LAFCam – Leveraging Affective Feedback Camcorder

Andrea Lockerd
MIT Media Lab
20 Ames St
Cambridge, MA 02139 USA
alockerd@media.mit.edu

Florian Mueller
MIT Media Lab
20 Ames St
Cambridge, MA 02139 USA
floyd@floydmueller.com

ABSTRACT
If a video camera recognizes and records affective data from the camera operator, this data can help determine which sequences will be interesting to the camera operator at a later time. In the case of home videos, the camera operator is likely to also be the editor and narrator of the final video. LAFCam is a system for recording and editing home video. We facilitate the process of browsing and provide automatic editing features by indexing where the camera operator laughed and visualizing the skin conductivity and facial expressions in the editing session.

Keywords
Laughter detection, video editing, galvanic skin response, camcorder, picture-in-picture, camera-on-camera, bio-signal, annotation, facial expression, browsing, index

INTRODUCTION
Digital appliances that sense more about the user and the environment have a better opportunity to enhance the user's experience. Take, for instance, the scenario of shooting home videos. It is easy for a video amateur to gather a large amount of video, which then needs to be edited since no one wants to watch the four hours of footage from your most recent holiday. This task of finding the few interesting pieces can be labor intensive and time consuming, requiring the user to review hours of video.

Working in this home video scenario affords a few assumptions about the user. First, that she/he has multiple roles in the content creation process: camera operator, editor, and narrator. Under this assumption, the LAFCam system uses information about the user from one stage (camera operator) to enhance the other (editor). We also assume that the camera operator uses the LCD on the camera.

APPROACH
The motivation of LAFCam is three-fold:

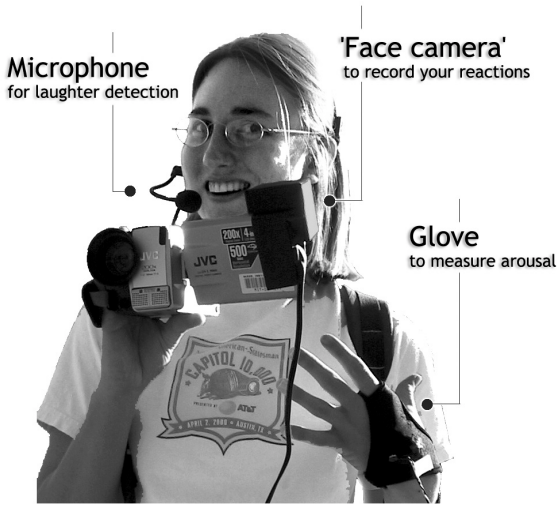

Figure 1. Three affective input channels

- Marking salient sequences: Indicating salient points in the video enhances browsing through the video. LAFCam visualizes two affective channels in the editing phase, suggesting parts the operator found interesting during the shoot.
- Automatic editing: An automatic edited movie can provide an easy overview and serve as a base for further editing.
- Adding perspective: "One long shot is the reason vacation videos are so boring!" [1]. LAFCam adds an additional perspective, the camera operator's face, which can be edited into the final video.

Using bio-signals to start/stop a camera originated with the startle cam [2]. We took this further and use three different affective channels as an input to a video editing application.

IMPLEMENTATION
LAFCam (Figure 1) detects laughter of the camera operator, records skin conductivity, and films the camera operator's face.

Laughter Recognition
We trained Hidden Markov Models (HMMs), one for laughter and one for all other speech using Expectation Maximization of spectral coefficients of the audio signal [3]. For testing, we used a data corpus of 40 laughter segments and 210 speech segments (collected over two weeks, from a single user). We used 70% for training and tested with the

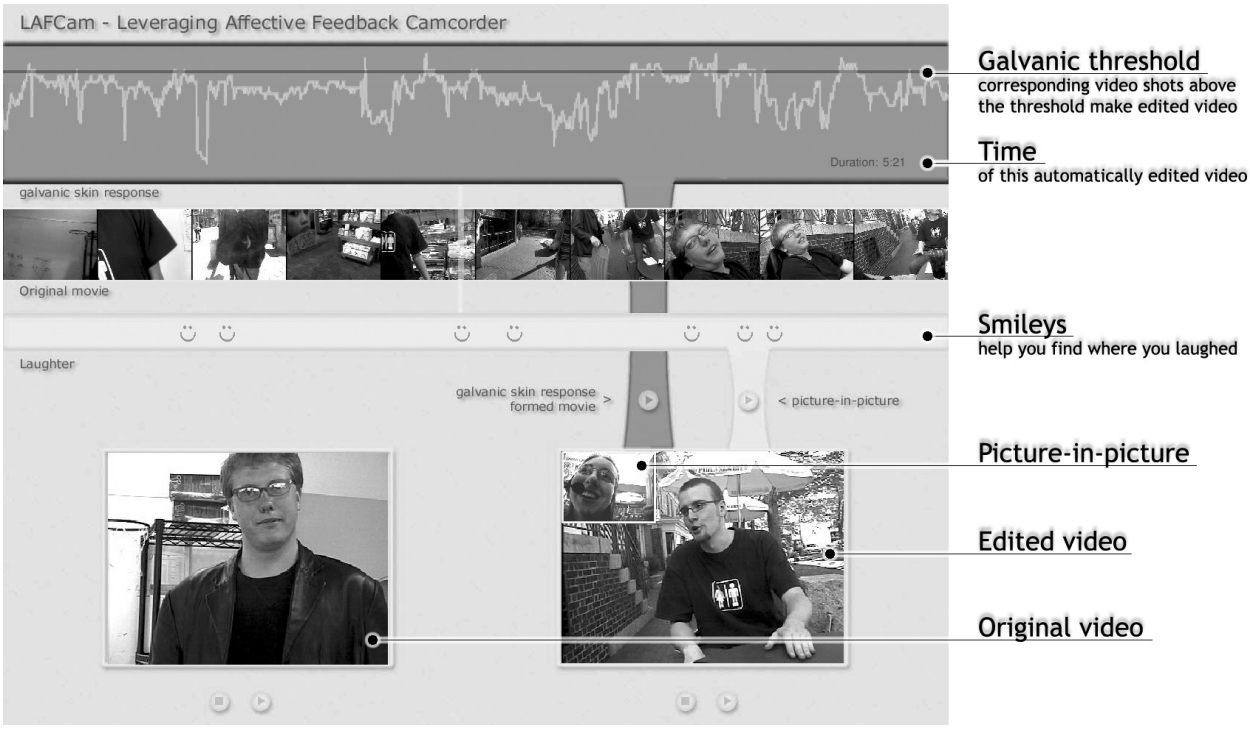

Figure 2. Video editing with LAFCam

remaining 30%. The models identified laughter and speech correctly in 88% of the test segments. This analysis used data segmented by hand, but for video editing, the audio track is segmented automatically and consecutive two-second windows are tested against the HMMs. This realistic method of segmentation and classification identified segments as laughter correctly 65% of the time; moreover, most mislabeled segments were sounds that could be considered out of the model vocabulary like coughs, cars and trains.

Skin Conductivity

The camera operator wears a glove that senses skin conductivity [4], giving LAFCam an indication of arousal.

Facial Expression

In order to provide an additional perspective, we use the fact that the camera's LCD is pointing towards the operator, at approximately a right angle and constant distance. Knowing this, we mounted a second camera on the LCD, assuring a constant stable capturing of the camera operator's face. Having the ability to watch the camera operator's face is also supportive in the editing process. In future work, we could analyze the facial expressions as an additional input.

APPLICATION - VIDEO EDITING SOFTWARE

Figure 2 shows the prototype editing application. The video is displayed using a storyboard metaphor typical in video editing software. The user can browse through the video using a slider along a filmstrip representation. The laughter detection output, displayed directly under the filmstrip, visually indicates the points of laughter as smiley face icons. The skin conductivity output, displayed directly above the filmstrip, correlates the camera operators skin response with the video. The "face camera" perspective is displayed either separately or as a "picture-in-picture".

An automatically edited version of the video can be created from the skin conductivity data. The assumption is that since it is a measure of arousal, higher conductivity indicates greater interest. A threshold bar on the skin conductivity graph allows the user to select a minimum value. The video is then edited into a movie containing all sequences in which the skin conductivity was above the threshold value, allowing the user to create automatically edited movies of variable length with little effort. This alternate version is displayed in a different video window.

CONCLUSION

Through detecting laughter, recording skin conductivity, and filming the camera operator's face, LAFCam is a system that provides an enhanced user interface augmenting the experience of editing home videos.

REFERENCES

1. http://www.wviff.org/Spring_Film_Contest_Tips_and_Help.htm
2. Healey, J., Picard, R. StartleCam: A cybernetic wearable camera. *Proc. Intl. Symposium on Wearable Computers* (Pittsburgh, PA, 1998)
3. Rabiner, L. A tutorial on Hidden Markov Models and selected applications in speech recognition, IEEE 77 (2): 257-286, 1989
4. Picard, R., Scheirer, J. The Galvactivator: A glove that senses and communicates skin conductivity. *Proc. 9th Intl. Conf. on HCI* (New Orleans, 2001)