New Interfaces for Video Browsing and Editing

Heimo Müller, Maja Pivec, Konrad Baumann

Abstract

In the field of video browsing, editing and processing a lot of software tools are on the market. Their user interface archetypes are predominantly based on early technological developments in non linear video editing (NLE) and follow the timeline concept. Typical users of non linear video editing software are either well trained professionals or experienced home users who have a high involvement with these software products for years.

Several end user studies showed a great lack of understanding and disorientation of NLE solution in the (mass) home market. We realized a strong users need for self-explanatory and easy to use products. This was the starting point for the design of a radically new interface for digital video editing.

We developed a new way of video editing based on content visualization (OM images) and the introduction of layout information to the video editing process. We propose to redefine the mental model of video pre-production and video editing by applying the typographic elements to the description and manipulation of video objects. The so-called Video Book concept affords editing like text and therefore can be used intuitively by users familiar with today's word processing software.

Video Book

Most of us know how to create a report, write a letter or edit a book. But what would happen, if we edit textual content according to the same principle as used in video editing? The (not so serious) answer to this question is the "Timeline based book" shown in Figure 1.

Looking at a page in a classical book, we have a lot of layout information in addition to the plain text, e.g. font type, character size, paragraph shape, white space, page border, title elements, and graphical objects, as illustrated in Figure 2. The employment of layout elements, e.g. "use a heading for the beginning of a new chapter" or "use italic font style for citations" is well known to all literate people.

Figure 1 – "Timeline based" Book

As well, people who use word processors know how to use these elements, a knowledge which has spread across the western world during the last decades.

paragraph

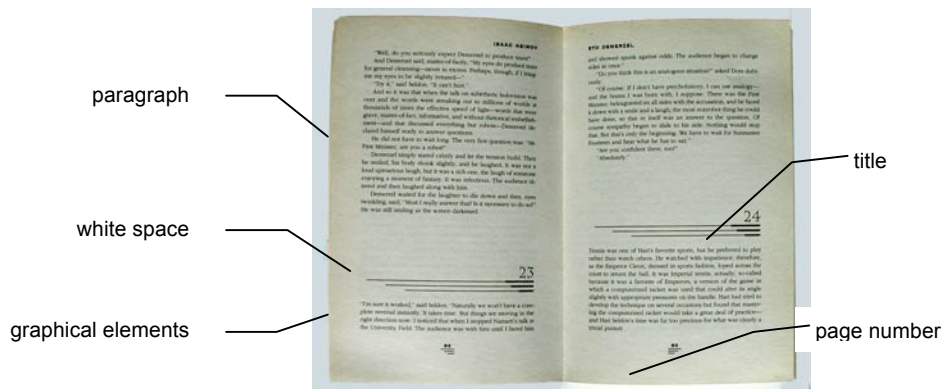white space

graphical elements

title

page number

Figure 2: A sample book layout using paragraphs, white space, titles, page headings, page numbers and white space

This example illustrates, that we can learn from the long history of typography and book design when developing a new video publishing tool.

If we apply the principles of typography to the description and manipulation of video objects, we have to redefine the user's (mental) model of video pre-production and video editing. We follow the idea of applying the principles of typography to the description and manipulation of video objects. In a first step we break the well known (linear) time-line into a number of lines and add layout information, see Figure 3.
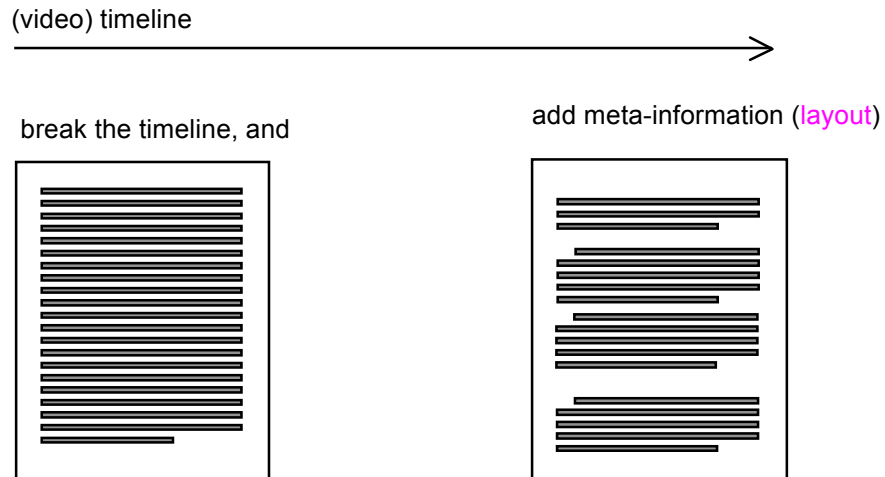
(video) timeline

break the timeline, and

add meta-information (layout)

Figure 3: Breaking the time-line in video content (left) and adding meta-information

OM-Images and Content Elements

In order to break a timeline we have to find the points in time where to interrupt the continuous stream of images. A video consists of shots, i.e. sequences that have been recorded continually. A natural structure of video content is to break at shot borders and to hyphenate a shot at a fixed temporal grid, e.g. every 4 seconds. The keyframes (pictures) that form the grid are called "object images" or O-images. In addition we use a visualization of the movement called M-image, defined by the plane of maximum image flow in a 3-dimensional image volume. (Müller, Tan 2000), Ardizzone, Cascia 1996)

The units generated by this subdivision of the timeline are called *Content Elements.* They consist of a series of large O-images (larger time grid, time difference e.g. 4 seconds), a series of small O-images (time difference e.g. 1 second), and a corresponding M-image. All three representations of the video are superimposed. Browsing across this content element is easy and intuitive, as the user can find a specific position in the video either by looking at the keyframes (O-images) or at the continuous time scale visualized by the M-image. Figure 4 shows an example content element taken from the movie "The Wizard of Oz".
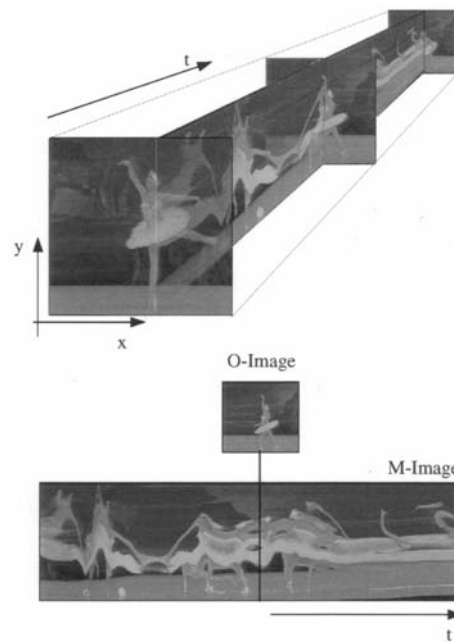
Figure 3: A visual representation of O-images (keyframes of the video) and an M-image (time section of the video); (Müller & Tan, 2000)

Figure 4: An example *content element* consisting of large O-images (top level), small O-images and an M-image (bottom level)

A *content element* is the basic building block representing a video clip, which can be manipulated by the user. Between text editing and video editing we can identify the following equivalences:

character	–	frame	
word	–	sequence of frames	– part of a Content Element
sentence	–	shot, clip	– Content Elemtent
paragraph	–	scene	– seq. of Content Elements
section	–	episode	
chapter	–	act	

If we use *Content Elements* mixed with text and graphics we come to the concept of a Video Book. Such a compound document is made-up by text,

graphics, content elements and layout information, e.g. headings, subheadings and line breaks. Figure 5 shows an example "sentence" from the "Vizard of Oz" movie and a draft lay-out of a Video Book.
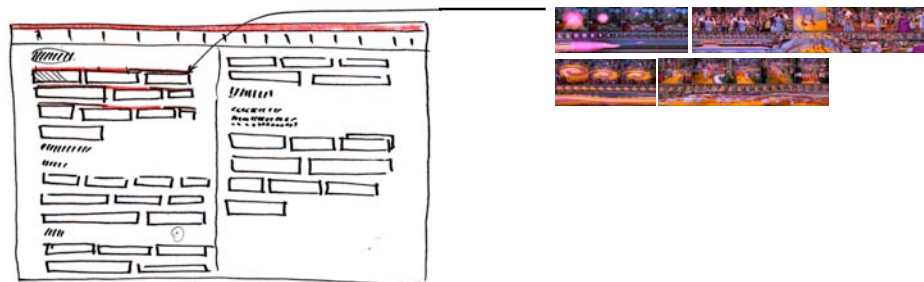

Figure 5: Video sentence consisting of video content elements (right side), Video Book (left side)

We can use the Video Book concept either for the visualization of the structure of an existing video and/or to support the whole creative process in video production. The definition of the first idea, a visual or textual sketch, the development of a script, the structuring of footage, editing, reviewing and documentation of the production process can be supported by the Video Book idea. (compare also ideas in van Dijk, 1980).

In a printed version of a Video Book, a director could even apply simple correction marks similar to the editorial marks used on a printed text document to communicate desired changes to the video production department, see figure 6.

Figure 6: Correction marks in a Video Book

The Video Book concept allows home and professional users to start right at the beginning of the creative process when writing a video outline and script. The same tool can be used later when structuring of video content and video editing. For consumers the Video Book concept supports video document styles and predefined video elements (templates) in order to enable a simple and very intuitive creation of audio-visual content.

Implementation

Several user interface studies have been created using OM-images and the Video Book concept. Figure 7 shows an interactive browser for movie and video content. In the lower middle part of the window there are three O-images and one M-image which is zoomed with a lens effect around the current position of the video cursor.
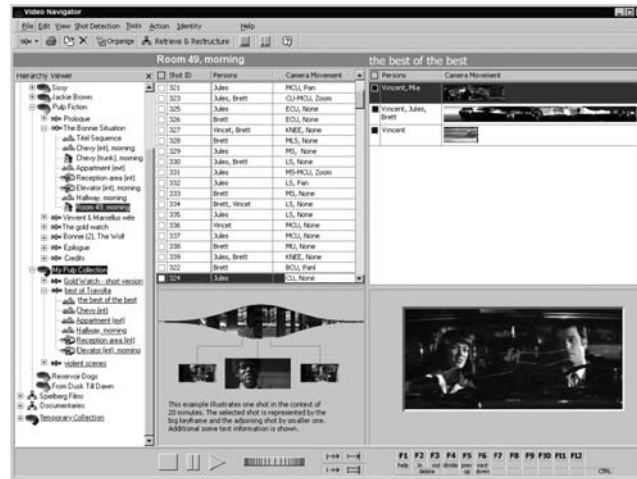

Figure 7: User interface concept for a video browser (Müller & Tan, 2000)

In the VIZARD project (funded by the European commission) a user interface based on the Video Book idea has been developed. Based on well known techniques in content based browsing and feature extraction (Aigrain et al. 1996) (Yeung et al. 1995) in the concept proof phase the prototype software *VideoPublisher* was implemented. Figure 8 shows a screen shot of the software prototype dealing with a template for a wedding video including a content element.

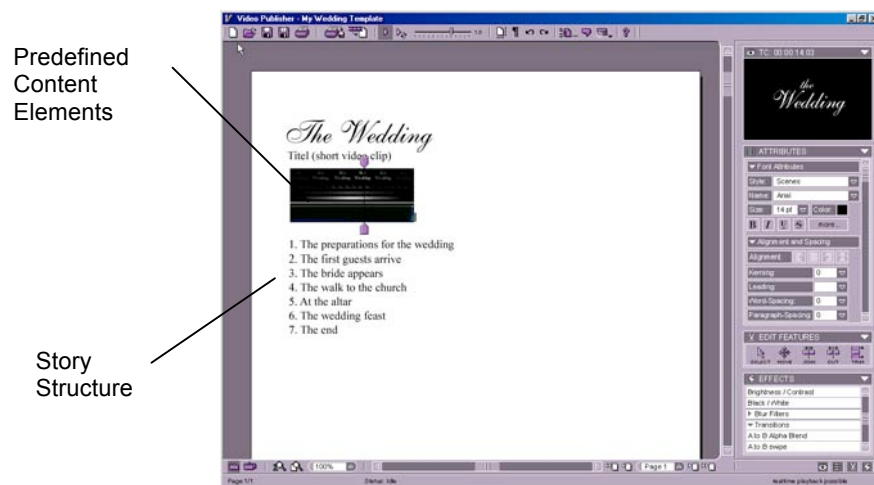Predefined Content Elements

Story Structure

Figure 8: A Video Template

Using the predefined structure an end user can create a simple script and import his raw material, see Figure 9.
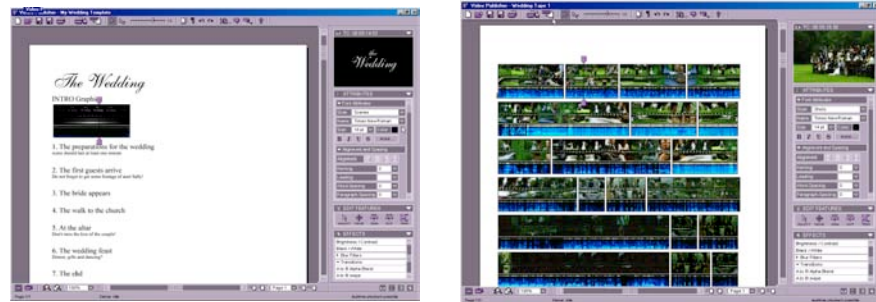

Figure 9: Writing a script and importing raw material.

In order to see more detailes within a Content Element the zoom factor can be increased. In the editing process layout information is added to the content elements. (Figure 8)
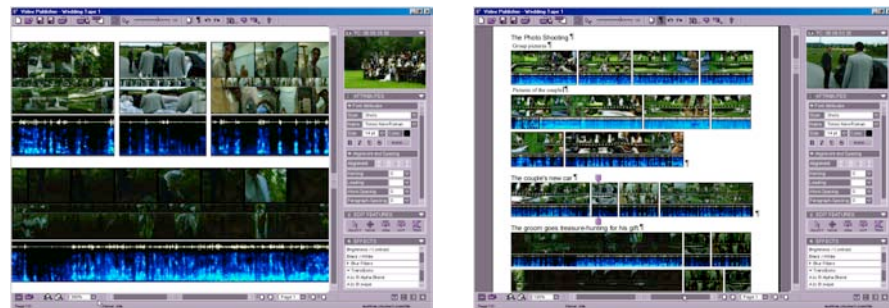

Figure 8: Zoom and Layout Elements

If, for example, the document contains lengthy shots where nothing much happens, an user may consider it convenient to display the paragraphs containing these shots in a very small size, allowing us to view more information on a single page and providing him with a sense of the importance of separate paragraphs. Alternatively, if he decides that some Content Elements contain images with a lot of detail, difficult to make out at a small size, scaling up these Content Elements can be of help.

Video Book – Evaluation

In the validation phase the *VideoPublisher* software was evaluated in a beta testing programme for home users (SONY Vaio Club) and by distributed testing and interviews with different end user groups. (Auffret and Boulanger 2003)

Beta testing Programe

The beta testing program was divided into 4 phases

1. Pre-test phase (6 weeks) : All preparations were done and the Club Vaio site informed potential testers on the scope of the program:

2. Test phase (4 weeks): During this time the following was set up. 196 end users applied, 21 testers were chosen.

3. Feedback evaluation phase: a mailing was sent to the testers to thank them and announce the end of the test phase

4. Feedback implementation phase: After an analysis of the feedback and the decisions which modifications to implement the *VideooPublisher* software was updated.

The response of the beta testers was mostly very positive. Recapulating their statements they said "Nice innovative concept", "Unusual but efficient", "Simpler and more intuitive than Premiere" and "Targeting to non professional users".

End User Interviews

In the evaluation phase several end user groups tested the usability of the Video Book approach. All partners of the VIZARD project performed usablity studies, e.g. thinking aloud. The following end user groups were involved.

- Librarians: professionals describing and indexing audio-visual documents from audio-visual libraries or archives (forum des images, Paris)

- Film editors/directors : professionals working on films to prepare a new feature or documentary movie (forum des images, Paris; ORF, Vienna)

- Producing documentalists and key researchers: are retrieving and collecting following the demands of a client. The key researcher prepares an EDL (Edit Decision List, cf. glossary) or a rough-cut and hands it over to the client. The final cut is done by a professional video editor. (ORF Vienna)

- University teachers, with no former experience in video editing (FH Joanneum, Graz)

- Professionals, but not directly related to editing: (two journalists, two graphic designers and one assistant producer (DuVideo, Lisboa).

As a conclusion and after intensive testing and validation by all user groups, it appears that the Video Book paradigm is considered as a new innovative and attractive concept by all users and that it is, in its current state, mostly a basic editing tool for non professional users.

For these users, it makes really video processing and editing easier and the transposition of textual manipulation habits to video seems to work perfectly.

For professional users with experience in film direction or video editing a great potential was identified in storyboarding.

Conclusion

We developed a new model for visualization and manipulation of digital video called the Video Book. This first user tests with the Video Book concept are very promising. The ultimate goal which we will pursue in the next future is to meet user needs of the mass market and to propose an entrance-level tool for video editing, helping to make this technology as simple and understandable as word processing.

Acknowledgment

References

Aigrain P., Zhang H.J., Petkovic D. (1996) Content-Based Representation and Retrieval of Visual Media – A State-of-the-Art Review, „Multimedia Tools and Applica-tions 3(3), pp. 179-202.

Auffret G., Boulanger P. (2003) – Vizard Deliverable D7.2, Validation Report, Forum des Images (Paris).

Ardizzone E., Cascia M.La, Video Indexing using Optical Flow Field (1996) Proc. of ICIP 96, vol 3, pp. 831-834.

Müller, H., Tan, E. S. (2000). Visualizing the Semantic Structure of Film and Video, in Visual Data Exploration and Analysis, Robert F. Erbacher, Philip C. Chen, Jona-than C. Roberts, Craig M. Wittenbrink, Editors, Proceedings of SPIE Vol. 3960, pp. 393-400.

van Dijk T.A. (1980), Macrostructures: An interdisciplinary study of global structures in discourse, interaction and cognition. Hillsdale NJ, L. Erlbaum.

Yeung M. et al. (1995) Video Browsing using Clustering and Scene Transitions on Compressed Sequences, Multimedia Computing and Networking, San Jose.