The Relation between Gaze Behavior and the Attribution of Emotion: An Empirical Study

Brent Lance and Stacy C. Marsella

University of Southern California
Information Sciences Institute
4676 Admiralty Way Suite 1001
Marina Del Rey, CA 90034
{lance,marsella}@isi.edu

Abstract. Real-time virtual humans are less believable than hand-animated characters, particularly in the way they perform gaze. In this paper, we provide the results of an empirical study that explores an observer's attribution of emotional state to gaze. We have taken a set of low-level gaze behaviors culled from the nonverbal behavior literature; combined these behaviors based on a dimensional model of emotion; and then generated animations of these behaviors using our gaze model based on the Gaze Warping Transformation (GWT) [9], [10]. Then, subjects judged the animations displaying these behaviors. The results, while preliminary, demonstrate that the emotional state attributed to gaze behaviors can be predicted using a dimensional model of emotion; and show the utility of the GWT gaze model in performing bottom-up behavior studies.

Keywords: Gaze, Nonverbal Behavior, Emotional Expression, Character Animation, Procedural Animation, Motion Capture, Posture, Virtual Agent.

1 Introduction

Animated characters in feature films function at a high level of believability, appear to come alive, and successfully engage the film's audience; as do characters in many video games, although arguably to a lesser extent. Unfortunately, virtual embodied agents struggle to achieve this goal. However, the animation methods, used to create the film and video game characters are expensive and time consuming, and only allow for limited interaction in dynamic environments, making them unsuitable for the development of virtual embodied agents. Instead, real-time animation systems that express believable behavior are necessary. Our specific interest is in gaze behavior, which is expressive not only in terms of where the gaze is directed, but also in how the gaze is performed, its physical manner. As such, the goal of this paper is to find a model that maps between emotion and the physical manner of gaze. The purpose of this mapping is to allow for the generation of believable, emotionally expressive gaze shifts in an interactive virtual human, while using the minimum amount of motion capture data necessary to maintain realistic physical gaze manner.

We present an approach to this problem that consists of an exploratory empirical study of the mapping between a set of gaze behaviors and the emotional content

H. Prendinger, J. Lester, and M. Ishizuka (Eds.): IVA 2008, LNAI 5208, pp. 1–14, 2008.

attributed to gaze shifts performing those behaviors by observers. This study is similar to the "reverse engineering" approach used by Grammer et al. [7], to study emotional state and facial expression. In this context, "reverse engineering" is used to mean a bottom-up approach where nonverbal behavior expressions are generated through the combination of low-level physical behaviors, and then displayed to subjects who rate the expression on its emotional content. Specifically, Grammer et al. [7], use Poser to generate random facial expressions from the space of all possible combinations of FACS Action Units. Users then evaluated the resulting expressions using a circumplex model of emotion.

Similarly, we found a model that describes the mapping between gaze behaviors and the attribution of emotion to gaze shifts displaying those behaviors by first determining how the model describes emotion. We used two different representations of emotion, a set of emotional categories, such as anger or fear, and the PAD dimensional model of emotion [13]. Then we determined the space of possible gazes and the physical manners which they perform. To do this, we have culled a set of low-level, composable gaze behaviors from the nonverbal behavior literature, such as a bowed head during gaze. We then generated all possible gazes allowed by our space of low-level behaviors using our model of expressive gaze manner based on the Gaze Warping Transformation (GWT) [9], [10].

We use this model because it is capable of displaying an arbitrary selection of gaze behaviors while directed towards an arbitrary target with a minimal library of motion capture data. In [9], we first described and evaluated the Gaze Warping Transformation (GWT) a method for producing emotionally expressive head and torso movement during gaze shifts. We then provided a neuroscience-based eye model, and integrated it with GWTs [10].

Finally, we collected data to determine what emotional states subjects attributed to animated characters displaying these behaviors during gaze shifts. As a result of this reverse engineering study, we were able to demonstrate that composition of these low-level gaze behaviors preserved the PAD dimensional ratings. These results, while promising, are still preliminary. However, the study clearly demonstrates the utility of the GWT as a research tool beyond generating animations, and points out several areas for future research.

While these results have the most application to our GWT-based gaze model, any procedural gaze model with sufficient control over the animation curves used to generate gaze shifts should be able to take advantage of this mapping.

2 Related Work

There have been many implementations of gazing behaviors in real-time applications such as embodied virtual agents. Several of these gaze implementations in virtual characters are based on communicative signals (e.g. [2], [16]). Other gaze models have been developed for agents that perform tasks in addition to dialog, such as [6], [17]. There are also models of resting gaze, which simulate eye behavior when the eye is not performing any tasks [4] [11]. Additionally, there are attention-based models of gaze that perform eye movements based on models of attention and saliency [18], [19].

There are several trends which can be seen in these implementations of gaze. First, the models focus on when and where the character looks, not on how the gaze shift occurs. Second, these models, with few exceptions, focus on communicative or task-related gaze behaviors, not on how gaze reveals emotional state.

In addition to the previous research on implementing models of nonverbal gazing behavior, there has been recent work focused on the manipulation of parameters describing the way in which general movement is performed. This concept is referred to as manner or style. This research can provide methods for manipulating the way in which movements are performed, or to obtain the style from one movement and transfer it to another [1], [3], [22]. This research was inspirational to the development of the Gaze Warping Transformation, but does not deal with the constraints specific to gaze movement, nor does it identify specific styles and their expressive meaning, which is the purpose of this study.

3 Expressive Gaze Model

We used our previous work on gaze to generate the gaze shifts for this study. Our gaze model combines two parts: first, a parameterization called the Gaze Warping Transformation (GWT), that generates emotionally expressive head and torso movement during gaze shifts [9]. The GWT is a set of parameters that transforms an emotionally neutral gaze shift towards a target into an emotionally expressive gaze shift directed at the same target. A small number of GWTs can then produce gazes displaying varying emotional content directed towards arbitrary targets.

The second part is a procedural model of eye movement based on stereotypical eye movements described in the visual neuroscience literature [10]. The procedural eye movement is layered framewise onto the GWT-generated body movement. Emotion is expressed using the GWT, while the procedural eye model ensures realistic motion.

3.1 Gaze Warping Transformation

A Gaze Warping Transformation, or GWT, is found by obtaining two motion captures of gaze shifts directed from the same start point to the same target, one emotionally expressive, the other emotionally neutral, and finding a set of warping parameters that would convert the animation curve representing each degree of freedom in the emotionally neutral animation into the animation curve for the corresponding degree of freedom in the emotionally expressive movement [9].

This works by transforming the keyframes of an animation curve. The keyframes of an animation are a subset of that animation's frames, such that the values of the motion curves for intermediate frames are found by interpolating between the keyframes. We select the keyframes for each gaze by aligning it to a "stereotypical" gaze shift with known keyframe locations [10]. The gazes are aligned using the ratio of movement that occurred by each frame to that throughout the entire curve [1].

The result of this is a set of keyframes $x(t)$, defined as a set of value, frame pairs, (x_i, t_i). These keyframes are transformed to those of a new motion $x'(t')$, defined as the set of pairs (x_i', t_i') through the use of two functions [21]. The first function, given a frame in the emotional curve t_i', calculates the frame t_i in the neutral motion curve to obtain the corresponding amplitude x_i. For the GWT, we use the function

$$t_i=g(t_i') \ , \tag{1}$$

$$g(t_i')=c(t_i')*(t_i'-t_{i-1}') \ , \tag{2}$$

where given a frame time in the emotional movement t_i', $g(t)$ determines the corresponding frame t_i in the neutral movement through a scaling parameter $c(t_i')$, which scales the time span between two adjacent keyframes. The second function is

$$x'(t_i')=x(t_i)+b(t_i) \ , \tag{3}$$

where $b(t_i)$ is a spatial offset parameter that transforms the neutral curve amplitude $x(t_i)$ into the corresponding emotional amplitude $x'(t_i')$. The final GWT is an $m * n$ set of (c, b) pairs, where m is the number of degrees of freedom in the animated body, and n is the number of keyframes in the animation.

As the GWT is based on a technique of simple geometric transformations [21], the generated animations can move outside the physical limits of a human body. To solve this, we use an inverse kinematics system implemented using nonlinear optimization. This system simulates a rigid skeleton, keeping our animated movement within the limits of the human body [10].

Table 1. List of Gaze Types

Gaze Type
Eye-Only Gaze Shift
Eye-Head Gaze Shift
Eye-Head-Body Gaze Shift
Head-Only Movement
Head-Body Movement

3.2 Procedural Model of Eye Movement

In addition to the GWT, which describes head and torso movement during gaze shifts, we developed an integrated procedural model of eye movement [10]. This model of eye movement is based on the visual neuroscience literature, specifically on research describing the different movements eyes perform during gaze, and the way in which eye movement and head movement are integrated during gaze shifts [12]. It generates several classes of gaze shifts (Table 1) using the following building blocks:

- Saccades. The saccade is a very rapid, highly-stereotyped eye movement which rotates the eye from its initial position directly to the target;
- Vestibulo-Ocular Reflex (VOR). Through the VOR, the eyes rotate within their orbit so that the gaze maintains the same target while the head moves. It produces the Head-Only and Head-Body movements; and
- Combined Eye-Head Movement. This is used to integrate eye movement and head-torso movement, and generates the Eye-Head and Eye-Head-Body gaze shifts;

4 Approach

We performed an empirical study to determine a preliminary mapping between a space of gaze behaviors and emotion attributed to the gaze behaviors by subjects. To obtain this mapping, we first selected appropriate emotional models and the space of gaze behaviors to map between. To determine the mapping between a particular gaze and the attribution of emotion to that gaze, we use a “reverse engineering” approach [7]. Specifically, we generate all the gazes allowed by our space of gaze behaviors, and collect data of subjects attributing emotion to these gaze shifts.

4.1 Structure of Model

Selected Emotion Model. There are many potential models of emotion we could have mapped to the gaze behaviors. We selected two: the first is the PAD model [13]; a model of emotion that views emotion as a space described with a three dimensions: pleasure / displeasure, arousal / non-arousal, and dominance / submissiveness.

The categories of emotion, such as anger or happiness, are represented in this model by subregions in the space defined by the emotional dimensions. For example, anger is defined as negative valence, high arousal, and high dominance, while fear is defined as negative valence, high arousal, and low dominance.

We are also using a categorization of emotion to map gaze behaviors to a set of intuitive emotional descriptors. Rather than using an existing categorical model, this categorization is derived from observer responses to the animations.

Table 2. Gaze Behaviors

Hypothesized Behaviors
Head Raised
Head Bowed
Faster Velocity
Slower Velocity
Torso Raised
Torso Bowed

Selected Gaze Behavior. In addition to the emotional model, we had to determine a space of gaze behaviors, due to the lack of a descriptive set of known gaze behaviors analogous to the FACS system. We identified a set of “emotional behaviors” from the psychology and arts literature that are likely to be used to reveal emotional state. This set of behavior guidelines can be seen in Table 2.

These guidelines are simplifications of the actual literature [5, 8]. Our guidelines are that users will view the character as more dominant when its head is turned upwards than when its head is turned downwards [15], that the perception of arousal is strongly related to velocity [9], and that vertical posture of the body will display emotional pleasure [20]. While there are many alternative gaze behaviors that could also be modeled using the GWT, such as subtle variations in dynamics, or wider variations on posture, this limited set provides a starting point for this research.

4.2 Motion Capture Collection

For the head and torso behaviors, we asked the actor to perform "raised," "neutral," and "bowed" versions of the behavior, and collected data from the resulting movement. We also collected "fast," "neutral," and "slow" velocity movements. However, the "raised" torso posture was indistinguishable from the neutral torso posture, due to the limitations of the motion tracking system we used, resulting in the set of physical behaviors shown in Table 3. All captured gaze shifts consisted of the desired behavior being displayed in a gaze aversion that started gazing straight ahead in a neutral position and posture, and ended gazing 30 degrees to the right displaying the intended gaze behavior. From this motion data, we produced eight behavior GWTs, one for each behavior listed in Table 3.

We also collected motion capture of the different gaze types (Table 1), and produced GWTs for each gaze type as well as the gaze behaviors. The gaze types were captured as gaze aversions that began gazing straight ahead and ended gazing 30° to the right, and gaze attractions that began 30 degrees to the right and ended gazing straight ahead. This resulted in 10 GWTs – one aversive and one attractive gaze shift for each of the different types of gaze in Table 1.

Table 3. Discretization of Gaze Behaviors

Behavior Dimension	Possible Values
Head Posture	Raised, Neutral, Bowed
Torso Posture	Neutral, Bowed
Movement Velocity	Fast, Neutral, Slow

4.3 Animation Generation

From these 8 GWTs representing the discretized physical behaviors (Table 3) and 10 GWTs representing the various gaze types (Table 1), we generated 150 animations for use in our empirical bottom-up study. We combined the gaze behaviors in Table 3 in all possible ways, leaving out combinations of a raised head with bowed torso due to the physical implausibility of the behavior, resulting in 15 total behavior combinations. Then, these combined gaze behaviors were applied to the 10 gaze type GWTs, resulting in 150 GWTs. Finally, to generate the animations, we applied these 150 GWTs to neutral gaze shifts, with the resulting output rendered using Maya. These animations can be seen at:
http://www.isi.edu/~marsella/students/lance/iva08.html

4.4 Category Formation

In order to determine the categories for our primary experiment, and obtain a picture of how well the animated gaze behaviors covered the emotional space defined by the emotion models, we performed a preliminary category formation study.

Approach. 31 informally selected people each viewed 20 animations randomly selected from the set of 150 animations with no duplicates, giving us 620 views, or approximately 4 per animation, and provided an open-ended written response to the question "What emotional state is the character displaying?" We then categorized the affective responses based on the hierarchical model of emotion described in [14].

Results. We used the hierarchical model as a sorting guideline, to divide the individual responses into ten categories (Table 4); for example categorizing “expression of contempt” as Contempt, or “terrified” as Fear. However, we utilized additional categories not described by the hierarchical model. After categorizing the responses, we then selected categories where at least one video had 50% of the subjects rate it with that category. We then discarded those categories that were related to attention, discarding responses such as “change in attention,” “displaying strong interest,” and “distracted.” Finally, we discarded the responses indicating “uncertainty,” as we were concerned that it would be applied when the subject was uncertain of the character’s state, not when the character was displaying uncertainty.

Table 4. Emotional Categories

Emotional Categories
Anger
Contempt
Disbelief
Excitement
Fear
Flirtatious
Guilt
Sadness
Secretive
Surprise

4.5 Emotional Attribution Experiment

After selecting the low-level behaviors, generating the animations, and setting the emotional categories, we performed the empirical study. The animations were placed online, and subjects rated the animation in two ways: first by selecting the emotional category (Table 4) that most closely approximated the emotion that they perceived in the animation, and second by locating the animation’s perceived emotion along the emotional dimensions of the PAD model. One hundred subjects selected through social networking rated fifteen unique, randomly selected animations each, resulting in ten ratings for each of the 150 animations. Subjects rated the animation’s location within the PAD model by using five-point Likert scales to indicate their agreement with two statements representing each dimension, seen in Table 5. The Likert scales were 1 = Strongly Disagree, 2 = Disagree, 3 = N/A, 4 = Agree, 5 = Strongly Agree. Emotional categories and rating statements were displayed in random order.

Table 5. Emotional Dimension Rating Scales

Emotional Dimension	Rating Statement
High Dominance	The character is dominant.
Low Dominance	The character is submissive.
High Arousal	The character is agitated.
Low Arousal	The character is relaxed.
High Valence	The character is pleased.
Low Valence	The character is displeased.

5 Results

We uncovered the mapping between emotion models and physical behaviors, in order to answer the following questions:

1. How did the subjects rate gaze shifts containing the low-level gaze behaviors in Table 3 along the PAD dimensions?
2. Does composition of low-level gaze behaviors in Table 3 preserve the PAD dimensions? For example, if a gaze shift displays low Dominance and low Pleasure behaviors, are low Dominance and low Pleasure attributed to it?
3. How did the subjects rate gaze shifts containing the low-level gaze behaviors in Table 3 using the emotional categories in Table 4?

While we had originally intended to find which of the 150 individual animations varied across emotional state, ten ratings per animation was too few to perform a reliable statistical analysis. Instead, we combined the gazes across gaze type (Table 1), giving us 50 ratings for each of the 15 combinations of gaze behaviors.

5.1 Dimensional Results

How reliable were the dimensional ratings scales?
Before exploring the dimensional results, we tested how well our dimensional rating scales measured the emotional dimensions they were intended to by calculating the correlation and Cronbach's Alpha between each pair of rating scales from Table 5.

The Pleased and inverted Displeased scales performed well. The correlation between the two was 0.615, and the standardized Alpha score indicating scale reliability was high, with $\alpha = 0.7610$, ($\alpha > 0.7$ is considered a reliable scale). Dominant and inverted Submission also did well, with a correlation of 0.6649, and a high Alpha ($\alpha = 0.7987$). Therefore, we averaged Pleased and inverted Displeased into one Pleasure scale, and combined Dominant and inverted Submission into one Dominance scale.] Correlations between the Dominance and Pleasure scales were low, (0.1569), indicating no overlap.

However using the ratings of Relaxed and Agitated as a scale for Arousal was less reliable, as both correlation (0.3745) and Alpha ($\alpha = 0.5449$) were low. In addition, correlations between Relaxed and Pleased (0.5353) and between Agitated and Displeased (0.4889) were higher than between Relaxed and Agitated. There are several possible explanations for this, and further research will be necessary to determine the actual reason, but for the remainder of this paper, we will be using the two scales separately as Relaxed and Agitated.] As we used 5-point Likert scales, but only animated 3-point scales of physical behavior, we condensed the collected data into 3-point scales by combining "Strongly Disagree" and "Disagree", as well as "Strongly Agree" and "Agree", leaving Neutral ratings unchanged.

How did the subjects rate gaze shifts containing the low-level gaze behaviors in Table 3 along the PAD dimensions?
To answer this question, we performed a series of MANOVAs (multivariate analysis of variance) and t-tests to determine whether or not the mean emotion dimensions ratings differed across to the low level behaviors found in Table 3. Results of this analysis can be seen in Table 6.

Table 6. Significant Relationships between PAD Dimension and Gaze Behaviors

Emotional Dimension	Head	Body	Velocity
High Dominance	Raised	Bowed	Fast
Low Dominance	Bowed	Neutral	Non-Fast
Relaxed		Bowed	
Agitated	Non-Bowed		Fast
High Pleasure	Neutral	Bowed	
Low Pleasure	Non-Neutral	Neutral	

Four MANOVAs were performed, each with one dimension (Pleasure, Agitation, Relaxation, or Dominance) as the dependent variable, and Head Orientation, Torso Orientation, Velocity, and Subject as the independent variables, while testing for second degree factorial interaction between the independent variables.

The MANOVA results for Dominance showed significant effects ($N = 1500$, $DF = 18$, $F = 14.51$, $p < .001$) for head orientation ($F = 24.08$, $p < .001$), torso orientation ($F = 82.55$, $p < .001$), and velocity ($F = 7.38$, $p < .001$), with a significant interaction between head and torso orientation ($F = 6.47$, $p < .05$). The t-tests showed clear differences between group means, with raised head corresponding to higher Dominance, and bowed head with lower. In addition the t-tests revealed that a bowed posture was rated higher than a neutral posture, and that the Dominance rating for fast was higher than for slow or for neutral (all significant to $p < .01$).

The Relaxed results showed significant differences ($N = 1500$, $DF = 18$, $F = 1.89$, $p < .05$) across the torso orientation ($F = 11.41$, $p < .001$) and the velocity ($F = 3.78$, $p < .05$) , with a significant interaction effect between torso and velocity ($F = 3.68$, $p < .05$); and the t-tests revealed that a bowed body was rated more Relaxed than a neutral body ($p < .01$). However, the t-tests did not reveal useful information about the velocity, indicating that the significant difference found by the MANOVA was likely related to the interaction between torso and velocity.

The MANOVA for Agitation found significant differences (N = *1500*, DF = *18*, $F = 4.60$, $p < .001$) across the head orientation ($F = 19.61$, $p < .001$), the velocity ($F = 6.04$, $p < .01$), and the subject ($F = 17.12$, $p < .001$), and a significant interaction effect between the head and the velocity ($F = 7.17$, $p < .05$). The t-tests showed that raised and neutral head were rated as significantly more Agitated than bowed head, and the rating for high velocity was higher than for slow or neutral ($p < .05$).

Finally, the analysis revealed that Pleasure significantly differed ($N = 1500$, $DF = 18$, $F = 5.93$, $p < .001$) across both the vertical orientation of the head ($F = 6.58$, $p < .05$) and the torso ($F = 77.57$, $p < .001$), with no significant interaction effects. The ratings for Pleasure also differed significantly ($F = 4.06$, $p < .05$) across subject. T-tests ($p < .01$) showed that the Pleasure rating for a neutral head orientation was significantly higher than those for bowed and raised head orientations, and that a bowed posture was rated higher than a neutral posture.

Does composition of low-level gaze behaviors preserve the PAD dimensions?

In order to determine whether the low-level behaviors can be combined according to the PAD model of emotion, we performed a second analysis. We performed six MANOVAs, each using an emotional dimension (High Dominance, Low Dominance, Relaxed, Agitated, High Pleasure, and Low Pleasure) as the dependent variable. We then used the number of behaviors associated with that emotional dimension, and the subject as the two independent variables. This tested whether or not gaze shifts displaying different numbers of behaviors attributed to a specific emotional dimension would have different values attributed to them. The results of this analysis showed that mean attributed ratings for an emotional dimension increased as the number of gaze behaviors associated with that emotional dimension increased, as seen in Figure 1. This indicates that physical gaze behaviors, when combined according to PAD dimensions will be rated as predicted by the combined behaviors.

The specific results for dominance show significant differences ($N = 1500$, $DF = 6$, $F = 32.24$, $p < .01$) across the number of both low ($F = 14.17$, $p < .001$) and high ($F = 26.39$, $p < .001$) dominance behaviors displayed in a gaze shift, and a significant interaction effect ($F = 6.93$, $p < .01$) between low and high dominance. T-tests showed that as the number of dominance gaze behaviors increased, the rating of dominance significantly increased ($p < .01$) for high dominance behaviors, and significantly decreased ($p < .05$) for low dominance behaviors.

The MANOVA for Agitated revealed significant differences ($N = 1500$, $DF = 3$, $F = 18.31$, $p < .001$) across the number of behaviors displayed in a shift, and showed significant differences across subjects ($F = 20.50$, $p < .001$), with no interaction effects. T-tests demonstrated that gaze shifts with no Agitated behaviors were rated significantly less agitated than those with 1 or 2 Agitated behaviors ($p < .01$).

Both low and high pleasure showed significant differences across the number of behaviors ($N = 1500$, $DF = 3$, $F = 22.96$, $p < .001$), although there were also significant differences across subjects ($F = 4.87$, $p < .05$), and no interaction effects. Subsequent t-tests showed that mean ratings of pleasure significantly differed ($p < .01$) across all numbers of pleasure-associated behaviors, and that as low pleasure behaviors increased, pleasure decreased and vice versa for high pleasure behaviors.

As the relaxed dimension only had one behavior associated with it, no further testing was performed.

5.2 Categorical Results

How did the subjects rate gaze shifts containing the low-level gaze behaviors in Table 3 using the emotional categories in Table 4?

To answer this question, we generated a cross tabulation of the 15 combinations of gaze behaviors against the emotional categories (Table 4), and used Pearson's chi squared (X^2) test to examine relationships in the data. We then performed further tests on the residuals to determine which had significant differences.

Table 7. Emotional Categories and Significantly Related Behavior Combinations

Emotional Categories	Significantly Related Behavior Combinations
Contempt	Head Raised, Body Neutral, Velocity Neutral
Excitement	Head Neutral, Body Bowed, Velocity Fast
Fear	Head Neutral, Body Neutral, Velocity Neutral Head Neutral, Body Neutral, Velocity Slow
Guilt	Head Bowed, Body Neutral, Velocity Neutral Head Bowed, Body Neutral, Velocity Slow
Sadness	Head Bowed, Body Neutral, Velocity Fast Head Bowed, Body Neutral, Velocity Neutral

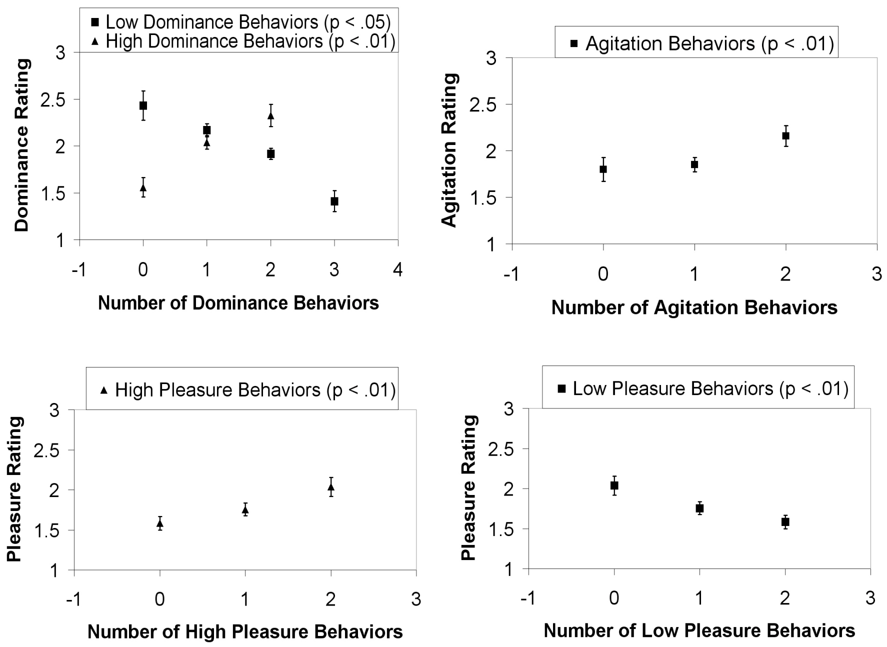

Fig. 1. Plots of Mean Ratings vs. Number of Behaviors for Dominance, Pleasure and Agitation

Results of this analysis can be seen in Table 7. The X^2 test showed that gaze combinations and emotional categories were not randomly related (N = 1500, DF = 126, $X^2 = 775.817$, $p < .01$). The table rows show behavior combinations with a significant number ($p < .05$) of ratings for that emotional category.

While only 5 of the 15 gaze behavior combinations had significant associations to emotional categories, it was clear through examination of the residuals that further analysis of the relationship between the emotional categories and the low-level behaviors from Table 3 could be useful. For example, while no individual gaze behavior combination was rated significantly high for Flirtatious, all gaze shifts with the bowed head behavior had more Flirtatious ratings than did the gaze shifts without bowed head. To examine this, we generated crosstabs of individual gaze behaviors against emotional categories, and performed additional X^2 tests (Table 8).

Table 8. Significant Relationships between Emotional Categories and Gaze Behaviors

Emotional Category	Head	Torso
Contempt	Raised	Neutral
Excitement		Bowed
Fear		Neutral
Flirtatious	Bowed	
Guilt	Bowed	Neutral
Sadness	Bowed	
Surprise	Neutral	

We found significant interactions between head vertical orientation and emotional categories, (N = 1500, DF = 18, $X^2 = 329.47$, $p < .001$). Testing the residuals showed that the Contempt category was more likely ($X^2 = 70.35$, $p < .05$) to be attributed to a gaze shift with the head raised behavior, while Flirtatious ($X^2 = 73.41$, $p < .01$), Guilt ($X^2 = 81.33$, $p < .01$), and Sadness ($X^2 = 42.51$, $p < .01$) were all more likely to be attributed to bowed head gaze shifts. Finally, Surprise was significantly less likely ($X^2 = 55.30$, $p < .01$) to be attributed to bowed head gazes. Anger, Disbelief, Excitement, Fear, and Secretive do not relate to head vertical orientation significantly.

Torso posture was not randomly related to emotional category ($N = 1500$, $DF = 9$, $X^2 = 187.49$, $p < .001$). Excitement was more likely to have a bowed torso ($X^2 = 62.94$, $p < .01$), while Contempt ($X^2 = 24.24$, $p < .05$), Fear ($X^2 = 29.19$, $p < .01$), and Guilt ($X^2 = 19.88$, $p < .01$) were attributed more often to neutral torso animations.

We also found, despite our expectations, no strong relationships between the emotional categories and the velocity of the gaze using a crosstab of emotions by velocity. While, the chi squared test showed that emotional category and velocity are not randomly related ($N = 1500$, $DF = 18$, $X^2 = 42.36$, $p < .001$), upon examination of the residuals, no emotional categories significantly differed across velocities.

6 Discussion

As a result of this reverse engineering study, we were able to demonstrate that composition of low-level gaze behaviors in Table 3 preserved the PAD dimensions (Figure 1). In addition, this preservation of PAD dimensions through composition can even be extended to some emotional categories. For example, Guilt can be mapped into the PAD space as low Pleasure, low Arousal, and low Dominance. By combining the behaviors associated with low Dominance, low Pleasure, and Relaxation (Table 6) we generate a movement with a bowed head, neutral torso, and slow velocity. Table 8 reveals that Guilt is attributed to gaze shifts displaying a bowed head, neutral torso, and slow velocity, just as predicted by the PAD model.

We also have a partial mapping between emotional categories and gaze behaviors. For example, in Table 7, we can see that the gaze categories Contempt, Excitement, Guilt, and Sadness are all clearly associated with specific combinations of gaze behaviors, although there is some overlap between Guilt and Sadness. In addition, through examining the relationships between individual low-level gaze behaviors and emotional categories (Table 8), we obtain reinforcement of this mapping, and additional information about how Flirtatious behavior is displayed as well.

This indicates that virtual embodied agents with disparate models of emotion should be able to make use of this mapping between gaze behaviors and attributed emotional state. If the agent uses a categorical model of emotion with emotional categories beyond those used in this study, then by mapping those categories into the PAD space, appropriate gaze behavior may still be generated.

A more thorough description of the low-level behavior components for gaze, similar to FACS for facial expression, would be very valuable to this type of research. While we determined our own space of gaze behaviors for this study, there are other possible ways to structure the gaze behavior space that may provide better results.

7 Conclusion

In this paper, we have provided the results of a reverse engineering study resulting in a preliminary mapping between gaze behaviors and emotional states that could be used with a variety of gaze or emotion models. In addition, we have shown that combining low-level behaviors associated with emotional dimensions in accordance with those dimensions generates a gaze shift that subjects attribute the combined emotional state to. These results, while promising, are still preliminary. However, this study demonstrates the utility of the GWT as a nonverbal behavior research tool, and points towards several directions for future research.

Many of the results of the mapping were not surprising, such as the link between increased dominance and increased vertical orientation of the head. However, there were unexpected results; for example, the link between high Pleasure and a bowed forward body. This indicates the need for a broader selection of gaze behaviors to determine why these unexpected results occurred. This work would also benefit from a more complete exploration of the way in which emotional state is attributed to different combinations of head and eye behavior, as well as a real-time implementation of the gaze mapping in a virtual embodied agent for evaluation.

Acknowledgements. Dr. Bosco Tjan, Dr. Skip Rizzo, and Tiffany Cole provided invaluable assistance, without which this paper would not have been possible. This work was sponsored by the U.S. Army Research, Development, and Engineering Command (RDECOM), and the content does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

References

1. Amaya, K., Bruderlin, A., Calvert, T.: Emotion From Motion. In: Proceedings of the Conference on Graphical Interface, pp. 222–229 (1996)
2. Bickmore, T., Cassell, J.: Social Dialogue with Embodied Conversational Agents. In: Bernsen, N. (ed.) Natural, Intelligent and Effective Interaction with Multimodal Dialogue Systems. Kluwer Academic Publishers, Dordrecht (2004)
3. Brand, M., Hertzmann, A.: Style Machines. In: Proceedings of SIGGRAPH 2000 (2000)
4. Deng, Z., Lewis, J.P., Neumann, U.: Automated Eye Motion Using Texture Synthesis. IEEE Computer Graphics and Applications 25(2) (2005)
5. Exline, R.: Visual Interaction: The Glances of Power and Preference. In: Weitz, S. (ed.) Nonverbal Communication: Readings with Commentary, Oxford University Press, Oxford (1974)
6. Gillies, M.F.P., Dodgson, N.A.: Eye Movements and Attention for Behavioral Animation. The Journal of Visualization and Computer Animation 13, 287–300 (2002)
7. Grammer, K., Oberzaucher, E.: Reconstruction of Facial Expressions in Embodied Systems. Mitteilungen, ZiF (2006)
8. Kleinke, C.: Gaze and Eye Contact: A Research Review. Psychological Bulletin 100(1) (1986)
9. Lance, B., Marsella, S.: Emotionally Expressive Head and Body Movement During Gaze Shifts. In: Pelachaud, C., et al. (eds.) IVA 2007. LNCS (LNAI), vol. 4722. Springer, Heidelberg (2007)

10. Lance, B., Marsella, S.: A Model of Gaze for the Purpose of Emotional Expression in Virtual Embodied Agents. In: Padgham, Parkes, Müller, Parsons (eds.) Proc. of 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008) (2008)
11. Lee, S., Badler, J., Badler, N.: Eyes Alive. ACM Transactions on Graphics 21(3) (2002)
12. Leigh, R.J., Zee, D.: The Neurology of Eye Movements, 4th edn. Oxford Press (2006)
13. Mehrabian, A.: Silent Messages: Implicit Communication of Emotions and Attitudes, 2nd edn. Wadsworth Publishing Company (1981)
14. Metts, S., Bowers, J.: Emotion in Interpersonal Communication. In: Knapp, M., Miller, G. (eds.) Handbook of Interpersonal Communication, 2nd edn. Sage, Thousand Oaks (1994)
15. Mignault, A., Chaudhuri, A.: The Many Faces of a Neutral Face: Head Tilt and Perception of Dominance and Emotion. Journal of Nonverbal Behavior 27(2) (2003)
16. Pelachaud, C., Bilvi, M.: Modeling Gaze Behavior for Conversational Agents. In: Rist, T., Aylett, R., Ballin, D., Rickel, J. (eds.) IVA 2003. LNCS (LNAI), vol. 2792. Springer, Heidelberg (2003)
17. Rickel, J., Johnson, W.L.: Animated Agents for Procedural Training in Virtual Reality: Perception, Cognition, and Motor Control. Applied Art. Int. 13(4-5) (1999)
18. Peters, C., Pelachaud, C., Bevacqua, E., Mancini, M.: A Model of Attention and Interest Using Gaze Behavior. In: Panayiotopoulos, T., Gratch, J., Aylett, R.S., Ballin, D., Olivier, P., Rist, T. (eds.) IVA 2005. LNCS (LNAI), vol. 3661. Springer, Heidelberg (2005)
19. Picot, A., Bailly, G., Elisei, F., Raidt, S.: Scrutinizing Natural Scenes: Controlling the Gaze of an Embodied Conversational Agent. In: Pelachaud, C., et al. (eds.) IVA 2007. LNCS (LNAI), vol. 4722. Springer, Heidelberg (2007)
20. Schouwstra, S., Hoogstraten, H.: Head Position and Spinal Position as Determinants of Perceived Emotional State Perceptual and motor skills 81(22), 673–674 (1995)
21. Witkin, A., Popovic, Z.: Motion Warping. In: Proceedings of SIGGRAPH 1995 (1995)
22. Zhao, L., Badler, N.: Acquiring and Validating Motion Qualities from Live Limb Gestures. Graphical Models 67(1) (2005)