Automated Registration of Multimodal Optic Disc Images: Clinical Assessment of Alignment Accuracy

Wai Siene Ng, FRCOphth, Phil Legg, PhD,†*
Venkat Avadhanam, FRCOphth, Kyaw Aye, MRCS,**
Steffan H. P. Evans, MBBCh,‡ Rachel V. North, PhD,§
Andrew D. Marshall, PhD,† Paul Rosin, PhD,†
and James E. Morgan, FRCOphth§

Purpose: To determine the accuracy of automated alignment algorithms for the registration of optic disc images obtained by 2 different modalities: fundus photography and scanning laser tomography.

Materials and Methods: Images obtained with the Heidelberg Retina Tomograph II and paired photographic optic disc images of 135 eyes were analyzed. Three state-of-the-art automated registration techniques Regional Mutual Information, rigid Feature Neighbourhood Mutual Information (FNMI), and nonrigid FNMI (NRFNMI) were used to align these image pairs. Alignment of each composite picture was assessed on a 5-point grading scale: "Fail" (no alignment of vessels with no vessel contact), "Weak" (vessels have slight contact), "Good" (vessels with $<50\%$ contact), "Very Good" (vessels with $>50\%$ contact), and "Excellent" (complete alignment). Custom software generated an image mosaic in which the modalities were interleaved as a series of alternate 5×5-pixel blocks. These were graded independently by 3 clinically experienced observers.

Results: A total of 810 image pairs were assessed. All 3 registration techniques achieved a score of "Good" or better in $>95\%$ of the image sets. NRFNMI had the highest percentage of "Excellent" (mean: 99.6%; range, 95.2% to 99.6%), followed by Regional Mutual Information (mean: 81.6%; range, 86.3% to 78.5%) and FNMI (mean: 73.1%; range, 85.2% to 54.4%).

Conclusions: Automated registration of optic disc images by different modalities is a feasible option for clinical application. All 3 methods provided useful levels of alignment, but the NRFNMI technique consistently outperformed the others and is recommended as a practical approach to the automated registration of multimodal disc images.

Key Words: multimodal registration, glaucoma, imaging, optic nerve, diagnostic tests/investigation

(*J Glaucoma* 2015;00:000–000)

The digitization of fundal imaging has transformed the clinical assessment of retinal disease. Although the digitization of retinal photography has been a major step forward, some of the most exciting developments have come from the introduction of laser-based imaging modalities which have provided clinically relevant information on retinal structure. Scanning laser ophthalmoscopy (SLO) was first to generate high-resolution topographic plots of the retinal surface.[1–3] More recently, optical coherence tomography (OCT) has entered routine use for the quantitative assessment of retinal structure at near cellular resolutions.[4–6] Other modalities such as scanning laser polarimetry have provided complimentary information related to the thickness of the retinal nerve fiber layer[7–10]

Despite these advances, the clinical assessment of retinal disease still relies on indirect ophthalmoscopy and in digital form, fundus photography. Clinicians are therefore presented with multimodal views of the retina for which the derivation of a combined image data set can be problematic. The registration of images obtained by different modalities would appear to offer considerable advantages[11] and is routinely used in other areas of clinical imaging such as computed toography-magnetic resonance imaging. For example, photographic views of the optic disc can provide clear views of the optic disc margin which are important for the derivation of optic disc parameters in the diagnosis of glaucoma. Similarly, retinal drusen may be more apparent in photographic than in coherence-based views.[12,13] The use of a common photographic platform for image registration would also be beneficial for the follow-up of patients. There is, therefore, a recognized need for efficient and accurate image registration to combine information from images acquired using different modalities.

Several investigators have reported methods for the registration of images.[1,14] Best results are generated when the observer selects retinal landmarks (fiduciary points) that are on the basis of known anatomic features, judged to arise from the same points in the retina.[1,3] Unfortunately, fully automated registration can result in lower levels of alignment accuracy, which might limit the use of these algorithms in the clinical setting, particularly when image clarity might be an issue.

As the automated registration of retinal images generated by differing modalities would facilitate clinical implementation, we have explored novel methods for the automated registration of these images. As a test data set that is representative of the types of images to be registered, we have taken extended focus images derived from the Heidelberg Retina Tomography II (HRT-II) and aligned these to digital optic disc photographs of the same field of view. On the basis of preliminary work, we selected 3 algorithms for the automated alignment of the image pairs. All were based on a mutual information (MI)

Received for publication June 10, 2014; accepted February 15, 2015.
From the *University Hospital Wales; †School of Computer Sciences and Informatics; §Cardiff Centre for Vision Sciences, Cardiff University, Cardiff; and ‡Royal Glamorgan Hospital, Llantrisant, Wales, UK.
Disclosure: The authors declare no conflict of interest.
Reprints: James E. Morgan, FRCOphth, Cardiff Centre for Vision Sciences, Cardiff University, Maindy Road, Cardiff, Wales CF24 4HQ, UK (e-mail: morganje3@cardiff.ac.uk).

DOI: 10.1097/IJG.0000000000000252

approach[4,6] which is widely used to register images of biological structures of different modalities. It is an entropy-based measure that maximizes the interdependence of the 2 images under consideration.

Although MI methods have been used successfully, we hypothesized that the accuracy of alignment could be further improved by the inclusion of the spatial distribution of neighboring pixels and intensity mapping [Feature Neighbourhood Mutual Information (FNMI)]. FNMI algorithms were applied in 2 forms. In the first, the images maintained their respective spatial arrangements and were not deformed (rigid FNMI). We then used a recently developed nonrigid FNMI method (NRFNMI) in which registration was enhanced locally using a local MI window. We report the performance of both the rigid and NRFNMI algorithms as well as the widely used MI approach in the registration of optic disc photographs with HRT scans.

MATERIALS AND METHODS

Registration was performed on 135 matching image pairs of color fundus photographs and SLO images centered on the optic disc. The images were obtained during the course of routine glaucoma care at the University Hospital of Wales with all patients providing written consent for images to be used for subsequent analysis and publication. The image data set included a range of normal, early glaucomatous, and advanced glaucoma subjects as determined by an experienced ophthalmologist. Fundus photographs were acquired using a TOPCON TRC50-EX with a Nikon digital camera providing images with a resolution of 3008×1960 pixels (24 bits per pixel). SLO images were captured with a HRT-II (Heidelberg Retina Engineering). The field of view for the SLO images subtended at 15×15 degrees and comprised of 384×384 pixels. As the HRT-II takes images at 32 focal depths, the intensity value at each pixel location was summed to generate a single (normalized) intensity value (8 bits per pixel). Images were viewed with a bitmap palette that approximated the color of the retina. Images were procured by an experienced medical photographer where only clear images of good quality were included and images with blur were excluded. All image pairs; fundus photograph and SLO image, were acquired at the same sitting.

For the purposes of grading the quality of the alignment of these image pairs, they were presented as a static image in which the registration results were presented using a mosaic formation (5×5 grid of pixels for the area covered by the HRT image) to show the overlap between the 2 images (Fig. 1C). In the first set, the pixel block at X[0] Y[0] represented the HRT image and in the second set, this pixel block started with the fundus photograph image. Each image was therefore graded twice but with alternating areas of each image used for comparison thereby generating 270 images to test each algorithm. The following registration techniques were evaluated: FNMI, NRFNMI, and Regional Mutual Information (RMI). The registration time was reduced by presenting the fundus images as a 564×367-pixel array and the HRT image as a 288×288-pixel array. The fundus photograph was used as the reference image with the location of the HRT image adjusted to provide the registration.

Image Grading

The grading system is illustrated in Figure 1. The alignment was deemed as "Fail" if vessels with the same point of origin had no point of contact. Alignment was "Weak" if the vessels contacted but overlapped $< 20\%$. Above this (up to 50% overlap) the alignment was judged to be "Good." With $> 50\%$ of overlap, the alignment was judged to be "Very Good." When the vessels showed complete alignment (no vessel border discontinuity), the overall quality of the alignment was judged as "Excellent." An example of an image pair with "Very Good" alignment is shown in Figure 1C. Assessors were masked to the technique used for alignment and images were presented in a random sequence, which were assessed at a single sitting.

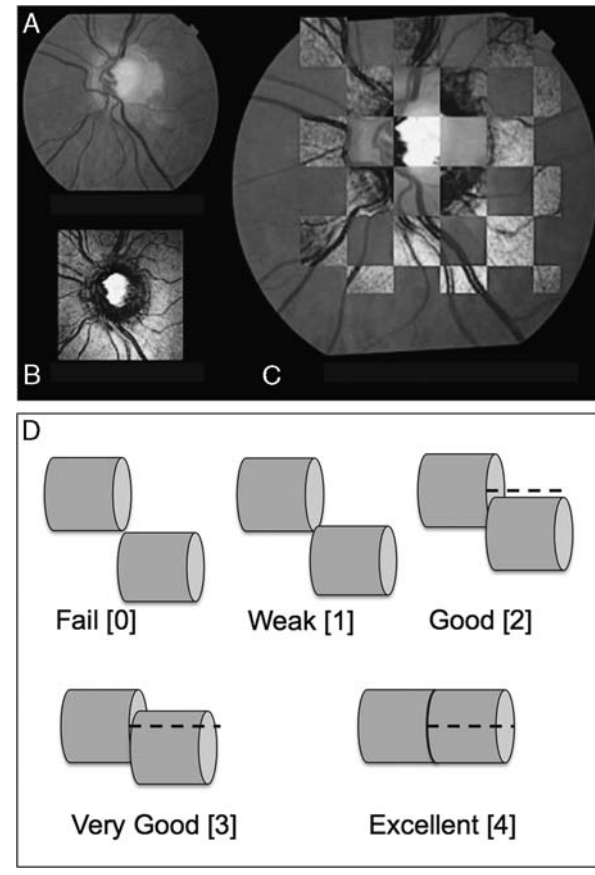

FIGURE 1. A, Fundus photograph of optic nerve head. B, Corresponding scanning laser ophthalmoscopy image of the same optic nerve head. C, Chequer-board arrangement used to grade alignment. The chequer-board pattern was counter phased so that each image pair was assessed on the base of 2 images. D, Illustration of alignment scores. "Fail" [0] = no point of contact, "Weak" [1] = vessels contacted but overlap <20%, "Good" [2] = up to 50% vessel overlap, "Very Good" [3] = > 50% vessel overlap, "Excellent" [4] = complete vessel alignment.

Image Alignment Techniques

The transformation space is considered to be the full space that captures some degree of overlap between the 2 images. The registration process is typically initialized with the images centered on each other, and the optimization algorithm performs an optimal search to find the transformation that maximizes alignment measure; there is no restriction in terms of practical limitations of how the images are collected as no assumptions are made within the alignment process.

RMI extends the standard MI method by incorporating at each pixel the additional intensities of all pixels within a fixed-size neighborhood. To avoid the enormous joint histograms that would arise when using the standard MI approach, the neighborhood was described by its covariance matrix, from which it is possible to efficiently compute MI.[15] Secondly, a registration search for transformation space was performed by incorporating a multiresolution image pyramid. The translational search was optimized using the Nelder-Mead simplex algorithm.[10]

As with RMI, rigid FNMI alignment[7,9] computes an MI score that incorporates more information than just individual pixels from the image pair. This was achieved by including the gradient magnitudes of each image that were calculated at multiple spatial scales to detect structural and spatial aspects of the images.

The aim of the NRFNMI is to refine the results of rigid FNMI by allowing for some deformation of the image. Within the area of interest, the 2 images to be registered were subdivided into a set of 4×4 windows. The centre points from these 16 windows defined our initial control points for performing the nonrigid deformation. Each window was subdivided into another set of 4×4 windows, and a normalized MI score computed for each of these smaller windows. The overall MI score was given by the sum of individual smaller window MI scores. The translational search of each subwindow was limited to a 3-pixel radius in the central 4 windows and a 5-pixel radius in the outer 12 windows to prevent window misplacement. Finally, deformation of the SLO image was performed by taking the central points of the large 16 windows as controls. A thin-plate spline warp algorithm was applied to determine the amount of deformation.

Statistical Analysis

For simplicity, the scale of registration assessment was treated as linear. The linear mixed-effects model was therefore employed in which patient image and assessors were treated as random effects. Registration modes and disc grade of disease were treated as fixed effects in the model. Block replicates of registered images were inserted into the model as an additional fixed effect. P-value of < 0.05 is taken as significant. Statistical analysis was carried out with IBM SPSS Statistics, version 20 software.

P-value of < 0.05 is taken as significant. Statistical analysis was carried out with IBM SPSS Statistics, version 20 software.

RESULTS

A total of 2430 evaluations were recorded as a result of 135 unique image pairs multiplied by 2 as block replicates, by 3 different alignment techniques (RMI, NRFNMI, and FNMI) and by 3 different assessors. All images were judged to be of suitable quality for the purposes of grading. These comprised of 45(33.3%) normal discs, 65(48.1%) early glaucomatous discs, and 25 (18.5%) advanced glaucomatous discs. On general observation, the assessors did not notice any systematic bias in images for regions that were consistently misaligned although this was not formally assessed in this study. An image grade was assigned on the basis of an area of misalignment within the region of interest; as shown in Figure 1C, the overall level of grading would be "Very Good" due to vessel border misalignments seen on the inferotemporal arcade vessels. The distributions of alignment scores for the 3 observers are shown for the 3 methods of registration in Figures 2 to 4. The proportion of alignment graded as failing was $< 1\%$ (mean value across observers) for all registration methods. All 3 techniques had $> 96\%$ of the images marked as "Good" or better by all the 3 independent observers. The proportion of alignments scoring "Excellent" was highest for the NRFNMI at 96.8% (range, 99.6% to 95.2%); higher than FNMI and RMI. This was followed by RMI at 81.6% (range, 86.3% to 78.5%) and FNMI at 73.1% (range, 85.2% to 54.5%) (Figure 5). NRFNMI also had the highest percentage of "Good" or better grading comprising 99.8% of its total score. Differences between modes of registrations and image grading were statistically significant ($P = 0.00$).

Block replicates and disease severity did not have a statistically significant effect on grading of images ($P = 0.181$, 0.972). Figures 6 to 8 demonstrate the high levels of "Excellent" grading throughout all subsets of disc disease severity. The estimate of covariance between the assessors was very low (0.007) indicating low levels of disagreement.

DISCUSSION

Ophthalmology is increasingly reliant on the integration of digital information from a range of modalities and devices. In the management of glaucoma, these device confer the greatest clinical power when used in combination.[2,16–18] Further diagnostic accuracy is also possible when consideration is given to the registration of visual field test locations and retinal structure.[19] Our study advances this by demonstrating the feasibility of registering retinal image data sets from 2 commonly used imaging modalities. The NRFNMI algorithm generated "Excellent" levels of alignment in 96% of cases across all observers which is likely to be sufficient for translation to clinical practice. Importantly, these alignments were fully automated and did not require users to mark fiduciary points to initiate the process.

Multimodal alignment technologies address an important need in ophthalmology to develop a common platform for the interrogation of retinal images. Fundus photographs remain the mainstay for the assessment of retinal diseases that may outlast the development cycle of

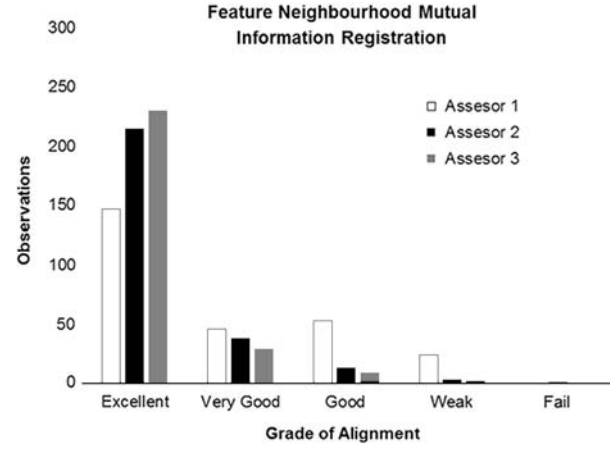

FIGURE 2. Bar chart of alignment scores from the 3 observers for the Feature Neighbourhood Mutual Information (FNMI) registration method.

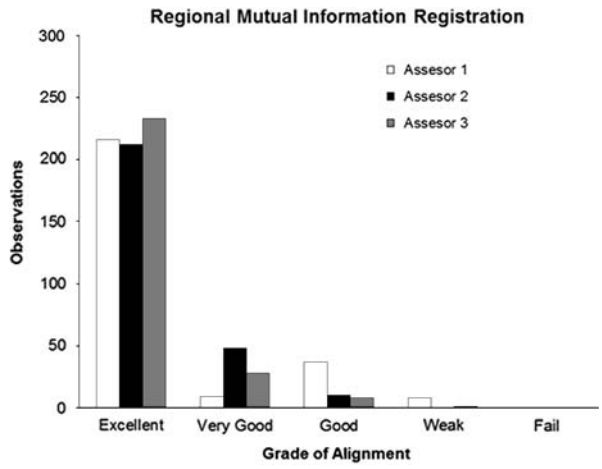

FIGURE 3. Bar chart of alignment scores from the 3 observers for the Regional Mutual Information (RMI) registration method.

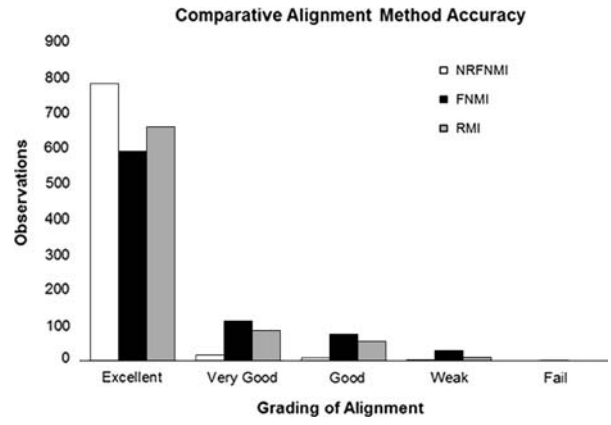

FIGURE 5. Bar chart of comparative alignment accuracy scores between rigid Feature Neighbourhood Mutual Information (FNMI), Regional Mutual Information (RMI), and nonrigid Feature Neighbourhood Mutual Information (NRFNMI) registration methods.

many proprietary imaging technologies. Automated multimodal registration can provide complimentary views for the demarcation of regions of interest (eg, demarcation of optic disc boundaries) which can be problematic in scanning laser images. The detection of other optic disc features indicative of disease such as retinal nerve fiber layer hemorrhages is improved with the inclusion of fundus photographs.[5,20]

Our data confirm the importance of including local spatial information in the design of a retinal image alignment algorithm. The reflectivity distribution needs to be taken into consideration in a fundus and tomographic image. For example, the neuroretinal rim, which overlies the highly reflective lamina cribrosa corresponds to a region of high reflectivity. By contrast, the signal in a confocal image is low as the angle between the incidence laser beam and the retinal surface undergoes considerable changes at the margin of the optic disc resulting in light scatter.[8,21] In view of these considerations, the relatively lower score with the MI method is not surprising as it uses pixel intensities rather than extracting structural features between images. Only individual pixel intensities are considered in this method but no spatial information is taken into account. Although MI is widely used in the field of medicine for registering images of different modalities, improvements to MI have been proposed. For example, with gradient MI, the standard MI is computed and is then multiplied by a gradient term.[22] A higher order MI proposed by Rueckert et al[23] calculates entropy for intensity pairs rather than individual intensities.

All registration algorithms are vulnerable to the effects of local maxima in the registration surface, which can adversely affect search optimization. Inclusion of the radius of the neighborhood pixels improved the accuracy of registration with RMI, but also increased the matrix size and computational time. FNMI, a novel similarity measure technique developed by Legg and colleagues[7,11,9,12,13] incorporates multiscale feature derivatives along with spatial neighborhood knowledge into the MI framework. The FNMI technique is advantageous in that excellent registration is provided for even blurred images, a drawback

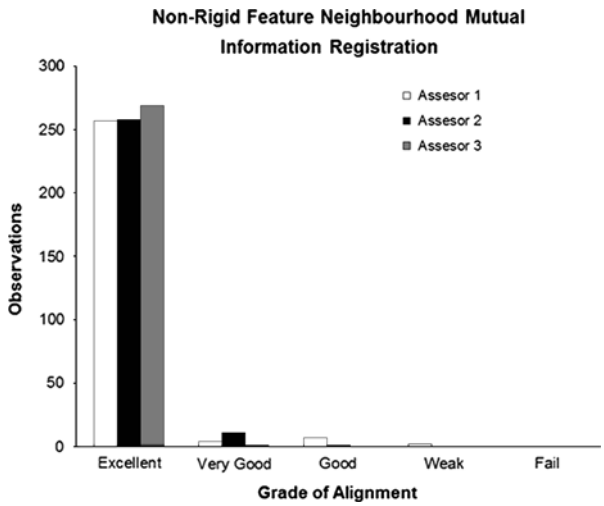

FIGURE 4. Bar chart of alignment scores from the 3 observers for the nonrigid Feature Neighbourhood Mutual Information (NRFMI) registration method.

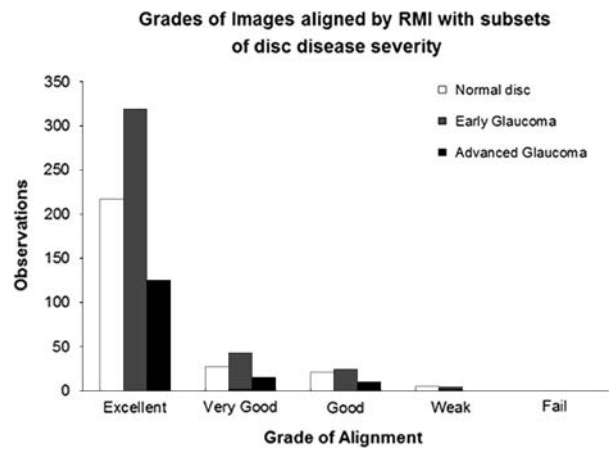

FIGURE 6. Bar chart of alignment scores from Regional Mutual Information (RMI) with subsets of images with disc disease severity.

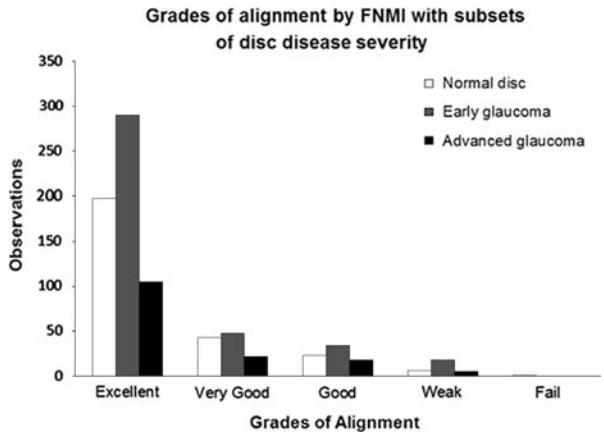

FIGURE 7. Bar chart of alignment scores from rigid Feature Neighbourhood Mutual Information (FNMI) with subsets of images with disc disease severity.

which often arises in clinical practice in the presence of cataract, high degrees of astigmatism, or movement.

Although rigid FNMI provides accurate registration for most of the image pairs, subtle misalignments were apparent in certain sets. Differences in the acquisition techniques of SLO and fundus photography coupled with the alterations in eye position during image acquisition can lead to subtle elastic deformations in the images that will become manifest during rigid registration. Kubecka and Jan[14] recommended a generalized elastic registration for the registration of fundus and SLO images as there can be some deformation between the 2 images.[1]

The alignment algorithms could potentially be used for tracking purposes over time, although this was not necessarily the focus of this work. The clinician can control the amount of deformation that is introduced into the registration; therefore this could be restricted to a $+1$ pixel range for each window. The clinician can also choose to compare against the rigid FNMI result. Each of the algorithms is designed to register the images where there is the greatest similarity shared between the images, however, in a case where over time the structure could well change significantly, then the clinician would most likely choose to make a judgment to avoid overfitting of the deformation model, hence reducing the chances of masking true anatomic change from disease progression for the sake of better alignment.

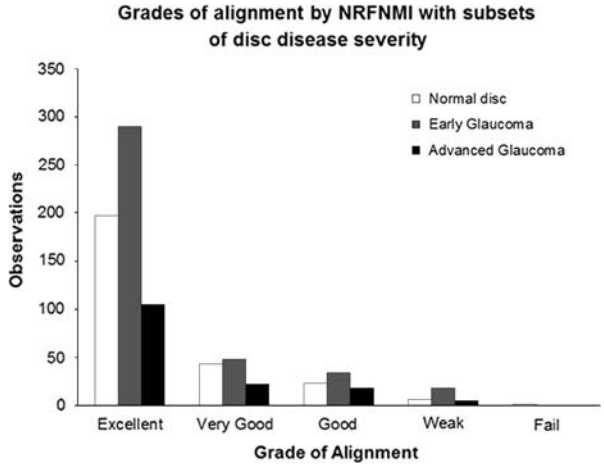

FIGURE 8. Bar chart of alignment scores from nonrigid Feature Neighbourhood Mutual Information (NRFNMI) with subsets of images with disc disease severity.

All 3 techniques perform relatively fast, and within a timely manner for practical clinical use, yet we recognize that the runtime could be improved further with optimized production code. For instance, in our paper, Legg et al[24] described that RMI computational times varies between 14.02 and 96.83 seconds depending on r. The mean FNMI time was 119.18 seconds, with NFRNMI taking an additional 30 seconds. Overall, registration for all algorithms was complete within 2 to 3 minutes, with NRFNMI giving the best registration accuracy. It is important to note that these intervals were based on debug (test) code rather than a fully optimized development code.

With regards to the robustness of the registration techniques, the search optimization scheme is used to avoid the algorithm becoming trapped by local maxima. The offset of an initialization point does not greatly impact on the registration runtime, as the search optimization scheme will take larger steps toward the maximal solution, which is first performed at a coarse level of the image pyramid. Having found the registration at the coarse level, the algorithm then refines the registration at finer resolution levels, until the result is derived for the full resolution. To enhance robustness, multiple initialization points could be utilized to avoid the possibilities of exceeding a maximum offset in terms of translation or rotation. These approaches are designed to limit dependency of requirements at the acquisition stage, as the algorithms were tested on a variety of image pairs captured under different conditions by the operating clinician. We also conducted registration convergence tests on a variety of registration approaches, including FNMI and RMI, to understand the "surface" of the registration function across the full image space. We found that FNMI provided the greatest coverage of convergence points, and also provides a clear incline toward the global maximum.[24]

In summary, NRFNMI-based alignment is a novel time-efficient method that requires a fraction of the data needed for RMI-based alignment and delivers excellent registration accuracy. Further studies are required to evaluate the clinical validity of this technique for registration of other image modalities such as stereophotography and OCT. We anticipate that these algorithms will be suited to the alignment of same modality images and therefore for the detection of change in images taken over time, an essential paradigm for clinical practice.

REFERENCES

1. Chrástek R, Skokan M, Kubecka L, et al. Multimodal retinal image registration for optic disk segmentation. *Methods Inf Med*. 2004;43:336–342.
2. Kamal D, Garway-Heath D, Hitchings R, et al. Use of sequential Heidelberg retina tomograph images to identify changes at the optic disc in ocular hypertensive patients at risk of developing glaucoma. *Br J Ophthalmol*. 2000;84:993–998.
3. Kolar R, Kubecka L, Jan J. Registration and fusion of the autofluorescent and infrared retinal images. *Int J Biomed Imaging*. 2008;2008:513478.
4. Paul V, William M. Wells III. Alignment by maximization of mutual information. *Int J Comput Vis*. 1997;24:137–154.
5. Drexler W, Fujimoto JG. State-of-the-art retinal optical coherence tomography. *Prog Retin Eye Res*. 2008;27:45–88.
6. Collignon A, Maes F, Delaere D, et al. Automated multimodality image registration based on information theory. In: Bizais Y, Barillot C, Di Paola R, eds. *Information Processing*

in Medical Imaging (Brest). Dordrecht (The Netherlands): Kluwer Academic; 1995:263274.
7. Legg PA, Rosin PL, Marshall D, et al. Incorporating neighbourhood feature derivatives with Mutual Information to improve accuracy of multi-modal image registration. *Proc Image Understand Anal*. 2008:39–43.
8. Boehm MD, Nedrud C, Greenfield DS, et al. Scanning laser polarimetry and detection of progression after optic disc hemorrhage in patients with glaucoma. *Arch Ophthalmol*. 2003; 121:189–194.
9. Legg PA, Rosin PL, Marshall D, et al. A robust solution to multi-modal image registration by combining mutual information with multi-scale derivatives. *Med Image Comput Comput Assist Interv*. 2009;12(Pt1):616–623.
10. Lagarias JC, Reeds JA, Wright MH, et al. Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM J Optim*. 1998;9:112–147.
11. Bernardes R, Lobo C, Cunha-Vaz JG. Multimodal macula mapping: a new approach to study diseases of the macula. *Surv Ophthalmol*. 2002;47:580–589.
12. Sohrab MA, Smith RT, Salehi-Had H, et al. Image registration and multimodal imaging of reticular pseudodrusen. *Invest Ophthalmol Vis Sci*. 2011;52:5743–5748.
13. Forte R, Querques G, Querques L, et al. Multimodal imaging of dry age-related macular degeneration. *Acta Ophthalmol*. 2012;90:e281–e287.
14. Kubecka L, Jan J. Registration of bimodal retinal images—improving modifications. *Conf Proc IEEE Eng Med Biol Soc*. 2004;3:1695–1698.
15. Russakoff DB, Tomasi C, Rohlfing T, et al. Image similarity using mutual information of regions. *ECCV*. 2004;3:596–607.
16. Sanchez-Galeana C, Bowd C, Blumenthal EZ. Using optical imaging summary data to detect glaucoma. *Ophthalmol*. 2001; 108:1812–1818.
17. Reus NJ, de Graaf M, Lemij HG. Accuracy of GDx VCC, HRT I, and clinical assessment of stereoscopic optic nerve head photographs for diagnosing glaucoma. *Br J Ophthalmol*. 2007;91:313–318.
18. Greenfield DS, Weinreb RN. Role of optic nerve imaging in glaucoma clinical practice and clinical trials. *Am J Ophthalmol*. 2008;145:598–603.
19. Asaoka R, Russell RA, Malik R, et al. A novel distribution of visual field test points to improve the correlation between structure-function measurements. *Invest Ophthalmol Vis Sci*. 2012;53:8396–8404.
20. Budde WM, Mardin CY, Jonas JB. Glaucomatous optic disc hemorrhages on confocal scanning laser tomographic images. *J Glaucoma*. 2003;12:470–474.
21. Bartsch DU, Freeman WR. Axial intensity distribution analysis of the human retina with a confocal scanning laser tomography. *Exp Eye Res*. 1994;58:161–173.
22. Pluim J, Maintz JA, Viergever MA. Interpolation artefacts in mutual information-based image registration. *Comput Vis Image Und*. 2000;77:211–232.
23. Rueckert D, Clarkson M, Hill D, et al. Non-rigid registration using higher-order mutual information. *Proc SPIE Medical Imaging 2000: Image Processing*. 2000:438–447.
24. Legg PA, Rosin PL, Marshall D, et al. Feature neighbourhood mutual information for multi-modal image registration: an application to eye fundus imaging. *Pattern Recog*. 2015;48: 1937–1946.

