Towards Retro-projected Robot Faces: an Alternative to Mechatronic and Android Faces

Frédéric Delaunay   Joachim de Greeff   Tony Belpaeme
School of Computing,
Communications and Electronics
University of Plymouth, UK
Email: frederic.delaunay@plymouth.ac.uk

***Abstract*— This paper presents a new implementation of a robot face using retro-projection of a video stream onto a semi-transparent facial mask. The technology is contrasted against mechatronic robot faces, of which Kismet is a typical example, and android robot faces, as used on the Ishiguro robots. The paper highlights the strengths of Retro-projected Animated Faces (RAF) technology (with cost, flexibility and robustness being notably strong) and discusses potential developments.**

I. INTRODUCTION

Within the human-robot interaction field it has been generally accepted that robots will need some form or shape that reminds the user of a human face. Designers never felt the need for service robots (such as vacuum cleaners or lawn mowers) to appear anything other than functional, but as soon as some form of affective, emotional or personal interaction is involved, robots always have been designed to have distinctive facial features. The minimum set of facial features necessary seem to be a head-like volume, often contained within a sphere or ovoid, with a vertically symmetric setup of eyes and a mouth. It has indeed been known for some time that from a very early age we are sensitive to faces [1]. The general morphology of the face —a round shape where eye and mouth features are in the correct position— has a profound effect on how much affinity we have towards it (for an application of this in avatars see [2]). Although the Human Computer Interaction community successfully argues that affective computation does not necessarily require an anthropomorphic device (for example [3]), the HRI community generally does not describe to this view and feels that for affective human-robot interaction to succeed a robot needs an anthropomorphic face. Moreover, there is evidence that a robot face is more persuasive; Kidd [4] for example, studied different support devices for weight loss programme and showed how a robotic weight loss assistant with a robot face has a higher success rate than alternative support methods.

Traditionally, expressive facial animation in robots has been implemented using mechatronic devices. Kismet [5] is one of the earliest and most classic mechatronic expressive robot, with all features –such as eye lids, eye brows, lips and ears– being physically implemented and controlled by electric motors. Other examples are the Philips iCat [6] which has a cat-like head and torso with motorised lips, eye lids and eye brows, and the MDS (mobile, dexterous and sociable) robots, which have motorised eyes, eye lids and a mouth [7]. A range of androids have been demonstrated as well, which have a larger number of mechatronic actuators controlling a flexible plastic skin; these technologies originated from animatronics. Examples are the Hanson robot faces, such as the Albert Hubo or Joey Chaos robot heads [8] or Ishiguro's androids [9]. Android heads, due to the number of actuators and the nonlinear interaction with the plastic skin, are typically more expressive than the above mentioned mechatronic heads. However, they are mechanically very complex and as such expensive to design, construct and maintain.

As an alternative, some robots carry a flat screen monitor displaying a synthetic character. While the hardware cost of these robots is considerably lower due to the use of off-the-shelf components, it is often felt that these attempts to endow the robot with an affective character are not as successful as the above described mechatronic solutions. Another approach to achieving facial expression is used in the iCub face which uses a line of LEDs to implement a mouth and eye brows.

An issue related to how successful specific designs of robots are within human robot interaction is the notion of uncanniness [12]. In short, Mori argued that as the design of robots gradually progresses towards robots which more and more resemble humans, at a certain point an uncanny valley is encountered; in which humans perceive the robot as not familiar at all. This feeling of strangeness could potential hinder human-robot interaction. The robots of the non-android class mentioned above are clearly mechanic with stylised features, and hence do not suffer from the uncanny valley. Androids however, and especially their movement, have not yet reached a level of performance which is perceived as convincing by humans, and thus feelings of uncanniness can play a role again. In general, it can be argued that for a robot to be successful in its interaction with humans, its mode of presentation needs to be clear: either it is presented as a non-human character which makes no pretension on being human-like, or the robot does resemble humans enough (in all relevant aspects) to overcome the uncanny valley. As the latter type is not achieved at present, and indeed may take a while longer, it can be argued that an approach in which a robot does not try to be as human as possible is more feasible to achieve effective human-robot interactions. This

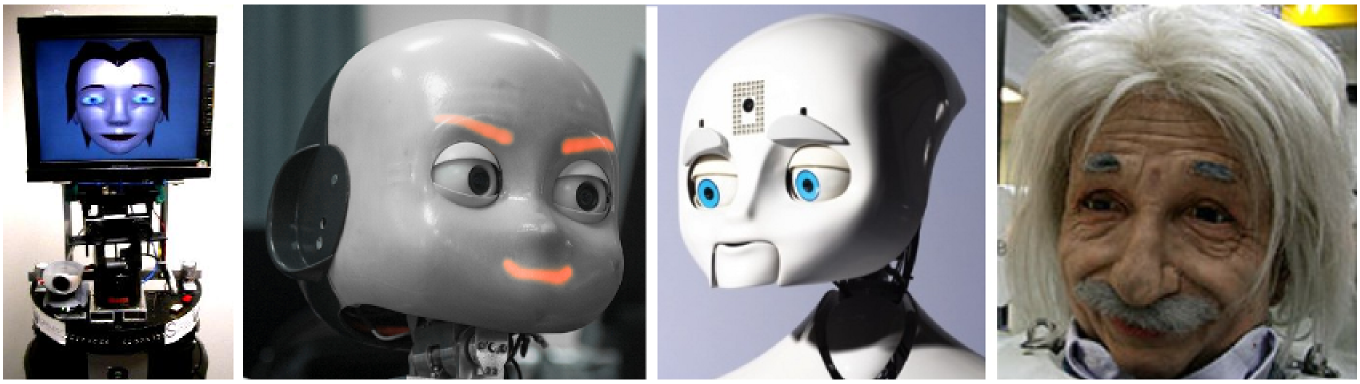

Fig. 1. Some examples of contemporary robot heads: flat screen head GRACE [10], i-Cub mechatronic and LED head[11], Nexi MDS mechatronic robot head [7] and Albert Hubo android robot head [8].

is of course not to say that the type of interactions should be limited because of this design choice.

An alternative to mechatronic and android heads is to retro-project a robot face into a face-shaped translucent surface. This has been tried by artists using video records of human face[1] but it seems the work did not lead to a publication. However, a publication focused on dome-shaped display [13] developped image correction and facial expression using line drawings.

This paper presents the design, implementation and preliminary tests of an implementation using an opal mask (with some facial features). We highlight some issues and offer solutions, opportunities and future directions for this technology.

II. A NEW APPROACH

Building a mechanically animated face (either for mechatronics or androids) takes a considerable amount of resources and the technologies involved have a number of drawbacks which are discussed below along with possible alternatives:

- Common actuator motion is not convincing: slow or jerky moves are to be avoided as most of human facial muscles fully stretch in a matter of tenths of seconds. Latency is also an issue as most of the time humans change from one expression to another in a very short amount of time. In contrast, video is able to display any kind of movement, faster than common servos (stepped motors), with smoothness of motion being only constrained by the video framerate.
- Expression range is limited: facial systems can be implemented with few mechanical parts (for the face, the Nexi and BARTHOC robot seem to have only 7 DOFs), but applications requiring more human-like animated faces need more actuators. Hence, more actuators requires a bigger volume, add weight and require more power. On the other hand, projected faces can be made very light, with the weight remaining constant whatever the animated facial features.
- Androids are more prone to uncanniness: as their physical appearance is closest to humans, actuators and skin quality is still an issue with current technology (see MacDorman [14] for a number of important insights). Computer graphics however are constantly improved and we are closer than ever to movie quality realtime character animation (see [15] for example).
- A flat screen is unnatural as a face. A key aspect is the geometry of the face itself, the half-spherical overall shape lets the audience view the face within a 180 wide area. Mimicking the geometry of the eyes, which remains the most important part in the face (see Maurer [16] about eye-contact) helps to interpret the robot's gaze, thus improving interaction.

Addressing the issues described above would yield a new design without the use of common servos nor unrealistic synthetic skin. As an answer to the limitations of current technology, we propose Retro-projected Animated Faces (RAF), bringing new interesting possibilities. Clearly, a complete HRI solution should support head movements (moving and rotating) as this helps bring the robot to life (see [17]). As a consequence, some mechanical parts are still required for proper interaction with the user; our setup is no exception and features a robotic neck.

III. PROTOTYPE

For our approach we opted for an on-line animation retroprojected onto a mask, which is not flat but has the shape of a human face. The geometry of the mask expresses some of the invariant facial features found on a face (ovoid shape, nose, eyeballs) while others like mouth and eyebrows are moving most of the time, and thus are not geometrically expressed by the mask. This improves on the Kamin-FA1 robot face [18], in which a system is described where a line drawing is projected into an semisphere, which seems to be the frosted shell of a light fixture. For the face of the robot we have opted for an adaptation of the iCub face cover[2]; the iCub face is rather simple and elegant, resembling a young infant through the large size of its eyes and its high forehead. The original face design has been reworked in a CAD suite to make into a mold; for this the inner volume of

[1]http://www.holoart.co.jp/cho01.html

[2]www.robotcub.org

the face mask was filled and the eyes closed. The model was printed on a rapid prototyping machine (a ZPrinter 310 by Z Corp) printing a high-performance composite. The material is sufficiently heat resistant at over 150°C to allow vacuum forming. In a next phase a sheet of 1.5mm opal high impact polystyrene (HIPS) thermoplastic was vacuum formed over the mold. The facial mask was then mounted onto a bracket which can be attached to the robot's neck.

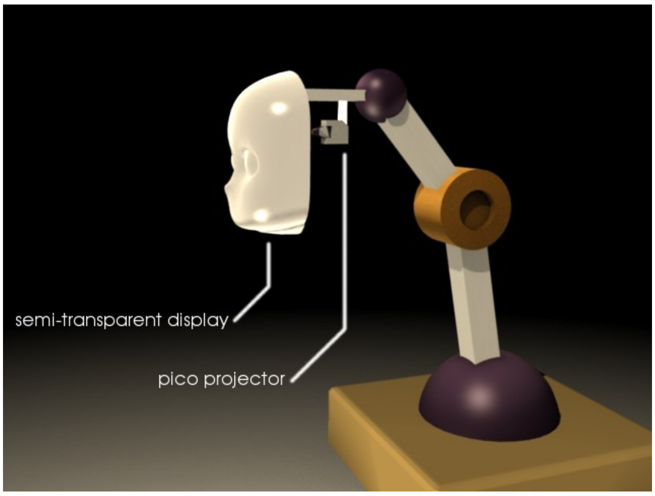

Fig. 2. Sketch of the robotic setup, the image is retro-projected into the opal face mask. The face is mounted onto an articulated neck. The current prototype described in this paper uses a ViewSonic XGA projector, however a picoprojector is to be used in a later version.

IV. FACIAL ANIMATION

As reported by psychologists, human communication relies a lot on facial expressions to convey emotion[19]. Ekman's facial action coding system (FACS) has been refined over the years and yielded the newest version in 2002 upon which computer graphics/vision researchers based their work ([20] for modeling, [21] for recognition). Briefly, FACS divides the face in 44 basic Action Units (AU) that are involved in facial expressions. Each AU stands for a muscle or set of muscles visually modifying a specific facial feature. FACS precisely describes all these modifications per AU. As for now, we decided not to implement exactly all of the AUs but to focus on a smaller (expandable) subset, mandatory for a basic range of expressions and convincing enough to generate an emotional response from humans.

Many techniques can be applied as an underlying drawing method to support facial animation (for a comprehensive survey see [22] or [23]). However, nowadays cheap 3D video-cards are powerful enough for complex 3D models to be displayed and animated in realtime. In this regard, we opted for 3D drawing but without high quality graphics in mind as the main focus of this work is the interaction itself and not near-realistic rendering. To set up a 3D face we tested two methods: using a template model featuring all AU effects on which a texture is applied (more likely to be used for mapping real faces) or an original 3D model scaled to fit the proportions of the template model upon which AUs effects are modeled. Creating a new face from templates is certainly quicker, however the methods described below apply to both.

Fig. 3. A picture of a human child is retro-projected into the facial mask, displaying some of the uncanny valley effect.

As can be seen Figure 3, using a picture of a human may yield to some of the uncanny valley effect. In order to avoid this, a non-realistic face was adapted from an original 3D model[3]. Thus, the 3D head was slightly modified to fit the display proportions by relocating eyes and nose according to the face mask in which the image was projected. The original 3D model was defined with a neutral face so this is equivalent to all AUs set at an intensity of 0. Each AU effect is then modeled on the face with a maximum activation defined as 1. This normalization allows us to precisely blend AUs together (similar to Wojdel [20]) and keep consistent values across the AUs involved in a facial expression, thus differing from FACS intensity classification. Consequently, we are using a weighted AU approach to define a facial expression E:

$$E = \sum_{i=0}^{S} w_i AU_i$$

where S is the AU subset mentioned previously.

Modeling an AU effect on the 3D face was done by defining the relative movement of all vertices involved. It is possible (and often the case) that some vertices belong to more than one AU. While this method is similar to what 3D modellers do, it allows selecting the smallest number of vertices with regard to the 3D model used. Defining movement of vertices may later be automated using a muscular based selection of vertices (similar [24]) as a pre-processing step but this is currently beyond the scope of this paper.

EMFACS is a later work from Friesen and Ekman bridging emotions and facial expressions based on AUs. The CMU-Pittsburgh AU-Coded Face Expression Image Database [25] stands as the current reference. Future developments of

[3] "maid-san" 3D model from author FEDB http://fedb.blogzine.jp/BA/body.zip

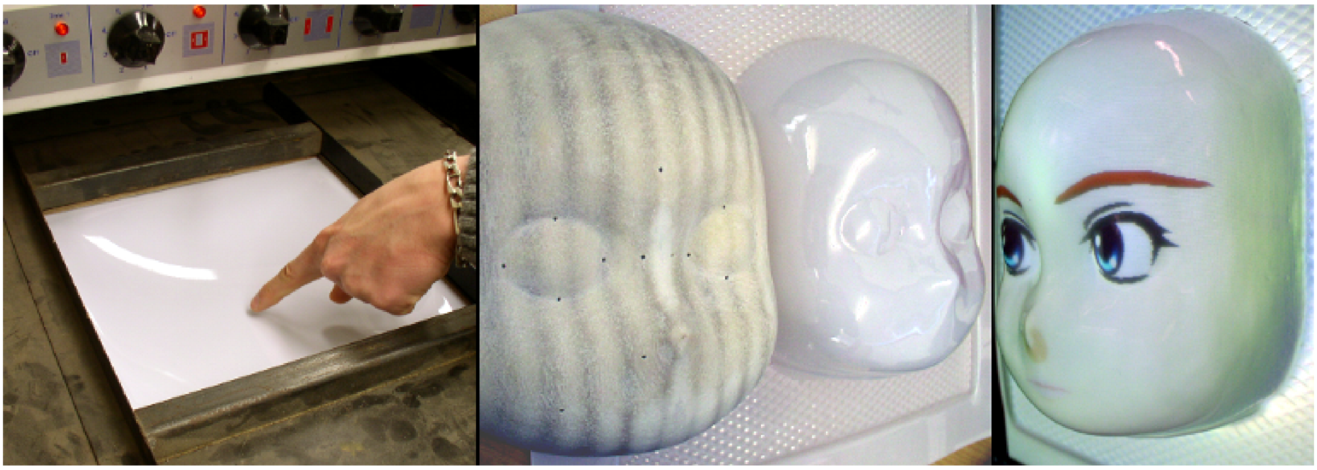

Fig. 4. Different stages of the process. Left: vacuum forming, middle: the mold and the resulting facial mask, right: an animated face is retro-projected into the facial mask.

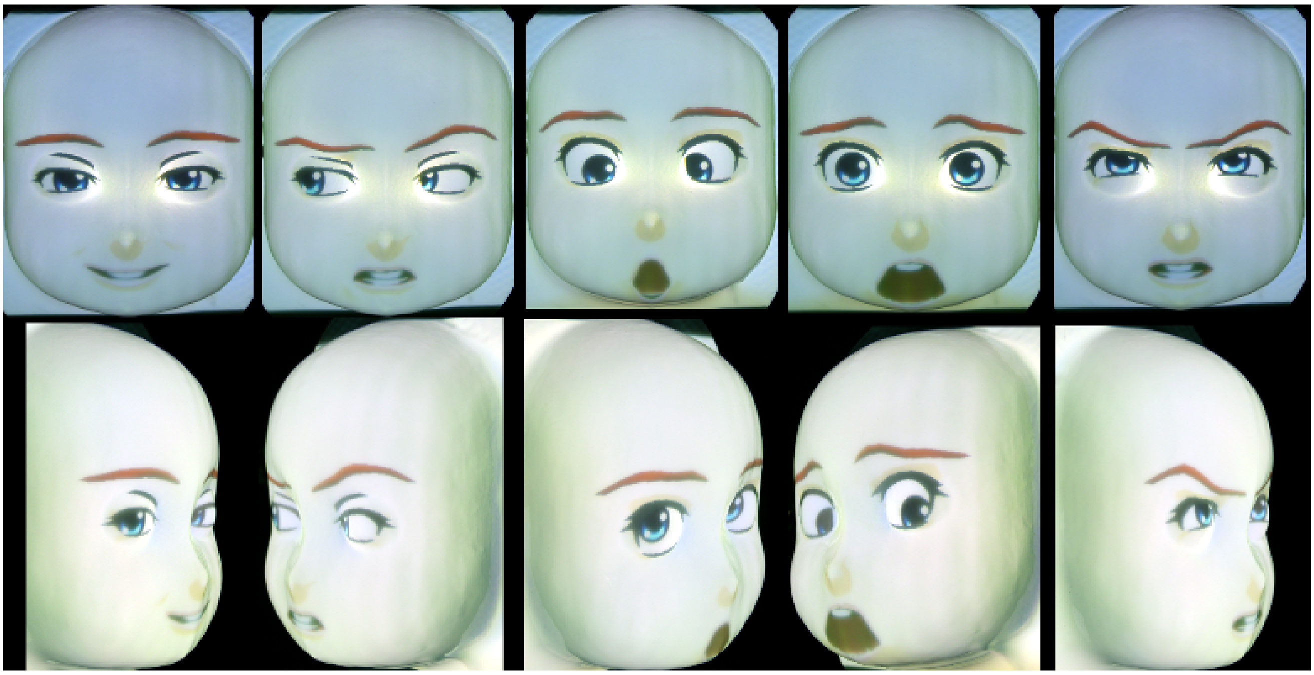

Fig. 5. Faces displaying basic five expressions. From left to right, both top and bottom: happiness, disgust, surprise, fear, anger. The neutral expression is not shown.

our facial animation and expression recognition systems are likely to be based on this resource.

The underlying emotional model responsible for facial control is a work in progress and as such is not described here.

V. DISCUSSION

Next to the practical evaluation of the previous section, we present an analysis of the technology, highlighting advantages, issues and future opportunities. To our knowledge, this technology has not been previously evaluated; for an evaluation of mechatronic robot faces, see [26].

A. *Advantages*

The main strengths of RAF technology are its cost and flexibility. The cost of designing and building a face is reduced considerably, mainly due to rapid prototyping and vacuum forming production techniques used for making the translucent mask and due to off-the-shelf projection technology. It is expected that projection technology will become more affordable and improve steadily, as opposed to the electromechanical components used in the mechatronic and android faces, which seem to have stagnated technologically and economically. The running cost of projected face technology is minimal compared to other technologies: the mean time between failure for LED projectors is about 20,000 hours and as no mechanical components are involved in the face animation, there are virtually no parts prone to failure. As such, hardly any maintenance is needed, making projected face technology ideal for commercial applications. Hybrid Laser/LED projective solutions (Explay Colibri Compact Mobile Module) exist as well, sharing the same robustness level. Moreover, these have the interesting property to keep the light beam focused for any distance within the projection range (0.2 to 2 meters) as opposed to manual focusing for standard projectors. The other advantage of projected face technology is flexibility: not being bound by electromechanical constraints, this technology allows designers to implement an almost unlimited range of emotional expressions. This can include the traditional gamut of actuating facial features, such as moving eye brows or animating a mouth, but is not limited to this. Earlier studies have shown that exaggerated features can be very effective in conveying emotions. Large eyes, for example, have been employed to create faces with which humans readily sympathise [26]. Sweating and blushing are easily programmable features and add a noticeable amount of emotional information to the user. Whilst these latter artefacts are rarely used in HRI, they may be well suited with other features for specific robotic issues where the user should be informed of the robot's affective state such as long processing times (sweating) or task failure (blushing), replacing more distracting (and perhaps meaningless) conversational fillers. Additional effects such as changing the colour of the face to convey the robot's emotional state are straightforward and have a lot of potential[4]. Retro-projected Animated Faces technology also allows for on-line changing of facial morphology, this could be used to change the character depending on the user interacting with the robot (for example, male users might prefer a female character).

The face animation is entirely implemented in software and this creates design opportunities of which a number have been explored in this paper. In essence, the face can range from a simple cartoon-like animation (perhaps as simple as the line drawings in [18]) to a playback of canned video-recorded faces. Although our prototype is controlled from a standard PC, the computational power of contemporary embedded devices is probably sufficient for the purpose of animating the projected face.

Other advantages include the absence of acoustic noise: the actuators (electric or pneumatic) of mechatronic and android faces make a very noticeable – and too often distracting – noise. The speed with which the face can respond is also high; this is crucial in HRI applications where responsiveness is key to achieving successful interaction. Because there is no moving mechanical part, the setup is not only less prone to breaking, but can also be used in close interaction with naive users. Finally, the weight of the setup can potentially be under 300gr, leaving behind other technologies.

[4]For an example of effective use of facial colour - and other facial features - see the RoboThespian from EngineeredArts, see http://www.engineeredarts.co.uk

B. Issues and potential problems

A potential issue with small pico projectors is that these projectors use LED technology instead of an incandescent light source. LEDs are widely used as a lighting solution, mainly promoted by high-tech industries, but their brightness constrains most of their applications to indoor environments. Specifications for pico projectors are still preliminary, but light power will range from 7 to 40 lumens; compared to micro and normal projectors, which typically have a light power of over 1000 lumens, this limited power poses a challenge. We therefore expect the robotic setup to be operated under dim lighting conditions; currently face projection with pico projectors will be unfit for bright conditions.

The technology and design of the robot requires free space between the face display and the projector, hence this space cannot be used to fit any sensors as these would block the projected video beam. While most robot designs allow a camera in the eyeballs, and some of them allow tactile interaction (for example Hishiguro's Geminoid, and Aiko), the display and projection area in projected face technology must remain clear. We consider this as relatively small issue as the robot is not supposed to be an android, and tactile interaction is not pursued. The camera can be mounted to the side of the face, for example below the chin. However, by moving the camera to an unnatural position (under or on top of the head), humans may have difficulty presenting objects very close to the robot as they would expect it to "see" through the displayed eyes. As an alternative, the robot could show its difficulty seeing the target object by moving it's head backwards or away from the object and displaying an appropriate expression.

C. Opportunities

New developments in projector technology promise cheaper, smaller, lighter and more energy efficient projectors, through for example using integration of electronics and using LED instead of incandescent light. Recent micro projectors weigh about 500 gr and pico projectors, expected in 2009, can weigh as little as 20 gr. This would allow projectors to be built into the robot and pico projectors using LED illumination could run from the robot's batteries. The optics of micro and pico projectors also allow for projecting a focused image at a close distance; current projectors still require a considerable distance between the lens and the projection surface. Recent small projectors can throw a focused image at distance as small as 20 cm, allowing for the projector to be mounted at the back of a robot head.

As a summary, the technology discussed here is compared with other technologies in Table I.

A full evaluation of the technology with naive users is planned for the near future. However, initial informal evaluations with naive visitors and students are very promising. We plan to use this technology to study how robots can engage with humans in a tutoring interaction and expect that the projected face mounted on a controllable neck will be an improvement over existing technologies.

TABLE I

COMPARISON OF RAF TECHNOLOGY AGAINST THE TWO ESTABLISHED MECHATRONIC ROBOT FACE AND ANDROID ROBOT FACE TECHNOLOGIES FOR A NUMBER OF KEY PROPERTIES.

	Retro-projected Animated Faces	Mechatronic face	Android face
Development cost	Relatively low	High	Very high
Maintenance cost	Very low	High	High
Robustness	High	Medium, due to mechanical parts	Low, due to mechanical parts and wear on flexible skin
Flexibility	High (software controlled)	Limited by hardware setup	Limited by hardware setup
Expressive range	Software controlled	Limited	Limited
Reality	Low	Low	High
Acceptance by users	Unknown	Relatively high	Relatively high (but see uncanny valley)
Uncanny valley risk	Medium (depends on facial model)	Low	High
Power consumption	Low (LED and/or Laser)	Medium	High
Texture	Unnatural	Unnatural	Closest to human skin
Acoustic noise	None	High	High
Weight	Potentially very low	Relatively high	Relatively high
Reactiveness	Fast	Medium	Medium
Ambient light restrictions	High	None	None

VI. ACKNOWLEDGMENTS

This work is supported by the CONCEPT project (EPSRC EP/G008353/1) and the EU FP7 ITALK project. We wish to thank Martin Woolner from Innovate (University of Plymouth) for the rapid prototyping support.

REFERENCES

[1] S. Bråten, *Intersubjective Communication and Emotion in Early Ontogeny*. Cambridge, UK: Cambridge University Press, 1998.

[2] H. Prendinger and M. Ishizuka, "Human Physiology as a Basis for Designing and Evaluating Affective Communication with Life-Like Characters," *IEICE Trans Inf Syst*, vol. E88-D, no. 11, pp. 2453–2460, 2005. [Online]. Available: http://ietisy.oxfordjournals.org/cgi/content/abstract/E88-D/11/2453

[3] R. W. Picard, *Affective computing*. Cambridge, MA, USA: MIT Press, 1997.

[4] C. Kidd, "Designing for long-term human-robot interaction application to weight loss," Ph.D. dissertation, Massachusetts Institute of Technology, 2008.

[5] C. Breazeal, *Designing Sociable Robots*. Cambridge, MA, USA: MIT Press, 2002.

[6] A. van Breemen, X. Yan, and B. Meerbeek, "icat: an animated user-interface robot with personality," in *AAMAS '05: Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*. New York, NY, USA: ACM, 2005, pp. 143–144.

[7] "Mds project at the personal robots group, mit media lab," 2008. [Online]. Available: http://robotic.media.mit.edu/projects/robots/mds/overview/overview.html

[8] D. Hanson, "Expanding the aesthetics possibilities for humanlike robots," in *Proceedings of IEEE Humanoid Robotics Conference, special session on the Uncanny Valley*, Tskuba, Japan, Dec. 2005.

[9] D. Matsui, T. Minato, K. MacDorman, and H. Ishiguro, "Generating natural motion in an android by mapping human motion," *Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ International Conference on*, pp. 3301–3308, Aug. 2005.

[10] R. Gockley, R. Simmons, J. Wang, D. Busquets, and C. DiSalvo, "Grace george: Social robots at aaai," in *Proceedings of AAAI'04. Mobile Robot Competition Workshop (Technical Report WS-04-11)*. AAAI Press, Aug. 2004, pp. 15–20.

[11] R. Beira, M. Lopes, M. Praa, J. Santos-Victor, A. Bernardino, G. Metta, F. Becchi, and R. Saltarn, "Design of the robot-cub (icub) head," in *IEEE International Conference on Robotics Automation*, Orlando, May 2006, pp. 94 – 100.

[12] M. Mori, "Bukimi no tani [the uncanny valley] (k. f. macdorman & t. minato, trans.)," *Energy*, vol. 7, no. 4, pp. 33–35, 1970.

[13] M. Hashimoto and D. Morooka, "Facial expression of a robot using a curved surface display," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2005, pp. 2532–2537.

[14] K. F. MacDorman, "Androids as an experimental apparatus: Why is there an uncanny valley and can we exploit it?" in *Proceedings of the CogSci 2005 Workshop: Toward Social Mechanisms of Android Science*, 2005, pp. 106–118.

[15] E. dEon, "Nvidia real-time graphics research: The geforce 8 demo suite," in *SIGGRAPH '07: ACM SIGGRAPH computer animation festival*, New York, NY, USA, 2007, p. 102.

[16] D. Maurer, "Infant's perception of facedness," in *Social perception in infants*, T. Field and M. Foxx, Eds. Norwood, NJ: Ablex, 1985.

[17] C. L. Sidner, C. Lee, C. D. Kidd, N. Lesh, and C. Rich, "Explorations in engagement for humans and robots," *CoRR*, vol. abs/cs/0507056, 2005.

[18] M. Hashimoto and H. Kondo, "Effect of emotional expression to gaze guidance using a face robot," in *Proceedings of the 17th IEEE International Symposium on Robot Human Interactive Communication (RoMan 2008)*, Tamatsu, Y, Aug. 2008, pp. 95–1.

[19] P. Ekman and H. Oster, "Facial expressions of emotion," *Annual Review of Psychology*, vol. 30, pp. 527–554, 1979.

[20] A. Wojdel and L. Rothkrantz, "Parametric generation of facial expressions based on facs," *Computer Graphics Forum*, vol. 4, no. 24, pp. 1–15, 2005, blackwell Publishing.

[21] M. Pantic and L. J. M. Rothkrantz, "Expert system for automatic analysis of facial expressions," *Image and Vision Computing*, vol. 18, no. 11, pp. 881 – 905, 2000. [Online]. Available: http://www.sciencedirect.com/science/article/B6V09-40HV0KG-2/2/3470d95d5e478bc1c480811d57604186

[22] J. Noh and U. Neumann, "A survey of facial modeling and animation techniques," Al- USC, Tech. Rep., 1998.

[23] B. Schroeder, "Facial animation: overview and some recent papers," 2008. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.94.5043

[24] K. Waters, "A muscle model for animation three-dimensional facial expression," in *SIGGRAPH '87: Proceedings of the 14th annual conference on Computer graphics and interactive techniques*. New York, NY, USA: ACM, 1987, pp. 17–24.

[25] T. Kanade, Y. Tian, and J. F. Cohn, "Comprehensive database for facial expression analysis," in *FG '00: Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000*. Washington, DC, USA: IEEE Computer Society, 2000, p. 46.

[26] C. F. DiSalvo, F. Gemperle, J. Forlizzi, and S. Kiesler, "All robots are not created equal: the design perception of humanoid robot heads," in *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, techniques*, London, England, 2002.

[27] M. Hackel, S. Schwope, J. Fritsch, B. Wrede, and G. Sagerer, "A humanoid robot platform suitable for studying embodied interaction," in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots Systems*, Edmonton, Alberta, Canada, Aug. 2005, pp. 56–61.