defined as follows:³⁹

$$\text{ERGAS} = 100 \frac{h}{l} \sqrt{\frac{1}{K} \sum_{k=1}^K \left[\frac{\text{RMSE}(k)}{\mu(k)} \right]^2},$$

where h/l is the ratio of pixel sizes between the monochrome and color images, K is the total number of color channels (4 for

Table 1 ERGAS quality metric for fusion results.

	α -blending	IHS	WF	Wav-IHS	CLS
PET/MR	0.09	0.14	0.18	0.075	0.061
Fluorescence	0.65	0.88	1.3	0.34	0.35
PET/CT	0.66	1.1	1.3	0.51	0.43
PET/CT and holes	0.66	1.1	1.3	0.51	0.43
MR flow	0.042	0.072	0.064	0.041	0.038
WLC + BLC	0.12	0.19	0.23	0.068	0.09

Table 2 MI comparing fused image with color image.

	α -blending	IHS	WF	Wav-IHS	CLS
PET/MR	0.65	0.39	0.61	0.84	0.76
Fluorescence	0.72	0.53	0.63	1.1	1.1
PET/CT	0.88	0.70	0.58	1.0	1.1
PET/CT and holes	0.87	0.69	0.58	1.0	1.1
MR flow	0.98	0.61	0.83	1.4	1.2
WLC + BLC	1.7	1.4	0.96	2.5	2.4

the problems described in this paper), $\text{RMSE}(k)$ is the root-mean-square error of the k 'th color channel and the fused image, and $\mu(k)$ is the mean value of the k 'th color channel. Lower ERGAS values indicate higher quality fusion results. Table 1 shows ERGAS values for each fusion algorithm on all the data presented in this paper. The CLS fusion algorithm consistently yields low ERGAS values.

MI is a measure of how predictable the joint distribution of two random variables is from their marginal distributions random variables from their marginal distributions. Since the relevant quantitative information is often encoded into the hue of the imagery in medical imaging, we computed the MI between the original color image and the fused image. Higher MI values indicate higher quality fusion results. Table 2 shows MI values for each fusion algorithm on all the data presented in this paper. The CLS fusion algorithm consistently yields high MI values.

With respect to both quality metrics, the CLS algorithm is a top performing fusion algorithm.

4 Conclusion and Future Work

This paper presents an embarrassingly parallelizable algorithm to fuse a color image with a monochrome image. Even though the fused image is the result of a CLS problem with nondifferentiable constraints, this problem can be solved with an efficient noniterative algorithm: solve a least squares problem and project the solution on the $[0,1]^3$ cube. The algorithm retains color information very well while managing to incorporate spatial context and signal strength present in a monochrome image. This is important for medical imaging applications where the color often portrays a physical quantity, e.g., the amount of PET tracer

present or the amount of fluorescent light emitted. Results were shown from a variety of medical imaging applications, and the ERGAS quality metric demonstrated that the CLS algorithm consistently produced fused images of high quality. MATLAB® files for this project can be found in Ref. 40.

With the development of CLS, there are many opportunities for future work. Since it is a pixel-based fusion, CLS is well suited for implementation on a GPU that could become a real-time implementation. In the case of WLC + BLC fusion, a real-time implementation would require development of an optical system that could capture aligned images. With its high quality, an investigation of the clinical significance of the CLS algorithm for different applications is appropriate. It may be possible to improve the quality of CLS by creating a hybrid technique with a color-altering algorithm. Finally, there may be applications of CLS outside the realm of medical imaging, e.g., surveillance.

Disclosures

The authors declare no conflicts of interest in relation to the work in this paper. John Pauly is a scientific advisor to Heart Vista, Inc.

Acknowledgments

N.D. would like to thank his students John Dunlop, Tanner Hallett, Jacob Miller, Henry Purdum, and Takayuki Yamakoshi for their early efforts in this project. The authors would like to thank Joseph Y. Cheng for providing the MR flow data. The authors would like to thank Joseph C. Liao for providing bladder cystoscopy data. The authors would like to thank Jennifer T. Smith for her expertise in bladder imaging. Finally, the authors would like to thank Daniel O'Connor for his advice regarding convex optimization. N.D. has been supported by the National Institute of Health's Grant No. NIH T32 EB009653 "Predoctoral Training in Biomedical Imaging at Stanford University," the National Institute of Health's Grant No. NIH T32 HL007846, The Rose Hills Foundation Graduate Engineering Fellowship, the Electrical Engineering Department New Projects Graduate Fellowship, and The Oswald G. Villard Jr. Engineering Fellowship. J.P. has been supported by NIH P41 EB015891.

References

1. G. Antoch and A. Bockisch, "Combined PET/MRI: a new dimension in whole-body oncology imaging?" *Eur. J. Nucl. Med. Mol. Imaging* **36**(1), 113–120 (2009).
2. R. Bar-Shalom et al., "Clinical performance of PET/CT in evaluation of cancer: additional value for diagnostic imaging and patient management," *J. Nucl. Med.* **44**(8), 1200–1209 (2003).
3. I. Mathworks, "Colormap help page for Matlab version R2016a," <http://www.mathworks.com/help/matlab/ref/colormap.html> (5 September 2016).
4. A. Drzezga et al., "First clinical experience with integrated whole-body PET/MR: comparison to PET/CT in patients with oncologic diagnoses," *J. Nucl. Med.* **53**(6), 845–855 (2012).
5. V. Ntziachristos et al., "Looking and listening to light: the evolution of whole-body photonic imaging," *Nat. Biotechnol.* **23**(3), 313–320 (2005).
6. N. C. Delioliannis et al., "Performance of the red-shifted fluorescent proteins in deep-tissue molecular imaging applications," *J. Biomed. Opt.* **13**(4), 044008 (2008).
7. F. Nooshabadi et al., "Intravital fluorescence excitation in whole-animal optical imaging," *PLoS One* **11**(2), e0149932 (2016).
8. F. Nooshabadi et al., "Intravital excitation increases detection sensitivity for pulmonary tuberculosis by whole-body imaging with β -lactamase reporter enzyme fluorescence," *J. Biophotonics* (2016).
9. J. A. Witjes et al., "Clinical and cost effectiveness of hexaminolevulinate-guided blue-light cystoscopy: evidence review and updated expert recommendations," *Eur. Urol.* **66**(5), 863–871 (2014).
10. G. T. Smith et al., "Multimodal 3D cancer-mimicking optical phantom," *Biomed. Opt. Express* **7**(2), 648–662 (2016).
11. M. I. Smith and J. P. Heather, "A review of image fusion technology in 2005," *Proc. SPIE* **5782**, 29 (2005).
12. C. Pohl and J. L. Van Genderen, "Review article multisensor image fusion in remote sensing: concepts, methods and applications," *Int. J. Remote Sens.* **19**(5), 823–854 (1998).
13. B. Yang, Z.-L. Jing, and H.-T. Zhao, "Review of pixel-level image fusion," *J. Shanghai Jiaotong Univ. (Sci.)* **15**, 6–12 (2010).
14. T. Porter and T. Duff, "Compositing digital images," *ACM SIGGRAPH Comput. Graphics*, **18**, 253–259 (1984).
15. P. Baudisch and C. Gutwin, "Multiblending: displaying overlapping windows simultaneously without the drawbacks of alpha blending," in *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*, pp. 367–374, ACM (2004).
16. M. Choi, "A new intensity-hue-saturation fusion approach to image fusion with a tradeoff parameter," *IEEE Trans. Geosci. Remote Sens.* **44**(6), 1672–1682 (2006).
17. J. Nunez et al., "Multiresolution-based image fusion with additive wavelet decomposition," *IEEE Trans. Geosci. Remote Sens.* **37**(3), 1204–1211 (1999).
18. T.-M. Tu et al., "A fast intensity-hue-saturation fusion technique with spectral adjustment for IKONOS imagery," *IEEE Geosci. Remote Sens. Lett.* **1**(4), 309–312 (2004).
19. K. Amolins, Y. Zhang, and P. Dare, "Wavelet based image fusion techniques—an introduction, review and comparison," *ISPRS J. Photogramm. Remote Sens.* **62**(4), 249–263 (2007).
20. G. Pajares and J. M. De La Cruz, "A wavelet-based image fusion tutorial," *Pattern Recognit.* **37**(9), 1855–1872 (2004).
21. A. Garzelli, "A review of image fusion algorithms based on the super-resolution paradigm," *Remote Sens.* **8**(10), 797 (2016).
22. P. C. Hansen, J. G. Nagy, and D. P. O'leary, *Deblurring Images: Matrices, Spectra, and Filtering*, Vol. 4, SIAM, Basel, Switzerland (2006).
23. Z. Li and H. Leung, "Fusion of multispectral and panchromatic images using a restoration-based method," *IEEE Trans. Geosci. Remote Sens.* **47**(5), 1482–1491 (2009).
24. S. Li and B. Yang, "A new pan-sharpening method using a compressed sensing technique," *IEEE Trans. Geosci. Remote Sens.* **49**(2), 738–746 (2011).
25. S. Li, H. Yin, and L. Fang, "Remote sensing image fusion via sparse representations over learned dictionaries," *IEEE Trans. Geosci. Remote Sens.* **51**(9), 4779–4789 (2013).
26. T. Goldstein and S. Osher, "The split Bregman method for l_1 -regularized problems," *SIAM J. Imaging Sci.* **2**(2), 323–343 (2009).
27. M. Nikolova, *Optimization, Applications to Image Processing*, Citeseer, Philadelphia, Pennsylvania (2014).
28. S. Setzer, "Operator splittings, Bregman methods and frame shrinkage in image processing," *Int. J. Comput. Vision* **92**(3), 265–280 (2011).
29. S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press (2004).
30. N. Parikh and S. P. Boyd, "Proximal algorithms," *Found. Trends Optim.* **1**(3), 127–239 (2014).
31. L. N. Trefethen and D. Bau III, *Numerical Linear Algebra*, Vol. 50, SIAM, Boston, Massachusetts (1997).
32. W. H. Press et al., *Numerical Recipes in C*, Vol. 2, Cambridge University Press, Philadelphia, Pennsylvania (1996).
33. M. S. Judenhofer et al., "Simultaneous PET-MRI: a new approach for functional and morphological imaging," *Nat. Med.* **14**(4), 459–465 (2008).
34. A. Martinez-Möller et al., "Tissue classification as a potential approach for attenuation correction in whole-body PET/MRI: evaluation with PET/CT data," *J. Nucl. Med.* **50**(4), 520–526 (2009).
35. J. Y. Cheng et al., "Comprehensive motion-compensated highly accelerated 4D flow MRI with ferumoxytol enhancement for pediatric congenital heart disease," *J. Magn. Reson. Imaging* **43**(6), 1355–1368 (2016).
36. C. Harris and M. Stephens, "A combined corner and edge detector," in *Alvey Vision Conference*, Vol. 15, p. 50, Citeseer (1988).
37. J. Lewis, "Fast normalized cross-correlation," in *Vision Interface*, Vol. 10, pp. 120–123 (1995).

38. R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed., Cambridge University Press, Cambridge (2003).
39. L. Alparone et al., "A global quality measurement of pan-sharpened multispectral imagery," *IEEE Geosci. Remote Sens. Lett.* **1**(4), 313–317 (2004).
40. N. Dwork, "CLS fusion Git repository," www.github.com/ndwork/CLS_Fusion (16 February 2017).

Nicholas Dwork is a PhD candidate of electrical engineering at Stanford University; his advisor is John Pauly. He received his BS degree in electrical engineering from California Polytechnic State University, San Luis Obispo, in 2001 and he received his MS degree in biomedical engineering from the University of California, Los Angeles, in 2004. His current research interests include reconstruction and quantification in medical imaging.

Eric M. Lasry was a visiting student at Stanford's Pre-Collegiate Summer Institute in 2016. He is currently a senior attending École Jeannine Manuel.

John M. Pauly is the Reid Weaver Dennis Professor of Engineering at Stanford University. He is the codirector of the Magnetic Resonance Systems Research Laboratory, which is devoted to MR instrumentation research. He has extensive experience in the development of MRI technology, including radio-frequency pulse design, pulse sequence design, and image reconstruction algorithms. He is currently pursuing research in device safety in MRI, developing wireless flexible receive arrays, and MR-guided minimally invasive RF ablation.

Jorge Balbás is an associate professor at California State University, Northridge, and an associate director of the Institute of Pure and Applied Mathematics, University of California, Los Angeles. His research focuses primarily on the analysis and implementation of high-resolution, nonoscillatory numerical schemes for hyperbolic system of conservation laws and related problems. In particular, he is interested in the computational and mathematical challenges posed by the equations of ideal magnetohydrodynamics, a hyperbolic model that describes the dynamics of electrically conducting fluids in the presence of a magnetic field.