Storyboards from Video (online id 0020)

Ciaran Wills
Department of Computing Science, University of Glasgow
c.j.wills@acm.org

1 Introduction

This sketch presents our work towards the automatic summarisation and annotation of video. A video clip is analysed to produce a static 2d summary in the form of an annotated storyboard. This storyboard conveys salient information about the content of a clip in a form that can be easily displayed or printed, and understood in a fraction of the time it would take to view the video itself. Rather than attempting to have the computer analyse and understand the video we instead aim to present the information in a form that aids the human user, who is far better equipped for such high-level understanding.

Our summary displays each shot as a single storyboard frame that quickly conveys the essence of that shot. It should be immediately obvious to the user what the principal objects are, their essential shapes and colours, and how both they and the camera are moving.

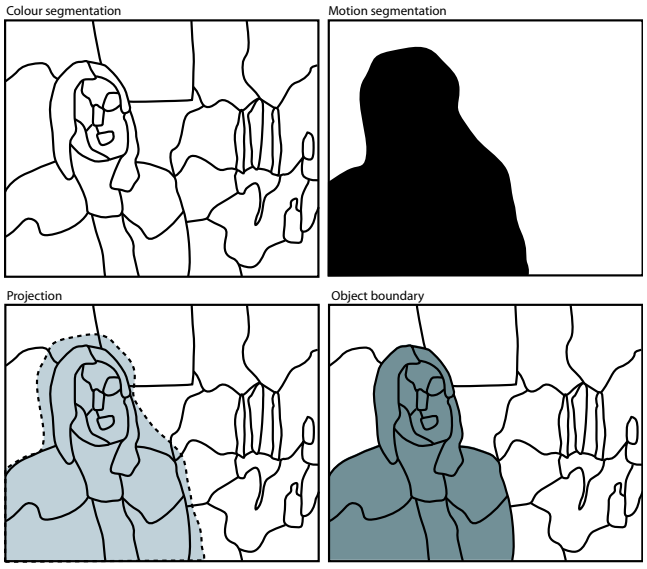
This work was carried out in the context of a commercial post-production facility and so the nature of the material we are handling is quite specialised. The techniques we have developed are tailored specifically for this type of video.

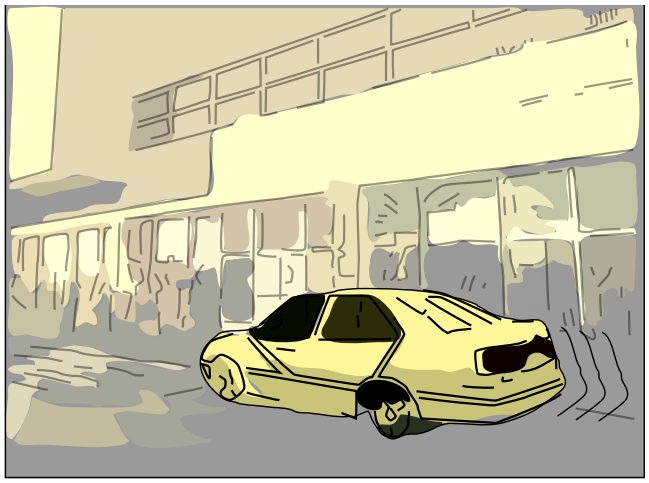
2 Video Analysis

Given an unstructured video stream we need to first break it down into shots, and then identify the objects and movements within each shot.

Scene change detection has received a lot of research attention but we found that the best-of-breed algorithms performed poorly because of the highly dynamic nature of commercials. A new algorithm was developed to address these issues, providing improved performance on such material and comparable performance for other material. The new algorithm can also operate directly on DCT compressed video, making it very efficient.

To identify the objects within each shot we make an important assumption; that objects have coherent motion independent of the background. We identify objects using a technique based on the work of the European COST 211 group on MPEG-4 encoding [Alatan et al. 1998]. Colour and motion based segmentations are performed on the shot simultaneously. The colour segmentation finds accurate boundaries but is unable to differentiate important edges from other detail, while the motion segmentation can identify the moving regions but with inaccurate borders. By projecting the motion segmentation onto the colour segmentation, as shown in the figure, we can find an accurate boundary for the moving object.





3 Presentation

We wish to present the storyboard frames in a visually appealing way that quickly and clearly conveys the important information. These are not new problems and so we borrow techniques from comic (and indeed storyboard) illustration, where visual techniques and conventions for depicting such information have developed over many years.

A bézier outline is fitted to each object and tracked throughout the shot using a snake algorithm. By following the bézier control points over time we can find the path of the object. The trails of the control points are used to draw either repeating contours or speedlines that indicate the motion of the object [Masuch 1999].

Any camera motion is apparent through the motion of the background region. Rather than trying to classify all the possible camera movements we instead annotate each corner of the frame with an arrow proportional to the direction and magnitude of the background motion at that corner. Although simple, this makes the background motion immediately obvious to the viewer.

Additional detail of the objects and background is found through a controlled colour segmentation of each region. The segmented regions are then represented as bézier shapes and used to fill each region. Further detail is found using an edge detector and fitting black lines to the detected edges. The background is desaturated to draw attention to the foreground objects.

The author would like to acknowledge the support and assistance of Chris Eborn and Smoke & Mirrors Ltd. in this work.

References

Alatan, A., Onural, L., Wollborn, M., Mech, R., Tuncel, E., and Sikora, T. 1998. Image Sequence Analysis for Emerging Interactive Multimedia Services—The European COST 211 Framework. *IEEE Trans. on CSVT* 8, 7, 802–813

Masuch M. 1999. Speedlines: Depicting motion in motionless pictures. In *Sketches & Applications, SIGGRAPH 1999*, 277.