

# Nonnegative Mixed-Norm Preconditioning for Microscopy Image Segmentation

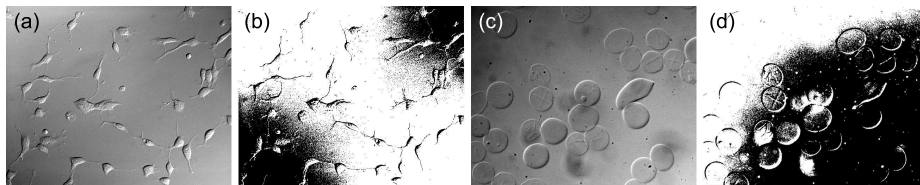
Kang Li and Takeo Kanade

Carnegie Mellon University  
5000 Forbes Ave, Pittsburgh, PA 15213, USA  
{kangli, tk}@cs.cmu.edu

**Abstract.** Image segmentation in microscopy, especially in interference-based optical microscopy modalities, is notoriously challenging due to inherent optical artifacts. We propose a general algebraic framework for *preconditioning* microscopy images. It transforms an image that is unsuitable for direct analysis into an image that can be effortlessly segmented using global thresholding. We formulate preconditioning as the minimization of nonnegative-constrained convex objective functions with smoothness and sparseness-promoting regularization. We propose efficient numerical algorithms for optimizing the objective functions. The algorithms were extensively validated on simulated differential interference (DIC) microscopy images and challenging real DIC images of cell populations. With preconditioning, we achieved unprecedented segmentation accuracy of 97.9% for CNS stem cells, and 93.4% for human red blood cells in challenging images.

## 1 Introduction

Microscopy image segmentation lays the foundation for shape analysis, motion tracking, and classification of biological objects. Despite its importance, automated segmentation remains challenging for several widely used non-fluorescence, interference-based microscopy imaging modalities, such as phase contrast microscopy and differential interference contrast (DIC) microscopy. These modalities employ interference optics to convert phase shifts of light induced by otherwise transparent objects into visible intensity variations. While being visually



**Fig. 1.** DIC microscopy images of unstained cell populations. (a) CNS stem cells. (c) Human red blood cells. (b) and (d) are the thresholding outputs of (a) and (c), respectively, showing nonuniform background and unsymmetrical pixel intensity distribution in cells.

contrastive, images generated by these modalities are often unsuitable for direct computer segmentation or measurement, owing to inherent optical artifacts.

The noninvasiveness of interference-based microscopy makes them uniquely suitable for long-term monitoring of living biological specimens. Recent advances in biology intensified the interest in using them for quantitative measurement of ongoing biological processes, e.g., for recognition and tracking of stem cell behaviors. These emerging applications call for novel algorithms that facilitate automated segmentation of interference-based optical microscopy images.

We propose in this paper a general algebraic framework for *preconditioning* interference-based microscopy images, which dramatically facilitates automated segmentation and analysis. Based on physical principles of microscopy image formation, we formulate preconditioning as the minimization of nonnegative-constrained convex objective functions with smoothness and sparseness-enhancing regularization. Taking advantage of recent advances in convex optimization [1, 2, 3], we propose efficient algorithms to solve the minimization problems.

The effectiveness of the proposed algorithms will be demonstrated for preconditioning DIC microscopy images. The dual-beam interference optics of DIC microscopes introduces nonuniform shadow-cast artifacts (Fig. 1), making direct segmentation notoriously difficult. Existing techniques for DIC image segmentation either are application-specific, relying on template matching and edge detection [4], or require *ad hoc* image preprocessing, e.g., line integration [5] or the Hilbert transform [6]. Attempts have also been made to derive exact imaging models, resulting in computationally expensive algorithms that are impractical for routine utilization [7, 8]. In contrast, our preconditioning algorithms efficiently reconstruct images according to well-defined optimality criteria, enabling high-quality segmentation using global thresholding. Moreover, our approach can be easily extended to other modalities (e.g., phase contrast microscopy) by incorporating appropriate imaging models, or to higher dimensions.

## 2 Algebraic Image Model

Our generic model for microscopy images consists of three components: 1) an *imaging* model  $h(\cdot)$  that represents the image formation process of the microscope; 2) an additive *bias*  $b(x, y)$  that compensates for a nonzero background level, nonuniform illumination, and spatial sensitivity variations of the detector; and 3) a *noise* model  $n(\cdot)$  that accounts for imaging and detection noise. The model can be written as:

$$g(x, y) = n(h(f(x, y)) + b(x, y)), \quad (1)$$

with  $g(x, y)$  being the *observed* image, and  $f(x, y)$  being the ideal *object* image that we want to retrieve, which could represent the optical path length distribution in the object, fluorescence intensities, or an *phenomenological* image that simply facilitates object segmentation.

Under the assumption of additive noise, we can express our model succinctly in a linear algebraic framework, given by

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{b} + \mathbf{n}. \quad (2)$$

Here  $\mathbf{g}$  denotes a vector representation of the observed image, which is formed by concatenating the image pixels in raster order. Specifically, given an image with  $N = N_x \times N_y$  pixels  $\{g_{x,y}\}$ <sup>1</sup>, where  $x = 1, \dots, N_x$  and  $y = 1, \dots, N_y$ , the corresponding vector  $\mathbf{g}$  is defined as:

$$\mathbf{g} = (g_1, \dots, g_i, \dots, g_N)^T = (g_{x_1, y_1}, \dots, g_{x_i, y_i}, \dots, g_{x_N, y_N})^T, \quad (3)$$

where  $(\cdot)^T$  denotes vector (or matrix) transposition, and  $(x_i, y_i)$  are the spatial coordinates of pixel  $g_{x_i, y_i}$ , or equivalently, of the  $i$ th element of vector  $\mathbf{g}$ . The vectors  $\mathbf{f}$ ,  $\mathbf{b}$ ,  $\mathbf{n}$  are defined likewise for the object, bias, and noise, respectively. The imaging model is expressed as a matrix-vector multiplication between the  $N \times N$  transfer matrix  $\mathbf{H}$  and  $\mathbf{f}$ , which is adequate for representing a wide range of microscopy image formation processes. Since  $\mathbf{H}$  can naturally represent shift-variant transfer functions, our expression is more flexible than the conventional convolution formulation in terms of a point spread function (PSF).

In the next section, we will elaborate on the preconditioning framework under the assumption of additive Gaussian noise. We will return to the discussion of alternative noise models in Section 6.

### 3 Nonnegative Mixed-Norm Preconditioning

With the image model specified in Eq. (2), we need to compute the ideal object image  $\mathbf{f}$  given an observed image  $\mathbf{g}$ . We tackle this inverse problem through a two-step process. First, we estimate and subsequently eliminate the bias from an observed image. Second, we reconstruct the object image  $\mathbf{f}$  from the bias-corrected image by minimizing a constrained mix-norm objective function. Collectively, we refer to this two-step process as *preconditioning*, which transforms an observed image that is unfriendly for computer analysis into an image that facilitates automated object segmentation and measurement.

#### 3.1 Bias Elimination

As the first step of preconditioning, we estimate the bias field from an image and obtain a bias-corrected image. This process is also known as flat-field correction or background subtraction. While many methods exist for this purpose, we present a simple approach that is sufficient for most microscopy images.

Under the assumption that the bias field is smooth and spatially slowly varying, we model it as a  $K$ -th order polynomial surface:

$$b(x, y) = \sum_{k=0}^K \sum_{j=0}^k p_{j+\frac{k(k+1)}{2}} x^{k-j} y^j = p_0 + p_1x + p_2y + p_3x^2 + p_4xy + p_5y^2 + \dots \quad (4)$$

<sup>1</sup>  $g_{x,y}$  denotes the discrete value sampled from the continuous function  $g(x, y)$  at  $(x, y)$ .

Algebraically, we can express the polynomial as  $\mathbf{b} = \mathbf{X}\mathbf{p}$ , where  $\mathbf{p} = (p_0, p_1, p_2 \dots)^T$  is the coefficient vector, and  $\mathbf{X}$  is a matrix of  $N$  rows and  $(K+1)(K+2)/2$  columns with the  $i$ -th row being  $(1, x_i, y_i, x_i^2, x_i y_i, y_i^2, \dots)$ . To eliminate the bias, we first compute the optimal coefficients by solving the over-determined linear system  $\mathbf{X}\mathbf{p} = \mathbf{g}$ . This amounts to solving the least-squares problem  $\mathbf{p}^* = \arg \min_{\mathbf{p}} \|\mathbf{X}\mathbf{p} - \mathbf{g}\|_2^2$ , which has a closed-form solution  $\mathbf{p}^* = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{g}$ . The bias-corrected image is then computed as  $\mathbf{g}^* = \mathbf{g} - \mathbf{X}\mathbf{p}^*$ .

### 3.2 Object Reconstruction via Convex Optimization

The second step of preconditioning reconstructs the object from the bias-corrected image. Ideally, the background pixels of the reconstructed image should be uniformly zero while the foreground pixels are positive, facilitating foreground-background separation. This goal is achieved by minimizing:

$$\mathbf{O}(\mathbf{f}) = \|\mathbf{g}^* - \mathbf{H}\mathbf{f}\|_2^2 + \gamma \text{Smoothness}(\mathbf{f}) + \beta \text{Sparsity}(\mathbf{f}), \quad \text{subject to } \mathbf{f} \geq 0, \quad (5)$$

where  $\|\cdot\|_2$  denotes a  $L_2$  norm.

The objective function  $\mathbf{O}(\mathbf{f})$  consists of three terms. Their relative importance is controlled by the positive coefficients  $\gamma$  and  $\beta$ . The first term penalizes the sum-of-squares difference between the reconstructed and observed images, promoting data *fidelity*. The second and third terms encourage the spatial *smoothness* and *sparseness* of the reconstructed image, respectively, which collectively provide *regularization*. Regularization ensures the well-posedness of the objective function, and is essential for high-quality reconstruction.

Next, we will introduce two specific formulations of the objective function and discuss the corresponding optimization algorithms.

**Case 1:  $L_2$  Smoothness + Weighted  $L_1$  Sparsity** The first formulation employs an  $L_2$  (Tikhonov) smoothness term and a weighted  $L_1$  sparseness term:

$$\mathbf{O}_1(\mathbf{f}) = \|\mathbf{g}^* - \mathbf{H}\mathbf{f}\|_2^2 + \gamma \|\mathbf{R}\mathbf{f}\|_2^2 + \beta \|\mathbf{W}\mathbf{f}\|_1, \quad \text{s.t. } \mathbf{f} \geq 0. \quad (6)$$

The smoothness term penalizes the  $L_2$  norm of the Laplacian of  $\mathbf{f}$ , where the  $N \times N$  matrix  $\mathbf{R}$  represents an algebraic Laplacian operator with symmetric boundary condition. In particular, the  $i$ -th element of  $\mathbf{R}\mathbf{f}$  is computed as:

$$(\mathbf{R}\mathbf{f})_i = f_{x_i, y_i} - \sum_{j, k \in \{-1, 1\}} \frac{f_{x_i+j, y_i+k}}{8}, \quad (7)$$

with  $x_i + j = x_i - j$  if  $x_i + j < 1$  or  $> N_x$ , and likewise for  $y_i$ . The sparseness term penalizes the weighted  $L_1$  norm of  $\mathbf{f}$ , where  $\mathbf{W}$  is a diagonal matrix with positive weights  $w_1, \dots, w_N$  on the diagonal and zeros elsewhere.

By rewriting  $\mathbf{O}_1(\mathbf{f})$  in terms of the symmetric positive definite matrix  $\mathbf{Q} = \mathbf{H}^T \mathbf{H} + \gamma \mathbf{R}^T \mathbf{R}$  and the vector  $\mathbf{l} = -\mathbf{H}^T \mathbf{g}^*$ , and letting  $\mathbf{w}$  denote the weight

vector  $(w_1, \dots, w_N)^T$ , we can express the minimization problem as the following nonnegative-constrained quadratic program (NQP):

$$\mathbf{f}^* = \arg \min_{\mathbf{f}} \frac{1}{2} \mathbf{f}^T \mathbf{Q} \mathbf{f} + (\mathbf{l} + \beta \mathbf{w})^T \mathbf{f}, \quad \text{s.t. } \mathbf{f} \geq 0. \quad (8)$$

We propose a simple and efficient iterative algorithm, the *sparseness-enhanced multiplicative update* (**SEMU**) algorithm, which is well-tailored for the NQP. Our algorithm exploits a nonnegativity-preserving multiplicative update rule [2] and the sparseness-enhancing effect of an iteratively reweighted  $L_1$  scheme. To introduce the algorithm, we express the matrix  $\mathbf{Q}$  in Eq. (8) in terms of its positive and negative components. Specifically, we let  $\mathbf{Q} = \mathbf{Q}^+ - \mathbf{Q}^-$ , where  $\mathbf{Q}^+$  and  $\mathbf{Q}^-$  are *nonnegative* matrices given by:

$$Q_{i,j}^+ = \begin{cases} Q_{i,j}, & \text{if } Q_{i,j} > 0, \\ 0, & \text{otherwise,} \end{cases} \quad \text{and} \quad Q_{i,j}^- = \begin{cases} |Q_{i,j}|, & \text{if } Q_{i,j} < 0, \\ 0, & \text{otherwise.} \end{cases}$$

The algorithm, defined in terms of  $\mathbf{Q}^+$  and  $\mathbf{Q}^-$ , alternates between updating  $\mathbf{f}$  and (optionally) recomputing  $\mathbf{w}$ , as outlined below:

**Initialize** Set iteration number  $t = 0$ , and  $f_i^{(0)} = 1$ ,  $w_i^{(0)} = 1$ ,  $\forall i \in \{1, \dots, N\}$ .  
**Repeat** Update  $\mathbf{f}$  and  $\mathbf{w}$  alternately according to:

$$f_i^{(t+1)} \leftarrow \left[ \frac{-(l_i + \beta w_i^{(t)}) + \sqrt{(l_i + \beta w_i^{(t)})^2 + 4(\mathbf{Q}^+ \mathbf{f}^{(t)})_i (\mathbf{Q}^- \mathbf{f}^{(t)})_i}}{2(\mathbf{Q}^+ \mathbf{f}^{(t)})_i} \right] f_i^{(t)}, \quad (9)$$

$$\text{(optional)} \quad w_i^{(t+1)} \leftarrow \frac{1}{f_i^{(t+1)} + \alpha}, \quad \text{where } \alpha \text{ is a positive constant,} \quad (10)$$

**Until**  $\|\mathbf{f}^{(t+1)} - \mathbf{f}^{(t)}\|_2^2 \leq \epsilon$ , where  $\epsilon$  is a small positive constant.

The multiplicative updates as given by Equations (9) and (10) are applied independently to each pixel, and can be trivially parallelized.

To understand the algorithm, we first exclude Eq. (10) and consider  $\mathbf{w}$  being constant at unity. In this case, the sparseness term in  $\mathbf{O}_1(\mathbf{f})$  is reduced to an ordinary  $L_1$  norm. If we define  $a_i = (\mathbf{Q}^+ \mathbf{f})_i$ ,  $b_i = l_i + \beta w_i$ , and  $c_i = (\mathbf{Q}^- \mathbf{f})_i$ , it is instantly recognizable that the multiplicative factor in Eq. (9) is the larger root of the quadratic equation  $a_i \nu_i^2 + b_i \nu_i - c_i = 0$ . This factor, which we denote as  $\nu_i^+$ , is guaranteed to be real and nonnegative as long as  $\mathbf{f}$  is nonnegative, which in return preserves the nonnegativity of  $\mathbf{f}$ . Moreover, using the fact that  $\partial \mathbf{O}_1 / \partial f_i = a_i + b_i - c_i$ , we can verify that the update rule have fixed points at  $f_i = 0$  and  $\partial \mathbf{O}_1 / \partial f_i = 0$  (i.e.,  $\nu_i^+ = 1$ ), which are consistent with the Karush-Kuhn-Tucker (KKT) condition for the NQP. While a rigorous convergence proof is nontrivial, it is evident that  $\partial \mathbf{O}_1 / \partial f_i < 0$  implies  $\nu_i^+ > 1$ , and  $\partial \mathbf{O}_1 / \partial f_i > 0$  implies  $\nu_i^+ < 1$ . Hence, the updates increase or decrease each  $f_i$  along the opposite direction of its partial derivative. Thanks to the convexity of  $\mathbf{O}_1(\mathbf{f})$ , the iterations convergence monotonically to the unique global minimum [2].

If we iteratively reweight  $\mathbf{w}$  according to Eq. (10), the  $L_1$  sparseness term in  $\mathbf{O}_1(\mathbf{f})$  is in effect replaced with a *log-sum* penalty function  $\sum_i \log(f_i + \alpha)$  [3]. The log-sum penalty is much more sparseness-encouraging than the  $L_1$  norm, because it charges increasingly larger penalties for smaller nonzeros. Despite the change,  $f_i = 0$  and  $\partial \mathbf{O}_1 / \partial f_i = 0$  remain as fixed points of the updates, and the monotonic convergence of the algorithm is not violated. However, since the log-sum penalty function is concave, we cannot expect the algorithm to always achieve a global minimum. Nevertheless, achieving global optimality is immaterial for preconditioning. In practice, iterative reweighting encourages background pixels to quickly converge to zero, achieving faster and superior reconstruction.

**Case 2: Total Variation Smoothness +  $L_1$  Sparsity** As an alternative to the  $L_2$  smoothness term in Eq. (6), total variation (TV) regularization [9] can be incorporated into the objective function:

$$\mathbf{O}_2(\mathbf{f}) = \|\mathbf{g}^* - \mathbf{H}\mathbf{f}\|_2^2 + \gamma \|\mathbf{f}\|_{\text{TV}} + \beta \|\mathbf{f}\|_1, \quad \text{subject to } \mathbf{f} \geq 0. \quad (11)$$

The TV norm, denoted as  $\|\cdot\|_{\text{TV}}$  in Eq. (11), is defined as:

$$\|\mathbf{f}\|_{\text{TV}} = \sum_{i=1}^N \sqrt{(\mathbf{D}_x \mathbf{f})_i^2 + (\mathbf{D}_y \mathbf{f})_i^2}, \quad (12)$$

where  $\mathbf{D}_x$  and  $\mathbf{D}_y$  are matrix representations of the forward-difference operators with homogeneous Neumann boundary conditions, i.e.,

$$(\mathbf{D}_x \mathbf{f})_i = f_{x_i+1, y_i} - f_{x_i, y_i}, \quad \text{and} \quad (\mathbf{D}_y \mathbf{f})_i = f_{x_i, y_i+1} - f_{x_i, y_i}, \quad (13)$$

with the exception that  $(\mathbf{D}_x \mathbf{f})_i = 0$  if  $x_i \geq N_x$ , and  $(\mathbf{D}_y \mathbf{f})_i = 0$  if  $y_i \geq N_y$ . The TV norm is essentially the  $L_1$  norm of the image gradient, which promotes the sparseness of gradients in the reconstructed image. It is well-known for its advantage in preserving discontinuities (i.e., sharp edges) in an image.

The minimization of  $\mathbf{O}_2(\mathbf{f})$  can be solved as a second-order cone program (SOCP) [10]:

$$\mathbf{f}^* = \arg \min_{\mathbf{t}, \mathbf{f}} (\mathbf{1}^T \mathbf{t} + \beta \mathbf{1}^T \mathbf{f}), \quad (14)$$

$$\text{s.t. } \mathbf{D}_x \mathbf{f} - \mathbf{u} = \mathbf{0}, \quad \mathbf{D}_y \mathbf{f} - \mathbf{v} = \mathbf{0}, \quad \mathbf{H}\mathbf{f} + \mathbf{w} = \mathbf{g}^*, \quad w_0 = \tau,$$

$$\text{and } \sqrt{u_i^2 + v_i^2} \leq t_i, \quad \sqrt{w_1^2 + \dots + w_N^2} \leq w_0, \quad f_i \geq 0, \quad t_i \geq 0, \quad \forall i = 1, \dots, N,$$

where  $\mathbf{t}$ ,  $\mathbf{u}$ ,  $\mathbf{v}$ ,  $\mathbf{w}$  and  $w_0$  are slack variables;  $\mathbf{1}$  and  $\mathbf{0}$  denote  $N$ -dimensional vectors of ones and zeros, respectively; and  $\tau$  is an adjustable noise-tolerance parameter. The SOCP as given in (14) can be solved robustly using interior-point methods (*a.k.a.* barrier methods) with general-purpose convex optimization software [10, 11]. We will refer to this approach as “**TV-SOCP**”. It alleviates the slow convergence and instability issues that plague traditional gradient-based algorithms [9], and allows easy incorporation of additional constraints.

## 4 Differential Interference Contrast Microscopy

DIC microscopy is widely used to provide contrast of unstained, transparent specimens, such as cells and microorganisms. Such specimens, known as *phase objects*, cause no detectable amplitude change to the light that passes through them. They are essentially invisible under an ordinary transmitted-light microscope. However, they diffract light due to the difference of their refractive indices with respect to the surrounding media, causing phase shifts to the incident light waves. DIC microscopy employs dual-beam interference optics to transform the invisible phase shifts into intensity variations in the observed image.

### 4.1 Physics of DIC Image Formation

The DIC image formation process can be summarized by the following equations. First, consider an illuminating light beam with amplitude  $A$  that is entering a DIC microscope parallel to its optical axis. Upon transiting through a polarizer, the wavefront of the light becomes coherent and plane-polarized, represented by:

$$u_0(\mathbf{x}) = A \exp(-if_0(\mathbf{x})). \quad (15)$$

Here  $\mathbf{x} = (x, y)$ , and  $f_0(\mathbf{x})$  represents the phase of the light field. A condenser prism splits the polarized light into two mutually coherent beams, spatially separated by a minute shear  $\mathbf{s} = (s \cos(\theta), s \sin(\theta))$ , where  $s$  is the *shear distance* and  $\theta$  is the *shear angle*. The two beams pass through the condenser and interact with the object, resulting in an altered phase  $f(\mathbf{x})$  containing object information:

$$u_1(\mathbf{x}) = A \exp(-if(\mathbf{x})), \quad (16)$$

$$u_2(\mathbf{x}) = A \exp(-i(f(\mathbf{x} + \mathbf{s}) + f_b)). \quad (17)$$

In the above equations,  $f_b$  is a constant relative phase shift between the two wavefronts, known as the *bias retardation*. After being focused by the objective lens, the two beams are recombined into a single beam by an objective prism in the objective back aperture. Finally, another polarizer, called the analyzer, transmits plane-polarized light from the objective that is able to interfere and generate an image  $g(\mathbf{x})$ . For ideal imaging,  $g(\mathbf{x})$  can be expressed as:

$$g(\mathbf{x}) = 4A^2(1 - \cos(D_\theta f(\mathbf{x}) + f_b)), \quad (18)$$

where  $D_\theta f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{s} \approx f(\mathbf{x} + \mathbf{s}) - f(\mathbf{x})$  is the first derivative of phase shift taken along the shear axis. Therefore, the intensity of a DIC image is a nonlinear function of the phase shift introduced by the object.

**Linear Approximation** The phase shift introduced by an object is related to its optical path length distribution, which is the product of refractive index and thickness distributions of the object relative to the surrounding medium. Since most biological specimens have refractive indices that differ little from that of

water (or the medium in which they are immersed), they introduce sufficiently small phase shift differences. If a microscope is properly adjusted such that  $f_b = 90^\circ$ , the image that it produces can be approximated by

$$g(\mathbf{x}) \approx 4A^2(1 + D_\theta f(\mathbf{x})), \quad (19)$$

which is a linear function of the phase shift derivative.

## 4.2 Algebraic Imaging Model

Based on the previous derivation, we define the DIC imaging model as the first derivative of the phase shift distribution  $f(x, y)$  along the shear angle  $\theta$ . Specifically, we define  $h(f(x, y)) = D_\theta f(x, y) \approx f(x, y) * \text{epsf}(x, y)$ , where  $\text{epsf}(x, y)$  is an *effective* PSF, defined as a steerable first-derivative-of-Gaussian kernel:

$$\text{epsf}(x, y) \propto -xe^{-\frac{x^2+y^2}{\sigma^2}} \cos(\theta) - ye^{-\frac{x^2+y^2}{\sigma^2}} \sin(\theta). \quad (20)$$

Rather than assuming fixed shear angles [5, 6], our definition can be adapted to arbitrary angles according to particular microscopes. In practice,  $\text{epsf}(x, y)$  is discretized as an  $M \times M$  matrix. We define the transfer matrix  $\mathbf{H}$  such that

$$(\mathbf{H}\mathbf{f})_i = \sum_{j=1}^M \sum_{k=1}^M \text{epsf}(i - \frac{M}{2}, j - \frac{M}{2}) f_{x_i+j-\frac{M}{2}, y_i+k-\frac{M}{2}}. \quad (21)$$

## 5 Experiments

The proposed algorithms were implemented in ISO C++. Experiments were carried out on a computer with a 2.53GHz Intel®Core™2 Duo processor and 8 gigabytes of memory, running 64-bit Ubuntu Linux. The large-scale optimization library MOSEK [11] was utilized as an SOCP solver.

### 5.1 Data

Our method has been validated on three different types of data:

1. *Computer-simulated DIC images*, including 12 simulated images of block structures with sizes ranging from  $16^2$  to  $1024^2$ . The images were generated according to the model given in Section 4.2, and were corrupted with randomly-generated nonuniform bias and additive Gaussian noise.
2. *Real DIC images of CNS stem cell populations*, consisting of a sequence of 1795 images of central nervous system (CNS) stem cells. The images are  $640 \times 512$  pixels each, captured every 5 minutes using a 12-bit Orca ER (Hamamatsu) CCD camera mounted on a Zeiss Axiovert 135TV microscope with a  $40\times$ , 1.3 NA oil-immersion DIC objective. 51 images were manually segmented by an expert biologist, and utilized as the ground truth.
3. *Real DIC images of human red blood cells*, including three extremely challenging images from the Broad bioimage benchmark collection<sup>2</sup>, for which ground-truth segmentations are available. The images are of size  $800 \times 600$ .

<sup>2</sup> Available online at <http://www.broad.mit.edu/bbbc>.



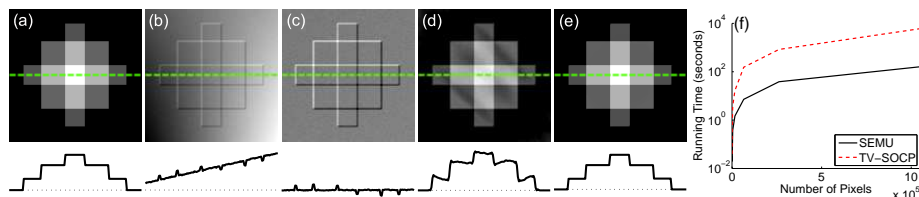
## 5.2 Parameter Settings

Our algorithms involve up to seven parameters. The common ones include the bias order  $K$ , sparseness coefficient  $\beta$ , DIC shear angle  $\theta$ , and the DIC kernel standard deviation  $\sigma$  (Eq. 20). Parameters specific for SEMU include the smoothness coefficient  $\gamma$ , the reweighting coefficient  $\alpha$ , and the stopping threshold  $\epsilon$ . TV-SOCP requires the noise factor  $\tau$  instead of  $\gamma$ .

We set  $K = 2$ ,  $\sigma = 1$ , and  $\epsilon = 10^{-6}$  for all experiments. The shear angle  $\theta$  is defined to be zero along the three-o'clock direction, increasing clockwise. It is usually multiples of  $45^\circ$ , and can be easily estimated by observation. Specifically, we set  $\theta = 225^\circ$  for the simulated DIC images and the images of red blood cells, and we set  $\theta = 45^\circ$  for the images of CNS stem cells. The parameters  $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\tau$  were manually adjusted for the simulated images. For each dataset of cell populations, the optimal parameters were determined using grid search based on the receiver operating characteristic (ROC) of one image. The specific values will be reported in the next section along with the results.

## 5.3 Results

Fig. 2 shows the preconditioning results for one of the simulated DIC images. The pixel values of the input images was normalized to the range of  $[0, 1]$ , and was superposed with severe bias and zero-mean Gaussian noise of standard deviation  $10^{-4}$ . The figure shows that the image was effectively corrected for bias, and the object was reconstructed with high fidelity. In particular, the TV-SOCP algorithm, which is well-suited for regularizing “blocky” structures, reconstructed the object near perfectly. The overall mean squared error (MSE) of the reconstructed images in the dataset using SEMU and TV-SOCP are  $1.95 \times 10^{-3} \pm 1.02 \times 10^{-4} \text{SD}^3$  and  $4.48 \times 10^{-5} \pm 3.40 \times 10^{-6} \text{SD}$ , respectively. The running times of both algorithms are roughly linear with respect to the number of pixels, as depicted in Fig. 2(f).



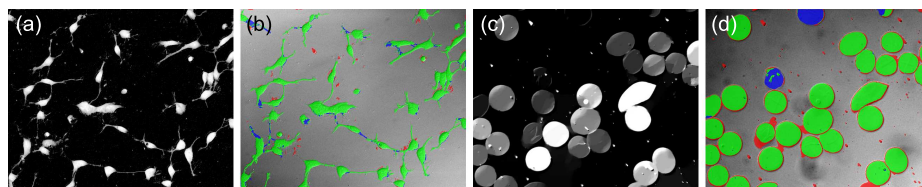
**Fig. 2.** Preconditioning of simulated DIC images. (a) Ground truth. (b) Observed image. (c) Bias-corrected image. (d) Preconditioned image with SEMU. (e) Preconditioned image with TV-SOCP. (f) Running time of the preconditioning algorithms in logarithmic scale.

Next, we present the results for the two datasets of living cell populations. Each of the datasets possesses unique challenges. The CNS stems cells have long,

<sup>3</sup> SD stands for (one) standard deviation.

thin processes that extend from the cell body, which sometimes interconnect different cells (see Fig. 1(a)). Segmenting the processes is of significant biological value, and yet extremely challenging. The images of the red blood cells are of very low contrast, and contain various visual contaminations (see Fig. 1(c)).

Despite the various artifacts in the DIC images, the preconditioned images became dramatically clean-cut. In particular, Fig. 3(a) shows an image of CNS stem cells corresponding to Fig. 1(a), which was preconditioned using the SEMU algorithm with parameters  $\alpha = 0.2$ ,  $\beta = 0.0005$ , and  $\gamma = 0.5$ . The pixel intensities are displayed in decibel scale to increase the visibility of thin structures. The average processing time was 32.5 seconds/image. Preconditioning preserved the fine details in the original images, especially the thin processes. Fig. 3(c) shows an image of red blood cells corresponding to Fig. 1(c), which was preconditioned using TV-SOCP with  $\beta = 0.05$  and  $\tau = 9$ . The average processing time was 102.9 minutes/image. Preconditioning removed the nonuniformity in the original image, while preserving details of the internal structures of cells.



**Fig. 3.** Preconditioned images of cell populations. (a) CNS stem cells (brightness in decibel scale). (c) Human red blood cells. (b) and (d) are color-coded thresholding outputs of (a) and (c). Green stands for true positive, red for false positive, and blue for false negative.

The preconditioned images enabled us to achieve high-quality cell segmentation using global thresholding. To evaluate the performance of preconditioning and not the separate problem of thresholding, we plotted the ROC curves for each dataset in Fig. 4 by trying every possible threshold on the preconditioned images. We use “positive” (or  $P$ ) to refer to the set of pixels belonging to cells according to the ground truth, and “negative” (or  $N$ ) for the background pixels. For pixels that are labeled as cells by thresholding, the subset that indeed belongs to cells is *true positive* (or  $TP$ ), and the remaining subset is *false positive* (or  $FP$ ). True positive rate and false positive rate are defined, respectively, as  $TPR = |TP|/|P|$ , and  $FPR = |FP|/|N|$ . In addition, we define *accuracy* as  $ACC = (|TP| + |TN|)/(|P| + |N|)$ , where  $TN = N - FP$  is the *true negative*.

As shown in Fig. 4(a), both SEMU and TV-SOCP achieved excellent performances for preconditioning the images of CNS stem cells. The areas under the ROC curves (AUC) for SEMU and TV-SOCP are 96.3% and 93.6%, respectively. By visual inspection, we found that TV-SOCP was less capable in preserving long, thin processes of CNS stem cells. It was also slightly more sensitive to parameter settings, i.e., the optimal settings vary according to cell density. Since we used identical parameters to process all images, the performance was sub-

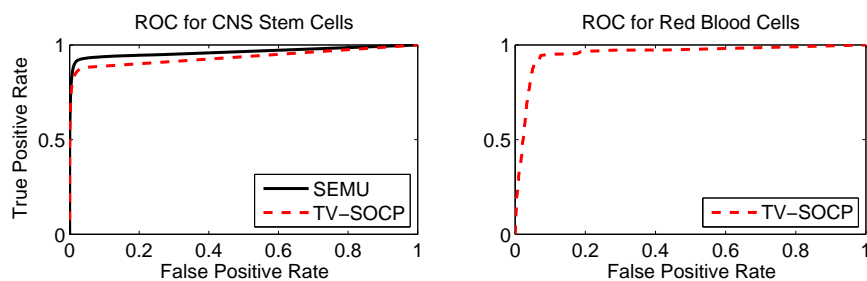


Fig. 4. ROC curves for the proposed preconditioning algorithms.

optimal for some. Contrarily, SEMU was relatively insensitive to cell density changes, which contributed to its superior performance. Excellent performance was also achieved for images of red blood cells, for which we tested TV-SOCP only. The ROC curve is shown in Fig. 4(b), and the corresponding AUC is 95.4%.

Finally, we segmented each image using the threshold value that maximizes the segmentation accuracy for that image. The results that correspond to the images shown in Fig. 1 are presented in Fig. 3 (b) and (d). Overall, we achieved a 97.9% accuracy for segmenting CNS stem cells, and 93.4% for red blood cells.

## 6 Discussions and Conclusion

**Alternative Noise Models** The long-tailed *Laplacian* distribution is often used for modeling impulsive noise. This model can be imposed by replacing the relevant  $L_2$  norms in the objective functions by  $L_1$  norms. Another important noise model is the *Poisson* model, which is non-additive, and is well-suited for describing quantum-limited noise. The Poisson noise model can be imposed by using Csiszár's I-divergence [12] as the data fidelity measure. A well-known minimizer for the I-divergence is the Richardson-Lucy (RL) algorithm [13]. The algorithm is nonnegativity-preserving under the assumption that all elements of the transfer matrix and input images are nonnegative, which does not hold for DIC microscopy. In contrast, the SEMU and TV-SOCP algorithms preserve nonnegativity for arbitrary matrices and inputs.

**Conclusion** We proposed a general framework for preconditioning microscopy images for facilitating segmentation. The approach reconstructs the intensities of an image utilizing principles of microscopy image formation under realistic assumptions. The reconstruction is performed by minimizing convex objective functions with nonnegativity constraints. Two specific objective functions were proposed and corresponding efficient algorithms, namely SEMU and TV-SOCP, were presented. SEMU is well suited for images of small, smooth objects, while TV-SOCP is better for relatively large, flat, and blocky objects. The preconditioned images can be effortlessly segmented using thresholding. The algorithm

achieved excellent performance on both simulated and real DIC images. It can be straightforwardly extended to other imaging modalities or higher dimensions.

## Acknowledgements

We would like to thank Drs. Rea Ravin and Daniel J. Hoepfner (NINDS) for providing DIC images of CNS stem cells with ground truth. This work was supported partially by NIH grants R01 EB007369-01, R01 EB0004343-01, and a fellowship from Intel Research Pittsburgh. Special thanks also goes to Drs. Mei Chen, Lee E. Weiss, and Phil G. Campbell for their selfless support, and to Steven Kang for his help. This work would not have been accomplished without the computing resources provided freely by Intel Research.

## References

1. D. Goldfarb and W. Yin, "Second-order cone programming methods for total variation-based image restoration," *SIAM J. Sci. Comput.*, vol. 27, no. 2, pp. 622–645, Oct. 2005.
2. F. Sha, Y. Lin, L. K. Saul, and D. D. Lee, "Multiplicative updates for nonnegative quadratic programming," *Neural Computation*, vol. 19, pp. 2004–2031, 2007.
3. E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted  $\ell_1$  minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5, pp. 877–905, Dec. 2008.
4. D. Young, C. Glasbey, A. Gray, and N. Martin, "Identification and sizing of cells in microscope images by template matching and edge detection," in *Proc. IEEE Image Proc. App.*, Jul. 1995, pp. 266–270.
5. B. Heise, A. Sonnleitner, and E. P. Klement, "DIC image reconstruction on large cell scans," *Microsc. Res. Tech.*, vol. 66, pp. 312–320, Jul. 2005.
6. M. R. Arnison, C. J. Cogswell, N. I. Smith, P. W. Fekete, and K. G. Larkin, "Using the Hilbert transform for 3D visualization of differential interference contrast microscope images," *J. Microsc.*, vol. 199, no. 1, pp. 79–84, Jul. 2000.
7. C. Preza, D. L. Snyder, and J.-A. Conchello, "Theoretical development and experimental evaluation of imaging models for differential interference contrast microscopy," *J. Opt. Soc. Am. A*, vol. 16, no. 9, pp. 2185–2199, Sep. 1999.
8. F. Kagalwala and T. Kanade, "Reconstructing specimens using DIC microscope images," *IEEE Trans. Syst., Man, Cybern.*, vol. 33, no. 5, pp. 728–737, Oct. 2003.
9. L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, pp. 259–268, 1992.
10. S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, Mar. 2004. [Online]. Available: <http://www.stanford.edu/~boyd/cvxbook/>
11. "The MOSEK optimization software," Mosek ApS, Copenhagen, Denmark. [Online]. Available: <http://www.mosek.com/>
12. I. Csiszár, "Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems," *Ann. Stat.*, vol. 19, no. 4, pp. 2032–2066, Dec. 1991.
13. N. Dey, L. Blanc-Féraud, C. Zimmer, P. Roux, Z. Kam, J.-C. Olivo-Marin, and J. Zerubia, "Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution," *Microsc. Res. Tech.*, vol. 69, pp. 260–266, Apr. 2006.