Registration of Nevi in Successive Skin Images for Early Detection of Melanoma

Juha Röning, Marcel Riech
Machine Vision and Media Processing Group
Infotech Oulu and Department of Electrical Engineering, University of Oulu, FIN-90570 Oulu, Finland
email: jjr@ee.oulu.fi

Abstract

The only cure for malignant melanoma (skin cancer) is early detection. Surgical removal of a newly developed melanoma will result in complete cure. In this study, the first steps towards the development of a skin cancer detection system are reported.

One possible way to detect early the occurrence of melanoma is to screen the body of a patient at regular intervals for changes or new lesions. To increase the accuracy of this laborious and painstaking task, a computer vision system could be used. One of the most important problems involved in a vision system of this kind is the need to determine which lesions in successive skin images taken over a given period, represent the same lesions. Repeated registrations of skin images also detect new lesions that do not have a counterpart on the previous image. After registration, the lesions in successive images are compared for alterations in size, shape, colour and so on, to detect changes that are suggestive of melanoma.

In this paper, we introduce a new algorithm for the registration of lesions in successive skin images. The baseline algorithm requires two initial matches to register the other lesions in the images. The initial matches are provided by a physician or an algorithm that selects the most likely initial matches. The test suggests that the baseline algorithm determines 99% of the matches correctly, and this performance is largely independent of the number of lesions in the skin images.

Key Words: Image registration, Melanoma screening, Point pattern matching

1. Introduction

The only cure for malignant melanoma is early detection. Surgical removal of a newly developed melanoma will result in complete cure. Much research has been devoted to discriminating between benign and malignant lesions on the basis of certain features [1, 2, 3], such as asymmetry [4], border irregularity [5], colour variation [6] and texture inside the mole [7, 8]. These features can be measured objectively using a vision system and can help to increase the accuracy of the diagnosis done by a physician [9]. Early melanoma is typically discovered when a new skin lesion appears or changes are seen in a preexisting lesion. For this reason, the patients with an increased risk for developing melanoma are screened on a regular basis. The physician performs the screening by comparing the current lesions on the patient's body to pictures taken during a previous visit. This way of screening is a painstaking task that requires remarkable concentration and patience by the physician. The patients with large numbers of lesions are at the highest risk of developing new melanoma, but they are also the most difficult to screen. A few lesions are easily overlooked among many.

A computer vision system could, in principle, screen the patient in an objective and systematic way without omitting any lesions. The vision system compares the images taken at the baseline visit (reference images) to the images taken of the same body parts during the follow-up examination (match images). It compares each lesion to its counterpart in the other image and determines whether or not it has changed over the period between the baseline and the follow-up visits. Since the features mentioned earlier are specially intended to discriminate between benign and malignant lesions, they are very useful in detecting these changes. If a lesion has changed and becomes malignant, one of these features will certainly indicate this change. By comparing the images, new lesions will also be discovered, because a new lesion does not have a counterpart in the other image.

One of the most important steps in comparing lesion images is the need to determine which lesions in the first image represent the same lesions in the follow-up image. This step is called registration or matching of lesion images.The determination of which lesions in the reference and match images represent the same lesion is no trivial task. The registration algorithm must be able to cope with differences between the images. These differences manifest themselves as rotation, translation, scaling, contour distortion and elastic deformation due to differences in the patient-camera positions and the posture of the patient's body between the two photo sessions.

Before an actual comparison of images can take place, the reference and match images have to be segmented to extract lesions and skin areas. After that, the location of the nevi and their features are determined. After segmentation, each lesion in the reference and match images is rep-

resented as a point and a vector of feature values that describe its size, shape, colour, texture, etc.

Some work on the registration of lesion images has been done by Perednia and White. In [10] they describe three algorithms that perform this registration by means of point matching. To reduce the dimensionality of the matching problem, these algorithms require some initial matches, which are provided by the attending physician or an initial match point selecting algorithm [11]. The best registration algorithm in [10] requires three initial matches, and it operates according the following principle. First, the reference and match images are overlaid. Then the three initial matches are used to calculate an affine transformation that transforms the lesion pattern of the reference image in such a way that the initial matches of the reference and match images coincide. The other matches or point pairs are now found by determining which reference and match lesions are each other's mutually closest neighbours. The lesions that do not have a mutually closest neighbour are assumed to be new ones.

One drawback of this algorithm is that is does not take into account the elastic deformation caused by differences in the patient's posture in the reference and match images. Another drawback is that the algorithm becomes unstable when the distribution of the initial matches is almost or fully colinear. The minimum angle of the corners of the triangle formed by the initially matched points plays an important role in the performance of the algorithm. Perednia and White report that the algorithm determines 97% of the matches correctly if only the initially matched point distributions are taken into account that have a minimum angle larger than 10°. When tested on our data, using the same minimum angle criterion, the algorithm detected only around 88% of the matches correctly. The reason for this is that the images of Perednia and White cover a skin area of 150×150 mm^2, whereas our images cover the complete upper trunk. When a small area of skin is registered, the differences can be modelled as a rigid body motion, but when larger areas are involved, elastic deformation is an important consideration that cannot be neglected.

In this paper, a baseline algorithm is introduced. This algorithm does not have the aforesaid drawbacks. Up to a certain extent, it takes into account elastic deformation, and the performance is independent of the distribution of

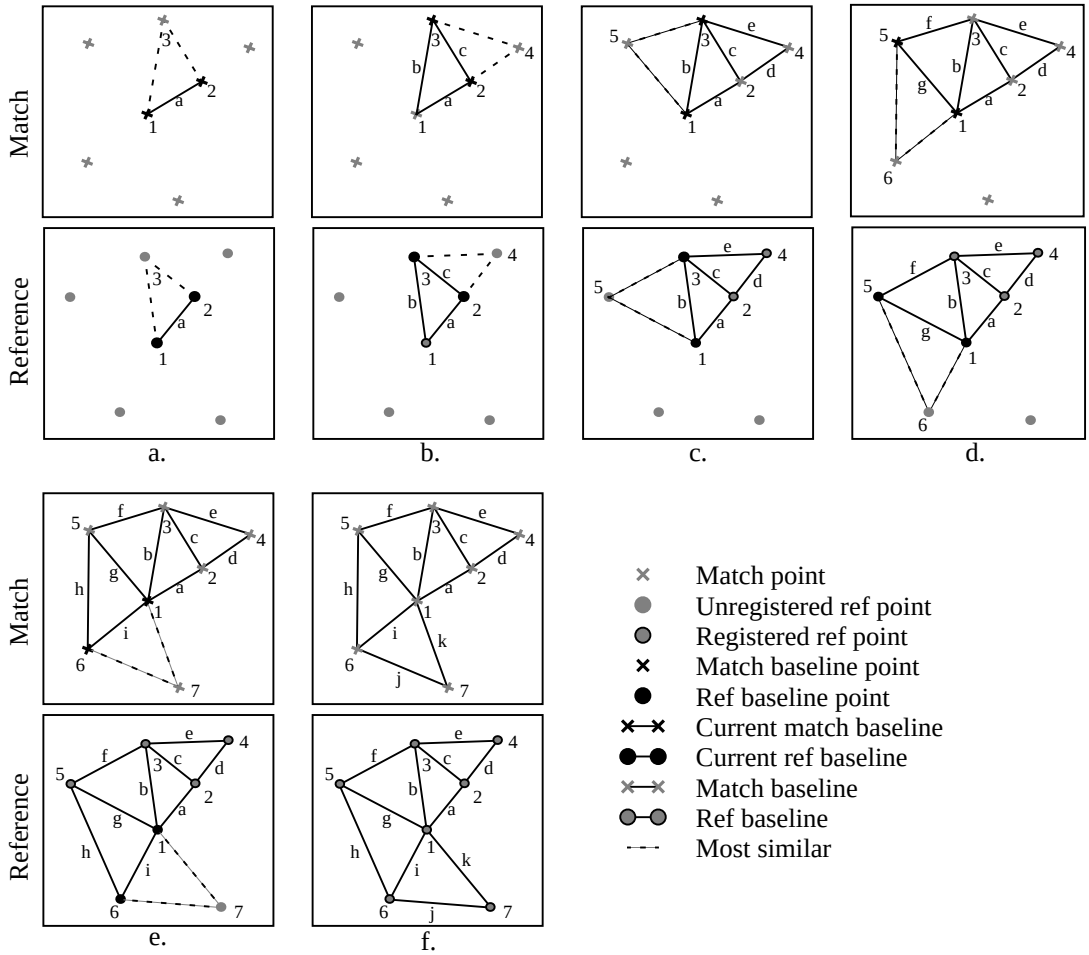

Figure1. The baseline algorithm in action.

the initial matches, of which it only requires two. Tests show that this new algorithm identifies an average 99% of the matches correctly. The baseline algorithm is described in detail below, after which it is tested on our own images and compared to the best performing algorithm of Perednia and White [10]. In the last section, the test results are discussed and conclusions given.

2. Baseline algorithm

The principle of the algorithm is explained by means of the example shown in Figure 1. In this figure, the top images represent the match images and the bottom ones the reference images. The baseline algorithm requires only two initially matched points, which are provided by the attending physician or by a preprocessing stage by selecting the most likely initial matches [11]. The matching starts by defining the line between the two initial point pairs as a baseline. The line between the points in the reference images is called the reference baseline, that in the match image is called the match baseline, and both are indicated by a letter a in Figure 1a. In the reference image the point that is closest to the midpoint of the reference baseline is taken as the third point of the triangle which is formed by the reference point and the two points of the baseline (the corners are labelled 1, 2 and 3 in Figure 1a). Certain characteristic geometrical properties of this triangle are calculated and used in a similarity metric, which will be discussed in the next section. In the match image these geometric properties are calculated for every point, and the points of the match base form the other corners of the triangle. The point in the match image that has the most similar geometric properties compared to those of the point in the reference image, is said to be the matched point, see point 3 in Figure 1a. These reference and match points are defined as a point pair. Together with the two points of the baseline, two new baselines can be defined, namely the baselines b and c in Figure 1b. The reference point is labelled as "registered" and will never again be selected as being closest to a baseline. Next, the point in the reference image that is the closest to the midpoint of one of the available baselines is taken, which yields reference point 4 and baseline c in Figure 1b. For this reference point and the selected baseline, the most similar matching point in the match image is determined on the basis of the similarity metric. In Figure 1b, match point 4 was selected as most similar. This results in a new point pair, with which two new baselines are again formed, namely the baselines d and e in Figure 1c. Again, the reference point is labelled as "registered" and excluded from the process. This process is then repeated over and over again until all reference points have been matched (paired) to points in the match image, Figure 1d, e and f. Finally, all points in the match image that are not paired with a reference point are defined as unmapped and represent new points that did not occur in the reference image.

The algorithm described above performs well as long as every reference point has a corresponding match point, but a problem arises if this is not the case. The algorithm described above would just define the best-fitting point in the match image as a point pair. This problem arises when, for example, a lesion in the reference image was diagnosed as melanoma and excised before the match image was made in the follow-up photo session. To overcome this problem, an extra step takes place after identification of the most similar point in the match image; see Figure 2a, where match point 3 is found to be most similar to reference point 3. The step of finding the most similar match point is now reversed. For all the points in the reference image the geometric properties are calculated with respect to the reference baseline (in this case baseline a in Figure 2b), and by using the similarity metric the reference point is found that is most similar to the point in the match

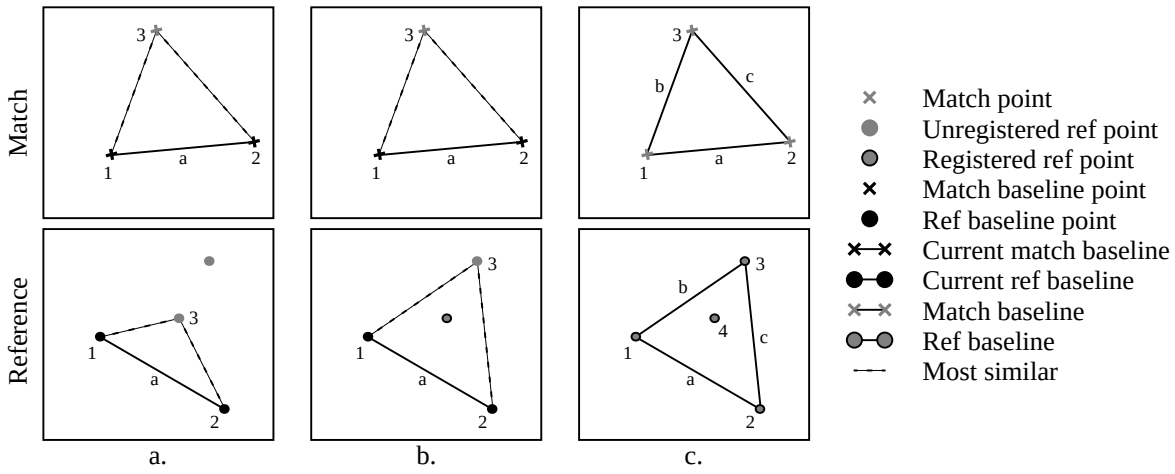

Figure 2. Finding point pairs when a reference point does not have a corresponding point in the match image.

image found in the previous step (in this case reference point 3 in Figure 2b). If this reference point is the same as the one with which the match point was found, these reference and match points are defined as a match, while otherwise they do not form a match (which is the case in Figure 2, since reference point 3 in Figure 2a is different from the reference point 3 in Figure 2b). The reference point from which the process was started (reference point 3 in Figure a) is labelled as "registered" and will never again be selected as being the closest to any baseline. The next reference point selected as closest to the midpoint of a baseline is point number 3 in Figure 2b. For this point, the match point labelled number 3 in Figure 2b is identified as most similar. Now, the most similar reference point for this point is found in the match image, which is point number 3 in Figure 2b. We thus returned to the same reference point that we started from. This means that these reference and match points form a point pair. The point pair results in two new baselines, namely baseline b and c in Figure 2c. And, finally, when all reference points have been labelled as "registered", all points in the reference and match image that are not paired with a point in the other image are defined as unmapped and represent new points that did not occur in the other image.

The computational requirements of the baseline algorithm mainly depend on the part of the algorithm that detects the closest possible combination of reference baseline and reference point. At the beginning of the registration process, there is only one baseline and N points that need to be registered. As the process advances, the number of baselines grows and the number of unregistered reference points decreases. At the last step of the registration process there are more than N baselines and only one unregistered reference point left. Hence the average computational requirements of this search process of comparing the distances between all the possible unregistered reference points and baselines are $O(N^2)$, and since this process is repeated N times to register all the N reference nevi, the whole algorithm is $O(N^3)$.

3. Similarity metric

A similarity metric is needed to determine whether or not the combination of a baseline and a point in the reference image corresponds to a baseline-point combination in the match image. This similarity metric is based on spatial information only. A way of characterizing a point with respect to a base is shown in Figure 3a. The distances l_1, l_2 and the angles α, β are characteristic geometric properties of the point P with respect to the baseline. If all the differences between the reference and match images could be accounted for with a rigid-body motion, then only the angles or only one distance and one angle would be enough to characterize a point with respect to the baseline. In practice, however, it is not a rigid-body motion because of such effects as elastic and contour distortion. If all the four geometric properties are used, the algorithm is more robust. There is one problem, however. To be able to compare angles, the angles must be between 0 and $\frac{\pi}{2}$ (always the smallest angles between two lines), which means that the kind of characterizing presented in Figure 3a is ambiguous. This is shown in Figure 3b. A point that is almost the mirror image, when mirrored in the baseline, of point P will produce approximately identical values for and l_1, l_2, α and β. This problem is solved by adding a third point, which is derived from the baseline, by rotating the baseline pair over an 90° angle around the baseline point 2, see Figure 3c. The extra properties l_3, γ will ensure unique characterization. If the positions of the baseline points 1 and 2 are represented by the vector $\vec{b}_1$ and $\vec{b}_2$, then the third point $\vec{b}_3$ can be derived in the following way:

$$\vec{b}_3 = \vec{b}_2 + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \cdot (\vec{b}_2 - \vec{b}_1) \qquad (1)$$

When the geometric properties are derived in the reference image they are called l_{R1}, l_{R2}, l_{R3}, α_R, β_R and γ_R, while for the match image they are called l_{M1}, l_{M2}, l_{M3}, α_M, β_M and γ_M. The similarity metric is now defined as an error:

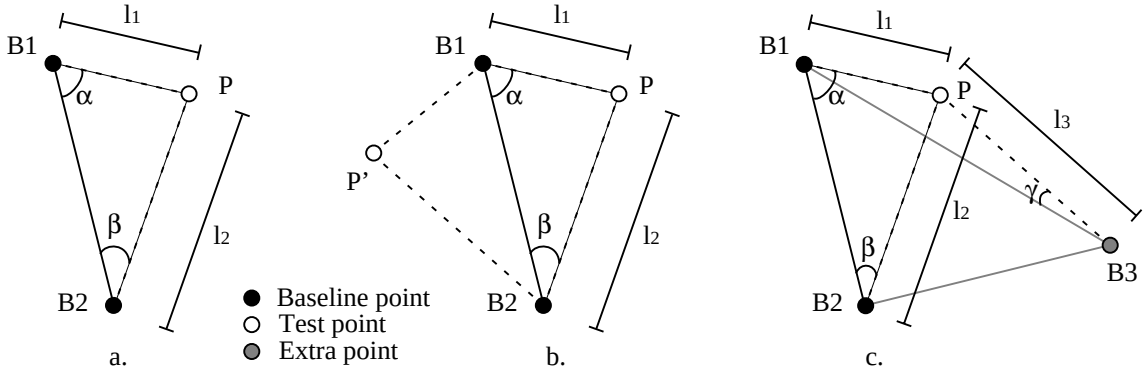

Figure 3. Characterizing point P with respect to baseline B1-B2.

$$Error = \frac{|l_{R1} - l_{M1}|}{l_{R1} + l_{M1}} + \frac{|l_{R2} - l_{M2}|}{l_{R2} + l_{M2}} + \frac{|l_{R3} - l_{M3}|}{l_{R3} + l_{M3}} + \qquad (2)$$

$$\frac{|\alpha_R - \alpha_M|}{\alpha_R + \alpha_M} + \frac{|\beta_R - \beta_M|}{\beta_R + \beta_M} + \frac{|\gamma_R - \gamma_M|}{\gamma_R + \gamma_M}$$

The error value is smaller if the reference and match baseline-point combinations are closer to each other.

4. Results

The algorithm of Perednia and White and the baseline algorithm are tested on random point (lesion) patterns generated by using two original image pairs (an image pair is a combination of a reference and the corresponding follow-up or match image of the same body part of a single patient, e.g. Figure 4 is one of the image pairs used) of the upper trunk of two different patients. Each original image pair contains 60 reference lesions (reference points) and an equal number of corresponding lesions in the match image (match points). Each image pair is used to generate 3000 random image pairs by randomly selecting n reference points and an equal number of corresponding match points from the original image pair. To simulate new lesions in the match image, m reference points are removed from the reference point pattern. In this way the two original image pairs result in 6000 random image pairs containing n-m matches and m new points that must be registered by the registration algorithms.

In the Figures 5 and 6 the results are for the algorithm of Perednia and White and the baseline algorithm tested on random images containing 45 reference point and 50 match points (n = 50, m = 5).

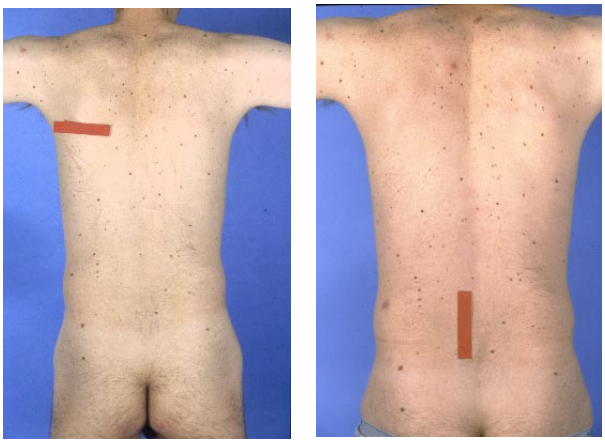

Figure 4. Example of one of the image pairs used to generate random images. The left image is the reference image and the right one is the match image.

Figure 5 shows that the performance of the algorithm of Perednia and White, depends heavily on the distribution of the initial matches. The performance of the baseline algorithm could be influenced by the distance between the two initial matches, but according to Figure 6 this is not the case.

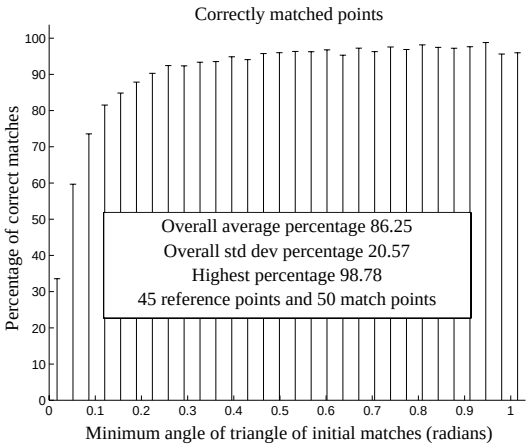

Figure 5. Correctly matched points as a function of the minimum angle of the triangle formed by the initial matches.

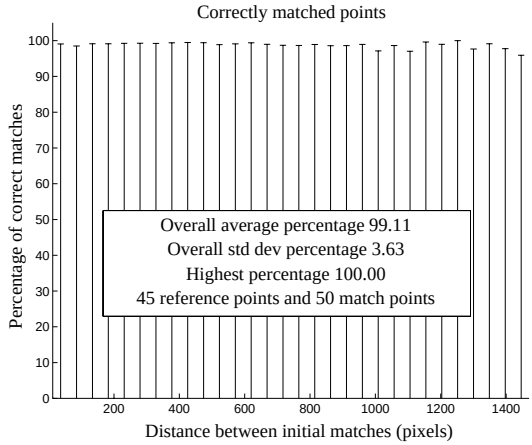

Figure 6. Correctly matched points as a function of the distance between the initial matches.

The Figures 7 and 8 show the performance of both algorithms as a function of the number of points to be matched and the number of removed reference points. When testing the algorithm of Perednia and White in Figure 7, no attention was given to the minimum angle criterion, and all angles were allowed. For this reason, the performance looks worse than in Figure 5, but it can still be seen that the percentage of correctly matched points is strongly dependent on the number of points in the reference and match images. The performance of the baseline algorithm, however, does not depend much on the number of points to be matched. The percentage of correctly matched points remains almost constant at 99%.

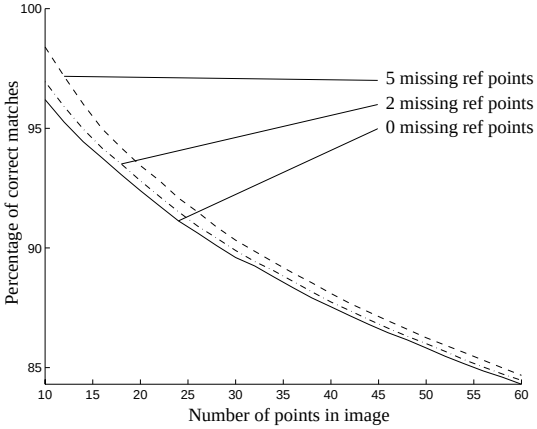

Figure 7. Performance of the algorithm of Perednia and White as a function of the number of points to be matched and the number of removed reference points. No restrictions are posed on the minimum angle of the triangle formed by the three initial matches.

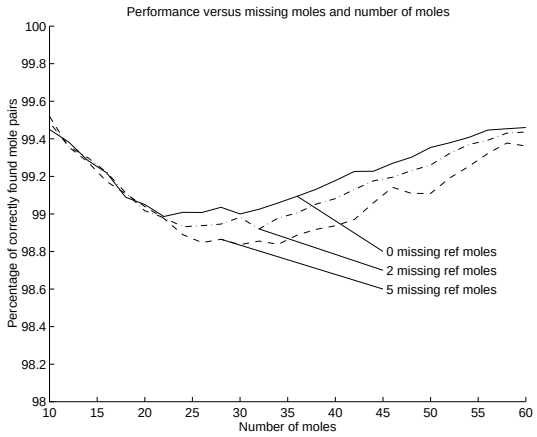

Figure 8. Performance of the baseline algorithm as a function of the number of points to be matched and the number of removed reference points.

5. Discussion and conclusions

We introduced the baseline algorithm for point-matching a large number of lesions in two successive skin images. This registration process is an important step in the automatic screening of patients to detect changes in lesions that are suggestive of melanoma. The tests suggest that the baseline algorithm is capable of determining 99% of the matches correctly when given two correct initial matches. The two initial matches are provided by a physician or an initial match-selecting algorithm. Unlike in the registration by the algorithm of Perednia and White, the distribution of the initial matches does not influence the performance of the baseline algorithm. This improves the overall performance of the registration process, i.e. the baseline algorithm combined with an initial match-selecting algorithm. The reason for this is that the initial match-selecting algorithm can select the most likely initial matches and does not have to select the less likely initial matches that comply with a certain distribution. Furthermore, the ability to find the matches correctly was found to be largely independent of the number of lesions in the images, and new lesions do not pose a problem for the baseline algorithm.

6. References

[1] Ercal, F.; Chawla, A.; Stoecker, W.V.; Hsi-Chieh Lee; Moss, R.H. Neural network diagnosis of malignant melanoma from color images. IEEE Transactions on Biomedical Engineering, Vol. 41 (Sept. 1994), No. 9, p. 837-845.

[2] Stoecker, W.V.; Moss, R.H.; Ercal F.; Umbaugh, S.E.Non-dermatoscopic digital imaging of pigmented lesions. Skin Research and Technology, Vol. 1 (1995), p. 7-16.

[3] Ganster, H.; Gelautz, A.; Pinz, A. Initial results of automated melanoma recognition The 9th Scandinavian Conference on Image Analysis, Uppsala, Sweden, June 1995.

[4] Stoecker, W.V.; Li, W.W.; Moss, R.H. Automatic detection of asymmetry in skin tumors. Computerized Medical Imaging and Graphics, Vol. 16 (1992), No. 3, p. 191-197.

[5] Claridge, E.; Hall, P.N.; Keefe, M.; Allen, J.P. Shape analysis for classification of malignant melanoma Journal of Biomedical Engineering, Vol. 14 (May 1992), p. 229-234.

[6] Schindewolf, T.; Stolz, W.; Albert, R.; Abmayr, W.; Harms, H. Classification of Melanocytic lesions with color and texture analysis using digital image processing. Analytical and Quantitative Cytology and Histology, Vol. 15 (Feb. 1993), No. 1.

[7] Stoecker, W.V.; Chiang, C.; Moss, R.H. Texture in skin images: comparison of three methods to determine smoothness Computerized Medical Imaging and Graphics, Vol. 16 (1992), No. 3, p. 179-190.

[8] Kontinen, J.; Röning, J.; MacKie, R.M. Texture features in the classification of melanocytic lesions Proceedings of International Conference on Image Analysis and Processing (ICIAP'97), (17-19 Sept 1997).

[9] Hall, P.N.; Claridge, E.; Morris Smith, J.D. Computer screening for early detection of melanoma - is there a future? British Journal of Dermatology, Vol. 132 (1995), p. 325-338.

[10] Perednia, D.A.; White R.G. Automatic registration of multiple skin lesions by use of point pattern matching. Computerized Medical Imaging and Graphics, Vol. 16 (1992), No. 3, p. 205-216.

[11] White, R.G.; Perednia, D.A. Automatic derivation of initial match points for paired digital images of skin. Computerized Medical Imaging and Graphics, Vol. 16 (1992), No. 3, p. 217-225.