33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Joint Spatio-Temporal Registration and Microvasculature Segmentation of Retinal Angiogram Sequences

Shuoying Cao*, Anil Bharath, Kim Parker, Jeffrey Ng, John Arnold, Alison McGregor and Adam Hill

Abstract— **We discuss the problem of 2D+t intra- and inter-sequential registration of retinal angiograms. A joint spatio-temporal registration algorithm is presented based on a RANSAC (RANdom SAmple Consensus) approach incorporating a quadratic model to describe "pairwise" image homography. This is incorporated into a local-to-global hierarchical joint registration framework. After registration, vessel centrelines are segmented to subpixel accuracy by applying multi-scale steerable complex wavelet filters. Frame-by-frame microvascular centrelines in Regions-of-Interest (ROIs) are evaluated against segmented centrelines of the temporal average of the registered sequences. The microvascular centrelines in registered sequences can be compared intra-sequentially and inter-sequentially, allowing non-invasive clinical monitoring of micro-circulation. This has the potential to detect the presence of microemboli and pathological structural alterations.**

I. INTRODUCTION

The retinal vascular system has spawned a wide range of clinical and pre-clinical research and diagnostic techniques since it provides unique *in vivo* access for studying the characteristics of the human vascular bed in a minimally invasive manner.

Fluorescein angiography is a well-established technique for clinical assessment of the retina. The passage of fluorescein dye through the retinal vessels reflects both the vessel structure and the rate of retinal blood flow. A fundus camera continuously photographs the retina from the onset of dye injection over a period of 3 to 5 minutes [1]. This captures filling ("wash-in") and the subsequent elimination ("wash-out") of dye in the retinal vessels. These angiogram sequences are roughly divided into the arterial phase (filling of retinal arteries), arteriovenous phase (complete filling of retinal capillaries with laminar flow exhibited in retinal veins), venous phase (complete filling in large retinal veins, leading to the maximum vessel fluorescence) and recirculation phase (approximately equal brightness in the veins and the arteries, with gradual elimination of dye from the retinal vasculature).

Measurements on retinal blood vessels have linked alteration of human (retinal) vasculature with diseases such as hypertension, diabetes and age-related macular degeneration [2]. For early stage detection, subtle changes in the retinal microvasculature require sufficiently fine-resolution imaging and dependable sub-pixel precision of registration to study vessels at the micrometer scale. Current literature focuses on establishing correspondence between global microvasculature measures and retinal blood flow [3]. Suggested parameters are blood flow velocity, arteriovenous passage time, difference of arterial and venous times to maximum intensity, and time to maximum image [3]. All of these fail to capture microvascular vessels individually.

Blauth *et al* [4] first suggested that comparison of pre- and post- operative retinal fluorescein angiograms might indicate the existence of microemboli that could be associated with cognitive impairment or even morbidity. However, only the macula region of one pre-op and one post-op image were used [4].

The analysis of sequential retinal angiograms has not been widely exploited; the dynamics of blood flow and the diffusion of the injected tracer introduce a low-frequency variation that is both difficult to compensate for and computationally demanding. However, fluorescein angiography allows visualisation of the microvasculature less than $30\mu m$ in diameter, which is not yet achievable by either funduscopy or color photography. By recruiting all frames in both sequences, it is possible to study retinal microvasculature dynamics and identify small, but potentially significant embolic events. In this paper, we address and solve three technical problems.

1) ***Image registration***: We need to estimate and model the distortion between frames in order to map each angiogram onto one common coordinate system (*the reference*). Current feature-based methods include global weak affine model with Bayesian matching [5], hierarchical model refinement [6] and dual-bootstrap iterative closest point (DB-ICP) [7]. Often, vascular bifurcation points are extracted as landmark points to estimate the transformation model. Inaccurate landmarks can heavily distort the transformation estimate, especially in [6], where each image only contains about 30-50 landmarks. This is improved in [7], where landmarks in local "bootstrap" regions are iterated to establish correspondence and to fit and refine the transformation estimate. This method also relies on the accurate initialization of corresponding landmark points.

2) ***Objective validation***: We need an objective error measure to evaluate the performance of our registration algorithm. Registration error can be determined by the extent of misalignment between the registered image and the reference image. Conventional ground truth is obtained from manual registration [5]. This form of reference standard is subject to inter- and intra-observer variability. Therefore, [6]&[7] use the "centreline" (linked lines between landmarks) locations of the original images as relatively "unbiased" ground

This work was supported by Hammersmith Hospital Trustees.

*S. Cao, A. Bharath, K. Parker, J. Ng and A. Hill are with Bioengineering Department, Imperial College London, South Beddington Campus, London SW7 2AZ (email: shuoying.cao@imperial.ac.uk).

John Arnold is with Charing Cross Hospital, Imperial College Healthcare NHS Trust.

Alison McGregor is with Faculty of Medicine, Surgery and Cancer, Imperial College London.

truth. However, both the linking algorithm that interpolates subpixel locations and the similarity measure that optimally matches subpixels between different frames vary from case to case in establishing error metrics. Therefore, [8] casts an evenly spaced virtual grid that intersects with vessels to obtain ground-truth pixels. This, however, disregards the location-dependent nature of pixel information. Pixels located near the capillary-rich macula region generate significantly more clinical interest than those near the retinal border (field-stop). Furthermore, their proposed "error tracing" route [8] theoretically favors strategies that pair both the forward and the inverse registration functions to optimize the "net offset" rather than a true assessment of the registration algorithm on its own. Additional processing is required to address the uneven global illumination.

3) ***Vessel segmentation***: We need to differentiate and identify microvascular segments from the background in each of our angiograms. Extensive research focused on this area includes: use of morphology with Laplacian-of-Gaussian filtering [5], exploratory vessel tracing [6], region-growing on Hessian matrix maxima in scale-space [9] and maximum likelihood estimates from multi-scale filter outputs in scale-space [10]. Yet, most algorithms are developed on fundus images, and are not necessarily able to capture capillaries in fluorescein images at a much finer scale.

II. METHODS

A. Pairwise Registration

Geometric distortion, radiometric degradation, and additive noise corruption contribute to the difficulty in registration. Problems specific to clinical retinal angiography are: the photographer's bias in image capture, the patient's involuntary movement, and the 3D to 2D warping between retina and the camera. Existing techniques, for example, approximate the retinal surface by a sphere [6], predict a perspective distortion [8], or model the intensity variation to address uneven illumination [11].

At the pairwise level, we combine projective RANSAC [12] with a quadratic homography transformation, preconditioned by contrast-enhancement within the field-stop. The latter is required because the low dye concentration in the retinal blood vessels both at the beginning and the end of the sequence requires temporary adjustment of the dynamic range of pixel intensities. To disregard the edge of the frame and the patient record information, a circular Hough Transform is used to detect the field stop. Local contrast within the field-stop is histogram equalized and vessel bifurcations or crossings (landmarks) are selected based on Harris corner measure. False corners near the borders are eliminated with a distance criterion. Landmarks are then putatively matched by windowed normalized cross-correlation. RANSAC iteratively estimates the "best-fit" projective model applicable to most putative matches. We use the *inliers-to-outliers* ratio (the comparison between the number of pairs that can be described by the "best-fit" model within an offset threshold and those that cannot) as a criterion on whether to apply a higher order quadratic transformation model (with 12 degrees of freedom). Although the quadratic model is more tolerant to pairwise distortion, the refinement is only sensible if the lower order registration has been successful.

B. Joint Registration

To analyse temporal information both within a consecutive sequence of retinal angiograms (intra-sequence) and across two different sequences taken before and after the operation (inter-sequence), we need a stepwise systematic framework that first aligns each pixel intra-sequentially then cross-aligns the same pixel inter-sequentially. Multi-temporal registration requires maximizing the point correspondence between similar structural features, while still allowing us to differentiate, detect or even monitor pathological changes. For each patient, the last post-op frame was acquired several hours after the first pre-op frame. This increases the chance that the vasculature alters either in width or in curvature, in addition to the natural variability of blood vessels.

For the n^{th} frame in a sequence S of length N, let (x_n, y_n) denote the location of each pixel in the frame coordinates at time t_n. The image is denoted by the function:

$$f_n := f_n(x_n, y_n; t_n), \;\; n \in [1, N] \tag{1}$$

Consider two unregistered angiogram sequences, acquired with frame-specific spatial coordinates relative to the camera lens, at unknown points in time relative to the cardiac cycle, and with non-uniform frame-to-frame intervals (from less than a second to tens of seconds) along each sequence:

$$S^{(A)} = \left\{ f_n^{(A)} \right\}_{n=1,2,3\ldots N_A} \text{and } S^{(B)} = \left\{ f_n^{(B)} \right\}_{n=1,2,3\ldots N_B}$$

We first spatially register each individual image $f_n^{(A)}$ in $\mathrm{S}^{(A)}$ to a *local reference* spatial coordinate system defined by frame $f_{lr}^{(A)}$. This **individual-to-local reference** registration is also applied to sequence $\mathrm{S}^{(B)}$ with frame $f_{lr}^{(B)}$ as its local reference. We then register the two local references $f_{lr}^{(A)}$ and $f_{lr}^{(B)}$ separately to one *global reference* f_{gr}. This **local-to-global reference** transformation is further combined with prior individual-to-local reference transformations to give the **individual-to-global reference** transformation that allows $\mathrm{S}^{(A)}$ and $\mathrm{S}^{(B)}$ to co-register to one global reference f_{gr}.

We adopt the clinical practice of selecting the darkest image as the *local reference*. Our algorithm computes the sum of the pixel intensities within the field-stop and selects the frame at the peak of the dye-time course (when the image is the darkest) in each sequence:

$$f_{lr} = \{f_{n^*}\}, \quad n^* = \underset{n \in [1,N]}{\arg\min} \langle f_n, M_n \rangle \tag{2}$$

with $\langle \cdot, \cdot \rangle$ denoting a spatial inner product and M_n a spatial weighting function (binary mask) that is unity for points (x_n, y_n) within the field-stop region and 0 outside.

The ***individual-to-local reference*** registration is given as:

$$f'_n := f'_n(x'_{(lr)}, y'_{(lr)}; t_n) = R_{n(lr)}(f_n), \;\; n \in [1, N] \tag{3}$$

where individual frame f_n is mapped to the spatial coordinate system of local reference f_{lr} by function $R_{n(lr)}$.

The ***local-to-global reference*** transformation is given as:

$$f'_{lr} := f'_{lr}(x'_{(gr)}, y'_{(gr)}; t_{lr}) = R_{lr(gr)}(f_{lr}) \tag{4}$$

where the local reference f_{lr} is mapped to the spatial coordinate system of global reference f_{gr} by function $R_{lr(gr)}$.

Lastly, the ***individual-to-global reference*** registration can be combined as:

$$f_n'' := f_n''(x_{(gr)}'', y_{(gr)}''; t_n) = R_{lr(gr)}(f_n'), \quad n \in [1, N] \quad (5)$$

where individual frame f_n is mapped to the spatial coordinate system of the global reference f_{gr}.

C. Vessel Segmentation

The estimate for the centreline location of the vessels requires subpixel resolution accuracy. In practice, fine vessel structures may not be captured in all frames within a sequence. It is commonly observed that capillaries may "disappear" from the previous frame then "re-emerge" in the following one. Lastly, due to the time delay in the passage of dye, angiograms at the beginning and the end of a sequence reveal significantly less information on the detailed microvasculature.

The temporal average of the registered sequence(s) is a more perceptually accurate representation of the retinal microvasculature as it takes into account previous and subsequent frames in one or more sequences. For globally registered sequence(s) with N number of frames, $\{f_n''\}_{n\in[1,N]}$, the temporal average across the sequence(s) is defined as:

$$f_{average} = \frac{1}{N}\sum_{n=1}^{N} f_n'' \quad (6)$$

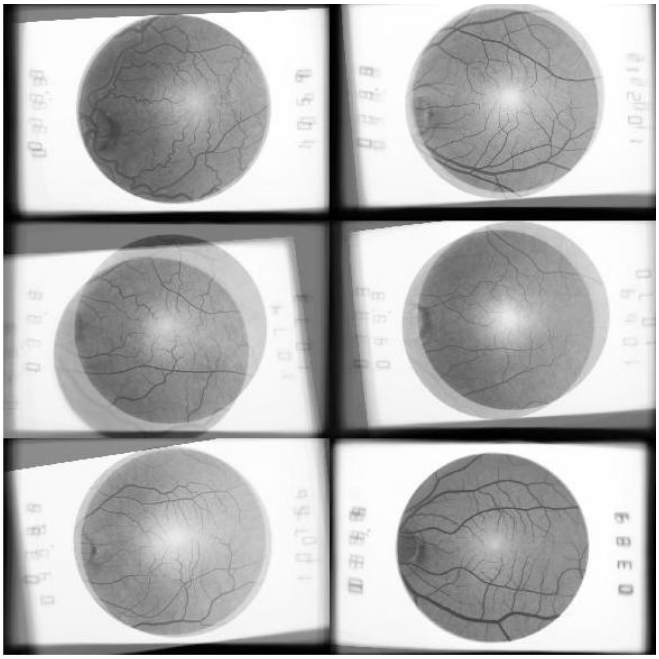

Fig. 1. Temporal average $f_{average}$ of registered sequences (pre- and post-pulmonarycardiac bypass operation) for each patient (a total of 6 patients)

Our segmentation algorithm uses prior work [10], [13] and [14]. We apply steerable complex wavelet filters at multiple scales on each frame f_n, and use the filter outputs to infer the presence of vessels and centreline locations.

A local orientation map can be constructed as:

$$\mathrm{O}_n^{(l)}(x_n, y_n) = \frac{\sum_{k=0}^{K/2-1} |g_k^{(l)}(x_n, y_n)| e^{j2\phi_k}}{p + (\sum_{k=0}^{K/2-1} |g_k^{(l)}(x_n, y_n)|^2)^{\frac{1}{2}}} \quad (7)$$

for $g_k^{(l)}(x, y)$, $k = 0, 1, 2, ..., K\text{-}1$ denotes the output of the k^{th} order oriented bandpass complex analysis filter from image f_n at level l, and p is a conditioning constant [13].

A local phase estimate, $\Psi_n^{(l)}$, is obtained from filer steering by the polynomial functions $s_p(\phi, k)$ and $s_q(\phi, k)$ on f_n:

$$\Psi_n^{(l)} = \angle\Big(\sum_{k=0}^{K/2-1} s_p(\phi, k) f_k^{(l)}(x_n, y_n) + \sum_{k=0}^{K/2-1} s_q(\phi, k)(f_k^{(l)}(x_n, y_n))^*\Big)$$

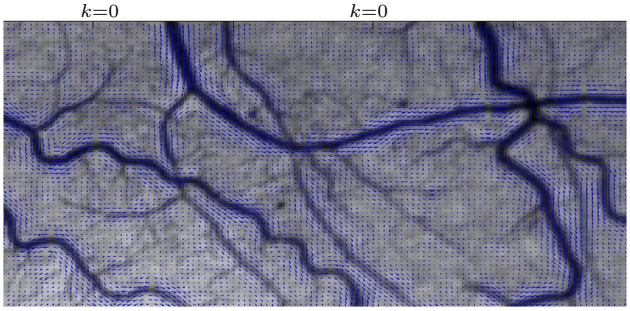

Fig. 2. Orientation dominance field of ROI near macula on an angiogram

For each pixel, we match local orientation $\mathrm{O}_n^{(l)}(x_n, y_n)$ (Figure 2) with phase estimate $\Psi_n^{(l)}$ in an 8-connected neighbourhood. The extracted centreline locations are refined by the subpixel information held in the phase shift between pixels. In scale-space, we link and weight the *candidate* locations at different scales to filter out the noise.

D. Objective Validation

For the validation, we take a patch (ROI) near the macula region with a high density of microvasculature. First, we establish our ground-truth as the centrelines $v_{average}(x, y)$ within the ROI on $f_{average}$ (Figure 1). Then we segment the vessel centrelines $v_n(x, y)$ (within the same ROI) frame-by-frame in the co-registered pre- and post-op sequences. For a given ROI area with vessel length L, we define the centerline error measurement (CEM) as:

$$CEM = \frac{1}{L}\sum_{n=0}^{L-1} \|v_n(x, y) - v_{average}(x, y)\| \quad (8)$$

where $\|\cdot\|$ is the Euclidean distance. This is a fairer assessment of registration quality as the clinically interesting fine-scale microvasculatures contribute more strongly to the error metric than a spatially-averaged global measure.

III. EXPERIMENTS

We tested our proposed method on 384 retinal angiograms (4288-by-2848 pixels per frame, captured by a Zeiss retinal camera at 30 degrees field) from 6 patients. Each patient has one pre-op and one post-op sequence, with ∼65 frames per sequence. We processed our sequences using MATLAB 2009 on an Intel Dual Core CPU with 3.48G RAM. Our segmentation evaluation used ROIs of size 101×101 pixels near the macula containing complex image structures, and with clearly displayed capillaries ranging from 5-10 pixels in width. We compared our algorithm with an affine model, that accounts for pairwise rotation, scale, translation and shearing. The validation and segmentation procedures were controlled, so the same ROIs were used in both models. Both algorithms ran on the same machine, and both algorithms used the same local and global references to give a fair

comparison. Figure 3 presents the comparison results of centreline error measure (CEM) between the outputs of the two algorithms. For fine structures, the affine model has on average 2-3 pixel misalignments and large error bars, up to 13-14 pixels. Our model, on the other hand, has a much more stable performance at 0.1-0.15 averaged pixel misalignment and no more than 0.3 pixel in all registered frames.

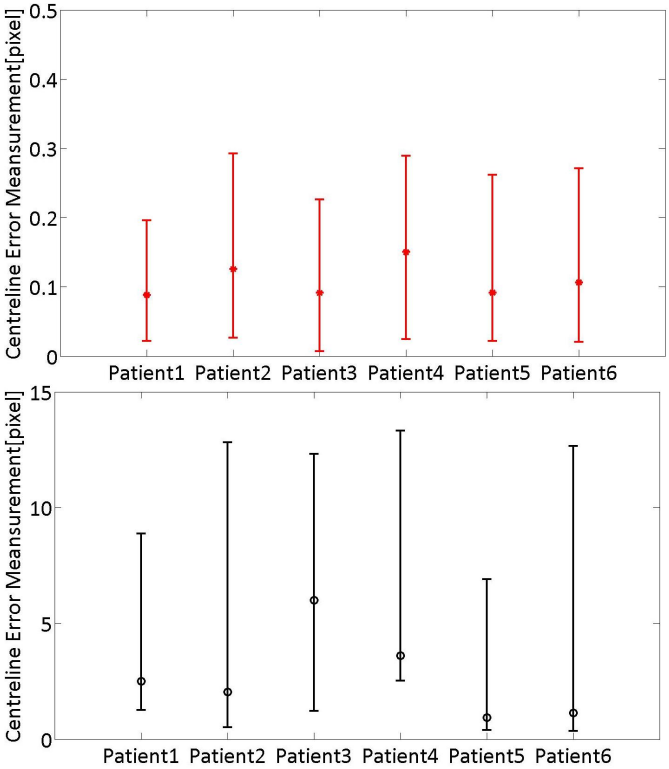

Fig. 3. Top: CEM of registered sequences per patient from our algorithm; Bottom: CEM of registered sequences per patient using affine model. Note the different scales.

IV. CONCLUSIONS

In this paper, we first suggested a novel joint registration procedure with promising results in both the spatial and temporal domains. We then presented a novel pixel-wise approach for centreline segmentation without thresholding or region-growing, which is used for evaluation of registration accuracy. This allows comparison and non-invasive monitoring of fine-resolution microvasculature from an existing well-established technique. It provides the potential for detecting temporal changes in the circulation (possibly caused by microembolism, see Figure 4) in real-time. This allows early preventative measures to be taken to reduce aggravated blood-clotting, thus improving patient's post-operative recovery. Future development includes extensive validation, real-time performance evaluation and clinical trials.

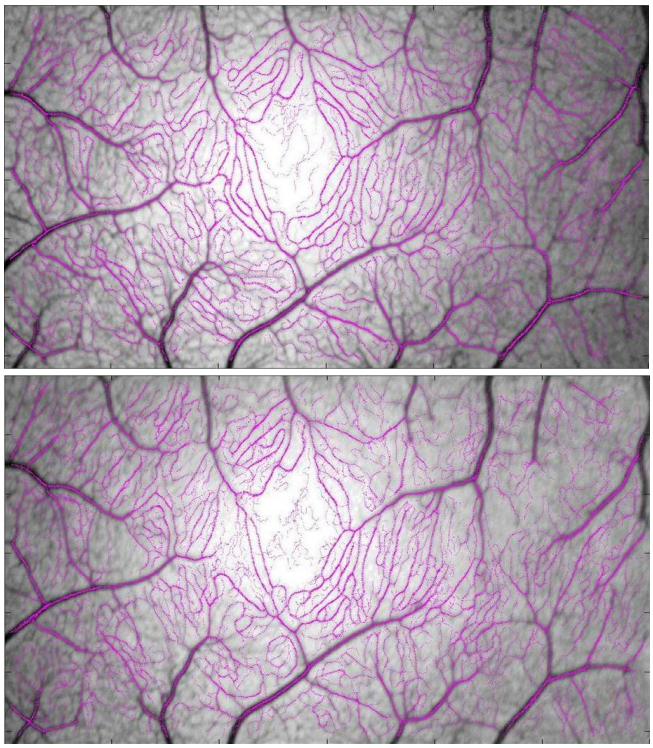

Fig. 4. Top: Macular microvasculature centreline of $f_{Pre-op\ average}$; Bottom: Macular microvasculature centreline of $f_{Post-op\ average}$ with visible disappearance of microvessels compared to the pre-op centrelines above, possibly occluded by microemboli.

REFERENCES

[1] G. Richard, "Fluorescein Angiography. Technique, Interpretation and Application", Oxford University Press, 1990.

[2] A. V. Stanton, B. Wasan, A. Cerutti, S. Ford, R. Marsh, and P. S. Sever, "Vascular network changes in the retina with age and hypertension", J. Hypertens., vol. 13, no. 12, pp. 1724-1728, 1995.

[3] J. H. Hipwell, A. Manivannan, P. Vieira, P.F. Sharp, and J. V. Forrester, "Quantifying changes in retinal circulation: the generation of parametric images from fluorescein angiograms", Physiol. Meas., vol. 19, no. 2, pp. 165-180, 1998.

[4] C. Blauth, J. Arnold, E. M. Kohner, and K. M. Taylor, "Retinal microembolism during cardiopulmonary bypass demonstrated by fluorescein angiography", Lancet, vol. 328, no. 8511, pp. 837-839, 1986.

[5] F. Zana, J. C. Klein, "A multimodal registration algorithm of eye fundus images using vessels detection and Hough transform", IEEE Trans. Med. Imaging, vol. 18, no. 5, pp. 419-428, 1999.

[6] 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. Machine Intell., vol. 24, no. 3, pp. 347-364, 2002.

[7] C. V. Stewart, C. L. Tsai, and B. Roysam, "The Dual-Bootstrap Iterative Closest Point Algorithm with Application to Retinal Image Registration", IEEE Trans. Med. Imaging, vol. 22, no. 11, pp. 1379-1394, 2003.

[8] S. Lee, J. M. Reinhardt, P. C. Cattin, and M. D. Abrmoff, "Objective and expert-independent validation of retinal image registration algorithms by a projective imaging distortion model", Med. Image Anal., vol. 14, no. 4, pp. 539-549, 2010.

[9] M. E. Martinez-Perez, A. D. Hughes, A. V. Stanton, S. A. Thom, A. A. Bharath, and K. H. Parker, "Retinal blood vessel segmentation by means of scale-space analysis and region growing", MICCAI, Lect. Notes Comput. Sci., vol. 1679, pp. 90-97, 1999.

[10] J. Ng, S. T. Clay, S. A. Barman, A. R. Fielder, M. J. Moseley, K. H. Parker, C. Paterson, "Maximum liklihood estimatation of vessel parameters from scale space analysis", Image Vision Comput., vol. 28, no. 1, pp. 55-63, 2010.

[11] J.C. Nunes, Y. Bouaoune, E. Delechelle, and Ph. Bunel, "A multiscale elastic registration scheme for retinal angiograms", Comput. Vis. Image Underst., vol. 95, no. 2, pp. 129-149, 2004.

[12] M. A. Fischler, R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", Comm. of the ACM, vol. 24, no. 6, pp. 381-395, 1981.

[13] A. A. Bharath, J. Ng, "A Steerable Complex Wavelet Contruction and Its Application to Image Denoising", IEEE Trans. Image Process., vol. 14, no. 7, pp. 948-959, 2005.

[14] J. Ng, A. A. Bharath, "Steering in Scale Space to Optimally Detect Image Structures", ECCV, Lect. Notes Comput. Sci., vol. 3021, pp. 482-494, 2004.