A Tablet Annotation Tool for Endoscopic Videos

Marco A. Hudelist
Institute for Information Technology
Klagenfurt University
Klagenfurt, Austria
marco@itec.aau.at

Sabrina Kletz
Institute for Information Technology
Klagenfurt University
Klagenfurt, Austria
sabrina@itec.aau.at

Klaus Schoeffmann
Institute for Information Technology
Klagenfurt University
Klagenfurt, Austria
ks@itec.aau.at

ABSTRACT

We present a tool for mobile browsing and annotation tailored for endoscopic videos. Professional users can utilize this tablet app for patients debriefings or educational purposes. It supports text input, free-hand and shape drawing as well as angle measurements, e.g. for comparing instrument orientation. It is possible to annotate single frames as well as user-defined video sections. Moreover, it provides easy and efficient navigation via a zoom-able navigation bar that is based on frame stripes. Frame stripes are created by extracting a single, one pixel wide vertical stripe from every keyframe. The stripes are then arranged next to each other to form a uniform bar. This gives users great overview of the content of a given video. Furthermore, the app supports creation of custom reports based on the entered annotations that can be directly mailed or printed for further usage.

CCS Concepts

•**Information systems** → **Video search;** *Retrieval on mobile devices;* •**Human-centered computing** → **Mobile computing;** *Mobile devices;*

Keywords

Video Browsing; Mobile; Endoscopic Videos; User Interface

1. INTRODUCTION & RELATED WORK

Mobile devices like tablets are an integral part of the day to day work routine in the health care system. Doctors and other medical staff are using them to easily plan, review and communicate medical data between each other, the management as well as with the patients. For example, in case of endoscopic videos doctors are in need of an easy-to-use, mobile application that can be used in surgery debriefings and for educational purposes.

Intuitive and effective inspection of endoscopic videos is challenging, as traditional mobile video players are not optimized for this specific task. They can not support users in navigating such videos fast and effortless. Moreover, if surgeons would like to quickly annotate certain areas in a video, e.g. to highlight specific elements for patients or when training other surgeons, they may even have to copy the frame and switch to a different application.

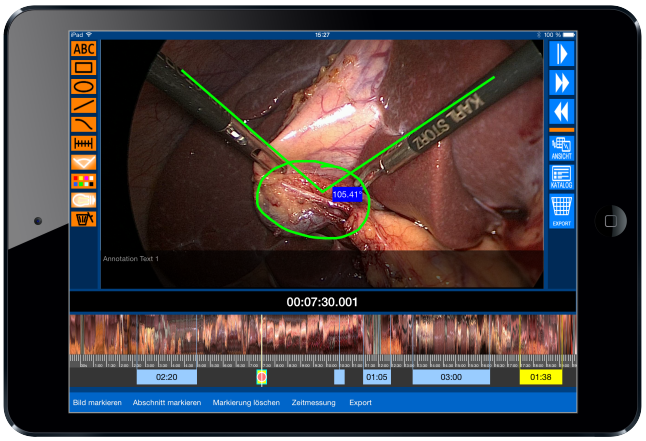

Figure 1: Screenshot of the interface.

In this paper we present an interface tailored for the special needs of surgeons and endoscopists. Users are able to annotate single frames as well as user-defined segments in a given endoscopic recording. They have options to add textual descriptions, shapes, free-hand drawings and angle measurements. Moreover, it provides capabilities to generate reports based on all or only selected annotations that can be shared via e-mail or directly printed. To make navigation easier the tablet app provides an enhanced navigation bar that utilizes *frame stripe visualization* [15, 12] to let users directly identify interesting sections of the video. The navigation bar is also zoom-able to give users better control on the level of detail that is displayed.

Although mobile video browsing tailored for endoscopic purposes is in its early stages there are of course other comparable approaches for mobile video browsing in case of other, more general content. With ProPane Ganhör [3] shows a mobile video navigation tool for fast and very precise browsing. Huber et al. [4] show Wipe'n'Watch, an interface for browsing interrelated video collections. It is similar to the approach of De Rooij et al. [2] for desktop computers. With *HiStory*, Hürst and Darzentas [8] show a tablet video browser that uses a hierarchical storyboard. Moreover, Hürst et al. also show various approaches for

MM '16 October 16–19, 2016, Amsterdam, Netherlands

DOI: 10.1145/1235

timeline-based or one-handed interaction on mobile devices [9, 10]. Hudelist et al. [5] present a browser that uses a metaphor of a 3D filmstrip for browsing videos on a tablet device. Also, they show a thumb-optimized interface [6] and a combination of multiple navigation bars of keyframes with different granularity levels based on sub-shot segmentation [7]. In addition, Schoeffmann et al. present browsing via a *scrubbing wheel* [13] and by using wipe gestures for controlling playback speed [14]. Furthermore, Karrer et al. [11] propose using direct manipulation of objects in a scene for navigation instead of traditional seeker bars.

2. INTERFACE

For navigation in a video users can either use the traditional play/pause, fast forward and fast rewind buttons on the right hand side or interact with the frame stripe navigation bar at the lower part of the interface by simple tap gestures. The navigation bar shows on one hand a visualization of the current videos' content as well as annotation markings below the frame stripes. The content is visualized by uniformly sampling frames from the video and extracting a single, one pixel wide stripe from the center of the image. This *frame stripe* visualization is inspired by the works of Schoeffmann et al. [15] - also known as MO-images as shown by Mueller-Seelich and Tan [12]. Of all keyframes an one pixel wide vertical strip is extracted. The extracted stripes are arranged consecutively in a vertical manner. This visualization enables users to quickly identify content changes in the video, e.g., the beginning of a new phase of a surgery. It is also very easy to identify scenes with a unique color pattern, e.g., out-of-patient scenes where the camera records shots from the operation room. In contrast to typical scenes with a color palette ranging from orange to red, such scenes typically generate green or gray patterns, as can be seen Fig. 2.

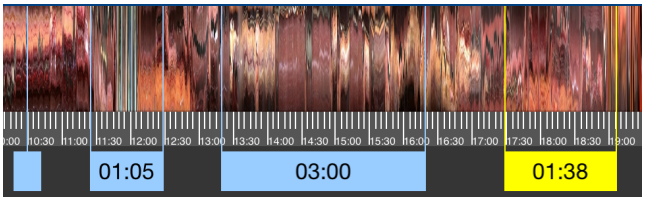

Figure 2: Zoomed screenshot of a portion of the navigation bar.

As the stripe can become quite dense when used with longer videos it is also possible to use pinch gestures to zoom in and out. In such a case the stripe becomes Below the stripe, ruler-like markings give an indication about time. Furthermore, below the time marks an area is reserved for visualizing single as well as range annotations. When selected via a single tap the player re-positions accordingly and loads drawings, measurements and text annotations.

New annotations can be added by first navigating to the appropriate position and then either using dedicated buttons to annotate the frame or a range at the lower part of the interface. A frame or a range can be annotated textually as well as with drawings and measurements. For instance, users can add shapes like circles and rectangles, freehand drawings but also measurements of angles. Such drawings are displayed directly in the player window as an overlay, while textual annotations are displayed at the bottom of the player window in a semi-transparent box (see Fig. 1).

It is also possible to measure time between two events in the video. For this, users have to activate the appropriate button and adjust a displayed range marking on the navigation stripe accordingly.

Moreover, the app supports exporting frame and range annotations for generating a report with screenshots. The report can be directly printed by using an AirPrint-ready[1] printer. To add an annotation to the report users need to select the appropriate annotation in the navigation bar and select the export button at the bottom of the interface. It is also possible to add any current frame of the player window by using the same method. A preview of the report and options for printing are accessible via a dedicated export button on the right hand side of the interface.

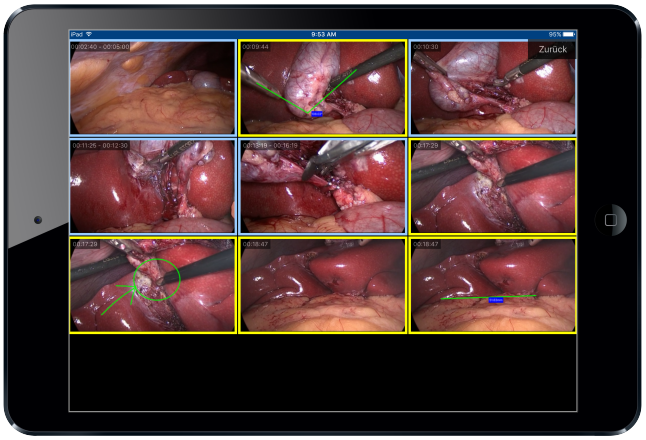

Figure 3: Grid visualization of all added annotations by their respective keyframes.

To quickly access annotations it is further possible to change the interface to a grid visualization of all already added annotations (see Fig. 3). For this, another dedicated button is available on the right side of the screen.

3. CONCLUSIONS

In this paper a tablet tool for browsing and annotation of endoscopic videos has been presented. It supports surgeons and other medical staff to easily debrief patients after procedures as well as it can be used for educational purposes. Textual annotations are supported as well as free-hand and shape drawings. Moreover, users are able to quickly navigate in the video by using a zoom-able navigation bar that utilizes frame stripe visualization. It is able to provide visual access to a videos' content in a very compact way.

4. ACKNOWLEDGMENTS

This work was supported by Universität Klagenfurt and Lakeside Labs GmbH, Klagenfurt, Austria and funding from the European Regional Development Fund and the Carinthian Economic Promotion Fund (KWF) under grant KWF 20214 u. 3520/ 26336/38165.

[1]Apple AirPrint [1] is a technology that enables iOS devices to directly send data to supported printers via WiFi connections.

5. REFERENCES

[1] About airprint - apple support. https://support.apple.com/en-jo/HT201311. Accessed: 2016-05-22.

[2] O. de Rooij, C. G. M. Snoek, and M. Worring. Mediamill: Semantic video search using the rotorbrowser. In *Proc. of the 6th ACM Int. Conf. on Image and Video Retrieval*, CIVR '07, pages 649–649, New York, NY, USA, 2007. ACM.

[3] R. Ganhör. Propane: Fast and precise video browsing on mobile phones. In *Proc. of the 11th International Conference on Mobile and Ubiquitous Multimedia*, MUM '12, pages 20:1–20:8, New York, NY, USA, 2012. ACM.

[4] J. Huber, J. Steimle, R. Lissermann, S. Olberding, and M. Mühlhäuser. Wipe'n'watch: Spatial interaction techniques for interrelated video collections on mobile devices. In *Proc. of the 24th BCS Interaction Specialist Group Conference*, BCS '10, pages 423–427, Swinton, UK, UK, 2010. British Computer Society.

[5] M. A. Hudelist, K. Schoeffmann, and L. Boeszoermenyi. Mobile video browsing with a 3d filmstrip. In *Proc. of the 3rd ACM Conference on Int. Conf. on Multimedia Retrieval*, ICMR '13, pages 299–300, New York, NY, USA, 2013. ACM.

[6] M. A. Hudelist, K. Schoeffmann, and L. Boeszoermenyi. Mobile video browsing with the thumbbrowser. In *Proceedings of the 21st ACM International Conference on Multimedia*, MM '13, pages 405–406, New York, NY, USA, 2013. ACM.

[7] M. A. Hudelist, K. Schoeffmann, and Q. Xu. *Proc. of the 21st Int. Conf. on MultiMedia Modeling*, chapter Improving Interactive Known-Item Search in Video with the Keyframe Navigation Tree, pages 306–317. Springer, Cham, 2015.

[8] W. Hürst and D. Darzentas. History: A hierarchical storyboard interface design for video browsing on mobile devices. In *Proc. of the 11th Int. Conf. on Mobile and Ubiquitous Multimedia*, MUM '12, pages 17:1–17:4, New York, NY, USA, 2012. ACM.

[9] W. Hürst, G. Götz, and M. Welte. A new interface for video browsing on pdas. In *Proc. of the 9th Int. Conf. on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '07, pages 367–369, New York, NY, USA, 2007. ACM.

[10] W. Hürst and P. Merkle. One-handed mobile video browsing. In *Proc. of the 1st Int. Conf. on Designing Interactive User Experiences for TV and Video*, UXTV '08, pages 169–178, New York, NY, USA, 2008. ACM.

[11] T. Karrer, M. Wittenhagen, and J. Borchers. Pocketdragon: A direct manipulation video navigation interface for mobile devices. In *Proc. of the 11th Int. Conf. on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '09, pages 47:1–47:3, New York, NY, USA, 2009. ACM.

[12] H. Mueller-Seelich and E. Tan. Visualizing the semantic structure of film and video, 2000.

[13] K. Schoeffmann and L. Burgstaller. Scrubbing wheel: An interaction concept to improve video content navigation on devices with touchscreens. In *IEEE Int. Symposium on Multimedia (ISM)*, pages 351–356, Dec 2015.

[14] K. Schoeffmann, K. Chromik, and L. Boeszoermenyi. Video navigation on tablets with multi-touch gestures. In *IEEE Int. Conf. on Multimedia and Expo Workshops (ICMEW)*, pages 1–6, July 2014.

[15] K. Schoeffmann, M. Taschwer, and L. Boeszoermenyi. The video explorer: A tool for navigation and searching within a single video based on fast content analysis. In *Proceedings of the First Annual ACM SIGMM Conference on Multimedia Systems*, MMSys '10, pages 247–258, New York, NY, USA, 2010. ACM.

