Evaluation of the Uncanny Valley in CG Characters

Laura M. Flach
Rafael H. de Moura
Soraia R. Musse
Virtual Humans Simulation Lab
FACIN/PUCRS

Vanderson Dill
Márcio S. Pinho
Virtual Reality Group
FACIN/PUCRS

Christian Lykawka
FACIN/PUCRS

Abstract

This paper revisits the Uncanny Valley subject in order to evaluate its effects on the perception of characters that currently use computer graphics, animation and computer simulation. It is based on surveys that generated hundreds of samples and showed preliminary results about new criteria and correlations regarding the familiarity of the characters with the public. The analysis of familiarity in those characters showed great agreement with the original curve of the uncanny valley. Future enhancements are then suggested to add new dimensions to the original graph.

Keywords:: Uncanny Valley, Computer Graphics, Modern Media, Movies, Games

Author's Contact:

{laura.flach,vanderson.dill}@acad.pucrs.br

1 Introduction

This paper studies the effects of the uncanny valley caused by CG characters, first presented by Mori in 1970 [Mori 1970] for robots. According to his work, robots made too similar to real humans can fall into the "uncanny valley", where too high degree of human realism evokes an unpleasant impression in the viewer. The increasing exposure of virtual characters to the general public motivated our group to test the original hypothesis using more up-to-dated samples from movies and games[1].

Our study is an evaluation on how people perceive CG characters that are presented in modern digital media. We try to answer: Does the uncanny valley exists in CG characters? Does previously know the character affects the empathy with it? Moreover, does adding movement to this characters changes the shape of the uncanny valley curve, like Mori [Mori 1970] suggested?

[1]Images from movies and games where used here for scientific purpose.

To answer these questions, we proposed an evaluation methodology based on a questionnaire with subjects. First, we selected representative characters from recent and well known movies and games, and other characters that are unknown. In order to evaluate the public's empathy with the chosen characters, we conducted a two-stages questionnaire. In the first stage we presented still images and, in the second stage, videos with the performance of the same characters in the same scenes.

As the results, more than two hundred of analysed samples leaded to conclusions that indicated the uncanny valley hypothesis is valid for CG characters. The following sections cover the related works on the subject, the model for evaluation (including the details of both surveys), obtained results and final considerations.

2 Related Work

In 1970, the roboticist Masahiro Mori [Mori 1970] published a theory on how humans react emotionally to artificial beings. According to Mori, robots should not be made too similar to real humans because such robots can fall into the "uncanny valley", where too high degree of human realism evokes an unpleasant impression in the viewer. Mori plotted the industrial robots on a hypothetical graph of familiarity versus appearance (Figure 1).

For creatures, including robots, movement is generally a sign of life [Mori 1970]. As show in Figure 1, adding movement changes the shape of the uncanny valley by exaggerating the peaks and valley. To reproduce human like movements, similarity of velocity and acceleration are required. For example, a robot has 29 artificial muscles in the face to make human like facial expressions, and laughing is a kind of sequence of face distortions, where the distortion speed is an important factor. If we cut the speed in half, laughing looks unnatural. This illustrates how slight variations in movement can cause a robot, puppet, or prosthetic hand to tumble down into the uncanny valley.

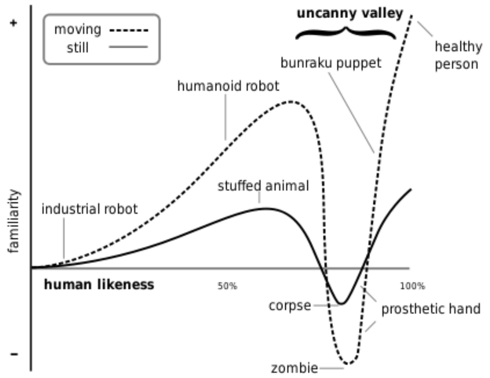

Figure 1: *This graph represents the hypothesized emotional response of human subjects plotted against anthropomorphism of a robot. The uncanny valley is the region of negative emotional response towards robots that seem "almost human". Also, the movement amplifies the emotional response [Mori 1970].*

There are several theories that were proposed to explain the cognitive mechanism underlying the uncanny valley phenomenon. Some of the possible answers found in the literature ([MacDorman and Ishiguro 2006], [Karl F. MacDorman 2009]) are:

- Mate selection: Automatic, stimulus-driven appraisals of uncanny stimuli elicit aversion by activating and evolved cognitive mechanism for the avoidance of selecting mates with low fertility, poor hormonal health, or ineffective immune systems based on visible features of the face and body that are predictive of those traits [Green et al. 2008], [Donovan 2002].
- Mortality salience: Viewing an "uncanny" robot elicits an innate fear of death. Plays on our subconscious fears of reduction, replacement, and annihilation [MacDorman and Ishiguro 2006].
- Pathogen avoidance: Uncanny stimuli may activate a cognitive mechanism that originally evolved to motivate the avoidance of potential sources of pathogens by eliciting a disgust response. "The more human an organism looks, the stronger the aversion to its defects" [Donovan 2002], [Karl F. MacDorman 2009].
- Violation of human norms: The uncanny valley may "be symptomatic of entities that elicit a model of a human other but do not measure up to it" [MacDorman and Ishiguro 2006]. If an entity looks sufficiently non-human, its human characteristics will be noticeable, generating empathy. However, if the entity looks almost human, it will elicit our model of a human other and its detailed normative expectations. The non-human characteristics will be noticeable, giving the human viewer a sense of strangeness.

Several studies have started empirical testing of the uncanny valley theory. Both Hanson [Hanson 2006] and MacDorman [MacDorman 2006] created a series of pictures by morphing a robot to a human being. This methods appears useful, since it is difficult to gather enough stimuli of highly human-like robots. However, it can be very difficult, if not impossible, for the morphing algorithm to create meaningful blends.

With the same idea, Seyama and Nagayama [Seyama 2007] investigated the uncanny valley by measuring observers' ratings of pleasantness of static facial images, whose degree of realism was manipulated by morphing faces of dolls, masks, and CG characters into real human faces. Their results showed that the uncanny valley emerged only when the face images involved abnormal features such as bizarre eyes.

MacDorman *et. al.* [MacDorman et al. 2009], presents four empirical studies that explores these issues related to the uncanny valley in still images. In each one, they tested different features in the face of a CG character, like skin texture, level of detail, proportions, eye positions. The main findings of the studies were that the more human the photorealisticaly-textured CG face looked, the easier it was for people to agree on its degree of human likeness. Also, as human likeness increased, the best-looking facial proportions were generally closer to the original proportions of the human model. Another interesting finding is that facial proportions that are far from ideal look eerier at higher levels of details that facial proportions that are nearly ideal. Besides, the CG face looked less eerie when the texture of the eyes and skin were at a similar level of photorealism that when their level of photorealism differed greatly. (For more details, see [MacDorman et al. 2009]).

Wallraven *et. al.* [Wallraven et al. 2008], provide an experimental framework for evaluating facial expressions, and presents two detailed evaluation results, using their framework, of facial animation techniques. Also does a investigation of important information channels in the visual processing of facial expressions, and demonstrate how the experimental results can be used to improve the perceptual quality of animation techniques. The most important insights which can be gained from their research are that accurate simulation of facial motion (rigid head motion and non-rigid deformations) is important for recognizing animated faces fast and accurately, and for perceiving intensity, sincerity and typicality of animated faces. The perceived intensity of animated facial expressions can be increased by enhancing shape and motion information and including eye motion. As long as motion information is available, degrading shapes and texture does not have a noticeable effect on recognition (at least not for the experimental settings used).

Chaminade *et. al.* [Chaminade et al. 2007], used a biological motion classification task to investigate the influence of discrete variations of animated characters' appearance on the perception of its actions. The participants had to classify a short running motion as 'Biological' or 'Artificial', corresponding to the motion capture and key-framing techniques used to animate the characters. They reported that the response of perceiving a motion as biological decreases with anthropomorphism of the characters used to render the motion, and a fMRI investigation found that the response bias towards "biological" is correlated with an increase of activity in regions of the brain involved in mentalizing and a decrease of activity in regions belonging to the mirror system. This results suggests that the actions of characters favouring mentalizing, not motor resonance, are perceived as biological.

Our study is a evaluation on how people perceive CG characters that are present in daily routine of using modern digital media, like watching movies and playing electronic games. The difference from previous evaluations of uncanny valley is the use of CG characters that are "normal human beings", i.e. without morphing or any other operation that could lead in abnormal features. The characters used here were created independently and chosen for cover the widest range possible in terms of human likeness level.

3 Model for Evaluation

The study was guided by a main question: How do people perceive Computer Graphics-made characters? To answer this question, a survey based on a questionnaire was taken. The questionnaire was made to answer some primitive questions: Can we perceive the uncanny valley effect in CG characters? Does it make any difference in a person's perception to previously know the character? Does static images and video lead different perceptions? Additionally, in case a person felt discomfort by looking to a character, which parts of the face most influenced in that discomfort? The following topics presents detailed descriptions of how the characters and the questions were chosen, and how the survey itself was conducted.

3.1 The Characters

A central issue in the construction of an appropriate question form is the choice of the characters being presented. Considering the questions that must be covered, we tried to choose a not very large number of characters to avoid a too long form. Criteria for choosing the characters were set in order to obtain a group that allowed us to cover all the questions above. The first criterion was the human likeness of each character. This feature is necessary because it is present in the uncanny valley graph (the horizontal axis). The second criterion was if the characters are or not well-known, to answer the question about previously knowing the character. Is important to note here, that it wasn't necessary to precisely define these features nor to previously order the characters according to them. The goal was to cover the widest range of features possible.

To cover the human likeness feature, we chosen characters intended to represent accurately a human, like "Emily" (C.7 in Figure 2). Cartoon characters were chosen to represent the minimum of human likeness, like "Mr. Incredible" (C.3). To cover many intermediate possibilities in this range, were chosen characters like "Obama's" cartoon (C.2) and the moustache man from the movie "Cloudy with a Chance of Meatballs" (C.13).

To evaluate if the character is well-known or not it was taken into account the origin of the character. So, we choosed characters from cinema, like the character of the "Beowulf" movie (C.6), games, like C.11 from "Heavy Rain" and C.4 from "Half Life", and others from the Computer Graphics community. In general, the population doesn't know the characters C.1, C.5, C.7, C.14, C.15, C.16 and C.17. We assume here that cinema is a more comprehensive media than games, older people play less electronic games than watch movies. Also we include in the set characters based on real people, like "Clint Eastwood" (C.12) and "Obama" (C.2).

Some constraints were considered when choosing the characters. The expression in the video needed to be neutral, i.e. no strong emotions should be shown by the character. The character had to be a human being and must be in a

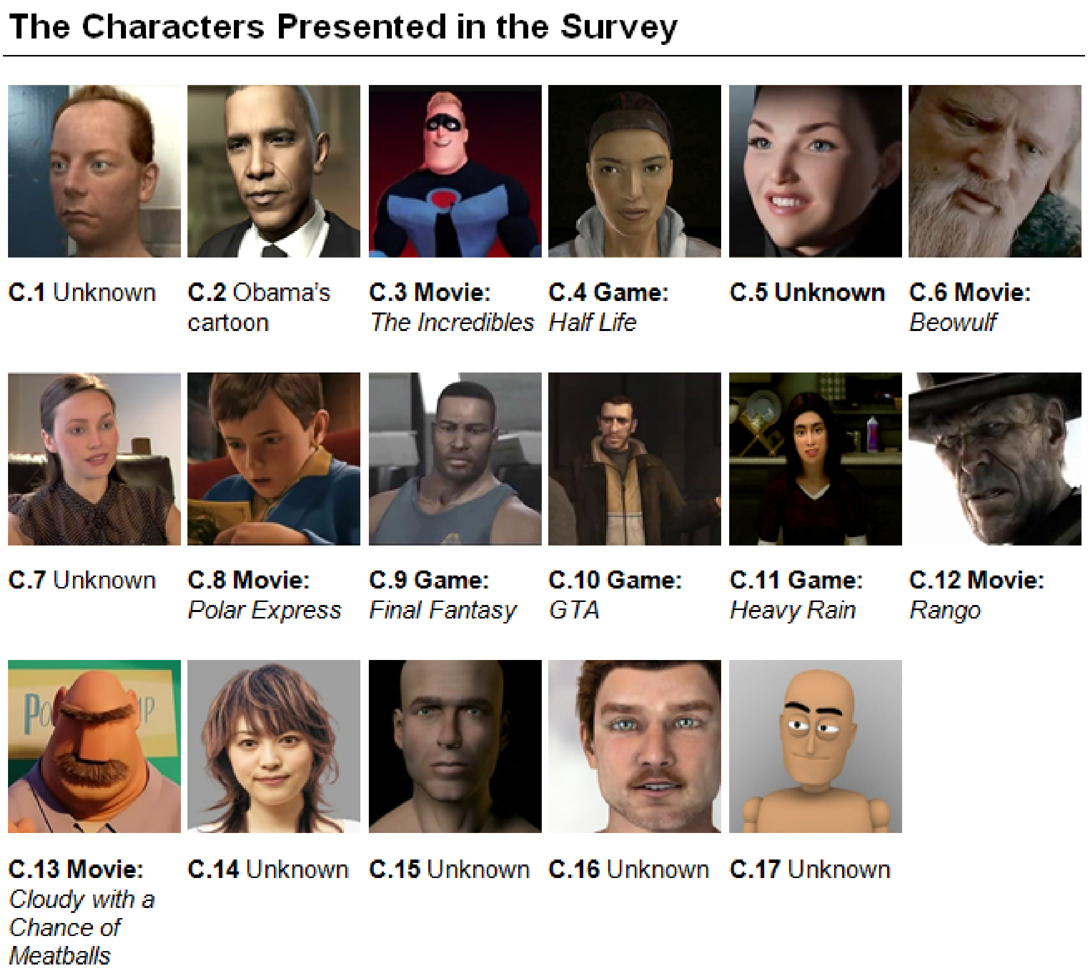

Figure 2: *The Characters presented in the survey and the origin of them.*

natural environment, in other words, had to be in some feasible place at the real world. Moreover, it should be wearing normal clothes for human-beings, without costumes or belongings that could distort the perception. All of these things could bias the perception of the character, and we tried to avoid that. Figure 2 shows the complete list of characters used in the survey. In each one is described an additional information about the origin of the character. This can be a movie, a game or an unknown origin.

3.2 The Questionnaire

Once the characters were defined, a properly questionnaire about them had to be created. The questionnaire must cover the questions that guided our study presented on Section 3. The participant had to indicate its feelings about the character in a subjective manner, since highly "psychological" issues were intended to be gathered. This indicates, firstly, that the participant should not know about the original intention of the survey. On the other hand the questions needed to be the more specific as possible, so that the participant wouldn't misunderstand it. The entire questionnaire was presented in the native language of the participants. In Table 1 the questions are presented in English.

Based on the goals of our study, each participant was asked about: if it is known or not, if some strangeness or discomfort (uncanny) was felt when looking to the character. In the last

Question	Alternatives	Type Of Question
1. Do you think that the character in the picture above is:	a. A real person b. Created with CG c. Don't know	Single choice
2. If created with CG, how realistic does it seem?	a. Very realistic b. Moderately realistic c. Unrealistic d. Don't know	Single choice
3. Do you know this character?	a. Yes b. No c. Don't know	Single choice
4. How do you would describe it?	a. Sympathetic b. Antipathetic c. Don't know	Single choice
5. Do you feel some discomfort (strangeness) looking to this character?	a. Yes b. No c. Don't know	Single choice
6. In which parts of the face do you feel more strangeness?	a. Region of the eyes b. Region of the mouth c. Region of the nose d. Hair d. Other	Multiple choice

Table 1: *The questions of the survey's form*

question also was asked which parts of the face caused the most strangeness.

The questions 1 and 2 were asked with the intention of capture the level of human likeness (how realistic is the character) in the user's perception. The order of the characters in the horizontal axis of the uncanny valley graph, that will be presented in Section 4, was defined by the answers of this two questions.

The question 4 was asked to evaluate if the perception of the characters personality (if it was a villain, antipathetic, or if it was sympathetic) influences in the uncanny valley effect. The answers of this question didn't lead us to any conclusion, so we discarded this possibility.

3.3 The Survey

The survey was conducted in two steps: the first with still images and the second with videos, to evaluate the difference of the perceptions that the two obtained. The images of the first step were taken from the videos used in the second step. Both steps were performed using the questionnaire presented. The online form was released broadly via email, social networks and with colleagues. Each form was made available during about one week. The first form obtained 115 responses and the second obtained 95 responses.

In the second form, some of the characters were removed to make it smaller and less tiring, in response of many suggestions in the first survey. The elimination was made in such a way that all the features of interest kept represented in the second set of characters. For each of the features: realism, popularity, empathy and strangeness at least one character was preserved, based on the responses of the first survey. So, the second form was reduced to 10 characters, who were: C.1, C.2, C.3, C.5, C.6, C.11, C.12, C.13, C.14 and C.16. Since the original set had more realistic characters than cartoons, the removed ones were all highly realistic.

4 Obtained Results

In the next paragraphs the main conclusions reached in this study are highlighted. The resulting answers of the two stages of the survey are condensed in the graphs presented below. To facilitate the comparison, the characters are ordered in all the graphs by human likeness (based on the answers of question 4 from the questionnaire in the Table 1).

The graph presented in Figure 4 shows the percentage of people that felt some discomfort in each character, for still images and videos. People were asked for "Do you feel some discomfort looking to this character?" the possible answers

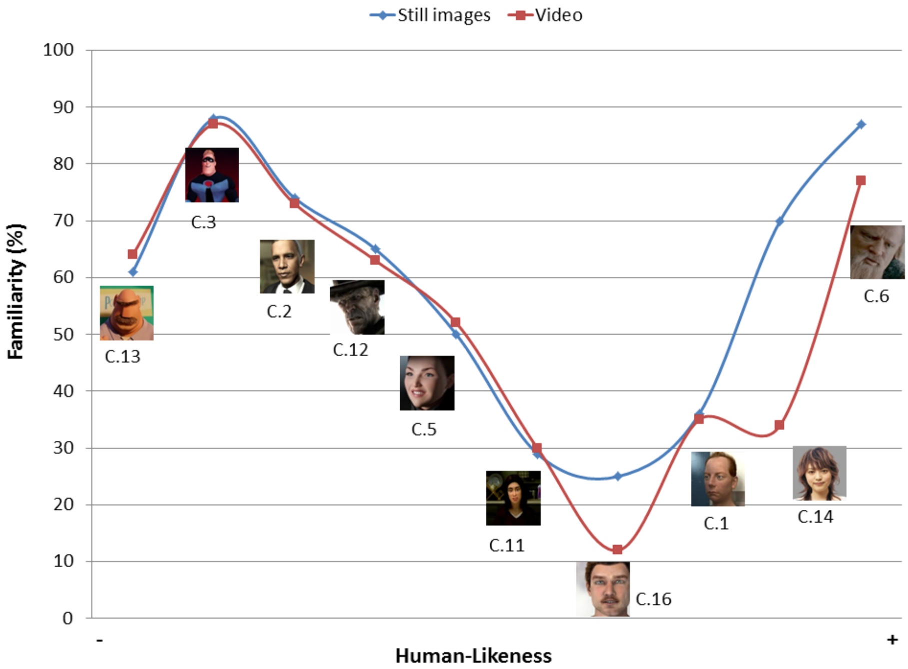

Figure 3: *Graph reproducing the original Uncanny Valley curve from Mori [Mori 1970]. In vertical axis, the percentage of people that didn't feel discomfort in each character. In horizontal axis the "human likeness" of the character.*

were "Yes", "No" or "Don't Know". The answers of both surveys show good distribution in terms of discomfort level, i.e. characters with low, high and intermediate levels of discomfort are present. This confirms that the choice of the characters selected for analysis was good, since we cover all the range of discomfort.

The level of discomfort felt was similar in still images and videos, except for C.14, where much greater level of discomfort was felt in the video. This character shows a high level of human likeness and, in contrast, was poorly animated (its movements are far from natural). This can lead us to a hypothesis that this contrast of a highly natural human model and strange movements can generate discomfort.

The graph of Figure 3 follows the original hypothesis proposed by Mori[Mori 1970], where the horizontal axis represents the human likeness of each of the twelve samples presented during the survey. The vertical axis indicates the familiarity level, here calculated by the inverse of the results of the specific question about the "discomfort level" (the negative emotional response) in order to provide an equivalent representation to be compared with the original uncanny valley.

Although the original model doesn't mention explicit values and the graph here shows exact values extracted from the survey (along the vertical axis), it is possible to find a similar function. It is especially noticeable that samples C.11, C.16, C.1 and C.14 suggest the same pattern of a negative emotional response as in result of an "almost human but unnatural" character representation. These specific samples caused more discomfort or, in other words, the inverse effect: less familiarity.

The uncanny valley original hypothesis also covered the effects of a moving against a still character. Even with a limited number of samples of this research, it was possible to detect a similar behaviour, more noticeable in characters C.16 and C.14. When these characters were displayed in movement (in this case, as a computer animation) the curve shape exaggerates the valley.

The analysis of both stages of the survey, in the attempt to verify it against the uncanny valley, leads to a positive result according to the original hypothesis of Mori, here represented by moving and still images of CG generated characters.

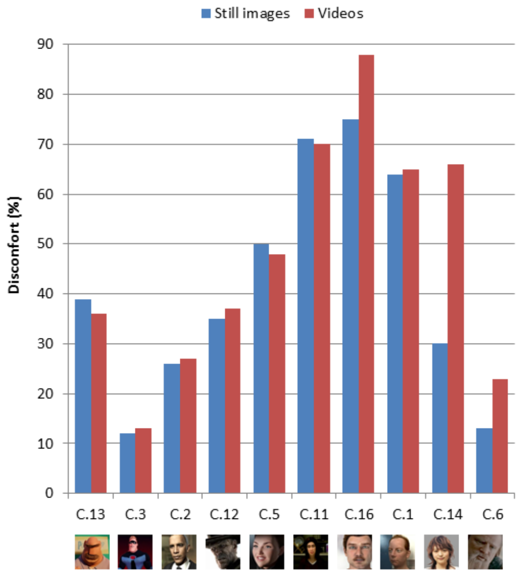

Figure 4: *Percentage of people who felt discomfort on each character.*

In the survey also was asked for the participants to inform in each character, in case they felt discomfort looking at it, which parts of the face they felt more this discomfort. This question allowed multiple answers and wasn't mandatory. Analysing the answers obtained, we could highlight two specific parts of the face: The eyes and the mouth. The answers of the two stages of the survey, being the first with images and the second with videos, had similar results that are demonstrated in Graph a and Graph b, respectively, in Figure 5. In the images survey were obtained 854 answers adding all the characters answers), being approximately 35.25% to the eyes and 29.98% to the mouth, completing a total of approximately 65.23% of the answers. In the videos survey the results were the following: 811 answers (adding all the characters answers), being approximately 32.18% to the eyes and 37.85% to the mouth, completing a total of approximately 70.03% of the answers. Noting that the two stages of the survey were launched on different periods, consequently with different participants answering.

When comparing these two graphs (a and b) with the graph of Figure 4, which represents the discomfort, it can be observed an apparently direct relation. This is, the number of people who felt discomfort and the number of people that indicated the strangest parts of the face where similar to each character. This observation demonstrates the consistency of the gathered data.

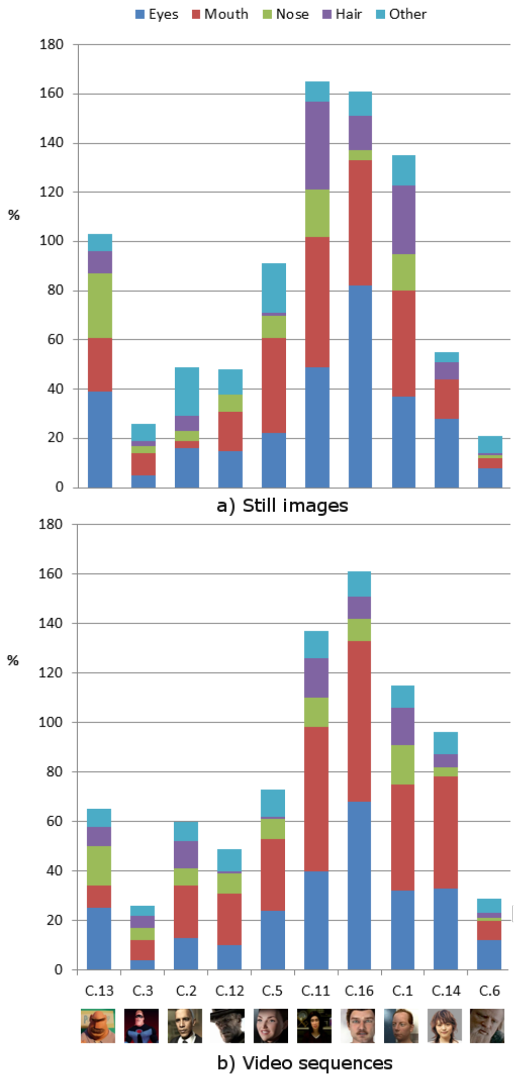

Figure 5: *For each character, what parts of the face people choose the most strange. In graph A the answers on the still images survey and in B in the videos survey. The question had multiple choices.*

Another question asked in the survey was whether the person knew or not the character, whose answers are shown by graph of Figure 6. Analysing the referred chart, we highlight four characters that are relatively well known from the selected set, which are: C.2, C.3, C.6 and C.13. The other characters are barely known by the participants, according to the graph.

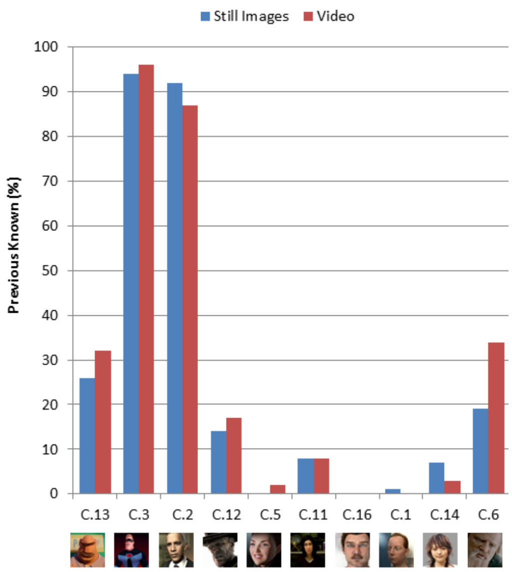

Figure 6: *How many people know each character before the survey.*

From the answers obtained, we can address the following question: "Does it make any difference in a person's perception to previously know the character?". To answer this question, we analysed the graph of Figure 6, which indicates how known the character is, and the graph of Figure 3, which indicates the level of familiarity generated by him. Thus, we hypothesize that well-known characters suggests greater levels of familiarity, since they are located in the higher parts of the curve presented in the graph of familiarity.

5 Final Considerations

Within this update on the uncanny valley subject, it was possible to validate its persistence under a contemporary scenario: CG generated characters and animations are a rich environment to observe the phenomena. Also, during the study we found an interesting correlation between the effects of the uncanny valley and the familiarity a given character has between the public, where the curve tends to be moderate for those characters previously known by the subject.

The subjectivity inherent to this type of evaluation leads us to deal with certain imprecisions that must be addressed by a considerable number of participants and generated samples. During this study, the total number of obtained samples allowed us to work with over two-thousand answers in each of the two stages, which showed to be plain enough to detect correlations and cover the criteria defined by the scope of this research.

Future work may be able to revisit the uncanny valley subject under a similar methodology, in order to distinguish better the effects of the technological evolution over the population against the human instincts considered as underlying cognitive mechanisms proposed as explanation of the phenomenon by the early studies. As a suggestion, for instance, the feature "popularity" here studied may be used to compound a third variable in the original uncanny valley graph, and so the other variables, such as empathy and regions of the face most significant for the human perception for the effect. This study didn't distinguish between gender, age, neither familiarity with animated movies nor computer games, which is also an interesting direction to be evaluated more deeply.

References

Chaminade, T., Hodgins, J., and Kawato, M. 2007. Anthropomorphism influences perception of computer-animated charactersâĂŹ actions. *Social Cognitive and Affective Neuroscience 2*, 3, 206–216.

Donovan, J. M. 2002. *Facial Attractiveness: Evolutionary, Cognitive, and Social Perspectives*. Ablex Publishing, Westport.

Green, R. D., MacDorman, K. F., Ho, C.-C., and Vasudevan, S. 2008. Sensitivity to the proportions of faces that vary in human likeness. *Comput. Hum. Behav. 24* (September), 2456–2474.

Hanson, D. 2006. Exploring the aesthetic range for humanoid robots. *Proceedings of the CogSci Workshop Towards social Mechanisms of android science*, 206–216.

Karl F. MacDorman, S. K. Vasuvedan, H. C. C. 2009. Does japan really have robot mania? comparing attitudes by implicit and explicit measures. *AI & Society 23*, 485Ű510.

MacDorman, K. F., and Ishiguro, H. 2006. The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies 7*, 3, 297–337.

MacDorman, K. F., Green, R. D., Ho, C.-C., and Koch, C. T. 2009. Too real for comfort? uncanny responses to computer generated faces. *Comput. Hum. Behav. 25* (May), 695–710.

MacDorman, K. F. 2006. Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the uncanny valley. *ICCSCogSci2006 Long Symposium Toward Social Mechanisms of Android Science Vancouver,*, 26–29.

Mori, M. 1970. Bukimi no tani (the uncanny valley). *Energy*, 33–35.

Seyama, Junichiro, N. R. S. 2007. The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence: Teleoperators and Virtual Environments*, 337–351.

Wallraven, C., Breidt, M., Cunningham, D. W., and Bülthoff, H. H. 2008. Evaluating the perceptual realism of animated facial expressions. *ACM Trans. Appl. Percept. 4*, 4 (Feb.), 4:1–4:20.