Christine Stachowicz. The Effectiveness of Storyboard Surrogates in the Subject Indexing of Digital Video. A Master's paper for the M.S. in L.S. degree. April, 2002. 61 pages. Advisor: Jerry D. Saye.

This study was designed to assess whether an indexer working from automatically generated storyboard surrogates could assign subject keywords to digital video, saving significant time while sacrificing little retrievability. The researcher assigned keywords to six videos selected from the Open Video Project repository. Three were indexed after viewing the videos in full; three were indexed after viewing their storyboard surrogates. A record of time consumed was maintained. Study participants viewed all six videos in full and listed what they believed to be the main topics addressed. Participant terms were compared to the researcher-assigned keywords. Retrievability was calculated as the percentage of participant terms that matched these keywords. Overall, time consumed in indexing the three videos based on their surrogates was 82% less than the time consumed in downloading and indexing the three full videos. Results reflected only 6% less retrievability for the three videos indexed based on their surrogates.

Headings:

Indexing – Video recordings

Information Retrieval

Information Systems – Special subjects – Video recordings

Internet Video

Subject Access

Video Surrogates

THE EFFECTIVENESS OF STORYBOARD SURROGATES IN THE SUBJECT INDEXING OF DIGITAL VIDEO

by
Christine Stachowicz

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Library Science.

Chapel Hill, North Carolina

April, 2002

Approved by:

_____
Advisor

Contents

Tables ii

Introduction 1

Motivation for the Project 2

Methodology 6

Compilation of Results

- Approach 1 – Strict Agreement of Terms 13
- Approach 2 – Conceptual Agreement of Terms 17

Analysis of Results 21

- Time Consumed 23
- Retrievability – Quantitative Analysis 28
- Retrievability – Qualitative Analysis 32
  - Pair 1 33
  - Pair 2 36
  - Pair 3 41

Summary of Results 43

Conclusion 46

Appendices 48

Works Cited 57

Tables

Table 1: Videos Selected for Testing 8

Table 2: Videos Included in Study 13

Table 3: Overall Strict Agreement Match Rates 17

Table 4: Overall Conceptual Agreement Match Rates 21

Table 5: Ti

er

video. For the most part, the correlations between familiarity rating and conceptual match rate did tend to be somewhat better than those between familiarity rating and strict match rate, but there were no strong relationships with respect to either.

Because no statistically significant relationships were noted between the participants' background information and their match rates, it can be surmised that the match rate results achieved in this study were not unduly influenced by outside factors. This means that similar results could be expected no matter the sex or adult age range of the participants, and regardless of their assessment of their searching skills. Likewise, since the participants' familiarity with the content of these videos had no significant bearing on their match rates under either the strict agreement or conceptual agreement approach, similar results might be expected irrespective of the subject matter of the videos. Based on these results, the pool of participants is considered to be fairly representative of the general public. Since this research study was based on a small sample of videos and a small group of participants, and given that there were no strong quantitative relationships between the data regarding either the videos indexed based on their surrogates or the videos indexed based on viewing them in full, qualitative analysis of the match rates of video pairs was deemed more constructive.

Retrievability – Qualitative Analysis

For purposes of this research study, it was determined that the conceptual agreement of terms approach would be more informative than the strict agreement of terms approach, and that the conceptual agreement match rate should be used as the primary measure of retrievability for each video. Therefore, qualitative analysis of paired videos will be limited to their conceptual agreement match rates, which will henceforth be referred to strictly as "match rates." **Table 10** lists match rates by video

pairs, along with the duration of each segment and the average familiarity rating noted by the participants.

Table 10
Conceptual Agreement Match Rates by Video Pairs

Segment Title	V/S **	Segment Duration m:ss	Average Familiarity/ Knowledge	Conceptual Agreement Match Rate
Pair 1				
A Wonderful New World of Fords (1960 Ford Spot)	V	3:00	3.1	96.5%
Roads to Romance: Coral Gables	S	3:00	2.7	75.3%
Pair 2				
SearchKids: A Digital Library for Children	V	6:00	2.5	65.6%
Television Remote Control (Tuner)	S	5:46	3.7	58.7%
Pair 3				
The Corvair in Action	V	6:25	2.8	72.2%
The Safest Place	S	6:23	3.5	90.6%

** Indexed based on full video **(V)** or surrogate **(S)**

Pair 1

The shortest videos tested were the 3:00 works in Pair 1. It is acknowledged that the shorter the video, the less time there would be to save, such that retrievability might be the key factor in judging the effectiveness of surrogates used in indexing shorter videos. *A Wonderful New World of Fords (1960 Ford Spot)*, indexed based on viewing the full video, earned a 96.5% match rate, versus a 75.3% match rate for *Roads to Romance: Coral Gables*, indexed via its surrogate.

A Wonderful New World of Fords, produced in 1960 by the Ford Motor Company, is described in the Open Video Project repository as a "Ford commercial linking new compact cars to futurism and the space frontier." Ironically, neither this researcher nor any of the 13 participants noted the concepts of space or futurism among their terms, nor did anyone mention COMPACT cars specifically. Nonetheless, all

seven researcher-assigned terms, plus four additional terms found in the title (WONDERFUL, NEW, WORLD, and SPOT) were conceptually matched 83 times out of 86 participant terms. The only unmatched participant terms were two participants' references to the music or jingle in the video and participant I's mention of U.S. with respect to automobile history. This could not be considered a broader term as previously defined. All 13 participants matched the researcher-assigned term FORD, while 11 matched 1960, and nine matched each CARS/AUTOMOBILES, ADVERTISEMENT, and either the proper names of the three models depicted in the video or the term MODELS.

This video garnered the highest match rate, but as indicated in **Table 10**, this high match rate did not correspond with the highest familiarity rating. The familiarity rating for this video averaged 3.1. Individual participant ratings ranged across the spectrum, with one participant circling 1 and one selecting 5. The exceedingly high match rate could be attributable to the fact that the video was a straightforward promotional piece and the researcher-assigned terms were simple. Although this video is over forty years old, today's public is generally familiar with car advertisements. Ford is a recognizable manufacturer, and one of the cars featured – the Thunderbird – is still in Ford's product line. Several participants alluded to difficulty in answering the question regarding their familiarity with or knowledge of what was covered in the videos. The lack of correlation between match rate and familiarity rating could have to do with the nonspecific nature of that question.

For *Roads to Romance*, indexed based on its surrogate **(Appendix A)**, all seven researcher-assigned terms, plus two additional terms found in the title (ROMANCE and ROADS), and CHEVROLET, found in the source organization field, were conceptually matched 61 times out of 81 participant terms. Four participants included a reference to the UNIVERSITY OF MIAMI, depicted in the video as one of the possible destinations in the Coral Gables area. In fact, the building shown in the first frame of the storyboard

surrogate is at the University of Miami. From the context of the rest of the frames, this researcher mistakenly considered that building to be a hotel or resort. Clearly, there is little, if any, chance an indexer would have accurately identified that building in the storyboard and assigned UNIVERSITY OF MIAMI as a keyword without reference to the audio track. Participant terms included three other references to specific locations in the video, which were unmatched by researcher-assigned terms. It seems reasonable to conclude that the use of storyboard surrogates could prove problematic in identifying specific people and places, except for the very recognizable, and proper names in general.

It is interesting to note that while four participants mentioned the University of Miami, two others did so but later crossed off the phrase before submitting their results. This implies that these two participants decided that the university was not among the main topics of the video, although they had identified the concept as present. Based on this researcher's viewing of the full video, UNIVERSITY OF MIAMI would indeed be a pertinent descriptor. Furthermore, members of the general public or K-12 educational community could realistically be interested in finding early footage of locations such as this.

The match rate for *Roads to Romance* was affected most by there being no reference to ADVERTISING among the researcher-assigned terms; eight participants mentioned this concept. Every participant also mentioned CHEVROLET or CHEVY. Had "source organization" not been considered one of the fields that would be scanned in a keyword search, these 13 terms would have gone unmatched, and the conceptual match rate for this video would have decreased from 75.3% to 59.3%. However, both concepts would have been addressed had the title frame of *Roads to Romance*, or any frames with text, been included in its storyboard surrogate.

Roads to Romance had two title frames, neither of which was represented in its surrogate. The first title frame read, "Chevrolet presents *Roads to Romance*." Based on the phrase "Chevrolet presents," this researcher would have assigned both CHEVROLET and ADVERTISEMENT as descriptors. (Similarly, the title frame for the Pair 3 video, *The Safest Place*, read "Chevrolet presents *The Safest Place*." This title frame was included in the surrogate for *The Safest Place*, which prompted this researcher to record both CHEVROLET and ADVERTISEMENT as descriptors for that video.) Had these terms been recorded for *Roads to Romance*, the match rate for this video would have been 85.2% rather than 75.3%. When compared to the 96.5% match rate for its pair member video, *A Wonderful New World of Fords*, the 85.2% match rate would have represented a difference of only 11.3 percentage points, rather than 21.2.

One participant noted BISCAYNE BAY among her terms. The second title frame in *Roads to Romance* read, "Coral Gables Florida – Sunland on Biscayne Bay." Had this frame been reflected in the storyboard surrogate, BISCAYNE BAY might also have been included among the researcher-assigned terms. This example suggests that storyboard surrogates might be most useful, for both indexing and browsing purposes, if they consistently included title frames, and ideally, any text frames.

Pair 2

The Pair 2 titles had the lowest retrievability rates of all three pairs. *SearchKids: A Digital Library for Children*, indexed based on viewing the full video, earned a 65.6% match rate, versus a 58.7% match rate for *Television Remote Control (Tuner)*, indexed via its surrogate. Although the raw scores are lower, the difference between the two rates is only 6.9 percentage points. This supports the hypothesis that the difference in potential retrievability between the two indexing methods could be relatively insignificant.

SearchKids, produced in 2001 by the Human-Computer Interaction Lab at the University of Maryland, presents results of research sponsored by the lab. Its description in the Open Video Project repository reads:

> An interdisciplinary, intergenerational team developed SearchKids, a zoomable digital library that contains multimedia information about animals. Our digital library supports collaboration by enabling several children to navigate the same information on the same computer at the same time. The design process as well as the technology is presented.

As might be expected due to its more technical nature, this title earned the lowest overall familiarity rating, 2.5. Of the 11 participants who viewed *SearchKids*, four circled a familiarity rating of 1 and another four circled 4. Interestingly, and reinforcing the lack of correlation between familiarity and retrievability, participants who indicated a familiarity level of 4 had an average match rate of 60%, while those who indicated a familiarity level of 1 averaged a 70% match rate.

Of the 13 index terms, including the title term, SEARCHKIDS, eight were conceptually matched 42 times out of 64 participant terms. The five researcher-assigned terms that were not listed by any participants included the names of two organizations featured in the video. HUMAN-COMPUTER INTERACTION LAB (HCIL) was the producer of the video and the setting for much of its action. HCIL partnered with YORKTOWN ELEMENTARY SCHOOL to create the SearchKids digital library, and the school was also the setting for some of the testing depicted in the video. No participants mentioned either of these institutions. Other terms not mentioned by any participants were INTERFACE, EFFICIENCY (specifically, SEARCH EFFICIENCY), and ZOOMABLE. This researcher was especially doubtful that ZOOMABLE would be noted by any participants, but the term was included as a keyword because it was mentioned frequently in the video. In retrospect, INTERFACE and EFFICIENCY may have been included as keywords because of this researcher's familiarity with these concepts in terms of online searching, but the terms may not be widely used among the general public.

At first glance, this example might seem to reflect a propensity to over-index upon viewing a video in full. The notable gaps in the indexing bring that into question, however. Six participants mentioned COMPUTER, PROGRAM, or ONLINE, five noted RESEARCH/DEVELOPMENT, and four listed EDUCATION. Three participants referred to the GRAPHIC nature of the SearchKids digital library, and three alluded to PROMOTION of the product. This researcher considered that last reference attributable, in part, to the fact that the other five videos shown were promotional in nature. The remainder of the unmatched terms seemed reasonable for *SearchKids*. While participant familiarity was not a good predictor of match rate, the discrepancy between researcher-assigned terms and participant terms for this video might have been a function, instead, of this researcher's knowledge of the topic.

The unmatched participant terms tended to be straightforward, as in the case of COMPUTER and EDUCATION, while the researcher-assigned terms INTERFACE and EFFICIENCY may reflect some assumptions grounded in this researcher's familiarity with digital libraries. For example, where digital libraries are concerned, COMPUTER would be taken for granted, as to a lesser degree, might EDUCATION. RESEARCH/DEVELOPMENT could likewise be assumed to be components of DESIGN, a researcher-assigned term which was matched by only two participants. The index terms that were matched most often were the simpler ones – SEARCH and ANIMALS, each listed by five participants – and those included in the video title – DIGITAL LIBRARY and CHILDREN, listed by nine and ten participants, respectively. It is worth noting that the two participants who did not list the obvious DIGITAL LIBRARY among their search terms were participants **C** and **D**, both of whose responses consisted primarily of longer phrases that were more conceptual than descriptive in nature.

Overall, the results for *SearchKids* reinforce the need for a strong keyword search facility to be supported by a robust thesaurus that highlights broader, narrower,

and related terms. In addition, the discrepancy between researcher-assigned terms and participant terms should serve as a reminder that within the indexing function, the potential audience for the material must be considered. While the research reported in this paper is ultimately focused on facilitating the public's access to moving image material, this researcher's exposure to the topic of digital libraries in the context of graduate study may have led to the assignment of some keywords that were inadvertently directed more toward the research community. Attempts to accommodate both ends of this spectrum may indeed lead to over-indexing and, therefore, to potential noise within keyword search results. This may be an inevitable cost of trying to maximize access for a broad range of potential users, a subject worthy of further research employing a greater number and wider variety of videos and participants.

Television Remote Control, produced in 1961 by RCA Victor, earned the highest overall familiarity rating (3.7), but the lowest match rate of all six titles tested. Its description in the Open Video Project repository reads simply, "early technology to enable the channel-surfer." In the video, a voiceover explains the features of the new RCA Victor console television and its accompanying remote control, while a woman demonstrates their use. Four of the six researcher-assigned terms, plus TUNER, found in the title, and RCA VICTOR, the source organization, were conceptually matched 37 times out of 63 participant terms. Had "source organization" not been considered one of the fields that would be scanned in a keyword search, the seven participant mentions of RCA VICTOR would have gone unmatched, and the conceptual match rate for *Television Remote Control* would have decreased from 58.7% to 47.6%. Two participants mentioned INSTRUCTIONS and "HOW TO," matching researcher-assigned terms, but credit for the match rate is due primarily to the 25 incidents of TELEVISION and REMOTE CONTROL among the participant terms. These terms would have been matched by the keyword

search's inclusion of the title field. They were also specifically included as keywords because it appeared from the surrogate **(Appendix B)** that little else was covered in the video. This researcher also included as keywords both FUNCTION and OPERATION with respect to the remote control, but no participants listed these or similar terms.

Even after viewing the video in full, this researcher questioned whether *Television Remote Control* was an advertisement. It should be noted that it was not described as one in the repository. The participants' interpretation of the video as an ADVERTISEMENT was clear, however, as nine of 13 noted the term in their responses. Five participants also specified that this was a COLOR television, and six used a variety of terms to make reference to this being a display of NEW TECHNOLOGY. Unlike in the case of *Roads to Romance*, no title or text appeared anywhere in *Television Remote Control*. Thus, this surrogate could not have been supplemented by the addition of any such frames. Perhaps the woman's appearing to model the television should have implied the promotional aspect of the piece, but this researcher simply did not pick up that notion. Since the storyboard and the whole video were primarily sepia-toned, it might have been difficult to discern that a color television was being depicted. Inferring from the images themselves the sense of this being a recent invention or new technology would likewise have been problematic.

Based on the participant responses, then, the core topics of *Television Remote Control* were addressed in the indexing, but the full perspective of the video was not. This failure to convey the tone or character of a work could be expected as a weakness of storyboard surrogates, and could render completely ineffective their use with respect to abstract works. Again, further research with a wider variety of videos may offer an opportunity to qualitatively consider the tradeoff between this potential shortcoming and the time savings afforded by video surrogates.

Pair 3

Results for the Pair 3 videos were perhaps the most surprising. *The Safest Place*, indexed via its surrogate, earned a conceptual match rate of 90.6%, which was 18.4 percentage points higher than the 72.2% match rate of *The Corvair in Action*, indexed based on viewing the full video. When compounded by the fact that it took over four times as long to download, view and index *The Corvair in Action*, the use of video surrogates indeed seems promising. It should be noted that this anomaly was evident even under the strict agreement of terms approach, where the match rate for *The Safest Place* exceeded that of *The Corvair in Action* by 9.8 percentage points.

The Corvair in Action, produced in 1960 by the Chevrolet Division of General Motors Corp., is described in the Open Video Project repository as a "promotional film for the controversial Chevrolet Corvair." All seven researcher-assigned terms were conceptually matched 57 times out of 79 participant terms. The match rate for this video was affected most by there being no reference to time period among the researcher-assigned terms; ten participants mentioned either 1960, 1960s, or HISTORY. The time period covered in *The Corvair in Action* could be correctly inferred from the production date of the video, but since that relationship cannot always be assumed, the creation date field was not considered one that would be scanned in a keyword search. Failure to note the year of the Corvair depicted in the video was likely the result of researcher oversight. Had the time period been reflected in the researcher-assigned terms, the match rate for this video would have risen to 84.8%.

It seems that a difference in perspective, rather than researcher error, accounted for the other main discrepancy between researcher-assigned terms and participant terms. Seven participants referred to the concept of TESTING in *The Corvair in Action*. The video depicted the Corvair in a variety of situations, some clearly engineered for promotional purposes. Rather than consider the general testing aspect of these scenes,

however, this researcher noted terms related to the results of that testing, namely the SAFETY FEATURES and superior ROAD HANDLING of this compact car. The discrepancies between researcher-assigned terms and participant terms can ultimately be viewed much like the inconsistencies among the participants themselves. No matter the medium, the same content can be interpreted somewhat differently by any number of people. The key remains that the core topics of this video, as determined by the degree to which participants conceptually agreed on them, were reflected in the researcher-assigned terms.

The same is true for *The Safest Place*, a work whose indexing was based on its storyboard surrogate. This video, produced in 1935 by the Chevrolet Division of General Motors Corp., is described in the Open Video Project repository as "how the automobile is the safest place a person can be." In it, the dangers of boats and airplanes and even the home are compared to the relative safety of the Chevrolet automobile, referred to as a "living room on wheels." Of the 12 researcher-assigned terms, ten were conceptually matched 58 times out of 64 participant terms. The two researcher-assigned terms not mentioned by any participants were admittedly vague references to LAND and WATER safety. Each of the six unmatched participant terms was listed only once. The lack of agreement among participants regarding these terms implied that the terms had little significance to the video as a whole. While *The Safest Place* enjoyed a high overall familiarity rating, it was yet another example of how this participant assessment proved a poor predictor of match rate. Indeed, the lowest participant match rate, 66.7%, was calculated for the one individual who indicated a familiarity of 5.

The variety and number of frames in the storyboard surrogate for *The Safest Place* led to this researcher spending significantly more time indexing it than was spent indexing the other two surrogates. It is possible that the process of piecing together a

story from 44 frames resulted in more attention being paid to each individual frame. The researcher-assigned terms were not finalized until after several close reviews of the entire sequence of frames in the storyboard surrogate. This type of effort was not undertaken with the surrogate for *Television Remote Control* because there was so much similarity between the scenes it depicted. For the most part, these frames simply alternated between the remote control and the television itself (see **Appendix B**). Likewise, with the ten frames of the *Roads to Romance* storyboard surrogate (**Appendix A**), a quick and incomplete determination of the video's content was made largely because there were fewer frames to interpret.

Simply requiring more frames in each storyboard surrogate is clearly not the answer, however. That would not have changed the indexing for *Television Remote Control*, for example. The MERIT keyframe extraction technology, based on scene changes, is still being tested. As research into this and other related technologies continues, the resulting improvements in the video surrogates created will undoubtedly make them more useful for manually subject indexing digital video. Alternate methods of surrogate presentation, such as slide shows that would allow for the individual frames to be larger, can also be applied to this purpose.

Summary of Results

Motivating this research study was the theory that a human indexer could use an automatically generated storyboard surrogate to assign subject-oriented keywords to a video, saving significant time while sacrificing relatively little in the way of retrievability. The results of this study are summarized in **Table 11**.

Table 11
Summary of Results
Time Saved and Retrievability Lost

Segment Title	V/S **	Time Consumed mm:ss	Time Saved %	Conceptual Agreement Match Rate	Retrievability (Lost) Gained
Pair 1					
A Wonderful New World of Fords					
(1960 Ford Spot)	**V**	9:25		96.5%	
Roads to Romance: Coral Gables	**S**	1:29	**84%**	75.3%	**(22%)**
Pair 2					
SearchKids: A Digital Library for Children	**V**	9:43		65.6%	
Television Remote Control (Tuner)	**S**	1:20	**86%**	58.7%	**(11%)**
Pair 3					
The Corvair in Action	**V**	17:32		72.2%	
The Safest Place	**S**	3:52	**78%**	90.6%	**26%**
Overall Average			**82%**		**(6%)**

** Indexed based on full video **(V)** or surrogate **(S)**

The "Time Saved %" column indicates that for each pair of videos, the time consumed in subject indexing the video based on its surrogate was 84%, 86%, and 78% less, respectively, than the time consumed in indexing the full video in that pair. In the aggregate, the time consumed in indexing the three videos based on their surrogates was 82% less than the time consumed in indexing the three full videos.

For Pair 1, the match rate of the video indexed via its surrogate was 78% of the match rate of the title indexed upon viewing the video in full. The resulting 22% relative loss in retrievability is reflected in the "Retrievability (Lost) Gained" column. For the Pair 2 titles, the relative loss in retrievability was 11%. Due to the anomaly in the Pair 3 data, the use of the surrogate there resulted in a 26% relative gain in retrievability. To calculate an overall match rate for the three videos indexed via their surrogates, their collective matched and unmatched terms were counted, then added together to arrive at a collective total number of terms. The collective matched terms sum was divided by the collective total number of terms, yielding the overall match

rate. The overall match rate for the other three titles, indexed based on viewing the videos in full, was calculated in the same manner. A comparison of these two overall match rates indicates only a 6% net relative loss in retrievability for the videos indexed based on their storyboard surrogates.

Analysis of the results for the individual videos gives rise to suggestions for further study and for improvements in storyboard surrogates. The results also reveal some potential shortcomings of relying on surrogates in the subject indexing of digital video. For a digital video repository employing video surrogates as an indexing tool, there is likely to be a tradeoff between the time consumed to make material accessible and the ultimate retrievability of that material. The promising results of this research study indicate that it would be worth pursuing the topic further by exploring a variety of video surrogates and assessing which types of surrogates lend themselves best to human interpretation for subject indexing purposes. A larger study, utilizing a more diverse range of videos, would also be useful for evaluating how effectively surrogates might convey different subject material.

For the study reported in this paper, this researcher viewed each video or surrogate and assigned keywords to describe its subject content. Rather than rely on this one interpretation as the basis for testing, it would be worthwhile to conduct a two-phase study in which various experienced indexers viewed each video or surrogate and assigned the keywords. In phase one of the study, the indexers' results could be evaluated against each other. In phase two of the study, participants representing the general public could view each video in full and describe what they believe to be its main topics. In order to gauge the potential retrievability of each video, the participant results could then be compared to the terms assigned by the multiple indexers. In addition, future studies on this topic could benefit from having more participants and from specifically requesting those participants to provide terms by which they might

search for a particular video. Results in this stricter format, versus the variety of participant responses in the study reported in this paper, could be analyzed more objectively within the structure of an existing keyword search thesaurus.