Check for updates

Frequency-constrained robust principal component analysis: a sparse representations approach to segmentation of dynamic features in optical coherence tomography imaging

JAMES P. MCLEAN, YUYE LING, AND CHRISTINE P. HENDON[*]

Department of Electrical Engineering, Columbia University, 530 West 120th Street, New York, NY 10025, USA

[*]*cpf2115@columbia.edu*

Abstract: Sparse representation theory is an exciting area of research with recent applications in medical imaging and detection, segmentation, and quantitative analysis of biological processes. We present a variant on the robust-principal component analysis (RPCA) algorithm, called frequency constrained RPCA (FC-RPCA), for selectively segmenting dynamic phenomena that exhibit spectra within a user-defined range of frequencies. The algorithm lacks subjective parameter tuning and demonstrates robust segmentation in datasets containing multiple motion sources and high amplitude noise. When tested on 17 *ex-vivo*, time lapse optical coherence tomography (OCT) B-scans of human ciliated epithelium, segmentation accuracies ranged between 91-99% and consistently out-performed traditional RPCA.

OCIS codes: (100.2000) Digital image processing; (170.4500) Optical coherence tomography.

References and links

1. J. Bang, T. Dahl, A. Bruinsma, J. H. Kaspersen, T. A. Nagelhus Hernes, and H. O. Myhre, "A new method for analysis of motion of carotid plaques from RF ultrasound images," Ultrasound Med. Biol. **29**, 967–976 (2003).
2. Y. Jia, S. T. Bailey, T. S. Hwang, S. M. McClintic, S. S. Gao, M. E. Pennesi, C. J. Flaxel, A. K. Lauer, D. J. Wilson, J. Hornegger, J. G. Fujimoto, and D. Huang, "Quantitative optical coherence tomography angiography of vascular abnormalities in the living human eye," Proc. Natl. Acad. Sci. U. S. A. **112**, E2395–402 (2015).
3. W. Tvaruskó, M. Bentele, T. Misteli, R. Rudolf, C. Kaether, D. L. Spector, H. H. Gerdes, and R. Eils, "Time-resolved analysis and visualization of dynamic processes in living cells," Proc. Natl. Acad. Sci. U. S. A. **96**, 7950–7955 (1999).
4. T.-W. Su, L. Xue, and A. Ozcan, "High-throughput lensfree 3D tracking of human sperms reveals rare statistics of helical trajectories, Supp. Mat," Proc. Natl. Acad. Sci. U. S. A. **109**, 16018–22 (2012).
5. R. Zareian, M. E. Susilo, J. A. Paten, J. P. McLean, J. Hollmann, D. Karamichos, C. S. Messer, D. T. Tambe, N. Saeidi, J. D. Zieske, and J. W. Ruberti, "Human Corneal Fibroblast Pattern Evolution and Matrix Synthesis on Mechanically Biased Substrates," Tissue Eng. Part A **22**, 1204–1217 (2016).
6. K. Jaqaman, D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S. L. Schmid, and G. Danuser, "Robust single-particle tracking in live-cell time-lapse sequences," Nat. Methods **5**, 695–702 (2008).
7. J. Ophir, I. Céspedes, H. Ponnekanti, Y. Yazdi, and X. Li, "Elastography: A quantitative method for imaging the elasticity of biological tissues," Ultrason. Imaging **13**, 111–134 (1991).
8. E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" J. ACM **58**, 1–37 (2011).
9. J. P. Haldar and Z. P. Liang, "Spatiotemporal imaging with partially separable functions: A matrix recovery approach," 2010 7th IEEE Int. Symp. Biomed. Imaging From Nano to Macro, ISBI 2010 - Proc. pp. 716–719 (2010).
10. R. Otazo, E. Candès, and D. K. Sodickson, "Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components," Magn. Reson. Med. **73**, 1125–1136 (2015).
11. H. Gao, J.-F. Cai, Z. Shen, and H. Zhao, "Robust principal component analysis-based four-dimensional computed tomography," Phys. Med. Biol. **56**, 3181–3198 (2011).
12. A. Baghaie, R. M. D'Souza, and Z. Yu, "Sparse and low rank decomposition based batch image alignment for speckle reduction of retinal OCT images," Proc. - Int. Symp. Biomed. Imaging **2015-July**, 226–230 (2015).
13. L. Fang, S. Li, D. Cunefare, and S. Farsiu, "Segmentation Based Sparse Reconstruction of Optical Coherence Tomography Images," **36**, 407–421 (2017).
14. Y. Sun, S. Li, and Z. Sun, "Fully automated macular pathology detection in retina optical coherence tomography images using sparse coding and dictionary learning," J. Biomed. Opt. **22**, 016012 (2017).
15. J. Barton and S. Stromski, "Flow measurement without phase information in optical coherence tomography images," Opt. Express **13**, 5234–5239 (2005).

16. B. K. Huang, U. A. Gamm, S. Jonas, M. K. Khokha, and M. A. Choma, "Quantitative optical coherence tomography imaging of intermediate flow defect phenotypes in ciliary physiology and pathophysiology," J. Biomed. Opt. **20**, 030502 (2015).
17. K. Baker and P. L. Beales, "Making sense of cilia in disease: The human ciliopathies," Am. J. Med. Genet. Part C Semin. Med. Genet. **151**, 281–295 (2009).
18. K. E. Tipirneni, J. W. Grayson, S. Zhang, D.-Y. Cho, D. F. Skinner, D.-J. Lim, C. Mackey, G. J. Tearney, S. M. Rowe, and B. A. Woodworth, "Assessment of acquired mucociliary clearance defects using micro-optical coherence tomography," Int. Forum Allergy Rhinol. **00**, 1–6 (2017).
19. D. L. Marks, T. S. Ralston, and S. A. Boppart, "Data Analysis and Signal Postprocessing for Optical Coherence Tomography," Optical Coherence Tomography (Springer, 2008).
20. A. L. Oldenburg, R. K. Chhetri, D. B. Hill, and B. Button, "Monitoring airway mucus flow and ciliary activity with optical coherence tomography," Biomed. Opt. Express **3**, 1978 (2012).
21. S. Wang, J. C. Burton, R. R. Behringer, and I. V. Larina, "In vivo micro-scale tomography of ciliary behavior in the mammalian oviduct," Sci. Rep. **5**, 13216 (2015).
22. Y. Ling, X. Yao, U. A. Gamm, E. Arteaga-Solis, C. W. Emala, M. A. Choma, and C. P. Hendon, "Ex vivo visualization of human ciliated epithelium and quantitative analysis of induced flow dynamics by using optical coherence tomography," Lasers Surg. Med. **49**, 270–279 (2017).
23. S. Boyd, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," Found. Trends Mach. Learn. **3**, 1–122 (2010).
24. X. Yao, Y. Gan, C. C. Marboe, and C. P. Hendon, "Myocardial imaging using ultrahigh-resolution spectral domain optical coherence tomography," J. Biomed. Opt. **21**, 061006 (2016).
25. Y. Ling, X. Yao, and C. P. Hendon, "Highly phase-stable 200 kHz swept-source optical coherence tomography based on KTN electro-optic deflector," Biomed. Opt. Express **8**, 3687 (2017).
26. K. Dabov, R. Foi, V. Katkovnik, and K. Egiazarian, "BM3D image denoising with shape-adaptive principal component analysis," Proc. Work. Signal Process. with Adapt. Sparse Struct. Represent. p. 6 (2009).
27. M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," Opt. Lett. **33**, 156 (2008).
28. Y. Zhao, Z. Chen, C. Saxer, Q. Shen, S. Xiang, J. F. de Boer, and J. S. Nelson, "Doppler standard deviation imaging for clinical monitoring of in vivo human skin blood flow," Opt. Lett. **25**, 1358–60 (2000).
29. Y. Jia, S. T. Bailey, D. J. Wilson, O. Tan, M. L. Klein, C. J. Flaxel, B. Potsaid, J. J. Liu, C. D. Lu, M. F. Kraus, J. G. Fujimoto, and D. Huang, "Quantitative optical coherence tomography angiography of choroidal neovascularization in age-related macular degeneration," Ophthalmology **121**, 1435–1444 (2014).
30. H. Y. Lee, P. D. Raphael, J. Park, A. K. Ellerbee, B. E. Applegate, and J. S. Oghalai, "Noninvasive in vivo imaging reveals differences between tectorial membrane and basilar membrane traveling waves in the mouse cochlea," Proc. Natl. Acad. Sci. **112**, 3128–3133 (2015).
31. N. C. Lin, C. P. Hendon, and E. S. Olson, "Signal competition in optical coherence tomography and its relevance for cochlear vibrometry," J. Acoust. Soc. Am. **141**, 395–405 (2017).

1. Introduction

Segmentation, detection, tracking, and analysis of motion are challenges in medical imaging that span multiple applications and modalities [1–5]. Understanding dynamic biological phenomenon necessitates quantitative analysis, which relies on tools for separation of the motion of interest from other static components in the scene [6,7]. Such separation of static and dynamic components is in general difficult due to variety in types of motion (particle flow, physical deformation, periodic motion, etc...).

1.1. Sparse representations

Sparse Representation theory offers many favorable tools for solving the kind of dynamic segmentation problems encountered in imaging. Sparse Representation is based on the claim that by utilizing the sparsity of some signal, we can solve an underdetermined system of linear equations. Many different formulations and solutions of this problem have been developed, one of which is the sparse and low-rank matrix decomposition problem as posed by Robust Principal Component Analysis (RPCA) [8]. Assuming some signal $Y \in \mathbb{R}^{m \times n}$ contains both static and dynamic components ($L, S \in \mathbb{R}^{m \times n}$, respectively), RPCA says that L and S can be estimated from the observation Y by solving the optimization problem:

$$
\begin{aligned}
&\text{minimize} \quad \|\mathbf{L}\|_* + \lambda\|\mathbf{S}\|_1 \\
&\text{subject to} \quad \mathbf{L} + \mathbf{S} = \mathbf{Y}
\end{aligned}
\tag{1}
$$

where $\|\cdot\|_*$ denotes the nuclear norm. Intuitively, this optimization problem seeks a sparse-valued matrix (S) and a matrix which is sparse in its singular values (L) whose sum is consistent with the observed matrix Y.

This technique has recently gained popularity in medical imaging applications, particularly Magnetic Resonance Imaging (MRI). Haldar and Liang recognized that in video MRI, temporal frames could be modeled as space-time columns where correlations in time yield a low-rank matrix [9]. More recently, this model was used to separate static and dynamics components of undersampled, video MRI [10]. Similar methods for separation of static and dynamics components have also been employed with success in 4D Computed Tomography imaging [11].

Despite the success it has demonstrated in applications like MRI, the scope of studies employing Sparse Representation theory in OCT imaging is relatively narrow. The majority of studies have focused on denoising, where generally speckle noise is reduced by modeling the denoised image as a low-rank version of the observed imaged [12]. Some studies have also explored dictionary learning methods for more targeted denoising approaches [13] and structure/texture based segmentation [14].

1.2. Segmentation in OCT

Optical Coherence Tomography (OCT) is a medical imaging technique which pairs millimeter-scale Field of View (FOV) with micron-scale spatial resolution. OCT's quick scanning speeds enable visualization of temporal characteristics of biological phenomenon by capturing multiple frames (B-scans) of the same scene over time. A time-lapse B-scan acquisition is often paired with Speckle Variance (SV) imaging, a post-processing technique that derives contrast from variation in the amplitude signal over time. An SV image is produced from a time-lapse B-scan dataset by calculating the standard of deviation over time at each pixel location [15, 16].

SV imaging is a useful technique for extracting additional functional information in biological applications, particularly in the study of ciliated epithelium. Motile cilia are found in the linings of specific human organs, such as trachea, fallopian tubes, and the epididymis, and have a hair-like morphology which beats in a periodic motion to induce fluid flow across the epithelium's surface [17]. Individual cilia are too small to directly observe their motion with most OCT systems [18]; however, their motion produces speckle-like fluctuations in the OCT image intensity that can be detected using SV imaging [19]. As a result, this property has been utilized in many studies to extract quantitative information on cilia behavior [20–22]. SV imaging, however, lacks sensitivity to other types of motion, and relies heavily on subjective processing which is prohibitive in studies involving large volumes of data. Both of these challenges remain un-addressed by current solutions.

We present a novel method based in Sparse Representation theory for segmenting dynamic contents of OCT data that addresses the issues of selectivity and user-subjectivity prevalent in conventional methods. In this manuscript, we first present a model for the datasets our method targets and a qualitative description of the algorithm's key features. This is followed by a brief explanation of the underlaying theory and a derivation of the algorithm. The remainder of the manuscript describes the employed methods, results, and implications of applying this algorithm to *ex-vivo*, time-lapse OCT B-scans of motile cilia from human trachea samples.

2.   Theory

2.1.   Data model

A time-lapse B-scan OCT dataset can be modeled as a composite of a static component corresponding to a low-rank matrix and a dynamic component corresponding to a sparse matrix. In Eq. (1), the sparse matrix S is obtained minimizing its ℓ^1 norm. In the context of an OCT B-scan, this implies that S represents the parts of the image which are sparse in the spatial-temporal dimension or basis.

Dynamic signals, particularly in biological applications, are often sparse in frequency, meaning they have a frequency domain representation with very few non-zero Fourier coefficients. This observation has motivated numerous other applications such as JPEG compression. To obtain an S that is sparse in the temporal-frequency domain, Eq. (1) can simply be modified to minimize the ℓ^1 norm of SF^T, where F^T is the transpose of the Discrete Fourier Transform matrix. Multiplying this by S is equivalent to taking the Discrete Fourier Transform of each pixel along the temporal dimension.

In addition to modeling Y as containing features which are sparse in frequency, we assume that it may contain multiple such features with distinct frequency spectra. A common example of this situation might be that Y contains both speckle noise and features which represent important biological motion. These are both time-lapse signals, however, we would like the ability to exclusively segment the biological motion of interest. We address this by introducing a second constraint on the sparse component's support that allows the user to choose which frequencies are allowed in S. These changes to Eq. (1) result in a new optimization problem

$$
\begin{aligned}
\text{minimize} \quad & \|\mathbf{L}\|_* + \lambda \|\mathbf{S}F^T\|_1 \\
\text{subject to} \quad & \mathbf{L} + \mathbf{S} = \mathbf{Y}, \\
& \mathbf{S}\mathbf{F}^T\mathbf{\Gamma} = 0
\end{aligned}
\tag{2}
$$

where $\Gamma \in \mathbb{R}^{n \times n}$ is a diagonal matrix of 1's and 0's. The user constructs Γ such that multiplying SF^T by Γ picks out the unwanted frequency components of S.

In summary, solving this optimization problem produces two matrices, L and S, which are sets of images corresponding to static background and dynamic motion of interest in the original time-lapse OCT B-scans Y, respectively. L and S are constrained to enforce consistency with the original dataset ($L + S = Y$) and S is constrained to only contain non-zero frequency components at frequencies specified by the user through construction of the matrix Γ. A visualization of this model is provided in Fig. 1. In the next section, we derive the necessary equations for solving Eq. (2).

2.2.   Algorithm

To handle the challenges of solving a constrained optimization problem with a non-smooth objective function and constraints, the Alternating Direction Method of Multipliers (ADMM) is employed [23]. The important feature of ADMM is that it allows us to solve for the *primal* variables (in this case L and S) by converting a constrained optimization problem into a series of unconstrained subproblems. Despite this convenience, Eq. (2) is not in a form where a solution can be easily derived using the tools of ADMM. Instead, we make an additional modification to the optimization problem using a variable splitting technique that will produce solvable update steps in our algorithm:
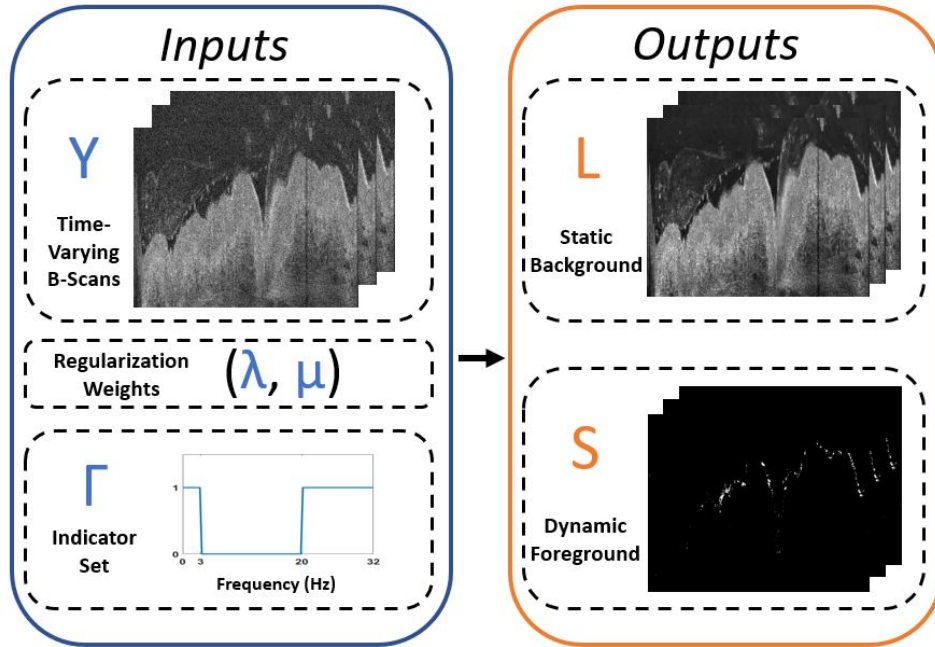
Fig. 1. Visualization of the input and output variables for the proposed Frequency Constrained RPCA algorithm. Inputs (blue box) include an image stack of time-lapse B-scans (Y), two scalar weights (λ, μ) which provide tuning of the ℓ^1 and ℓ^2 regularizers, respectively, and the indicator set (Γ) which allows the user to select frequencies to reject from the sparse component during the decomposition. The algorithm outputs (orange box) are two image stacks (L, S) of the same size as input Y that correspond to the low-rank background and sparse foreground, respectively.

$$
\begin{aligned}
\text{minimize} \quad & \|\mathbf{L}\|_* + \lambda \|\bar{\mathbf{S}}\mathbf{F}^T\|_1 \\
\text{subject to} \quad & \mathbf{L} + \mathbf{S} = \mathbf{Y}, \\
& \mathbf{S}\mathbf{F}^T\boldsymbol{\Gamma} = 0, \\
& \bar{\mathbf{S}} - \mathbf{S} = 0
\end{aligned}
\tag{3}
$$

where $\bar{S}$ is a new primal variable and a new constraint is added to enforce consistency between S and $\bar{S}$. ADMM can be used to solve Eq. (3) by maximizing the dual function g over the *dual* variables which are matrices of prices associated with each constraint. The dual function is defined as

$$
g(\boldsymbol{\nu}, \boldsymbol{\rho}) = \inf_{\mathbf{L}, \mathbf{S}} \mathcal{L}_\mu(\mathbf{L}, \mathbf{S}, \boldsymbol{\nu}, \boldsymbol{\rho})
\tag{4}
$$

where ($\boldsymbol{\nu}, \boldsymbol{\rho}, \boldsymbol{\gamma}$) are the dual variables and $\mathcal{L}_\mu$ is the augmented Lagrangian. The augmented Lagrangian associated with Eq. (3) is defined as

$$
\begin{aligned}
\mathcal{L}_\mu(\mathbf{L}, \mathbf{S}, \bar{\mathbf{S}}, \boldsymbol{\nu}, \boldsymbol{\gamma}, \boldsymbol{\rho}) = {} & \|\mathbf{L}\|_* + \lambda \|\bar{\mathbf{S}}\mathbf{F}^T\|_1 \\
& + \langle \mathbf{L} + \mathbf{S} - \mathbf{Y}, \boldsymbol{\nu} \rangle + \langle \bar{\mathbf{S}} - \mathbf{S}, \boldsymbol{\gamma} \rangle + \langle \mathbf{S}\mathbf{F}^T\boldsymbol{\Gamma}, \boldsymbol{\rho} \rangle \\
& + \frac{\mu}{2}(\|\mathbf{L} + \mathbf{S} - \mathbf{Y}\|_F^2 + \|\bar{\mathbf{S}} - \mathbf{S}\|_F^2 + \|\mathbf{S}\mathbf{F}^T\boldsymbol{\Gamma}\|_F^2)
\end{aligned}
\tag{5}
$$

where $\| \cdot \|_F$ denotes the Frobenius norm. The augmented Lagrangian penalizes the constraints in Eq. (3) through a Frobenius inner product and a quadratic penalty. We can maximize g by solving a series of iterates that update the primal and dual variables one at a time. These iterates take the following form:

$$\mathbf{L}^{(k+1)} \in \arg \min_{\mathbf{L}} \mathcal{L}_\mu(\mathbf{L}, \mathbf{S}^{(k)}, \bar{\mathbf{S}}^{(k)}, \boldsymbol{\nu}^{(k)}, \boldsymbol{\gamma}^{(k)}, \boldsymbol{\rho}^{(k)}) \tag{6}$$

$$\bar{\mathbf{S}}^{(k+1)} \in \arg \min_{\bar{\mathbf{S}}} \mathcal{L}_\mu(\mathbf{L}^{(k+1)}, \mathbf{S}^{(k)}, \bar{\mathbf{S}}, \boldsymbol{\nu}^{(k)}, \boldsymbol{\gamma}^{(k)}, \boldsymbol{\rho}^{(k)}) \tag{7}$$

$$\mathbf{S}^{(k+1)} \in \arg \min_{\mathbf{S}} \mathcal{L}_\mu(\mathbf{L}^{(k+1)}, \mathbf{S}, \bar{\mathbf{S}}^{(k+1)}, \boldsymbol{\nu}^{(k)}, \boldsymbol{\gamma}^{(k)}, \boldsymbol{\rho}^{(k)}) \tag{8}$$

$$\boldsymbol{\nu}^{(k+1)} = \boldsymbol{\nu}^{(k)} + \mu(\mathbf{L}^{(k+1)} + \mathbf{S}^{(k+1)} - \mathbf{Y}) \tag{9}$$

$$\boldsymbol{\gamma}^{(k+1)} = \boldsymbol{\gamma}^{(k)} + \mu(\bar{\mathbf{S}}^{(k+1)} - \mathbf{S}^{(k+1)}) \tag{10}$$

$$\boldsymbol{\rho}^{(k+1)} = \boldsymbol{\rho}^{(k)} + \mu(\mathbf{S}^{(k+1)}\mathbf{F}^T\boldsymbol{\Gamma}) \tag{11}$$

While the dual variable updates have a straight forward solution, the primal updates require additional tools to solve. Solving Eq. (6) and Eq. (7) involves minimizing a convex function $h(x)$ ($\| \cdot \|_*$ in the case of Eq. (6)) and a separable quadratic. This operation is termed the *proximal operator* of $h(x)$:

$$\text{prox}_{h/\mu} = \arg \min_x \left\{ h(x) + \frac{\mu}{2}\|x - w\|_F^2 \right\} \tag{12}$$

Representing the minimization this way is desirable since for many convex (even non-smooth) functions, its proximal operator has a simple numerical solution. Eq. (6) and Eq. (7) can be solved by rearranging them to be the same form as Eq. (12), where the proximal operator is applied to the nuclear norm and ℓ^1 norm, respectively.

$$\begin{aligned}\mathbf{L}^{(k+1)} &= \arg \min_{\mathbf{L}} \|\mathbf{L}\|_* + \langle \mathbf{L} + \mathbf{S} - \mathbf{Y}, \boldsymbol{\nu} \rangle + \frac{\mu}{2}\|\mathbf{L} + \mathbf{S} - \mathbf{Y}\|_F^2 \\ &= \text{prox}_{\|\cdot\|_*}(\mathbf{Y} - \mathbf{S}^{(k)} - \mu^{-1}\boldsymbol{\nu}^{(k)})\end{aligned} \tag{13}$$

$$\begin{aligned}\bar{\mathbf{S}}^{(k+1)} &= \arg \min_{\bar{\mathbf{S}}} \lambda\|\bar{\mathbf{S}}\mathbf{F}^T\|_1 + \langle \bar{\mathbf{S}} - \mathbf{S}, \boldsymbol{\gamma} \rangle + \frac{\mu}{2}\|\bar{\mathbf{S}} - \mathbf{S}\|_F^2 \\ &= \arg \min_{\bar{\mathbf{S}}} \lambda\|\bar{\mathbf{S}}\mathbf{F}^T\|_1 + \frac{\mu}{2n}\|\bar{\mathbf{S}}\mathbf{F}^T - (\mathbf{S} - \mu^{-1}\boldsymbol{\nu})\mathbf{F}^T\|_F^2 \\ &= \text{prox}_{n\lambda\|\cdot\|_1}([\mathbf{S}^{(k)} - \mu^{-1}\boldsymbol{\gamma}^{(k)}]\mathbf{F}^T)(\mathbf{F}^T)^{-1}\end{aligned} \tag{14}$$

The proximal operator for the nuclear norm is computed by element-wise soft-thresholding of the singular values of the argument. Similarly, the proximal operator for the ℓ^1 norm is computed by direct element-wise soft-thresholding of the argument. The soft-thresholding operation $soft()$ of matrix element u_i by threshold λ is defined for complex valued entries as $\text{sign}(u_i)\max(|u_i| - \lambda, 0)$

The expression for the S-update cannot be represented as an easily computed proximal operator; however, because all of the terms in $\mathcal{L}_\mu$ that contain S are differentiable, the update expression can be derived directly using the gradient. This method produces the following S update:

$$
\begin{aligned}
\mathbf{S}^{(k+1)} = \mu^{-1}[(\boldsymbol{\gamma}^{(k)} - \boldsymbol{\nu}^{(k)} - \rho^{(k)}\boldsymbol{\Gamma}^*(\mathbf{F}^T)^*) \\
+ (\mathbf{Y} - \mathbf{L}^{(k+1)}) + \bar{\mathbf{S}}^{(k+1)}] \times [2I + (\mathbf{F}^T\boldsymbol{\Gamma}\boldsymbol{\Gamma}^*(\mathbf{F}^T)^*)]^{-1}
\end{aligned}
\tag{15}
$$

At the end of each iteration, the Lagrangian is calculated using the latest variable updates. This update process is repeated until the value of the Lagrangian changes between iterations by some negligible amount, at which point the algorithm is assumed to have converged. The derived algorithm is summarized in Table 1.

Table 1. Frequency-Constrained RPCA Algorithm

Input: Observation matrix $\mathbf{Y} \in \mathbb{R}^{m \times n}$, weights μ, λ, frequency constraint matrix $\boldsymbol{\Gamma}$
Initialize: $\mathbf{L}_0 \leftarrow \mathbf{Y}, \mathbf{S}_0, \bar{\mathbf{S}}_0, \boldsymbol{\nu}_0, \boldsymbol{\gamma}_0, \boldsymbol{\rho}_0 \leftarrow 0$
while not converged
$(\mathbf{U}, \tilde{\mathbf{S}}, \mathbf{V}) \leftarrow svd(\mathbf{Y} - \mathbf{S}_k - \frac{1}{\mu}\boldsymbol{\nu}_k)$
$\mathbf{L}_{k+1} = \mathbf{U}[\text{soft}(\tilde{\mathbf{S}}, \frac{1}{\mu})]\mathbf{V}$
$\bar{\mathbf{S}}_{k+1} = \text{soft}([\mathbf{S}_k - \frac{1}{\mu}\boldsymbol{\gamma}_k]\mathbf{F}^T, (\lambda/\mu)n)(\mathbf{F}^T)^{-1}$
$\mathbf{S}_{k+1} = \frac{1}{\mu}[(\boldsymbol{\gamma}_k - \boldsymbol{\nu}_k - \rho_k\boldsymbol{\Gamma}^*(\mathbf{F}^T)^*) + (\mathbf{Y} - \mathbf{L}_{k+1}) + \bar{\mathbf{S}}_{k+1}] \times [2I + (\mathbf{F}^T\boldsymbol{\Gamma}\boldsymbol{\Gamma}^*(\mathbf{F}^T)^*)]^{-1}$
$\boldsymbol{\nu}_{k+1} = \boldsymbol{\nu}_k + \mu(\mathbf{L}_{k+1} + \mathbf{S}_{k+1} - \mathbf{Y})$
$\boldsymbol{\gamma}_{k+1} = \boldsymbol{\gamma}_k + \mu(\bar{\mathbf{S}}_{k+1} - \mathbf{S}_{k+1})$
$\boldsymbol{\rho}_{k+1} = \boldsymbol{\rho}_k + \mu(\mathbf{S}_{k+1}\mathbf{F}^T\boldsymbol{\Gamma})$
end while
Output: L, S

3. Methods

3.1. Time varying B-scans of motile ciliary beating

Five trachea samples from five different human subjects were obtained from the Columbia University Medical Center as the discarded regions of healthy, donated tissue from surgical lung transplantation for imaging. Samples were kept in GIBCO Medium 199 (Fischer Scientific, Waltham, USA)and imaged at 37° C to preserve tissue integrity.

Data was acquired using a custom Ultra-High Resolution OCT (UHR-OCT) system designed in house that has previously been used for imaging human trachea samples [22, 24, 25]. The UHR system operates at a frame rate of 64 Hz with 2.72 μm axial resolution and 5.52 μm lateral resolution. Each dataset covers a 21-second period of time (1350 frames) and a 1.44×1.80 mm FOV (800×500 pixels). Images from each dataset are digitally saved at double precision and log-power transformed upon acquisition. All samples were placed in formalin for histology processing with hematoxylin and eosin (H&E) staining following imaging.

3.2. RPCA decomposition

All images are denoised and co-registered before decomposition. Image stacks are denoised using VBM3D, an OCT speckle noise reduction technique, with an estimated noise standard of deviation of 12.0 (on a [0, 255] intensity range) [26]. Image frames from each dataset are co-registered using an "efficient sub-pixel co-registration" algorithm developed by Guizar-Sicairos *et al* [27]. To reduce the computational load, images from each dataset are cropped into three patches of equal size (300×150 pixels) which are individually processed by the FC-RPCA algorithm. Images are then vectorized such that each pixel is processed as a different dimension/feature and each frame in time is a distinct observation of that feature. Parameters μ and λ are chosen by

running a grid search over a range of values and choosing the pair that produces the best balance between a strong cilia signal and rejection of noise. This method yielded a value of 0.5 for μ and 0.0013 for λ and these parameters remained fixed for processing all datasets. The convergence of FC-RPCA was measured by calculating the change in $\mathcal{L}_\mu$ over subsequent iterations of the algorithm. Convergence was reached when this change was less than 10^{-2} or the algorithm exceeded 30 iterations. All processing was done using an Intel Core i7™3.4 GHz processor in MATLAB R2017a™. For comparison to the existing, state of the art RPCA technique, all datasets were processed twice, once with a frequency constraint of 3 to 14 Hz, and a second time with a constraint that only rejects the DC component. All other algorithm parameters were kept the same for both processing variations.

3.3. Quantitative accuracy/selectivity metric

Many potential applications of FC-RPCA related to segmentation involve eventually using the sparse output to create a mask that extracts the object of interest from the original dataset. To quantitatively measure the performance of FC-RPCA, a metric was developed for a mask-based segmentation task to calculate false positive rates. A ground truth was first established by performing blind, manual segmentation on each dataset. A researcher familiar with OCT images was chosen to perform the manual segmentation and was instructed to draw in an outline of the ciliated layer in an image from each dataset. Next, a second mask was created from the FC-RPCA results by taking the Maximum Intensity Projection (MIP) of the sparse component S and applying a 3 by 3 pixel spatial median filter. Segmentation accuracy was determined by calculating the percentage of non-zero pixels in the MIP mask that were outside the manually segmented region. In short, this metric quantifies how many pixels in the FC-RPCA segmentation that are associated with mucus or other sources of noise are falsely assumed to be associated with cilia activity. This metric additionally provides a means for quantitative comparison between traditional RPCA (which uses no frequency constraint) and our FC-RPCA.
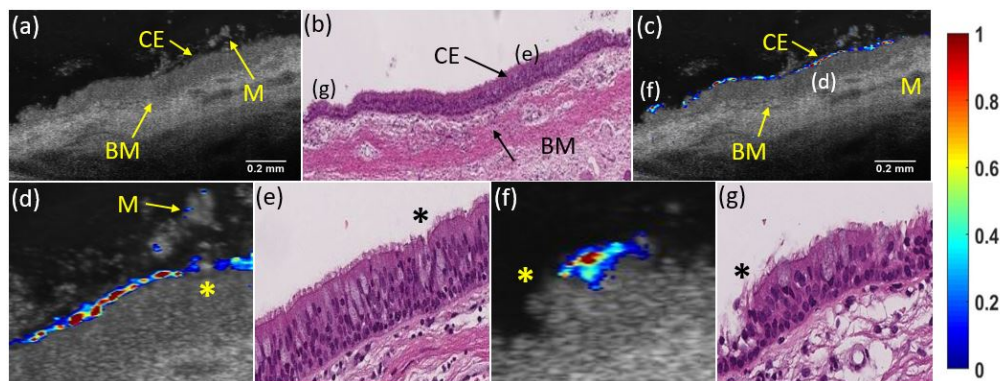

Fig. 2. (a) *Ex-vivo* OCT B-scan image of ciliated epithelium from human trachea (Sample 5) and (b) corresponding histology. (c) The Maximum Intensity Projection of the FC-RPCA sparse output painted over the image in (a) using MATLAB's *jet* colormap. (d) A closer look at a dense area of cilia located directly under a mucus cloud from (c). In contrast, (f) examines a sparsely populated area of cilia where the cells may be damaged or non-functioning. (e) and (g) are closer views of the histology corresponding to (d) and (f), respectively. The locations of these two regions are marked with the corresponding figure letter in the full view histology (b) and OCT B-scan (c) images. Colorbar indicates normalized intensity of the FC-RPCA sparse component. The arrows and corresponding labels mark key physiological regions (CE = ciliated epithelium, BM = basement membrane, M = mucus).

4. Results

Multiple datasets were acquired from imaging each sample at a number of different locations. In total, 17 datasets spanning 5 different samples were collected and processed using FC-RPCA.

One dataset from Sample 5, its corresponding histology, and FC-RPCA results are shown in Fig. 2. Arrows in Figs. 2(a)-2(d) are provided to highlight matching physiological features between the OCT B-scan and histology images. The segmentation results are visualized in Fig. 2(c) which was created by calculating the MIP image from the FC-RPCA sparse component output and overlaying it on the B-scan image shown in Fig. 2(a). The FC-RPCA overlay in Fig. 2(d) shows a dense area of cilia directly under a thick mucus cloud. With the exception of the small region directly under the thickest portion of mucus (see asterisk), the algorithm identifies the ciliated area with practically zero false positives from the surrounding mucus. Figure 2(f) examines another interesting area where the epithelial cells are partially denuded. In the corresponding histology (Fig. 2(g)), the asterisk marks a small area of which appears partially denuded with some potentially non-functioning cilia.

Figure 3 compares two frames from the FC-RPCA sparse component produced by processing a dataset from Sample 1 using two different frequency constraints. Figure 3(b) was produced using a 3 to 14 Hz constraint (the expected CBF range) while Fig. 3(c) was produced by blocking only the DC component. The frequency constraint allowed for more selective segmentation of ciliated tissue by rejecting high frequency noise and other low frequencies signals that corresponded to noise or other biologically processes like the slow moving mucus clouds at the tissue's surface. Videos of the corresponding constrained and unconstrained sparse components are provided in Visualization 1 and Visualization 2 , respectively.

Algorithm performance and the effect of the FC-RPCA frequency constraint were quantitatively evaluated using the segmentation accuracy metric described in Part C of the Methods section. Table 2 compares the ensemble average segmentation accuracy and standard deviation under both frequency constraint conditions for each imaged dataset. FC-RPCA performance was very high for the frequency constraint condition with accuracies ranging from the worst at 91.82% for Sample 4 to the best at 99.66% for Sample 1. Unsurprisingly, targeting the expected CBF of the cilia with the 3 to 14 Hz frequency constraint improved segmentation accuracy over the unconstrained condition in all samples.

Table 2. **Comparison of average cilia segmentation accuracy across datasets for multiple frequency constraint conditions**

Sample #	Datasets	Parameters	Segmentation Accuracy ± SD	
	(N)	(λ, μ)	Freq. Range [3, 14] Hz	Freq. Range (0, 32] Hz
1	4		99.66 ± 0.45%	99.12 ± 0.69%
2	3		92.94 ± 4.10%	87.95 ± 5.05%
3	5	(0.0013, 0.5)	96.08 ± 1.96%	93.17 ± 2.20%
4	2		91.82 ± 0.74%	88.40 ± 0.55%
5	3		94.36 ± 4.97%	89.98 ± 7.08%

5. Discussion

A new version of the RPCA algorithm with a user-defined frequency constraint on the sparse component was formulated and its solution was derived using an ADMM based approach. The effectiveness of the algorithm was demonstrated on time varying OCT B-scans of ex-vivo human trachea samples. The performance of FC-RPCA was evaluated both qualitatively and quantitatively and showed that the new frequency constraint feature increased segmentation selectivity and accuracy in 17 tested datasets over RPCA.
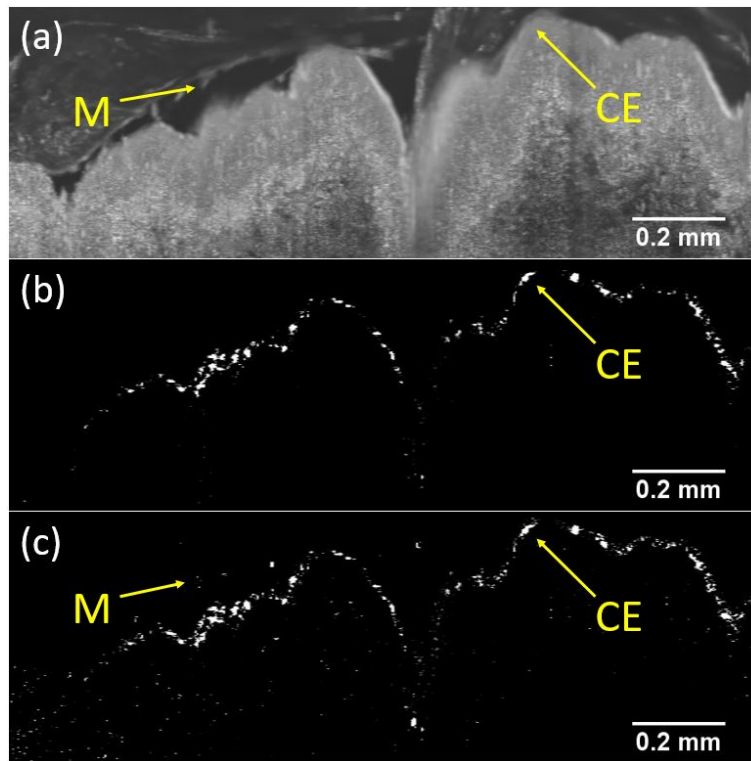
Fig. 3. Comparison between FC-RPCA sparse components created using two different sets of frequency constraints on data from Sample 1. (a) Single OCT B-scan before FC-RPCA processing. (b) Cilia region segmented using frequencies constrained between 3 and 14 Hz (see Visualization 1). (c) Cilia region segmented with only rejection of the DC component (see Visualization 2). Arrows in (a) and (c) point to a bright mucus cloud that appears in the unconstrained output (c), but was successfully rejected from the frequency constrained output (b). Additionally, many bright pixels in (c) associated with undesirable noise from static tissue were rejected in (b) due to the frequency constraint. The arrows and corresponding labels mark key physiological features (CE = ciliated epithelium, M = mucus).

Though not explicitly compared, the results additionally demonstrate clear advantages of the technique over the presently used Speckle Variance method. The major drawbacks of SV processing are that the technique is blind to the source of any temporally varying signal, is influenced by noise artifacts and low SNR regions, and significant parameter tuning is required to produce adequate identification of the moving object of interest [17-22]. Inspection of the quantitative results shown in Table 1 and the Fig. 3(b) clearly demonstrate the selectivity of FC-RPCA. Further evidence is provided in the supplementary materials (Visualization 3) which displays the FC-RPCA overlay (left) from Fig. 2(c) side-by-side with the Speckle Variance overlay of the same image (right). It is clear from visual inspection that the Speckle Variance technique is severely influenced by the mucus clouds at the surface of the tissue and the low SNR regions deep beneath the tissue, as expected. Additionally, parameter tuning in FC-RPCA is incredibly minimal in comparison to SV and is discussed in more detail in the next paragraph.

Reducing elements which introduce user-subjectivity to the processing chain was an important goal in the design of this algorithm. While FC-RPCA is significantly more objective and robust than the Speckle Variance technique, there are still aspects of the processing chain which could be considered subjective, in particular, the choice of FC-RPCA parameter λ. Surprisingly, a practical

range of λ's that will work for any L and S is directly given by the theory and a reasonable value within that range can be approximated as $\min(m, n)^{-1/2}$ [8]. We found, however, that obtaining the best segmentation results required testing the full range of possible values and selecting one value based on qualitative assessment of the result. While this choice is clearly subjective, the result was still general and robust in the sense that the grid search only needed to be run once to determine a value of λ that worked for all of the processed datasets. Based on the results of this study and the proposed theory in [8], we believe that λ need only be calibrated once for a given imaging system to continually produce meaningful segmentation results.

The results of the quantitative segmentation performance evaluation (Table 2) confirm that the sparse spectrum characterizing beating cilia can be leveraged using FC-RPCA in a segmentation task. The performance metric specifically identifies how well FC-RPCA rejects signals from thick mucus clouds and high amplitude noise in the static tissue. We chose to evaluate the FC-RPCA performance using this metric because rejection of these signals is a major drawback of the currently used speckle variance method. It should be emphasized that this metric does not reflect the ability of FC-RPCA to separate ciliated and denuded regions since this is very difficult to quantitatively evaluate against ground truth, even with corresponding histology and fluorescence microscopy images [21]. Despite this, the results provide significant qualitative evidence that FC-RPCA distinguishes between ciliated and denuded regions (see Fig. 2(f)). Drawing from first principles, we also expect FC-RPCA to be sensitive to this feature because denuded regions, even those with high amplitude noise, will not exhibit a sparse temporal-frequency spectrum within the expected range of CBF's.

While this manuscript only explores the applications in cilia imaging, it is clear that FC-RPCA has the potential to become a valuable tool in a much broader range of applications. Even within biomedical OCT imaging, the proposed data model fits for many new and exciting applications like OCT angiography [28, 29] and vibrometry [30, 31]. More generally, FC-RPCA could have applications in any segmentation task that targets objects which are sparse in the temporal frequency domain, particularly in images with high amplitude noise and/or multiple dynamic sources.

6. Conclusion

A novel, alternative form of the RPCA problem that targets sparsity in the temporal-frequency domain and features a user defined frequency constraint, termed FC-RPCA, was formulated and its solution via an ADMM based approach was derived. The method's feasibility was demonstrated for multiple time-lapse OCT B-scans through segmentation of motile cilia in a ciliated region of human trachea. The sparse component generated by FC-RPCA was shown to effectively rejects pixels associated with spectral features outside the specified frequency constraint such as nearby mucus clouds. The performance of FC-RPCA was quantitatively assessed for 17 different datasets and was shown to produce robust segmentation results in all cases. Performance was additionally evaluated for the unconstrained frequency case, which demonstrated the algorithm's decreased sensitivity to noise and other dynamic features when using the frequency constraint. The user defined frequency constraint ultimately gives the this new alternative to the traditional RPCA technique the potential to be a valuable tool for separating dynamic foreground features from static background in the noise heavy OCT imaging environment.

Funding

Acknowledgments