

Response to Reviewers' Comments

#AURO0645: Human Motion Database with a Binary Tree and Node Transition Graphs

Katsu Yamane, Yoshifumi Yamaguchi and Yoshihiko Nakamura

We would like to thank the editors and reviewers for their thorough and constructive comments, which were very helpful in revising the paper. Below is the list of changes we have made in response to the reviewers' comments.

REVIEWER #1

- 1) *Comment:* The scientific contribution of the paper is high. The experiments are nicely designed and described. However, additional experiments using the second database would increase the quality of the paper.
Response: We have added another example where the goal heading is turned 1 rad to the left, with the same position goal (1.0 m, 2.0 m). We have also added a brief description of the planning algorithm.
- 2) *Comment:* Minor errors:
 - Page 1: data aset → dataset
 - Page 3: difficult compute → difficult to compute
 - Page 5: Eq.(7) → Eq. (7)
 - Page 6: Explain IK before using it
 - Reference [14] is now published (remove “in press”)

Response: The errors are corrected.

REVIEWER #2

- 1) *Comment:* The section on virtual marker sets in the introduction of the paper should be dropped. Lucas Kovar has already used an equivalent idea to help cluster similar motion clips and identify possible transitions.
Response: We agree that Kovar's representation is equivalent to ours, and therefore the paragraphs that describe the rationale behind the choice have been removed. We have, however, kept the part of the section that gives a formal definition of the variables used in the rest of the paper.
- 2) *Comment:* The related work section should probably cite the work by Hertzmann and Brand on Style Machines in the section on HMM's.
Response: Thank you for pointing us to the paper. The new reference is added to the related work section.
- 3) *Comment:* In general, I thought the Related Work section of this submission does a very poor job of placing this work in the context of existing approaches. For example, the last paragraph cites a work in vision that sounds very similar to the proposed method. How is this work different?
Response: We have added a couple of sentences that briefly states the difference between our work and the cited papers including Sidenbladh et al., which is similar to our work but does have significant differences in the clustering algorithm and applications.
- 4) *Comment:* The proposed approach uses PCA to reduce the input data to a small dimensionality. I wasn't really sure in the end how many dimensions were used (3???) but there is no discussion about how much variance is contained in those first k dimensions.
Response: Please note that we only use the first principal component to cluster the data, not to reduce the dimensionality, although it is possible to use the results of PCA for dimensionality reduction. The low-dimensional spaces shown in the results section are only for visualization purpose. This point has been clarified in the third paragraph of Section III-B.
- 5) *Comment:* The last paragraph of section III-B comes out of nowhere and sounds like a complete hack. At the very least, you should motivate the need for the hack. My guess is that you wish to keep the tree balanced?
Response: As the reviewer pointed out, this is indeed a simple “hack” required to make all branches in the tree have the same depth so that all frames appear in every layer. This motivation has been clarified in the last paragraph of Section III-B.

- 6) *Comment:* The hierarchical dynamic programming algorithm could be described more clearly. Also, there should be some explanation of why the proposed method is better than standard approximations to DP.
Response: A more detailed description of Fig. 4 has been added to the last paragraph of Section IV-A.
- 7) *Comment:* Equation 9 is backwards.
Response: Corrected.
- 8) *Comment:* Many of the figures in this submission are terrible. The text in the results section does a decent job of describing the results, but the figures are incomprehensible. For example, the probability graphs are illegible. What are the axes? What are the different time series? These are not clearly labeled. Why not combine figure 10 and 7 and make the character a little more visible? I didn't get much from the other visualizations of the data set. Figure 13 makes a reference to a blue node, but it is not clear which of the hundred blue nodes you are talking about.
Response: We apologize for the bad quality of the figures. The graphs and Fig. 7 (with old Fig. 10 embedded) have been redrawn with larger fonts and thicker lines. The *blue* nodes in the figures currently numbered 11 and 12 actually mean the *dark blue* nodes. This error has been corrected as well.
- 9) *Comment:* It's pretty hard to evaluate the quality of the motion synthesis component of this approach without seeing an animation.
Response: Please note that these examples are shown only to demonstrate the possibility of the application to motion planning, and making realistic animations would require a lot of future work such as smoothing and contact resolution. Nevertheless, we will submit a supplemental movie corresponding to Fig.16 to give the idea of quality of the animations directly obtained from our database.