

# Interface Design of Video Scout: A Selection, Recording, and Segmentation System for TVs

John Zimmerman<sup>†</sup>, George Marmaropoulos\*, Clive van Heerden\*

Philips Research<sup>†</sup> and Philips Design\*

345 Scarborough Road, Briarcliff Manor, NY, USA, 10510

{john.zimmerman, george.marmaropoulos, clive.vanheerden}@philips.com

## 1 Abstract

Video Scout is a prototype retrieval application that allows Personal Video Recorders to actually *watch* the TV programs they record. By analyzing the visual, audio, and transcript data, Scout can segment and index TV programs, finding and recording specific video clips that match requests in users' profiles. For example: if users request information on Philips, Scout will watch news programs and capture any stories it finds on Philips. The Scout interface offers a familiar TV environment where users can interact with whole TV programs and video clips organized by topic. Scout also provides users with tools for managing their profiles. This paper captures the Video Scout interface design process, from concept sketches to user testing to final prototype design.

## 2 Background

PVRs, like TiVo® [1], have changed the way people interact with television. The user's task has changed from (a) finding something to watch from 100+ channels to (b) finding something to record from 10,000+ weekly shows. By monitoring shows users watch and by providing an interface for rating shows, PVRs construct user profiles that allow them to recommend and automatically record programs. This is the first smart TV appliance.

Video Scout represents a future direction for PVRs in that it actually watches the programs it records. Scout analyzes the visual, audio, and transcript data and integrates the results in a Bayesian engine. By searching for matches between users' profiles and the indexed clips, Scout can provide users with the clips they are looking for. For example: if users want information on an earthquake that just happened, Scout can watch the news programs and extract just the stories on this earthquake. (For a more detailed description of the Scout engine, please see [2]). This paper details the Video Scout interface design process, from concept sketches to user testing to final prototype design. It illustrates how feedback from the user test influenced the look and feel of the final prototype.

## 3 Concept Sketches

Our design process started with three concept sketches. (i) Raindrops: offers users tools for finding events within whole TV programs. (ii) Seeker: displays a playback system for topic-based video clips that are automatically recorded according to users' profiles. (iii) TV Magnets: presents an interface for specifying the kinds of content users want to attract and repel. We implemented these concepts with simple interactivity in Macromedia Director and demonstrated them to test subjects using a video projector to create the illusion of a futuristic TV.

The Raindrops concept (figure 1) presents a TV program guide where current and future shows fall like "rain". Stored programs are displayed at the bottom of the screen. Over time the stored programs slowly sink into the drain, indicating to users that they are going to be deleted. The system provides tools for finding specific segments from whole TV shows. The example below displays all of the fieldgoals from a stored football game.

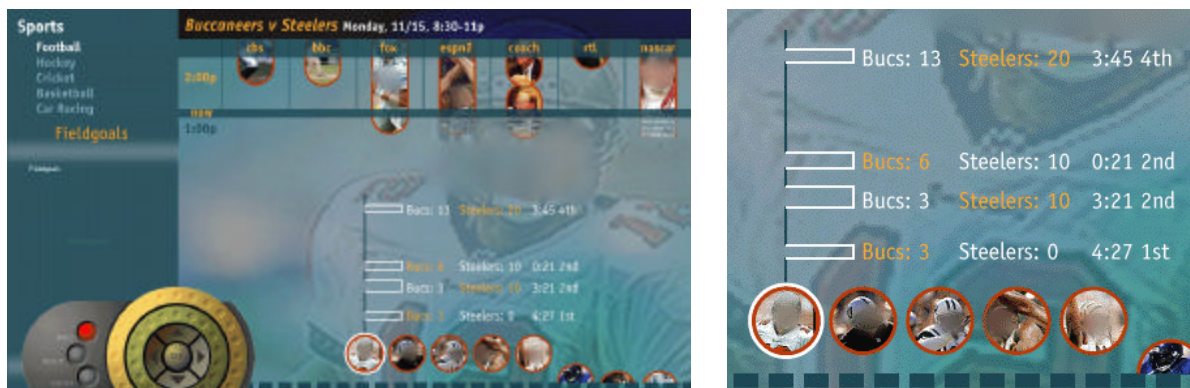


Figure 1. Raindrops sketch and detail of a stored and segmented football game.

The Seeker concept (figure 2) finds video clips on specific topics based on users' profiles. In the example below, the system employs users' TV viewing histories to infer which *celebrities* they like. It then records all video clips it finds with these celebrities. When users begin watching TV, they see all of the clips the system has collected. A graphical scroll bar located on the right of the list reveals the relative length of the individual clips. Users can build a playlist on the right from the list of all captured clips on the left.



Figure 2. Celebrity Seeker Sketch

The TV Magnets concept (figure 3) lets users choose topics to be recorded. Users create different kinds of magnets—event magnet, news magnet, people magnet, etc.—that attract and repel specific story segments. Users can individually adjust the strength of attraction or repulsion for topics. In the example below the user has specified an event magnet that attracts stories about plane crashes, earthquakes, tornados, and floods. The same magnet repels TV commercials and stories on weather.



Figure 3. TV Magnet Sketch

### 3.1 Reactions

Of the three concepts, Raindrops was best received. Subjects liked selecting TV shows by image or face as well as by text. They also liked the new feature that let them identify specific events from within a program.

In general, they were more confused by Seeker and TV Magnets. *At first, almost all subjects had trouble understanding why they would want only a piece of a TV show instead of a whole TV show.* They often expressed anxiety that if they only had a video clip, they might be missing something better. They said that both systems were cumbersome, offering too many options that got in the way of trying to watch TV.

With Seeker, subjects indicated that they did not want narrative content cut into pieces unless these pieces worked as a preview, a task beyond the semantic ability of the Scout engine. Several also stated that this technology would be great for tracking financial news.

For TV Magnets, users liked specifying the kind of information to be recorded. They especially liked the idea of repelling advertisements. Their main concern, however, was the fear of a *need to constantly manage the magnets* to get the right content.

#### 4 Prototype Design

We based the Scout interface on the look and feel of Raindrops because this design had received the best feedback, and *it allowed us to lead users into the idea of video clips* by initially presenting whole TV shows. Our main task was to take the best pieces of the other designs and incorporate them into Raindrops.

We divided the prototype into two main sections called Program Guide and TV Magnets. The Program Guide section works a lot like Raindrops, allowing users to interact with whole TV shows. The TV Magnets section offers users access to their profiles and video clips organized by topic.

##### 4.1 Grid

Building on the Raindrops design, we divided the screen into three zones. The left side, called the tool zone, lets users adjust parameters, changing the results displayed on the right side of the screen. The right side was divided into the broadcast zone (top) and the stored zone (bottom). The broadcast zone displays live and future shows. The stored zone displays recorded TV shows and video clips. Each zone is flexible and can dynamically expand, allowing users more space. We designed these zones in order to free users from “tree” menu structures they found cumbersome in many TV products available today.

##### 4.2 Program Guide

The Program Guide is the first screen users see (figure 4). Our guide differs from traditional TV guides in two ways. First, *we display live and stored content in a single interface*. Second, we changed the orientation of the time/channel grid to support the metaphor of *raining* content. This rotation also helps explain the relationship between broadcast and stored content. Live shows fall across the current time boundary—the dark line separating the broadcast and stored zones—and join other stored shows at the bottom of the screen.

We chose to display stored and broadcast content together because they represent similar values to users. When users select something to watch, they are indicating how important this item is in terms of their time.



Figure 4. Program Guide showing a stored talk show's program and commercial segments.

The tools zone (left) contains category filters. By selecting a filter, users can reduce the number of broadcast and stored shows that appear. Below the categories are the TV magnet indicators. These feedback devices inform users when the system has stored new topic-based video clips.

The broadcast zone (top right) displays all shows being broadcast that match the highlighted category. The shows appear as pill-shaped images with their heights indicating the length of the program. Shows selected to be recorded have a circle in the upper left corner. When a recorded show crosses the current time barrier, it transforms

into a circle and joins the other recorded shows at the bottom of the screen. Shows not selected to be recorded dissolve as they cross the current time barrier.

Users can expand stored shows to see an initial segmentation based on TV commercials. Dark segments indicate program content separated by TV commercials. In this example, a talkshow, summary labels indicate guest and host segments. Organization of segments matches the time axis for broadcast shows. The lowest segment labeled “David Letterman” represents the beginning of the show, and the top segment labeled “joni mitchell” represents the end of the show. For talk shows, users can select two additional segmentations. They can see all of the jokes in the program—the system listens for a laugh and searches back for the end of the previous laugh and calls this a joke—and they can see a segmentation based on musical performance. In the case of figure 4, this allows users to quickly play a video clip of Joni Mitchell singing. The initial segmentation by TV commercial works for all TV shows; however, the additional segmentations are specific to the recorded show’s genre.

### 4.3 TV Magnets

The TV Magnets section displays video clips organized by topic. Figure 5 below shows the Financial TV Magnet screen. The tool zone lists all of the topics and includes company names, stock markets, and high-level financial topics. The black circles with the white Ns indicate topics that have attracted new content. As users navigate up and down this list, stored clips matching the highlighted topic appear on the right in the stored zone. Each clip represents an individual story on the topic.

TV shows to be analyzed in the future for financial information appear in the broadcast zone. However, *not all programs that appear will be recorded*. For example: Scout does not know if a financial news broadcast will contain information on Microsoft before it is actually broadcast. Instead, Scout selects shows that are likely to contain desired content. It then records these shows and searches for stories within the recorded content. Any segment that matches a topic is stored, and the rest is deleted.

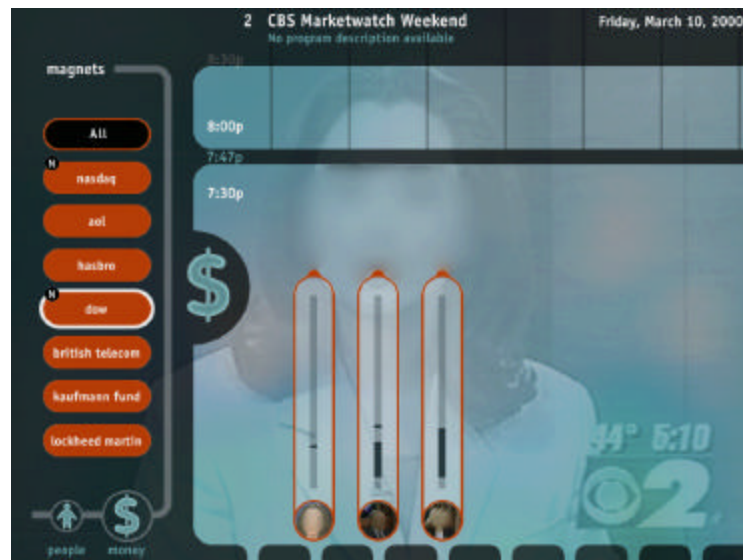


Figure 5. TV Magnet showing Financial Magnet results.

Figure 5 shows four clips that have been attracted on the topic “dow”. Faces pulled from the video stream help indicate what shows the clips come from. In this example, one clip comes from *Marketwatch* and three clips come from two different episodes of *Wall Street Week*. We chose to display the video clip in relation to the whole program for three reasons. First, this display is consistent with the segmentations users see in the program guide section. Second, the length of the clip in relation to the whole show offers a quick preview as to how much detail is covered in the clip. Third, topics that appear at the beginning or appear several times within a program have been given a higher priority by the broadcaster and this information may be important to users who select one clip over another.

TV Magnets also provide users with tools for refining the content that gets recorded. We created magnets for financial news and celebrities. For celebrities, Scout automatically records video clips on the most watched celebrities that appeared in users’ TV viewing histories. For financial news, Scout records video clips on companies and financial topics that match the users’ investment portfolios. Adding the investment portfolio and viewing history to the user profile allows Scout to begin recording right out of the box.

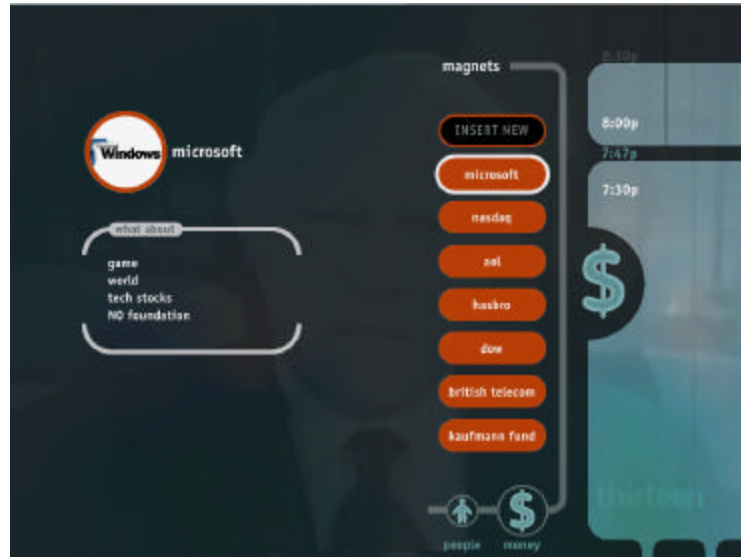


Figure 6. Modify TV Magnets

By default, Scout records all information on a listed topic. However, users can adjust what is recorded by adding keyword filters between the “what about” brackets (figure 6). In figure 6, only stories about Microsoft that mention “game”, “world”, or “tech stocks” will be recorded. Additionally, no stories that match these keywords but also mention Bill Gate’s Foundation will be recorded. The “NO” indicates the repel feature of the magnet. By playing with the keywords, users can adjust the system over time, until it records exactly what they want.

#### 4.4 Implementation

We built the Scout user interface using Macromedia Director, and we used Tabuleiro’s DirectMedia Xtra for controlling MPEG files. We recorded approximately 30 hours of programming and analyzed it using our video analysis engine [2]. As users navigate the prototype, the Director executable reads in text files output by this engine and dynamically lays out screens.

In order to create a more TV-like experience, we used a Philips Pronto remote control for user input. We developed the prototype to respond to keyboard commands and trained the Pronto using a wireless keyboard.

#### 5 Conclusion

We created the Video Scout prototype using a traditional user-centered design process of concept sketches, user testing, and design. Our main challenge was to create a new TV interface that comfortably allows users to interact with both whole TV programs and with video clips extracted from TV programs. Scout represents a future direction for PVRs because it gives the user much more control of the content that gets recorded. User profiles are automatically generated, but can be also be modified by the users themselves. In this way, Scout can quickly be trained to retrieve only the content users really want. And as the amount of video content people have to choose from increases with digital cable, digital satellite, and downloadable video from the Internet, the need for Scout to control information overload becomes clearer.

In the future we hope to use this prototype with users to better assess what applications they would like this technology to perform in their homes.

#### 6 References

- [1] <http://www.tivo.com>
- [2] Dimitrova, N., McGee, T., Agnihotri, L., Dagtas, S., & Jasinschi, R. (2000). On Selective Video Content Analysis and Filtering, *Proceedings of SPIE Storage and Retrieval for Media Databases, 2000, San Jose, CA, USA*, 359-368.
- [3] Jasinschi, R., Dimitrova, N., McGee, T., Agnihotri, L., Zimmerman, J., & Li, D. (2001). Integrated Multimedia Processing for Topic Segmentation and Classification, *Proceedings of International Conference on Image Processing, 2001, Thessaloniki, Greece*.