DESIGNING AN INTERACTIVE SONIC MEDIA ENVIRONMENT FOR STROKE REHABILITATION

Diana Siwiak
Victoria University of Wellington
New Zealand School of Music

ABSTRACT

Immersive media environments assist in solving the real-world issues of presenting multimodal information and of responding to user reactions, such as with patients during stroke rehabilitation therapy. Embedding media in the experience creates a new perspective for both the therapist and the patient. It is in this perspective that creative implementation of the arts within technology flourishes. A sonic media environment designed for stroke therapy provides the patient with the ability to produce an accurate trajectory. Based on self-assessment, focus is on the linearity of movement and speed is not crucial. Gradually reducing the amount of real-time media information and promoting a more constructivist learning approach towards therapy, is a cornerstone of this paper.

1. INTRODUCTION

Media arts is an effective medium to assist users, such as school children or rehabilitation patients, in devising individual learning strategies for themselves using constructivist learning techniques, see (Duff *et al*. 2010; Sundaram & Rikakis 2006). These are two of the many approaches for communicating this concept. In an adaptive, mixed reality rehabilitation (AMRR) system, each level of feedback is designed to communicate the appropriate type of movement performance information, while maintaining engagement and understanding through an unfolding, interactive narrative. This knowledge blossoms by synthesizing perception, modeling, interaction, feedback, and sensing into a precariously balanced, mutually beneficial relationship.

This paper discusses the iterative, rapid prototype design of a real-time, customizable generative sound engine and toolkit, including the extensive prior work that fueled this research. As part of a three-year project, the research team set out to prove the meaning, which arises from the interrelationships, connects action to music and to propose changes to a developed system. The generative sound engine, as part of a more complex, multi-modal system architecture, provides an interactive, multi-temporal sonic media environment (SME) for patients recovering from movement restrictions present after stroke. However, only the audio component of the system is addressed. The toolkit includes adaptable features to make adjustments to the soundscape in real-time, thus affording possibilities for a plethora of intricate sound designs.

2. PRIOR WORK

Current research within the field of rehabilitation investigates the best therapeutic strategies for stroke rehabilitation. This includes the use of virtual and augmented reality in order to leverage sensory-motor integration for interactive learning. Embedding media in the experience creates a whole new perspective for both the therapist and the patient. The Mixed Reality Rehabilitation (MRR) Research Lab has been in the process of developing an adaptive, mixed reality rehabilitation (AMRR) system for stroke survivors recovering from right-handed hemi-paresis since 2006. The introductory system is embedded in a hospital. Comparative trials of 11 patients have already been completed by the lab and are currently under assessment by graduate students in the lab. Preliminary results prove the expected outcome, which is an improvement in movement performance through mediated physical therapy versus traditional physical therapy.

Experiential media systems explore the relationship between digital technologies, which can sense human action, and cultural processes, whereby human actions transform media information through the use of these digital technologies. The MRR Research Lab has shown that experiential media systems, or interactive environments, can be used to provide effective feedback and feed-forward information and to promote positive motor learning skills, including retaining and applying these skills. (Duff *et al*. 2010) The immersive SME (in this paper, feed-back and feed-forward information with respect towards audio) is connected to both observable and unobservable movement activities. These encourage active physical and cognitive participation, with the end goal being the learning of generalizable movement strategies. The movement activities and media environments created here are adaptable to a patient's individual ability and progress. This process allows the patient to be challenged physically and cognitively without frustration. The following two sub-sections have been heavily adapted, edited, summarized, or directly quoted from seminal manuscripts written by the MRR Research Lab, in order to lay substantial groundwork for this research.

2.1. Real-Time Sonic Media Environment

An interactive SME (designed to imitate a form of structured music) is a powerful tool for movement training, particularly utilizing its temporal aspects to assist in pacing the speed of the movement activity. (Jaques-Dalcroze 2007) discusses how music helps the brain connect body, space, and action in an intuitive manner. (Dourish 2004) notes that multimodal media compositions, such as films, rely on music to communicate implicit messages and emotional states. The same methodologies are employed in this system in order to encourage natural interaction with the patient. The harmonic progressions commonly found in popular tonal music can encourage a feeling of forward movement, giving a natural affordance for use within an interactive training environment or rehabilitation system to convey positive feedback.

The first version of the interactive sound engine, embedded within the hospital scenario, capitalizes on the inherent powers of music, such as harmonic structure and rhythm, by mapping movement activities onto the features of an abstract sonic environment. "The resulting immersive environment has two main goals: a) to encourage the patient to perform the required movement activity and b) to offer to the patient an intuitive way to self-assess their movement performance, understand the cause of error, and develop an improvement strategy," (Duff *et al*. 2010). Spatial and temporal information are not mutually exclusive. However, spatial thresholds hold the key to the harmonic framework of chord progressions. Temporal thresholds predominantly drive the rhythmic framework of note densities.

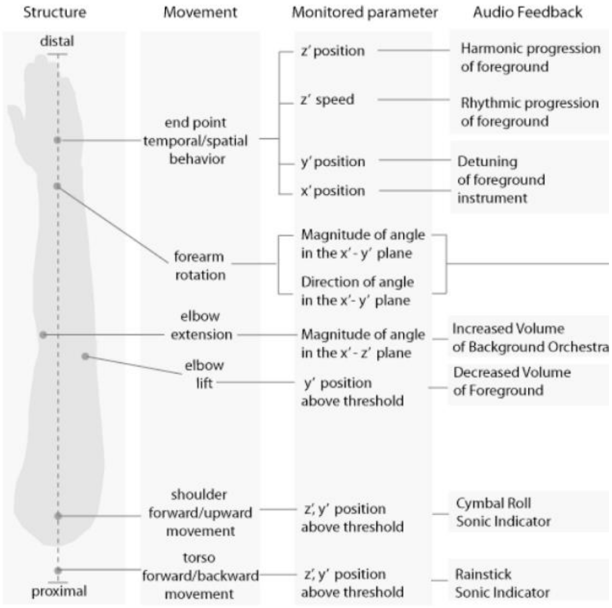

Figure 1. Parallel movement components and feedback mapping

Figure 1, depicts how various movements, monitored parameters, and media environments (both audio and visual) co-exist with one another in the mixed reality rehabilitation system. The patient's arm is laden with several passive motion capture rigid bodies.

Eleven infrared cameras capture these rigid bodies in real-time. The data from the movement of these rigid bodies within the capture space is sent to a processing program. The processing program then provides information to determine various parameters, such as position data and spatial information. The parameters are managed and distributed to the sonic media environments, which create an immersive and interactive media space. The real-time media environment provides explicit and implicit information from which the patient can self-assess his or her movement strategies. It is through this self-assessment that the patient can plan his or her next movement activity and make necessary improvements, creating a cohesive media environment.

This real-time sonic media environment provides prescriptive knowledge about the movement performance. In order to accomplish a quick learning curve towards interpreting the media environments, the audio scene does not change instruments nor chordal progressions, but does sustain a structured composition. This rigid sonic media environment provides an opportunity to expand the possibilities of sound designs, such as in the manipulation of an ocean soundscape, where the patient is affecting the density of waves or bubbles, as opposed to note densities.

2.1.1. Harmonic Content

In the hospital system, harmonic progressions sonify real-time movement activities through both temporal and spatial information. A key signature is stochastically chosen from either a current chord or from one of the adjacent chords within the circle of fifths. This creates sequences of related progressions with smooth key changes between movement activities. At the same time, one of three possible chord progressions, all of which are common harmonic cadences, is randomly chosen. Each progression has four chords and the chords are mapped to thresholds in space, as shown in **Table 1**. Each note is selected by pulling pitches from the chosen chord. The intent is that the patient be intuitively motivated to complete the reaching movement activity so he or she can hear the entire musical phrase performed to completion. It is through this initial basis, set forth by the hospital-based system, that the home-based system (Siwiak *et al*. 2011) was provided with the opportunity to build upon using harmonic content to sonify temporal and spatial information.

Table 1. Harmonic progressions & spatial thresholds

State (normalized velocity)	Prog 1	Prog 2	Prog 3
Resting (0.0-0.1)	I	I	I
Reach (0.1-0.9)	V7/IV (I7)	vi	ii
Grasp (0.9-1.0)	IV	ii	V7
Return (0.9-0.1)	V7	V7	IV

2.1.1. *Rhythmic Content*

Musical rhythm sonifies the speed of the movement activity performed by the patient. The notes per beat progress from top to bottom through the durations depicted in **Table 2**. As the patient's movement becomes faster, the notes per beat increase. The mapping of spatial information to notes per beat is scaled by a desired maximum movement speed, predetermined by the therapist. If this maximum speed is reached, the notes per beat peak, with individual notes merging into a continuous stream of seemingly obscurely defined notes.

This change from individual notes to a streaming series provides an implicit notification for the desired maximum speed. Through these mappings, the patient is encouraged to perform a smooth acceleration and deceleration pattern. The peak speed occurs in the middle of the reach, so that a smooth rhythmic pattern is heard. The rhythmic information conveyed in the hospital-based system laid further groundwork for iterative design within the home-based system. (Siwiak *et al.* 2011) The patient focuses on successfully completing the movement activity while the music implicitly trains the timing structure.

Table 2. Rhythmic patterns & temporal thresholds

% of velocity with respect to desired velocity peak	Foreground note subdivision (notes per beat)
0-20% (resting)	Quarter notes (1.0)
20-40%	Triplet eighth notes (1.5)
40-60%	Eighth notes (2.0)
60-80%	Quintuplet eighth notes (2.5)
80-100% (streaming)	Triplet sixteenth notes (3.0)

2.1.1. *Musical Context*

The musical progressions present in the hospital system function to assist in learning a timing structure to help the patient achieve ideal results (i.e. proper and consistent movement performance). The interconnection between training musical rhythm and training movement in time and space has been effectively utilized under the Dalcroze method and other music pedagogy methods (e.g. Kodaly, Orff). It is important to note that, within these pedagogies, the learning of movement and music is subconscious. The complex synchronies and phrasing structures are learned intuitively. Because of this intuitive connection of music to movement activity training, SMEs have been used for several therapeutic purposes, (Ghez *et al.* 2000; Headlee, Koziupa, & Siwiak 2010; Siwiak, Berger & Yang 2009).

2.1.2. *Emotional and Cognitive Connections*

A patient's emotions play an important role in the success of his or her stroke recovery. Feelings of tension can exacerbate other physiological symptoms and the side effects of stroke, such as tremor or compensation techniques. For this reason, these MRR systems are designed to provide a calming SME and a continued motivation throughout the movement activity. (Gabrielsson & Lindstrom 2001; Seitz 2005)

As described in (Ghez *et al*. 2000), the brain maintains a strong memory for musical constructs. The patient connects successful movement activities with pleasing musical phrases. The memory of the pleasant musical phrase is then used to plan future successful movements. The majority of this type of music-assisted movement activity learning happens subconsciously, similar to the intuitive learning that happens during dance. (Prinz 1990, 1997)

2.2. AMRR: Integrative, Interactive Learning

Abstract feedback mapped to movement activity re-contextualizes a performance of the movement activity into a performance of an interactive SME The patient receives the knowledge of his or her performance and the results of his or her movement activities in terms of his or her interaction with the SME Within this context, the patient is able to intuitively develop evaluative measures of his or her performance on key aspects of movement activities in the SME While the patient is learning to control the SME, he or she is intuitively developing more efficient movements.

Another effective aspect of using abstract feedback is the promotion of active learning. Abstract feedback requires the patient to determine the causality between components of movement activities and individual attributes of the SME The patient must connect how the overall form of the SME reflects the overall form of the movement activities. The metaphorical reconstruction between movement activity and media environment requires active engagement, and eventually results in parallel cognitive and motor learning by the patient. This approach to learning is at the core of constructivist learning theories. (Cirstea & Levin 2007; Gardner 1999; Papert 1980)

By taking a constructivist approach to the system design, these methods decrease the likelihood that the patient develops dependencies characteristic of prescriptive learning, which may not allow for generalization or maintenance past the therapy session. (Schmidt 1991) Therefore, abstract media mappings are consistently applied to key aspects of movement across a variety of movement activities.

3. CURRENT WORK

The adaptive, home-based, interactive sonic media environment (SME) evolved based on the prior work that was completed and assessed in a more comprehensive version of the software, the hospital-based system. As the home-based system strives to be the next generation of an interactive multimodal rehabilitation system, its design, goals, and implementation begin to grow. The goals of this new SME are to 1) reduce the amount of detailed information given in real-time as the movement activity is happening, 2) provide summaries of performance on a per movement activity level and across groups of movement activities, 3) provide a media environment that correlates to a reconfigurable space, and 4) begin to integrate media environments into the patient's physical space. In other words, this new, simpler rehabilitation system endeavors to focus on economy, scalability, and a more long-term engagement. The innovative approach towards gradually reducing the amount of real-time media information, therefore promoting a more constructivist learning approach towards therapy, is the cornerstone of the home-based system when compared to the hospital-based system. (Siwiak *et al.* 2011)

The ecological coupling (Sundaram & Rikakis 2006) between a patient's movement speed and the SME relies on rhythm, so therefore the system maps hand speed to density of notes. Burkholder clarifies this in [**Figure 1**]. During the reaching task, the harmonic intervals and attributes of the sound design are altered accordingly, based on movement performance. The movement activity becomes more consistent in speed as a consequence of the patient retaining motor learning skills by practicing movements. The soundscape displays encouraging and peaceful music. The new sounds designed for the home-based system have been created to allow for changing soundscapes without changing the semantic meaning they convey. The patient does not have to repetitively listen to the same droning sound throughout the training period, but can, instead, interact with waves in an ocean scene or with fallen leaves in a wooded scene. Sound suites such as these each convey similar semantics and can thus relieve the monotony for the patient during therapy.

The research on audio design for the home-based system works towards a more predictive (or feed-forward) model, rather than a reactive (or feedback) model. This is because the reaction to a triggered media stream, such as an audio cue, has inherent delay. This could cause confusion during real-time movement activities if the delay happens to be substantial. This research uses a probabilistic way of analyzing and predicting action to iterate through a predictive model for feed-forward information, thus creating an interactive SME containing both feed-forward and feedback information about movement performance.

3.1. Generative Sonic Content

In order to transition to a constructivist learning approach, the first version of the sound engine (from the hospital-based system) needed a redesign. The motion capture and analysis engines in the home-based system create a large amount of processing overhead. In order to reduce this processing, significant improvements within the interactive audio environment were made to provide a more efficient system. The processing now occurs at the level of the operating system, which has reduced the necessity for proprietary (and processor-heavy) software. It also opened the door to create custom software that is more efficient and effective. This redesign also created a more integrative interaction between the various system components.

With the background, training, and experience in digital audio signal processing and sound design, the proposed changes to the original sound engine included developing a real-time customizable, physically modeled sound design toolkit. The C++ library used to develop this toolkit is The Synthesis ToolKit in C++ (STK). "STK is a set of open source audio signal processing and algorithmic synthesis classes written in the C++ programming language. STK was designed to facilitate rapid development of music synthesis and audio processing software, with an emphasis on cross-platform functionality, real-time control, ease of use, and educational example code."[1] It is with this library through which a custom software wrapper could be implemented within the larger system architecture.

Using synthesized sounds in the home-based system allows for detailed manipulation of each sound. This action opens the auditory landscape while significantly reducing the overhead of computer processing and storage requirements when compared to the hospital-based system, which uses a dedicated sample-based software synthesizer application, Kontakt. Some popular audio processing programs such as Max/MSP, Pd, and ChucK host proprietary wrappers for STK. However these applications are all processor-heavy and have the innate characteristic of significant delay from an action to a reaction. Since the infrastructure of the home-based system is a proprietary Mac-based program, it afforded the opportunity to create an Audio Unit wrapper for the STK and a custom integrated Audio Unit host.

In order to ease the transition from the hospital-based scenario (where the primary instrument is a marimba), the home-based scenario utilized the banded waveguide modeling class included in the STK. Essl and Cook describe in their documentation that "this class uses banded waveguide techniques to model a variety of sounds, including bowed bars, glasses, and bowls." (Essl & Cook 1999) The sonic content of a plucked or bowed wooden bar is not unlike the sonic content of a marimba. The parameters and presets within the banded waveguide class are easily manipulated to create a marimba sound.

3.2. Multi-Layered Sonic Media Environment

The research team employed a different approach for designing compositions consistent with some standard movements for a stroke survivor. Once the compositions have been determined and designed, the proximate data streams are analyzed and a mapping function within this interaction space is then determined. This afforded the opportunity to be creative with the music compositions and to not be constrained by mapping functions or streaming data. A higher-level narrative emerged from this unbounded design scheme, where the various interaction layers (real-time, post-activity, and post-set) could be simultaneously individual, yet connected.

---

[1] Synthesis Toolkit. CCRMA.
https://ccrma.stanford.edu/software/stk

In the hospital-based system, there are two layers of interactive media environments: real-time and post-activity. In the home-based system, there are several layers of interactive media environments, including: real-time, post-activity, and post-set (where a set is a group of movement activities without a real-time media environment).

In this paper, the concept of dovetailed layers describes the transitions that occur between prescriptive (real-time movement performance) information and reflective (summary movement performance) information. The information contained within each layer of the SME is dovetailed in order to facilitate the patient's understanding as the system transitions between layers. Because of the flexibility of the methodology, different scenarios within these layers may be developed to suit the needs and preferences of different stroke rehabilitation patients. However, the form and content of each scenario must be able to adapt and accommodate the appropriate type of movement performance information being communicated.

3.2.1. *Real-Time Activity Media Environment*

The basic concept from the hospital-based system carries over to the home-based system where note density maps to movement activity speed and harmonic changes map to spatial information. The parameters in **Figure 1** have been reduced in the home-based system, especially since the sensing capabilities have been reduced (to encourage and economically feasible system), which means the SME contains a fraction of the original sonic content. The sound still encourages the patient to move with a smooth, natural speed, but less information is given regarding body function movements (such as elbow openness, which in **Figure 1**, is signified by orchestra volume). The main differences in the sound design are the back-end system architecture, the communication protocols between the audio engine and the system engine, and, most importantly, the soundscapes now possible with the implementation of Audio Unit wrappers for STK. To date, a handful of wrappers (one for each class in the STK) have already been designed and implemented for use within the SME for the home-based system.

3.2.2. *Post-Set Activity Media Environment*

The sound design concepts for the post-set layer build upon the real-time SME The post-set layer creates an affective summary, or sonic images. Compositional constraints based on well-known paradigms associate speed to rhythm, as previously mentioned in Section 2. The design concept applied herein ventures to progress from musical composition to data mapping to motion capture data stream, rather than the opposite approach. The scheme includes individual movement activities, where one-to-one mappings of data to sonic information is displayed, and the larger narrative picture, where comprehensive scenarios provide a less prescriptive and more constructivist perspective on media environment to movement activity - or, coupling feedback to action, which is contrary to concepts discussed in (Burkholder, Grout, & Palisca 2010).

Within each sonic image, both reactive and predictive models of performance are conveyed through dual-layered compositions, with the lower voice aurally depicting an ideal movement strategy and the upper voice aurally depicting deviation from the ideal. The ultimate goal is to be rewarded with the voices accompanying each other in a pleasing, non-dissonant composition. If after a set, the patient has consistently performed ballistic movements, the upper voice would contain short, tin-sounding notes, dissonant intervals, and fast-moving density of notes when compared to the lower voice, which would contain the sound of a gentle, rolling ocean. The two voices would not match in speed or in timbre, and the patient would then be able to assess performance and plan future tasks accordingly.

In the current version of the system, there are 4 sonic images, each with sequences containing 8 “polychords” (in this paper, polychords describe groups of notes). Polychords were designed so that there is not necessarily a prominent progression, unlike the well-known harmonic progressions in the hospital-based system. This affords the opportunity for variance without the need for immediate familiarity, in order to gather proper meaning from the SME The polychords within each sequence are paired together, providing a higher-level composition over groups of movement activities, or sets. To toggle between pairs of polychords, a spatial threshold of at least halfway must be attained. This carries over from the concepts learned from and tested within the hospital-based system. (Duff *et al.* 2010) Rather than sonifying specific and prescribed parameters over an individual movement activity, depicted in **Figure 1**, a more generalized communication of movement performance can be attained. This gives us the opportunity to decrease the amount of real-time media information, to allow the patient to practice groups of efficient movement activities, and to create an interactive media narrative.

3.2.2.1 Sonic Image Tags

An “ideal” sequence signifies a proper movement for a group of activities. As shown in **Figure 2**, this composition has consistent pedal tones (in this case, C major without the third) throughout the sequence. The upper voice of each polychord progresses from D to B*b*, then A*m* to G, then F to E*m*, then E*b* to C. The reason behind this sound design is to create a directional, yet not forceful or abrasive, progression that is pleasing and is not an explicitly known or well-established sonic memory of chordal progressions. Similar to the chordal progressions embedded within the hospital-based system, this new compositional tactic provides the feeling of forward movement. This sequence becomes a recognizable signature for the patient to understand that recent groups of movement activities exhibited ideal movement performance.

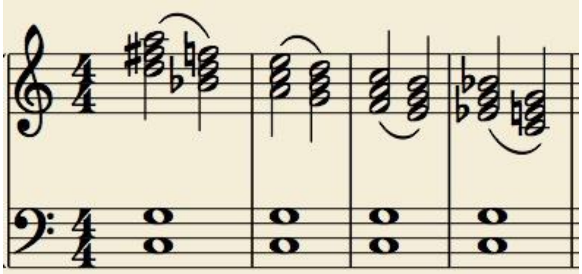

Figure 2. Composition for "Ideal" Sequence

A “too slow” sequence signifies an overall too slow movement for a group of activities. As shown in **Figure 3**, this composition has decrementing pedal tones for each pair of polychords starting at C and ending at G. Its upper voice progression is D*bm* to B*m*, C*m* to B*bm*, B*m* to A*m*, then A*bm* to G. The reason behind this sound design is to enforce that there is movement, but also to convey a tone that is “dark” and "foreboding". This gives the forewarning: “Something is not correct”. The reason behind the pedal tones in these first two sequences is to provide consistency throughout the entire progression, in order to tie the polychords together. This sequence becomes a recognizable signature for the patient to understand that recent groups of movement activities exhibited movement performance of being too slow.
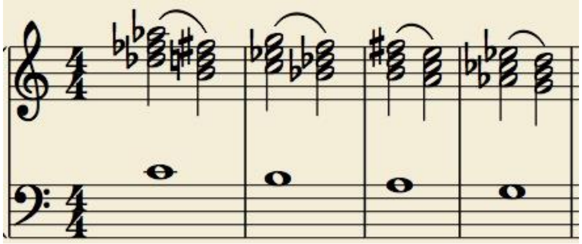
Figure 3. Composition for "Too Slow" Sequence

A “not smooth” sequence signifies an overall inconsistent movement for a group of activities. As shown in **Figure 4**, this composition does not have pedal tones, which, contrary to the first two sequences, enforces the concept that the movement activity is not consistent. It instead moves through “tetrachords” which are all based on C major. It begins: D to G and G to C, then C to F and F to B, then B to E and E to A, then A to D and D to G. The reason behind this sound design is to create a cyclical motion similar to how someone moving in a hesitant manner would travel. This sequence becomes a recognizable signature for the patient to understand that recent groups of movement activities exhibit an inconsistent movement performance.

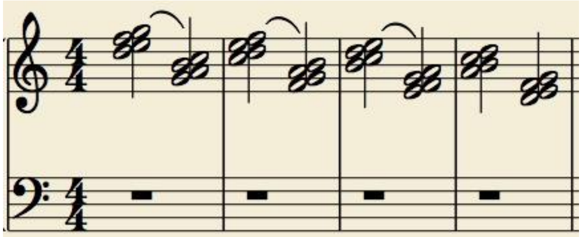

Figure 4. Composition for "Not Smooth" Sequence

A “too fast” sequence signifies an overall too fast movement for a group of activities. As shown in **Figure 5**, this composition also does not have any pedal tones. It progresses through pentatonic scales. Each pair is one or two minor seconds away from each other, creating noticeable, jarring, fast-moving and non-resolute successions. This sequence becomes a recognizable signature for the patient to understand that recent groups of movement activities exhibited movement performance of being too fast.

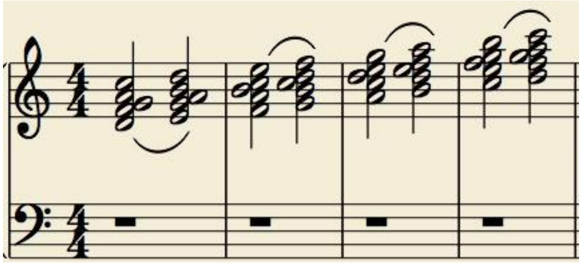

Figure 5. Composition for "Too Fast" Sequence

3.2.2.2 *Sonic Rhythmic Distributions*

The sonic image tags listed above form a cohesive sound suite that can be generalized while implementing different sound samples without changing the semantic meaning of the SME The three primary semantics are consistency of movement activity duration, ideal acceleration and deceleration speeds during movement, and the peak velocity during movement. While sonic content is an important factor in relaying these semantics, it is rhythm that holds the key towards maintaining useful and consistent sonic media and movement information.

Table 3. Rhythmic density distribution per sonic image tag

	Density 4	**Density 3**	**Density 2**	**Density 1**
Normal	340ms	180ms	120ms	90ms
Too Slow	500ms	400ms	320ms	256ms
Hesitant	400ms	250ms	349ms	198ms
Too Fast	300ms	150ms	90ms	62ms

Another sonic attribute that can be manipulated is the speed filters. Mapping constraints help determine the most important attributes from which to focus and to pull the most valuable pieces of information. **Table 3** shows note durations for each of the densities within each of the speed filters for each of the sonic image tags. The densities are broken down as ratios of each other starting with the base (density #4). As acceleration continues and adaptable spatial thresholds are crossed, the system traverses through the densities below (from left to right). These densities fit within the spatial thresholds much like the harmonic progressions do within the hospital-based system, creating bell-like note densities.

Other sonic attributes that can be manipulated are the register range, change, and the probability of change. A low probability of change would keep notes within one or two registers, which allows for repeated notes.

Repeated notes convey to the patient that “beautiful, expressive melodies” are not being accomplished, therefore a low Brownian variable would be needed. Constraining to one or two primarily high (or primarily low) registers conveys the feeling of “being stuck”, which is not an ideal situation. The patient must break out of the rut by correcting movement activities in order to create collections of more pleasing melodies.

Although much research has already gone in to the conceptual design of this composition model (see Section 2), defining the sonic space within the framework of a rehabilitation system that provides *post-set* sonic media information is relatively novel.

3.2.2.3 Assessment of the Sonic Media Environment

The efficacy of the affective summaries for the SME discussed in Section 3.2 was evaluated with 11 unimpaired subjects in an IRB-certified user study. The purpose of the study was to determine if there is a discontinuity between understanding sonic media environments given in real-time versus sonic media environments given after groups of movement activities. It helped inform and validate the sonic media designs used within the home-based system. The music conveyed a set of movements that were either all fluid, inconsistent, all too fast, or all too slow. During the user study, the subjects matched the sonic tags to the provided affective descriptors with over 90% accuracy.

4. CONCLUSION

The purpose of this paper is to document three years' worth of research on the sonic media environment for the home-based mixed reality rehabilitation system and to exhibit validated proof that meaning connects action to music. Although the home-based system has not yet been used or tested in real scenarios, there is a significant ground-truth basis that conveys the need for such systems. Proposed within this paper was to change the dialogue surrounding sonic media environments for rehabilitation purposes by researching viable solutions and designing a system to provide an immersive, multi-layered sonic media environment. Discussed, herein, is the motivation behind the project. Herein, the previous sound paradigms are defined alongside the proposed sound design, as well as the system design and the quantitative and qualitative solutions that target a new discussion for audio biofeedback.

5. REFERENCES

Burkholder, J. P., D. J. Grout, & C. V. Palisca. 2010. *A history of western music*. New York: Norton & Company.

Cirstea, M. C., & M. F. Levin. 2007. *Improvement of arm movement patterns and endpoint control depends on type of feedback during practice in stroke survivors*. Neurorehabilitation & Neural Repair, vol. 21, pp. 398.

Dourish, P. 2004. *Where the Action Is: The Foundations of Embodied Interaction.* Cambridge, MA: The MIT Press.

Duff, M., et al. 2010. *An Adaptive Mixed Reality Training System for Stroke Rehabilitation*. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 18(5), pp. 531-541.

Essl, G. & P.R. Cook. 1999. *Banded waveguides: Towards physical modeling of bowed bar percussion instruments*. Proceedings of the International Computer Music Conference (ICMC). Beijing, China. October 22-27.

Gabrielsson, A., & E. Lindstrom. 2001. *The influence of musical structure on emotional expression*. Music and Emotion: Theory and Research, vol. 223, pp. 223-248.

Gardner, H. 1999. *Intelligence reframed: Multiple intelligence for the 21st century*. New York: Basic Books.

Ghez, C., R. Scheidt, & H. Heijink. 2007. *Different learned coordinate frames for planning trajectories and final positions in reaching*. Journal of Neurophysiology, 98(6), pp. 3614-3626.

Ghez, C., T. Rikakis, R.L. DuBois, & P.R. Cook. 2000. *An Auditory display system for aiding interjoint coordination*. Proceedings of the International Conference on Auditory Display. Atlanta, GA. April.

Headlee, K., T. Koziupa, & D. Siwiak. 2010. *Sonic Virtual Reality Game: How Does Your Body Sound*. Proceedings of the International Conference on New Interfaces for Musical Expression. Sydney, Australia. June 15-18.

Jaques-Dalcroze, E. 2007. *The Eurhythmics of Jaques-Dalcroze.* Rockville, MD: Wildside Press.

Papert, S. 1980. *Mindstorms: Children, Computers, and Powerful Ideas*. New York: Basic Books, Inc.

Prinz, W. 1990. *A common coding approach to perception and action.* Relationships between Perception and Action, pp. 167-201.

Prinz, W. 1997. *Perception and action planning*. European Journal of Cognitive Psychology, vol. 9, pp. 129-154.

Schmidt, R.A. 1991. *Motor learning principles for physical therapy*. In M.J. Lister, editor. Contemporary management of motor control problems: Proceedings of the II STEP Conference. Alexandria, VA: Foundation for Physical Therapy. pp. 49-63.

Siwiak, D., et al. 2011. *A home-based adaptive mixed reality rehabilitation system.* Proceedings of the ACM International Conference on Multimedia. New York, pp. 785-786.

Siwiak, D., J. Berger, & Y. Yang. 2009. *Catch Your Breath - musical biofeedback for breathing regulation*. Proceedings of the International Conference on New Interfaces for Musical Expression. Pittsburgh, PA.

Seitz, J. 2005. *Dalcroze, the Body, Movement, and Musicality.* Psychology of Music, vol. 33(4), pp. 419-435.

Sundaram, H., & T. Rikakis. 2006. *Experiential Media Systems.* In B. Furtht, editor. Encyclopedia of Multimedia (Vol. XXVIII, p. 989).

Thaut, M. 2007. *Rhythm, Music, and the Brain: Scientific Foundations and Clinical Applications*. New York: Routledge.