

# A Graph-based Approach to Skin Mole Matching Incorporating Template-Normalized Coordinates

Hengameh Mirzaalian<sup>1</sup>, Ghassan Hamarneh<sup>1</sup>, Tim K. Lee<sup>1,2,3</sup>

<sup>1</sup>Medical Image Analysis Lab, Simon Fraser University, BC Canada

<sup>2</sup>Cancer Control Research, BC Cancer Agency, BC Canada

<sup>3</sup>Department of Dermatology and Skin Science, University of British Columbia, BC Canada

{hma36, hamarneh}@cs.sfu.ca, tlee@bccrc.ca

## Abstract

*Density of moles is a strong predictor of malignant melanoma. Some dermatologists advocate periodic full-body scan for high-risk patients. In current practice, physicians compare images taken at different time instances to recognize changes. There is an important clinical need to follow changes in the number of moles and their appearance (size, color, texture, shape) in images from two different times. In this paper, we propose a method for finding corresponding moles in patient's skin back images at different scanning times. At first, a template is defined for the human back to calculate the moles' normalized spatial coordinates. Next, matching moles across images is modeled as a graph matching problem and algebraic relations between nodes and edges in the graphs are induced in the matching cost function, which contains terms reflecting proximity regularization, angular agreement between mole pairs, and agreement between the moles' normalized coordinates calculated in the unwarped back template. We propose and discuss alternative approaches for evaluating the goodness of matching. We evaluate our method on a large set of synthetic data (hundreds of pairs) as well as 56 pairs of real dermatological images. Our proposed method compares favorably with the state-of-the-art.*

## 1. Introduction

Melanoma is one of the fastest growing cancers among the white population in the world with an average 3% increase in incidence for the last four decades. In the USA and Canada alone, it is estimated that there will be 67,080 new cases of melanoma in 2008 [17, 1].

The mechanism of melanoma development is not fully understood. Nevertheless, mole density (number of moles per unit area of skin) has been reported as the strongest risk factor, and about 50% of melanoma originates from pre-

existing moles. Early diagnosis of melanoma may lead to potentially life-saving therapy. To allow for early diagnosis, patients are full-body scanned periodically and digital color images of the skin are collected during the process. One example of the back images at two different times is shown in Figure 1(a). During a dermatological examination, a physician compares the skin images at different time instances to observe changes in the number of moles and their appearance, such as size, color, shape, and texture. However, relying on visual inspection of moles is costly, time consuming, and may be error prone due to fatigue [8]. Therefore, an automatic method for matching corresponding moles would have significant health benefits, especially for tracking moles in patients who are at a high risk of developing melanoma and, hence, require regular mole examinations.

To the best of our knowledge, there exists limited previous work on skin mole or lesion matching. In [9], Huang and Bergstresser proposed to utilize the area of the voronoi cells surrounding moles in the similarity term for mole matching. Then, a dynamic programming approach was used to find corresponding moles. To favor preserving the topology of the constellation of moles, moles were sorted based on their distance to a reference mole. In [15], Perednia and White proposed a two-step process to match moles. First, they assumed some known mole matches and solved for an affine transformation that brings the known matches into alignment. Then, they corresponded each unmatched mole to its closest neighbor after alignment. Similarly, the approach in [18] also relies on an initial matching of moles to establish a course alignment. The lines between the initial matches are then used to construct baselines. Certain geometrical properties for each mole with respect to these baselines are then used in a similarity metric to find the corresponding moles. Compared with [15], [18] is more robust to non-rigid transformation caused by changes in the patients posture.

The mole matching problem can be seen as a general point cloud or graph matching problem. We, therefore, highlight some key related works in this latter area. In [11], a spectral method for finding consistent correspondences between two sets of features is proposed by Leordeanu and Hebert. After constructing an adjacency matrix representing the pairwise agreements between edges, the correct assignments are recovered based on how strongly they belong to the main cluster of the matrix, using eigenvector decomposition. In [14], Coherent Point Drift (CPD) algorithm was proposed by Myronenko et al. CPD establishes correspondence between two point sets based on a maximum likelihood estimate that enforces motion (or velocity field) coherence. CPD simultaneously finds both the non-rigid transformation and the correspondence between two point sets without making any prior assumption about the transformation model except that of motion coherence. In [2], shape context features are used in the similarity measurement to find corresponding points and the best non-rigid alignment of two shapes is provided by thin-plate splines. Zheng and Doermann [23] formulate point matching as an optimization problem to preserve local neighborhood structures during matching and use shape context distance to initialize the graph matching, followed by relaxation labeling updates for refinement. Maciel and Costeira [12] formulate the matching problem as an integer constrained minimization and find global optimal solution via a relaxed concave objective function. Gold et al. [7] define the pose term (the affine transformation between points) into the point matching objective and use *soft assign* to find good suboptimal solutions. Zass and Shashua [22] derive the hyper-graph matching problem in a probabilistic setting, which is solved via convex optimization and based on an algebraic relation between the hyperedges; the global optimum of the matching is found via an iterative successive projection algorithm.

We note that previous mole matching approaches (e.g. [18, 9, 15]) do not benefit from state of the art point and graph matching techniques. Further, for improved mole matching, we argue that matching algorithms (e.g. [22]) must be extended to incorporate high-level, domain knowledge about the anatomy, e.g. the human back. In this paper, we propose to augment anatomical knowledge into the state-of-the-art matching algorithms and develop an anatomy-aware mole matching technique. We are inspired by works on human brain atlases, which have benefited the medical image analysis research in many ways; mainly by providing a common frame-of-reference that facilitates cross-sectional and longitudinal comparisons of brain anatomy [19, 20, 5, 13]. We propose the first human back template (or atlas) to encode spatial mole coordinates in a common normalized coordinate system (sections 3 and 4). After a discussion on graph representation of the constellation of moles in section 5, we extend the hy-

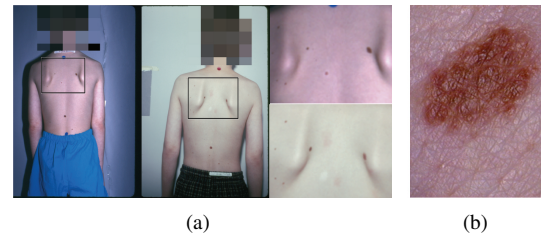


Figure 1. (a) Constellation of moles on a back image [6], (b) A close up on a big mole using a dermoscope with non-polarized light.

per graph matching algorithm [22] to incorporate proximity regularization, angular agreement between mole pairs, and template-normalized coordinates (section 6). We evaluate encoding these criteria into two alternative objective function formulations. Furthermore, we propose and compare different approaches for evaluating the goodness of matching relative to ground truth correspondences and for learning the optimal weighting between the terms of the objective function (section 6). We apply our algorithm on a large synthetic data set (hundreds of pairs) as well as 56 pairs of real dermatologic images and demonstrate improved mole matching results compared to the other methods (section 7). We conclude the paper with a discussion and closing remarks (section 8).

## 2. Back Skin Segmentation and Mole Detection

In order to have a fully automatic mole-tracking system, we first need to extract the contours of the human back skin from the dermatological color images. Skin segmentation can be performed using a variety of methods, e.g. using color [16] and texture similarity [21]. The extracted human back boundaries and several *Type I* [3] anatomical landmarks (e.g. left armpit, right hip) are used to map (or warp) the subject-specific back coordinates into the normalized back template (as described in section 4). The output of the back segmentation step is also used as a region of interest (or mask) for the mole detection stage. Mole detection and segmentation methods are divided into two main classes. The first class of methods segments a single mole, occupying hundreds of pixels and centered in the image with a skin background [4] (Figure 1(b)). The second class detects one or typically more moles in an image capturing a larger field of view, in which each mole occupies only a few pixels, such as those appearing in the full back images of interest in this current work (Figure 1(a)). In [10], for example, back image moles are extracted using a variant of the mean shift algorithm, whereas in [16], facial skin irregularities (nevi) are localized using a multi-scale template matching and texture analysis procedure. The focus of our current work is on mole matching and not on the skin or mole segmentation steps. The latter are currently done using a semi-automatic,

user-assisted algorithm to rule out errors in matching due to incorrect segmentation (see future work in section 8).

### 3. Human Back Template

In order to have an anatomically meaningful comparison between moles in the human back images, we need to perform spatial normalization (similar to works on brain atlases [19, 20, 5, 13]). In this work, we propose the first human back template (atlas) to obtain this normalization. The template is constructed as follows. Four pairs of points corresponding to anatomically meaningful (i.e. Type I) landmarks are used as reference points in the template (Figure 2(a)): (i) the point where the left side of the neck meets the left shoulder, or neck-left ( $n_l$  for short) and similarly neck-right ( $n_r$ ); (ii) the point where the left shoulder meets the left arm or shoulder-left ( $s_l$ ) and similarly shoulder-right ( $s_r$ ); (iii) the left ( $a_l$ ) and right ( $a_r$ ) arm pit; and (iv) the points corresponding to the left ( $h_l$ ) and right ( $h_r$ ) hip. The template is then constructed as a square patch with a pre-defined number of longitudes and latitudes (Figure 2(b)). The central latitude connects  $a_l$  and  $a_r$ , whereas the central longitude connects the medial points  $n_m$ ,  $s_m$ ,  $a_m$ , and  $h_m$ , where the subscript  $m$  indicates the midpoint between the corresponding left and right reference points, e.g.  $n_m = (n_l + n_r)/2$ . Note that in the template space these medial longitudes and latitudes are two straight line segments perpendicular to each other. Finally, based on the anatomical landmarks and the reference central latitude and longitude, a template-based rectilinear coordinate system is defined and a complete Cartesian grid is overlaid on the template.

### 4. Spatial Normalization of Back Coordinates

To perform spatial normalization, each subject-specific back image is registered into the template coordinate system as follows. The anatomical landmarks in the dermatologic image are identified and mapped to their respective reference points in the template, e.g. the left shoulder point in the image space,  $s_l$ , is mapped to  $S_l$  in the template space (Figure 2(b)). The neck, shoulder, and armpit points in the image are identified as inflection points, i.e. based on the curvatures of the back boundary obtained in section 2 (Figure 2(a)). However, it is more challenging to define the hip points segmentation (see future work in section 8).

As the focus of this work is on mole matching, we rely on the user to confirm or edit the extracted position of these anatomical landmark points in the image, and therefore do not provide any evaluation of the accuracy or repeatability of this anatomical landmark detection step (see section 8). Note that, in the image space, all latitudes and longitudes are no longer straight lines but rather smooth curves fit to the locations of the anatomical (and the medial) land-

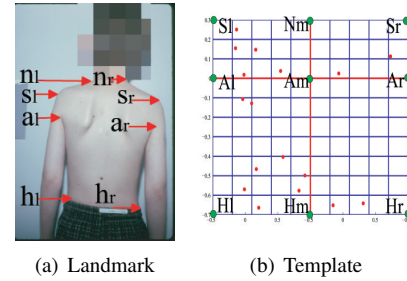


Figure 2. Human back template. (a) Landmarks are shown in the image. (b) Template of the back image. The red and green points correspond to the normalized coordinates of the moles and the reference anatomical landmarks, respectively. The vertical and horizontal red lines correspond to the central longitude and latitude.

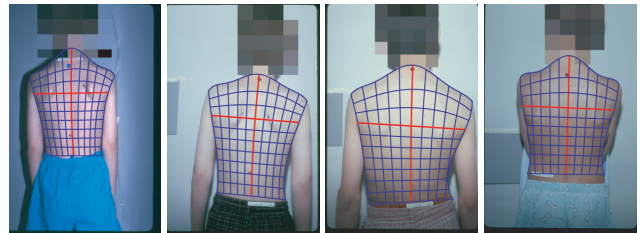


Figure 3. Examples of back images with overlaid grid.

marks in physical coordinates. Consequently, the grid in the image space is no longer rectilinear but rather curved. Equal arc length sampling is used to correspond the grid lines in the image space with those in the template. Table 1 contains the details of how the latitudes and longitudes are defined in the image domain. Finally, with the latitudes and longitudes defined in both the image and the template domain, the intersection points of the latitudes and longitudes now correspond across both domains. In order to establish correspondence between any point (not only latitude and longitude intersections points) in the two continuous domains, we evaluated two interpolation methods: Barycentric coordinates (BC) and Thin plate splines (TPS). TPS is more accurate than BC but it is more computationally expensive. The result is a function  $f : \Omega \rightarrow \hat{\Omega}$  that establishes a bijective mapping between points in any location within the two continuous domains: the image space  $\Omega \subset [0, x_{max}] \times [0, y_{max}]$  and the template space  $\hat{\Omega} \subset [-0.7, 0.3] \times [-0.5, 0.5]$ , where  $x_{max} \times y_{max}$  is the size of the image and the limits of  $\hat{\Omega}$  cover a unit-square patch. Given this mapping  $f$ , any coordinate in the image, mainly mole locations, are mapped to spatially-normalized coordinates in the template. In Figure 3, some examples of a grid overlaid on the back images are shown.

### 5. Graph Representation

After calculating the normalized coordinates of two sets of moles from two different images captured at different

Left/Right longitudes:	Degree 3 polynomials least-squares fitted to the lateral edge points of the back's left silhouette and constrained to pass through $a_l$ , similarly for the right side.
Central longitude:	Medial curve between the left and right longitudes.
Central latitude:	Line connecting $a_l$ to $a_r$ .
Superior latitude:	Degree 5 polynomial least-squares fitted to the superior edge points of both shoulders' silhouettes.
Inferior latitude:	Line connecting $h_l$ and $h_r$ .

Table 1. Definitions of latitudes and longitudes for the back image grid

times, we turn our attention to matching the moles across time. Finding correspondences between moles is modeled as a graph matching problem. Let  $G = (v, e)$  and  $G' = (v', e')$  be two graphs with vertices  $v$  and  $v'$  and edges  $e$  and  $e'$ , respectively. Vertices in  $G$  and  $G'$  are related to the moles from the two images and the edge weights represents some features related to the vertices connected by that edge. Based on these definitions, a matching between  $G$  and  $G'$  is a vertex to vertex mapping  $\Pi_v : v \rightarrow v'$ . This may be viewed as a bipartite graph matching, where edges from one set to other set represent the likelihood of matching nodes from  $G$  to  $G'$ . Note that for dermatological studies, it is important to not only match moles but also to detect appearing and disappearing moles.

## 6. Mole Matching Objective function

Assuming moles do not change their anatomical position with time and a perfect construction of the back anatomy based coordinate system, corresponding moles from two different times should be mapped into the exact same coordinate in the template. In that ideal scenario, perfect matching can be obtained by simply using the Euclidean distance between vertices of  $G$  and  $G'$ , i.e. corresponding moles will have a nil Euclidean distance in the template space. However, due to different sources of error, including limitations of the imaging system, back skin segmentation errors, inaccuracies in mole detection, imperfect localization of the anatomical landmark positions, etc, it is unavoidable that this ideal situation will not take place. Instead, corresponding vertices in one set can be obtained through a nonlinear smooth warp (or nonlinear spatial transformation) that preserves local neighborhoods. This desired regularization motivates edge-to-edge rather than vertex-to-vertex matching, i.e.

$$\Pi_e(e_i) = \Pi_e(v_{i1}, v_{i2}) = (v'_{j1}, v'_{j2}) = e'_j \in e' \quad (1)$$

$$v_{i1}, v_{i2} \in v, v'_{j1}, v'_{j2} \in v'$$

where  $\Pi_e$  maps an edge in  $e$  to an edge in  $e'$ . A favorable mapping is one which maps edges with the same length and the same anatomical direction (e.g. inferior-superior, medial-lateral, or other oblique direction), and maps the vertices of those edges to the proximate normalized spatial coordinates. These criteria describe how well the rela-

tive pairwise geometry is preserved after mapping. Consequently, we combine three terms into the objective function, which the sought optimal mapping should minimize: Euclidian distance between individual vertices of edges measured in the anatomy-based template space; difference between the lengths and between the directions of pairs of edges.

There are two main approaches to combine these terms into an objective function:

$$E_{\Pi_e} = (P_\alpha)^{\omega_\alpha} (P_\beta)^{\omega_\beta} (P_\theta)^{\omega_\theta} \quad (2)$$

$$E_{\Pi_e} = \omega_\alpha P_\alpha + \omega_\beta P_\beta + \omega_\theta P_\theta \quad (3)$$

where the  $P_\alpha$ ,  $P_\beta$ , and  $P_\theta$  measure the reward of matching two edges based on Euclidean distance, difference in edge length, and difference in edge direction, respectively, and are given by:

$$P_\alpha(e_i, e'_j) = e^{-(dist(v_{i1}, v'_{j1}) + dist(v_{i2}, v'_{j2}))}$$

$$P_\beta(e_i, e'_j) = e^{-(abs(dist(v_{i1}, v_{i2}) - dist(v'_{j1}, v'_{j2})))}$$

$$P_\theta(e_i, e'_j) = e^{-(arccos((\vec{v}_i \cdot \vec{v}'_j) / (|\vec{v}_i| |\vec{v}'_j|)))}$$

$$\vec{v}_i = (v_{i1}, v_{i2}); \vec{v}'_j = (v'_{j1}, v'_{j2}) \quad (4)$$

where  $dist(p, q)$  calculates the Euclidian distance between points  $p$  and  $q$ . Although it is more common in the literature to use an objective function of the form (3), we argue that (2) is more suitable in our current situation since the optimal solution should ideally maximize *all* three terms (i.e. an *AND* operation), whereas in (3) maximizing *any* of the terms will produce a larger value for the objective function (i.e. an *OR* operation). In section 7, we demonstrate quantitatively that this is indeed the case. In section 6.2, we describe how we optimize the weights  $(\omega_\alpha, \omega_\beta, \omega_\theta)$ . We now seek the matching  $\hat{\Pi}_e = argmax_{\Pi_e} (E_{\Pi_e})$ .

In order to extract a vertex to vertex mapping from  $E_{\Pi_e}$ , we make use of hypergraph matching method proposed in [22], which we summarize in the next section for notational completeness.

### 6.1. Hypergraph Matching

Hyperedges connect  $d$  nodes, where  $d \geq 2$  is the degree of the hyperedge. In order to find vertex to vertex matching,

matching hyperedges is considered in [22]. For instance, for  $d = 4$ , the ratio between the area of the triangles described by  $v_1, v_2, v_3$  and  $v_2, v_3, v_4$  ( $v_i$  is a vertex of  $G$ ) is used as a similarity measure for matching hyperedges. A probability matrix reflecting the likelihood of an edge-to-edge matching is constructed as  $E_{\Pi_e} = E(e, e') = P(\Pi_e(e) = e' \mid G, G')$ .

According to [22], under the assumption of  $\Pi_v(v_i) \perp \Pi_v(v_j)$ ,  $E_{\Pi_e}$  is written as  $d$  times Kronecker product of  $X(v, v')$ , i.e.  $E_{\Pi_e} = \otimes^d X$ , where  $X$  is a doubly semi-stochastic ( $X \geq 0, X\mathbf{1} \leq \mathbf{1}, X^T\mathbf{1} \leq \mathbf{1}$ ) whose entries store the probability of matching  $v_i$  and  $v'_j$ . Therefore, we need to find  $\hat{X} = \operatorname{argmin}_X \operatorname{dist}(E_{\Pi_e}, \otimes^d X)$ .

This function can be reduced to a simpler form by defining a marginalization matrix  $Y$ :  $Y_{v,v'} = \sum_{i=1}^d \sum_{e|v_i=v, e'|v'_i=v'} E(e, e')$  where  $e = (v_1, v_2, \dots, v_d)$  and  $e' = (v'_1, v'_2, \dots, v'_d)$ . Therefore, the minimization can be rewritten as  $\hat{X} = \operatorname{argmin}_X d(Y, X)$ . A successive projections algorithm is used to find the globally optimal solution of the matching probability matrix  $X$  and a linear assignment problem is solved to obtain a hard matching, if desired. For  $d = 2$ , the matching cost function in [22] simplifies to the difference between the lengths of edges. However, in the presence of noise, relying on edge-lengths is not sufficient to find correct corresponding points and, according to [22], hyperedge matching is recommended. In this case, the time complexity of the algorithm would be:  $O(|v|^d \cdot |v'|^d)$ , which is impractical when the number of nodes is high. Zass and Shashua reduced the algorithm complexity to  $O(|v| \cdot |v'| \cdot z^{2(d-1)})$  by sampling the  $z$  closest hyperedges per vertex and used only the correlation between the sampled Hyperedges.

However, in our mole-matching problem, with the presence of appearing and disappearing moles,  $z$  could not be far from  $\min(|v|, |v'|)$ , which means that high running time problem still exists. In this work, we use  $d = 2$  and adopt Zass and Shashuas approach to solve for  $X$  from  $E_{\Pi_e}$ .

## 6.2. Learning the optimal weights

Given a specific choice of weights of  $\omega_\alpha, \omega_\beta, \omega_\theta$ , the objective function  $E_{\Pi_e}$  is evaluated according to (2) or (3) and the matching probabilities  $X$  are computed as in [22]. Clearly, different choices for the weights will result in a different objective function and, hence, a different optimal solution. It is, therefore, imperative to use appropriate weights.

In order to find the best choice of weights, we rely on generating comprehensive training data with ground truth correspondence and evaluating the matching performance given different choices of weights. In particular, we generate multiple reference training data sets, where each training set consists of two groups of points: a reference set  $V$  and

a target set  $V'$ , mimicking the moles to be matched across two images. The total number of points in  $V$  and  $V'$  are  $n_V = n_V^i + n_V^o$  and  $n_{V'} = n_{V'}^i + n_{V'}^o$ , where the  $n_V^i$  points in  $V$  correspond to the  $n_{V'}^i$  points in  $V'$  (i.e.  $n_V^i = n_{V'}^i$ ) and the positions of the latter set of points are generated by displacing the former by random distances drawn from uniform distributions with different support intervals  $\sigma_n$ . These perturbations in positions mimic the aforementioned errors (e.g. mole detection, anatomical landmark localization, etc.).  $n_V^o$  is the number of moles that exist only in  $V$  but not in  $V'$ , i.e. mimicking disappearing moles, whereas  $n_{V'}^o$  exist in  $V'$  but not in  $V$ , i.e. new appearing moles.

As we control the generation of the two sets of points, the ground truth matrix of matching between points,  $X_{Real}$ , is known. Next, the *dissimilarity* between  $X$ , resulting from the matching algorithm, and  $X_{Real}$  is calculated for different weights. Then, the best weights are chosen as the set giving the minimum dissimilarity between  $X$  and  $X_{Real}$ .

$$(\hat{\omega}_\alpha, \hat{\omega}_\beta, \hat{\omega}_\theta) = \operatorname{argmin}_{(\omega_\alpha, \omega_\beta, \omega_\theta)} d(X, X_{Real}) \quad (5)$$

There are different ways to measure the dissimilarity between  $X$  and  $X_{Real}$ . We evaluate the following measures for the goodness of matching: Hamming distance error, *HDE*, between defuzzified (crisp mapping)  $X$  and  $X_{Real}$ ; the number of incorrect matching, *NIM*; and the spatial Euclidian distance, *SED*, between the mole position mapped to by  $X$  and by  $X_{Real}$ . The distance between  $X$  and  $X_{Real}$ ,  $d(X, X_{Real})$  for each of these measurements is given by:

$$\begin{aligned} d_{HDE} &: \sum_{i,j} X(i, j) \oplus X_{Real}(i, j) \\ d_{NIM} &: \sum_{i,j} X_{Real}(i, j) - X(i, j) \wedge X_{Real}(i, j) \\ d_{SED} &: \sum_{\{(i,j) | X(i,j) - X(i,j) \wedge X_{Real}(i,j) = 1\}} SED(i, j) \end{aligned} \quad (6)$$

Figure 4 visualizes  $SED(i, j)$  for different cases. Note that *HDE* and *NIM* are only defined for hard (crisp) matching whereas *SED* is defined for both hard and soft matchings. As it can be seen in Figure 5, there is a high correlation between *NIM* and *HDE*, but they are not perfectly correlated,  $\rho_{NIM, HDE} = 0.89$ , whereas  $\rho_{SED, NIM} = 0.46$  and  $\rho_{SED, HDE} = 0.61$ .

## 7. Results

The real data used in this study were obtained from color slides taken from an epidemiologic study concerning the use of broad-spectrum sunscreen and mole development. The images were digitized with 24-bit color at 2000 dpi, with a final resolution of about 0.25 mm/pixel. A set of 56 pairs of digitized images containing moles was chosen to



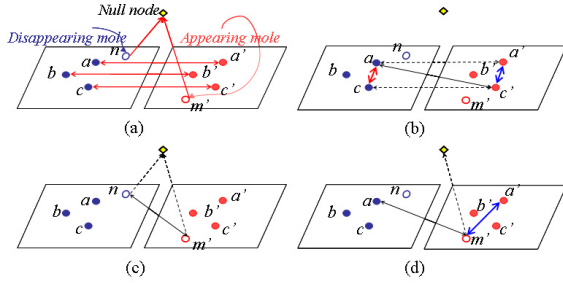


Figure 4. Visualization of *SED* for different cases. (a) Two sets of moles  $\{a, b, c\}$  and their ground truth corresponding moles  $\{a', b', c'\}$  are connected by red lines.  $n$  and  $m'$  are disappearing and appearing moles, respectively and are connected to the null node. Let  $A$ ,  $D$  and  $C$  be the subsets of Appearing, Disappearing and Common moles in  $G$  and  $G'$ , respectively. Assume that mole  $v$  is mapped incorrectly to mole  $v'$ . With respect to the position of  $v$  and  $v'$  in  $A$ ,  $D$  and  $C$ , *SED* measures different distances. In this example:  $a, b, c, a', b', c' \in C$ ,  $m' \in A$  and  $n \in D$ , (b) when  $v = a \in C$ ,  $v' = c' \in C$  and  $\Pi_v(a) = c'$  then  $SED(v, v') = dist(a', c') + dist(a, c)$  (c) when  $v = n \in D$ ,  $v' = m' \in A$  and  $\Pi_v(n) = m'$  then  $SED(v, v') = dist(n, m')$ , (d) when  $v = a \in C$ ,  $v' = m' \in A$  and  $\Pi_v(a) = m'$  then  $SED(v, v') = dist(a', m') + dist(a, m')$ , similar error is measured when  $v = n \in D$  and  $v' = a' \in C$ .

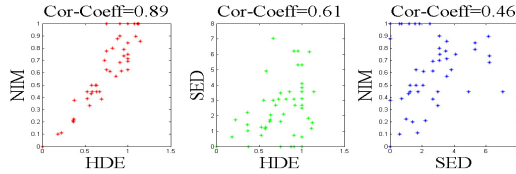


Figure 5. Scatter plot and correlation between different distance measures.

evaluate the proposed method. Ground truth expert correspondence was provided for all the moles in these pairs of images [6].

Table 2 shows the best weights obtained for the synthetic data (100 pairs of point sets times 3 cases in which  $\sigma_n$  varies from 0 to 10% of the size of the unit-square template). As expected (Figure 5), the best weights based on *HDE* and *NIM* are similar. In table 2, for cost function (2), the best weights based on *SED* are similar to the weights obtained by either *HDE* or *NIM*. Therefore, we prefer (2) as our cost function to find corresponding moles given its robustness to the evaluation metric. The final weights used in our experiments are fixed equal to the average weights (i.e. the average of the left column in Table 2). Intuitively, the cost function (2) is more preferable than (3) because (2) and (3) are related to *AND* and *OR* of different energy terms, respectively, and the best matching should ideally optimize the three energy terms, i.e. *AND-ing* them.

Next, we performed a cross validation study, in which 10 real training images [6] were used to find the best weights, and repeated for all permutations of these 10 images. The results are shown in Table 3. These weights (approximately

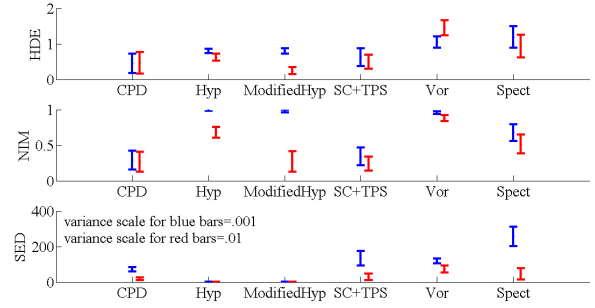


Figure 6. Point matching evaluation (using *HDE*, *NIM*, and *SED* criteria) of different methods (along the horizontal axis) when spatial image coordinates (blue bars) or normalized coordinates (red bars) are used in the matching. Refer to the text for the meaning of acronyms. Note that *SED* values for normalized coordinates will generally be smaller because of the smaller range of normalized values (i.e.  $[-0.7, 0.3] \times [-0.5, 0.5]$  as opposed to 100s for pixel coordinates). Note also that using normalized coordinates does not affect the performance of CPD, whereas it improves the accuracy of SC+TPS. Observing *HDE* and *NIM* for our proposed ModifiedHyp, we note a substantial improvement in accuracy when normalized coordinates are used.

1/3, 1/3, 1/3) are close to the weights resulting from the synthetic data experiment (Table 2), which reinforces the adoption of weights obtained from synthetic to matching moles in real images data.

In order to evaluate the proposed algorithm on real data, moles and anatomical landmarks were identified manually. A comparison between our proposed method (ModifiedHyp) and (Vor) [9], Coherent Point Drift (CPD) [14], shape contexts (SC+TPS) [2], the spectral technique of (Spect) [11] and original Hypergraph matching ( $d = 2$ ) (Hyp) [22] is presented in Figure 6. Blue and red bars show the results when *spatial image space coordinates* and *normalized space coordinates* are used, respectively. As it can be seen, the performance of CPD does not improve when normalized coordinates are used. However, CPD fails when the number of common moles is not large enough as shown in Figure 7. As it can be seen in Figure 6, the modified version of hypergraph matching (with the three different energy terms) performs worse when it is fed by image space mole coordinates. However, when our proposed anatomy-based normalized coordinates are used, mean and variance of error measurements for this method are better than the other methods. Some examples of finding corresponding moles in real image data are shown in Figure 8.

## 8. Conclusions

An automatic method for matching moles on human back images is of utmost importance for early detection of potential malignancies. Mole matching can be treated as a graph matching problem. However, general state-of-the-art point matching methods do not incorporate anatomical in-

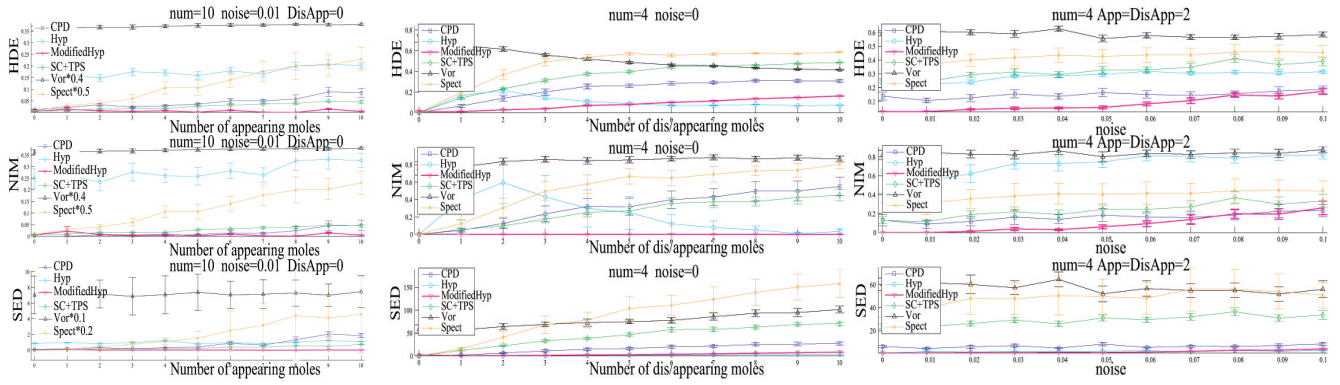


Figure 7. Synthetic experiment with known ground truth.  $num$  and  $DisApp$  denote size of  $n_V^i = n_V^i$ , and  $n_V^o$  respectively. (left) errors for fixed value of  $n_V^i$ ,  $n_V^o$  and  $noise$  for increasing values of  $n_V^i$ . (middle) errors for fixed value of  $n_V^i$  and  $noise$  for increasing values for  $n_V^o = n_V^i$ . (right) errors for fixed value of  $n_V^i$  and  $n_V^o = n_V^i$ , for increasing values for  $noise$ . As it can be seen, Modified Hypergraph matching gives better results.

Equation (2)					Equation (3)				
Case	Error	$(\omega_\alpha, \omega_\beta, \omega_\theta)$	$\bar{e}$	$\sigma^2(e)$	Case	Error	$(\omega_\alpha, \omega_\beta, \omega_\theta)$	$\bar{e}$	$\sigma^2(e)$
case1	HDE	(.30, .39, .30)	0.47	0.02	case1	HDE	(.10, .48, .40)	0.44	0.02
	NIM	(.32, .37, .30)	0.52	0.04		NIM	(.07, .50, .42)	0.43	0.04
	SED	(.34, .28, .36)	31.9	282.6		SED	(.32, .28, .39)	36.51	566.16
case2	HDE	(.31, .38, .29)	0.38	0.03	case2	HDE	(.08, .49, .41)	0.38	0.034
	NIM	(.31, .38, .29)	0.61	0.03		NIM	(.08, .49, .42)	0.60	0.03
	SED	(.34, .30, .35)	12.23	72.37		SED	(.31, .32, .35)	16.01	166.77
case3	HDE	(.35, .32, .32)	0.36	0.03	case3	HDE	(.09, .49, .40)	0.30	0.04
	NIM	(.35, .32, .32)	0.64	0.03		NIM	(.10, .49, .40)	0.62	0.07
	SED	(.34, .31, .33)	16.72	157.47		SED	(.32, .32, .34)	12.64	207.66

Table 2. The best weights resulted from training on synthetic data for the two cost functions (in (2) and (3)). Synthetic data are generated under different cases (case1-3) and each case is related to different values of the variables:  $case1 = \{\sigma_n \in [0, .1], n_V^o \in [0, 7], n_V^i \in [0, 7], n_V^i = 10\}$ ,  $case2 = \{\sigma_n \in [0, .1], n_V^o \in [0, 7], n_V^i = 0, n_V^i = 10\}$ ,  $case3 = \{\sigma_n \in [0, .1], n_V^o = n_V^i = 0, n_V^i = 10\}$ . (100 random generations for each case).  $\bar{e}$  and  $\sigma^2(e)$  represent mean and variance of the error.

Equation (2)		Equation (3)	
Error	$(\omega_\alpha, \omega_\beta, \omega_\theta)$	Error	$(\omega_\alpha, \omega_\beta, \omega_\theta)$
HDE	(.29, .40, .29)	HDE	(.01, .55, .42)
NIM	(.30, .40, .29)	NIM	(.01, .55, .42)
SED	(.32, .36, .31)	SED	(.23, .35, .41)

Table 3. The best weights resulting from the cross validation experiment with 10 training images [6].

formation. We proposed the first human back template and demonstrated, on large sets of both real and synthetic data, that the use of template-normalized spatial coordinates and additional matching cost terms outperform several state-of-the-art matching algorithms. Further, we provided an extensive study on choosing the form of the objective function, learning its optimal weights, and evaluating the resulting matching; issues often marginally addressed in the literature. Our human back atlas is the first in the field and we are

already reaping the benefits of it by substantially improving the mole matching. Nevertheless, we plan on exploring other template alternatives and anticipate other groups will do too. One of the challenges, as observed in this study, will be to develop a robust definition of the hip anatomical landmark, which we plan to address through discussions and consultations with dermatologists. The mole matching results we obtained are very encouraging from the computational and clinical perspectives. This encouraged us to pursue a fully-automated back skin segmentation and mole detection algorithms to provide an end-to-end system applied in clinical practice.

## References

- [1] Canadian cancer society / national cancer institute of canada: Canadian cancer statistics 2008. Toronto, Canada, 2008.
- [2] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE TPAMI*,

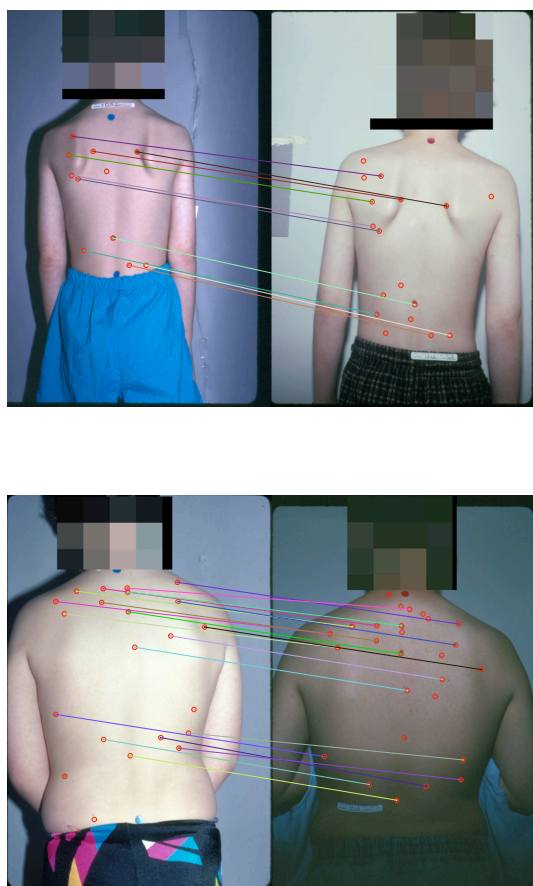


Figure 8. Each row is a different subject. The images in the right column were captured 3 years later than the left column. The resulting mole matching indicated by lines.

- 24(4):509–522, Apr 2002.
- [3] T. Cootes, C. Taylor, D. Cooper, and J. Graham. Active shape models: Their training and application. *CVIU*, 61(1):38–59, Jan 1995.
- [4] F. Ercal, M. Moganti, W. Stoecker, and R. Moss. Detection of skin tumor boundaries in color images. *IEEE TMI*, 12(3):624–626, Sep 1993.
- [5] A. Evans, M. Kamber, D. Collins, and D. MacDonald. An mri-based probabilistic atlas of neuroanatomy. *Magnetic Resonance Scanning and Epilepsy, Plenum Press*, pages 263–274, 1994.
- [6] R. P. Gallagher, J. K. Rivers, T. K. Lee, C. D. Bajdik, D. I. McLean, and A. J. Coldman. Broad-Spectrum Sunscreen Use and the Development of New Nevi in White Children: A Randomized Controlled Trial. *Journal of American Medical Association*, 283(22):2955–2960, 2000.
- [7] S. Gold, A. Rangarajan, C. ping Lu, and E. Mjolsness. New algorithms for 2d and 3d point matching: Pose estimation and correspondence. *Pattern Recognition*, 31:957–964, 1998.
- [8] E. A. Holly, J. W. Kelly, S. N. Shpall, and S. H. Chiu. Number of melanocytic nevi as a major risk factor for malignant melanoma. *J. Am. Acad. Dermatol*, 17:459–468, 1987.
- [9] H. Huang and P. Bergstresser. A new hybrid technique for dermatological image registration. *IEEE BIBE*, pages 1163–1167, Oct 2007.
- [10] T. Lee, M. Atkins, M. King, S. Lau, and D. McLean. Counting moles automatically from back images. *IEEE TBE*, 52(11):1966–1969, Nov 2005.
- [11] M. Leordeanu and M. Hebert. A spectral technique for correspondence problems using pairwise constraints. *ICCV*, 2:1482–1489, Oct 2005.
- [12] J. Maciel and J. Costeira. A global solution to sparse correspondence problems. *IEEE TPAMI*, 25(2):187–199, Feb 2003.
- [13] J. C. Mazziotta, A. W. Toga, A. Evans, P. Fox, and J. Lancaster. A probabilistic atlas of the human brain: Theory and rationale for its development : The international consortium for brain mapping (icbm). *NeuroImage*, 2(2, Part 1):89–101, Jun 1995.
- [14] A. Myronenko, X. Song, and M. Carreira-Perpinan. Non-rigid point set registration: Coherent point drift. In B. Schölkopf, J. Platt, and T. Hoffman, editors, *Advances in Neural Information Processing Systems*. MIT Press, 2007.
- [15] D. Perednia and R. White. Automatic registration of multiple skin lesions by use of point pattern matching. *Computerized Medical Imaging and Graphics*, 16:205–216, May 1992.
- [16] J.-S. Pierrard and T. Vetter. Skin detail analysis for face recognition. *CVPR*, pages 1–8, Jun 2007.
- [17] L. Ries, D. Melbert, M. Krapcho, D. Stinchcomb, N. Howlader, M. Horner, A. Mariotto, B. Miller, E. Feuer, S. Altekruse, D. Lewis, L. Clegg, M. Eisner, M. Reichman, B. Edwards, and M. Bethesda. Cancer statistics review, 1975–2005. National Cancer Institute, 2008.
- [18] J. Roning and M. Riech. Registration of nevi in successive skin images for early detection of melanoma. *ICPR*, 1:352–357, Aug 1998.
- [19] J. Talairach and P. Tournoux. *Co-Planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System : An Approach to Cerebral Imaging*. Thieme Medical Publishers, Jan 1988.
- [20] L. Thurfjell, C. Bohm, and E. Bengtsson. Cba—an atlas-based software tool used to facilitate the interpretation of neuroimaging data. *Computer Methods and Programs in Biomedicine*, 47(1):51 – 71, 1995.
- [21] V. Vezhnevets, V. Sazonov, and A. Andreeva. A survey on pixel-based skin color detection techniques. In *Proc. Graphicon-2003*, pages 85–92, 2003.
- [22] R. Zass and A. Shashua. Probabilistic graph and hypergraph matching. *CVPR*, Jun 2008.
- [23] Y. Zheng and M. Doermann. Robust point matching for nonrigid shapes by preserving local neighborhood structures. *IEEE TPAMI*, 28(4):643, Apr 2006.