Easy-To-Use Interface for a Blind Person to Control a Media System

Last update: March 27, 2015
Author: Alan McNaughton alan@mcnaughton.org Cell 403 202.4747

General Description

This paper describes a physical and logical interface to enable a person with certain visual and tactile challenges to access various forms of entertainment as well as perform some other limited functions. The input from the user is by pressing, as appropriate in context, one of five buttons and turning a volume control. The user is presented with 'menu' choices audibly in spoken words. The system initially controls a TV and speakers, and later may control other devices such as a phone.

Target User

This system is designed for someone with limited vision or no vision and limited tactile ability. The choices for the user are presented audibly in spoken works so the user must have sufficient hearing and cognition to hear and understand the choices presented.

The person who initially inspired the development has visual agnosia, is unable to distinguish letters or numbers, is unable to distinguish an item the size of keys on a regular keyboard, and has tactile agnosia which in his case may be thought of as being very clumsy and without fine motor control. The user does not know Braille and doesn't have enough sense of touch to learn it. The person has a great knowledge and love of music and the system was conceived to enable him to select his own choice of music and movies.

Audible Prompting

Audible prompting and feedback is provided at each progression though a list. For example when selecting "Live Radio" from the top list by pressing the select button (typically the green button), the system says the name of the first item in the "Live Radio" list, which maybe something like "CBC Radio One Edmonton". If the user pressed the forward button (typically the blue button) the next item in the list would be said. The audio feedback is provided by a text-to-speech feature such as Cepstral Personal http://www.cepstral.com/. The text-to-speech ability allows meta-data from the content to drive what is populated in some lists. This meta-data is information like the name of the song or album and the artist who performed it. The Cepstral system provides control over the speed of the voice, the pitch and the choice of several voices.

Volume Control

Volume control is provided by a large rotary volume control such a model from TC Electronic called a Level Pilot. It was chosen for its physical size and mechanical characteristics. http://www.tcelectronic.com/level-pilot/ . An alternate being explored is the use of a rotary encoder which should provide more flexibility in how volume control is handled. Once example of

such a rotary encoder is the Griffin PowerMate USB
http://store.griffintechnology.com/laptops/powermate

Turning the volume control clockwise increases the volume and clockwise decreases it. With the Level Pilot there is a physical stop a zero volume and at a full volume, while with any rotary encoder the knob can be turned infinitely in either rotation.

Buttons

The system uses five buttons to navigate through nested lists, to make content selection and to control the experiencing of the content. The buttons are large, 64 mm (2.5 Inches) in diameter, and each button is a different colour. The buttons make an audible click when pressed and provide a small amount of vertical travel to help the user know the button has been pressed. The buttons are from Origin Instruments, a model called Orby (http://www.orin.com/access/orby/)

The buttons are arranged in a grid as follows:

BLUE (forward)	GREEN (select/resume)
YELLOW (backward)	RED (stop)
WHITE (information)	

This layout is assumed to be best for a right handed person where the buttons are placed to the right of the user.

Whether five buttons is optimal for the task has not been studied. It is not known if, for example, having six buttons may be simpler cognitively for the users. An extra button could, for example, be dedicated to the "Go to the top and start over" function.

Button Usage

The meaning of each button depends on the context. The usage has been made as intuitive as possible. For example when listening to content the red button is used to stop and the green button would resume (go) building on the STOP/GO of traffic lights. Within a single context each button can have several uses depending upon whether the button is pressed once or twice in quick succession. The simplest use is a single press and the system has been designed to be useable with only single presses.

The purpose of the buttons remains as consistent as possible between different contexts. For example, the Blue button is always for some type of next or forward.

List Mode

When in List Mode the buttons are used to move through the list, to select an item in the list or to move to a different level of the nested lists.

	SINGLE PRESS	DOUBLE PRESS
BLUE	FORWARD in the list	Same as two single presses as hard to distinguish from trying to rapidly move through the list
YELLOW	BACKWARD in the list	Same as two single presses as hard to distinguish from trying to rapidly move through the list
GREEN	SELECT the current item if it is content and go to Content Mode and begin playing, or SELECT and move into the list if the item is a list	
RED	Up a level	Future: Directly to top level
WHITE	Explain/Help	Future: Longer explanation

Content Mode

In the Content Mode content is being presented to the user and the buttons are used to move between pieces of the content, to pause and resume the presentation of content.

	SINGLE PRESS	DOUBLE PRESS
BLUE	FORWARD a logical amount (EG. a song)	Future: FORWARD a larger logical amount (EG. an album)
YELLOW	BACKWARD a logical amount	Future: BACKWARD a larger logical amount (EG. an album)
GREEN	if paused, continue	
RED	Pause, or if paused, up a level	Future: Directly to top level
WHITE	Describe the content	Future: Longer explanation

Random Mode

The Random Mode appears at many levels within the lists. Choosing it presents a random piece of content within the scope of where the random mode was selected.

	SINGLE PRESS	DOUBLE PRESS
BLUE	Suggest a piece of content at random	
YELLOW	BACKWARD to the previously proposed item	
GREEN	Select this item (goes to Content Mode & begins playing)	
RED	Up a level	Future: Directly to top level
WHITE	Describe the content	Future: Longer explanation

Nested Lists

The system will make use of a hierarchical set of lists, a set of nested lists. The user hears one item at a time and can scroll forward or backwards to another item in that list. After hearing an item it can be selected. The item selected can be another list.

The following is an example of the nested lists:

Live Radio (select from a list of live audio streams that are radio)
- CBC 1
- CKUA
- NPR
- [Other live radio stations, a fairly small list]

Live Television
- Space Channel
- History Channel
- [Others, all named, not all channels, just the ones of interest to the user]

Music (select recorded music in various ways)
- Random
- Recently Added
  - Random
  - Most recent not-yet-listened-to
  - Most recent
  - List of recently added
    - [Content driven list of items]
- By Genre
  - Jazz
    - Random
    - List
      - [List albums by name from meta-data in content]
  - Classical
  - [A small list of other identifiable genre's such as Rock, SoundTrack, etc]
- By Artist
  - [List artists from meta-data in content]
- By Album
  - [List albums from meta-data in content]

Recorded Audio
- Recorded Radio shows
  - Resume Last
  - Recently Added
    - Random
    - Most recent not yet listened to

Most recent
List of recently added
[Content driven list of items]
Long List
[Content driven list of items]
Lectures/Talks
Resume Last
Recently Added
Random
Most recent not yet listened to
Most recent
List of recently added
[Content driven list of items]
Long List
[Content driven list of items]
Comedy
[Content driven list of items]

Email
Read New
Send a Reply Read
Reply with Audio message
Read Old

Make a phone call
[List of names eg. Alan's Cell]

Feature Presentation

Not all features would be implemented initially; features would not show up until they were functional. The focus in software development will be to do the basics well and then add other functionality. The basics are music, live TV and movies.

The order of the items in the list and whether or not a particular feature would be presented would be easily configurable. The lists can be configured to either wrap around to the top from the bottom (and vice versa) or to stop when the top or bottom is reach.

The system will require configuration by someone capable of using a full keyboard and with moderate computer skill. There will be no attempt to make the system configurable by the target user.

Goals of the System, Software and Hardware

The goals of system are:

- To be as simple and intuitive to use as possible. Ideally the features would be discoverable.
- To be mechanically robust as the system is designed for someone who is clumsy.
- To develop it as an open source project so that other can benefit from it and, hopefully, share in the long term maintenance effort. This is not to preclude rapid development by use of complete commercial components (Eg: a commercial text-to-speech system this is easy to integrate and is of higher quality than open source alternatives).
- To be highly flexible via configuration. The assumption is that the end user won't be able to configure the system; it will be done by someone with full sight and using a full keyboard.
- To be developed with as little custom software as possible. It is to be built as much as possible using existing open source packages and commercially available software.
- To be developed with as little custom hardware as possible. For example the buttons, volume control, and computer should be existing commercial items. It is expected that the overall enclosure will be a purpose built custom arrangement.

Proposed Computing Approach

A Raspberry PI 2 (http://www.raspberrypi.org/products/raspberry-pi-2-model-b/) would be the brains of the system to run all the software: nested list application, KODI media system, text-to-speech program, etc. The PI would be connected to a TV, the buttons, the volume control, to speakers for the audio prompting and an IR blaster. An IR Blaster would be used to change source and channels on the TV. Content will be stored on a LAN accessible file system accessed by the PI over WIFI.

The system will leverage the KODI entertainment system (formerly call XBMC), a large open source project for entertainment systems that provides all the infrastructure for playing audio and video. http://en.wikipedia.org/wiki/Kodi_%28software%29

The operating system on the PI will be Linux; the exact flavour is not important and will be chosen for development convenience.

Software Languages for Custom Software

The programming language that the custom software will be written in hasn't been chosen. Python and Ruby are likely candidates. Some C++ may be necessary to interface to the low level IO.

The custom software needs to interface to a number other components:

COMPONENT	MFG	PROGRAM INTERFACES AVAILABLE
IR Control	Phidgets	Via USB communications, examples available in Python, Ruby, C#, C++ and many others
Buttons	Phidgets	Via USB communications, examples available in Python, Ruby, C#, C++ and many others
Rotary Encoder (Volume control)	tbd	The approach has not been decided. Possibilities include USB communications using either Phidgets (see above) or PoKeys (examples available in Python, C#, C++, VB.Net, Delphi) or other.
Text-to-Speech	Cepstral	Closed Source, has C, C++ interface
Media System	KODI (XBMC)	Open source and has several extension approaches see: http://en.wikipedia.org/wiki/Kodi_%28software%29#Addons_Manager.2C_addons_and_plugins Includes Python or JSON-RPC. Ruby via: https://rubygems.org/gems/xbmc-client

Mechanical and Electrical Interconnections

The buttons and volume control are mounted on the top of a plate. A shoebox sized enclosure under the plate houses the InfraRed (IR) control, the custom interconnection box that contains the button I/O interface, the Raspberry PI, speakers and the interconnecting cables.

The Raspberry PI is connected to the TV via an HDMI cable. The Raspberry PI is powered by a power adapter that connects to the mini USB port of the PI. The HDMI cable and the power cable are the only two cables to exit the 'shoebox', all other connections are internal.

The buttons, from Origin Instruments, (http://www.orin.com/access/orby/), are wired into a custom interface box that connects the wires on the buttons into the terminal strips of a model 1018 I/O interface kit from Phidgets (http://www.phidgets.com/products.php?product_id=1018) The Phidgets 1018 connects via USB to a Raspberry PI 2. (http://www.raspberrypi.org/products/raspberry-pi-2-model-b/). The cable is a Type Mini B USB on the Phidgets end to a Type A USB on the Raspberry PI end.

The InfraRed control is provided by a Phidgets model 1055_0 (http://www.phidgets.com/products.php?category=27&product_id=1055_0) that is connected via USB to the Raspberry PI. The cable is a Type Mini B USB on the Phidgets end to a Type A USB on the Raspberry PI end.

The Raspberry PI is connected to a LAN and onward to the internet via a USB connected Wifi dongle (an example: http://www.buyapi.ca/product/gold-edition-raspberry-pi-mini-usb-wireless-network-card-ep-n8508gs/). The dongle occupies a Type A USB on the Raspberry PI.

The volume control is provided by a Griffin PowerMate USB (http://store.griffintechnology.com/laptops/powermate) which has an integrated cable that ends in a Type A USB on the Raspberry PI. Alternate rotary encoders are being explored that provide better physical feel and better mounting possibilities.

All four available type A USB ports on the Raspberry PI are therefore occupied (by IR blaster, WIFI, volume, button interface). Further USB based functionality will require either a USB hub or moving from WIFI to the use of the wired Ethernet port. Some I/O expansion is possible via the Phidgets 1018 interface device which already connects the Buttons.

The headphone jack on the Raspberry PI is connected to a pair of mains powered speakers. The system could use any powered speakers, the Logitech S-120 2.0 Speakers for their small size so that they can be mounted inside the shoe box sized enclosure.
http://www.memoryexpress.com/Products/MX19568

Prototype Hardware

The prototype hardware only had four buttons, as the decision to add a fifth was taken after this prototype was developed. The "Plate" that the buttons are mounted to is a wood fiber cutting board. The "Shoe Box" enclosure is a plastic kitchen basket. Add some twist ties, tie wraps, velro, cable tie downs and fours small bolts and you have it.

Figure 1, Intended Orientation, facing TV

Figure 2, Top view Showing 4 Buttons and Volume Control

Figure 3, ISO View of System

Figure 4, View Inside "Shoebox" from below

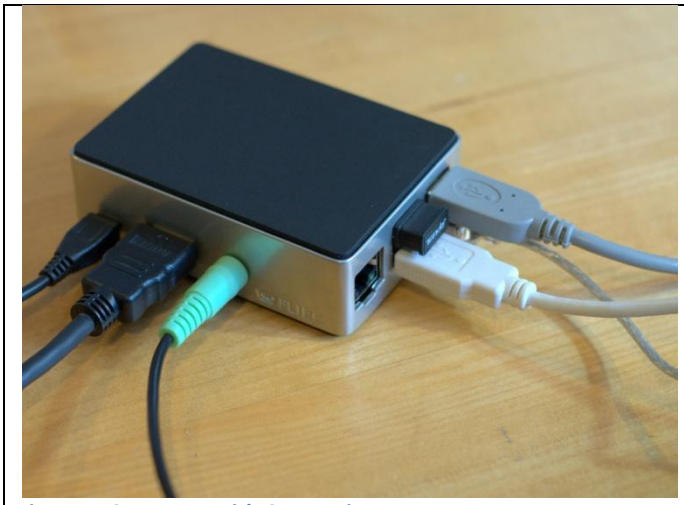

Figure 5, Computer with Connections

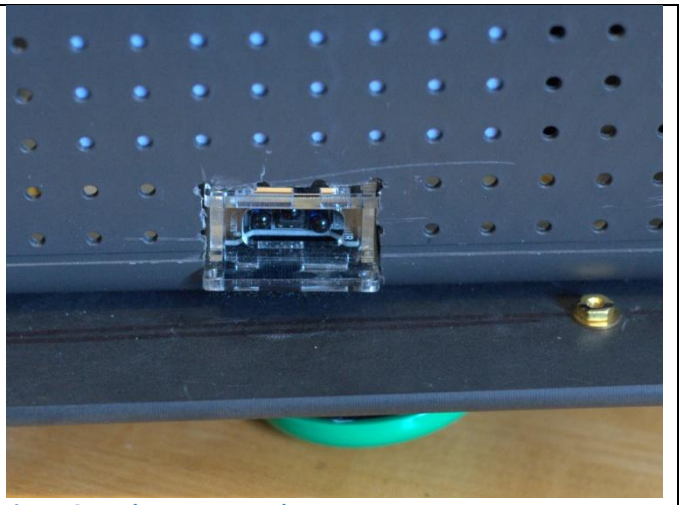

Figure 6, IR Blaster Mounted

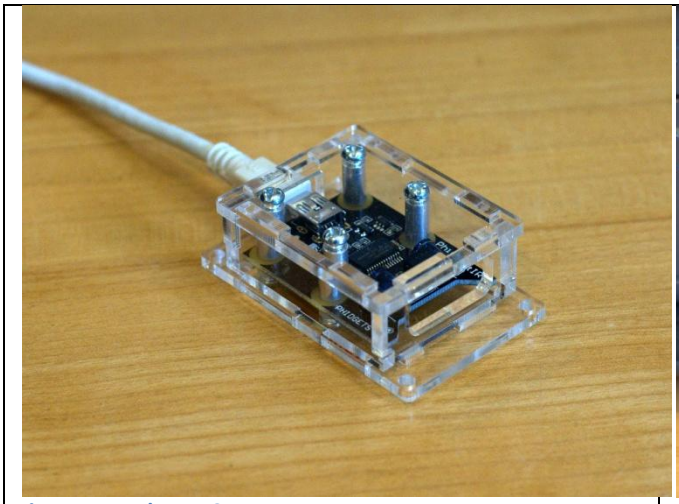

Figure 7, IR Blaster Component

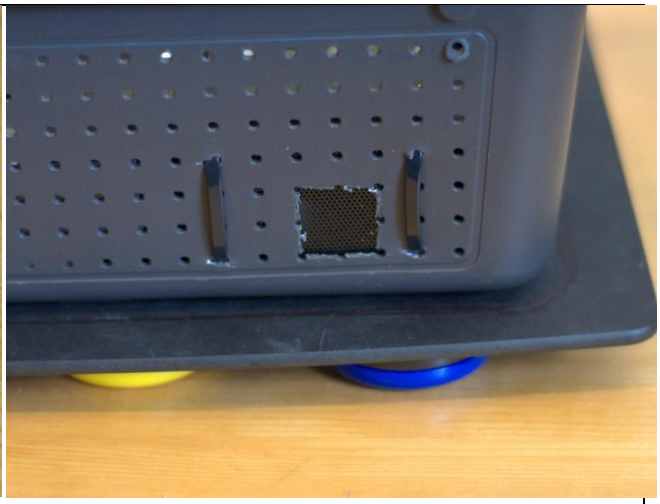

Figure 8, Speaker Grill

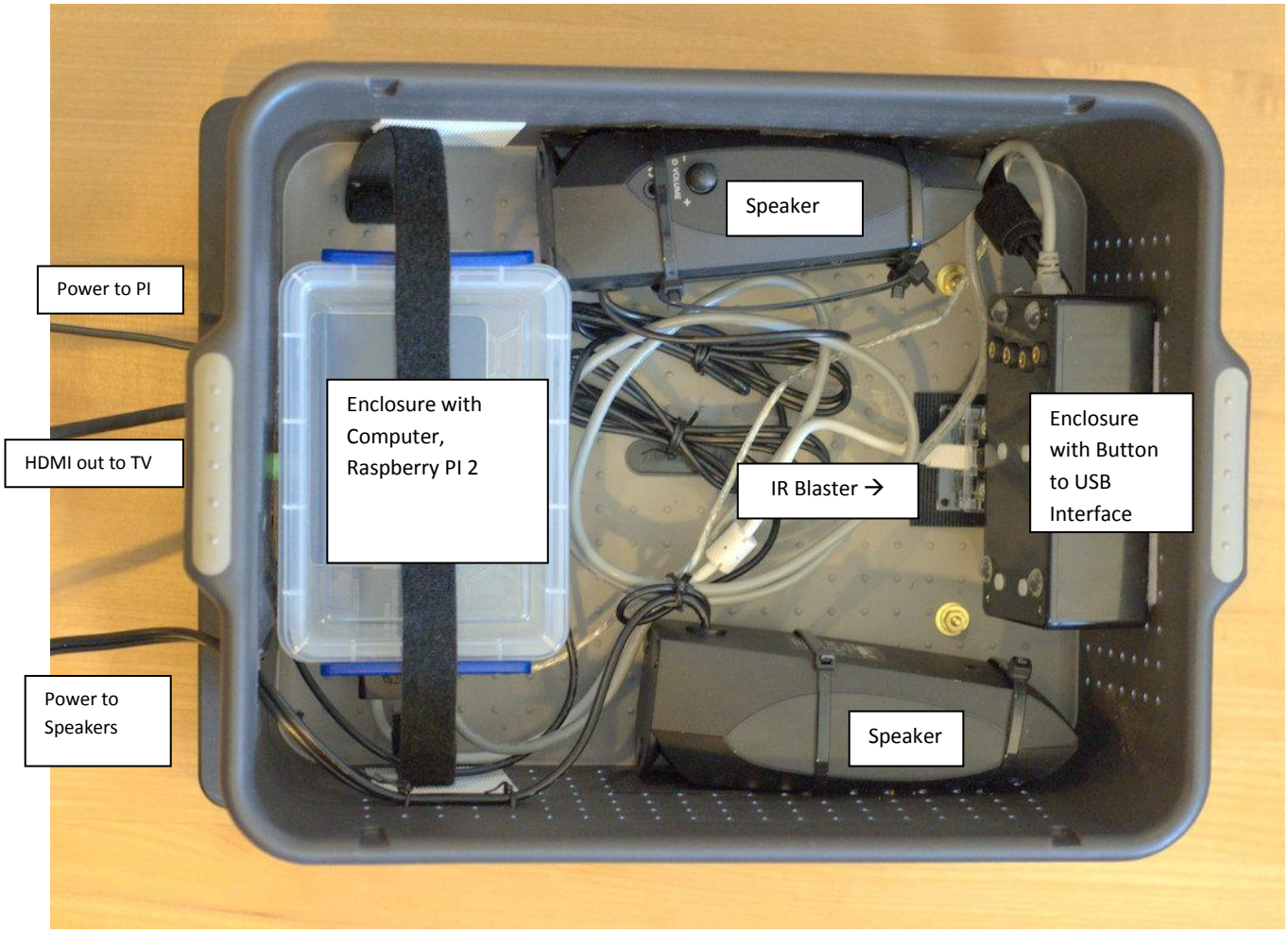

Figure 9, Label Mounted Components viewed from Below

Comments and Suggestions

Comments and suggestions for this system are very welcome. I'm also happy to answer any questions you might have. Please contact me, Alan McNaughton, on my cell at 1 403 202-4747 or by email at alan@mcnaughton.org

Licensing of this document