Interactive Montages of Sprites for Indexing and Summarizing Security Video

Chris Pal
University of Massachusetts
Dept. Computer Science
Amherst, MA, USA
Email: pal@cs.umass.edu

Nebojsa Jojic
Microsoft Research
Redmond, WA
USA
Email:jojic@microsoft.com

DESCRIPTION

In this video we present a new model of interaction for indexing and visualizing video in the context of security applications. We wish to index security video that contains relatively rare but important events, such as security video taken from large public or industrial facilities (e.g. a nuclear power plant) or exterior home security video from a residential neighborhood. We present a method of indexing video by arranging irregularly shaped icons or sprites into a montage representing motion events or security events within the original video scene. The sprites in the montage are used as an index into the original video. We also generate video montages to summarize video in which motion events are compressed and overlayed in a video of shorter time duration. This summary video also acts as an index into the original video stream. We use a simple, novel method of extracting sprites for the image and video montages based on incrementally building a Gaussian mixture model with conjugate priors for the background. We then use fast morphological operators to extract foreground elements. Our approach can be viewed as a fast maximum a posteriori (MAP) inference procedure in a layered image model. The contributions of this work are new interaction and summary schemes that allow viewers to potentially survey hours of security video in the order of minutes.

In our approach we first determine a set of initial clips or frames in the video that constitute motion events. To obtain the indices of frames for these events we perform an initial low complexity pass of the video data. In this pass we look for segments of video where groups of pixels in the image undergo a coherent and rapid change from the background. We build a background model of the security scene and adaptively update this model to account for more slowly occurring lighting changes and alterations of camera position. To achieve this, we compute a MAP assignment of each pixel to the Gaussian mixture class or a uniform distribution foreground class. This allows us to quickly determine a binary assignment into foreground vs. background for pixels that do not "fit" the current background model. We then applied a binary median filter to remove point "classification noise" in the (binary) foreground class assignment image. Finally, we applied morphological closing and opening operations to isolated connected "blobs" of pixels. We detect "changes" when a fraction of pixels in the resulting image were different from a threshold. See [1] for an empirical comparison of variations of this type of procedure for event detection in surveillance and security applications. Once we have a set of motion events, we treat each event as a "video sprite".

Fig. 1
A STATIC SUMMARY OF HALF AN HOUR OF SECURITY VIDEO

Figure 1 shows a screen shot of a static montage of sprites which acts as an index back into the original security video video. Here we use a single key frame from the middle of each video sprite to act as a "clickable" index. In our video we present a novel method for indexing the original video based on an animated, video montages of sprites. In these video representations, we overlap events in time so as to compress the action of a video sequence into a much shorter, animated index video. This allows the user to review much larger quantities of video in a given time period.

ACKNOWLEDGEMENTS

We thank Sumit Basu for assistance recording the video.

REFERENCES

[1] E. Stringa "Morphological change detection algorithms for surveillance applications," *Proceedings of the British Machine Vision Conference (BMVC) 2000, Bristol, UK, 11-14 September*, 2000.