Automatic Hand-Over Animation using Principle Component Analysis

Nkenge Wheatland
UC Riverside*

Sophie Jörg
Clemson University

Victor Zordan
UC Riverside

Abstract

This paper introduces a method for producing high quality hand motion using a small number of markers. The proposed "hand-over" animation technique constructs joint angle trajectories with the help of a reference database. Utilizing principle component analysis (PCA) applied to the database, the system automatically determines the sparse marker set to record. Further, to produce hand animation, PCA is used along with a locally weighted regression (LWR) model to reconstruct joint angles. The resulting animation is a full-resolution hand which reflects the original motion without the need for capturing a full marker set. Comparing the technique to other methods reveals improvement over the state of the art in terms of the marker set selection. In addition, the results highlight the ability to generalize the motion synthesized, both by extending the use of a single reference database to new motions, and from distinct reference datasets, over a variety of freehand motions.

Keywords: character animation, motion capture, hand motion, dimensionality reduction, PCA

1 Introduction

Producing quality whole-body motion involves the movement of the hand in relation to the rest of the body. However, using a motion capture system, it can be difficult to record the full body of a moving person while also capturing the hand and all of its detail because the whole-body and hand appear at largely different scales. While it is possible to record a high-resolution capture of the hand through a comprehensive set of markers (typically 13-20 markers), this is often only possible in a small capture region, isolating the motion of the hand. However, in a larger, full-body capture region, the complete set of markers becomes difficult to discern, and so this approach is usually abandoned in lieu of the capture of a smaller set of markers (2-6 markers) coupled with a "hand-over" process for reconstructing the full hand animation [Kang et al. 2012]. In this paper, we propose a robust technique to accomplish the latter that both automatically selects the "sparse" marker set to record, and subsequently produces joint trajectories for a full hand from the sparse marker set.

Our technique employs a combination of principle component analysis (PCA) [Bishop 1995] to construct a low-dimensional representation of the hand data along with linearly weighted regression (LWR) [Atkeson et al. 1997] to aid in the reconstruction. Starting from a reference database that is recorded using a full-resolution marker set, we first determine the best sparse marker set to record based on the PCA representation of the data. We experiment with different test sizes for the marker set to record, specifically reduced marker sets of six and three markers, and we compare our selection method with different ones proposed for selecting the markers, including manual selection, following Hoyet et al. [2012], and a method that uses representative cluster-based search for selection [Kang et al. 2012]. In contrast, the technique in this paper computes the marker set directly from the PCA, and our findings show that this marker set is superior to the other methods of selection for the reconstruction technique we propose. For reconstruction, our method employs a second PCA in a synthesis step combined with LWR. Starting from a test query that records only the sparse marker set, we use LWR to build a locally sensitive model between the markers and the principle components.

We use American Sign Language (ASL) as our primary testbed. ASL is an important and interesting freehand application of hand motion. Further, it includes a rich, diverse set of configuration poses for the hands. We show that we can construct new (unseen) ASL signs with high-visual quality using a simple, generic ASL database. Generalization of the database reveals that we can use our technique to capture other motions, such as counting. Our effort holds close similarities to previous work, especially the full-body motion control of Chai and Hodgins [2005]. In contrast, our main contributions include the distinct exploration of rich hand data, such as ASL, as well as our method for determining the best reduced marker set to take advantage of the power of dimensionality reduction realized by PCA. Further, our approach is far simpler and lends itself to ease-of-use and re-implementation. Our approach also has notable advantages over other related papers for hand-over animation, such as the work of Hoyet et al. [2012] and Kang et al. [2012] in that we compute the best reduced marker set directly, rather than selecting it manually or through brute-force search. Compared to other techniques, ours is both simple to implement and fast to compute, striking a valuable compromise which is likely to lead to greater adoption for commercial use.

2 Related work

The detailed and subtle motions of the hands are hard to capture. Several approaches for recording have been suggested, each with advantages and disadvantages. In particular, optical motion capture systems, while being very accurate, can require substantial post-processing to handle occlusions and mislabelings. Cyber Gloves [2013] and the like are robust to captures in larger spaces, but they require regular calibration [Wang and Neff 2013] and do

*e-mail:wheatlan@cs.ucr.edu

not provide high enough accuracy for many applications [Kahlesz et al. 2004]. For other camera-based or range-scan type systems, the hand needs to be in a confined space and the body can not captured synchronously [Wang and Popović 2009; Zhao et al. 2012; Wang et al. 2013]. The lack of robust recording solutions has lead to a practice of "hand-over" animaton. Specifically *hand-over* refers to the (general) manual post-production of hand motion, following a lower-detail or no capture of a talent's hand. Algorithms have been proposed to generate hand motion automatically based only on the motion of the body [Jörg et al. 2012] or also on contacts with objects [Ye and Liu 2012]. A different approach suggests to capture the body and hand motions separately and to combine them afterwards [Majkowska et al. 2006]. However, none of these ensure that the resulting hand motion the same as the one performed during the initial body capture, a goal of our work.

Our aim specifically focuses on facilitating the quality capture of hand motions, together with full-body motions, in a motion capture system. To reach this goal, we investigate the most effective way to capture accurate hand motions using the smallest possible number of markers and suggest a corresponding, specialized hand-over technique to reconstruct the full hand from the markers. Other researchers have analyzed finger motions and found strong correlations between different degrees of freedom. Rijpkema and Girard [1991] report that the relationship between the flexion of the distal and the proximal interphalangeal joint (DIP and PIP, respectively) is approximately linear, with $DIP = 2/3 * PIP$. Jörg and O'Sullivan [2009] show how to reduce the degrees of freedom of the hands by eliminating irrelevant and redundant information. These approaches reveal that finger motion is highly redundant. We take advantage of correlations between different degrees of freedom of the hand to optimize the capturing and construction of high quality hand animation.

Principal component analysis, as a standard technique to analyze and reduce high-dimensional data, has been used to effectively synthesize body motions [Safonova et al. 2004] and also to study hand movement. Braido and Zhang [2004], show that for the hand the two first principal components of a PCA accounted for over 98% of all variance in the joint angles. However, their motion database did not take into account the thumb and involved only two types of tasks - cylinder grasping and voluntary flexion of individual fingers. Santello et al. [1998] studied a variety of grasp poses and found that over 80% of the measured degrees of freedom could be described by their first two principal components. Chang et al. [2007] used supervised feature selection to find a set of five optical markers that could classify grasps with a 92% prediction accuracy. However, these studies, applied to grasps, do not require specific motions from individual fingers. In contrast, we present a method applied to American sign language (ASL), which exhibits impressive dexterity and variety of finger motions [Courty and Gibet 2010]. We hypothesize that there is less redundancy in typical finger motions of ASL than in standard grasping motions.

One of our goals is to determine which is the most effective set of markers for capturing. Previous work has studied the best marker sets for hand motions, for example, by testing and comparing marker sets chosen manually (with reconstruction done using inverse kinematics) [Hoyet et al. 2012], or through a brute-force approach, that compares the error of similar poses found in a database [Kang et al. 2012]. Chai and Hodgins [2005] studied full-body motion with similar goals to the ones described in this paper, but use a manually selected marker set.

3 Overview

Our overall technique is divided into two stages: 1) the computation of the sparse marker set; and 2) the reconstruction of the full-resolution, skeleton-driven hand animation from the sparse marker set. For our study, we collect full-resolution motion capture data of hand motions in a small capture area. The actor wears 13 small (6mm) markers directly on the hands as well as three markers on the lower forearm. The lower forearm acts as the root link for our hand skeleton with the assumption that these same three markers will appear in full-body captures. To account for gross body hand motion, marker positions in the database are put into the same coordinate frame by computing the transformation of each marker relative to the root link. Our hand model consists of 18 joints. For the results in this paper, we construct two databases, one for sign language and the other freehand gesture data. The sign language database includes only the alphabet, but we are able to construct novel *word* signs such as "girl" and "walk". Beyond, ASL, we also employ the same gesture database as Kang et al. [2012]. That gesture database includes a series of expressive hands motions and is used to reconstruct novel sequences with similar qualities.

In the first phase, we perform PCA over the markers of the reference data to derive a rank ordering for the markers based on their influences over the principle components. From this rank-ordered list, we select the top markers to act as our sparse marker set. For the second phase, reconstruction, we set up a locally weighted regression (LWR) model to map from the sparse marker positions to estimated principle components. In this case, PCA is applied to the joint angles. The LWR model is built for each test query based on the input markers for the query and their proximity to the analogous markers in the reference data, after correcting for the lower arm (root) movement. Joint angles for the low dimensional input are reconstructed by reversing the PCA, going from principle components derived from the LWR to a newly computed full set of joint angles.

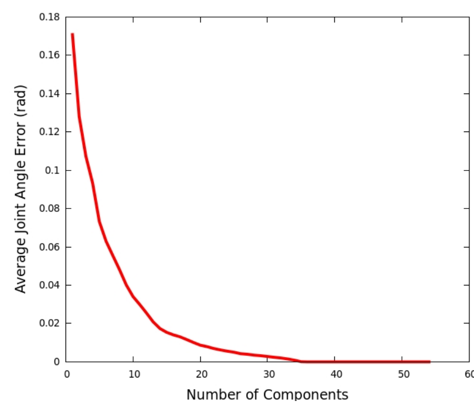

Figure 1: *Dimensionality reduction for sign language database. PCA is capable of using as few as ten components with relatively small average errors.*

4 Hand Motion Dimensionality and PCA

At the core of our technique is the assumption that hand motion is relatively low-dimensional. Even though a full resolution skeleton of the hand can have several dozen degrees of freedom (DOF), many of the DOFs of the hand show correlations while others have barely any motion. Thus, the inherent dimensionality of the hand motions is much lower [Santello et al. 1998; Braido and Zhang

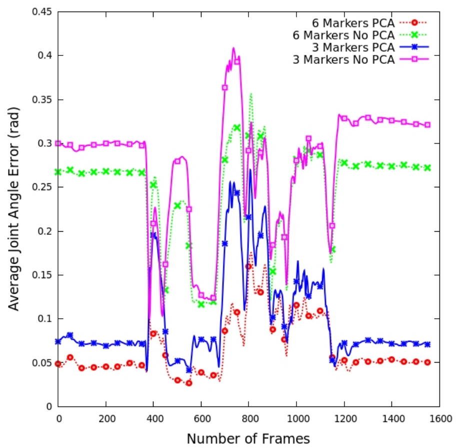

Figure 2: *Sign language sample motion with and without PCA employed. Note the error for six markers without PCA is larger than that of three markers with it.*

2004; Jörg and O'Sullivan 2009]. In our approach, PCA is used to exploit this low dimensionality, as we assume that PCA will allow us to capture the important features of the whole-body hand motion in a small number of principle components.

To support our assumptions, we performed various tests to study the power of PCA for capturing the desired reduced dimensionality of hand motion. In Figure 1, we show that PCA is indeed capable of accurately describing hand motions in a lower dimensional space. This figure shows errors applied to our ASL database, which represents a diverse expression of poses for the hand. We see that PCA shows significant reduction in reconstruction error after around 10 components. While this is larger than reported findings for finger motion, the complex hand gestures of ASL are still well-represented with a relatively small number of components.

Next, to compare the power of PCA for our particular application we experimented with two reconstruction methods with and without PCA. The details of the reconstruction appear in Section 6, however, we include the plot in Figure 2 here to support that PCA is very effective in producing higher quality hand motion. In the figure, we clearly see the benefit of employing PCA as a go-between from markers to joint angles. When we attempt to reconstruct without it (i.e. fitting markers to joint angles directly) the error remains large even as the number of markers employed to inform the handover process is doubled.

5 Sparse Marker Selection

To construct an effective sparse marker set, our method starts from the full set of 13 markers recorded in the reference database, and evaluates each marker's contribution to the whole-hand motion. In constrast to the exhaustive search proposed by Kang et al. [2012], our technique computes the markers directly using PCA.

To this end, we conduct PCA with the *Cartesian positions* of the markers relative to the root link. With 13 markers, this leads to a PCA with 39 dimensions. The results of the PCA is a covariance matrix and the eigenvectors of this matrix which we use to rank-order the markers. Specifically, each eigenvector has 39 coefficients that describe the influence of each marker's Cartesian coordinate on the principle component. By adding up the total contribution of each marker (x, y, z coordinates) to all of the principle components, we produce a convenient way to rank-order the total influence of each marker on the principle components. Further, from the eigenvalues we know the relative importance of each principle component with respect to each other. By weighting the contribution of each marker based on this importance, we can also account for this bias. In our technique, we use the eigenvalue importance, PCA_value, as a weighting to bias each eigenvector coefficient's influence, PCA_coeff, which are taken from the elements of the covariance matrix. We summarize this procedure in Algorithm 1.

Algorithm 1 Ordering markers based on influence.

procedure MARKER_RANK_ORDER(PCA_coeff, PCA_value)
  Vector $marker\_influence$
  for $i =$ each marker **do**
    $x, y, z = 0$
    for $j =$ each component **do**
      $x \mathrel{+}= |PCA\_coeff(3*i+0, j) * PCA\_value(j)|$
      $y \mathrel{+}= |PCA\_coeff(3*i+1, j) * PCA\_value(j)|$
      $z \mathrel{+}= |PCA\_coeff(3*i+2, j) * PCA\_value(j)|$
    end for
    $marker\_influence(i) =$ **sum**(x, y, z)
  end for
  sort($marker\_influence$)
end procedure

In our results, we highlight sparse marker sets of three and six markers, as those form the range of what can be captured and post-processed easily based on our experience. Given the number of markers desired for the sparse set, we select the set simply as the top markers based on the rank-ordering described. We experimented with two methods of producing this rank-ordering, one with the eigenvalues acting as a weighting bias and the second treating all of the top-N principle components as equally important and simply ignoring the remaining components. Conservatively experimenting with N to be between one fourth and three fourths of the full dimensionality, these two approaches produced similar results. However, if we selected N to be the value of the full dimensionality, we see a drop in the quality of the final solutions. In practice, we employ the eigenvalue weighted ranking for all results showcased.

A nice feature of selecting the marker set in this fashion is that the rank-ordering simply adds subsequent markers from smaller sets to produce the larger sets. Thus, the described priority ranking reveals which are the definitively *most* influential markers regardless of the size of the sparse marker set. And so, in practice, adding more markers for higher quality recordings does not require a complete change of markers, only the addition of the desired number of markers to the ones employed in the lower quality recording.

6 Reconstruction

The reconstruction process takes as input a recorded sequence of the sparse marker set. It produces joint angle trajectories that estimate the full hand motion. To this end, we build a regression model to construct joint angle measurements for a full motion sequence. Specifically, our locally weighted regression (LWR) model maps marker positions in the recorded sequence to principle components. Next, the principle components are converted into joint angles using the PCA covariance matrix to produce the final motion.

An LWR model is built for each individual frame, or query, taken from the recorded sequence. In this step, each instance in the database is weighted and this weighting is used to bias the model. The weighting is computed as the inverse of the Euclidean distance from the (root-link corrected) marker positions between the query and the samples in the database. Then, standard regression

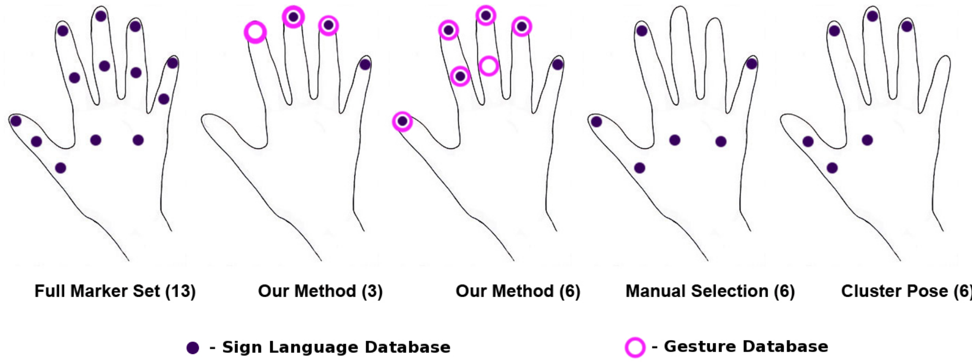

Figure 3: *Marker sets (Left to right).* ***Full Marker Set (13):*** *The full set of thirteen markers used in the recording of the motions in the reference database.* ***Our Method (3)*** *and* ***(6):****The sparse sets of three and six markers selected by our approach. Markers for the sign language database are solid and markers for the gesture database are open circles. Note the considerable amount of overlap between the marker sets for the two databases which indicate that the fingertips are best for reconstructing using our method.* ***Manual Selection(6):*** *A manually selected set of six markers proposed by Hoyet et al [2012] based on perception studies. While intuition may lead us to believe one marker placement is superior to another, this marker set revealed itself to be particularly poor for ASL, clearly because the lack of markers on the middle digits lead to problems when reconstructing sign language poses.* ***Cluster Pose (6):*** *This set of six markers selected by the cluster pose error method reported by Kang et al [2012]. Reported for "freehand" motions, the visible errors from this dataset reveal how sensitive the motion can be to the choice of reference data.*

is performed with each element given its individual weighting as described. The LWR result is a regression model that places importance on the reference samples that are close to the test query, while also down-weighting the influence of reference samples which are distant from the query.

At run-time, we introduce an input sequence recorded from the sparse marker set. The input data is put through the regression modeling step to predict the principle components. To ensure smoothness, the trajectories of the principle components are filtered before they are converted into joint angles. In our results, we use a cone filter with a size of seventeen (with our sample rate for the motion recordings set at 120 hz.) We also experimented with filtering the joint angles to produce smoothness, but found more visually appealing results when we filtered the principle components. Our assumption for this finding is that the principle components combine to produce more "crisp" motion even when they are filtered, while the joint angle filtering dilutes the unique features of individual poses over time. Further study of this phenomena is likely to reveal some interesting findings.

7 Results

With ASL as a primary goal for us, we first describe the use for our technique in producing ASL animations before describing our forays into other motion classes. Our ASL database is comprised of only 52 'letter' sign instances, specifically two continuous runs of the letters of the alphabet signed by the same actor. We test the database on various sequences that include "word" signs (e.g. single signs for words as in "girl" or "walk"). Note, no word signs appear in the reference database.

For our sparse marker set, we choose to use three and six markers as our baseline in order to show both the power of our approach and also to compare our technique to existing solutions. Using the method described in Section 5, we derive the marker sets of three and six as seen in Figure 3. In our analysis of results, we compare this marker set of six to those derived from the manually selected set proposed by Hoyet et al. [2012] and the one found by the cluster pose error method by Kang et al. [2012]. Using the reconstruction method described here, our marker set produces a smaller aver-

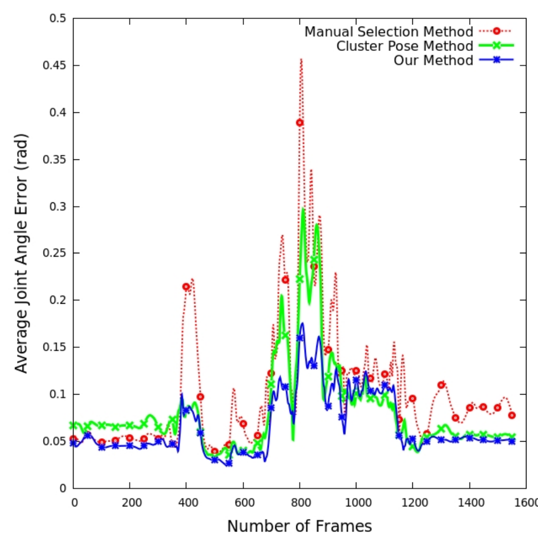

Figure 4: *Comparison of marker-set selection methods.*

age joint angle error per frame for several sign language sequences (shown in the video). Also, Figure 4 shows differences for an examplary sign language clip. Note, the manual selection process of Hoyet et al. relies on an IK-based reconstruction and as such, our reconstruction method is not a fair assessment of the quality of their approach. Instead, their result merely provides an objective alternative marker set from which we can compare the importance of marker selection within the scope of our reconstruction method.

Our reconstruction method uses regression to predict principle components for a sequence of motion. In Figure 5, we compare the estimated components from the regression of our sign language example with the computed components derived from the original joint angle motion. To evaluate the regression's power at estimating the principle components, we use the PCA covariance matrix from the ASL database to convert the joint angles of the test sequence to principle components. We treat this as the "ground truth"

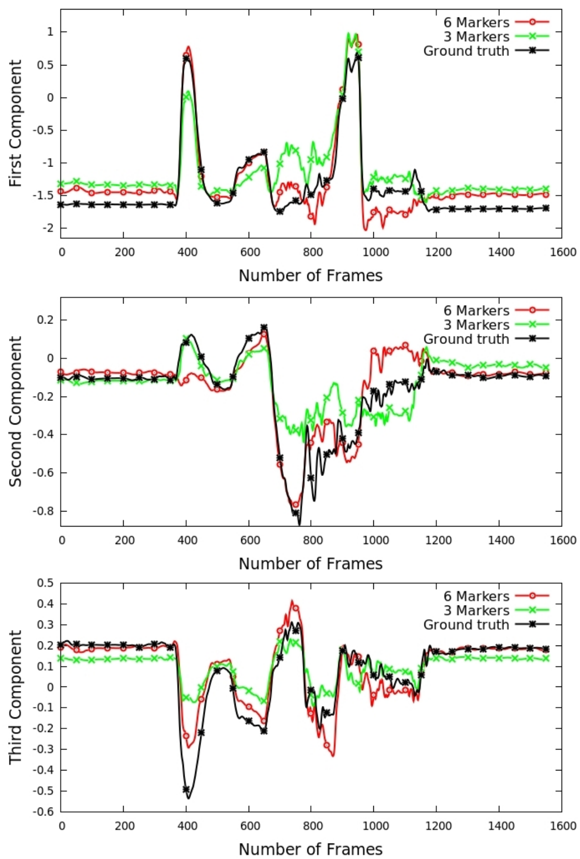

Figure 5: *Comparison of the components of a reconstructed clip using 6 markers and 3 markers. Ground truth is the original clip recorded with 13 markers.*

for the principle components of this motion. Though there are differences, the motion of each component closely follows that of the ground truth. This can also be seen in a reconstructed animation in the accompanying video. Further, three distinct poses are also compared against the various marker selection methods in Figure 6. Our marker selection approach is consistently closer to the original pose.

To test robustness, we attempt to reconstruct motions that are not sign language. The motions we test include counting and general gesticulations. Our sparse marker set of six fairly successfully reconstructs counting the numbers 1 through 5, but the marker set of three fails to reconstruct the number 5. For gesture motion, many of the general poses in the sequence appear to be reached, but the accuracy of the joint angles is not as good as for the sign language motions, as seen in the video. When we test the gesture motion against a more similar "gesture" database, we see drastic improvement in the gesture animations synthesized. We note, the selected sparse marker sets are different than those reported for the ASL database. The marker sets found with the gesture database are shown in Figure 3. Using the gesture database results in high quality gesture reconstructions for both marker sets of three and six.

8 Discussion

Qualitatively, PCA appears to be a good choice for capturing the hidden structure in our hand data input. In contrast, we tested fitting marker positions directly to joint angles and, as seen in Figure 2,

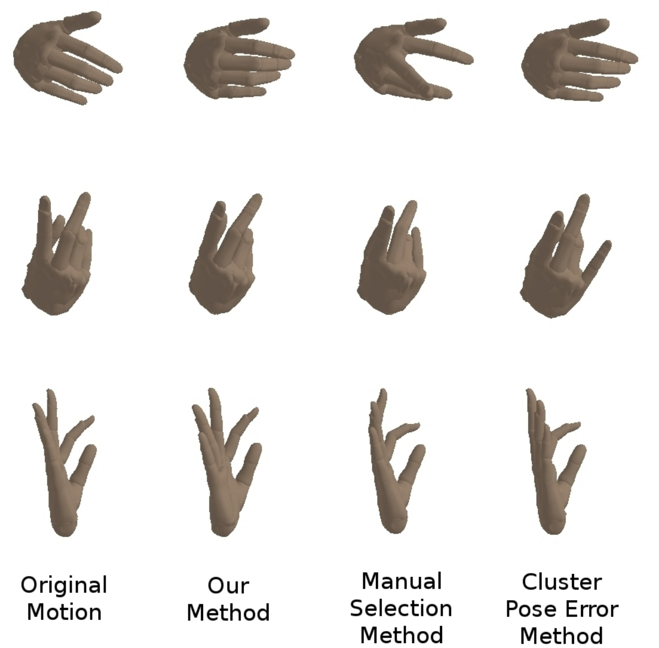

Figure 6: *Three signs not present in the ASL database, reconstructed with the different marker sets, compared to original poses.*

the average joint angle error per frame was notably higher. Also, in the animations produced with these reconstructed joint angles, the hand does not reach the extrema of the poses in the motion. That is, the hand looks much less clean. In ASL, meaning is derived from the end poses, and PCA, while it included error, produces higher quality poses over direct joint angle reconstruction. From this we hypothesize that there is a quantifiable and clear benefit to producing and using principle components to reconstruct the joint angles of the hand, even though the findings reported here are still preliminary.

The specific three markers the system prefers to reflect the motion is a surprising finding, especially for ASL since it does not include index or thumb markers at all. However, we are encouraged to see that the reduced marker set of three performs as well as it does. Although the set of three has a larger average error than the marker set of six, it still produces acceptable results in the majority of cases. We also see when looking at the top principle components of the reconstructed motion, and comparing then to the top principle components of the original motion, that indeed the three marker regression is powerful enough to glean the main trends from the hand motion. Following the findings of previous work [Jörg et al. 2012], we anticipate that we can push even further improvement by exploiting the motion of the full-body, which has been largely overlooked in the current technique. A hurdle that lies ahead is dealing with the non-homogeneity of a database with both hand and body markers. We feel this represents a good direction for future work.

When performing the regression we map marker positions to principle components. In our reported technique, the regression computes the full complement of principle components, regardless of the number of input markers. We experimented with a smaller number of components but found the full set would produce a better reconstruction of the joint angle data. Specifically, we found that mapping to the full 54 components produces the smallest average error, although we can map as low as 35 components with very little degradation from a full component set. We contrast this result to the described technique of [Chai and Hodgins 2005] where they drastically reduce the number of coordinates to simplify (and speed up) the optimization they use to perform reconstruction. While we do not employ such an optimization and thus have the luxury of choosing the full component set, our finding seems to imply that reducing the dimensionality in this step of the process will lead to degraded motion quality.

Lastly, as reported, when we reconstruct motions that are different from the original database, we get mixed results. For example, the motions for counting were close enough to ASL assumably, because we find reasonably acceptable results from counting synthesis using the ASL databse. However, it is not completely clear why the seemingly simpler gesture motion was not equally easily reconstructed by the same database. While the general poses in the gesture sequences appear to be reached, the motion itself was not of very high quality. From this finding, questions arise regarding the intricacies of overlap in motion styles, between basic and more complex, between trained and more "natural" and so on. Similarly, investigation of the effect of different subjects on the final data, as is the case here, also remains for future work.

9 Conclusion

In this work, we present a method to capture subtle hand motions with a sparse marker set consisting of three to six markers. Our method first specifies an appropriate set of markers using PCA to exploit the redundancies and irrelevancies present in hand motion data. It then reconstructs the full hand motion based on the sparse marker set found and a LWR mapping from marker positions to PCA's components, via a reference motion database.

We show that our technique can reconstruct complex finger motions based on only three markers per hand and outperforms recent similar methods, such as those presented by Hoyet et. al [2012] and Kang et. al [2012], based on the marker sets they report. Our findings also clearly indicate that using a regression model for mapping marker positions to principle components leads to better results for reconstruction of the full hand motion than using regression for mapping marker positions directly to joint angles, indicating that PCA is notably effective at exploiting the redundant dimensionality of the hand.

Acknowledgements

We kindly acknowledge the help and input from the School for the Deaf, Riverside. Also, the primary author was the recipient of a GEM fellowship which helped to make this research possible.

References

ATKESON, C., MOORE, A., AND SCHAAL, S. 1997. Locally weighted learning. *Artificial Intelligence Review 11*, 1–5, 11–73.

BISHOP, C. M. 1995. *Neural Networks for Pattern Recognition*. Oxford University Press, Inc.

BRAIDO, P., AND ZHANG, X. 2004. Quantitative analysis of finger motion coordination in hand manipulative and gestic acts. *Human Movement Science 22*, 6, 661–678.

CHAI, J., AND HODGINS, J. K. 2005. Performance animation from low-dimensional control signals. *ACM Transactions on Graphics (SIGGRAPH) 24*, 686–696.

CHANG, L. Y., POLLARD, N., MITCHELL, T., AND XING, E. P. 2007. Feature selection for grasp recognition from optical markers. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2944–2950.

COURTY, N., AND GIBET, S. 2010. Why is the creation of a virtual signer challenging computer animation? In *Motion in Games*, Springer-Verlag, MIG'10, 290–300.

CYBERGLOVE, 2013. http://www.cyberglovesystems.com.

HOYET, L., RYALL, K., MCDONNELL, R., AND O'SULLIVAN, C. 2012. Sleight of hand: perception of finger motion from reduced marker sets. In *ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, I3D '12, 79–86.

JÖRG, S., AND O'SULLIVAN, C. 2009. Exploring the dimensionality of finger motion. In *Proceedings of the 9th Eurographics Ireland Workshop (EGIE 2009)*, 1–11.

JÖRG, S., HODGINS, J. K., AND SAFONOVA, A. 2012. Data-driven finger motion synthesis for gesturing characters. *ACM Transactions on Graphics (SIGGRAPH Asia) 31*, 6, 189:1–189:7.

KAHLESZ, F., ZACHMANN, G., AND KLEIN, R. 2004. 'Visual-fidelity' dataglove calibration. In *Computer Graphics International*, IEEE, CGI '04, 403–410.

KANG, C., WHEATLAND, N., NEFF, M., AND ZORDAN, V. 2012. Automatic hand-over animation for free-hand motions from low resolution input. In *Motion in Games*, MIG '12. Springer Berlin Heidelberg, 244–253.

MAJKOWSKA, A., ZORDAN, V. B., AND FALOUTSOS, P. 2006. Automatic splicing for hand and body animations. In *ACM SIGGRAPH / Eurographics Symposium on Computer Animation*, SCA '06, 309–316.

RIJPKEMA, H., AND GIRARD, M. 1991. Computer animation of knowledge-based human grasping. In *SIGGRAPH '91*, ACM, 339–348.

SAFONOVA, A., HODGINS, J. K., AND POLLARD, N. S. 2004. Synthesizing physically realistic human motion in low-dimensional, behavior-specific spaces. In *ACM Transactions on Graphics (SIGGRAPH)*, 514–521.

SANTELLO, M., FLANDERS, M., AND SOECHTING, J. F. 1998. Postural hand synergies for tool use. *The Journal of Neuroscience 18*, 23, 10105–10115.

WANG, Y., AND NEFF, M. 2013. Data-driven glove calibration for hand motion capture. In *ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '13, 15–24.

WANG, R. Y., AND POPOVIĆ, J. 2009. Real-time hand-tracking with a color glove. *ACM Transactions on Graphics (SIGGRAPH) 28*, 3, 1–8.

WANG, Y., MIN, J., ZHANG, J., LIU, Y., XU, F., DAI, Q., AND CHAI, J. 2013. Video-based hand manipulation capture through composite motion control. *ACM Transactions on Graphics (SIGGRAPH) 32*, 4, 43:1–43:14.

YE, Y., AND LIU, C. K. 2012. Synthesis of detailed hand manipulations using contact sampling. *ACM Transactions on Graphics (SIGGRAPH) 31*, 4, 41:1–41:10.

ZHAO, W., CHAI, J., AND XU, Y.-Q. 2012. Combining marker-based mocap and RGB-D camera for acquiring high-fidelity hand motion data. In *ACM SIGGRAPH / Eurographics Symposium on Computer Animation*, SCA '12, 33–42.