IMEKO, IEEE, SICE
2nd International Symposium on Measurement, Analysis and Modeling of Human Functions
1st Mediterranean Conference on Measurement
June 14-16, 2004, Genova, Italy

PERCEIVING ANIMACY AND AROUSAL IN TRANSFORMED DISPLAYS OF HUMAN INTERACTION

Phil McAleer, Barbara Mazzarino, Gaultiero Volpe*, Antonio Camurri*, Helena Paterson, Kirsty Smith, Frank E. Pollick*

Department of Psychology, University of Glasgow, Glasgow, Scotland
*Infomus Lab, D.I.S.T., University of Genova, Genova, Italy

1. INTRODUCTION

When viewing a moving abstract stimulus, people tend to attribute social meaning and purpose to the movement. The classic work of Heider and Simmel [1] investigated how observers would describe movement of simple geometric shapes (circle, triangles, and a square) around a screen. A high proportion of participants reported seeing some form of purposeful interaction between the three abstract objects and defining this interaction as a social encounter. Various papers have subsequently found similar results [2,3] and gone on to show that, as Heider and Simmel suggested, the phenomenon was due more to the relationship in space and time of the objects, rather than any particular object characteristic. The research of Tremoulet and Feldman [4] has shown that the percept of animacy may be elicited with a solitary moving object. They asked observers to rate the movement of a single dot or rectangle for whether it was under the influence of an external force, or whether it was in control of its own motion. At mid-trajectory the shape would change speed or direction, or both. They found that shapes that either changed direction greater than 25 degrees from the original trajectory, or changed speed, were judged to be "more alive" than others. Further discussion and evidence of animacy with one or two small dots can be found in Gelman, Durgin and Kaufman [5]

Our aim was to further study this phenomenon by using a different method of stimulus production. Previous methods for producing displays of animate objects have relied either on handcrafted stimuli or on parametric variations of simple motion patterns. It is our aim to work towards a new automatic approach by taking actual human movements, transforming them into basic shapes, and exploring what motion properties need to be preserved to obtain animacy. Though the phenomenon of animacy has been shown for many years, using various different displays, very few specific criteria have been set on the essential characteristics of the displays. Part of this research is to try and establish what movements result in percepts of animacy, and in turn, to give further understanding of essential characteristics of human movement and social interaction.

In this paper we discuss two experiments in which we examine how different transformations of an original video of a dance influences perception of animacy. We also examine reports of arousal, Experiment 1, and emotional engagement in Experiment 2.

2. EXPERIMENT 1

In this experiment we transform the original video to reduce the visual information available from a display of two dancers to examine if this will result in animacy displays. Furthermore, we investigate the change in arousal across four displays with changing levels of information. We expect that the four display conditions will result in free response descriptions using animate terms, and that arousal levels will decrease as available visual information is reduced.

2.1 Method

In a between subjects design, with each participant only viewing one stimulus, participants watched a short excerpt with a duration of 1 minute 46 seconds. Using a slider subjects rated the excerpt on arousal and afterwards completed a free response task. Four visual displays, of the same interaction between two dancers performing a modern dance, were created using the EyesWeb open platform for multimedia application and motion analysis [6] (www.eyesweb.org). These conditions (Figure 1), in decreasing order of available visual information were: (1) full video recording; (2) body silhouettes; (3) motion of the barycentre of each dancer represented by a single small rectangle changing in size in accordance with the respective motion energy [6] of the dancer; (4) motion of the barycentre of the dancer represented by a small rectangle that did not change in size.

The body silhouette in condition 2 was obtained by removing color information and applying a background subtraction technique to input video frames. In condition 3 the size of the small rectangles were related to the Quantity of Motion (QoM) as measured by algorithms included in the EyesWeb Expressive Gesture Processing Library [7]. QoM is computed as the change in area of the dancer, in the silhouette format, from one frame to the next, summed on the last few frames (4 frames in this experiment). QoM can be assumed as a measure of the global amount of detected motion and it can be thought as a first rough approximation of the physical momentum [8]. This is done separately for each dancer. Condition 4 uses techniques for the tracking of the barycentre of the respective dancer [9].

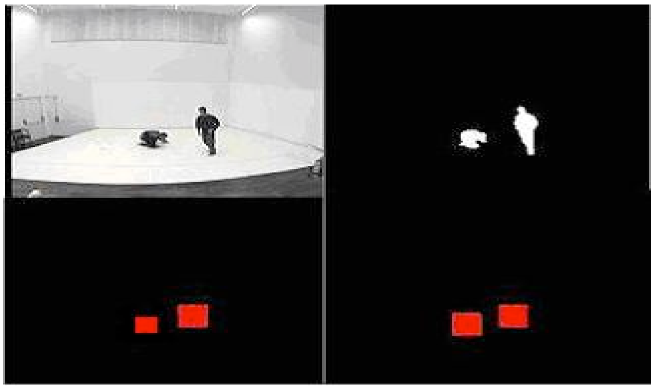

Figure 1: Four experimental conditions 1-4 in clockwise order from the top left.

The experiment was run using EyesWeb on a Dell Precision 450 computer, and presented at full screen on a 20'' monitor. Screen resolution was set at 1600x1200. Participants were sat at a distance of approximately 1m giving a starting visual arc of the dancer in conditions 1 and 2 of 6.05', and 4.6' in condition 3 and 4.

Whilst watching the clip, subjects were instructed to move the computer mouse towards and away from themselves as they thought arousal changed: towards for high arousal, away for low arousal. This vertical position of the mouse cursor on the screen was tracked and used to give an approximate measurement of arousal. The cursor was centred on the screen at the start and the position was recorded at a sample rate of approximately 25 samples/second. A similar method was used by Krumhansl and Schenk to measure emotional engagement via a foot pedal [10]. Afterwards, giving as much detail as possible, they were instructed to write a description of what they saw.

2.2 Participants

36 individuals participated in the experiment, obtained from the undergraduate subject pool, each receiving course credit for their participation.

2.3 Results

The aims of the experiment were to examine the change in arousal as visual information became more reduced, and to examine the free responses for words relating to animacy. We hypothesised that arousal should drop off with decreasing information and that all conditions would result in perception of animacy. Data was normalised and averaged [10], and the changes in average arousal for each stimulus are shown in Figure 2. The values plotted are the averages adjusted to the range from 0 to 1.

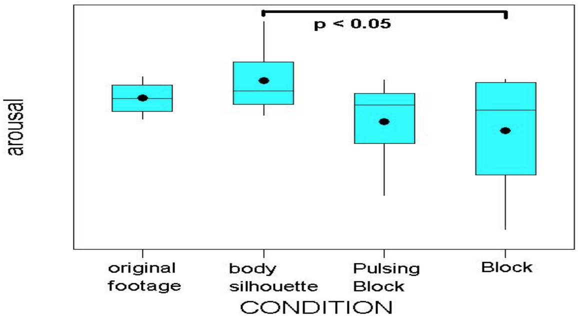

Figure 2: Average Arousal of clip for each condition, showing means and variances.

As shown in figure 2, average arousal between stimuli is similar. An ANOVA reveals there is a significant difference over viewing conditions, $F(3,32) = 3.46$, $p < 0.05$. Posthoc analysis shows that the only significant difference is between the body silhouettes and the abstract block which does not change in size, $F(1,16) = 5.93$, $p < 0.05$. Differing from the expectation of the real video footage being the most arousing, the body silhouette condition had the highest mean arousal.

In the free responses, similar to Heider and Simmel [1], we looked for terms and statements that indicated that subjects had attributed human movements and characteristics to the shapes. These were terms such as touched, chased, followed, and emotions such happy or angry. Other guides to animacy were the shapes generally being described in active roles, as opposed to being controlled in a passive role. By looking for such terms we conclude that conditions 1 and 2 were described in animate terms but conditions 3 and 4 were not. Condition 1 was generally seen and described for what it was, a dance. Condition 2, body silhouettes, resulted in confusion as to the purpose of the movement. Though still seen as human figures, a proportion of participants reported it as some form of fight between the two men or some form of martial arts. Conditions 3 and 4 were not reported using animate words, but were instead reported merely as squares moving passively around in varying degrees of speed.

2.4 Discussion

The experiment was designed to measure changing arousal levels between the varying visual presentations of a human interaction and the use of animate terms to describe them. It was shown that arousal did vary between conditions but, unexpectedly, the body silhouette condition was the most arousing. In relation to animate terms being used, only the full video and body silhouette conditions showed animacy. It was expected that arousal should decrease in accordance with decreasing visual information. Thus the full body image would have been most arousing, then silhouette, etc. Yet, the fact that the body silhouettes were viewed as more arousing may be explained by the free response task. With the silhouette being viewed as a fight, this would coincide with the higher arousal than the full video, described as an abstract dance that did not conform to any commonplace social schema, or the two conditions showing only moving squares. The nature of the dance may also have played an important part in the lack of animate terms in describing conditions 3 and 4.

3. EXPERIMENT 2

In a follow-on from the previous experiment, we investigate if using footage of a solo modern dancer, rather than the original dance footage of the complex interaction, would facilitate the use of animate terms when describing the display. We also examine the emotional engagement between the subject and the display across the four same experimental conditions. We expect that animate terms will be used to describe the real footage and body silhouette conditions and there will be occurrence of animate terms for the remaining conditions. We also expect that emotional engagement will decrease along with available visual information.

3.1 Methods

The same between subject experimental design and technique of stimulus production from the first experiment was repeated. Using a short excerpt of a solo dancer performing a modern dance of duration 37 seconds, the four experimental conditions were created: (1) full video recording; (2) body silhouette; (3) motion of the barycentre of the dancer represented by a single small block changing in size in accordance with the motion energy of the dancer; (4) motion of the barycentre of the dancer represented by a small block that did not change in size (Figure 3).

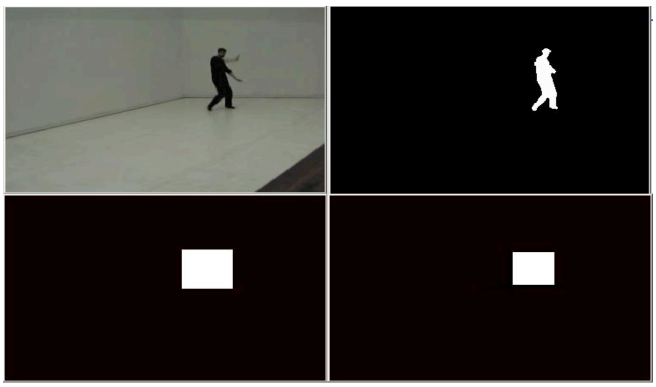

Figure 3: Four experimental conditions 1-4 in clockwise order from the top left.

Screen resolution was again set at 1600x1200, with participants sat at a distance of approximately1m giving a starting visual arc of the dancer in conditions 1 and 2 of 5.73' and 3.72' in conditions 3 and 4.

Similar to the measurement of arousal in the previous experiment, subjects were instructed to move the mouse and therefore the cursor on the screen, towards and away from themselves as their emotional engagement with the clip changed: towards for low engagement, away for high engagement. The vertical position was again used as the approximate measure of emotional engagement. Following the clip, the subjects were again instructed to write a free response of what they perceived.

3.2 Participants

32 individuals were used in the experiment, obtained from the undergraduate subject pool. No incentive was given for participation.

3.3 Results

The aims of the experiment were to examine the change in emotional engagement between the subject and the clip as the visual information available is reduced, and to examine if animate results are achieved when the display is a single dynamic block. Similar to arousal we hypothesised that emotional engagement should reduce with decreasing visual information, and that all conditions of display would result in animacy. Data was again normalised and averaged, and the changes in average emotional engagement for each stimulus are shown in Figure 4. To make the scale consistent with the previous experiment, the data was also reversed.

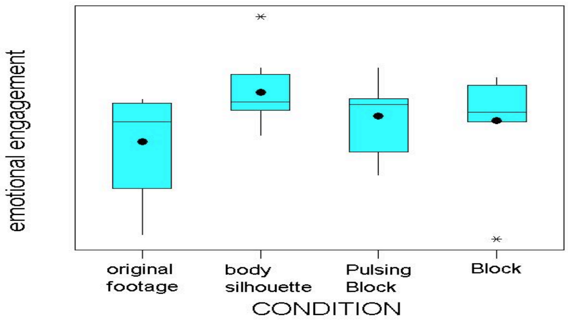

Figure 4. Average emotional engagement of subjects for the clip across conditions, showing means and variances.

Average emotional engagement for subjects viewing the clip was similar across all viewing conditions. An ANOVA showed there was no significant difference between the conditions. As was the case in arousal, body silhouette had the largest mean emotional engagement.

The free response data indicated that the real footage and body silhouette conditions were both described in animate terms. The remaining two conditions were not. Conditions 1 and 2 were quite clearly seen as a man dancing by himself. Conditions 3 and 4, where the figure had been reduced to a white rectangle, were reported as a block moving around the screen in altering speeds and directions. The movement of the block in neither condition 3 or 4 was given any purpose or intention.

3.4 Discussion

One purpose of this experiment was to examine changes in emotional engagement across the varying visual presentations. It was shown that emotional engagement was not significantly affected by visual presentation. However, similar to arousal, highest emotional engagement was found when the clip was presented in the body silhouette format.

The main purpose of the experiment was to examine if changing the display from an abstract complicated interaction to a presentation of a solo dancer resulted in animate terms being used to describe all display conditions. As was the case in experiment 1, only original footage and the body silhouette were described in animate terms, with the two abstract conditions, the pulsing and non-pulsing blocks, being described in geometric terms. Changing the display to the movement of one dancer did not result in the use of animate terms for describing the abstract conditions 3 and 4.

4. GENERAL DISCUSSION

Heider and Simmel [1] showed that a display depicting the movement of three geometric shapes would be described in animate terms, with subjects ascribing thoughts, feelings, emotion and intentions to these blocks in order to make sense of the display. Subsequent research showed that characteristics of the shapes were not so much important, and that it was more the movement of the shapes and their relation to each other that resulted in the phenomenon. Nearly fifty years later, research [4,5] showed that this same phenomenon could be shown using a solitary shape if its trajectory was altered mid-course.

It was our attempt to try and produce similar stimuli that would result in animate responses using a new method of stimulus production. We created the stimuli directly from human movements and interactions by representing the humans as rectangular boxes and altering their sizes in accordance with the motion energy of the respective dancers. The experiment also looked at changes in arousal and emotional engagement but results showed only small changes in both of these measures across the displays, with body silhouettes having the highest average of both.

No animate terms were used to describe the displays of human movement when they were reduced to rectangles. It was thought that this may be due to the complicated nature of the dances in the dyad display, however no animate terms were used when the display showed only one rectangle.

In previous research, Tremoulet and Feldman (2000) showed that movement of a single dot or rectangle would be described more alive if during the course of a trajectory, the shape underwent a change in direction or a change in kinetic energy. The displays used in this experiment and the previous one both contained moments of these two descriptors, yet failed to result in animacy. These results do not argue against the validity of previous research, as quite clearly an object that suddenly and unexplainably changes direction and speed, is likely to be determined controlled by an unseen internal source. Perhaps it is plausible then that, in the case of the displays used in the current research, some human actions are much more subtle than the motion of animated objects such as those used in the research reviewed in the previous introduction. It rather seems the case that when an object lacks human-like qualities, the attribution of animacy relies on an exaggerated display of social behaviour or a limited large deviation in behaviour that cannot be the explained by an external source. Subtle and continuous behaviours, such as that depicted on our displays, do not result in the same kinds of attributions. This is not an unreasonable assumption, since an object blowing in the wind will change direction and speed many times, though when viewed through a window, with no sound or visual cues to the wind, it will still be appropriate to categorise such an object as something blowing in the wind rather than a self-propelled being. Our results therefore suggest that while motion might represent an important cue to animacy, the quality and context of this motion is also important. These features will form the focus of our future research.

5.CONCLUSION

Although the displays share characteristics of displays that have produced animate responses in previous research, it is suggested that previously unidentified characteristics of the displays such as quality and context of the motion also play an important role.

Though animacy was not seen in all conditions, using this new technique of starting with a full body video and transforming it into basic geometric shapes may result in important information being realised about the attribution of animacy and social meaning. This work serves as a springboard for the development of study into creating stimuli that result in animacy using this automatic approach.

ACKNOWLEDGEMENTS

We thank Matteo Ricchetti and the choreographer Giovanni Di Cicco for their help in selecting the dances and the useful discussions.

6. REFERENCES

[1] Heider, F., & Simmel, M., An Experimental Study of Apparent Behavior, American Journal of Psychology, 57:2, (1944).

[2] Bloom, P. & Veres, C. The Perceived Intentionality of Groups, Cognition, 71, B1-B9, (1999).

[3] Scholl, B. J., & Tremoulet, P. Perceptual causality and animacy. Trends in Cognitive Sciences, *4(8)*, 299 - 309. (2000).

[4] Tremoulet, P.D., & Feldman, J. Perception of animacy from the motion of a single object, Perception, 29, 943-951, (2000).

[5] Gelman, R., Durgin, F. & Kaufman., L., Distinguishing between animate and inanimates: not by motion alone. In Casual Cognition (ed. Sperber., D., Premack, D. & Premack, A.J.) Oxford: Clarendon Press, pp. 151-184, (1995),

[6] Camurri, A., Trocca, R., &Volpe, G., Interactive systems Design: A KANSEI-based Approach. Submitted full paper for NIME 2002, Dublin. (2002).

[7] Camurri A., Mazzarino B. & Volpe G. Analysis of Expressive Gesture: The EyesWeb Expressive Gesture Processing Library, in A. Camurri, G. Volpe (Eds.), Gesture-based Communication in Human-Computer Interaction, LNAI 2915, Springer Verlag. (2004).

[8] Camurri A., Lagerlöf I & Volpe G., Recognizing emotion from dance movement: Comparison of spectator recognition and automated techniques, International Journal of Human-Computer Studies, vol. 59, 213-225. (2003).

[9] Camurri A., Krumhansl C. L., Mazzarino B. & Volpe G. An exploratory study of anticipating human movement in dance, in the Proceedings of this same Conference. (2004).

[10] Krumhansl C. L. & Schenk, D.L., Can dancr reflect the structural and expressive qualities of music? A perceptual experiment on Balanchine's choreography of Mozart's Divertimento No. 15. Musicae Scientiae, 1, 63-85. (1997).

Authors

(1)Phil McAleer, Department of Psychology, University of Glasgow, 58 Hillhead St, G12 8QB, (0044)(0)141-330-3395, phil@psy.gla.ac.uk

(2)Barbara Mazzarino, Infomus Lab, D.I.S.T., University of Genova, Viale Causa, 13, 16145 Genova, Italia, +39-010-3532201, bunny@infomu.dist.unige.it

(3)Gaulteiro Volpe, Infomus Lab, D.I.S.T., University of Genova, Viale Causa, 13, 16145 Genova, Italia, +39-010-3532201, volpe@infomus.dist.unige.it

(4)Antonio Camurri, Infomus Lab, D.I.S.T., University of Genova, Viale Causa, 13, 16145 Genova, Italia, +39-010-3532201, toni@infomus.dist.unige.it

(5)Helena Patterson, Department of Psychology, University of Glasgow, 58 Hillhead St, G12 8QB, helena@psy.gla.ac.uk

(6)Kirsty Smith, Department of Psychology, University of Glasgow, 58 Hillhead St, G12 8QB, 9904460S@student.gla.ac.uk>

(7)Frank E. Pollick, Department of Psychology, University of Glasgow, 58 Hillhead St, G12 8QB, frank@psy.gla.ac.uk