Access via Features versus Access via Transcripts: User Performance and Satisfaction[1]

Barbara M. Wildemuth, Meng Yang, Anthony Hughes, Rich Gruss,
Gary Geisler, and Gary Marchionini
Open Video Project, School of Information and Library Science
University of North Carolina at Chapel Hill

TREC VID 2003 Notebook Paper
SILS Technical Report 2003-05

Abstract

The Open Video Project is specifically concerned with the surrogates that can represent the objects in a digital video collection and the mechanisms through which people can manipulate those surrogates. In TREC VID 2003, we compared the effectiveness of a transcript-only search system, a features-only search system and a search system combining transcript and feature searching. We also presented several different views for users to browse the results pages: a horizontal view, a vertical view, a "before & after" view, and an extra-keyframe view. A within-subjects research design was used, so that each of the 36 participants was exposed to all three search systems. Each participant searched half (12) of the assigned topics. The user satisfaction measures recommended by NIST were augmented by measurements of participants' perceived usefulness, perceived ease of use, and flow. Results indicated that, with the transcript-only system and the combined system, users were able to achieve higher recall in less time per search. The results from the measures of satisfaction indicate that the users found the transcript-only and combined systems to be more useful and easier to use, and their use resulted in stronger perceptions of enjoyment and concentration than the features-only system. It is concluded that, as users gain experience with features searching, it will be a welcome supplement to transcript searching.

1 Introduction

The overall purpose of the Open Video Project is to investigate people's interactions with digital video collections. Specifically, we are concerned with the surrogates that can represent the objects in the collection and the mechanisms through which people can manipulate those surrogates. In addition, we are concerned with methods for studying people's interactions with video collections and ways to measure their performance and satisfaction with those interactions.

Through our participation in TREC Video Retrieval Evaluation 2003, we compared the effectiveness of a transcript-only search system, a features-only search system and a search system combining transcript and feature searching. We also presented several different views for users to browse the results pages: a horizontal view (a storyboard with brief annotations), a vertical view (a keyframe and full text of the transcript for each shot retrieved), a "before & after" view (keyframe and transcript from the shots immediately preceding and following a selected shot) and an extra-keyframe view (the extra poster frames available, shown on mouseover). Both user performance with these three systems and their satisfaction with the systems were investigated.

---

[1] This work is supported by National Science Foundation (NSF) Grant IIS 0099638.

2 Background

The design of the three search systems compared in this study is based on two types of evidence: the theoretical literature related to people's understanding of video and results from empirical studies of video searching systems. Each of these two types of evidence is briefly reviewed here.

Past theoretical and empirical work related to video/image retrieval suggests that people interact with images/videos at three levels [5, 9]. At the most basic level, primitive features of the image (e.g., color, shape) are perceived. At a second level, logical features (e.g., people, things, places, actions) are perceived. At this level, people draw on their existing knowledge to identify the objects perceived. The third level requires inductive interpretation of the image/video, with inferences being made about its abstract attributes, including emotional cues and atmosphere [9]. This three-level hierarchy is remarkably similar to Panofsky's earlier [13, 14] description of three levels of comprehension for visual images: pre-iconographical, iconographical, and iconological. In addition, it is consistent with Grodal's [10] four-stage model of the process by which people understand film: basic perception, followed by memory-matching, cognitive-emotional appraisal, and reactions at a high level of arousal. It seems likely that a retrieval system needs to "work" at multiple levels of the user's understanding in order to be completely successful.

A retrieval system that addresses multiple levels of understanding might incorporate both transcript searching and features searching. Transcript searching may be related to the more abstract higher-level aspects of video understanding, with words representing particular objects in the video (e.g., Mount Rushmore) or classes of objects in the video (e.g., mountains). Depending on the source of the text, it may also represent emotional aspects of a video (e.g., the words "I love you" appearing in a transcript). Features searching may be used to represent the more basic perceptual aspects of the video, such as bright colors or the camera zooming in. However, as features detection becomes more sophisticated and more accurate, features searching may also be able to access more abstract aspects of the video, overlapping with some aspects of text searching (e.g., identifying mountains or streets).

The TREC VID 2002 studies conducted at Dublin City University [2, 6] made the first comparison of transcript-only searching with a system combining transcript and features searching. Six people searched each system, for the 25 topics from TREC 2002. Study participants were not required to use either features or transcripts in the combined system; features were incorporated in 75% of the searches on that system and transcripts were used in 89% of the searches. There were no differences in performance (based on precision at different levels of recall), but there was more variability in user performance with the transcript-only system. No formal measures of user perceptions were gathered.

The current study extends the 2002 work from Dublin City University in several ways. A third system, that allowed only features searching, was added to the study design. In addition, with the combined system, participants were required to enter at least one term and at least one feature in each search. Finally, formal measures of user perceptions of the three systems were added to the research protocol. Our goal was to evaluate user performance with and perceptions of three systems: one supporting transcript searching, one supporting features searching, and one supporting a combination of transcript and features searching.

3 The Search Systems

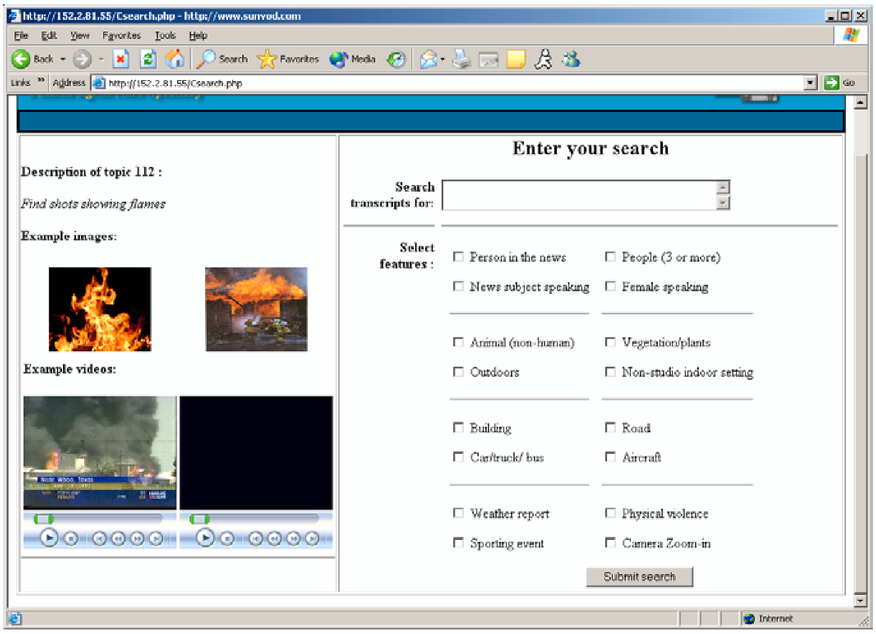
In our system designs, we focused on two components: the ways in which people could enter a search and the ways in which people could view the search results. The ways in which people enter their searches (and the data to which those searches are addressed) distinguish the three systems being compared. Each of the three systems made available four different displays of the search results, with easy navigation between them. These two components of the system design and the basic design of the search system itself will be described in this section.

3.1 Entering a search

The transcript-only system allowed users to search the ASR transcripts of the video collection [7], via a text box for search entry. Participants were instructed that they could enter any number of terms, and that phrase searching and searching of short words (i.e., less than three characters) was not supported. The MySQL full text search engine was used for this study; their default list of stopwords was accepted and the research team set the minimum word length at four characters. In computing a relevance score, MySQL takes into account the number of words in a record, the number of unique words in that record, the total number of words in the collection, and the number of records that contain a particular word. The search results were ranked based on the relevance score computed by MySQL.

The features-only system allowed users to search the features provided from ten groups' results of the TREC VID 2003 features extraction task. The groups' results were aggregated by generating a "features score" on each feature for each shot; the score was the proportion of the runs that identified that feature in a particular shot. The 17 features were represented to users as semantically-related groups of items with checkboxes. The meanings of the features were provided in a training handout, and users were allowed to check as many features as they liked. The results from this system were ranked based on the average feature score for each shot, across all features included in the search (i.e., all the features included in the search were weighted equally in aggregating them).

The third system provided both transcript and features searching (see Fig. 1), and required that users enter at least one term and check at least one feature. They received the instructions combined from the other two systems. The text box for transcript searching appeared at the top of the screen, followed by the checkboxes for the features. The results from the system included shots that matched at least one feature or at least one text word searched. They were ranked by taking into account both transcript and feature scores,


http://152.2.81.55/Csearch.php - http://www.sunvod.com
File Edit View Favorites Tools Help
Back Search Favorites Media
Links Address http://152.2.81.55/Csearch.php Go
Internet

Description of topic 112 :

Find shots showing flames

Example images:

Example videos:

Enter your search

Search transcripts for:

Select features :

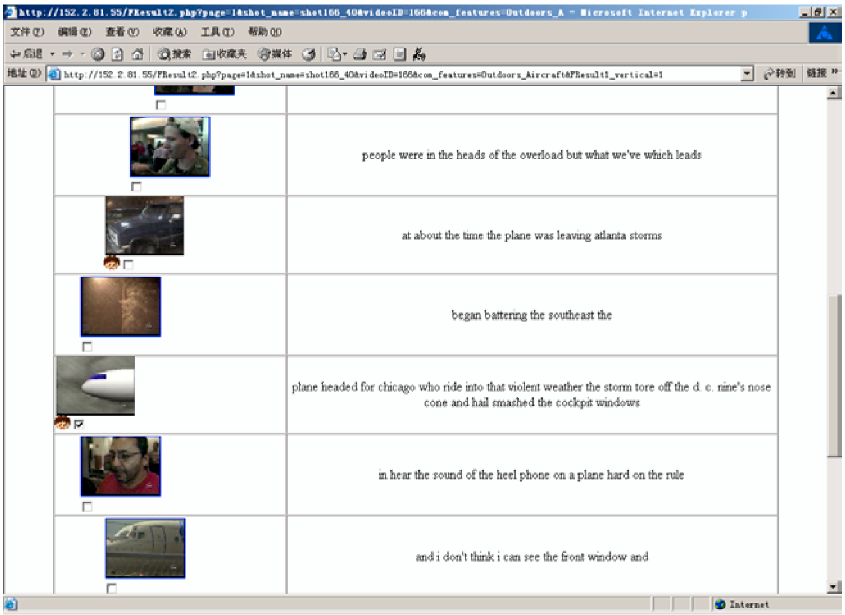
☐ Person in the news
☐ People (3 or more)
☐ News subject speaking
☐ Female speaking
☐ Animal (non-human)
☐ Vegetation/plants
☐ Outdoors
☐ Non-studio indoor setting
☐ Building
☐ Road
☐ Car/truck/ bus
☐ Aircraft
☐ Weather report
☐ Physical violence
☐ Sporting event
☐ Camera Zoom-in

Figure 1. Search entry screen for combined system

weighting the transcript-based scores twice as heavily as the aggregated feature scores. This weighting scheme was used because Dublin City's 2002 TREC VID results [6] indicated that the transcript searches were more consistent and because text words tend to be more specific than the features available for searching.

The results were displayed, by default, in a horizontal view. This view looks like an annotated storyboard, and includes a keyframe from each shot plus a few words from the transcript, selected in a window surrounding the search terms. The user could easily click on a link to the vertical view, which displayed the keyframes in a column on the left. At the right of the keyframe was the entire transcript associated with that shot.

By clicking on the keyframe in either of these basic views, the user could go to a before-and-after view. This view was provided because so many of the relevant shots appear just before or after a shot retrieved with a transcript search [16]. For example, a newscaster may say something about President Bush, but the pictures of the President appear in the *next* shot rather than the shot synchronized with the transcript. Six shots preceding and six shots following the selected shot are represented in this view by their keyframes and full transcripts (see Figure 2). The keyframe of the selected shot is aligned at the left side of the column, with the before and after shots indented slightly.


http://152.2.81.55/FResult2.php?page=1&shot_name=shot166_40&videoID=166&con_features=Outdoors_A - Microsoft Internet Explorer p
文件(F) 编辑(E) 查看(V) 收藏(A) 工具(T) 帮助(H)
后退 搜索 收藏夹 媒体
地址(D) http://152.2.81.55/FResult2.php?page=1&shot_name=shot166_40&videoID=166&con_features=Outdoors_Aircraft#FResult1_vertical=1
转到 链接
Internet

☐
☐ people were in the heads of the overload but what we've which leads
☐ at about the time the plane was leaving atlanta storms
☐ began battering the southeast the
☑ plane headed for chicago who ride into that violent weather the storm tore off the d. c. nine's nose cone and hail smashed the cockpit windows
☐ in hear the sound of the heel phone on a plane hard on the rule
☐ and i don't think i can see the front window and

Figure 2. The before-and-after view, with the selected shot aligned at the left

By mousing over any of the keyframes in the basic views, the user could see additional keyframes for that shot, if there were any available. A mark on the screen indicated when such additional keyframes were available.

4 Study Methods

Thirty-six study participants were recruited from among students, faculty and staff at UNC, by posting flyers in several buildings on campus, as well as email announcements within the School of Information and Library Science. The participants were scheduled for individual or small group sessions, with a research assistant monitoring each session.

A within-subjects research design was used, so that each of the 36 participants was exposed to all three search systems, as described above. After giving informed consent, each person completed the following activities:

- a pre-session demographic questionnaire, including both questions suggested for TREC studies and questions used in previous Open Video studies,
- 5 search tasks on system A (including a training task), each followed by a post-search questionnaire using questions suggested by TREC VID,
- a questionnaire about system A, using questions suggested by TRECVID and measures of perceived usefulness, perceived ease of use, and flow,
- 5 search tasks on system B (including a training task), each followed by the post-search questionnaire,
- a post-system questionnaire about system B,
- 5 search tasks on system C (including a training task), each followed by the post-search questionnaire,
- a post-system questionnaire about system C, and
- a brief post-session questionnaire, using questions suggested for TREC VID.

Each participant searched half (12) of the assigned topics, i.e., four topics with each search system. The order of the system and the topics were counter-balanced among the 36 subjects.

User performance on the three systems will be compared by calculating the average precision and the average recall achieved with each system, across all topics and all users. Relevance assessments provided by NIST (augmented with assessments made by the Open Video team, using the same procedures) will be compared with the shots selected by the study participants. Calculation of recall [15] was based on the assumption that the full set of relevant items in the collection is represented by the relevant items identified by NIST assessors plus any additional relevant items identified by Open Video assessors. While this approach to performance measurement is somewhat unusual within the context of TREC and other traditional information retrieval experiments, we believe that it is more able to take into account the variability in searcher performance, which can be as great as an order of magnitude [1]. In addition to the overall comparisons, analyses of the differences by topic will be conducted.

The user satisfaction measures recommended by NIST were augmented by measurements of participants' perceived usefulness, perceived ease of use, and flow (enjoyment and concentration), measured after use of each of the three search systems. Each of these measures, and evidence of its reliability and/or validity, is described briefly here:

> *Perceived usefulness and perceived ease of use*: Each of these measures, as adapted for this study, is a six-item questionnaire, using a five-point Likert scale to gather responses. They were originally developed by Davis [4] for use in studying the adoption of personal computer applications, to measure two constructs in the Technology Acceptance Model. They have been applied in hundreds of studies of technology adoption. The reliability of the usefulness measure is consistently high (e.g., alpha = 0.98 [4]; alpha = 0.88 [11]). Usefulness is a strong predictor of a person's adoption of a technology, over a wide range of technologies including Web applications [11, 12, 17]. The reliability of the ease of use measure is not quite as high, but still quite acceptable (alpha = 0.94 [4]; alpha = 0.84 [11]). While not as strong a predictor of adoption, ease of use is consistently found to be a statistically significant construct within the Technology Acceptance Model.
>
> *Flow (enjoyment and concentration):* Two aspects of flow were evaluated in this study, each with four seven-point semantic differential scales, developed by Ghani, Supnick, and Rooney [8]. These scales drew on Czekszentmihalyi's [3] ideas about flow and some of Webster's [18] work on microcomputer playfulness. Ghani et al. [8] reported reliability (alpha) of 0.88 and 0.82 for enjoyment and concentration, respectively.

Each of the three systems was evaluated, using both performance and satisfaction data. Data analysis methods included calculation of descriptive statistics, and comparison of systems with analysis of variance, using Bonferroni t Tests for post hoc analysis.

5 Results

5.1 Characteristics of the participants

Twenty-seven women and nine men participated in the study. Their mean age was 22.6 years (s.d. = 3.3 years). Twenty-two of them were undergraduate students, and the other 14 were graduate students. Twelve were students in the School of Information and Library Science, 12 were in communication studies, and the other twelve were from other disciplines.

On average, the participants had 6.5 years of online searching experience (s.d. = 1.6 years). All of them user computers daily, and 29 of the 36 conduct online searches daily. They considered themselves very experienced with point-and-click interfaces, online library catalogs, and Web search, but less experienced with searching CD ROMs and commercial online systems (see Table 1).

Table 1. User experience with searching (1, no experience, to 5, a great deal of experience)

	Mean	s.d.
Point-and-click interface	4.8	0.5
Searching an online catalog	4.0	0.8
Searching on CD ROM	3.1	1.1
Searching on commercial online systems	2.7	1.1
Searching on the web	4.8	0.5

The participants are also moderately heavy users of video. Twenty-three (64%) of them watch videos daily or week. However, they search for videos less frequently; 22 (61%) of them search for videos only occasionally or never and only five search for videos daily. Those who search for videos primarily search online or in newspapers and magazines. They most often search by title, but also search by author, actor, or topic.

5.2 Performance results

Performance on the three systems was compared in terms of the average precision the study participants achieved with each system, the average recall the study participants achieved with each system, and the average amount of time used per search on each system. These results are summarized in Table 2.

Table 2. Summary of performance, by system

	Precision		Recall		Time per search	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
Transcript-only	0.81	0.31	0.11	0.16	4.0	3.8
Features-only	0.77	0.31	0.04	0.09	5.8	4.2
Combined	0.82	0.27	0.09	0.12	5.2	4.2

Note: 144 searches contributed to each of the means, except for precision, where precision could not be calculated if no items were selected. For the transcript-only system, precision was calculated for 127 searches; for the features-only system, 84 searches; and for the combined system, 128 searches.

There was not a statistically significant difference between the systems, in terms of the precision achieved by study participants (F (2, 336) = 0.70, p=0.4968). The differences in recall were statistically significant (F (2, 429) = 11.89, p<0.0001), and post hoc Bonferroni t tests indicated that the recall achieved with the features-only system was lower than that achieved on the other two systems. It should also be noted that the variability in the recall measure was extreme. This might be attributed either to individual differences between participants or differences in the difficulty levels of the topics to be searched. These possibilities will be explored further.

The amount of time spent per search was also analyzed and compared across systems (see Table 2). There was significant variability across participants, as indicated by the large standard deviations in Table 2. The effects of topic difficulty on this variability were explored, but found to be not statistically significant. Thus, it can be attributed to individual differences between participants, an effect that was found to be statistically significant. Much of this variability can be attributed to the difference between participant 1, who averaged 11.4 minutes per search across all systems, and those who were much faster, averaging less than three minutes per search (five participants). In spite of this variability, the difference between systems in the length of time spent per query was statistically significant (F (2, 428) = 6.93, p = 0.0011), and post hoc analysis indicated that the study participants spent less time per search on the transcript-only system than on the other two systems.

A subset of search results (nine runs) was submitted to NIST to be evaluated in terms of mean average precision. Each of the nine runs represented one of the three systems (three runs for each system) and included the searches that had the highest number of selected shots for each topic. The results of those analyses, as reported by NIST, were then averaged by system, and are shown in Table 3.

Table 3. Results reported by NIST, aggregated by system

	Average precision	Hits at depth 10	Hits at depth 30	Hits at depth 100	Hits at depth 1000
Transcript-only	0.14	4.5	5.1	5.3	5.3
Features-only	0.06	2.9	3.1	3.5	3.5
Combined	0.12	4.3	4.7	5.0	5.0

These results mirror the results obtained through analysis of the entire set of searches conducted by the 36 participants. The transcript-only and combined systems outperformed the features-only system.

5.3 Satisfaction results

In addition to the measures suggested by NIST, measures of user perceptions of usefulness, ease of use, and flow (enjoyment and concentration) were taken in relation to each of the three search systems. Each set of user perceptions will be described, in turn.

The questionnaire recommended by NIST was used to collect data on user perceptions immediately after each search. In addition, one question was added to address users' perceptions of their ability to search by video features. The results from this questionnaire, aggregated by system, are shown in Table 4.

Table 4. User perceptions, based on post-search questionnaire (1, not at all, to 5, extremely)

	Transcript-only		Features-only		Combined	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
Are you familiar with this topic?	3.5	1.1	3.3	1.2	3.5	1.2
Was it easy to get started on this search?	3.9	1.1	3.0	1.2	3.9	1.2
Was it easy to do the search on this topic?	3.7	1.3	2.4	1.4	3.5	1.4
Was the ability to search by particular features of the video useful?			2.3	1.3	3.1	1.4
Are you satisfied with your search results?	3.4	1.4	2.4	1.5	3.3	1.4
Did you have enough time to do an effective search?	4.3	1.1	3.3	1.4	4.1	1.1

There was no statistically significant difference between systems on the searchers' familiarity with the topic being searched. It should be noted, however, that topic did have an overall effect on responses to the first item; participants were less familiar with topic 118, Congressman Mark Souder, than with any of the other items. For the remaining items on this questionnaire,[2] analysis of variance indicated that the features-only system was not perceived as positively as the other two systems. There were also statistically-significant relationships between topic and the responses (main effect of topic); and the interaction between topic and system was statistically significant.

After completing the four assigned searches for each system, each participant completed measures of usefulness (6 items), ease of use (6 items), and two dimensions of flow (4 items each). The results from these measures are shown in Table 5.

Table 5. User perceptions, based on post-system measures (1-5; lower scores indicate more positive attitudes

	Transcript-only		Features-only		Combined	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
Perceived ease of use	2.2	1.0	3.2	0.9	2.7	1.0
Perceived usefulness	2.3	1.0	4.1	0.9	2.8	1.1
Flow (enjoyment)	3.8	1.5	5.4	1.4	3.7	1.3
Flow (concentration)	3.2	1.4	4.4	1.7	3.2	1.2

The systems differed on perceived ease of use (F (2, 70) = 10.35, p<0.0001). Post hoc analyses indicated that the features-only system was perceived as harder to use than the other two systems. The systems differed on perceived usefulness (F (2, 70) = 29.17, p < 0.0001), with the same results from the post hoc analyses. Parallel results were obtained for enjoyment (F (2, 70) = 24.43, p < 0.0001) and concentration (F (2, 70) = 15.37, p < 0.0001). Thus, it can be concluded that these participants found the features-only system to be less useful, harder to use, and resulting in less positive perceptions of enjoyment and concentration than the other two systems.

After working with all three systems, the participants completed one additional questionnaire. Most (81%) of the participants indicated that they "completely" understood the search task and most (78%) found it "somewhat" similar to other searching tasks. Some (28%) found the systems to be "completely" different from each other and the remainder found the systems "somewhat" different from each other. The participants were then asked for direct comparisons of the three systems; their responses are shown in Table 6.

Table 6. Comparison of systems in post-session questionnaire (number of participants giving each response)

	Transcript-only	Features-only	Combined	No difference
Easier to learn to use	12	3	14	7
Easier to use	20	1	14	1
Liked the best overall	19	2	15	

These results clearly indicated that the features-only system was least preferred by the study participants. Each of the other two systems had some proponents, and several people found the systems to be equivalent in terms of how easy it was to learn to use them.

6 Discussion

The goal of our TREC VID participation this year was to fully take into account (1) users' perceptions toward information retrieval systems and (2) the variability in users' retrieval performance and perceptions. We accomplished the first goal by augmenting the basic NIST questionnaires with additional measures of user perceptions of system usefulness, ease of use, and flow in relation to the three systems being compared. We accomplished the second goal by increasing the number of study participants, developing a research design suitable for such typical social science research studies, and analyzing the data with methods that explicitly examine the variability of user performance/perceptions. The result of achieving these two goals is a data set that provides a rich picture of participants' interactions with and perceptions of three particular retrieval systems.

---

[2] The fourth item, related to features, was asked only when participants were searching the features-only or combined systems. Thus, the contrast tested for this item was only between those two systems; the difference was statistically significant.

The clearest finding from our analyses is that the features-only system is weaker than the other two systems. While users were able to achieve the same average level of precision with the features-only system, there were a large number of searches conducted with that system for which no shots were selected and, thus, no precision value was calculated. The participants were able to achieve more recall with the transcript-only and combined systems in less time per search.

The data on user perceptions also support this conclusion. With the features-only system, participants found it harder to get started on a search and harder to do the search; they were not as satisfied with the outcomes of their searches and did not feel they had enough time to complete their searches. When asked directly, they did not find the features searching capability to be useful. Similarly, participants' perceived ease of use, perceived usefulness, and flow all supported the conclusion that perceptions of the features-only system were more negative than for the other two systems. The post-session questionnaire, asking for a direct comparison of the systems, also corroborated this conclusion.

This conclusion, however, is not particularly interesting, in that the features-only system was developed for research purposes, and it was not expected that it could compete with two other systems that were more similar to the Web searching conducted daily by most participants. The interesting comparison is between the transcript-only system and the combined system. The transcript-only system is very similar to a "simple" Web search engine; the combined system augments the available features by allowing the users to also search on features extracted from the videos. This is a novel capability, applicable only to video content, and so it's of great interest to see how it fares alongside the more traditional transcript-only system.

When looking only at the comparison between the transcript-only system and the combined system, we find that there is no difference in precision achieved and no difference in recall achieved. The time per search was slightly longer for the combined system (5.2 minutes per search versus 4.0 minutes for the transcript-only system). There were no differences between the two systems on the post-search questionnaire results or the measures of usefulness, ease of use, or flow. Twelve people found the transcript-only system easier to learn, but 14 people found the combined system easier to learn. More people found the transcript-only system easier to use (20 versus 14) and liked it better overall (19 versus 15). In summary, people's performance with the combined system and their perceptions of it were remarkably positive, given that this experiment was their first exposure to the concept of features searching. It seems highly likely that the remaining differences in time per search and preferences will disappear as people become familiar with features searching and as the data underlying features searching becomes more robust. These results certainly warrant continued work to make features searching even more useful as a supplement to text searching.

7 Conclusion

In the current study, we have emphasized people's interactions with video retrieval systems. Our results indicate that, while a search system based on features alone is unlikely to be successful, a search system that combines transcript searching and features searching holds promise for future development.

Our results also indicate that, if user behaviors are to be taken into account, future TREC VID research designs should incorporate several changes. First, TREC participants should be encouraged to use a larger number of participants. People's behavior varies and the selection of just a few participants may lead to results that are not generalizable to larger populations. Second, the TREC VID search track should place more emphasis on *interactive* searching. There is ample evidence that people formulate and reformulate their search strategies during the search process, and our systems must be able to accommodate and support this type of iterative search process. Third, TREC studies should incorporate reliable and valid measures of user perceptions of systems. It is often the case that performance and attitudes are not highly correlated; but both must be taken into account in designing effective search systems. Incorporating these changes into TREC VID studies can only help increase our understanding of how people search video collections and how those searches can be effectively supported.

8 References

[1] Borgman, C. L. All users of information retrieval systems are not created equal: an exploration into individual differences. Info. Proc. & Mgmt., 25(3), 237-251, 1989.

[2] Browne, P., Czirjek, C., Gurrin, C., Jarina, R., Lee, H., McDonald, K., Smeaton, A. F., & Ye, J. Dublin City University Video Track Experiments for TREC 2002. http://www.cdvp.dcu.ie/Papers/TREC2002_Video_report.pdf. Last accessed October 31, 2003.

[3] Czekszentmihalyi, M. Beyond Boredom and Anxiety. San Francisco: Jossey-Bass, 1975.

[4] Davis, F. D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319-340, 1989.

[5] Eakins, J. P., & Graham, M. E. Content-Based Image Retrieval: A Report to the JISC Technology Applications Programme. Institute for Image Data Research, University of Northumbria at Newcastle, January 1999. http://www.unn.ac.uk/iidr/report.html. Last accessed May 14, 2003.

[6] Gaughan, G., Smeaton, A. F., Gurrin, C., Lee, H., & McDonald, K. Design, Implement, and Testing of an Interactive Video Retrieval System. Paper presented at MIR 2003 - 5th ACM SIGMM International Workshop on Multimedia Information Retrieval, Berkeley, CA, 7 November 2003. http://www.cdvp.dcu.ie/Papers/MIR2003.pdf. Last accessed October 31, 2003.

[7] Gauvain, J. L., Lamel, L., & Adda, G. The LIMSI broadcast news transcription system. Speech Communication, 37(1-2), 89-108, 2002.

[8] Ghani, J. A., Supnick, R., & Rooney, P. The experience of flow in computer-mediated and in face-to-face groups. ICIS Proceedings, 229-237, 1991.

[9] Greisdorf, H., & O'Connor, B. Modelling what users see when they look at images: a cognitive viewpoint. J. of Documentation, 58(1), 6-29, 2002.

[10] Grodal, T. Moving Pictures --- A New Theory of Film Genres, Feelings, and Cognition. Oxford: Clarendon Press, 1997.

[11] Lin, J. C.-C., & Lu, H. Towards an understanding of the behavioural intention to use a Web site. Internat. J. of Info. Mgmt., 20, 197-208, 2000.

[12] Moon, J.-W., & Kim, Y.-G. Extending the TAM for a World-Wide-Web context. Info. & Mgmt., 38, 217-30, 2001.

[13] Panofsky, E. Meaning in the Visual Arts: Meanings in and on Art History. Doubleday, 1955.

[14] Panofsky, E. Studies in Iconology: Humanistic Themes in the Art of the Renaissance. New York: Harper & Row, 1972.

[15] Saracevic, T., Kantor, P., Chamis, A. Y., & Trivison, D. A study of information seeking and retrieving. I. Background and methodology. J. of the Amer. Soc. for Info. Sci., 39(3): 161-176, 1988.

[16] Smeaton, A. F., & Over, P. The TREC-2002 Video Track: Overview [slides]. 2002.

[17] Teo, T. S. H., Lim, V. K. G., & Lai, R. Y. C. Intrinsic and extrinsic motivation in Internet usage. Omega-Internat. J. of Mgmt. Sci., 27(1), 25-37, 1999.

[18] Webster, E. J. Playfulness and Computers at Work. Ph.D. dissertation, New York University, 1989. UMI AAC8911170.