Rigid Image Registration based on Pixel Grouping

Demetrios Gerogiannis * Christophoros Nikou Aristidis Likas

University of Ioannina,
Department of Computer Science,
PO Box 1185, 45110 Ioannina, Greece,
{dgerogia,cnikou,arly}@cs.uoi.gr

Abstract

We propose a pixel similarity-based algorithm enabling accurate rigid registration between single and multimodal images. The method relies on the partitioning of a reference image by a Gaussian mixture model (GMM). This partition is then projected onto the image to be registered. The main idea is that a Gaussian component in the reference image corresponds to a Gaussian component in the image to be registered. If the images are correctly registered the total distance between the corresponding components is minimum. An advantage of the proposed method is that it may handle multidimensional (vector valued) images where histogram-based methods such as the widely used mutual information is not tractable due to the high dimension of the data. Also, experimental results indicate that, even in the case of images presenting low SNR, the proposed algorithm compares favorably to the histogram-based mutual information method that is widely used in a variety of applications.

1 Introduction

The goal of image registration is to geometrically align two or more images in order to superimpose pixels representing the same underlying structure. Image registration is an important preliminary step in many application fields involving, for instance, the detection of changes in temporal image sequences or the fusion of multimodal images. For the state of the art of registration methods we refer the reader to [3, 22]. Medical imaging, with its wide variety of sensors (MRI, nuclear, ultrasonic, X-Ray) is probably one of the first application fields [14, 7]. Other research areas concerned by image registration are remote sensing, multisensor robot vision and multisource imaging used in the preservation of artistic patrimony. Respective applications include the following of the evolution of pathologies in medical image sequences [17], the detection of changes in urban development from aerial photographs [11] and the recovery of underpaintings from visible/X-ray pairs of images in fine arts painting analysis [8].

The overwhelming majority of change detection or data fusion algorithms assume that the images to be compared are perfectly registered. Even slightly erroneous registrations may become an important source of interpretation errors when inter-image changes have to be detected. Accurate (i.e. subpixel or subvoxel) registration of single modal images remains an intricate problem when gross dissimilarities are observed. The problem is even more difficult for multimodal images, showing both localized changes that have to be detected and an overall difference due to the various responses of multiple sensors.

Since the seminal works of Viola and Wells [21] and Maes *et al.* [13] the maximization of the mutual information measure between a pair of images has gained an increasing popularity as a criterion for image registration [18]. The estimation of both marginal and joint probability density functions of the involved images is a key element in mutual information based image alignment. However, this method is limited by the histogram binning problem. Approaches to overcome this limitation include Parzen windowing [21, 10], where we have the problem of kernel width specification, and spline approximation [20, 15]. A recently proposed method relies on the continuous representation of the image function and develops a relation between image intensities and image gradients along the level sets of the respective intensity [19].

Gaussian mixture modeling [2, 16] constitutes a powerful and flexible method for probabilistic data clustering that is based on the assumption that the data of each cluster has been generated by the same Gaussian component. In [12] GMMs were trained off-line training to provide prior infor-

*This work was partially supported by Interreg IIIA (Greece-Italy) grant I2101005.

mation on the expected joint histogram when the images are correctly registered. GMMs have also been successfully used as models for the joint [6] as well as the marginal image densities [9], in order to perform intensity correction.

Inspired by the application of GMMs to image segmentation and intensity correction, we apply GMM modeling to image registration. More specifically, we train a GMM model once for the reference image and obtain the corresponding partitioning of image pixels into clusters. Each cluster is represented by the parameters of the corresponding Gaussian component. The main idea is that a Gaussian component in the reference image corresponds to a Gaussian component in the image to be registered. If the images are correctly registered the sum of the distances between the corresponding components is minimum.

It is well-known that GMMs overcome the binning problem of histogram-based methods and provide a continuous model of the image density. When successfully trained, they produce a sensible approximation of the pdf of image intensity, by placing Gaussian components in a sensible *data-driven* way (i.e on intensity regions exhibiting high density). Although there is still the problem of specifying the number of components in GMM modeling, experimental results indicate that our GMM-based method is quite robust from this point of view, provided that the number of components is neither very big (overfitting) nor very small (underfitting).

In the remainder of this paper, we present the proposed registration method in section 2. Experimental results are provided in section 3 and conclusions are drawn in section 4.

2 Registration by minimization of density distances

Consider the multivariate normal distributions $N_1(\mu_1, \Sigma_1)$ and $N_2(\mu_2, \Sigma_2)$ and denote $\Theta_i = \{\mu_i, \Sigma_i\}$, with $i = \{1, 2\}$, their respective parameters (mean vector and covariance matrix). The Chernoff distance between these distributions is defined as [5]:

$$C(\Theta_1, \Theta_2, s) = \frac{s(1-s)}{2} \Delta\mu^T [s\Sigma_1 + (1-s)\Sigma_2]^{-1} \Delta\mu + \frac{1}{2} \ln \left(\frac{|s\Sigma_1 + (1-s)\Sigma_2|}{|\Sigma_1|^s |\Sigma_2|^{1-s}} \right), \tag{1}$$

where $\Delta\mu = \mu_2 - \mu_1$. The Bhattacharyya distance is a special case of the Chernoff distance with $s = 0.5$:

$$B(\Theta_1, \Theta_2) = \frac{1}{8} \Delta\mu^T \left[\frac{\Sigma_1 + \Sigma_2}{2} \right]^{-1} \Delta\mu + \frac{1}{2} \ln \left(\frac{|\frac{\Sigma_1+\Sigma_2}{2}|}{\sqrt{|\Sigma_1||\Sigma_2|}} \right) \tag{2}$$

The Bhattacharyya distance may still be used even if the underlying distributions are not Gaussian. However, for distributions deviating markedly from a Gaussian, the distance will not be informative.

Let I_{ref} be an image of $N \times N$ pixels with intensities denoted as $I_{ref}(x^i)$, where x^i, $i = 1, ..., N^2$, is the i^{th} pixel. The purpose of rigid image registration is to estimate a set of parameters $\mathcal{S}$ of the rigid transformation $T_{\mathcal{S}}$ minimizing a cost function $E(I_{ref}(\cdot), I_{reg}(T_{\mathcal{S}}(\cdot)))$ that, in a similarity metric-based context, expresses the similarity between the image pair. In the 2D case the rigid transformation parameters are the rotation angle and the translation parameters along the two axes. In the 3D case, there are three rotation and three translation parameters. Eventually, scale factors may also be included, depending on the definition of the transformation.

Consider, now, a partitioning of the reference image I_{ref} into K clusters (groups) using a GMM obtained via the EM algorithm [2]. Therefore, the reference image is represented by the parameters $\Theta_k^{ref} = \{\mu_k^{ref}, \Sigma_k^{ref}\}$, $k = 1, \dots, K$ of the GMM components. The partitioning of the image is described using the function $f(x) : [1, 2, ..., N] \times [1, 2, ..., N] \rightarrow \{1, 2, ..., K\}$, where $f(x) = k$ means that pixel x of the reference image I_{ref} belongs to the cluster defined by the k^{th} component. Let us also define the sets of all pixels of image I_{ref} belonging to the k^{th} cluster:

$$P_k = \{x^i \in I_{ref}, i = 1, 2, ..., N^2 / \delta(f(x^i) - k) = 1\},$$

for $k = 1, ..., K$, where $\delta(x)$ is the Dirac impulse:

$$\delta(f(x^i) - k) = \begin{cases} 1, & \text{if } f(x^i) = k \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

The above GMM-based segmentation of the reference image is performed once, at the beginning of the registration procedure. The reference image I_{ref} is, thus, partitioned into K clusters, generally, not corresponding to connected components in the image. This spatial partition is projected on the image to be registered I_{reg}, yielding the same partition of this second image (i.e., the partitioning of the reference image acts as a mask on the image to be registered). Then, we assume that the pixel values of each cluster k in I_{reg} are modeled using a Gaussian component with parameters the sample mean μ_k^{reg} and the sample covariance matrix Σ_k^{reg}:

$$\mu_k^{reg} = \frac{1}{|P_k|} \sum_{i=1}^{N^2} I_{reg}(T_{\mathcal{S}}(x^i)) \delta(f(x^i) - k) \tag{4}$$

and

$$\Sigma_k^{reg} = \frac{1}{|P_k|} \sum_{i=1}^{N^2} (\Delta I_k^i)(\Delta I_k^i)^T \delta(f(x^i) - k), \tag{5}$$

where $\Delta I_k^i = I_{reg}(T_{\mathcal{S}}(x^i)) - \mu_k^{reg}$ and $|P_k|$ is the cardinality of set P_k. The role of $\delta(f(x^i) - k)$ in eq. (4) and (5) is to determine the support (the pixel coordinates) for the calculation of the mean and covariance. These parameters are computed *on the image to be registered* for the pixel coordinates belonging to the k^{th} class *on the reference image*. This also implies a Gaussian generative model for the components of I_{reg}.

The energy function we propose, is expressed by the sum of Bhattacharyya distances (2) between the corresponding GMM components in I_{reg} and I_{ref}:

$$E(I_{ref}(\cdot), I_{reg}(T_{\mathcal{S}}(\cdot))) = \sum_{k=1}^{K} \pi_k B(\Theta_k^{ref}, \Theta_k^{reg}) \qquad (6)$$

where π_k is the mixing proportion of the k^{th} component:

$$\pi_k = \frac{|P_k|}{\sum_{k=1}^{K} |P_k|}.$$

If the two images are correctly registered the criterion in (6) assumes that the total distance between the whole set of components would be minimum.

For a given set of transformation parameters $\mathcal{S}$, the total energy between the image pair is computed through the following steps:

- segment the reference image $I_{ref}(\cdot)$ into K clusters by a Gaussian mixture model.
- for each cluster $k = 1, 2, ..., K$ of the reference image:
  - project the pixels of the cluster onto the transformed image to be registered $I_{reg}(T_{\mathcal{S}}(\cdot))$.
  - determine the mean μ_k^{reg} (4) and covariance Σ_k^{reg} (5) of the projected partition of I_{reg}.
- evaluate the energy in eq. (6) by computing the Bhattacharrya distances (2) between the corresponding densities.

2.1 Optimization

The Iterated Conditional Modes (ICM) [1] algorithm was implemented for the minimization of the energy function as (6) is highly non-linear. ICM is a deterministic Gauss-Seidel like algorithm, that only accepts configurations decreasing the cost function and has fast convergence properties. The overall optimization algorithm may be summarized as follows:

- ICM $(\mathcal{S}_0, E, max)$

$\mathcal{S}_0$	initial parameters
E	energy function
ϵ	minimum change in energy between iterations
max	maximum number of iterations

  - all parameters are declared unvisited ;
  - $\mathcal{S} \leftarrow \mathcal{S}_0$; $k \leftarrow 0$ (counter) ;
  - **while** ($k \leq max$ or $\Delta E > \epsilon$) **do**:
    - ∗ **while** there are unvisited parameters **do**:
      - · randomly chose un element $\mathcal{S}_k(i)$ of $\mathcal{S}_k$ at iteration k ;
      - · construct several configurations different from $\mathcal{S}_k$ only by the element $\mathcal{S}_k(i)$;
      - · keep the configuration giving the minimum energy;
      - · declare parameter $\mathcal{S}_k(i)$ visited
    - ∗ **end do**
  - reduce parameter search space
  - declare all parameters unvisited
  - $k \leftarrow k + 1$
  - **end do**

A large number of interpolations are involved in the registration process. The accuracy of the rotation and translation parameter estimates is directly related to the accuracy of the underlying interpolation model. Simple approaches such as the nearest neighbor interpolation are commonly used because they are fast and simple to implement, though they produce images with noticeable artifacts. More satisfactory results can be obtained by small-kernel cubic convolution techniques. In our experiments, we have applied a bilinear interpolation scheme, thus preserving the quality of the image to be registered.

Finally, let us notice that the energy in (6) may be applied to both single and multimodal image registration. In the latter case, the difference in the mean values of the distributions in (6) should be ignored, as we do not search to match the corresponding Gaussians in position but in shape. This also stands for the single modal case if the intensities of the image pair have significantly different contrasts.

3 Experimental results

In order to evaluate our method, we have performed a number of experiments in a relatively difficult registration problem by aligning a range image to its intensity counterpart. The image in fig. 1(a) is a synthetic stereo image composed of natural objects and the image in fig. 1(b) its corresponding range image (these images are reproduced and used in the experiments by courtesy of the Computer Science department, University of Bonn, Germany). The complimentary but not redundant information carried by the image pair augments the difficulty of the registration process.

The intensity image (fig. 1(a)) underwent several rigid transformations by different rotation angles and translation

14th International Conference on Image Analysis and Processing (ICIAP 2007)
0-7695-2877-5/07 $25.00 © 2007 IEEE

Table 1. Statistics on the registration errors for the compared methods. A total number of 250 registrations was performed for each method. The errors are expressed in pixels.

Registration error statistics		
	MI	GMM
Mean	0.63	0.20
St. dev.	0.78	0.02
Median	0.40	0.21
Max	2.85	0.31

parameters and was corrupted by white Gaussian noise in order to obtain signal to noise ratios (SNR) varying between 5 dB and 1 dB. The resulting images were provided as input to our GMM-based registration algorithm with a predefined number of Gaussian components ($K = 2, ..., 10$ and $K = 16$). The total number of registrations is 250 (5 different rigid transformations, 5 levels of SNR, 10 different clusters).

Registration errors were computed in terms of pixels and not in terms of transformation parameters. Registration accuracies in terms of rotation angles and translation vectors are not easily evaluated due to parameter coupling. Therefore, the registration errors are calculated at the corners of the image frame where their values are larger with respect to their values in the center of the image. More precisely, these values represent the average error in pixels of the four corners of the image frame with respect to the ground truth position.

Table 1 summarizes the registration errors for the different configurations of the registration experiments. For comparison purposes, the registration errors provided by the mutual information (MI) method are also shown. As it can be observed, our GMM-based registration clearly outperforms the standard method. This is more pronounced in the case of low SNR. Table 2 shows the mean values of the registration errors when the intensity image was corrupted by white Gaussian noise, thus, obtaining a SNR of 1 dB. A representative registration example, for both methods, with the corresponding registration errors is illustrated in fig. 2.

An important advantage of the proposed method is that it may handle multidimensional (vector valued) images where mutual information is not tractable due to the high dimension of the data. Modern imaging techniques provide an array of imaging modalities which enable the acquisition of complementary information, representing for instance an underlying anatomy in the case of multimodal medical images (MRI, SPECT, PET etc.). Also, into the same modality, an image pixel (or voxel) may be vector valued. For instance, a possible application of our registration technique is

Table 2. t Mean value of the registration errors for different numbers of clusters (K) in the case of noise with SNR of 1 dB.

Mean error (SNR = 1 dB)		
Clusters	MI	GMM
$K = 2$	0.87	0.55
$K = 3$	2.10	0.49
$K = 4$	0.83	0.45
$K = 5$	0.75	0.41
$K = 6$	0.67	0.38
$K = 7$	0.64	0.38
$K = 8$	0.73	0.38
$K = 9$	0.75	0.44
$K = 10$	0.78	0.47
$K = 16$	0.86	0.48

the registration of diffusion tensor magnetic resonance images (DT-MRI) where each pixel is 6-dimensional. In that case, the computation of image histograms and the application of mutual information is prohibitive. Other applications include the registration of color (RGB) images or images where the extraction of multiple features is necessary (e.g. textured images).

4 Conclusion

We have presented a method for the registration of single and multimodal images. The method relies on the minimization of distances between probability density functions defined by partitioning the two images. The first image is partitioned by a GMM through the EM algorithm. This partition is then implied onto the second image. We have shown the effectiveness and accuracy of the method especially with images presenting dissimilarities, such as the range image pair, where the mutual information method fails to correctly register the two images.

Important open questions for GMM-based registration are how the number of model components can be selected automatically and which features, apart from image intensity, should be used. These questions are still the subject of on going research in our group [4]. Another future work direction is the employment of mixtures with robust density functions that are expected to provide better results in the case of noisy data.

References

[1] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society*, 48(3):259–302, 1986.
[2] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

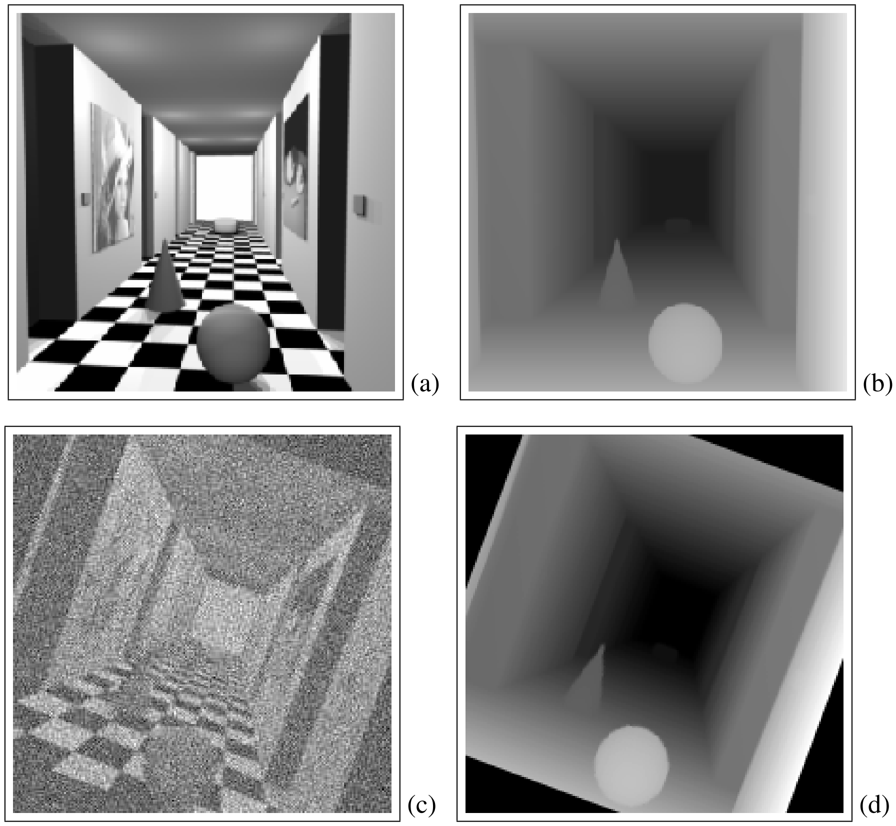

Figure 1. (a) A grey level image and (b) its disparity map. (c) The image in (a) rotated by 20.3 degrees and corrupted by white Gaussian noise in order to obtain a SNR of 1 dB. (d) The corresponding range image rotated by the same angle of 20.3 degrees.

[3] L. G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, 1992.

[4] C. Constantinopoulos, M. K. Titsias, and A. Likas. Bayesian feature and model selection for Gaussian mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(6):1013–1018, 2006.

[5] K. Fukunaga. *Introduction to statistical pattern recognition*. Academic Press, 1990.

[6] A. Guimond, A. Roche, N. Ayache, and J. Meunier. Three-dimensional multimodal brain warping using the demons algorithm and adaptive intensity corrections. *IEEE Transactions on Medical Imaging*, 20(1):58–69, 2001.

[7] J. V. Hajnal, D. L. G. Hill, and D. J. Hawkes. *Medical image registration*. CRC Press, 2001.

[8] F. Heitz, H. Maître, and C. de Couessin. Event detection in multisource imaging: application to fine arts painting analysis. *IEEE Transactions on Acoustic, Speech and Signal Processing*, 38(4):695–704, 1990.

[9] P. Hellier. Consistent intensity correction of MR images. In *Proceedings of the 2003 IEEE International Conference on Image Processing (ICIP'03)*, volume 1, pages 1109–1112, Barcelona, Spain, 2003.

[10] G. Hermosillo and O. Faugeras. Dense image matching with global and local statistical criteria. In *Proceedings of the 2001 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'01)*, volume 1, pages 73–78, Los Alamitos, USA, 2001.

[11] J. R. Jensen. *Introductory digital image processing: a remote sensing perspective*. Prentice Hall, 3rd edition, 2004.

[12] M. Leventon and E. Grimson. Multi-modal volume registration using joint intensity distributions. In *Proceedings of the Medical Image Computing and Computer Assisted Intervention Conference (MICCAI'98)*, pages 1057–1066, Cambridge, Massachussetts, USA, October 1998.

[13] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997.

[14] J. B. A. Maintz and M. A. Viergever. A survey of medical image registration techniques. *Medical Image Analysis*, 2(1):1–36, 1998.

[15] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank. Nonrigid multimodality image registration. In K. Hanson, editor, *Proceedings of the 2001 SPIE Medical Imaging Conference*, volume 4322, pages 1609–1620, San Diego, USA, 2003.

[16] G. McLachlan. *Finite mixture models*. Wiley-Interscience, 2000.

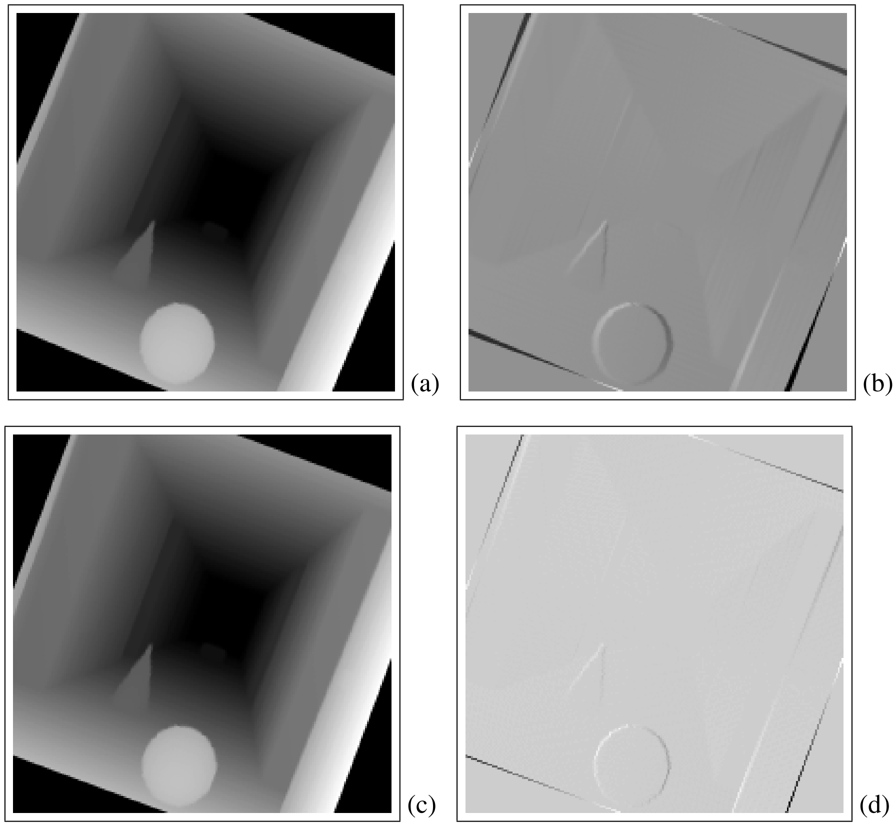

Figure 2. The range image in fig. 1(b) was registered to the grey level image in fig. 1(c) by the standard mutual information and our GMM-based methods. The registered images and the difference between the ground truth image in fig. 1(d) and the registered images are presented. (a) Registration by mutual information and (b) its difference from the ground truth. (c) Registration by our GMM method and (d) its difference from the ground truth. In the difference images, the intensity varies between 0 and 255 with positive values being lighter, negative values being darker. Grey level zero is represented by a value of 128.

[17] C. Nikou, F. Heitz, and J. P. Armspach. Robust voxel similarity metrics for the registration of dissimilar single and multimodal images. *Pattern Recognition*, 32:1351–1368, 1999.

[18] J. Pluim, J. B. A. Maintz, and M. Viergever. Mutual information based registration of medical images: a survey. *IEEE Transactions on Medical Imaging*, 22(8):986–1004, 2004.

[19] A. Rajwade, A. Banerjee, and A. Rangarajan. A new method of probability density estimation with application to mutual information based image registration. In *Proceedings of the 2006 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'06)*, New York, USA, 2006.

[20] P. Thévenaz and M. Unser. Optimization of mutual information for multiresolution image registration. *IEEE Transactions on Image Processing*, 9(12):2083–2099, 2000.

[21] P. Viola and W. W. III. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2):137–154, 1997.

[22] B. Zitova and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):1–36, 2003.

14th International Conference on Image Analysis and Processing (ICIAP 2007)
0-7695-2877-5/07 $25.00 © 2007 IEEE