FUSION OF MULTISPECTRAL RETINAL IMAGES

Ing. Libor KUBEČKA, Doctoral Degree Programme (2)
Dept. of Biomedical Engineering, FEEC, BUT
E-mail: xkubec01@stud.feec.vutbr.cz

Supervised by: Prof. Jiří Jan

ABSTRACT

This contribution refers to the fusion of 2D multi-spectral images. Short review of the currently used methods is given and an algorithm using the wavelet representation of vector-valued is discussed. The pyramidal dyadic wavelet transform is modified according to the definition of maximal length of the multi-valued gradient and a solution of problems with orientation of the computed orthogonalized gradient is given. Several applications of the method including multimodal ophtalmic images fusion and multi-focused color photographs creation are presented and the suitability of the fusion for subsequent processing is discussed.

1 INTRODUCTION

As a result of rapid development of new imaging modalities, the demand for simultaneous processing of several of these images has been arisen in past few decades. Especially the area of medical imaging produces many types of images of the same scene e.g. MRI images taken by means of various parameters (T1, T2, …), CT and PET images, images enhanced by contrast injection. Another type of multi-spectral images is provided by remote-sensing or infrared imaging. Nowadays, processing of these multi-spectral images is usually done by band by band processing, but this approach has several disadvantages as if the inter-band dependencies are not taken into account. For this reason, the fusion can be done to obtain scalar image and then usual processing can be performed or processing of the image is carried out by generalized vector methods. In [1] it is shown that principle component analysis (PCA) is a relevant method for merging remotely sensed imagery because of its ability to reduce the dimensionality of the original data from n to 2 or 3 transformed principal component images, which contain majority of information. In [2] the wavelet-transform-based image fusion is presented, where to-be-fused images are first decomposed using discrete wavelet transform (DWT), then a fusion decision map is generated based on a set of fusion rules and the fused wavelet coefficient map is constructed from the wavelet coefficients of the source images according to the fusion decision map. Finally the fused image is obtained by performing the inverse wavelet transform. The method for merging high-resolution panchromatic image (SPOT) and a low-resolution multi-spectral image (LANDSAT) presented in [3] consists of adding the wavelet coefficients of the high-resolution image to the multi-spectral image data. The method is compared to standard intensity-hue-saturation (IHS) method. The algorithm using the orthogonal gradient representation of multi-spectral images and its applications in image merging and color image demosaicing is described in [4]. In this paper, we present this slightly modified method for automatic fusion of images acquired from two imaging modalities, namely of the confocal scanning laser ophtalmoscope (CSLO) image and color photographic image, sometimes called fundus photographs. This fusion is done with expectation of benefits to consequential segmentation of the optic disk in fused vector valued image data. This segmentation is necessary pre-step for evaluating retinal images and also for better diagnosing of glaucoma and cerebro-vascular diseases. Nowadays, this segmentation is performed by medical experts but there are several serious problems. The first problem is the inter- and intra-operator variability, the second one is unsatisfactory sensitivity and specificity of individual diagnostic methods and finally there is a problem of increasing costs caused by ineffective burdening of medical experts. For this reason the automatic segmentation is necessary especially for glaucoma screening where large amount of people should be examined.

1.1 METHODS

Image fusion is defined as creating a single image from a set of input images. The fused image should have more complete information which is more useful for human or machine perception. Image fusion can improve reliability (by redundant information) and capability (by complementary information) of subsequent processing.

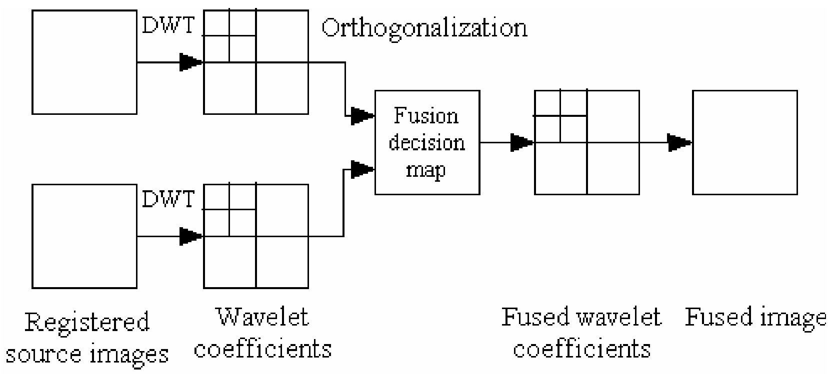


Fig. 1: *Illustration of image fusion making use of wavelet transform*

The image fusion can be done making use of discrete wavelet decomposition, where images are fused by merging their wavelet coefficients. This merging can be based on creating sc. Fusion decision map [2] or the wavelet representation for multi-valued images is made by combining detail (edge) information contained in vector-image bands into one set of details using DiZenzo's first fundamental form [4]. It means that subsequent processing of the scalar image will take into account an advantage of processing all the vector image bands simultaneously. This algorithm is exploited and described in this paper. For obtaining the image in image domain the pyramidal reconstruction scheme can be applied (see **Chyba! Nenalezen zdroj odkazů.**).

The first fundamental form

Let $\mathbf{I}(x,y)$ be a vector valued image with components $I_n(x,y)$, $n = 1, \ldots, N$ and let $\mathbf{L}$ is a vector operator whose components are both linear operators L_x, L_y defined as isotropic convolution operators applied in direction x or y. Examples of such a operator can be gradient operator $\nabla = (\partial / \partial x, \partial / \partial y)$ or Gaussian gradient filter. Some proposals of this operator have been done like $\mathbf{L}(\mathbf{I}) = \frac{1}{N}\sum_n \mathbf{L}(I_n)$ or $L(I) = L(I_j),\ j = \max_n \|\mathbf{L}(I_n)\|^2$. However both procedures do not take the bands into account simultaneously hence there is problem with averaging opposite vectors, which could annihilate despite that they should have some information. For that reason multi-valued 'maximal length' and 'direction of maximal length' were defined making use of quadratic form of the image differential [4]. For the norm of the differential of scalar product $\mathbf{L}(\mathbf{I})\cdot d\mathbf{x} = L_x(\mathbf{I})dx + L_y(\mathbf{I})dy$ can be written:

$$\begin{aligned}(\mathbf{L}(\mathbf{I})\cdot d\mathbf{x})^2 &= \begin{pmatrix} dx \\ dy \end{pmatrix}^T \cdot \begin{pmatrix} \sum (L_x(I_n))^2 & \sum (L_x(I_n))(L_y(I_n)) \\ \sum (L_y(I_n))(L_x(I_n)) & \sum (L_y(I_n))^2 \end{pmatrix} \cdot \begin{pmatrix} dx \\ dy \end{pmatrix} \\ &= \begin{pmatrix} dx \\ dy \end{pmatrix}^T \cdot \begin{pmatrix} G_{xx} & G_{xy} \\ G_{yx} & G_{yy} \end{pmatrix} \cdot \begin{pmatrix} dx \\ dy \end{pmatrix} = d\mathbf{x}^T \cdot G \cdot d\mathbf{x}\end{aligned} \tag{1}$$

The matrix G is symmetric and semi positive-definite hence its eigenvalues are real and non-negative. The quadratic form represents changes in vector-valued image. The direction of maximal change is defined by the eigenvector θ^1 of the matrix G corresponding to the maximal eigenvalue λ^1 representing the maximal squared length of the vector $\mathbf{L}(\mathbf{I})$ whereas the second eigenvector lies in the orthogonal direction of maximal length and the corresponding eigenvalue denotes minimal length the matrix.

For the purpose of fusing images it can be defined:

$$\mathbf{L}(\mathbf{I}(x,y)) \equiv \begin{pmatrix} \sqrt{\lambda^1}\theta_x^1(x,y) \\ \sqrt{\lambda^1}\theta_y^1(x,y) \end{pmatrix} \tag{2}$$

For the fusion of the vector-valued image we used method based on discrete orthogonal wavelet decomposition [4] described on the (**Chyba! Nenalezen zdroj odkazů.**). By this transform, an image $f(x,y)$ is decomposed at different scales j into a lower resolution image $A_{2^j} f$ (sc. approximation) and three detail images $D_{2^j}^1 f$, $D_{2^j}^2 f$, $D_{2^j}^3 f$. This is done by uniform down-sampling of 2D convolution products. At each scale $j \in [-1, J]$, $(n,m) \in Z^2$ we can write:

$$\begin{aligned} A_{2^j} f &= f(x,y) * \phi_{2^j}(-x)\phi_{2^j}(-y)(2^{-j}n, 2^{-j}m) \\ D_{2^j}^1 f &= f(x,y) * \phi_{2^j}(-x)\psi_{2^j}(-y)(2^{-j}n, 2^{-j}m) \\ D_{2^j}^2 f &= f(x,y) * \psi_{2^j}(-x)\phi_{2^j}(-y)(2^{-j}n, 2^{-j}m) \\ D_{2^j}^3 f &= f(x,y) * \psi_{2^j}(-x)\psi_{2^j}(-y)(2^{-j}n, 2^{-j}m) \end{aligned} \tag{3}$$

Where ϕ and ψ are separable low and high bandpass filters. This wavelet transform can be computed by gradual applying of 1D quadrature mirror filters Lo_D (lowpass) and Hi_D (highpass) for decomposition and symmetric filters Lo_R, Hi_R for reconstruction (see **Chyba! Nenalezen zdroj odkazů.**). Let L be a linear vector operator:

$$\mathbf{L} = \begin{pmatrix} Hi\_D_x \\ Hi\_D_y \end{pmatrix} * \tag{4}$$

Vector image I is then decomposed into wavelet coefficients according to the schema on **Chyba! Nenalezen zdroj odkazů.**. Detail images on each scale level are computed by filtering relevant approximations using Hi_D filter, then the orthogonalization (1) is done and a new gradient representation is obtained using (2), then final down-sampling and filtering is performed defining fused detail images at current scale level. The approximation can be computed as averages of each band approximations depending on given application. There is a problem with diagonalization caused by not uniquely specified signs of the eigenvectors. An extensive study and solution of this problem is given in [5]. In the [4] the direction of maximal length is compared to the average direction, obtained by applying gradient operator on all bands separately and if directions of these vectors are different, the maximal length direction is flipped into average direction. Due to problems with annihilation of opposite bands while averaging we have modified this approach. We compare the direction of maximal length with the direction of the gradient of the first component acquired from the PCA analysis of approximations of each band at the current position. For this case the scalar product of these two vectors is evaluated and the direction of the eigenvectors is maintained if the product is positive, else the eigenvectors with the opposite direction are applied in subsequent computations. The final image representation is obtained with inverse wavelet transform.

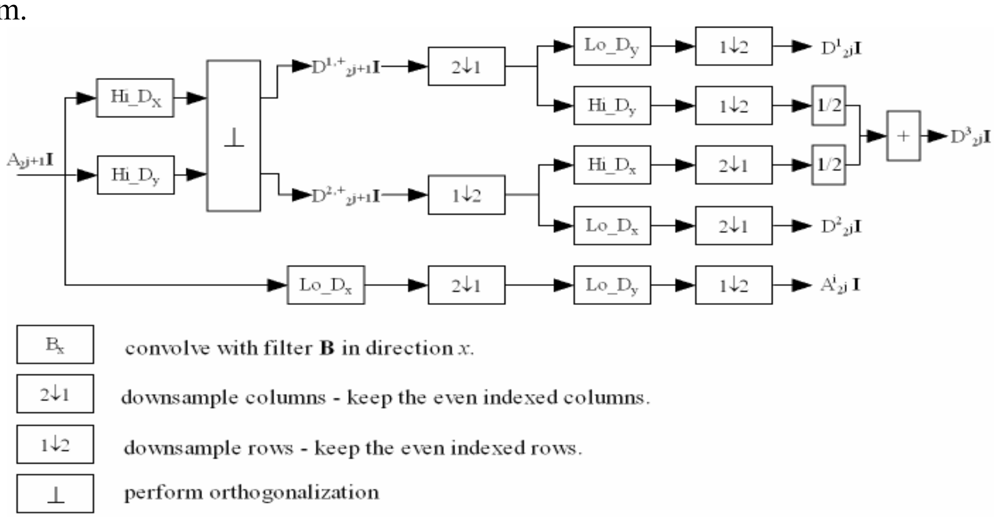


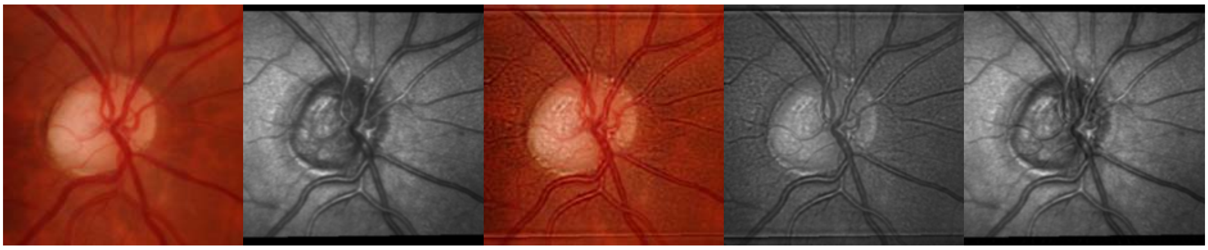
Fig. 2: *Orthogonal wavelet decomposition schema.*

2 EXPERIMENTS AND RESULTS

The method was tested on images acquired from the Heildeberg Retina Tomograph II (HRTII) and the standard color fundus camera Kowa. Both images were previously registered. Several results of the fusion depending on the final representation of the approximation coefficients can be seen on the Fig 3. The suitability of these results has to be consulted with our medical experts, but it seems that some less distinct but important objects like neuroretinal-rim (a dark ring within the optic nerve head) have disappeared during the fusion. This makes the subsequent segmentation of these structures impossible. For that reason the processing of separate-bands or the direct vector processing of such images seems to be better choice. The application of the method for merging of multi-focused color photographs is presented on the Fig 4. This can also be very useful for the case of increasing depth of focus of the microscopic images.

3 CONCLUSION

A method for merging vector images into scalar image has been described in this paper. Two application of the method have been discussed and it was observed that image fusion seems not to be appropriate method for subsequent processing of vector-valued retinal images. On the other hand the successful application of this method - merging of multi-focused images has been described.



A B C D E

Fig. 3: *A - color fundus photograph; B - HRT image; C - HRT image fused into the color photograph; D - image created by fusion of 4 bands image (R, G, B, HRT) , final approximation coefficients are computed as average of particular bands; E - fusion of the four bands image into the scalar image where the HRT image was taken as the final approximation coefficients.*

Fig. 4: *On the left is foreground focused image, in the middle is background focused image and on the right is the image fused using proposed algorithm.*

ACKNOWLEDGEMENTS

The paper has been prepared as a part of the solution of MŠMT project No. 3118/2005 and with the support of the research plan MSM 2622000111.

REFERENCES

[1] Chavez, P. S., Kwarteng, A. Y.: Extracting spectral contrast in Landsat Thematic Mapper image data using selective principal component analysis, Photogrammetric Engineering and Remote Sensing, 55(3): 339-348.

[2] Zhang, Z., Blum, R. S.: Multisensor Image Fusion using a Region-Based Wavelet Transform Approach, Proc. of the DARPA IUW, pp. 1447-1451, 1997.

[3] Nunez, J., Otazu, X.: Multiresolution-based image fusion with additive wavelet decomposition, Trans. on Geoscience and Remote Sensing, vol.: 37 Is.: 3, 1999, pp: 1204

[4] Scheunders, P.: An orthogonal wavelet representation of multivalued images, IEEE Transactions on image processing, Volume: 12, Issue: 6, June 2003, Page(s): 718- 725.

[5] Cumani, A.: Edge detection in multispectral images, CVGIP Graphical models and Image Processing, vol. 53, pp. 40-51, 1991

[6] Kubecka, L., Jan, J.: Registration of Bimodal Retinal Images - improving modifications In Proc. 26th Annual Int. Conf. of the IEEE EMBS, San Francisco, 2004, pp. 1695