Ambiguity – A User Quality

Collaborative Narrative in a Multimodal User Interface

Anders-Petter Andersson and Birgitta Cappelen

School of Arts and Communication / Interactive Institute

Malmo University

S-20506 Malmo, Sweden

Anders-Petter.Andersson@kk.mah.se   Birgitta.Cappelen@kk.mah.se

Abstract

In this paper we describe an application model for a collaborative narrative in a multimodal tangible user interface. We investigate how a story can be developed in negotiation between three users. By collaborating the users create a soundtrack and negotiate the next video sequence. We developed rhetorical principles for editing moving image material for a node based interactive system. We present a new concept of quality in the user experience that we define as ambiguity, user experience caused by the tension between complexity, control and risk.

Background

Complex multimodal user interfaces is an important area for future research and development [1, 2, 3, 4]. Particularly we think that physical attributes like proportions, weight, resistance and tactility will be important qualities in the future of user interfaces.

Collaboration as a part of the user experience has a huge potential as a physical real-time experience, establishing and developing relations over time. Collaboration in this sense has both an aesthetically and motivating potential [5], that has not been fully explored in CHI and CSCW research [6, 7, 8]. Structuring and organising information of various media types is a major and complex field, where we consider narrative structuring principles to become increasingly more important [9, 10, 11]. Narrative structures make it possible to create a more meaningful relation between sequences and diverse content built on a deeper understanding of use over time, where the rhetoric inherent to the material drives the use further, and creates a flow in the use process. We think this will develop a new rhetoric language for visuals, audio and other media in the field of interactive systems.

Goals

Our goal was to explore how to develop a collaborative narrative in a multimodal tangible multiuser interface. We define a collaborative narrative as a story developed in collaboration between several users. In construction of the narrative we wanted to investigate the rhetoric of linear media, such as editing techniques in film and music, and investigate their relevance for interactive media. By this we wanted to contribute to the development of a rhetoric for multimodal interactivity.

Further we wanted to investigate what qualities made the installation aesthetically satisfying to use in the dynamic construction of a narrative. In other words, what made the users aesthetically satisfied when interacting individually and by collaborating.

The Installation

As an answer to de goals formulated we created the art installation "Moving Paths" which offered up to three users to collaborate in the development of a film narrative. The intention from the design point of view was to design an installation that offered the user a complex communication situation with several layers of meaning. Pertinent layers, which offered the users potential experiences of risk, control and collaboration while interacting. This in order to be aesthetically satisfying over time so a narrative structure could be developed. The installation was placed in a dark room. Three large blue painted interaction objects were placed in the middle of the room in front of a large video screen. The objects produced sounds when interacting with them as when playing on a musical instrument. The interaction objects were placed in such manner that the users were both able to see each other and the screen while interacting with the instruments. Moving image sequences were projected onto the screen in a sequence depending on the collaboration between the users. The audience could stand around the "stage like" centre of the room watching the screen and the interactors on the stage.

Physical structure

The three large interaction objects were a lever, a drum and a stick. The lever was built on top of a standing cylinder as illustrated in figure 1. We chose to develop three large and clearly distinguishable instruments that gave the user an experience of being a character with a clear identity and having to use their whole body in the interaction.

Figure 1 - The hardware structure
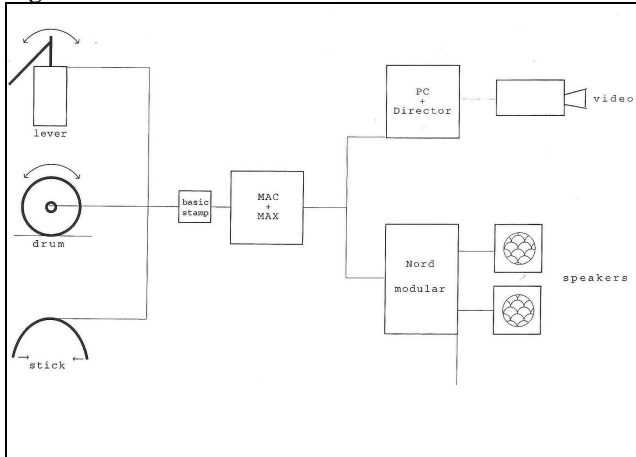
The user interacted with the instrument by turning the lever as the arrows indicate. The drum was a big cylinder about one meter in diameter, which was to be rolled back and forth on the floor. The stick was a two-meter long plastic bar that could be held and bent with ones arms. On all instruments analogue sensors were attached with wires connected to a micro controller who sent the sensor data as MIDI-signals to a Mac and a MAX program. The MIDI-signals were also sent to a PC running a Director program controlling the video sequences to be projected on a screen. The MIDI-signals were sent to a Nord Modular synthesiser controlling parameters in the sound algorithms in the synthesiser. Speakers were connected to the synthesiser.

Program structure

The program structure consisted of a "database" containing the video clips, a story engine that controlled the sequence of the video clips, a tone generator algorithm that produced sound according to the user interaction and a negotiation algorithm that interpreted the collaboration and controlled the sound and video output. See figure 2.

Figure 2 - The program structure
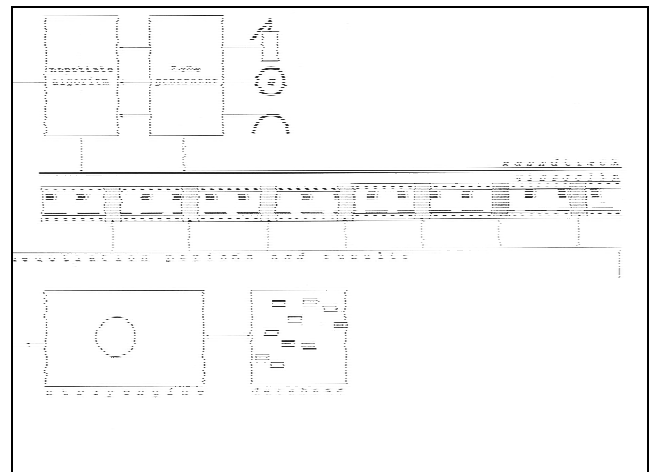
The program registered the user interaction as frequency of the activity (how quick, how many times) and the turning angle (how much), all attributes related to the identity of the instrument. When playing together each instrument controlled a parameter in the sound algorithms, and the users thereby formed the soundtrack together. One instrument gave the value to the pitch variable that controlled the sound's pitch. Another controlled the volume, and the third controlled a filter that gave the sound additional characteristics. These design choices were made to increase the complexity in the user situation by both giving the user a clear experience of being a distinguished character on the stage as well as being part of the playing team.

The negotiation, or collaboration, between the users decided in what direction the story developed, by choosing the next video sequence from the database. See figure 2.

The system was mode based which in the application meant that the negotiation periods were limited to the last 4-6 seconds of the 15 to 30 second long video sequences.

The set design

All the instruments were painted in the same dark Cobol blue colour in order to express the heaviness of the objects and create a unity between the instruments and the interactors. The installation as a whole where placed in a closed room and the set design was designed to connote a kind of stage with blue lighting directed at the stage so the interactors could see each others faces while playing. This was done to create an experience of risk for the players on the stage. The players where standing rather close to each other (approx. 2 m). Their faces were directed so that they could see each other as well as the screen while playing. This was done to make a more complex and open input structure, so the users could choose to concentrate on his own interaction and the sound he produced, the collaboration with the other users or the cinematic development of the narrative on the screen. Also audience was allowed to enter the room and they were standing behind the users so that they could watch the screen and

the players. This was chosen to make the user situation even more dramatic.

One could say that the whole installation was a hybrid between a cinema, a gaming arcade and a theatre stage.

Collaboration

We created a three-user installation in order to make a minimal dramatic structure of conflicting wills to create higher level of excitement and complexity in the collaboration.

By playing with each other and against each other the interactors could collaborate in developing the narration. The degree of harmony and conflict in the collaboration and the intensity level of each user's activity made different narratives both in content and atmosphere in the moving images. A high intensity level and a high degree of competition and conflict between the players gave a sequence of moving images with a lot of action.

A less intense collaboration gave a more harmonic development of the narration both in action and atmosphere.

User generated sound

We used sound synthesis instead of manipulation of sound samples because of the synthesis aesthetically more satisfying performance qualities in real-time processing. We made pre-composed sounds by building sound methods from LFO filters, FM synthesis and sequencer components. A MIDI controller value was assigned to each of the parameters for pitch, volume and rhythmic/chaotic ordering, which made them controllable from an external source. This gave the possibility of maximum real-time manipulation.

The role of the synthesised soundtrack was to give a direct stimuli-response feedback in order to give the user feeling of control of the situation. Whenever the user turned, bent or rolled his particular input device he would get a response that varied with frequency (how many times and how fast) and angle (how much). The sound should invite the interactor to explore movement in space with creation of sound. The interactors were in a sense creating their own soundtrack, using the interaction interface as a musical instrument. On the other hand the composer/designer had limited the possibilities by partly choosing the form (synthesised sound) and the patterns (collaborative, competitive, dialogue) programmed in the algorithms.

Moving image

We selected Jacques Tati's movie "Les Vacances des monsieur Hulot" from 1953 as the raw material for the video sequences. This black and white movie belongs to a category, often referred to as absurd movies. The story is about monsieur Hulot's strange and funny summer holiday by the sea on French Atlantic coast.

Our reason for choosing a movie from this category in general and this movie in particular is it's loosely structured narrative, consisting of short, closed, funny scenes. We thought the material was well suited for the editing of node based video sequences that in principle needed to be combinable and re-combinable in any order.

The most important research outcome of this project was what we found during the video editing experiments. Our goal was to investigate in what way traditional, linear continuity editing principles could be used when editing video sequences for a node based database. We discovered that most of the traditional editing principles could be adopted, even if they came in conflict with principles in user interface design.

In cinematic editing it is normal to cut from one visual point in one scene to the same visual point in the next scene, while the visual point in the user interface follows the cursor movement on the screen. The first was a cinematic editing principle that we found could be used in node based editing, along with other principles like "shot-reversed shot", cutting from "close up" to "totals" and back, looking at something outside of the frame and cutting to the next scene [12, 13]. We found that the optimal length of a sequence was from 15 seconds to 30 depending on the level of action in the scene.

A genuine node based editing principle we developed during our experiments was that it seemed to be necessary for each clip to start and end in almost the same way. This principle contradicted traditional continuity editing, but was instead consistent with the principle of closureness in interface design. Further, it seemed to be a good solution if the action in one sequence took place in the same room or place. We could therefore also talk about a spatial closureness.

Test and Results

In the user tests conducted in the project we tried to investigate what qualities made the installation aesthetically satisfying to use in the dynamic construction of a narrative. What could make the interactors aesthetically satisfied when interacting individually and/or by collaborating in a multimodal interface? The method used was inductive in the sense that we generalised from observing users and their activities over time. The parameters investigated during the observation were: One user situation, Two user situation, Three user situation, Stage in broad daylight, Stage in dim performance lighting, Extensive active participation, Brief active participation, Occasional participation, Repeated participation.

One user situation

The active user that entered the stage was cautious, concentrating on the relation between movement in the interaction and output in audio. The video projection with the looped sequence functioned more as a "living" background to the physical interaction. In the dim performance lighting the user became aware of his

vulnerability caused by the risk of responding to a new and unfamiliar situation. When entering the stage in broad daylight the user was even less focused on his own activity and more aware of the audience and environment. The audio feedback was crucial for getting the user engaged in the installation for a shorter period of time. Resemblance in form (all objects were painted in a Cobol blue colour) and the objects exaggerated size made it easier for the user to be able to separate the interface from the environment and made it attractive to engage physically. This also made the situation more risky. To roll the big drum on the floor, to pull the lever, or bend the two-meter long stick gave clearly recognisable and caricatured movements.

Two user situation

In a brief active participation the user still was concentrating on functionality, exploring the system. At a longer active time the users acted against each other. The experience of risk and of being exposed to an audience decreased in the development of a common language and in sharing the same experience. The co-operation could start like a call-and-respond game between two voices where the user reciprocally responded to each other in a collective improvisation by imitating and responding to each other's movements and sound. The game often went through a phase of competition where the users first tried to "shout" louder than the other persons but soon led the game back to a more harmonic form of dialogue and/or game.

Figure 4: User testing with two users

Three user situation

The three user situation developed into a battle between contending wills, but soon the movements and gestures of the interaction was subordinate to the group's collaborative activity.

The focus moved from the risk of being observed on a stage to a human-to-human communication consisting of co-ordination of movements, glances and exchange of meaning looks between the users. Unlike the situation with

one or two users the three-user situation was more dynamic in that it always was a third part that the other two could play off against.

Exploration, Conversation, Collaboration

In the tests we found that the user activity developed over time and could be described in three main phases – exploration, conversation and collaboration:

Exploration was characterised by the user's experience of self-consciousness being attentive on the sounds he created in a direct stimuli-response situation, i.e. an experience of being in control. The experience of risk and vulnerability came from the situation of being on a stage.

Conversation was characterised by awareness of the other users built on human-to-human communication consisting of co-ordination of movements and rapid glances between the users. The communication continuously changed from being an activity where the users had the same goal in mind such as call-and-respond and conversation, as well as being a competition where the users had different goals. The activities could develop into sequences of improvisation where the physical objects worked as musical instruments.

Collaboration was characterised by the users' negotiation of the film narrative, i.e. an external entity on the video screen. By working as a team towards a common goal they continuously could control the development of the story by cutting from one video sequence to another. In negotiating the user found that he could chose to collaborate by increasing or decreasing his activity in the interaction, or simply by making the negotiation impossible by putting himself outside the activity or pretending ignorant. Further, while engaging in the collaborative negotiation of the next video sequence the participants forgot about the technical-conceptual aspects of the system. The strength in the collaboration in the third phase was to be found in the development of what we would like to define as a *second narrative*.

Second Narrative

The 15-30 seconds long video sequences had a narrative curve, creating a peripeteia (dramatic peak) towards the end of the sequence. It was built from the rhetoric principles in continuity editing. The second narrative could be characterised as a pattern developed in collaboration describing it's own dramatic curve and peripeteia separated from the narrative inherent to the film.

After about 2-3 minutes the users learned to estimate the length of the sequences. The narrative curve in the film helped the users to understand when to be active or passive. The interaction thereby created a new second narrative level. Yet, this level was not only perceived as a new level, but also as stronger and more pertinent than the narrative inherent to the moving image. For example to be thrown back to the beginning of the same film-loop was experienced as if the group had failed and therefore less

satisfying than to collaboratively reach a new video sequence.

When the users understood that they could control the ordering of sequences in the moving image the dynamics of their collaboration changed. The interaction, sound, physical movement and the moving image was all at once the result of the collaborative activity as well as a catalyst for new potentially coming video sequences, movements, sounds and interactions. From the audience point of view in extensive active participation the sound created an audio-visual relation in the narrative curve that resulted in a continuous soundtrack [14]. The audio-visual relation functioned either as "on-location sound" (sound source visible to the audience) or "pit music" (sound source hidden in the orchestra pit) depending on if the focus was on the three users or on the moving image projection on the wall. Both ways it created a cause-and-effect experience. Yet, focusing more on the users and less on the screen made a stronger impact since the audience immediately could follow the collaborative narrative and get a direct correlation between efforts made in the interaction (movement/on-location sound) and the result in a new video sequence. In repeated participation when the narrative structure inherent to the moving image was made clear to the individual user, it was to the collaborative activity of negotiating the video sequences that the group returned.

Conclusions

Complexity in the communication, control in the interaction and actual risk in the collaboration altogether gave a quality to the user situation that we have chosen to call ambiguity. We define ambiguity as a user experience that changes over time, where the user switches between feeling satisfaction and control in the response, risk and tension, desire and expectation.

We think it was the experience of ambiguity as a user experience that made the "Moving Paths" installation satisfactory for the participants for shorter as well as longer periods of active interaction. It made the users motivated to collaborate and build narrative sequences together. We believe that ambiguity will become an important design principle in future user applications and user situations, both in work oriented areas and in entertainment.

The desire and expectations can in large be said to operate on two rhetoric levels: (i) on the forms of each isolated content material, e.g. the editing of the video sequences, or the narrative inherent to the moving image, or the design of the sound synthesis algorithms; (ii) on the whole polysemic multimodal output created in the collaboration. Yet, it was the complexity in the overall communication that can be said to represent the rhetoric potential of the installation.

Further research

We believe there is a big potential yet to be explored in the field of multimodal user applications, both on the level of content material, i.e. what can be coded in a database, and on the level of individual expression. On the content material level it is possible to investigate the construction of meaningful sequences, i.e. rhetoric principles that the user experience as continuous, creating desires and expectations in a node based material. We intend to continue investigating moving image material, graphics and sound in isolation, e.g. how to build meaningful sequences in node based material. We also intend to study polysemic "texts" and multimodal interfaces in order to reveal how different media work in combination, e.g. layers of sound, sound effects and different physical forms of interaction.

We believe that further research on what we have defined as ambiguity, built on tension between control, risk and desire, has great potential as new dimensions in future user interface designs. This is both related to use in general, but also to a deeper understanding of individual user experience. To develop this potential in user interface design, AI research areas like adaptivity, user modelling and learning ought to combine research from aesthetically disciplines like rhetoric, art and design.

Acknowledgements

References

1. Ishii, H., and Ullmer, B. 1997. Tangible Bits: Towards Seamless Interfaces between People Bits and Atoms. In Proc. ACM CHI '97, Atlanta GA, USA

2. Rekimoto, J. 1997. Pick-and-Drop: A Direct Manipulation Technique for Multiple Computer Environments. In Proc. UIST 97, Banff, Alberta, Canada

3. Schneiderman, B. 1997. Designing the User Interface: Strategies for Effective Human-Computer Interaction. Addison Wesley Publishing Company, US

4. Simon, P. 1999.Texts and works on robotics and interactive art installations. Web link.

http://www-art.cfa.cmu.edu/Penny/index.html

5. Susani, M. 1999. Mobile, Ubiquitous and the Sense of Space. Invited Keynotes, First International Symposium on Handheld and Ubiquitous Computing, HUC '99, Karlsruhe, Germany

6. Butz, A., Höllerer, T., Feiner, S., MacIntyre B., and Beshers, C. 1999. Enveloping Users and Computers in a Collaborative 3D Augmented Reality. In Proc. International Workshop on Augmented Reality, IWAR '99, IEEE Computer Society Press

7. Holmer, T., Lacour, L., and Streitz, N. 1998. i-LAND: An interactive landscape for creativity and innovation. In Proc. Conf. on Computer-Supported Cooperative Work, ACM CSCW '98

8. Aiken, R., Abramski, A., Bates, J., and Blackadar, T. 1999. Middleware for Ubiquitous Computing. In Panel Session, First International Symposium on Handheld and Ubiquitous Computing, HUC '99, Karlsruhe, Germany

9. Davenport, G., Evans, R., and Halliday, M. 1993. Orchestrating Digital Micromovies. Leonardo, Vol 26, No. 4, pp. 188-283, the MIT Press

10. Aarseth, E. 1997. Cybertext: Perspectives on Ergodic Literature. Johns Hopkins University Press

11. Jennings, P. 1996. Narrative Structures for New Media. Leonardo, Vol. 29, No. 5, pp. 345-350, the MIT Press

12. Bordwell, D., Thompson, K. 1996. Film Art: an Introduction. McGraw-Hill Publishing Company.

13. Bordwell, D. 1990. Narration in the Fiction Film. Routledge

14. Chion, M. 1994. Audio-Vision: Sound on Screen. Colombia University Press