Registration of Confocal Fluorescence Endomicroscopy Images Using Phase Correlation

Feng Zhao and T. M. McGinnity
Intelligent Systems Research Centre
University of Ulster, Magee Campus
Londonderry, UK, BT48 7JL

Abstract— *The emerging confocal fluorescence endomicroscope is capable of imaging living tissues in a non-invasive way using a probe to continuously scan the surface and subsurface tissue structures. Due to possible tissue movement and tissue expansion/contraction, the acquired images contain various noises and distortions. It is necessary to align these images in order to obtain a better 3D reconstruction of the tissue's microstructure for clinicians. In this paper, we present an automatic image registration method using the phase correlation technique, which uses a fast frequency-domain approach to estimate the relative transformation parameters between two consecutive endomicroscopy images.*

Keywords: Confocal Fluorescence Endomicroscope, FFT, Phase Correlation, Image Registration

1. Introduction

The confocal fluorescence endomicroscope (Fig. 1) is a newly developed endoscopic tool that makes it possible to carry out *in vivo* microscopic observations of living subjects with about 1000-time magnification and subcellular resolution [1], [2], [3]. An endomicroscope is more powerful than a microscope or an endoscope which generally needs biopsy and carries the risk of causing bleeding, infection, perforation, or mechanical agitation that may lead to the spread of tumor cells through the blood and lymphatic vessels [4]. In addition, a microscope or an endoscope can only see the surface layer without depth resolution.

As illustrated in Fig. 2, the endomicroscope operates in a non-invasive way. By placing the probe on the surface of the target subject, it enables direct observation of molecular mechanisms by continuously scanning the surface and subsurface tissue structures without removing tissues from the body or sacrificing animals. The fluorescence imaging parameters are optimised for a wide range of tissues including brain, intestine, lungs, colon, kidneys, muscle, heart, liver, and pancreas. Thus, molecular imaging of different types of tissues and diseases is becoming feasible, and thus has the potential to facilitate early diagnosis of cancers. Compared with other multi-million pounds imaging instruments such as magnetic resonance imaging (MRI) scanners [5], [6] and X-ray computed tomography (CT) scanners [7], [8], the endomicroscope is of much lower cost (£100K). While MRI and CT are widely used in disease diagnosis by acquiring global information from the scanned subjects, the new technique of endomicroscopy can provide complementary local information in detail for the clinicians, thus further improving disease diagnosis accuracy. This will provide profound health benefits to society. The technique will also enable further advancement in the field of basic cell biology, aid our understanding of the mechanism of disease progression, and allow monitoring of drug effects at the cellular level.

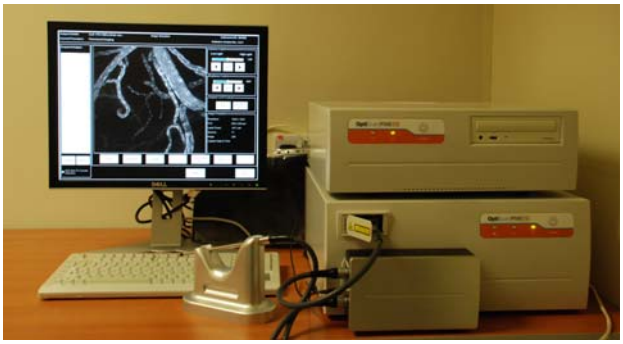

Fig. 1: The *in vivo* cellular imaging system using an endomicroscope.

Before the imaging process is carried out, a fluorescence dye is injected into the target subject. After half an hour, the endomicroscope probe is placed on the subject's surface. The probe begins to scan an area of $475\mu\mathrm{m} \times 475\mu\mathrm{m}$ (field of view) from the surface layer. Once an image is acquired and saved, it continues to scan the subsurface layer ($4\mu\mathrm{m}$ below the surface) by adjusting the laser illumination within the probe. This process continues until the laser light reaches the deepest layer ($250\mu\mathrm{m}$ below the surface). Finally, a stack of 60 slice images is obtained with a resolution as high as 1024×1024. Then the probe may be moved to a new site to capture another image stack.

The acquired endomicroscopy images are quite different from natural images in several aspects: (1) the images are molecular imaging of the living tissues across a $475\mu\mathrm{m} \times 475\mu\mathrm{m}$ area, (2) they are usually magnified by 1000 times by the microscopic probe, (3) the images are labelled with photosensitisers that selectively accumulate within the tissue,

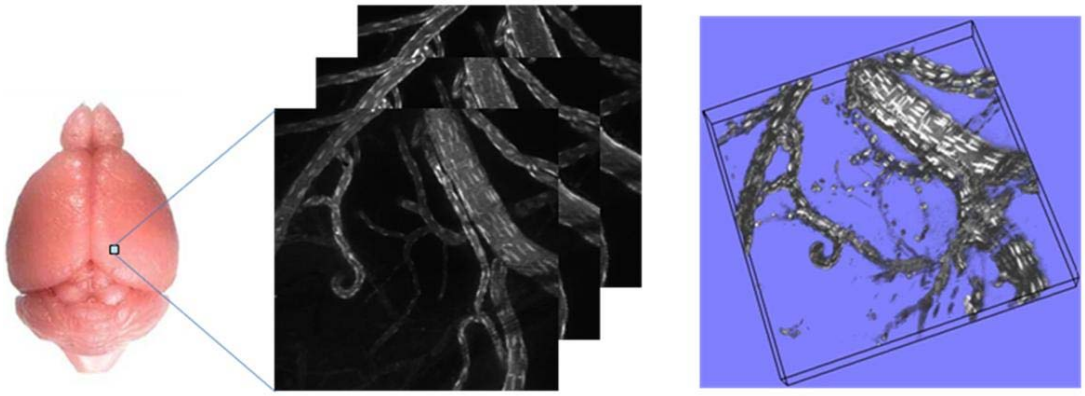

Fig. 2: The endomicroscopy images of mouse brain microvasculature at different z-depths and its 3D reconstruction.

(4) the fluorescence images are much noisier due to low signal to noise ratio, (5) the images are non-uniformly illuminated, and (6) the images are translation-, rotation-, and scale-variant. To reconstruct the 3D microstructure of the living tissue, we need to first register/align consecutive slice images. The challenges are twofold. On the one hand, each of the arbitrarily taken slice images suffers from various distortions due to possible tissue movement and tissue expansion/contraction. On the other hand, beyond a certain time frame, the 3D volumetric images may be different due to physiological changes.

Image registration or image alignment algorithms can be classified into two categories: spatial-domain methods and frequency-domain methods. One of the images is referred to as the reference and the second image is referred to as the target. In this work, we present a phase correlation-based image registration algorithm, which finds the transformation parameters while working in the frequency domain. Applying the phase correlation method to a pair of images produces a third image which contains a single peak. The location of this peak corresponds to the relative translation between the images. Compared with the spatial-domain algorithms such as intensity-based correlation methods [9], feature-based methods [10], and graph-theoretic methods [10], the phase correlation method is resilient to noise, occlusions, and other defects typically in the biomedical images. Additionally, the phase correlation uses the fast Fourier transform (FFT) to compute the cross-correlation between the two images, generally resulting in large performance gains. The method can be extended to determine rotation and scaling differences between two images by first converting the images to log-polar coordinates. Due to properties of the Fourier transform, the rotation and scaling parameters can be determined in a manner invariant to translation.

2. Theoretical Analysis

Assume two images $I_1(x, y)$ and $I_2(x, y)$ with a displacement (x_0, y_0), i.e., $I_2(x, y) = I_1(x - x_0, y - y_0)$. Applying the Fourier transform, we have,

$$I_2(u, v) = e^{-j2\pi(ux_0+vy_0)} I_1(u, v). \tag{1}$$

The cross-power spectrum of the two images is defined as,

$$\frac{I_1(u, v)I_2^*(u, v)}{|I_1(u, v)I_2^*(u, v)|} = e^{j2\pi(ux_0+vy_0)}, \tag{2}$$

where $I_2^*(u, v)$ is the complex conjugate of $I_2(u, v)$. The Fourier shift theorem guarantees that the phase of the cross-power spectrum is equivalent to the phase difference between the images.

By applying the inverse Fourier transform to the above phase difference, we have an impulse function $r(x, y) = \delta(x - x_0, y - y_0)$. The location of its peak value corresponds to the displacement that is needed to optimally register the two images. Fig. 3 shows the flowchart of the phase correlation technique.

The advantage of this method is that the discrete Fourier transform and its inverse can be performed using the fast Fourier transform, which is much faster than intensity-based correlation for large images. In practice, it is more likely that $I_2(x, y)$ will be a simple linear shift of $I_1(x, y)$, rather than a circular shift as required. In such cases, $r(x, y)$ may not be a simple delta function, which can possibly reduce the performance of the method. Therefore, a window function such as the Hamming window [11] should be employed during the Fourier transform to reduce edge effects, or the images should be zero padded so that the edge effects can be ignored. If the images consist of a flat background, with all detail situated away from the edges, then a linear shift will be equivalent to a circular shift, and the above derivation will hold exactly. For periodic images such as a chessboard, phase correlation may yield ambiguous results with several peaks in the resulting output.

The method can be extended to determine the rotation and scaling differences between two images by first converting the images to the log-polar coordinates. Assume $I_2(x, y)$ is a translated, rotated, and scaled replica of $I_1(x, y)$ with displacement (x_0, y_0), rotation θ_0, and scale s, according

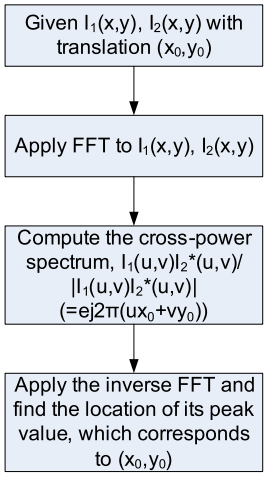

Fig. 3: The phase correlation technique.

to the Fourier translation, rotation, and scale properties, we have,

$$\begin{aligned} I_2(u,v) &= e^{-j2\pi(u'x_0+v'y_0)} I_1(su'\cos\theta_0 - sv'\sin\theta_0, \\ &\quad su'\sin\theta_0 + sv'\cos\theta_0). \end{aligned} \tag{3}$$

Assume $F_1(u',v') = |I_1(u',v')|$ and $F_2(u,v) = |I_2(u,v)|$ are their Fourier magnitude spectra, we have,

$$F_2(u,v) = F_1(su'\cos\theta_0 - sv'\sin\theta_0, su'\sin\theta_0 + sv'\cos\theta_0), \tag{4}$$

i.e.,

$$\begin{aligned} u &= s(u'\cos\theta_0 - v'\sin\theta_0), \\ v &= s(u'\sin\theta_0 + v'\cos\theta_0). \end{aligned} \tag{5}$$

In the polar coordinate system, we have,

$$\begin{aligned} u &= \rho\cos\theta, \\ v &= \rho\sin\theta, \end{aligned} \tag{6}$$

and

$$\begin{aligned} u' &= \rho'\cos\theta', \\ v' &= \rho'\sin\theta', \end{aligned} \tag{7}$$

By combining Eqs. (5)-(7), we have,

$$\begin{aligned} u &= s(\rho'\cos\theta'\cos\theta_0 - \rho'\sin\theta'\sin\theta_0), \\ &= s\rho'\cos(\theta'+\theta_0), \\ &= \rho\cos\theta, \end{aligned} \tag{8}$$

$$\begin{aligned} v &= s(\rho'\cos\theta'\sin\theta_0 + \rho'\sin\theta'\cos\theta_0), \\ &= s\rho'\sin(\theta'+\theta_0), \\ &= \rho\sin\theta, \end{aligned} \tag{9}$$

i.e.,

$$\begin{aligned} \rho = s\rho' &\Rightarrow \rho' = \rho/s, \\ \theta = \theta' + \theta_0 &\Rightarrow \theta' = \theta - \theta_0. \end{aligned} \tag{10}$$

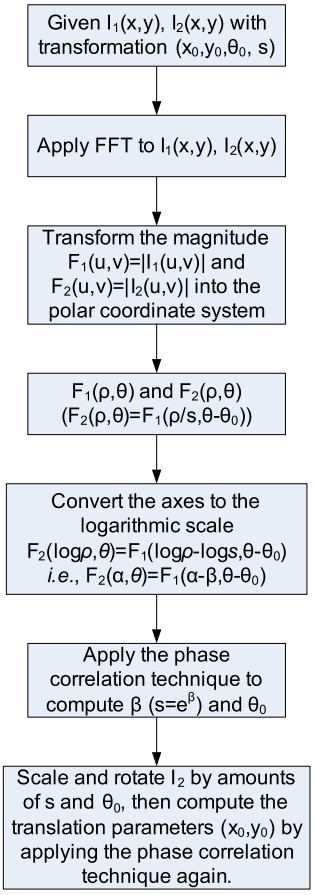

Fig. 4: The flowchart of the phase correlation-based image registration.

Thus, the Fourier magnitude spectra $F_1(u',v')$ and $F_2(u,v)$ in the polar representation are related by,

$$F_2(\rho,\theta) = F_1(\rho/s, \theta-\theta_0). \tag{11}$$

By converting the axes to logarithmic scale, we have,

$$F_2(\log\rho,\theta) = F_1(\log\rho - \log s, \theta-\theta_0), \tag{12}$$

i.e.,

$$F_2(\alpha,\theta) = F_1(\alpha-\beta, \theta-\theta_0), \tag{13}$$

where $\alpha = \log\rho$, $\beta = \log s$. Thus, the problem becomes one with relative translation only. Applying the phase correlation technique, we can find the scale $s = e^{\beta}$ and rotation θ_0.

After scaling and rotating $I_2(x,y)$ by the amounts of s and θ_0 respectively, the translation parameters x_0 and y_0 can then be obtained using the phase correlation technique. Fig. 4 summarises the process of the phase correlation-based image registration approach.

3. Experimental Results

To evaluate the performance of the algorithms, we perform a series of experiments on several sets of confocal fluorescence endomicroscopy images. Note that all the images in

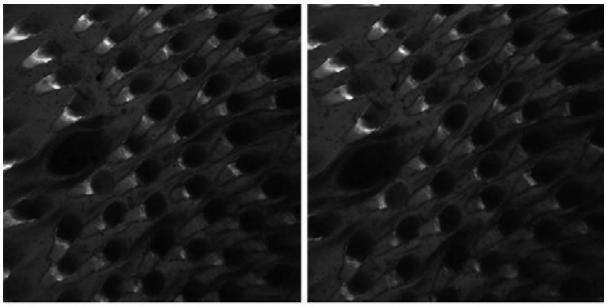

Fig. 5: Mouse tongue images (1024×1024) with displacement. (a) The original reference image and (b) the target image.

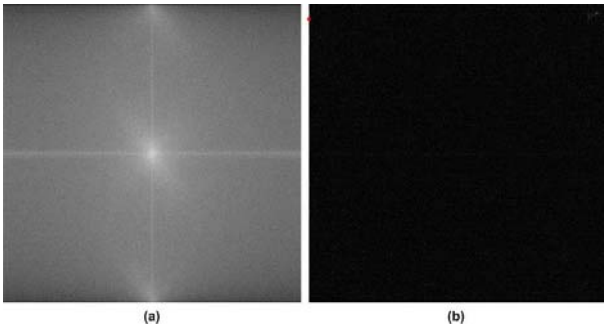

Fig. 6: The Fourier spectrum of the original reference image and the phase correlation image of the tongue images, where the translation parameters is estimated as: x=1, y=49.

the following figures are largely reduced for display purpose. First, we apply the phase correlation technique to a pair of mouse tongue images, as shown in Fig. 5. Fig. 6 shows the Fourier spectrum of the original reference image and the inverse Fourier transform of the cross-power spectrum of the tongue images. We can see a peak in the phase-correlation image approximately at (1, 49). Theoretically, the peak value should be equal to 1.0. However, the presence of dissimilar parts and the noise in images reduce the peak value. The aligned target image is illustrated in Fig. 7. Experimental results conducted on a pair of mouth images are shown in Fig. 8. From these results, we can see that the phase correlation technique does not work without preprocessing the endomicroscopy images.

In order to obtain a reasonable estimation of the translation parameters, we filter the original endomicroscopy images by a Laplacian filter to remove high-frequency components in the frequency domain. Fig. 9 shows the Fourier spectrum of the reference tongue and mouth images after applying the Laplacian filter and their aligned target images. The improved results demonstrate that by applying the Laplacian filter, the phase correlation is applicable to the endomicroscopy images with displacement.

Fig. 10 shows registration results on the mouse tongue (another pair) and mouse brain microvasculature images.

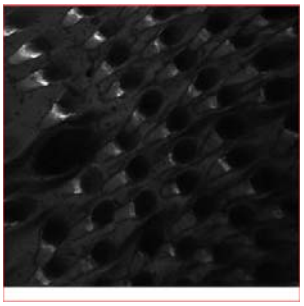

Fig. 7: The aligned mouse tongue images with white-pixel padding.

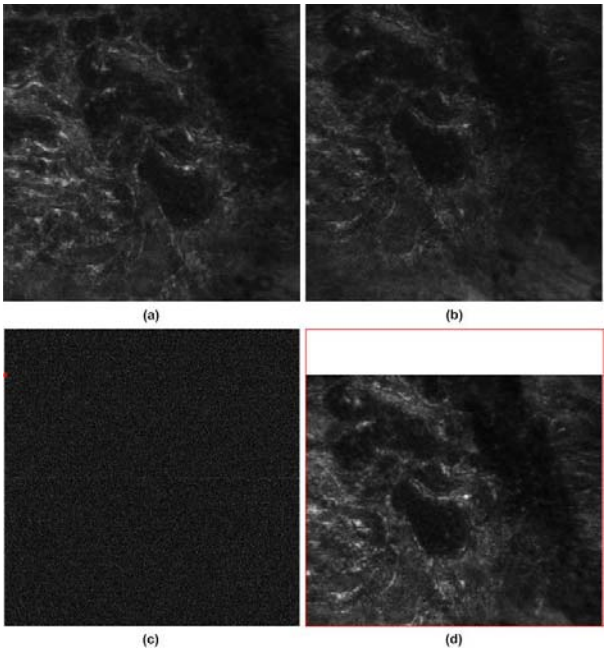

Fig. 8: The mouth image pair (1024×1024) with displacement. (a) The original reference image, (b) the target image, (c) the phase correlation image with the estimated translation parameters: x=2, y=159, and (d) the aligned tongue images with white-pixel padding.

From these experimental results, we can see that the phase correlation method is a robust approach for the estimation of the transformation paraments, leading to good image registration results.

4. Conclusions

In this work, we have developed a phase correlation-based registration approach for estimation of the relative transformations in the consecutive endomicroscopy images. The experimental results conducted on different sets of images reveal that the phase correlation-based alignment can be performed in real time and is robust to noise, occlusions, and other defects existing in the images. The good alignment between consecutive slice images will directly benefit the subsequent 3D reconstruction and visualisation of the

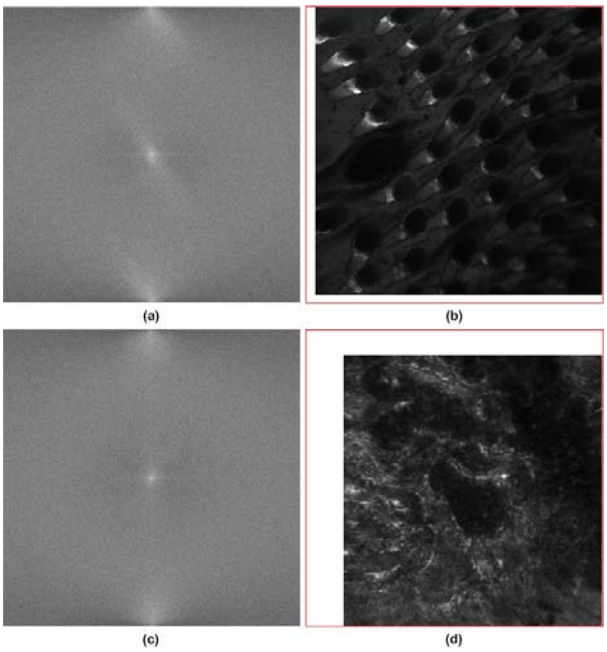

Fig. 9: The phase correlation results after applying the Laplacian filter. (a) The Fourier spectrum of the filtered reference tongue image, (b) the aligned tongue image using the estimated translation parameters (x=36, y=27) by phase correlation, (c) the Fourier spectrum of the filtered reference mouth image, and (d) the aligned mouth image using the estimated translation parameters (x=133, y=89).

living tissue's microstructure. It will enable clinicians to navigate within the living tissue freely, leading to a much more clinician-friendly imaging tool, and more definitive diagnostic results of various diseases including early-stage cancers, in a non-invasive way. This will provide profound health benefits to society.

Acknowledgment

This work is supported by the "Strengthening the All-Island Research Base" project, funded by the Northern Ireland Department of Education and Learning. The authors would like to thank M. Goetz, C. Schneider, et al. (University of Mainz, Germany), who provided the mouse brain microvasculature images for this study. We are also thankful to Prof. Hock Soon Seah, Dr. Feng Lin and Dr. Kemao Qian (Nanyang Technological University, Singapore), Prof. Soo Khee Chee, Dr. Malini Olivo and Dr. Patricia Thong (National Cancer Centre Singapore), for sharing the mouse tongue and mouth data with us.

References

[1] M. Goetz, C. Fottner, E. Schirrmacher, P. Delaney, S. Gregor, C. Schneider, D. Strand, S. Kanzler, B. Memadathil, E. Weyand, and M. Holtmann et al., "In-vivo confocal real-time mini-microscopy in animal models of human inflammatory and neoplastic diseases," *Endoscopy*, vol. 39, pp. 350–356, 2007.

[2] M. Goetz, S. Thomas, A. Heimann, P. Delaney, C. Schneider, M. Relle, A. Schwarting, P. R. Galle, O. Kempski, M. F. Neurath, and R. Kiesslich, "Dynamic imaging of microvasculature and perfusion by miniaturised confocal laser microscopy," *Eur. Surg. Res.*, vol. 41, pp. 290–297, 2008.

[3] P. S.-P. Thong, M. Olivo, K.-W. Kho, W. Zheng, K. Mancer, M. Harris, and K.-C. Soo, "Laser confocal endomicroscopy as a novel technique for fluorescence diagnostic imaging of the oral cavity," *J. Biomedical Optics*, vol. 12, no. 014007, pp. 1–8, 2007.

[4] F. Koenig, J. Knittel, and H. Stepp, "Diagnosing cancer in vivo," *Science*, vol. 292, pp. 1401–1403, 2001.

[5] D. B. Parente, E. L. Gasparetto, L. C. H. da Cruz, R. C. Domingues, A. C. Baptista, A. C. P. Carvalho, and R. C. Domingues, "Potential role of diffusion tensor MRI in the differential diagnosis of mild cognitive impairment and alzheimer's disease," *American J. Roentgenology*, vol. 190, no. 5, pp. 1369–1374, 2008.

[6] A. G. Filler, "The history, development, and impact of computed imaging in neurological diagnosis and neurosurgery: CT, MRI, DTI," *Internet J. Neurosurgery*, vol. 7, no. 1, 2010.

[7] G. T. Herman, *Fundamentals of Computerized Tomography: Image Reconstruction from Projection*, 2nd ed. Springer, 2009.

[8] R. Smith-Bindman, J. Lipson, R. Marcus, K.-P. Kim, M. Mahesh, R. Gould, A. B. de González, and D. L. Miglioretti, "Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer," *Arch. Intern. Med.*, vol. 169, no. 22, pp. 2078–2086, 2009.

[9] J. Kim and J. A. Fessler, "Intensity-based image registration using robust correlation coefficients," *IEEE Trans. Medical Imaging*, vol. 23, no. 11, pp. 1430–1444, 2004.

[10] A. A. Goshtasby, *2-D and 3-D Image Registration for Medical, Remote Sensing, and Industrial Applications*. Wiley Press, 2005.

[11] Y. Song and X. Peng, "Spectra analysis of sampling and reconstructing continuous signal using hamming window function," in *Proc. IEEE Fourth Int'l Conf. Natural Computation*, 2008, pp. 48–52.

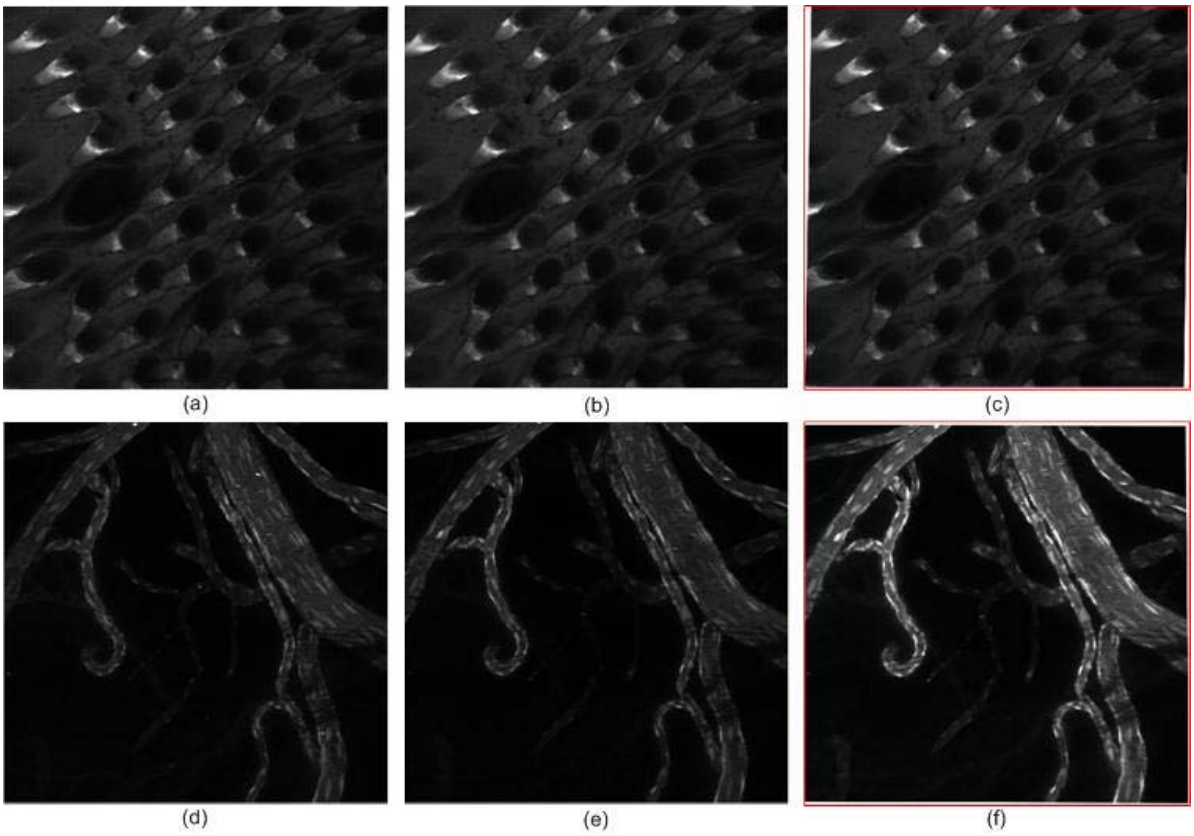

Fig. 10: The alignment results of mouse tongue and brain microvasculature images (1024×1024). (Left column) The original reference images, (middle column) the target images, and (right column) the aligned images with white-pixel padding.