Soundly Located: Programmable Sounds to Assist Localization in Visual Impaired Sport

Surya P. N. Singh[1], Hanna Kurniawati[1], Paul E. I. Pounds[1], Craig Freakley[1], Wayne Wilson[2], and Louise Arvier[3]

Abstract— **The Interactive Ball ("I-Ball") is a programmable tonal soccer ball equipped with a speaker and microphone. Compared to rattle or beeper-based tonal balls, it allows for arbitrary and sensor-adaptive outputs. As a sporting aid for low-vision, this makes participation in team sports more accessible. By being programmable, the participation can be tailored to the individual. It then begs the question of "what tone to play?" This paper presents an exploitative evaluation of tone bandwidth (i.e., a narrow-band "ping" or a broad-band "rustle" sound) and tone volume (quiet to load) in environments with increasing with levels of white background noise with a focus on the ball (source) localization. This was evaluated with blindfolded subjects in an acoustic source localization task (audible angle perception). Initial show that while the narrow-band tone had better (more accurate angle localization) performance in some select cases, the broad-band "rustle" sound had better performance in general, especially in high noise, which is in keeping with a cross-correlation mechanism for localization. Output levels are more nuanced than direct signal-to-noise ratio (SNR) maximization. They show that louder sounds (higher SNR) is not always better and similarly more noise (lower SNR) is not always worse. This suggests that to aid interaction, that the nature of the tones generated by an interactive aid should factor external factors so as to better aid perceptual localization performance.**

I. Introduction

Sport is a key form of social interaction. Moreover, as the World Cup this year serves to highlight, sport is also a part of the social fabric. For those with a visual impairment, the solution has been to augment the ball with devices (such as bells or buzzers or even plastic shopping bags) that provide spatial cues via auditory signals so that the play may locate the ball and engage in game play.

Typically these tonal balls have had a fixed tone set by the chosen sound mechanism [1]. In this paper, we introduce a speaker-based programmable tonal ball, which we term the "I-Ball" or "Interactive Ball" (see also 1). This extends on previous work on adaptive (or smart) balls that modulate the tone to the rate of motion [2].

This work was supported by support from the Ian Potter Foundation, UQ Faculty of Engineering, Architecture and Information Technology Advancement Program, The Biomedical and Electromagnetics Engineering group, and Vision Australia. It was conducted at the Robotics Design Lab (RDL) at The University of Queensland (UQ) in Brisbane, Queensland, Australia, `http://robotics.uq.edu.au`

[1]S. Singh (corresponding author), H. Kurniawati, P. Pounds, and C. Freakley are with the UQ Robotics Design Lab, {`spns, hannakur, paul.pounds, c.freakley`}`@uq.edu.au`

[2]W. Wilson is with the Audiology program in the UQ School of Health and Rehabilitation Sciences, `w.wilson@uq.edu.au`

[3]Louise Arvier is a physiotherapist with Vision Australia, `Louise.Arvier@visionaustralia.org`

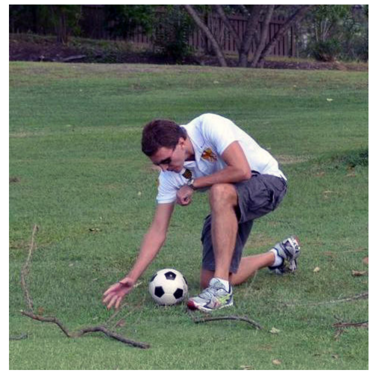


Fig. 1. A visually impairment soccer player locating the I-Ball

The inclusion of a speaker and microphone electronics, over a peizeoelectric element, allows for a diversity of tones to be played and for the tones to be adapted to motion and background noise. This poses the issue of what tones best support interaction and localization, particularly for visually impaired subjects. While the use of audio has been considered in human-computer interaction [3], this is typically associated as being complementary (as compared to replacing) the visual channel (particularly in the realm of virtual reality) [4].

This paper explores this topic by varying the type of tone and the its volume in environments with varying levels of noise. While the visually-impaired generally have heightened acuity in hearing and touch [5], this study is tested with blindfolded subjects performing an audible angle recovery task as the heightened acuity is variable and not universal and as this pilot is seeking to identify general trends for human/robot interaction.

II. Related Work

Sport is an inherently dynamic and spatial exercise. While it may be contended that vision is the spatial system par excellence, people with visual impairments are capable of acquiring a fully global conception of space [6] and may build up a set of spatial relations that are functionally equivalent to those of sighted people [7]; however, due to the limited sensory channels, they do so more slowly and by different means [8].

While a host of electronic aids has been developed [9], these have mostly concentrated on mobility, and primarily towards sensing obstacles in or near the traveler's path and include "robotic" devices such as sensorized canes [10] and range-measuring glasses [11]. Even within the realm of sports, current aids are limited. Beeper balls [12] produce

a fixed tone regardless of the balls motion and thus can lead to confusion or frustration [13]. Balls with an embedded bell [14] or those covered in a plastic bag [15] can not be heard when stationary and effect the motion and game-play of the ball as they are often rigidly attached to the ball's exterior. Despite this, there is evidence that these are less frustrating to sighted team-mates and thus allow for a richer form of interaction.

Previous work on "smart" assistive devices for people with visual impairment focused on detecting objects that may obstruct the person and providing navigation guidance [16]. Such capabilities alone are not sufficient to play soccer; players need to identify and track the ball while it is moving through the air, bouncing on the ground, etc.

There is a more global problem of absolute orientation and navigation that is separate from mobility. The wide-availability, compact size, and self-contained operation of commercial inertial sensing elements has led to diverse applications in robotics [17] and embedded systems in general. Yet, reliable in-field motion tracking of agents over extended durations (while close) is not available due to drift and model variation over time.

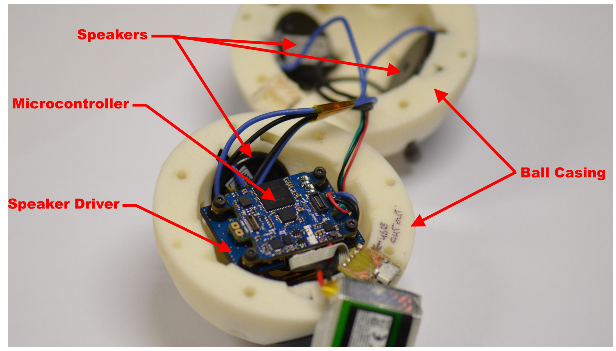
The perception (or psychoacoustics) suggest as compared to the hearing of sound suggests that sound localization of both range and azimuth (orientation) are sensitive to the frequency. For orientation there are two primary cues – Interaural Time Difference (ITD) and Interaural Level Difference (ILD) [18]. The ITD is a time difference effect and is frequency dependent – at low frequencies, where the wavelength of the sound is long relative to the head diameter, the differences in sound at the two ears is small. Thus ITD suggests a high-frequency tone. Furthermore, human hearing is more sensitive to loudness values of approximately 1000 Hz to 4000 Hz (i.e., the "isophon" or equal-loudness effect). However, producing high-frequency sounds can be shrill and unpleasant; thus, suggesting the use of ILD.

The differences in level suggest (initially) that the goal should be to produce sounds of a sufficient level as to exceed a minimum signal to noise ratio or SNR so as to have high enough SNR to be heard [19]. While this would suggest narrow-banding the tone, this comes at the risk of having a signal that might be easy to hear and difficult to perceive. Together these seem to suggest a loud, high frequency tone that maximize the SNR. However, there is a small paradox as such tones are "attention-getting" as compared to "information-giving" and are the equivalent of a bright strobe to the visually impaired. In general, the preference seems to be for "broad-band" scratching or rustling sounds as compared to "narrow-band" beeping or ping sounds [2], [20]. Thus, the most informative sound to play may not simply be the loudest one with the highest SNR.

While the ball does not move on its own, its sound does! In this case, robotics provides a basis for a programmable interface to facilitate adaptive modulation of the (sound) output so as to the encode sensory information needed to conceptualize ball position and spatial motion in a manner that is robust to noise from bystanders and the environment.

III. The Interactive Ball

An interactive ball (or "I-Ball") has been developed to assist people with visual impairment play sports (see Fig. 2). It comprises a hollow foam soccer ball and uses a micro-controller to vary the tones emitted by a speaker based on a motion sensor and signals from a microphone, which can be located with the ball or separately (e.g., as in a watch like bracelet). Since it is programmable, it provides a richer and more diverse form of interaction by allowing the tones to be changed easily.



(a) I-Ball Adaptive Tonal Interface

(b) I-Ball Remote Microphone

Fig. 2. (a) The I-Ball features a smart embedded circuit, a speaker, a rechargeable battery, wireless communications, and motion sensors + microphone to provide an adaptive tonal assistance. (b) A remote, wearable microphone board with radio communications to the I-Ball

IV. What Tone to Play?

Now that we can emit any type of sound from I-Ball, the question is what type of sound to play? This question can be further refined into what type of tone to play and how loud the tone should be.

In this preliminary work, we focus on identifying the type of tone to play, such that with minimum energy consumption, a visually impaired person is able to track an object with ease. Let's first define this objective more precisely. Since energy consumption is dominated by the amplitude of the signal, we seek the tone that maximizes the likelihood of sound source identification at low amplitude. Suppose n is ambient noise of the environment, t is the type of tone being played, A is the amplitude of the sound being played. Then, for a particular tone t and noise n, the amplitude $A_e(t, n)$ that should be emitted by I-Ball can be defined as:

$$A_e(t, n) = \text{the minimum amplitude} \quad (1)$$

$$A \in [dB_{\text{mixer min}}, dB_{\text{mixer max}}] \quad (2)$$

$$\text{subject to } P(X \mid t, A, n) > h$$

where P is the probability density function, X is a random variable indicating whether a person is able to identify the sound source or not, and h is a constant indicating the minimum acceptable level of performance. The range (dB_{mixer}) is determined by the capabilities of the computer's sound card mixer (n.b., in our case this was 50 dB of background noise at minimum mixer output (0%) and 83.5 dB of "noise" at maximum (100%) mixer output while playing white noise). The tone t_e we will use is then the tone that on average (over various types of noise) minimizes $A_e(t, n)$, i.e.,

$$t_e = \arg\min_{t \in T} \mathbf{E}[A_e(t, n)] \tag{3}$$

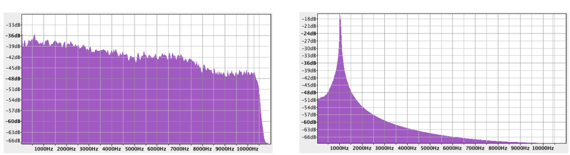
where T is the set of possible tones. $\mathbf{E}$ is the first expectation.

To determine the set of possible tones T, we use two sources. First is based on the experience of visually impaired people in using various types of rattle balls that has been used for playing soccer among the visually impaired. Second is based on signal processing. Interestingly, the type of tones that the various rattle ball emits are wide-band, which is the opposite of the good tones to overcome noise —narrow-band— according to the theory of signal processing.

V. Experiments

A. Experimental Setup

We conducted experiments with 10 volunteers to identify $A_e(t, n)$ for two types of tones. The first one is the rustle sound, denoted as t_{rustle}. It is the most commonly used sound for balls that have been specifically designed for visually impaired people. The second one is a ping tone, denoted as t_{ping}, and represents the narrow-band signal and should be easy to differentiate from noise according to signal processing. The sound spectrum of each tone is shown in Figure 3.



Rustle tone (t_{rustle}) Ping sound (t_{ping})

Fig. 3. Sound spectrum of the tested tones

For each tone, we vary the volume of I-Ball into three different levels, i.e., the baseline (threshold of hearing), medium volume (baseline +6dB), and high volume (baseline +12dB). Notice that we use relative volume with respect to each volunteer threshold of hearing, so as to minimize the variety in the volunteers' hearing capabilities. For each tone and each volume level, we asked the volunteers to identify the position of I-Ball under three different levels of background noise, i.e., no noise, low noise (66 dB SPL [Sound Pressure Level]), and high noise (80 dB SPL). In total, each volunteer were asked to identify the position of I-Ball under 18 different scenarios.

To test the ability of each volunteer in identifying the position of I-Ball under various aforementioned scenarios, each volunteer was blindfolded and stand in the middle of an empty room of size 9m×10m. The room was equipped with a Public Address (PA) sound system. This system was used to amplify ambient noise. The I-Ball was placed randomly at a distance of approximately 4m from the volunteer. An illustration of this setup is presented in Figure 4. For each tone, each volunteer was asked to identify where the I-Ball is under each of the 9 possible combinations of volume and noise levels, as described in the aforementioned paragraph (see also Figure 5). At the beginning of the experiment, each volunteer were given sample sounds of each tone they need to identify, were tested for their threshold of hearing. To minimize the effect of correct identification due to learning the position of I-Ball rather than due to perceiving the I-Ball sound, the volunteers were asked to rotate in-place several times (while being blindfolded) between each identification. Furthermore for each tone, the combination of sound volume and noise levels were presented randomly. We consider a volunteer successfully identifies the position of I-Ball under a certain scenario whenever he/she is able to identify the relative direction of I-Ball with respect to him/herself with accuracy $\pm 10^0$.

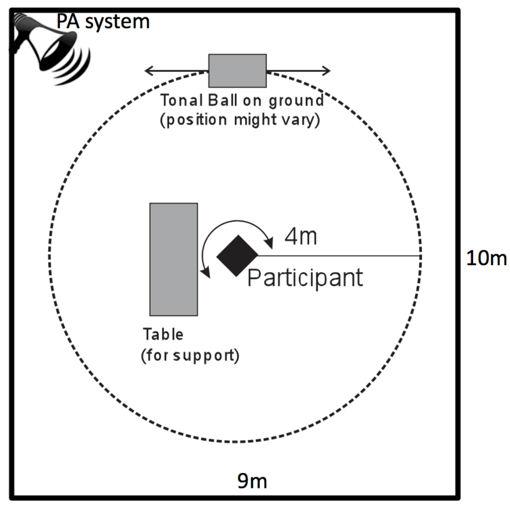


Fig. 4. Experimental setup

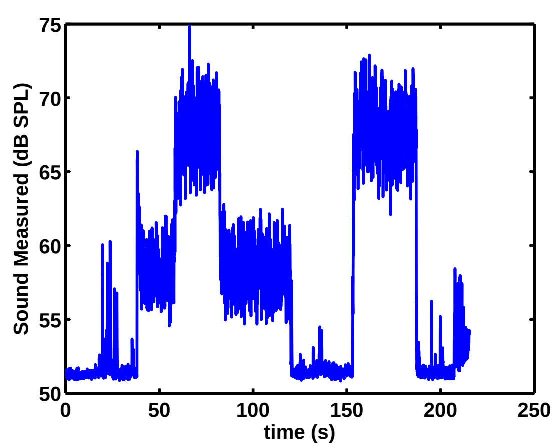


Fig. 5. A typical example of the variation in sound and noise as measured by the remote microphone located near the subject

B. Experimental Results

Table I presents the percentage of successful identification with respect to various noise levels for each tone, i.e., the marginal probability of success $\sum_A P(X \mid t, A, n)$. These results indicate that contrary to signal processing theory, at high noise level, the volunteers' performance in identifying rustle tone is significantly less than identifying the ping tone. An explanation for this is that the ping tone has a very narrow-band (see Figure 3 – right) that often requires higher human concentration to "search" for the particular signal. However, this type of signal is generally too "sharp" to the ear, that it often disturb concentration. As a result of these contradictory properties, people tend to respond better to wide-band sound signal to guide operation that requires concentration for a significant amount of time, such as playing soccer.

	No added noise (50 dB SPL)	Low noise (66 dB SPL)	High noise (80 dB SPL)
t_{rustle}	0.77	0.63	0.77
t_{ping}	0.83	0.60	0.57

TABLE I

LIKELIHOOD OF SUCCESSFUL IDENTIFICATION OF THE RUSTLE (t_{RUSTLE}) AND PING (t_{PING}) TONES (MARGINALIZED FOR THE THREE SPEAKER OUTPUT LEVELS TESTED).

	Baseline (Hearing threshold)	Medium volume (+6 dB)	High volume (+12 dB)
t_{rustle}	0.73	0.70	0.81
t_{ping}	0.57	0.80	0.61

TABLE II

LIKELIHOOD OF SUCCESSFUL IDENTIFICATION WITH INCREASING VOLUME. THE BASELINE IS SET TO THE PARTICIPANT'S THRESHOLD OF HEARING. THE RUSTLE (t_{RUSTLE}) AND PING (t_{PING}) TONES ARE MARGINALIZED FOR THE ADDED NOISE CONDITIONS (IN TABLE I).

VI. SUMMARY AND FUTURE WORK

The results shed light on the paradox of what tone to play. Clearly the sound needs to adaptive as louder is not always better and likewise more noise is not always worse. A speaker based I-Ball was developed as a programmable tonal soccer ball equipped with a speaker and microphone (that can additionally be worn remotely). As a sporting aid for children with blindness and low-vision it makes participation sport more accessible and interactive.

The results seem to confirm the view that a broad-band "rustle" sound is better than a sharp, "narrow-band" ping sound, especially in high noise environments. Again, this suggests that a more nuanced solution is needed for human-robot interaction than simply a SNR maximization.

Many avenues are possible for future work. In the immediate future, our goal is to equipped I-Ball with adaptive volume control. The idea is to enable I-Ball automatically learn the right volume for (a group) of individuals in the presence of various background noise. This is necessary to enable visually impaired children play soccer with normal children, where family and friends are free to clap and shout to support the players.

ACKNOWLEDGMENTS

The authors thank Gerrard Gosens of Vision Australia and Ben McFie and Ian Jones of Sporting Wheelies for logistical support, testing, and user feedback coordination.

REFERENCES

[1] L. Lieberman and E. McHugh, "Health-related fitness of children who are visually impaired," *Journal of Visual Impairment & Blindness (JVIB)*, vol. 95, no. 05, 2001.

[2] S. Singh, P. Pounds, and H. Kurniawati, "I-ball: A programmable sporting aid for children with a visual impairment to play soccer," in *HCI International Conference*, 2013.

[3] D. N. Zotkin and R. Duraiswami, "Signal processing for Audio HCI," in *Handbook of Signal Proc. Sys.* Springer, 2010, pp. 243–265.

[4] B. E. Riecke, A. Väljamäe, and J. Schulte-Pelkum, "Moving sounds enhance the visually-induced self-motion illusion (circular vection) in virtual reality," *ACM Transactions on Applied Perception (TAP)*, vol. 6, no. 2, p. 7, 2009.

[5] B. N. Schenkman and M. E. Nilsson, "Human echolocation: Blind and sighted persons ability to detect sounds recorded in the presence of a reflecting object," *Perception*, vol. 39, no. 4, p. 483, 2010.

[6] S. Ungar, M. Blades, C. Spencer, and K. Morsley, "Can visually impaired children use tactile maps to estimate directions?" *Journal of Visual Impairment & Blindness; Journal of Visual Impairment & Blindness*, 1994.

[7] J. Fletcher, "Spatial representation in blind children. 1: Development compared to sighted children." *Journal of Visual Impairment and Blindness*, vol. 74, no. 10, pp. 381–85, 1980.

[8] S. Millar, "Models of sensory deprivation: The nature/nurture dichotomy and spatial representation in the blind," *International Journal of Behavioral Development*, vol. 11, no. 1, pp. 69–87, 1988.

[9] J. A. Brabyn, "New developments in mobility and orientation aids for the blind," *Biomedical Engineering, IEEE Transactions on*, vol. BME-29, no. 4, pp. 285 –289, april 1982.

[10] J. Borenstein and I. Ulrich, "The guidecane-a computerized travel aid for the active guidance of blind pedestrians," in *Robotics and Automation, 1997. Proceedings., 1997 IEEE International Conference on*, vol. 2. IEEE, 1997, pp. 1283–1288.

[11] D. Dakopoulos and N. Bourbakis, "Wearable obstacle avoidance electronic travel aids for blind: a survey," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 40, no. 1, pp. 25–35, 2010.

[12] P. Ponchillia, "Accessports: A model for adapting mainstream sports activities for individuals with visual impairments," *RE: view*, vol. 27, no. 1, pp. 26–35, 1995.

[13] T. Çolak, B. Bamaç, M. Aydin, B. Meriç, and A. Ozbek, "Physical fitness levels of blind and visually impaired goalball team players," *Isokinetics and exercise science*, vol. 12, no. 4, pp. 247–252, 2004.

[14] M. Schilling, "Aids to develop throwing and catching skills," *PAM Repeater*, vol. 11, p. 5, 1982.

[15] S. Lewis and J. Tolla, "Creating and using tactile experience books for young children with visual impairments," *Teaching Exceptional Children*, vol. 35, no. 3, pp. 22–28, 2003.

[16] L. Ran, S. Helal, and S. Moore, "Drishti: An integrated indoor/outdoor blind navigation system and service," in *Proceedings of the Second IEEE International Conference on Pervasive Computing and Communications (PerCom'04)*, ser. PERCOM '04. Washington, DC, USA: IEEE Computer Society, 2004, pp. 23–.

[17] C. Ellum and N. El-Sheimy, "Inexpensive kinematic attitude determination from MEMS-based accelerometers and gps-derived accelerations," *Navigation*, vol. 49, no. 3, pp. 117–127, 2002.

[18] B. C. Moore, *An introduction to the psychology of hearing*. Brill, 2012.

[19] M. D. Good and R. H. Gilkey, "Sound localization in noise: The effect of signal to noise ratio," *The Journal of the Acoustical Society of America*, vol. 99, no. 2, pp. 1108–1117, 1996.

[20] B. Yuan and E. Folmer, "Blind hero: enabling guitar hero for the visually impaired," in *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 2008, pp. 169–176.