

SHREC'14 Track: Automatic Location of Landmarks used in Manual Anthropometry

A. Giachetti^{1,*}, E. Mazzi^{1,*}, F. Piscitelli^{2,*}, M. Aono³, A. Ben Hamza⁴, T. Bonis⁵, P. Claes⁶, A. Godil⁷, C. Li⁷, M. Ovsjanikov^{5,8}

V. Pătrăucean⁵, C. Shu⁹, J. Snyders⁶, P. Suetens⁶, A. Tatsuma³, D. Vandermeulen⁶, S. Wuhrer¹⁰, P. Xi⁹

¹ Department of Computer Science, University of Verona, Italy

² Department of Neurological and Movement Science, University of Verona, Italy

³ Department of Computer Science and Engineering, Toyohashi University of Technology, Japan

⁴ Concordia University, Canada

⁵ Ecole Polytechnique, France

⁶ KU Leuven, ESAT/PSI - UZ Leuven, MIRC - iMinds, FH Dept.

⁷ National Institute of Standards and Technology, USA

⁸ INRIA Geometrica, France

⁹ National Research Council, Canada

¹⁰ Saarland University, Germany

* Track organizers

Abstract

In this paper we report the results of the SHREC 2014 track on automatic location of landmarks used in manual anthropometry. The track has been organized to test the ability of modern computational geometry/pattern recognition techniques to locate accurately reference points used for tape based measurement. Participants had to locate six specific landmarks on human models acquired with a structured light body scanner. A training set of 50 models with manual annotations of the corresponding landmarks location was provided to train the algorithms. A test set of 50 different models was also provided, without annotations. Accuracy of the automatic location methods was tested via computing geodesic distances of the detected points from manually placed ones and evaluating different quality scores and functions.

Categories and Subject Descriptors (according to ACM CCS): I.3.8 [Computer Graphics]: Computer Graphics—Applications

1. Introduction

The availability of whole body scanner devices is rapidly changing the way anthropometric data are collected. Sets of accurate measurements on high quality geometrical models of the subjects bodies can, in fact, be automatically obtained without the need of difficult and time-consuming manual procedures. However, to validate digital measurements and to integrate manually and digitally acquired data in longitudinal studies, it could be important to compare the classical tape-based anatomical measurements with automatic or semi-automatic measurements that can be performed on the digital models.

Manual antropometry is based on the localization by specific landmarks on the body (as indicated in the International

Society for the Advancement of Kinanthropometry manual, www.isakonline.com), that are located typically by palpation and using specific knowledge of the experienced anthropometrist. The identification of these points on digital models is not very easy, due to the variability of human subjects' features and the poor or missing geometrical characterization of the points.

However, several novel and effective methods for salient points recognition have been recently proposed in the literature thanks to the effort of the scientific groups involved in the SHREC contests. Therefore, it is interesting to see if, exploiting annotated training data and using state-of-the-art point description and pattern recognition techniques it is possible to locate automatically anthropometrical land-

marks with a reasonable accuracy. The anthropometric research community would strongly benefit from the availability of an effective information retrieval tool for this task.

2. Dataset creation

Models were originally acquired during standard anthropometric analysis in the Department of Neurological and Movement Science, University of Verona with a structured light 3D body scanner (Breuckmann BodyScan). In the original textured models, anthropometric landmarks are visible as white crosses on the skin texture, marked with a dermatographic pen. For the recognition task we selected six landmarks located in the right part of the body (Acromiale, Radiale, Stylium, Iliocristale, Trochanterion, Tibiale laterale, see figure on the left). The coordinates of the original points have been manually annotated with the Meshlab [CCC*08] Pick-Point tool and saved before the further processing steps.

These steps are:

- **Floor and noise removal.** Floor and isolated points have been removed with Meshlab scripts.
- **Remeshing.** Models have been remeshed with Poisson method, preserving high resolution detail while creating watertight triangulations
- **Simplification.** Models have been simplified in order to have approximately 50000 nodes, avoiding the creation of non-manifold edges.
- **Anonymization.** In order to avoid face recognition, we automatically identified and selected the head region exploiting the annotated landmarks, and applied to it an iterated smoothing process.
- **Coordinate transform.** In order to have an easy discrimination of left and right landmarks, we transformed the reference systems of the models in order to have the x axis approximately representing the mediolateral direction of the body with negative coordinates representing the right size. This allows an easy disambiguation of left/right symmetric salient points (we are interested in points with $x < 0$ (right part of the body)). An example of final models with the corresponding reference system is shown in Figure 1.

Annotated anthropometric points were mapped on the closest points in the new mesh.

We created a training set of 50 models. For each model of the training set and for each of the six points we provided participants with the triangle mesh in .ply format, a simple text file with extension ".pts" storing the xyz coordinates, the index of the closest vertex, the index of the closest triangle. A test set with the same number of models was similarly created, but only the meshes were available to the participants. Both training and test sets included males and female examples. All data will be available on the contest web site <http://www.andreagiachetti.it/shrec14/>.

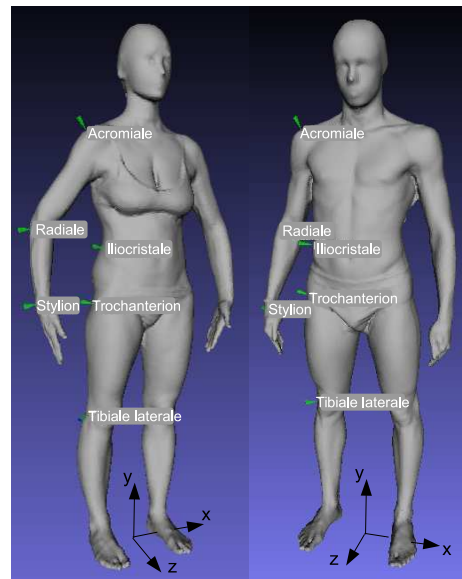


Figure 1: Examples of preprocessed models with landmarks' annotations.

3. Proposed task, participants and proposed techniques

Participants were asked to localize the six specific landmarks in the 50 models of the test set, given the corresponding points localization in the training set. They could provide either all the three fields included in the original annotations (3D coordinates, closest node, closest face) or only one of the fields (in this case we estimated the coordinates of the points from nodes or face indexes or the indexes from the 3D coordinates for evaluation).

Six points localization methods have been proposed by five different research groups. In the following, these methods are briefly described.

3.1. Graphical Models (GM) by C.Shu, P. Xi and S. Wuhrer

This method is based on the observation that anthropometric landmarks are usually located on the skeletal joints of the human body. The way a human expert locates landmarks is by palpating the subject. This suggests that landmarks are where the bones are close to the skin. Therefore, the surface geometry local to the landmarks may have distinguishable characteristics. On the other hand, the relative locations of the landmarks reflect the human body structure. This information can constrain the locations of the landmarks. Authors used a Markov network to represent the structure of the landmarks. The nodes of the graph represent the landmarks. The edges of the graph encode the distance relationship between neighboring landmarks (Figure 2). Finding the locations of the landmarks is defined as finding a set of vertices on the

scan such that its configuration is most compatible with the Markov network. In this way, the landmark locating problem is casted as a classification problem.

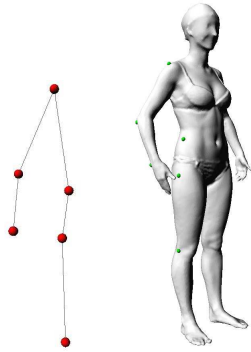


Figure 2: Landmark graph

The graphical model is trained using the manually marked models. Once trained, belief propagation method is used to find the optimal solution [YFW01].

When the Markov network is applied, the nodes of the graph represent landmarks and the structure of graph captures the relative location of the landmarks on the human body. In this case there are only six landmarks and they are located on a linear curve. In general, the graph can be more complex, even with loops.

Node features chosen by authors were curvatures and SPIN images. They are both popular surface descriptors for 3-D models. To compute a SPIN image at a point on the surface, a cylindrical-like coordinate system is created using the surface normal at the point. Every other point of the model is projected onto this coordinate system and an image is created by counting the number of points falling into each bins [Joh97]. As a surface descriptor, SPIN image has a few advantages. It is invariant to rigid transformation. Depending on the range of projection, it can be both local and global. It is also relatively robust to noise. The drawback of SPIN image is that it is not a compact descriptor – it is necessary to store an image for every vertex of the model. Therefore, it is computationally expensive to use it for comparison. We remedy this problem by compressing the SPIN images using Principal Component Analysis.

Curvature has also been added to the local surface description. Certain landmarks, like the one on the elbow, have clear distinguishable curvatures, invariant to rigid transformation. However, curvature is defined on continuous surfaces and the estimation of curvature on a mesh is sensitive to noise.

Based on the surface descriptors, a potential function $\phi_i(l_i)$ for each landmark l_i is defined, representing the likelihood of a vertex to be the i th landmark. For each edge $\{l_i, l_j\}$ a compatibility potential $\psi_{ij}(l_i, l_j)$ is defined, representing the spatial constraints of the two variables to be assigned

consistently. The joint probability over the Markov network is given by:

$$p(L) = \frac{1}{Z} \prod_i \phi_i(l_i) \prod_{i,j} \psi_{ij}(l_i, l_j) \quad (1)$$

where Z is a normalizing factor.

To train the graphical model, the node potential function and edge potential function are modeled as Gaussian distributions of the feature vectors and the parameters of the distribution functions are computed. For the node potential, the feature vector consists of the SPIN image and the curvature values at each landmark location. The edge potential is modeled by a Gaussian distribution of the Euclidean distance between the two neighboring landmarks. In the training stage, the parameters of the distribution functions are estimated from the data.

Finding the landmarks amounts to maximizing the function defined by equation 1. This is a large-scale optimization problem. Authors used belief propagation [YFW01] to solve this problem.

The implementation proposed is based on the method described in Ben Azouz et al [ASM06]. It has been extended to predict landmarks for arbitrary postures [WAS10].

Running time: According to the authors, the training time for the method on a 4-core Pentium CPU was six hours. The labelling time for each test case about three minutes.

3.2. Spectral geometry based methods (SM) by C. Li, A. Godil and A. Ben Hamza

Authors of this contribution adapted the spectral geometry based methods [Li13] for landmark location in anthropometry. These methods are based on the eigendecomposition of the Laplace-Beltrami operator (LBO). A course-to-fine procedure was applied to detect a given type of point. It consists of two stages: (1) candidate estimation of landmarks in the real line [LH13a], and (2) landmark detection via the minimal sum dissimilarity of spectral graph wavelet signature [LH13b].

Candidate estimation Given a triangle meshed shape X , the second eigenfunction of LBO is the smoothest mapping from the manifold to the real line, resulting in the ordering of the vertices by their second eigenfunction values quite stable [LH13a]. Therefore, Authors use the index of ordered vertices as the intrinsic coordinate, and normalized them in the interval $[0, 1]$. For each training sample, its landmarks can be represented as $P = (p_1, \dots, p_6)$, where p_i is the intrinsic coordinate of a specific landmark. To solve the two-sign possibilities problem in the ordering, Authors used k-means to separate all P into two sets, within which the orderings are consistent with manifolds. Authors inverted the vertices ordering of the landmarks in one cluster, and the corrected coordinate align well with the the ones in the other cluster.

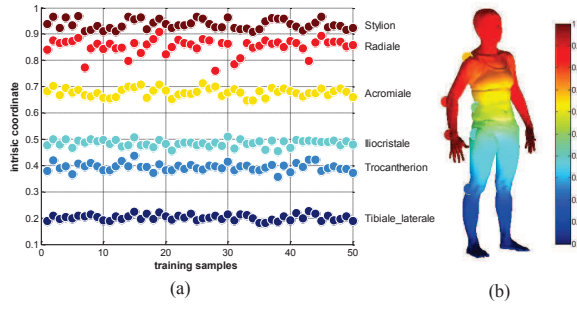


Figure 3: Statistic of landmarks on training set. (a) The corrected intrinsic coordinates of 50 training samples. (b) Body shape colored in the second eigenfunction

Figure 3 illustrates the statistic of corrected landmarks' coordinates on training samples.

For each specific type of landmark ℓ , the mean u_ℓ and standard derivation σ_ℓ of its intrinsic coordinates were obtained. The candidate vertices were estimated as the ones with the coordinate within $[u_\ell - \frac{1}{2}\sigma_\ell, u_\ell + \frac{1}{2}\sigma_\ell]$.

Landmark detection: The second stage aims to finding the optimal hypothesis from the candidate vertices. It is implemented via the computation and comparison of the spectral descriptor $h(x)$ at each candidate vertex on the testing samples and labeled landmarks on the training samples. In general, any one of spectral descriptors with the eigenfunction-squared form reviewed in [LH13c] can be used. Authors used the spectral graph wavelet signature (SGWS) $S_x(t, x) = \sum_{i=1}^m g(t, \lambda_i) \phi_i^2(x)$ as the local descriptor; $g(t, \lambda_i)$ is set as a cubic spline wavelet generating kernel and the scaling function is considered. The resolution level is set as 5.

Finally, each vector-valued descriptor is normalized in L_2 norm, and dissimilarity between two descriptors is computed as the Euclidean distance. The optimal hypothesis is determined as the the vertex with minimal sum dissimilarity to all the labeled landmarks in that type.

Running time: The method is implemented in MATLAB and is influenced by the computational complexity of the LBO. For the proposed meshes with 50K vertices, it takes about 250 seconds to compute the first 30 eigenpairs of LBO. To compute the signature on a mesh, it takes less than one second. To find the specific points on a testing sample, it costs about 1.3 seconds. This means that the training time was about 3.5 hours and finding points on a new model takes about 4 minutes.

3.3. Augmented Point Feature Histograms (APFH), by A. Tatsuma and M. Aono

Authors estimated specific points using a new local 3D model descriptor and a support vector machine (SVM) clas-

sifier. The new local 3D model descriptor is known as augmented point feature histogram (APFH). APFH expands point feature histogram (PFH) [RMBB08] by adding the statistics of its geometric features. PFH is a local feature vector for 3D point clouds that construct a histogram of geometric features extracted from neighboring oriented points. To improve the discriminant power of PFH the mean and covariance of its geometric features have been added.

An overview of the proposed APFH is illustrated in Figure 4. With APFH, we first randomly generate oriented points on a triangular surface of a 3D model using the method proposed in [OFCD02]. To generate random point \mathbf{p} on an arbitrary triangular surface comprising vertices \mathbf{v}_a , \mathbf{v}_b , and \mathbf{v}_c , we employ the following formula:

$$\mathbf{p} = (1 - \sqrt{r_1})\mathbf{v}_a + \sqrt{r_1}(1 - r_2)\mathbf{v}_b + \sqrt{r_1}r_2\mathbf{v}_c. \quad (2)$$

In the implementation proposed, two random variables, r_1 and r_2 in the above equation, are computed using the Niederreiter pseudo-random number generator [BFN94]. The oriented point is generated by inheriting the normal vector of the surface as an orientation of the point.

Next, a PFH is built for each oriented point. A PFH locates the k -neighborhood for each oriented point and calculates the four-dimensional (4D) geometric feature $\mathbf{f} = [f_1, f_2, f_3, f_4]^T$ as proposed in [WHH03]. The 4D geometric feature is defined for each pair of points \mathbf{p}_a and \mathbf{p}_b in the k -neighborhood and for their normal vectors \mathbf{n}_a and \mathbf{n}_b as follows:

$$\begin{aligned} f_1 &= \arctan(\mathbf{w} \cdot \mathbf{n}_b, \mathbf{u} \cdot \mathbf{n}_a), \\ f_2 &= \mathbf{v} \cdot \mathbf{n}_b, \\ f_3 &= \mathbf{u} \cdot \frac{\mathbf{p}_b - \mathbf{p}_a}{d}, \\ f_4 &= d, \end{aligned}$$

where $\mathbf{u} = \mathbf{n}_a$, $\mathbf{v} = (\mathbf{p}_b - \mathbf{p}_a) \times \mathbf{u} / \|(\mathbf{p}_b - \mathbf{p}_a) \times \mathbf{u}\|$, $\mathbf{w} = \mathbf{u} \times \mathbf{v}$, and $d = \|\mathbf{p}_b - \mathbf{p}_a\|$. The PFH collects the 4D geometric features in a 16-bin histogram \mathbf{f}_h . The index of the histogram bin h is defined as follows:

$$h = \sum_{i=1}^4 s(t, f_i) \cdot 2^{i-1},$$

where $s(t, f)$ is a threshold function defined as 0 if $f < t$ and 1 otherwise. We set the threshold value of f_1 , f_2 , and f_3 to 0, and f_4 to the average value of f_4 in the k -neighborhood.

Furthermore, means and covariances of the 4D geometric features are computed. Here, let \mathbf{f}_i be a 4D geometric feature of an oriented point in the k -neighborhood. The mean feature \mathbf{f}_m and covariance feature \mathbf{f}_c in the k -neighborhood are defined as follows:

$$\mathbf{f}_m = \frac{1}{k} \sum_{i=1}^k \mathbf{f}_i,$$

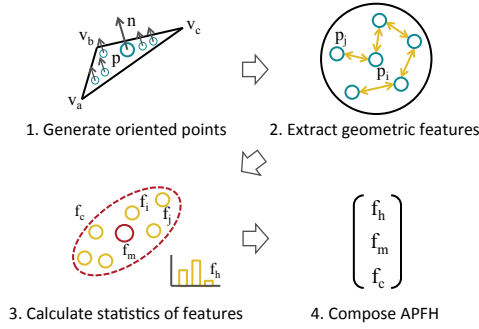


Figure 4: Overview of the Augmented Point Feature Histograms (APFH)

$$\mathbf{f}_c = \text{Upper} \left(\frac{1}{k-1} \sum_{i=1}^k (\mathbf{f}_i - \mathbf{f}_m)(\mathbf{f}_i - \mathbf{f}_m)^T \right),$$

where $\text{Upper}(\cdot)$ concatenates the upper triangular part of the matrix. APFH \mathbf{f}_{APFH} comprises \mathbf{f}_h , \mathbf{f}_m , and \mathbf{f}_c .

Coordinates of the oriented point are also added to the APFH feature. Finally, the APFH vector is normalized by ℓ_2 normalization.

For the recognition of specific points, a probability estimation algorithm using the SVM classifier [CL11, WLW04] was applied. Set of APFHs extracted from the oriented points located in the four neighborhoods in each labeled point are used as the training dataset. In addition, randomly select distant oriented points were selected from the labeled points and their APFHs were added to the training dataset as miscellaneous labeled data.

In the test phase, APFHs extracted from the oriented points (with $x < 0$) were given to the classifier assigning the label to the oriented point with the maximum probability for each label, except for miscellaneous labels.

APFH parameters were set empirically. The number of points chosen was 16384 and that of neighborhoods to 160. For the SVM implementation, LIBSVM [CL11] was used and optimization of the parameters was achieved using an automatic script in the LIBSVM tools.

Running time: On a PC with Intel(R) Xeon(R) CPU E3-1275 V2 @ 3.50GHz CPU, 32 GB of memory, the feature extraction (average running time per 3D model) took 6.616s, the SVM training, 0.544s, the SVM predict 0.231s.

3.4. Landmark detection using ICP by T. Bonis, M. Ovsjanikov, V. Pătrăucean

Due to the fact that the variability in the test data is well represented in the training models, a *brute force* approach can be considered. Specifically, authors first ran the *rigid* ICP algorithm between every pair $\langle M_{train}^i, M_{test}^j \rangle$,

$i = 1..N_t, j = 1..N_t$, and for every test model kept the most similar training model. The degree of similarity between two models is given by the distance error between M_{train}^i and RM_{test}^j , where R is the rigid transform (i.e. translation, rotation) reported by ICP. The smaller the error, the better the alignment.

Once obtained for every test model the most similar model from the training set, a *non-rigid* ICP was performed for each pair, and finally the landmark annotations were transferred from the training model to the (non-rigidly deformed) test model.

The success of this approach relies on the assumption that the training set contains a fairly similar model for every test model.

Running time: A significant drawback is represented by the complexity of the approach. Whilst the rigid ICP takes a few seconds for each pair, hence few minutes to find for each test model the most similar training model, the non-rigid ICP can take up to 90 minutes to complete one pair. Hence a rough estimate of the computational time needed to detect the landmarks for one model can go as high as 100 minutes on a regular (quad core) laptop. Thus, the practicality of the approach is limited, since a human could label the landmarks in only few minutes.

3.5. Landmark detection using local features (LF) by T. Bonis, M. Ovsjanikov, V. Pătrăucean

In this *machine learning* approach, authors extracted descriptors for the sought landmarks from the training models, and used them to detect the landmarks on the test models.

This was accomplished by first extracting some reliable stable points on the mesh, called by the authors *anchor points*, and then by characterizing the sought landmarks with respect to these anchor points. 8 different anchor points were used, which correspond to: extremities of the hands and of the legs, top of the head, axillae and pelvis region. Authors chose these points as they represent local maxima and minima of the body curvature and have high detection rates in terms of repeatability. To detect them, they computed HKS (Heat Kernel Signature) descriptors with a small time step, and then identified the most prominent local maxima and minima using persistence.

Subsequently, the sought landmarks were characterized through the geodesic distances with respect to these anchors, and through HKM (Heat Kernel Maps) signatures that use the anchors as source points. Additionally, HKS (Heat Kernel Signature) descriptors were computed for each landmark. This yielded 38-dimension descriptors. For each landmark, authors considered the corresponding descriptors from all training models, and applied PCA to reduce the dimension of the descriptor space to 20.

At test time, to label a new mesh, authors computed

the 38-dimension descriptor for every point of the mesh, and projected the obtained descriptors in the 20-dimension spaces associated to the sought landmarks. Then they computed the sum of the distances from each descriptor to the 15 nearest neighbours of the training set, and the vertex having the smallest cumulative distance is declared to be the corresponding landmark.

Running time: This approach takes less than one minute to label a shape.

3.6. Surface to surface registration (STS) by J. Snyders, P. Claes, D. Vandermeulen, P. Suetens

This method is based on surface-to-surface registration. One sample in the training set was taken as an initial template and registered to each of the other training samples. The resulting objects represented the shape of the training samples, but with corresponding points. Subsequently, geometric averages for both the surface and the sample markers were calculated. Remeshing was performed to ensure a good quality mesh. Once the template was built, it was registered to the test set. The template markers were transformed along, yielding an estimation of the anatomical landmarks that were targeted.

The surface-to-surface registration was based on the original work of TPS-RPM by Chui et al. [CR03] In an iterative process, a correspondence search followed by the computation of a non-rigid spatial mapping was performed until convergence. In this framework, as originally proposed, a deterministic annealing process decreased the need for parameter tuning.

Determining correspondences was done using weighted k-Nearest Neighbours. The weighing is inversely proportional to the Euclidean distance in a 6-dimensional space. The first three dimensions account for the x-, y- and z-coordinates of each node's location, the last three for the (scaled) x-, y- and z-coordinates of each node's surface normal. The mapping of the template towards its corresponding points was performed by a visco-elastic transformation model. The reader is referred to [SCVS14] for further details.

Running time: On a Mac with 2.7GHz Intel Core i5 processor and 16GB 1333 MHz DDR3 memory, both training and obtaining the results each took around 40 minutes total running two processes in parallel. Surface-to-surface registration takes about 1.5minutes per registration.

4. Evaluation

Only one method proposed (SM) annotated precise nodes as output, all the others provided free coordinates on the surfaces. In any case we computed estimated closest nodes and faces for all the annotated points and checked when these points/faces corresponded to the manually annotated ones.

As shown in Tables 1 and 2, this happens in very few cases, and this indicates that the triangulation used is sufficiently dense for our analysis.

Geodesic distances of extracted points, added to the meshes, from the manually annotated corresponding ones have been then computed, and different validation measurements have been evaluated:

- Average, standard deviation and median of the geodesic distances between corresponding points (see Table 3)
- False negative error vs geodesic radius (percentage of missed point detections considering correctly detected only the landmarks within one geodesic radius from the target one, see Figure 5).

Errors are higher than inter-human operator ones. Reference values found in literature for the median of inter-observer manual operators errors are 11.5 mm. for Acromiale, 5.0 mm. for Radiale, 5.3 mm. for Stylion, 10.8 mm. for Trochanterion, 12.3 mm. for Iliocristale, 10.3 mm. for Tibiale [KM11]. These inter-observer errors are considered high and the cited study shows that they can make derived measures performed by the two operators not comparable. This means that current automatic landmarking methods cannot locate the points with sufficient accuracy and that digital anthropometry should currently rely on specific methods or geometrically salient landmarks. The only landmark where the lowest automatic error is close to the human one is Acromiale, that is probably the one with the highest geometric saliency.

There is not a single method outperforming all the others, but STS and GM are globally the best ones. STS seems able to locate better Acromiale and Tibiale laterale. GM outperforms all the other methods in the arms (Stylion and Radiale). These two methods provide similar results for Iliocristale, while GM is not so accurate for locating Trochanterion, where the other registration-based approach, ICP provides results very close to STS.

STS and GM have also a reasonable efficiency and can compute landmarks in a few minutes, a time suitable for practical applications. APFH is clearly the faster method.

These results suggest that registration-based methods (or methods based on global optimization) provide in general better results than those based on local features, even if context-aware. Surface registration seem to be less effective for the landmarks in the arms, where the graph based method is more accurate, but performs better in the other regions.

It should be noted that these methods should be less accurate in case of varying poses (GM is based on SPIN images, not pose independent, and registration is harder with non-rigid transforms). Spectral methods like SM and LF should instead provide similar accuracy with different poses, but their errors were, however, rather high in this test.

The reason for the poor performance of spectral methods

probably consists in the fact that the dataset of this contest includes a variety of different subjects, and transforms among these subjects are not isometric. A simple method suggested by the authors to improve SM labelling performances would be to estimate the candidates and label the testing model according to the several most similar subject body shapes in the training set or the training shape with most similar 2nd eigenfunction.

A larger training set would be certainly useful to increase the accuracy of these modified methods and of other methods as well.

	Acr.	Rad.	Styl.	Troc.	Iliocr.	Tib.
SM	0	0	0	0	0	0
GM	3	1	3	1	1	1
APFH	1	1	0	1	1	0
ICP	0	0	0	0	1	2
LF	0	0	0	0	0	0
STS	2	0	3	0	3	4

Table 1: Number of "closest nodes" corresponding to ground truth.

	Acr.	Rad.	Styl.	Troc.	Iliocr.	Tib.
SM	0	0	0	0	0	0
GM	1	2	1	0	0	1
APFH	0	1	2	0	0	0
ICP	0	0	3	0	0	0
LF	0	0	0	0	0	0
STS	2	2	1	0	2	2

Table 2: Number of "closest faces" corresponding to ground truth.

	meas.	Acr.	Rad.	Styl.	Troc.	Iliocr.	Tib.
SM	mean(mm)	84.47	57.93	50.05	109.03	73.25	64.55
	st.dev.(mm)	53.96	47.20	20.03	56.13	38.48	47.37
	median(mm)	70.62	38.14	49.66	114.87	71.36	49.44
GM	mean(mm)	17.74	17.70	14.21	34.26	22.82	22.98
	st.dev.(mm)	11.47	13.32	10.49	17.83	12.25	14.06
	median(mm)	15.45	14.14	12.11	30.87	20.70	20.80
APFH	mean(mm)	37.63	30.49	25.44	46.83	36.63	31.76
	st.dev.(mm)	65.58	27.37	16.35	38.31	22.91	26.46
	median(mm)	19.36	21.25	20.45	40.19	29.93	23.50
ICP	mean(mm)	22.19	24.45	25.86	28.17	29.42	26.39
	st.dev.(mm)	16.74	15.26	12.64	12.55	14.86	10.53
	median(mm)	18.02	21.45	22.13	26.69	25.40	25.59
LF	mean(mm)	20.00	33.50	31.81	38.56	41.00	32.33
	st.dev.(mm)	10.93	23.07	21.62	22.46	23.28	17.48
	median(mm)	17.80	24.50	25.33	32.63	39.69	29.55
STS	mean(mm)	12.40	27.22	24.70	25.78	22.90	15.61
	st.dev.(mm)	6.68	19.55	21.36	13.82	12.96	8.66
	median(mm)	11.60	23.05	17.21	25.24	19.99	15.46

Table 3: Average geodesic distance / standard deviation of geodesic distance / median geodesic distance of detected points from manually annotated ones for the different methods proposed.

5. Discussion

The limited accuracy obtained by automatic methods to identify anatomical landmarks on human body meshes does not mean that it will be impossible to have an automatic localization with accuracy similar to an expert anthropometrist in the near future. It must be considered, in fact, that participants had a limited amount of time to optimize their methods and the training set was relatively small.

Methods based on registration procedures performed better thanks to the use of global optimization processes and could be improved by using more complex and task-specific registration models. Methods based on spectral point descriptors were penalized by the non isometric transforms mapping points across different body types. They could greatly benefit from the availability of a large training set, by combining the retrieval of most similar training shapes to the mapping of the landmarks annotated on them on the new shape. Landmarks localization is strictly related to the problem of retrieving the closest shapes in the training set or finding a good, continuous interpolation between training shapes optimally fitting the test models.

We believe that the availability of contest data and results will be extremely useful for researchers interested in improving the automatic landmarking accuracy and in comparing manual and digital anthropometry.

References

[ASM06] AZOUZ Z. B., SHU C., MANTEL A.: Automatic locating of anthropometric landmarks on 3D human models. In *International Symposium on 3D Data Processing, Visualization and Transmission* (2006). 3

[BFN94] BRATLEY P., FOX B. L., NIEDERREITER H.: Programs to generate Niederreiter's low-discrepancy sequences. *ACM Trans. on Math. Software* 20, 4 (Dec. 1994), 494–495. 4

[CCC*08] CIGNONI P., CALLIERI M., CORSINI M., ET AL.: Meshlab: an open-source mesh processing tool. In *Eurographics Italian Chapter Conference* (2008), The Eurographics Association, pp. 129–136. 2

[CL11] CHANG C.-C., LIN C.-J.: LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2 (2011), 27:1–27:27. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>. 5

[CR03] CHUI H., RANGARAJAN A.: A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding* 89, 2 (2003), 114–141. 6

[Joh97] JOHNSON A.: *SPIN-Images: A Representation for 3-D Surface matching*. Tech. rep., PhD thesis, Robotics Institute, Carnegie Mellon University, 1997. 3

[KM11] KOUCHI M., MOCHIMARU M.: Errors in landmarking and the evaluation of the accuracy of traditional and 3d anthropometry. *Applied ergonomics* 42, 3 (2011), 518–527. 6

[LH13a] LI C., HAMZA A. B.: Intrinsic spatial pyramid matching for deformable 3d shape retrieval. *International Journal of Multimedia Information Retrieval* 2, 4 (2013), 261–271. 3

[LH13b] LI C., HAMZA A. B.: A multiresolution descriptor for deformable 3d shape retrieval. *The Visual Computer* (2013), 1–12. 3

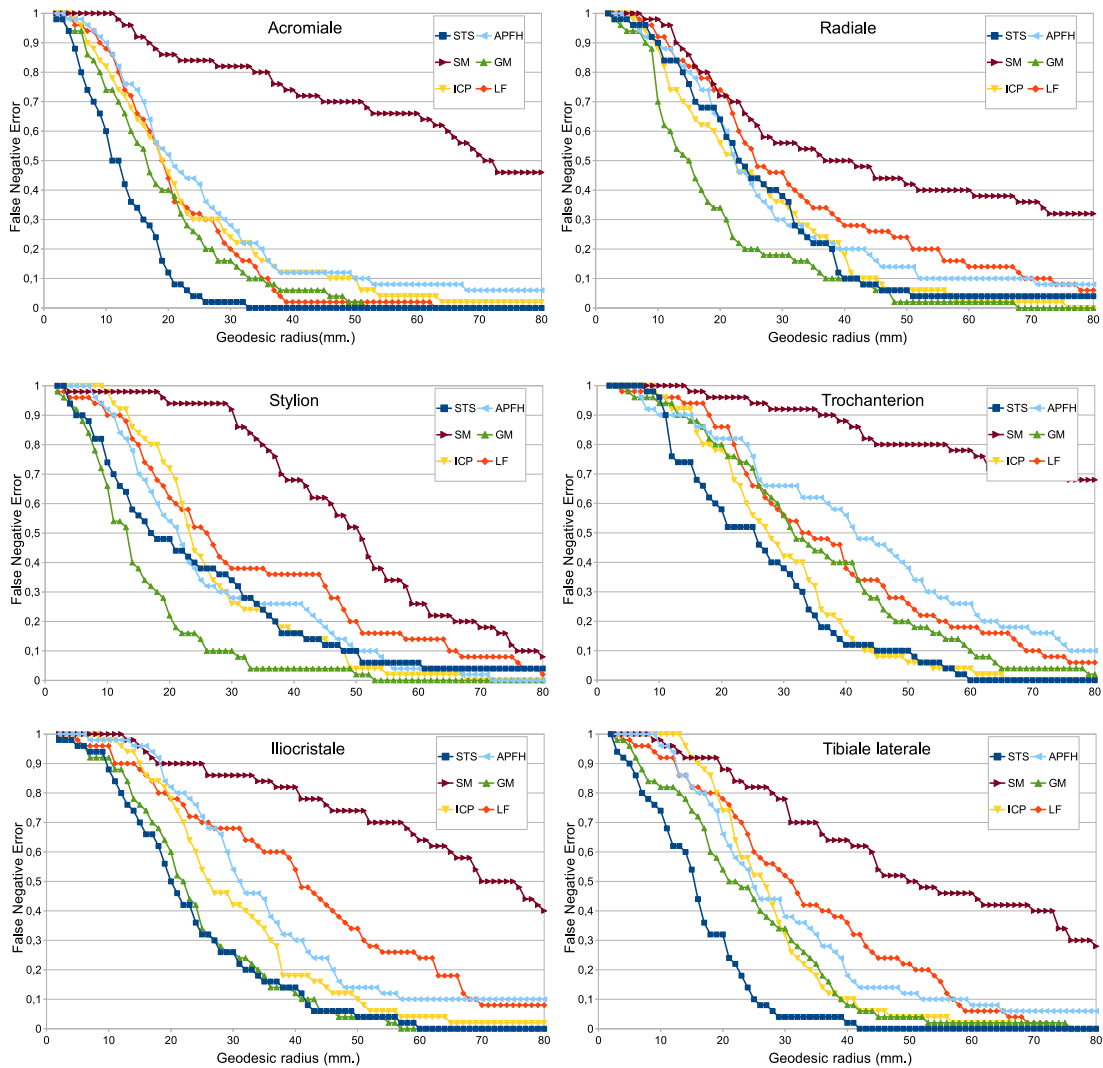


Figure 5: False negative error (percentage of non detected points) vs geodesic radius for the six points with the six methods proposed.

[LH13c] LI C., HAMZA A. B.: Spatially aggregating spectral descriptors for nonrigid 3d shape retrieval: a comparative survey. *Multimedia Systems* (2013), 1–29. 4

[Li13] LI C.: *Spectral Geometric Methods for Deformable 3D Shape Retrieval*. Master’s thesis, Concordia University, 2013. 3

[OFCD02] OSADA R., FUNKHOUSER T., CHAZELLE B., DOBKIN D.: Shape distributions. *ACM Transactions on Graphics* 21 (2002), 807–832. 4

[RMBB08] RUSU R. B., MARTON Z. C., BLODOW N., BEETZ M.: Persistent point feature histograms for 3D point clouds. In *Proceedings of the 10th International Conference on Intelligent Autonomous Systems* (2008). 4

[SCVS14] SNYDERS J., CLAES P., VANDERMEULEN D., SUETENS P.: *Development and Comparison of Non-Rigid Surface Registration Algorithms and Extensions*. Tech. rep., KU Leuven, Department of Electrical Engineering, ESAT/PSI, Med-

ical Image Computing, UZ Leuven, Medical Imaging Research Center, iMinds, Future Health Department, January 2014. 6

[WAS10] WUHRER S., AZOUZ Z. B., SHU C.: Semi-automatic prediction of landmarks on human models in varying poses. In *Canadian Conference on Computer and Robot Vision* (2010). 3

[WHH03] WAHL E., HILLENBRAND U., HIRZINGER G.: Surflet-pair-relation histograms: A statistical 3D-shape representation for rapid classification. In *Proc. of Int. Conference on 3D Digital Imaging and Modeling* (2003), pp. 474–482. 4

[WLW04] WU T.-F., LIN C.-J., WENG R. C.: Probability estimates for multi-class classification by pairwise coupling. *The Journal of Machine Learning Research* 5 (2004), 975–1005. 5

[YFW01] YEDIDIA J., FREEMAN W. T., WEISS Y.: Understanding belief propagation and its generalizations. In *International Conference on Artificial Intelligence (IJCAI 2001), Distinguished papers track* (2001). 3