

Facing Interaction

Tomas Laurenzo

School of Creative Media, City University of Hong Kong
tomas@laurenzo.net

Abstract

In this paper, we present four interactive artworks –all of them in progress– that explore the use of face-based interaction in art installations investigating the possibilities of non-verbal communication.

Introduction

In this paper, we present a set of works in progress that explore the use of face-based interaction in art installations. These artworks offer different approaches to the question of how can we communicate sensations and emotional states, in non-verbal ways, to others, to ourselves, to objects, or even to places?

Under the assumptions that facial gestures could constitute a window to somebody's emotion, the art pieces aim to help reflecting on the poetics of non-verbal communication.

In the following sections we will describe five prototypes of artworks: *Walrus*, *Traces*, *Facial Pentatonic*, *Facial Sounds*, and *Look at me!*

We will conclude by presenting a short discussion, and our future related works.

Walrus

I am he as you are he as you are me and we are all together. John Lennon, lyrics to The Beatles' song "I am the walrus", 1967.

Walrus is an interactive installation consisting of a oval framed "magic mirror" that only reflects the interactor's face, with the reflected image substituted in real-time for a previous interactor's face in similar position and facial expression.

The system, for every frame, captures the user's face and stores it in a database. It then searches for a similar pre-stored face and displays it instead.

Using a depth camera and face tracking software, *Walrus* creates and manages a database of faces where each frame is classified according to its three-dimensional rotation plus some gesture descriptors. This classification of the stored faces allow for the real-time substitution so that, in every frame, the user's reflection consists of a different person's face.

Walrus creates a sense of awe in the interactor that arises from the fact that even if the facial features in the mirror are completely different to the users', the identification with the displayed image is natural, unavoidable, and immediate. The unnatural fact of a mirror that only reflects the face and does not follow optics rules (the reflected face is always centered in the mirror) creates a tension that interactors systematically alleviate by selecting a "physically correct" position.



Fig 1. *Walrus*, Prototype installed in Microsoft Research, Redmond, WA. 2012.

"Magic mirrors" (also called "augmented-reality mirrors" or "mixed-reality mirrors", among other similar names), that is, computational mirrors that behave in creative ways are very common in new media art, and with the advent of depth cameras a resurgence of this ever-

present type of installation has been seen. Among all the existent works, it is perhaps Chris O'Shea's *Body Swap* [1] the one closely related to *Walrus*.

Technical details

The installation is composed of a Microsoft Kinect Sensor (first generation), a computer running custom software, a projector, and an oval-shaped picture frame.

As with the rest of the pieces presented in this paper, we utilize the depth camera to track the interactor's head, and Microsoft's Face Tracker to locate the face and extract some gestural features: mouth openness, rising of eyebrows, mouth general shape, among others.

The computer stores each new face and its associated data into a database, and returns an existing equivalent one from the database. We organize the database as a hash table, with similar faces stored under the same hash entries. Face similarity is defined by a L^∞ norm of the head rotation plus similar gestural features.

When a new face is detected, it is stored into the hash bucket with the most similar representative. To avoid running out of storage, we cap a maximum size of each hash entry, and randomly kick out an existing entry when this limit is reached. We then randomly pick another face from the same hash entry. This can be seen as a cheap way of finding similar faces to the input via hashing.

The projector is mounted either on the ceiling or on top of a tripod that allows the returned face to be projected onto the oval picture frame without the interactor casting a shadow onto it.

Traces

What is our relationship with the architectural spaces we inhabit? *Traces* is an interactive installation that tries to help in answering this question.

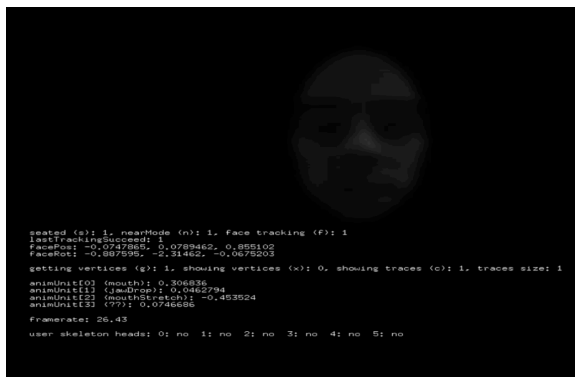


Fig 2. Screen capture of *Traces*' software, showing an acquired face and running information and parameters.

To experience *Traces*, the interactor arrives to a dark area within the gallery space. In a wall, the interactor will find several faces projected on the walls. Subtle ambient sound is played. Every face is drawn with its eyes closed.

When a blink is detected, the computer extracts a bitmap corresponding to the user's face, desaturates it and slightly blurs it. It then adds to the collection of faces that is projected. The spectator then becomes part of the installation.

The artwork becomes, then, a testimony of the visitors to the room, inhabiting it but not seeing it. In *Traces* visitors become subjects of the room, recipients of the spatial communication.

Traces reflects on the relationship between people and the spaces they inhabit: is a space changed because we have been there? Do we leave any trace on the places we have been to?



Fig 3. *Traces* prototype as installed in the School of Creative Media, Cityu University of Hong Kong, 2014.

The piece also questions what do we actually see and experience from a specific space. *Traces* is a log of people not seeing the space where it is exhibited, a rendering of some traces we might be unaware that we leave behind.

Traces also becomes a communicational vector between different visitors, as every spectator contributes – albeit passively – to how the piece looks at any time. However, the piece is always changing, and every interactor contribution, every trace, fades out with time.

The piece stores every participant's faces, becoming a witness of all its visitors in the moment of helplessness that their momentary blindness generates.

Technical details

Traces is composed of, depending the specific space where it is installed, one or various depth cameras one or various projectors, a computer and custom software.

After one spectator has being tracked for thirty to forty-five seconds, the installation enters into blink-detection

mode for that spectator. If there are more than the maximum of images – dependent of the specific gallery space – the oldest projected face is slowly faded out.

Our custom blink detector utilizes a computer vision library (OpenCV [2]) to extract one RGB bitmap per eye and raise a blink event when the bitmap changes more than a certain threshold. Change is measured by binarizing the images, applying a Sobel filter, and verifying that the resulting images shows no more than one continuous blob.

The Sobel operator performs a 2-D spatial gradient measurement on an image emphasizing regions of high spatial frequency that correspond to edges.

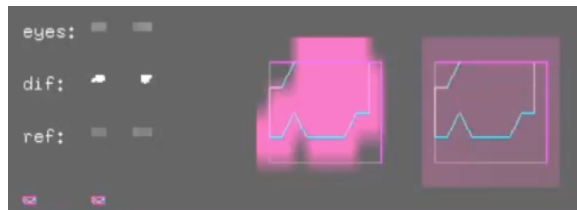


Fig 4. Our custom blink detector. On the left the extracted eyes are drawn. On the right the detected blob is drawn, signaling the detection of a blink.

The collection of faces stores the extracted images, and displays them. The size of the projected faces is configured depending on the installation space.

Also based on the installation space the maximum amount of projected faces is selected.

Facial Pentatonic and Face Sounds

Facial Pentatonic and *Face Sounds* are two musical instrument prototypes that map the user's face onto sound production.

In *Face Sounds* the user's head orientation and facial expression are mapped onto continuous parameters of a MIDI synthesizer instrument running in Ableton Live (a digital audio workstation specialized in real-time operation [3]).

Users trigger the sound by opening their mouth. The instrument embodies a virtualized voice that is controlled by the head's orientation.

Faces Pentatonic is a similar musical instrument, also triggered by the users mouth, with the difference that the interactor's head orientation is used to directly select the note being played instead of modifying timbral parameters.

The user head's pitch selects the octave, while the head's yaw selects the note within the scale (pitch

corresponds to left-right rotation, as in the western "no" gesture; and yaw corresponds to up-down rotation, as in the western "yes" gesture).

The system provides real-time visual feedback, showing the selected note. Its hands-free interaction allows the user to play another instrument at the same time (again, it becomes a virtualized, always-on-tune voice).



Fig 5. Screenshot from *Facial Pentatonic*, showing the tracked face and the selectable octaves and notes.

The system allows the user to select one note of the A minor pentatonic (five notes per octave) scale, which comprises the notes A, C, D, E, and G.

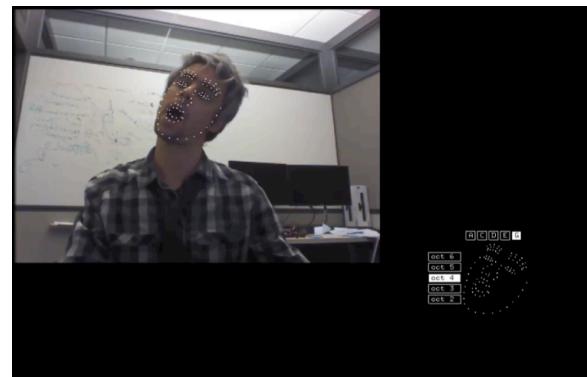


Fig 6. A user performing with the Face Pentatonic, with the G4 note selected.

Look at me!

Another line of work consists on investigating vibration as feedback. Vibrating motors are very cheap and easy to control providing an interesting (and popular) opportunity for artistic appropriation.

Look at me is a prototype of vibration-based feedback: a small installation that attempts to forceful capture the

user's attention. It consist on a small object (in our prototype a plastic cup). After the user looks at it, the installation attempts to not allow the user look away. When the interactor starts turning his or her head away, it lights a LED up and emits a soft high-pitched tone. If the user looks further, it vibrates and produces a loud buzz.

The installation aims at subverting the power relationship between the observed and the observant, between consumer and product.

Discussion

In spite of face tracking being widely distributed in consumer electronics, it is still not very frequently used in interactive artworks. The pieces described in this paper try to show that there are many possible uses for interaction based in face tracking within new media art.

In *Look at me!* face tracking is not used directly but, instead, works as an estimation of the user's attention. Even if more precise methods exist (e.g. gaze estimation), *Look at me!*'s heuristic is accurate enough so that the users' can correctly understand the installation's working and intention.

More widely explored is the use of face tracking for directly controlling music instruments, with perhaps the first example using a personal computer would be David Merrill's 2003 guitar effect controller [4]. Our prototypes show that designing this type of interaction is easily

feasible with available devices, opening interesting interaction design alternatives.

The artworks *Walrus* and *Traces* combine spatial augmented reality with face-tracking, showing that it is possible to investigate in alternative modes of communication within media arts and architectural research. The ability of detecting blinking or to "understand" a face gesture in real time offer the possibility to create behaviors that inspire feelings of awe in the audience while opening new modes of interaction with the spaces users inhabit.

Walrus spatially augments the oval frame granting it the affordances of a mirror. With its radical change of functioning, it aims at reflecting on self-perception, artistic exhibition, surveillance, control, and public entertainment.

Traces shares the same conceptual space but explicitly includes the architectural space in its interaction proposal.

In both artworks the inner working is not apparent to the users, allowing to reflect on the contemporary panopticism.

References

1. O'Shea, C. 2011. Body Swap.
2. Intel, Garage, W., and Itseez 2013. OpenCV.
3. Ableton What is Live?
4. Merrill, D. 2003. Head-tracking for gestural and continuous control of parameterized audio effects. Proceedings of the 2003 conference on New interfaces for musical expression, 218-219.