

Robust Automatic Labelling of Anatomical Landmarks on 3D Body Scans

Andrea GIACHETTI^a, Umberto CASTELLANI^a, Christian LOVATO^b, Carlo ZANCANARO^b

^aDepartment of Computer Science, University of Verona, Verona, Italy;

^bDepartment of Neurological, Neuropsychological, Morphological and Movement Sciences, University of Verona, Verona, Italy

Abstract

In this paper we analyze the performance of a pipeline for the extraction and semantic labelling of geometrically salient points on acquired human body models, improving the quality of the results and discussing its robustness against pose and body type variations.

Following the approach introduced in [10], we use a heat diffusion approach automatically detecting points as maxima of the autodiffusion function and using supervised classification to assign them a semantic label related to the anatomical part where the point is located. The resulting map can be used to perform measurements or to detect pose. The motivation of this approach is related to the fact that landmarks used in traditional anthropometry are not easily identified in digital models because they are localized by palpation and are not geometrically salient. An anatomical measurement system should be instead based on purely geometrical or image based landmarks. It is therefore interesting to find points of this kind that can be recognized in different subjects.

The use of heat diffusion analysis enables a robust salient point detection at different levels of detail and the creation of rich point descriptors not depending only on local geometry but also on global context, invariant with respect to articulated deformations (pose variation) and sufficiently stable against changes in body type.

In this work we improved our previous semantic labelling approach based on heat diffusion by selecting optimal point descriptors and feature-space distances and applying a hierarchical coarse to fine correction performing the classification at different scales and propagating the assigned labels from the coarser scale to the finer ones. Furthermore we tested the method on a collection of models representing different body types and with approximately fixed or largely different poses in order to demonstrate the robustness of the pipeline.

Experimental results show that this approach can be used to recognize robustly at least a selection of landmarks on subjects with different body types and independently on pose and could therefore be applied for automatic anthropometric analysis.

Keywords: Salient points, Heat diffusion, 3d body scanning

1. Introduction

In the near future reliable and flexible tools for the automatic analysis of human body models, as those acquired by 3D scanners, will become extremely important for anthropometric and medical applications, due to the necessity of efficiently handling the huge amount of information captured by these devices.

In order to automatically extract relevant indicators and measurements of interest, these tools should be able, for example, to identify meaningful locations on the body surface and to characterize them with geometrical descriptors, to compute geodesic distances between them or performing other measurements, possibly doing this in a standardized way and independently on the acquisition device.

Different approaches have been applied to achieve these goals. Classical segmentation based approaches apply generic [1,2] or specifically designed [3,4,5] methods to extract the different body parts that can be then characterized with local measurements.

Accurate human body partitioning in anthropometric applications is, however, usually obtained by systems exploiting specific protocols, poses, machines, etc. and are therefore not good for examining heterogeneous data or stored archives, while generic methods tested on human meshes are do not usually provide an accurate partitioning in specific anatomical regions.

A complete point to point correspondence between human meshes can be obtained by registering SCAPE models encoding pose and principal body type variations [6,7] with the scan data. However they require input correspondences and do not provide precise matching of relevant structure.

* andrea.giachetti@univr.it ; <http://profs.sci.univr.it/~giachetti>

A completely different approach consists of searching for anatomical landmarks on acquired models recognizing them on different subjects. The task is not easy due to the fact that the local characterization of points is poor and non local descriptors or graph based methods are needed to perform it. To simplify the task, methods used for anthropometric-like applications use posture constraints, as in the case of commercial measurements systems.

Furthermore, if the task required is the identification of classical anthropometric locations that are not necessarily geometrically salient [8,9] the localization is even more difficult or impossible.

In this paper we follow the approach presented in a previous work [10] showing the possibility of using heat diffusion based methods to detect and semantically label specific points on human models. We improved the original pipeline by introducing a multi-scale hierarchical classification and selecting with specific tests features and metrics to measure point similarity. Furthermore, we tested the performances of the system on a dataset including male and female, overweight and pose-varying subjects in order to verify the robustness of the system and the influence of different factors on classification results. Results show that some anatomical locations can be robustly detected and labelled automatically without relying on a previous segmentation or graph matching.

2. Salient point detection and semantic labelling

The basic idea of the system consists of extracting salient points using the autodiffusion function (ADF) and characterizing them at corresponding scales in different models using ADF related descriptors. ADF is a linear combination of the eigenfunctions of the Laplace-Beltrami operator that has a simple physical interpretation as how much heat remains in a location x after time t if a heat impulse has been applied in the same position (see [11,12] for details)

The computation of this function on scanner data requires a preprocessing of the acquired models, consisting mainly in cleaning, remeshing the data with the Poisson method [13] and decimating the resulting model (we perform the final processing on 15K vertices meshes).

After this step we perform the Laplace-Beltrami decomposition of the models, approximating models with the first 200 eigenvalues and computing their autodiffusion function (ADF) at roughly corresponding sampled scales (determined by the time variable in the heat diffusion process). This is done by taking 101 equally spaced samples in the logarithmic scale in the interval $[t_{\min}, t_{\max}]$ where $t_{\min}=4 \log(10/\lambda_{200})$ and $t_{\max}=4 \log(10/\lambda_2)$ being λ_i the eigenvalues of the ADF. It is also possible to introduce a new variable s that has values corresponding to the sample indexes (0,1,... 100) at the corresponding locations in order to simplify the analysis. We can in this way represent the ADF of different body shapes in a common scale space, obtaining a function $ADF(s)$ that, being the function invariant for isometric transforms, should be approximately corresponding in different subjects for similar s . This allowed us to compare ADF and derived descriptors in the range of interest using standard measures of dissimilarity.

Salient points can be extracted as local maxima of the ADF at different scales. At $s=0$ (finest scale in the considered range), the set of salient points computed includes the most relevant anatomical locations. Salient points extracted at $s=100$ (coarsest scale in the domain) are only those related to fingers, toes, and head. Salient points are stable in the sense that those that have been detected at a particular value of s are approximately constant when the parameters s is decreased. When s is decreased new points, possibly with new labels appear. At $s=25$ approximately half of the total labels are present, so that we decided to test supervised classification of the salient points also at this scale.

2.1. Semantic labels of automatically extracted points

To obtain the semantic labels we asked to experienced anthropometrists to give semantic labels to the points detected at the three scales chosen for the analysis ($s=1$, $s=25$, $s=100$). This is a critical aspect, because, according to the experts, not all the points correspond to a specific unambiguous anatomical location and have been in those cases labelled as belonging to a large body region. Furthermore, some of the point classes defined can have only one instance in a specific body while other classes can have multiple instances. In any case, the set of labels used for the reported experiments include 24 salient points/regions that are: head tip, ear, nose, chin, scapular region, shoulder, pectoral, abdomen, pubic, crotch, gluteal, trochanterion, femoral region, knee, crural, heel, toe, arm, elbow, forearm, finger, vertebral, other (see details in [10]). These labels have been used to train the automatic classifiers developed and to test their accuracy.

2.2. Automatic labelling procedure and proposed improvements

The procedure proposed in [8] for the automatic semantic labelling of the points consists of using ADF related descriptor, e.g. heat kernel signature [10] and using a nonlinear supervised classifier (nearest neighbor or dissimilarity based) to label points of a subject on the basis of examples of other subjects. This classification task proposed is not simple, due to the fact that, even if the HKS-based descriptor of a point should not depend on pose, it clearly varies with individual body features both local (affecting HKS at low t) and global (affecting HKS values at high s).

This means that corresponding points are probably not concentrated in small regions of the feature space and that classes are not easily separable. For this reason we performed several tests in order to improve the accuracy of the classification by changing features and learning algorithm. In particular, we considered different descriptors obtained from the HKS, and different metrics to compare them.

As alternative descriptors that can replace the HKS we propose the first order derivative of the heat kernel with respect to the index s (DHKS) and the second order derivative (D^2 HKS). The reason of this choice is that these functions should be less influenced by eigenvalue normalization and are easy to be computed using finite differences from the original HKS.

As alternative metrics we applied some of the dissimilarity measures widely used in histogram comparison, e.g. histogram intersection distance, chi square dissimilarity, angular distance, Minkowski and diffusion distances [14]. Furthermore, we introduced an improvement of the method in order to exploit the multiscale nature of the salient points. Having, in fact, the possibility of extracting and automatically labelling the points at the three levels of detail ($s=0$, $s=25$, $s=100$) where the annotation are available, and being, according to the experts, salient points stable across scales, we correct the labels estimated at the finer scales by propagating labels detected at the coarser one to the points with the the lowest geodesic distance.

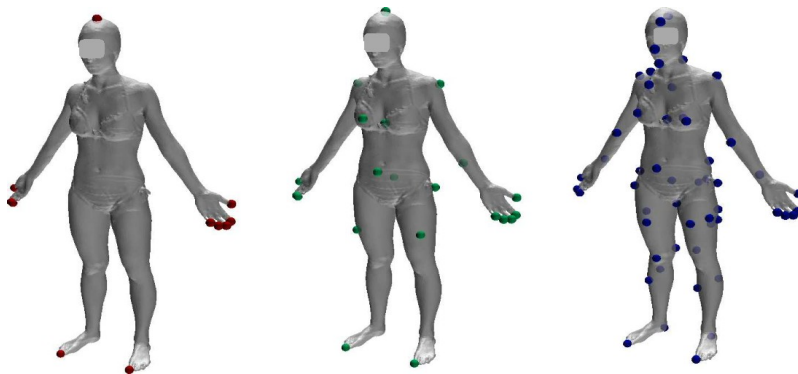


Fig. 1. Salient points extracted on one of the acquired models at $s=100$ (left), $s=25$ (center), $s=1$ (right).

This procedure is expected to reduce errors thanks to the more accurate classification at coarser scales. All these variations of the method have been tested on a dataset with models of different subjects with varying body types and pose and poor quality due to hair and holes in order to test the robustness of landmark recognition. Results are presented in the next section.

3. Experimental testing

The dataset used in our experiments has been provided by the anthropometry laboratory of the Department of Neurological, Neuropsychological, Morphological and Movement Sciences of the University of Verona. The laboratory performs 3D body scanning during normal anthropometric routine using a structured light based body scanner device (Breuckmann BodyScan).

This scanner creates high resolution (400k vertices, 1mm. resolution) meshes with an acquisition time of 5 seconds. The complete dataset consists of 80 human body meshes that can be divided in four categories equally represented: normal women (20), obese women (20), normal men in similar pose (20), normal men in largely varying pose (20). This choice allowed us to test not only the detection and recognition of landmarks, but also the effect of pose and body type on it. Models present also relevant problems related to unmasked hair, holes and artifacts.

3.1. Detection of points at different scales

For all the model we performed the mesh processing pipeline described before, computing ADF values at the corresponding values of the parameter s and detecting salient points at different scales and making them label at $s=1$, $s=25$ and $s=100$. Fig. 1 show examples of detected salient point at these three scales. If we analyze points extracted at these three scales and assigned labels, we can

observe that at $s=100$ only points related to fingers, toes and head tip appear, only 3 semantic labels are given and at least one point with each label are detected. Not all the five fingers/toes are always detected due to limited mesh quality. At $s=0$ we have the complete set 24 semantic labels described before, even if not all the points are detected for all models. A number of relevant salient points are, however, always detected, such as knee, heel, shoulder, crural, femural, gluteal, elbow. At the intermediate scales approximately half of the labels are assigned, but most of these labels are not assigned in all the models.

Automatic labelling of points

With the same approach described in [10] we tested a supervised classification approach to give automatic labels to the salient points extracted on models using a leave subject out approach: e.g. we use for each experiment points extracted on one subject as test set and all the others as training set, summing/averaging then the errors to compute average classification error rates.

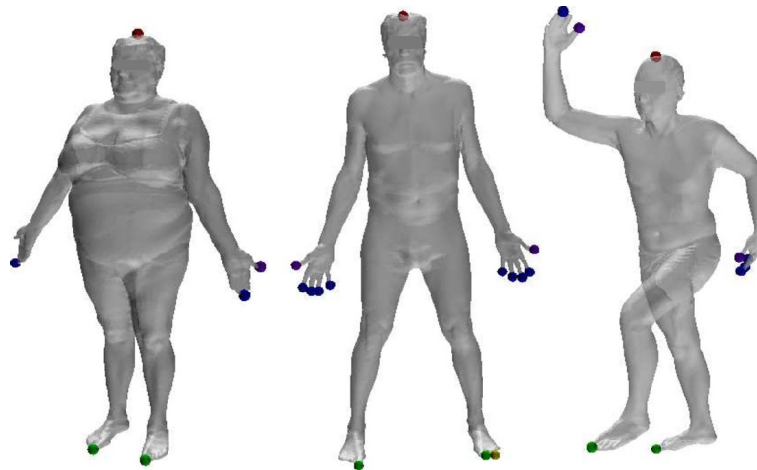


Fig. 2. labelling at $s=100$ is always correct independently on pose or body type.

Points are normally described by the so-called heat kernel signature that is the function $ADF(s)$ sampled in the previously defined range [11,12]. As in [10] we tried also the sampled first order derivative of ADF with respect to the variable s and, in this new work also the sampled second order derivative of the ADF as a new descriptor. For the classification step we tested different methods. In this way we verified that the use of dissimilarity based algorithms and other nonlinear methods do not provide accuracy improvements with respect to the nearest neighbor approach that seemed in this case the best option due to the high complexity and dimensionality of the feature space. Relevant improvements in accuracy can be obtained by changing the descriptors and the metric used to compare them. In particular, we found that first and second order derivatives of the ADF with respect to s improves the average accuracy and also the use of different distance functions is relevant. Table 1 shows the accuracy obtained for the recognition task at the three different scales.

Labels given at the coarsest scales are always correct, while at the finer scales errors appear.

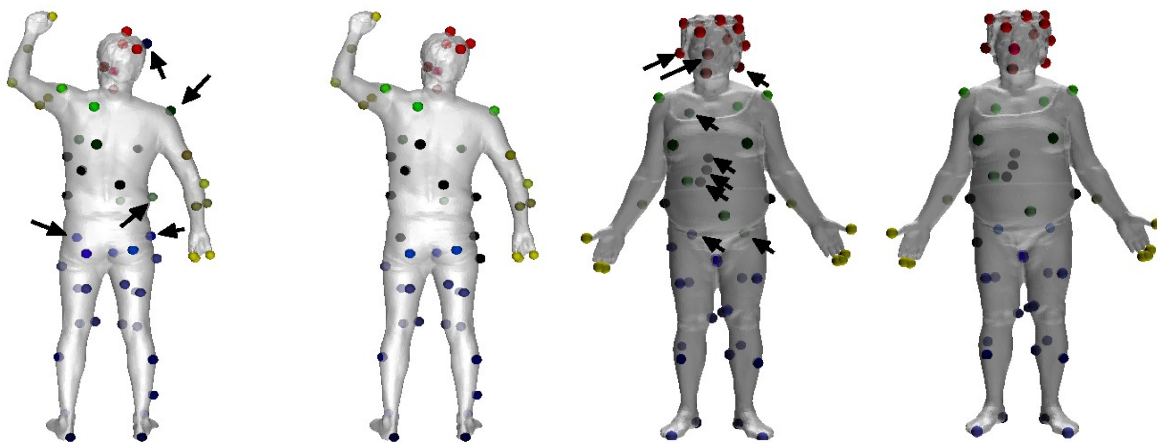


Fig. 3. Salient points labelled with the automatic method at $s=1$ (left, different colors characterize different points/regions) and true labels (right). Wrong labels are indicated by the arrows.

Table 1. Nearest Neighbor cross validation (leave subject out) accuracy of the salient point classification at different scales using different descriptors and dissimilarity functions.

distance	Crossvalidation error								
	S=1			S=25			S=100		
	HKS	DHKS	D ² HKS	HKS	DHKS	D ² HKS	HKS	DHKS	D ² HKS
Euclidean	0,236	0,175	0,156	0,142	0,123	0,113	0,000	0,000	0,000
Intersection	0,253	0,696	0,427	0,164	0,743	0,376	0,000	0,000	0,000
Chi square	0,232	0,385	0,223	0,142	0,235	0,147	0,000	0,000	0,000
Angular	0,203	0,168	0,157	0,116	0,121	0,104	0,000	0,000	0,000
Minkowski	0,237	0,171	0,159	0,145	0,122	0,115	0,000	0,000	0,000
Diffusion	0,236	0,175	0,152	0,115	0,115	0,113	0,000	0,000	0,000

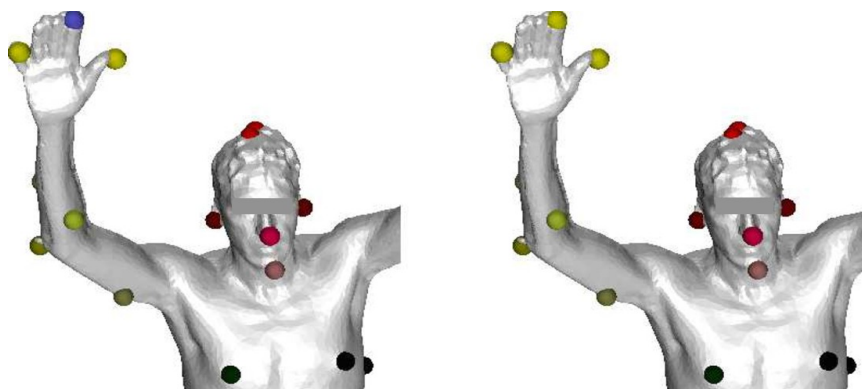


Fig. 4. Coarse to fine propagation of labels can remove wrong results. Left: multi-class labelling at $s=0$ with an error in a finger point (blue sphere) probably due to poor local characterization created by acquisition artifacts. Right: correct result (yellow) obtained propagating across scales labels extracted at $s=100$.

Being possible to link salient points across scales we tested finally the hierarchical labelling of the points, propagating the label assigned at the coarsest scale to the finer one. Table shows that these procedure is able to correct some errors that may occur (see Fig. 4).

Table 2. Best average crossvalidation errors are reduced using the coarse to fine label propagation.

distance	Crossvalidation error					
	S=1			S=25		
	HKS	DHKS	D ² HKS	HKS	DHKS	D ² HKS
Single scale	0,203	0,169	0,152	0,115	0,115	0,105
Multiscale	0,187	0,161	0,146	0,108	0,112	0,099

The acquired dataset allowed us also to test the influence of body type on the semantic classification. In fact, performing tests on subset of only men in fixed poses (20 models), men in varying poses (20), normal women (20) and overweight women (20) we found (see table 3) pose does not affect relevantly the accuracy, while female subjects are more difficult mainly due to the more relevant body variability and the fact that we used raw models with hair not covered, creating spurious salient points often badly classified (see Fig. 3).

To verify the possible use of the salient point recognition for practical applications it is interesting to see the accuracy obtained for particular classes, e.g. points that are frequently/always detected and can be recognized robustly. The analysis reveals that most of the feature points that classes are always represented are recognized with accuracy equal to 1 or larger than 0.9. This is true for toes, fingers, head tip, knee, crural region, chin, shoulder. This means that the method could be used in practical applications, to compare features or measurements across different subjects or just to initialize registration method or complete mesh segmentation procedures directly exploiting connectivity information (e.g. graph cut).

Table 3. Best average crossvalidation errors on single model classes.

	minimum average error
fat women	0,145
women	0,204
men varying poses	0,116
men standard pose	0,108

4. Discussion

Improving the approach presented in a recent work, we analyzed the possibility of using heat diffusion to perform a semantic labelling of the geometrically salient points extracted on routinely acquired scans of human bodies not subject to quality, pose or body type constraints. The improvements here proposed are related to the use of new descriptors, distance measures and a coarse to fine reasoning. Furthermore the method has been tested on a new set of models including subjects with relevantly varying body types and pose. Results, despite clear limitations due to spurious salient points and difficulties in experts labelling, show that it is possible to classify correctly all the points detected at the coarsest scale and about 85% of the points extracted at the finest scale, and that some of the geometrically salient points can be extracted and correctly classified in almost all acquired models, even independently on sex, body type and pose, just using heat diffusion based descriptors even without using neighboring salient point information or graph matching.

References

1. Attene, M., Katz, S., Mortara, M., Patane, G., Spagnuolo, M., Tal, A.: Mesh segmentation - a comparative study. In: Proceedings of the IEEE International Conference on Shape Modeling and Applications, p. 7. IEEE Computer Society (2006)
2. Kalogerakis M, Hertzmann A., Singh K. Learning 3D Mesh Segmentation and Labeling, TOG 29 3, Siggraph 2010
3. Mortara, M., Patanè, G., Spagnuolo, M.: From geometric to semantic human body models. Computers & Graphics pp. 185–196 (2006)
4. Lovato, C., Castellani, U., Giachetti, A.: Automatic segmentation of scanned human body using curve skeleton analysis. In: Proceedings of Mirage, pp. 34-45. LNCS (2009)
5. Xiao, Y, Siebert, J.P., and Werghi, N. (2003) *A discrete Reeb graph approach for the segmentation of human body scans*. In: Fourth International Conference on 3-D Digital Imaging and Modeling, 6-10 October 2003, Banff, Alberta, Canada.
6. Anguelov, D., Srinivasan, P., Koller, D., Thrun, S., Rodgers, J., Davis, J.: SCAPE: shape completion and animation of people. ACM Trans. Graph. 24(3), 408-416 (2005).
7. Hasler, N., Stoll, C., Sunkel, M., Rosenhahn, B., Seidel, H..P.: A statistical model of human pose and body shape. Computer Graphics Forum 28(2), 337– 346 (2009).
8. Azouz, Z.B., Shu, C., Mantel, A.: Automatic locating of anthropometric landmarks on 3d human models. 3D Data Processing Visualization and Transmission, International Symposium on 0, 750-757 (2006)
9. Wuhler, S., Shu, C., Xi, P.: Landmark-free posture invariant human shape correspondence. The Visual Computer pp. 1–10 (2011). DOI 10.1007/s00371-011-0557-z.
10. Lovato C., Zancanaro C., Castellani U, and Giachetti A. Heat Diffusion Approach for Feature-based Body Scans Analysis proc. Eurographics 2011 Workshop on 3D Object Retrieval pp. 33-40
11. Gebal, K., Baerentzen, J.A., Aanaes, H., Larsen, R.: Shape analysis using the auto diffusion function. In: In SGP (2009)
12. Sun, J., Ovsjanikov, M., Guibas, L.: A concise and provably informative multi-scale signature based on heat diffusion. In: Proceedings of the Symposium on Geometry Processing, pp. 1383-1392 (2009)
13. Kazhdan M, Bolitho M, and Hoppe H.. Poisson surface reconstruction. In *Proceedings of the fourth Eurographics symposium on Geometry processing* (SGP '06). Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 61-70.
14. Puzicha, J., Buhmann, J., Rubner, Y., Tomasi, C.: Empirical evaluation of dissimilarity measures for color and texture. In: Computer Vision, 1999. IEEE International Conference on, vol. 2, pp. 1165-1172 vol.2 (1999)