Image registration of follow-up examinations in digital dermoscopy

Christos Nikolaos E. Anagnostopoulos, Dimitrios D. Vergados and Panagiotis Mintzias

Abstract—**Automated assessment tools for dermoscopy images play an important assistive role mainly due to inter- and intra-observer variations in human interpretation. Prior to classification, a robust registration module is necessary. In this paper we propose a fully automatic method for successful registration (matching) of melanocytic nevi images acquired from digital dermoscope to ensure reliable quantitative comparison of the image later. A modified SIFT algorithm (called ROI-SIFT), followed by RANSAC to identify outliers, is used to localize and match correspondence interest points, that will be used to compute affine transformations for the registration of a dermatological image with the respective follow-up examination image. Experimental evaluations of the registration method are reported for image pairs acquired during follow-up examinations. Finally, a short discussion of scheduled future work is proposed.**

I. INTRODUCTION

Dermoscopy is an imaging technique dermatologists use to better visualize pigmented skin lesions (PSLs) and determine their malignancy. In digital dermoscopy, polarized light is used to cancel out skin surface reflections. When the images or video clips are digitally captured or processed under standardized conditions, the instrument can be referred to as a "digital epiluminescence dermatoscope". Moreover, sophisticated software for body mapping of lesions is used that allows easy storage, retrieval and management of examinations.

In order to provide a diagnosis for skin lesions, physicians follow a two-step procedure: (1) classify the lesion in to melanocytic and non-melanocytic type and (2) for the melanocyticones, classify in to benign and malignant lesions. In order to perform the above, various methods are commonly used, i.e. the ABCDE rule of dermoscopy, the7-point check list, and the Menzies method [1].

The acronym of the first method (ABCDE) offers a good mnemonic rule for malignant lesions, where A corresponds to asymmetry, B to borders (irregularity), C to color (non-uniform, more than one), D to diameter (greater than 6 mm – 0.24 inch), and E to evolution over time. The last parameter is equally important to ABCD, while the evolution (expansion or regression) is difficult to be assessed automatically. This is due to the fact that the registration of follow-up images is a necessary procedure for the detection of differences between the lesions being examined. The successful regitration of the two under comparison images will allow shape, area measurements and direct comparisons. This is essential for evolving regions exhibiting even the slightest color and shape differences.

In this paper, we propose a fully automatic method for successful registration (matching) of melanocytic nevi images acquired from a digital dermoscope. Obtaining digital images that are usually for reporting reasons, our aim is to create an algorithm that aligns consecutive images from follow-up examinations in order to measure, in a quantitative manner, possible alterations in melanocytic nevi.

Reviewing the related literature, image registration methods of various skin lesions have been presented in the literature. Hybrid methods using the log-polar transformation for estimating scaling and rotation parameters and a sign change similarity criterion in combination with a hill climbing optimization scheme for translation estimation are proposed in [2]. Maleti et al [3] work with statistical shape analysis after psoriasis segmentation to do a first rigid alignment under the assumption of image scale constancy. Afterwards, small internal displacements are corrected by a combined registration and alignment scheme. Delgado et al [4] segment the lesions in the images and use first and second order moments of the resulting binary images to assess rotation and translation. SIFT-based registration for hemangioma images is presented in [5]. Hemangiomas change appearance during the period of follow-up examinations, thus a robust approach is necessary. The authors use classical SIFT features only in the hemangioma area after manual elimination of SIFT features in the healthy skin area. Then, a homography estimation follows that eventually leads to successful registration.

Our problem has many common aspects with [5], since the evolution assessment over time for melanocytic nevi, implies that reliable matching (and thus registration) is only possible for a part of the nevi, due to difference in their appearance between examinations (see Fig.1). Contrary to [5], we choose not to exclude SIFT keypoints that lie outside the nivus area, since the digital camera of the dermatoscope offers standardized imaging conditions.

C.N.E. Anagnostopoulos is with the Cultural Technology and Communication Dpt., University of the Aegean, Mytilene, Lesvos, GR 81100, Greece, (Tel: +30 2251036624; fax: +30 2251036637; e-mail: canag@ct.aegean.gr).

D.D. Vergados is with the Informatics Dpt., Univeristy of Pireaus, 18534, Greece (e-mail: vergados@unipi.gr).

P. Mintzias is Director of the Dermatological Department of the General Hospital of Kalamata, 24100, Greece, (Email: petermintzias@gmail.com).

II. MATERIALS AND METHODS

More than 200 images from a total of 48 subjects were selected from the Dematological Unit of the General Hospital of Kalamata. For this study, a Mediscope digital videomicroscope with lens that provide manual focus and polarized LED light and the Optipix Dermatoscopy software

were used, which are built for management of dermatoscopy images. The system includes a sophisticated imaging platform for body mapping of nevi. The user may link dermatoscopy images to either a predefined body map or to a macro image. All nevi can be marked with a circle. The focus is to visualize the locations of the suspicious moles in an easy and accessible way (Fig. 2). One of the advantages of using this software is to store and manually compare follow up images with the previous image. By flagging a location as follow up the physician can easily access the nivus the next examination. However, the automatic comparison of follow-up images is not possible if those images are not registered accordingly [6]. Only this way, a reliable quantitative comparison method could be possible, where even the slightest variation can be identified and assessed. All of the magnifications of the instrumentation are pre-calibrated and can be used to measure important dimensions such as area, circumference, color variation etc.

Fig. 1 The same nivus captured in two consequtive examinations. The box in the follow-up image indicates difference in the nivus. The images are cropped for illustration purposes.
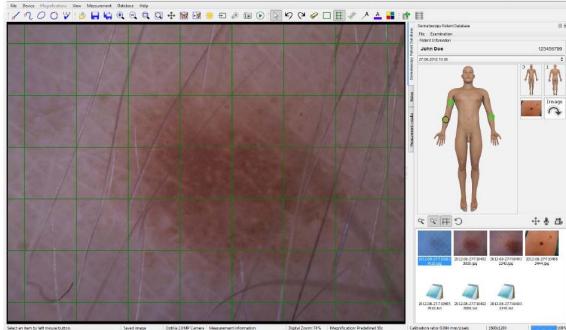

Fig. 2. The imaging softare (Optipix, Optilia).

III. METHOD

In this paper, a modified SIFT algorithm (called ROI-SIFT) is used to localize and match correspondence interest points, which will be used to compute affine transformations to map a dermatological image to the respective follow-up examination image. But, since the mapping procedure can be misled by erroneous correspondences, RANSAC is used to identify outliers and maximize the registration accuracy. Then follows rigid image registration, based on the best homography found.

We use rigid transformation method since it is global in nature and it does not model local geometric differences

between images. Thus, local geometric differences in nevi between follow-up images will be exposed successfully.
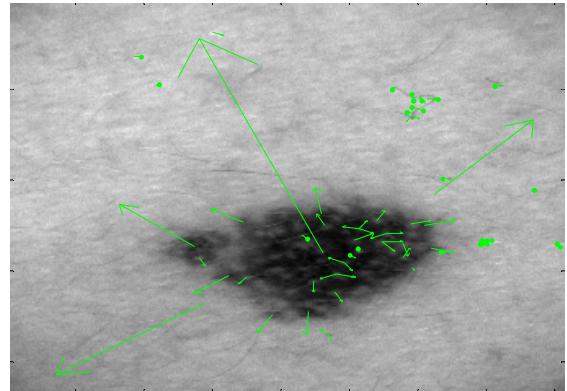

Fig. 3. Initial keypoints with h=0.01 in the left image of Fig. 1. The length of the arrow represent the scale of keypoint and the direction represents the orientation.
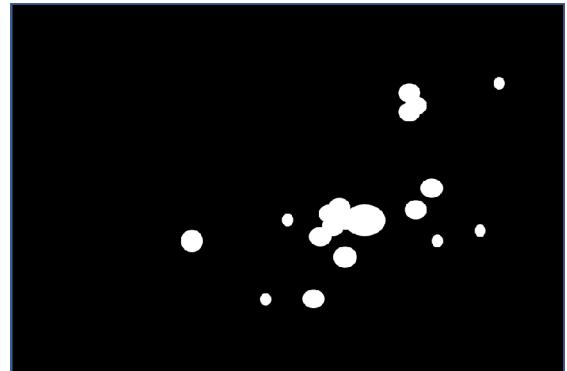
Fig. 4. Candidate RoIs for retrieving keypoints with h=0.03 in SIFT algorithm
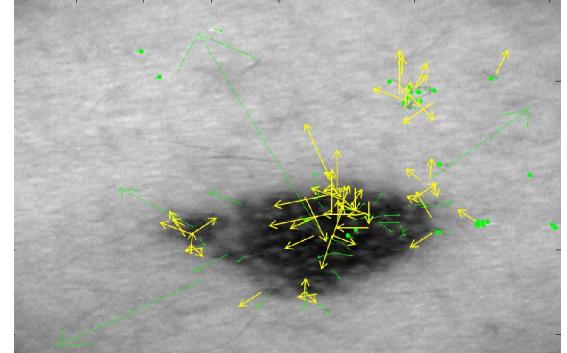
Fig. 5. Keypoints with: h=0.01 (green) and h=0.03 (yellow).

A. Salient feature detection using ROI-SIFT

SIFT [7] is the state-of-the-art, in the field of image recognition and is also used in a wide range of image registration and matching applications. It exploits the idea of replacing images by a set of scale and orientation invariant keypoint descriptors, using gradient orientation histograms. The invariant features are detected and extracted, exploring the scale-space structure of the image [8]. Features are localized and filtered, keeping only those that are likely to remain stable over affine transformations, having adequate contrast and are not along edges. The presence of keypoints not lying in the edges and having adequate contrast is

ensured with the appropriate selection of the parameters h in the Difference of Gaussians (DoG) function, as described analytically in [9]. The keypoint descriptor is created by sampling the magnitudes and orientations of the image gradient in a patch around the detected feature, resulting in a 128-D vector of direction histograms.

In this paper, the parameter h is not set with a constant value as in original SIFT. Specifically, the algorithm begins with a "hard" value for h (h=0.01) and the respective feature points are identified (Fig. 3). Then, according to the position, the scale and the orientation of the keypoints, we define the expanded Region of Interests (RoIs) (see Fig. 4). The next step implements the execution of the SIFT algorithm with a "soft" value for h (h=0.03) producing a larger number of keypoints (Fig. 5). In our method, only the keypoints belonging to the ROIs are kept, while those that fall outside the ROIs areas are eliminated.

B. Feature matching, validation and registration

Once the interest points have been localized, their SIFT description has to be computed and they have to be matched in a robust way in order to proceed to the registration. However, with SIFT descriptors we have correct and incorrect matches between the pair of images. A good way to tackle with this problem is to implement the RANdom SAmpling Consensus method [10] (RANSAC). The method achieves its goal by iteratively selecting a random subset of the original data. These data are hypothetical inliers and this hypothesis is then tested as follows: (a) a model is fitted to the sample of hypothetical inliers, (b) all other data are then tested against the fitted model and, those points that fit the estimated model well are considered as part of the consensus set, (c) the estimated model is reasonably good if sufficiently many points have been classified as part of the consensus set and (d) the model may be improved by re-estimating it using all members of the consensus set.

Despite the fact that RANSAC is capable of handling a large portion of incorrect matches, a pre-selection of the matches with highest confidence can improve the performance. Therefore, before RANSAC validation, for each feature in the follow-up image, the descriptor is used to search for its nearest neighbor (NN) matches among all the features from the image of the initial examination. The nearest neighbors are selected satisfying a minimum Euclidean distance. The main idea is to randomly and repeatedly choose three matches out of the set of potential matches, then compute a homography for each set of three matches and acquire the homography H that has the largest number of inliers in the remaining set. Those keypoints that fit well to the homography transformation are called inliers, while those that are inconsistent to it are called outliers (Fig.6). A keypoint is considered to be an inlier or outlier according to its conformity to an affine geometrical transformation, which takes into account small differences in translation, rotation and scaling between the initial examination image and the respective follow-up image. On our sample images a maximum distance of 5 pixels for the inlier decision is allowed.

IV. EXPERIMENTAL RESULTS

In order to assess the performance of the proposed registration procedure, we tested our method on two image sets. The first set consists of 64 pairs of images only with the nevi under examination, while the second one (33 pairs of images) contains visible artifacts (mostly hairs). The evaluation in two distinctive sets was decided due to the fact that, unfortunately, the proposed registration method is distracted by the presence of dense hair (straight or intersecting to each other). However, the presence of thin and sparse hair segments do not affect the performance of the method, since the ROI-SIFT method and the RANSAC algorithm usually are tolerant to this kind of artifacts.

It has to be noted also that pairs of follow-up images that represent visible aggravation of the nivus (such as in the pair of fig.1, see Fig. 7) present higher global registration errors due to the fact that there is severe difference between the differences in these areas (part of the images change during the two consecutive examinations). However, when local registration metrics are applied (block-based methods), it is shown that local registration errors are negligible.
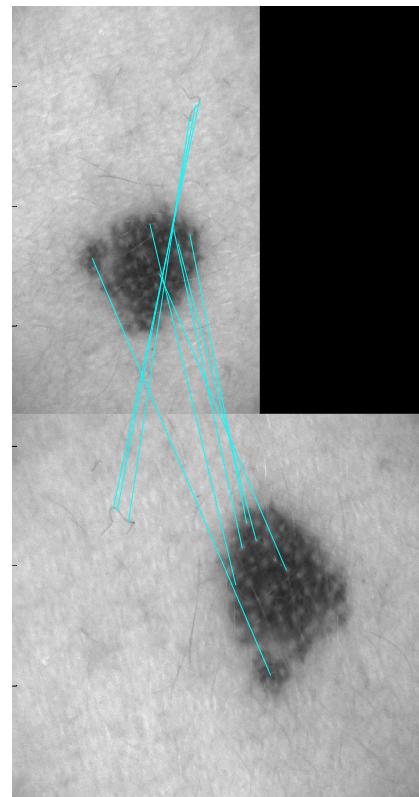

Fig. 6. Final nine keypoints (inliers) after RANSAC in the pair of images of Fig. 1. These keypoints are used for registration.

The registration error is measured by defining manually 10 pairs of matching points for each follow-up pairs of images and measuring the average pixel distance achieved with the estimated homography. The error is computed as: (i) the average Euclidean distance between the matched points by RANSAC (Table 1, column 4), (ii) average

Euclidean distance between the 10 manually selected pair of points (Table 1, column 5) and (iii) global image metrics like Peak signal-to-noise ratio (PNSR) and Mean Squared Error (MSE) of the initial and the registered image (Table 1, column 6).

TABLE I
REGISTRATION PERFORMANCE IN 2 SETS
(ALL VALUES ARE AVERAGE VALUES)

Set	Initial key-points (ROI-SIFT)	Final key-points after RANSAC	Average distance error of final keypoints	Average distance error of 10 manually selected pair of points	PNSR/ MSE
Set 1	67.4	9.8	4.2	6.7	36.6032/ 14.2156
Set 2	97.2	4.3	5.8	14.1	12.2569/ 45.4361
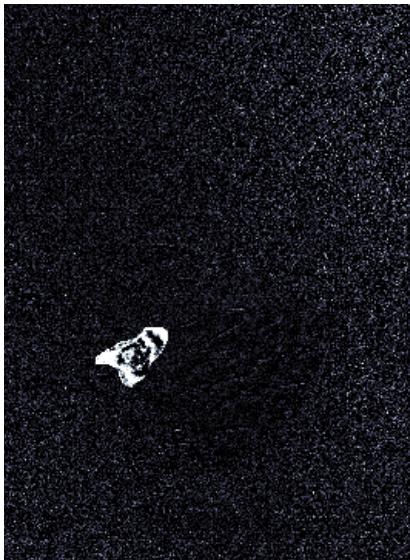
Fig. 7. Difference in area of skin nivus after follow-up examination (the image is accentuated for illustration purposes).

V. CONCLUSIONS

Automated assessment tools for dermoscopy images have become an important research field mainly because of inter- and intra-observer variations in human interpretation. One of the most important difficult issues in dermoscopy image analysis is automated comparison of images taken in follow-up examinations.

At present there is great interest in the prospects of an early screening system for teledermatoscopy, based on the automatic analysis of dermatoscopic images. The benefits of such systems are two-fold: (1) to provide quantitative information about a nivus that can be relevant for the clinician; (2) to be used as a stand-alone early warning tool, with the inherent advantages of time effectiveness and low cost procedures of diagnosis and treatment.

To this end, our future work aims to enhance the proposed system, by: (a) improving the performance of the registration module by adding an efficient hair removal algorithm and (b) assessing the remaining 4 factors of ABCDE rule.

Specifically, for (a) we examine the work reported in [11] since DullRazor [12] has not provided sharp images after execution. For (b) we examine adaptive snakes [13] watershed segmentation [14], fuzzy c-means [15] and phase congruency [16] for quantitative measurements in the affected skin. These methods would probably cover adequately the requirements of A-B-C-D factors and the Total Dermatoscopy Score (TDS) index [17].

REFERENCES

[1] R.P. Braun, H.S. Rabinovitz, M. Oliviero, A.W. Kopf, J.H. Saurat, Dermoscopy of pigmented skin lesions, J. Am. Acad. Dermatol. 52, pp. 109–121, 2005.
[2] S.A. Pavlopoulos. New hybrid stochastic-deterministic technique for fast registration of dermatological images. Medical & Biological Engineering & Computing, 42(6), pp. 777–786, 2004.
[3] G. Maletti, B. K. Ersbøll, and K. Conradsen. A combined alignment and registration scheme of lesions with psoriasis. Information Sciences, 175(3), pp. 141–159, 2005.
[4] D. Delgado, B. K. Ersbøll, and J. M. Carstensen. S.H.A.R.P: A smart hierarchical algorithm to register psoriasis. In Proc. of Int. Workshop on Systems, Signals and Image Processing, 2004, pp. 43–46.
[5] S. Zambanini, G. Langs, R. Sablatnig, H. Maier, Automatic Robust Registration of Cutaneous Hemangiomas for Follow-up Examinations, Proc. of 31st AAPR/OAGM Workshop, 2007.
[6] Optipix Video Dermatoscopy, Optilia, Sweden, www.optilia.eu,
[7] D. Lowe, "Object recognition from local scale-invariant features", International Conference on Computer Vision, Corfu, Greece, 1999 pp. 1150-1157.
[8] T. Lindeberg, "Scale-space theory: A basic tool for analysing structures at different scales", Journal of Applied Statistics, 21(2), pp. 224-270, 1994.
[9] D. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, 60 (2), pp. 91-110, 2004.
[10] Fischler M., Bolles R.,"Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", Comm. of the ACM, vol. (24), pp. 381–395, 1981.
[11] H. Zhou, M. Chen, R. Gass, J.M. Rehg, L. Ferris, J. Ho, L. Drogowski, Feature-preserving Artifact Removal from DermoscopyImages, Proceedings of the SPIE, 6914, 2008, pp. 9 – 14.
[12] T. Lee, V. Ng, R. Gallagher, A. Coldman, D. McLean, DullRazor: A software approach to hair removal from images. Computers in Biology and Medicine, 27, pp. 533-543, 1997.
[13] J. Nascimento, J. S. Marques, Adaptive Snakes Using the EM Algorithm, IEEE Trans. Image Processing, pp. 1678-1686, 2005.
[14] Jean Cousty, Gilles Bertrand, Laurent Najman, and Michel Couprie. Watershed cuts: thinnings, shortest-path forests and topological watersheds. IEEE Transactions on Pattern Analysis and Machine Intelligence. 32 (5), pp. 925–939, 2010.
[15] G.E. Tsekouras, A. Mamalis, C. Anagnostopoulos, D. Gavalas, D. Economou , Improved batch fuzzy learning vector quantization for image compression, Information Sciences, 178 (20), pp. 3895-3907, 2008.
[16] A. Psyllos, C.N. Anagnostopoulos, E. Kayafas, Vehicle model recognition from frontal view image measurements, Computer Standards & Interfaces, 33, pp. 142–151, 2011.
[17] Nachbar F, Stolz W, Merkle T, Cognetta AB, Vogt T, Landthaler M, Bilek P, Braun-Falco O, Plewig G., The ABCD rule of dermatoscopy. High prospective value in the diagnosis of doubtful melanocytic skin lesions, Journal of the American Academy of Dermatology, 30 (4) pp. 551-559, 1994.