

# Investigating Political And Demographic Factors In Crowd Based Interfaces

Tom Barker, M. Hank Haeusler, Frank Maguire, Jason McDermott

University Of Technology Sydney, Faculty of Design Architecture & Building

Tom.Barker@uts.edu.au, Frank.Maguire@uts.edu.au & Jason.mcdermott@uts.edu.au

## ABSTRACT

Techniques that enable groups of people to control or influence digital system applications collectively have been greatly facilitated through the emergence of faster and better image processing and sensing technologies. This paper considers design issues that relate to crowd or group based user interfaces. One key difference when comparing group interface design with one-on-one user interfaces, is that a group format raises issues of digital political determinism within the system algorithms. These include the impact of an individual's weighting within the group; problems relating to inclusivity across certain user groups; and communication of appropriate user interaction to a diverse audience. These issues were explored by the authors' research using an anamorphic, anthropomorphic experimental display screen in a public location. An input mechanism was developed employing human facial expression analysis, to deliver emotionally expressive visual feedback.

## Author Keywords

User interface design, image recognition, emotion recognition, cybernetics, digital political determinism and anamorphic anthropomorphic display.

## ACM Classification Keywords

Design, Experimentation, Human Factors.

## INTRODUCTION

The authors define a crowd based interface (see figure 1) as a cybernetic system (Weiner, 1948) in which several participating human players have a partially or wholly collective communication with a Digital Control System that causes an intervention, which manifests itself back to the crowd as a tangible or visible transition through a feedback process. The degree to which the communication is collective is subject to a "Democratic Algorithm", the parameters of which include the duration of influence of the group's collective decision (time-span), and the polling technique used (eg: majority, first past the post, proportional representation).

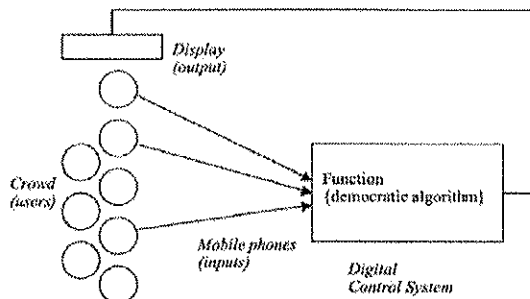


Figure 1. Crowd based interface

Internet sites allowing sharing of social data and images are growing fast in popularity (Nielson, 2005-6). Up to this point these shared images and emotions have typically have not been expressed in open public space informing of the collective changing mood.

Brian Massumi's cross-connecting perceptual change through collectivised emotions is a notable exception in this context. Massumi refers there to Lars Spuybroek's / NOX and Q.S. Serafijn's *D-tower* ([www.d-toren.nl](http://www.d-toren.nl): Oct 2007) a sculpture in the town of Doetinchem in the Netherlands. The sculpture abstracts the emotions of the town's inhabitants via answers to an online questionnaire. Each evening it then transmits the 'emotional state' of the town by assigning a correlating colour to the most predominant emotion. Massumi (2006) explains:

"Affect has been given visual expression. The predominant affective quality of people's interactions becomes visible. This can undoubtedly reflect back on the interactions taking place in the town by making something that was private and imperceptible public and perceptible. A kind of feedback loop has been created between private mood and public image that has never existed in quite this way before ... between different perceptual modes, different phases of perception formation and between perception and affect."

In order to study the phenomena of a crowd-based interface with an embedded Democratic Algorithm, the authors constructed *Janus*, an outdoor prototype. *Janus* is described below.

## Project experiment summary

The project proposed a pixel façade generated out of 183 grayscale-controlled light spheres (pixels) arranged as a giant human face hanging above a street – people could MMS pictures of their faces to it or submit images via a website, thus accumulating facial emotions of the

Permission to make digital or hard copies of part or all of this work or personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee.

'09, - , 2009,

© ACM 2009 ISBN: 978-1-60558- -4/09/1 ...\$10.00

participants – after which the present leading emotion could modify the face animation.

The project was inspired by *Janus*, the roman god with two faces and Greek theatre masks and through emotional expressions informs a continual feedback loop between collective mood and visitors of a public space.

Contrary to conventional screens, that display media content on a flat display surface, the '*Janus* screen' is anthropomorphic and draws upon the complexity of form in a face including a slight asymmetry. The screen was installed, suspended above Kendall Lane in The Rocks area of Sydney, Australia for the inaugural Smart Light Sydney festival in May 2009 (see figure 4.) ([www.smartlightssydney.com](http://www.smartlightssydney.com)).

The backend of the system, which was to drive the matrix utilised Phidget LED64 'Pulse Width Modulation' output hardware (see [www.phidgets.com](http://www.phidgets.com)), driven by software written in Max/MSP 5. The emotion videos were chosen and presented based on the highest polling emotion, stored in a database populated by the emotion recognition system.

#### Facial Expression and Analysis

The 19<sup>th</sup> Century French neurologist G. Duchenne (1862) identified sixteen facial expressions. He concluded that, although there are dozens emotions that we can feel only sixteen are visible (see figure 2) which can thus lead to misreading between people. In 1972, Ekman argued that six universal emotional expressions were commonly used and understood across cultures (happy, sad, angry, fearful, disgusted and surprised, in addition to a neutral face). These are referred to as the so-called *basic emotions*, and were codified into the Facial Action Coding System (FACS) (Ekman & Friesen, 1978).



Figure 2. Facial expression research by Duchenne

Numerous computer-based techniques have been developed utilizing FACS in analysis of video images, to determine outwardly exhibited emotions (Essa & Pentland, 1995). While the accuracy of automated techniques has significantly increased in recent years the analysis of facial expression is still limited in its ability to capture hidden or non-visual cues to human emotion (Fasel & Luettn, 2002).

The technique employed by *Janus* uses a pattern-matching algorithm for observing and extracting

symbolic information from video frames. These algorithms typically recognize a number of visible points on a face, which are identified as codified features and analyzed for visible movement and are compared to the FACS. The seven FACS expression set is analysed by "eMotion" (Valenti, et al. 2008), which forms one part of the software used in the experiment. The basic emotions established the focus of abstraction and engagement with the crowd of participants.

#### Democratic Algorithm

To allow the participation of many users, a democratic algorithm was devised which polled and proportioned the votes as interpreted before display.

The impact of any one participant becomes proportional to the number of participants and the length of time each spends interacting. Users were not restricted to a single vote, rather a longer engagement with any one emotion during the polling period resulted in an increased effect on the output. Any one person using the interface is able to influence the results more greatly if they choose to.

#### METHOD

Through the use of a face shaped screen, displaying emotionally expressive content, a deeper psychological connection between the viewer and the screen is suggested. Humans have a fascination with faces even being drawn to them only minutes after birth though our eyes can barely focus (McNeill, 1998). Upon viewing the recognizable form of the *Janus* face and the emotional expressions of an actor, the user is compelled to interact with it, this draws on our innate desire for human interaction and capacity for social cognition (Adolphs, 2003).

The interface sought to avoid the risk of banality arising from incomprehensible abstraction of expressive input. The full range of the 'basic emotions' being expressed by the audience and interpreted by the democratic algorithm were expressed in the video content.

The *Janus* face was represented in a two-dimensional array arranged within a non-rectangular matrix of pixels, 13 wide by 19 tall, a low-resolution face was represented (see figure 3). It has been observed that humans are good at recognizing facial characteristics from sparse visual information cues (Adolph, 2003). While there is a certain resolution below which it is no longer possible to discern a human face or the emotions being displayed on it human observers are able to recognize more than half of an unprimed set of familiar faces with image resolutions of merely 7 x 10 pixels (Sinha et al., 2006).



Figure 3. Reducing the display resolution to a minimum for facial recognition

### Anamorphic Form

Through the use of its three-dimensional form, the screen gains a non-static anamorphic relationship to its' viewers location, changing the experience of the screen as users move around it. The screen is only readable as flat from one privileged perspective.

There arose an interplay between the two parameters of low resolution – leading to better image viewing from a distance – and the three-dimensionality – leading to a better reading of the facial features from up close. The privileged perspective effect was exploited by the choice of a public space allowing the possibility to view and perceive it from different angles. Upon entering Kendall Lane viewers experience this privileged perspective (see figure 4).

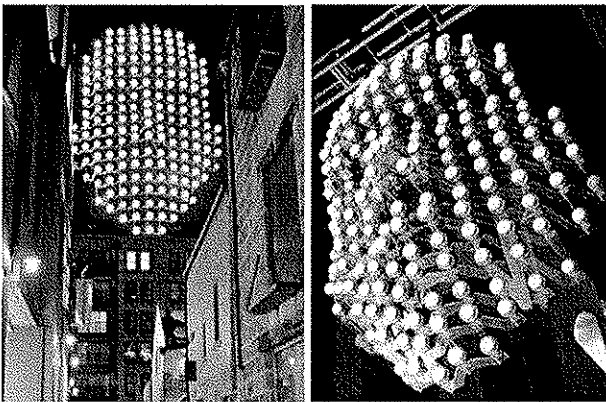


Figure 4. The visibility of features depends on aspect to screen, (left: the “privileged perspective”)

*Janus* employed an invisible interface, without recognizable machinic elements or standard GUI design semiotics. It therefore required new affordances not evident in traditional flat or screen based interfaces. An invisible interface produces minimal system feedback to the users, such as program status or immediate responses to input. Whilst there is accumulated feedback for a multitude of users, individualized feedback for each user is not given. There are no loading bars, status messages or decision points for users to be able to visualise the inner workings of the system. The users' engagement with the system arises through their participation and observation of the democratic algorithm.

### Employing the Democratic Algorithm

To have *Janus* smile, the participants themselves must collectively smile. As polling continues, individuals may be inclined to encourage others to express particular emotions deliberately, to change the effect of their votes.

The *Janus* interface did not identify individuals in the process of analyzing faces, nor did it look for relationships between recognized expressions and their context within the image. Differences in image context included multiple individuals, varied lighting conditions, and background information such as time or place. Such parameters were ignored; the *Janus* system simply recognized the expressive facial features in each image.

The process of user input polling occurred over a fifteen-minute period, with the three most prevalent observed expressions being displayed to the audience during the following period. The highest polling emotion was displayed proportionally more than following two emotions. Users witnessed the results displayed as a response to their votes, the correlation between the display and their vote being mediated by the Digital Control System.

Due to this fifteen-minute polling and display interval the user feedback provided by the interface was not immediate, nor directed toward any individual user. However given the public nature and variations of participation levels the polling method offered system flexibility. For the *Janus* experiment, tests indicated that the fifteen-minute interval was sufficient to allow for variation in the number of active users and to mediate the influence of any one participant.

Both direct and indirect voting is allowed. Users were able to interact with the system from the site of the installation via mobile phone, or from any other location through the Internet. The demographic of participants appeared to be predominantly young or middle aged (see figure 5). It could be suggested that to participate, an affinity for use of digital imaging, mobile and Internet communications technologies would have been of benefit.

There was a many-to-one relationship between users and the system, and a one-to-many relationship between the display and its' viewers. System factors such as the number of photos received, how polling decisions were made, who had voted or the timing of voting periods, were not revealed to users. Inference of these factors was left to each viewer.

### RESULTS

The festival participants were provided with an explanation of the experiment and guidelines on how to participate, on a separate Bluetooth link, at the display. Not all visitors were aware of this and therefore the individual participation levels into the cybernetic system were reduced. The *Janus* interface received engagement from an interested public and had many visitors who enjoyed the emotional experience.

### Complexities and Limitations

The software system was built by using prototyping tools and repurposed software. As a “rapid prototyping” method this was effective. The display hardware was built from ‘first principles’ with a design and construction time of only three months. This process proved a complex challenge in part due to the size and bulk of the screen (weighing 160 kg, 2.5m x 4m), as well as the suspension of the display three meters above the ground.

The animated expressions displayed on the *Janus* face were not always clear to all viewers. The non-static relationship between animated image and anamorphic form may have been complicated by the use of video footage. The use of static images or even simplified animations of emotional expressions may have been

easier to interpret for users, resulting in greater engagement with the interface.

The implementation of an invisible interface restricted the user's ability to use existing technology, flexibility of input from different user scenarios, and the ease of adaptability for new users of the system. These require an increase in the complexity and flexibility of all parts of the system. In order to provide a shallow learning curve for new users to an invisible interface, the design must cater for a wide range of expectations and abilities.

To interact with *Janus*, participants were required to have access to a digital camera or camera phone, as well as a capability to send the image to the system. This imposed restrictions on who could engage with the system to 'vote'. Children too young to have mobile devices, people who had older phone models and those who have no knowledge of technology would have had difficulty in influencing the output.

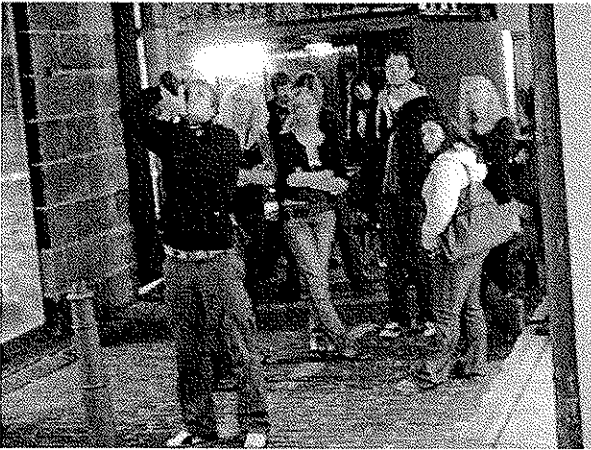


Figure 5. Audience engagement: Reflecting a young demographic

## CONCLUSIONS

The research was able to investigate several issues of crowd-based interfaces. The time-span of the prototyping experiment restricted the testing of different democratic algorithm weightings for the various parameters within the algorithm. This raises important user engagement questions that arose from this experiment, regarding the perceived agency of a single user in a multi-user system. Such as the amount of visual information provided to give users an understanding of what impact they are having, and how they might learn and adjust their actions.

Further work would consider these questions: (1) might a user be inspired to vote repeatedly, or encourage others in the space to do so as well? (2) Do users coordinate their emotions to test their impact on the system? (3) What other methods could be used to improve upon the inclusivity issues discovered, and what other forms of interaction could be designed with this kind of interface mode and methodology in the future?

Other areas for investigation are (4) controlled study of different types democratic algorithms and their impact on the participants, (5) improved recognition of the facial animations used on the display, possibly through simplified computer animations instead of video of an actor, and (6) a more exact way of measuring or recording the audience reaction, ie: the feedback loop, would be desirable.

The anamorphic anthropomorphic display sub-component, a solar powered "digital autonomous pixel", has been filed as an Australian patent application and will also be further developed by the researchers.

## ACKNOWLEDGMENTS

The team was greatly assisted by research assistant Mona Lieb who worked tirelessly on all aspects of the project. Late assistance for project management on site was by Elisa Hudson. Additional thanks go to the Smart Light festival for financial support and publicity.

## REFERENCES

- Adolphs, R., Cognitive Neuroscience Of Human Social Behaviour, *Nature Reviews, Neuroscience*, Vol. 4, pp 165-178, March 2003
- Duchenne G. *Mécanisme de la Physionomie Humaine*. 1862.
- Ekman, P & Friesen, W.V., *Facial Action Coding System: A Technique for the Measurement of Facial Movement*, Consulting Psychologists Press, Palo Alto, 1978.
- Essa, Irfan A. & Pentland, Alex P. *Coding, Analysis, Interpretation, and Recognition of Facial Expressions in IEEE Transactions on Pattern Analysis and Machine Intelligence* 1995.
- Fasel, B & Luetin, J, *Automatic facial expression analysis: a survey*, *Pattern Recognition*, 2003.
- Gevers, T. of Intelligent Systems Lab Amsterdam (ISLA) *eMotion Software*, [www.visual-recognition.nl](http://www.visual-recognition.nl) (2008)
- Massumi, Brian, *Transforming Digital Architecture from Virtual to Neuro*, An interview with Brian Massumi by Thomas Markussen & Thomas Birch, published online: [http://www.intelligentagent.com/archive/Vol5\\_No2\\_massumi\\_markussen+birch.htm](http://www.intelligentagent.com/archive/Vol5_No2_massumi_markussen+birch.htm), (accessed November 2006)
- McNeill, D., *The Face*, Little-Brown, Boston, 1998.
- Nielson. [http://news.cnet.com/8301-13577\\_3-9761150-36.html](http://news.cnet.com/8301-13577_3-9761150-36.html) (2005-06).
- Sinha, P. Balas, B. Ostrovsky, Y. Russell, R., *Face Recognition By Humans in Face Processing: Advanced Modelling and Methods*. San Diego. Academic Press. 2006
- Valenti, R. Sebe, N. Gevers, T. Cohn, I. *Machine Learning Techniques for Multimedia*, 2008.
- Weiner, Norbert. *Cybernetics: Or the Control and Communication in the Animal and the Machine*. Paris, France: Librairie Hermann & Cie, and Cambridge, MA: MIT Press. Cambridge, MA: MIT Press. 1948.