Aesthetic Forms of Expression as Information Delivery Units

Paul Nemirovsky and Glorianna Davenport

Interactive Cinema Group, MIT Media Laboratory

Cambridge, MA, USA

{pauln, gid}@media.mit.edu

Introduction

In this paper we explore the hypothesis that aesthetic forms of expression – such as music, painting, video – can be used for direct information delivery. In contrast to text or verbal narrative techniques, which require a conscious act of transcoding, these aesthetic forms stimulate a more direct, emotional response. If shown viable, such a hypothesis could open new channels for the delivery of various types of information, providing us with a background information channel in situations of information overload, leaving our foreground concentrated on the more thought-demanding tasks.

To develop a system based on the notion of using aesthetic forms of expression for direct information delivery, we need to develop its core elements. In this paper we define a core element called "emon", a small discrete unit of aesthetic expression, which generates an expected emotional response affecting human behavior. The study is currently restricted to the domain of music, with candidate emons being 1-15 seconds long loops of audio that are currently assumed to be the only audio source perceived by the user. The emons are characterized as units of an independently describable value, without the necessity of connection / abstraction to / from other pattern units – i.e. if a specific emon is played we will be able to relate to its qualities without accessing our knowledge about other emons. In the first half of this paper we discuss the guidelines for emons' creation, describe the categorizations process, and report the results of emons' testing performed by a group of fourteen users.

Given the hypothesis that certain musical patterns (emons) can be used to provide cues that affect behavior, we need a system that can provide a further validity to the usefulness of that approach. In the "Implementation" chapter we report the ongoing development of the GuideShoes wearable system, which assists a user in navigating an open space such as streets, by sequencing musical emons as navigational cues. It is then followed by a discussion of the navigation tools written for this project and future research directions.

Imagine…

…You are in Tokyo, facing the travelers' everyday problem of getting from point A to point B. It's an even harder problem in a city that doesn't have street names, with your ability to speak Japanese equal to your ability to compose a symphony. This time the situation is different – rather than panicking, you put your GuideShoes and a headset on, and tell it where you'd like to go. Upon your request, GuideShoes pack connects to the web, and finds out your current and desired locations. As you start walking down the street, your headset starts playing music. There are no signs or language to assist you - but you are language-independent. Musical patterns (emons) provide you with information regarding the correctness of your direction – they evolve in ways that you find natural and in correspondence with the emotional states of "right", "wrong", and the gray area in between. The patterns may indicate the duration of your journey and proximity to the destination. You no longer need to remember the map or think about how to get where you are going. The only thing that is required is your ability to hear. Have you entered an unfriendly neighborhood? GuideShoes will tell you. The GuideShoes try to do it by considering your musical preferences, merging the invisible and inadvertent emons into a meaningful tangible interface utilizing background information channels.

Finding your direction is only one example of a frustrating and time-consuming task that can be addressed using emons. People who can't or won't use printed or spoken instructions – small children, the visually impaired, users occupied with other, more urgent tasks – can be helped in new efficient ways instead of being left alone to deal with situations that require a significant amount of cognitive and perceptual skills.

Hypotheses

Defining the relationship between pattern and meaning has been the objective for a few generations of musicologists, cognitive scientists, and other researchers concerned with the interpretation of things played, drawn, and otherwise acted. Most theories are based on sequential systems of interpretation of artistic mediums – an artistic expression is first perceived, then recognized cognitively, and then referenced or given meaning beyond its initial domain. Theorists are doubtful regarding the ability of music to deliver meaning as it lacks the precise semantics present in verbal language. According to this widespread view, music cannot convey meaning as every listener may derive very different "meanings" from the same musical piece. Even among the heteronomists – the theorists that agree on the possibility that artistic expressions are valuable beyond their aesthetic appeal – only a few believe in the possibility that artistic mediums can be used as self-contained information containers using their emotional capacity. If proven possible, it could lead us to a design of new informational meta-mediums with fewer cognitive steps required during the recognition process of an incoming information stream.

Similar to the studies conducted by Konrad Lorenz et al that proved the possibility of information imprinting on animal behavior, nowadays an extensive research is being conducted on how imprinting a specific combination of emotions on information can enhance that information's appeal for humans (Losee, 1998). By studying the existing mappings of aesthetic forms onto behavior models of the brain as well as imprinting our own, we can possibly achieve higher efficiency of information and make a better use of users' emotional abilities. To use these mappings we need to know what the individual components of an emotion are. The whole – an aesthetic expression – should be dissected into affective (i.e. emotionally charged) units.

The GuideShoes project, described in the second half of this paper, represents our work on incorporating such units into a wearable device, which will provide users with information regarding their travel. The system provides users with emotional cues by employing a system of aesthetic information fragments, called *emons*.

The emon approach is based on the following hypotheses:

1. Aesthetic forms can be used as direct information delivery mediums.

   We hypothesize that human perception abilities are underused and can be enhanced with the emon approach. This paper focuses on the musical segment of that approach and the use of musical patterns as the building blocks. Musical emons are appropriate because of music's ability to communicate emotion in an immediate and efficient way. Music is also convenient because its high level of abstraction allows us to test the principles of emons' construction and find whether certain candidate emons get a consistent response/ratings in a series of tests. Candidate emons that demonstrated a consistent rating should then be considered the true emons and employed in the emon-based systems. System users who listen to the emons would then be able to use their musical and emotional judgment to recognize the emons' meaning.

2. Aesthetic information can be isolated into small autonomous elements (emons).

   According to the traditional approach, aesthetic expression consists of a continuous stream of elements. We hypothesize that by dividing the stream into discrete emo-informational elements it will be possible to open a new channel of information delivery, and achieve a more efficient perception process without increasing the learning curve.

   To provide an initial testing for this hypothesis we have created a library of candidate emons, and developed an evaluation application to check the efficiency and the consistency of emons' ratings (i.e. how high different emons scored, and how stable these scores are). A statistical tool to allow easy visualization/estimation of the testing results has also been developed. A further testing of this hypothesis is provided using the GuideShoes; a passive I/O device for a real-time delivery of emotion-based information content (emons). GuideShoes' aim is to help the wearer to navigate through an open space (such as streets), making navigation personalized and less cumbersome by maximizing information input, by incorporating emotional disposition and reaction, and by combining artistic and informative communication.

3. Emons can be recombined to produce predictable emotional/informative responses.

   The exploration of this hypothesis – defining the laws of "emons' combinatorics" – is a part of future emon research beyond the scope of this paper. It would be naïve to regard emotion as a mere sequence of positive/negative states; and indeed, this work does not have that goal. We do hypothesize however that the emotional power provoked by aesthetic means can be utilized in order to achieve a more efficient means of information delivery – granted the laws of emons' combinatorics have been defined. An initial examination of this hypothesis is provided using the GuideShoes system. We are also working on an application that will serve as a front end to the emons' database, providing a further functional framework for the research of combination-related functions. More projects are envisaged to provide the necessary validity for that hypothesis.

Prior Research

The emon approach is based on a number of concepts proposed by scholars in various fields. Here is an overview of a few of the ideas that inspired the theoretical approach and the practical application of emons. The ideas addressed in this chapter are relevant to emons' action, meaning, parameters' definition, and composition process.

Marvin Minsky (1985) in his "Society of Mind" proposes a theory of how the brain learns/memorizes a concept. To explain this process, he introduces the concept of *paranomes*: "The idea is that, typically, what people call a "concept' is represented in the brain in several different ways. However, these will usually be cross-connected so that the rest of the mind can switch easily from one representation to another. The trick is that although each representation-method has its own type of data-structure, many of the "terminals" of their frames (or whatever else they might use) are activated by the same "pronome" signals". In this context emons can be viewed as triggers for emotionally charged representations.

William Buxton (1995) in his "New Taxonomy of Telematics" proposes a two-dimensional model of human-computer interaction (HCI). In his model, the interaction happens on *either* background or foreground levels, so that any medium populates one of the two cells. He is aware of the power of the background information channels and proposes "…a means of sharing the periphery, the background social ecology, by means of appropriate technological prostheses". However, his model lacks flexibility to address the notion of switching between the foreground and background cells within the action frame of the same object. It is interesting to see whether the emons' approach could address the need of tangible objects to be easily transferable between those states. It would make them more adequate for use in a real world situation, where switches of our attention continuously cause dynamic reassignment of attention weights in regard to the surrounding artifacts. These new tangible objects would address the frame of action in more intelligent and modal ways, bringing us closer to the creation of natural forms of computational objects.

HCI researchers frequently address the question of modality. Hiroshi Ishii (1997) relates to it as one of the topics of his research in tangible interfaces. He writes, "…subconsciously, people are constantly receiving various information from the "periphery" without attending to it explicitly. If anything unusual is noticed, it immediately comes to the center of their attention". Ishii's view of a tangible user interface design is to "employ physical objects, surfaces, and spaces as tangible embodiments of digital information. These include foreground interactions with graspable objects and augmented surfaces, exploiting the human senses of touch and kinesthesia" and also "background information displays which use "ambient media"—ambient light, sound, airflow, and water movement." He seeks to "communicate digitally-mediated senses of activity and presence at the periphery of human awareness". Inspired by these ideas, we explore the notion of musical emons as basic elements of a new tangible approach – processed in the background, in parallel with other media sources, and reconfigurable to reflect the current state of the user and of the environment.

The combination of audio and wearable computing is explored by Peter Meijer (1997) in his "Auditory Image Enhancement". His vOICe system "translates arbitrary video images from a camera into sounds. This means that you can see with your ears, whenever you want to". While achieving a solution for an interesting technical challenge, it seems that mapping of the visual domain onto the auditory domain would be much more effective and easy to perceive if an emotional component played a more significant role. Meijer's approach may work well in static situations, such as in the interpretation of still pictures. However, in dynamic open-space environments, such as streets, his solution for the cognitive problem of navigation seems inefficient as it overloads the users' perception channels with vast amounts of unfiltered information and is problematic for tasks involving items of varying importance or priority.

The Anatomy of an Audio Emon

Music has an infinite number of properties to play with. In this chapter we give an overview of our decisions during the emons' design process and discuss the axes that can be populated. It has to be noted that while this paper addresses the music domain, a similar deconstruction could possibly be performed with other forms of aesthetic expression.

Approach

In order to build an information delivery system based on aesthetic forms of expression, we have to define the elements of which this system consists, and create a personalizable mapping system. "Colors, sounds, odors, tastes, tactile experiences, all may be "heavy" or "light" or have "volume" and dozens of other psychological similarities" (Faber Birren (1992), defining a correspondence between color segments and emotional states). We start with an attempt to demonstrate a correspondence between certain musical solutions and emotional states. Music has a variety of interesting properties to play with – rhythm, timbre, texture, and so forth. As a complex medium, music offers us great degrees of freedom – its high degree of abstraction allows us to manipulate and adjust it in almost any imaginable way. It also presents us with the challenge of inventing a technique to reliably convey an emotional state using a unified method.

Definition

Discrete units of emotional expression, or *emons*, are aesthetic information containers, which are capable of generating an expected emotional response that can affect human behavior.

The emons are characterized as units of an independently describable value, without the necessity of connection / abstraction to / from other pattern units – i.e. if a specific emon is played we will be able to relate to its qualities without accessing our knowledge about other emons.

Design & Evaluation

In order to design a valid set of emons for various life situations, we first selected an emotional model and defined criteria for emons' construction. Subsequently, we composed a set of candidate emons, defined the evaluation methods, wrote the evaluation software, and started to conduct a line of tests – first with a group of subjects determining the real emons among the candidate ones and then with a different group using the selected emons for real-time navigation. A brief overview of this process follows.

Affective Categories

We chose to base the emon categorization on a partial implementation of the circumplex model of the emotions, developed by Robert Plutchik (1994). The model is structured in a relatively simple (and therefore applicable) way, while being adequate in its perspective on the possibilities of emotion synthesis from individual components. The implementation described in this paper is limited to the primary emotions (as outlined by Plutchik), namely: Acceptance, Anger, Anticipation, Disgust, Fear, Joy, Sorrow, and Surprise. This is an initial attempt at exploring a practical application of Plutchik's vision of a possibility of emotional combinatorics, which may prove very helpful in dealing with discrete emotional elements such as musical emons.

Construction Principles

In order to test the emon approach a library of musical emons has been created. We classified the emons' inter-relationships into the following categories to aid the composition process:

1. Major ←→ minor as positive ←→ negative. The binary notion of right/wrong is not meant to imply a binary relationship, but rather a tendency in the emotional reaction to a composed emon. In this form, the right/wrong scale can be applied to all the following pairs, however the actual assignment of the emotional extremes to the positive/negative ends of the emotional scale is up to the actual users.
2. Loud ←→ quiet as positive ←→ negative.
3. Defined fast tempo ←→ slow tempo as positive ←→ negative.
4. Continuous sound with a difference in pitch. High ←→ low pitch as positive ←→ negative.
5. Separate sounds unified by certain rhythmic patterns; steady ←→ unsteady rhythmical pattern as positive ←→ negative.
6. Instrumental density; the amount of simultaneously heard instruments mapped to the positive ←→ negative scale.
7. Melodic density; the notes/time ratio mapped to the positive ←→ negative scale.
8. Rhythmic and melodic repetitiveness versus [pseudo]randomness; the repetitiveness ratio mapped to the positive ←→ negative scale.

In order to test the emon approach, a number of libraries of candidate musical emons have been created. The emons were created with no restrictions on their style or form, basing their design on the composition sense as well as trying to utilize the components of composition techniques from various musical styles. The candidate emons were created with no specific emotional category in mind.

A total of 200 emons have been composed, each one being a 1-15 second long loop. After recording, the emons were saved as separate MIDI file loops, ready to be streamed in real time upon request.

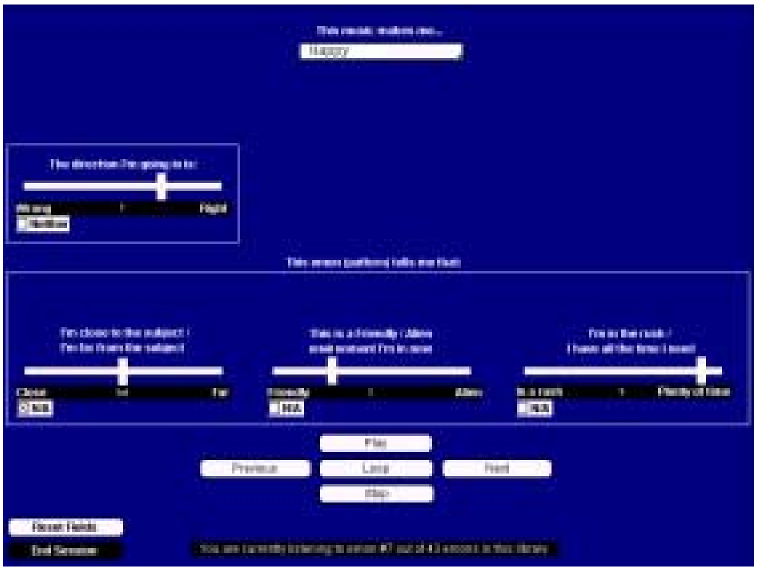
Technical Issues

All the emons were composed using Korg Trinity V3, Roland SoundCanvas SC-88, and NordLead synthesizers, edited in Steinberg Cubase VST/24, recorded onto the hard disk using MOTU 2408 & DigitalAudio CardD+, and converted to WAV-formatted loops in Steinberg Wavelab. The categorization tool (fig. 1), used to classify the emons into various emotional/action categories, was written in Lingo, with additional Visual C++ modules for database access (with the MS Access connectivity package), MIDI files playback, drawing capability, and visual effects. Multiple clients can be run simultaneously via the network, with the database located on a server, hence allowing for a parallel and more efficient testing process.

Tests / Evaluation of results

The testing process is aimed at gathering data for future exploration of various ways of mapping emotion/music to physical spaces. More specifically, the emons found in the testing are to be used in the GuideShoes system (aimed at providing us with a navigational means, as well as enhancing our awareness of the surroundings).

Figure 1: Audio emons' testing tool.



The emon categorization test gathers the data by asking a group of subjects to respond to a number of questions at each stage of the test. The test starts with a demographic questionnaire, asking for subject's age, gender, favorite musical styles, level of musical literacy, up to four personal emotional characteristics, current mood, and the arousal level. The subject then proceeds to do the main part of the test. The test is performed by listening to each of the candidate emons, which are presented to each subject in a different order, to avoid a possible bias. During the test the subject: classifies each of the emons as related to one of the emotional categories; sets the intensity of the relationship; rates the candidate emon's environmental scales (as defined in software shown in fig. 1). There is also an option of contributing "custom" emotional states. After the end of the experiment, the subjects are asked to provide their comments in a free form essay.

Results of the Emons Testing

It is too early to state the results in definite terms, but taking the limitations of a small group testing, the overall results are surprisingly promising. Despite the fact that the emons were composed with *no* intent of addressing particular navigation-related qualities of the emons (their directional, proximity and other ratings), a surprisingly high number of candidate emons were rated as either "completely right" or "completely wrong" (values in [6.5000-7.0000] & [1.0000-1.5000] ranges respectively) on all the scales (direction, proximity, etc.)

The characteristics most accurately conveyed by the emons were directionality and time-rush. The accuracy can be seen across the subject groups, with 8% of the emons rated in the highest ("most positive") range (top 15%) by the directionality factor and 6% by the time-rush factor (2% for proximity & 1% for environment factors). The highest value for the navigation factor (as rated by *all* subjects) was [7.00], and the lowest was [1.33]. While presenting a basis for an interesting hypothesis, these results should be tested with a far larger group of users. It does seem possible however to state that users seem to agree on certain musical fragments in relation to their direction- and time- related characteristics. As soon as the user group is narrowed demographically down from the "all" criteria, even more extreme values show up.

Figure 2: Full output of the testing procedure for 'all users' selection query:

	%by Direction	**by Time Rush**	**by Proximity**	**by Environment**
Top/Bottom 15%	*8/1*	*6/2*	*1/0*	*2/0*
Top/Bottom 30%	*36/11*	*20/22*	*20/4*	*26/9*
Top/Bottom 45%	*57/27*	*42/48*	*50/27*	*61/19*
Neutral	*21*	*15*	*27*	*25*

This paper represents a work in progress. While the current user base is small and the received results cannot be generalized beyond more than a limited group of users, a good basis can be seen from the initial tests for a future research regarding the use of emons as information cues affecting users' behavior. According to the initial tests, emons have the potential to become useful as information delivery units. As of now, ~10% of the candidate emons were defined by the subjects consistently as either "very positive" or "very negative", and that finding holds the premise of filling the first part of an emon grammar. More work is required to design additional libraries of emons and to understand the correspondence between compositional factors and users' ratings.

Implementation

The GuideShoes project was chosen as an initial platform for the test of the emons' approach. Our primary goal with the GuideShoes system is to test the emon-related hypotheses through musical emons-based information delivery, while hoping to design a new navigational tool. Navigational control was chosen as an example of a rather simple and expandable task that has opportunities for exploring both binary (right/wrong) and fuzzy (better/worse) relationships. Using the GuideShoes wearable interface is useful as it allows us to come up with multiple test scenarios to provide the validity for the emon approach (as outlined in the "Hypotheses" section).

Current Scenario of Interaction

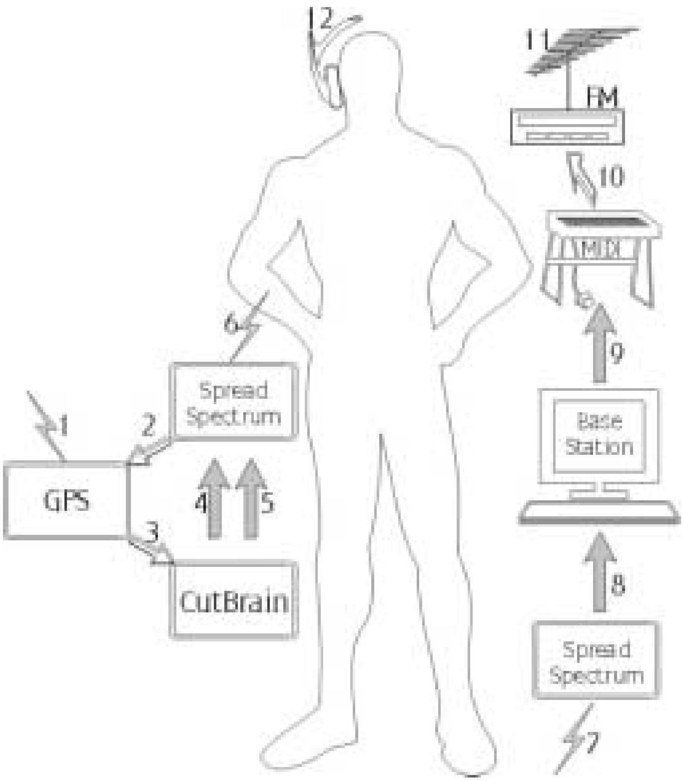
A user comes to the base station, picks a destination point on the map, specifies a few optional profile details (such as age, gender, and favorite musical style), puts the GuideShoes on, and starts walking. Depending on the correctness of the direction, and compliance with the properties specified for the travel, the user hears different musical patterns that provide the necessary navigational cues to get to the destination, while presumably leaving the cognitive foreground open for more creative tasks.

System Overview

The GuideShoes system consists of a wearable part – a leg-mounted pack equipped with a GPS, a wireless spread spectrum radio, a custom-built motherboard plus a pair of FM radio headphones – and a base station that acts as the central unit for path selection, data processing, and emons' retrieval/playback. The emons are produced by a MIDI synthesizer, and delivered to the user using an FM transmitter – at the place and time where they are needed. The placement of the wearable has been chosen based on several factors: an unobtrusive location, an otherwise unused strong spot of the human body, and closeness to the body – thus allowing accurate readings for the attached sensors.

Structure Of GuideShoes

Figure 3: System architecture. Numbers show the flow of data.



The client-server interaction schematic is shown in fig. 3. The client consists of a DGPS (Ashtech SK-8), custom-built CPU (CutBrain), spread spectrum radio (FreeWave OEM), and a pair of radio headphones. The base station runs a custom-built map UI with path selection features (MapTool, implemented in Tcl/Tk), and emons' retrieval engine (C++). It also controls the FreeWave base station, an SBX-2 differential beacon receiver, an additional SK-8 DGPS for detection of base station location, and a Veronica Kits FM transmitter used to deliver the audio to the GuideShoes user.

System Operation

Every second, the base station sends differential corrections to the client's differential GPS, which sends the corrected user position back to the base station. The spread spectrum modules on both ends hold these and other data exchanges. The base station processes the corrected DGPS data and, based on the correctness of the movement, retrieves and sends one of the emons from its library back to the GuideShoes, where it is played through a wireless headset.

Navigational Applications

GuideShoes is a navigational tool designed for a wide range of people that can be described as users with navigational difficulties, and users for whom the emons' approach presents a plausible alternative for the otherwise cumbersome task of classical navigation.

These groups of people include:

1. Children: The problem of children's safety (such as not getting lost while walking through a city) cannot be fully solved by GuideShoes; however the GuideShoes can make this problem more manageable. If a child has been exposed to the proper use of emons, the GuideShoes could provide him with a fun and helpful aid for getting to the place he needs to go and give parents the knowledge of his current location / alert them to potential direction problems.*
2. Orientation for disabled (blind & psychological orientation impairment): The GuideShoes, if wisely implemented, may become an important navigational aid for people with severe vision problems. As the work of Peter Meijer (1997) aims to show, people can be conditioned to fairly complex patterns of movement expression using audio information. The melodic structure of the emons should make them easier to relate to than to a purely algorithmically produced audio. Mapping emons to the validity of a walking direction (as well as potentially to the location of objects) can enhance their navigational abilities.
3. People with brain damage: According to experiments conducted by Martha Farah (1989), people with certain brain damage in the area of visual perception also had visual memory deficits "directly comparable to their perceptual deficits". Therefore, it will be interesting to test whether people with spatial disabilities (such as being unable to see relative locations of objects in a scene) can rely on GuideShoes as a complementary device to receive the same information through the alternative channel of music. **
4. Tourists: While not having the navigational difficulties similar to the three groups before, tourists have to find their way around cities without signs in their native languages. They will prefer to concentrate on the experience and not on the navigational task.
5. There is a substantial group of people who, while not having the difficulties of the groups described before, prefer not to be overwhelmed with unnecessary information. Among these are the young people for whom the GuideShoes could represent an easy to learn alternative, and people with well-developed auditory perception, that otherwise goes unused in their daily navigation process.
    - * As of today, the DGPS precision is limited to 2 - 5 meters.
    - ** Testing with narrowly defined groups of population is beyond the scope of this study.

Continued development of the GuideShoes system

Technical

GuideShoes is a research project in progress. To further develop & demonstrate the emons' foundational principles, we plan to develop a larger library of emons for navigation-related purposes. The library, combined with a larger amount of test subjects, will allow us to further study the relationship between the emons' compositional parameters, their navigational ratings by the test subjects, and the effect of such on the GuideShoes' users.

We are also planning to increase the robustness of the GuideShoes system operation. We hope to achieve this by designing a custom enclosure for the GuideShoes, securing the hardware elements in their positions, and developing an easier method for the battery charging process.

Conceptual

This paper describes the process of selecting musical emons to deliver navigational cues and describes the ongoing development of the GuideShoes system as the first artifact that utilizes the idea of emons. The question that drives the future development of the emon approach is: What other kinds of information can be delivered by emons and what kind of emons would these be? We have defined the four main related research directions as:

1. Merging between the emon approach and the recent psycho-physiological research.
2. Enhancing the emons approach to create complete aesthetic musical-informational spaces by assigning emons to physical objects.
3. Developing additional applications based on the approach.
4. Expanding the emon approach to olfactory domains.

This research will hopefully lead us to the creation of an emon-enriched space, capable of monitoring and affecting the emotional state of its inhabitants and will evolve into a support system for educational and stress-relief purposes.

Acknowledgements
We would like to thank Dan Overholt, Dennis Kwon, Richard Tibbetts, Hristo Bojinov, and Raymond Luk for their help in creating the hardware and software components of the GuideShoes system.

References

Birren, Faber (1992). *Color Psychology and Color Therapy*. Citadel Press.

Buxton Bill (1995). *Integrating the Periphery and Context: A New Taxonomy of Telematics.* In proceedings of Graphics Interface '95.

Gardner, Howard (1983). *Frames of Mind: The Theory of Multiple Intelligences*.

Gray, Peter (1994). *Psychology*. Worth Publishers.

Ishii, Hiroshi., Ullmer, Brygg (1997). *Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms.* In proceedings of CHI '97, ACM Press.

Losee, Paul – "The Creation of Memory and the Imprint Process", M+ research group, 1998

Meijer, Peter (1997). *vOICe System*. http://ourworld.compuserve.com/homepages/Peter_Meijer/voice.htm

Meyer, Leonard (1956). *Emotion And Meaning In Music.* University of Chicago Press.

Minsky, Marvin (1985). *Society of Mind*. Simon & Schuster.

Picard, Rosalind (1997). *Affective Computing*. MIT Press.

Plutchik, Robert (1994). *The Psychology and Biology of Emotion*. Harper Collins.