FREQUENCY SELECTIVE HEIGHT-FROM-STEREO RECONSTRUCTION: APPLICATION TO RETINAL OEDEMAS DETECTION

Slaven Marusic¹, Petr Dokladal², Ali Erginay³, Jean-Claude Klein² and M. Palaniswami¹

1) Department of Electrical & Electronic Engineering, The University of Melbourne, Australia
2) Centre of Mathematical Morphology, School of Mines, Paris, France,
3) Service d’Ophtalmologie, Hôpital Lariboisière, Paris, France.
Email: s.marusic@ee.unimelb.edu.au, petr.dokladal@ensmp.fr

ABSTRACT

The wavelet transform decomposes a given signal into a time-frequency representation, preserving the locality of both. Given that the local relief elevation is correlated to the phase difference in the stereo pair of images, we demonstrate a frequency selective relief reconstruction from a stereo pair of images, utilizing the dual-tree complex wavelet transform to provide phase information.

The dual-tree complex wavelet decomposition enables the hierarchical measurement of the phase difference in situations where the relief elevation is only observable in limited components of the image. Incorporating a data specific image registration technique, this paper presents an new application of the reconstruction of the macular oedema from a stereo pair of eye fundus photographs, offering the potential to serve as a diagnostic aid in the process of detecting macular oedema.

1. APPLICATION AND MEDICAL CONTEXT

Macular oedema may occur in many ocular diseases, like diabetic retinopathy, uveitis, vascular diseases or after eye surgery. The retina is a very fragile, photoreceptive layer covering the rear side of the eye ball, behind the vitreous humor (as displayed in Figure 1). The swelling of the retina, induced by the formation of liquid-filled cavities, represents a risk of visual loss if not treated. The fovea, situated in the centre of macula, has the highest density of photoreceptors, and is responsible for central vision and losing sight in this area is extremely disabling. In presence of exudates often associated with oedema [2], the patient is asked to have an optical coherence tomography (OCT) examination performed to measure the foveal thickness [1]. However, the OCT is not available everywhere, and other alternative techniques might be helpful in proving/excluding or quantifying an oedema. Image processing techniques have been applied for the detection of exudates [2]. This method, however, does not adequately address the particular challenges associated with directly detecting macular oedema.

One approach to the detection of retinal oedemas is through the use of stereo retinal images. A pair of photographs of the retina is taken, with 6 to 9 mm horizontal displacement of the retinograph. The 6 mm displacement can be achieved with the natural pupil dilation in darkness. A 9 mm displacement needs a dilation of the pupil using drugs (eye drops).

Measurement of the height of the retinal relief is a challenging task for several reasons:

- the stereo angle is weak
- the relief elevation is extremely small (below 500um)
- the displacement has to be measured on almost uniformly coloured objects, while the most visible, contrasted objects are not affected by the relief elevation

Visual evaluation of stereo pairs by trained practitioners presents the same challenges, while also causing significant eye strain and consequently limiting the effective evaluation time. Traditional automated techniques, e.g. block matching by correlation, are not selective, and measure the displacement of the most contrasted objects. In application to the retinal oedema detection, the relief elevation measure must be selective, since not all objects can be used.

Presented here is a novel approach to perform frequency selective height-from-stereo reconstruction for the detection of retinal oedemas. The technique is based firstly on a data specific image registration technique, presented in Section 2, that is critical in preserving the respective locations of the retinal oedema in the stereo pair, while aligning unrelated image features. This enables the height evaluation from stereo reconstruction of the oedema. The stereo displacement estimation, which reflects the increased foveal thickness associated with the oedema, is performed using a complex wavelet based, hierarchical framework described in Section 3. This is followed by the results of the proposed method, evaluated using retinal images with corresponding OCT data. The proposed method could also be applied to three-dimensional reconstructions of optic disc deformation for detection of glaucoma as proposed in [3].

2. IMAGE REGISTRATION

The stereo pair of images (see Figure 2) were taken with a horizontal displacement of the retinograph, with natural dila-
tion of the pupil. The objectives of registering is to align the images so that the useful stereo information (inside the macula) is preserved. The registration is performed by a block matching algorithm on the periphery of the images. The reasons are two: i) the stereo information should not be distorted in the center of the image, i.e. the macula, and ii) the registration outside the macula is easier due to the presence of well-contrasted vessels.

The registration proceeds in several steps.

1. Mask out the black, outer portion of the image and the non-uniformly illuminated data on the periphery. Crop the useful data inside the circular support (circle $m$ at Figure 3). This can be done manually (once for a given retinograph), or using the two following steps:
   - Let $f_1^i$, $f_2^i$ denote the two images. Threshold $f_1^i$, $f_2^i$ on the intensity to mask the black part outside the circular useful data.
   - Erode the mask $m'$ (see Figure 3) of the useful image data to eliminate the illumination disparity around the border,
     
     $$ m'_i = \begin{cases} 1 & \text{if } f_1^i(x) > 0 \\ 0 & \text{elsewhere} \end{cases} \quad \text{for } i = 1, 2 $$

     with $B$ being a circular structuring element of $\approx 1/100^2$ diameter of $m'$. The operation $\varepsilon_B$ is a morphological erosion with a structuring element $B$. For the basics of Mathematical Morphology see [4].
   - For the following consider $f_i = f_i^i \cap m$, for $i = 1, 2$.

2. Smooth the data with a linear filter. The used filter is a linear, lowpass filter with a 2-D, rotationally symmetric, impulse response. For these experiments, we have used a gaussian 3x3 window lowpass filter. A rectangular window 3x3 can also be used (even though it introduces a slight, unobservable anisotropy in the filtered signal).

3. The displacement estimation for registering is estimated by searching the maxima of normalized cross-correlation between $f_1$ and $f_2$. The block matching is performed locally in a rectangular window $w$.

Recall that cross-correlation $r(f, g)$ of two functions $f, g$ can efficiently be computed by utilizing the Fourier transform property of convolution $\mathcal{F}(f*g) = \mathcal{F}(f) \mathcal{F}(g)$ where the relation between correlation and convolution is $r(f, g) = \mathcal{R}e\{\mathcal{F}^{-1}[\mathcal{F}(f(x)) \mathcal{F}(f(-x))]\}$ (1) $\mathcal{R}e$ denotes the real part of a complex number. Recall that, if $\mathcal{F}$ is implemented with FFT, the right-hand side of Eq. (1) has a lower computation complexity than the naive computation of the convolution. For FFT we need to use a window with $2^n$ lateral size. Let $w^n \subset \mathbb{Z}^2$ denote a rectangular window of lateral size $2^n$.

The displacement estimation is performed iteratively, in a coarse-to-fine manner, with increasing resolution. In every iteration, one takes into account the displacement detected in the previous step.

The algorithm sums to the following:

- Given the diameter of the useful data support $m$, find maximum $N$ such that the window $w^N$ fits inside $m$.
- Define a grid of coordinates $B^N = \{b^n\} \subset \mathbb{Z}^2$, such that placing its centre at $b^n$, the window $w^N$ fits inside $m$ and touches the border of $m$. One may allow a small, one half or less, overlap between adjacent window positions, see Figure 3.
- At coordinates $b^n$, initialize to zero the displacement map $d(b^n) = \{0\}$.
- With increasing resolution, for $n = \{N, N-1, N-2, \ldots \}$ repeat:
  - Compute the local displacement map:
    * Compute the x-correlation
      
      $$ r(f_1, f_2)(x) = r(f_1^{m^w}(x+b^n), f_2^{m^w}(x+b^n + d(b^n))) $$

      where $f(x+b^n)$, with $x \in \mathbb{Z}^2$, denotes the translation of $f$ at coordinate $b_i$. For the sake of easy reading, $f^{m^w}$ denotes $f$ $\cap w$, and $f^{m^w}(x+b)$ translation and windowing of $f$.
    * and find
      
      $$ d(b^n) = \arg \max_{x \in w^N} r(x,y) $$

      - Refine the grid of coordinates $B^{n-1} = \{b^{n-1}\}$.
      - Refine the displacement map $d(b^n)$ by to-nearest-neighbor interpolation to $d(b^{n-1})$.
    - End repeat

In our experiments, letting $n = \{N,N-1,N-2, \ldots \}$ proved to be sufficient.

4. The registration is performed by a linear transformation of $f_2$. Linear transformations include rotation, translation, scaling and shearing. For basics to image registration, see a survey [5]. The transformation is determined with Least Squared Errors (LSE) from the displacement coordinates $d(b^n)$ obtained with last $n$ in the serie $\{N, N-1, \ldots \}$.

Figure 2: A stereo pair of retina photographs.

Figure 3: Image registration.
3. RETINAL OEDEMA DETECTION

The method utilized here to estimate the increased foveal thickness resulting from the retinal oedema is a wavelet-based approach that measures the displacement of image features between registered stereo images of the retina. Recognizing that the determining characteristics of retinal oedemas are often only subtle and gradual variations in intensity in the region of interest between the two images, a frequency selective approach is well suited to the problem. Limitations of intensity based stereo displacement and motion estimation solutions include a sensitivity to noise and illumination differences in the respective images, as well as the potentially high computational complexity associated with region based matching approaches. A hierarchical framework addresses the latter concern. A phase rather than intensity based solution was proposed in [8] that partly addresses the former concern though still sensitive to noise and computationally intensive.

Thus a complex wavelet based approach is applied, utilizing the phases of the complex filtered images as a key measure. The use of phase information can also provide real valued estimates for the stereo displacement of image features by virtue of the close relationship between phase differences and small shifts at given scales. The specific approach is based on a hierarchical technique for motion and displacement estimation using complex wavelets [6]. An adaptation of this hierarchical framework is displayed in Figure 4. The dual-tree complex wavelet transform forms the basis of the scheme that incorporates directionally selective Gabor-like filters [7].

3.1 Dual Tree-Complex Wavelet Transform

The Dual-Tree Complex Wavelet Transform [7] applies two sets of wavelet filters to the same data giving a 2:1 redundancy, which applied separably to two-dimensional data produces a 4:1 redundancy. It is the relationship between these filter pairs that delivers many useful properties to the DT-CWT, including Gabor-like directionally oriented filters and approximate shift invariance. The design approach, of which there are a number, requires that the wavelet of the upper tree is the approximate Hilbert transform of the wavelet in the lower tree. This allows the two trees to be regarded as respectively real and imaginary, providing the set of complex coefficients. The analysis filter framework of the 1D DT-CWT is shown in Figure 5. The filters used in the first stage are different from the remaining stages to ensure that the remaining stages provide an approximately analytic behaviour and single-sided frequency responses. Applied separably in two-dimensions, alternating between rows and columns, the two-dimensional DT-CWT produces subbands oriented in six directions (at approximately 15°, 45°, 75°, 105°, 135°, 165° respectively) in both the real and imaginary wavelet trees.

In addition to reducing the potential aliasing effects of real 2D DWTs, the DT-CWT produces directionally selective filters, encompassing six evenly spaced orientations. The dyadic sampling applied to each tree (beyond the first level) offers the scale selectivity that enables a coarse to fine representation of the data. The filter orientations are displayed in Figure 6 at the first two scales.

3.2 Displacement Estimation

The estimation of the stereo displacement is based on [6], where each subband is considered the output of a downsampled 2D Gabor-like complex bandpass filter. A key property is the finding in [8] that enables the interpolation between scales by modulating a low-pass interpolation kernel.
to the centre frequency of the band pass filter and convolving with the modulated kernel. The wavelet transformed images are represented by $F^{(k,\alpha)}(n)$ and $F^{(k,\beta)}(n)$, $k$ is the subband orientation, $s$ is the decomposition level and $n$ is the current coefficient location.

Given a delta function interpolation kernel (which produces reliable displacement estimates up to half the sampling bandwidth at the current level, $0.5 \times 2^s$ pixels), the link between phase difference and spatial translations is given by

$$F^{(k,\alpha)}(n + c) \approx F^{(k,\alpha)}(n)e^{i\theta(c)}$$

where $c$ is the interpolated sub-pixel index at level $s$, such that $c = d/2^s$ where $d$ is the actual displacement in pixels.

The subband squared difference (SSD) between coefficients in respective images $F_1$ and $F_2$ is given by

$$SD^{(k,\alpha)}(n, c) = |F_1^{(k,\alpha)}(n + c) - F_2^{(k,\alpha)}(n)|^2$$

The SSD surface can be approximated as a quadratic function around its minimum line, which can then be combined with those from the remaining orientations to find the surface minimum $(c_{\text{min}}, y_{\text{min}})$ produced by the squared difference surface, given by

$$SD^{(j)}(n, c) = \sum_{k=1}^{5} SD^{(k,\alpha)}(n, c)$$

$$SD^{(j)}(n, c) \approx \alpha (c_x - c_{\text{min}})^2 + \beta (c_y - c_{\text{min}})^2 + \gamma (c_x - c_{\text{min}})(c_y - c_{\text{min}})$$

The surface minimum location is the stereo displacement estimate for each pixel at the corresponding scale. The curvature parameters $\alpha$, $\beta$ and $\gamma$ produce a curvature matrix associated with each displacement estimate which forms the basis of a confidence measure obtained from the corresponding eigenvectors.

The same process is repeated at increasing resolutions to refine the estimates, by producing a cumulative squared difference (CSD) surface, with the estimates from level $s$ as a starting point.

$$CSD^{(j)}(n, c) = \begin{cases} 
SD^{(j+1)}(n, c) + SD^{(s)}(n, c), & s < s_{\text{max}} \\
SD^{(j)}(n, c), & s = s_{\text{max}}
\end{cases}$$

The level $s_{\text{max}}$ must be large enough to encompass the largest expected displacement. The CSD is scaled at each iteration to compensate for the required interpolation between scales.

4. RESULTS

The proposed method was evaluated for a number of retinal images with oedemas, the locations of which were known as a result of optical coherence tomographic (OCT) data which provides a measurement of the retinal thickness at the fovea. A number of enhancements were applied to the core algorithm based on various thresholds, including: maximum displacement estimates up to 0.35, a confidence measure associated threshold of 0.9 and bilinear interpolation.

The images displayed here are associated with the green colour plane of the RGB format retinal images (indicated in Figure 7). It must be noted that the three colour planes...
produce different outcomes, corresponding to the strength and visibility of the oedema in the respective planes. From our evaluations the green and blue planes produce the clearest outcomes, and the combination of results from all three planes can also be advantageous.

The results displayed in Figures 9 and 10 correspond to the cropped area indicated in the original retinal image (Figure 7). As seen in Figures 9 and 10, the consistency of the raised area corresponds to the known location of the oedema as identified in the OCT image in Figure 8. The detected oedema is represented as a disjointed mass and also by the border of the oedema rather than an entirely raised mass. This is due to the way in which the oedema is observed in the retinal images as a raised semi-transparent region covering a significant area. With respect to the detection scheme, the appearance at the centre of the oedema does not necessarily vary in the respective stereo images. The border regions are most easily detected as these will show the greatest detectable variations between the images. Also, the two images only differ in their acquisition by a horizontal shift (approximately horizontal, assuming minimal eye movement between image captures). So the edges of the oedema approximately parallel to the shift are expected to be less observable than the perpendicular edges.

The results presented here demonstrate the potential for such an approach to provide useful and accurate information with respect to the presence and location of retinal oedemas. The approach applied here takes advantage of the scale and spatial localization properties of the wavelet domain, while also utilizing the phase properties observable through the use of the DT-CWT. This phase based approach overcomes problems such as edge detection, which are responsible for the presence of retinal oedemas and progressively refines these initial estimates with higher spatial resolution information, thereby incorporating low pass filtering in the detection process to reduce the impact of high frequency image structures such as blood vessels and noise artifacts. This is significant as blood vessels can be observed both above and below the oedema, and so cannot be utilized as a robust detection feature.

5. CONCLUSIONS AND FUTURE WORK

This paper presents a selective relief height measure (corresponding to retinal thickness) from a stereo pair of retinal images. The technique is frequency selective and can thus be tailored to features of interest that are represented by specific components of the image in order to detect the stereo displacement. For the accurate detection of macular oedema, the image registration step is critical in correctly aligning the images while preserving the displacement resulting from the oedema itself.

Future investigation is necessary to evaluate the precision of the displacement measure, i.e. the maximum error obtained for a given displacement of the photographs. A database of images with OCT ground-truth is necessary. Further work is also required to improve the robustness of the detection technique, particularly in addressing the effect of localized illumination variations that can occur during image acquisition. In such cases the shape and position of such an effect can be misconstrued as an oedema.

The proposed technique effectively represents height based on stereo displacement, where different frequency components are necessary to measure the displacement. Importantly, it was demonstrated on the detection of macular oedema in stereo pairs of retinal photographs, where block matching algorithms fail.

REFERENCES