

Towards The Ultimate Aesthetic Experience

Haakon Faste

PERCRO Perceptual Robotics Laboratory
Scuola Superiore Sant' Anna
Via Martiri, 11
56127 Pisa, Italy
haakon_faste@hotmail.com

Massimo Bergamasco

PERCRO Perceptual Robotics Laboratory
Scuola Superiore Sant' Anna
Via Martiri, 11
56127 Pisa, Italy
bergamasco@sssup.it

ABSTRACT

Robotic interfaces combined with virtual reality provide an unparalleled platform for cognition research. Using a combination of design strategy, trend analysis and the programming of immersive virtual 3D worlds, we have developed a low-cost aesthetic multimodal experience. Our intent has been to design the “ultimate interface” for creative interaction. This paper discusses the process, content, results, and impact on our engineering laboratory that this research has produced.

Keywords

Perceptual robotics, multimodal cognition, virtual reality, 3D environments, design strategy, aesthetics, creativity.

INTRODUCTION

Robots are extensions of human activity. Just as we use our bodies to create and experience art, robotic bodies may be used as expressive and experiencing tools. Robots, like humans, are processes of interaction and learning that operate within the context of pre-existing systems. The robotic body is a purposeful medium, not an end in itself.

Robots are often sophisticated and expensive tools, and may be designed or appropriated for artistic use. When designed to satisfy specific objectives, the use of robotic technology requires a significant investment of resources before a working system is produced. For robotic art to have a serious cultural impact, it can utilize design processes that allow imagined or prototyped visions or fantasies of robotic expression to be realized according to a plan. Engineering strategy thus becomes an artistic process by which the expressive potential of the robot is defined.

To touch, feel, and empathize with our surroundings are central aspects of human experience that are widened through the use of robotic technology. When coupled with technological advances in virtual reality (VR), robotic interfaces displaying vibrotactile or force-feedback provide an unparalleled platform for research on human creativity and cognition. Not only do such interfaces enable novel possibilities for creative expression, they provide a

“multisensory laboratory” with which to examine the mind. The result is a fully immersive “fantasy medium” within which humans can explore, imagine and learn.

As advances in artificial intelligence and robotics become increasingly sophisticated, human-machine systems become collaborative minds. Our involvement in the field of perceptual robotics indicates that creative expression is central to human memory and learning, and of particular interest for the study of artificial cognition. We have therefore developed we are calling “*The Ultimate Aesthetic Experience*,” a high-impact virtual environment that can be immersively experienced via relatively low-cost interface technologies including stereoscopic projection, 3D audio, robotic feedback and optical tracking. Our intent has been to explore the possible behaviors of and relations to this technological platform while maximizing its emotive potential. Throughout the design process, the specific configuration and behavior of elements of the system has been left intentionally ambiguous such that they may emerge through the process of designing and interacting with it. The system architecture includes numerous parallel computational frameworks that are reactive to interactions between the robot and human with the intent of creating a profound sense of magic and wonder.

This paper outlines the development process and preliminary results of this project. Prior to discussing the system in detail, a brief overview of the field of perceptual robotics has been provided to give appropriate context with which to frame our research.

BACKGROUND: PERCEPTUAL ROBOTICS

Perceptual robotics is the use of robotic and multimodal display systems as a platform for sensory interface with real or virtual worlds. Well implemented perceptual robotic interfaces provide:

- Fully immersive and believable interaction with virtual or tele-operated worlds, including the sense of touch, force feedback, and presence in that world.
- Intuitive perception of the robot as an extension of the user's body and mind.
- Insight into human perception through this interaction, which may be used in the development of more intelligent interface systems that are capable of perceiving and learning autonomously.



Figure 1. Existing prototypes of perceptual robotic technology. These include, from left to right: the PERCRO Body Extender, an EEG-controlled robotic exoskeleton for manipulating heavy loads in unstructured environments [1]; a novel encountered haptic interface for the finger [3]; the GRAB system, which is easily employed as a general purpose robotic force feedback device [19]; and Passages [11], an enactive interface for artistic expression.

An overview of perceptual robotic technology is shown in figure 1. When combined with advances in pervasive and ubiquitous computing, technologies of this nature will become increasingly central to “everyday creativity.” Low-cost gaming systems, context awareness and machine intelligence have already altered radically how we interact with our surroundings. Advances in robotics will further revolutionize human-computer interaction. Key emerging technologies include the use of artificial neural networks and machine vision to simulate the functions of human perception, for example. Typically these are adaptive computational models that adjust their structures based on external or internal information flowing through the network during its learning phase. Using these techniques, complex relationships and patterns in data can be located and used to provide interactive feedback. For example, body tracking and exoskeletal robotic devices coupled with intelligent software can be used to recognize a user’s gestures or analyze skill performance. In the case of scalp-mounted brain-computer interfaces, the simulation of human cognition is a direct interpretation of what the user perceives. A comprehensive discussion of perceptual robotics must therefore touch on the basics of multimodal cognition, virtual reality, and methodologies for the design of such interface systems.

Multimodal Cognition

From a scientific perspective, the study of perception is central to an improved understanding of human brain function, for psychology, the cognitive- and neurobiological sciences. All human learning begins with perceptual stimulus, from which memories are formed and behaviors are learned. Although neurophysiologists still cannot say precisely where and how memory formation takes place, it is clear that bodily experience is central to learning. In discussing the various stages of human development, for example, the psychologist Jean Piaget is noted for his identification of three principal stages of growth: a body-based stage (in which children explore the world using predominantly kinesthetic senses), a visual stage that begins at about the age of five, and the symbolic stage we associate with adult cognition [13]. Because each stage builds on the innate knowledge and wiring of the

previous one, an increasing trend in philosophical and cognitive models of the human mind understands all linguistic and iconic knowledge in terms of “embodiment” [16,30]. The study of the embodied mind is a rich unification of all academic fields, touching on computation, robotics, linguistics, developmental psychology, philosophy, religion, sport, craftsmanship, etc.

At a cognitive level, our capacity to remember both skills and concepts relies on our ability to organize the world around us into categories, some general, some specific. According to recent theories, information in the brain is distributed among many maps, and incessant reference back and forth among them is necessary for categorization to occur [7]. Remembering is not the re-excitation of innumerable fixed, lifeless and fragmentary traces, but rather an imaginative reconstruction, or construction, built out of the relation of our attitude towards a whole active mass of organized past reactions or experience [25]. Recent experiments with vision, for example, demonstrate that perceptual properties are constructed and grouped through phenomena strongly related to perceptual constancy, such as binocular depth perception, lightness constancy, amodal completion, and illusory contours [26]. They show also that grouping may be a ubiquitous, ongoing aspect of visual organization that occurs for each level of representation, rather than as a single stage that can be definitively localized relative to other perceptual processes [22]. Through the systematic correspondences among dimensions of categorization, memories are constructed and modified through time.

Our ability to perceive the world around us requires the categorization of perceptual cues into generalized concepts. Bodily experiences that engage multiple sensory modalities at the same time, such as vision, hearing and touch simultaneously, have been shown to be a particularly fruitful area for cognitive research. Although diverse perceptual attributes such as color and shape are processed in different parts of the brain and over an interval of different times (the perception of color preceeding that of form by 40 ms and of motion by 80 ms) we experience objects in the world as single, unitary entities [31]. Studies of synesthetes (individuals possessing a cognitive abnormality of the

perceptual system causing a “blurring” across senses) demonstrate that the combination of sensory fragments (qualia) to categories of mental concepts is evoked at a preconscious sensory level. Synesthetic color, for example, arises after binocular fusion and appears to be bound to a form as the form is being recognized [24]. Interestingly, synesthesia goes beyond pure sensory-sensory pairings to include the binding of qualia to categories of mental concepts. In his investigations of synesthesia, Richard Cytowic has shown that the brain’s transmodal modules (those modules that don’t pertain to any single sense) serve three neurobiological functions: they “construct multisensory representations of the world, they provide memory and affect the experience, and they critically participate in establishing categories via groups of coarsely tuned neurons” [5]. This supports the distributed system theory of brain organization described above in the sense that there are multiple mappings of a given function, and indicates that multisensory stimulus acts as an aid to memory stimulation and the formation of concepts.

Multimodal VR

Most VR systems make use of stereoscopic visual displays in addition to spatial tracking of the users body combined with sensory feedback in other modalities (such as audio or haptic displays). The combination of these technologies is intended to heighten the user’s sense of *immersion* and *presence* in a virtual space, where presence is the (psychological) state of consciousness corresponding to “being there” [27], and immersion refers to a quantifiable description of the technology, that is, the extent to which the computer display is extensive, surrounding, inclusive, vivid and matching [28]. Participants who are highly present in immersive VEs consider their experiences as places they have visited rather than as images seen. The feeling of presence enabled by VR is one of its unique characteristics, and allows human cognition to be examined in regards to an individual’s ability to act and react in specific experimental or therapeutic contexts. Slater et al. [29] note that although increased immersion may well improve performance in certain tasks due to the higher quality and quantity of information available, it is not presence per se that facilitates task performance but rather that presence brings into play “natural” reactions to a situation, in turn enhancing the sensation of presence and so on.

Multimodal VR systems intrinsically allow multi sensory data to be simulated and handled in structurally similar ways, superposed, and substituted across modalities [20]. Not only does combined sensory data arriving in parallel provide more convincing virtual experiences [15, 17], the human cognitive apparatus has also been shown to operate more rapidly given multiple, simultaneous sensory stimuli. Comparing simple detection times of unimodal (auditory, visual, and haptic) events with bimodal and trimodal combinations, for example, Hecht, Reiner and Halevy have shown that increased modality is correlated with faster mental processing [12]. This is why, despite supplying users with an impoverished, incomplete, and often inconsistent set of sensory cues, well

implemented multimodal systems consistently enable a rich experience of presence by providing additional time for perceptual integration. It also indicates that machine systems with limited senses in different modalities could gain awareness more rapidly through the study of cross-modal pattern recognition and learning.

The Ultimate Aesthetic Experience is conceived as a training ground for the study of improved human and machine cognition through their engagement in a dynamic and collaborative interaction experience. The mind of the robot is a virtual world, one that humans can experience first-hand from within. Fully immersive robotic minds allow the experience of unlimited fantasy. Possibilities are boundless. We have therefore been employing design methodologies to inspire and guide the imaginative process. This process has included a combination of ethnographic research, trend analysis, drawing, sculpting, computer programming and creative speculation. The intent is to use the resulting experience as a platform for further research on creativity and cognition.

Design as a Strategy for Aesthetic Experience

The field of design offers numerous methodologies for addressing the desirability, usability, and viability of new technologies. Due to the high costs of engineering in industry, for example, design strategy is increasingly fundamental to insure that the right problems are being addressed. By engaging with users early in the process and emphasizing the creative skills of synthesis, pattern recognition, aesthetic awareness and hands-on involvement, the processes of design can play a central role in determining where and how engineering resources should be focused.

Design can be considered a three step process consisting of identifying needs, giving form to possible solutions, and communicating these solutions to others. The first phase, also called “needfinding” [21], deals with uncovering opportunities for potential innovation. This process typically entails ethnographic observation, qualitative synthesis, and a focus on human factors [23]. It values the designer’s innate sense of feeling and aesthetics in addition to the analytical skill of reason and the study of function.

Aesthetics, from the Greek *aisthetikos*, deals with the ability to perceive with our senses impart feelings to others. It is only recently, as a result of literary criticism of the nineteenth century, that it has come to be associated with surface beauty rather than the perception of feelings. Aesthetics actually has to do with an overall perception of quality and is therefore central to both creative thinking and engineering practice at large [10]. In particular, to touch, feel, and empathize with our surroundings are central aspects of human nature, and at the core of VR’s fundamental goal of creating believable perceptual simulations. It is our contention that the most compelling human experiences will result from interfaces that maximize aesthetic potential. We have therefore employed design process as a tool to guide aesthetic experience and create an experimental platform for “cognitive art.”

DESIGN PROCESS OVERVIEW

Applying design strategy to engineering problems requires that the designer be attuned to the subtleties of human behavior and have synthesis skills by which to discern cultural trends, pain-points, blocks, and opportunities. Once identified, this information can provide a valuable roadmap for the accomplishment of project goals.

The design process we have used in the development of *The Ultimate Aesthetic Experience* is shown in figure 2. This process has involved rapid, iterative cycles of ethnographic research, concept synthesis, brainstorming and rapid prototyping. An early trend-finding exercise has identified nine major opportunity areas that represent a cross-section of what we call “*The Technocultural Condition*.” Based on these findings, innovative VR experience concepts have been generated, entered into a database, and mapped onto an easy/hard low/high impact matrix. The most promising concepts have then been expanded through iterative cycles of ideation, scenario storyboarding, synthesis, further brainstorming and software prototyping. In addition, the demo is being evaluated in cycles by a small group of test users in an immersive environment, feedback from whom has been incorporated into the final design.

Needfinding

The needfinding process we have pursued has included several introspective activities, beginning with a clear overview of today’s “technocultural trends.” Asking questions about technocultural interests (such as what fascinates us, what we enjoy doing, what people are talking about, what captures the “zeitgeist” or avant-garde of the moment, etc.) and developing a strong point-of-view on which trends to pursue provides both clarity of intent and guiding principles for the development of powerful multimodal experiences.

For three weeks at the beginning of the project, the authors took extensive notes (on sticky-notes) everywhere they went. These notes focused on the following areas of interest: intriguing concepts and discussion topics; university seminars and activities; personal fascinations, brainstorms and streams of thought; advertisements, publicity, films, movies, magazines, online experiences and other forms of mass media content; repeated conversations with interested professors, colleagues and friends; and observations of things that people such as ourselves seemed to enjoy doing. These notes were placed unorganized on boards, and then synthesized into “clusters” based on similarity. Nine major

trend areas were identified, representing opportunities for perceptual robotic experience design. They are (in no particular order): *[In]finite Reality*; *Subliminal Perception*; *Networked Agency & Mediation*; *Systems Fusion & Influence Mapping*; *Cyber Resistance & Reactivity*; *Multimodal Fantasy Futures*; *Meaning Making (The Utility of Theory)*; *Technocultural Neurostimulus*; and *Posthuman Artificial Life*. An in depth examination of these trends has been summarized in previous research [8].

These nine groups of trends represent a cross-section of *The Technocultural Condition*. They represent opportunity areas for the design of innovative VR systems and experiences. Based on these findings, brainstorm topics have been identified, concepts have been generated, and prototypes have been constructed. For example, in the case of *Multimodal Fantasy Futures*, the following questions were identified: How could perceptual robotic technologies be used to... probe experiential narrative fictions?; design for the kids of tomorrow?; incubate fantasy?; and create surreal, transcendent, breathtaking experiences? A second round of synthesis grouped these 52 key questions into clusters indicating seven major areas of action. In this way, the identified trends serve as brainstorm topics intended to do more than simply inspire the creation of new concepts—they provide a concrete point of reference by which the engineering of multimodal VR experiences may be guided towards a relevant, useful, and high-impact realization.

Concept Mapping and Elaboration

The identification of brainstorm topics drawn from *The Technocultural Condition* allows concepts to be generated that address clearly defined areas of opportunity. For example, the question “How could perceptual robotic technology be used to *speak with aliens*?” (an aspect of *[In]finite Reality*) has been asked during a brainstorm session. Each such question has resulted in around 100 raw ideas, which have then been pared down to 5-10 refined favorite concepts and elaborated through storyboards and sketches. Contributions from each of these brainstorm topics has resulted in the generation of over 250 refined concepts so far.

The process of concept ideation has been conducted both independently by members of our laboratory and as an ongoing series of group brainstorm sessions. The rules of brainstorming we have employed are: gleefully suspend judgment (don’t apply conventional wisdom), leapfrog off the ideas of others (don’t make logical connections), and go for quantity (don’t focus) [9]. All of the ideas generated during

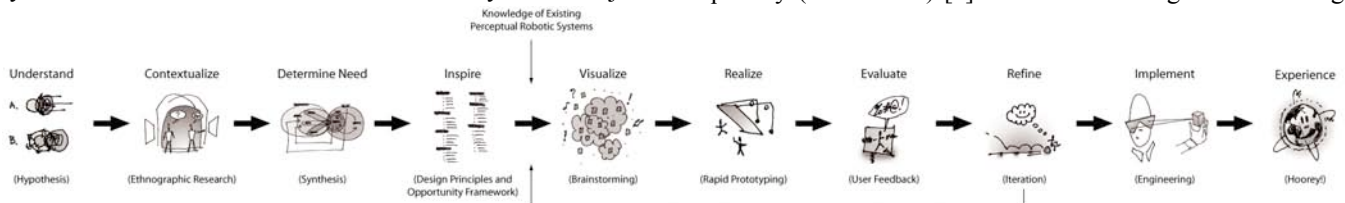


Figure 2. Design process overview for *The Ultimate Aesthetic Experience*. At a high level this may be divided into three phases: Needfinding (understand, contextualize, determine need, inspire), Formgiving (visualize, realize, evaluate, refine), and Communication (both as self-guidance and to specify details of the system to future engineers who may be implementing the system). These are loose and iterative stages (the project is currently somewhere between Formgiving and Communication).

There are many criteria by which concepts may be evaluated. Ultimately the decision of which concepts to pursue is a matter of feasibility, presumed impact, personal interest, and willingness to invest. By allowing archived concepts to be ranked, tagged, and filtered by multiple

Figure 3 shows two-dimensional output of the system, where concepts have been ranked by potential aesthetic impact vs. implementation difficulty. This map makes the strategic relationship between generated concepts clear. The five large circles on the diagram identify concepts as being out of scope, heavy investments, worth a try, easy impact, or cheap thrills. Most of the concepts classified

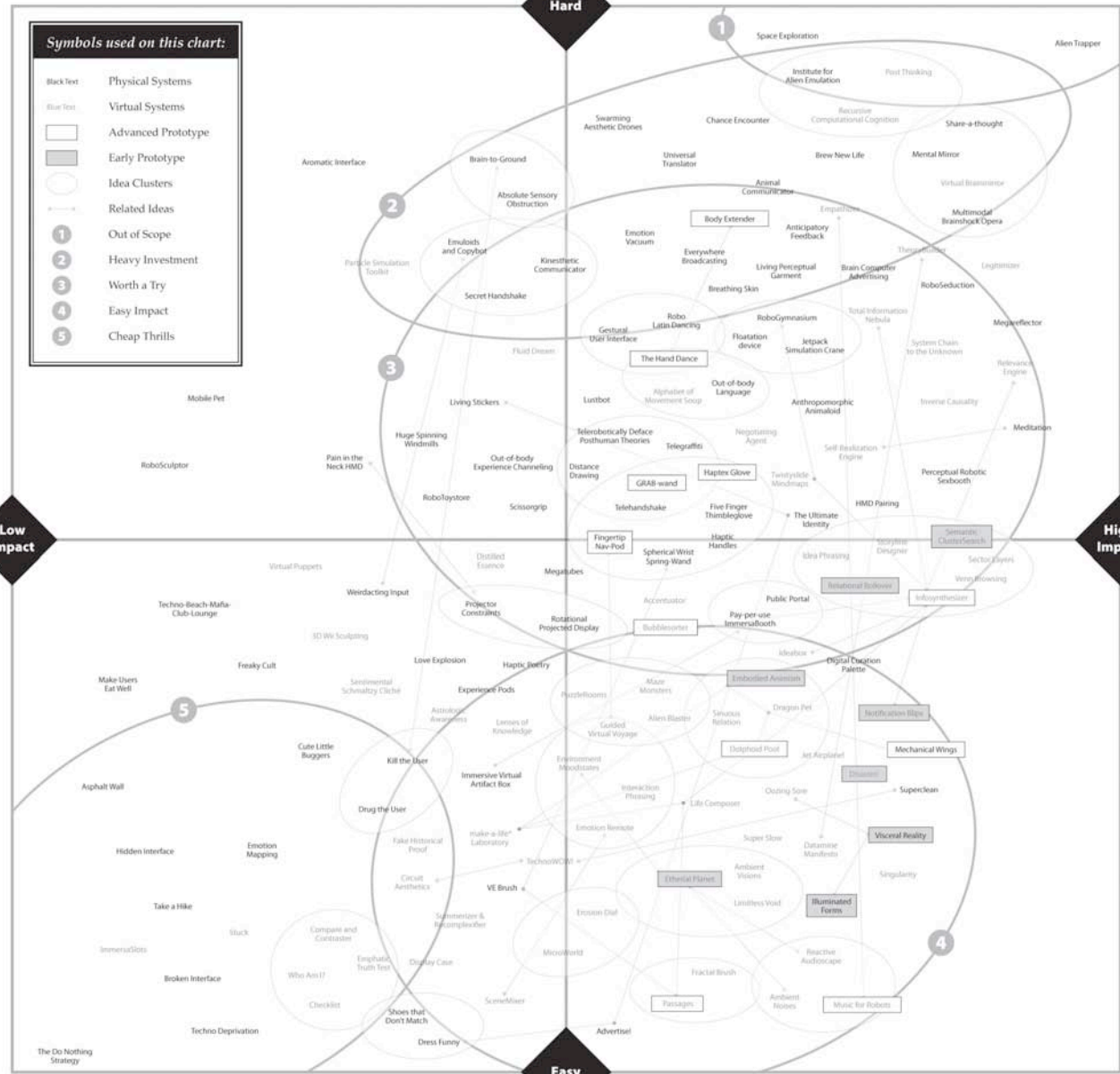


Figure 3. Over 250 of the concepts generated by our design process so far, plotted by potential aesthetic impact (low to high) vs. implementation difficulty (easy to hard).

“easy impact” may be quickly prototyped in software, while those identified as being “worth a try” involve more complicated engineering. With regards to *The Ultimate Aesthetic Experience*, our intent has been to implement a majority of the concepts in the lower-right quadrant (high impact, easy to implement), and a selected assortment of the more difficult concepts in the upper right quadrant.

The most promising concepts from this map have been expanded in depth. This process has included defining each concept and its intent more clearly, sketching possible implementations (see figure 4), and identifying potential blocks to development. The concept map has also revealed clusters of ideas that are related to one another (small circles), and suggested connections between other concepts that may not have been evident when the idea was first thought of (connecting lines). This information has proven useful to concept refinement, allowing sets of concepts to be combined or distinguished from one another. Because many of the interaction concepts have potentially overlapping features and possible implementations, storyboards have proven to be a particularly useful means by which to integrate ideas and define their differences. The public nature of our system has also been useful in this regard.

SYSTEM EVALUATION AND REFINEMENT

The ranking and evaluation of proposed concepts for *The Ultimate Aesthetic Experience* is an ongoing and iterative process. During the first round of prototyping, approximately 30 favorite ideas were selected for advancement, which were then subdivided into 5 basic groups: virtual scenes and applications; virtual features; environment aesthetics; interface hardware; and physical installations. Our main emphasis with this project has been programming everything in the first three categories. Because developing new interface hardware and physical installations requires significant implementation resources, we have limited our initial experimental setup to a basic VR system capable of realizing the majority of virtual scenes and interactions. This comprises an immersive display system with stereo vision, 3D audio, and a spatially tracked vibrotactile wand. Other interface available systems may also be used—such as exoskeletal or force-feedback robotic devices—but the experience has been designed to function well without them.

The experience has been evaluated by a small group of test users who have provided feedback by which to improve the design. At the moment these individuals have been limited to PhD students and researchers working in our laboratory. Observation findings at each stage of development have been synthesized to determine possible modifications to the interface based on real-world physical constraints, and the database system is updated accordingly after each round of evaluation. Refinement of the system is thus an iterative cycle based on evaluation of information stored in a central networked location as well as real-world user feedback via direct immersive experience. Because many of the generated concepts are extensions or modifications to existing prototypes or systems, user feedback from these sessions has also proven to be an excellent venue for communication and “cross-pollination” between researchers and projects. One of the principal aims of *The Ultimate Aesthetic Experience* has been to develop integrated systems using existing resources and to guide the coordination of research efforts at large, and in this regard the development process has been quite successful.

Experimental Setup

The system is installed in a darkened room containing a modified Nintendo Wii vibrotactile wand, immersive 3D virtual environment, audio speakers, Infitec glasses and a stereoscopic projection screen. When robotic interfaces are used, they are sensitive to input forces and internally actuated in relation to the environment such that they cooperate as a tool to provide interactive force feedback. The users’ glasses are tracked as they move in real time by a VICON motion tracking system, allowing spatial exploration of the virtual world as it is being created. Visual display is projected within an CAVE-like [4] rear-projection system driven by two pairs of superimposed high resolution projectors connected to two PCs rendering the virtual world. Real-time positional information of the user’s eyes is processed by the tracking system and passed to the rendering engine to generate the appropriate perspective for each of the user’s eyes. This allows immersive exploration of the virtual space. The environment is programmed in XVR (eXtreme VR), a C++-based scripting language designed for mechatronic device integration, high-speed graphics and online network rendering [2] (the projected stereo environment runs within Internet Explorer). MatLab

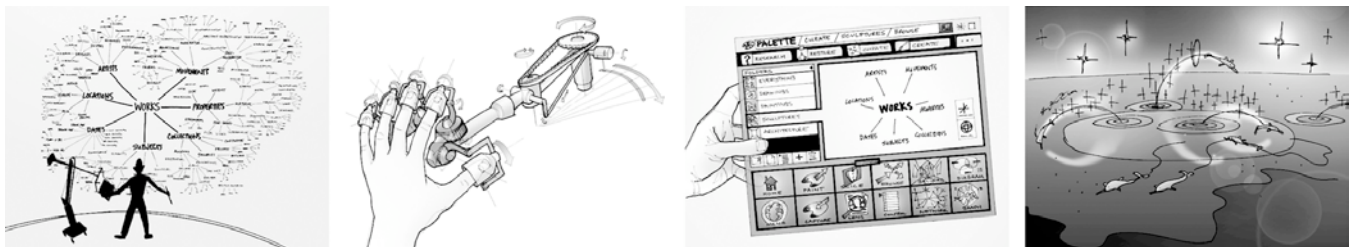


Figure 4. As new ideas have been documented they are entered into a database containing their name, a brief description, and sketches of the concept. From left to right, the concepts shown here include: *Total Information Nebula* (for 3D interactive mapping of data), *Five Finger Thimbleglove* (for grabbing and squeezing virtual objects), *Digital Curation Palette* (a hand-held digital menu with which to choose options and paint), and *Dolphoid Pool* (in which users can swim with intelligent “dolphoids” and teach them tricks).

software is also used to drive certain virtual computations (i.e., those requiring the use of artificial neural networks).

Experience Architecture

The software architecture of the currently implemented system is shown in figure 5. The environment is designed to create a smoothly transitioning “intelligent landscape” that reacts to user activity in beautiful, shocking, and unexpected ways. Three principal landscape concepts have been developed (*Cubeland*, *MicroWorld*, and *Limitless Void*), each of which can be “rendered” in one of six aesthetic modes (or combinations thereof). Each mode controls the graphic style of the landscape (shading effects, colors, contrast, luminosity, etc.), audio output (music, sound effects, etc.) and “velocity” of the scene. Three “transition mechanisms” interpret user activity over time and express the behavior of the landscape correspondingly.

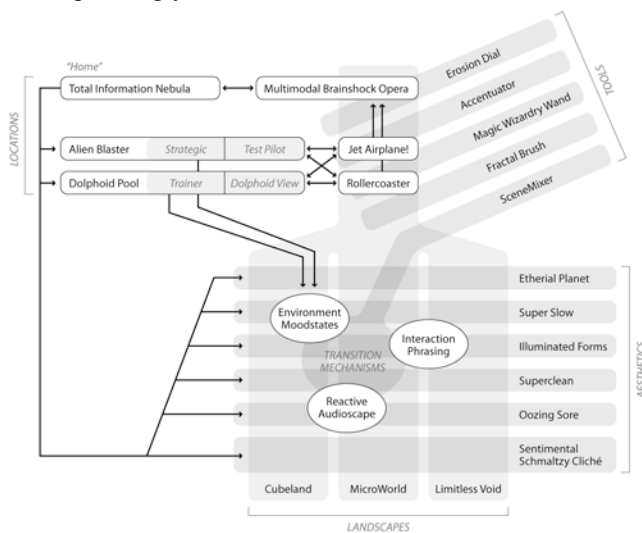


Figure 5. The virtual experience architecture of the system. Brainstormed concepts from the previous phases (black text) define the scenes and interaction paradigms of a continuously shifting aesthetic landscape that responds “emotionally” to user activity.

As changes are taking place in the environment, the user is occupied creatively experiencing and interacting with aspects of the space using the vibrotactile wand (figures 6 and 7). A variety of virtual “tools” are available for user expression, allowing creation and manipulation of objects in the environment. These include *Erosion Dial* (which “melts” aspects of the world), *Accentuator* (which intensifies color and sounds), *Magic Wizardry Wand* (which allows the user to trace colorful lines and sparkles through space with the wand), *Fractal Brush* (which creates geometric polygon structures), and *SceneMixer* (which allows limited control over the aesthetic landscape itself). Control over tool behavior changes in response to the environment, such that in certain locations of the space the tool can be used for specific manipulative tasks, such as clicking a navigation hyperlink (*Total Information Nebula*), interacting with fleet of alien spaceships (*Alien Blaster*), or

feeding dolphoids (*Dolphoid Pool*). User performance in these locations, such as training dolphoids so that they are happy and do tricks, for example, correspondingly effects the “moodstate” of the aesthetic landscape at large. Happy dolphoids may make the scenery “enthusiastic” and colorful, while scared dolphoids (or destroyed alien civilizations) will make the environment “angry” or “depressed.” Finally, the user has limited control over his or her movement through the space. Pressing the “home” key on the wand returns the user to the *Total Information Nebula*, for example, a landscape of hyperlinks referring to each of the actively implemented concepts in the experience. This scene operates as a three-dimensional menu, allowing the user to navigate directly to an aspect of the world of particular interest. Other options include “riding” on a dolphoid, spaceship, rollercoaster or jet airplane. In each of these cases, wand-based activity is used to control and influence the passing audiovisual landscape, such that the user may paint traces with the *Magic Wizardry Wand* while riding on the rollercoaster, for example.

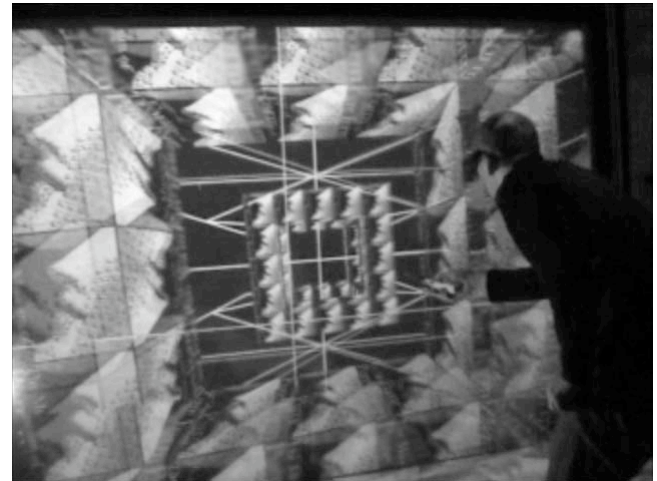


Figure 6. A scene from *The Ultimate Aesthetic Experience*.



Figure 7. A modified Nintendo Wii vibrotactile wand serves as the principal mechanism for user interaction.

Preliminary Findings

Programming the virtual architecture described in the previous section is a complex, time-consuming, and ongoing activity. Although the fully realized experimental platform we envision is not yet complete, many observations can be drawn from its development so far. Evaluation of the system has been carried out through non-structured interviews during the interaction, allowing users to express their emotions and reactions to the installation as their knowledge of it grows. The interaction paradigm we have implemented expressly avoids traditional interaction metaphors such as mouse/keyboard interaction in favor of a simple 3D pointer with two principal buttons (the Wii's "A" and "B" keys). Participants have been provided with a very basic explanation of the installation setup, limited to instructions such as "wear the stereoscopic glasses," "move your hand to trace lines in the space," and "push these two buttons to interact with the world." Most users have interacted with the installation for an average of 10 minutes, during which time more instructions were provided if the user seemed stuck (i.e., "you can ride on the dolphoid if you click it"). Direct observation of users experiencing the system for the first time has been particularly insightful since it provides evidence as to whether aspects of the interface are intuitive.

The Ultimate Aesthetic Experience has been designed to be stimulating, provocative, adaptive, exciting and fun. In this regard the demo has been extremely successful. For users who have never experienced stereoscopic virtual reality systems, the simple thrill of being allowed to walk through even the most basic virtual landscape can be a powerful sensation. The "magical" quality of the evolving and interactive scenery has provided further curiosity and motivation to explore. In particular, creatively interactive aspects of the interface, such as those that involve "painting" objects and sounds into the space (e.g., spirals) and then walking within them, have been especially enjoyed by the majority of users.

Most individuals have been surprisingly comfortable exploring the virtual environment in spite of its non-representational and ambiguous spaces. The notion of VR as an entertainment/game seems intuitive and familiar. Because the environment keeps changing, many users have commented that they are motivated to keep exploring because they might be "missing something" somewhere else. In this regard, the correlation between the environment's "mood" and a user's activity is a subtle characteristic requiring additional work.

Flying through the space on "autopilot" has been a particularly enjoyable pastime for most users. The virtual track followed by the rollercoaster, for example, is constructed of numerous discrete sections (or "rooms") placed in sequence. The sequence of these rooms is generated automatically such that when experienced the demo gives the illusion of a continuous ride through a series of the "emotional" aspects of the landscape. This

modular setup allows for various combinations of rooms to be experienced, and for carefully orchestrated multimodal feedback to be displayed. Positive reactions to this aspect of the experience has led to the inclusion of many near-miss obstacles and unexpected turns along the camera's path. These have been designed to create "misaligned mental models" with regard to expected behavior. There is no need for the viewpoint to follow the tracks, for example, and the same "segment" of graphic elements can occur twice with a different behavior. Basic sound effects triggered at intervals along the course of the journey contribute to the user's sense of moving through the environment and have been shown to radically influence the viewer's perception of speed. The enjoyability of these experiences (or extreme cases of surprise, fear, vertigo, etc.) would seem to be correlated directly with a user's sensation of presence in the space.

In summary, the installation has been demonstrated to encourage and motivate spatial exploration and creative expression. The added value of being "fun" to interact with, as noted by the majority of the users who experienced it, seems a deceptively simple yet remarkably valid foundation for cognitive research. Previous research has shown, for example, that aspects of the experience would be particularly suitable as a tool for rehabilitation, especially for children, due to its playful and colorful aesthetics and "magic wand" interaction metaphor [11]. Indeed, environments in which users have "immediate," movement-based control over sensual feedback of interesting content have been shown capable of providing significant therapeutic value, since they supply individuals with a level of self-motivation and coordination that may not otherwise be expressed [18].

CONCLUSIONS

This article has discussed the concept and implementation of *The Ultimate Aesthetic Experience*, an ongoing project exploring the expressive potential of multimodal environments. Most experimental setups for cognitive or engineering research involve controlling an experience to be specific and limited. By holding as many variables as possible fixed, for example, single parameters (such as spatial memory in relationship to the perception of time) may be tested in a variety of situations. This project has taken a different approach, one which engages human cognition more "realistically" within multimodal worlds of creative interaction. In particular, we maintain that immersive experiences of this nature can and should be used to intensify perceptual and cognitive processes. The interface we have developed is based on human interaction with the world as a continuous, body based activity. As the future of computing becomes increasingly multimodal, continuous, overlapping, and multifaceted, virtual skills will become increasingly central to "everyday creativity." The truly ultimate aesthetic experience would be one which, like our experience of childhood, is a continuous and incessantly stimulated learning experience.

Char Davies has noted that immersive virtual art experiences can lead to profound sensations of joy or euphoria, verbal indescribability, and a paradoxical sense of being both in and out of the body [6]. A limit to these effects, however, is that “the medium’s perception-refreshing potential is possible only to the extent that a virtual environment is designed to be unlike those of our usual sensibilities and assumptions.” It would thus seem that, in creating an aesthetic experience capable of continuously evolving, situations familiar enough to provide a strong sense of presence would be allowed to remain “consistently strange.” This hypothesis seems to be supported by our current experiments although much further work in this area is needed, especially as the experience becomes one that can remain continually novel for lengthy durations. We believe that the realization of such “learning environments” would be best addressed by exploring beyond the limitations of discrete applications and domains, focusing instead on the skill-potential of natural movement and the expressions capable through uninhibited gesture, particularly in intelligent and/or networked multimodal environments.

Creativity deals with the experience of reality, not specific instances, methods, media or tools. Perceptual robotic art, such as installation described in this paper, inherently concerns itself with the perception and cognition of and by robots, the behavior of robots and humans in their presence, and the fantastic visions of both that result. Artists can become the parents of robots, manipulating technology in the creation of life, or collaborators with them as tools to help creative cognition emerge. In this project we have pragmatically avoided a premature definition of the robotic body and mind such that its behavior may emerge from the cultural, psychological, and physical constraints we experience through the sharing of artistic cognition.

Using robotics and virtual technologies as a medium, today’s artists have an unprecedented opportunity to design and express new kinds of emotion, perception, and consciousness. Because issues relating to the emotional sensitivities of these new forms of artificial life will become increasingly important areas of research—their feelings, dreams and self perceptions—we must teach our robots empathy, self expression, and aesthetic awareness. Our robots should be nurtured with new kinds of emotional experiences and artistic environments, ones which—like the analogy of playing Mozart to newborns—are capable of touching the robotic heart.

REFERENCES

1. Bergamasco, Massimo, Antonio Frisoli, and Carlo Alberto Avizzano, “Exoskeletons as Man-Machine Interface Systems for Teleoperation and Interaction in Virtual Environments,” in *Advances in Telerobotics*, Berlin: Springer/Heidelberg, 2007.
2. Carrozzino, M., Tecchia, F., Bacinelli, S., Cappelletti, C. and Bergamasco, M., “Lowering the development time of multimodal interactive application: the real-life experience of the XVR project,” *Proc. ACM SIGCHI International Conference on Advances in computer entertainment technology*, 2005, pp. 270 - 273.
3. Cini, G., A. Frisoli, S. Marcheschi, F. Salsedo and M. Bergamasco, “A novel fingertip haptic device for display of local contact geometry,” *Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, First Joint Volume, 2005, pp. 602-605.
4. Cruz-Neira, C., Sandin, D.J., DeFanti, T.A., Kenyon, R.V., Hart, J.C., “The CAVE: audio visual experience automatic virtual environment,” *Communications of the ACM*, v.35 n.6, 1992, pp.64-72
5. Cytowic, Richard E., *Synesthesia: A Union of the Senses*, MIT Press, 2002
6. Davies, Char, “Changing Space: Virtual Reality as an Arena of Embodied Being,” in John Beckmann (ed.), *The Virtual Dimension: Architecture, Representation, and Crash Culture*, Princeton Architectural Press, New York, 1998
7. Edelman, G. and G. Tononi, *A Universe of Consciousness: How Matter Becomes Imagination*, Basic Books, New York, 2001
8. Faste, Haakon and Massimo Bergamasco, “A Strategic Map for High-Impact Virtual Experience Design,” *proc. SPIE*, vol. 7238, 2009, pp. 723-805.
9. Faste, Rolf A., “A Visual Essay on Invention and Innovation,” Keynote article, *Design Management Journal*, vol. 6, no. 2, 1995.
10. Faste, Rolf A., “The Role of Aesthetics in Engineering,” *Japan Society of Mechanical Engineering*, 1995.
11. Ghedini, Fiammetta, Haakon Faste, Marcello Carrozzino and Massimo Bergamasco, “Passages: an Immersive 3D Interface for ArtAbilitation,” in Athanasios Vasilakos and Tony Brooks (eds.), *International Journal of Arts and Technology*, in press, 2009.
12. Hecht, D., M. Reiner and G. Halevy, “Multimodal Virtual Environments: Response Times, Attention, and Presence,” *Presence*, vol. 15, no. 5, 2006, pp. 515-523.
13. Inhelder, B. and J. Piaget, *The Growth of Logical Thinking from Childhood to Adolescence*, New York: Basic Books, 1958.
14. Kelley, Tom, *The Art of Innovation*, Doubleday, 2001.
15. Kohlrausch, A., and S. van der Par, “Audio-visual interaction in the context of multimedia applications,” In: Blauert, J. (ed.), *Communication Acoustics*, Berlin, Germany: Springer Verlag, pp. 109–138, 2005.
16. Lakoff, George, and Mark Johnson, *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*, Basic Books, New York, 1999.

17. Lieberman, Jeff, and Cynthia Breazeal, "Development of a Wearable Vibrotactile Feedback Suit for Accelerated Human Motor Learning," *Proc. IEEE International Conference on Robotics and Automation*, Roma, Italy, 2007.
18. Lincoln, Y.S., and Guba, E.G., *Naturalistic Inquiry*, Sage Publications, Beverly Hills, CA, 1985.
19. Magennis, Mark, Bláithín Gallagher, John Wood, Elena Cano Arias, José Muños Sevilla, Keith Gladstone, Helen Graupp, Teresa Gutierrez, Carlo Avizzano, and Massimo Bergamasco, "GRAB—A New Haptic and Audio Virtual Environment Enabling Vision Impaired People to Access the Three-Dimensional Computer Graphic World," in *Assistive Technology—Shaping the Future*, G. Craddock et al. (eds), IOS Press, 2003.
20. Malina, R., 2008, Online Yasmin Mailing List, available at: <http://www.media.uoa.gr/yasmin/viewtopic.php?t=3875>, [accessed 11/11/2008].
21. McKim, Robert, *Experiences in Visual Thinking*, Brooks/Cole Publishing Co, Monterey, California, 1972.
22. Palmer, Stephen E., Joseph L. Brooks and Rolf Nelson, "When Does Grouping Happen?," *Acta Psychologica*, vol. 114, no. 3, 2003, pp. 311-330
23. Patnaik, Dev and Robert Becker, "Needfinding: The Why and How of Uncovering People's Needs," *Design Management Journal* (1999).
24. Ramachandran, V. S., and E. M. Hubbard, "Synesthesia: A Window into Perception, Thought, and Language," *Journal of Consciousness Studies*, vol. 8, no. 12, 2001, pp. 3-34
25. Rosenfield, I., *Neural Darwinism: A New Approach to Memory and Perception*, The New York Review of Books, October 9, 1986.
26. Schulz, Melissa F., and Thomas Sanocki, "Time Course of Perceptual Grouping by Color," *Psychological Science*, vol. 14, no. 1, 2003 pp. 26-30
27. Sheridan, T., "Musings on Telepresence and Virtual Presence," *Presence: Teleoperators and Virtual Environments*, vol. 1, no. 1, 1992, pp. 120-126.
28. Slater, M., Usoh, M. and Steed, A., "Taking Steps: The Influence of a Walking Metaphor on Presence in Virtual Reality," *ACM Transactions on Computer Human Interaction (TOCHI) Special Issue on Virtual Reality*, September 1995
29. Slater, M., Linakis, V., Usoh, M. and Kooper, R., "Immersion, Presence, and Performance in Virtual Environments: An Experiment using Tri-Dimensional Chess," *proc. ACM Virtual Reality Software and Technology*, 1996
30. Valera, Francesco, Evan Thompson and Eleanor Rosch, *The Embodied Mind: Cognitive Science and Human Experience*, MIT Press, 1991.
31. Zeki, Semir, *A Vision of the Brain*, Wiley-Blackwell, 1993