Research Paper Digests

bowtie

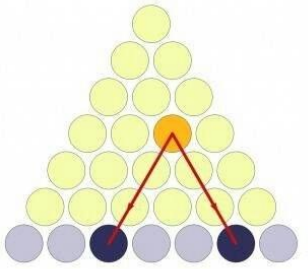
Entertainment Technology Center

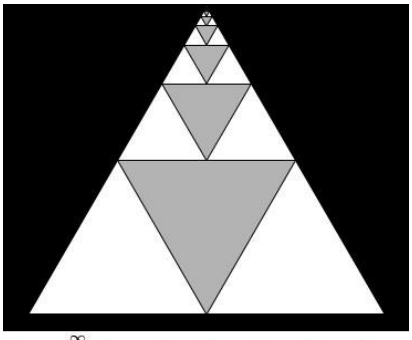
Contents

Preface

Before delving into robotics, just want to warm up by sharing two interesting ‘proof without words’. If you are to prove the two math equations below, what method will you use? The two graphs smartly proved it without using a single word, and more importantly, making it seems so self-evident.

$$1+2+\cdots+(n-1)=\binom{n}{2}$$





$$\frac{1}{3}=\sum_{i=1}^{\infty}\frac{1}{4^i}=\frac{1}{4}+\frac{1}{4^2}+\frac{1}{4^3}+\frac{1}{4^4}+\frac{1}{4^5}\cdots$$

Like a Chinese proverb goes, “A picture is worth a thousand words”, visual expression is sometimes much stronger than verbal expression. And yet the mechanism human visual channel interpret things is always mysterious even in today’s psychological and computational forefront studies, especially involving interpretations of human emotions and robotic motions.

After researching on a lot of papers, it is found that some concepts are frequently mentioned in the realm of study around robotic expression of emotions. For example the Uncanny Valley , Six Basic Emotions are frequently being cited among all papers in this field. In this summary, you will be able to learn about these concepts.

Also you will be able to learn a lot of methodologies of scientific research. For example using questionnaire test and error matrix for emotion classification in ‘The effect of posture and dynamics on the perception of emotion’, the histogram in ‘Emotional Body Language Displayed by Artificial Agents’ and so on.

1. The effect of posture and dynamics on the perception of emotion

The six basic emotions, upper body expression, and blending with neutral pose

Six basic emotions have great implications on HERB. On ETC bowtie's HERB project, we try to translate the six basic emotions anger, disgust, fear, happiness, sadness, and surprise into HERB movements. Although HERB doesn't have facial expression and has limited lower body movements, The papers studied partial occlusion of part of the body, and corresponding effects for emotion perceptions. For each of the six emotions, they chose the two clips with the highest recognition rates. They then occluded different parts of the body: the head motion (or NH for "No Head motion"), the lower body motion (NL), and the upper body motion (NU). The unaltered motion is labeled OR for "original". We did not alter the root motion for any of the conditions to avoid very unnatural motions that could affect the ratings in unintended ways.

In this paper, three experiments were performed to gain a better understanding of how motion editing might affect the emotional content of a motion-captured performance, particularly two most important factors shown to be important bodily emotions-posture and dynamics. In these studies, they analyzed the properties and perception of a varied set of full-body motion clips representing the six emotions, namely anger, disgust, fear, happiness, sadness, and surprise. It is found that emotions are mostly conveyed through the upper body, that the perceived intensity of an emotion can be reduced by blending with a neutral motion, and that posture changes can alter the perceived emotion but subtle changes in dynamics only alter the intensity.

Figure 1: *Characteristic frame for each emotion from the clips with the best recognition rates.*

The author investigates which aspects of body language are important for conveying emotions with two goals in mind. The first goal is to understand how changes introduced by motion editing might alter the emotional content of a motion, so we can ensure that important aspects of a performance are preserved. The second goal is to gain insight on how we may edit captured motion to change its emotional content, further increasing its reusability. The author studies six basic emotions shown to be readily recognized

across cultures: anger, disgust, fear, happiness, sadness, and surprise (Figure 1). The motion of a character's body can effectively express all basic emotions.

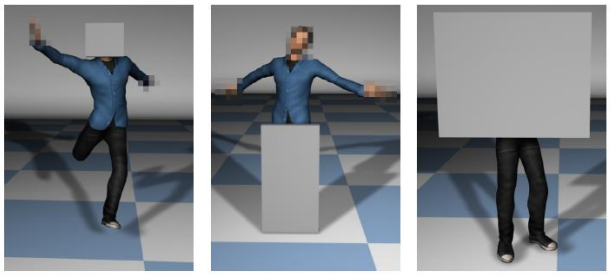


Figure 5: *Stimuli examples from our second experiment in which we hid either the head, the lower body, or the upper body.*

The paper reached the conclusion that the upper body is crucial for the perception of emotions. The lower body or the head alone were not relevant in our set of clips to recognize the emotion. The irrelevance of the head for all emotions except sadness could have been due to our blurring of the head in the baseline clips: when we occluded the head entirely, there was no considerable impact. Alternatively, this finding might have to be due to the head being unimportant for recognizing the emotion in nearly all our clips. However, the relatively high error rate for sadness when the head was occluded complies with previous work that head motion is particularly important for displaying sadness. The research showed that differences between the occluded recognition rates were smallest for two of the emotions that displayed very distinct lower body motions, namely fear and anger.

The paper takes the research further by investigating the Body Blend and its effects. (Figure 8) Although animation blending is not directly applicable to a physical robot like HERB, but this idea leads to our teammates Jon Lew's design of Openness as a metric of movement trajectories.

Figure 8: *Body blend (BB) condition. The upper body joints are blended with a neutral motion having the arms at the side. The above example shows the result of blending 50% of the original motion with 50% of the neutral motion. Joint rotations are represented using quaternions and blended with slerp. This condition changes both pose and velocity.*

From these, It's learned that the upper body motion is most crucial for the recognition of emotions, that changes to posture can change the perceived motion type whereas changes to dynamics can change the perceived intensity, and that the perceived intensity of an emotion can be reduced by blending with a neutral motion.

2. Emotional Body Language Displayed by Artificial Agents

The uncanny valley and its implications to HERB

The uncanny valley, frequently discussed in the realm of robotic expressions of emotions, is an phenomenon, that the closer a robot looks like human, there exists a sudden drop of believability . HERB is quite non-human like, compared to humanoids like Sony's QRIO, but we should avoid totally translating human upper body movement onto HERB's two arm movements, rather, it will be good if we put what's people's perception of an HERB animation first. To put it simple, we should create the way HERB is sad, than directly translating the way human is sad onto HERB.

Since complex and natural social interaction between humans and artificial agents necessitates the display of rich emotions in order to be believable, socially relevant, and accepted, and to generate the natural emotional responses that humans show in the context of social interaction, such as engagement or empathy. Whereas some robots use faces to display (simplified) emotional expressions, for other robots such as Nao or out robot HERB, body language is the best medium available given their inability to convey facial expressions.

Indeed, it is not evident that expressions displayed by a human and by an agent are interpreted in a similar way. Encouraging results have, however, been found from a much weaker stimulus. Using restricted technology, it was found that humans tend to interact with computers as they do with real people [Nass and Moon 2000]. However, it cannot be simply assumed that greater fidelity would improve the interaction. As agents become more visually realistic, they are confronted with the well-known Uncanny Valley problem [Mori 1970]. The Uncanny Valley (Figure 1) models a drop in believability as agents acquire greater visual similarity with humans [Brenton et al. 2005]. The concept was first introduced in robotics, where it was reported that highly realistic humanoid robots tend to be found repulsive [Mori 1970].

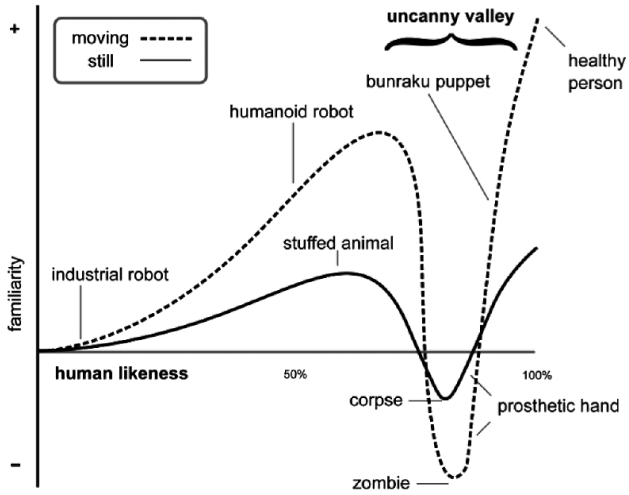


Fig. 1. The Uncanny Valley (adapted from Mori [1970]).

In recent years, improvement in the field of animation technology has increased the level of visual realism that can be achieved. In the context of animation, visual realism is defined in terms of physical realism (i.e., how similar to a human the character looks) and behavioral realism (i.e., how similar to a human the movements are). The expectation was that as visual realism increases so should believability.

This phenomenon also exists in the world of film and animation industry. As discussed in the paper, animated characters from the film industry have also been confronted with the same drop in believability as described by the Uncanny Valley. For example, the characters from The Polar Express [Zemeckis 2004] or Final Fantasy [Sakaguchi and Sakakibara 2001] have failed to be convincing [MacDorman et al. 2009]. Unrealistic characters, such as the ones used in traditional cel animation, were not confronted with this issue as they generated empathy. Further discussions on the concept of the Uncanny Valley can be found elsewhere [Dautenhahn and Hurford 2006]. However, the Uncanny Valley is not grounded on systematic studies and its very existence is still subject to debate. Moreover, existing studies seem far too simple to fully explore the full complexity of the problem. The causes of it might not be straightforward, and might involve a complex combination of all sorts of contextual, cultural, social, and other factors [Ca ˜namero 2006]. Moreover, part of the uncanny effect could be due to poor design or disappointing storylines. The Uncanny Valley may also result from issues with the display of body language or, as suggested by Brenton et al. [2005], from poor facial animation. Traditional animation which created a wide range of credible characters, avoiding the Uncanny Valley, highlights the importance of displaying appropriate emotional behaviors [Bates 1994; Thomas and Johnston 1995].

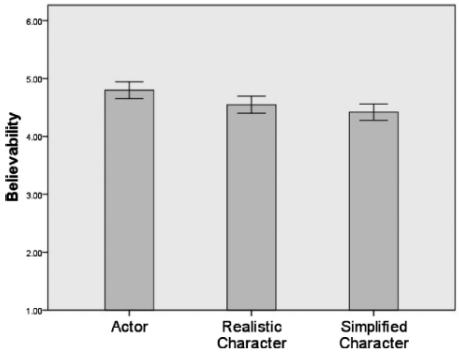


Fig. 5. The overall believability by character type (means + 2*standard error).

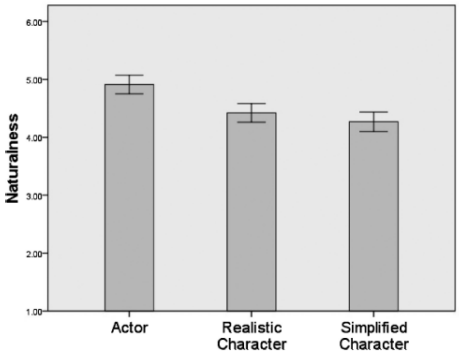


Fig. 6. The overall *Naturalness* by *Character Type* (Means + 2*Standard Error).

The study was advertised as a study on emotional body language, and no mention of the presence of animated characters was made to participants prior to the experiment. All participants were tested by the same experimenter in individual sessions. Each session began by obtaining consent, followed by the emotional intelligence questionnaire. After completion of the questionnaire, participants were told how to use the software and given an explanation of the GEW. The term "believability" was clearly defined as "to what extent do you think the character is feeling the emotion" and naturalness was defined as "the quality of the way the character moves." Participants were informed that faces and hands were blurred before they watched and assessed the 63 video clips—(10 emotions +1 neutral state) ∗ 3 character types. Each video was played through only once. Then participants responded, which triggered the next video. When all video clips were interpreted, the post study questionnaire automatically started. Finally, participants were fully debriefed regarding the purpose of the study. The whole procedure took less than one hour.

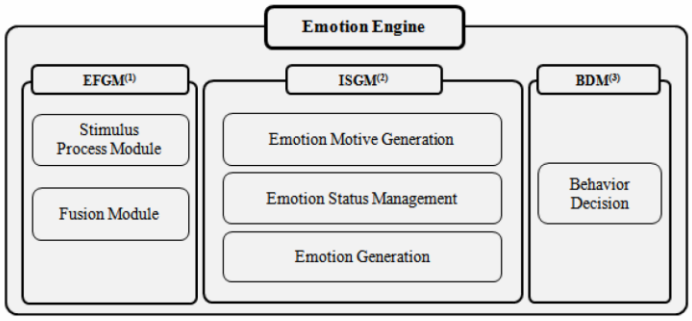
3. The Emotion Expression Robot through the Affective Interaction: KOBIE

A novel emotion expression system

Although in the scope of this project, HERB's move is pretty puppetry, but the research in this paper provides an example of how to further the responsive mechanism into a intelligent being by building a emotion expression system, comprising need model, emotion engine, mood model and so on.

In this paper, the researchers proposed the method for expressing emotion for the emotional robot through the affective stimuli such as ‘hit,’ ‘stoke,’ ‘tickle,’ ‘poke,’ ‘embrace,’ etc. They developed the experiment platform, KOBIE, for experiment the emotion expression system. The KOBIE is an emotional robot which is made for interaction between a human being and a robot through the affective interaction. The system can be used in developing an avatar having an emotion or developing the apparatus having an emotion in the ubiquitous environment.

The emotion expression system is comprised of emotion feature information collector component, internal status management component, and action expression component. Figure below shows the architecture of emotion engine. The emotion engine is the main component of the emotion expression system. The emotion engine is composed of the EFGM (Emotion Feature Generation Module), the ISGM (Internal Status Generation Module) and BDM (Behavior Decision Module). In brief, the EFGM, generate emotion features, is composed of the Stimulus Process Module and Fusion Module. The ISGM generates emotions and manages the internal status information. The BDM determines the behavior for emotion expression.



EFGM(1) : Emotion Feature Generation Module
ISGM(2) : Internal Status Generation Module
BDM(3) : Behavior Decision Module

Needs model is based on the Maslow’s Hierarchy of Needs. The emotion need parameters are constituted by a 5-step layer including physiological needs, safety needs, love and belongingness needs, esteem needs, and self-actualization needs. As shown in Fig.2, so that each of the emotion needs parameters can maintain the equilibrium range of an emotion in the activation regions of an emotional needs motive, the emotion motive generator satisfies emotional needs motives while changing a satisfaction level through expression of emotions and emotional actions.

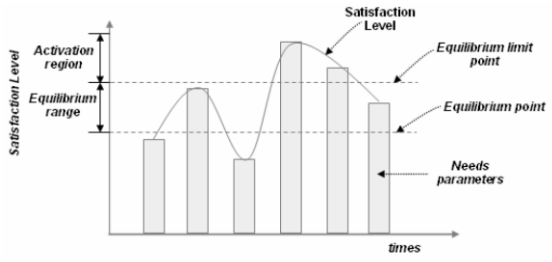


Fig.2 An example of needs-based emotion motive generation

The mood model has six parameters to determine the change of moods through internal/external stimuli. The mood parameters are as follows, ‘happy,’ ‘gloomy,’ ‘comfort,’ ‘irritable,’ ‘listless’ and ‘depressed’. The moods parameters affects the emotional motive generation. Also, moods parameters maintain previous robot status information.

The emotion motive is affected by the results that keep to changes of the emotional parameters. First, the needs parameter based emotions are generated in the priority order of physiological needs, safety needs, and love and belongingness needs, and the emotion parameter based emotions are generated sequentially and repeatedly in a predetermined priority order. The emotional motive generator determines the motive of a corresponding emotion on the basis of the current state information about the intelligent robot that is received from the EFGM. That is, the emotional motive generator determines an emotional motive according to the level of satisfaction with needs-based needs parameter, which is based on the internal stimulus information received from the internal sensor, and with a state-information-based emotion parameter, which is based on the external stimulus information received from the external sensor.

4. Investigating the Role of Body Shape on the Perception of Emotion

Forms, Motions and Emotions

Although HERB has a strong physical ‘form’ which it cannot get rid of, this paper’s study about the presence/absence of forms and motions , and their relationships could give us more insights about human perception of humanoid or non-humanoid motions and how emotion are generated.

In order to analyze the emotional content of motions portrayed by different characters, we created real and virtual replicas of an actor exhibiting six basic emotions: sadness, happiness, surprise, fear, anger, and disgust. In addition to the video of the real actor, his actions were applied to five virtual body shapes: a low- and high-resolution virtual counterpart, a cartoon-like character, a wooden mannequin, and a zombie-like character (Figures 1 and 2). In a point light condition, we also tested whether the absence of a body affected the perceived emotion of the movements. Participants were asked to rate the actions based on a list of 41 more complex emotions. It’s found that the perception of emotional actions is highly robust and to the most part independent of the character’s body, so long as form is present. When motion alone is present, emotions were generally perceived as less intense than in the cases where form was present.

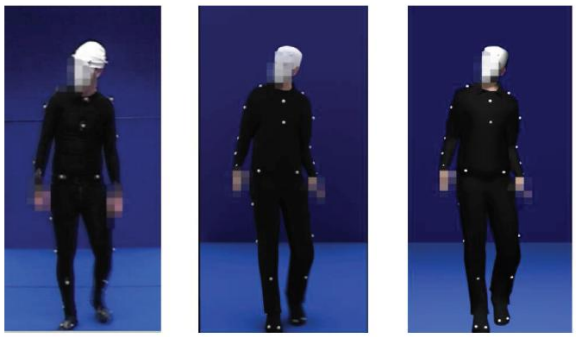


Fig. 1. Image taken from real video, high-resolution virtual male, and low-resolution virtual male.

Fig. 2. Image taken from wooden mannequin, toon, zombie, and points.

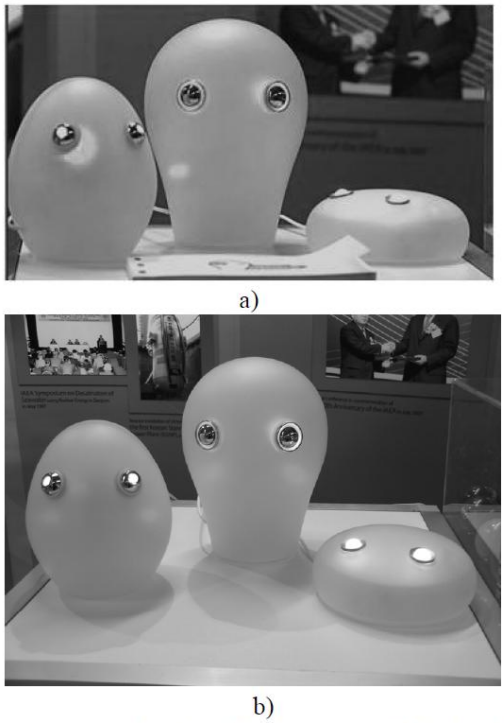
A cartoon-like figure was chosen in order to determine if his friendly appearance would lead him to be perceived as exhibiting more positive emotions, whereas we felt that a zombie-like character would be perceived in a more negative light. A wooden mannequin model was chosen as a completely neutral character that we felt may be perceived to be exhibiting less intense emotions than the others.

Virtual characters in animated movies and games can be very expressive and have the ability to convey complex emotions. However, it has been suggested that the more anthropomorphic a humanoid robot or virtual character becomes, the more eerie it is perceived to be [Mori 1970]. In this article,the focus is on one potential factor that could affect this “uncanny valley” theory, by examining how emotional actions on real and virtual characters are perceived.

5. Evaluation of the expressions of robotic emotions of the emotional robot, "Mung"

Colored LED which helps with expression of emotions

Although in the scope of this project, HERB does not have different color LED light, but HERB is very extensible, you can add almost any desired hardware and software via the basic services by the ROS. This papers provides and interesting look on the color of LED, bruises (blue or dark blue) and complexions (red or pink), as a means for expressing emotions, just like the robot QUASI colorful antenna. So some of the conclusions could be useful if in future we are to add LEDs to HERB.



a)

b)

Figure 1: The developed robots as exhibited in the IAEA 51st Annual General Conference: a) bruised robots such as having a blue eye b) blushing robots such as red cheeks.

This study verifies the effectiveness of the robotic emotional expressions developed using non-verbal languages. An emotional robot known as "Mung" was developed by applying the nonverbal expressions of bruises (blue or dark blue) and complexions (red or pink) using full color LEDs, to express robotic emotions. With these emotional expressions, Mung functioned as a language purifier by responding to humans' use of language emotionally. In order to evaluate emotional expressions used in Mung, an experiment was performed to compare the robotic emotions using a verbal language with those using non-

verbal languages. A 2x2 between-groups factorial design was used for this experiment. The between-groups factors were expression type and emotion type. According to the results of a post-experiment survey, emotional expressions through bruises and complexion are as effective as speech for the delivery of robotic emotions. This result emphasizes the effectiveness of the developed robotic emotional expressions.