Analysis and Generation of Emotionally-Charged Animated Gesticulation

Bożena Kostek and Piotr Szczuko

Multimedia Systems Department,
Gdansk University of Technology
{bozenka, szczuko}@sound.eti.pg.gda.pl

Abstract. Computer-animated sequences of emotionally-charged gesticulation are prepared using keyframe animation method. This method consists in creating an animation by changing the properties of objects at key moments over a time sequence. Such a sequence is analyzed in terms of locations and spacing of the keyframes, shapes of interpolation curves, and emotional features present in them. In the paper the keyframe method serves for creating animated objects characterized by differentiating emotions. On the basis of the analysis of these computer-animated sequences several parameters are derived. Then decision tables are created containing feature vectors describing emotions related to each object at key moments. This system serves for derivation of rules related to various categories of emotions. Rules are analyzed, and significant parameters are discovered. Conclusions and future experiments are also outlined.

1 Introduction

Traditional animation and computer keyframe animation are governed by their own rules and principles [2]. Both aim at animating a character with a highly human appearance, personality and emotions though using different artistic and technical means [9,11]. The simplest definition of the animation process is creation by changing the properties of objects over time. A keyframe is a time point when a property has been set or changed. In traditional animation master animators draw the keyframes and assistants do the in-between frames. In computer animation in-between frames are calculated or interpolated by the computer. Although computer animation process is much faster than traditional one, the quality of computer animations still lacks naturalness.

In our work an attempt is done to design an expert system based on animation rules used in traditional animation, capable of creating realistic animated characters according to the animator's requirements. One can analyze and parameterize features of hand-made animation and then translate them into the description of emotional features. These data can be utilized for the creation of the knowledge base containing feature vectors derived from the analysis of animated character emotions. In such a way the user's tasks could be limited to designing a simple animated sequence, and to delivering a description of the desired emotions to the expert system. These data could next be processed in the fuzzy logic module of the system, in which the animation parameters are modified, and as a result an emotionally charged animation is generated. The last phase is the evaluation of the conveyed emotion rendered into the animation.

2 Animation of Realistic Motion of Character

In the beginning of XX century, simultaneously with the invention of the movie, the animation was born. At the very early stage of animation, to create the animated motion, a small black-and-white picture was drawn on several layers of celluloid. Continuity of motion was achieved by introducing very small changes between each frame (cel), drawn one after one. Later the color drawing on papers and on foils attached to the peg was invented. Therefore it was possible to create frames in varied order - first the most important and powerful poses were designed, then transitional frames were filled in, called also in-betweens [2]. The same approach is utilized in the computer animation systems with keyframes [11]. An animator first sets up main poses in time and space then the system fills in the transitional frames by interpolating locations, rotations and torque of objects and characters' bones. The animator can influence the interpolation process, therefore the acceleration, slow-in and slow-out phases can be changed, and keyframes inserted or transposed.

2.1 Animation Rules

Correct utilization of poses and transitions between them can implicate different emotional features of the character's motion. As a result of experience derived from the first years of traditional animation the animation rules were created in 1910, by the art designers from the Walt Disney studio. These rules state how to achieve specific features, utilizing posing, keyframes, and phases of motion [9,11].

One of the basic rules is anticipation. This refers to preparation for an action. For every action there is an equal and opposite reaction. Before the main action starts there should always be a preparation for it, which is a slight movement in a direction opposite to the direction of the main action. If the preparation phase is long, the performing character will be perceived as weak, or hesitating. Short anticipation gives effect of a self-confident, strong character.

Follow-through is a rule related to physics of the moving body. The motion always starts near the torso: first the arm moves, then the forearm, and the last is hand. Therefore keyframes for bones in the forearm and hand should be delayed comparing to the arm bone. The delay adds a whip-like effect to the motion, and the feeling of the elasticity and flexibility.

Another rule connected to the physical basis of motion is overshoot. The last bone in the chain, for example hand, cannot stop instantly. It should overpass target position, and then goes back to it and slowly stops.

Stops are never complete. A rule called moving hold is related to keeping character in a pose for some time. Very small movements of head, eyes, and limbs should be introduced to maintain the life-like character.

Natural motion almost always goes along curve. Only a vertically falling object moves absolutely straight. Joints in human body implicate curve moves, but also head movement from left to right if straight, will not appear to be natural.

Other rules, not mentioned here, are related to staging, i.e. posing a character in front of the camera, exaggeration of motion, squashing and stretching for achieving cartoon-like effects, and so on. However, these rules do not fall within the scope of the objectives of this research study.

The know-how of hand-made animation and the rules being passed through generations of animators, have a subjective nature, and have never been analyzed by scientific means.

2.2 Computer Animation Methods

Many attempts were done to achieve a realistic motion with computer methods. Physical simulations of a human body motion were created, resulting in a realistic motion during flight, free fall, etc. [4]. It is necessary to assign boundary conditions such as how the motion starts, when and where it should stop, and on this basis the transitional phase is calculated. Such a method makes the animation realistic but neglects emotional features.

An attempt was also done to connect emotions with energy, therefore a very energy consuming motion was assumed as happy and lively, and a low energy consuming motion as tired and sad. In self-teaching algorithms the energy consumption was assumed to be the target function. The method was tested on 4-legged robot creature, and is still in research phase [6].

Genetic algorithms and neural networks were applied to create a model of human, and to teach it to move and react in a human-like way [10]. Data related to physiology of the human body were gathered and the target functions were composed related to keeping vertical position, not to falling over, and reaching a desired location on foot. A system developed for this purpose was trained to accomplish these tasks. Various behaviors were implemented, such as walking, falling, running, limping, jumping, reacting to being hit by different forces, but the method developed actually lacks emotional acting.

There were also some efforts to create new controllers for the animation process. For example a recorded motion of the pen drawing on the computer tablet is mapped to some motion parameters, like changes of location, rotation and speed of selected character's bone, or the movement of character's eyes. It gives the animator new technical means to act intuitively with the controller, and map that action onto the character [8].

A very similar and well-known method is the motion capture [3], consisting in the registration of sensor motions, which are attached to the key parts of the performer's body. It is possible to use a single sensor as a controller for the selected motion parameter [7]. Unfortunately this method is expensive, and the processing of the captured data is very unintuitive and complex, and editing by hand is then nearly impossible.

3 Case Study

In this paper it is assumed that animated motion could be described in a similar way as the animator does it. This means that using data of character's bones in keyframes along with the interpolation data is sufficient to comprehensively analyze and generate the realistic and emotionally charged motion. It is assumed that the animation rules presented earlier are related to these data, thus it is desirable to treat them as a knowledge base, and to evaluate both their correspondence to emotional features of a motion, and the effectiveness of the parametrization performed in terms of its categorization.

4 Preliminary Emotional Feature Evaluation

For the purpose of this research work one of the authors, a semi-professional animator, basing on traditional animation rules, created a series of animations. These animation objects present two arm gestures expressing picking up hypothetical thing and pointing at hypothetical thing with fear, anger, sadness, happiness, love, disgust and surprise. Animations consist of rotations of the joints as shown in Figure 1. In such a case none of other bone parameters are changed. 36 test animations were prepared.

It is assumed that each animation should clearly reproduce the emotion prescribed according the the rule applied. However taking into account subjectivity of emotions, an additional evaluation of perceived emotions is needed. This may help to verify the emotional definition of each sequence.

The following subjective test was conducted. For each animation, presented in a random order, viewers were supposed to give an answer as to what kind of emotion is presented by indicating only one emotion from the list. This was done by assigning answers from 3-point scale (1 - just noticeably, 2 - noticeably, 3 - definitely). Additional features were also evaluated: strength of emotion, naturalness of motion, its smoothness, and lightness, all of which have a 5-point scale assigned to them (e.g. 1 - unnatural, 3 - neutral, 5 - natural). 12 non-specialists, students from the Multimedia Systems Department of the Gdansk University of Technology took part in the tests. This was the typical number

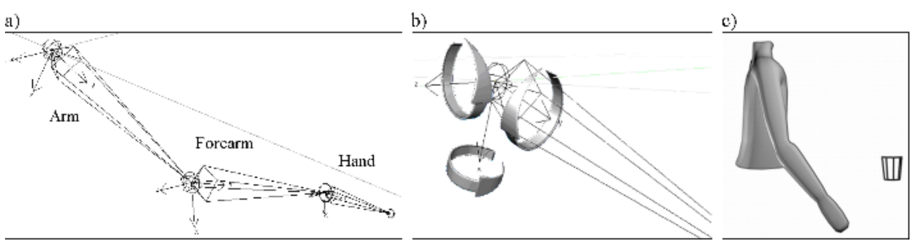

Fig. 1. Presentation of a chain of bones (a), possible bone rotations (b), torso with animated arm (c)

of subjects taking part in such an evaluation [5]. Students were not trained to make these evaluations.

In the first attempt of evaluation, 17 of 36 animations were described by participants as having different emotion than the animator aimed at. Therefore in the next stage all gathered answers were analyzed, but not taking into account the prior assumptions about emotions. The results obtained revealed also some difficulty in selecting only one description. Participants reported similarity of gestures, e.g. similarities between 'negative' emotions like fear, surprise, and anger. 27.9% of acquired data describes anger emotion, 19.4% fear, 16.5% happiness, 13.3% love, and others like disgust, sadness and surprise were the remaining emotions indicated. It was also checked that after training, subjects were more consistent in their answers, thus in future tests a training session will be carried out prior to the evaluation tests.

5 Parametrization of Animation Data

For the purpose of analysis of animation data various parameters of the keyframed motion are proposed. These parameters are related to the process of creating animation. The starting point of the animation process is very simple

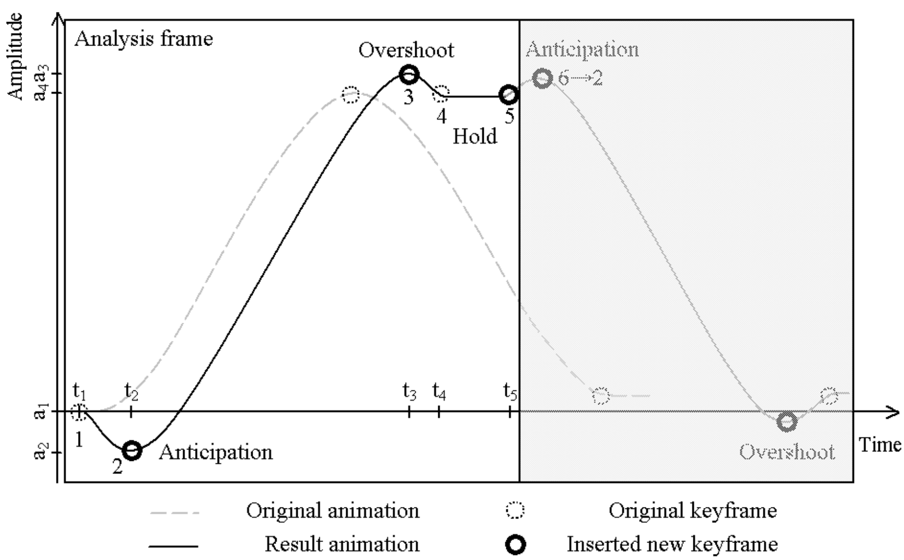

Fig. 2. Presentation of animation parameters. Original animation is without important elements such as anticipation and overshoot. Result animation is a variation of the original one, with anticipation and overshoot inserted. For that purpose it is necessary to add new keyframes, which change the curve of motion amplitude. Keyframes times are marked as t_i, and values of amplitude associated with them as a_i.

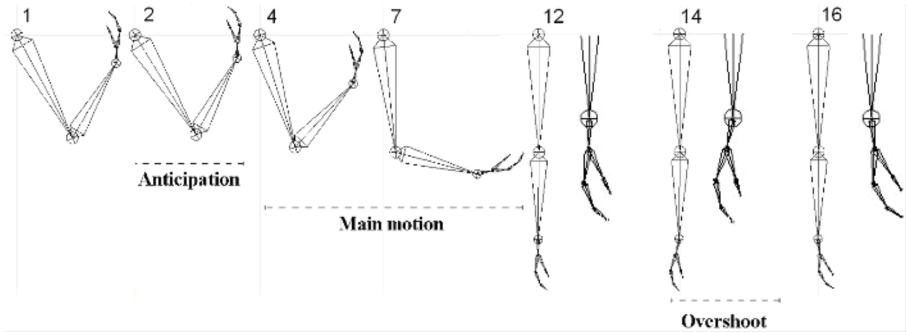

Fig. 3. Animation utilizing bone rotations with frame numbers and motion phases marked. Anticipation is very short (2 frames long) and subtle motion in direction opposite to main motion. Main phase usually extends across many frames and changes of rotation for bones are distinct. The overshoot is a short phase before the complete stop, when last bones in chain overpass target position (hand bones are magnified to visualize the overpass in frame 14).

Table 1. Decision table

Bone	Rot	A_m	t_m	V_m	A_a	t_a	V_a	A_o	t_o	V_o	t_h	A_a/A_m	t_a/t_m	Decision
Arm	Rot_X	1.2	2	0.6	1	1	1	0.5	1	0.5	2	0.83	0.5	Surprise
Hand	Rot_Y	1	2	0.5	2	2	1	2	1	2	3	2	1	Fear
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Hand	Rot_Z	2	2	1	1	2	0.5	0.2	1	0.2	2	0.5	1	Happiness

and the animator adds important elements such as anticipation and overshoot later (see Figure 2).

Animations for the test were prepared utilizing only rotation of bones, therefore there are separate keyframes for rotations along X-, Y-, Z-axes, and in such a case amplitude literally means the rotation angle (see Figure 3).

Long animated sequences are segmented into parts with one main motion phase and one hold phase (see Figure 2, 'Analysis frame'). Each segment is analyzed and included as a pattern in the decision table (Table 1). For each animation segment, values of animation parameters are calculated related to amplitudes of particular phases ($A_a = a_2 - a_1, A_m = a_3 - a_2, A_o = a_4 - a_3$), their lengths ($t_m = t_3 - t_2, t_a = t_2 - t_1, t_o = t_4 - t_3, t_h = t_5 - t_4$), and speeds ($V_m = A_m/t_m, V_a = A_a/t_a, V_o = A_o/t_o$). Variables a_i and t_i are depicted in Figure 2. The decision parameter for an animation segment is the emotion assigned most often by the participants of a test while rating the animation.

This table serves as an input to the rough set system. The system task is to evaluate rules describing interrelations between the calculated parameters and features of motion.

6 Data Mining

For the generation of rules based on the decision table, the Rough Set Exploration System can be used [1]. During the processing, the automatic discretization of parameter values is performed. In the case of experiments performed the local discretization method was used. Some parameters were excluded at this level as not important for defining the searched relations. 12 parameters were left in the decision table: *Bone*, *Rotation axis*, amplitude for first keyframe a_1, length and amplitude of anticipation phase t_a, A_a, amplitude for anticipation keyframe a_2, length and speed of main motion phase t_m, V_m, time for overshoot keyframe t_3, length of hold phase t_h, speed of overshoot phase V_o, and time for ending keyframe t_5.

With utilization of the genetic algorithm available in the Rough Set Exploration System there were 1354 rules generated containing the above given parameters. Total coverage was 1.0, and total accuracy in classification of objects from decision table was equal to 0.9. The shortening of the rules resulted in a set of 1059 rules, giving total coverage 0.905, and accuracy for classification was equal to 0.909.

An attempt was done to lower the number of used parameters. Therefore from 12 parameters, only 6 were chosen. There were such as follows: *Bone*, and ones having more than three discretization ranges: a_1, t_a, t_m, t_3, t_h. This resulted in generation of 455 rules with total coverage 1.0 and accuracy equal to 0.871. After additional shortening, 370 rules were left, resulting in lower coverage (0.828) and accuracy of 0.93.

After discarding a_1, when only *Bone* and time parameters have been left, results not changed much: there were 197 rules, with coverage 1.0 and accuracy of approx. 0.871. Also after shortening, when only 160 rules left, coverage was the same as before (0.828), and accuracy was equal to 0.933. The resulting confusion matrix is presented in Table 2. It should be noted that the a_1 parameter is discretized into 10 ranges, and consecutively t_a into 12, t_m into 6, t_3 into 7, and t_h into 5 ranges.

Further attempts to reduce the number of parameters resulted in great accuracy loss.

Table 2. Results of classification of objects with derived set of 160 rules

		Predicted:						
		Anger	Sadness	Love	Happiness	Fear	Surprise	Disgust
Actual:	Anger	103	0	0	0	0	0	0
	Sadness	4	31	0	0	0	0	0
	Love	0	0	37	0	0	0	0
	Happiness	0	0	0	50	3	0	0
	Fear	0	0	8	0	57	0	4
	Surprise	0	0	0	0	0	16	0
	Disgust	0	4	0	0	0	0	24

It may be said that these results are very satisfying, namely it was discovered that time parameters are especially important for defining emotions in motion. In the derived 160 rules generated for 5 parameters, 100 of them use only time parameters. 102 rules have a size of 2, and 58 rules have a size of 3. Maximal support for the rule is 29, minimal 3, and mean is 5.7.

For rules, related to each class, parameter values were analyzed, resulting in creation of representative sets. For example for *surprise* emotion, most objects in decision table have t_a=(7.5,8.5), t_m=(17.5,Inf), t_3=(45.5,51.5), t_h=(18.5,19.5) for all bones, and t_a=(16.5,19.5), t_m=(17.5,Inf), t_3=(36.0,45.5), t_h=(19.5,21.0) for *love*. This information could be utilized later in a fuzzy logic module, for generating variations of animations introducing desired emotional features.

Examples of the rules derived are: IF (Bone=Forearm) AND (t_m=(-Inf,5.5)) THEN (Decision=Anger), and IF (t_a=(16.5,19.5)) AND (t_m=(-Inf,5.5)) THEN (Decision=Anger), which can be rewritten in natural language as follows: *"if the forearm main motion phase is very short then emotion is anger"*, and *"if anticipation in motion is long and the main motion phase is very short then emotion is anger"*.

Next for each discretization range appropriate linguistic description should be subjectively selected, and used as membership function name. Tests will be needed to choose functions for ranges of e.g. *short*, *medium* and *long* anticipation.

7 Conclusions and Future Experiments

It is assumed that most of the rules generated in the rough set analysis are closely congruent to the animation rules derived from the traditional animation. The presented here work will be continued in order to create tools for realistic and emotionally charged animation generation.

Results presented verify a very important fact known to the animators, namely that timing of keyframes is important for defining the emotion in animation. Interrelation between emotions and amplitudes gives possibilities to introduce many constrains to motion, such as exact specification of target position for grabbing, or walking sequence. In such case the amplitudes may remain unchanged to satisfy constrains, but timing may vary, resulting in different emotional features.

The rules generated could be used in the expert system containing a fuzzy logic module. The rough set measure can be applied as the weight in the fuzzy processing. The input parameters fed to the system may consist of data obtained from keyframes of a simple animation created for easy manipulation along with the description of the desired emotional feature. Output parameters are the ones needed for creating realistic motion sequence: positions of keyframes and lengths of anticipation and overshoot phases. For each parameter, membership functions should be defined, correlated to discretization cuts acquired in the rough set rule generation process (e.g. for animation parameters), or triangle membership functions covering ranges from 0 to 100 (e.g. for animation features).

For any simple animation, a set of modifications could be created, utilizing different values of emotional descriptions. With the output parameters from the fuzzy logic module, changes can be introduced to the animation, resulting in new keyframes, insertion of anticipation and overshoot phases. It is planned to generate sequences with the same emotional features as the ones prepared by the animator in the first stage, and verify their emotional quality and naturalness. The results obtained may be utilized for derivation of better rules, and the effectiveness of the system can increase.

The methodology outlined can also be extended to other parts of the human body, and this is planned as the future aim. The practical utility of this research is to enhance computer-based animation features in order to create animation more realistic and human-like.

References

1. Bazan, J.G., Szczuka, M.S., Wróblewski, J.: A new version of rough set exploration system. In: Alpigini, J.J, Peters, J.F., Skowron, A., Zhong, N. (eds.): RSCTC'2002. Lecture Notes in Artificial Intelligence, Vol. 2475. Springer-Verlag, Malvern (2002) 397–404
2. Blair, P.: Cartoon Animation. Walter Foster Publishing, Laguna Hills (1995)
3. Bruderlin, A., Williams, L.: Motion signal processing. Computer Graphics **29** (1995) 97–104
4. Fang, A.C., Pollard, N.S.: Efficient Synthesis of Physically Valid Human Motion. ACM Transactions on Graphics, Vol. 22. ACM (2003) 417–426
5. Kostek, B.: Soft Computing in Acoustics, Applications of Neural Networks, Fuzzy Logic and Rough Sets to Musical Acoustics, Physica Verlag, Heidelberg, NY (1999).
6. Park, J., Kang, Y., Kim, S., Cho, H.: Expressive Character Animation with Energy Constraints. Proc. Edu+Compugraphics '97. Vilamoura, Portugal (1997) 260–268
7. Popović, J., Seitz, S.M., Erdmann, M.: Motion Sketching for Control of Rigid-Body Simulations. ACM Transactions on Graphics, Vol. 22. ACM (2003) 1034–1054
8. Terra, S.C.L., Metoyer, R.A.: Performance Timing for Keyframe Animation. Proc. Eurographics/SIGGRAPH Symposium on Computer Animation (2004) 253–258
9. Thomas F., Johnston O.: Disney Animation - The Illusion of Life. Abbeville Press, New York (1981)
10. Whitepaper: Dynamic Motion Synthesis. NaturalMotion Ltd., Oxford (2004)
11. Williams, R.: The Animator's Survival Kit: A Manual of Methods, Principles, and Formulas for Classical, Computer, Games, Stop Motion, and Internet Animators. Faber & Faber (2002)