Painting into Music

An interactive multimedia edutainment tool for painting

Adam Jareh, Tom Parish, Cansin Rathge, Elina Vasina, and Kia Ng
ICSRI M – University of Leeds
School of Electronic and Electrical Engineering and School of Music
Leeds LS2 9JT, UK
mme2011@icsrim.org.uk, www.icsrim.org.uk

Abstract—Interactivity is an integral part of engaging children with what they are learning. We address this engagement by creating an interactive entertainment and learning tool for children to use in the classroom or at home using a painting scenario. We do this by coalescing music, multimedia and electronics together through the real-time mapping of painting gestures into musical output and feedback. The system analyses and objectifies aspects of the painting process, measuring painting gestures with multiple sensors and using the data to create a meaningful musical feedback that is controlled by the actions of the painter on the interactive “canvas”.

The system integrates various software modules, sensor inputs and hardware systems. The paper describes several different mapping strategies, from simple one to one relationships, to complex webs of interaction between gestural data and sonic processing and controls. Input features include brush stroke analysis such as speed and position, shape and angles of drawn items. They can be configured and mapped to offer different mapping strategies, focusing on varied characteristics of the painting gesture, depending on the application context.

Keywords—multimedia; interactive; painting; mapping; sensor

I. INTRODUCTION

The bridge between artistic expression and mathematically mapped music is perhaps a difficult one to build as they are different forms of expression – visual and sonic. This project aims to explore the relationship between the two expressive domains by systematically analysing several different aspects of painting gestures and “translating” the painting movements into a musical output whilst providing an entertaining and interactive tool for children. The idea being that the user paints onto a canvas, whilst the system measures appropriate physical motion created by the action of painting (e.g. shape of the painter’s strokes, speed of strokes, paint colour, etc.). Gestural data is then captured and transformed into multimedia feedback. The mapping strategy follows a set of rules to recreate similar sounds when similar painting actions are performed that can convey some insights into the strokes that are not easy to determine by eye, such as, the consistency and variations of stroke speed.

The system offers two different ways of interaction for the user. With this system, children are able to listen to their painting in addition to just looking. They can either paint as usual and listen to the music produced, or paint strategically to try and create a certain style of music.

Significant control over the system is given to the user. For example, a teacher would be able to configure appropriate mapping (or select one from a preset library) for the student in accordance to the exercise, and the user/student has the ability to choose colours and mix them in a realistic paint-based fashion.

The paper discusses related background, presents the design and development of the system and focuses on the mapping of various geometric features of paint strokes that are used to affect the musical output.

II. RELATED RESEARCH

Significant developments within trans-domain mapping and sonification of movement have been made in the last decade with a lot of research and industrial applications [1]. The sonification of painting gestures for technology-enhanced learning has not been fully explored. We have adopted and modified methods and existing research for our system from the following related systems.

Knoerig et al. [2] examined the conducting gesture, paying particular attention to the relationship between gestural expression within painting and the playing of a musical instrument. By using a physical brush on a canvas that was linked to a conductor program, they were able to show the importance of physical movement in relation to music. The focus was on a metaphorical visual representation of sound and the opportunity to learn a musical instrument without the need to comprehend the complexities of an actual instrument. This idea became the focus of the early stages of this system.

Rozin [3] proposes a method of capturing the painting process into a digital format. He used rear-projection with an easel that gradually revealed a video stream as the user painted onto the canvas, with movement of the “paint brush” being
captured with the use of video cameras. Rozin’s setup is adapted for our project.

Motoglyph [4] used similar methods of input, however, unlike Rozin’s project, it provided the user with a musical output. Pseudo-spray cans were implemented to graffiti three glass panels, each of which had its own unique sample library. The graffiti created on the panels were measured and mapped in order to generate ringtones. Motoglyph used ultrasound as a tracking method.

The initial idea for our system was to bridge painting and music for children in an edutainment application context. [5] described a study on infants’ perception of consonance and dissonance which shows the importance of the musical output. It examined the “universal” perceptions of music and the reactions of infants to two different versions of a melody, one consonant and one dissonant. The article concluded that infants looked significantly longer at a source of sound that played a consonant passage, compared to a dissonant passage, where they tended to turn away and fret. The overall results suggested that infants are biologically inclined to treat consonance as perceptually more pleasing than dissonance. Whilst this project is aimed at slightly older children, who will be more susceptible to a consonant output, thus suggesting that it would be best to avoid using dissonant or unpleasant tones as the main output of our system [5]. However, the usefulness of dissonant outputs cannot be ignored, therefore, they will be implemented within the system, e.g. as a way of “rewarding” the user for consistency.

When considering the musical output, ‘In B Flat’ [6] gives an interesting example of how numerous sound sources can be combined to sound pleasing and avoid dissonance. ‘In B Flat’ uses a large number of pre-recorded videos with music played in b flat on a wide variety of instruments. The instruments can be played in any order, at any time, and sound coherent and musical. This is one of the mapping strategies that we adapt for this project, i.e. by using a set of samples in the same key which will ensure a pleasant output when played simultaneously.

III. DESIGN AND DEVELOPMENT

A. Requirements

User requirements were sought from a group of expert users consisting of 5 artists and 5 musicians. A formal survey was carried out and a summary of key points is provided. The artists believed that receiving any feedback from the process of painting would be more engaging, as it could provide a real-time feedback feature to painting that is not normally associated with the process. They were also interested in the final piece of music being composed by an artist; this was seen as a potential bridge between those art forms. The musicians were more prone to believe that a musical output could be a distraction whilst in the process of painting. However, there was interest in the detailed aspects of composing a piece of music using painting techniques and the idea of using the process of painting as a musical instrument. Overall, both groups were positive in regards to the interaction between the two art forms and would be personally interested in using or testing out the final device. We used this information and setup our program in a strategic way so that a musical piece could be easily composed using artistic gestures. This was achieved by creating a mathematical grid across the canvas. The frequency of the sample changes dependant upon the area in which the user paints with the idea that this visual representation of the notes within a scale would be interactive and logical.

The musicians were also asked about their perceptions of colour in accordance to the musical output, with a general consensus of warm shades (reds, yellows, etc.) corresponding to major, happier and more energetic sounds and instruments, and cool shades (blues, greens, etc.) to minor and mellow sounds. We used this information to setup our samples in accordance to these results. Warmer shades of colours used in the system trigger happier and more energetic sounds where as the cooler shades trigger mellow samples and ambience.

Another survey was carried out to provide some understanding of the techniques used by children when painting. 60 preschoolers were asked to paint on a piece of A3 paper whilst monitored by a teacher. Different parameters were then noted. It was found that in regards to the initial placement of the drawing (i.e. where the child starting drawing) a larger percentage of children began at the bottom of the page. Using this information, we positioned the first note of the scale at the bottom of the canvas with the following notes increasing higher up the canvas.

In terms of initial colour selection, yellow, blue and green were the most popular colours. A tally was also made for the number of times the children dipped their paint brushes into specific colour, with the results showing that yellow was the colour used most frequently and for the longest periods of time. This information was used in such a way that the more popular colours trigger the more common instruments/samples. It also triggers the more simple samples so that the piece is less likely to sound too complex.

The averaged time spent painting (continuous) was 5 minutes. Comments on techniques used were also provided by the teachers, showing that singular long lines and dabbing were used most frequently. The audio effects controller and mapping was designed to ensure that these more common artistic gestures would create common and easily noticeable effects; delay and filtering. This meant that the child may be more aware that their gestures are producing the relative effects.

B. System Overview

The end product of the design will closely replicate an artist’s actual utensils and environment (see Figure 1). This is done to ensure that the end user has a familiar experience when compared with using real paint.
The design uses a wooden easel, measured to be comfortable for the average height of a 7 year old (122cm [7]) which is also usable for a seated adult. A projector is mounted onto the back of this easel, pointing at a perpendicular angle toward a sheet of smoked Perspex which is used as the screen. Using rear projection the user is presented, on screen, with the interactive “canvas” to draw on (see Figure 2). The draw tool is a modified paint brush containing sensors that enable it to interact with the canvas and software. Using a Nintendo Wiimote and an infra red LED embedded on a paint brush, the canvas is turned into a “touch screen” [8]. Colours are selected via the graphical user interface and then placed into “pots” which act as real paint pots, in that, the user dips the paintbrush in to cover it with the colour selected. The main software is developed using Max/MSP (Max hereafter) [9], which controls all the mappings, processing, analysis and feedback (visual and audio feedback), see Figure 3. The program can end at any time by selecting the ‘Finish’ option. Colours can be selected and if necessary mixed dependant on the number of colours chosen. If clear is selected/activated the colour choices are deactivated. If it has not been chosen the colour or colours selected will remain stored. If there is a load on the bend sensor the relevant colour image will appear on the screen and the relevant musical output is produced dependant of the point based and geometric based mapping. This is done through an audio effects patch that is manipulated relative to the output values from the mapping. These audio parameters are also effected from the shape-based, similarity-based and feature-based mapping. The visual data is sent through a projector that projects the image onto the screen. The audio data is sent to an audio output. The sensors communicate with Max using an Arduino board [10] (see Figure 4).

C. Inputs

Physical movement of the paint brush provide the key inputs to the system. A bend sensor is embedded inside the brush to measure the “pen-down” and “pen-up” action which triggers the paint and the amount of the paint, as in, controlling the thickness of the stroke depending on applied pressure.

The position of the pain brush is measure by a Wiimote tracking an infra red LED embedded on the paint brush. The Wiimote is clamped to the back of the easel and pointed toward the Perspex. After calibration, it is used to provide the coordinates of the infra red LED to a computer via Bluetooth.

When the bend sensor is triggered, the infra red LED activates. This acts as the activity trigger, which starts the position tracking and enables drawing onto the canvas. This input gives a variable value of pressure on the brush to use within mapping.

In the faux paint pots there are photocell resistors, with LEDs, that send the software a signal through Arduino when the light beam is broken. This signals that the colour picked has been added to the paint brush. The option is also present to mix the two coloured pots together to create different shades.

The position tracking of the painting gestures are derived from measuring x and y co-ordinates, when the activity trigger is on. All of these values are stored in a stream of data which we interpret with the mapping strategies and eventual output. With the 2D co-ordinates of the paint brush, its speed and acceleration are computed to objectify the speed of the paint stroke. Stroke/line length is also tracked and used as one of the key input elements to the mapping module.
D. Data analysis and mapping strategy

The mapping of the data is managed through a master patch in Max, which is further divided into several sub-patches. It has been designed in such a way so that the input data streams are setup independent of the output (audio) parameters that are to be manipulated. This provides more flexibility and allows us to manipulate various audio parameters by any visual ones. In doing so, it is easy to produce various system setups and mapping strategies to create a mixture of mapping possibilities. The mapping techniques that apply in the system fall under four categories.

The first category is a point-based mapping strategy, where the output is dependent on the acting point of the brush, i.e. wherever the brush is touching at any given time. This includes the coordinates and, in turn, a mathematical grid that can be determined as another input variable.

The second category is stroke-based mapping which includes variables including speed, acceleration and direction. These values are worked out within Max by using the 2D positional coordinates of the LED on the paint brush as they move over time.

The third category is feature-based mapping. This area focuses on textural analysis. The painting techniques that the user produces are measured in a variety of ways, i.e. the number of angles and their size within a single brush stroke are quantified. In doing so, it is determined whether circular motions or sharp ‘saw tooth’-like motions are being produced. The line distance is also measured to determine whether a longer or shorter line is produced, this will also pick up on any techniques such as “splodges” onto a page. Other textural techniques are noted, e.g. compact blocks of one colour being created or the size of a particular part of the painting. This mapping strategy is particularly important for pedagogical applications working on specific techniques.

The final category is class-based mapping, this focuses on analysing the shapes and sizes created on the canvas. Various regular shape forms such as squares or rectangles can be discerned from more obscure irregular shapes which can give musical feedback relevant to the types of shapes created. All of these variables can work both independently and also with other variables to produce data outputs that are then fed
accordingly into a separate system that contains all of the audio parameters. This link between the two systems is easily manageable and interchangeable allowing more control over the desired output.

E. An example multimedia mapping setup

The overall system can be configured to support different mapping using different input, analysis and features depending on the application context and personalised for the individual student/user. For example, the overall length of a line (or the area of a drawn object) creates reverb, sustain and release effects. The average value of the angles within a stroke affects high and low pass filtering (with a large overall angle creating a low pass filter and a smaller angle creating a high pass filter), the idea being that a larger angle creates a larger “arc” which would let through lower and “stronger” frequency range. The number of curves in a line creates either delay or vibrato effects, i.e. if a user paints a large number of sharp angles, the visual lines are representative of the consistent frequency fluctuations within vibrato effects. The angle of the line relative to x and y axis, in relation to the length of a line, creates real time reverb and echo effects. Similarly to the high and low pass filtering effects, the visual aspects of a large angle create a larger space, therefore, the idea that the sound needs to be enlarged or elongated is appropriate. The circularity of a line alters vibrato effects and dotting on the canvas creates pizzicato effects.

IV. Prototype

A. Interactive canvas

Two main components make up the interactive canvas interface. The first component is in the form of a large white box. This is the drawing environment (the canvas) that the user draws onto to create a painting. The second component is a menu bar located underneath the drawing environment. This contains a number of options for the user to choose. There are two virtual paint palettes each containing all eight of the main colours accessible to the user. Choosing a colour will store that colour into the desired paint pot. There are also two colour indicators located on the screen above each paint pot to show the user which colour is currently stored in each pot. Also, on the menu bar is a master volume controller and a restart option. Once a colour has been assigned to each paint pot, the user can select an individual colour by dipping the paintbrush into a single pot. To mix the colours together, the user must dip the brush into both of the paint pots. The third pot is a clear function that will remove the current colour stored on the brush.

B. Output

For accessibility and to avoid musical dissonance mapping (for start up initial stage), the system applies a non-genre specific output by default, although the user can of course configure it to other behaviour as required.

The main intention of the system is not to “compose” a musical score, but to keep a child engaged for the duration of their painting and to offer meaningful feedback.

Outputs involve musical instruments that have been recorded and re-sampled to comply with requirements. The output will also contain synthesised instruments that are controlled and processed based on the mapping strategies and configuration of the system.

V. Evaluation

Qualitative evaluation of the system was conducted, internally and externally. A questionnaire was designed to capture participants’ expectations for, and their experience with, the system.

A pre-usage questionnaire asked participants to rate their interest in the idea of the system, their predicted enjoyment of it and their predictions for the way they would use the system. The results showed that people were generally interested in the idea and thought they would find it enjoyable. They did not choose to attempt to predict how consonant or pleasant the output would be or whether it would accurately represent inputted painting gestures, which shows that, whilst there is understanding about what the system does, its actual output would be unclear to a first-time user until they used the system. The subjects said that they were very likely to mix colours and use a larger number of them (5-6 colours on average).

A post-usage questionnaire asked the subjects to rate their experience with the system, say how they thought they used the system and provide further comments. Questions were asked regarding how easy each tool was to use, how realistic they were and how they enhanced the experience, also, questions related to the operation of the system.

The results showed that, in terms of realism, the easel was unanimously rated as the most realistic. The paint pots, canvas and brush were rated neutrally, but, considering the additional comments, it is clear that the paint pots were a particularly distinctive feature. Whilst some noted the need to both choose a colour in the GUI and dip the brush into a paint pot, it was also noted as an interesting feature of the system. The physical tools were also, on average, rated both comfortable and easy to use, as was the overall system.

In terms of the comparison of the two surveys, participants found the system more interesting and more enjoyable than they had anticipated. They also found the output to be both more pleasant and more consonant than expected and to represent their visual inputs more accurately. It can also be noted that the time they thought they would spend using the system, was, on average, significantly shorter than the actual time they spent using it and, furthermore, their understanding of the time spent using the system and the actual timed result...
differed, which would suggest that the system is engaging and interesting. In terms of usage of colour, there was a clear trend for using more colours than the subjects had predicted in the pre-usage survey.

It is clear that whilst the sample group used in this questionnaire is not of a significant size, there are already some clear results emerging and this system, in its present state, has shown its potential to be interesting, engaging and realistic. It would be exceedingly beneficial to not only enlarge the sample group, but to test this system with its target user, i.e., children, and compare the results.

VI. CONCLUSION AND FUTURE DIRECTION

In this paper we present an interactive multimedia painting interface with musical feedback for technology-enhanced edutainment. It discusses the application context, requirements, a brief overview of the system design and development, and mapping strategies.

In addition to optimisation, future developments include additional sensors, such as accelerometer and pressure sensors on the paint brush to offer more detailed understanding of the painting gesture and the development of a wireless paint brush for better accessibility.

REFERENCES