Flexible Navigation in Smartphones and Tablets using Scalable Storyboards

Shuai Zheng[1,2,3], Luis Herranz[1,2], Shuqiang Jiang[1,2]
[1]Key Lab of Intelligent Information Processing, Chinese Academy of Sciences, Beijing 100190, China
[2]Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China
[3]School of Software Engineering, Beijing Jiaotong University, Beijing 100044, China
zs910504@yahoo.cn, luis.herranz@vipl.ict.ac.cn, sqjiang@ict.ac.cn

ABSTRACT

In this demo paper we present a multiscale browsing interface for handheld devices, in which the user can interactively change the scale of the storyboard to easily adjust the amount of information desired. Conventional and hierarchical storyboards provide one or very few possible lengths. In contrast, scalable storyboards allow the number of images and the storyboard itself to be adapted to the device constraints (e.g. aspect ratio, resolution) and navigation state with much finer granularity. Several levels and modes, including segment of interest, are provided for more intuitive and convenient navigation.

Categories and Subject Descriptors

H.3.1 [**Information Storage and Retrieval**]: Abstracting methods

General Terms

Algorithms, Design

Keywords

storyboards, browsing, navigation, adaptation, scalable summaries

1. INTRODUCTION

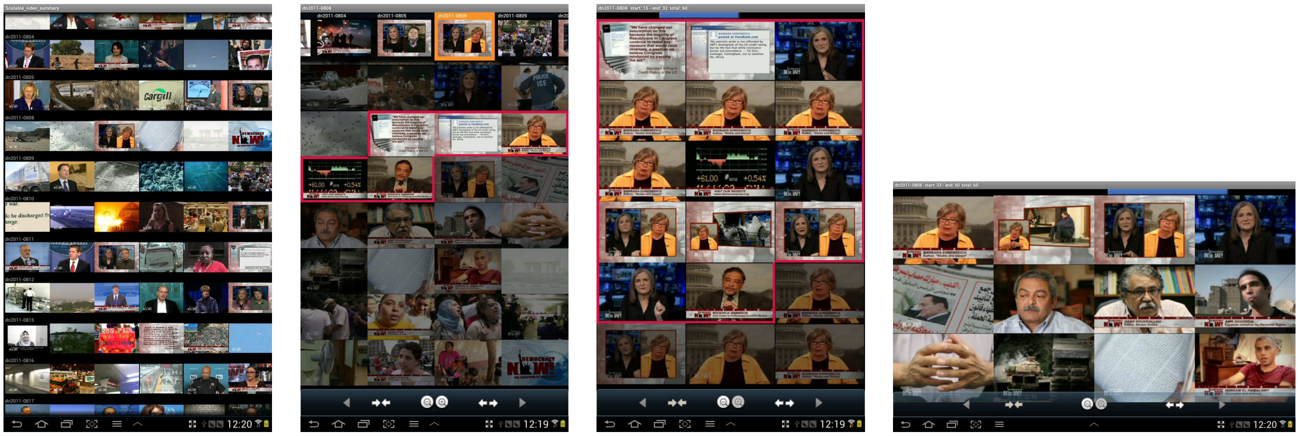
In recent years, mobile devices have become the main tool to access multimedia content. Typically, users use their phones or tablets to browse through hundreds of videos from diverse sources, ranging from online digital libraries (e.g. YouTube) to local video collections stored in the device. Current smartphones and tablets have enough storage capacity for hundreds of videos. For those reasons, intuitive and effective browsing interfaces are critical. When exploring large lists of videos, users cannot spend too much time visualizing the actual video. Instead, a quick idea about the content is enough, so compact visual abstracts are often used in search and browsing applications. In the case of mobile devices, these abstracts are often limited to the title and one keyframe. However, longer abstracts (e.g. storyboards, video skims) can provide more details about each video at the cost of slower browsing time. Conventional video summarization techniques[2] generate summaries with a certain length (e.g. number of keyframes in storyboards, duration in video skims). This length is fixed and cannot be adjusted on demand, which is not very suitable for flexible navigation and visualization in dynamic interfaces. In this demo we present a flexible browsing interface in which storyboards can be adapted to different screen sizes and to the specific navigation state, based on the idea of scalable storyboards[1]. Compared with the other family of multiscale summaries, i.e. hierarchical summaries[3], scalable storyboards adaptation is much more flexible. In hierarchical summaries, users just can browse images in fixed navigation levels such as chapters or scenes, and navigate within units with strict boundaries. In contrast, scalable storyboards enable a more flexible adaptation with finer granularity. Users can select an arbitrary segment (segment-of-interest) to explore in more detail, and then the system will provide a more detailed storyboard focusing on that particular segment.

Compared with text and image processing, video analysis requires much more computational effort. Due to limited resources in mobile devices, such as processing capabilities or battery life, it is not feasible to include too complex analysis or analysis requiring to process all the frames (e.g. shot detection). To address this problem, we first extract a set of keyframes by sampling the video at constant intervals, obtaining a long storyboard, but approximately covering most of the content. Then we perform content analysis over the keyframes to organize them with a scalable structure (i.e. scalable storyboard), stored in a scalable description. When a certain number of keyframes is required in the storyboard, the system just needs to check the scalable description to retrieve the list of keyframes to create that particular scale of the storyboard. This simple adaptation strategy is very fast and requires very little computational effort.

2. SCALABLE STORYBOARDS

In general, when a new storyboard with a different length is required, the summarization process needs to be executed again, with appropriate parameters. In contrast, with scalable storyboards the analysis algorithm is executed only once, and extracting summaries with different lengths on demand is very simple and fast. We use a scalable representation based on hierarchical clustering[1]. The keyframe rate is one keyframe per minute, which is often more than enough for explorative browsing, and also not too high to allow fast videos analysis (e.g. the tablet used in the tests can process one hour videos in just a few seconds). During the interaction, the system solves an adaptation problem, in which the number of images required is dynamically computed based

ICMR'13, April 16–20, 2013, Dallas, Texas, USA.
ACM 978-1-4503-2033-7/13/04.



(a) (b) (c) (d)

Figure 1: Navigation using scalable storyboards: a) collection browser, b) storyboard of the full video with a segment-of-interest highlighted, c) storyboard of the segment-of-interest conveniently scaled, d) landscape interface after rotation.

on device constraints (mainly display size/resolution) and the current navigation state (number of columns, temporal span, segment of interest boundaries). Then, the corresponding scale of the storyboard is displayed.

3. INTERACTION MODEL

The system provides navigation in three different levels (*collection*, *video* and *segment*), that can be dynamically changed by the user during the navigation process. Within each of these levels, the user can navigate through different adaptation dimensions:

- *Summarization detail*, related to the amount of information presented per time or area unit. Typically the user interface allows to adjust this scale indirectly by increasing or decreasing the number of columns of the storyboard.
- *Temporal span*, corresponding to the temporal interval presented in the display area.
- *Temporal position*, related with the position in the timeline of the temporal interval.

The user interface and an example of navigation is shown in Fig. 1. The control bar enables navigation by zooming in summarization detail, to different videos/segments and changing the temporal span. Pinch and flick gestures are also implemented to provide intuitive multitouch interaction. The collection browser (see Fig. 1a) displays a list of videos, in which each row represents the storyboard of a certain video item in the collection. Using the navigation controls, users can dynamically change the number of images in the storyboard to show fewer but larger images, or more result videos. When the user clicks on certain item, a full page storyboard is shown (video level, see Fig. 1b). In addition to summarization zooming, the user can change the temporal span and the temporal position of the segment, which provides navigation at segment level (see Fig. 1c and d). The duration and the temporal position of the segment is shown as a blue rectangle. A more flexible tool is the segment of interest navigation, which allows the user to select several consecutive images and zoom to that particular segment (shown highlighted and with a red frame in Fig. 1b and c). Using scalable storyboards, the length of the storyboard is selected automatically to fit as many images as possible in the available display area with a very fine granularity of one image. The same mechanism also enables adaptation to different devices or adaptation to other orientation when the device is rotated (see Fig. 1d).

4. CONCLUSIONS

In this demo we have presented a browsing interface of video collections which allows users to have quick navigation at different levels, and providing functionality to dynamically change the detail scale. It also enables zooming to an arbitrary segment of interest to get more detailed information. The core of the interface is based on scalable storyboards, which enables displaying suitable storyboards regardless of the display area and navigation state.

Acknowledgments

This work was partly supported by the National Natural Science Foundation of China under Grants 61150110480 and 61070108, by the Chinese Academy of Sciences Fellowships for Young International Scientists under Grant 2011Y1GB05 and by the National Basic Research Program of China (973 Program): 2012CB316400.

5. REFERENCES

[1] Luis Herranz. Multiscale browsing through video collections in smartphones using scalable storyboards. In *ICME Workshops*, pages 278–283, 2012.

[2] Ba Tu Truong and Svetha Venkatesh. Video abstraction: A systematic review and classification. *ACM Trans. on Multimedia Computing, Communications and Applications*, 3(1):3, 2007.

[3] Xingquan Zhu, Jianping Fan, Ahmed K. Elmagarmid, and Xindong Wu. Hierarchical video content description and summarization using unified semantic and visual similarity. *Multimedia Systems*, 9(1):31–53, July 2003.