Regression-Based Locating Landmark on Dynamic Humans

Deok-Kyeong Jang
Korea Advanced Institute of Science and Technology
291 Daehak-ro, Yuseong-gu
Daejeon 34141, Republic of Korea
shadofex@kaist.ac.kr

Sung-Hee Lee
Korea Advanced Institute of Science and Technology
291 Daehak-ro, Yuseong-gu
Daejeon 34141, Republic of Korea
sunghee.lee@kaist.ac.kr

ABSTRACT

We present a novel framework that consists of two-level regressors for finding correlations between human shapes and landmark positions in both body part and holistic scales. To this end, we first develop pose invariant coordinates of landmarks that represent both local and global shape features by using the pose invariant local shape descriptors and their spatial relationships. Our body part-level regression deals with the shape features from only those body parts corresponding to a certain landmark. For this, we develop a method that identifies such body parts per landmark, by using geometric shape dictionary obtained through the bag of features method. Our method is nearly automatic, requiring human assistance only once to differentiate the left and right sides, and shows the prediction accuracy comparable to or better than those of existing methods, with a test data set containing a large variation of human shapes and poses.

CCS CONCEPTS

• **Computing methodologies** → *Learning linear models*; *Shape analysis*;

KEYWORDS

KCCA, regression, segmentation, landmark detection

ACM Reference format:
Deok-Kyeong Jang and Sung-Hee Lee. 2017. Regression-Based Locating Landmark on Dynamic Humans. In *Proceedings of SCA '17, Los Angeles, CA, USA, July 28-30, 2017,* 2 pages.
DOI: 10.1145/3099564.3106645

1 INTRODUCTION

Analyzing and understanding human body shapes are important problems in computer graphics research, with a number of applications such as registration, retargeting, and shape retrieval. Anatomical landmarks on human bodies are essential features for obtaining anthropometric information, but detecting the anatomical landmarks remains a challenging problem due to high variedness of human shapes and poses.

A conventional approach for landmark detection is by finding correspondences between a template body model with annotated landmarks and a particular body shape, typically through mesh registration methods [Chui and Rangarajan 2003]. This registration-based approach is effective for the body shapes similar to the template, but has a limited capability in generalizing to the whole range of human shapes with different poses. A more principled and potentially powerful approach would be learning and predicting the relationship between various body shapes and their landmark locations. In addition, a good method should not require complex preprocessing, such as alignments, on the input data.

In this paper, we solve the anatomical landmark detection problem on human models with dynamic poses, by training a statistical regression model that learns connections between human body shapes and landmarks. When trained with a data set of a wide range of human shapes and poses, the regression-based method achieves a higher performance than registration-based approaches.

2 OUR APPROACH

Specifically, we show that the kernel canonical correlation analysis (KCCA) method successfully models the correlation between human body shapes and landmarks.

We develop several key ideas that enable robust landmark detection against severe variations in shape and pose. First, we develop a novel method that serves as pose invariant coordinates of landmarks. This is achieved by representing the position of a landmark with the feature vector in the pose invariant, local descriptor space. In addition, since the mapping from vertices to local descriptor space is non-injective (e.g., left and right Stylions have the same feature vector), we augment this representation with the spatial relationship information between landmarks, which is also described in a pose invariant manner.

Second, in order to increase the regression accuracy, we develop a two-level regression method that separately regresses body part-level features and holistic features. The part-level regressor models the connection between the local feature of a landmark and the shape of the body parts that are related with the landmark. To identify such body parts per landmark, we develop a method that uses geometric shape dictionary obtained through the supervised bag of features (S-BoF [Litman et al. 2014]) method. The segmentation of body parts allows for training the lower-level regressor with respect to the only body parts that are related with the landmark, and thus blocks the effect from other unrelated body parts. The global level regressor relates the holistic shape characteristics with the spatial relationship between landmarks, and this overcomes the non-injectivity of local features and allows for finding landmarks nearly automatically.

SCA '17, Los Angeles, CA, USA

DOI: 10.1145/3099564.3106645

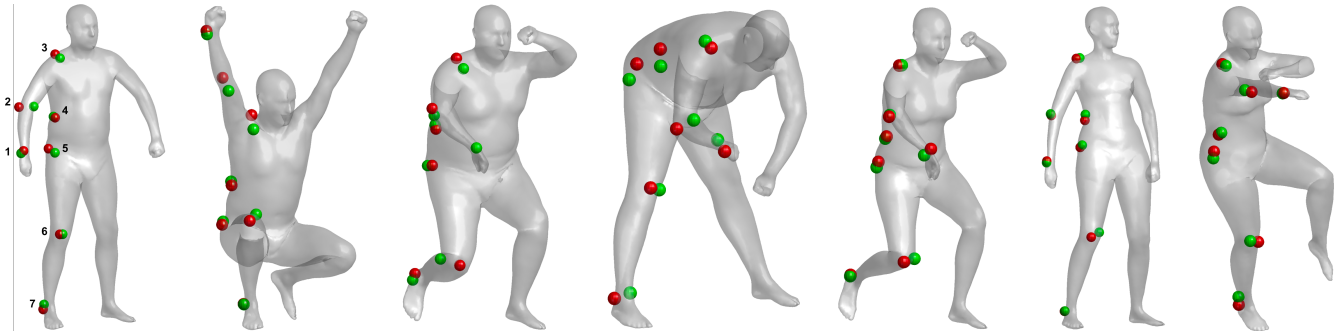

Figure 1: Results of landmark detection on human models with dynamic poses. Red markers are ground truth landmarks annotated manually(tylion (1), Radiale (2), Acromiale (3), Iliocristale (4), Trochanterion (5), Patella (6) and Malleolus (7). See leftmost figure) and green markers are estimated landmarks using our method.

3 EXPERIMENT

We constructed a training data set of 160 models by randomly selecting four poses from each subject. After that, we manually marked seven landmarks (i.e., stylion, radiale, acromiale, iliocristale, trochanterion, patella, and malleolus. See Fig. 1) on the right side of human body. We created a test data set of 40 models by randomly selecting one pose from each subject.

We applied our method to 40 test human models. Figure 1 shows the results of located landmarks for dynamic poses of a variety of human models. One can see that our method estimates landmark positions with reasonably good accuracy. However, among the 40 test human models, the four human models show that the landmarks are detected in the other symmetric side.

These four human models were classified as error data and the remaining 36 test models were analyzed. For each landmark, the difference between the ground truth and the predicted landmarks was calculated using both geodesic and Euclidean distances.

Table 1: Average, standard deviation, and median of geodesic (left) and Euclidean (right) distances between the estimated landmarks and the ground truth data.

Landmark	mean(mm)		st.dev.(mm)		median(mm)	
Sty.	17.80	15.81	16.79	12.61	12.35	12.10
Rad.	78.51	63.17	38.43	28.03	82.64	65.37
Acr.	63.03	58.95	36.05	31.87	59.87	56.55
Ilio.	50.85	48.32	32.60	30.28	42.76	40.80
Troc.	51.09	48.94	29.27	27.56	47.77	45.47
Pate.	91.14	82.25	46.22	38.51	89.57	84.56
Malle.	41.98	29.55	43.53	24.09	22.85	21.62

Table 1 shows the average, standard deviation, and median of distance errors for each landmark. The average error of all landmarks is less than $9.1cm$ in geodesic distance and $8.2cm$ in Euclidean distance. However, we also have four invalid test data which stamped wrong side (left side) out of the 40 test data sets. Overall, the performance of our method is better than [Wuhrer et al. 2010, 2011] that shares the similar purpose with our paper. Although [Wuhrer et al. 2010, 2011] tested against less challenging landmarks for less varying poses while ignoring the symmetry problem , our approach generally produces better results with respect to performance and computation time.[1]

4 CONCLUSION AND FUTURE WORK

In this paper, we introduced novel methods for landmark detection using S-BoF and KCCA. The main contributions of our work are the segmented body part descriptor using geometric dictionary from S-BoF, and the pose invariant landmark coordinates combining the landmark position in the local descriptor space with the spatial relationships between landmarks. Our method allows for efficient detection of landmarks independently of alignment, vertex ordering, and pose changes.

There are many additional future research directions. While we only deal with the landmarks that are located on the surface, our method can be used for the regression of points inside the body as well. A good application would be to predict the locations of joints from the body shape. Lastly, our landmark detection method can be used to various applications, such as global shape alignment that can be performed around the detected landmarks, and registration and one-to-one correspondence detection using the matched landmarks. Also, it may help subject-specific skinning through finding joint positions through regression.

REFERENCES

Haili Chui and Anand Rangarajan. 2003. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding* 89, 2 (2003), 114–141.

Roee Litman, Alex Bronstein, Michael Bronstein, and Umberto Castellani. 2014. Supervised learning of bag-of-features shape descriptors using sparse coding. In *Computer Graphics Forum*, Vol. 33. Wiley Online Library, 127–136.

Stefanie Wuhrer, Zouhour Ben Azouz, and Chang Shu. 2010. Semi-automatic prediction of landmarks on human models in varying poses. In *Computer and Robot Vision (CRV), 2010 Canadian Conference on*. IEEE, 136–142.

Stefanie Wuhrer, Chang Shu, and Pengcheng Xi. 2011. Landmark-free posture invariant human shape correspondence. *The Visual Computer* 27, 9 (2011), 843–852.

[1] We cannot make exact comparison with ours and [Wuhrer et al. 2011] due to the differences in database and landmarks. Instead, we made a rough comparison with respect to three landmarks (radiale, acromiale and patella) that overlap with our landmarks. Euclidean distance errors of radiale, acromiale and patella from [Wuhrer et al. 2011] are 76.78, 122.3 and 97, whereas our approach marks 63.17, 58.95 and 82.25, respectively.