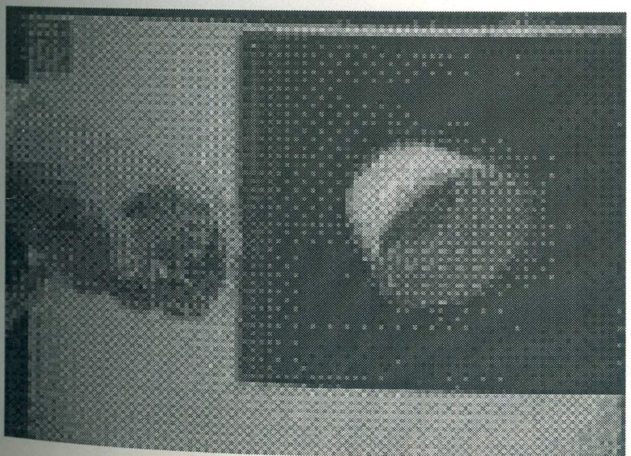
2

Report on paper for-

Asia Pacific Computer Human Interaction 1998 (APCHI 98)

Title: Castle in the air- by computer model.



Author: Shikha Varshney
student (industrial design)
Industrial Design Centre,
Indian Institute of Technology
Bombay 400076,India.

Under guidance of
Prof.U.A. Athavankar

Contents of the paper

Approval Sheet

Special project entitled 'Castle in air -by computer model' by Ms. Shikha Varshney
under the guidance of
Prof. U. A. Athavankar
is approved in partial fulfilment of the requirements for the Master's Degree in Design.

Guide ____________________

Internal Examiner-1 ____________________

Internal Examiner-2 ____________________

Acknowledgement:

This report is a part of the paper submitted for APCHI'98. I could never have been able to write such a paper if Prof. Athavankar was not with me in this project. I feel this paper is a good example of a collaborative work between a student and a professor. I'll take this oportunity to thank Prof. Athavankar for guiding me and devoting his valuable time for this paper. In fact it was he who suggested to write the paper...Thanks for pursuing me to complete the paper on time.

I'll also like to thanks all my friends who readily helped me during the pilot experiments, especially Sonu, Suparna and Devesh. Above all I'll like to thank Mr. Pankaj Sapkal and Mr. RajGopal who helped me during the final experiments.

Shikha Varshney.

Author's Note:

This special project is a paper sent for Asia Pacific Computer Human Interaction 1998 (APCHI '98) .Therefore this report is in the form of a final paper sent for the conference. All the experiments and the studies done are briefly described in the report as they are written in the paper. The paper might be forwarded for publication in IEEE journal.

Shikha Varshney

Title : Castle in the air- by computer model.

Abstract

During the early design explorations, designers and architects tend to treat freehand sketching as a mirror of their mind. The spontaneity of conventional sketching is undoubtedly important for the designer's thinking process. CAD systems and the contrived interfaces do not offer the advantage that sketching does. But, can the modern technology offer a tool that the designers would readily accept for their creative explorations?

The study is based on a visualization of a hypothetical system, where the designer develops a representation of 3D shapes of reasonable complexity, sitting in front of a terminal. Using his gestures and limited speech, he inputs into the computer an online representation of shape. This feedback is available to the designer live on his monitor. The experiments reported here test this hypothetical system using simulation.

Gestures involve spatial movements and they could be more convenient tools for the designers to express spatial ideas. Keeping modeling of shapes in the computer during the act of design as a focus of this study, this paper aims to identify the potentials and limitations of the natural gestures as a source of shape information.

1.0 Aim: A Designer Friendly Interface

This research study is a simulated exploration of use of hand gestures as a source of shape information. It aims to influence development of more designer friendly and cognitively compatible interface for entering shape data in the computer. The focus is mainly on the work of industrial designers during the early creative phase, who often sculpt a 3D shape while evolving a product form. Typical shapes that they conceive often use surfaces in 3D space that flow into each other. We would begin by briefly reviewing design behavior during the early creative phase, so that the current study can be viewed in that context.

Keywords- *Gestures as interface, Design, Creative process, Sketching, Shape description.*

1.1 Some Notes on Design Behavior

Industrial Designers, like architects, must typically evolve a new product shape as a response to the problem. The process does not start with a preconceived shape, but like sculpting, it emerges in the process of iterative interactions with some form of representation, typically a sketch or a 3D mockup. Traditionally, Designers have been using sketching as a tool to develop early shape ideas and to support their thinking process. Sketching is often viewed as a part of an emotional outburst during the creative phase. With the result, such sketches tend to be somewhat personal and have a notation like qualities. It is no wonder that designers consider sketching as a fundamental aspect of design and inseparable from their thinking process[1], [2]. As a result, much of the design research is focused on the role of sketching in the creative phase.

There is also another side to the story. Sketching may not be an ideal tool for representation of ideas. It requires extensive formal learning and involves conscious sensory motor operations to convert 3D ideas into 2D representations. This is a major difficulty that designers and particularly beginners face. The conscious efforts required to get a good sketch tends to divert the designer's attention from the developing thought process in his mind[3].

Computer graphics offers an alternative form of representation of 3D shape that is easy to manipulate. However, designers don't nor use computers during the idea generation phase, but after the idea is concretized. The literature on this topic shows that the advantages of sketching can not be replaced by computers during the early creative phase. Computer interfaces and commands do not compliment the cognitive processes naturally, the way sketching does. For example, primitives like cube, cylinder or operations like boolean are rarely used by the designers in their thinking process. This paper addresses the question 'Can we conceive of a new interface on a computer, that is easier to learn and is cognitively closer to the thinking process, thus replacing sketching as a form of representation?'

2.0 Gestures and Communication

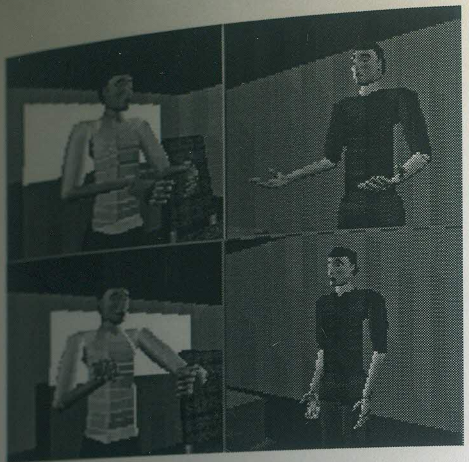
Observation that during the creative phase designers tend to argue with themselves by using sub-vocal speech and/or use hand gestures to get the feel of the evolving product shape suggests a possibility of using hand gestures and limited speech as a source of shape information for computers. Gestures continue to naturally accompany general conversations and manifest more vigorously when dealing with spatial problems such as giving directions. Unlike sketching, gestures operate directly in 3D space and consequently, they are more convenient when communicating spatial information. Studies on design behavior during the early creative phase indicate that gestures like sketching can potentially help the thought process [4].

The idea of using gestures as a mode of input is not new. The vision based gesture understanding is accepted as a step towards making computers more human like. By providing these machines with the ability to understand gestures, speech and facial expressions, we can move computer-human interaction closer to human-human interaction. Several studies have focused on use of gestural inputs for multi-modal applications such as interacting with virtual environment. Gestural analyzer enables natural gesture input to any appropriate application [5]. Recent development in this area are haptic interface input devices. Haptic interactions may be accompanied by man or machine hands and the environment can be real or virtual [6].

2.1 Gestures as Shape Information

Gestures have opened up exciting interface possibilities in various application areas, such as entertainment, education, industry, graphic arts and industrial design. With recent advancements in interfaces and data entry devices, gestures could be thought of as a tool for entry of shape data for computers. Keeping modeling of shapes in the computer as a focus of this study, the paper aims to identify the potentials and limitations of the natural gestures as a source of shape information during the act of design.

3.0 Visualization of a New Interface



The two animated agents interact with one another, producing speech with appropriate intonational contours, hand gestures, and facial movements such as head turns and nods

The study is based on a visualization of a hypothetical system, where the designer is expected to sit in front of a terminal and using his gestures and limited speech, he develops a representation of 3D shapes of reasonable complexity. It is assumed that the computer is able to convert his gestures and speech inputs into an online representation of shape and this feedback is available to the designer live on his monitor. But before major expenses are committed on implementing such an idea, can the system be simulated to see if such an interaction would be useful to the designers, and if yes, explore what difficulties may be encountered in the interaction This study is part of the ongoing research efforts to simulate this idea.

3.1 Decoding Gestures by Human Machine

The series of experiments on use of gestures are in fact a spin-off of explorations of use of 'mental imagery as a design tool', where the designer was blindfolded and was asked to speak aloud while designing series of products [4]. Because the designer couldn't sketch, he used hand gestures extensively to evolve his product shapes. All throughout the session, the designer almost 'modeled' the entire shape with hand gestures, altered and even manipulated it, as if it existed in front of him. These sessions were videotaped, which recorded gestural and verbal protocols. Videotaped protocols appeared so rich with information that it prompted new studies on possibilities on decoding the gestures.

cyber grasp-

In the next series of studies, the question that we addressed was 'Could another designer viewing the video protocols understand these gestures and reconstruct the product shape?' So, these video protocols were shown to another set of designers and they were asked to guess the final shape of the product that the original designer had conceived and reproduce it as a sketch [7]. The experiment treated the designers decoding the protocols as a simulated 'expert system'. Surprisingly, the results showed a fairly close correspondence between the ideas as conceived by the original designer and the sketches at the end of the decoding session. These results showed a distinct possibility of a future 'expert system' that could model a 3D shape in a computer as a response to the hand gestures and accompanying speech strings. But before we could hazard such a guess, it is important to reflect on the possible reasons for the success of these experiments.

The correspondence between the designer's ideas and the decoder's sketches could have been due to any one or more of the following factors ...

1) Availability of decision logic in the verbal protocols, which could have helped the other designers to reconstruct the shape.

2) Visible relationship between the speech strings and the hand gestures because of the simultaneity.

3) Ability of the experts to develop, their own hypothesis on the conventions used by the original designer in generating hand gestures, which of course implies use of higher level intelligence.

4) Ability of human intelligence to eliminate the irrelevant gestures and speech strings

3.2 New Hypothesis

The current experiment tries to control some of the above factors to ascertain whether the designers acting as 'experts' can still decode the shapes. This paper addresses following specific questions 'Is the logic used by the designers in taking decisions (as recorded in verbal protocols) a complimentary input for the decoding process? What if this logic is eliminated or restricted? To what extent can we minimize the complementary verbal strings ?

The present study is a pilot project based on a simulation of computer-human interaction and continues to use humans as an 'expert system' decoding the shape. It explores if such an expert system is worth building and what precautions will have to be taken while developing it. This limited data may not permit generalization. Yet, the exciting possibilities of using hand gestures as an interface in design is worth reporting.

4.0 Design of Experiment

The experimental setup simulates a Computer-human interaction based on natural gestures and limited speech. The experiment is almost like a dumb-charade without a direct contact between the two people involved. It requires the designer, referred as a coder, to describe the shape using hand gestures and limited speech. A video camera is focused on coder but the frame avoids the coder's face from the camera view. The second designer, referred as a decode, watches this video data on his monitor and simultaneously develops a sketch of the shape. A ceiling mounted camera records his sketching process and relays it live to the coder on his monitor, which acts like a simulated computer screen.

Preliminary experiments suggested that viewing of facial expressions for reassurance or to communicate confusion etc. which had favorable effect on the decoding process. So, only restricted live views of each others actions are relayed on the respective monitors.

4.1 Sequence of Stages

The sequence of experimental stages evolved was as follows...

1. The coder who was required to descrbe the shape using hand gestures, was presented with orthogonal/perspective views of the shapes. He was given time to study the views till he felt confident of recalling it with fair amount of accuracy. He was also allowed time to evolve the strategy of describing the shape and if necessary rehearse it in privacy. He was then asked to sit in front of the video camera and explain the shape mainly through hand gestures and with the help of limited speech. For instance he could indicate part's name/ colour/ texture/ dimensions or give commands like scale/ erase/ repeat etc.

2. The decoder sat away from the coder, and could hear the coder and see the hand gestures live on monitor in front of him. He was asked to interpret this data and simultaneously sketch the shape ideas as the session progressed. Thus the decoder developed several on-line sketches of parts and consolidated the idea incrementally.

   In case the decoder had difficulty understanding particular set of gestures, he could signal by pressing a bell, so that the coder can go back a few steps and repeat the explanations and the gestures.

3. Coder could see the progress of the decoder's sketch online on his monitor and if it showed a wrong direction, he could always ask the decoder to erase and repeat his actions to remove ambiguity. Thus the coder remained in command of the process completely and decoder was like a machine responding to his actions.

4. At the end of the session, decoder was asked to consolidate all the previous sketches and produces his final sketch. Decoder's final sketch was then compared with the original sketch shown to the coder.

4.2 Choice of Visual Material

Preliminary studies showed that, the decoders tended to guess the product category first and use their knowledge to reconstruct the shape. To eliminate familiarity with the product category, nonsense shapes were developed for the final experiment. This strategy also ensured that the coders had no access to the logic that influenced the development of the shapes.

4.3 Choice of Subjects and Pairing

The idea is to simulate designer interacting with a computer using hand gestures and limited speech. So, the subjects who coded the shape had Industrial design background. Use of design student as well as a design professional in the coding process in some way simulated the difficulties that a learner would face. The other side is represented by an intelligent computer (an expert system with professional knowledge and expertise). So, it made sense to again use professional designers to act as 'experts'.

4.5 Analysis Planned

The original sketch given to the coder and the sketch produced by the decoder was then compared for similarity using like:

1. Similarity in overall shape.
2. Existence of specific shape features, which included parts like main body and add-ons, attributes like colours, texture etc.
3. Spatial relationship of parts in space
4. Correctness of overall proportion and not of actual size.

5.0 Level of Correspondence

Experiment 01: In this experiment the coder was a design student and the decoder was an experienced designer. Coder was given sufficient time after his study of the views of the shape was complete. The total time taken for the session was 31 minutes. The figure 01 shows the sketch presented to the coder and figure 02 shows the view produced by the decoder at the end of the session. Comparison of these figures reveals the level of correspondence.

Overall shape was depicted correctly by the decoder. All the shape features and the attributes in the original sketch are also present in the final sketch and their relative proportions show a fair degree of correspondence.However, the wrap was mistakenly assumed by the decoder to be a thin thread and its loose end was slightly displaced in the sketch. The colours and the texture assigned to various parts were same as the original sketch.

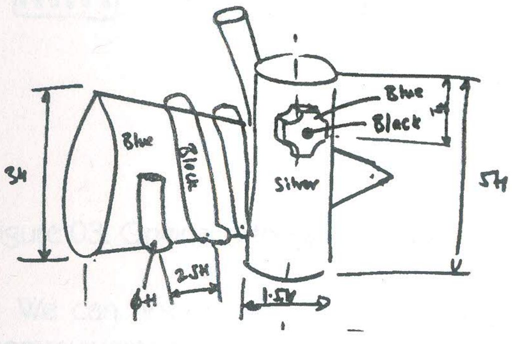


fig. 01 Coder's sketch

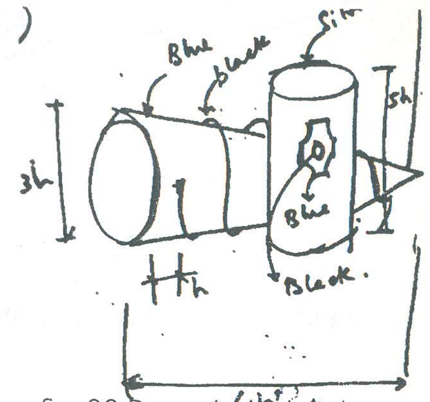


fig. 02 Decoder's sketch

Although the viewing angle was not an important criterion, in this particular experiment, considering that the gestures were shot from the front and that the decoder had an opposite view of the virtual shape, the angle of view he drew was surprisingly close to the original sketch.

Experiment 02: In this experiment, both the subjects were experienced designers and the shape used was also more complex. Coder was given sufficient time to study the orthographic views and plan his communication strategy. For the actual session, coder took little over 43 minutes. Figure 04 shows the orthographic views presented to the coder and figure 05 shows the view produced by the decoder at the end of the session.

Though the original sketch is in orthographic views, a closer look would reveal that, except for the intersection angle, the overall configuration shown by the decoder in the final sketch is similar to the original shape. The decoder may have assumed a default value of 90 degrees because the intersection angle of the extruded square section was not specified and was difficult to guess from the frontal camera angle. While the overall feel of the shape is correct, there is a slight shortening of the length in the decoder's final sketch. Besides, the number of vertical grooves around the cylindrical face were fewer and also slightly disproportionate.

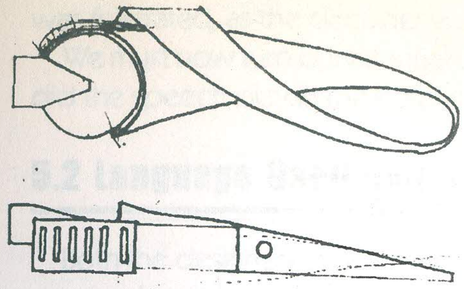


Figure 03: Original orthogonal view

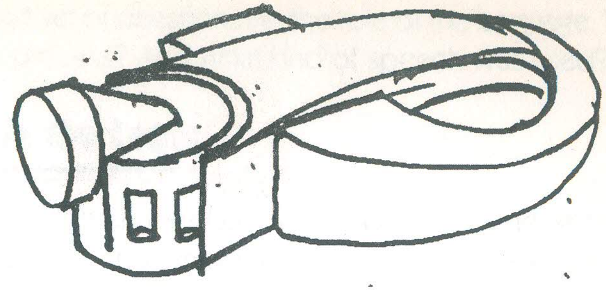


Figure 04. Decoder's sketch

We can now return to the first question that we started with i.e. 'Can gestures successful communicate shape information?'

5.1 Effectiveness of Gestures and Words?

A quick glance at the sketches shown above suggests that the answer is affirmative. There is a perceptible resemblance between the shape depicted in the original views and the sketches that the decoders produced based on the gestures and limited speech. One can perhaps attribute the correspondence in the two sketches to the online feedback on the monitor. It offered opportunities to the coder to correct the shape till it more all less reached what he had in mind. The success can be also attributed to the fact that the coders were able to quickly come up with a strategy of breaking the complex shape, then enacting it part by part and later attaching them to each other to form the final shape. The fact that a design student was also able to come up with such a strategy quickly, suggests that it would not need much learning. There are also other things worth noting.

The increase in the complexity of the shape between experiment 01 and 02 was deliberately. It allows us to speculate on the level of complexity that a decoder can handle. Higher level of complexity of the shape in experiment 02 has not deterred the decoder and he was reasonably successful in reproducing the shape. Not that there were no communication problems at all. During the experiment, there were phases when the coder just could not get his ideas across, tried alternative strategies and was frustrated, as the decoder was not able to understand the actions.

We must now turn our attention to the next set of questions i.e. the role of the language. What role did the speech play in the communication process? And what kind of speech was used?

5.2 Language Used and Access to the Logic

Both the designers did not evolve their product shapes but used specially developed nonsense shapes shown to them just before the experiment. Thus, they had no access to the logic (?) used for developing the shape and the possibility of the coder's speech strings reflecting the logic was totally eliminated.

In both the experiments, following the instructions, the coders used limited words to described the shapes. Even when they encountered difficulties in explaining, they used speech only to the extent it was permissible. Their words could be roughly classified in four categories i.e. Operators, Locators, Indicator labels and Numerals. It is difficult not to notice the similarity of some of these words with CAD commands. Both the coders were familiar with the CAD systems, which perhaps prompted them to use typical CAD commands like save, mirror, intersect, extend, etc. to explain their ideas. A quick look at what these categories include will throw some light on the nature of communication.

Operators:

Most often used words were 'operators' or 'commands' that expected the decoder to execute the operation. These included command like words such as merge, fuse, rotate, mirror, move, turn, bend, intersect, extend etc. Also included were editing commands like select, save, save as, cut, erase, repeat, duplicate, start again, next, remove, place, stop, etc.

Locators:

The second group involved 'locators' which suggested placement and alignments of parts. Examples include words like centre, base, edge, end, level or vertical, horizontal, right/left, along the axis etc.

Indicator Labels:

The third type could be called as indicators labels which included nouns, often those referring to either standard shapes or colours. If the part did not suggest a name quickly, coders turned to using alphabets or decimals as well as Roman numbers, as substitutes. Sometimes even views were named or numbered so that they can be referred to later.

Numerals:

The last group include numbers which were used to dimension the parts.

It is encouraging to observe that the descriptions were not too dependent on the language nor on processing of grammer. It suggests that such a system may work with a limited vocabulary of words and may not have to encounter the difficulties involved in language processing.

6.0 Gesture Based Modeling: A Reality?

Though the study is only a simulation of a new form of Computer-human interaction, results from the experiments clearly point to the potentials of using gestures, supported by limited descriptive utterances, to enter shape information. Indeed, gestures appear to be a viable tool for shape data entry. The fact that the technologies like gesture analysis in real-time exist, make the idea of developing such an expert system not too remote.

The description process was dependent on the language in a limited way, but restricting the words and eliminating sentences from the descriptions has not proved to be an impediment in decoding. This is a good news as it suggests that such an expert system may not need to process sentences and grammar and a modest dictionary with a limited vocabulary would be sufficient.

6.1 Some Useful Observations

The simulation process threw-up some interesting ideas that would be useful in developing an expert system. For instance, the decoders tended to assume missing data, for example the intersection angle was assumed to be 90 degrees. This suggests that the system could have 'default settings' for all geometric aspects, which the system would use when the information is not complete. The screen will then give a feedback to the designer that he has missed out some dimensions/values and prompt him to correct the assumption immediately.

Although direct communication was not allowed, subjects developed their own strategies for communication exploiting the contact through the cameras. For instance, decoders sketched marks like ?, !, or signaled by keeping their pen on the sketch at the point of confusion. Coders tended to use hand gestures to make the decoder stop at certain point. These strategies were used for reassurance, confirmation or clearing the ambiguities. It suggests that when the gestural communication is ambiguous to the 'expert system', it could also offer feedback through similar icons.

Lastly, human decoders sometimes missed part of the gestural information while they were busy sketching. This may not be a problem when machines decode the gestures. But there are other problems that have remained unsolved at this point of time.

6.2 Unresolved Problems and Challenges

The system may have to overcome some of the problems for which there are no immediate solutions to offer. For instance, in the current task, the coders were able to respond to the feedback from the monitor and take corrective actions without getting too distracted. It is difficult to guess if this interaction will hinder the creative process, when designers are actually working on a live design problem. At the moment this question remains unanswered. Similarly, coders sometimes closed their eyes to concentrate on the virtual models in their mind and thus missed looking at the feedback that was available on the monitor in form of sketching. When the communication became very tedious, the coder used gestures of frustration, such as fidgeting of hands in the space, crossing fingers while thinking etc. Human decoders could easily discern and eliminate processing of such gestures. It is not clear how machine could handle such situations. Lastly, like language, gestures are also subject to considerable cultural influences and vary from place to place. Can machine account for cultural influences? These problems are left open for future explorations.

It must be conceded that the idea needs to be tried out on more subjects with varying levels of training and expertise. But even at the pilot study level, the potentials that the idea offers is proved beyond doubt. I am sure designers would eagerly look forward to such a friendly interface, which allows them to 'gesture and talk' to the computer in future.

For further interaction on this topic, you can direct your communication to Prof. Athavankar.

e.mail-uaa@idc.iitb.ernet.in

References

[1] Goldschmidt, G., Serial Sketching: visual problem solving in Designing. Cybernetics and Systems, 23: pp 191-219,1992.

[2] Goldschmidt, G., On Visual Design Thinking : the vis kids of architecture. Design Studies, 15(2): pp 158-174,1994

[3] Athavankar U., Rediscovery the Act of Sketching: Implication of its Support to the Creative Thought Process, Design Recherche, No. 2, pp 45-60, Sept. 1992.

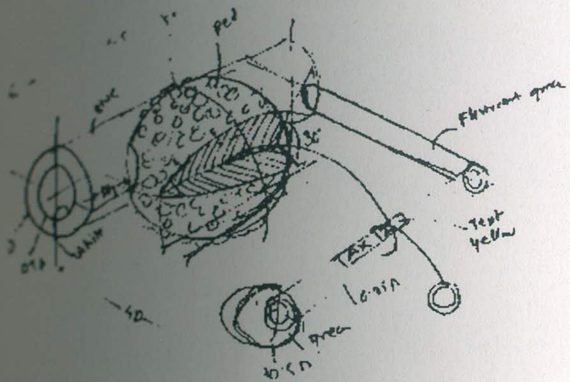
[4] Athavankar, U., Mental Imagery as a Design Tool, Cybernetics and Systems, Vol 28, No 1, Jan-Feb, pp 25-42, 1997. (Initially published in Proceedings of the Thirteenth European meeting on Cybernetics and Systems Ed. Robert Trappl. Austrian Society of Cybernetics Studies and University of Vienna, Austria, April 19-12, 1996, Vol II, pp 382-387)

[5] Wexelblat, A., ACM Transactions on Computer-Human Interaction, Vol.2, No.3, pp 179-200, Sept. 1995

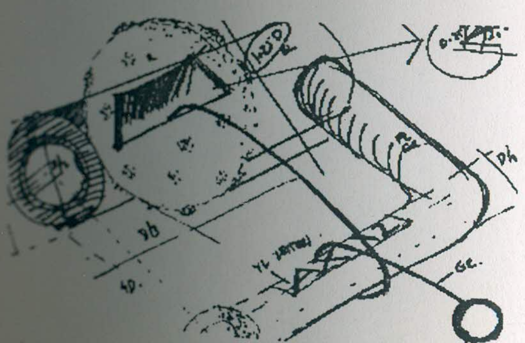
[6] Srinivasan M. & Basdogan C., Haptics in Virtual Environments: Taxonomy, Research Status and Challenges, Computer & Graphics, 21,4, pp 393-404,1997.

[7] Athavankar, U., Learning from the way Designers Model Shapes in their Mind, Cognitive Systems : from Intelligent Systems to Artificial life? ed. J.R. Issac and V. Jindal, Tata McGraw-Hill, New Delhi, pp 221-232, 1997. .

Few pilot experiments sketches



enacter's sketch



decoder's sketch

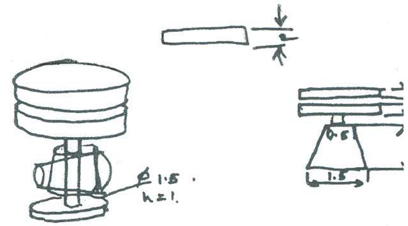
8.0 Annexures:

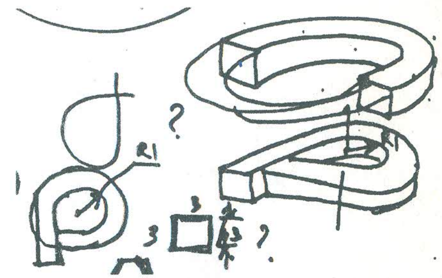
Today much of the research and study has been done on hand gesture recognition system which is able to recognise hand gestures in real time. Hand gestures is one of the most natural and expressive way of interaction- especially for hearing impaired. However, because of the complexity of dynamic gestures, most of the researches focussed either on static gesture, postures or small set of dynamic gestures. To solve this problem gesture is decomposed as a sequence of postures and postures can be quickly recognised using hidden Markov model. Similar type of researches has been made for hand recognition applicable in many areas of application.
In this paper, we tried to est the idea of using gesture recognition for developing a more expressive way of interface for the designers. This paper is not a research done in technical aspect of gesture recognition, but is a behavioural study to test whether a designer will be facilitated by such interface or not and also identify the limitations of such an interface for developing the virtual model.
Therefore to do the final experiment few pilot experiments were done.

Points considered for pilot experiment:

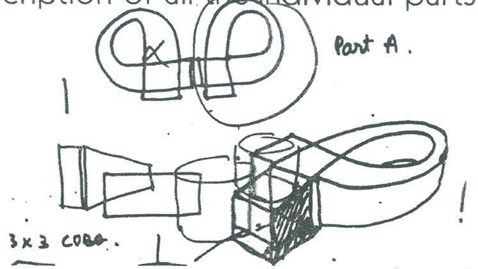
- decoder of high intellect
- product new to enacter as well as decoder, else guess or assume the missing information.
-product should be of moderate scale
- no direct eye contact between the two or else communicate through expressions
- an indicator query button to ask questions.

Sequence of development of sketches in the final experiment 02:

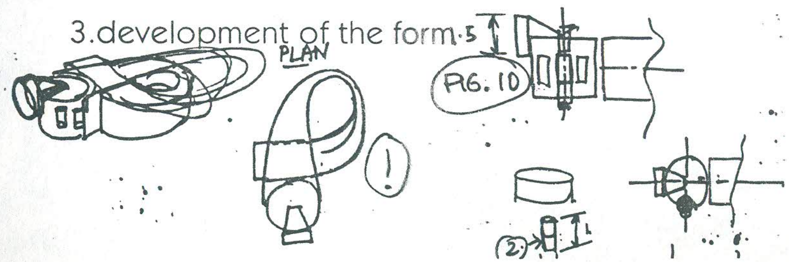


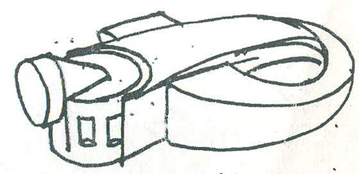


1. description of all the individual parts 2.initial sketches of curvilinear form



3.development of the form





4.final form