Int. J. Human-Computer Studies 82 (2015) 31–47

Contents lists available at ScienceDirect

Int. J. Human-Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

Using affective and behavioural sensors to explore aspects of collaborative music making ☆

Evan Morgan *, Hatice Gunes, Nick Bryan-Kinns

School of Electronic Engineering and Computer Science, Queen Mary University of London, Mile End Road, London E1 4NS, UK

ARTICLE INFO

Article history:
Received 6 August 2014
Received in revised form
1 May 2015
Accepted 2 May 2015
Communicated by Paul Van Schaik
Available online 15 May 2015

Keywords:
Psychophysiology
Affect
Collaboration
Creativity
Improvisation
Music

ABSTRACT

Our research considers the role that new technologies could play in supporting emotional and non-verbal interactions between musicians during co-present music making. To gain a better understanding of the underlying affective and communicative processes that occur during such interactions, we carried out an exploratory study where we collected self-report and *continuous* behavioural and physiological measures from pairs of improvising drummers. Our analyses revealed interesting relationships between creative decisions and changes in heart rate. Self-reported measures of creativity, engagement, and energy were correlated with body motion; whilst EEG beta-band activity was correlated with self-reported positivity and leadership. Regarding co-visibility, lack of visual contact between musicians had a negative influence on self reported creativity. The number of glances between musicians was positively correlated with rhythmic synchrony, and the average length of glances was correlated with self-reported boredom. Our results indicate that ECG, motion, and glance measurements could be particularly suitable for the investigation of collaborative music making.

1. Introduction

Modern technologies and interfaces for musical expression have provided new ways for humans to collaborate creatively and make music together (Bevilacqua et al., 2013). These advances are being adopted by a wide range of people, from professional musicians, to those less able to play traditional instruments. The devices can be controlled by various means, such as touch surfaces, gesture recognition, and accelerometers (Kuhara and Kobayashi, 2011; Mitchell et al., 2012; Kuyken et al., 2008), allowing sounds to be created through novel modes of expression. Despite their heightened ability to sense human input, these devices predominantly focus on single-user operation and do very little to sense and support the important emotional, interpersonal, and communicative elements of collaborative music making (Carlile and Hartmann, 2005; Fencott and Bryan-Kinns, 2010).

Our research is motivated by a desire to develop interfaces that specifically take into account the presence, behaviours, and emotions of musical collaborators. Such an interface could be integrated into the musical instrument itself, or it could function alongside both traditional and modern instruments. The underlying aim is that it should support and enrich the affective aspects of collaborative performance, and in turn, the creative musical outcomes. In order to achieve this, we must first gain a better understanding of the behavioural and affective processes that accompany collaborative music making and creativity. There is very little existing research in this area, and our work is largely influenced by the fields of Affective Computing and Psychophysiology; both of which deal with the detection of people's inner psychological states based upon externally observable and quantifiable measurements. In the following section we provide an overview of related research in these fields. The remainder of the paper documents a study, where we collected continuous behavioural and affective measurements from pairs of improvising drummers. We use these measures to explore various aspects of collaborative music making.

2. Related work

Our multidisciplinary approach is influenced by existing research in a number of different fields. We can broadly organise these studies into two categories: (i) studies relating to the measurement of behaviour and affect, and (ii) studies dealing with the analysis of co-present, collaborative music making activities. This section discusses relevant work within these two categories.

☆This paper has been recommended for acceptance by Henrik Iskov Christensen.

* Corresponding author. Tel.: +44 7815 481 041.
E-mail addresses: e.l.morgan@qmul.ac.uk (E. Morgan), h.gunes@qmul.ac.uk (H. Gunes), n.bryan-kinns@qmul.ac.uk (N. Bryan-Kinns).

http://dx.doi.org/10.1016/j.ijhcs.2015.05.002
1071-5819/

2.1. Measuring behaviour and affect

A major component of our research is the use of sensor technologies to measure and quantify aspects of human behaviour and affect. The apparatus and techniques that we employ are influenced by research across a range of disciplines, but most prevalently *Psychophysiology*, *Affective Computing*, and *Social Signal Processing*. We will briefly introduce these disciplines before highlighting literature that is specifically relevant to this study.

Psychophysiology is the study of how psychological experiences (thoughts, feelings, emotions) relate to the physiological activity of the body. The typical approach to psychophysiological studies is to measure physiological variables in the lab, using equipment developed for medical diagnostics. These measurements are then compared to qualitative and quantitative measures of behaviour and experience, according to the specific focus of the study. Over recent years equipment for physiological measurement has become increasingly non-invasive, miniaturised and affordable, making it easier to conduct studies, not just in the lab but also in naturalistic settings (Morgan et al., 2013). Furthermore, these developments are leading towards the integration of physiological sensors in everyday technologies such as phones, smartwatches, and computer game consoles. An advantage of adopting physiological measurement as a means for inferring psychological states is that it often requires no conscious effort on behalf of the user. The development of a car seat that can sense the driver's heart rate and detect tiredness (Edwards, 2013) is great example of this paradigm; involving technology development (Walter et al., 2011), research in psychophysiology (Patel et al., 2011), and an application that revolves around the need for subconscious measurement.

The field of *Affective Computing* also concerns the measurement of psychological states. However, it focuses more specifically on the development of technologies that are able to recognise, react to, and/or express emotions. The work in this field has mainly focused on categorising the discrete emotional responses of individuals who are presented with pre-recorded, static or virtual stimuli, usually in a laboratory setting (Sariyanidi et al., 2014).

Social Signal Processing (SSP) (Pentland, 2005) is a relatively new domain, which aims to provide computers with the ability to sense and understand human social signals (Vinciarelli et al., 2009). In this context a social signal is defined as a "communicative or informative signal that, either directly or indirectly, provides information about 'social facts', that is, about social interactions, social emotions, social attitudes, or social relations" (Pantic et al., 2011, p. 8). SSP researchers develop tools and techniques for the sensing and machine analysis of behavioural and psychological constructs. A thorough survey of current work in SSP can be found in Vinciarelli et al. (2012).

A common theme within these disciplines is that human experience and affect can be segmented into three components; cognitive (thoughts), behavioural (expressions and actions) and physiological (biochemical and electrical changes in the body). Measuring each of these components presents varying challenges, requiring distinct technologies and processing techniques. The following sections discuss relevant research relating to each component, since our study incorporates measurements of all three.

2.1.1. Cognitive measurement

Neuroimaging techniques allow us to obtain information about cognitive processes inside the brain. The most commonly used techniques are EEG, which involves placing electrodes on the subject's scalp; and fMRI, which involves the subject lying still in a large magnetic scanner. Both techniques have been used to identify felt emotions by analysing the brain's response to affective stimuli such as images (Schneider et al., 1997) and music (Schmidt and Trainor, 2001). Neuroimaging studies have also sought to uncover links between brain activity and creative behaviour. A comprehensive review of such studies is provided by Dietrich and Kanso (2010), where the authors highlight that the literature is, on the whole, fragmented and inconclusive. They broadly conclude that tasks involving creative cognition induce changes in prefrontal activity of the brain. On top of this they suggest that creativity may not even be localisable in the brain, given limitations of current neuroimaging systems.

Regarding the practicalities of neuroimaging, more user-friendly EEG systems have been developed in recent years, however they still suffer from high susceptibility to noise, and a poor spatial imaging resolution. fMRI has far better spatial imaging resolution, but the requirement for participants to be stationary inside a large and immobile scanner make it impossible to conduct studies in naturalistic settings. For example, in a study of the neural aspects of musical improvisation, jazz pianists were asked to play whilst lying in an fMRI scanner (Limb and Braun, 2008).

2.1.2. Behavioural measurement

In the context of our research we are interested in measuring relatively short-term behaviours (in the order of seconds and minutes), many of which can be categorised as non-verbal communicative acts. Argyle (1978) outlines seven forms of non-verbal communication (NVC): *facial expressions, gaze, gestures, bodily posture, bodily contact, spatial behaviour* (e.g. proximity), and *appearance* (e.g. clothing). He models NVC as a simple, communications theory-inspired sequence, whereby a sender encodes a 'social signal', which is subsequently decoded by a receiver. It is implicit that this signal transfer is not error free. One person will never perfectly interpret the non-verbal communicative act of another person. Furthermore, the process of NVC does not always involve conscious awareness. This makes it a particularly interesting parameter in the study of human interactions, as it indicates subtle features of the interaction that cannot be revealed through self-report based measures.

Because we are investigating co-present musical interactions, we will be most concerned with gaze, bodily posture, and spatial behaviour. Gaze can reveal a great deal about the dynamics and nature of co-present human interactions. Numerous studies have shown how gaze is closely synchronised with speech during conversations (Kendon, 1967; Cummins, 2012; Oertel et al., 2012). Additionally, the amount of time people spend looking at each other has been shown to relate to dominance and rapport (Argyle, 1978). Mutual gaze has also been shown to be physiologically arousing (Mazur et al., 1980). Gaze is commonly measured by manually annotating video footage, which is a time-consuming task. However, modern eye-tracking glasses are able to continuously track where someone is looking within a scene captured from a head-mounted camera.

With respect to bodily posture and spatial behaviour, accurate measurements can be obtained using a marker-based motion tracking system, which uses multiple cameras to detect small reflective markers positioned on the body. Glowinski et al. (2013) used such a system to study the bodily movements of a string quartet. Their results suggest that head movement features can be used to distinguish between an engaging and non-engaging performance (as rated by the performers). Healey et al. (2005) examined the spatial behaviour of a group of seven improvising musicians. They observed how the use of space played a complex role in maintaining the coherence of the performance, and drew a number of parallels with conversational interactions. The trade-off with marker-based systems is that they take some time to set up and are not particularly portable. With the advent of the Microsoft