International Journal of Computational Science
1992-6669 © Global Information Publisher
2007, Vol. 1, No. 1, 40-57

Mouthbrush: A Multimodal Interface for Sketching and Painting

Chi-Ho Chan [1*], Michael J. Lyons [2],

[1] Centre for Vision, Speech, and Signal Processing
School of Electronics and Physical Sciences
University of Surrey
Guildford, Surrey, GU2 7XH, UK
[2] College of Image Arts and Sciences
Ritsumeikan University
56-1 Tojiin Kitamachi, Kita-ku Kyoto 603-8577, Japan
lyons@im.ritsumei.ac.jp

Abstract. We present a novel multimodal input method permitting users to draw or paint using coordinated gestures of hand and mouth. A head-worn camera and computer vision algorithm provides a noninvasive, comfortable and low-cost user interface. Combining colour thresholding and image morphology techniques permits segmentation of the region of the open mouth. An adaptive algorithm, based on Fisher's discriminant, automatically determines colour and intensity thresholds to separate the inside of the mouth from surrounding skin regions. We have designed two new controllability measures, the circle control task and the line width control task which we use to compare the users' performance using the Mouthbrush with their performance using previously existing input devices: a table with pressure-sensitive stylus and a stylus with a finger wheel. The results of the user studies show that the Mouthbrush offers advantages over the pressure-sensitive stylus and finger wheel controller. Trials of the Mouthbrush system by professional and amateur artists show that it is ready for use in creative artistic expression.

Keywords: human computer interaction, painting, gesture interface.

[*] Corresponding Author, Email lyons@im.ritsumei.ac.jp.

1 Introduction

Drawing and painting generally involve the application of motor skills by an artist to express feelings and intentions. Broadly speaking this requires the manipulation of tools and media to transform a schema from the artist's imagination or memory to a medium such as paper, canvas, or an electronic image. Innovative drawing tools that are easy to use and control or encourage new forms of creative artistic expression are therefore highly desirable.

In most cases, paintings or drawings are created using the hands. This is partly a consequence of the fact that manual interaction allows fine motor control, however it is also related to the nature of drawing tools which have been invented so far. From this viewpoint, it is very interesting to imagine how other human motor capabilities might be useful for artistic expression. We have been interested for some time in using action of the human face, especially the mouth, for human machine interaction. This is partly inspired by the observation, a half century ago, by the neurologist Wilder Penfield and his colleagues [1] that relatively large areas of sensory and motor cortex are devoted to the control of the muscles of the face, especially the mouth and other organs of speech. The mouth is used for both verbal and non-verbal forms of communication, such as in speech, singing and facial expression. What if the mouth could be used directly, in the creation of visual art? While this may seem a strange or unusual question, there are some precedents for the use of mouth controllers in art such as the breath atomizers, which may be thought of as a simple and ancient version of the Mouthbrush interface we describe here. Indeed, some of the earliest cave paintings were made by blowing pigment through tubes onto a surface. The mouth is controlled by a different group of muscles than those which control the hands, so it is reasonable to assume that simultaneous coordinated control of both groups may be relatively facile. Such considerations inspired us to consider using coordinated gestures of mouth and hand in artistic expression, and we have designed and implemented a prototype tool, called a Mouthbrush, shown in figure 1, for the digital sketching and painting.

Our desire for a non-invasive, comfortable, and low-cost solution led us to select computer vision as a promising candidate for real-time capture of mouth gesture. We designed and implemented a simple vision algorithm to measure the area of the open mouth. This single parameter can be mapped to control various properties of the pen or brush in real-time as the artist sketches or paints digitally with a hand held stylus.

In the remainder of this article we first briefly describe related works, and then discuss our design considerations. The details of a hardware and software implementation are presented next, followed by the introduction of novel usability tasks, the circle control task and line width control task which we used to study the static and dynamic controllability of the Mouthbrush interface by normal users. Finally, we describe use of the system by artists and provide samples of artwork created using the interface.

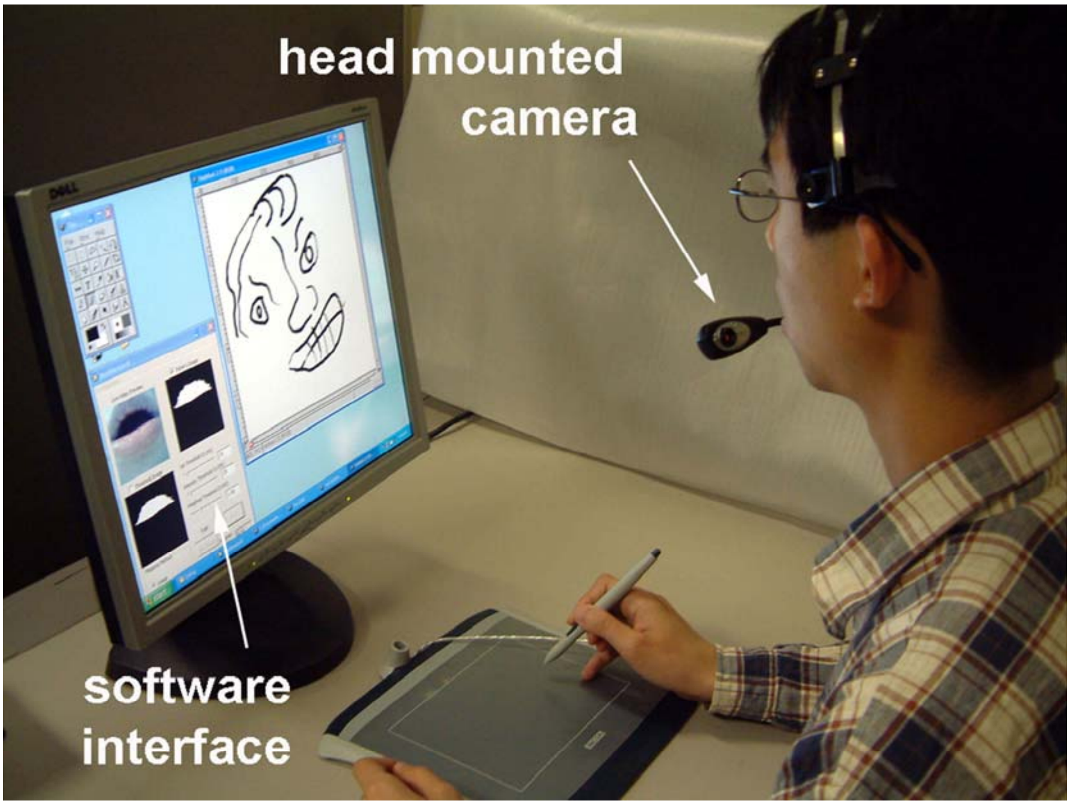


Fig. 1. Sketching with the Mouthbrush

2 Related Work

Table 1 gives a summary of recent work on novel interfaces for digital sketching and painting. Action of the mouth has not been previously been used for digital painting. However, the mouth has been used for human-machine interaction in other contexts. For example, Salem and Zhai [9] developed *Tonguepoint,* a tongue pointing device in which a small pressure sensitive isometric joystick is mounted on a mouthpiece held with the teeth. This allows the user to control a cursor with the tongue. Lyons [10] developed the *Mouthesizer,* a vision-based system for extracting shape parameters of the external appearance of the mouth cavity such as height, width, and aspect ratio and mapping these to musical control parameters. Vogt et al.[11], with their *Tongue 'n' Groove* system, used ultrasound imaging to measure the tongue position and movement in real time for real-time sound synthesis control.

Table 1. Related Work on Multimodal Interfaces for Digital Sketching and Printing

Authors	Input Device(s)	Mapping	Application(s)
Baxter et al. [2]	PHANTOM: 6 degree of freedom input, including position and orientation, with haptic feedback.	The 6 DOF are used to model the bristles of the brush, while force feedback enhances the sense of realism of the virtual brush.	Virtual oil and acrylic painting, Japanese calligraphy.
Browne et al. [3]	A multi-touch surface providing position and pressure data.	Light pressure triggers pointing, heavy pressure triggers painting.	Collaborative finger painting for children.
Keefe et al. [4]	Painting props, Digital glove.	Two handed interaction technique. The preferred hand holds a brush for paining. The non-preferred hand wears a digital glove to select colour, move and scale the drawing object and change the brush size.	Virtual 3D scene creation.
Ip et al. [5]	Infrared illumination and vision system used to capture whole body motion.	Body movement is transformed to visual attributes such as colour value, hue, saturation and brush size.	3D art creation by movement of the whole body.
Greene [6]	Vision-based system and semi-transparent prism. Allows any brightly coloured object to be used.	Objects can be used to continuously adjust line quality and texture in drawing.	Drawing table for 2D painting.
Hornof et al. [7]	Infrared eye-tracking system to capture eye movements.	Drawing mode is triggered by gaze fixation. The user can then move the brush by moving the eyes.	Computer drawing system for children with severe disabilities.
Ryokai et al. [8]	A digital stylus and tablet-display (Wacom Cintiq) with an attached camera.	The camera is used to pick up colours, textures, or frame sequences, which are applied to the digital canvas with the brush. Brush speed determines the stroke translucence.	Engaging tool which encourages sensitivity to the visual environment and playful creation of art.

3 Mouthbrush Implementation

3.1 Hardware

Figure 1 shows the setu p of the mouth controller. We use a head mounted camera to acquire an image of the mouth region. This reduces the complexity of the vision task by eliminating the need to detect and track the head. Moreover, it affords the user greater mobility and comfort than they would have if they had to face a fixed camera mounted on the desk or computer. We modified a Shure SM10A headset microphone by replacing the bean and microphone assembly with a miniature USB web camera on a flexible arm. The flexible arm allows users to easily adjust the camera position for optimal alignment with the mouth. The FlexiC@m B100 USB camera used in this work is very small: the 1/5" CMOS imaging sensor and the USB interface are built into a single circuit board having dimensions 45×18 mm. The camera is powered by the USB port, eliminating the need for an extra power cable. The simple, minimal, and lightweight design we describe here represents the outcome of a long process of experimentation with options for mounting the camera.

3.2 Software

The image processing algorithm first reflects the image about the vertical axis to allow for intuitive adjustment of camera position by the user. We use a colour and intensity thresholding and image morphological operators to convert the input image into an array of a binary variable which indicates whether a pixel belongs to the open area of the mouth (the mouth cavity) or other regions. We then compute the area of the mouth cavity simply by pixel summation. The algorithm normalizes this variable and sends it via TCP/IP to an application such as a drawing program. The system block diagram is shown in Figure 2.

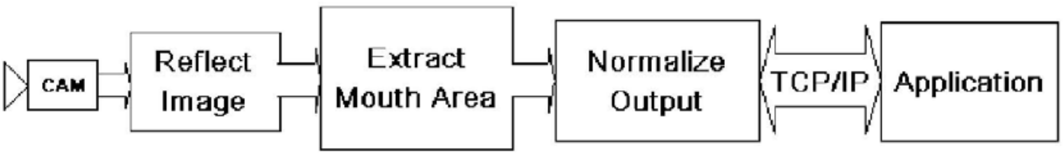


Fig. 2. Block diagram of the Mouthbrush System

The software was developed in Visual C++, under Windows XP, using DirectShow. The Mouthbrush system runs as a self-contained application under Windows XP or 2000.

Our task is not specifically to recognize or classify a facial action, but to devise a system which provides a smooth and repeatable function of mouth shape which seems natural and intuitive for the user. Hence we choose to extract the apparent area of the mouth cavity in the image using a segmentation algorithm based on colour.

Colour is known to be a useful feature for segmenting human skin, as for example in detecting the face. Our approach is to use Fisher Discriminant analysis [12] in a two dimensional colour space to determine a decision boundary separating mouth cavity and other regions of the input image. We first empirically determined that the red component *(R)* and total intensity *(I=R+G+B)* provided a representation of the input colour information suitable for detecting mouth cavity pixels. Intensity is useful because the mouth cavity appears as a shadow under usual lighting conditions. The inside of the mouth is also characterized by a relatively high red component.

The system has two modes: a training mode in which the discriminant analysis is initialized and run and output normalization facts determined, and an operating mode during which the system acts as a TCP/IP server outputting mouth cavity area in real-time.

When training mode is entered, the user positions their mouth cavity region in a central rectangular area such that skin areas occupy two flanking rectangles (see Figure 3a). Red (R) and Intensity (I) values from pixels in the two classes (cavity and non-cavity) are used to calculate a discriminant vector using Fisher analysis [12]. The position of the decision boundary between cavity and non-cavity pixels is decided using a threshold determined from the mean, μm, and standard deviation, sm, of projection of the mouth cavity *(R, I)* values onto the discriminant vector and the corresponding mean μs for non-cavity pixels.

In operating mode, *(R, I)* values of pixels in the input image are classified as cavity or non-cavity using the Fisher discriminant classifier. Each pixel in the image is then labeled as cavity or non-cavity. The labeled mouth cavity image is shown in Figure 3b.

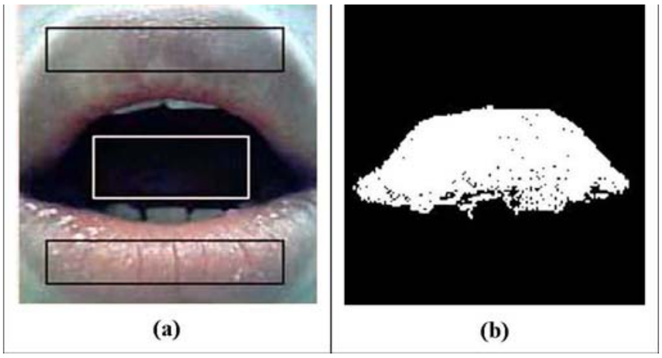


Fig. 3. a) training image and b) segmented image

We calculate the area of the mouth cavity using the total number of pixels, *M*, in the labelled image. Moderate translations or rotations of the camera leave this parameter unchanged, making it insensitive to small movements or vibration of the camera. The algorithm normalizes the area according to the maximum and minimum mouth cavity areas, *Mmax* and *Mmin:*

$$A = \begin{cases} 0 & M < M_{min} \\ \dfrac{M - M_{min}}{M_{\max} - M_{min}} & \\ 1 & M_{max} < M \end{cases} \tag{1}$$

So that the output parameter A, lies between 0 and 1. The user inputs maximum and minimum mouth positions during training mode.

The normalized area, *A*, is mapped to a pen parameter, *P*, which can be used to control size, hardness, opacity, colour etc⋯

$$P/P_{\max} = gA \tag{2}$$

where *Pmax* is the maximum value of that pen parameter. (*P/Pmax*), A, and the gain, *g*, are therefore dimensionless parameters having values between 0 and 1. With a larger value for *g*, the user need not open the mouth as much to input a given pen parameter, *P* and the system should require less effort to operate. However, if g is too large we expect it will be difficult to control P accurately because involuntary movements of the mouth will be amplified.

4 Mouthbrush Controllability

Because humans use their mouths in tasks requiring complex motor control, such as speech, it seems reasonable to suggest that it should be possible for us to control the open area of the mouth with a fair degree of accuracy. Nevertheless, we considered it important to verify this experimentally. Therefore we designed two new objective evaluation tasks to measure the degree of controllability afforded with the Mouthbrush system. The first task, the "circle control task", evaluates the ability of a user to hold the mouth area statically at a fixed level for a few seconds. We compared several users' performance on this task using the Mouthbrush controller with their performance when using a tablet and pressure-sensitive stylus. With the second evaluation task, called the "line width control task", we measure how well the user can control the width of a line while drawing. Here we compared users' performance using the Mouthbrush with their performance using the tablet and pressure-sensitive stylus and also a stylus with a finger-wheel controller. In all of the studies, we used a Wacom Intuos 12"x18" tablet, stylus having a 10-bit range of pressure sensitivity, and a Wacom airbrush stylus with finger-wheel controller. Five volunteer subjects participated in two experiments. Three of the subjects had some prior experience using the Mouthbrush controller and all of them had some prior experience using a Wacom tablet and stylus. The subjects sat comfortably at a distance of approximately 60 cm from a 19" LCD with a resolution of 1280x1024 pixels. Our subjects completed the circle control task first, followed by the line width control task. All subjects could take a break anytime between individual trials of the task.

4.1 Circle Control Task

First, we compare the mouth controller and the tablet and grip pen having a pressure sensitive tip. The subject adjusts the radius of a circle shown on the computer display by changing the open area of the mouth or by controlling the tip pressure of the stylus until the circle is exactly tangent to a target square. When the target size is matched, the subject clicks the space bar and attempts to hold the same position for 3 seconds, whence the next trial is prompted.

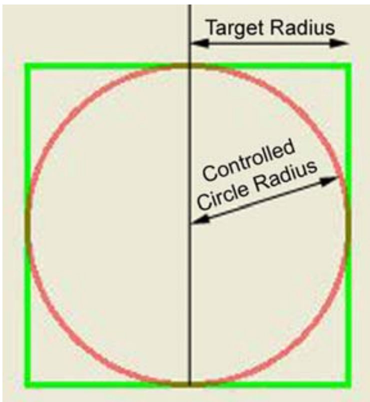


Fig. 4. Circle Control Task

The independent variables for this task are the gain, g, which can take values g = (0.25, 0.5, 0.8) and the size of the target circle, with radius = (5, 23 42, 61 and 80). The maximum of the input circle radius is P_{max} = 100 pixels. We repeated each combination of gain and target size with test devices three times for a total of 90 trials. The order of presentation of trials was randomized for each participant. The experiment took about 20 minutes to complete.

For each trial, pressing the space bar causes the computer to record the radius of the controlled circle at 67 Hz for the subsequent 3 seconds. We calculated the standard deviation, , from the mean position as a measure of precision of user control and the mean absolute error, *E*, (difference between target and actual radius) as a measure of the accuracy of control.

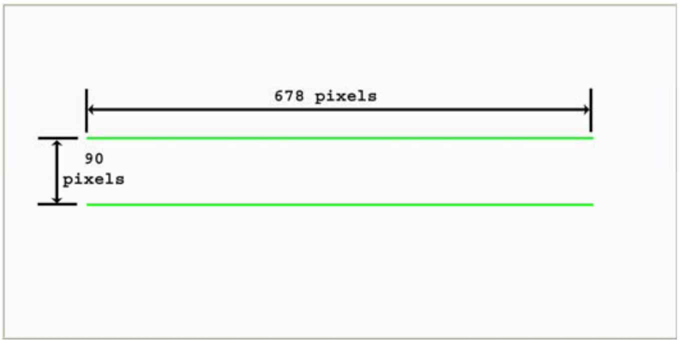
$$E(x,y)=\sum_{i=0}^{n}\left|C(x,y)_i-T(x,y)\right| \Big/ n \tag{3}$$

$$\sigma(x,y)=\sum_{i=0}^{n}\left(C(x,y)_i-\overline{C}(x,y)\right)^2 \Big/ n(n-1) \tag{4}$$

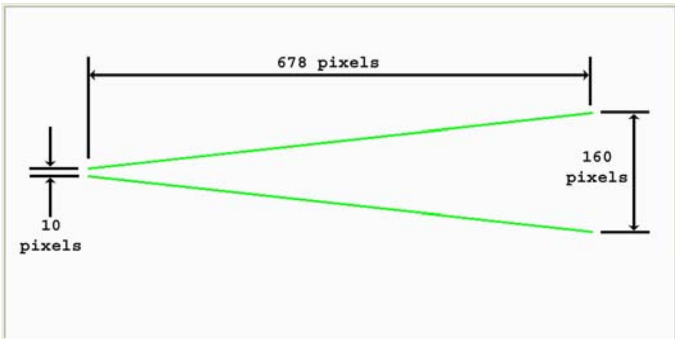
where x and y are indices of gain and target parameters, n is the total sample number in each target test. T is the target radius and *C* is the input circle radius. *C* is the mean of input circle radius.

4.2 Line Width Task

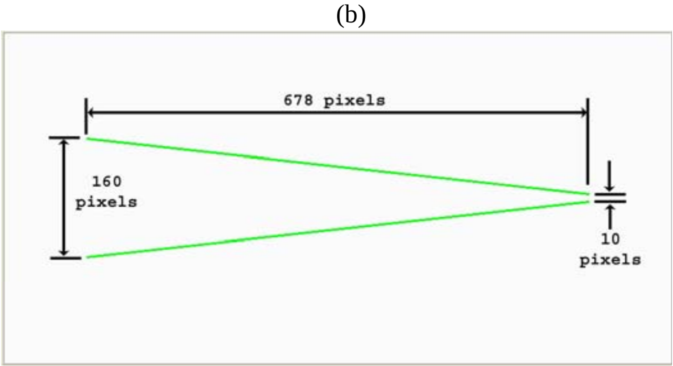
We designed this task in order to measure the users' degree of control while drawing a line of constant or variable width. At the beginning of each trial, the user is presented with a target consisted of a channel made of two green lines, which are either parallel, diverging, or converging, (see Figure 5). Participants used the tested input device to adjust the size of an I-shaped cursor to match the green guidelines while tracing the channel from left to right. After reaching the right-most extreme of the target channel, participants press the space bar to proceed to the next trial. The participants are free to take a break between trials.



(a)



(b)



(c)

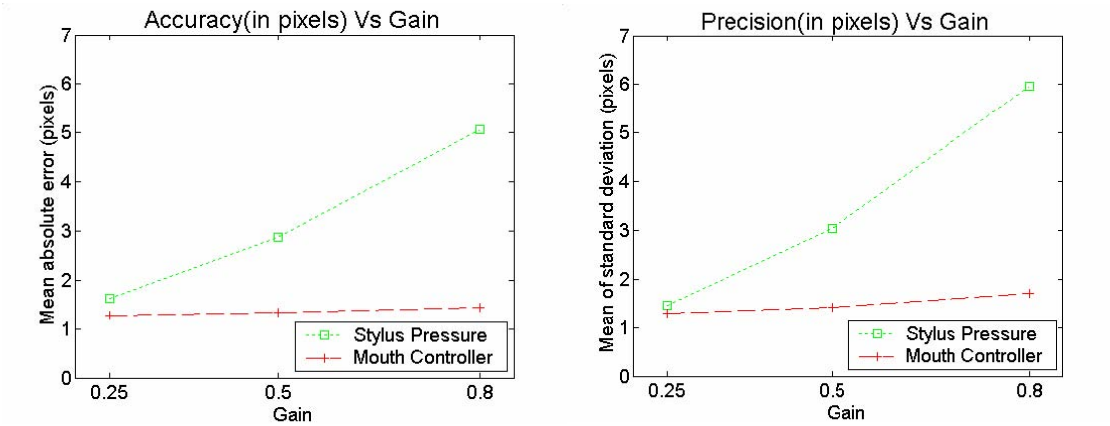
Fig. 5. Line width control task: (a) straight, (b) divergent and (c) Convergent targets

The independent variables in this experiment are the channel type (T = straight, converging, diverging) and path width (W = 10, 90, 160 pixels). Fixed parameters are: the gain (g = 0.25), path length (L = 678 pixels), and the maximum size of the I-shaped cursor (P_{max} = 600 pixels). Input devices studied included the Mouthbrush controller, the pressure sensitive stylus, and the finger-wheel airbrush stylus. Each combination of independent variables was repeated three times for a total of 27 trials. In addition, participants were allowed a practice stage for each trial. During testing, the length of the I-shaped cursor and the time taken to trace the target between the leftmost and rightmost extremes of the channel. We calculated the standard deviation of the length of the cursory as a measure of the precision of user control and calculated the mean absolute error (distance from the edge of the cursor to the green guide line) as a measure of the accuracy of control. Further, we calculated the average task completion time as a measure of efficiency.

5 Results

Figure 6 shows the mean user accuracy and precision on the circle control task, using the two controllers. We observed lower values of the mean absolute error and standard deviation with the Mouthbrush controller, indicating that higher accuracy and precision of control are obtainable with the mouth controller than with stylus pressure. Accuracy and precision were not strongly dependent on gain for the mouth controller. However, with the stylus pressure controller, higher gain values made it very difficult to control the circle size accurately.

As expected, we observed that precision, a measure of how well the user can keep the input size of the circle at a fixed value, was better at lower gain for both of the devices studied. The highest accuracy and precision of control of the circle size was obtained for the lowest gain setting, g = 0.25. At this setting, users were able to attain, with little effort, near single-pixel control of the circle radius using the Mouthbrush controller.



(a) (b)

Fig. 6. (a) Accuracy and (b) precision for the circle control task

Figure 7 shows the experimental results for the line width control task using the three controllers. Overall, tracing the convergent and divergent channels was more difficult than tracing the straight channel, and users took significantly longer and showed lower accuracy and precision on these two tasks. The finger wheel stylus controller offers an advantage for the task of tracing a straight channel with parallel guidelines: it may be carefully set at the width of the channel at the beginning of the task then left unchanged while the user quickly traces the channel. The other two controllers do not allow the user to fix the input value. Accordingly, the fastest completion time for the task with the straight channel was recorded with the finger wheel airbrush stylus, however the Mouthbrush controller was only very slightly slower. For the more difficult tasks with varying channel width, the Mouthbrush offered significantly greater ease of control and users recorded faster completion times – by as much as about 10 seconds compared to the pressure-sensitive stylus. The latter was particularly slow because the user must use the dominant hand to carefully coordinate changes in pressure while the moving the stylus – with the Mouthbrush these are controlled independently by different motor groups. Likewise, users again showed consistently lower accuracy and precision on all three line width control tasks with the pressure-sensitive stylus. The finger wheel, by contrast, showed some advantage over the Mouthbrush in accuracy and precision for the divergent but not the convergent line width control task. We have so far not been able to determine the reason for this difference between the tasks.

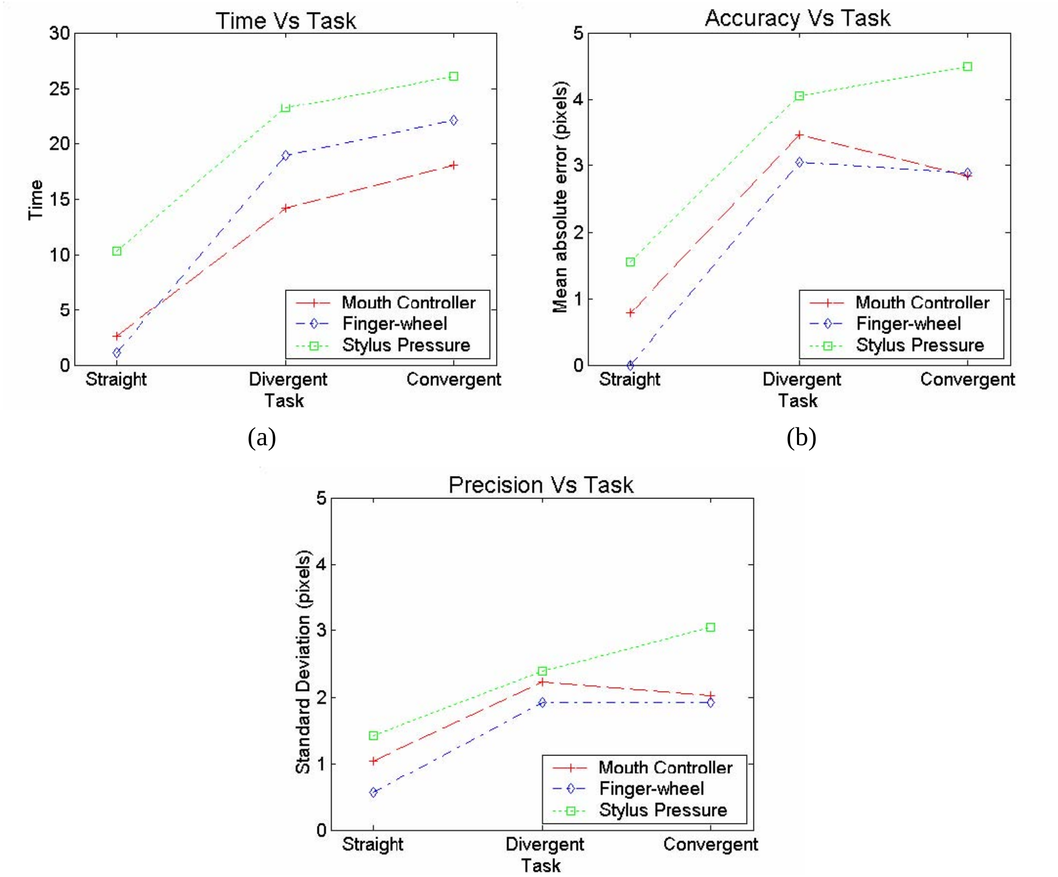
Taken together the results show that the Mouthbrush controller is superior to the pressure-sensitive stylus in accuracy, precision, and efficiency of use, and more efficient than the finger wheel stylus controller. Overall, the finger wheel airbrush stylus offered somewhat superior accuracy and precision of control than the Mouthbrush controller, probably because this controller allows the input value to be set at a fixed value.

5 Sketching and Painting with the Mouthbrush

To evaluate the suitability of the Mouthbrush as a tool for artistic expression, we asked several artists to use it to sketch and paint creatively. We used Windows sockets and the TCP/IP protocol to create communications between the Mouthbrush system and the open-source GNU Image Manipulation Program (GIMP), which has basic functionality for digital drawing and painting. We modified the GIMP code so that the program could act as a TCP/IP client which connects to the Mouthbrush system (acting as a TCP/IP server) to read the open mouth area control parameter, gA.

Four artists participated in our study: three amateurs and one professionally trained commercial artist. Each artist sketched and painted for approximately two hours using the Mouthbrush and the Wacom Graphire pressure-sensitive tablet and stylus. We show a one typical sketch from each of the four artists in Figure 8.

In Figure 8(a), the artist used their mouth to vary the color of the T-shirt, and to vary the size of the brush cross-section while drawing the texture on the pants and the blades of grass. This artist also used their mouth to control the opacity of the bright green colour on the grass, though it may not be readily visible in the figure. In figure 8(b), another artist used their mouth to simultaneously control color, size and opacity. This artist expressed his enthusiasm for the ability of the system to allow selection of a rich, continuous range of colour while sketching. He also commented that the Mouthbrush system has some similarity with airbrush technique, with which he has some experience. The professional, commercial artist in our study sketched figure 8(d). She noted that the Mouthbrush seemed to her a stimulating new technique for expressive drawing.



(a) (b) (c)

Fig. 7. (a) Completion time (b) accuracy and (c) precision for the line width control task

Fig. 8. Digital paintings created using the Mouthbrush

5 Concluding Remarks

The Mouthbrush is a novel interface, based on simple computer vision methods, for drawing and painting via coordinated hand and mouth movements. By making use of a basic technique from pattern recognition to set colour and light intensity thresholds, the Fisher discriminant, our system automatically adapts for use under various lighting conditions. To measure the controllability of the new interface we designed two new usability tasks. With the circle control task, we could measure the static controllability of the mouth and compare it to a pressure sensitive stylus. Results showed that the mouth offered superior control of an input variable. In fact, we were able to control the size of an input shape to a single pixel using the Mouthbrush. We also measured dynamic controllability of the Mouthbrush using the line width control task and compare it with a pressure-sensitive stylus, as well as one with a finger wheel. Again, the results showed that the mouth controller offered advantages over the other two input devices. Sketching and painting trials by several amateur and professional artists were very encouraging. While the idea of using the mouth for digital sketching may sound unusual at first, we would like to point out again that the mouth has played a role in the creation of some of the earliest human art with the paint blown through blow pipes. Moreover, many musical instruments involve use of the mouth. Overall, we conclude that the action of the mouth could offer a valuable new input method for digital sketching and painting and we would like to encourage further exploration of this new interaction paradigm.

References

1. Penfield, W., and Rasmussen, T., The Celebral Cortex of Man: A Clinical Study of Localization of Function. New York: Macmillan, 1990.
2. Baxter, B., Scheib, V., Lin, M.C., and Manocha, D., “DAB: Interactive Haptic painting with 3D Virtual Brushes”, Proceedings of SIGGRAPH 2001, 461-468, 2001.
3. Browne, H., Bederson, B., Druin, A., Sherman, L., and Westerman, W., “Designing a Collaborative Finger Painting Application for Children”,2000.
4. Keefe, D.F., Feliz, D.A., Moscovich, T., Haidlaw, D.H., and LaViola, J.J., “CavePainting: A Fully Immersive 3D Artistic Medium and Interactive Experience”, Proceedings of the 2001 Symposium on Interactive 3D Graphics, 85-93, 2001.
5. Ip., H.S., Young, H., Tang, A., “Body Brush: Generating Virtual 3D Painting through Body Motion”, Emerging Technologies, SIGGRAPH, July, 2003.
6. Greene, R., “The Drawing Prism: A versatile Graphic Input Device”, Proceedings of SIGGRAPH 1985, Vol 19 Issue3, 1985.
7. Hornof, A., Cavender, A., and Hoselton, R. “EyeDraw: A System for Drawing Pictures with the Eyes”, Proceeding of CHI 2004, pp. 1251-1254, 2004.
8. Ryokai, K., Marti, S., and Ishii, H., “I/O Brush: Drawing with Everyday”, Proceeding of CHI 2004, pp. 303-310, 2004.
9. Salem, C., Zhai, S. “An Isometric Tongue Pointing Device”, Proceedings of CHI'97, 538-539.
10. Lyons, M. J. and Tetsutani, N. “Facing the Music: A Facial Action Controlled Musical Interface”, in CHI 2001 Extended Abstracts, ACM Press, 309-310.
11. Vogt, F., McCaig, G., Ali, M.A. and Fels, S. “Tongue ‘n’ Groove: An Ultrasound based Music Controller”, in Proc of the 2002 Conference on New Interfaces for Musical Expression (NIME-02), 60-64.
12. Duda, R. and Hart, P. Pattern Classification and Scene Analysis. New York: Wiley, 1973.