# PocketDRAGON: A Direct Manipulation Video Navigation Interface for Mobile Devices

Thorsten Karrer
RWTH Aachen University
Ahornstrasse 55
52074 Aachen
+49 (0) 241 80-21064
karrer@cs.rwth-aachen.de

Moritz Wittenhagen
RWTH Aachen University
Ahornstrasse 55
52074 Aachen
+49 (0) 241 80-21067
wittenhagen@cs.rwth-aachen.de

Jan Borchers
RWTH Aachen University
Ahornstrasse 55
52074 Aachen
+49 (0) 241 80-21050
borchers@cs.rwth-aachen.de

## ABSTRACT

We present PocketDRAGON, a demonstrator prototype that allows direct manipulation video navigation on mobile touchscreen devices. In contrast to traditional video navigation techniques, PocketDRAGON does not require any overlay UI elements that occupy valuable screen real estate and obstruct the users' view on the video. Also, direct manipulation video navigation techniques have been shown to compare favorably to the established timeline slider interfaces in terms of performance times, intuitiveness, precision, and perceived ease of use. Our demonstrator system still uses a backend server for the computationally expensive parts of the algorithms but delivers the full-fledged user experience on the mobile device.

## 1. INTRODUCTION

One of the prime requisites for mobile devices is high portability. The tradeoff for meeting this demand is that the surface area and screen real estate on the device are extremely limited. Since this resource is needed for both input and output, there is a need for interaction techniques that are applicable without requiring too much space for buttons or other input elements. This applies to traditional mobile devices with dedicated hardware input elements as well as to those that utilize touch screens with virtual input elements. One application domain that requires an especially rich visual output channel and thus an unobstructed view on the data is watching video on the mobile device.

Users want the video to occupy as much area of the screen as possible [3] leaving little or no space for control UI elements such as play and pause buttons or a timeline slider for navigating the video. Because applications cannot help with this problem on mobile devices with dedicated hardware controls we investigated several possibilities for touch screen devices. Current solutions implemented in commercially available mobile devices mostly use virtual control elements that take up valuable screen real estate. We implemented a version of our direct manipulation video navigation technique, DRAGON [4], on a mobile device to allow controlling video playback without occluding the video with overlay UI elements or reducing its visible size.


*MobileHCI'09*, September 15 - 18, 2009, Bonn, Germany.
ACM 978-1-60558-281-8.

Our demonstrator, PocketDRAGON, lets people experience the differences in look-and-feel, performance, and ease of use between a traditional timeline slider overlay, a video skimming interface, and direct manipulation video navigation on a mobile device.

## 2. RELATED WORK

Video browsing interfaces implemented on commercially available mobile devices usually use the established metaphors known from desktop computers. For example, the mobile video player for the Apple iPhone[1] provides an interface that is similar in principle to that of the QuickTime or Windows Media player (see Figure 1).

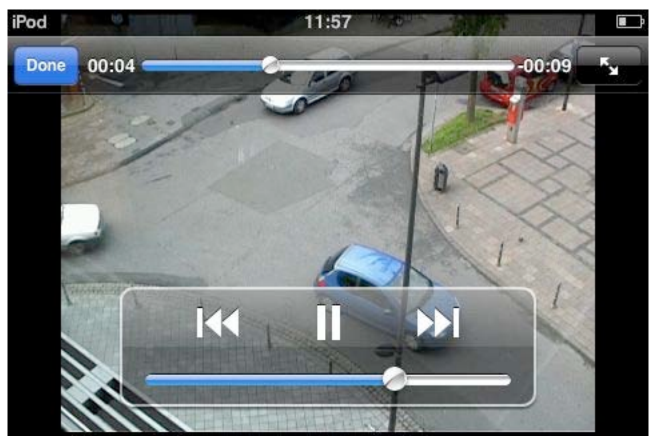


**Figure 1: Current video navigation interfaces on mobile devices require a large amount of screen real estate and obstruct the view on the video while being active.**

To make the required buttons and sliders accessible with a finger the video is occluded in large parts while the device is showing the controls. When the user is done manipulating the controls the on-screen overlays are faded out after a timeout. We argue that this technique is not only problematic for fine-grained video navigation but also obstructs the users view on the video while the controls are shown on the screen.

[1] www.apple.com/iPhone

Recently, the idea of direct manipulation video navigation systems for desktop environments has become a topic of active research [1,2,4,5]. These systems employ an interaction technique where objects in the video can directly be dragged along their movement trajectories, thereby navigating the video timeline. An example video that demonstrates the DRAGON system can be found at http://hci.rwth-aachen.de/dragon.

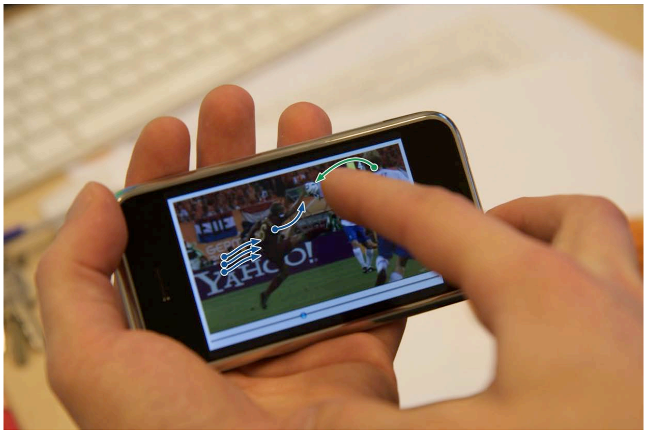


**Figure 2: PocketDRAGON in use. The user taps on the ball in the video scene and drags it to the hand of the soccer player to navigate to the exact frame where the player committed the foul.**

Several studies [1,4] have shown independently that direct manipulation video navigation improves the ability to navigate within a scene in terms of performance times and ease of use. All of these systems, however, are designed for desktop use and need substantial processing power far beyond anything available on mobile devices today.

# 3. DESIGN

We extended our existing direct manipulation video navigation system, DRAGON, to work on the iPhone or iPod Touch (see Figure 2). The interaction is mostly identical to the desktop version but does account for the reduced accuracy that results from using the finger for touch input by picking the most dominant motion trajectory in the touch area. Also, we make use of the multi-touch capabilities of the iPhone by allowing the user to navigate directly to the previous / next scenes by multi-finger swiping gestures. While any one-finger gesture is interpreted as a direct manipulation interaction, swiping the touch screen horizontally using two fingers results in a scene jump. This allows not only for quick and precise inter-scene and intra-scene navigation but also adds an additional semantic navigation layer in comparison to traditional timeline slider interfaces. To give the users of our demonstrator the possibility to directly compare PocketDRAGON to other existing video navigation techniques we also implemented two reference video players, one that is using the standard timeline-based interface provided by Apple's iPhone SDK and one that employs the video skimming technique known from Apple's iMovie and iPhoto.

# 4. IMPLEMENTATION

Bringing direct manipulation video navigation capabilities to a mobile device faces several issues. Foremost, today’s mobile devices still lack the computational power of standard desktop computers, thus making the computation of the movement trajectories for the objects in the video infeasible. Even with pre-rendered trajectory metadata, current video players are not able to provide fluid interaction since decoding video frames in full resolution takes too long for random access in the video. This, however, is required for direct manipulation video navigation.

In our prototype demonstrator we work around these problems to focus on the interaction aspects of the system. Instead of creating the objects' trajectories on the mobile device itself, we offload the expensive parts of the computation to a dedicated server on a local network. Whenever the user touches the screen, the mobile device sends a trajectory request for the indicated coordinates to the server. This approach introduces a small delay but allows us to demonstrate the technique with today's mobile devices. Also, informal user tests have shown that the influence of the added network latency on the flow of the interaction is negligible.

To overcome the problem of accessing random frames in encoded videos, we developed an adaptive image loading and caching algorithm. Video frames are pre-fetched in different resolutions to ensure quick responses to large changes of the requested frame number while presenting high quality results for small changes. At the same time the algorithm constantly monitors the available memory on the device and adjusts the caching strategy accordingly.

# 5. DEMONSTRATOR

Our demonstrator system (see Figure 3) consists of two iPod Touches and a Mac Pro server. The iPods are running PocketDRAGON and the two reference video player applications.

**Figure 3: Sketch of the PocketDRAGON demonstrator setup. The two mobile devices are running PocketDRAGON, the stationary computer is the trajectory server which can also show the desktop version of DRAGON for comparison with PocketDRAGON.**

Conference attendees can experience direct manipulation video navigation on mobile devices and directly compare it to traditional video navigation techniques. Additionally, we show the current desktop version of DRAGON on the server as a reference.

## 6. FUTURE WORK

While user tests on the desktop system [4] supported our claim that direct manipulation video navigation is superior to timeline based in-scene navigation, we still have to verify these claims for the domain of mobile devices in a formal user study.

Regarding the input precision issues (e.g., finger occlusions) that are commonly observed with mobile touch screen devices, we are looking into experimenting with advanced input techniques like [6,7] to overcome these problems.

## 7. ACKKNOWLEDGMENTS

This work was partially funded by the German B-IT Foundation and by the German Government through its UMIC Excellence Cluster for Ultra-High Speed Mobile Information and Communication and its HumTec Human Technology Center at RWTH Aachen University.

## 8. REFERENCES

[1] Dragicevic, P., Ramos, G., Bibliowitcz, J., Nowrouzezahrai, D., Balakrishnan, R., & Singh, K. (2008). Video browsing by direct manipulation. *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems* (pp. 237–246). Florence: ACM.

[2] Goldman, D. B., Gonterman, C., Curless, B., Salesin, D., & Seitz, S. M. (2008). Video object annotation, navigation, and composition. *IST '08: Proceedings of the 21st annual ACM symposium on User interface software and technology* (pp. 3–12). Monterey: ACM.

[3] Hürst, W., Götz, G., & Welte, M. (2007). Interactive video browsing on mobile devices. *MULTIMEDIA '07: Proceedings of the 15th international conference on Multimedia* (pp. 247–256). Augsburg: ACM.

[4] Karrer, T., Weiss, M., Lee, E., & Borchers, J. (2008). DRAGON: a direct manipulation interface for frame-accurate in-scene video navigation. *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems* (pp. 247–250). Florence: ACM.

[5] Kimber, D., Dunnigan, T., Girgensohn, A., Shipman, F., Turner, T., & Yang, T. (2007). Trailblazing: Video playback control by direct object manipulation. *2007 IEEE International Conference on Multimedia and Expo* (pp. 1015–1018). IEEE.

[6] Vogel, D., & Baudisch, P. (2007). Shift: a technique for operating pen-based interfaces using touch. *CHI '07: Proceeding of the twenty-fifth annual SIGCHI conference on Human factors in computing systems* (pp. 657–666). San Jose: ACM.

[7] Wigdor, D., Forlines, C, Baudisch, P., Barnwell, J. & Shen, C. (2007). Lucid touch: a see-through mobile device. *UIST '07: Proceeding of the twentyth annual ACM symposium on User interface software and technology* (pp. 269–278). Newport: ACM.