ABSTRACT

In this work, we present Video Collage system, which automatically constructs a compact and visually appealing synthesized collage from a video sequence for efficient video browsing. Given a video, Video Collage is able to select the most representative images, extract salient regions of interest (ROI) from these images and resize ROI according to their saliencies, and seamlessly arrange them on a given canvas while preserving the temporal structure of video content. Furthermore, Video Collage provides a novel user interface that enables users to browse video content in a variety of more efficient ways in contrast to many existing approaches to video browsing.

Categories and Subject Descriptors

H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—video; I.2.10 [Artificial Intelligence]: Vision and Scene Understanding—video analysis

General Terms

Algorithms, Experimentation, Human Factors.

Keywords

Video content analysis, video presentation, video collage.

1. INTRODUCTION

The quantity of video content is increasing dramatically with the popularity of consumer digital capturing devices. While online delivery of video content surged to an unprecedented level in current years, today’s users face a daunting volume of videos. Highly compact and visually appealing representation of video content is a key to effective and efficient browsing of video sequence, be it an unedited home video, a professional video program, or an online video clip. Although there have been a lot of efforts on video representation such as Booklet [2] and pictorial [5], most of them are lack of satisfying presentation layout and compact form, which make them not very efficient for quick browsing and understanding of the whole video content.

We demonstrate in this work a more efficient system for video browsing, named Video Collage, which is built upon our previous work on static video representation [4]. This novel system is able to not only automatically generate a highly compact and visually appealing synthesized collage, but also provide an effective user interface for browsing content in a variety of efficient ways. Fig. 1 and 2 shows a sample 2D and 1D collage of a home video. Video Collage has the following advantages.

2. GENERATION OF COLLAGE

Given a video sequence V, Video Collage first selects N (N can be given by the user or predefined) representative...
images from \( V \) and arrange the ROI of these images on a collage. For the sake of simplicity, we assume that each image only has one ROI. Let \( \lambda = \{ I_i, R_i \}_{i=1}^N \) denote one possible solution, where \( I_i \) and \( R_i \) denote selected image and its ROI, we formulate the generation of collage as finding the best \( \lambda \) to minimize an energy or cost \( E(\lambda) \)

\[
E(\lambda) = \omega_1 E_{rep}(\lambda) + \omega_2 E_{trans}(\lambda)
\]  

(1)

where \( E_{rep}(\lambda) \) denotes the cost from representativeness of \( \lambda \), \( E_{trans}(\lambda) \) denotes the cost of any transition that is not visually smooth, \( \omega_1 \) and \( \omega_2 \) are two predefined weights. The representativeness cost is measured by the combination of visual saliency, image quality and distribution of selected images \([3]\), while the transition cost is measured by the sum of color differences across the seams of neighboring ROI. The cost \( E(\lambda) \) is minimized by a two-step heuristic searching algorithm. For more details, please refer to \([4]\).

3. BROWSING WITH COLLAGE

We have extended our work on collage generation to build a novel video browsing system. Figure 3 shows Video Collage user interface. In addition to the four panels in Figure 3, users can change collage resolution (i.e., the number of ROI in collage) by moving the marker on slide bar (i.e., the bar between A and B) vertically to view the collage content in different resolution. Video Collage supports the following video browsing manners.

- By 2D static collage – watch 2D collage in A and left-click a specific ROI to access the corresponding video content in panel B.
- By 2D dynamic collage – watch 2D collage in panel A and right-click a specific ROI to select playing the corresponding clip in A or playing all the clips in A (on a pop-up menu). It is similar to a “Video Wall” consisting of many video thumbnails in which each thumbnail corresponds to a short clip \([1]\). The differences between 2D collage and Video Wall are threefold from the viewpoint of representation: first, collage is composed of ROI instead of images in Video Wall, which makes collage more compact; second, the thumbnails in collage are resized according to saliencies while those in Video Wall have the same size; third, collage is designed for a single video while Video Wall for multiple videos.
- By 1D static collage – similar to the first manner except for 1D collage in panel C.
- By 1D dynamic collage – similar to the second manner except for 1D collage in panel C.
- By key-frames – watch key-frames in panel D and click a specific key-frame to access the corresponding video content in panel B.

By the above five manners, users can browse the video content very efficiently.

4. CONCLUSIONS

We have demonstrated Video Collage which supports a very compact and visually appealing representation and a variety of ways for browsing of video content. Although we mainly applied Video Collage into home video and movie, it is very easy and interesting to integrate collage into online video scenario. Our future work include combining the textual information such as title and closed caption to a more rich representation of video content, and adding more interactive browsing functionalities into current system such as enabling users to drag and resize ROI manually.

5. REFERENCES