Emotional Body Expression Parameters In Virtual Human Ontology

A. García-Rojas, F. Vexo, D. Thalmann,[1]
A. Raouzaiou, K.Karpouzis and S. Kollias[2]

[1]Virtual Reality Laboratory (VRlab)-EPFL
[2]Image, Video and Multimedia Systems Laboratory
National Technical University of Athens

Abstract. To create expressive body animation in Virtual Humans one uses motion captured sequences, because the results are more realistic and credible. This technology is expensive in time and cost. Its reusability is not evident and it depends on the quality of the descriptors one can adhere to animations. We propose an ontology where one can look for animations that can be adapted inside an emotional context. These animations are classified within a representation of emotion. Since we use MPEG-4 animations it is also possible to provide parameters that can enhance the animation to a more expressive performance.

1 Introduction

Body gestures in Virtual Humans (VH) are important to increase realism and believability, moreover the body also communicates. A good performance of body movements in VHs is achieved by pre-constructed motion captured sequences. However, motion capture is an expensive and exhaustive work, consequently the reusability of animations is highly desired. For reusing specific animations it is necessary to annotate them with extra meta-data. This information should provide not only the kind if movement it performs, but also the way that it is performed in order to fit it in a context. In [7], authors presented a behavior library based on MPEG-4 animations where reuse can be made depending on the content, but emotional expressions are not contemplated. Even if we can find the desired animation, we would like to have the same animation with some variations to avoid repetitions or to adapt it better into a expressive context.

Raouzaiou et al. have presented in [1] approaches to synthesize a MPEG-4 based expressions using discrete and dimensional emotion representation models [2]. They have provided high level descriptors to communicative gesture animations and the possibility of enhancing animations with expressivity parameters. Expressive animation comprehends many fields of knowledge like psychology, animation control, standards, etc. using ontology suggest a good solution for a formal specification of a shared knowledge [5].

In this paper we aim at structuring inside an ontology concepts involved in the mentioned work. We annotate animations with emotional information and also we register the knowledge to enhance animation's expressivity. Reusing and adapting animations using semantics facilitates the construction, functionality and control of VHs. This work will be part of the Virtual Human Ontology [6] created in the framework of AIM@SHAPE, where the knowledge of virtual human construction has been posed.

2 Ontology Domain

The goal of ontologies is to lay out the knowledge and express it in a way that is understandable for both humans and computers. The format of animations we deal with in the paper is the Face and Body Animation object (FBA) of MPEG-4 framework. FBA is for describing the animation and structure for human-like characters. Concerning body animation, Body Animation Parameters (BAP) provide transformation values for the VH's skeletal structure for each frame to produce the animation. Under this structure we can establish a relation between the animations control and expressive parameters. This relation is made using psychological models of emotions that are also described in the ontology. The model we use to represent the emotions is Whissel's wheel activation-evaluation space [16]. This model is a simple circular representation that captures a wide range of emotions and simplifies them in two dimensions: activation and evaluation. In figure 1 we present this model with an example of some emotions and their position and values in the wheel.

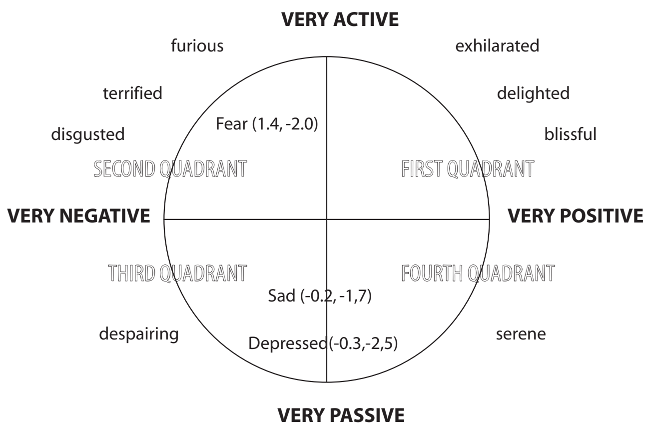

Fig. 1. Whissel's activation-evaluation space model of emotions.

In next sections we present the ontology under the described domain: MPEG-4 animation framework and Whissel's wheel model of emotion. The ontology was created using Ontology Web Language [14]. OWL is a formal language that precisely specifies the semantic relationships among entities. OWL facilitates greater machine interpretability of web content than that supported by XML, RDF, and RDF Schema (RDF-S) by providing additional vocabulary along with a formal semantics. This language offers more ways of entity relation, cardinality relation between entities and more others logical operations when compared with older languages. The ontology design was made using Protégé ontology editor [12] which is an open source and knowledge-base framework.

3 Body Expressive Parameters

Body expressions are hard to synthesize because they are context dependent. Unlike face expressions it is impossible to identify emotions outside a context and without the face. Moreover, we cannot synthesize body movement due to the complexity of its structure. We need to make use of pre-created animation sequences to achieve a realistic performance. For reusing animations we incorporate emotional information in order to extract animations that can fit in an expressive context.

In the ontology we classify animations gestures by associating them to emotions. For example, hand clapping gesture can be associated to Joy and excitation, arms crossing can be associated to anger or unconformity, etc. This process of gesture classification was derived from captured videos of acted session presented in [1]. Those sessions included 7 actors (students of our lab), where each one performed 7 gestures. Each gesture was performed several times with the student-actor impersonating different situations. Afterwards we made a classification presented in the table 1. This classification has been translated into the Whissel's model, where they have been assigned within one or more quadrants of the activation-evaluation space; this means that those gestures can be considered as part of all that emotional scenarios.

Gesture	**Whissel Quadrant**
explain	(0,0), (+,+),(-,+),(-,-)
oh my God (both hands over head)	(+,+), (-,+)
leave me alone	(-,+), (-,-)
raise hand (draw attention)	(0,0), (+,+),(-,-)
bored (one hand under chin)	(-,-)
wave	(0,0), (+,+),(-,+),(-,-)
clap	(0,0), (+,+),(-,+),(-,-)

Table 1. Classification of gestures in a Whissel's wheel quadtrant.

The classification of gestures is one part of the goal. With this we are able to retrieve gestures that can be applied inside an emotional state context.

To be able to enhance animations with more expressive gestures we have found through the literature of perception studies [9], [15] that expressive communicative behavior (only) can be described in six dimensions. The dimensions of expression are: Overall activation, Spatial extent, Temporal, Fluidity, Power/Energy and Repetitivity. Overall activation is considered as the quantity of movement during a conversational turn. Spatial extent is modeled by expanding or condensing the entire space in front of the agent that is used for gesturing. The temporal parameter of the gesture determines the speed of the arm movement of a gesture's meaning carrying stroke phase and also signifies the duration of movements (e.g., quick versus sustained actions). Fluidity differentiates smooth/graceful from sudden/jerky ones. The power, actually is identical with the first derivative of the motion vectors calculated in the first steps.

Each dimension of expression can take place during the different phases of the gesture. For an arm gesture, expressivity works at the level of the phases of the gesture: for example the preparation phase, the stroke, the hold as well as on the way two gestures are co-articulated [8], [4]. For each emotion we have several values of each dimension. The table 2 shows an example of the values assigned to the emotion Joy.

Overall Activation	Spatial Extent max	Spatial Extent mean	Temporal	Fluidity	Power	Repetitivity
3164.57	2.52	1.8236	71	0.93	44.57	0
2389.68	3.07	1.59	71	1.03	33.66	0
3599.99	8.53	4.14	81	0.91	44.44	0
...						

Table 2. Classification of gestures in a Whissel's wheel quadrant.

The tested application used for comparing the emotionally enriched gestures is GRETA [3]. Details about the actual implementation can be found in [10]. In this example we need a particular gesture "wave" in order to complete an animation. If we want synthesize the emotion joy, we will use the BAPs values of gesture "wave" multiplied by the appropriate values of expressivity parameters. The frames resulting of applying this mechanism is showed in the figure 2.

4 Ontology of Body Expressions

The representation of the animations within their classification and the expressive parameters is presented inside the ontology in the Figure 3. Each animation is categorized inside one or more Whissel's wheel quadrant; and it is the same for each emotion, where each emotion belongs to a one quadrant. Each emotion has values associated of the six parameters of expressivity. Therefore an animation that belongs to a quadrant can be enriched with the expressive parameters of the emotions that are in the same quadrant.

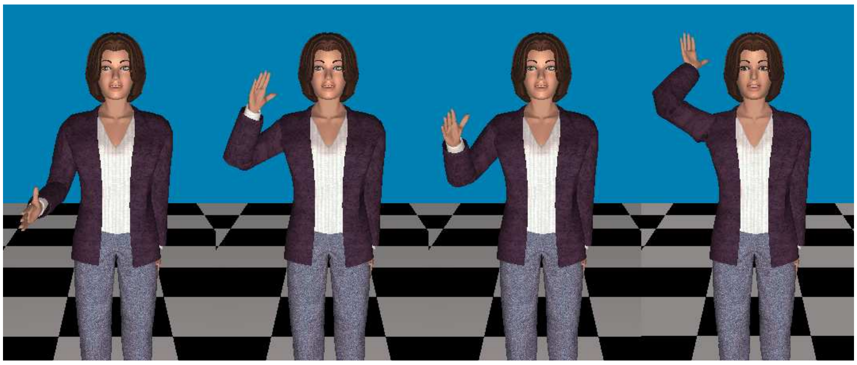

Fig. 2. Example of "wave"gesture enhanced with joy.

For the population of the ontology we instanced the animation gestures presented in table 1: explain,"oh my Go" (both hands over head), "wave", etc. Each animation is associated to their corresponding quadrant of Whissel's wheel. Also we created the corresponding dimensions of expressivity for the joy emotion. In the next section we will explain the results we can obtain once we have this population.

5 Ontology Usage Scenarios

The way to extract information from ontologies is by making queries. To do this we have used the plug-in new Racer Query Language interface [11] for OWL ontologies in Protg. This query language is close to natural language. nRQL language is provided by RacerPro [13] which is a system for managing semantic web ontologies based on OWL. RacerPro is necessary to be run for providing reasoning when querying the ontology using nRQL.

To illustrate the usage of the ontology, we provide the following questions that the ontology is be able to answer:

What are the animations that express joy? (Animations that belongs to the same quadrant as joy emotion)

$(retrieve\ (?c)\ (and\ (?a\ |WhisselWheel|)$
$(?a\ ?b\ |isLocatedInQuadrant|)$
$(?b\ |WhisselWheelQuadrant|)$
$(?b\ ?c\ |hasAnimationRelated|)$
$(?c\ |BodyAnimation|)$
$(?a\ (some\ |modelsEmotion||CONCEPT-FOR-Joy|))))$

Result:

claping
explain
oh_my_God
right_hand_wave

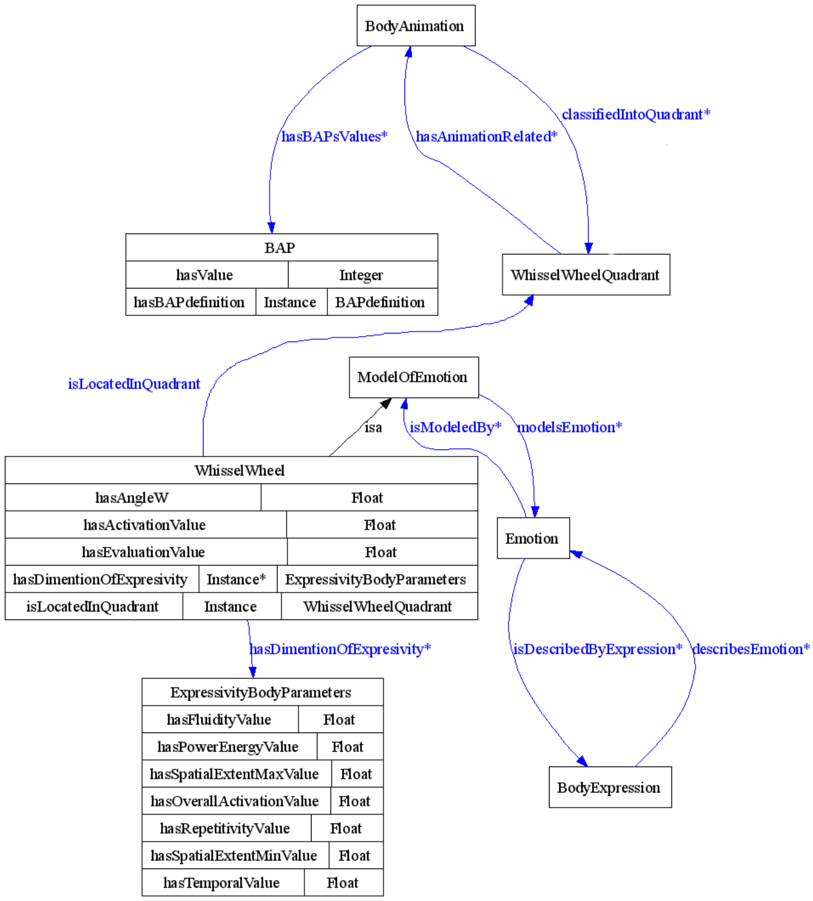

Fig. 3. Diagram of the body expression ontology.

What are the parameters of expression to enhance this animation with joy?
$(retrieve\ (?b)\ (and\ (?a\ |WhisselWheel|)$
$(?a\ ?b\ |hasDimentionOfExpresivity|)$
$(?b\ |ExpressivityBodyParameters|)$
$(?a\ (some\ |modelsEmotion||CONCEPT-FOR-Joy|))))$
Result:
$ExpressivityBodyParameters\_Joy2$
$ExpressivityBodyParameters\_Joy1$
Each of these last elements has a set of the six parameters of expressivity.

6 Conclusions and Future work

In this paper we have presented an ontology that lays out the knowledge of previous work on body animation expressions within MPEG-4 framework. The animations are classified into a model of emotion and some parameters are presented in order to enhance the expressivity. The structure of this ontology allows us to retrieve animations that have been annotated with emotional descriptors under the presented psychological model of emotion; and also to retrieve the expertise of the emotional expressive enrichment of animations. The main issue that is normally faced when adding semantics to a knowledge domain is the extraction of the meta-data and the interconnection with existing knowledge sources, such as audiovisual databases with expressive material. This issue needs to be solved as part of our future work.

Another issue that is currently being investigated in the integration of the ontology with an MPEG-4 compliant animation systems such as Greta [3]. In this scenario, the animation would query the ontology for the representation of an avatar with specific expressivity, e.g. a woman showing extreme joy. The result of this query would be the definition of the avatar geometry and a set of facial expressions and body gestures, ready to use in an animation.

7 Acknowledgments

This research has been funded by the EC and the Swiss Federal Office for Education and Science in the framework of the European Networks of Excellence IST-HUMAINE (http://emotion-research.net) and IST-AIM@SHAPE (http://www.aim-atshape.net).

References

1. G. Caridakis, A. Raouzaiou, K. Karpouzis, and S. Kollias. Synthesizing gesture expressivity based on real sequences. In *Workshop on multimodal corpora: from multimodal behaviour theories to usable models, LREC 2006 Conference*, Genoa, Italy, May 2006.
2. R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. Taylor. Emotion recognition in human-computer interaction. In *IEEE Signal Processing Magazine, no. 1, January 2001, pp.32-80*, 2001.
3. F. de Rosis, C. Pelachaud, I. Poggi, V. Carofiglio, and B. D. Carolis. From greta's mind to her face: modelling the dynamics of affective states in a conversational embodied agent. *Int. J. Hum.-Comput. Stud.*, 59(1-2):81–118, 2003.
4. P. Gallaher. Individual differences in nonverbal behavior: Dimensions of style. *Journal of Personality and Social Psychology*, 63, 1992.
5. T. R. Gruber. The role of a common ontology in achieving sharable, reusable knowledge bases. In *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning*, pages 601–602, 1991.

6. M. Gutiérrez, D. Thalmann, F. Vexo, L. Moccozet, N. Magnenat-Thalmann, M. Mortara, and M. Spagnuolo. An ontology of virtual humans: incorporating semantics into human shapes. In *Proceedings of Workshop towards Semantic Virtual Environments (SVE05), March 2005*, pages 57–67, 2005.
7. A. Guye-Vuilléme and D. Thalmann. Specifying mpeg-4 body behaviors. In *CA '02: Proceedings of the Computer Animation*, page 126, Washington, DC, USA, 2002. IEEE Computer Society.
8. J. A. Harrigan. Listener's body movements and speaking turns. In *Communication Research*, volume 12, pages 233–250, April 1985.
9. B. Hartmann, M. Mancini., and C. Pelachaud. Implementing expressive gesture synthesis for embodied conversational agents. In L. Springer, editor, *Inproc Gesture Workshop*, 2005.
10. M. Mancini, B. Hartmann, and C. Pelachaud. Non-verbal behaviors expressivity and their representation. Technical report, PF-star report 3.
11. nRQL. new racer query language (nrql) interface to owl ontologies in protégé. http://www.cs.concordia.ca/k̃_bhoopa/nrql.html.
12. Protégé. (c) 2005 stanford medical informatics. http://protege.stanford.edu/index.html.
13. RacerPro. Renamed abox and concept expression reasoner. http://www.racer-systems.com.
14. W3C. Owl web ontology language. http://www.w3.org/tr/owl-features/.
15. H. G. Wallbott. Bodily expression of emotion. *European Journal of Social Psychology*, 28(6):879–896, December 1998.
16. C. M. Whissel. *The dictionary of affect in language R. Plutchnik and H. Kellerman (Eds) Emotion: Theory, research and experience: vol 4, The measurement of emotions.* Academic Press, New York, 1989.