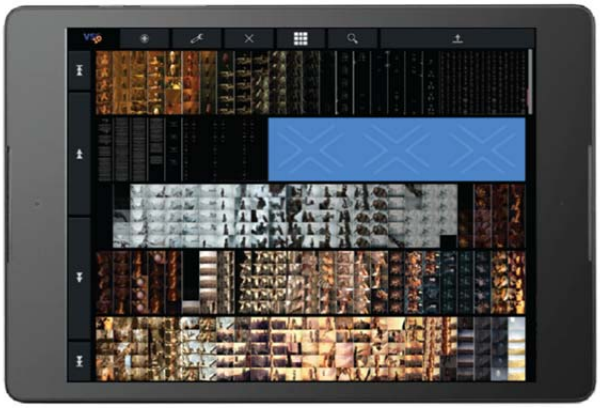

Figure 2. Tablet-based visual inspection tool.

2.2 Large-scale Visual Inspection

The traditional way of representing potentially relevant parts of a video as search results is efficient if and only if the target location is part of these visualized segments. Other situations though might require more details. Navigating individual clips using the small playback window to explore a particular video is possible, but often cumbersome and not flexible enough due to the enforced sequential navigation structure. Fig. 2 shows a screenshot of our tablet-based tool for visual inspection that is optimized for quick, efficient visual skimming of large amounts of video at detailed levels. The whole database is visualized as a storyboard representing all videos. The blue bar in Fig. 2 illustrates the end of a single video file. For each video, a key frame is extracted every second and shown as a small thumbnail. Thumbnail sizes are kept extremely small to make maximum usage of the limited screen estate, yet being large enough to ensure a good human classification performance with respect to visual perception [3]. In order to better identify scenes in such a miniaturized visualization, thumbnails are not arranged in the common line-by-line representation, but instead they are shown from left to right in columns of five thumbnails. Fig. 2 illustrates that this results in "clusters" of scenes that are easier and faster to spot. This visualization approach has proven to perform extremely well for known item search tasks (KIS). In the VBS 2015 competition, a similar interface was used with a storyboard containing all 100 hours of the database in a fixed but random order. Although it did not feature any content-based indexing or querying at all but enforced the user to visually skim the entire database, it achieved a remarkable 3rd place in the end [4].

2.3 Combined Collaborative Approach

The results from [4] have proven that an optimized visualization method performs extremely well even at database sizes of up to 100 hours. Yet, it does not scale indefinitely and there are obvious limits to how much data can be visually inspected in a given amount of time (e.g., at VBS 2015, the time limit per task was 5 minutes). Thus, we present an integrated system that combines both approaches in a way that copes with the respective disadvantages and maximizes the benefits of each individual search concept (Fig. 3). In particular, results from the CBVR tool are sent to the tablet and used to update the order of videos in the storyboard. Likewise, files that have been inspected on the tablet and classified as irrelevant are excluded from future retrieval request thus reducing the database and increasing the likelihood of finding relevant results quickly. In our research [8], we investigated different ways to resort the visualizations in the storyboard on the tablet based on the CBVR tool results as well as how the information sent via the back channel from the tablet to the PC tool can be used in optimal ways.

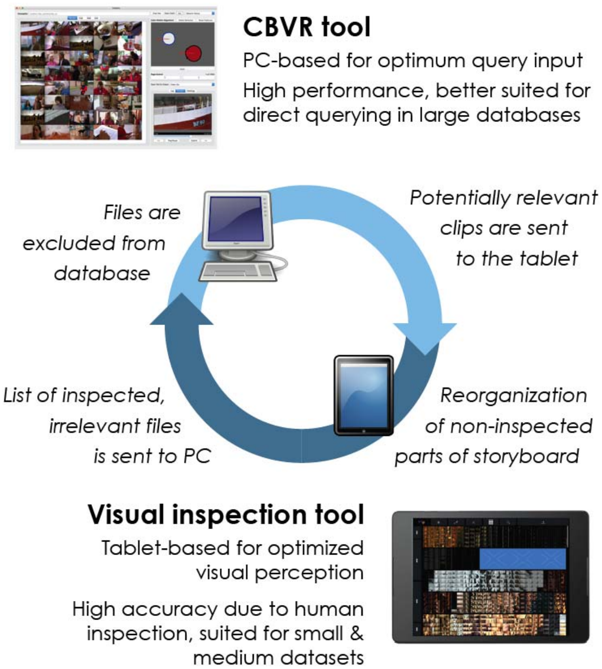

Figure 3. Collaboration between CBVR & tablet tool.

3. REFERENCES

[1] Christian Beeks, Steffen Kirchhoff, Thomas Seidl. 2013. Signature matching distance for content-based image retrieval. In *Proceedings of the 3rd ACM International conference on multimedia retrieval* (ICMR'13), 41-48.

[2] Marco A. Hudelist, Claudiu Cobârzan, Christian Beecks, Rob van de Werken, Sabrina Kletz, Wolfgang Hürst, Klaus Schoeffmann. 2016. Collaborative Video Search Combining Video Retrieval with Human-based Visual Inspection. *MultiMedia Modeling*, Vol 9517 of Lecture Notes in Computer Science, 400-405.

[3] Wolfgang Hürst, and Rob van de Werken. 2015. Human-based Video Browsing – Investigating interface design for fast video browsing. In *Proceedings of IEEE International Symposium on Multimedia* (ISM 2015), 363-368.

[4] Wolfgang Hürst, Cees G.M. Snoek, Willem-Jan Spoel, and Mate Tomin. 2010. Keep moving!: revisiting thumbnails for mobile video retrieval. In *Proceedings of the 18th ACM international conference on Multimedia* (MM '10). ACM, New York, NY, USA, 963-966.

[5] Jakub Lokoč, Adam Blažek, and Tomáš Skopal. 2014. Signature-Based Video Browser. *MultiMedia Modeling*, Vol 8326 of Lecture Notes in Computer Science, 415-418.

[6] Klaus Schoeffmann. 2014. A User-Centric Media Retrieval Competition: The Video Browser Showdown 2012-2014. *IEEE Multimedia*, Vol. 21, No. 4, Oct.-Dec., 2014, 8-13.

[7] Arnold W. M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, and Ramesh Jain. 2000. Content-Based Image Retrieval at the End of the Early Years. *IEEE Trans. Pattern Anal. Mach. Intell.* 22, 12 (December 2000), 1349-1380.