The Emoting City

Designing feeling and artificial empathy in mediated environments

Sayjel Vijay Patel¹, Raffi Tchakerian², Renata Lemos Morais³, Jie Zhang⁴, Simon Cropper⁵
¹,2,3 Dubai Institute of Design and Innovation ⁴ OPT Architects ⁵ University of Melbourne
¹,2,3 {sayjel.patel|raffi.tchakerian|renata.morais}@didi.ae ⁴ jiezhang5@gmail.com ⁵ scropper@unimelb.edu.au

This paper presents a theoretical blueprint for implementing artificial empathy into the built environment. Transdisciplinary design principles have oriented the creation of a new model for autonomous environments integrating psychology, architecture, digital media, affective computing and interactive UX design. ‘The Emoting City’, an interactive installation presented at the 2019 Shenzhen Bi-City Biennale of Urbanism/Architecture, is presented as a first step to explore how to engage AI-driven sensing by integrating human perception, cognition and behaviour in a real-world scenario. The approach described encompasses two main elements: embedded cyberception and responsive surfaces. Its human-AI interface enables new modes of blended interaction that are conducive to self-empathy and insight. It brings forth a new proposition for the development of sensing systems that go beyond social robotics into the field of artificial empathy. The installation innovates in the design of seamless affective computing that combines ‘alloplastic’ and ‘autoplastic’ architectures. We believe that our research signals the emergence of a potential revolution in responsive environments, offering a glimpse into the possibility of designing intelligent spaces with the ability to sense, inform and respond to human emotional states in ways that promote personal, cultural and social evolution.

Keywords: Artificial Intelligence, Responsive Architecture, Affective Computation, Human-AI Interfaces, Artificial Empathy

INTRODUCTION

Formalist approaches have dominated architecture for the past century, but have been criticized for a geometric rationality which distances the city from “the shapeless and dynamic acts of life and the ephemeral feelings evoked by architecture” (Pallasma, 2015). The introduction of ambient data-generating technologies such as sensors, wearables,
and personal robots in mediated environments creates the possibility of a new architecture, one with the power to perceive and elicit human emotions in non-anthropomorphic ways.

Currently, autonomous technologies are pervasive in cities, and have been used extensively to regulate human behavior. From medical wearables and robotic vacuums to digital thermostats, autonomous technologies operate in a continuous loop of monitoring, analysis, and feedback. Machine Learning (ML) will soon provide even greater autonomy through adaptive programming. The ability of autonomous systems to operate with no human intervention raises new design challenges and significant ethical concerns. These include unintended or aggravating behavior (Picard 2008) and issues surrounding risk and responsibility for that risk; exemplified by the controversy surrounding driverless vehicles and their safety. Furthermore, autonomous systems utilizing facial recognition are often designed in a way that objectifies, limits and seeks to control human behaviour in a way that contributes to skepticism surrounding their use (Gates 2011).

This paper presents a theoretical blueprint for implementing artificial empathy into the built environment. ‘The Emoting City’, an interactive installation presented at the 2019 Shenzhen Bi-City Biennale of Urbanism/Architecture (Figure 1) represents an initial steps into this emerging field. The installation explores the possibilities of human-AI mediated self-empathy in environmental design. It also investigates the technical implementation of these concepts by proposing new architectural elements capable of artificial perception and the ability to express artificial empathy. ‘The Emoting City’ brings to light new modes of human-human and human-AI interaction that are emerging in the context of autonomous environments. It brings forth a new proposition for the development of spaces that go beyond surveillance systems and anthropomorphic designs coming from social robotics, and that move forward to innovate in the design of interfaces that are transparent and self-effacing and therefore conducive to seamless levels of interactivity.

By connecting emotion sensing software (McDuff et al, 2016) with kinetic architecture, ‘The Emoting City’s interactive model facilitates the design of new elements in the built environment which respond to the facial expressions of occupants using motion and colour as semiotic tools for the purposes of personal insight. By perceiving and decoding human emotions, it is possible to create environments which facilitate the integration of autonomous technologies in ways that radically expand what is possible within traditional formalist environments.

**CONTEXT**

Empathy is defined as the ability to share someone else’s feelings or experiences by imagining what it would be like to be in that person’s situation. This definition is generally applied to human-human interaction but as we interact with non-human systems more and more the concept of empathy requires some modification to include human-AI interaction. Indeed, the role of human-AI ‘empathy’ in interactive design has never been so important. This modified application of Empathy has its roots in the concept of Cybernetics originally formulated by Norbert Wiener (Wiener, 1948), Pask’s (1976) proposal for human-machine, teach-learn feedback in ‘Conversation Theory’, and is at the core of embodied interaction (Dourish 2001) and user-centred design (Black 1998, Woodcock et al. 2018, Wright & McCarthy 2008).
In the context of affective and aesthetic interactive experiences that blend human and artificial intelligences, our understanding of empathy acquires new connotations. Traditional approaches to empathy-based design have been focused on human to human empathy processes that rely on cultural and symbolic systems of emotion expression, perception and interpretation. Empathy among human actors is derived from audiovisual cues such as dialogue and expression, and depends upon a subjective interpretation that relies on a particular individual’s ability to decode subtle signs of emotion such as tone of voice and facial expressions.

Human-human interactions that are empathic generate affective diversity and give origin to emotional bonds. However in the context of human-AI interactions, empathy becomes elusive. A different type of question emerges: can empathy exist in the context of human-AI interaction? If yes, how different is it from human-human interactions and what makes it unique? There are many initiatives currently researching possible answers to these questions.

Even though there are many current human-AI applications that are dedicated to creating new artificial empathy experiences, initiatives that address the design of embedded sensing capabilities in responsive environments with the specific purpose of promoting self-empathy and insight could not be found in our research for this project. ‘Emoting City’ is our response to this gap in human-AI interface design.

**EMBEDDED CYBERCEPTION**

Cyberception is defined as “the convergence of perceptive and conceptual processes in which the connectivity of telematic networks plays a constructive role” (Cipolletta 2018). When considering environmentally embedded processes of AI, facial-recognition technologies present themselves as powerful tools for cyberception. However, these technologies have acquired a bad reputation for their use as social surveillance apparatuses (Gates 2011) that collect personal information that is unidirectional: it is farmed directly from individuals into the databases of institutions, however individuals have no access to their own information.

The potential facial-recognition systems show in the context of artificial empathy has been researched extensively by the futurist Pamela Pavliskac (2018, 2019), who has coined the term Design Feeling to refer to the new possibilities brought about by emotional artificial intelligence (McStay 2018). The term recognizes the increasingly interdependent nature of how technology affects human emotions and how emotion-detection informs the design of new affective technologies.

Design Feeling reinforces the vision for a transdisciplinary architecture that facilitates new experiences which connect virtual, environmental, corporeal, and psychological dimensions. This transdisciplinary approach to architecture has been well established by several seminal architects and artists. Ascott (1995) argues that the current state of western architecture does not support “the human need for transformation”. Cyberception can enable a new form of architecture to support post-biological life, characterized by “artificially enhanced interactions of perception and cognition”. This approach was pioneered by Derrick de Kerckhove (2002), who has connected Ascott’s concept of cyberception as architecture directly to environmental applications of AI. The architecture of intelligence (de Kerckhove 2002) describes a two-way road bridging material levels of reality to cognitive levels of experience.

**Architectural Precedence**

The interconnected relations between psychology, digital technologies and architecture are clarified by deCOi Architects in their seminal paper Technological Latency: from autoplastic to alloplastic (2000). This work marks the beginning of the application of digital sensing technologies to built environments to design transformative cognitive experiences. Embedded cyberception expands the transformational potential of architecture causing “a shift from an autoplastic (a self-determinate operative strategy) to an
alloplastic (a reciprocal environmental modification) mode of operation” (deCOi 2000).

Phillip Beesley and Mark Goulthorpe are two key architects who have sought to leverage emerging technologies to evoke and bring to light various psychological responses through their work. In “Hypo-surface” (2003), Mark Goulthorpe designed a faceted metallic surface that deforms physically as a real-time response to environmental stimuli. The architect defines this work as “alloplastic” - a term attributed to Freud by Ferenczi (1994) to describe an individual’s psychological attempt to adapt to a situation by changing their external environment. The concept of alloplastic adaptation [the outer-oriented adaptive response of an individual to a challenging reality] has been established in classical psychological theory in a parallel to autoplastic adaptation [the inner-oriented adaptive response to the same reality].

In his architecture, Goulthorpe enables the “possibility of a reciprocal transformation in which both subject and environment negotiate interactively” (Goulthorpe, 1999). In the context of embedded cyberception, processes of reciprocal transformation that are mediated by reflective surfaces presuppose an interplay between human and artificial empathy. Even though, from a classical psychological perspective, all transformational processes are initially autoplastic and originate from individual attempts of self-transformation, ultimately the result of autoplastic adaptation is to generate collective processes of alloplastic adaptation. Alloplastic social interventions come from “advanced technological societies” and “are generally characterized by ‘alloplastic’ relations with the environment, involving the manipulation of the environment itself” (Malmgren 1991).

Alloplastic adaptation in architecture has been connected to processes of broad cultural and social evolution (deCOi 2000, Malmgren 1991) that when mixed with digital technologies that allow real-time interactivity can simulate human empathy. Phillip Beesley is a strong example of alloplastic architecture, leveraging technology to simulate empathy in physical spaces. In his “Hylozoic Series”, Beesley designs instability, fragility and weakness into the structural meshworks of the installations. By absorbing and dissipating environmental stresses, the structures inspire a new architectural lexicon. Beesley (2012) uses the words “shuddering”, “resonating”, and “oscillating” to describe the creation of a “felt space of empathy and exchange” in this work.

While these examples focus on alloplastic adaptation using real-time transformation and manipulation of environments to elicit and assess collective psychological responses, ‘The Emoting City’ focuses instead on autoplastic adaptation at an user’s cognitive level, enabling processes of self-transformation that are based on artificial empathy generated by cyberception across reflective surfaces.

**Layers of The Emoting City**

‘The Emoting City’ brings forward a multi-layered conceptual model that uses intelligent environmental design to create artificial empathy. It operates across three main layers of interactivity (Figure 2): 1. Computational, 2. Environmental and 3. Human. This simultaneous interactive layering of computational, environmental and human dimensions of experience embodies Ascott’s concept of cyberception by creating moist environments that blend the dry dimension of technology and artifacts (AI, hardware, built objects) to the wet dimension of living multispecies (humans, plants, natural life).

- **The first layer** represents the dry computational part of interactivity and contains artificial technologies such as various kinds of AI systems: affective computing systems, data storage and processing, sensing, synthesis of automated behaviour.
- **The second layer** represents the moist tangible spaces that bring together physical artifacts and living non-human entities forming the threshold of interactivity. This layer is composed of a mesh of connected devices and elements which hunt and gather affective data from the environment and its inhabitants.
Figure 1. “The Emoting City” conceptual diagram depicting applications and components across 3 layers: computational, environmental, and human.

Figure 3. ‘The Emoting City’ installation: an affective computing network of smart mirrors.
The third layer represents the wet dimension of human cognition, composed of neurological and psychological elements that determine the emotional meaning and transformational potential of interactivity.

REFLECTIVE SURFACES

In the emerging research field of artificial empathy, human-AI interfaces have mostly been used to generate useful data for machine learning that aims to simulate human-empathic responses for the purpose of user engagement and therapy. Current research on artificial empathy has been restricted to machine learning systems that are trying to mimic human empathy in their user interfaces. This approach is unidirectional: human-AI interaction happens for the purposes of improving machine-simulations of human empathy. What we lack is a multidirectional approach to artificial empathy that leaves behind social robotics and moves forward into non-anthropomorphic embedded cyberception (Ascott 1995).

By placing robotic emotion-sensing mirrors in public spaces to reflect passerby’s emotions back to themselves, ‘The Emoting City’ enables multidirectional cyberception (Figure 3: the information being sensed is reflected back to the individual user in an immediate and transparent way, and can be used to generate self-empathy and insight. This cyberceptive informational loop can then generate artificial empathy.

A network of robotic mirrors act as interactive surfaces that allow autoplastic emotional self-insight (Figure 4). The mirrors use visual cues to discern emotion states by analyzing a real-time video feed captured by an internal camera. We have used two real-time facial recognition toolkits to enable cyberception: AffdexSDK[2] and PoseNet [3]. AffdexSDK provides real-time facial coding of multiple human expressions. It is trained on the world’s largest dataset of facial expressions (McDuff, 1996). It works by tracking facial features in order to determine facial actions (ie. smile, frown). The combination of facial actions are then used to infer seven categories of emotion expression (anger, disgust, fear, joy, sadness, surprise and contempt), to which it assigns a score from 0 to 100. PoseNet then complements this categorization and measuring with a Google-developed real-time machine learning model that performs human pose estimation (Papandreou et. al, 2018). PoseNet is able to detect and track the motion of the human body by decoding “key-points” from images and video.

The data collected from this apparatus of emotion sensing is then used to inform robotic responses using movement and light. Robotic movement mimics an user’s motions at the same time as the user receives visual stimuli from different colored lights that match the user’s facial expressions to a colour cate-
category based on emotions. A pan-tilt mechanism allows each mirror to swivel in two axis, providing the ability for the robot to perform mimicry of body gestures; a common approach applied in social robots to elicit responses in users (Shazwani et al, 2019). In addition, seeing their emotions decoded visually generates emotional self-awareness which in turn brings about a process of artificially-mediated self-empathy. Therefore, ‘The Emoting City’ is able to connect mimicry-based empathy to reflective processes of self-empathy. In human societies, it is common for individuals to mimic the facial expressions of others to communicate empathy, preference and affiliation. The mirrors elicit responses from people using gestual mimicry and improves processes of self-awareness by using colored light and motion as emotional tools for self-insight. Because the mirrors do not possess anthropomorphic embodiments, we chose to use color as a visual reflection of emotional states. This decision was based on the observation that human subjects do tend to consistently ascribe the relative amount of redness and yellowness in a face (defined along the two colour axes of CIELAB color space) to the exemplars of six classic emotion categories (Thorstenson, Elliot, Pazda, Perrett and Xiao, 2018).

By focusing on environmental-based artificial empathy experiences and mediated individual processes of self-insight, this approach contrasts other developments such as social robots. Unlike most social robotic applications that use anthropomorphic embodiments to simulate human communication processes, and other virtual implementations of emotion sensing that aim to mirror emotional states by mimicking human empathy, we do not try to give a human likeness to AI (Figure 5). We move beyond the anthropomorphic mimicking of human emotions by social robotics into using environmental robotics as self-effacing surfaces that can decode and replicate a user’s emotions for the purposes of self-empathy and personal insight.

**CYBERCEPTIVE ARCHITECTURE**

The cyberceptive architecture of the emoting city installation is illustrated in figures 6 and 7. The software, hardware, and social components of ‘The Emoting City’ have been designed to be deeply intertwined and able to function, adapt, and scale to changing spatial and temporal dimensions. A custom web-based application was developed to be run on a high-performance laptop, through a local NodeJS server. On the back-end, each mirror uses two SDK’s to support artificial perception. Firstly, PoseNet provides human pose tracking, returning key points on the human face, such as the eyes and nose. Secondly, Affedex provides facial coding and maps facial expressions to human emotion states. The software architecture is configured to simultaneously run multiple instances of the mirror program asynchronously. In each instance, face position and expression data captured by the webcam are translated into dynamic motion and lighting behaviours. Using WiFi WebSockets, the computer communicates unique instructions to each mirror.
**Robot Assembly**

Figure 8 describes the assembly of each mirror. They are designed with an automated pan-tilt mechanism composed of 3D printed as well as off-the-shelf components (see Figure 3). Each assembly is equipped with a SQ11 micro-usb camera (J) for facial recognition and a 1.5m WS2812B RGB LED strip (O) to display the colors in relation to the emotions. The assembly is designed to be mounted on generic tripods using a 1/4"-20 D socket (G). The pan-tilt mechanism is governed by two 15kg torque servos (B,H). Both servos and LED strip are powered by an external 5 Volt power adapter. The movement and LED color of each mirror is controlled by an ESP8266 wi-fi module (F) connected to a central WIFI-hub. Additionally, the pan-tilt mechanism hosts the “smart mirror” consisting of two 1mm thick round acrylic disks (L,P) covered with a one-way mirror film. The disks sandwich the LED strip at the perimeter pointing inwards with a spacer (N), creating the infinity-mirror illusion once activated. The infinity mirror assembly is directly attached to the tilt bracket (K) through an opening, and fastened by two 3mm screws from the back to a 3D printed fastener (M).

**Adaptive Behaviour**

Each mirror is encoded with the ability to change its behaviour based on the surrounding context. Distance is estimated by calculating and comparing the pixel distance between facial features. The mirrors apply intelligence to modulate levels of interactivity based upon oscillations in the number of users in a given environment. When their context is crowded, and multiple faces are identified, the mirrors can limit the distance in which faces will be tracked and analyzed. When there are few users around, the mirror either enters a sleep mode assuming a neutral color (search mode), exhibiting a continuous pre-set motion pattern in order to try to find new users and encourage engagement.

Over the course of the day, and responding to variations in user flow and movement, the behaviour of the robots oscillates between “shy” or “extroverted” modes of operation. When operating on “extroverted” mode, the mirrors actively seek out interaction with human faces. When operating on “shy” mode, each mirror assumes an avoidance behaviour, moving in an opposite direction in relation to the user, encouraging a higher level of engagement from the user.

**CONCLUSION**

‘The Emoting City’ explores the emerging field of artificial empathy (McStay 2018) in the context of environmental design and alloplastic architecture. The conceptual inquiry and technical approaches dis-
Humans process information based on their surrounding context and recognize emotions in an integrated manner (Barrett, Lindquist, & Gendron, 2007; Masuda et al., 2008). As a result, this work can be improved by continuing to explore and develop new systems to engage a wider range of sensory modalities tied to emotion beyond facial movements, such as voice prosody (Larsen et al., 2008) and body movement. To achieve this, multisensory increments can be added to future iterations of the mirrors.

An initial idea for technical innovation would be to incorporate auditory elements for emotion perception and elicitation (Larsson et al., 2009). The auditory feature would include earcons devoid of emotional meaning (noise and tone complexes) rather than identifiable sounds to give the mirrors their own unique identity, detached from any association with real-world entities. In fact, the ability of non-musical and non-vocal sounds to induce emotions is proven by research (Todd 2001, Vitz 1973), with the latter being associated with a pleasant-unpleasantness dimension of experience (Wundt 1924). Beyond expanding sensing capabilities to afford multisensory data processing, a further avenue for experimentation would be to integrate autonomous intelligent agents such as therapeutic bots (ie Replika) to the devices.

Questions for Further Research
If intelligent machines can learn how to simulate human empathy based on human-AI interaction, what can humans learn about their own selves when interacting with Als? How can we build autoplasic environments? Can human-AI interactions serve as mirrors to evaluate how humans can achieve higher levels of self-empathy? If yes, what would be the ethical and psychological implications of designing responsive environments that can sense, reflect and influence emotions?
REFERENCES

deCoi Architects, n/a 2000, 'Technological latency: from autoplastic to alloplastic', Digital Creativity, 11:3, pp. 131-143


Giorgio, Cipolletta 2018, 'Ubiquitous Bodies: A "Metrobodily" Transition from Real to Virtual', in Guazzaroni, G (eds) 2018, Virtual and Augmented Reality in Mental Health Treatment, IGI Global, Hershey

e Kerkhove, Derrick 2002, The Architecture of Intelligence, Zeitschrift Für Interaktive Und Kooperative Medien


Mcduff, D 2016 'AffDex SDK,' 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems


Papandreou, T, Zhu, T, Chen, LC, Gidaris, S, Tompson, J and Murphy, K 2018 'PersonLab: Person Pose Estimation and Instance Segmentation with a Bottom-Up, Part-Based, Geometric Embedding Model,' ECCV 2018 LNCS

Pask, G 1976, Conversation Theory: Applications in Education and Epistemology, Elsevier, Amsterdam

Pavlicsk, p 2019 'When your Internet Things Know How You Feel,' Interact Conference, London


Vitz, P 1972, 'Preference for Tones as a Function of Frequency (Hertz) and Intensity (Decibels),' Perception & Psychophysics, 11 (1)(10.3758/BF03212689), p. 84–88


Wright, P and McCarthy, J 2008 'Empathy and experience in HCI: CHI