Distance transform for automatic dermatologic images composition

C. Grana[a], G. Pellacani[b], S. Seidenari[b], R. Cucchiara[a]
[a]Department of Information Engineering and [b]Department of Dermatology
University of Modena and Reggio Emilia

ABSTRACT

In this paper we focus on the problem of automatically registering dermatological images, because even if different products are available, most of them share the problem of a limited field of view on the skin. A possible solution is then the composition of multiple takes of the same lesion with digital software, such as that for panorama images creation.
In this work, to perform an automatic selection of matching points the Harris Corner Detector is used, and to cope with outlier couples we employed the RANSAC method. Projective mapping is then used to match the two images. Given a set of correspondence points, Singular Value Decomposition was used to compute the transform parameters.
At this point the two images need to be blended together. One initial assumption is often implicitly made: the aim is to merge two rectangular images. But when merging occurs between more than two images iteratively, this assumption will fail. To cope with differently shaped images, we employed the Distance Transform and provided a weighted merging of images.
Different tests were conducted with dermatological images, both with standard rectangular frame and with not typical shapes, as for example a ring due to the objective and lens selection. The successive composition of different circular images with other blending functions, such as the Hat function, doesn't correctly get rid of the border and residuals of the circular mask are still visible. By applying Distance Transform blending, the result produced is insensitive of the outer shape of the image.

Keywords: mosaicing, dermatologic images, distance transform

1. INTRODUCTION

Image registration is a process that allows aligning two or more images of the same scene, acquired from different points of view and in different temporal moments. The alignment is done by means of a geometrical transform that allows changing an image (*sensed image*) in the space of another (*reference image*).
Typically image registration is required by *remote sensing* applications (multispectral classification, ambient analysis, mosaicing, super resolution images creation), in *medicine* or in *cartography*. Different application fields are shown in literature [1,2,3,4,5].
In this paper we will focus on the dermatological images area, because in recent years the availability of cheaper video cameras and digital storage allowed for a larger diffusion of digital instruments for the memorization and diagnosis of cutaneous lesion, such for example malignant melanoma. Different solutions have been devised, but most of them share the physical problem of a limited field of view. A possible solution could be the composition of multiple takes of the same lesion with digital software such as that available for panorama images creation. Unfortunately most of these commercial products lack the flexibility of managing rotations, often due to changing patient position, or other differences not typically found in panorama creations.
Most of the techniques proposed in literature share these common points:

- Salient point detection (feature): distinctive objects are detected (e.g. sides, corners, intersections) in manual or automatic ways. In the next steps of computation these features are represented by their centre of mass or representative points, and take the name of *control points*.
- Correspondence detection: at this step, the features are matched between sensed and reference images. Spatial relations and similarity measures are employed to detect these correspondences.
- Transform estimation: by means of the previous parameters a transform is computed to map features from an image to the other. Then the transform is used to align the two images.
- Resampling and blending: the two images are merged together; in case of partial superimposition of the two images, blending techniques have to be used to decide which color has to be assigned to the final pixel.

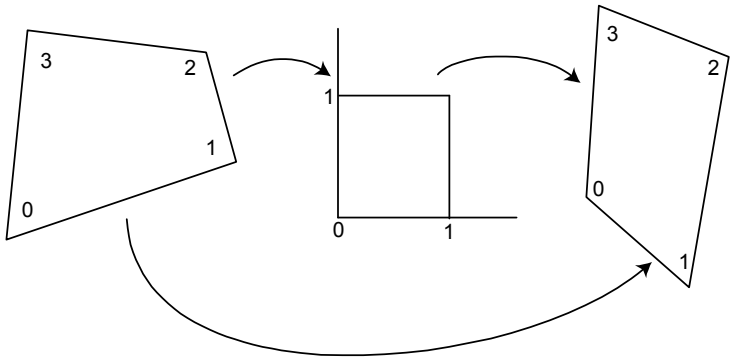

Fig. 1. Projective mapping

In this paper we will present a complete approach to the mosaicing of dermatological images both applicable in case of manual and automatic point detection.

2. PROJECTIVE TRANSFORM

We will work with projective geometry, because one of the main advantages in this case is that the link between points is linear, in particular, any transformation can be expressed by a 3x3 square matrix.
Starting from the nine possible parameters, it is easy to realize that one of these is always redundant because it manages the scale, so a generic transformation is defined by the following notation:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \tag{1}$$

and to better show the advantage of the use of projective coordinates we can rewrite the generic transform in Cartesian coordinates:

$$\begin{aligned} x &= \frac{x_1}{x_3} = \frac{t_{11}X + t_{12}Y + t_{13}}{t_{31}X + t_{32}Y + t_{33}} \\ y &= \frac{x_2}{x_3} = \frac{t_{21}X + t_{22}Y + t_{23}}{t_{31}X + t_{32}Y + t_{33}} \end{aligned}, \tag{2}$$

which shows its non linear nature.

2.1. Projective mapping

In this work we choose to use the projective mapping, which transforms quadrilaterals in other quadrilaterals by keeping straight lines (Fig. 1). Simple affine mapping could have been used while merging images coming from an approximately flat surface, but skin is not rigid, so the different pressure strength and direction produced images that needed shearing. The standard formulas for projective mapping are provided in matrix product form:

$$\begin{pmatrix} x' \\ y' \\ w \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} u' \\ v' \\ q \end{pmatrix} \tag{3}$$

that can be rewritten in compact form:

$$\mathbf{p}_\mathrm{d} = \mathrm{M}\,\mathbf{p}_\mathrm{s} \tag{4}$$

It is possible to write $(x, y)^T = (x'/w, y'/w)$, and $(u, v) = (u'/q, v'/q)$ to go in Cartesian form. The transform is thus defined by the following equations:

$$x = (au + bv + c)/(gu + hv + i)$$
$$y = (du + ev + f)/(gu + hv + i) \tag{5}$$

In this case the equations that describe the transform are not linear and in particular are called linear rational, to indicate the linear nature of numerator and denominator. One of the more interesting features of this type of transform is that also its inverse is a projective mapping [6]. The inverse mapping matrix is given by $\mathbf{M}^{-1} = adj(\mathbf{M})/\det(\mathbf{M})$, and using the equivalence defined in the projective space we can compute the adjoint of the matrix. The inverse transform gets the following form:

$$\mathbf{p}_s = M^{-1}\mathbf{p}_d \tag{6}$$

$$= \begin{pmatrix} u' \\ v' \\ q \end{pmatrix} = \begin{pmatrix} A & B & C \\ D & E & F \\ G & H & I \end{pmatrix} \begin{pmatrix} x' \\ y' \\ w \end{pmatrix} \tag{7}$$

$$= \begin{pmatrix} ei - fh & ch - bi & bf - ce \\ fg - di & ai - cg & cd - af \\ dh - eg & bg - ah & ae - bd \end{pmatrix} \begin{pmatrix} x' \\ y' \\ w \end{pmatrix} \tag{8}$$

that is:

$$u = \frac{Ax + By + C}{Gx + Hy + I}, \qquad v = \frac{Dx + Ey + F}{Gx + Hy + I} \tag{9}$$

2.2. Parameters Estimation

Given four correspondences is possible to recover the mapping parameters by using these equations:

$$x_k = \frac{au_k + bv_k + c}{gu_k + hv_k + 1} \Rightarrow u_k a + v_k b + c - u_k x_k g - v_k x_k h = x_k \tag{10}$$

$$y_k = \frac{du_k + ev_k + f}{gu_k + hv_k + 1} \Rightarrow u_k d + v_k e + f - u_k y_k g - v_k y_k h = y_k \tag{11}$$

where $(u_k, v_k)^T$ and $(x_k, y_k)^T$ with $k = 0,1,2,3$ point to corresponding couples. The system assumes the following form:

$$\begin{pmatrix} u_0 & v_0 & 1 & 0 & 0 & 0 & -u_0x_0 & -v_0x_0 \\ u_1 & v_1 & 1 & 0 & 0 & 0 & -u_1x_1 & -v_1x_1 \\ u_2 & v_2 & 1 & 0 & 0 & 0 & -u_2x_2 & -v_2x_2 \\ u_3 & v_3 & 1 & 0 & 0 & 0 & -u_3x_3 & -v_3x_3 \\ 0 & 0 & 0 & u_0 & v_0 & 1 & -u_0y_0 & -v_0y_0 \\ 0 & 0 & 0 & u_1 & v_1 & 1 & -u_1v_1 & -v_1y_1 \\ 0 & 0 & 0 & u_2 & v_2 & 1 & -u_2v_2 & -v_2y_2 \\ 0 & 0 & 0 & u_3 & v_3 & 1 & -u_3v_3 & -v_3y_3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{pmatrix} = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ y_0 \\ y_1 \\ y_2 \\ y_3 \end{pmatrix} \tag{12}$$

that can be seen in the classic form of a linear system with eight equations and eight unknowns:

$$A\mathbf{x} = \mathbf{b} \tag{13}$$

with

$$\mathbf{x}^T = \begin{pmatrix} a & b & c & d & e & f & g & h \end{pmatrix}$$
$$\mathbf{b} = \begin{pmatrix} x_0 & y_0 & x_1 & y_1 & x_2 & y_2 & x_3 & y_3 \end{pmatrix}. \tag{14}$$

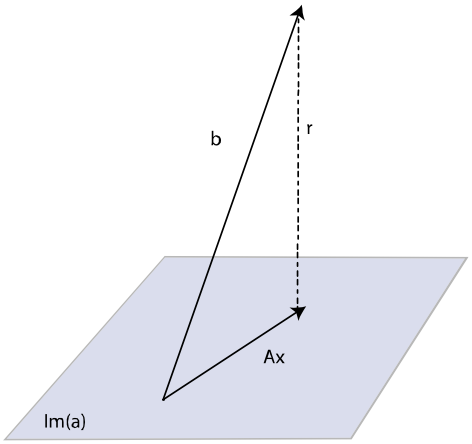

Fig. 2. Geometric interpretation of the normal equation

The system can be solved by the Gauss elimination method, or by LU decomposition, since A it is a standard square matrix. Of all the possible solutions, a possible approach is to search the one that minimizes the error, i.e. the least squares solution of the system. Since to obtain this solution it is necessary to compute $A^T A$, it can be shown that the conditioning number is squared, which means that the solution is very sensitive to slight variations. For this reason a technique called Singular Value Decomposition (SVD) is used.
In case more points have been selected, A and **b** get larger because more rows are available in the system. It is said that the system is over determined [12,7,13]. Of all the possible solutions, a possible approach is to search the one which minimizes the error. In practice, we want to compute $\hat{x}$ such that:

$$\left\|A\hat{\mathbf{x}}-b\right\|_2 = \min_x \left\|A\mathbf{x}-b\right\|_2 = \min_x \left\|r(\mathbf{x})\right\|_2 \tag{15}$$

where $r(x) = b - Ax$ takes the name of *residual*. The condition which allows minimizing the function is that:

$$\frac{d}{dx}\left\|r(\mathbf{x})\right\|_2^2\Big|_{x=\hat{x}} = 0 \tag{16}$$

Using this expression, we can write:

$$\left\|r(\mathbf{x})\right\|_2^2 = r^T r = (b-A\mathbf{x})^T(b-A\mathbf{x}) = b^T b - 2\mathbf{x}^T A^T b + \mathbf{x}^T A^T A\mathbf{x} \tag{17}$$

and computing its first derivative:

$$A^T A\hat{\mathbf{x}} - A^T b = 0 \tag{18}$$

This equation is called normal equation and its solution is the least squares solution of the system.
We can give a graphical interpretation of this equation by rewriting it in this way:

$$A^T(A\hat{\mathbf{x}} - b) = A^T r(\mathbf{x}) = 0 \tag{19}$$

That means that the residual $r(\mathbf{x})$ has to be orthogonal to the columns of A, that is the image space of A (Fig. 2).
To obtain this solution it is necessary to compute $A^T A$, it can be shown that the conditioning number is squared, that means that the solution is very sensitive to slight variations. For this reason a technique called Singular Value Decomposition (SVD) is used. For a detailed analysis see [14, 15].

3. SUCCESSIVE IMAGE BLENDING WITH DISTANCE TRANSFORM

During the transformation we have a number of overlapping points between the different images (Fig. 3). The problem with this area is that we have to decide which RGB color we should use, starting from the color of the two original images. Of course it is likely that the two images will show (hopefully slightly) different values. Our aim is to get uniform and smooth color variations in the output image. The simplest approach to merging, that is using only one of the

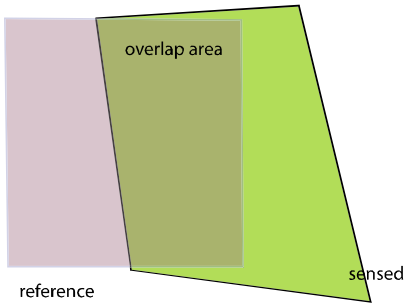

Fig. 3. Overlap area: common area between the two images

two images, produces very sharp edges at the end of the two images contours. The solution is to combine, with different weights, the two values and many proposals are available for this [8,9].
Among the different strategies proposed for image blending, one initial assumption is often made: the aim is to merge two rectangular images. When two images are aligned, the composition of the two will be the superimposition of a rectangle (the reference image) and of a quadrilateral (the sensed image).
When only two images have to be merged, bilinear blending or hat function weighting may be sufficient, because they weight pixels proportionally to how near they are to the image centre. If we want to add a new image to this ensemble, a mask is necessary to denote which points contain *valid* information. We can divide the image in two subsets O and O^t, where the points that belong to the first subset are the valid ones, while the others are to be discarded. At this point, the composed image (the set O) is likely not to be rectangular, so weighting function assuming that the lowest importance is at the rectangle edge, will assign wrong weights to the real edge points. The fact is that what is really needed is to progressively weight more the pixels with respect to their distance to the contour. We decided to compute the distance of each point O to the nearest point of O^t as a proportional weight for the blending. The distance transform is the mathematical tool that allows for this. To implement the Distance Transform, we chose to use the raster scan algorithm proposed by Danielsson [10] in the fast implementation proposed by Leymarie [11]. After computing the distance transform, the reference ensemble and the transformed sensed image are weighted according to the value of the distance transform:

$$I_{blend}(x,y) = \frac{\sum_i w_i(x,y)^2 \cdot I_i(x,y)}{\sum_i w_i(x,y)^2}. \tag{20}$$

In Fig. 4 an example of the weighting of the points in the image is shown.

4. AUTOMATIC REGISTRATION

The approach chosen for the registration phase is divided into four points: feature search (corner detector), potential match list creation, relaxing and ambiguity elimination, RANSAC fitting for the final match.

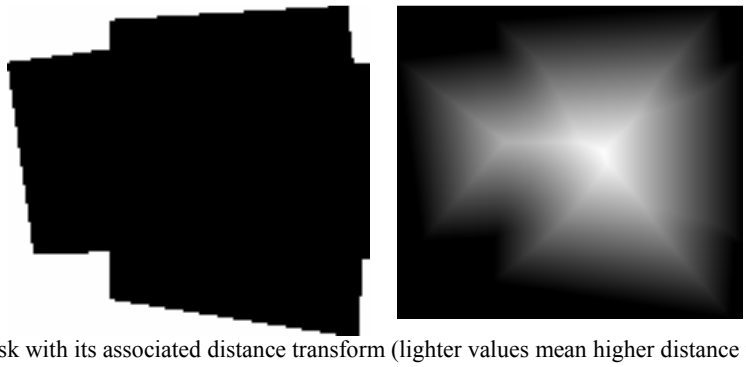
Fig. 4. Example of a mask with its associated distance transform (lighter values mean higher distance from the outside region).

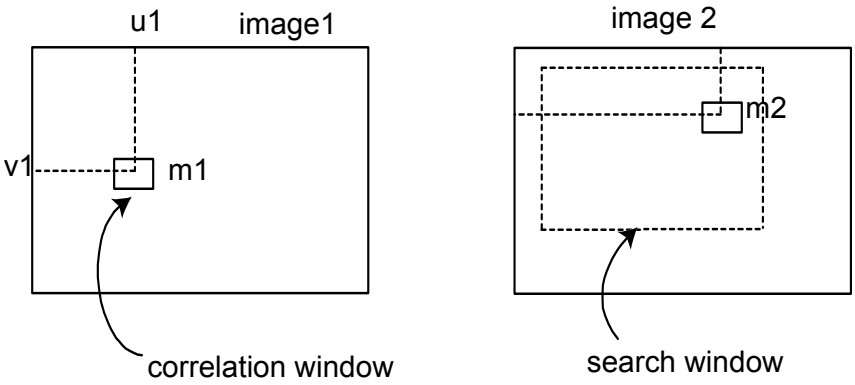

Fig. 5. Correlation and Search Window.

4.1. Corner detector

The feature selection is performed with the *Harris Corner Detector* that defines for each point of the image a *corner strength*:

$$H(x,y) = \det C - \alpha(trace\ C)^2 \tag{21}$$

where:

$$C = \begin{bmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{bmatrix} \tag{22}$$

is the covariance matrix of the gradient vector I_x, I_y in a square neighborhood of the current pixel. α is a parameter between 0 and 0.25, while $H \geq 0$. A corner is found if:

$$H(x,y) > H_{thr} \tag{23}$$

where h_{thr} is a threshold on the corner strength. All the corners in a neighborhood of a stronger corner are deleted. If we computed the eigenvalues of C, and recalling that $\det(A) = \prod_i \lambda_i$ and $trace(A) = \sum_i \lambda_i$, we obtain:

$$H = \lambda_1\lambda_2 - \alpha(\lambda_1 + \lambda_2)^2 \tag{24}$$

4.2. Corner Matching by image correlation

After a corner finding phase, is necessary to compute the matching corners. Given a corner in the first image, m_1 , a correlation window is defined as a window of dimensions $(2n+1)\times(2m+1)$ centered on the same point. The window is a template that is searched in the second image (Fig. 5).

Then a search window of dimensions $(2d_u+1)\times(2d_v+1)$ around the point in the second image is defined and all the correlation values between corner m_1 and all the other corners m_2 that lay in the search window.

The similarity function used was:

$$Score(m_1,m_2) = \frac{\sum_{i=-n}^{n}\sum_{j=-m}^{m}\left[I_1(u_1+i,v_1+j) - \overline{I_1(u_1,v_1)}\right]\left[I_2(u_2+i,v_2+j) - \overline{I_2(u_2,v_2)}\right]}{(2n+1)(2m+1)\sqrt{\sigma^2(I_1)\sigma^2(I_2)}} \tag{25}$$

where:

$$\overline{I_k(u,v)} = \sum_{i=-n}^{n}\sum_{j=-m}^{m}\frac{I_k(u+i,v+j)}{(2n+1)(2m+1)} \tag{26}$$

Fig. 6. Typical example of images for the mosaicing procedure.

$$\sigma(I_k) = \sqrt{\frac{\sum_{i=-n}^{n}\sum_{j=-m}^{m}(I_k(u,v) - \overline{I_k(u,v)})^2}{(2n+1)(2m+1)}} \tag{27}$$

With $\overline{I_k(u,v)}$ we denote the average value of the correlation window and with $\sigma(I_k)$ the standard deviation. The correlation value ranges between 1 and -1 and this value is thresholded to retain only the best matches. This value was set between 0.7 and 0.8.
A winner take all approach was than used to obtain a 1 to 1 matching of the corner detector.

4.3. Random Sample Consensus (RANSAC)

The set of couples obtained cannot be directly used for the transform estimation, because in this set some couples could be *outlier*, i.e. wrong matches. Techniques that can detect and remove outliers are called *robust statistics* and one of the most commonly used is the RANSAC method.
The RANSAC method is an iterative algorithm: a subset of couples is selected and used to compute the transform, then the remaining points are analyzed and it is checked if these are *valid* for the chosen transform. If the number of inliers is big enough the algorithm finishes, otherwise the process is repeated until a good set is found or a number of iteration is reached. The selection of the threshold for validity, of the maximum number of iteration and the number of samples that have to fit the estimated transform is crucial and depends on the problem.
After determining a set of inliers, these are used to compute the transform with the least squares fitting method.

5. RESULTS

To verify the performances of the algorithm different tests have been conducted with dermatological images (Fig. 6 and 7), both with standard rectangular frame and with not typical shapes, as for example a ring due to the objective and lens selection. As shown in Fig. 9 the successive composition of different circular images of Fig. 8 with an Hat function blending doesn't correctly get rid of the border and residuals of the circular mask are still visible (lower right corner for example). By applying Distance Transform blending, the result produced is insensitive of the outer shape of the image and global light differences or local light gradients effects due to the imaging instrument are much reduced (Fig. 10). Intensive testing of the automatically produced transformations were performed, and the overall results are rather good. The main problems encountered are due to the feature search stage so, if the proposed solution is not acceptable, manual correction was performed.

Fig. 7. Final result based on the distance transform

Fig. 8. Original set of images presenting a strong ring effect that needs to be masked away.

6. CONCLUSIONS

In this paper we have shown a complete approach to produce mosaics of generic planar surface images, with particular application to dermatological images, providing solution for the automatic detection of control points, with robust statistics, for the effective blending of successively added images, even if their contours are not of usual shapes.

REFERENCES

1. B. Zitova and J. Flusser, *Image registration methods: a survey*. Image and Vision Computing, 2003. **21**: p. 977-1000.
2. D.Hill, et al., *Medical image registration*. 2000, Radiological Sciences, King's College London: London.
3. M. Audette, F. Ferrie, and T. Peters, *An Algorithmic Overview of Surface Registration Techniques for Medical Imaging*. Medical Image Analysis, 2000. **4**.
4. Fonseca, M.G. and H.M. Costa. Automatic Registration of Satellite Images. in Brazilian Symposium on Graphic Computation and Image Processing. 1997.
5. J.B.A. Maintz, M.A.V., *A survey of Medical Image Registration.* Medical Image Analysis, 1997. **2**(1).
6. Z. Zhang, et al., A robust technique for matching two uncalibrated image through the recovery of the unknow epipolar geometry. 1994.
7. W. Moss, *Linear Algebra*, Department of Mathematical Sciences, Clemson University.
8. A.Rocha, R.Ferreira, and A.Campilho. Image Mosaicing using Corner Detection. in SIARP2000 - V Ibero-American Symposium on Pattern Recognition. 2000.
9. P.Burt and E.Adelson. A multiresolution Spline with application to Image Mosaics. in Proceedings of SPIE. 1983.
10. P.E. Danielsson, "Euclidean distance mapping", Computer Graphics and Image Processing, 14:227-248, 1980.
11. F.Leymarie and M.D.Levine, *A Note on Fast Raster Scan Distance Propagation on the Discrete Rectangular Lattice.* CVGIP: Image Understanding, 55, 1992. **55**.
12. D.Kammer, Linear Algebra for Test and Analysis, University of Winsconsin, Engineering Physic Course.
13. C.Fuhren, Numerical Analysis Advanced Course, in Course Material, Lund Uni-versity. 2000.
14. W.H.Press, S.A.Teukolsky, and W.Vetterling, Numerical Recipes in C, ed. C.U. Press. 1997.

a) b)

Fig. 9. Composition of three images with a) Hat and b) Distance Transform blending

Fig. 10. Final composition of all the image set with Distance Transform blending. Non uniform lighting effects are partially solved by the blending strategy, but still perceivable.

15. D.Whittum and F.Zimmermann, Use Singular Value Decomposition in Ten Min-utes, or Your Money Back. ARDB Technical Note 163, 1997.