Available online at www.sciencedirect.com
ScienceDirect
Interacting with Computers 20 (2008) 386–405
Interacting with Computers
www.elsevier.com/locate/intcom

Non-visual game design and training in gameplay skill acquisition – A puzzle game case study

Tatiana V. Evreinova, Grigori Evreinov *, Roope Raisamo

Department of Computer Sciences, University of Tampere, Kanslerinrinne 1, Pinni B, FIN-33014 Tampere, Finland

Received 14 July 2007; received in revised form 28 January 2008; accepted 23 February 2008
Available online 5 March 2008

Abstract

This paper reports the results of a study on the design and evaluation of the game and techniques which allow puzzles to be played in the absence of visual feedback. We have demonstrated that a camera-mouse can be used successfully for blind navigation and target location acquisition within a game field. To gradually teach the players the sequential learning method was applied. Blind exploration of the gamespace was augmented with sticky labels and overview sound cues, verbal and non-verbal, which can significantly reduce the cognitive load and facilitate mental matching and integration. The full-sticky labels technique does not require fine motor skills and allows a user to gain control over the game with a minimum level of skills. With the vertical sticky labels technique training was focused on the development of accurate head movements only on a horizontal plane. With practice, the players can use the non-sticky labels technique. After 240 trials (3–4 h), the cumulative experience of the blindfolded players was increased 22.5–27 times compared to the initial 10 trials.

Keywords: Non-visual puzzle game; Sonification; Overview sound cues; Sticky labels; Sequential learning; Skills training

1. Introduction

A visual-based culture has a strong impact on human intellectual and creative potential, the development of perceptual and motor abilities (Segall et al., 1968). Nowadays, in different forms and formats, digital video is becoming increasingly ubiquitous. Video images and audio messages contain emotional components and specific patterns which can easily be perceived, and, due to empathic arousal, have a strong influence on people viewing, listening and reading by forming their social responses and behaviour (Grèzes and Decety, 2001; Prinz and Meltzoff, 2002). Watching is the easiest way to learn through stimulation of the 'theoretical imagination'. Being deprived of access to information with emotionally rich content, blind and visually impaired children have experienced a significant emotional distress which can lead to depression and delay in cognitive development (Barresi and Moore, 1996; Chartrand and Bargh, 1999). Furthermore, for children with partial sight image-based games are considered to be a supporting therapy to encourage residual vision and a way to continue the development of the visual cortex (Eriksson and Gärdenfors, 2004; Guhl, 2005).

On the other hand, the information capacity of visual and auditory systems in humans is comparable (Fjeldsenden, 2000; Shimojo and Shams, 2001; Wada et al., 2003). Nevertheless, perceptual mechanisms of different modalities have evolved to complement the human perception of different dimensions to disambiguate the afferent flow in a sensory conflict situation and not to duplicate each other. Audition is perfect and dominates in processing temporal information and verbal constructions; vision dominates in the exploration of spatial information, social interaction, and cognitive skill acquisition (Bertelson and Aschersleben, 1998; Kitajima and Yamashita, 1999; Kitagawa and Ichihara, 2002; Repp and Penel, 2002; Guttman et al., 2005).

* Corresponding author. Tel.: +358 3 3551 8549; fax: +358 3 3551 6070.
E-mail addresses: etv@cs.uta.fi (T.V. Evreinova), grse@cs.uta.fi (G. Evreinov), rr@cs.uta.fi (R. Raisamo).

0953-5438/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.intcom.2008.02.008

Sounds are only partly complementary to tactile and not to the kinaesthetic or proprioceptive signals. To appropriately comprehend and interpret spatial relationships blind people have to employ cognitive modelling of the external space which must be well coordinated with their motor activity. Cognitive mapping in the total absence of visual feedback is quite different from visual mapping for representation of the spatial locations (Kurze, 1996; Ungar, 2000). The absence of visual perception creates the conditions when the sensory-motor cortex and other associative brain regions will be involved in the parsing and integration of all accessible information from the intact modalities.

According to Millar (1994), in pointing tasks cognitive spatial encoding can be done as egocentric, allocentric, and be based on 'movement memory'. In fast implicit tasks for immediate goal-directed movements egocentric coding is more appropriate (Milner and Goodale, 1995) while visual and haptic experience will affect the mental model in a different extent. As it was demonstrated recently for early-blind, late-blind, and blindfolded-sighted individuals "the need to rely more exclusively on haptic inputs stimulates haptic dexterity in the blind" (p. 1263, Postma et al., 2007) while sighted people strongly rely on vision to represent peripersonal space. In order to be able to play the game, blind and visually impaired people mostly have to rely on sound and haptic/tactile cues. Thus, auditory and haptic inputs will dominate in interaction scenario.

So far there is not a common method for image-to-sound conversion as a natural and intuitive substitute for the vision sense acquired and used by a number of blind people. It is apparent that metadata (Metadata, 2007) can simplify the processing of the information the video images are intended to communicate to the person. Such a technique is very suitable for creating accessible mainstream games where a designer and a player should a priori follow the game scenario. But the problem is how to communicate to the blind person a spatial layout of static objects within a short time frame. Should it be some kind of a sound equivalent of the visual frame (Meijer, 1992) or would another way be more natural?

2. Games and sounds

Designing audio games is a very specific field of sound application. When gameplay should occur exclusively in an auditory space, any sound events will have an impact on the player's behaviour in terms of perceptual, motor and cognitive deviations (Whitmore, 2003; Collins, 2007). Sounds can be used to convey different types of information. Sound cues can guide a player's attention in different ways: by holding and attracting attention, by diffusing attention or distracting. Short messages and specific speech effects, such as non-verbal vocalizations, (Carey, 1980) can influence decision-making and indirectly affect player activity. For instance, a derisive laugh or ironic remarks can induce the person to continue playing to improve the results. Acoustic attributes such as pitch and timbre are used to stress the positive or negative features of the game characters. The temporal nature of action sounds, pace, rhythm and timing are efficiently employed to stimulate perceptual–motor coordination in action games. Quiet background music or appropriate environmental sounds are not only ornamental sounds, this trick is intended to enhance player immersion and to create a feeling of the unique atmosphere of the game.

Sounds used for audio-based interaction in games can be categorized from the viewpoint of listening comprehension as causal, semantic and reduced (Chion and Gorbman, 1994). The causal sound means that the listener can recognize auditory events only to the extent of the nature of their source. Semantic listening refers to the comprehension of the encoded information from an auditory message. Reduced listening implies that the perceiver will only pay attention to a sound event as an adequate auditory stimulus without recognizing of specific sound features, physical or cognitive (Blesser and Salter, 2007).

Conventional board games provide a strong external memory aid through the option to overview the game field at any moment. In contrast, non-visual presentation of the board game by means of sonification implies converting most of the game features into sound parameters (Andresen, 2002). The great challenge is how to present efficiently a spatial layout of several objects, their actions, features and spatial relationships, employing the temporal dimension of audition. By the term 'efficiently' we mean that a time for the recognition and comprehension of the sound pattern would be compatible with the time for the mental processing of the visual prototype (an overview of the layout). However, when all the features of the game sound simultaneously and continuously, e.g., in a loop manner, sonification becomes more annoying than useful (Winberg and Hellström, 2001). It is reasonable to ask: why should the task of image-to-sound transformation be implemented by simulating a concept of the visual frame in a manner reminiscent raster (Meijer, 1992)?

In fact, due to the 'cocktail-party effect' the listener can recognize and distinguish a specific sound pattern in the noise-like sound mixture (Cocktail party effect, 2007). However, gameplay sonification should be implemented in such a way as to give the blind player the opportunity for total control over auditory information and sound events, like focus of attention in human vision. Instead of a 'passive' listening and a sophisticated parsing for sound mixture we propose to present information in a different audio format (wave-files, MIDI sounds, synthesized speech) coordinated with the exploratory activity of the player. We therefore introduced *overview sound cues* - events which can inform the player through a descriptive message as the result of the specific analysis of an object's features, actions or an overview of the particular situation in a local area of the game field. Below we will demonstrate how the overview sound cues, verbal and non-verbal, can