

Random Methods for Fast Exploration of the Raw Video Material

Tijana Janjusevic*, Sergio Benini†, Ebroul Izquierdo* and Riccardo Leonardi†

**Multimedia and Vision Research Group*

School of Electronic Engineering and Computer Science

Queen Mary, University of London, Mile End Rd, London, E1 4NS UK

Email: {firstname.lastname}@elec.qmul.ac.uk

†*DEA-SCL, Università di Brescia*

Via Branze 38, 25123 Brescia, Italy

Email: {firstname.lastname}@ing.unibs.it

Abstract—In this work we propose a visualisation and access method for exploring a database of raw video material, where the span of information stored can be huge and difficult to understand at a glance. Developed visual mining tool presents the overview of the hierarchically structured repository in a intuitive way and provides interactive navigational support for the user. Browsing through the un-annotated content is enabled by two complementary interactive methods, a direct selection and a random access, both supported by visual display of the preview nodes. User evaluation aims at demonstrating how the hierarchical random visualisation assists the process of accessing and retrieving content relevant to the user.

Keywords—Rushes, Database, Exploration, Browsing, Visualisation, Hierarchical Organisation

I. INTRODUCTION

The usage of digital equipment in the contemporary society leads to a constant generation of large digital media repositories. Due to the technological evolution, the problem of the storage space is solved revealing the issue of how to efficiently explore, understand and find items of interest in huge databases.

In most cases of professional video archives, broadcasters need to index, access and visualise huge professional archives of raw, unedited audio-visual footage, known as *rushes*. Most traditional forms of indexing, searching and retrieval methods are often not suitable for dealing with video material. On the other side, when exploring a huge rushes database, even a professional-user is not always aware of what is present in the collection, or he might not have a precise idea of the material he is looking for. Therefore there is the need of providing the user with useful tools for understanding and efficiently exploring the available content in order to accomplish his task.

To enhance the user's understanding of the database content, this work proposes a visualisation of a hierarchical structure of rushes previews. In general, abstract representation (as previews for videos, thumbnails for images, etc.) allow the user to quickly estimate the relevance of the information. By visualising hierarchical previews, the user is able to quickly extract high-level information glancing through the database content and structure.

Solutions such as Hyperbolic Tree [1], Treemaps [2][3] and Cone Trees [4][5], are information visualisation approaches for presenting hierarchy of data. Complexity of these displays can not always be justified since they can induce an additional obstacle for performing the browsing task. One of the issues we deal with is: how to keep low level of interface complexity while preserving the information about the hierarchical structure of the repository.

Hierarchical video/image browsing has been previously addressed in several works. In [6] the authors propose a system where visually similar images are grouped together using hierarchical clustering, while the zooming functionality enables the user to follow the path of a selected key-frame from the root to the nodes. In hierarchical visualisation approaches browsing is mostly initialised by selecting one of the number of displayed key-frames [6][7][8]. In such case due to the limited display space, we must either limit the number of displayed key-frames or reduce the image sizes. Either choice negatively affects the initialisation of the browsing since the user might not find the appropriate key-frame or the size is too small to understand the content. What we found as an important matter, not adequately addressed by these systems, is the way to initialise the query if no significant or query-similar key-frame is in the first visualised set. These systems do not consider alternative browsing or query initialisation option for the user in such cases.

In [9] authors propose video browsing system based on similarity threads between video fragments. The user can follow one of the four threads (textual, visual, time and semantic) in order to find content of interest. However the user does not have the freedom to explore the content in the initial interaction stage. Following the thread can be useful in case the user finds some content of interest but the important question is how to get to this stage in content exploration?

To address the previously stated concerns, we propose a visualisation of hierarchical video previews with a novel access method for browsing initialisation and content understanding. The proposed novel random access to a hierarchy node allows the user to explore the repository in case the

desired content can not be identified among the initially displayed key-frames. This way the user becomes aware of the material that the examined repository contains. Accompanying sequential access method enables the user to direct the query on his own, once a significant key-frame is found or similar content is identified. On each currently selected node, limited content set is displayed to avoid overflowing the user with information. This set is a random set of key-frames which can be refreshed in case the user asks to see additional content of the same node.

The paper is organised as follows: in Section II the algorithm for hierarchical organisation of the database is presented; in Section III the random and sequential access methods employed in user interaction are described; Section IV contains the user evaluation approach and the discussion about the obtained results. Conclusion and future work are stated in Section V.

II. HIERARCHICAL PREVIEWS

In the case of a large video database, it can be helpful for browsing to structure the given video material into a hierarchy, splitting the content into smaller subsets on different levels. For example, a professional user could be interested in browsing all “sport” segments separately from the “news” (see Figure 1). Then, on a lower level, it might be relevant to separate video material coming from different sports, such as “tennis” and “football”; similar considerations can be iteratively repeated through lower levels. Such arrangement of videos in a hierarchical tree where each node is labelled with a semantic category enables fast access and user understanding of the database structure.

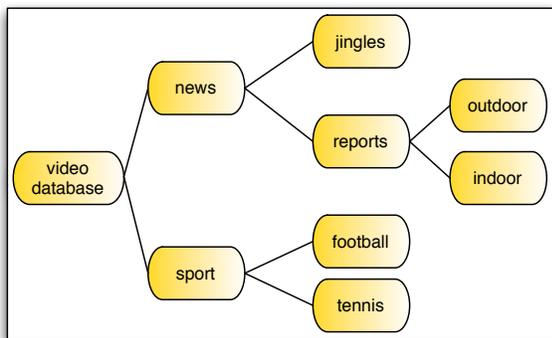


Figure 1. Hierarchical arrangement of a video database.

Unfortunately, in the case of rushes, most material lies unannotated in huge unorganised databases, so that a grouping based on semantic labels is not applicable. The proposed hierarchical summary is therefore obtained by a visual clustering of key-frames extracted from shots, where visual content is represented through a dictionary of visual words, as in [10]. Even if the grouping of similar content is based on visual similarity rather than semantics, the proposed

arrangement assists the browsing process by reducing the semantic gap between low-level features and high-level concepts familiar to the user.

A. Preview structure

Video summarisation techniques present collections of key-frames highly representative of the database content, which are arranged in a layered structure. From top to bottom, each layer of the hierarchy contains an increasing number of clusters of visually similar key-frames. Such organisation enables structured exploration: once the user identifies an interesting key-frame, he can interactively request more content from the same cluster, or refine his search by descending into the hierarchy.

B. Visualisation of the hierarchical structure

The developed user interface provides a visual navigation tool, forming a thread between the concealed structure and the user needs. In the upper window the hierarchical organisation is presented by using a tree view, as shown in Figure 2. The tree metaphor is chosen since it is a common way of representing a hierarchy and the majority of users are familiar with it.

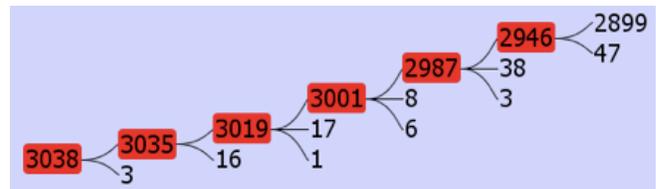


Figure 2. Tree view of the database structure (upper window).

The chosen tree metaphor in the upper window fulfils the following tasks:

- provides the visualisation overview at all times;
- describes the parent-child relations between content by visualising the fusion of nodes;
- provides the information of the number of key-frames in each node;
- facilitates the comprehension of the selected node current position within the database by colour encoding;
- supports the user moving forward, backward and making a progressive refinement by the access methods described in Section III.

In the lower window the preview key-frames belonging to the currently visited node are shown in a grid. Similar key-frames are placed sequentially in the grid in order to reduce the time needed by the user to grasp the node content. Combining both parts of the interface, the user has both global and local overview at all times, which is an important issue when developing an user interface for this type of data. Global, by seeing the structure and the current position at each step and local by observing the visual summary of the selected part of the repository.

III. RANDOM ACCESS AND DISPLAY

Since the entire database is too large to be explored in a single step, we need to provide some sort of “virtual tour” through the repository. Browsing is the commonly used method for examining the content: it begins with an entry point and consequently develops according to the retrieved results. Most hierarchical visualisation solutions use one single key-frame to represent the node content and initialise the browsing, as in [6]. In the work here presented, by displaying the tree view and the related preview key-frames, the user can initialise the browsing either by selecting a single key-frame or by clicking on any node of the hierarchy to grasp its content preview.

A. Random access

Starting the browsing, the user might be unaware of the what kind of data database contains. In traditional user interfaces which initially show set of key-frames, in case an interesting key-frame is not presented at first, the user most likely performs some attempts to find something by “randomly” moving within the database. The aim of these attempts is to bring the user to the content of interest or items similar to the user interest. We aim at showing that the best way to model the start of the “virtual tour” is to provide a novel system functionality which imitates the user “random” behaviour.

This browsing strategy is modelled by a statistical law, whose density function represents the probability to access one node of the hierarchy and display its summary. When the user selects this navigation modality, the algorithm randomly selects one level of the hierarchy with uniform probability; then, inside the chosen level, the probability of accessing one specific node is shaped by the following elements:

i) Entropy of the node: the probability of selecting a node is proportional to the quantity of information in the node itself, so that more informative nodes are more likely to be chosen with respect to less informative ones;

ii) Node dimension: the probability of selecting a node is proportional to the quantity of information in the node itself, so that nodes containing more images are promoted with respect to smaller ones;

iii) Browsing user’s history: the probability of jumping to a node containing key-frames that stimulated in the past the user’s interest is higher than nodes which have been disregarded in the past;

iv) User’s feedback: if the currently visited node somehow interests the user, the random search is statistically directed to close nodes, while if the shown content is totally different from the searched one, the algorithm takes the user on far branches of the hierarchy.

Elements from i) to iv) can be combined with different weights in order to shape the probability law which determines the access to nodes inside one level of the hierarchy.

If the visual preview of the selected node, which consists of a limited number of key-frames belonging to the related cluster, is of no interest for the user, he can continue his random walk to another node. This access method takes the user randomly through the different content sets at different levels of hierarchy, until the user is able to find an “interesting” key-frame and choose to follow the information scent.

B. Sequential browsing

From the moment the user finds a key-frame or a set of key-frames that attract him, he can refine his quest by using a sequential access. This method is supported both by key-frame and node selection. By clicking on the key-frame, the application takes the user to the deepest node of the tree which contains the selected item. In case the user wants to explore the video of the selected key-frame playing is supported by the integrated player. In the upper window the complete path of nodes containing the selected key-frame is highlighted by colour encoding (as shown in Figure 2), while the lower window presents the deepest node content preview. By clicking on nodes the user can move backward and forward through the tree to refine his search, while constantly being aware of the current position inside the database.

C. Random display of key-frames

Both random and sequential explorations are assisted by the display of the visual content previews in the bottom application window. Our initial solution displays a grid of maximum 21 key-frames, where appropriate image sizes allow the user to see the visual content summary without any additional interaction (zooming, etc.).

The aim of a visual summary is to show a set of representative key-frames for each node in the hierarchy. However there are no semantic labels that could assist us in defining the appropriate representative set for the selected cluster. Even already proposed solutions, such as showing the image centroid [7], or showing one key-frame for child node [8], can not guarantee an adequate selection of key-frames. Therefore we adopt a similar approach to the one presented in [6], and we randomly extract from each node a limited set of key-frames to be displayed.

Although a limited number of images can not completely represent the content of the cluster, this approach avoids information overflow or inadequately re-sized key-frames and benefits from low computational complexity. In case the user wants to perform further exploration of the same node, he can request for additional content. In this case a new random set is extracted from the node and displayed. A screenshot of the whole user interface is shown in Figure 3.

The selected node is the last red node on the right belonging to the coloured search-path in the hierarchy. The set of keyframes belonging to this node is displayed in the

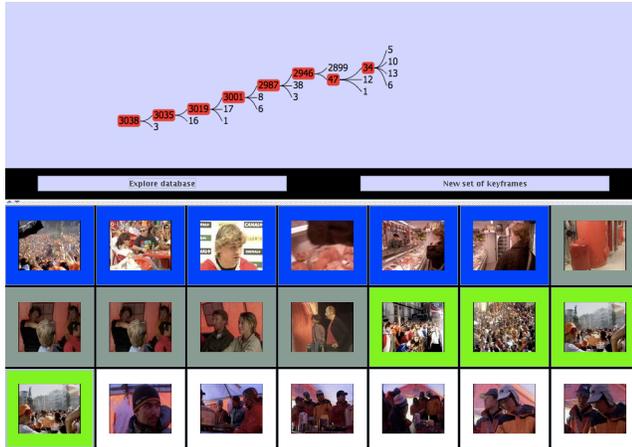


Figure 3. Screenshot of the interface: tree view of the database structure (upper window); “Random access” and “Show more keyframes” buttons; preview key-frames (lower window).

grid below. The colour encoded borders of the grid cells differentiate content of the different child nodes so that the user can potentially define further directions to follow through the tree.

IV. EVALUATION ON RUSHES

The proposed application is dedicated for *rushes* and the repository used for tests and evaluation contains 80 rushes which were provided by the main Basque broadcaster EiTB [11]. The evaluation of basic application implementation was based on professional use cases and performed by 5 journalists from EiTB. The application uses more than three thousands extracted key-frames, which are automatically organised in hierarchical previews based on visual similarity. The evaluation was performed by using a random access algorithm where the probability of accessing a node within one level is uniformly distributed within all nodes of the level itself.

A. Evaluation methodology

For the usability evaluation of the proposed application, tests were conducted in a realistic user scenario with an approach known as *simulated work task situation* [12]. In line with this approach, the application is evaluated by the journalists the application is built for, under real usage conditions and in their working environment. Targeted users were given two professional user-cases chosen from the ones defined within the EU funded project RUSHES [13], *i.e.*:

- **User-case 1:** Find key-frames containing aeroplanes.
- **User-case 2:** Find key-frames showing a flood scene.

Examples of content related to the tasks are given in Figure 4.



Figure 4. Examples of flood and plane key-frames from the dataset.

B. Usability evaluation

After performing the defined tasks, EiTB journalists were asked to fill a questionnaire (given in Table I) for rating their satisfaction with the most important aspects of the proposed solution.

No.	Statement
1.	It is easy to use the interface
2.	Interaction with the interface is comfortable
3.	It is easy to understand the functionality of the interface
4.	It is easy to understand the functionality of each button
5.	The tree in the upper window is helpful for understanding the database structure
6.	The interface is well organised
7.	The functionality of the interface is intuitive
8.	The organisation of the visualisation display is clear
9.	The tree in the upper window is helpful for browsing and accessing content
10.	The key-frames displayed in the bottom window provide a good overview of the node content
11.	Random exploration is helpful for browsing when I don't know where to find the desired content
12.	The interaction method is intuitive
13.	The colour of the tree nodes helps me to understand my position in the database
14.	The interface is pleasant to use

Table I
SELECTED QUESTIONS FROM USER QUESTIONNAIRE.

The design of the questionnaire enables the user to mark their level of agreement with the provided statements. Marks were given using the typical format of five-level Likert scale (1 - Strongly disagree; 2 - Disagree; 3 - Neither agree nor disagree; 4 - Agree; 5 - Strongly agree). Obtained evaluation results are given in Figure 5 as a diagram of the marks per each question.

After filling the questionnaire the journalists stated positive and negative aspects of the application, and gave personal comments on potential improvements. The most frequent positive and negative aspects are listed in Table II.

As it is evident in Figure 5, the novel random access exploration has been highly appreciated by the journalists as an useful tool for browsing when they didn't know where to find the desired content. The same positive comment on random access was also one of the most cited among the reported beneficial aspects of the application (in Table II).

Several journalist recommended to provide more information about the video such as title, date and duration.

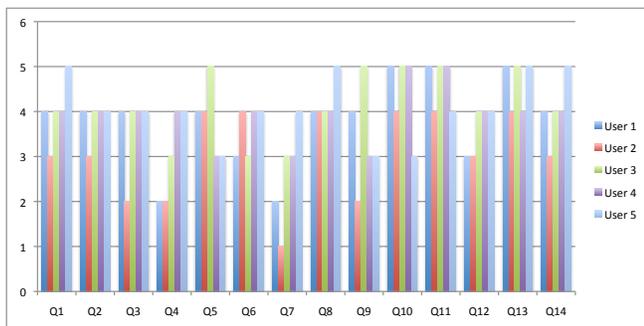


Figure 5. Marks per question indicate positive evaluation results for the proposed visual mining interface.

They also emphasised the importance the option of going a step backward in the navigation process, in case the chosen direction of exploration is not fulfilling their expectation or they navigated too fast.

No.	Positive remarks
1.	Comfortable tool for browsing clip databases
2.	Random exploration is helpful to find content without text-based search
3.	Preview key-frames are good representation of the node content
No.	Negative aspects
1.	Add more information about the clips
2.	Buttons should have icons and tool-tips
3.	Displayed key-frames should be adaptive to node size

Table II
MOST FREQUENTLY CITED POSITIVE AND NEGATIVE REMARKS.

V. CONCLUSIONS

The hierarchical visualisation solution proposed in this paper addresses the task of browsing and/or searching through a rushes repository. In order to reduce the overload of the displayed information rushes are organised by a hierarchical clustering approach. Apart from dedicated integration of several visualisation methods, the novelty of the presented solution lies in modelling the user navigational behaviour with a combination of a random and a sequential access for exploring the database. Navigation support is provided by visualising the database tree structure and displaying the content preview for each tree node. User tests and evaluation were performed on the basic application version, where random jump probabilities are uniformly distributed over all nodes within a tree level. These tests showed that the proposed solution helps professional users to understand what type of content is contained in the repository. There are reasonable expectations that further evaluation of the random application with advanced statistical modelling for accessing nodes will show improved evaluation results.

ACKNOWLEDGMENT

This research work has been partially supported by EU project *RUSHES (FP6-045189)*. We would also like to thank Leticia Fuentes Ardeo, Mikel Frutos Hernandez and journalists at EiTb for priceless help during the experimental evaluation.

REFERENCES

- [1] J. Lamping and R. Rao, "Laying out and visualizing large trees using a hyperbolic space," in *Proc. of UIST '94*. New York, NY, USA: ACM, 1994, pp. 13–14.
- [2] B. Johnson and B. Shneiderman, "Tree-maps: a space-filling approach to the visualization of hierarchical information structures," in *Proc. of VIS '91*. Los Alamitos, CA, USA: IEEE Computer Society Press, 1991, pp. 284–291.
- [3] B. Bederson, "Photomesa: a zoomable image browser using quantum treemaps and bubblemaps," in *Proc. of UIST '01*. New York, NY, USA: ACM, 2001, pp. 71–80.
- [4] G. G. Robertson, J. D. Mackinlay, and S. K. Card, "Cone trees: animated 3d visualizations of hierarchical information," in *Proc. of CHI '91*. New York, NY, USA: ACM, 1991, pp. 189–194.
- [5] M. A. Hearst and C. Karadi, "Cat-a-cone: an interactive interface for specifying searches and viewing retrieval results using a large category hierarchy," in *Proc. of SIGIR '97*. New York, NY, USA: ACM, 1997, pp. 246–255.
- [6] D. Borth, C. Schulze, A. Ulges, and T. Breuel, "Navigator-Similarity Based Browsing for Image & Video Databases," in *Proc. of KI 2008*, vol. II, Kaiserslautern, Germany, September 2008, pp. 22–29.
- [7] G. Schaefer and S. Ruszala, "Hierarchical image database navigation on a hue sphere," in *Proc. of ISVC06*, 2006, pp. II: 814–823.
- [8] S. Krishnamachari and M. Abdel-Mottaleb, "Image browsing using hierarchical clustering," in *Computers and Communications, IEEE Symposium*, vol. 0. Los Alamitos, CA, USA: IEEE Computer Society, 1999, p. 301.
- [9] O. de Rooij, C. G. M. Snoek, and M. Worring, "Query on demand video browsing," in *MULTIMEDIA '07: Proc. of the 15th international conference on Multimedia*. New York, NY, USA: ACM, 2007, pp. 811–814.
- [10] S. Benini, A. Bianchetti, R. Leonardi, and P. Migliorati, "Extraction of significant video summaries by dendrogram analysis," in *Proc. of ICIP'06*. Atlanta, GA, USA, 8-11 October 2006.
- [11] EiTb, "Euskal irratia telebista," <http://www.eitb.com/>.
- [12] P. Borlund, "The IIR evaluation model: a framework for evaluation of interactive information retrieval systems," *Information Research*, vol. 8, no. 3, 2003.
- [13] R. FP6-045189, "Retrieval of multimedia semantic units for enhanced reusability," <http://www.rushes-project.eu/>.