Robotic synchronized to human gesture as a virtual coach in (re)habilitation therapy

Brooks, Tony

Published in:
3rd International Workshop on Virtual Rehabilitation (IWVR2004)

Publication date:
2004

Document Version
Publisher's PDF, also known as Version of record

Robotic synchronized to human gesture as a virtual coach in (re)habilitation therapy.

Associate Professor Tony Brooks

Software and Media - Computer Science Department,

Aalborg University Esbjerg, Niels Bohrs vej 8, 6710 Esbjerg, Denmark.

Telephone: +45 79 12 77 16, Fax: +45 79 12 77 10

E-mail: tonybrooks@.cs.aue.auc.dk,

Abstract

This paper describes the latest developments of the SoundScapes [Brooks 2004] body of research which is utilizing technology in assistive rehabilitation and habilitation from Virtual Interactive Space (VIS) [Brooks 1999]. This paper exhibits the implementation of robotic physical movement synchronously manipulated from sourced data movement information of a human. The user groups are children with severe physical/multi disabilities. The sourced capture of the human data is from enhanced virtual interactive space created from sensors. The results highlight the positive effect of physical control by those with limited facility and conclude at the potential of implementation as a supplement to traditional therapy techniques.

Keywords

Robotics, Memory, Cognition, Gesture, Motivation, Consciousness, Play, Interactive, Empowerment, Control, Sensors.

1. Introduction

If one searches for the term <virtual coach> on the World Wide Web it brings up links that offer '*telephone coaching*'; '*CyberCoaching*' (by e-mail); and – "*combinations using state-of-the art conference call systems for interactive and dynamic TeleForums!*"

This paper is suggesting a somewhat different approach for the investigation of a virtual coach. An approach away from The Web where through empowering human gesture to control synchronous robotic movement and multimedia feedback severe disabled children are encouraged to "play" utilizing whatever physical ability they may have. This "play" element is the apparent and upper-most level which is perceived and which encourages immersive state through interaction. Other not so easily perceived levels are inherent which through this study suggest at potentials as a new tool for therapeutic supplement as a virtual coach. This presented study is contained within a larger body of research titled SoundScapes which is based on the stimulation of the human consciousness through proactive interactive "play" experiences and empowerment. In SoundScapes new technologies are utilized to enable controlled manipulation of physical robotic devices, computer programs and other interesting feedback. The research has been expertly appraised as giving potential in human afferent efferent (often called sensorimotor) neural loop closure (see figure 1).

A "virtual coach" for therapy should be complimentary so as to contribute as a supplement to traditional practices. It should in no way deter from the therapy or interfere with the patient's well-being. It should be motivational and inspiring for both therapist and user such that in certain instances, non-verbal communication and expressive cues, which are inherent through fun interactions, are encouraged and utilized in the therapy.

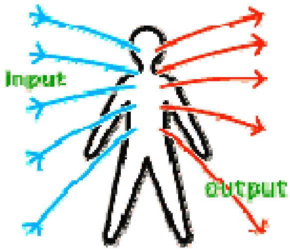
Afferent (sensory)

Vision

Hearing

Smell

Touch



Efferent (motor)

Voluntary Corticospinal

Basal Ganglia

Cerebellar

Autonomic

Figure 1: Human Neural Loop.

When working within the human sciences certain individuality must be adopted so as to account for varying faculty, preferences and limitation. An ability by the therapist to improvise within the 'patient-therapist' sessions is an obvious advantage and this comes about through knowledge of the tools and the subject. The content of this paper details the use of 'intelligent' robotic devices which respond directly to human gesture and which can supplement the "tool box" of the contemporary therapist. Through empirical testing the author has developed a theory where such devices are suggested as being a potential interesting 'virtual coach' for therapeutic activities. This theory, while utilizing robotic devices that are very different from the state of the art (see section 2.1), shares common keywords to related research in the field. The added real-time synchronized response, however, that is achievable through the mapping of captured human gesture in 'activated air' to robotic movement gives an opportunity for a new body of interdisciplinary investigation to explore and advance the field.

2. Background

The concept and methodology of SoundScapes began through interactions with severe disabled family members. It was developed into a research body of work from around 1993 following its initial seeding period from 1986 where proof of concept experiments began. The author claims tacit knowledge in the field from almost half a century of interactions with his disabled relatives and subsequently through his research in the field. This paper is a legacy from sessions with the same group of children that were investigated in a Panorama Virtual Reality environment in 2001 where they were studied controlling a space ship through gestured head movements [Brooks et al 2002]. That study titled TWI-AYSI (The World Is As You See It) was the feasibility of a resulting European funded project titled CAREHERE[1] (**C**reating **A**esthetic **R**esonant **E**nvironments for **H**andicapped, **E**lderly and **RE**habilitation) [Brooks et al 2004] which was based around the work. However no robotic devices were researched in those projects as is evident from the publications.

2.1. Interaction, Therapy, and Theory

Studies on the interaction between humans and human, and human and machines point to the fact that through interaction a level of understanding is achievable that is often relative to the length of the time of interaction. Longer interactions usual lead to a desire for subsequent interactions which stimulate; resulting in motivations of behavior being generated.

In the last decade robotic devices have been created specifically for human interaction so as to motivate subsequent behavior [Shibata 1997]. These devices have been used in children's wards and elderly wards in hospitals as well as at senior-citizens homes in 'robot-therapy' sessions. In Japan, where positive results from the inquiry is reported with many users referring to the devices as "cute" and having a preference for time passing by interactive "play with robots" over the usual switching on the radio or television.

Scientists are positive - towards the ideal results of potential huge savings in medical costs, reduced burdens on families and caretakers, and healthier and happier aged - as forecasts predict of an overwhelming elderly population in the future[2]. The device utilized in this study was not designed for use in hospitals, or homes, or created specifically for human interaction such as researched here. The device is a stage light robot.

Robot therapy is not new. Hogan [1996] details the training of an arm of a person disabled by a stroke which is interacting with a device resembling a robotic arm which could measure the force being exerted by the user.

[1] http://www.bris.ac.uk/carehere/

[2] *In gadget loving Japan, robots get hugs in therapy sessions.* USA Today 11/4/2004 (The Associate Press, journalist Yuri Kageyama)

This study had a different objective which was to create a wireless non tangible link between the human and robotic coach and to create an interaction where the user feels in control. Empowerment of all users to be able to establish and maintain such a sycophantic relationship with a robotic agent without being encumbered with attachments and wires is paramount to the study. The theory offered from this research is towards the virtuality of potential opportunities in therapeutic coaching from a simple interaction – a human having fun controlling an interactive synchronous responsive robotic device.

2.2. Play, Interaction, Creation and empowerment.

A child moves his head 5 millimeters and three robotic devices physically respond and rotate their projection heads in a synchronized whirling 'ballet' of gestured interaction resulting in a "dance" of lights across the walls and ceiling of the laboratory. Correspondingly the same gesture "plays" a musical sequence through a sound system. The child follows the lights with his eyes and moves his head accordingly; he stills, looks and listens, smiles and continues to explore the invisible sensor space of the system exhibiting playful immersed interactive control of the robotic lights until exhausted.

The definition of play, as offered by the great Danish scholar of child play, Erik Erikson [1977] in his Harvard Godkin Lectures, is "*the illusion of mastery over life's circumstances.*" Play is closely related to creation, and when watching children at play it is often amazing how they fantasize through role creation, scene creation and interactions where they create together. As adults, the creativity and play are often mostly associated to artistic pursuits such as in music, theatre, sculpture and painting or to sports. The aesthetic philosopher and play commentator Karl Groos (1898) is credited with recognizing the comparisons between play and artistic activity. He regarded play as a simplified form of artistic creation. Although the research was deemed misguided, naïve and incomplete [Beach 1945, Loizos 1966] - the author questions if Groos' findings were so unfounded due to the findings in his own research that one may reflect upon in respect to play, creation and artistic activities.

This paper is written nearly 30 years after E. O. Wilson (1975 p. 164) had observed, "no behavioral concept has proved more ill-defined, elusive, controversial, and even unfashionable than play," the author reflects on the fact that "play" has evolved from being "unfashionable," to being a learned subject of inquiry and of great interest for scientists possibly explicitly due to its elusive and controversial qualities. Robert Fagan (1981 p. 5) in his contributing publication purposed to the definition of a biological approach to animal play in relation to scientific knowledge about social behavior and development, states the case of "*play behaviors representing structural transformations and functional rehearsals or generalizations of behaviors or behavioral sequences*" depending on the characterization used and on underlying assumptions about behavior classification and how, relative to context, play behaviors "*yield relatively specific and immediate beneficial effects*." Behavior indicating play sequences and structured transformations are in our case based on gesture of the user where it is clearly observed that a deliberate exploration of the created 'virtual interactive space' is subjected to transformations between sequences whereby the user is able to detect the causality of affecting interaction with the robotic device. Fagan further suggests that '*using a tool*' is an example of such effects and within this study we focus upon the use of our robotic device centered system as the "tool" for the therapist to utilize as a virtual coach towards the yielding of beneficial effects. Such *beneficial effects* are specific to each individual user.

3. Objectives

The objective of this study was to observe interaction between children with severe disabilities and responsive 'synchronous to gesture in real-time' robotic devices to determine if there is a potential as a virtual coach. In this way it may be possible to define characteristics of a robotic virtual coach for use in rehabilitation and habilitation, as a supplement to traditional therapy. Such characteristics are suggested as an inherent layered approach which has at the top-most level an interactive causality of "play" environment that is experienced proactively by the user. The environment has feedback qualities which are tailored to each individual user's faculties, preferences and desires so as to be conductive to immersive interaction. Creativity is an element of the concept and methodology. So too are the task oriented interactions which can explore the limit of the user (by proximity) relative to the user's therapy goal. An example of this strategy is where a hand movement is being trained and it is focused as being interactive to a feedback which achieves a level of success for the user. Such as playing a series of music tones, painting a color or manipulating a robotic agent. As the sessions progress the interactive space which enables the success factor is correspondingly more difficult to reach as the therapist sequentially directs. In this way the user is self-motivated towards pushing their limit to reach a previously attained goal, by extending their limit, which is the goal of the therapy. Thus the levels, content and strategy co-exist and the objective of the research is to ascertain if such a theory has valid grounding for further inquiry.

4. Methods

The institutes involved were asked to volunteer children who were all able to see and hear. It was not determined how much they could see but only that they were able to follow objects with their eyes. The children were to be selected from the group that is classified in Scandinavia as multi-disabled. At all sessions there was a knowledgeable helper/assistant close by to ensure no discomfort for the child. Four children were selected in Denmark and eight were selected in Sweden at the start of the full study. The range of ages was from 3 and half years old to seventeen years old. The four children in Denmark were the only ones that used robotic synchronized control to their gestures, their age ranged from 3 and a half years to 5 years of age and they were all male gender. The extent of this study focuses on these Danish children and the interaction to the robotic interaction as a potential "virtual coach". Previously the same Danish children were investigated in a Virtual Reality (VR) panorama environment corresponding to the children in Sweden who had only sound and visuals and vibrations. A reflection to the earlier sessions with the Swedish children is referenced in this paper. The earlier report on the study on the Swedish children without robotics and the Danish children with the VR is referenced in publication [Brooks *et al* 2002].

4.1. Session method

It was decided to set up a basic child assessment coding scale for each session by asking the assistant how the child was at the start of every session, how he was perceived by her during the session, and how he was following return to the institute. In the analysis of the session video tapes a coding was set up to establish during the sessions what was observed. These indicators of the user's sense of awareness were compiled into giving an insight to the use of the robotic device as a potential virtual coach for therapy. The knowledge of the user who is involved in each session is also important as a predictive estimate can be visualized for what range of signals will be obtained from the body gesture in the sensor space. This determines sensor selection and set up of the sensors as well as the programming of the range of the light pan and tilt to ensure that a full turn of the head facility can view the full extent of traverse of the light pattern. This encourages the virtual coach to motivate the user to exercise the neck by turning of the head as well as the gesture selected being exercised. Sound selection was also as issue as to whether the child liked aggressive or smoothing sounds. The assistants were able to give advice on the child's preferences and dislikes as well as abilities and limitations.

Hardware –

An 8 channel moving light controller (Elektralite CP10) capable of translation of MIDI to DMX 512 was central to the system. The 'intelligent' robotic devices that project light patterns are controlled by the DMX 512 protocol. Testing prior to the user groups being involved was with robotic intelligent lighting devices supplied by the Martin Group, a global leader in the field of stage lighting, so as to choose the optimal units for the study.

Unit reference	Dichroic Colors	Gobos	Type	Pan (x)/Tilt (y) micro-step
Roboscan 812	11 (+w)	11	Scanner	175 x 83
Roboscan 518	17 (+w)	5	Scanner	180 x 90
Roboscan Pro 218	17 +2 (+w)	18	Scanner	180 x 90
Robocolor Pro 400	32 (+w)	3 (7/14/21)	Effect	0
Destroyer X250	7 (+5w)	12	Effect	0
Robocolor 11	11 (+w)	0	Effect	0
MiniMAC Profile	**12**	**7**	**Moving head**	**540 x 270**

Table 1: Characteristics of the tested robotic devices.

The robotic device chosen was the moving head MiniMAC Profile intelligent lighting unit that projects multi-color light patterns of high contrast at up to 540 degrees of pan and up to 270 degrees of tilt. Such light patterns for a child may relate to a Tivoli, firework display or other fun experience so as to achieve the desired "play" feel of the interaction. However in this study the child (as user) was empowered to control the "fireworks!" For the research sessions the Martin group supplied three MiniMAC Profile moving head units.

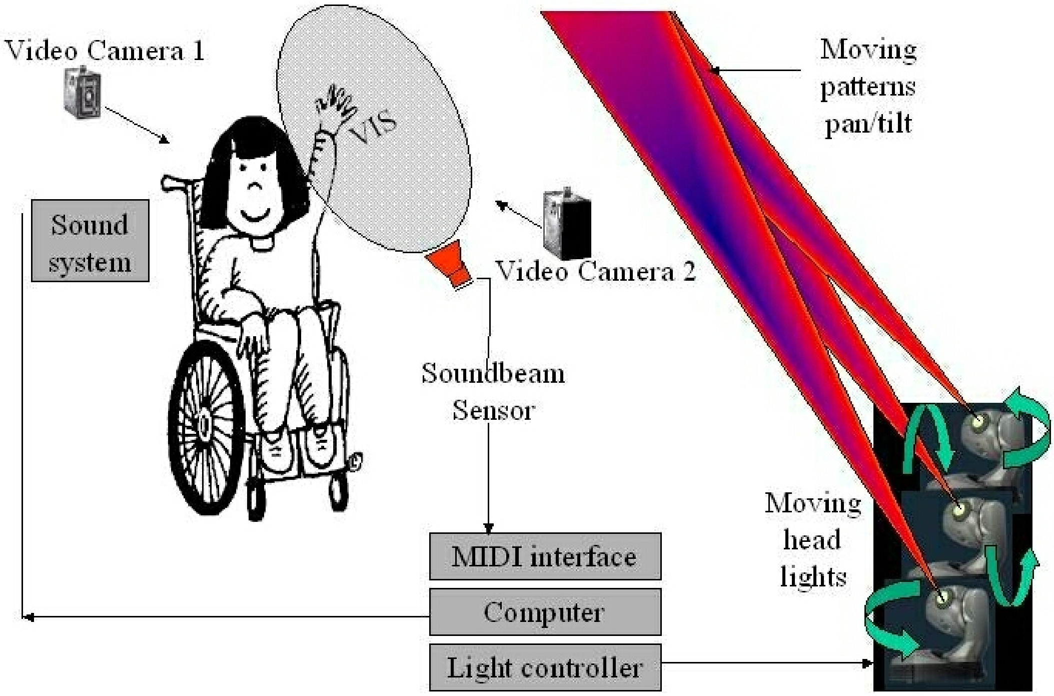


Figure 2: Set up of system. Human gesture captured by sensor for feedback manipulation.

The set up of the sessions was in a large empty room so that the full range of the robotic moving heads could be programmed to extend beyond the users peripheral field of view so that a head movement, (= neck exercise) was required to observe at full extremities. As shown in figure 2, the user was positioned near the center of the room with a camera behind to capture the whole scene. A second camera was positioned in front to capture gesture and facial expression. A sensor (a Soundbeam ultrasound unit – see also figure 3) is set up according to the user faculty of movement – shown in figure as hand and head gesture capture. Other sensors used were infrared volumetric sensors (author prototypes – figure 4). Gesture signal from the sensor capture is transmitted as a MIDI[3] signal to an interface, PC and a sound module and speaker array. The software algorithm used for mapping of the MIDI signal is sent to the lighting controller for translation into the light protocol of DMX 512.

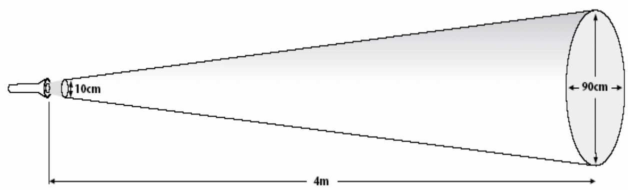


Figure 3: Soundbeam linear ultrasound sensor.

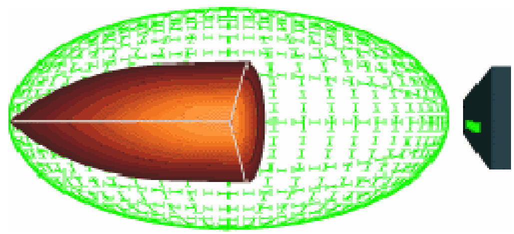


Figure 4: Infrared volumetric sensor.

Software –

Programming the units to respond to MIDI was with the use of an object oriented graphical programming environment which is used for music, audio and multimedia. This software (Max / MSP) is an industry standard for the mapping of MIDI data stream information. The mapping of the data information to the robotic movement and end lighting effect is a main task when setting up the system. Knowing the range of each lamp and each specific channel of communication to each unit is imperative so as to avoid confusion. Each time the units are set up in a new location a new system set up has to be programmed. On average at least half a day is needed to set up and then another half day for fine tuning the system.

[3] MIDI = Musical Instrument Digital Interface, an industry standard for communication between machines.

4.2. Analysis method

To analyze the videos from the sessions both camera views were synchronized so that it was possible to correlate the manipulation of the robotic heads, the manipulation of the projected light patterns on the facing walls and the gesture and facial expression of the user. A coding system was established for each child with help from the helper who was familiar with the various responses. As best as possible we tried to generalize common expressions accounting for each child having individual faculty limitations and abilities, typical features such as a smile, a mouth opening, a quieting, an eye focus, a frown, hand or lower torso movement.

5. Results

The three robotic moving heads are manipulated by a 5 mm gesture of the disabled child's head. The projected light patterns on the facing white wall move approximately 20 meters from their prior position. The ratio and sensitivity of gesture to resulting feedback is totally programmable and limited only by the physical constraints of the hardware and room location. The child is immersed in the dynamic exploration of what is happening under his control and is motivated to explore further through varying the range, the speed and direction of his gesture – he indicates awareness of a direct correspondence and control to the physical movement of the robot head and the subsequent movement of the lights.

The twelve picture sheet in the appendix (Figure 5) illustrates moments from the sessions where total immersion was apparent. These pictures alone cannot tell the whole story but only hint at the explorations and new experience from controlling robots that the young children with severe disabilities succeeded in achieving.

Most of the sessions lasted around 30 minutes, with the shortest at 11 minutes and the longest at 46 minutes. Each session timeline involved dynamic interaction and response showing a recurring pattern. One child was not comfortable in his chair so we accommodated by placing him of a floor mat that was adjusted for him to be able to see the light patterns. One male was asleep on entering the VIS in one session. Slowly he woke and explores the play of moving his head which is moving the physical robotic device, the projected light patterns and playing the sounds. The four young friends exhibited immersion in every session showing a distinct consciousness of intent.

As mentioned, the robotic devices used for this study are not intended for such use. They are adapted by the author for this study. This enabled an otherwise prohibitive investigation viable through corporate sponsorship of the devices. The choice of the MiniMac profile robotic intelligent light devices was astute. The units responded with a latency of around half a second which was acceptable but not optimal. We tested various gobos (patterns) and colors but could not ascertain if they were making a difference to the child.

Observations of the child following the patterns gave us an indication of his awareness. His facial expressions and exploration of the virtual interactive space suggesting that to some extent he was aware that he was empowered to manipulate the robotic device which resulted in the projected light pattern, however the author suggests further that his only conscious awareness was of the light patterns and the audible movement of the robotic device was a more sub-conscious element to be an associated element.

The sessions followed a recurring pattern similar to that statement accredited to Hutt (1966) in his publication on the subject of exploration and play in children, where it was quoted -"*inspection gives way to play, the emphasis changes from the question of "what does this object do?" to "what can I do with this object?"* The study shows that along similar lines the sequence is hereby extended with the author's subjective appraisal with further emphasis change of *"when I move - the light patterns move:" "when I move - I hear sounds:"…" when I stop moving I hear neither sounds or see the light patterns moving!"* … *"Hey I'm in control here- and its fun!"* And further to suggest that the emphasis is extended to *"well nobody told what I should do, or for how long, so I will just have more fun with what I have learnt I can do with this object!"* In the sessions all children continued to explore until exhausted. The report from the assistant was positive from sessions and after return to the institute.

This is significant in respect to the goals of the author's research. By giving a means of expression, control or authority to a user, one is giving empowerment over an entity; such characterization attributes applies to an envisioned virtual coach system. In this study the empowerment can be observed through the physical signs as indication of the child's awareness of causality and this is what is useful for the exploration of validity of the author's virtual coach theory with the utilization of robot agents which with this study has reached its objectives.

6. Discussions and Conclusions

On reflection of the sessions, the set ups could have been improved, especially the VIS sensor locating for gesture capture from the human user. The sensor needs to be remotely controlled from a distance so as not to interfere with the user. In these sessions two set ups were used (1) three infrared sensors were used one for head and two for upper torso/hands according to faculty (2) a single ultrasound for head. An overhead sensor positioning is preferable. The optimal set up was with the single sensor for direct associations. This was because it gave a predictable space of interaction and avoided any confusion for the user. Being linear it also gave a progressive response to the gesture (light patterns, corresponding robotic moving heads and sound) as opposed to the volumetric data space of the infrared sensor which was unpredictable for the user as the same data is produced from different movements across the various axes.

With the current technology it is possible for the movement 'point of interference' to be between two events such that to the naked eye there is no perceived action yet two events are triggered. This is accounted for in the summary, however, after much analysis of the videos, occasionally some of the children hinted, through facial expression, at finite control whereby they could have been controlling without it being observable. The author questions if a heightened perception was possible from them in such instances of finite control.

The SoundScapes body of research utilizes cameras as well as sensors (infrared, ultrasound etc.,) to capture user body movement. As the robotic device was projecting light patterns to achieve optimal effect for the users the rooms were in darkness for the sessions. The light changes in the dark space made it not viable to use camera systems for user capture. However such a future set up is expected to give additional information for the study.

Of note for readers who would wish to emulate the sessions is that the set up should always be prepared before the user enters into the space as often it is boring for them to wait and occasionally they fall asleep.

Acknowledgements

The author wishes to thank institutes involved namely: Fenrishus, Aarhus, Denmark; Emaljskolan, Landskrona, Sweden; CAVI (Center for Advanced Visualization and Interactivity), Aarhus, Denmark. Research assistant Stefan Hasselblad, Sweden and the wonderful boys and girls that were involved in the SoundScapes research.

References

Beach, F. A., **Current concepts of play in animals**, Am. Nat. 79: pp 523-541. 1945.

Brooks, A.L. Inhabited Information Spaces 'Living with your data' – ***"Soundscapes"*** (pp. 89-100) Springer ISBN 1-85233-728-1 (2004)

Brooks, A.L., Hasselblad, S., Camurri, A., Canagarajah, N. ***'Interaction with shapes and sounds as a therapy for special needs and rehabilitation'*** - ICDVRAT 2002 - The 4th International Conference on Disability, Virtual Reality and Associated Technologies In proc. (pp. 205 – 212) 2002.

Brooks, A.L. '**Virtual Interactive Space**,' Proc. World Confederation for Physical therapy WCPT, Yokohama, Japan. In proc. pp. 66, 1999.

Erikson, E. **Toys and Reasons.** Harvard Godkin Lectures, MIT, Boston, USA. 17-18, 1977.

Fagan, R. **Animal Play Behavior**. Oxford University Press ISBN 0-19-512760-4, 1981

Groos, K., **The play of animals**. Trans. E.L. Baldwin. D. Appleton, & Co., N.Y. 1898,

Hogan, N., MIT, Boston, USA "**MIT-Manus robot aids physical therapy of stroke victims"**. 2000.

Loizos, C., 1966. **Play in mammals**. Symp. Zool. Soc. Lond. 18: 1-9.

Shibata, T, and Irie, R., **Artificial Emotional Creature for Human-Robot Interaction – A New Direction for Intelligent System**, IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM'97), paper number 47 in Proc. 1997.

Wilson, E.O., **Sociobiology**. The Belknap press of Harvard University Press, Cambridge, Mass. 1975.

7. Appendix:

Figure 5: Images from sessions with the four children in Denmark[4].
A 1 – Smiling at a success; A2 – Hard work getting the head back up to play again; **A3** – Exhausted.
B1-3 – Floor manipulation of the overhead sensors with lights on ceiling. Large extended stretching.
C1-3 – Close observation of the light patterns. A controlled 'investigation – success – play, timeline.'
D1-2 – Full immersion as head, hands and arms are working. D3 - Open jaw indication of a success.

[4] The photographs are low quality and this is through being captured from a video camera which was set on night vision in the dark sessions. The author apologizes but believes the content of interest for the study results.

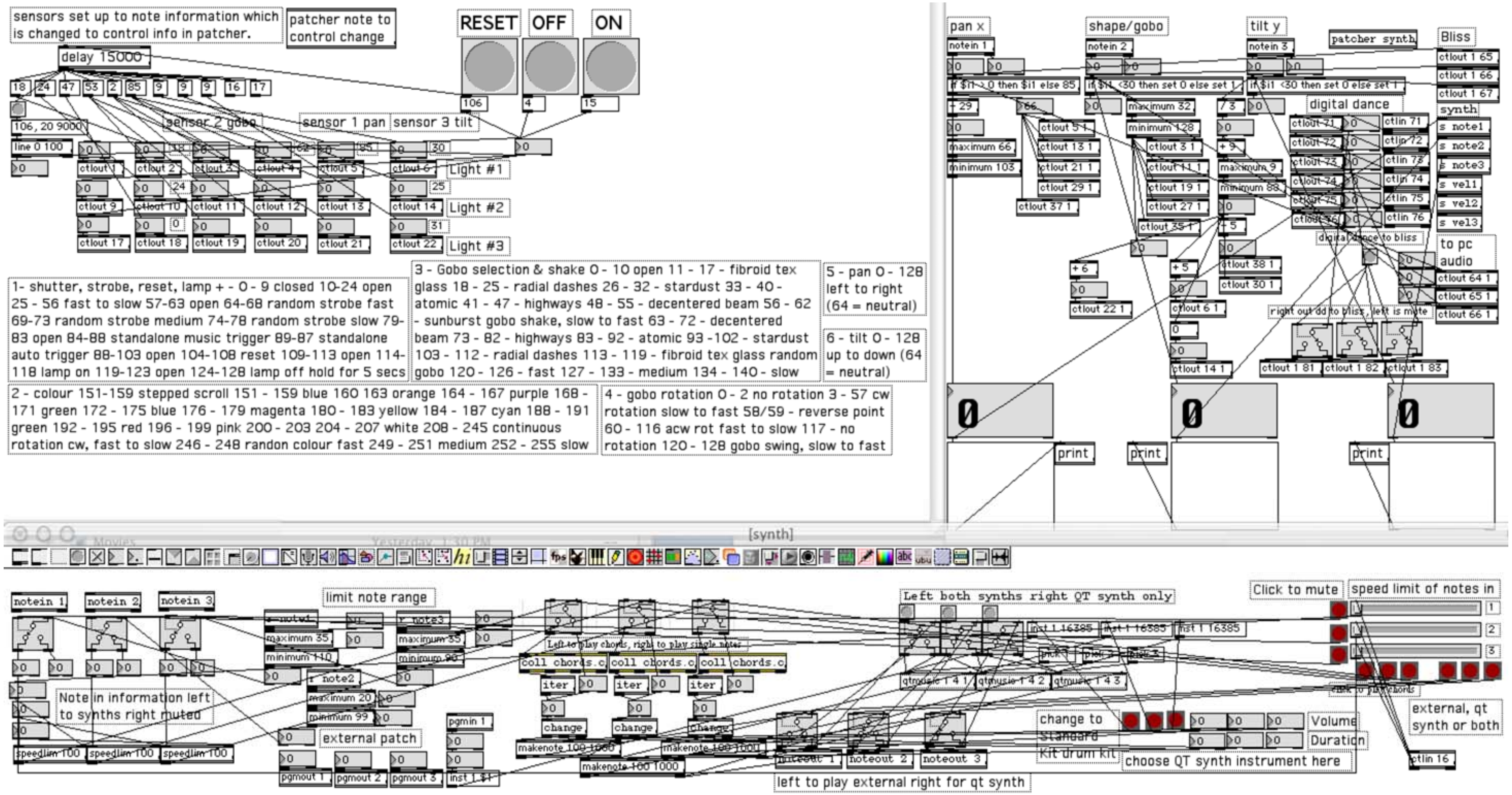


Figure 6: – The software MAX patch for real-time control of the robotic lights in edit mode showing the routing of the signal flows. The upper left window gives numerical read outs on the various parameters of control via the lighting controller of the robotic light devices. The top right window is the sensor input routing and limiting parameters. The lower window is the sound.

Figure 7: (Left) A ceiling mounted MiniMAC Profile intelligent light unit. **Figure 8** (Right) A scanner lighting device. The difference between these units of concern to the research was that a physical interaction between gesture and robotic movement was possible with the Mini Mac Profile. The early trials showed that the physical moving head of the Mini Mac Profile units gave a higher awareness for the user than the scanner devices which only enabled a moving of a mirror assembly to direct the light patterns relative to the gestures. The range of movement was also a determining factor between 540 x 270 and 180 x 90 degree of Pan (x)/Tilt (y) in micro-step increments so that a higher ration of mapping could be programmed.

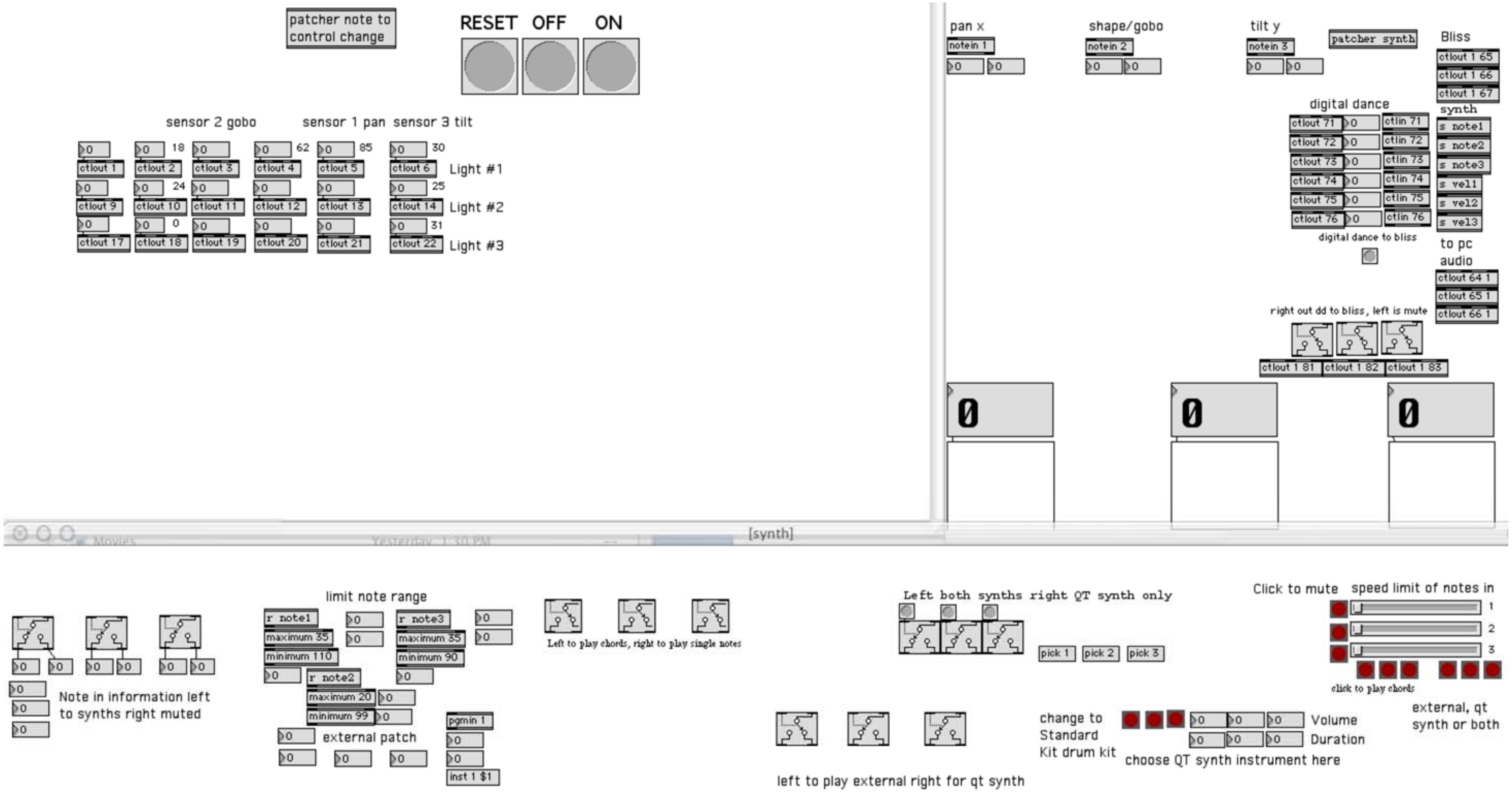


Figure 9: The graphical user interface (GUI) for the control of the virtual coach system with robotic light devices controlled from the upper two windows and the sound synthesizer controlled from the lower window. Upper left window shows six vertical columns of number boxes giving indication of the output to controller and correspond to the sliders on the MIDI to DMX 512 controller CP10. Three rows of number boxes correspond to the number of lights – in this case 3 x Martin Mini MAC profile at 8-bit resolution (Mode 1). Three sensors 1, 2, and 3 were used in this case to correspond to columns 5, 3, and 6 respectively so that column 5 was pan control, column 6 was tilt control and column 3 was gobo (pattern) control. Upper right window shows the real-time data stream from the sensor Virtual Interactive Space (VIS) in the three large graphic and number boxes. The right hand side shows mapping options to paint, music and external PC. Digital Dance is an alternative sensor input device used for calibration and performance trials. The text boxes give read out of the various parameters that can be programmed for range.

The lower window is the mapping of signals to the sound modules in the PC (such as Apple QuickTime synthesizer) or external to hardware modules or both. In this window selection of sounds that respond to the user’s gesture can be programmed with separate mapping resolution in respect of constraints that may otherwise be enforced upon the robotic light range limits. Mute buttons for all sensors are able to quiet the sound so that the user can explore the visual interactions only.