IDENTIFYING AND COMMUNICATING 2D SHAPES USING AUDITORY FEEDBACK

Javier Sanchez

Center for Computer Research in Music and Acoustics (CCRMA)
Stanford University
The Knoll, 660 Lomita Dr.
Stanford, CA 94305, USA
jsanchez@ccrma.stanford.edu

ABSTRACT

This research project shows a technique for allowing a user to "see" a 2D shape without any visual feedback. The user gestures with any universal pointing tool, as a mouse, a pen tablet, or the touch screen of a mobile device, and receives auditory feedback. This allows the user to experiment and eventually learn enough of the shape to effectively trace it out in 2D. The proposed system is based on the idea of relating spatial representations to sound, which allows the user to have a sound perception of a 2D shape. The shapes are predefined and the user has no access to any visual information. While exploring the space using the pointer device, sound is generated, which pitch and intensity varies according to some given strategies. 2D shapes can be identified and easily followed with the pointer tool, using the sound as only reference.

1. INTRODUCTION

The aim of this research project is to use sound as feedback with the aim of recognizing shapes and gestures. The proposed system has been designed with the idea of relating spatial representations to sound, which is as a way of sonification.

Sonification can be defined as the use of nonspeech audio to communicate information [6]. Basically, our proposal consists on relating some parameters of the 2D shape that we want to communicate, with some sound parameters as pitch, amplitude, timbre or tempo between others. By nature, sonification is an interdisciplinary field, which integrates concepts from human perception, acoustics, design, arts, and engineering.

The best-known example of sonification is the Geiger counter, invented by Hans Geiger in the early 1900's. This device generates a "beep" in response to non-visible radiation levels, alerting the user of the degree of danger. Frequency and intensity vary according to the existing radiation level, guiding the user.

Another example of sonification is given by the Pulse-oximeter, which was introduced as medical equipment in the mid-1980's. This device uses a similar concept that the Geiger counter. It outputs a tone, which varies in frequency depending on the level of oxygen in the patient blood.

Other known example of sonification is the Acoustic Parking System (APS) used for parking assistance in many cars. It uses sensors to measure the distance to nearby objects, emitting an intermittent warning tone inside the vehicle to indicate the driver how far the car is from an obstacle.

Sonification has been used to develop navigation systems for visually impaired people [8] allowing them to travel through familiar and unfamiliar environments without the assistance of guides.

Other works [2],[11] are focused on creating multimodal interfaces to help blind and impaired people to explore and navigate on the web. The design of auditory user interfaces to create non-visual representations of graphical user interfaces has been also an important research activity [1], [9].

Some systems have been developed to present geographic information to blind people [5], [7], [10]. It allows the user to explore spatial information.

In some works the aural feedback is added to an existing haptic force feedback interface to create a multimodal rendering system [3], [4].

Although our system would be used to assist visually impaired people in the recognition of shapes and gestures, we do not want to limit its scope to this field of application.

2. SYSTEM DESCRIPTION

In this section is described our proposal, that consists on using auditory feedback to help users in the identification and communication of 2D shapes in those situations where they have no access to any visual feedback.
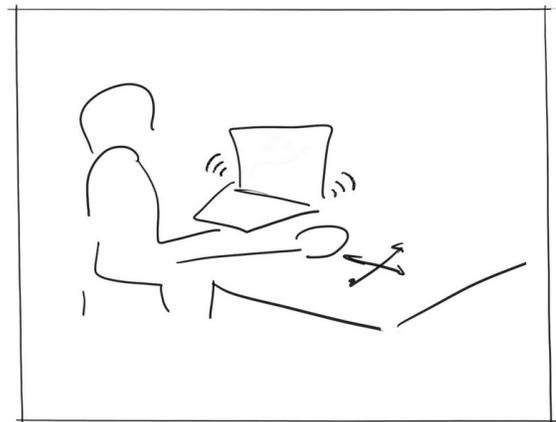
Figure 1. Using an universal pointer device to interact with the system.

Although the system would be conceived as a stand-alone product, the first prototype is designed as a piece of software that runs in any computer.

As the idea of the proposed system is to communicate a 2D shape to other users using auditory feedback, the first thing that has been implemented is a simple drawing interface to generate a 2D shape.

Once the 2D shape has been created or imported into the system, the system is ready to communicate the shape to the user. This communication is made possible by emitting some sounds while the user gestures using a universal pointer device as a mouse, a pen tablet, a pen display or a touch screen of a mobile device. This has been an important design specification of the system, which allows the user to interact with the system using any universal pointer device.

Figure 1 shows how the user interacts with the system using a pointer device. Although the user is sitting in front of a computer, it must be clearly stated again that the user has no access to any visual information.

In order to identify the 2D shape, the user should start exploring the space around him by moving the pointing device. The movement of the user pointer tool is directly associated with the movement of a virtual point in a virtual 2D space where the shape is located.
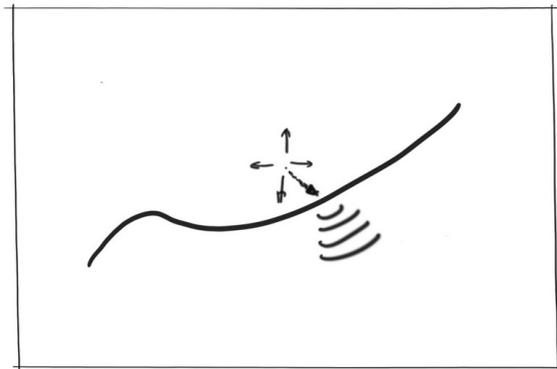
Figure 2. The user has not access to any visual information. A sound is generated when the user approaches to the shape.

As the user approaches to the shape, a sound is generated which pitch, timbre and intensity can vary according to a specific spatial to sound mapping strategy. Figure 2 shows how sounds are generated when the user approaches to the shape.

Once the user has located the 2D shape, the following step consists on trying to follow the shape using the sound as only feedback. If the user moves away from the curve, the sound disappears and the user can get lost into the silence. Anyway, the user can easily move the pointer back to the last position where the sound appeared, to continue tracking the position of the shape.

The size of the user workspace while moving the pointer matches with the size of the screen where the shape is located. When the user moves the pointer further the limits of the workspace, a different sound tells him that he has reached the workspace limits. This is very useful when using the mouse as pointer device.

The proposed system is based on the proprioception sense, which provides a relation between the gestures made by the

user while following the sound, with the spatial representation of these gestures.

Thanks to the proprioception sense, the hand gesture made while following the sound is transformed into a spatial representation of the shape. Figure 3 shows how the user can reconstruct mentally the 2D shape using the auditory feedback.
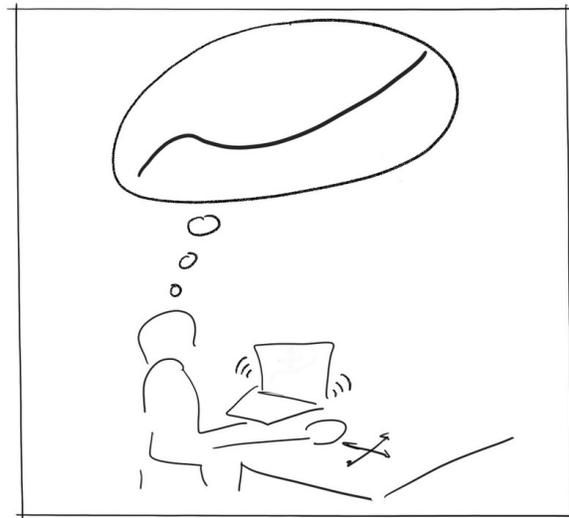
Figure 3. Users transform the gesture made while following the sound, into a spatial representation of the shape.

There are several ways of identifying the 2D shape using the pointer device. Some users would prefer to keep following the 2D shape slowly without loosing the sound. On the other hand, other users would prefer to start moving around the whole workspace, from side to side, getting some scattered points, which can be later connected mentally to form the 2D shape (see Figure 4).
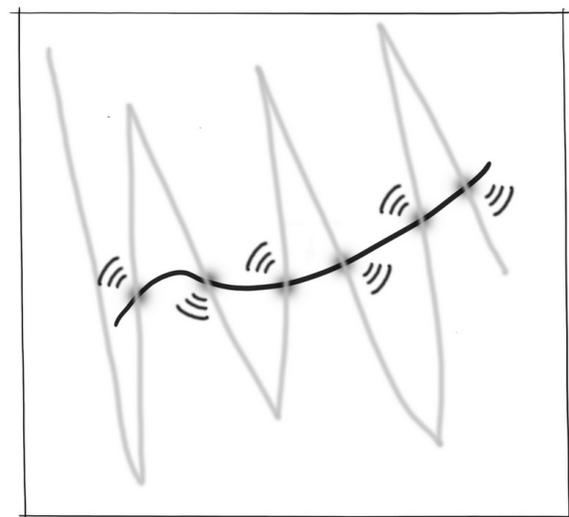
Figure 4. User movements from side to side of the screen trying to find a 2D shape using the sound as feedback

In order to provide a relation between the gesture made and the sound feedback, a perfect synchronization of perceived audio events with expected tactile sensations is needed.

The user workspace is divided into two different areas: *sound* areas and *no sound* areas. Figure 5 shows how the limits between sound and silence are located at certain distances at both sides of the 2D shape. The transition between silence and sound is made gradually, as shown in figure 5 where the sound intensity increases as the distance to the curve decreases.

The value given to this distance is not trivial and its appropriate selection will ensure that the user will be able to identify adequately the 2D shape using auditory feedback. If the distance were greater than needed, the sound area would be too wide. This would imply the possibility of finding multiple solutions, which would be far from the 2D shape that the user was trying to identify.

In the other hand, if the distance were too close to the 2D shape, it would be difficult for the user to locate a 2D shape, due to its thin thickness. The 2D shape would become invisible.

The value of this distance depends also on the pointer device used. For example, it is not the same to use the small track pad of the laptop that using a 15'' pen tablet. In this example, the ratio between the size of the finger and the track pad area is much bigger that the ratio between the stylus diameter and the area of the 15'' pen tablet. The value of this distance is also related to the resolution of the pointer device.

So, further studies should be carried out to find the optimum distance that delimits the sound area around the 2D shape.
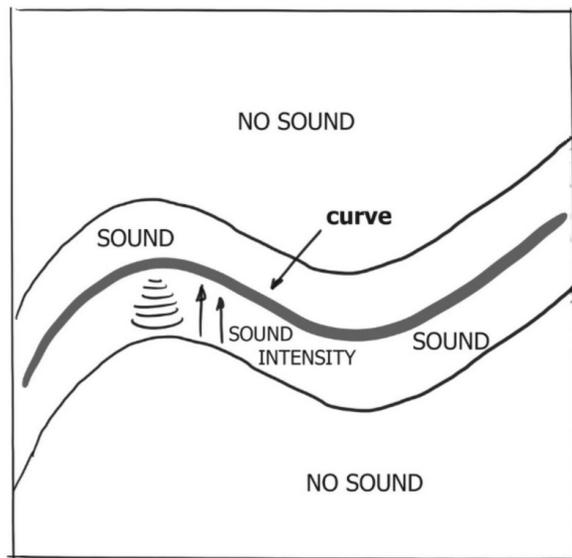
Figure 5. Sound to spatial relationship. Sound intensity increases as the distance to the curve shape decreases.

When working with pointer devices, it is necessary to be aware of the differences between relative and absolute referencing. In our system, it is much better to work with absolute references. Most of the pen displays and touch screens use absolute references. It is not the same case when dealing with a mouse or a track pad, where the referencing system is relative.

As example, if the user lifts the mouse, moves it away, and places it again on the surface, the pointer stays in the same position on the screen. This is not useful for our system, since the user would lose the spatial reference while trying to locate a 2D shape.

On the other hand, if the user uses a pen tablet, the whole area of the tablet is mapped to the whole area of the screen. So, if the user lifts the stylus, moves it away and places it again on the surface, the pointer moves to another position on the screen. This is exactly what we need.

3. STRATEGIES TO MAP GEOMETRY TO SOUND

An application has been built with the aim of studying how easy would be for a user to identify a 2D shape using the sound as feedback. Some parameters will be set to adjust the process. In this section are given some technical details and strategies used to develop the application.

As stated in the previous section, the sound intensity increases as the distance from the pointer to the 2D shape decreases. In addition to this, some parameters of the 2D shape, as position, slope or curvature, are used to enrich the sound information given to the user.
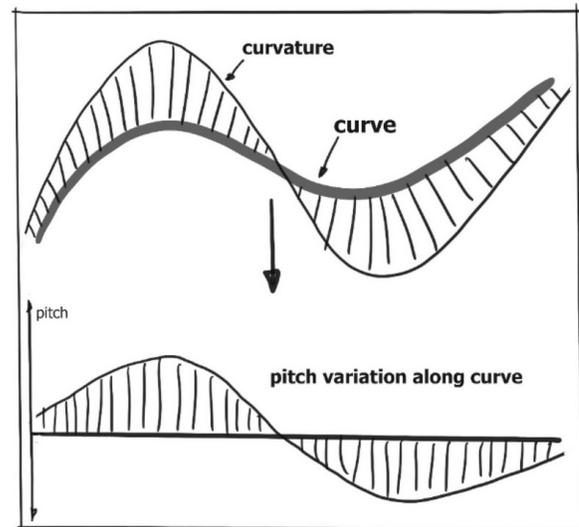
Figure 6. Sound to spatial relationship. Some properties of the 2D shape as slope or curvature are associated with the sound parameters to enrich the sound feedback.

As example, a pitch variation of the sound feedback would tell the user about the curvature of the shape at each point. So, a possible strategy would consist on varying the sound pitch along the 2D shape, according to the curvature at each point of the shape.

According to this, a straight line will generate a constant pitch. The curve represented in figure 6 has a variable curvature, so the user will have different pitch perceptions while moving along the shape.

Another useful strategy would be to use the slope of the 2D shape at each point to generate different sound pitches along the shape. Depending on the shape, it would be more appropriate to use one or other strategy.

So, the position of the pointer together with some geometric properties of the 2D shape would help to enrich the sound information given to the user.

There are other sound parameters that would be used to enhance the auditory feedback. As example, the duration of the sound can be related to the thickness of the 2D shape. This strategy would allow the users to make a difference between shapes with different thickness. It would be even possible to identify changes in thickness within the same shape, using the sound as feedback.

Depending on the pointer device used, it would be more convenient to relate the thickness of the shape to the loudness of the sound generated.

What about adding some effects to the original sound to express other variations that would appear on the geometry? We would distort the original sound using some filter, as a reverbs or an echo, to relate the new sound to the style of the pencil used to draw the 2D shape. Other parameters of the 2D shape as the transparency or the applied pressure while creating the stroke would be associated with some distortion of the generated sound.
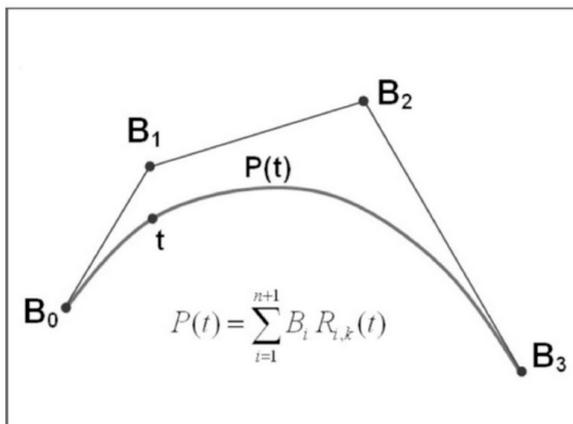
Figure 7. Parametric curves are used to define shapes.

Color is another property that would be associated with some sound property. We can start thinking in a system with 8 basic colors, which are associated with 8 different sound timbres. This relation has been established since timbre is considered as the color of music. Both terms timbre and color are used indistinctly traditionally to represent the sound quality.

Other possibility that have been included in the system is to represent a 2D closed shape. Imagine that the user is trying to follow a 2D rectangular shape. We can use the same previous strategies to identify the edges of the rectangle, relating them to a specific sound, and add a new sound to the area that is contained inside the rectangle. This strategy will enrich the sound feedback and will help the user to identify a shape.

Some primitive shapes as, circles, ovals, rectangles, triangles, etc, can have a secondary sound associated to the shape, which would indicate the user that he is trying to identify

one of these singular shapes. This secondary sound doesn't need to be always active; it can appear slightly every few seconds to avoid excessive noise in the scene.

The idea of including several channels at the same time to express several shape properties would really facilitate to the user the identification of 2D shapes and enrich the sound feedback. Other sound parameters that would be used in the proposed system can be panoramization effects, changes in tempo and rhythm, or fade-in and fade-out transitions between sounds.

Auditory feedback should not be reduced only to sound. Music, voice or noise would be also used in the proposed system. A voice can be mapped to a linear shape and be triggered depending on the position of the pointer along the shape. The user would control the voice or some music back and forward at the desired speed as if controlling a music player. Following a music score would be also associated with the movement of the pointer device.

Special care should be taken with the selection of the generated sound. Using the same kind of sounds can be hard and tedious for the user, or even painful, depending of the ranges of pitches used. A library of sounds can be included to allow users to choose their own sounds. Random sound selection is another option. Ambient sound can be used to fill the background and some atmosphere sounds can be associated with the internal area of closed shapes. Textures can be associated with some noise added to the original sounds.

The 2D shapes are represented by means of parametric curves, which are a standard in 2D drawing representation. Since Drawing Exchange Format (DXF) is used to store the graphic information, it is very easy to generate curve shapes using any commercial CAD application and import them into our system. Figure 7 shows an example of a parametric curve.

Multiple curve shapes can be defined into the same scenario using different sound for each curve (see figure 8). Distances to the curves are evaluated as the user interacts with the model. Including too many entities in the same scene can be not the best idea, especially if using the track pad or the mouse as pointer device. A bigger workspace would be needed. A pen tablet or a pen display are preferred when working with multiple shapes.
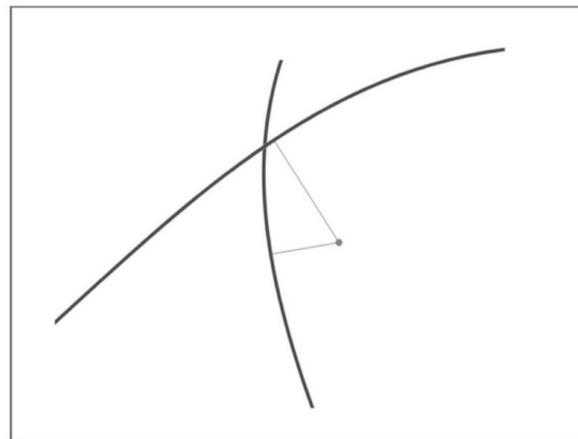
Figure 8. Multiple shapes are associated with different sounds.

4. SYSTEM IMPLEMENTATION

The analysis of the user motion, the curve representation and the output sound has been computed using MAX/MSP, a visual programming environment specifically designed to simplify the creation of acoustic and control the application.
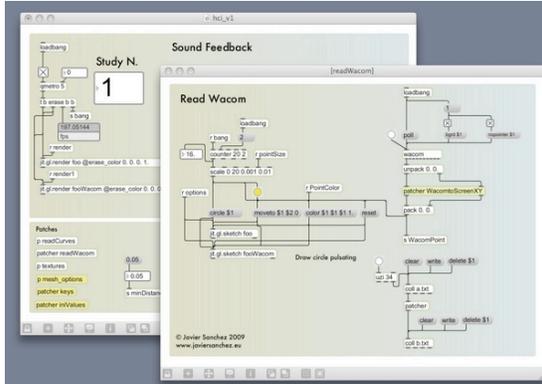
Figure 9. MAX/MSP is an excellent programming environment to test a prototype system, adjust sound parameters or communicate with any universal device.

Controlling external devices as the mouse, a pen display, an iPad or and iPhone is very easy to do using MAX/MSP. The visual programming environment facilitates the control of the process and the communication with other systems. Figure 9 shows a MAX/MSP snapshot.

The Processing programming environment has been chosen for building the visuals of the application (see Figure 10). Processing is an open source programming language and environment to work with images, animation, and interactions. It is also an ideal tool for prototyping.
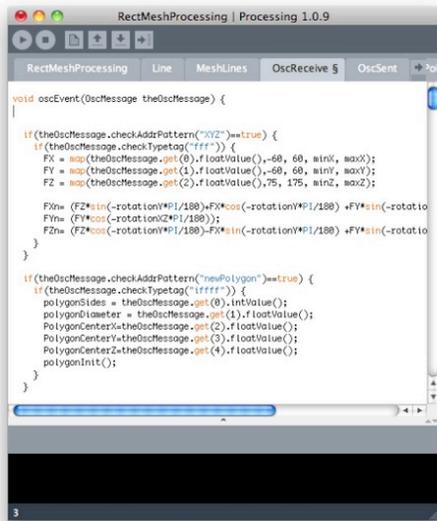
Figure 10. Processing is the programming environment used to control the application visuals.

The connection between MAX/MSP and Processing is made using the OSC (Open Sound Control) protocol, which bring the benefits of using modern networking technologies. It provides also everything needed for real time control of sound and other media.

Some other devices as the iPhone or the iPad Touch can be used as pointer devices. The OSC protocol can be used to communicate the mobile device with MAX/MSP using the wireless network. The TouchOSC application [12] has been used to connect the iPhone with MAX/MSP.

Figure 11 shows the appearance of the implemented application. As the idea of the system is to recognize shapes using the sound as feedback, the first step consists on drawing something on the screen. A schematic shape of a car has been represented using 5 lines: one for the external profile, two for the wheels, one for the door and another one for the bottom line. This drawing can be drawn by another user or can be loaded from a collection of drawings stored in the computer.

Once the drawing is completed, the next step consists on recognizing the shapes using the sound as feedback. It is evident that the user has no access to any visual information. As the user moves the pointer device, some lines appear on the screen, which represent the shortest distance from the pointer device to the drawing lines. These lines are updated as the user navigates around the screen.
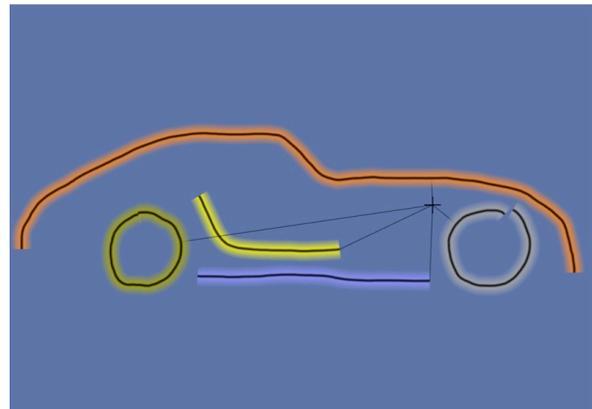
Figure 11. Snapshot of the implemented application, showing a schematic shape of a car, which is recognized by the user using the sound as feedback.

When the user approaches to any of the lines, a sound appears. This sound is related to the geometry by means of some mapping strategies, which are described in the previous section.

A new mapping strategy consists on the use of music as auditory display, instead of sound. The reason of this is that it is much more comfortable for the user to use his own library of music, that synthesized sound. Each curve can be related to a different music theme of the user library. So, when the user approaches to a line on the screen, a specific music theme is played.

In Figure 11 can be shown how each curve is made of two different sub-curves: a thin black curve inside and a ticker colored curve outside. These two different curves are associated with two different audio channels: music and white noise. Let's explain this. When the user approaches to the curve and the

pointer is touching the colored area, a white noise appears, which tells the user that he is approaching to the curve. As the user moves closer to the black curve, the white noise disappears gradually, and the music appears clearly. When the user moves away of the thin black line, the music disappears gradually, and the white noise appears again.

The metaphor used in this system is based on the idea of tuning a radio. When the user approaches to a radio station, a clear sound appears. The white noise is telling the user to move the dial until he gets the desired radio station. So, our system can be seen as 2D radio tuner. The user can navigate in the 2D space identifying the curves and following them using the music and the white noise as feedback.

Figure 12 shows a control panel in which the user can associate each curve with a specific theme from his music library. The color and the width of each of the two sub-curves associated with each curve can be also adjusted easily from this control panel.
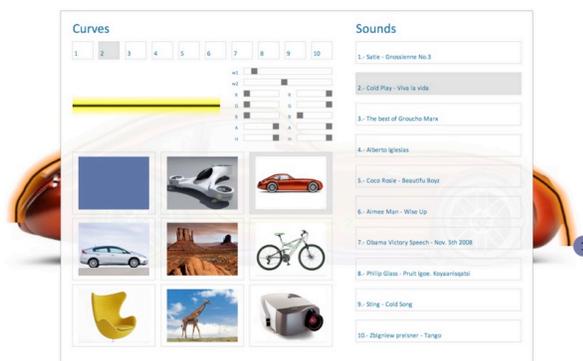
Figure 12. Control panel of the implemented system.

A background can be used as reference to trace easily the curves of the model. Figure 13 shows how a picture of a car has been used to sketch the five curves of the model. Users can have their own library of pictures to be used as background.

Finally, it is important to emphasize that each line has an identifier and can be edited or deleted if desired.
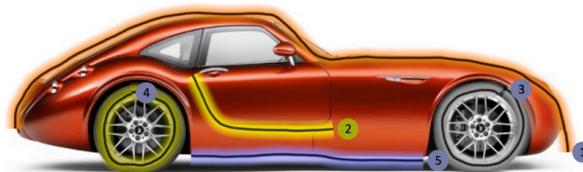
Figure 13. Users can use their own picture library as background to trace the curves of the model.

5. CONCLUSIONS

This paper proposes a novel method that consists in the use of auditory feedback to identify a 2D shape while the user gestures using a pointer device.

Several universal pointer devices, as a mouse, a pen tablet or a mobile device can be used to interact with the system, facilitating the human computer interaction.

Parametric curves are used, as they are a standard in 2D drawing representation. Some of the curve parameters, as slope, curvature or position, are related to the sound output, helping the user to identify the 2D shape.

Other parameters of the 2D shape as color, thickness can be associated to different timbres or loudness. Multiple sound channels can be included to add extra information of the background or to identify some closed areas.

Multiple 2D shapes can be defined in the same scenario using different sounds for each shape.

As it occurs in any interaction device, the user needs certain time to become familiar and confident with the new environment. Users can become skilled in a short time since the application is very intuitive and easy to use.

Current work is related with the use computer vision techniques to track the hand movement of the user. By means of this, the user can interact directly with the system, using the webcam of the computer.

It is also being evaluated the possibility of using the system as an extension (add-on) of some existing computer application.

Other applications are also been studied in which the sound can be related to a gesture to assist the user in common tasks.

The overall low cost of the system and its easy implementation is also an important point in favor.

A collection of applications based on the idea of using sound as feedback has been implemented for the new iPad. Applications for visually impaired people and collaborative games are the most important.

6. REFERENCES

[1] W. Buxton, "Using Our Ears: An Introduction to the Use of Nonspeech Audio Cues" in Extracting meaning from complex data: processing, display, interaction, edited by E.J. Farrel, *Proceedings of the SPIE*, Vol. 1259, SPIE 1990, p. 124-127.

[2] H. Donker, P. Klante, P. Gorny, "The design of auditory user interfaces for blind users" *in Proc. of the second Nordic conference on HCI*, pp. 149-156 (2002)

[3] N.A. Grabowski, K.E. Barner, "Data visualization methods for the blind using force feedback and sonification." *in Prceedings of the SPIE Conference on Telemanipulator and Telepresence Technologies*, 1998

[4] [IFeelPixel: Haptics & Sonification http://www.ifeelpixel.com/faq/#whatitwill

[5] H. Kamel, J. Landay. "Sketching images eyes-free: a grid-based dynamic drawing tool for the blind." *In Proc. of ACM SIGCAPH Conference on Assistive Technologies* (ASSETS). pp. 33-40. (2002)

[6] G. Kramer, B. Walker, T. Bonebright , P. Cook, J. Flower, N. Miner, J. Neuhoff "Sonification Report: Status of the

Field and Research Agenda." *In International Community for Auditory Display*, ICAD (1997)

[7]   M. Krueger, "KnowWare™: Virtual Reality Maps for Blind People." *SBIR Phase I Final Report, NIH Grant #1 R43 EY11075-01*, (1996)

[8]   J.M.Loomis, G. Reginald, L.K. Roberta "Navigation System for the Blind: Auditory Display Modes and Guidance. " *in Presence*,V.7,N.2,193–203(1998)

[9]   E. Mynatt, G. Weber. "Nonvisual Presentation of Graphical User Interfaces: Contrasting Two Approaches." *in Proc. of the Computer,* CHI 94. (1994)

[10] P. Parente, G. Bishop "BATS: The Blind Audio Tactile Mapping System." ACMSE. (2003)

[11] W.Yu, R. Kuber, E. Murphy, P. Strain, G.A. McAllister " Novel Multimodal Interface for Improving Visually Impaired People's Web Accessibility." *in Virtual Reality*, Vol 9: 133-148 (2006)

[12] Touch OSC.

http://www.creativeapplications.net/iphone/iosc-iphone/