Appealing Virtual Humans

Rachel McDonnell

Graphics, Vision, and Visualisation Group,
Trinity College Dublin
Rachel.McDonnell@cs.tcd.ie
http:gv2.cs.tcd.ie/mcdonner/

Abstract. Computer generated virtual humans play an important role in many 3d graphical applications, and are often used to evoke empathic responses from human users, due to their human form and behaviour. Most movie and game studios aim to produce appealing virtual characters that audiences will react to in a positive manner. However, humans are very experienced at observing the motion and appearance of other humans, and can be very unforgiving of minor errors in the form or movement of virtual characters. Negative reactions to virtual characters are common and have been attributed to many factors including incongruence between their motion and appearance. In our recent work, we investigated the complex interaction between the movement and appearance of virtual humans using perceptual experiments. The aim being to guide developers in the creation of plausible characters that are considered appealing to viewers.

Keywords: Motion Capture, Perception, Uncanny Valley.

1 Introduction

Animated virtual humans are needed for many applications in entertainment, education, and science. Their movements can be created by artists, by programs (procedural animation), or recorded from real people using motion capture technology. Motion capture is the process of transferring a live performance into a digital performance. It is widely used in graphics applications that require animations that mimic real human movement. In games, it is not always feasible or desirable to use virtual model replicas of the captured actors. Therefore, it can often be the case that the appearance of the captured actor does not match their in-game character. However, little is known about the effect of this mismatch on viewer perception.

Producing appealing virtual characters is typically the goal of most movie makers and game studios, and so avoiding negative reactions from audiences is of great importance. There are many other applications besides the movie and game industry that could benefit from guidelines on creating appealing characters. For example, in advertising it could be very advantageous to determine which human representations would best convince viewers to purchase their products.

M. Kallmann and K. Bekris (Eds.): MIG 2012, LNCS 7660, pp. 102–111, 2012.

Furthermore, in virtual training applications, using appealing virtual characters could help with training success.

People expect a virtual character to behave in a manner befitting its appearance and they find discrepancies in behaviour disturbing. [1]. There are a number of reasons why mismatching form and motion may appear incorrect or disturbing to a viewer. The style in which a human moves can be related to a multitude of factors, both physical and psychological. If the physical or emotional state of the virtual character does not match that of the captured actor, this may cause the viewer a conscious or unconscious feeling that there is something wrong with that character. If, for example, the captured actor was a large muscular, confident man, whose motions were applied to a character that appeared unhappy and thin, this mismatch could appear very strange. More subtle incongruence may cause unconscious effects that something is wrong, such as described in the Uncanny Valley (UV) theory. The UV has become a standard term for the hypothesis that virtual humans can appear unintentionally eerie or creepy when they do not appear how the viewer expects. This UV phenomenon was first hypothesized by robotics professor Masahiro Mori [5]. Mori predicted that as a robot looks more human it also looks more agreeable, until it appears so human that we begin to find subtle imperfections unsettling. More recently, the UV hypothesis has been transfered to human avatars in computer graphics, and has been explored directly in some studies [14,15].

The complex interaction between motion and appearance has been examined in previous work. Hodgins et al. [3] found that viewers' perception of motion characteristics is affected by the geometric model used for rendering. They observed higher sensitivity to changes in motion when applied to a polygonal model, than a stick figure. Chaminade et al. [2] investigated how the appearance of a range of characters influenced perception of their actions. They found that the perceived biological motion of a character decreased with anthropomorphism. More recently, Hodgins et al. [4] conducted perceptual experiments to determine how degradation of human motion affects the emotional response of participants to animation. They found that removing facial animation and/or sound changed the emotional content that was communicated to their participants.

Research in experimental psychology has demonstrated how finely tuned the human brain is at recognizing biological motions and distinguishing between very subtle characteristics of this motion. Johansson [7,8] reported that natural motion, in the absence of any spatial information, can give a rich perception of a moving human figure. His 'point-light' displays were designed to separate biological motion information from other sources of information that are normally intermingled with the motion of a human, such as form or outline. He showed that 12 moving light points suffice for motion perception within a very short space of time (200msec, or five frames of a movie).

In this paper, we discuss some of our recent research which investigates some factors which contribute to the appeal and plausibility of virtual characters. In a series of psychophysical experiments, we examine the plausibility of applying motion captured movements to a range of different virtual bodies. In Section 2,

we discuss experiments which vary the body shape of the characters, while determining plausibility of the applied motion. Locomotion, conversational and emotional body motions were tested. In Section 3, rather than changing the morphology, we change the render style, to determine if this also has an effect on viewer perception.

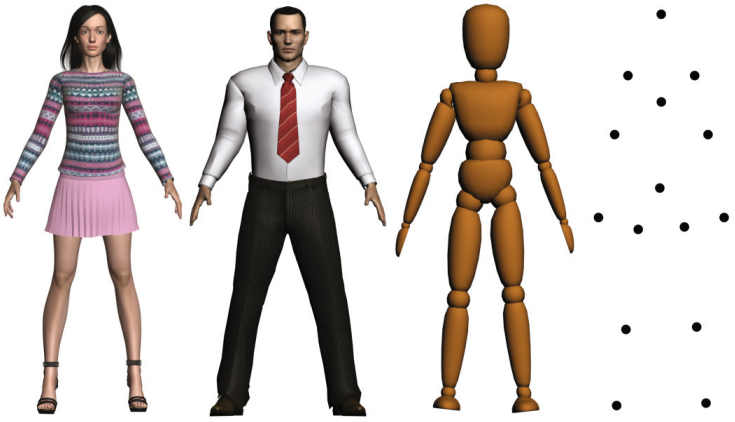

Fig. 1. Models used to examine the effect of applying male and female motion [11]

2 Body Shape

This first set of experiments examine the effect of viewing the same motion applied to a range of different morphologies, to determine the limits of plausibility. We start with the extreme case of perception of motion transfer between male and female bodies. Specifically, we ask the question whether, if there is a clear visual indicator of sex (i.e., an unambiguously female or male model, as shown in Figure 1), does the motion have to be congruent? Then, in Section 2.2, we investigate if different bodies (ranging from a pleasant cartoon character to a grotesque zombie) evoke different emotional reactions.

2.1 Female vs. Male

People's perception of the sex (whether the character is male or female) of a human representation has been studied in the experimental psychology literature, with respect to human motion with minimal shape information and the effect of body shape. It is widely known that the sex of a walker can be determined using only motion signals [16]. However, only recently have the effects of both shape and motion been considered together. Johnson and Tassinary [9] studied the effects of both shape and motion on the perception of sex, gender, and attractiveness. Instead of using point light walkers which give little body shape

information, they used silhouettes of human body shapes with varying waist to hip ratios, from exaggerated hourglass to tubular figures. Even though synthetic motion was restricted to two highly informative parameters (i.e., swagger for men and hip sway for women), they found that the shape of the character was more informative of the sex of the individual than its motion. They recently followed this up with a study in which they showed that both form and motion information, using these exaggerated feminine and masculine cues, contributed to participants' judgments of attractiveness [17].

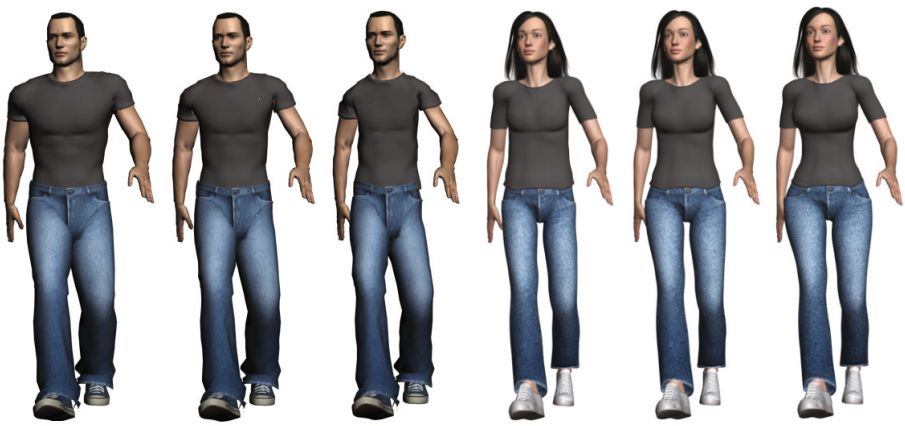

Fig. 2. Body shapes ranging from very male to very female [11]

Locomotion. In our work [11], we aimed to determine the factors that influence the perceived sex of virtual characters. We used natural female and male walking motions, as well as synthetic neutral motions, and applied them to realistic male and female models (Figure 2.1). Furthermore, we directly compared results with an androgynous shape and point light walker and also examined the effects of incongruent shape and motion combinations. Unlike previous work, our results indicated that both the appearance and the motion of the character influence perceived sex. We also found that for walk motions with no male or female characteristics, it is the appearance of the character that will dominate our perception of the character's sex, whereas it is the motion that dominates our perception for characters with an androgynous appearance. This information will be useful in computer graphics applications, as it is now clear that in order to have a convincing male/female character, you must apply a female walk to a female character, and a male walk to a male character. This implies that consistency is highly important when designing a virtual character. As discussed by Vinayagamoorthy et al. [1], people expect virtual characters to behave in a manner befitting its appearance, and the results of our study confirm this. From our results, it is certain that you cannot apply a female walk to a male character, or vice versa, as viewers will judge this conflicting information to be ambiguous.

In a follow up experiment, we further investigated the influence of body shape and motion on different male and female shapes to determine if additional indicators of sex in the shape will change the overall perception of the sex of the character. We manipulated the polygons of the woman and man models in order to create very exaggerated female and male body shapes, and also very androgynous versions. We then used a linear morph to create an in-between shape (Figure 2). We found that adding stereotypical indicators of sex to the body shapes influenced sex perception. Also, that exaggerated female body shapes influences sex judgements more than exaggerated male shapes.

Conversational Motion. It is clear from our work and from previous work that walking motions contain indicators of the sex of the walker. Presumably, this is related to the difference in shape between male and female bodies, which causes a difference in weight distribution while walking. However, it was unclear to us if this difference would translate to other motions where weight distribution would not be a factor. Therefore, we chose conversational body motion as our next case to test for character plausibility. We also investigated the effect of adding auditory cues. We showed that, based on body motion alone, participants were able to distinguish between male and female conversational body motions [12]. Also, as before, when incongruent pairs were presented, this was judged as ambiguous by participants. In order to create unambiguous, plausible conversing characters, we found that voice affects sex perception, but does not dominate. An ambiguous motion can be disambiguated with the addition of audio, but incongruent audio and motion will always be considered as ambiguous. The addition of extra appearance cues (female or male as opposed to androgynous character models) does not affect ratings. These results will be useful for computer graphics applications, since it is now clear that conversational motions carry sex information and should only be displayed with a congruent voice to avoid ambiguity. Also, pairing a voice with a motion is an effective way to increase how male/female a character appears.

2.2 Body Type

In recent games, virtual characters are becoming more and more expressive and have the ability to display complex emotions. Expressing emotion is one of the most important mechanisms for endowing virtual humans with personality and achieving a bond between that character and the viewer. In this experiment [10], we investigated whether or not different body types would alter the perception of the intended emotion.

Our aim was to explore how emotions were perceived when applied to anthropomorphic but nonhuman characters. A range of body types were tested: A cartoon-like figure was chosen in order to determine if his friendly appearance would lead him to be perceived as exhibiting more positive emotions, whereas we felt that a zombie-like character would be perceived in a more negative light. A wooden mannequin model was chosen as a completely neutral character that

Fig. 3. Body shapes used in Section 2.2. From left to right: virtual actor replica, wooden mannequin, cartoon character, zombie. Faces were blurred out to avoid any emotional bias due to facial expression. Hands were also blurred as we did not capture finger motion [10].

we felt may be perceived to be exhibiting less intense emotions than the others. In order to analyze the emotional content of motions portrayed by different characters, we captured an actor exhibiting six basic emotions: sadness, happiness, surprise, fear, anger, and disgust. His actions were then applied to five virtual body shapes: a virtual counterpart, a cartoon-like character, a wooden mannequin, and a zombie-like character (Figure 3). Participants were asked to rate the actions based on a list of 41 more complex emotions. We found that body representation had little or no effect on people's perception of the emotional content of the scene, which lead us to conclude that people's perception of emotional motion is very robust and unaffected by a character's physical appearance. Furthermore, that there was no difference between the emotion ratings for the zombie and the other characters. This would imply that, even though descriptions of the static body suggested that participants found him unappealing, a negative reaction did not occur when realistic motion was applied. This indicates that perhaps it could be the face and hand motions that cause negative reactions and not the body motions, as tested here.

3 Render Style

One major factor that influences appeal that we have not discussed before now is the style used to render the virtual character. Perceptual experiments were conducted in order to independently test the effect of render style on virtual human perception [13]. The focus was on the face of the characters in these experiments, since we previously found no difference in emotion perception for body motion alone. Also, research on the perception of human faces suggests that faces are processed in a different cognitive manner than other objects. Studies have indicated that humans possess a face-specific recognition system [18]. Therefore,

even the smallest anomaly in the face shape, movement, or skin texture can be detected and classified as incorrect. We ensured in this case that the motion captured actor matched their virtual counterparts, in order to avoid any unknown effects that the mismatch may cause.

3.1 Appearance Ratings

There has been much speculation about whether or not it is profitable to invest large resources into trying to achieve photorealistic characters, as the audience reaction to these characters can often be negative. For example in the movie *Polar Express*, characters were realistically rendered yet were frequently considered off-putting when compared to the stylized characters used in movies such as *The Incredibles* [6]. In this first set of experiments, we explicitly ask participants to rate different render styles applied to the same character, in order to determine the direct effect of render style on their impression of the character. Based on Mori's theory [5], we hypothesized that participants would give abstract characters more positive ratings than characters that they considered to have almost photo-realistic appearance.

We collected participants' direct impressions of the ten render styles (Figure 4), by asking them to rate how realistic, re-assuring, familiar, appealing, friendly, and trustworthy they found them to be. These ratings were collected separately on moving and still characters. We also tested regular motion and motion with anomalies present, to determine the effect of unnatural motion. These direct ratings allowed us to determine the relationship between realism and appeal for virtual characters, and whether motion had an effect on the ratings. We found that much of the information that we use to rate virtual characters is available in a still image. Movement changes only how familiar we find the characters, and also how appealing they are considered. Highly unappealing characters are considered more so when movement is applied, and motion anomalies are considered more unpleasant on human than on cartoon render styles. We believe that this was due to the fact that humans are inherently conditioned to analyze human faces, and are therefore less forgiving of anomalies when a human photograph is applied to the model. These results are very much in line with the Uncanny Valley theory [5]. Contrary to the theory, however, was the fact that our most realistic character was often rated as appealing or pleasant as our cartoon characters. The drop in appeal for our stimuli came from characters that were in the middle of our 'abstract to realistic' scale, which could be attributed to the fact that these characters were difficult for the brain to categorize due to their uncommon appearance, as suggested in [19].

Based on the results of our experiments, we can reassure developers that selecting realistic virtual characters is not as risky as often discussed. Negative reactions occurred mainly for characters that used human texture maps, but that were not rendered with realistic eye and skin shaders. Cartoon characters were also considered highly appealing, and were much more forgiving to motion artefact's than characters with human appearance.

Fig. 4. Three of the render styles used in Section 3, ranging from realistic to abstract [13]

3.2 Lie Detection

Social interaction with virtual humans is becoming popular in interactive drama video games, where players decisions and actions during the game affect the narrative. For example, in the video game *L.A. Noire*, gamers are asked to interact on a higher level with characters than ever before, by trying to determine if a virtual character is lying to them or not. So far, these types of games have chosen human-like characters, presumably to allow for more believable performances and interaction. In this experiment, we investigate if the choice of rendering style is important for the display of subtle cues such as those that occur when lying, and if different rendering styles will evoke different reactions to the same performance.

We chose a deception task for this experiment, where participants were asked to detect if a virtual character was telling the truth or lying. Previous work using videos of real people has shown that deception detection is difficult, with participants usually being able to detect a lie at just above chance level. We postulated that rendering style would unconsciously influence their decisions, where a negative reaction towards a character could increase the 'lie' responses. Using recorded truths and lies from two human actors and their virtual replicas, we determine if the style in which the characters are rendered has an effect on how truthful they are perceived to be. Also, if there are differences in ratings between real and virtual characters. Overall, we found that the audio track contained the most reliable cues for deception detection, and without this, participants were performing at chance level. With audio and visual cues, participants relied on visual more than auditory cues for the appealing virtual avatars. Furthermore, we found that rendering style did not bias 'truth' responses, and that lies and truths were perceived on real and virtual characters in the same way.

Our result is interesting since it shows that participants were so focused on the task, that the appearance of the character did not sway them. We found that the audio and animation contributed to the interpretation of the characters' intention rather than the render style. In the entertainment industry, one could speculate, based on our results, that blaming the render style for the box-office failure of photorealistic CG movies is not a valid argument. One could argue that it was the content (animation, audio, modeling) and not the render style

that contributed towards the failure, as we found that render style does not alter the interpretation of content in a positive or negative manner.

4 Conclusions and Future Work

In this paper, an overview has been presented of our current research into creating plausible and appealing virtual humans, by investigating the perception of form and motion. We have gained some valuable insights into this complex issue. We found that incongruent form and motion can cause undesirable effects. In particular, if the sex of the captured actor does not match that of the virtual character, participants will find this disturbing. Body shape did not affect ratings of emotional body motions, however, which implies that any virtual body can be used to convey emotion convincingly. This was true whether a cartoon character or a realistic virtual body replica of the captured actor was used. Similarly, we found that virtual faces with motion capture applied were equally able to convey subtle cues that come from lying as real humans. Also, that the render style did not alter the impression the performance gave to the viewer. This suggests that when motion cues are strong, visual cues will not alter participants' opinions.

Although some interesting insights have been gained from our work, we are still a long way from truly understanding how humans perceive virtual characters. However, we believe that implicit testing (such as our lie detection experiment) could prove very useful and practical to investigate the much discussed issue of the "uncanny valley". Future experiments will probe further into the effects of auditory, visual, and motion cues using different render styles, in an attempt to qualify the importance of each of the cues.

References

1. Vinayagamoorthy, V., Gillies, M., Steed, A., Tanguy, E., Pan, X., Loscos, C., Slater, M.: Building Expression into Virtual Characters. Eurographics State of the Art Reports (2006)
2. Chaminade, T., Hodgins, J., Kawato, M.: Anthropomorphism influences perception of computer-animated characters' actions. Social Cognitive and Affective Neuroscience 2, 206–215 (2007)
3. Hodgins, J.K., O'Brien, J.F., Tumblin, J.: Perception of Human Motion With Different Geometric Models. IEEE Transactions on Visualization and Computer Graphics 4(4), 307–316 (1998)
4. Hodgins, J.K., Jörg, S., O'Sullivan, C., Park, S.I., Mahler, M.: The saliency of anomalies in animated human characters. ACM Transactions on Applied Perception 7(4), 1–14 (2010)
5. Mori, M.: The Uncanny Valley. Energy 7(4), 33–35 (1970)
6. Levi, S.: Why Tom Hanks is less than human; While sensors cannot capture how humans act, humans can give life to digital characters. Newsweek 650, 305–306 (2004)
7. Johansson, G.: Visual perception of biological motion and a model for its analysis. Perception and Psychophysics 14(2), 201–211 (1973)

8. Johansson, G.: Spatio-temporal differentiation and integration in visual motion perception. Psychological Research 38(4), 379–393 (1973)
9. Johnson, K.L., Tassinary, L.G.: Perceiving sex directly and indirectly: Meaning in motion and morphology. Psychological Science 16(11), 890–897 (2005)
10. McDonnell, R., Jörg, S., McHugh, J., Newell, F., O'Sullivan, C.: Investigating the role of body shape on the perception of emotion. ACM Transactions on Applied Perception 6(3), 14:1–14:11 (2009)
11. McDonnell, R., Joerg, S., Hodgins, J.K., Newell, F., O'Sullivan, C.: Evaluating the effect of motion and body shape on the perceived sex of virtual characters. ACM Transactions on Applied Perception 5(4), 20:1–20:14 (2009)
12. McDonnell, R., O'Sullivan, C.: Movements and voices affect perceived sex of virtual conversers. In: APGV 2010: Proceedings of the Symposium on Applied Perception in Graphics and Visualisation, pp. 125–128 (2010)
13. McDonnell, R., Breidt, M., Bülthoff, H.: Render me real?: investigating the effect of render style on the perception of animated virtual humans. ACM Transactions on Graphics 31(4), 91:1–91:11 (2012)
14. MacDorman, K.F., Green, R.D., Ho, C.H., Clinton, T.K.: Too real for comfort? Uncanny responses to computer generated faces. Computers in Human Behavior 25, 695–710 (2009)
15. Bartneck, C., Kanda, T., Ishiguro, H., Hagita, N.: Is the uncanny valley an uncanny cliff? In: Proceedings of the 16th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2007, pp. 368–373 (2007)
16. Kozlowski, L., Cutting, J.: Recognizing the sex of a walker from a dynamic point-light display. Perception and Psychophysics 21, 578–580 (1977)
17. Johnson, K.L., Tassinary, L.G.: Compatibility of basic social perceptions determines perceived attractiveness. Proceedings of the National Academy of Sciences 104(12), 5246–5251 (2007)
18. Farah, M.J.: Is an object an object an object? Cognitive and neuropsychological investigations of domain specificity in visual object recognition. Current Directions in Psychological Science 1(5), 164–169 (1992)
19. Saygin, A.P., Chaminade, T., Ishiguro, H., Driver, J., Frith, C.: The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions. Social Cognitive and Affective Neuroscience 7(4), 413–442 (2012)