Workshop on Virtual Reality Interaction and Physical Simulation VRIPHYS (2011)
J. Bender, K. Erleben, and E. Galin (Editors)

A Study On Perceptual Similarity of Human Motions

Björn Krüger[1], Jan Baumann[2], Mohammad Abdallah[1], and Andreas Weber[1]

[1]Bonn University, Institute of Computer Science II
[2]Fraunhofer FKIE, Wachtberg
email: {kruegerb,baumannj,abdallah,weber}@cs.uni-bonn.de

Abstract

We perform a user study involving different classes of pre-recorded human motions displayed in abstract form either as stick figures or as point lights. Collecting data on more than 1000 user votes of various triples of short motion sequences asking whether a motion A is perceived to be "more similar" to a reference motion O than B or vice versa, we test for associations with numeric distance measures for human motions described in the literature. Our preliminary hypothesis that perceived similarities using stick figure representations are more highly associated to "joint angle based distance measures" than to "point cloud based distance measures" has to be rejected on grounds of the experimental data. We find that there are higher associations for "point cloud based distance measures" than for "joint angle based distance measures" both for the perceived similarities for point light representations as well as for stick figure representations.

Categories and Subject Descriptors (according to ACM CCS): Computer Graphics [I.3.7]: Three-Dimensional Graphics and Realism—[Animation] Computer Graphics [I.3.7]: Three-Dimensional Graphics and Realism—[Perception]

Keywords: motion data, perception, distance measure

1. Introduction

For data driven approaches of motion synthesis and content-based retrieval of motion data the task of searching for "similar motion segments" is of central importance. The notion of similarity has to be defined further, and in the last decade a wide variety of distance measures for poses and motions with different dimensionality and properties have been proposed in the literature: There are purely pose-based distance measures such as the one measuring distances on joint angles [CH05]. As the distance measure depends on the encoding of the joint angles, e.g. whether quaternion-based representations or Euler angle-based representations are used, different variants of these are feasible [KTWZ10]. PCA-based compression of pose-based feature sets [SHP04, CH05, BCvdPP08] is one way to reduce the dimensionality of the distance measures, using a selection of specific joints another one [KTWZ10]. In order to describe not only the properties of a pose statically but also to encode the kinematic properties of a motion sequence in the feature set of a frame, Kovar and Gleicher [KG04] introduced a point cloud distance measure on a normalized window of the previous and subsequent poses.

Whereas user studies have been performed to relate effects of using different distance metrics to naturalness of human motions, the quality of motion blending and other perceptual aspects [vBE09, RPE*05, RP03], the underlying problem still seems to be unsolved, as has been stated in [TWC*09, Sect. 2.2]

> As a matter of fact, finding an accurate and robust metric for human motion perception remains, to the best of our knowledge, an open problem.

In this paper we describe a user study in which the question of similarity of human full body motions is the topic of direct perceptual investigation. In order to rate the concept of "similarity" between motions we use the simplest possible setting: For a given triple of motions O, A, and B we ask whether a motion A is "more similar" to the reference motion O than B or vice versa. This question can be posed for different perceptual representations of motions, but also for all of the algorithmic distance measures described in the literature.

Our initial hypothesis was that there is a dependency on the used pose-representation for the perception of similarity of motions. More specifically, we conjectured that the com-

DOI: 10.2312/PE/vriphys/vriphys11/065-072

monly used joint angle based distance measure have a higher correlation to perceived similarity when using stick representations for the motions, as in these the joint angles are visually well exposed—whereas conversely we conjectured that when using point light representations a higher correlation to the point cloud representations will occur. As can be seen from the results of the user studies presented above, the initial hypothesis has to be rejected on grounds of the empiric data.

Due to these initial hypothesis we used point light representation and stick representation of the motions instead of "full flesh" geometric representations of the avatar motions.

2. Related Work

The influence of the visualization of motion data with respect to perceived properties of the motions is subject of several papers: Hodgins et al. [HOT98] show that viewers perception of motions as being different is affected by the geometric model used for rendering. Their experiments indicate that users were better able to observe changes viewing a polygonal model than a stick figure. McDonnell et al. [MJM*08, MJM*09] investigate how the body shapes of the rendered models influence the perception of emotions. The authors found, that the perception of emotions is robust and at most independent of the characters body. The result that facial anomalies are salient even in the case of body anomalies is concluded by Hodgins et al. [HJO*10]. The authors performed a study on the influence of differences of facial and body anomalies with respect to the emotional response of a viewer.

Reitsma and Pollard [RP03] come up with user studies on naturalness in ballistic human motion, and they develop a metric for measuring errors in such motion data. An examination of the user's sensitivity to errors in physically rigid body simulations is done by Reitsma and O'Sullivan [RO09].

In the field of creating realistic transitions and blending between motion sequences van Basten and Egges [vBE09] evaluated three metrics quantitatively and qualitatively, and Wang and Bodenheimer [WB03, WB08] compute optimal weights for a transition cost function. They performed a user study that demonstrates that results are more appealing using this weighted cost function.

To quantify the naturalness of a motion sequence Ren et al. [RPE*05] develop a measure for naturalness of human motions that is compared and evaluated by a user study. McDonnell et al. [MNO07] performed a study on the parameter-dependency of smoothness perceptions of motion sequences.

The influence of time warping of motion sequences on the perception of users is investigated by Pražák et al. [PMO10].

A special distance function for classification of motion capture sequences is developed by Onuma et al. [OFH08].

Figure 1: *Photo of the environment used for the experiments.*

Machine learning techniques were used by Tang et al. [TLKS08] to propose a similarity measure for motion sequences. In this work pairs of motions sequences were shown to the participants of a study which had to decide on similarity.

Nevertheless the perceived similarity of several motion sequences has obtained little attention: The only work we are aware of explicitly investigating perceived similarity is the study of Pražák et al. [PMKO09]. The participants were asked to select two out of four randomly chosen walking sequences which they felt to be most similar. Based on the data of this study a metric that combines joint angles, joint positions and joint velocities is developed. However in [PMKO09] no comparison of the new metric to established distance measures is given.

3. Materials and Methods

3.1. Perceptual Studies

Our perceptual studies involved thirty-nine participants (34 male, 5 female) ranging in age from 21 to 34. Each of the participant separately took part in a sequence of 20 trials of a perceptual experiment. In a single trial of an experiment three short sequences of human motions were simultaneously displayed on a 23-inch flat screen LCD monitor.

The displayed human motions were rendered using a fixed virtual camera perspective from 3D-motion data in ASF/AMC file format [MRC*07]. For a specification of the motions see Sect. 3.2. The rendering of the motion was done in one of the following two ways for all three motion clips for a single round:

1. Stick figure: each bone of the skeleton was rendered as a stick of red color, see Fig. 3 and Fig. 2.
2. Point lights [Joh73]: Only the joints of the skeleton were rendered as red spheres, see Fig. 2.

Even though point lights are a simplistic representation it is possible to recognize a walking subject [CK77] and even

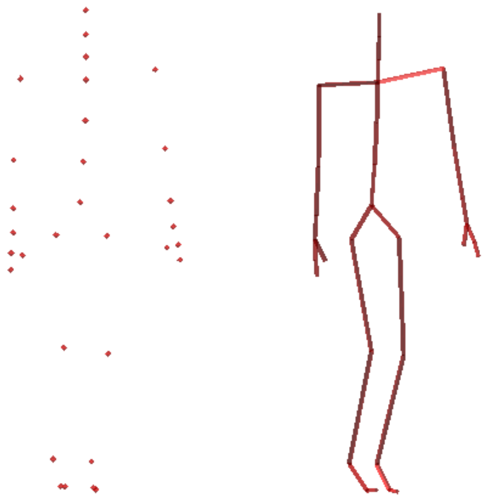

Figure 2: *Comparison of point light (left) and stick figure (right) rendering of a pose.*

a parametric model [Tro02] has been build on experiments based on point lights. Point lights can be the representation of choice if perceptual aspects of motion should be explored independent of other visual information [Joh73].

Each trial used a motion triple out of a collection of 48 rendered motion triples specified in Table 2. (The 24 motion triples given in the table were each rendered as point lights and stick figures.) Each participant was asked to perform 20 trials. For each trial a random selection of a triple of the collection of 48 motion triples were chosen also using randomization in the order of *A* and *B* to avoid ordering effects.

The participants were asked to answer two questions in each trial.

3.1.1. Perceptual similarity of motions

The main question was whether a participant perceived a motion *A* to be more similar to the reference motion *O* than the motion *B* or vice versa. The meaning of being "more similar" was not specified any further but left to a naíve interpretation to the participants. In addition to either labeling motion *A* or *B* as being more similar also the possibility that no decision can be done by the participant for the displayed triple was possible.

In Fig. 3 a screenshot of the interface that we used for this study is shown. The upper part of the interface shows the current $\mathbf{T}(\mathrm{A},\mathrm{O},\mathrm{B})$ of motions as video sequence. The buttons in the part of the screen below the videos can be used for control. Their functions are *play*, *stop* and *goto next triple*. The lower part of the interface is used for the input of the judgments the participants made. The users have to decide if motion *A* (left) or *B* (right) is more similar to motion *O* (middle). The button *unsure* was added for the reason that users might be unsure and to prevent them to choose left or right randomly.

We made no limitation for the participants how often they can view each triple before making their decisions. We stored the number of repetitions and the time they needed for the judgments.

Although repeated viewing has been allowed, nevertheless in 5 % of the trials an "uncertain" was marked.

An additional minor question was whether a participant believed that one question was synthesized from motion capture data. The evaluation of this question is given in the appendix, as in the context of this paper it only serves for a verification that the synthesized ones could not be distinguished from the natural ones, so that all results with respect to the perceived similarity can be used from trials involving a synthesized motion, too.

Each of the 39 participants performed 20 trials of the experiment subsequently in one day. Moreover, 28 participants repeated the experiment seven to ten days later. Since the triples of motion sequences in the experiments were chosen randomly, these experiments are counted as independent experiments, so that the data of experiments on $(39+28)\cdot 20 = 1340$ triples were collected.

3.2. Selection of Motion Sequences and Motion Classes

We wanted to perform tests on different motion classes, but each test should involve motions from one class only, so that each class should contain a sufficient number of different motions. So we decided to use motions from the HDM05 database [MRC*07]. HDM05 contains more than three hours of systematically recorded and well-documented motion capture data in C3D as well as in ASF/AMC data format. Furthermore, HDM05 contains 10 to 50 realizations for each of roughly 70 motion classes performed by various actors. In addition to recorded natural motions from the HDM05 database we also included synthesized motions in our tests. The synthesized motions were obtained by morphing of motions from the HDM05 database.

From the about 70 motion classes available in the HDM05 database we wanted to use a collection which should include simple locomotion as well as more dynamic motions sequences. Specifically we choose the following motion classes for the experiment. We use the naming convention of the HDM05 documentation [MRC*07].

- sneak2StepsRStart: sneaking two steps starting with the right foot.
- walk2StepsRStart: walk two steps forward.
- walkLeft2Steps: walk two steps to the left side.
- walkRightCircle4StepsRStart: walk four steps on circle into right direction.

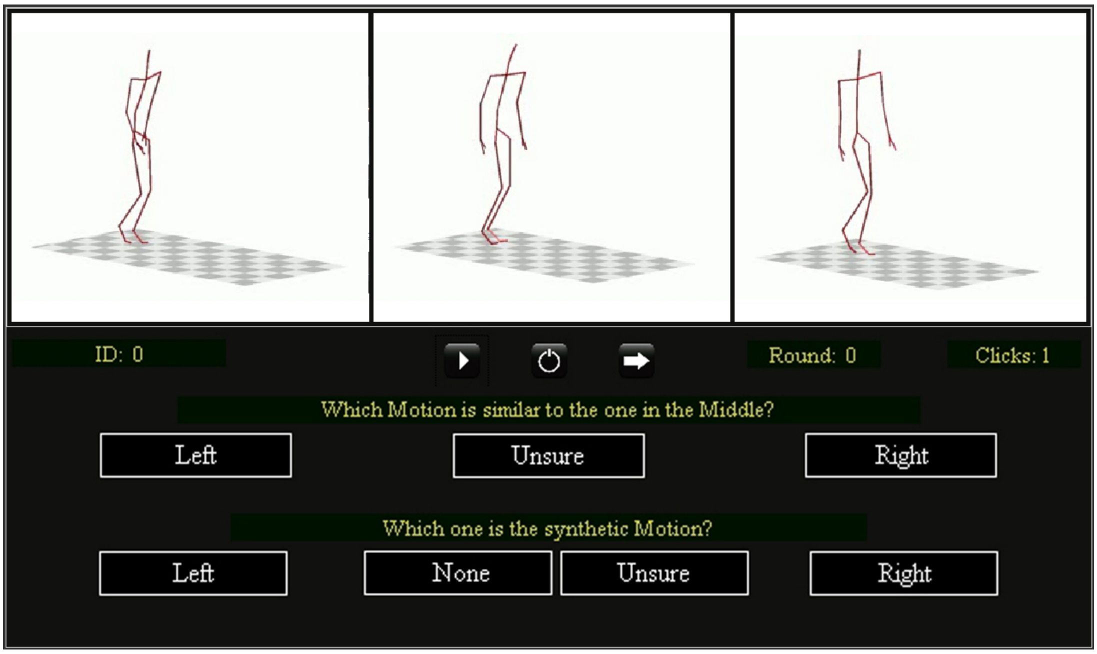

Figure 3: *The interface used for the experiments.*

- jogLeftCircle4StepsRStart: jog two steps on a circle into left direction.
- hopBothLegs1Hops: jumping with both legs simultaneously.

As a primary goal of our study was to test possible associations of various numerical distance measures we choose 4 motion triples of any of these classes for which the numerical distance measures (specified in Sect. 3.3) gave the widest variety to the questions whether a motion A is more similar to the reference motion O or a motion B. Moreover, 2 of the 4 triples for each class should consist of natural motions only, whereas in the 2 other triples two motions—the reference motion and one of motion A or B—should be natural motions whereas the other one should be synthesized.

A table with the exact specification of the used motion triples is given in Table 2.

3.3. Distance Measures

Using the notation of [KTWZ10] for various local distance measures on motions we use the following set of feature sets and induced distance measures in our comparisons.

$\mathcal{D}_{\mathbf{euler}}$ Encoding of the joint angles on Euler angle-based representation.

$\mathcal{D}_{\mathbf{quat}}$ Encoding of the joint angles on quaternion-based representation.

$\mathcal{D}_{\mathrm{E}}^{\mathbf{15}}$ Consists of the positions of hands, feet and head.

$\mathcal{D}_{\mathrm{E}}^{\mathbf{30}}$ All features of $\mathcal{D}_{\mathrm{E}}^{15}$; as well as the 5 positions of the elbows, knees and one chest joint.

$\mathcal{D}_{\mathrm{E}}^{\mathbf{39}}$ All features of $\mathcal{D}_{\mathrm{E}}^{30}$; in addition position of the shoulders and one lower-back joint.

$\mathcal{D}^{n\times l}$ Distance measures including several frames on a small window to represent the local evolution in time. The windows are sampled sparsely, using only 3 or 5 frames per window. The resulting distance measures will be denoted by $\mathcal{D}_{\mathrm{E}}^{15\times 3}$, $\mathcal{D}_{\mathrm{E}}^{15\times 5}$ and $\mathcal{D}_{\mathrm{E}}^{30\times 3}$.

$\mathcal{D}_{\mathbf{pca}}^{\mathbf{p}n}$ PCA-based distance measures [SHP04, EMMT04, CH05]. Here, n means the number of principal components on joint positions in body frame—pre-computed on a fixed database, which will be chosen to be the entire HDM05 database in all our experiments (n dimensions).

$\mathcal{D}_{\mathbf{pc}}^{n}$ Point cloud distance measure on a normalized window of the previous and subsequent $n/2$ poses— introduced by Kovar and Gleicher [KG04].

We lift these local distance measures to distance measures on motions by using the accumulated distances on the minimal cost time warping path, cf. [KTWZ10]. These lifted distance measures for two motions are denoted by the same symbols as their underlying local distance measures.

3.4. Statistical Tests

All of our statistical tests are computed using Matlab and the Statistics Toolbox [Mat11]. For the standard statistical tests and the non-parametric tests we refer to one of the many references on the subject, e.g. [KV07].

4. Results

4.1. Results of the perceptual studies

All experiments consisted of 20 trials involving one motion triple each and which had to be judged by the participants. The average measured time for participation was 20 minutes. So the average decision time for a one trial was one minute. For making the required judgments the motion triples were viewed 12 times in average in one trial.

For a detailed summary of the answers we refer to Table 1.

4.2. Validity of tests

In 5 % of the trials the subjects used "uncertain" as answer. When giving these answers equal probabilities for voting A or B (as the worst-case consideration in a forced-choice test) and using these results together with the outcome of the true votes, the hypothesis that the answers A and B were obtained randomly with equal chance can be rejected for all motion classes—except the sneaking motions—by Bernoulli tests on the 1 % level for the tests involving stick figure representations as well as point light representations. Thus our tests allowing "uncertain votes" and repeated viewings are valid.

4.3. Relating the results of perceived similarity for different representations

We correlate the results of the votes for A and B for the tests involving the stick figure representations with the ones using point light representations for each motion triple by using the value $A-B$ as a signed magnitude disregarding the uncertain votes U.

When combining all results we obtain a rank correlation coefficient $\tau = 0.58$ between the results of tests using the stick figure representations vs. the point light representations. The the hypothesis of independence can be rejected on the 1 % level.

4.4. Relating results of perceptual study to similarity measures

We examined the associations of the perceived similarity according to the perception tests to the numerical distance measures defined above. For robustness we use Kendall's rank correlation between the signed magnitudes $A-B$ for any motion triple and the differences of the distances

$$\mathcal{D}(A,O)-\mathcal{D}(B,O)$$

for any of the distance measures defined in Sect. 3.3. A visualization of the values of Kendall's rank correlation coefficient τ for the tests involving stick figure as well as point light renderings of the motions is given in Fig. 4.

Using the associated test statistics it can be concluded that the hypothesis that there is no association between the numerical distance measure and the perceived similarity using stick figure representation can be rejected on the 5 %-level for all distance measures defined in Sect. 3.3 *except* $\mathcal{D}_{\text{euler}}$ and $\mathcal{D}_{\text{quat}}$.

Using point lights the corresponding result on the 5 %-level can be obtained for all of these distance measures *except* $\mathcal{D}_{\text{euler}}$, $\mathcal{D}_{\text{quat}}$, $\mathcal{D}_{\text{E}}^{15}$, and $\mathcal{D}_{\text{pca}}^{\text{p8}}$.

5. Discussion

Our initial hypothesis that there is a dependency on the used pose-representation for the perception of similarity of motions has been falsified—at least for the two very different representations that we used (point lights and stick figures): In our experiments we found a very high correlation between the outcome of the tests when using stick figure representations and point light representations, a result that was not expected by the authors of the paper. Extending the user study to include 3D-geometric representations of the avatars will be a topic of future work—as we are now rather uncertain whether such representations will yield different results (as has been the case for other perceptional studies) or they will yield similar perceptual results to the two representations used in this study.

In this paper the presented study was restricted to six motion classes, where mostly locomotions were considered (except the hopBothLegs1Hops class). The extension of this experiments to additional, more dynamic, motion classes will be a strand of future research.

The higher dimensional feature sets such as $\mathcal{D}_{\text{E}}^{39}$ yielded slightly higher rank correlation coefficients than the lower dimensional ones—both the ones obtained by principal component analysis and the specifically designed $\mathcal{D}_{\text{E}}^{15}$, which was concluded by Krüger et al. [KTWZ10] to be the one of choice—especially for real-time applications. However, all of these feature sets based on point clouds yielded significantly higher rank correlations with the outcomes of the user studies than the ones based on joint angles.

So the outcome of our user study indicates that the use of distance measures for motions based on joint angles is also less preferable from a perceptional point of view than using point cloud based distance measures.

References

[BCvdPP08] Beaudoin P., Coros S., van de Panne M., Poulin P.: Motion-motif graphs. In *ACM SIGGRAPH/Eurographics Symposium on Computer Animation* (2008), Gross M., James D., (Eds.). 1

[CH05] Chai J., Hodgins J. K.: Performance animation from low-dimensional control signals. *ACM Trans. Graph. 24* (July 2005), 686–696. 1, 4

[CK77] Cutting J. E., Kozlowski L. T.: Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society 9*, 5 (1977), 353–356. 2

Table 1: *Results of the tests: For the question on "similarity" votes for A means that the subjects perceived the motion A to be more similar to the reference motion O than they perceived motion B; votes for B means that they perceived the motion B to be more similar to the reference motion O than A. Votes for U mean that for the displayed triple of motion the subjects felt that no such decision could be taken. For the question on "synthetic motion" votes for A mean that the motion A was believed to be a synthesized motion, votes for B mean that the motion B was believed to be a synthesized motion, votes for N mean that none of A and B was believed to be a synthesized motion, and votes for U mean that the subjects were not sure whether one of the motions A and B were synthesized or not.*

stick figures								
		votes "similarity"			votes "synthetic motion"			
motion triple	**#votes**	**A**	**B**	**U**	**A**	**B**	**N**	**U**
T hop 01	28	9	18	1	9	7	11	1
T hop 02	45	18	22	5	12	11	20	2
T hop 03 syn	21	0	21	0	14	4	2	1
T hop 04 syn	20	1	19	0	10	6	4	0
T jog 01	28	6	22	0	9	4	15	0
T jog 02	41	11	28	2	14	10	17	0
T jog 03 syn	20	4	16	0	7	5	8	0
T jog 04 syn	20	7	13	0	8	6	6	0
T sneak 01	44	23	17	4	5	16	20	3
T sneak 02	44	20	21	3	9	12	20	3
T sneak 03 syn	22	12	15	5	9	8	12	2
T sneak 04 syn	21	11	9	1	4	7	9	1
T walk 01	46	15	24	7	10	17	18	1
T walk 02	45	20	21	4	9	11	23	2
T walk 03 syn	20	1	19	0	9	3	8	0
T walk 04 syn	20	8	12	0	6	7	7	0
T walkLeft 01	28	10	18	0	13	6	9	0
T walkLeft 02	28	10	17	1	11	7	9	1
T walkLeft 03 syn	21	17	3	1	4	9	7	1
T walkLeft 04 syn	20	0	20	0	12	4	4	0
T walkRight 01	44	28	13	3	13	21	9	1
T walkRight 02	28	11	17	0	9	9	10	0
T walkRight 03 syn	20	11	9	0	7	8	5	0
T walkRight 04 syn	22	10	12	0	8	8	4	2

point lights								
		votes "similarity"			votes "synthetic motion"			
motion triple	**#votes**	**A**	**B**	**U**	**A**	**B**	**N**	**U**
T hop 01	28	6	21	1	10	5	12	1
T hop 02	42	10	29	3	10	7	23	2
T hop 03 syn	21	2	18	1	15	3	3	0
T hop 04 syn	20	2	18	0	9	3	8	0
T jog 01	40	6	32	2	15	9	16	0
T jog 02	28	3	24	1	10	5	12	1
T jog 03 syn	20	3	17	0	10	3	7	0
T jog 04 syn	21	5	15	1	5	4	12	0
T sneak 01	46	26	16	4	11	16	16	3
T sneak 02	21	11	9	1	5	11	5	0
T sneak 03 syn	25	11	13	1	5	12	7	1
T sneak 04 syn	21	8	12	1	4	6	11	0
T walk 01	28	11	16	1	7	7	13	1
T walk 02	27	14	9	4	5	8	14	0
T walk 03 syn	20	3	17	0	9	2	9	0
T walk 04 syn	20	13	7	0	4	9	7	0
T walkLeft 01	34	7	25	2	14	5	15	0
T walkLeft 02	24	1	22	1	7	4	12	1
T walkLeft 03 syn	21	16	4	1	8	5	8	0
T walkLeft 04 syn	21	3	17	1	9	2	10	0
T walkRight 01	41	26	13	2	11	14	15	1
T walkRight 02	22	9	11	2	6	5	10	1
T walkRight 03 syn	22	9	11	2	4	4	13	1
T walkRight 04 syn	22	8	13	1	4	4	14	0

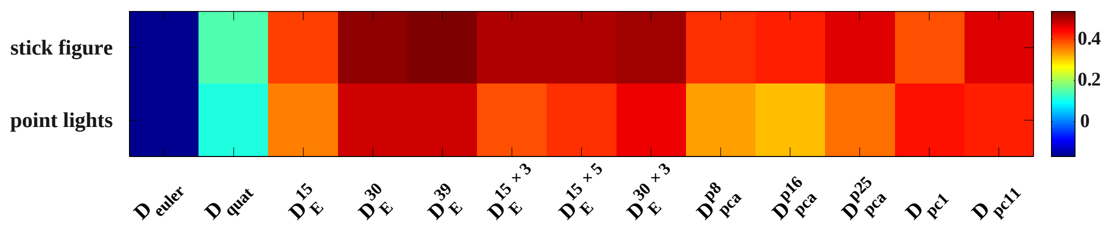

Figure 4: *Visualisation of Kendall's rank correlation coefficient τ between various distance measures and the results of the perceptual study for all motion triples.*

[EMMT04] EGGES A., MOLET T., MAGNENAT-THALMANN N.: Personalised real-time idle motion synthesis. In *12th Pacific Conference on Computer Graphics and Applications (PG 2004)* (2004), pp. 121–130. 4

[HJO*10] HODGINS J., JÖRG S., O'SULLIVAN C., PARK S. I., MAHLER M.: The saliency of anomalies in animated human characters. *ACM Trans. Appl. Percept. 7* (July 2010), 22:1–22:14. 2

[HOT98] HODGINS J., O'BRIEN J., TUMBLIN J.: Perception of human motion with different geometric models. *IEEE Transactions on Visualization and Computer Graphics 4*, 4 (1998), 307–316. 2

[Joh73] JOHANSSON G.: Visual perception of biological motion and a model for its analysis. *Perception And Psychophysics 14*, 2 (1973), 201–211. 2, 3

[KG04] KOVAR L., GLEICHER M.: Automated extraction and parameterization of motions in large data sets. *ACM Transactions on Graphics 23*, 3 (2004), 559–568. SIGGRAPH 2004. 1, 4

[KTWZ10] KRÜGER B., TAUTGES J., WEBER A., ZINKE A.: Fast local and global similarity searches in large motion capture databases. In *Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation* (Aire-la-Ville, Switzerland, Switzerland, July 2010), SCA '10, Eurographics Association, pp. 1–10. 1, 4, 5

[KV07] KVAM P., VIDAKOVIC B.: *Nonparametric statistics with applications to science and engineering*. Wiley series in probability and statistics. Wiley-Interscience, 2007. 4

[Mat11] MATHWORKS: Statistics toolbox. http://www.mathworks.com/help/toolbox/stats/, 2011. 4

[MJM*08] McDonnell R., Jörg S., McHugh J., Newell F., O'Sullivan C.: Evaluating the emotional content of human motions on real and virtual characters. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization* (New York, NY, USA, 2008), APGV '08, ACM, pp. 67–74. 2

[MJM*09] McDonnell R., Jörg S., McHugh J., Newell F. N., O'Sullivan C.: Investigating the role of body shape on the perception of emotion. *ACM Trans. Appl. Percept. 6* (September 2009), 14:1–14:11. 2

[MNO07] McDonnell R., Newell F., O'Sullivan C.: Smooth movers: perceptually guided human motion simulation. In *Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation* (Aire-la-Ville, Switzerland, Switzerland, 2007), SCA '07, Eurographics Association, pp. 259–269. 2

[MRC*07] Müller M., Röder T., Clausen M., Eberhardt B., Krüger B., Weber A.: *Documentation: Mocap Database HDM05*. Computer Graphics Technical Report CG-2007-2, Universität Bonn, May 2007. www.mpi-inf.mpg.de/resources/HDM05. 2, 3, 8

[OFH08] Onuma K., Faloutsos C., Hodgins J. K.: FMDistance: A fast and effective distance function for motion capture data. In *Short Papers Proceedings of EUROGRAPHICS* (2008). 2

[PMKO09] Pražák M., McDonnell R., Kavan L., O'Sullivan C.: A perception based metric for comparing human locomotion. In *Eurographics Ireland 2009: 9th Irish Workshop on Computer Graphics* (2009). 2

[PMO10] Pražák M., McDonnell R., O'Sullivan C.: Perceptual evaluation of human animation timewarping. In *ACM SIGGRAPH ASIA 2010 Sketches* (New York, NY, USA, 2010), SA '10, ACM, pp. 30:1–30:2. 2

[RO09] Reitsma P. S. A., O'Sullivan C.: Effect of scenario on perceptual sensitivity to errors in animation. *ACM Trans. Appl. Percept. 6* (September 2009), 15:1–15:16. 2

[RP03] Reitsma P. S. A., Pollard N. S.: Perceptual metrics for character animation: sensitivity to errors in ballistic motion. *ACM Trans. Graph. 22*, 3 (2003), 537–542. 1, 2

[RPE*05] Ren L., Patrick A., Efros A. A., Hodgins J. K., Rehg J. M.: A data-driven approach to quantifying natural human motion. *ACM Trans. Graph. 24*, 3 (2005), 1090–1097. SIGGRAPH 2005. 1, 2

[SHP04] Safonova A., Hodgins J. K., Pollard N. S.: Synthesizing physically realistic human motion in low-dimensional, behavior-specific spaces. *ACM Transactions on Graphics 23*, 3 (2004), 514–521. SIGGRAPH 2004. 1, 4

[TLKS08] Tang J. K. T., Leung H., Komura T., Shum H. P. H.: Emulating human perception of motion similarity. *Comput. Animat. Virtual Worlds 19* (September 2008), 211–221. 2

[Tro02] Troje N. F.: Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision 2*, 5 (2002), 371–387. 3

[TWC*09] Tournier M., Wu X., Courty N., Arnaud E., Revéret L.: Motion compression using principal geodesics analysis. *Computer Graphics Forum 28*, 2 (2009), 355–364. EUROGRAPHCS 2009. 1

[vBE09] van Basten B. J. H., Egges A.: Evaluating distance metrics for animation blending. In *Proceedings of the 4th International Conference on Foundations of Digital Games* (2009), ACM, pp. 199–206. 1, 2

[WB03] Wang J., Bodenheimer B.: An evaluation of a cost metric for selecting transitions between motion segments. In *SCA '03: Proceedings of the 2003 ACM SIGGRAPH/Eurographics symposium on Computer animation* (Aire-la-Ville, Switzerland, Switzerland, 2003), Eurographics Association, pp. 232–238. 2

[WB08] Wang J., Bodenheimer B.: Synthesis and evaluation of linear motion transitions. *ACM Trans. Graph. 27*, 1 (2008), 1–15. 2

Appendix

Natural vs. synthesized motion

A minor question was whether the participants believed that one of the displayed motion did not correspond to an unedited motion capture sample but was a synthesized motion. The participants were advised that the displayed reference motion O always corresponds to a original motion capture sequence, whereas one of the two other motions possible could be synthesized (and that with equal chance one of the two other motion A or B might be a synthesized motion). However, the participants were informed that at most one of the motions A and B was synthesized.

Motion Synthesis

Synthesized motion sequences used during our experiments were computed applying a simple morphing approach: First the original motion capture sequences were warped to the same length using dynamic time warping with the distance measure $\mathcal{D}^{6}_{\text{pc}}$. Second the root trajectory was synthesized by linear interpolation of the frame difference of the original trajectories. Third the quaternion based orientation data were interpolated using the slerp algorithm. Within the interpolation steps all motion sequences used for morphing were equally weighted.

Results for the question of natural vs. synthesized motion

Disregarding the number of the uncertain-votes U the following two-by-two tables can be extracted from Table 1.

point light renderings		
	natural motion	synthetic motion
vote natural	345	195
vote synthetic	293	57

stick figure renderings		
	natural motion	synthetic motion
vote natural	410	174
vote synthetic	352	75

Thus it can be concluded on the 1 % level by a χ^2 test that the synthesized motions could not be distinguished from the natural motions for the point light walker representations as well as for stick figure representations.

So in the tests on perceptual similarity we do not have to distinguish between triples containing only natural motions from those containing a synthesized motion.

Table 2: *Mapping of motion files from the HDM05 database to the motion triples used in the experiment. Used motions from the HDM05 database for the motion triples used in the tests. For the natural motions we use the names specified in [MRC*07]. The synthetic motions are obtained by morphing between HDM05 motions, and the names encode the motions used for morphing. The motions in the following table are given in a normalized order: if there are synthetic motions, the are listed as motion B; in the test there were random permutations between A and B.*

triple name	**motion** *A*	**motion** *O*	**motion** *B*
T hop 01	HDM_bd_hopBothLegs1hops_001_120	HDM_dg_hopBothLegs1hops_024_120	HDM_bk_hopBothLegs1hops_015_120
T hop 02	HDM_bd_hopBothLegs1hops_01_02_120	HDM_dg_hopBothLegs1hops_023_120	HDM_dg_hopBothLegs1hops_022_120
T hop 03 syn	HDM_bd_hopBothLegs1hops_01_02_120	HDM_dg_hopBothLegs1hops_024_120	hopBothLegs1hops_001_120_syn
T hop 04 syn	HDM_bd_hopBothLegs1hops_001_120	HDM_dg_hopBothLegs1hops_023_120	hopBothLegs1hops_001_120_syn
T jog 01	HDM_mm_jogLeftCircle2StepsRstart_007_120	HDM_tr_jogLeftCircle2StepsRstart_010_120	HDM_tr_jogLeftCircle2StepsRstart_009_120
T jog 02	HDM_dg_jogLeftCircle2StepsRstart_015_120	HDM_tr_jogLeftCircle2StepsRstart_010_120	HDM_tr_jogLeftCircle2StepsRstart_009_120
T jog 03 syn	HDM_mm_jogLeftCircle2StepsRstart_007_120	HDM_tr_jogLeftCircle2StepsRstart_010_120	jogLeftCircle2StepsRstart_09_10_120_syn
T jog 04 syn	HDM_dg_jogLeftCircle2StepsRstart_015_120	HDM_tr_jogLeftCircle2StepsRstart_010_120	jogLeftCircle2StepsRstart_09_10_120_syn
T sneak 01	HDM_mm_sneak2StepsRStart_012_120	HDM_tr_sneak2StepsRStart_015_120	HDM_tr_sneak2StepsRStart_016_120
T sneak 02	HDM_mm_sneak2StepsRStart_011_120	HDM_tr_sneak2StepsRStart_015_120	HDM_tr_sneak2StepsRStart_016_120
T sneak 03 syn	HDM_mm_sneak2StepsRStart_012_120	HDM_tr_sneak2StepsRStart_015_120	sneak2StepsRStart_14_15_16_120_syn
T sneak 04 syn	HDM_mm_sneak2StepsRStart_011_120	HDM_tr_sneak2StepsRStart_015_120	sneak2StepsRStart_14_15_16_120_syn
T walk 01	HDM_mm_walk2StepsRstart_023_120	HDM_tr_walk2StepsRstart_028_120	HDM_tr_walk2StepsRstart_027_120
T walk 02	HDM_dg_walk2StepsRstart_013_120	HDM_tr_walk2StepsRstart_028_120	HDM_tr_walk2StepsRstart_027_120
T walk 03 syn	HDM_mm_walk2StepsRstart_023_120	HDM_tr_walk2StepsRstart_028_120	walk2StepsRstart_023_021_120_syn
T walk 04 syn	HDM_dg_walk2StepsRstart_013_120	HDM_tr_walk2StepsRstart_028_120	walk2StepsRstart_023_021_120_syn
T walkLeft 01	HDM_mm_walkLeft2Steps_013_120	HDM_tr_walkLeft2Steps_014_120	HDM_tr_walkLeft2Steps_016_120
T walkLeft 02	HDM_mm_walkLeft2Steps_011_120	HDM_tr_walkLeft2Steps_014_120	HDM_tr_walkLeft2Steps_016_120
T walkLeft 03 syn	HDM_mm_walkLeft2Steps_013_120	HDM_tr_walkLeft2Steps_014_120	walkLeft2Steps_07_11_120_syn
T walkLeft 04 syn	HDM_mm_walkLeft2Steps_011_120	HDM_tr_walkLeft2Steps_014_120	walkLeft2Steps_07_11_120_syn
T walkRight 01	HDM_mm_walkRightCircle4StepsRstart_010_120	HDM_tr_walkRightCircle4StepsRstart_014_120	HDM_tr_walkRightCircle4StepsRstart_015_120
T walkRight 02	HDM_mm_walkRightCircle4StepsRstart_009_120	HDM_tr_walkRightCircle4StepsRstart_014_120	HDM_tr_walkRightCircle4StepsRstart_015_120
T walkRight 03 syn	HDM_mm_walkRightCircle4StepsRstart_010_120	HDM_tr_walkRightCircle4StepsRstart_014_120	walkRightCircle4StepsRstart_01_02_120_syn
T walkRight 04 syn	HDM_mm_walkRightCircle4StepsRstart_009_120	HDM_tr_walkRightCircle4StepsRstart_014_120	walkRightCircle4StepsRstart_01_02_120_syn