Emotionally Expressive Head and Body Movement During Gaze Shifts

Brent Lance and Stacy C. Marsella

University of Southern California
Information Sciences Institute
4676 Admiralty Way Suite 1001
Marina Del Rey, CA 90292
{brent,marsella}@isi.edu

Abstract. The current state of the art virtual characters fall far short of characters produced by skilled animators. One reason for this is that the physical behaviors of virtual characters do not express the emotions and attitudes of the character adequately. A key deficiency possessed by virtual characters is that their gaze behavior is not emotionally expressive. This paper describes work on expressing emotion through head movement and body posture during gaze shifts, with intent to integrate a model of emotionally expressive eye movement into this work in the future. The paper further describes an evaluation showing that users can recognize the emotional states generated by the model.

1 Introduction

The manner of a persons gaze, how it is performed, reveals much about their inner state and intent. In fact, a key role of gaze in human interaction is to express the feelings and attitudes of the individual gazing. This role is revealed in the rich vocabulary used to describe a person's gaze. Phrases such as glare, gawking, furtive glance, etc. all indicate different ways of looking based on the emotional and cognitive state of the individual gazing. In this research, we are interested in how to create a virtual human capable of revealing its emotional states through the manner of its gaze behavior. We define the manner of gaze behavior as changes in physical parameters of individual movements, such as the velocity of the head in a single gaze shift, as opposed to changes in specific properties such as the target or time of occurrence of a gaze shift.

While there are many potential influences on gaze behavior [2], [10], we will only be looking at a subset of these. Specifically, we will be looking at how emotion affects the manner of gaze behavior. We have chosen to examine emotional factors for a number of reasons. In our previous work on gaze behavior manner [13], we found a greater association between physical parameters and emotion than between physical parameters and the other gaze-affecting factors that we examined, such as speech-related gaze shifts. More importantly, we found that the manner of gaze is a highly expressive signal. Despite findings from psychological research that show the importance of gaze manner in displaying emotion [2], [12], its recognized importance

C. Pelachaud et al. (Eds.): IVA 2007, LNAI 4722, pp. 72–85, 2007.

in animation [23], and our own findings [13], little work in the virtual humans community has been done on using gaze manner to express emotion.

A key challenge that arises in developing such a model of gaze manner is that gaze is not simply eye movement. Gaze is a complex of behaviors that can include eye, head, posture, and even stepping or standing, and all of these components must be taken into account. In addition, these behaviors are not independent from each other [16]. The importance of both appropriate manner for gazing behaviors and of appropriate physical interrelations between distinct body components can be seen in a number of virtual human designs, as well as in some computer graphic animated films. The result is that independent behaviors seem robotic and unnatural, and their relationship appears random and disjointed. This effect is immediately disconcerting to the viewer. As an example, consider the film "The Polar Express," in which the characters were animated through motion capture, except for the eyes, which were separately hand-animated. One reviewer noted "Although the human characters look about 90% lifelike, it is that darn 10% (mostly the lifeless eyes) that winds up making them seem really creepy," [22]. Other reviewers agreed, describing unnatural eyes, and referring to the animated characters as "zombies" or "creepy."

This paper addresses the problem of emotionally expressive gaze manner by describing a preliminary approach for expressing emotion during gaze behaviors. This approach is described as follows: first, recordings of head, eye, and body movement are made of actors performing emotionally expressive and emotionally neutral gaze shifts. Then, parameters describing how the emotionally expressive gaze shifts differ from neutral gaze shifts are extracted. This parameterization is then applied to emotionally neutral gaze shifts at different targets from the emotional gaze shifts, transferring the physical properties of the emotional gaze to the neutral gaze. Finally, the emotional content of these generated shifts are evaluated. While this approach only describes a subset of a dimensional model of emotion and does not address the problem of emotionally expressive eye movement, the preliminary results described here show promise for future research.

While it may seem counterintuitive to discuss generating gaze while not discussing the generation of eye movement, we believe that modeling eye movement and modeling movement of the head and body require different approaches, due to the physical differences between the two types of movement. We are currently developing a model of emotionally expressive eye movement, derived in part from the eye data collected while performing this work, which will be integrated with this work once it is completed.

2 Related Work

Gaze has many uses in human interaction. Much research has been done on how gaze regulates interaction between individuals, as well as the use of gaze to signal communicative acts [10]. However, in this work, we are more interested in how gaze is used to display relational attitudes and affective states, specifically dominance/ submissiveness, arousal/relaxation, and pleasure/displeasure. Dominance is a signal sent through gaze [7], head movement [18], and posture [5]. For example, displaying increased gaze while speaking, a raised head, and upright posture all signal dominance, while the opposite behaviors signal submission.

Arousal is also closely related to gaze [2]. While there has been little work on how arousal and relaxation are specifically related to head movement, velocity has been shown to be an indicator of arousal [19]. Evidence for the relationship of gaze to the display of pleasure/displeasure is more limited. In fact, as reported in [12], some have argued that gaze is incapable of displaying emotional valence. However, the head [18], and the body [6], have both been shown to reveal pleasure/displeasure, although there is overlap with the behaviors which reveal dominance/submissiveness.

There have also been many implementations of gaze behaviors in real-time applications, such as Embodied Conversational Agents. Many of these implementations are based on communicative signals, such as [3], [20]. Other models of gaze have been developed for agents that perform tasks, interacting with an environment instead of with other characters or users [21]. Further models have simulated resting gaze, when the eye is performing no other tasks [14], or models of gaze based on realistic models of saliency [11].

In addition to models of gazing behavior, there has also been work focused on the manipulation of parameters describing the way in which movement is performed. This concept is referred to as “manner” or “style.” One of the primary works on style uses what are called “style machines,” combinations of statistical Hidden Markov Models, to allow for easily modifying a style, or learning a style, such as “angry,” from one movement and applying it to another movement [4]. Other style research includes applying different styles to walking behaviors [15], or using style to express emotion, although through gesture instead of gaze, as described in [25], and [1]. The research in [1] employs a similar technique to ours, but focuses on a simple door-knocking movement, as opposed to our work, which focuses on the greater emotional expressivity possible through gaze.

Despite the numerous models of gaze in virtual agents, and the work done on transferring manner from one movement to another, there currently has been no exploration of how changes in emotional state affect changes in the manner of gazing behavior. This work is intended to begin this exploration.

3 Approach

Our approach to realizing a model of emotionally expressive head and body movement during gaze shifts is based on deriving a Gaze Warping Transformation (GWT), a combination of temporal scaling and spatial transformation parameters that describe the manner of an emotionally expressive gaze shift. This transformation, when applied to an emotionally neutral gaze shift (created procedurally or through motion capture); will modify that neutral gaze shift into one which displays the same emotion as the original shift. A small number of transformations would then be used to produce gazes displaying different emotional content that vary in the directionality of the gaze. It is currently unclear to what extent emotional expression can be transferred between different categories of gaze. For example, if an individual is interacting with another individual on a catwalk high above her, how is her gaze behavior different from that of two people speaking face to face? Due to this, the scope of this paper is the display of a few emotional dimensions through head and body movement during gaze attractions and gaze aversions in face-to-face interaction.

In order to find a GWT we first use psychological research into expressive gaze manner to generate a series of guidelines describing how emotional state affects gaze behavior. These guidelines are provided to actors, whose performances of the behaviors result in three sets of collected motion capture data. The first set consists of emotionally expressive gaze shifts and emotionally neutral gaze shifts directed at a single target. From this data, we derive the GWT. This transformation is then applied to the second set of motion capture data, which consists of emotionally neutral gaze shifts averting from that target, transferring expressive manner to the neutral gazes. Finally, animations are generated from these modified gaze shifts, and compared to a set of emotionally expressive gaze shifts collected for evaluation.

3.1 Emotion Model

The first step in this procedure is to construct a set of guidelines for how emotional state affects expressive gaze manner. In order to do this, a model of emotion is used as a framework for the gaze behavior. We are using the model of emotion described in [17], which is a dimensional model of emotion, one that views the set of emotions as a space described with a small number of dimensions.

This model of emotion is called the Pleasure-Arousal-Dominance, or PAD model, which are the three emotional dimensions comprising the model. The intuitive categories of emotion, such as anger, fear, or happiness, are represented in this model by subregions in the space defined by the emotional dimensions.

For example, anger is defined as negative pleasure, positive arousal, and positive dominance (see [9] for a categorization of the PAD model for use in a computational model of emotion).

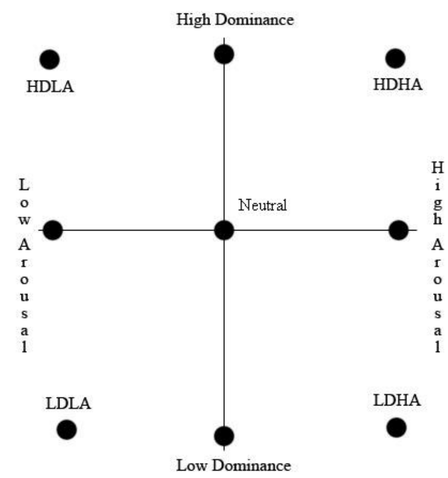

Fig. 1. Dimensional Model of Emotion

There are many alternative emotional models, such as models where emotions are viewed as discrete categories, or appraisal models, which use appraisal variables to define emotional states. We chose the PAD model because it is composed of a small number of dimensions, each of which have a background of research describing how gaze behaviors vary along the dimension [2], [7], [12]. This will allow us to develop a model of how expressive gaze behavior of a character varies based on the location of that character's emotional state in the PAD space.

We have so far used only the Arousal and Dominance dimensions of the model, dividing each dimension into discrete high and low values. This gives us nine distinct regions of the emotional space: Neutral Dominance and High Arousal (NDHA), Neutral Dominance and Low Arousal (NDLA), Neutral Arousal and High Dominance (NAHD), Neutral Arousal and Low Dominance (NALD), High Arousal and High Dominance (HAHD), High Arousal and Low Dominance (HALD), Low Arousal and High Dominance (LAHD), Low Arousal and Low Dominance (LALD), and the emotionally neutral origin, as shown in Fig. 1.

3.2 Data Description

The data for this work was collected by motion capture from gaze shifts performed by an actor who was provided with the set of performance guidelines shown in Table 1. For the basis of these guidelines, we use the findings that coders will rate an individual with upright head and posture as more dominant than one slouching forwards [5], [18], as well as results showing that velocity is clearly identified with arousal (for example in gesture [19]). Our interest is in arousal and dominance as signals, and thus the focus here is on those physical properties that are reliably decoded by observers, as opposed to how dominance and arousal are actually encoded in behavior, which is a more complex relationship [7].

Table 1. Predicted Physical Behaviors

Emotional Dimension	Behavior Guidelines
High Dominance (NAHD)	Upright Posture Head Turns Upwards Face Is Towards Other Individual
Low Dominance (NALD)	Hunched Forward Posture Head Turns Downwards Face Is Turned Away from Other Individual
High Arousal (NDHA)	Faster Movement Increased Blink Rate Body Moves Forward Slightly
Low Arousal (NDLA)	Slower Movement Decreased Blink Rate Body Moves Backward Slightly

Using these guidelines, the three sets of motion capture data are collected. For the first set, an actor is provided with the guidelines and demonstrates these behaviors while performing four to six gaze shifts from a target 90 degrees to the side of the actor to another target in front of the actor. The actor also performs this same gaze shift while displaying no emotion. The GWT is drawn primarily from this set.

The second set of motion capture data consists of emotionally neutral gaze shifts at different targets. For this, the actor also performs gaze shifts at three targets - one directly in front of the actor, one twenty degrees to the side of the actor, and one 45 degrees to the side - while demonstrating no predicted behaviors, and attempting to display no emotion. These are intended to simulate different levels of gaze aversion, where the character looks away from the user. These are the gaze shifts that will be converted into emotional gaze shifts using the GWT. This means that we will be emotionally transforming these aversion gaze shifts based on GWT's drawn from attraction gaze shifts, where the character looks at a user, in order to test the capability of the GWT's. Finally, the third set of data consists of the actor performing the combined behaviors for HAHD, HALD, LAHD, and LALD while shifting gaze from a target 90 degrees to the actor's side to a target to their front. This set provides additional data for determining the GWT, and will be compared to the converted emotional gaze shifts for evaluation purposes.

Recordings of the position of the head, body, and eyes are made throughout the gaze shifts. The data is collected as a set of time-series data points from three motion sensors: one on the head, one at the base of the neck, and one at the base of the spine. Each of the sensors records a time stamp, along with the position of the sensor in (x,y,z) coordinates, where the x axis runs laterally, with regard to the actor, the y axis runs from the actor's back to their front, and the z axis records vertical movement. This is followed by the orientation of the sensor as an Euler angle, resulting in eighteen total recorded degrees of freedom (DOF). Although the work reported here focuses on the head and body movements, the eyes were also recorded as an (x, y) coordinate representing the eye's angular position. Each DOF is represented as a separate two-dimensional curve (x_i, t_i), with the value of the sensor (x_i) plotted against the time of the reading (t_i). The head and body movement are captured with Ascension Flock of Birds electromagnetic sensors, while the eye movement is captured with an Applied Science Laboratories H-6 head-mounted eye tracker.

After capturing the data, it needs to be evaluated in order to avoid developing a model of gaze manner based on incorrect expectations. It is possible that our guidelines as to how emotional dimensions can be decoded from gaze behaviors are incorrect, or that the actor's performance is lacking, leading to an unclear emotional display, or a different emotion being displayed. For the evaluation, animations generated directly from first set of motion captures were displayed in pairs to eleven coders, who rated each animation individually on 5-point Likert scales of dominance and arousal. Each coder saw each animation three times.

We then performed two analyses of variance (ANOVAs), comparing the ratings of animations along the emotion scales. Ratings of NALD animations on the dominance scale ($M = 2.515$) were significantly different from ratings of NAHD animations ($M = 3.727$), and from ratings of animations neutral on the dominance scale ($M = 3.348$), $F(2, 131) = 12.1474$, $p < .01$, although NAHD animations were not significantly different dominance-neutral animations.

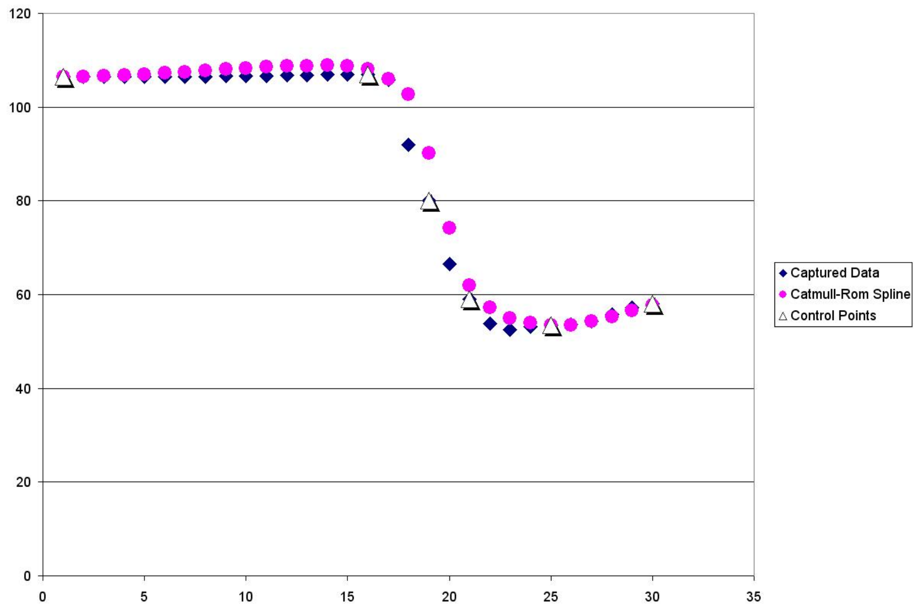

Fig. 2. Comparison of Data to Spline

Similarly, ratings of NDHA animations on the arousal scale (M = 3.909) were significantly different from ratings of NDLA animations (M = 2.727) and from animations neutral on the arousal scale (M = 2.803), $F(2, 131) = 12.2311$, $p < .01$, though NDLA and arousal-neutral animations did not significantly differ.

Because coders differentiated between low and high values for both dominance and arousal, and the neutral animations for each scale fell between low and high, we drew our GWTs from these animations. Had the coders been unable to distinguish either scale, we would have revised our behavior guidelines, and collected new data.

3.3 Gaze Warping Transformation

After evaluating the motion capture data, the GWT was derived from the first and third motion capture sets, which consisted of four to six motion captures of the actor looking from a target 90 degrees to their left to a target directly in front of them, for each point in the emotional space (Fig. 1), including an emotionally neutral gaze. The GWT is found through the following process:

1. The representation of animation curves in each gaze shift is changed to a spline, effectively down-sampling the curve to a set of control points
2. The point-to-point transformations between down-sampled emotionally expressive and emotionally neutral gaze shifts are found.
3. The transformations that significantly differ along emotional dimensions are found.
4. The GWT for an emotional category is assembled from the point-to-point transformations for an entire DOF where a plurality of the transformations for that DOF significantly differ along the emotional dimensions comprising that category.

First, we represent our data with cubic interpolating splines: parametric cubic functions defined using a set of control points. We do this because the spline functions as smoothing, cleaning up noise and outliers. It also effectively down-samples the data, giving us a sparser representation for determining the GWT, while keeping some of the time variability of the data. Finally, it allows the GWTs to be easily applied to new gaze shifts.

While splines are commonly used in animation, these splines are often multilevel B-splines. We use Catmull-Rom splines [8], which are mathematically simpler, but serve our purpose as they are interpolating splines that pass through all of their control points. If there are not enough control points, the spline will not closely fit the collected data, but if there are many, then the data is not adequately down-sampled. Fig. 3 shows a comparison of a data curve to a Catmull-Rom curve with six control points. While six points has shown adequate performance so far, modeling more complex gaze behavior may require either increasing the number of control points, or segmenting the complex behavior into simpler gaze shifts.

In order to determine the proper location for the control points, we first find the locations of the control points on the animation curve representing the direction in which head rotation is the largest, as the head performs the easiest to distinguish and largest amplitude movements during many gaze shifts. Since these are single gaze shifts, the curve representing the angle the heard turns through resembles an S-curve.

The control points are placed to minimize least squared error while meeting a number of constraints, so that of the six control points, one point represents the start, one the end, two points bracket the first curve, and two points bracket the second. Then, we place control points in all of the other animation curves in the gaze shift at the same temporal location as these six.

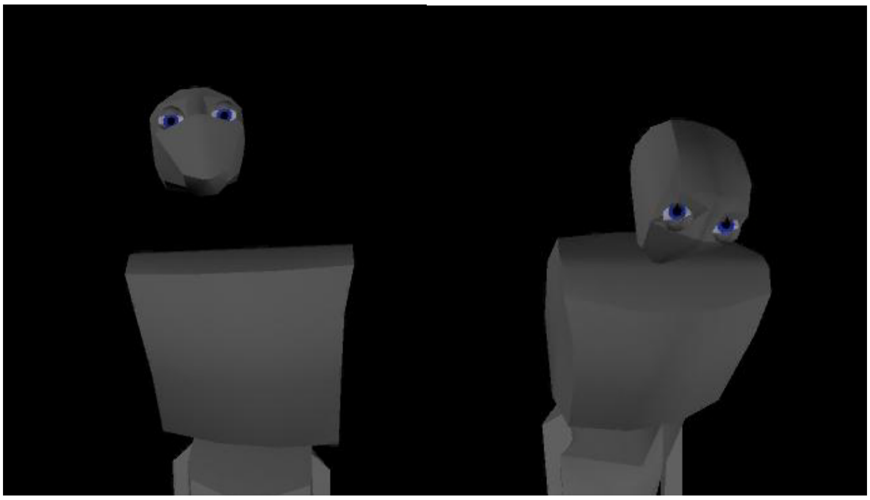

Fig. 3. Animated Model showing High (left) and Low (right) Dominance

As the GWT is a set of point-to-point transformations, the control points on each individual gaze shift need to align, and increasing the accuracy of alignment will result in increased accuracy in the transformation. Currently, the constrained least squares minimization method only produces an approximate alignment.

The next step is to find, for each emotionally expressive gaze shift, the motion warping [24] functions that transform the control points for that shift into the control points for an emotionally neutral gaze shift following the same path. A motion warping function provides a pointwise transformation from one animation curve $x(t)$ that consists of a set of (x_i, t_i) pairs to a new motion $x'(t')$. The first function is $t = g(t')$, where given an actual frame time t', g describes where in the unwarped motion curve to obtain x. We use $g(t')=ct$, where c is a time scaling parameter. The other function is $x'(t) = a(t)x(t) + b(t)$, where $a(t)$ is a scaling function, and $b(t)$ is an offset function. However, $a(t)$ and $b(t)$ are not uniquely determined for a single x'. In [24], the user selected one of the two values to use, and one to hold constant. We currently use the offset function $b(t)$, and hold $a(t)$ constant, but either would work.

So, to convert the six control points for a DOF in an emotionally expressive gaze shift into the six points for that DOF in an emotionally neutral shift, we find, for each pair of points, the spatial offset parameter $b(t)$, and the time scaling parameter c that would transform between them. By doing this for each DOF in the gaze shift, we end up with nineteen sets - one set for each DOF, and one set for the time scaling parameter - of six motion warping parameters - one for each control point - for each gaze shift.

The next step is to find which of the motion warping parameters are significant for which emotional dimensions. While all of the motion warping parameters could be included, we would prefer sparse GWTs, as this makes them easier to combine. To do this we perform a series of ANOVAs across multiple animations that compare the motion warping parameters for the different emotional states to each other. This tells us which of the motion warping parameters are significantly different across different states along an emotional axis. We compared twelve low arousal motion warping parameter sets to twelve high arousal parameter sets (n = 24), and twelve low dominance sets to fourteen high dominance sets (n = 26), running one ANOVA for each DOF and control point combination. A subset of the results of these ANOVAs can be seen in Table 2, which shows some significant relationships between motion DOF's, and emotional dimensions.

Each row in the table shows the emotional dimensions that had a difference significant to $p < 0.05$ between its low and high values for that degree of freedom. Additionally, the behavior guidelines described in Table 1 can be seen in the subset of parameters in Table 2. For example, the guidelines called for changes in speed and forwards/backwards position of the body to show arousal, which is what we found. In addition, dominance was significantly related to head pitch and body vertical position. The warping parameters thus provide specific values for the abstract, general guidelines expressed in Table 1.

The GWT was assembled from the motion warping parameters for each DOF that had a plurality of control points significant to that dimension. Entire DOF's were used to increase the smoothness of movement.

Table 2. Some Significant Emotional and Movement Relationships

Movement Dimension	Emotional Dimension	Observed Change
Speed	Arousal	Low Arousal Animations are 80% slower than High Arousal Animations
Head Rotation - Pitch	Dominance	Low Dominance Head Pitch is 25 degrees lower than High Dominance
Body Movement - Front/Back	Arousal	The body moves forwards 1.5 inches more in High Arousal than Low Arousal Animations
Body Movement - Vertical	Dominance	The body moves down 1 inch more in Low Dominance than High Dominance Animations

The final GWTs were then used to warp neutral gaze shifts into emotionally expressive gaze shifts. We note that using individual motion warping transformations in this way assumes independence between the transformations, which is currently an untested assumption

3.4 Producing Emotionally Expressive Animations

After obtaining the GWT for emotionally expressive head and body manner during gaze shifts, new movements can be generated. Because the GWT is a set of motion warping parameters, an emotionally neutral gaze shift is down-sampled to six control points, and then the parameters are used to warp the neutral gaze shift control points. Finally, interpolation generates an emotionally expressive gaze shift. In addition, GWTs describing different emotional dimensions can be combined. Currently, there is no overlap in the arousal and dominance parameters, simplifying combination. In the future, better methods for combining parameters will be needed.

For evaluation we generated animations of HDHA, HDLA, LDHA, and LDLA by warping three gazes: one straight ahead of the actor, one 20° to the actor’s left, and one 45° to the actor’s left. This results in manner parameters taken from attraction gazes, looking at a target directly ahead of the actor, placed onto aversion gazes looking away from this target. We also used combinations of gaze warping transformations, instead of individual transformations; for example, combining the GWTs from NAHD and NDHA instead of using the collected HDHA gazes. The purpose of this was to provide a broader evaluation of the performance of GWTs. The result of applying the gaze warping parameter to the Head Rotation – Pitch DOF can be seen in Figure 4, and the resulting change in generated gaze can be seen in Figure 3.

After applying the transformations, constraints are applied to ensure that the motion warping does not cause physically impossible motions. This is needed because motion warping is a geometric, not a physically-based, technique. We are currently using ad-hoc constraints drawn from motion capture where an actor performed motions to explore the limits of their movement. Finally, the new emotionally expressive gaze shifts are animated on a very simple model in Maya (Figure 3). The model is intended to ablate nonverbal signals that are not being examined in the course of this research, such as facial expression or hand gesture.

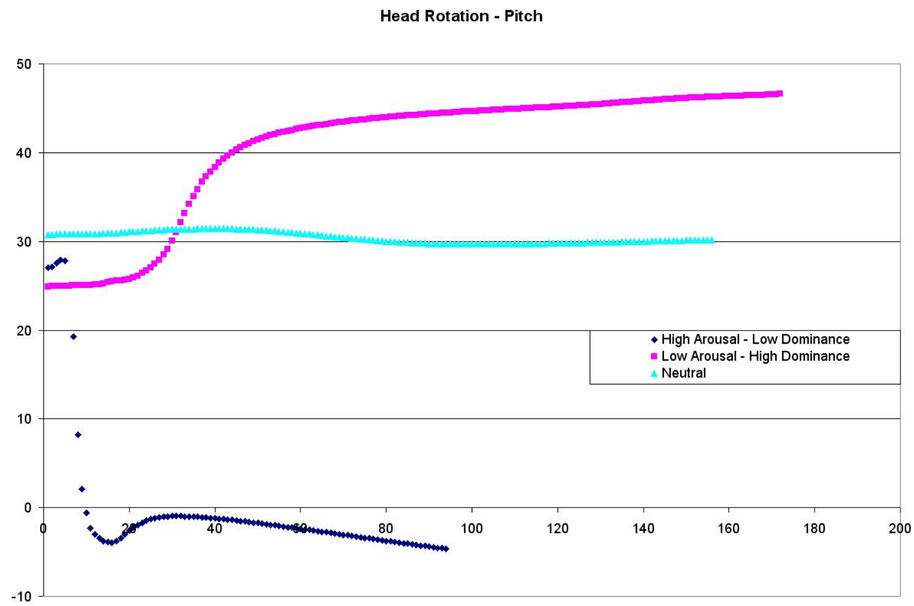

Fig. 4. Head Rotation –Pitch Showing Transformation from Neutral to HALD and LAHD

4 Evaluation

We then evaluated the animations to determine if the emotional signals in the generated animations for HDHA, HDLA, LDHA, and LDLA were coming through as strongly as the emotional signals in motion-captured animations of the same emotional states. To do this, we provided the videos and a questionnaire to 21 coders. Ten coders saw the set of twelve generated animations, while eleven coders saw the set of four motion-captured animations. Animations were displayed to coders in pairs, differing along a single emotional dimension, e.g. HDHA vs. HDLA. The set of animations is arranged according to a Latin Square. The coders chose the animation that showed higher arousal, and the animation that showed a more dominant character. Then, they rated the arousal and dominance of each animation individually on five-point Likert scales. Each coder saw each animation twice, but rated it only once on dominance, and once on arousal.

The result of this evaluation showed that coders significantly distinguished between low and high arousal; and between low and high dominance for both the generated and motion captured animations, as shown by a Chi squared test. The ability to distinguish arousal was similar for both the generated and captured animations (see Table 3). When asked to select the more highly aroused animation, the coders selected the generated animation intended to display high arousal 85% of the time, while the percentage was 86% for the captured animations.

Table 3. Evaluation Results - Arousal

Arousal		
Statistic	Generated Animations	Captured Animations
Number of Comparisons	60	22
Comparison Recognition	85%	86%
Comparison *p*	<.01	<.01
Number of Ratings	120	44
Low Arousal Mean Rating	3.000	2.80
High Arousal Mean Rating	3.817	3.70
F	9.045	15.2821
ANOVA *p*	<.05	<.01

The results for dominance can be seen in Table 4. Unlike arousal, the recognition for dominance, while still significant, was lower for the generated animations (66%) than for the captured animations (90%).

A within-subject Multivariate ANOVA showed that ratings on the Likert scales were also significantly ($p < .05$) different for low and high arousal animations. However, the Likert scale ratings were not significantly different, for low and high dominance in the animations generated with the GWT, although they were for the motion-captured animations. The interaction effects were also not significant.

The evaluation results show that we have transferred the manner of emotionally expressive gaze shifts to different emotionally neutral gaze shifts in such a way that the signal can still be recognized. It also reveals that there is room for improvement, specifically with regards to the signaling of dominance. The dominance animations that obtained the lowest recognition were those where the movement was quick and the animation did not end facing straight ahead. During debrief, coders explained that the character appeared to be looking at objects instead of expressing emotion.

One cause of this is the lack of a model of eye movement. We are currently working on a model of expressive eye movement, drawing from the same motion capture data, to merge with this model of head and body movement during gaze shifts. Another likely cause is the transfer of manner from the captured attraction gaze to the generated aversive gaze which was performed to test the ability of GWTs to generalize between different types of gaze. This transformation of a gaze, based on the emotional content of a very different type of gaze, could cause the decrease in recognition.

Table 4. Evaluation Results - Dominance

Dominance		
Statistic	Generated Animations	Captured Animations
Number of Comparisons	60	22
Comparison Recognition	66%	90%
Comparison *p*	<.01	<.01
Number of Ratings	120	44
Low Dominance Mean Rating	3.000	2.55
High Dominance Mean Rating	3.333	3.70
F	2.250	30.9729
ANOVA *p*	.168	<.01

5 Conclusion

There has been little prior work on using the expressive gaze manner of characters to display their emotional state. In this work we have described the Gaze Warping Transformation, a method for combining and transferring expressive gaze manner from emotional gazes to neutral gazes, and performed an evaluation showing that the encoded emotional dimensions can be recognized. There are improvements that need to be made, the most pressing of which is the integration of a model of expressive eye movement. In addition, the pleasure dimension of the PAD model needs to be integrated, and the relationship between expressive behavior and PAD space needs to be explored. Finally, implementation details such as improving point-to-point alignment and determining to what extent to down-sample animation curves must be addressed. Yet, the performed evaluation demonstrates the utility of our model, and we will continue to build upon it.

Acknowledgements

We would like to thank Mei Si and Sinhwa Kang for sharing their knowledge of statistical analysis, Tiffany Cole for her acting knowledge, and Dr. Bosco Tjan and Dr. Skip Rizzo for the use of their data recording hardware.

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 1996 Conference on Graphical Interface, pp. 222–229 (1996)
[2] Argyle, M., Cook, M.: Gaze and Mutual Gaze. Cambridge University Press, Cambridge (1976)
[3] Bickmore, T., Cassell, J.: Social Dialogue with Embodied Conversational Agents. In: Bernsen, N. (ed.) Natural, Intelligent and Effective Interaction with Multimodal Dialogue Systems, pp. 23–54. Kluwer Academic Publishers, Dordrecht (2004)
[4] Brand, M., Hertzmann, A.: Style Machines. In: Proceedings of SIGGRAPH, ACM Press, New York (2000)
[5] Carney, D.: Beliefs About the Nonverbal Expression of Social Power. Journal of Nonverbal Behavior. 29(2), 105–123 (2005)
[6] Coulson, M.: Attributing Emotion to Static Body Postures: Recognition Accuracy, Confusions, and Viewpoint Dependence. Journal of Nonverbal Behavior 28(2) (2004)
[7] Exline, R.: Visual Interaction: The Glances of Power and Preference. In: Weitz, S. (ed.) Nonverbal Communication: Readings with Commentary, Oxford University Press, Oxford (1974)
[8] Foley, J., van Dam, A., Feiner, S., Hughes, J.: Computer Graphics: Principles and Practice, 2nd edn. Addison-Wesley Publishing Company, Redwood City,CA, USA (1997)

[9] Gebhard, P.: ALMA: A Layered Model of Affect. In: Proceedings of AAMAS, ACM Press, New York (2005)
[10] Kendon, A.: Some Functions of Gaze Direction in Two-Person Conversation. In: Kendon, A. (ed.) Conducting Interaction: Patterns of Behavior in Focused Encounters (1990)
[11] Kim, Y., Hill. Jr., R., Traun, D.: A Computational Model of Dynamic Perceptual Attention for Virtual Humans. Proceedings of the 14th Conference on Behavior Representation in Modeling and Simulation. May (2005)
[12] Kleinke, C.: Gaze and Eye Contact: A Research Review. Psychological Bulletin. v. 100(1), 78–100 (1986)
[13] Lance, B., Marsella, S., Koizumi, D.: Towards Expressive Gaze Manner in Embodied Virtual Agents (2004)
[14] Lee, S., Badler, J., Badler, N.: Eyes Alive. ACM Transactions on Graphics 21(3), 637–644 (2002)
[15] Liu, C.K., Hertzmann, A., Popović, Z.: Learning Physics-Based Motion Style with Nonlinear Inverse Optimization. In: Proceedings of SIGGRAPH, ACM Press, New York (2005)
[16] Marsella, S., Gratch, J., Rickel, J.: Expressive Behaviors for Virtual Worlds. In: Prendinger, H., Ishizuka, M. (eds.) Life-Like Characters. Tools, Affective Functions, and Applications, Springer, Heidelberg (2003)
[17] Mehrabian, A.: Silent Messages: Implicit Communication of Emotions and Attitudes, 2nd edn. Wadsworth Publishing Company (1981)
[18] 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) (Summer 2003)
[19] Paterson, H., Pollick, F., Sanford, A.: The Role of Velocity in Affect Discrimination. In: Proceedings of the 23rd Annual Conference of the Cognitive Science Society (2001)
[20] Pelachaud, C., Bilvi, M.: Modelling Gaze Behavior for Conversational Agents. In: Proceedings of Intelligent Virtual Agents 2003. LINAI Series, Springer, Heidelberg (2003)
[21] Rickel, J., Johnson, W.L: Animated Agents for Procedural Training in Virtual Reality: Perception, Cognition, and Motor Control. Applied Artificial Intelligence 13(4-5), 343–382 (1999)
[22] Sobczynski, P.: Polar Express, The (November 2004), http://www.efilmcritic.com/review.php?movie=10918&reviewer=389
[23] Thomas, F., Johnston, O.: The Illusion of Life: Disney Animation. Walt Disney Productions (1981)
[24] Witkin, A., Popovic, Z.: Motion Warping. In: Proceedings of SIGGRAPH, ACM Press, New York (1995)
[25] Zhao, L., Badler, N.: Acquiring and Validating Motion Qualities from Live Limb Gestures. Graphical Models 67(1), 1–16 (2005)