What does touch tell us about emotions in touchscreen-based gameplay?

YUAN GAO, Computer Science, University College London
NADIA BIANCHI-BERTHOUZE, UCLIC, University College London
HONGYING MENG, Electronic and Computer Engineering Department, Brunel University

Nowadays, more and more people play games on touch screen mobile phones. This phenomenon raises a very interesting question: does touch behavior reflect the player's emotional state? If possible, this would be a valuable evaluation indicator for game designers but also for realtime personalization of the game experience. Psychology studies on acted touch behaviour show the existence of discriminative affective profiles. In this paper, finger stroke features during gameplay on an iPod were extracted and their discriminative power analysed. A system was built and tested to recognize four emotional states (Excited, Relaxed, Frustrated and Bored) and 2 levels of arousal and two of valence. Discriminant Analysis of the collected data shows that pressure features discriminate frustration states from the other three states. Stroke speed and directness features discriminate between different levels of arousal whilst stroke length features discriminate mainly boredom from a relaxed state. Three machine learning algorithms were used to build a person-independent automatic emotion recognition system based on touch behaviour. All three algorithms produced very interesting results in discriminating between 4 emotional states reaching between 69% and 77% of correct recognition. Higher results (~89%) were obtained for discriminating between two levels of arousal and between two levels of valence. These results highlight the potential of using touch behaviour as a non-obstructive way to measure users' emotional states in contexts where touch-based devices are used.

1.   INTRODUCTION

Mobile games are becoming very popular these days. According to [NilsenWire 2011], games are the most popular mobile app category at 64%, ahead of weather apps, social networking apps, and maps and search apps. Among all types of smart phones, iPhone is the most successful gaming device. The statistics on the blog report that iPhone users spend almost twice as much time playing games on the phone than the average smart phone gamer with 14.7 hours per month.

In parallel with this increasing use of touch-based devices to play games, there is also an increasing interest in the game industry in adding emotion recognition capabilities to games (see for example Microsoft Kinect) to increase the engagement experience of the player [Bianchi-Berthouze in press]. This capability could be used to either evaluate new games or to create games that make use of the player's affective states to create a personalized engaging game experience. For example, it could be used to adapt the difficulty level to maintain the right balance between challenge and player's skills [Csikszentmihalyi 1990], or to decide the evolution of the gameplay at run-time. An example of the latter is provided by [Plass-Oude Bos et al. 2010] where the players' character shape and its powers depend on the player's stress level. When the player's stress level is high, the player's character assumes the shape of an aggressive bear with combat capabilities, whereas it returns to be an elf with cast spell skills when the stress level is low. The stress level is measured through a brain-computer interfaces that captures the player's brain alpha waves.

Biosensors, facial and vocal expression recognition [Calvo and D'Mello 2010], body movement/posture [Kleinsmith and Bianchi-Berthouze in review] and EEG [Nijholt, Plass-Oude Bos, and Reuderink, 2009] have been explored to measure user's emotional states with interesting results in different contexts. However, these affective modalities can be intrusive and cumbersome in a mobile situation. Biosensors have the problem of being very sensitive to environment conditions, whereas facial expression recognition is computationally very costly. This raises the interesting question of investigating touch behavior as an affective communication modality. The monitoring of this modality would be less intrusive and at the same time less computationally expensive. Recent studies in psychology literature [Hertenstein, Holmes, McCullough, Keltner 2009] provide

support to this line of research as they have shown that touch behaviour may convey not only the valence of an emotion but also the type of emotion (e.g., happy vs. upset). Unfortunately, these studies have been carried out only in a social context (person-person communication) and only through acted scenarios. This paper aims at exploring the possibility of using this modality to capture a player's emotional state in a naturalistic setting of touch-based computer games.

This paper is organized as follow. First, we review the literature on touch behaviour as a modality to recognize emotions as well as on related modalities such as keystroke typing and body movement. Then we present our protocol to collect and label the data for building and testing our emotion recognition system. We then analyze the extracted touch behavior features and discuss their relationship with the emotional state of the players. Automatic recognition models are then build and tested on person independent context. We conclude by discussing the results.

2.   BACKGROUND

2.1   Emotion recognition by touch behavior

An underexplored modality for automatic emotion recognition is touch. This is possibly due to the fact that even if of great social importance, this modality has received very little attention from this perspective even in the field of affective science [Hertenstein et al. 2009]. Initial studies on touch behaviour as an affective modality argued that the role of touch was mainly to communicate the valence of an emotion (i.e., positive vs. negative emotions) and its intensity [Jones and Yarbrough 1985; Knapp and Hall, 1997]. More recently, two consecutive studies by Hertenstein et al. [Hertenstein et al. 2009; Hertenstein, Keltner, App, Bulleit and Jaskolka 2006.] have instead shown that touch communicate much more about emotions in interpersonal communication. In their first study [Hertenstein et al. 2006], they showed that participants could discriminate between different emotions when they were touched by another person (a stranger) on their forearm without seeing the hand or the person. The recognition of anger, fear, disgust love, gratitude, and sympathy reached performances between 48% and 83%. The subsequent study [Hertenstein et al. 2009] was carried out in a more ecological context but still with acted touch expressions. The results confirmed the findings of the previous study and showed that also happiness and sadness could be decoded with recognition rates higher than chance level. The recognition performances reached with touch alone were indeed comparable with the ones obtained with other affective modalities [Elfenbein and Ambady, 2002].

What are then the features that facilitate the discrimination of the different types of affective touch? In their recent study, Hertenstein et al. investigated 23 different types of tactile behavior (e.g., stroking, squeezing, patting). They found that the tactile behaviour alone was not sufficient to discriminate between different affective messages, i.e., the same type of tactile behaviour was shared by different emotions. For example, stroking was observed when communicating sadness but also when communicating love or sympathy. From a more detailed analysis of touch behaviour they found that two more tactile behavior qualities were necessary for an accurate discrimination between emotions: duration of the tactile behavior and intensity (i.e., "… *the amount of skin deformation created by the touch on the receiving person*").

Whereas these studies clearly promote touch behaviour as a fine grained affective modality, they have being carried in acted situations, and it is not clear how these findings would vary in a naturalistic context. Furthermore, these studies relate to inter-personal communication and it is hence necessary to understand if and how they translate to a context of interaction with a touch-based device and in a game play situation.

2.2   Touch as an affective communication channel in HCI

Studies in neuroscience have shown that our skin has special receptors dedicated to detect pleasurable touch. In [Essick, McGlone, Dancer, Fabricant, Ragin, Phillips, Jones, Guest 2010], Essick et al. describe the mechanisms underlining the skin sensors in charge of detecting the pleasurable experience of touch. There extensive experiments show that these sensors respond particularly to gentle dynamic touch. It has also being shown that hedonic touch can promote hormonal responses that can facilitate bond [Uvanas-Moberg, Arn, Magnusson 2005].

Given these properties, various studies have been carried out to explore if touch can enhance bond and engagement also with or through technology. This modality has been for example explored in Human-Robot interaction to facilitate bond. Robots have been integrated with touch sensors over their body to capture people touch behaviour and provide affective displays in response. Examples of such robots are Shibata's seal Paro [Shibata, Mitsui, Wada, Touda, Kumasaka, Tagami, Tanie 2001] and the Huggable [Stiehl, Lieberman, Breazeal, Basel, Lalla, Wolf 2005]. In these robots, the mapping between the touch qualities (e.g., intensity) are predefined (i.e., not automatically learnt) and the responses humans receive to their touch provides a sense of robot's agency and desire of communication.

Touch has been also explored through the use of haptic devices to facilitate distance communication [Noma and Miyasato 1997; Oakley, Brewster and Gray, 2000; Strong and Gaver, 1996]. Most of these works has shown through informal testing that virtual touch received positive reactions from the participants. In [Chang et al. 2002], the authors designed a vibrotactile communication device where the two interlocutors could enhance their verbal message through vibrotactile signals. It was mainly test to see how turn-taking cues, emphasis of message, tapping mimicry to signal presence. The preliminary results shows that this information could enhance communication and could also provide information in case the verbal communication was limited. Another interesting study is presented in [Ho, Srinivasan, Slater 2000], where haptic feedback is investigated as a possible mean to facilitate collaborative task in shared virtual environment. The results show that participants not only produced higher performance when touch feedback were provided but also reported an increase sense of togetherness. Another interesting study is presented in [Bonanni, Vaucelle, Lieberman, Zuckerman 2006] where a haptic scarf is used to send affectionate and nurturing touches to the wearer. Touch behaviour is captured and recorded by the scarf. The wearer can then play back the touches when apart from the person that recorded them. The recognition of the meaning of the touch behaviour is left to the receiver's interpretation of it.

A study more in line with ours is reported in [Bailenson, Brave, Merget, Koslow 2007]. The authors investigated if a two-degree of freedom force feedback joystick could be used to communicate emotions. Participants were asked to communicate through the use of the joystick the emotion of anger, disgust, fear, interest, sadness and surprise to another person (called here receivers). Using the x and y coordinates of the joystick recorded every 5ms, measures of distance, acceleration, jerkiness and direction of the movement. Using ANOVA, they investigated the discriminative power of each measure. The results showed that distance, speed, acceleration were much higher in joy and anger than for other emotions and showed the smallest values when expressing sadness. The direction measures were also quite informative with fear showing smaller use of the major axis and sadness showing smaller use of the minor axis. The measures were used to build seven classification models (one for each emotion) by using a SVM machine learning algorithm. The results were compared with the human receivers' discrimination performances. The receivers where hence asked to classify the type of emotion that had been communicated through the haptic devices. The results show that the receivers recognized the correct emotion 33% of time (chance level = 14%). The SVM obtained similar recognition performances (36%). The authors conclude that the emotions could be communicated well above chance level and that higher performances could be obtained with a more complex device. To this purpose they tested the ability of participants to recognize the mental state of other through handshake. The test reached 51% of correct recognition. These results are quite encouraging as they indicate the possibility of building machine with the ability to recognize the affective content of touch. However, as in the previous section, these messages were acted and hence strongly stereotyped and forcefully communicated. As the authors [Bailenson et al. 2007] of this study acknowledge as well as other researchers (e.g., [Wallbott and Scherer 1986; Nass and Brave 2005; Kleinmsith et al. in review]) in the field of emotion communication, natural expressions differs from acted ones.

Furthermore, whilst these studies do show the importance of this communication modality, they still focus on interpersonal touch and cannot be directly translated into game context where touch may not always be used to purposely convey emotion or to communicate with someone else. Before presenting our study, we first review works on automatic emotion recognition on two modalities that may share characteristics with touch behaviour: keystroke tapping and body movement.

2.3   Emotion recognition by keystroke

Initial studies on the existence of finger pressure profiles reflecting emotion were reported in [Clynes 1973.]. In his work on sentography, Clyne investigated and showed the existence of vertical and horizontal finger pressure profiles for different emotions. The data were collected in a controlled situation with no finger or arm movement (the finger is on a finger rest). Following this direction of investigation, a few studies in human-computer interaction explored the possibility to automatically detect emotions from typing behaviour and mouse hitting.

In the study reported in [Khanna and Sasikumar 2010], the authors found that keyboard typing behaviour is affected by user's emotional state. People reported not only that their emotional states affected the frequency of selection of certain keys (e.g., backspace) but that in a positive mood their typing speed tended to increase. They carry out further experiment to quantitatively characterize such behaviour. People were asked to type in either a positive or a negative mood. The results partially supported such self-report showing an increase in speed in positive states and a decrease in speed in negative states. However, the authors states that the patterns were not so clear and this could be due to the fact that they have used broader emotional classes: positive vs. negative state. Each class includes a large variety of emotions that may indeed have different keystroke behaviour (e.g. negative states include anger as well as sadness). Given the large variability they identified, they developed a system that discriminates only between positive and neutral states or between negative and neutral states. They obtained a maximum performance of 87.7%. In a recent paper [Lv, Lin, Yin, Dong 2008.], the authors further these investigations using novel pressure sensor keyboard. The keyboard produces a pressure sequence when the user enters characters. Six different emotions (neutral, anger, fear, happiness, sadness and surprise) were elicited in the participants by getting them to immerse themselves in a set of

predefined stories. It was then asked them to type a set of words not specified in the paper. Using three different types of features (global pressure sequences, dynamic time warping and keystroke dynamics) and using a fusion techniques they reach a 93.4% recognition performances over the six emotions. Unfortunately, no confusion matrices or discussion on emotion-keystroke pressure patterns were provided.

Matsuda et al. [Matsuda, Sakuma, Jimbo, Kobayashi, Arafune, Isomura 2010] investigated the possibility to use finger behaviour to automatically recognize the emotional state of deaf people when communicating through a finger Braille device. In this study, the authors use duration and acceleration of finger dotting to discriminate between neutral, joy, sadness and anger states. The participants were asked to input two different sentences multiple times by expressing each time one of the four emotions. By using ANOVA, their results show that the duration of the dotting behaviour in the joy condition was significantly shorter than in the other conditions, whilst the dotting behaviour in the sadness condition was significantly longer. The finger load was significantly stronger in the anger condition. These results differ from the results on touch behaviour presented in [Hertenstein et al. 2009] where joy is longer than anger. This could indicate that keyboard stroking and touch behavior may be characterized by different patterns. However, also in [Hertenstein et al. 2009], touch behaviour in sadness is longer than joy but of similar duration than in other positive states. Both studies show an increase on pressure for the Anger state. Using Discriminant Analysis, the automatic discrimination model reached an average of 55% on absolute data and 82.5% with standardized data (taking into account daily variations in dotting strength). Even if the collected expressions were acted and hence prone to prototypical emotional behaviour, these results are quite promising.

These studies confirm the importance of duration and pressure qualities to discriminate between emotional states. They also identify speed as an important discriminating modality. However, they investigate one type of touching behaviour (tapping). In our study, we aim to investigate finger stroking touch behaviour (as discussed in section 3) as it is an important type of touch behaviour with touch-based game devices. To complete our review on automatic emotion recognition, we draw on the relation between finger movement and body movement, an area much more explored than the one of touch. Touch and movement are also often coupled as part of the proprioceptive system.

2.4 Emotion recognition by body expressions

The literature shows that body movement is an important modality to decode the affective states of a person (for a review see [Kleinsmith and Bianchi-Berthouze In review]). Automatic recognition systems based on this modality have shown comparable performances to systems for emotion recognition in facial expressions. Most work has been done in acted situations by exploiting both body configuration and dynamic information. Gross level descriptions of body movement have been initially investigated by Camurri in dance context [Camurri, Mazzarino, Volpe 2003]. Inspired by Laban studies [von Laban 1971], they propose a set of gross features capturing the configuration of the body (shape) and its dynamics (i.e., shape and effort in Laban's terminology. Effort is described in terms of directionality, flow (bound vs. free), time (quick vs. sustained) and weight of the movement. Their emotion recognition system from dancers' body movement [Camurri, Lagerlof, Volpe 2003; Camurri, Mazzarino, Ricchetti, Timmers, Volpe 2004] aimed at discriminating between 4 discrete emotion categories (anger, fear, grief and joy) reached correct recognition performances (average of 40%) comparable to the ones of human observers (56%). Their analysis of body movement features showed that joy movements were characterized by high fluidity as opposed to grief movements that demonstrated more jerkiness. Using a more complex architecture of the emotion recognition system, Kapur et al. [Kapur, Virji-Babul, Tzanetakis, Driessen, 2005] obtained much higher performances with recognition rates ranging from 62% to 93%.

Another interesting work presented in [Bernhardt and Robinson 2007], highlighted the importance of idiosyncrasy in the modeling of emotion recognition from movement. Their system was built to discriminate between three discrete emotional states (i.e., angry, happy, sad) expressed by knocking arm movements. Using SVM classifiers, the correct recognition rate of affective states reached 50%. However, by taking into account individual idiosyncrasies in the description of the movement, the performances increased to 81%. The recognition performances were comparable to human observers' performances (varying between 59% and 71%) for the same set of stimuli, as discussed in [Pollick, Paterson, Bruderlin and Sanford 2001].

More recently, there has been a shift from acted expression to non-acted context where the expressions are less stereotypical and possibly more subtle. The shift has been dictated by the emergence of full-body technology and the need to use such recognition system in real-life human-machine interaction. The results are very promising and similar to the ones obtained in acted context [Kleinsmith et al. in review]. We report here about two works that have been developed in game context. The first one [Kleinsmith et al. 2011] aimed to detect the emotional state of the players when the player was looking at the replay of the game just after having score or missed a point (i.e., between game points). The game used was a Wii-Nintendo sport game. The results showed that, by exploiting low-level body configuration features (i.e., angle between body segments), the systems reached on

average correct recognition rates of 60% for four affective states (concentrating, defeated, frustrated and triumphant). Their results were comparable with the human observers' level of agreement (i.e., 67% recognition rate) reached for the same set of stimuli. Their work showed also very high performances for the recognition of affective dimensions levels (arousal=energy level, valence=level of positivity of the emotional state).

The expressions investigated in this study were quite static (i.e., between games) and hence body configuration features produced high results. A study more relevant to our work is presented by Savva et al. in [Savva and Bianchi-Berthouze 2012]. They proposed a system to continuously recognize the emotional states of people while playing Wii tennis rather than between games. Dynamic features describing the movement of the arm were used for this purpose. The best results were obtained by using angular velocity, angular speed and amount of movement. A recurrent neural network that takes into account the history of the classification was used to classify frame by frame, the collected motion capture data. Overall the system was able to correctly classify 64% of the high intensity negative expressions, 67% of the low intensity negative expressions, 58% of the 'happiness' expressions and only 36% of the 'concentration' expressions. An analysis of the results highlighted the large diversity of the players' playing styles. Some of the participants tended to play the game using only their hand/wrist [Pasch, Bianchi-Berthouze, van Dijk, Nijholt 2009] in comparison with the other group that used their arm and shoulder as well.

In this study, we investigate the possibility of using finger stroke behavior when interacting with touch-based devices to detect the emotional state of the user. We investigate this question in the context of touch-based computer games because of their spread and also because they do elicit quite strong emotions in a short period of time. In the following sections, we present the data collection, modeling process and testing. We then conclude discussing the results and the possibility to use this modality to create less invasive affective-aware technology that could be used on the move.

3.   DATA COLLECTION

This section presents the material and the procedure used to capture and analyse players' finger stroke behaviour with respect to their emotional state. The most discriminative features are then used in section 4 to build person-independent emotion recognition models.

3.1 Game

The iPhone game Fruit Ninja[1] was selected for this study. Fruit Ninja is an action game that requires to squish, slash and splatter fruits. To play, players swipe their fingers across the screen to slash and splatter fruit simulating a ninja warrior. Fruit Ninja has been a phenomenon in the app store with incredible sales, and has inspired the creation of many similar drawing-based games. Being the Fruit Ninja not an open source app, an open source variation of it developed by Ansca Mobile was instead used[2]. It is developed with the Corona SDK. In order to elicit different kind of emotions from players and be able to capture players' touch behaviour, a few modifications to the Ansca's game were implemented.

Our new version of the Fruit Ninja contains 20 sessions in total. Each session lasts 30 seconds and has a target point to reach. The goal is to reach this target point by slashing enough fruits. Each fruit is worth 1 point. The players have to pay attention and avoid the bombs since slashing a bomb makes the player lose 5 points. There are no constraints on how to slash the fruits. Players can slash the fruits one by one or wait for multiple fruits to be lined up and to slash them at once. The difficulty level of the game increases as the session number increases. The difficulty is based on the speed of the game (shooting speed and angular speed of the objects) and on the ratio between the number of fruits and the number of bombs. If the player reaches the target point within the time limit, they win the session otherwise they loose it. After having finished all the 20 sessions, a final score is given according to the number of sessions they won. Figure 1 shows the screenshots of New Fruit Ninja.

---

[1]http://www.fruitninja.com/
[2]http://maniacdev.com/2011/05/open-source-fruit-ninja-type-example-game-with-the-corona-sdk/

Figure 1. Screenshots of the adapted Fruit Ninja game: slashing of a single fruit (left) and of multiple fruits (right).

The developed New Fruit Ninja game software captures and stores players' finger stroke behaviour during game play: the coordinates of each point of a stroke, the pressure area of the finger at each point and the time duration of a stroke, i.e., the time occurring from when the finger touch the display to when the finger is lifted. Figure 2 shows an example of two different strokes. In order to take into consideration physical differences between individual, i.e., the dimension of the tip of a finger, a baseline touch area is computed before the game starts. Each participant is asked to perform a stroke. The average of the pressure area of each point of this stroke is used as a pressure baseline for that participant. The pressure area values collected during the game are hence normalized by subtracting the participant's baseline value.
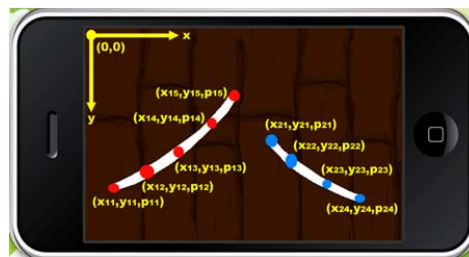

Figure 2. Coordinates of the points and their pressure areas for two different strokes.

3.2   Data collection and labelling procedure

Our experiment setup included the following items: a laptop with a camera to videotape the whole experiment process, an iPod touch, a self-assessment questionnaire used to capture player's emotional state. The self-assessment questionnaire consisted of a list of emotion words. This list was developed through a pilot study with 7 participants. During this pilot study, participants were allowed to use any words to express their emotional state. The most common emotional words used were: Delighted, Excited, Annoyed, Frustrated, Satisfied, Relaxed, Concentration and Bored. The list was further refined. First of all, concentration was eliminated as it was considered a neutral state and also because, as shown in other studies [Kleinsmith, Bianchi-Berthouze, Steed 2011] and from the result of our interview with the 7 participants, the term concentration was often used when participants were not sure of their emotional states. Secondly, the terms that belonged to the same quadrant of the valence-arousal (V-A) emotion space [Russell 1980] were grouped: Excited and Delighted, Annoyed and Frustrated, and Satisfied and Relaxed. The aim was to ensure to obtain a large set of data for each emotion class necessary to the analysis and modeling process. The final set of words was: Excited, Relaxed Frustrated and Bored.
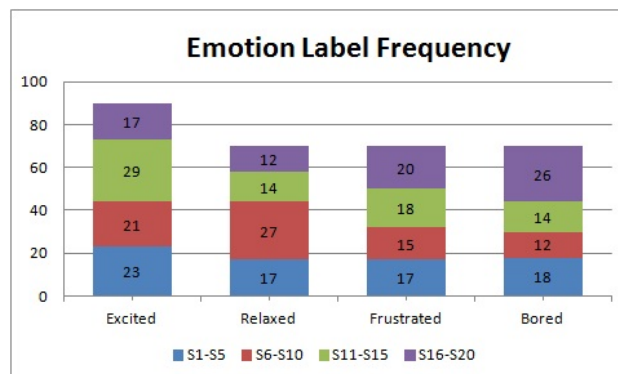
Figure 3. Distribution of the emotion labels over the 20 sessions of the game for all 15 participants. Each bin of the histogram is divided in 4 sections. Each section of a bin represent 5 consecutive sessions with increasing level of difficulty. The blue sector at the bottom of each bin represents the number of times an emotional state appeared in the first five sessions of the game (sessions 1-5); the second sector (red) from the bottom represents sessions 6-10 and so on.

The experiment procedure was as follows. We introduced and explained the experiment process, the iPod touch device and the game rules of New Fruit Ninja to participants and answered their questions. The participants were then asked to play the game and get familiar with it. After the training session, the participants were asked to play and try to complete 20 sessions of the game. As a form of motivation, they were informed that the participant with the highest final score was rewarded with an iTunes gift card worth £25. The game was video recorded. In order to decide the ground truth, participants were first asked to fill out a self-assessment questionnaire after every game session. To reduce the effect of the game results on the labeling, the Cued Recall Brief approach [Bentley, Johnston, von Baggo 2005] was used. At the end of the 20 sessions, the player was shown a recorded video of the gameplay s/he had just carried out. While the video is playing s/he is encouraged to recall his/her thoughts and emotional states and if necessary re-label the sessions.

Fifteen participants (9 male and 6 female whose average age varied between 18 and 40 years old) were invited to take part to the experiments and play the 20-session game. This provided data samples for a total of 300 sessions. During the relabeling phase, only 12% of sessions were relabelled. Figure 3 shows the distribution of the 4 emotion labels over the collected dataset: 90 excited, 70 relaxed, 70 frustrated and 70 bored.

Stroke Features	
Feature Name	**Function**
Number of strokes in a session	N
Average stroke length	$\bar{l} = \frac{\sum_{i=1}^{N} l_i}{N}$, where $l_i = \sum_{j=1}^{n_i-1} \sqrt{(x_{i(j+1)} - x_{ij})^2 + (y_{i(j+1)} - y_{ij})^2}$ and n_i is the number of points for the ith stroke.
Median stroke length	$l_{med} = $ median (l_i), for all $i = 1, 2, \ldots, N$
Max stroke length	$l_{max} = $ max (l_i), for all $i = 1, 2, \ldots, N$
Min stroke length	$l_{min} = $ min (l_i), for all $i = 1, 2, \ldots, N$
Average stroke speed	$\bar{v} = \frac{\sum_{i=1}^{N} v_i}{N}$, where $v_i = \frac{l_i}{t_i}$
Median stroke speed	$v_{med} = $ median (d_i), for all $i = 1, 2, \ldots, N$
Max stroke speed	$v_{max} = $ max (v_i), for all $i = 1, 2, \ldots, N$
Min stroke speed	$v_{min} = $ min (v_i), for all $i = 1, 2, \ldots, N$
Average stroke Directness Index (DI)	$\bar{d} = \frac{\sum_{i=1}^{N} d_i}{N}$, where $d_i = \sqrt{(x_{i1} - x_{in_i})^2 + (y_{i1} - y_{in_i})^2}$
Median stroke DI	$d_{med} = $ median (d_i), for all $i = 1, 2, \ldots, N$
Max stroke DI	$d_{max} = $ max (d_i), for all $i = 1, 2, \ldots, N$
Min stroke DI	$d_{min} = $ min (d_i), for all $i = 1, 2, \ldots, N$
Average Pressure	$\bar{p} = \frac{\sum_{i=1}^{N} p_i}{N}$, where $p_i = \frac{\sum_{j=1}^{n_i} p_{ij}}{n_i}$, p_{ij}: the size of the touching area of the jth point on the ith stroke and n_i is the number of points in the ith stroke. The touching area for each individual point is normalized with respect to a person-based baseline pressure value as discussed in the text.
Median Pressure	$p_{med} = $ median (p_i), for all $i = 1, 2, \ldots, N$
Max Pressure	$p_{max} = $ max (p_i), for all $i = 1, 2, \ldots, N$
Min pressure	$p_{min} = $ min (p_i), for all $i = 1, 2, \ldots, N$

Table 1: Stroke Features

4. FEATURE ANALYSIS

4.1 Visual Inspection

As discussed earlier, the coordinates and the pressure area of each point of each stroke as well as the time duration were recorded within each game session. For each stroke, sixteen features were extracted as listed in Table 1: length, speed, directionality index (DI) and pressure area. The features were based on the literature on touch and body movement reported in section 2. In particular, the set of features selected can be related to the effort characteristic of the movement proposed by Laban [von Laban 1971]. In Table 1, N denotes the total number of strokes in a recorded session, p_{ij} denotes the jth point on the ith stroke and x_{ij} and y_{ij} denote respectively its x-coordinate and its y-coordinate. n_i denotes the total number of points on the ith stroke and t_i denotes the time duration of the ith stroke.

At first, a visual inspection of these features was carried out to identify possible emotion-feature patterns. Figure 4 shows the boxplots of the average and median values of each type of feature. The boxplots show that the positive states (i.e., Excitement and Relaxed states) tend to have longer strokes with Relaxed having overall the longest ones. The strokes for Frustration present the largest variation. Boredom has the shorter strokes. The pressure features shows a clearer separation between Frustration and the other emotions with Frustration showing overall higher values. Speed seems instead to separate the states along the arousal dimension with Frustration and Excitement showing higher values than the other two states. The directness index (DI) is another very interesting feature showing that Excitement and Frustration have less direct stroke than the other two states with Boredom showing the most direct ones (i.e., lower values).
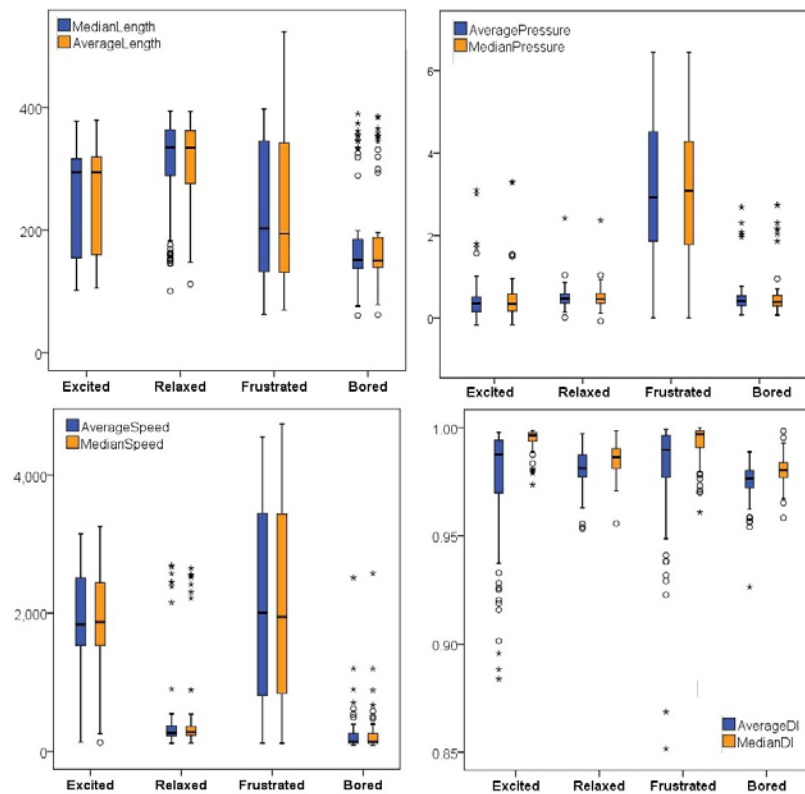

Figure 4. Boxplots of the average and median values of length, pressure, speed and DI features.

In order to verify that these patterns were independent from the difficulty of the sessions, we look at the marginal means of these features. Figure 5 shows the estimated marginal means graphs for average length, pressure, speed and directness index over different levels of difficulty of the game. Each 5-consecutive session are represented by one line. We can see that for all the four features, the lines show a similar trend independently of the difficulty of the session. Only the stroke length values for the first 5 sessions (S1-S5, i.e., the easiest sessions) shows a different trend with Frustration reaching the highest values. It is possible that normalizing these data with respect to the difficulty of the session may bring to a better classification results. However, the large difference is only for these first 5 sessions and it could be due to outliers.

The comparison of the observed emotion-touch behaviour with the one reported in section 2 is not so straightforward. Our results

seems to support the results on touch behavior reported in [Hertenstein et al. 2009] on the difference between excitement and frustration (respectively joy and anger in [Hertenstein et al. 2009]). In fact, the duration of the stroke was longer in joy states compared to anger. Also, higher pressure values were observed for anger in [Hertenstein et al. 2009]. Different results were observed for the low energy negative states: i.e. Sad in [Hertenstein et al. 2009] and Boredom in our study. Sadness in [8] has produced the longest stroke whereas in our study Boredom produced the shortest one. The difference could be that Boredom indicates the desire of disengaging from the game. This may not be the case for a feeling of Sadness in the social.

More difficult is the comparison with the characteristics of the typing and dotting behaviour reported in [Matsuda et al. 2010] and [Khanna et al. 2010]. In [Matsuda et al. 2010], differently from our study, anger and sadness produced longer dotting duration than the other two states. In [Khanna et al. 2010], speed appeared to discriminate along the valence dimensions rather than the arousal one as it seems the case in our study. However, in [Matsuda et al. 2010], frustration produced higher pressure values as in our study. This could mean that in the case of typing or dotting, the duration of the behaviour is more strongly affected by pressure actions bringing the time spent on the key/dot to be longer and hence slower. Unfortunately, a more detailed comparison of stroking versus typing is not possible as in [Khanna et al. 2010], the states are separated only along the valence dimensions.
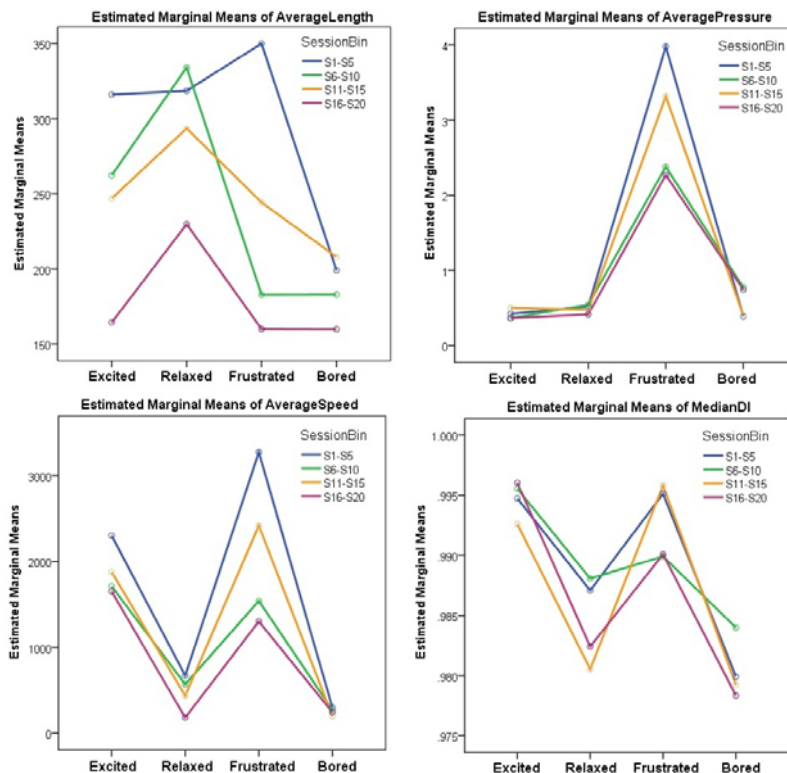

Figure 5. Estimated marginal means by session bins: average length, pressure, speed and directness index

4.2 Feature extraction

After the visual inspection aimed at investigating the relevance of the computed features, a further evaluation of their discriminative relevance was carried out. IBM SPSS version 20 was used to this purpose. Using the Bernouilli function of SPPS with parameter set to 0.4, the 300 instances were randomly split into two subsets. 173 instances were assigned to the training set and 127 to the testing set. Given the similarity and high correlation between the median and average values for length, speed and pressure features (see Figure 4), their median values were not used in this analysis.

A Discriminant Analysis (DA) was carried out to identify the most relevant features for the discrimination between the 4 emotions categories. The results showed that only the Average-Stroke-DI feature was not significant (p-value = .589). This feature was hence removed from further analysis. All the others features obtained p-values <.000 but for Min-Stroke-DI whose was still significant with p-values < .010).

Using the remaining features, the DA model was built. The results are reported in Tables 2-4. The results identified three discriminant functions whose Eigen values are reported in the top part of Table 3. The first discriminant function explain 57.6%

of the variance (canonical R^2=.823) and the second explained 34.2% (canonical R^2=.745). The third explained only 8.2% of the variance (canonical R^2=.479). The bottom part of Table 2 shows that, in combination, these three functions significantly differentiate between the four emotion groups (Λ=.111, χ^2 (48) = 259.351, p=.000).

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	2.065	57.6	57.6	.823
2	1.605	34.2	91.8	.745
3	.290	8.2	100.0	.479

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 3	.111	259.351	36	.000
2 through 3	.344	126.069	22	.000
3	.771	30.687	10	.001

Table 2. Eigenvalues and Wilk's Lambda for the first 3 discriminant functions

Table 3 shows the correlations between the stroke features and the three discriminant functions. The table reveals that pressure features loaded highly on the first discriminant function. Speed and DI features loaded highly on the second discriminant function. Finally, the length features loaded highly on the third discriminant function. The discriminant function plot in Figure 6 shows that the first function discriminates between the strokes performed in a Frustrated state from the strokes performed under the other affective states. The second function, loading on speed and DI features, discriminates Excited states from low arousal states (relaxed and bored) and from frustration. Finally, length features discriminate between boredom and relaxed states.

	Function		
	1	2	3
Max Pressure	**.856**	-.221	.321
Average Pressure	**.774**	-.228	.322
Min Pressure	**.658**	-.275	.309
Max DI	.360	**.575**	.469
Max Speed	.505	**.537**	.361
Average Speed	.463	**.529**	.348
Min Speed	.364	**.498**	.319
Min DI	-.050	**-.213**	.026
Average Length	-.095	.047	**.808**
MinLength	-.244	-.073	**.714**
Median DI	281	.522	**.533**
Max length	.077	.144	**.472**

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function.
Bold. Largest absolute correlation value between a variable and the discriminant functions

Table 3. Structure matrix for discriminant analysis for 4 emotions. In bold the most significant features.

The confusion matrices for the plot are represented in Table 4. Cross-validation results on the training set and validation results on the testing set show respectively 77.5% and 76.4% of correct classification. The best results are obtained for the classification of Frustrated and Excited strokes. 86% of Frustrated strokes are correctly classified and the performances for Excited are just below 80%.). The lowest performances (63.6%) are obtained for Relaxed samples. Most misclassifications occur between Relaxed and Bored (37.9%) and between Relaxed and Excited (23.9%).
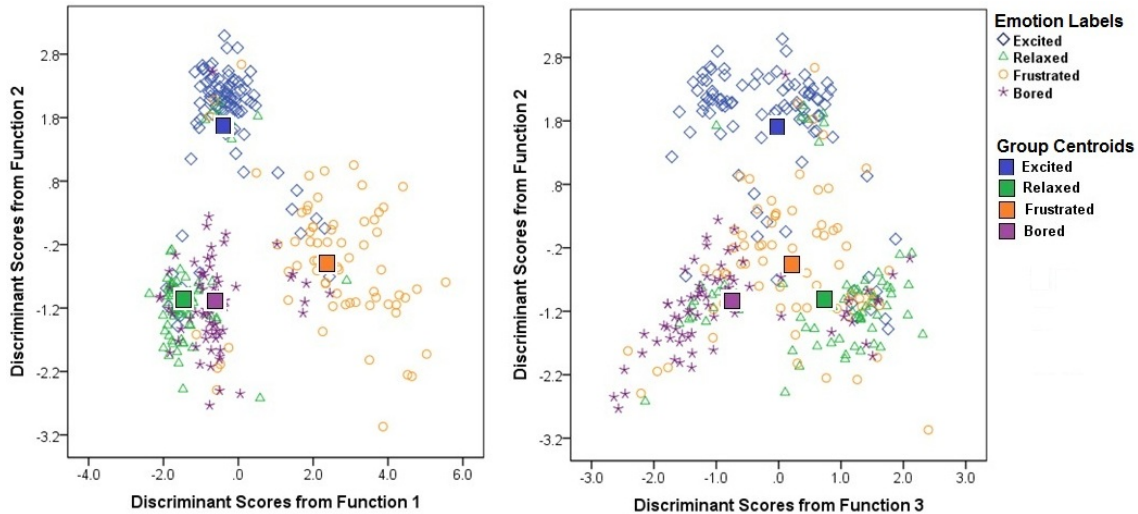
Figure 6. Results for the Discriminant Analysis for 4 emotions. Each stroke is represented by a symbol in the 3-dimensional space. The type of symbol indicates the predicted class for that stroke. The centroids of each emotion group is represented by a square. The centroids illustrate how each 3 discriminant function (and hence the features that load on them) support the discrimination of the four emotion classes.

A discriminant analysis was also performed to identify the most relevant features for arousal (low vs. high arousal state) and for valence (negative vs. positive valence states). The results are shown in Tables 5 and 6. Table 5 shows that the first discriminant function explain 100% of the variance for both arousal (canonical R^2=.793) and valence dimensions (canonical R^2=.717). It also shows the discriminant functions significantly differentiate between respectively two level of arousal (Λ=.372, χ^2 (13) = 162.773, p=.000) and two levels of valence (Λ=.486, χ^2 (13) = 118.846, p=.000).

		Classification Results					
			Predicted Group Membership				
		Emotion	Excited	Relaxed	Frustrated	Bored	Total
Cross-validation on Training set (173 Sample data)	Count	Excited	47	5	4	1	56
		Relaxed	3	27	0	7	37
		Frustrated	3	.0	33	4	40
		Bored	1	7	5	27	40
	%	Excited	83.9	8.9	7.1	.0	100.0
		Relaxed	8.1	73.0	.0	18.9	100.0
		Frustrated	7.5	.0	82.5	10.0	100.0
		Bored	2.5	17.5	12.5	67.5	100.0
Testing Set (127 Sample data)	Count	Excited	27	4	2	1	34
		Relaxed	4	21	1	7	33
		Frustrated	2	0	26	2	30
		Bored	0	5	2	23	30
	%	Excited	79.4	11.8	5.9	2.9	100.0
		Relaxed	12.1	63.6	3.0	21.2	100.0
		Frustrated	6.7	.0	86.7	6.7	100.0
		Bored	.0	16.7	6.7	76.7	100.0

76.4% of testing set cases were correctly classified.
77.5% of training set using cross-validated were correctly classified.

Table 4. Confusion matrix for discriminant analysis. Validation results using both cross-validation and testing set are reported

Table 6 shows the importance of speed and DI features to separate between low and high arousal states. Pressure appears to have less discriminative power than speed and DI for discriminating along the arousal dimensions. Length does not show any relevant weight on this affective dimension. Differently from arousal, pressure and length are mainly used to separate strokes along the valence dimension. The Discriminant Model built for Arousal obtained 89% of correct classification in both the cross-validation on the training set and in the validation on the testing set. The results for the Discriminant Model for Valence were slightly lower with 83% of correct classification in both the cross-validation on the training set and validation on the testing set.

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation

Arousal	1.653	100.0	100.0	.789
Valence	1.040	100.0	100.0	.715

Test of Function	Wilks' Lambda	Chi-square	df	Sig.
Arousal	.377	160.963	12	.000
Valence	.489	117.926	12	.000

Table 5. Eigenvalues and Wilk's Lambda for the first 3 discriminant analysis functions

5. EMOTION RECOGNITION SYSTEM

5.1 Modeling

The results of the Discriminant Analysis showed that finger stroke behaviour allows for the discrimination of the four affective states we investigated as well as for the discrimination between two levels of arousal and between two levels of valence. The results showed very high performance on both cross validation on unseen sample data. The next step in our analysis is to investigate if person-independent emotion recognition models could reach similar results. Creating person-independent model is generally harder as the data of the person on which the model is tested on are not used to build the model. This testing is very important as it shows the generalization capability of the stroke features.

Affective dimension	Features	Function	Affective dimension	Features	Function
	Max Speed	**.766**		Max Pressure	**.568**
	Average Speed	**.733**		Average Pressure	**.545**
	MaxDI	**.693**		Min Pressure	**.532**
	Min Speed	**.639**		Min Length	**-.335**
	Median DI	**.594**		Average Length	**-.328**
Arousal (2 levels)	Max Pressure	**.316**	Valence (2 levels)	Median DI	-.221
	Average Pressure	**.285**		Max DI	-.173
	Min Pressure	.211		Max Length	-.163
	MinDI	-.188		Min Speed	-.101
	Max Length	.180		Min DI	.096
	Min Length	-.175		Average Speed	-.056
	Average Length	.006		Max Speed	-.036

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant function. Variables ordered by absolute size of correlation within function. Only the higher correlations are reported.

Table 6. Structure matrix for discriminant analysis for each affective dimension. In Bold the features that load higher on the discriminant functions.

Parameter Name	Value
Number of hidden layers	3
Number of input layer nodes	equal to the number of features
Number of hidden layer nodes	same as input layer
Number of output layer nodes	Equal to the number of emotion classes (i.e., either 4 or 2)
Transfer Function (TF) for input layer to 1st hidden layer	Tansig
TF for 1st to 2nd hidden layer	Tansig
TF for 2nd to 3rd hidden layer	Tansig
TF for 3rd hidden layer to output layer	Purlin
Back-propagation transfer function (BTF)	Traing dx (Levenberg-Marquardt backpropagation)
Learning Rate	0.03
Training Epoch	1000
Momentum	0.2

Table 7. ANN Parameters

To this purpose, three modeling algorithms were selected: Discriminant Analysis (DA), Artificial Neural Network (ANN)

with Back Propagation, Support Vector Machine (SVM) classifiers. The last two learning algorithms were selected for their popularity and their great ability for generalization.

A 3-layered feed-forward back-propagation ANN was used. The parameters of used for the ANN are shown in Table 7. The selection of these parameters was based on [Wu, Jang, Kwak, Pu 2008]. The setting of the parameters for the ANN was the same for all the experiments. The inputs of ANN were the stroke features listed in Table 3. Three models were built: one for arousal with two output nodes corresponding to low and high level states; one for valence with two output nodes corresponding to positive and negative states; and one for the discrimination of the 4 emotional states with hence 4 output nodes. The ANN classifiers were implemented in MATLAB using the Neural Network toolbox.

Linear SVM is a very powerful and popular classifier and has achieved excellent performance for pattern recognition in many applications. It was used with both linear (default configuration) and non-linear kernel. The non-linear kernel function Radial Basis Function (RBF) [Plass-Oude Bos et al. 2010] was used:

$$K\big(x_i, x_j\big) = e^{-\gamma|x_i - x_j|^2} \qquad (2)$$

where γ is the parameter of the RBF kernel function. LIBSVM toolbox in MATLAB was used in the implementation. The tuning of the parameters of the kernel SVM was based on cross-validation result. Since SVM is a binary classifier, 4 SVM classifiers were trained, one for each emotion category: Excited, Relaxed, Frustrated, Bored. A winner-take-all approach was used for the multi-class classification, i.e., deciding the winner between the four emotions. Two binary SVM classifiers were also built for arousal and valence dimensions.

Classes	Features	Classifiers	Accuracy
4 Emotion labels	12 features listed in Table 3	DA	75.3%
		ANN	72.7%
		Linear SVM	**77.0%**
		SVM_kernel	69%
2 levels of arousal	12 features listed in Table 3	DA	89.3%
		ANN	88.3%
		Linear SVM	**89.7%**
		SVM_kernel	86.7%
2 levels of valence	Pressure, length and speed features	DA	83.0%
		ANN	84.7%
		Linear SVM	82.3%
		SVM_kernel	**86%**

Table 8. 15-fold cross-validation results. Best results in bold.

Valence: Length feat. **Arousal:** All features		**Predicted Label**			
		Excited	Relaxed	Frustr.	Bored
Actual Label (count)	Excited	**74**	9	6	1
	Relaxed	7	**47**	1	15
	Frustrated	5	0	**60**	5
	Bored	2	11	7	**50**
Actual Label (%)	Excited	**82.2**	10.0	6.7	1.1
	Relaxed	10.0	**67.1**	1.4	21.4
	Frustrated	7.1	0	**85.7**	7.1
	Bored	2.9	15.7	10.0	**71.4**

Table 9. Confusion matrix for 4-Label Classification (Excited, Relaxed, Frustrated, Bored) by linear SVM classifier.

5.2 Modeling

The 12 selected features listed in Table 3 were used as input feature to the classifiers. For 4-emotion classification and Arousal classification, the best results were obtained by using all these 12 features. For Valence classification, the best results were obtained by using only stroke length, pressure and speed features. Prior to the training of the models, all feature values were normalised to the interval [0,1] by using maximum and minimum values observed in the database.

Leave one person out 15-folds cross-validation was used to test the built models. Table 8 shows the average results for each algorithm. The models for Arousal produced the best correct recognition results ranging from 86.7% (kernel SVM) to 88.7% (linear SVM) to 86.7%). In the case of Valence, the best performances were obtained with the kernel SVM (86%) and the lowest with the linear SVM (82.7%). The results for the discrimination of the 4 emotion labels were also well above chance level with the best results obtained with the linear SVM (77%) and the lowest with kernel SVM (69%).

The confusion matrix for the 4-emotion classification with linear SVM is shown in Table 9. We observe the same pattern

observed in Table 4 for non-person-independent models. The best results are obtained for the classification of both Frustrated and Excited strokes. In each case more than 80% of the stroke samples were correctly classified. The lowest performances (67%) are obtained for Relaxed samples. Most misclassifications occur between Relaxed and Bored (37.1%) and between Relaxed and Excited (20%).

The results for person-independent model obtained with the DA algorithm (75%) are comparable to the results shown in Table 4 for non-person-independent models (77%). This indicates that the selected features have high generalization capability independently of people differences
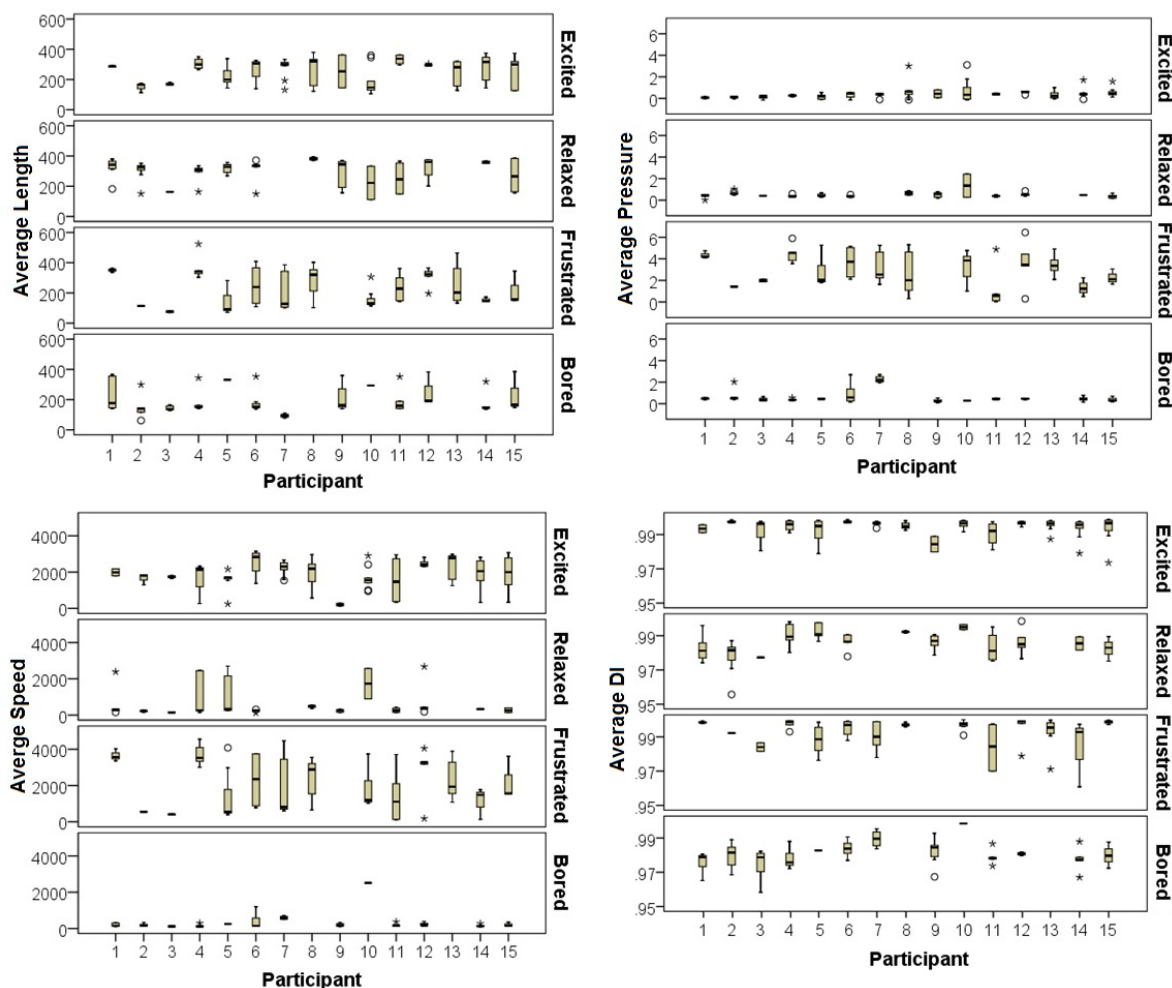

Figure 7. Boxplots showing the variation in average length, speed, pressure and DI values for the 15 participants over the 20 sessions.

6.   DISCUSSION AND CONCLUSIONS

The paper presented a non-obstructive approach to the detection of player's emotional states during gameplay with touch-based devices. The emotion recognition is based on the touch behavior (stroke) of the player. Length, pressure, direction and speed of finger strokes are used for this purpose. A visual analysis of the stroke features as well as the results of Discriminant Analysis showed that the length and the pressure of the stroke are particularly informative in discriminating between two levels of valence (positive vs. negative states) whereas the speed and the direction of the stroke strongly inform the discrimination of levels of arousal. Pressure also supported such discrimination.

Both person-dependent model and person-independent models were created. They showed similar correct recognition performances and similar misclassification patterns. In the case of person-independent model (most challenging one), arousal classification reached 89% of correct recognition rate using a linear SVM model and all selected features. In the

case of valence, the best performances were obtained with the kernel SVM algorithm and were slightly lower (86%). For the four emotion recognition case, the best performances (77%) were obtained with the linear SVM algorithm.

The results are comparable to the results on acted and naturalistic body movements (e.g., [Camurri et al. 2003, 2004; Kapur et al. 2005; Bernhardt et al. 2007; Kleinmsith et al. 2011; Savva et al. 2012]). Like for body movement, the best results are obtained for the arousal dimensions [Kleinmsith et al. 2011]. Differently from studies in body movement, the discrimination between Frustration and Excitement appear to be strongly facilitated by the use of the pressure feature. This is supported also by studies on typing and dotting [Khanna et al. 2010; LV et al. 2008; Matsuda et al. 2010]. These results support and extend previous studies by showing that touch behavior is indeed an interesting modality for automatic emotion recognition also in naturalistic setting when touch-based game devices are used.
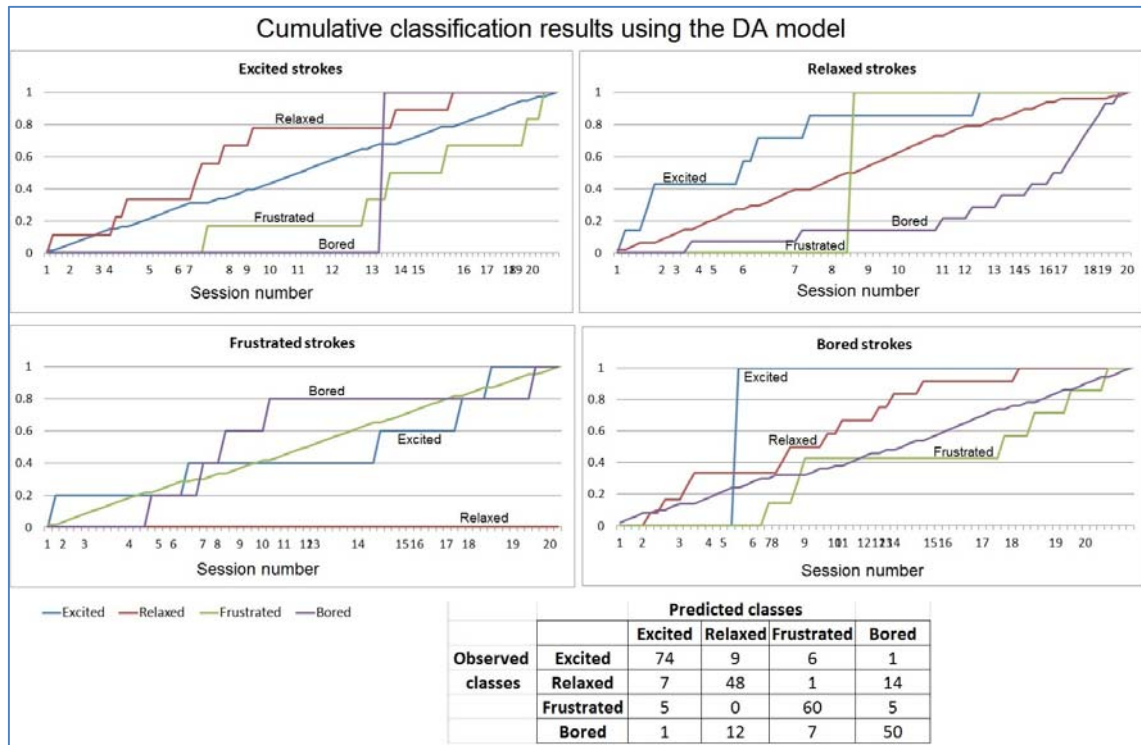

Figure 8. Classification results for all 300 stroke sample data according to the DA model of section 4. The results are shown in the confusion matrix at the bottom of the figure. In each panel, the labelled lines correspond to the cumulative percentage (vertical axis) of strokes that were misclassified with that label given the session number (horizontal axis). The unlabelled line denotes the cumulative percentage (vertical axis) of strokes that were correctly classified.

Improvement to the model could be obtained by considering the variability between players and the difficulty of the game. Finger dexterity as well as touch behaviour idiosyncrasies may affect people stroke patterns independently of their emotional state. The study reported in [Bernhardt et al. 2007] and discussed in section 2 showed that by removing idiosyncratic movements from knocking patterns, automatic emotion recognition rates improved significantly (from 50% to 80%). In [Khanna et al. 2010], it was shown that by considering daily variability in people's strength, the recognition performance could be improved. We could hence expect that personalized models may reach higher performance. However, differently from [Khanna et al. 2010; Bernhardt et al. 2007], our recognition performances are already quite high without taking into account such differences. The person-independent model did in fact reach performances similar to the person-dependent model.

Looking at Figure 7 we can also see that the variability is not only between participants but also and mainly within participants (long boxplots). This variability could be explained by the use of different strategies at different levels of difficulty of the game. In our game, the player could for example decide to slash many fruits at once, generating a less direct stroke, or using shorter and faster strokes to slash one fruit at the time. These strategies could be person-dependent but also depends on the level of difficulty of the game. Figure 3 showed that the four emotional states were triggered at any level of the game. However, the figure showed that certain states are more frequent in different level of difficulty of the game. For example, the number of boredom states was higher towards the end of the game, whereas, relaxed states increased after the first sessions and decreased again as the game got

faster.

To investigate a possible effect of game difficulty on stroke misclassification, we further analysed the results of the DA person-dependent model obtained in section 4. The classification results are provided in Figure 8. The results are shown in the confusion matrix at the bottom of the figure. Each graph shows the results for the set of strokes belonging to the class indicated in the graph title. For each graph, the labelled lines correspond to the cumulative percentage (vertical axis) of strokes that were misclassified with that label given the session number (horizontal axis). The unlabelled line denotes the cumulative percentage (vertical axis) of strokes that were correctly classified. The top two graphs show that the misclassifications between Excited and Relaxed strokes reach 80% within the first 10 game sessions. The graphs for Relaxed and Bored strokes show instead that the misclassification between these two classes occurs in the second part of the 20 session game. Given that Figure 5 shows that length and speed in particular decrease as the game become faster, the normalization of these features according to the speed of the game may remove some of these errors. Similar patterns can be observed between Frustrated and Excited and between Frustrated and Bored indicating that similar normalization for pressure and DI could also improve the results.

In this paper we have investigated the possibility of using touch behaviour to capture the affective states of the player in a touch-based game device. The results are comparable with automatic emotion recognition studies in acted key-pressing and dotting Braille devices and also with results on emotion recognition in interpersonal touch-based communication. The positive results shows that this modality has the potential to be used both for user-experience evaluation study in game context but also for on-line adaptation of the game play. Both the level of the game as well as the gameplay and the functions available to the player could be adjusted accordingly to maintain the engagement of the player. In fact, an emerging trends in computer games is the creation of algorithms that generate the content of the game at run time by taking into account the affective and cognitive experience that the player is going through [Yannakakis and Togelius, 2011; Liapis, Yannakakis, Togelius, 2012] with the goal to optimize it. Most of these generative algorithms infer the players' affective and cognitive states on the basis of choices the players make and their mistakes. More recently there has been an interest in facilitating the modelling of the player by also monitoring the player physiological signals [Yannakakis, G.N., and Hallam, J. 2007; Yannakakis, Martınez, Jhala 2010] and also other non-verbal modalities [Savva et al. 2012] or from explicit player reports when computational power is an issue. The use of touch as affective communicative channels would be an interesting modality to apply such generative algorithm to touch-display game device where facial expression and body movement recognition or bio-signal detection may not be feasible.

As the use of touch-based devices (e.g., ipad, tabletop) is spreading to different contexts, these results show also the potential for using this non-intrusive affective measure in other emotion-sensitive contexts. Not only the entertainment industry in general could benefit by such approach (e.g., mood-aware music player), also other areas where emotion and mood play a critical role could benefit by offering support to emotion enhancement, awareness and regulation. The psychology literature (e.g., [Clore and Palmer, 2009]) has shown that emotional states strongly affect our cognitive processes (e.g., by enhancing or inhibiting memory process). Of particular interest is the area of education where the emotional states' of students and their capacity to regulate their states are strongly related to the students' motivations, learning capabilities, and performances, as well as health [Schutz, Pekrun 2007]. By detecting the emotional state of the students during self-directed studies, the device could provide some form of support generally provided by a teacher as for example tailoring the teaching material and the pace to the students' needs [Conati, C., Maclaren, H. 2009] or offering empathic and encouraging behaviour to help an emotional shift. Studies in this area have shown that the ability of the teaching application to provide such support can indeed show positive results [Maldonado, H., Lee, J.R., Brave, S., Nass, C., Nakajima, H., Yamada, R., Iwamura, K., Morishima, Y., 2005]. However, most of these studies are based on facial expression detection or self-report and hence have their limitations. Facial expressions recognition requires high computational power to provide reliable results and it is quite intrusive as discussed in the introduction. It is also strongly affected by illumination conditions and by the visibility that the camera has on the face of the student. Whereas in the case of desktop conditions this may be easier to address, the situation is quite different with touch-based devices such as iPad or iPhone. Having to self-report on the other side is often seen as a intrusive and frustrating activity that often people try to skip.

Automatic recognition of emotional states trough touch could offer a non-obtrusive channel to facilitate awareness and hence reflection on emotional state in rehabilitation [Kapoor Burleson, Picard 2007]. In mental health rehabilitation, patients are often asked to maintain diary of their emotional experience. Various electronic palm-top diaries or diary applications for smart phone are now commercially available [XXX]. Unfortunately, it is often the case that patients are not aware of them [XXX]. Software that is aware of the patients' state and able to on-line adapt the rehabilitation therapy by taking into account the recognize state (provide high accuracy) appear to be welcome by patients [Swann-Sternberg, Singh, Bianchi-Berthouze, CdeC Williams 2012] as society become more at easy with such technology. On the other side, reflection may be also facilitated by such devices by simply presenting cues on their affective states and let the patients reflect on their meaning [Lindstrom, Stahl, Hook, K., Sundstrom, Laaksolahti, Combetto, Taylor, Bresin 2006].

Whereas our results show this potential, the modeling of the relationship between touch behaviour and emotional states may in part be affected by the context of use and by the activity performed. It is hence necessary to carry out further studies in a variety of contexts to better understand this relationship. It would also be interesting to investigate the possibility to integrate different

touch with other modalities when the use of multiple modality is a possibility. This could provide even more reliable results as shown in recent literature on the integration of facial, vocal and body movement expressions (e.g. [Gune and Piccardi 2006]).

REFERENCES

BAILENSON, N.J., BRAVE, N.Y.S, MERGET, D., and KOSLOW, D. 2007. Virtual interpersonal touch: expressing and recognizing emotions through haptic devices. Human-Computer Interaction, 22, 325–353.

BALAAM, M., FITZPATRICK., G., GOOD., J., and LUCKIN, R. 2010. Exploring Affective Technologies for the Classroom with the Subtle Stone", Proceeding of the Twenty-Eighth Annual SIGCHI Conference on Human Factors in Computing Systems, 1623 – 1632

BENTLEY, T., JOHNSTON, L., and VON BAGGO, K. 2005. Evaluation using cued-recall debrief to elicit information about a user's affective experiences. Proceedings of the 17th Australia conference on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future, 1–10.

BERNHARDT, D., and ROBINSON, P. 2007. Detecting affect from non-stylised body motions. Proceedings of International Conference on Affective Computing and Intelligent Interaction.

BIANCHI-BERTHOUZE, N. In press. Understanding the role of body movement in player engagement. Human-Computer Interaction.

BONANNI, L., VAUCELLE, C., LIEBERMAN, J., and ZUCKERMAN, O. 2006. Taptap: a haptic wearable for asynchronous distributed touch therapy", Proceedings of CHI 2006, 580–585.

CALVO, R.A., and D'MELLO, S. 2010. Affect Detection: An Interdisciplinary Review of Models, Methods, and their Applications. IEEE Trans. on Affective Computing, 1, 1, 18-37.

CAMURRI, A., LAGERLOF, I and VOLPE, G. 2003. Recognizing emotion from dance movement: Comparison of spectator recognition and automated techniques. International Journal of Human-Computer Studies, 59, 1-2, 213-225.

CAMURRI, A., MAZZARINO, B., RICCHETTI, M., TIMMERS, R., and VOLPE, G. 2004. Multimodal analysis of expressive gesture in music and dance performances. Gesture-based Communication in Human-Computer Interaction, 20-39.

CAMURRI, A., MAZZARINO, B., and VOLPE, G. 2003. Analysis of Expressive Gesture: The EyesWeb Expressive Gesture Processing Library. Gesture Workshop. 460-467

CHANG, A. O'MODHRAIN, S., JACOB, R., GUNTHER, E., and ISHII, H. 2002. ComTouch: Design of a vibrotactile communication device. Design of Interactive Systems Conference. London.

CLORE, G.L., and PALMER, J. 2009. Affective guidance of intelligent agents: How emotion controls cognition, Cognitive Systems Research, 10, 1, Modeling the Cognitive Antecedents and Consequences of Emotion, 21-30

CLYNES, M. 1973. Sentography: Dynamic Forms of Communication of Emotion and Qualities. Conputers in Biology and Medicine, Pergamon Press, 3, 119 -130

CONATI, C., and MACLAREN, H. 2009. Empirically building and evaluating a probabilistic model of user affect. User Modeling and User-Adapted Interaction, 19, 3, 267-303.

CSIKSZENTMIHALYI, M. 1990. Flow. New York: Harper Collins Publishers, 1990.

ELFENBEIN, H.A., and AMBADY, N. 2002. Is there an ingroup advantage in emotion recognition? Psychological Bulletin, 128, 243–249.

ESSICK, G.K., MCGLONE, F., DANCER, C., FABRICANT, D., RAGIN, Y., PHILLIPS, N., JONES, T., and GUEST, S. 2010. Quantitative assessment of pleasant touch. Neuroscience Biobehavioral Reviews, 34(2), 192-203

GUNES, H., and PICCARDI, M. 2006. A Bimodal Face and Body Gesture Database for Automatic Analysis of Human Nonverbal Affective Behavior. *ICPR*, 1148-1153.

HERTENSTEIN, M. J., HOLMES, R., MCCULLOUGH, M., and KELTNER, D. 2009. The Communication of Emotion via Touch. Emotion, 9, 4, 566-573,

HERTENSTEIN, M.J., KELTNER, D., APP, B., BULLEIT, B.A., and JASKOLKA, A.R. 2006. Touch communicates distinct emotions. Emotion 6, 528–533

HO, C.C., SRINIVASAN, M.A., and SLATER, M. 2000. An experimental study on the role of touch in shared virtual environments. ACM Transactions on Computer-Human Interaction, Special issue on human-computer interaction and collaborative virtual environments, 7, 4, 443-460.

JONES, S.E., and YARBROUGH, A.E. 1985. A naturalistic study of the meanings of touch. Communication Monographs, 52, 19-56.

KAPOOR, A., BURLESON, W., and PICARD, R.W. 2007. Automatic prediction of frustration. International Journal of Human Computer Studies, 65, 8, 724-736.

KAPUR, A., VIRJI-BABUL, A., TZANETAKIS, G. & DRIESSEN, P.F. 2005. Gesture-based affective computing on motion capture data. Proceedings of International Conference on Affective Computing and Intelligent Interaction, 1-7.

KHANNA, P., and SASIKUMAR, M. 2010. Recognising Emotions from Keyboard Stroke Pattern", International Journal of Computer Applications, 11, 9, 8975 – 8887

KLEINSMITH, A., and BIANCHI-BERTHOUZE, N. In review. Affective Body Expression Perception and Recognition: A Survey. IEEE Transaction on Affective Computing.

KLEINSMITH, A., BIANCHI-BERTHOUZE, N., and STEED, A. 2011. Automatic Recognition of Non-Acted Affective Postures. IEEE Trans. on Systems, Man, and Cybernetics Part B, 41, 4, 1027-1038.

KNAPP, M.L., and HALL, J.A. 1997. Nonverbal communication in human interaction (4th ed.). Harcourt Brace College Publishers.

LIAPIS, A., YANNAKAKIS, G.N., and TOGELIUS, J. 2012. Adapting Models of Visual Aesthetics for Personalized Content Creation**.** IEEE Transactions on Computational Intelligence and AI in Games, Special Issue on Computational Aesthetics in Games 2012.

LINDSTROM, M., STAHL, A., HOOK, K., SUNDSTROM, P., LAAKSOLAHTI, J., COMBETTO, M., TAYLOR, A., and BRESIN, R. 2006. Affective diary - Designing for bodily expressiveness and self-reflection. Proceedings CHI 2006, ACM Press, 1037-1042.

LV, H., LIN, Z., YIN, W., and DONG, J. 2008. Emotion recognition based on pressure sensor keyboards. IEEE International Conference on Multimedia and Expo, 1089-1092.

MATSUDA, Y., SAKUMA, I., JIMBO, Y., KOBAYASHI, E., ARAFUNE, T., and ISOMURA, T. 2010. Emotion recognition of finger Braille. International Journal of Innovative Computing, Information and Control, 6, 3(B), 1363-1377.

MALDONADO, H., LEE, J.R., BRAVE, S., NASS, C., NAKAJIMA, H., YAMADA, R., IWAMURA, K., and MORISHIMA, Y., 2005. We learn better together: enhancing eLearning with emotional characters. In: Koschmann, T., Suthers, D., Chan, T.W. (Eds.), Computer Supported Collaborative Learning 2005, 408–417.

NASS, C., and BRAVE, S. 2005. Wired for speech: How voice activates and advances the human-computer relationship. Cambridge, MA: MIT Press.

NILSENWIRE. 2011. Play Before Work: Games Most Popular Mobile App Category in US, http://blog.nielsen.com/nielsenwire/online_mobile/games-most-popular-mobile-app-category, 6 July 2001, downloaded on 10 October 2011

NIJHOLT, A., PLASS-OUDE BOS, D., and REUDERINK, B. 2009. Turning shortcomings into challenges: Brain–computer interfaces for games. Entertainment Computing, 1 (1), 85-94.

NOMA, H., and MIYASATO, T. 1997. Haptic communication for cooperative object manipulation. Proceedings of the International Workshop on New Media Technology.

OAKLEY, I., BREWSTER, S. A., and GRAY, P.D. 2000. Communicating with feeling. Proceedings of the First Workshop on Haptic Human-Computer Interaction. London: Springer Verlag.

PASCH, M., BIANCHI-BERTHOUZE, N., VAN DIJK, B., and NIJHOLT, A. 2009. Movement-based Sports Video Games: Investigating Motivation and Gaming Experience. Entertainment Computing, 9, 2, 169-180.

PLASS-OUDE BOS, D., REUDERINK, B., VAN DE LAAR, B.L.A., GÜRKÖK, H, MÜHL, C., POEL, M., HEYLEN, D.K.J., and NIJHOLT, A. 2010, Human-Computer Interaction for BCI Games: Usability and User Experience. inProceedings of the International Conference on CYBERWORLDS, A. Sourin (eds), IEEE Computer Society, Los Alamitos, 277-281, 2010

POLLICK, F.E., PATERSON, H.M., BRUDERLIN, A., and SANFORD, A.J. 2001. Perceiving affect from arm movement. Cognition, 82, 51-61.

RUSSELL, J.A. 1980. A circumplex model of affect. Journal of Personality and Social Psychology, 39, 1161-1178.

SAVVA, N., and BIANCHI-BERTHOUZE, N., 2012. Automatic recognition of affective body movement in a video game scenario. Proceedings of International Conference on Intelligent Technologies for Interactive Entertainment.

SCHUTZ, R. PEKRUN P.A. (Eds.) 2007. Emotion in education. Academic Press, San Diego, CA, 2007

SHIBATA, T., MITSUI, T., WADA, K., TOUDA, A., KUMASAKA, T., TAGAMI, K., and TANIE, K. 2001. Mental commit robot and its application to therapy of children. Proceedings of IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 2, 1053-1058.

STIEHL, W.D., LIEBERMAN, J., BREAZEAL, C., BASEL, L., LALLA, L., and WOLF, M. 2005. Design of a therapeutic robotic companion for relational, affective touch. IEEE International Workshop on Robot and Human Interactive Communication, 408-415

STRONG, R., and GAVER, B. 1996. Feather, scent and shaker: Supporting simple intimacy. Videos, Demonstrations, and Short Papers of CSCW'96: Conference on Computer Supported Cooperative Work. New York: ACM.

SWANN-STERNSBERG, T., SINGH, A., BIANCHI-BERTHOUZE, N., CDEC WILLIAMS, A. 2012. User needs for an interactive technology to support physical activity in chronic pain. Extended Abstracts of CHI'12.

UVANAS-MOBERG, K., ARN, I., and MAGNUSSON, D. 2005. The psychobiology of emotion: the role of the oxytocinergic system. International Journal of Behavioural Medicine, 12, 59–65.

VON LABAN, R. 1971. The mastery of movement. MacDonald & Evans Ltd.

WALLBOTT H.G., and SCHERER, K.R. 1986. Cues and channels in emotion recognition. J. Pers. and Social Psych., 51, 4, 690-699.

WU, G., JANG, S., KWAK, H., and PU, J. 2008. A New Parameter Selection Method of Neural Network", International Conference on Multimedia and Ubiquitous Engineering, 378 – 383.

Yannakakis, G. N. and Togelius, J. 2011. Experience-driven Procedural Content Generation, IEEE Transactions on Affective Computing, 2, 3, 147-161.

Yannakakis, G.N., and Hallam, J. 2007. Towards optimizing entertainment in computer games. Applied Artificial Intelligence, vol. 21, pp. 933–971.

Yannakakis, G.N., Martınez, H.P., and Jhala, A. 2010. Towards affective camera control in games. User Modeling and User-Adapted Interaction, 20, 313–340.

Author Statement:
None of the material presented in this paper was previously published.