Evaluating Emotional Content of Acted and Algorithmically Modified Motions

Klaus Lehtonen and Tapio Takala

Aalto University,
School of Science,
Department of Media Technology,
Espoo, Finland
klaus.lehtonen@tkk.fi,tapio.takala@tkk.fi

Abstract. Motion capture is a common method to create expressive motions for animated characters. In order to add flexibility when reusing motion data, many ways to modify its style have been developed. However, thorough evaluation of the resulting motions is often omitted. In this paper we present a questionnaire based method for evaluating visible emotions and styles in animated motion, and a set of algorithmic modifications to add emotional content to captured motion. Modifications were done by adjusting posture, motion path lengths and timings. The evaluation method was then applied to a set of acted and modified motions. The results show that a simple questionnaire is a useful tool to compare motions, and that expressivity of some emotions can be controlled by the proposed algorithms. However, we also found that motions should be evaluated using several describing dimensions simultaneously, as a single modification may have complex visible effects on the motion.

Keywords: computer animation, motion capture, emotional motion, evaluation of motion style, motion editing

1 Introduction

We can see emotions in facial expressions, poses and motions of animated characters. Motion capture is a convenient way to record the visible emotions. However, doing new motion capture for each different case requires time and money and is not practical in all situations [1]. New ways to reuse motion data can lower the costs. Reusing and modifying motions are especially beneficial in games, because game characters need to move and react in many different ways in real time.

Much research has been done to change the style of captured motions, but there has been less effort in validating the effects of the modifications. Many published modifications create changes in motions, but how significant the changes are to people viewing the motions has not been investigated [2–5]. In a case where motion style of modified motions was evaluated with many viewers, it was found out that the modification was not always noticed by the viewers [6].

We present a questionnaire based method for evaluating the emotional and stylistic content of motions with a large audience. This evaluation is used to compare the effects of three algorithmic modifications to motion data. These modifications change the posture, motion path length, and timing of the motions. The results show that a questionnaire is an effective method for measuring emotional content of motions and to compare the modification methods. The questionnaire can also reveal if a modification produces unwanted side effects to the style of the motions.

2 Related Works

2.1 Evaluating Motions

We found an earlier evaluation of modified motions that used valence, arousal and dominance as the dimensions in their questionnaire [6]. However, some stylistic properties like masculinity and femininity are not easily mapped to the three dimensions. In another study, dynamic and geometrical features of dancing motions were assessed using Laban motion analysis [7]. This approach requires expert knowledge from the evaluators. Enabling untrained people to evaluate everyday actions requires a simpler approach.

Evaluating emotions seen in facial expressions is a common practice [8]. Facial expressions are a simpler case than bodily motions, because the face is mainly used for communication and much less for other actions, whereas in bodily motion both appear. Questionnaires used to evaluate facial expressions often force labeling one face image with one emotion. This is a justified approach as prototypical faces are associated with basic emotions in commonly used models. Defining similar prototypical bodily motions is hard and, since motions can have many stylistic properties in addition to visible emotions, the task is even harder.

Based on these observations we decided to limit our study to assessing emotional content of bodily motions only. We animated motions with a simplified human model without face and fingers. In the questionnaire we asked evaluation of multiple simultaneous dimensions instead of a single choice.

2.2 Modifying Motions

Neff and Kim state that important components affecting motions are posture, transitions between poses, simplification and exaggeration [9]. We wanted to use modifications that would directly affect these components. Some of these aspects of motion can be changed in motions that are defined with a low number of keyframes per second (around 1 Hz) [10]. For example transitions between poses in keyframes can be easily edited without changing the keyframes themselves. Same methods cannot be used straight away with motion defined with a high number keyframes per second (over 10 Hz), such that motion capture produces, as transitions between the keyframes are too fast.

Many modification methods of motion styles are based on the difference between two recorded motion signals [2–4]. The way these methods treat the motion data differ greatly, but the results are all dependent on the styles of the input motions. In our study, we used a method based on differences between still poses for changing postures of characters. For other changes, however, we used procedurally generated modifications.

Bruderlin and Williams presented multiresolution filtering that we used for exaggerating and diminishing the length of motion paths [5]. The transitions between poses can be changed by the speed transform described by Amaya et. al. [2]. However, the speed transform is not a direct method as it requires two motions for defining the changes made to a third motion. To make the method more direct, we developed a new heuristic approach that defines the changes in timings based on the velocity of the input motion.

3 Implementation

3.1 Capturing Motions

To create suitable motions, we asked one female and one male actor to perform a short walk and a knocking motion with ten different styles that were used earlier by Pollick et. al.: *afraid*, *angry*, *excited*, *happy*, *neutral*, *relaxed*, *sad*, *strong*, *tired* and *weak* [11]. Both of the actors had acted in theater performances, but they did not have previous experience with motion capture. We also captured several standing poses. For recording the motions, we used Optitrack Full Body Motion Capture system, consisting of twelve cameras capturing motions with 100 frames per second. The system gives coordinates and orientation of the hips, and rotations for 18 joints.

3.2 Modifying Motion Paths

To change the length of the motion paths, multiresolution filtering described by Bruderlin and Williams was used [5]. The basic idea is to divide an original motion signal into frequency bands that are modulated separately in a way that adds no phase shift. Then we can return to the original signal format by summing up the modified frequency bands. We observed that multiplying the middle bands (between 1.0 and 12.5 Hz) makes the motion paths exaggerated or reduced.

With this method we can make a short and slow walk longer and faster or vice versa. During the process, the modification exaggerates or diminishes the poses as seen in Figure 1. Our hypothesis was that this modification would make motions look more *excited*, *tired* and *weak*. The hypothesis was based on the findings that high movement activity has been connected with *elated joy*, *hot anger* and *terror*, while low movement activity has been connected to *boredom* and *contempt* [12].

Multiresolution filtering suffers from stretching bones if applied to joint positions as input signals. Using joint rotations can also be problematic as unnatural rotations may happen when the angles are near a gimbal lock. To avoid these problems, we represented the joint rotations with two orthogonal vectors pointing the direction of the bones. This way the bone lengths are not affected by the modification and gimbal lock is not an issue.

Multiplying frequency bands can make the feet of the character slide. This was fixed in post-processing by recalculating the coordinates of the hips. The fix was done by measuring the original supporting periods of the feet and forcing the feet to stay still during these periods in the new motions.

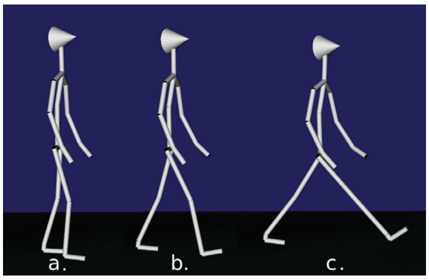

Fig. 1. Changes in pose when a walking motion (*b*) is modified by making motion paths shorter (*a*) and longer (*c*).

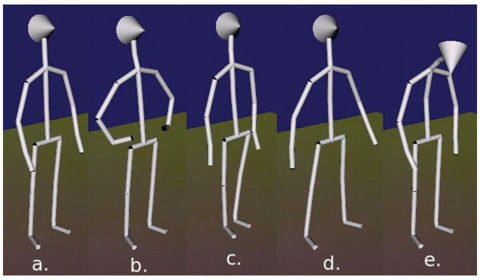

Fig. 2. Still poses used in changing posture of animated characters. Pose (*a*) is the neutral pose of the act

3.3 Modifying Posture

For creating motions with varying posture we need a motion with neutral posture, a neutral still pose and varying expressive poses (Fig. 2). Next we calculate the differences between joint rotations in the *neutral* pose and the other poses. Then we can add the desired difference to each frame of the motion to change the posture of the animated character. If we need to reduce or exaggerate the change of the posture, we can multiply all the changes with a constant. We made a hypothesis that modification to posture could make motions look more *sad* or more *confident*, as earlier studies show that those emotion types have stereotypical positions of upper body, shoulders and head [12].

With a motion like a regular walk the change of posture works quite well as it is. If the motion has parts where pose of the character is very different compared to the neutral pose, the technique can have unwanted side effects. For example, if the character is reaching forward with a straight arm, the arm can become twisted by the change of the posture. This can be fixed by fading the changes gradually out when the end of a limb goes too far from its position in the *neutral* pose. The fade-out must be done to all the joints that affect the position of the limb.

3.4 Modifying Timings

Amaya et al. adjusted the speed of motions by time warping the motions according to differences between two reference motions [2]. Their approach depends on the quality of the original motions and requires motion segmenting to find links between the reference motions. To bypass these requirements, we developed two heuristics that define the time warps without any reference motions. The heuristics aim to produce motions with either constant speed or added acceleration while keeping the total duration of the motion unchanged (Fig. 3). Our hypothesis was that these changes would make motions look more *relaxed* and *angry*, respectively. The hypothesis was based on findings that high movement dynamics are connected with *hot anger* [12].

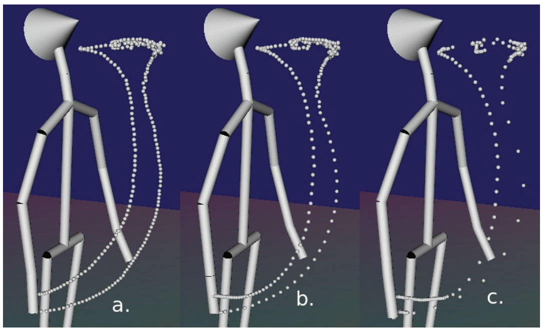

Fig. 3. Trajectories of the right hand during a knocking motion. Original motion (*b*) is modified to nearly constant speed (*a*) or with exaggerated acceleration (*c*).

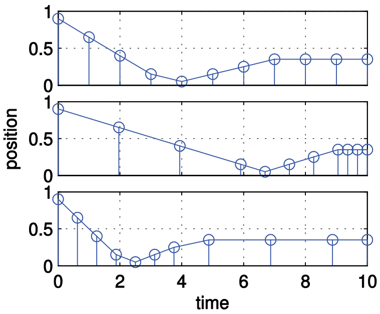

Fig. 4. Timewarping motion: evenly timed original frames (*top*) are repositioned in time to form constant speed (*middle*) or added acceleration (*bottom*).

Transforming a motion to constant speed can be done by taking evenly timed motion samples, spreading out frames with fast movement and compressing frames with slow movement, as depicted in figure 4. New motion is then formed by resampling between these as keyframes. The frames where the position does not change at all, would be compressed completely together to get constant speed, but this would cause all pauses of the motion to be skipped. More natural motion results if the pauses are preserved by enforcing a minimum duration between frames (as seen in the last four frames in the middle plot of figure 4).

In practice time warping is done separately for each limb of the character. When constant speed for each frame is desired, we can use the velocity values of the limb limited by a minimum value as the durations between frames and then scale all frames to preserve original total duration of the motion. When adding acceleration to a motion, we use original durations of the frames and increase the durations next to frames that are local minima of velocity and then scale as in the previous case. This makes the pauses of the motion last longer and other parts faster as seen in the lowest plot of figure 4.

3.5 Combining Modifications

To use the modifications interactively, the path lengths of an original motion are modified to produce a motion with short motion paths and another with long motion paths. Next, the sliding feet are fixed. Then, the timings of the three motions are modified to produce versions with constant speed and added acceleration. Once these steps are done we can create new motions by interpolating between the produced motions.

Modification of posture is done last to the interpolated motion. Modifying path lengths, fixing the sliding feet and modifying timings requires off-line processing. When they are done an animator controlling the system has real-time feedback as the interpolation of motions and changing the posture are both very fast operations.

4 Questionnaire

Evaluation of emotional and stylistic content of motions is necessary for comparing the motions and for testing methods that create modifications. We were also concerned with the validity of a questionnaire based evaluation. This led to three research questions: 1. Can acted styles and emotions be distinguished by viewing motions animated with a stick figure? 2. Do the three implemented modifications change emotions seen in the motions? 3. What are suitable dimensions to be rated when evaluating motions?

The captured acted motions were used as material for the first question. For the second question, we created pairs from the *neutral* motions to each of the acted motion styles. The pairs were created according to our hypotheses of the effects of the modifications and intended to have the same emotional and stylistic content as the acted motions. We also created motions with intended styles *masculine* and *feminine* that attempted to change the gender of the characters. Motion *energetic sadness* was created by combining modifications that we had hypothesized to increase *sadness* and *excitement*. The emotion *happy* was omitted because it was not known which modifications could affect *happiness*. The final combinations of the modifications are in figure 7.

40 videos were created from the acted and modified motions. The videos were shown in randomized order. The participants were able to play the videos many times, but it was instructed that viewing the videos once should be usually enough.

We included in the questionnaire all the emotional descriptions given to the actors. However, *neutral* was not included as it was assumed to fall in the middle of a scale. *Confident* was not part of the original set of emotions recorded with actors, but it was necessary to have a pair for *afraid*. To make the questionnaire simpler to answer, we combined opposite motion styles, making scales between pairs *sad-happy*, *tired-excited*, *angry-relaxed*, *weak-strong*, *afraid-confident* and *masculine-feminine*.

For each video, the participants were asked to evaluate how these adjectives describe the character in the video in a scale with five steps. The middle choice was the default and it was instructed to be used if neither of the alternatives feels good or if the participant is unsure which one is better.

The questionnaire was made with a server side PHP script and the videos were embedded in the web page as Flash objects. 28 non-paid participants were recruited through social media. 8 of them were female and 20 were male.

An answer to our first research question can be found by examining a confusion matrix between the intended styles and the perceived styles as shown in figure 5. Each acted style has the ratings of videos by both actors except the neutral male and neutral female columns, which show how the ratings for *neutral* motions differ between the actors.

Perceived style \ Acted style	weak	tired	afraid	sad	relaxed	angry	excited	happy	strong	neutral	neutral male	neutral female
masculine	,4	,5	,2	,4	,5	,6	,3	,5	,3	,3	,5	,1
feminine	,2	,0	,5	,2	,3	,1	,5	,4	,5	,4	,0	,7
confident	,1	,3	,1	,2	,8	1,4	1,2	1,2	1,1	,5	,6	,4
strong	,1	,2	,1	,1	,5	,8	,6	,6	,6	,3	,3	,3
happy	,1	,0	,0	,0	,1	,3	,8	,3	,4	,2	,2	,1
excited	,1	,0	,2	,0	,1	,8	1,1	,6	,5	,1	,1	,2
angry	,1	,2	,1	,1	,1	,6	,2	,2	,2	,1	,1	,0
relaxed	,2	,4	,1	,4	,6	,4	,5	,5	,7	,6	1,0	,3
sad	,9	,6	,3	,7	,1	,0	,0	,0	,1	,1	,1	,1
afraid	,9	,2	1,3	,6	,1	,0	,0	,0	,1	,2	,2	,3
tired	,6	1,3	,3	,7	,4	,0	,1	,1	,1	,3	,5	,1
weak	,9	,7	,9	,6	,3	,1	,1	,0	,1	,3	,2	,4

Fig. 5. Confusion matrix between acted and perceived styles, showing on average how much each style was seen in each video on a scale from 0 to 2. Underlined scores tell how well an intended style was recognized. Strongest perceptions are on dark background.

Perceived style \ Modified style	weak	tired	afraid	sad	relaxed	angry	excited	strong	female to masculine	female to feminine	male to feminine	energetic sadness
masculine	-,1	,0	-,1	,0	,0	,4	,2	1,0	,0	-,1	-,1	,3
feminine	,1	-,1	,1	-,2	,0	,1	,1	-,3	-,4	,9	,2	,0
confident	-,3	-,4	-,3	-,4	,1	,7	,6	,9	-,3	,1	-,1	,2
happy	-,1	-,1	-,1	-,2	,0	,3	,3	,1	,0	,0	,1	-,2
strong	-,2	-,2	-,1	-,2	,1	,5	,4	1,0	-,2	,0	-,1	,3
excited	-,1	-,1	,0	-,1	,1	,5	,4	,3	,0	,0	,0	,1
angry	,0	,0	,0	,1	,0	,1	,1	,3	,1	,0	,0	,8
relaxed	-,4	-,5	-,4	-,5	-,2	,1	,4	-,3	,0	,0	-,2	-,8
sad	,3	,7	,0	1,3	,0	-,1	-,1	-,1	,0	-,1	,1	1,2
afraid	,3	1,0	,4	,4	,0	-,2	-,2	-,1	,4	-,1	-,1	,1
tired	,4	,5	,3	,6	-,1	-,3	-,2	-,3	,0	,0	,4	,4
weak	,6	1,3	,3	,8	-,1	-,2	-,2	-,3	,1	,0	,2	,2

Fig. 6. Difference of scores between original and modified motions. Maximum possible change is 2. Positive changes higher than 0.25 are on white background and negative under -0.25 are on black background. Values for the intended styles are underlined.

The second research question calls for a comparison between ratings of the original *neutral* motions and the motions made by modifying them. The results are shown as differences of their scores in figure 6.

The third research question was about finding suitable characterizing dimensions to be rated when evaluating motions. We analysed how well the dimensions were chosen by estimating their common factors (using Matlab function *factoran*). Results in figure 8 show that the dimensions are not independent and some of them are effectively the same.

Modifications	
A	very short motion paths
B	short motion paths
C	long motion paths
D	increased acceleration
E	constant speed
F	hands close to body
G	feet close to each other
H	head looking down
I	feet more apart
J	elbows broadly

Intended styles	
afraid	B, F, G
angry	C, D
energetic sadness	C, H
excited	C, D
feminine	F, G
masculine	I, J
relaxed	E
sad	B, H
strong	C, J
tired	A, H
weak	A, F, G

Fig. 7. Intended styles and the modifications done to achieve them.

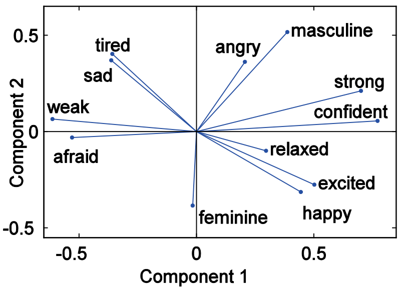

Fig. 8. Maximum likelihood estimate for common factors in the questionnaire. Loadings of the original dimensions are plotted into a two-dimensional model.

5 Discussion

The ratings of the acted videos in figure 5 tell that it is possible to see many styles from a motion that does not have hands or facial expressions. The styles are not seen one at a time, but rather a motion seems to fit a range of styles. This can be partly explained by close relations between the styles that are seen in figure 8, but it is also possible that motion alone does not separate all different styles and emotions as well as facial expressions do. Based on these findings, evaluation of motions requires multiple dimensions to be accurate and selecting more than one description should be allowed.

Figure 6 shows that the three modifications can change emotions and styles seen in motions. The effects of the modifications are not constrained to a single emotion or style. When we compare the intended effects and the perceived styles in figure 6, we can see that the *strong* and *sad* modifications worked well. A contributing factor could be that both modifications had postures that are easily identifiable. Modifications *weak*, *tired*, *afraid* and *excited*, created mainly with changes to motion paths, did add at least a little of the intended style, but the modifications *relaxed* and *angry* did not work as intended. *Anger* and *relaxedness* were hypothesized to be created with changes to the timings of the motions. It is possible that walking and knocking motions are not suitable material for the modification, but based on this data, the modification of timings has to be considered quite useless. A similar ineffectiveness of retiming motions has been noticed earlier when trying to change the emotional content of captured motions [12].

The modification *energetic sadness* in figure 6 created both *sadness* and *angriness*. The modification that was intended to create *sadness* did not create *angriness* and the only difference between the *sad* and the *energetic sad* modifications was in the length of motion paths. This suggests that the modifications do not have only one-to-one relations with emotions, but combinations of modifications can be used to create emotions that the modifications cannot produce separately. This also suggests that the effects of the modifications depend greatly on the styles and emotions of the input motion.

The evaluation shows that modifications to posture and motion paths are good tools for artists, but it does not reveal if the modifications can be used without a human checking the results. A systematic evaluation of different kinds of motions with all combinations of the modifications would be necessary for assessing the reliability of the modifications.

The dimensions in the questionnaire are not perfect as pairing two descriptions to one axis prevents those from being selected at the same time. A common factor analysis of the styles shown in figure 8 tells that there is redundancy in the descriptions we used. In this data set the pairs *tired-sad*, *weak-afraid*, *excited-happy* were closely related. We could simplify the questionnaire by removing one description from each pair without losing much information.

During the analysis two weaknesses were found from the questionnaire. One is related to calculating statistically meaningful variances, which is limited because we had only five steps in each dimension. The second weakness is that all descriptions cannot be used at the same time as one dimension joins two descriptions. A better solution could be to have dimensions that have continuous range instead of steps and descriptions that are in form 'happy - not happy'. However, this would increase the number of decisions a test participant has to do.

6 Conclusions and Future Work

We captured acted motions with different styles and emotions and then tried to produce similar stylistic effects with three algorithmic modifications applied to neutral motions. The motions were evaluated with a questionnaire to determine what styles and emotions are visible in them. This kind of evaluation was not done in any of the papers related to modifications to motion that we have found.

We found evaluating motions with a questionnaire to be a good tool for comparing motions. It was observed that many styles can be seen from one motion and that a single modification can affect many styles. This confirmed that comparison of motions is more meaningful, when examining several describing dimensions simultaneously, instead of only concentrating to one dimension. In this sense evaluating motions is different from evaluating facial expressions as forced choice of one description is not enough for motions. What are the best dimensions for evaluating emotions and styles in motions remains a question yet to be answered.

The results show that modifying posture allowed creating the *strong* and *sad* motions. Changing the length of the motion paths helped in creating *weakness*, *tiredness*, *fear* and *excitement*. Changing timings of the motions was not found to affect the content of the motions significantly. Adjusting timings might be helpful if different types of motions were studied.

When combining modifications, the effects on the style of the motion were not always the same as the sum of effects of the modifications separately. This suggests that modifying already emotional motions could reveal more about the modifications than modifying only *neutral* motions. Similar phenomena could also be present when interpolating emotional motions. Also, creating totally *neutral* motions by motion capture is hard as the physical appearance of actors always affects captured motions. In the end, bodily motions alone cannot express all emotions. Therefore, other methods such as facial expressions must also beused when making complete animations.

Acknowledgments. This work was partially funded by aivoAALTO project of Aalto University and by Academy of Finland, project Enactive Media (128132). Also, thanks to Timo Idänheimo for his help during the work.

References

1. Menache, A.: Understanding Motion Capture for Computer Animation and Video Games. Academic Press. 238 p. (2000)
2. Amaya, K., Bruderlin, A., Calvert, T.: Emotion from motion. Proc. Graphics interface GI'96, pp. 222-229. Canadian Information Processing Society. (1996)
3. Hsu, E., Pulli, K., Popovic, J.: Style translation for human motion. Proc. SIGGRAPH '05, ACM Transactions on Graphics 24 (3), pp. 1082-1089. (2005)
4. Shapiro, A., Cao, Y., Faloutsos, P.: Style components. Proc. Graphics Interface GI'06, pp. 33-39. Canadian Information Processing Society. (2006)
5. Bruderlin, A.,Williams, L.: Motion signal processing. Proc. SIGGRAPH '95, ACM Transactions on Graphics 14 (3), pp. 97-104. (1995)
6. Heloir, A., Kipp, M., Gibet, S., Courty, N.: Evaluating Data-Driven Style Transformation for Gesturing Embodied Agents. Proc. Intelligent Virtual Agents (IVA '08). In: Prendinger, H., Lester, J., Ishizuka, M. (Eds.), LNCS, Vol. 5208, pp. 215-222. Springer-Verlag, Heidelberg. (2008)
7. Hachimura, K., Takashina, K., Yoshimura, M.: Analysis and evaluation of dancing movement based on LMA. IEEE International Workshop on Robot and Human Interactive Communication, pp. 294-299. (2005)
8. Ruttkay, Z.: Cultural dialects of real and synthetic emotional facial expressions. AI & Society 24 (3), pp. 307-315. (2009)
9. Neff, M., Fiume, E.: From Performance Theory to Character Animation Tools. In: Klette, R., Metaxas, D., Rosenhahn, B. (eds.) Human Motion: Understanding, Modelling, Capture, and Animation. Springer. (2008)
10. Chi, D., Costa, M., Zhao, L., Badler, N.: The EMOTE model for effort and shape. Proc. SIGGRAPH '00, pp. 173-182. ACM Press/Addison-Wesley Publishing Co., New York, USA. (2000)
11. Pollick, F., Paterson, H., Bruderlin, A., Sanford, A.: Perceiving affect from arm movement. Cognition, Vol. 82, Issue 2. (2001)
12. Wallbott, H.: Bodily expression of emotion. European Journal of Social Psychology, Vol. 28, pp. 879-896. (1998)