Modulating movement with sound: comparing a target step trained by sound or vision.

Philip Jepson

Department of Electronics, University of York, Heslington, York YO10 5DD, UK
pj10@york.ac.uk

Adar Pelah

Department of Electronics, University of York, Heslington, York YO10 5DD, UK
adar.pelah@york.ac.uk

ABSTRACT

We experimented with acoustic feedback designs to explore intuitive mappings used to guide stepping movements of walking in healthy subjects. The accuracy with which subjects reproduced a target movement based on alternative pre-programmed movement to sound mappings was compared. The most effective mapping produced a mean percentage error in either step distance or duration of approximately 12%. In additional experiments, we measured the relative accuracy of such movements based on visual feedback as compared to auditory feedback. Mean percentage errors (±std) in visual and auditory feedback cues were: Vision 11.5 ± 7.0; Audition 12.9 ± 11.8. Results indicate that, when played to their respective sensory strengths, visual and auditory feedback can be comparably effective in entraining for stepping movements.

1. INTRODUCTION

Disruption to normal multisensory and motor processes following a neurological accident, such as a stroke, may lead to sensory and physical impairment including the loss of normal proprioceptive feedback from muscles and joints. Rehabilitation can be aided through augmented visual and/or auditory biofeedback that stimulates neuro-plasticity [1], but the design and effectiveness of feedback, particularly in the auditory domain, are non-trivial issues.

Design of acoustic feedback to help in the rehabilitation of gait and movement in general could take several forms. Rhythmic cuing is effective in improving gait training for stroke patients [2], [3] to help synchronise steps and improve frequency. Another approach is to provide acoustic spatial clues to help navigate 3-D environments [4] or distance to objects in obstacle based tasks [5]. Shea et al [6] showed that acoustic models could be effective in training timed key-pressing tasks, compared with a group that received only visual feedback. Ghez et al. [7] designed auditory feedback based on musical principles to investigate whether sonifying hand position or joint motion could provide a substitute for the loss of natural feedback from joint and muscle receptors. An initial spatial target matching exercise found that trying to use acoustic feedback to provide accurate location information for the hand could be learnt but required training and did not achieve the accuracy of vision. It was noted that visual feedback can only partially compensate for proprioceptive loss and, while the primary source for positional information, it is not as effective at capturing dynamic information such as velocity and acceleration. These, it was suggested, might be expressed more effectively with sound because of differences in neural processing between vision and audition - speed of processing for example - and its role in movement timing and synchronization for activities such as music and dance.

1.1. StroMoHab

StroMoHab (Stroke Mobility Rehabilitation), was developed as an aid for gait rehabilitation following stroke. At its core is an OptiTrack passive optical motion capture system (eight V100 cameras) used to track the position of reflective markers attached to a subject. A treadmill sited in front of the cameras allows a user's movement to be captured while walking on the spot. Feedback of the user's movements is by way of an avatar displayed on a TV monitor. The system as used in these experiments, without treadmill is shown in figure 1.

Besides providing visual feedback of actual movements, it was anticipated that the avatar could present target movements for the patient to follow. This suggested the idea of an acoustic avatar to provide complementary acoustic feedback and/or movement training cues. It was the latter the experiments set out to test.

By converting a particular movement into a structured sound, how well could a listener recreate the (unseen) movement by listening to the generated sound and then attempting to reproduce the sound?

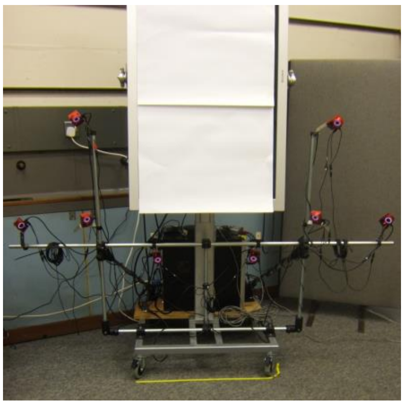

Figure 1: StroMoHab as used in the current experiments, showing motion capture cameras and screen (temporarily covered).

Initial attempts to find appropriate ways that sound might help guide movement were somewhat arbitrary. Discussion in [7] on representing velocity and acceleration in the auditory domain provided ideas for a way forward. Also, comparisons by Hunt and Kirk [8] on the pros and cons of simple one-to-one versus complex, multi-parametric mappings suggested that, for training test subjects in a novel task, a simple approach would probably work best. It was decided to take an elementary approach, examining principles by paring down the various components into simple representations. For example: using a simple gait related movement – just a single step; examining the step motion in one plane only (sagittal), tracked using a single toe mounted marker; and using a small subset of parameters for motion (velocity and acceleration) and sound (gain and pitch).

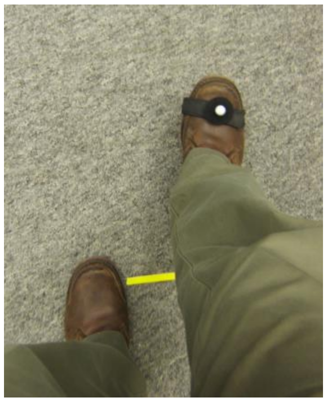

Figure 2: Example of a single step and marker attached to foot

2. COMPARING THE EFFECTIVENESS OF ALTERNATIVE MAPPINGS OF STEP VELOCITY AND ACCELERATION TO GAIN AND PITCH FOR USE IN A SIMPLE ACOUSTIC MODEL OF A STEP

2.1. Methods

Alternative ways of mapping movement to sound were compared using a training paradigm. Participants (n=7) were asked to listen to a model sound played over headphones and then attempt to recreate, as closely as possible, the sound they had heard by making a simple step movement. The model sounds were generated by mapping the velocity and/or acceleration of an emulated step to the amplitude and/or frequency of a sine wave. The distance and duration of each step attempt was recorded for comparison to the distance and duration of the emulated step. If they succeeded in recreating the sound then the distance and duration of the step taken matches the emulated step.

A target step involved keeping one foot rooted and moving the other forwards [Figure 2]. When looking at the motion only in the sagittal plane i.e the forwards motion, the pattern of acceleration is approximately sinusoidal. A sine wave of peak 2.5 m/s^2 and period 0.9 seconds was used as the basis of the emulated step. Position data was calculated from the acceleration so that in generating the model sounds the data would follow the same processing path as motion capture data generated by the participants. The length of the resulting emulated step was 0.322 meters.

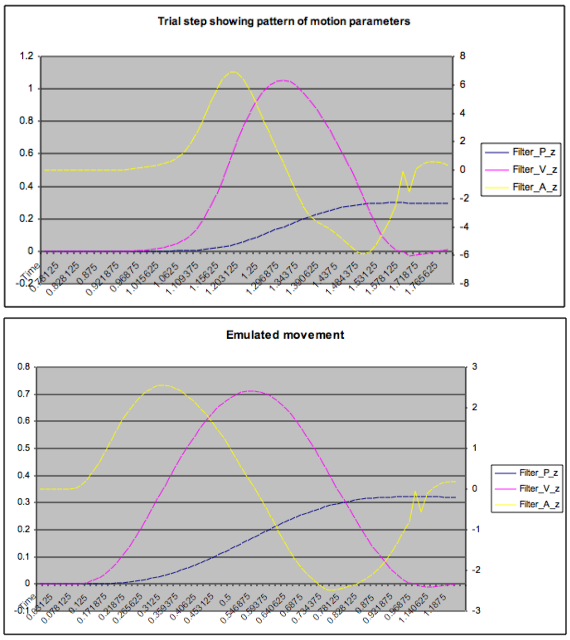

Figure 3: Motion of real step recorded as basis for emulation model (top); and bottom, the corresponding traces for the emulated step.

Velocity and acceleration were calculated by numerical differentiation of position data for both the motion capture and emulated data [Figure 3].

Position data were sampled by motion capture at 64 Hz and fed through a 4th order Butterworth IIR filter prior to numerically deriving velocity and acceleration. Values of velocity and acceleration were used to change gain and pitch parameters in OpenAL (Open Audio Library Application Programming Interface) which was used to render sound through a PC sound card. Gain and pitch are adjusted using a float value:

- gain = 0 (source is muted) to 1.0 (max volume permitted)

- gain = 0.5 (6dB attenuation)
- pitch = 1.0 (base frequency)
- doubling the pitch value raises the frequency by an octave, halving the pitch value lowers the frequency by an octave.
- the source was a sine-wave oscillator, base frequency = 250Hz, with a peak equal to the maximum PCM value (±32767, 16-bit mono samples)

Five different model sounds were created by mapping the velocity and/or acceleration of an emulated step to amplitude and/or frequency of the 250 Hz sine wave:

1. velocity to gain
2. velocity to gain and pitch
3. acceleration to gain
4. acceleration to gain and pitch
5. velocity to gain
6. velocity to gain, acceleration to pitch

These gain and pitch values were calculated for each of the mappings as follows:

1. gain = velocity (i.e. no scaling)
2. gain = velocity; pitch = velocity + 0.5 (i.e. f = 125 Hz as the foot just starts to move)
3. gain = |acceleration|*0.25
4. gain = |acceleration|*0.25; pitch = (acceleration*0.25) + 1.5
5. gain = velocity
6. gain = velocity; pitch = (acceleration * 0.25) + 1.5

Participants were presented with six model sounds (velocity to gain was repeated). Each sound was listened to twice before five 'on the spot' steps were made whilst listening to concurrent acoustic feedback through the headphones. This *two-five* pattern was repeated 10-times for each of the model sounds, by each participant. They were not given any information about the basis of the sounds and were simply presented with sounds and asked to recreate them by making a step-like movement.

2.2. Results

Statistical significance was not achieved with the small sample size. None of the motion to sound mappings tried showed any particular advantage in helping to reproduce the distance of the step. However, consistent with the temporal nature of sound, some of the mappings appeared to have the potential to modulate step duration more than others with velocity to gain appearing the best [Figure 4].

One subject appeared to find the tests difficult, which is borne out by looking at normalised scatter plots comparing performances across subjects where subject C appears as an outlier in both temporal and spatial domains [Figure 5]. Removing subject C reduces overall errors across all tests by around 30 per cent.

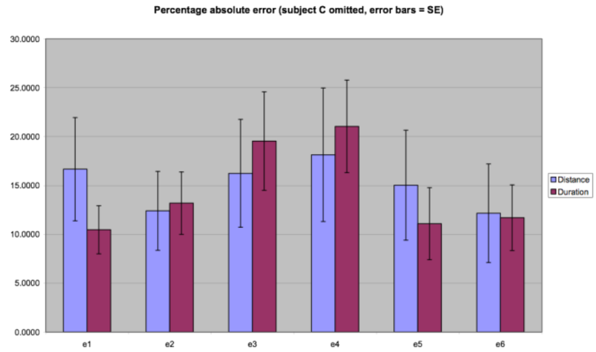

Figure 4: Bar chart showing percentage absolute errors for step distance and duration achieved using alternative mappings

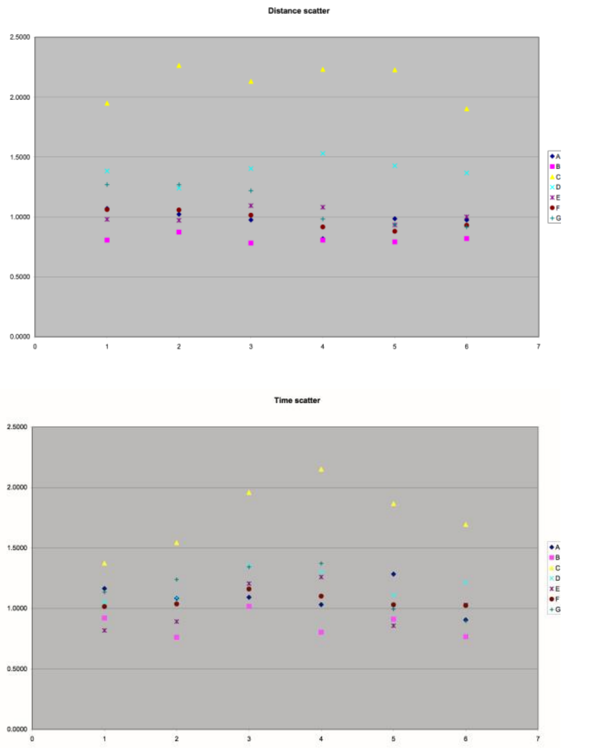

Figure 5: Normalised scatter plots for distance and duration illustrating outlier subject C.

The pilot experiment was run to compare alternative mappings to provide a guide to which might be useful to work on. As such a control was not included, so it is not possible to judge how good performance at the task was and how much benefit acoustic feedback provided.

A second experiment was therefore designed in order to provide some measure of the effectiveness of acoustic feedback as a training tool.

3. COMPARING AUDITORY/TEMPORAL WITH VISUAL/SPATIAL FEEDBACK BASED TRAINING

3.1. Method

Again, a single step was used as the target movement: keeping one foot rooted; moving the other forward; ending the step with the foot flat before returning to the start position. Each test subject (n = 20) was asked to make 20 such movements, one after the other, and to try to keep the steps consistent. The duration and distance of the forward part of the step were recorded using motion capture and formed a baseline measurement of the subject's preferred step distance and duration.

The volunteers then trained either a target step length or a step duration, which consisted of taking 20 steps based on their baseline measurement plus or minus 25 per cent. Whether subjects trained to a lengthened or shortened step measurement was done at random.

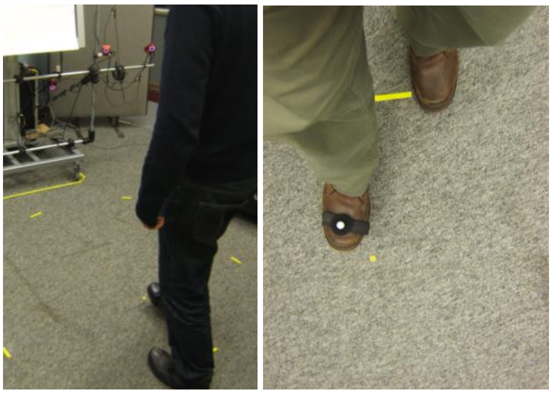

Figure 6: Visual/spatial step training using a target marked on the floor

Step length training used measured visual marks on the floor (Figure 6).

The auditory feedback for duration training was designed to provide roughly equivalent cues available implicitly in the visual condition (Figure 7). It used a tone (950 Hz sine wave) based on the velocity-to-gain mapping used in the first set of tests (but of fixed length equal to the target duration). It started playing when the subject started to move, with the bell-shaped envelop providing an indication of the progression of the target step. This approximated to being able to see the whole task in the visual condition. The visual training also provided implicit feedback of results - participants could judge how close to the target they were for every step. This was mirrored in the auditory domain by inserting 'clicks' - one when the target duration had elapsed and another when the participant had finished the step. Coincident clicks indicated a step of the correct duration.

The training phase was immediately followed by the attempt to replicate the target step (with no visual or auditory cues) again, with 20 step attempts.

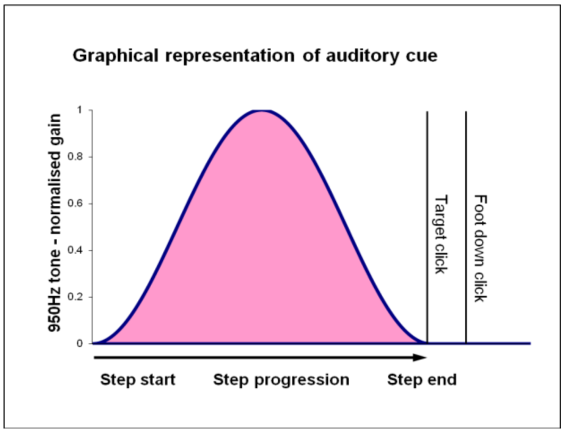

Figure 7: Graphical representation of the auditory training cue

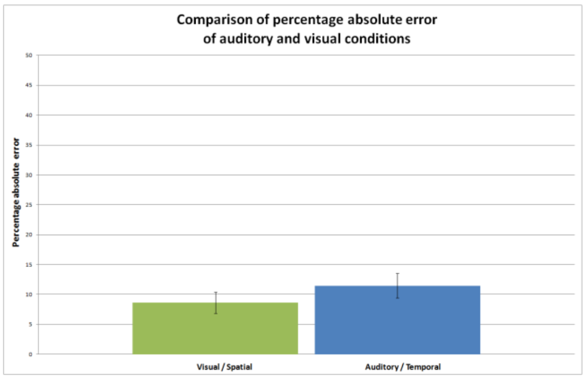

Figure 8: Comparison of mean absolute errors

3.2. Results

A graphic summarizing the results is shown in Figure 8.

Auditory temporal cues and feedback are at least half as good as visual spatial cues for a simple step-like task. The ratio of mean error rate, spatial/temporal, = 0.75. 95% confidence interval = 0.51 to 1.11, P = 0.15.

Hence there is no evidence that the spatial and temporal differ, and the important thing is that the ratio of the means is between 0.5 and 1.1.

3.3. Implications

Motor learning is often considered to be dominated by visual feedback [9] during self-guided movements in combination with proprioception. Contrary to expectations, the result of the experiment suggests that under comparable conditions auditory feedback can be at least half as effective, in terms of percentage error, in the temporal domain as visual feedback is in the spatial domain. This offers the possibility that auditory cues to motor learning can be effective to enhance rehabilitation in mobility in addition to visual feedback, or in situations where visual inputs may be limited (such as stepping under foot), restricted or entirely absent (such as for the visually impaired). The comparable levels of error between the visual (spatial) and auditory (temporal) modalities may indicate a shared or similar neural mechanism, or alternatively that motor learning in the two conditions may be bounded by similar sets of constraints.

4. REFERENCES

[1] Holden MK, "Virtual environments for motor rehabilitation: review", Cyberpsychol Behav. 2005 Jun;8(3):187-211; discussion 212-9

[2] Prassas S., Thaut M., McIntosh G., Rice, R. (1997). Effect of auditory rhythmic cueing on gait kinematic parameters of stroke patients. Gait and Posture, 6, 218-223

[3] Thaut MH1, Leins AK, Rice RR, Argstatter H, Kenyon GP, McIntosh GC, Bolay HV, Fetter M., "Rhythmic auditory stimulation improves gait more than NDT/Bobath training in near-ambulatory patients early poststroke: a single-blind, randomized trial", Neurorehabil Neural Repair. 2007 Sep-Oct;21(5):455-9. Epub 2007 Apr 10.

[4] Tapio Lokki, Matti Gr ohn, "Navigation with Auditory Cues in a Virtual Environment", IEEE MultiMedia, vol. 12, no. 2, pp. 80-86, Apr. 2005, doi:10.1109/MMUL.2005.33

[5] Mathias Wellner, Audrey Schaufelberger, Joachim v. Zitzewitz, Robert Reiner. "Evaluation of visual and auditory feedback in virtual obstacle walking". Presence, Vol. 17, No. 5, October 2008, 512-524

[6] Shea CH, Wulf G, Park JH, Gaunt B. "Effects of an auditory model on the learning of relative and absolute timing". J Mot Behav. 2001 Jun;33(2):127-38.

[7] Claude Ghez, Thanassis Rikakis, R. Luke DuBois, Perry R. Cook, "An auditory display for aiding interjoint coordination", Proc. Int. Conf. On Auditory Displays (ICAD) Atlanta, Georgia, April 2000.

[8] Andy Hunt and Ross Kirk.(2000), "Mapping Strategies for Musical Performance". In: M.M. Wanderley and M. Battier, (eds.) Trends in Gestural Control of Music, IRCAM - Centre Pompidou

[9] Michael I. Posner, Mary Jo Nissen and Raymond M. Klein, "Visual Dominance: An information-Processing Account of Its Origins and Significance", Psychological Review, Volume 83, No2, Pages 157-171