Groupwise registration of sequential images from multispectral imaging (MSI) of the retina and choroid

JIANWEI LIN,1 YUANJIE ZHENG,1,2,3,* WANZHEN JIAO,4 BOJUN ZHAO,4 SHAOTING ZHANG,5 JAMES GEE,6 AND RUI XIAO6

1School of Information Science and Engineering, Shandong Normal University, Shandong, China
2Institute of Life Sciences at Shandong Normal University, Jinan, China
3Key Lab of Intelligent Information Processing at Shandong Normal University, Jinan, China
4Department of Ophthalmology, Shandong Provincial Hospital, Shandong, China
5Department of Computer Science, University of North Carolina at Charlotte, Charlotte, NC, USA
6Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

*yzjzheng@sdnu.edu.cn

Abstract: Multispectral Imaging (MSI) produces a sequence of discrete spectral slices that penetrate different light-absorbing species or chromophores and is a noninvasive technology useful for the early detection of various retinal, optic nerve and choroidal diseases. However, eye movement during the image acquisition process may introduce spatial misalignment between MSI images. This potentially causes trouble in the manual/automatic interpretation of MSI, but still remains an unresolved problem to this date. To deal with this MSI misalignment problem, we present a method on the groupwise registration of sequential images from MSI of the retina and choroid. The advantage of our algorithm is at least threefold: 1) simultaneous estimation of landmark correspondences and a parametric motion model via quadratic programming, 2) enforcement of temporal smoothness on the estimated motion, and 3) inclusion of a robust matching cost function. As validated in our experiments with a database of 22 MSI sequences, our algorithm outperforms two state-of-the-art registration techniques proposed originally in other domains. Our algorithm is potentially invaluable in ophthalmologists’ clinical practice regarding various eye diseases.

© 2016 Optical Society of America

OCIS codes: (110.4234) Multispectral and hyperspectral imaging; (100.4145) Motion, hyperspectral image processing; (100.0100) Image processing; (100.2960) Image analysis; (170.4470) Ophthalmology.

References and links

1. P. Ghassemi, T. E. Travis, L. T. Moffatt, J. W. Shupp, and J. C. Ramella-Roman, “A polarized multispectral imaging system for quantitative assessment of hypertrophic scars,” Biomed. Opt. Express 5(10), 3337–3354 (2014).
2. N. T. Clancy, S. Arya, D. Stoyanov, M. Singh, G. B. Hanna, and D. S. Elson, “Intraoperative measurement of bowel oxygen saturation using a multispectral imaging laparoscope,” Biomed. Opt. Express 6(10), 4179–4190 (2015).
3. M. B. Bouchard, B. R. Chen, S. A. Burgess, and E. M. Hillman, “Ultra-fast multispectral optical imaging of cortical oxygenation, blood flow, and intracellular calcium dynamics,” Opt. Express 17(18), 15,670–15,678 (2009).
4. D. Hitchmoth, “Multispectral Imaging: A revolution in retinal diagnosis and health assessment,” Adv. Ocul. Care 4(4), 76–79 (2013).
5. D. L. Shechtman and P. M. Karpecki, “A look at MSI: multispectral imaging may help eye care providers diagnose retinal conditions earlier than conventional fundoscopy,” Rev. Opt. 149(1), 88–90 (2012).
6. R. Maharaj, “The clinical applications of multispectral imaging,” Rev. Opt. 148(11), SS19–SS19 (2011).
7. I. Diebele, I. Kuzmina, A. Lihachev, I. Kapostins, A. Derjabo, L. Valeine, and J. Spigulis, “Clinical evaluation of melanomas and common nevi by spectral imaging,” Biomed. Opt. Express 3(3), 467–472 (2012).
8. J. Zhang, Z. Yu, and L. Liu, “Multimodality imaging in diagnosing polypoidal choroidal vasculopathy,” Optom. Vis. Sci. 92(1), e21–e26 (2015).
9. C. Zimmer, D. Kahn, R. Clayton, P. Dugel, and K. Freund, “Innovation in diagnostic retinal imaging: multispectral imaging,” Retina Today 9(7), 94–99 (2014).
10. N. Everdell, I. Styles, E. Claridge, J. Hebdon, and A. Calcagni, “Multispectral imaging of the ocular fundus using LED illumination;” in European Conferences on Biomedical Optics, p. 7371C (International Society for Optics and Photonics, 2009).
11. I. B. Styles, A. Calcagni, E. Claridge, F. Orijuela-Espina, and J. Gibson, “Quantitative analysis of multi-spectral fundus images,” Med. Image Anal. 10(4), 578–597 (2006).
12. S. K. Balci, P. Golland, and W. Wells, “Non-rigid groupwise registration using B-spline deformation model,” Open source and open data for MICCAI pp. 105–121 (2007).
13. O. Arandjelic, D.-S. Pham, and S. Venkatesh, “Groupwise registration of aerial images,” presented on the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015) (2015).
14. S. Boyd and L. Vandenberghe, Convex Optimization (Cambridge University, 2004).
15. F. P. Oliveira and J. M. R. Tavares, “Medical image registration: a review,” Comput. Method. Biomech. 17(2), 73–93 (2014).
16. A. Ghodlouei, N. Kehtarnavaz, R. Briggs, M. Devous, and K. Gopinath, “Brain functional localization: a survey of image registration techniques,” IEEE Trans. Med. Imaging 26, 427–451 (2007).
17. J. Modersitzki, Numerical Methods for Image Registration (Oxford University, New York, NY, USA, 2004).
18. L. Zöldle, E. Learned-Miller, E. Grimson, and W. Wells, “Efficient population registration of 3D data,” in International Workshop on Computer Vision for Biomedical Image Applications, pp. 291–301 (Springer, 2005).
19. C. Wachinger and N. Navab, “Simultaneous registration of multiple images: Similarity metrics and efficient optimization,” IEEE Trans. Pattern Anal. Mach. Intell. 35(5), 1221–1233 (2013).
20. J. Kim, C. Liu, F. Sha, and K. Grauman, “Deformable spatial pyramid matching for fast dense correspondences,” in Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2307–2314 (2013).
21. T. Zhou, Y. J. Lee, X. Y. Stella, and A. A. Efros, “Flowweb: Joint image set alignment by weaving consistent, pixel-wise correspondences,” in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1191–1200 (IEEE, 2015).
22. E. G. Learned-Miller, “Data driven image models through continuous joint alignment,” IEEE Trans. Pattern Anal. Mach. Intell. 28(2), 236–250 (2006).
23. B. G. Huang, V. Jain, and E. Learned-Miller, “Unsupervised joint alignment of complex images,” in 2007 IEEE 11th International Conference on Computer Vision, pp. 1–8 (IEEE, 2007).
24. A. Rasoulian, R. Rohling, and P. Abolmaesumi, “Group-wise registration of point sets for statistical shape models,” IEEE Trans. Med. Imaging 31(11), 2025–2034 (2012).
25. P. Ghosh and B. Manjunath, “Robust simultaneous registration and segmentation with sparse error reconstruction,” IEEE Trans. Pattern Anal. Mach. Intell. 35(2), 425–436 (2013).
26. A. Wagner, J. Wright, A. Ganesh, Z. Zhou, H. Mobahi, and Y. Ma, “Toward a practical face recognition system: Robust alignment and illumination by sparse representation,” IEEE Trans. Pattern Anal. Mach. Intell. 34(2), 372–386 (2012).
27. J. Kutanina and P. Pedersen, “A Markov random field approach to group-wise registration/mosaicing with application to ultrasound,” Med. Image Anal. 24(1), 106–124 (2015).
28. X. Shen, L. Xu, Q. Zhang, and J. Jia, “Multi-modal and multi-spectral registration for natural images,” in European Conference on Computer Vision, pp. 309–324 (Springer, 2014).
29. D. G. Lowe, “Object recognition from local scale-invariant features,” in Computer Vision, 1999, Seventh International Conference on, vol. 2, pp. 1150–1157 (IEEE, 1999).
30. F. Pluim, J. A. Maintz, and M. A. Viergever, “Mutual-information-based registration of medical images: a survey,” IEEE Trans. Med. Imaging 22(8), 986–1004 (2003).
31. R. Kolar, L. Kubova, and J. Jan, “Registration and fusion of the autofluorescent and infrared retinal images,” Int. J. Biomed. Imaging 2008(2008).
32. M. Irani and P. Anandan, “Robust multi-sensor image alignment,” in Computer Vision, 1998. Sixth International Conference on, pp. 959–966 (IEEE, 1998).
33. J. Lewis, “Fast normalized cross-correlation,” vol. 10, pp. 120–123 (1995).
34. C. V. Stewart, “Robust parameter estimation in computer vision,” SIAM Rev. 41(3), 513–537 (1999).
35. A. Can, C. V. Stewart, B. Roysam, and H. L. Tanenbaum, “A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina,” IEEE Trans. Pattern Anal. Mach. Intell. 24, 347–364 (2002).
36. Y. Zheng, A. A. Hunter III, J. Wu, H. Wang, J. Gao, M. G. Maguire, and J. C. Gee, “Landmark matching based automatic retinal image registration with linear programming and self-similarities,” in Biennial International Conference on Information Processing in Medical Imaging, pp. 674–685 (Springer, 2011).
37. Y. Zheng, E. Daniel, A. A. Hunter, R. Xiao, J. Gao, H. Li, M. G. Maguire, D. H. Brainard, and J. C. Gee, “Landmark matching based retinal image alignment by enforcing sparsity in correspondence matrix,” Med. Image Anal. 18(6), 903–913 (2014).
38. H. Chui and A. Rangarajan, “A new point matching algorithm for non-rigid registration,” Comput. Vis. Image Und. 89, 114–141 (2003).
39. C. Rorden and M. Brett, “Stereotaxic display of brain lesions,” Behav. Neurol. 12(4), 191–200 (2000).
40. H. Chui and A. Rangarajan, “A new point matching algorithm for non-rigid registration,” Comput. Vis. Image Und. 89(2), 114–141 (2003).
41. W. M. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, “Multi-modal volume registration by maximization of mutual information,” Med. Image Anal. 1(1), 35–51 (1996).
1. Introduction

Multispectral Imaging (MSI) captures image data at specific frequencies across the electromagnetic spectrum and is a non-contact, optical technology used in various applications, including but not limited to remote sensing, aerial survey, astronomy, dentistry, dermatology and histopathology [1–3]. An emerging and advanced application of MSI is its use to dissect and visualize the inner limiting membrane all the way to the choroid of the ocular by producing a series of spectral images, e.g., the Retinal Health Assessment (RHA, Annidis Health System Corp.) [4–6]. MSI allows the clinician to differentiate and follow a wide variety of complex conditions and disorders of the eye [4–7]. Compared with conventional retinal fundus photography, which is limited to parameters of the visible spectrum, MSI is more sensitive in detecting features that are reflected in longer range of the spectrum, e.g., deep retinal layers and the choroid. With the ability to complement the clinician’s knowledge of retinal anatomy and alterations in physiology, MSI provides a novel way to examine the retina [4] and may help to diagnose retinal conditions earlier than conventional fundoscopy [5, 8].

MSI of the ocular employs an extensive range of discrete monochromatic LED-sourced wavelengths (ranging from 550nm to 950nm), with each one capable of creating a discrete spectral slice by penetrating/fluorescing different light-absorbing species or chromophores throughout the retina and choroid, as shown in Fig. 1. Each spectral slice represents a cumulative en face spectral view of the fundus and the clinicians’ diagnosis is commonly carried out by examining multiple slices. For example, early changes of retinal pigment epithelium (RPE) of RPE disruption are observed as normal structures in short wavelengths and a slight structural variation in pigment pattern in the red (620nm, 660nm, 690nm and 740nm) and infrared (760nm, 780nm, 810nm and 850nm) spectral slices [9]. Manual assessment of the variation of retinal/choroidal structures and functions remains the reference standard for MSI based diagnosis. However, automatic and quantitative analysis of MSI images is becoming more and more demanded because it can help to deal with the problems of the manual works, e.g., being tedious, labor-intensive, subjective and error-prone.

One main drawback of most MSI systems lies in the possible significant movement between frames in the spectral slice stack (as shown in Fig. 2), which is caused by the fact that the acquisition time of a complete set of spectral slices is often longer than the natural timescale of eye’s saccadic movement [10]. Although there exist a lot of efforts to increase the image acquisition speed, current MSI technologies (e.g., the RHA) still require seconds for taking a complete set of spectral slices. The movements between spectral slices might introduce troubles into manual/automatic interpretation and evaluation of MSI. Theoretically speaking, the movements can be handled by registering the images, however, this is not a trivial task in practice due to three major hardships. First of all, the anatomical and pathological retinal features may appear significantly different between different spectral slices. This results from the fact that different wavelengths are used for imaging different anatomic components of the retina and choroid. For example, shorter wavelengths are used to reflect anterior retinal features while longer ones are employed for deeper retinal features and the choroid. Besides this, different tissues may have different light absorption and reflection properties. For example, blood vessels appear as having more well-defined features in spectral slices corresponding to shorter wavelengths than longer wavelengths. The second hardship in registering MSI spectral slices is attributed to the illumination inhomogeneity caused by the curvature of the retina. As a result, the incident light intensity at a given spatial location may change significantly between images because of eye movement, introducing difficulties in interpreting the image data [11]. The third hardship lies in the requirement of registering a sequence of MSI images instead of the common registration tasks for which a pair of images are concerned. Moreover, MSI image sequence is different from a generic image group investigated by most groupwise image registration algorithms in the fact that eye motion changes smoothly during MSI imaging process.
Fig. 1. A collection of MSI images from Annidis RHA, in which from left to right and from top to bottom, the first 11 sequential images are captured with short wavelengths of MSI-550, MSI-580, MSI-590, MSI-620, MSI-660, MSI-690, MSI-740, MSI-760, MSI-780, MSI-810 and MSI-850, respectively.

There has been very little work on groupwise registration of MSI sequential images. Moreover, an examination of groupwise registration techniques designed for other images (e.g., MRI brain images [12] or aerial images [13] as to be summarized in Sec. 2) shows that they can’t be readily applied to resolving the problem of groupwise MSI registration. State-of-the-art techniques of groupwise image registration are mostly proposed for aligning intra-modality images. For these algorithms, image acquisition properties are consistent across different images, being different from MSI images that are acquired using different LED-source wavelengths in order to penetrate different species/chromophores as mentioned above. There exist techniques for jointly registering inter-modality images, however, most of them assume that motions obtained from different pairs of images are irrelevant. This simple assumption prevents these techniques from dealing with the smooth change of eye motion captured during MSI imaging process.

The goal of this work is to provide a tool for achieving a groupwise registration of sequential MSI images of the retina and choroid. We present a novel landmark matching approach which is validated to be capable of dealing with the challenges in registering MSI images (as shown in Fig. 2). Our method is superior to state-of-the-art groupwise registration algorithms due to several major novelties of significance. First, our approach measures similarity between landmark points by using a robust cross correlation function. Second, it establishes dense correspondences between MSI images by simultaneously building connections between landmark points across different images and estimating a predefined motion model. Third, a constraint of temporal smoothness is enforced when estimating the motions across the image sequence, which complies better with the circumstances of retinal/choroidal MSI imaging in practice. Finally, our optimization is accomplished by resolving all unknowns using a combinatorial optimization technique - quadratic programming [14].

The remainder of this article is structured as follows. In Sec. 2, we present a review of the works related to groupwise image registration. In Sec. 3, we describe the details of our approach. Experimental results are provided in Sec. 4 and we conclude in Sec. 5.
Fig. 2. Significant motion exists between sequential MSI images and our groupwise registration algorithm can handle it effectively. Top row, left to right: MSI images of MSI-550 and MSI-850 in Fig. 1, respectively. Bottom row, left to right: patches in the images created by overlaying MSI-550 and MSI-850 before and after applying our algorithm, respectively. Misalignment is eliminated in the right image in contrast to the left one.

2. Related Work

Image registration has been extensively involved in a wide variety of applications in fields of medical image analysis and computer vision. The corpus of relevant previous work is pretty rich and varied, which can be basically categorized as pixel/voxel-wise based techniques, landmark/feature based methods or their combinations [15–17]. They often bear a high degree of domain specificity and few papers can be found specifically for MSI applications. Most existing image registration techniques register images in a pair-wise manner, either aligning an image to another, or an image to a map. Groupwise registration operates on image sets instead, i.e. aligning a group of images simultaneously. Detailed review to all pairwise image registration techniques is beyond the scope of this paper and in this paper we only provide a brief review to methods of groupwise image registration.

Groupwise registration is an important tool in fields of both medical image analysis and computer vision. It can be used for establishing anatomical correspondences among subjects in a population in order for characterizing anatomical shape differences within/across populations [18,19]. It can also achieve building connections between similar points/regions across different images in order for representation, 3D modeling, synthesis, morphing, browsing of for example human faces [20,21].

Most of the existing groupwise registration techniques focus on aligning a collection of intra-modality images. The well-known congealing framework and its extensions [12,22,23] aim to align a large number of binary images from a database of handwritten digits, 2D face images or 3D medical images by summing up the entropy of pixel/voxel stacks over the image. Congealing warps each image via a parametric transformation and takes advantage of sequential optimization to gradually lower the entropy of the intensity distribution of the image. It identifies one image as a template and align all other images to it with a pair-wise approach. However, it leads to an undesired introduction of bias with respect to the a priori chosen template. The
point set based approach proposed in [24] attempts to create correspondences between groups of point sets by leveraging mean shape and its transformations to all training shapes. However, it requires the shape of objects of interests to be known in advance. The methods in [25, 26] exploit sparse learning and focus on face alignment. The techniques in [19, 27] assume a Markov property of the image and accumulate pairwise image matching in order for achieving a groupwise registration.

Groupwise registration of inter-modality images has recently attracted intensive research interests and various methods were proposed. For example, the method proposed in [13] focuses on aerial images and formulates group registration as a function of pairwise registration, in which image similarity is characterized as normalized cross-correlation of high-pass filtered images. The technique in [28] is designed for aligning multi-modal/multi-spectral natural images by using a robust matching cost function and a gradient descent optimization technique.

Our approach (to be detailed in Sec. 3) is distinguished from existing groupwise registration algorithms due to several of its characteristics, including a simultaneous estimation of landmark correspondences and motion model, an inclusion of temporal smoothness on the estimated motion, a robust matching cost function and an optimization with quadratic programming. These enable our approach to produce extremely accurate estimation of motion between MSI images.

3. Our approach

We believe that a successful technique for achieving an accurate group registration of sequential MSI images depends on several key factors: robustness to appearance variations between MSI images, superior optimization for obtaining more accurate estimations and ability to enforce temporal smoothness on the estimated motions. To target these, we developed a landmark matching based groupwise image registration algorithm. It exploits a robust matching cost function and a motion model designed specifically for eye motion. Moreover, it simultaneously optimize landmark correspondences and motion model parameters by leveraging a quadratic programming technique. Besides these, it achieves groupwise image registration by combining a series of pairwise image registrations, for which motion differences between temporally neighboring pairs are penalized.

In the rest of this section, we start with a description of our overall formulation, followed with an elaboration on the landmark detection, robust matching cost, motion model and optimization strategy, respectively.

3.1. Problem formulation

Given a collection of MSI images \{I_1, I_2, \cdots, I_N\}, we aim to find spatial correspondences jointly over the entire image set. We carry out the task by establishing connections between landmark points across pair of images \(((i, j), 1 \leq i \leq N, 1 \leq j \leq N)\) while enforcing a parametric model (denoted by its parameters \(\Theta_{i,j}\)) which represents the motion between \(I_i\) and \(I_j\). Our groupwise registration algorithm is then formulated as an optimization problem which is cast as minimizing the total matching cost

\[
\min \sum_{i=1}^{N} \sum_{j=1}^{N} \text{tr}(C_{i,j}^T E_{i,j}) \quad (1)
\]

while obliging a parametric motion model

\[
X_j = \chi_i \Theta_{i,j}. \quad (2)
\]

In Eq. (1), \(C_{i,j}\) and \(E_{i,j}\) denote the cost matrix and the correspondence matrix for matching landmark points of \(I_i\) to the ones of \(I_j\), respectively. For \(C_{i,j}\) and \(E_{i,j}\), rows correspond to the
landmark points of $I_i$ while columns are associated with the ones of $I_j$. $E_{i,j}$ is binary, and for which, each row contains exactly one element, meaning that $E_{i,j}$'s point specified by the row number of $E_{i,j}$ matches $I_i$'s point specified by the column number of $E_{i,j}$. These can be stated alternatively as the following two constraints:

$$E_{i,j} \in \{0, 1\}^{N_i \times N_j},$$

$$E_{i,j} \mathbf{1} = \mathbf{1},$$

where $N_i$ and $N_j$ denote the number of landmark points of $I_i$ and $I_j$, respectively, and $\mathbf{1}$ is a vertical vector of ones. In Eq. (1), $\text{tr}()$ means the trace of a matrix and $\text{tr}(C_{i,j}^T E_{i,j})$ amounts to calculate the total cost when matching $I_i$ to $I_j$.

In Eq. (2), the spatial transformation from $I_i$ to $I_j$ is expressed as a motion model. $X_j$ is a $N_j \times 2$ matrix formulated by the coordinates of all $N_j$ landmark points of $I_j$, with each row composed by the coordinates $[x, y]$ of the corresponding landmark point. $\chi_i$ is a matrix created using coordinates of the landmark points of $I_i$, for which each row corresponds to a landmark point and consists of components built with exponentiation and multiplication of this point's coordinate $x$ and $y$. $\Theta_{i,j}$ is a matrix containing parameters of the motion model, containing two columns for $x$ and $y$ coordinates, respectively. The specific expressions of all components in $\chi_i$, the number of columns of $\chi_i$ and size of $\Theta_{i,j}$ are determined by the motion model (to be described in Sec. 3.3) we will choose.

Unknowns in our groupwise registration algorithm as defined in Eqs. (1)-(4) include the correspondence matrices $\{E_{i,j}, 1 \leq i \leq N, 1 \leq j \leq N\}$ and the motion model’s parameters $\{\Theta_{i,j}, 1 \leq i \leq N, 1 \leq j \leq N\}$. In order to resolve these unknowns, we need to specify the landmark detection technique, the matching cost, an appropriate motion model for MSI of the retina and choroid and an optimization technique for computing the unknown parameters, which will be detailed in the follows.

### 3.2. Landmark detection and matching cost

The landmark points used for matching MSI images in our groupwise registration algorithm are specified as the key locations of SIFT [29] due to their prestigious benefit of being stable. These key locations are defined as maxima and minima of the result of difference of Gaussian function applied to different image scales, for which low contrast candidate points and edge response points are discarded in a postprocessing.

Determining the matching cost between landmark points in different images is vital in establishing correspondences but not a trivial due to the inconsistency of intensity, gradient or even structure across MSI images. Measures being widely used in state-of-the-art multi-modal image registration techniques include mutual information [30], cross correlation of image gradient [31], cross correlation of Laplacian energy map [32] and normalized cross correlation of image color and image gradient [28]. Inspired by the work in [28], we employ a robust cross correlation. Taking the $k$th point of $I_i$ being matched to the $l$th point of $I_j$ as an example, our matching cost is defined as

$$C_{i,j}(k,l) = w \cdot \rho \left( \frac{1}{\gamma (I_i(k), I_j(l))} \right) + (1 - w) \cdot \rho \left( \frac{1}{\gamma (\nabla I_i(k), \nabla I_j(l))} \right)$$

where $I_i(k)$ and $\nabla I_i(k)$ means the patch (15 × 15 window in our experiments) centered at landmark point $k$ in $I_i$ and its gradient space $\nabla I_i$, respectively, and $I_j(l)$ and $\nabla I_j(l)$ have similar definitions. $\gamma(\cdot, \cdot)$ denotes the normalized cross correlation [33]. $\rho(\cdot)$ represents a robust function and we define it as the Geman-McClure function

$$\rho(t) = \frac{t^2}{1 + t^2/a^2}$$

$$E_{i,j} \mathbf{1} = \mathbf{1},$$
where $a$ is a parameter in charge of adjusting the stringency of rejecting outliers and can be determined by the given image as explained in [34]. In Eq. (1), $w \in [0, 1]$ adjusts the contributions of the two terms defined respectively on image intensity and image gradient and its value can be empirically specified (e.g. $w = 0.5$ as used in all our experiments).

3.3. Motion model

We choose the 12-parameter inter-image transformation model in [35] as our parametric motion model for representing the spatial relation between MSI images. It was derived by modeling the retina as a rigid quadratic surface and has been validated to outperform other popular motion models (e.g. rigid and affine) for retinal image analysis. With this model, the $k$th row of $\chi_i$ in Eq. (2) corresponds to the $k$th landmark point of $I_i$ (with coordinate $[x_{ik} \ y_{ik}]$) and is written as a horizontal vector

$$\chi_i(j,:) = \begin{bmatrix} x_{ik}^2 & x_{ik} y_{ik} & y_{ik}^2 & x_{ik} & y_{ik} & 1 \end{bmatrix}$$

(7)

where $\chi_i(j,:)$ means the $j$th row of $\chi_i$. With the above definition of $\chi_i$, $\Theta_{i,j}$ in Eq. (2) is a $6 \times 2$ matrix for which the two columns correspond to motion parameters regarding $X_j$’s coordinates $x$ and $y$, respectively.

3.4. Optimization

In order to resolve the unknowns $\{E_{i,j}\}$ and $\{\Theta_{i,j}\}$ in our groupwise registration scheme, we need to optimize Eq. (1) based on the constraints in Eqs. (2)-(4). However, this optimization problem is very difficult because of at least three challenges. The first one lies in the requirement of binary value on $\{E_{i,j}\}$, which makes the optimization problem to be computationally intractable. The second one results from how to enforce motion’s temporal smoothness. The final one is attributed to the simultaneous optimization of $\{E_{i,j}\}$ and $\{\Theta_{i,j}\}$. We will below provide our solutions to these challenges.

3.4.1. Penalizing to-centroid spatial deviation

As shown in our previous work [36, 37], the hardship caused by the binary matrix $\{E_{i,j}\}$ can be circumvented with a minimization of the so called “to-centroid spatial deviation” if we relax the requirement in a fashion that the matrix’s elements can take a continuous value within $[0 \ 1]$. This relaxation can be mathematically expressed as the combination of Eq. (4) with the following equation

$$E_{i,j} \geq 0.$$  

(8)

This relaxation allows more than one landmark points in $I_j$ matched to each of the point in $I_i$ and is equivalent to the widely used “soft-assignment” strategy [38]. However, it may cause ambiguities in the matching process because of the non-sparse correspondence matrix it may produce. To deal with this, we combine the strategy of minimizing the to-centroid spatial deviation which measures the spatial concentration of the matched landmark points.

We first define

$$\bar{X}_{i,j} = E_{i,j}X_j$$

(9)

for which each row of $\bar{X}_{i,j}$ represents the coordinates of the centroid of the matched points in $I_j$ to the corresponding point in $I_i$. The to-centroid spatial deviation matrix is then written as

$$D_{i,j} = \sqrt{\left(\bar{X}_{i,j}(\cdot,1)1^T - 1(X_j(\cdot,1))^T\right)^2 + \left(\bar{X}_{i,j}(\cdot,2)1^T - 1(X_j(\cdot,2))^T\right)^2}$$

(10)

where the size of $D_{i,j}$ is $N_j \times N_i$ guaranteed by the length of $1$. In Eq. (10), $\bar{X}_{i,j}(\cdot,1)$ and $\bar{X}_{i,j}(\cdot,2)$ represent the first and second column of $\bar{X}_{i,j}$, respectively. $X_j(\cdot,1)$ and $X_j(\cdot,2)$ are defined similarly.
Penalizing the to-centroid spatial deviation when aligning \( I_j \) to \( I_i \) amounts to minimizing the summation over all deviations, which is expressed as

\[
\min \, \text{tr}\left( D^T_{i,j} E_{i,j} \right). \tag{11}
\]

Regarding our groupwise registration task, it is written as

\[
\min \sum_{i=1}^{N} \sum_{j=1}^{N} \text{tr}\left( D^T_{i,j} E_{i,j} \right). \tag{12}
\]

Eq. (1) in our optimization problem becomes

\[
\min \sum_{i=1}^{N} \sum_{j=1}^{N} (1 - \lambda) \text{tr}\left( C^T_{i,j} E_{i,j} \right) + \lambda \text{tr}\left( D^T_{i,j} E_{i,j} \right) = \min \sum_{i=1}^{N} \sum_{j=1}^{N} \text{tr}\left( (1 - \lambda) C^T_{i,j} E_{i,j} + \lambda D^T_{i,j} E_{i,j} \right) \tag{13}
\]

where \( \lambda \) balances the contributions of the two terms.

For the proposed optimization scheme as shown in Eqs. (13), (8), (4) and (2), image alignment is carried out within each pair of MSI images, and in this process, different pairs are processed independently.

### 3.4.2. Enforcing temporal smoothness on motion

As described in Sec. 1, MSI generates a set of spectral slices sequentially within a limited time period of image acquisition. Therefore, eye motion changes smoothly during the process of image acquisition. In other words, we should assume that eye motions are similar between temporally neighboring slices. This is in contrast to the unpredictable relations between motions in generic groupwise image registration problems, e.g. brain image registration [12] for which images are maybe acquired at different patient visits.

To enforce this temporal consistency of motion, we penalize \( \Theta_s \) of the temporally neighboring pairs and obtain the following minimization function

\[
\min \left( \sum_{i=1}^{N} \sum_{j=1}^{N} \text{tr}\left( (1 - \lambda) C^T_{i,j} E_{i,j} + \lambda D^T_{i,j} E_{i,j} \right) + \tau \sum_{k=2}^{N-1} \text{tr}\left( \left( \Theta_{k-1,k} - \Theta_{k,k+1} \right)^T \left( \Theta_{k-1,k} - \Theta_{k,k+1} \right) \right) \right) \tag{14}
\]

### 3.4.3. Formulating optimization as quadratic programming

As we discovered, the minimization in Eq. (14) constrained by the functions in Eqs. (8), (4) and (2) can be resolved via a quadratic programming (QP) [14]. We first vectorize the matrices \( E, C \) and \( D \) and denote the resulted vertical vectors as \( \mathbf{e}, \mathbf{c}, \) and \( \mathbf{d} \), respectively. We also write \( \theta = \begin{bmatrix} \Theta(:,1) \\ \Theta(:,2) \end{bmatrix} \). We then reformulate Eq. (4) as

\[
P_{e_{i,j}} \mathbf{e}_{i,j} = 1 \tag{15}
\]

where \( P_{e_{i,j}} \) is a binary matrix in size of \( N_i \times (N_i \cdot N_j) \), satisfying \( P_{e_{i,j}} \mathbf{e}_{i,j} = E_{i,j} \mathbf{1} \).

It is straightforward to create a matrix \( Q^1_{i,j} \) in size of \( N_i \times (N_i \cdot N_j) \) by leveraging \( X_f(:,1) \), such that

\[
E_{i,j} X_f(:,1) = Q^1_{i,j} \mathbf{e}_{i,j} \tag{16}
\]

Similarly, we can take advantage of \( X_f(:,2) \) to form \( Q^2_{i,j} \) and we have

\[
E_{i,j} X_f(:,2) = Q^2_{i,j} \mathbf{e}_{i,j} \tag{17}
\]
We also re-write Eq. (9) by replacing $X_j$ in Eq. (2) with $\bar{X}_{i,j}$ as

$$\chi_i \bar{\Theta}_{i,j} = E_{i,j} X_j.$$  

(18)

Combining Eqs. (17) and (18), we have

$$\begin{bmatrix} \chi_i & 0 \\ 0 & \chi_i \end{bmatrix} \begin{bmatrix} \theta_{i,j} \end{bmatrix} = \begin{bmatrix} Q^1_{i,j} \\ Q^2_{i,j} \end{bmatrix} e_{i,j}.$$  

(19)

Our optimization problem is finally expressed as the following QP

$$\min \sum_{i=1}^{N} \sum_{j=1}^{N} \left[ (1 - \lambda) c_{i,j} + \lambda d_{i,j} \right]^T \begin{bmatrix} e_{i,j} \\ \theta_{i,j} \end{bmatrix} + \tau \min_{k=2}^{N-1} \| \theta_{k-1,k} - \theta_{k,k+1} \|_2^2$$

(20)

For Eq. (20), $\forall (i, j) : 1 \leq i \leq N, 1 \leq j \leq N$, we have the following constraints

$$\begin{bmatrix} P_{e_{i,j}} & 0 \\ -Q_{i,j} & \chi_i \\ -Q_{i,j} & \chi_i \end{bmatrix} \begin{bmatrix} e_{i,j} \\ \theta_{i,j} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

(21)

and

$$-e_{i,j} \leq 0.$$  

(22)

In this quadratic programming formulation, $e_{i,j}$ and $\theta_{i,j}$ are unknowns which are solved simultaneously.

3.4.4. Algorithm overview

Our algorithm is summarized as the following five steps:

1. Solving the QP problem specified by Eqs. (20)-(22) without considering the to-centroid deviations (i.e. set $\lambda = 0$ in Eq. (20)).

2. Updating the to-centroid deviation matrix using Eq. (10) with current estimation of the correspondence matrix.

3. Increasing $\lambda$ by $\delta \lambda$ (e.g. 0.05).

4. Solving the QP problem specified by Eqs. (20)-(22) with considering the to-centroid deviation matrix.

5. Go to step 2 until the stopping criterion is satisfied.

Our algorithm is unique in group-wisely registering sequential MSI images because of at least six of its properties. First, correspondences of landmark points and the pre-defined transformation model are estimated simultaneously in step 1 and step 4. Second, the transformation model can help to guide the estimation of correspondences and resist noise/outliers better. Third, the motions estimated by our algorithm are guaranteed to change smoothly. Fourth, gradual increase of $\lambda$ in step 4 approaches a binary correspondence matrix with a less ambiguity in the assignment. Fifth, step 1 can offer a very nice initialization without any manual intervention. Finally, step 1 and step 4 can both result in an optimal solution based on the QP technique.

In our groupwise registration scheme, the entire set of images are assessed via a series of pairwise registrations which are connected by the constraints on motion’s temporal smoothness. With this strategy, information from pair-wise registration is propagated globally across the entire set of images being registered, resulting in an integration of available information in a meaningful manner which leads to a globally good solution by virtue of local constraints [21].
4. Experiment

4.1. Data

Our data set for validating the proposed algorithm comprises 22 sequences of MSI images acquired by using an Annidis RHA™ instrument (Annidis Health Systems Corp., Ottawa, Canada). These images were of oculus dexter (OD) and oculus sinister (OS) from 8 healthy subjects and 3 patients which were diagnosed with hypertensive retinopathy. These images are grayscale in a bit depth of 16, provided in format of dicom. All images are in size of 2048 × 2048. Each sequence has at least 8 images captured with wavelengths of amber, green, infrared, red and yellow. An ophthalmologist manually picked 15 points for each sequence and marked them in all the MSI images of the sequence by using MRcron [39]. These manual marks are treated as the ground truth for validating our algorithm and their locations were compared with the ones transformed by using the estimated motion model in our experiments.

4.2. Implementation details

We stopped our algorithm after the steps 2-5 of our algorithm are repeated for 3 times, and from our experiments, we found that an accurate estimation can be resulted for most cases. We performed a set of experiments to empirically choose an optimal value for the parameter $\tau$ in Eq. (20). As a postprocessing, each floating-landmark is finally assigned to the reference-landmark with the largest correspondence likelihood value. We used the the open-source Computational Geometry Algorithms Library (CGAL) (downloaded from http://www.cgal.org/) for solving the QPs. We used a Dell PC with 2.39 GHz Intel Core 2 CPU and found that the optimization process can be finished within 19 seconds to group-wisely match 10 MSI images, for each of which about 200 landmark points are detected.

4.3. Evaluation

In our experiments, for each of the manually marked points and each pair of MSI images in a sequence, we computed points’ distance (in pixels) to their location transformed by the estimated motion model. The final error measurement was treated as the the averaged distance value over all manually marked points and all image pairs.

We first ran our algorithm on our data set by taking advantage of various cost functions including sum of squared difference (SSD), mutual information (MI) and our proposed robust cross correlation (RCC). At the same time, we changed $\tau$’s value in Eq. (20). From the results shown in Fig. 3, we can see that our cost function outperforms SSD and MI obviously. Moreover, our algorithm performs best when $\tau = 0.2$, which will be used in our following experiments. Comparing the best resolutions at $\tau = 0.2$ with the ones at $\tau = 0$, we can also see that the enforcement of temporal smoothness on the estimated motions enables our algorithm to perform better than without this constraint.

We then evaluated our algorithm by comparing its performance with state-of-the-art image registration algorithms. We chose the EM like technique of Chui & Rangarajan [40], which performs as repeating the estimations of correspondences and motion model in contrast to our simultaneous estimation strategy. In order for comparison, we only employed the EM strategy for optimization and kept other components consistent with our algorithm, e.g. the landmark points, the matching cost and the motion model. We also chose the groupwise registration approach of Wells et al. [12] by incorporating the mutual information measurement [41] as the matching cost function. This technique resolves a pixel-wise registration by estimating an affine motion model followed by a free-form B-spline deformation model based on a gradient descent optimization strategy. For the B-spline, we set the distance between control points as 30 pixels. We compared the errors obtained from the healthy subjects, the patients and the complete data set, respectively.
From the results in Table 1, we can see that our algorithm generated accuracies superior significantly to the algorithms of Chui & Rangarajan and Wells et al. From this, we have validated several of our claims mentioned before. First, our strategy of simultaneous estimation on correspondence and motion model via quadratic programming is better than the EM technique which iteratively estimates these two unknowns. Second, our algorithm which is designed specifically for registration of MSI images outperforms generic image registration approaches (e.g. the well-known method of Wells et al.) possibly because of the integration of our motion model, cost function, temporal smoothness constraint on motion and global optimization technique.

From Fig. 4 and Fig. 5, similar findings as mentioned above can be concluded as well. Moreover, we can see from Fig. 5 that image variations between the pair of images, which were caused by the retinal degenerations, modality difference and inhomogeneous light, degraded significantly the method of Chui & Rangarajan. For this case, the method of Wells et al. even failed. In contrast, our algorithm produced a very promising result.

5. Conclusion and future work

In this paper, we provide a solution to an unresolved but important problem of aligning a sequence of images from multispectral imaging (MSI) of the retina and choroid in an automatic and joint fashion. We have four significant technical contributions. First, we leverage the quadratic programming (QP) to simultaneously estimate landmark correspondences and a predefined motion model, which is validated to help produce better accuracies due to the optimal optimization of QP and the mutual benefits of correspondences and motion model during the optimization process. Second, we enforce a temporal smoothness on the estimated motion,
resulting in a better consistency with the practical MSI acquisition. Third, our groupwise registration algorithm is able to achieve an optimization globally over the entire image set, which is accomplished by propagating information obtained from local pairwise image registration. Finally, the confounding variations caused by anatomical/pathological differences or illumination inhomogeneity in MSI images are handled by matching key points only and defining a robust matching cost function. Our algorithm is potentially invaluable in clinical practices of ophthalmologists and being able to extend to other multi-model/multi-spectral images.

Our future work would include not only further validations with more MSI images but also various clinical studies for examining various eye diseases by using our algorithm. In addition, we are interested in research on detection of landmark points specifically for retina in MSI images in order for better robustness and accuracy.

**Funding**

Natural Science Foundation of China (NSFC) (61572300); Natural Science Foundation of Shandong Province in China (ZR2014FM001); Taishan Scholar Program of Shandong Province in China (TSHW201502038); National Institutes of Health (NIH) (P30 EY001583).
Fig. 5. Registration results from a pair of MSI images of a patient diagnosed with hypertensive retinopathy. Left to right and top to down: MSI-550, MSI-850, overlapping MSI-850 to MSI-550, overlapping MSI-850 to MSI-550 transformed by the method of Chui & Rangarajan [40], Wells et al. [12] and ours, respectively.