

디지털 영상 초록의 설계와 평가에 관한 연구

Designing and Evaluating Digital Video Storyboard Surrogates

김 현 희(Hyun-Hee Kim)*

김 용 호(Yong-Ho Kim)**

고 수 현(Su-Hyun Ko)***

< 목 차 >

I. 서론	III. 영상 초록의 배열 모형 설계와 평가
1. 연구의 필요성과 목적	1. 설 계
2. 연구 방법	2. 평 가
II. 이론적 배경	3. 논 의
1. 선행 연구	IV. 결론과 후속 연구
2. 연구 가설	

초 록

본 연구에서는 먼저 영상 초록의 배열 모형을 영상물의 소통 구조 이론, FRBR 모형 및 용어간의 구문적 및 어의적 관계를 고려하는 PRECIS 색인 이론 등을 이용하여 설계하고 이 모형에 따라서 영상 초록을 구성한 후 순차적 영상 초록과 요약문과 색인이 추출의 정확도 측면에서 어떤 차이를 보이는지 6개의 표본 비디오와 26명의 피조사자를 이용하여 실험을 통해서 조사해 보았다. 조사 결과, 배열 모형 기반 영상 초록이 순차적 영상 초록 보다 더 정확한 요약문을 추출하는 것으로 나타났고 색인이 추출의 정확도는 큰 차이가 없는 것으로 나타났다. 또한 영상 초록의 반복 보기의 효과를 측정할 결과 순차적 영상 초록을 먼저 보고 배열 모형 영상 초록을 반복해서 본 경우에 요약문의 정확도가 크게 증가하는 것으로 나타났다. 이러한 실험 결과에 기초하여 디지털 비디오 도서관 환경에서 영상 초록의 두 가지 활용 방안 즉, OPAC 환경에서 영상 초록의 초기값으로 순차적 영상 초록을 제시하고 선택을 통해서 배열 모형 기반 영상 초록이 함께 출력되는 비디오 브라우징 인터페이스를 구성하는 것과 제안된 배열 모형 기반 영상 초록을 영상 질의의 구조화된 매칭 자료로 활용하는 것에 대해서 제안하였다.

주제어: 영상 초록, 영상 소통 이론, 순차적 영상 초록, 배열 모형 기반 영상 초록, 디지털 비디오 도서관

ABSTRACT

This study examines the design and utilization of video storyboard surrogates in the digital video libraries. To do this, first we constructed the arrangement model of key-frames for storyboard based on the FRBR model, image communication and PRECIS indexing theories and evaluated the model using 6 sample videos and 26 participants. The study results show that the video storyboard surrogates based on the arrangement model has a higher accuracy value in terms of summary extraction than that of the sequential video storyboard. Moreover, watching both types of video storyboard one after another, especially browsing the sequential video storyboard first and then the arrangement model-based one, produces a remarkable increase in accuracy value of summary extraction. The study proposes two methods of utilizing the video storyboard surrogates in the digital video libraries: Designing a video browsing interface where users can use the sequential storyboard as a default and then the arrangement model-based one for re-watching; and utilizing the arrangement model-based storyboard as structured match sources of image-based queries.

Key Words: Video Storyboard Surrogates, Digital Video Libraries, Arrangement Model of Key-frames, Image Communication, Video Browsing Interface

* 명지대학교 문헌정보학과 교수(kimhh@mju.ac.kr)(제1저자)

** 부경대학교 신문방송학과 교수(kimyho@pknu.ac.kr)(공동저자)

*** 명지대학교 교육대학원생(snow27@dreamwiz.com)(공동저자)

• 접수일: 2007년 11월 24일 • 최종심사일: 2007년 11월 26일 • 최종심사일: 2007년 12월 24일

I. 서론

1. 연구의 필요성과 목적

정보 기술의 발전과 인터넷 웹의 등장으로 방송국은 물론 대학, 학교, 연구기관 심지어는 공공기관의 도서관 환경에서도 기존 텍스트 기반의 자료가 멀티미디어 자료로 변화되고 있다.¹⁾ 이와 같은 정보 매체의 급격한 변화로 디지털 비디오 시스템 구축의 필요성이 높아지고 있다. 더 나아가 참여, 개방, 공유의 이념을 바탕으로 웹을 플랫폼으로 하여 이용자들이 다양한 멀티미디어 콘텐츠를 생산·유통·소비하는 웹2.0 시대에 들어서면서 UCC(User Created Content)가 주목받고 있다. 이러한 정보 환경의 변화는 멀티미디어 정보가 도서관의 핵심적인 콘텐츠가 될 것이라는 것을 어렵지 않게 예측하게 해 준다. 이러한 콘텐츠의 변화에 발 맞춰 특정 주제에 대한 비디오의 적합성을 평가할 수 있도록 해 주는 효율적인 내용 분석 방법에 대한 연구 수행은 더 이상 미룰 수 없는 과제가 되었다.

최근 정보시스템이나 인터넷 환경에서 비디오 영상 초록이 다른 메타데이터 요소들과 함께 사용되어 검색자의 적합성 판정을 돕는 지시적 초록의 기능을 하고 있으며 또한 비디오의 키프레임이 키워드와 유사한 기능을 하고 있다고 보고하고 있다.²⁾³⁾ 기존 연구들이 영상 초록의 중요성을 제안하면서도 효율적인 영상 초록 구성에 대해서는 심도 있게 다루지 않고 있음을 발견할 수 있다. 전통적인 텍스트 초록에는 키워드들이 다수 포함되지만 이들 키워드들은 문맥적으로 의미를 갖는 구 또는 문장으로 연결되어 있다. 이에 비해서 영상 초록은 너무 단순한 구조를 가지고 있다. 즉, 전체 비디오 클립에서 일정한 시간 간격에 따라서 키프레임들을 추출하거나 이미지 프로세싱을 이용하여 장면을 구분하고 각 장면의 중간에 나타나는 프레임들 키프레임으로 추출하거나 또는 정해진 알고리즘에 따라서 키프레임을 추출하여 이들을 단순히 순서대로 배열하고 있다.⁴⁾ 비디오의 영상 초록을 구성할 때 비디오의 제작자가 의도한 대로 키프레임들을 비디오에서 나타난 순서대로 배열하는 것 대신 문맥적인 측면을 고려하여 키프레임들을 배열하는 방법이 제안되기도 했지만 이에 대한 유용성은 검증되지 않았다.⁵⁾

1) Smeaton, A. F. "Indexing, browsing, and searching of digital video." In: *Annual Review of Information Science and Technology* Edited by Blaise Cronin, 2004, pp.371-372.

2) Hughes, A. et al. "Text or pictures? An eye-tracking study of how people view digital video surrogates." In: *Proceedings of CIVR*, 2003, pp.271-280.

3) 이태동. 통합된 인덱싱 방법을 이용한 내용기반 디지털 방송 검색 시스템의 설계와 구현(박사학위논문, 아주대학교 컴퓨터공학과, 2003), pp.9-10.

4) Browne, P. and Smeaton, A. F. "Video retrieval using dialogue, keyframe similarity and video objects." *ICIP 2005 - International Conference on Image Processing*, Genova, Italy(2007), pp.11-14.

5) Lyer, H. and Lewis, C. D. "Prioritization strategies for video storyboard keyframes." *Journal of the American Society for Information Science and Technology*, No.5(2007), pp.629-644.

본 연구에서는 영상 초록의 키프레임 배열 모형을 영상물의 소통 구조 이론, FRBR 모형 및 용어간의 구문적 및 어의적 관계를 고려하는 PRECIS 색인 이론을 기초로 하여 설계한다. 설계된 배열 모형에 따라서 영상 초록을 구성한 후 순차적 영상 초록과 요약문과 색인어 추출의 정확도 측면에서 어떤 차이를 보이는지 실험을 통해서 평가해 본 후, 이 배열 모형 기반 영상 초록을 디지털 도서관 환경에 적용하여 이용자들의 적합성 관정을 좀 더 용이하게 하고 더 나아가 영상 질의의 구조화된 매칭 자료로 활용할 수 있는 방안을 모색하고자 한다.

2. 연구 방법

배열 모형 기반 영상 초록의 유용성을 알아보기 위해서 4개의 연구가설을 설정하고 이를 검증하기 위해서 표본 비디오 6개를 이용하여 순차적으로 배열한 영상 초록과 배열 모형을 기반으로 한 영상 초록을 볼 수 있는 웹 사이트를 구축하였다. 구축한 두 가지 유형의 영상 초록을 통한 색인어와 요약문 추출의 정확도를 측정하기 위해 동일한 과목을 듣는 학부 학생과 대학원생을 피조사자 그룹으로 선정하였다. 즉, 피조사자는 D대학의 조사방법론과목을 수강하고 있는 신문방송학과 전공 학부 학생 8명 그리고 J대학의 정보검색과목을 이수하고 있는 문헌정보학과 전공 학부 학생 8명과 정보네트워크과목을 수강하고 있는 사서 전공 교육대학원생 10명, 총 26명으로 구성하였다. 실험은 두 대학의 컴퓨터실에서 따로 진행하였고, 각 피조사자가 한 대의 컴퓨터를 사용하도록 하였다. 즉, 피조사자에게 영상 초록이 있는 웹사이트 주소를 알려 주고 각 초록을 보게 한 후 4개 이상의 색인어와 두 문장 이상의 요약문을 기술하게 하였다. 두 가지 유형의 초록(순차적 영상 초록과 배열 모형 기반 영상 초록)의 요약문 및 색인어의 평균 정확도 차이 그리고 두 가지 유형의 반복 보기(유형 1: 순차적 초록->배열 모형 초록과 유형 2: 배열 모형 초록->순차적 초록)의 요약문 및 색인어의 평균 정확도 차이를 검증하기 위해서 두 집단 간의 차이를 분석할 때 사용하는 t-검증을 이용하였다.

II. 이론적 배경

1. 선행 연구

비디오의 영상 소통 구조와 영상 초록을 활용한 비디오 검색 시스템의 사례 연구에 대한 선행 연구를 살펴보면 다음과 같다.

가. 영상 소통 구조

Lyer와 Lewis는 비디오의 영상 초록을 구성할 때 키프레임의 순서가 비디오 의미를 파악하는데 영향을 미칠 수 있다는 제안을 하였다.⁶⁾ 그들은 비디오의 제작자가 의도한 대로 키프레임들을 비디오에서 나타난 순서대로 배열하는 것 대신 문맥적인 측면을 고려하여 키프레임들을 배열하는 방법의 유용성에 대해서 기술하고 있다. 그들은 키프레임을 3단계 즉, 제1단계는 1차 핵심 주제 정보(활동과 객체), 제2단계는 2차적인 배경 정보(활동과 객체) 그리고 제3단계는 3차적인 배경 정보(속성, 장소 및 시대)로 배열하는 것이 효율적이라고 제안하였다. 그러나 이들은 키프레임들이 각 단계로 분류되기 위해서 어떤 기준들이 적용되어야 하는지에 대한 구체적인 설명을 하지 않았고 이러한 배열 모형 기반의 영상 초록이 순차적 영상 초록 보다 비디오의 의미 추출에서 더 성공적으로 활용될 수 있음을 증명하지 않았다.

영상 자료의 의미 파악을 위해서 영상물의 소통구조를 좀 더 자세히 분석해 볼 필요가 있다. 이를 위해서 시퀀스 이론과 서사 이론에 대해 살펴볼 필요가 있다. 시퀀스 이론은 모든 형태의 소통(상호작용)은 시작 시퀀스와 종료 시퀀스를 거친다고 가정하고 시작에서는 채널 오픈, 신원 확인, 토픽 전개라는 일련의 시퀀스를 거쳐서 소통이 시작되고, 그 원래의 소통 목적이 달성되면 소통은 토픽 종료, 관계 종료, 채널 종료라는 시작 시퀀스와는 정반대의 수순을 거쳐서 종료된다는 이론을 기초로 하고 있다. 한편 영상물의 서사구조는 이야기와 담론으로 나뉜다. 서사의 내용은 이야기이고, 이 내용을 전하는 서사형식이 담론이다. 이러한 이야기와 담론이 수용자에게 전달되는 과정이 소통이다. 소통은 효율성과 정확성을 위해 관례처럼 구조화되는 경향이 있다. 영상서사의 전달에서도 이는 마찬가지이다. 짧은 시간에 적은 노력으로 정확하게 서사를 전달하는 것이 영상서사 소통 구조화의 목적이다. 영상서사의 이야기는 3요소로 구성된다. 사건, 인물, 배경이 그것이다. 핵심 사건과 주변 사건이 있고, 주인공과 주변 인물이 있으며, 이야기가 이루어지는 시간과 공간이라는 배경에서 사건들이 인물들에 의해 수행된다.⁷⁾

김재우는 TV나 영화로 대표되는 영상 미디어에 공통적으로 적용되는 구성 법칙과 이에 부가되어 있는 의미를 사례별로 추출해 내는 일이 영상 제작자와 이용자가 어떤 상호 주관성을 근거로 하고 있는지를 파악하는데 있어서 기초 작업이 될 것이라고 가정하였다.⁸⁾ 그는 영상을 하나의 기호 체계로 인식하고 그 구조가 지니는 인식적, 문화적, 사회적 상징성에 명백한 법칙이 있다고 제안하였다. 영상이 특정한 의미를 지니고 있고 또 그렇게 해석될 수 있다는 믿음은 영상이 기호로 작용할 수 있다는 기호학적 관점에 근거한다.

김무규는 이러한 기호학적 접근 방식은 영상의 본질적인 특성을 충분히 고려하지 못하는 문제점

6) Lyer, H. and Lewis, C. D. *Ibid.*, pp.629-644.

7) 존 버거 저, 강명구 역. *영상커뮤니케이션과 사회*(서울 : 나남출판사, 1991), pp.1-15.

8) 김재우. "영상 기호의 의미 작용에 관한 연구." *동서언론*, 제6권(2002), pp.41-44.

이 있다고 주장한다.⁹⁾ 그는 영상을 특정 기의(기표가 담고 있는 개념)에 대한 기표(언어로 표시되는 기호)로 작용한다고 전제하기보다 이 사이에 존재하는 기호화과정으로 간주하고 이러한 이해를 위해서 사이 공간(의미 작용이나 소통상황에서 발생하는 두 가지 구성 요소 즉, 기표와 기의, 서사 행위와 서사체, 정보의 생산자와 이용자와 같은 요소들의 차이가 존재하는 공간)이라는 개념을 제안하였다. 황민선은 영상물은 '보기 위한 것'이고 출판물은 '읽기 위한 것'이라는 매체상의 근본적인 차이 이외에 출판물과 영상물이 어떤 차이를 가지는지 분석하였다.¹⁰⁾ 영상물은 도상적 기호(iconic sign)를, 출판물은 상징적 기호를 더 용이하게 사용한다. 다시 말해서 영화의 상은 감각적, 현실적인 상이며 소설의 상은 관념적 심상이다. 감상 형식에 있어서 독자의 역할을 살펴보면, 자유로운 재창조의 과정을 갖는 출판물의 독자와 달리 영상물의 관객은 피상적이고 수동적임을 고찰할 수 있다. 또한 매체적 속성에 있어 출판물을 원작으로 하는 영상물은 출판물에 비해 더 대중적으로 나타났다.

나. 영상 초록을 활용한 비디오 검색 시스템의 사례 연구

Open Video 프로젝트(<http://www.open-video.org/>)는 노스캐롤리나 대학의 문헌정보학과에 의해서 수행되었는데 전자 환경에서 브라우징과 탐색을 이해하고, 디지털 도서관 연구와 개발을 위해서 Open Video Digital Library(OVDL)이라는 오픈 소스 디지털 비디오 테스트베드를 구축하였다.¹¹⁾ 현재 테스트베드는 3,968건의 비디오 자료가 올려져 있으며 메타데이터는 연도, 장르, 키워드, 저작권 사항 등 12개의 요소로 되어 있으며, 비디오의 영상 초록은 물론 고속 감기(fast-forward), 처음 7초 듣기(7-second excerpt), 전체 클립 등을 브라우징해 볼 수 있다. 또한 이 프로젝트 내에서 비디오 해제와 요약 도구 및 비디오 색인과 시각화 조직자와 같은 비디오 요약 데이터를 생성, 조직하는 실험적인 도구와 시스템 개발 등과 같은 연구들이 수행되었다.¹²⁾ 인포미디어(Infomedia) I과 II는 미국 카네기 멜론 대학에서 개발한 비디오 도서관 시스템으로 비디오내의 텍스트 정보, 객체 인식, 오디오의 내용 인식 등을 통해서 고수준의 특징값을 사용한 비디오 색인 기법을 사용함으로써 데이터베이스 구축과 검색의 효율을 높였다. 즉, 음성으로 질의를 하면 관련 자료들의 비디오 내용을 요약하여 보여주는 '요약 비디오(skim video)' 등을 보여 주며 이러한 메타데이터들을 통해서 정확하게 찾고자하는 비디오인가를 판단하게 된다.¹³⁾

9) 김무규. "영화 영상에 관한 기호학적 이해의 문제점과 대안: '사이공간' 개념을 중심으로," 언론과학연구, 제6권, 제3호(2006), pp.65-69.
 10) 황민선. "출판물과 출판물을 원작으로 하는 영상물의 매체 특성 비교 연구." 한국출판학연구, 제48권(2005), p.371.
 11) Marchionini, G. and Geisler, G. "The Open Video Digital Library." *D-Lib Magazine*, Vol.58, No.5(2006), pp.629-644. <<http://www.dlib.org>>[cited 2007.8.25].
 12) Marchionini, G., Wildemuth, B. M., and Geisler, G. "The Open Video Digital Library: A Möbius strip of research and practice," *Journal of the American Society for Information Science and Technology*, Vol.57, No.12(2006), p.1631.
 13) Hauptmann, A. G. "Lessons for the future from a decade of Infomedia video analysis research." *Lecture*

2. 연구 가설

본 연구에서는 영상 초록의 키프레임 순서가 영상자료의 의미 파악에 어떤 영향을 미치는지 조사하고자 한다. 본 연구에서 제안한 배열 모형은 전체 내용 전개에 대한 구조적 틀을 보여 주기 때문에 요약문 정확도는 순차적 영상 초록의 정확도 보다는 높게 나타날 것으로 가정한다. 또한 색인어의 정확도도 배열 모형 기반 영상초록의 색인어 정확도가 다소 높게 나타날 것으로 예상된다.

다음으로 영상 초록을 볼 때 대개 한번만 보는 것이 아니라 반복해서 화면들을 보게 되는데 이러한 반복 효과에 대해서 알아보하고자 한다. 반복 효과를 측정하기 위해서 두 개의 서로 다른 영상 초록(순차적 초록과 배열 모형 기반 초록)을 보게 하는 경우 어떤 유형의 초록을 먼저 보여 주는 것이 요약문과 색인어의 정확도를 증가시켜 주는지 알아보하고자 한다. 이와 같은 연구 문제들을 알아보기 위해서 다음과 같은 네 가지 가설을 설정하였다.

- 1) 순차적 영상 초록과 배열 모형 기반 영상 초록의 요약문 정확도는 차이가 있을 것이다.
- 2) 순차적 영상 초록과 배열 모형 기반 영상 초록의 색인어 정확도는 차이가 있을 것이다.
- 3) 영상 초록 반복보기의 유형 1(순차적 초록-)배열 모형 초록)과 유형 2(배열 모형 초록-)순차적 초록)간의 요약문 정확도는 차이가 있을 것이다.
- 4) 영상 초록 반복보기의 유형 1(순차적 초록-)배열 모형 초록)과 유형 2(배열 모형 초록-)순차적 초록)간의 색인어 정확도는 차이가 있을 것이다.

Ⅲ. 영상 초록의 배열 모형 설계와 평가

1. 설 계

영상 초록 배열 모형을 PRECIS 색인 이론, FRBR 개념 모형 및 영상물의 소통 구조 이론을 기초로 하여 설계하고자 한다. 영상 초록의 배열 모형은 개념간의 관계를 나타내는 관계적 기호와 개념의 범주를 나타내는 범주적 기호를 사용하여 용어들을 문맥의존적인 순서에 따라 조합하는 PRECIS 색인에서 아이디어를 얻은 것이다.¹⁴⁾ 이와 같이 PRECIS 색인 이론에서 시작하여 영상 초록의 키프레임의 문맥적 측면을 고려한 배열 모형 구성 요소들을 추출하기 위해서 그 다음으로 FRBR 모형과 영상물의 소통 구조 이론을 살펴보았다. 먼저 FRBR 모형에서는 서지정보개체를 3개 집단으로 정의

Notes in Computer Science, Vol.3568(2005), pp.1-10.

http://www.informedia.cs.cmu.edu/documents/CIVR05_Hauptmann.pdf[cited 2006.6.25].

14) 정영미. 정보검색론(서울 : 구미무역, 1993), pp.131-132.

하고 각 집단의 개체를 제1집단은 저작, 표현형, 구현형, 개별자료, 제2집단은 개인과 단체, 제3집단은 개념, 대상, 사건, 장소로 정의하였다. 특히 배열 모형의 구성을 위해서 제2집단과 제3집단을 살펴 보았다. 먼저 제2집단은 개인과 단체로 행위나 이벤트의 중심이 되는 행위자들이다. 제3집단의 개념은 추상적인 관념이나 이상으로 지식분야, 이론, 과정, 기법 등을 포함하고, 대상은 자연계에 존재하는 인간이외의 모든 생물 또는 무생물이 된다. 사건은 행위나 일어난 일로 시대, 시기 등과 같은 광범위한 기간도 포함되며, 장소는 현재의 지역뿐만 아니라 역사적인 지역, 지리적 특성도 포함된다.¹⁵⁾

영상물의 소통구조 분석을 위해서 선행연구에서 언급한 서사물의 이야기와 담론에 대한 분석은 물론, 소통행위자, 소통의 맥락(시간과 공간), 채널, 매체, 수용자에 관한 정보도 포함되어야 한다. 즉, 영상물의 내용 요약(초록 작성)을 위해서는 핵심적인 내용을 나타내는 사건, 인물, 배경이 담긴 프레임(이미지)을 추출하고, 영상 서사의 담론적 특성을 표현하기 위해서는 서술 행위가 드러나는 프레임 [예, 서사내 인물들의 대화나 독백; 서사내 해설, 카메라의 묘사(암시적 묘사, 명시적 묘사); 서사의 작가의 서술 행위(카메라워크, 화면편집, 배경음악, 화면밖 목소리)]들을 포함시키면 될 것이다.

앞에서 언급한 세가지 이론을 기초로 하여 영상 초록 배열 모형을 크게 세 요소 즉, 사건, 행위자(인물, 대상), 배경으로 구성하였다. 구체적으로 사건에는 핵심 사건과 주변 사건이 있고, 행위자에는 핵심 및 주변 행위자가 있다. 배경에는 다양한 요소들을 포함시켰는데, FRBR(Functional Requirements for Bibliographic Records) 제3집단의 개념의 일부(이론, 과정 및 기법)를 배경 요소에서 표현하기로 하였다. 이외에 사건들이 행위자에 의해서 수행되는 배경 정보인 시간과 공간을 포함시켰다. 또한 핵심 사건(또는 주제)을 기술하는 도구나 방법 그리고 비디오 내용에서 나레이터(해설자)가 있는 경우에는 배경 요소에서 표현하기로 하였다. 서사 형식, 카메라의 묘사, 서사의 작가의 서술행위 등 표현하기 어려운 것들은 메타데이터 요소로 수용할 수 있을 것이다. 영상 초록 배열 모형을 개념적으로 표현하면 <그림 1>과 같다.



<그림 1> 배열 모형

15) IFLA, *Functional Requirements for Bibliographic Records: final report*. UBCIM Publications, New Series Vol.19(Munich: K. G. Saur, 1998), pp.82-87.

2. 평 가

가. 실험 설계

(1) 비디오 표본 선정

비디오 표본 자료의 장르에 따라서 연구 결과가 달라질 수 있기 때문에 영상으로 의미를 전달하는 데 큰 문제가 없는 다큐멘터리 비디오를 선택하였고, 재생시간에 따라 영상 초록의 기능이 달라질 수 있기 때문에 2분 36초~10분 사이의 재생시간을 가진 비디오를 선정하였다. 표본 자료는 노스캐롤리나 대학의 문헌정보학과에서 구축한 Open Video Digital Library(OVDL, www.open-video.org)에 올려진 비디오 자료에서 이미지에 주로 의존하여 정보를 전달하는 총 6건을 선정하였다. OVDL 비디오 자료의 영상 초록은 하이브리드 방식을 통해서 구성된 것이다. 첫 단계에서 컬러 히스토그램과 같은 키프레임간의 차이를 이용하여 자동적으로 키프레임 후보 집단을 추출하고 추출된 키프레임 집단에서 최종적으로 사람들에 의해서 키프레임들이 선정되고 키프레임수는 최대 36개까지로 제한하였다. OVDL에 있는 키프레임의 크기가 너무 작아 판독하는데 어려움이 있어 본 실험에서는 비디오 플레이어인 KMplayer를 이용하여 OVDL 사이트의 영상 초록에 있는 키프레임대로 전체 클립에서 다시 추출하였다. 다만 본 연구는 이미지의 의미전달의 가능성을 조사하는데 초점을 맞추었기 때문에 OVDL 사이트에 있는 키프레임이라도 자막이 있는 키프레임들은 되도록 제외시켰다.

(2) 실험 시스템

(가) 초기 화면

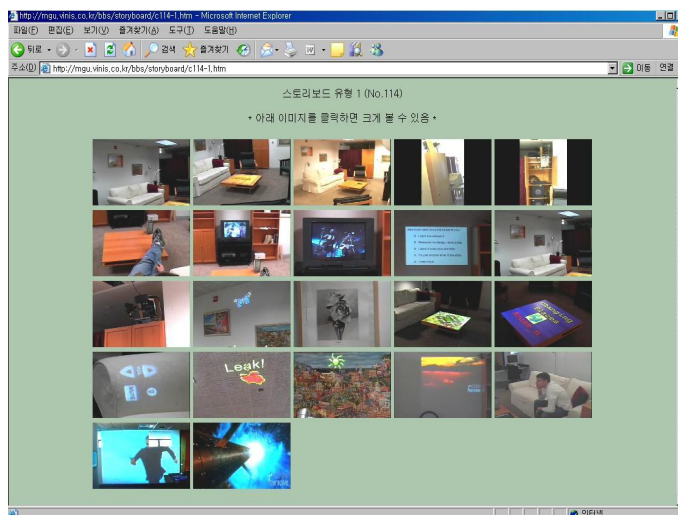
영상 초록 목록의 초기화면은 <그림 2>와 같다. 표제를 포함한 주제 파악에 증거가 되는 일체의 메타데이터를 제외시키고 오직 비디오코드만 보여주었다. 영상 초록 유형 중 어떤 유형을 먼저 사용하는냐에 따라서 가설 1과 2를 검증하기 위한 정확도에 영향을 받는다고 판단하여 순차적 영상 초록과 배열 모형 기반 영상초록을 번갈아 가며 먼저 볼 수 있도록 목록을 구성하였다. 즉, 선정된 6개의 비디오(테스트용 비디오 제외) 중 3개는 순차적 영상 초록을 본 후 배열 모형 영상 초록을 보게 하고 나머지 3개는 배열 모형 영상 초록을 본 후 순차적 영상 초록을 보게 하였다. 또한 이러한 제시 방법은 비디오 보기의 반복 효과를 측정할 때 어떤 유형의 비디오를 먼저 보았느냐에 따라 요약문 및 색인어의 정확도가 달라지는 것을 측정할 수 있어 나머지 두 개의 가설(가설 3과 가설 4)을 검증하는 자료로 이용된다.

비디오코드	스토리보드 유형 1	스토리보드 유형 2
테스트용: N0.103	유형 1	유형 2
테스트용: N0.110	유형 2	유형 1
N0.101	유형 1	유형 2
N0.104	유형 2	유형 1
N0.109	유형 1	유형 2
N0.113	유형 2	유형 1
N0.114	유형 1	유형 2
N0.115	유형 2	유형 1

〈그림 2〉 초기 화면

(나) 순차적으로 배열한 영상 초록

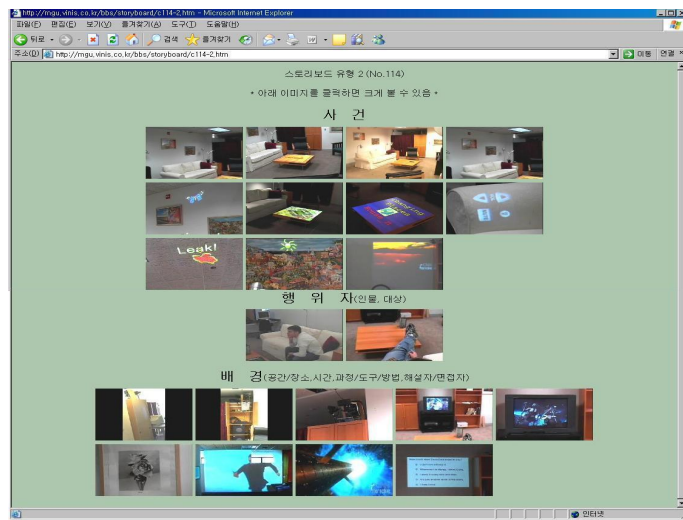
예로 든 표본 비디오는 '거실에서의 유비쿼터스'라는 표제를 갖고 있으며, 장소에 구애받지 않고 자유롭게 실생활 환경에서 컴퓨터 기능을 이용할 수 있다는 내용을 담고 있다. 즉, 쇼파에서도 컴퓨터 인터페이스를 구현할 수 있고, 텔레비전도 사람이 움직이는 것에 따라 화면이 움직이게 할 수 있다는 것으로 총 22개의 키프레임으로 구성되어 있다(그림 3 참조).



〈그림 3〉 순차적 영상 초록

(다) 배열 모형을 기반으로 한 영상 초록

앞에서 사용한 '거실에서의 유비쿼터스'의 영상 초록을 본 연구에서 제안한 사건, 행위자(인물, 대상), 배경으로 구성된 배열 모형에 따라 재구성한 것이 <그림 4>이다. '사건' 카테고리에는 유비쿼터스가 실제로 실행되어지는 모습들을 담은 키프레임들을 배치하고 '행위자' 카테고리에는 유비쿼터스를 직접 경험하고 있는 인물이 담겨있는 키프레임들을 포함시켰다. '배경' 카테고리에는 유비쿼터스가 실제적으로 적용이 가능한 집 내부의 부분들과 기기들을 포함한 이미지들을 포함시켰다.



<그림 4> 배열 모형 기반 영상 초록

(3) 설문지와 피조사자

설문지는 크게 두 종류로 구성된다. 첫 번째 설문지 유형 1은 순차적 영상 초록을 먼저 본 후 4개 이상의 색인어와 2 문장 이상의 요약문을 기록하게 하였고, 그 후에 배열 모형 기반 영상 초록을 보게 하여 색인어와 요약문을 기술하게 하였다. 두 번째 설문지 유형 2는 첫 번째 경우와 반대로 배열 모형 기반 영상 초록을 먼저 본 후 색인어와 요약문을 기록하게 하였다. 그리고 그 후에 순차적 영상 초록을 보게 하여 색인어와 요약문을 기술하게 하였다. 6개의 표본 비디오 중 3개의 비디오에 대해서는 설문지 유형 1을, 나머지 3개의 비디오에 대해서는 설문지 유형 2를 사용하였다.

피조사자는 학부 학생과 대학원생을 대상으로 하였다. 피조사자는 D대학의 신문방송학과 전공 학부 학생 8명 그리고 J대학의 문헌정보학과 전공 학부 학생 8명과 사서 전공 교육대학원생 10명, 총 26명으로 구성하였다. 실험은 두 대학의 컴퓨터실에서 따로 진행되었다. 실험에 대한 설명은

15분 정도 한 후 설문지를 배포하고 영상초록이 있는 웹사이트 주소를 알려주어 6개의 비디오에 대한 색인어 및 요약문을 기술하도록 하였다. 각 영상 초록(순차적 영상 초록 또는 배열 모형 영상 초록)에 대한 색인어와 요약문을 적는데 5분을 할당하였다. 즉, 1개의 비디오에 대해서 10분, 6개 비디오에 대해서 60분을 주어 총 실험시간은 약 75분 쯤 사용되었다.

(4) 색인어와 요약문의 정확도 측정 방법

Open Video Digital Library 사이트에 올려진 색인어와 텍스트 초록을 표준자료로 이용하고자 하였으나 색인어나 초록이 없거나 색인어와 초록이 있는 비디오도 대부분 색인어수가 너무 적거나 초록이 너무 짧아서 직접 전체 클립을 반복해서 듣고 표준 색인어와 요약문을 재구성하였다. 비디오에 4~20개의 표준 색인어와 두 문장 이상의 표준 요약문을 할당한 다음 정확도를 측정하였다. 색인어의 정확도 측정은 피조사자가 기술한 색인어가 표준 색인어 그룹에 있으면 0.25를 주었고, 제목과 같은 색인어를 할당한 경우에는 가중치를 주어 0.50을 주었으며, 그 합이 1을 넘는 경우는 측정치를 1로 제한하였다. 요약문의 경우 표준 초록과 내용이 어느 정도 일치하느냐에 따라서 0~1 사이의 값을 주었으며, 측정자의 주관적인 판단이 개입될 것을 우려해 두 명이 정확도를 측정하고 서로 논의하여 최종값을 결정하였다.

나. 데이터 분석과 가설 검증

(1) 순차적 및 배열 모형 기반의 영상 초록의 요약문과 색인어의 정확도 차이 분석

(가) 통계 분석: T-검증

순차적 영상 초록과 배열 모형 영상 초록의 색인어와 요약문 정확도 차이를 검증하기 위해 t-검증을 하였다. 분석결과, 요약문의 정확도는 배열 모형 영상 초록(0.47)이 순차적 영상 초록(0.41)보다 높게 나타났고 유의확률수준 0.01에서 통계적으로 유의미한 차이를 보였다. 색인어의 정확도는 배열 모형 영상 초록(0.47)이 순차적 영상 초록(0.46)보다 다소 높게 나타났으나 통계적으로 유의미한 차이를 보이지는 않았다(표 1 참조).

<표 1> T-검증(순차적 영상 초록 vs. 배열 모형 영상 초록)

기술 통계값(항목별 평균)					
요약(순차)		요약(배열)		색인어(순차)	
0.41		0.47		0.46	
				0.47	
t-검증 결과					
대응	T값	유의확률	대응	T값	유의확률
요약(순차) vs. 요약(배열)	-3.67	0.00**	색인어(순차) vs. 색인어(배열)	-0.46	0.64

(나)가설 검증: 연구 가설 1과 2

“가설 1: 순차적 영상초록과 배열 모형 기반 영상 초록의 요약문 정확도는 차이가 있을 것이다.”와 “가설 2: 순차적 영상초록과 배열 모형 기반 영상 초록의 색인어 정확도는 차이가 있을 것이다.”라는 가설들을 검증하기 위해서 t-검증을 한 결과, 가설 1은 검증되었으나 가설 2는 검증되지 못했다(표 1 참조). 본 연구에서 제안한 배열 모형은 비디오의 줄거리를 요약하는데 도움을 주는 것으로 나타났다.

(2) 영상 초록 반복보기의 유형 1과 유형 2간의 요약문 정확도 차이 분석

(가)통계분석: T-검증

영상 초록 반복보기의 유형 1(순차적 초록-)배열 모형 초록)과 유형 2(배열 모형 초록-)순차적 초록)간의 요약문의 최종 정확도의 차이를 분석하기 위해 t-검증을 하였다. 유형 1의 경우 요약문에서 최종 정확도가 0.51로 나타났고(0.42->0.51), 유형 2의 경우는 요약문의 최종 정확도는 첫 번째 정확도 보다 오히려 낮아져 0.40으로 나타났다(0.42->0.40)(표 2 참조). 이 두 유형의 최종 요약문 정확도는 유의확률 수준 0.01에서 통계적으로 차이가 있는 것으로 나타났다.

〈표 2〉 T-검증(유형 1의 요약문 정확도 vs. 유형 2의 요약문 정확도)

기술 통계값(항목별 평균)			
순차적 초록->배열 모형 초록(유형 1)		배열 모형 초록->순차적 초록(유형 2)	
요약(순차)	요약(배열)	요약(배열)	요약(순차)
0.42	0.51	0.42	0.40
t-검증 결과			
대응	T값	유의확률	
유형 1(최종 요약문 정확도) vs. 유형 2(최종 요약문 정확도)	3.02	0.00**	

(2)가설 검증: 연구 가설 3

“가설 3: 영상 초록 반복보기의 유형 1(순차적 초록-)배열 모형 초록)과 유형 2(배열 모형 초록-)순차적 초록)간의 요약문 정확도는 차이가 있을 것이다.”라는 가설들을 검증하기 위해서 t-검증을 한 결과, 유형 1의 요약문의 최종 평균 정확도가 0.51이고, 유형 2의 요약문의 최종 평균 정확도가 0.40으로 나타났다. 유의확률수준 0.01에서 이 두 유형의 최종 정확도는 통계적으로 차이가 있는 것으로 나타나 가설 3은 검증되었다(표 2 참조). 이러한 실험 결과는 사람들에게 순차적 영상 초록을 한번 보여준 후 구조화된 배열 모형 영상 초록을 다시 한번 보게 하는 것이 비디오의 줄거리를 훨씬 용이하게 파악할 수 있음을 보여주고 있고 이러한 결과는 앞으로 영상물 데이터베이스 구축에 참조될 수 있을 것이다.

(3) 영상 초록 반복보기의 유형 1과 유형 2간의 색인어 정확도 차이 분석

(가) 통계분석: T-검증

영상 초록 반복보기의 유형 1(순차적 초록->배열 모형 초록)과 유형 2(배열 모형 초록->순차적 초록)간의 색인어의 최종 정확도의 차이를 분석하기 위해 t-검증을 하였다. 유형 1의 경우 색인어의 최종 정확도가 0.48로 나타났고((0.45->0.48), 유형 2의 경우는 색인어의 최종 정확도는 0.47로 나타났다(0.46->0.47)(표 3 참조). 이 두 유형의 최종 색인어 정확도는 통계적으로 차이가 없는 것으로 나타났다.

<표 3> T-검증(유형 1의 색인어 정확도 vs. 유형 2의 색인어 정확도)

기술 통계값(항목별 평균)			
순차적 초록->배열 모형 초록(유형 1)		배열 모형 초록->순차적 초록(유형 2)	
색인어(순차)	색인어(배열)	색인어(배열)	색인어(순차)
0.45	0.48	0.46	0.47
t-검증 결과			
대응	T값	유의확률	
유형 1(최종 색인어 정확도) vs. 유형 2(최종 색인어 정확도)	0.04	0.97	

(나) 가설 검증: 연구 가설 4

“가설 4: 영상 초록 반복보기의 유형 1(순차적 초록->배열 모형 초록)과 유형 2(배열 모형 초록->순차적 초록)간의 색인어 정확도는 차이가 있을 것이다.”라는 가설을 검증하기 위해서 t-검증을 한 결과, 이 두 유형의 최종 색인어 정확도는 통계적으로 차이가 없는 것으로 나타나 가설 4는 검증되지 못했다(표 3 참조).

3.논 의

제안된 배열 모형 영상 초록이 순차적 영상 초록 보다 비디오 내용 파악에 있어서 더 유용한 것으로 나타났다. 특히 초록의 유형을 달리하여 반복하여 브라우징하는 것 즉, 순차적 영상 초록을 먼저 본 후 배열 모형 영상 초록을 보면 요약문의 정확도가 현저하게 증가하는 것을 확인할 수 있었다. 또한 순차적 초록 또는 배열 모형 기반 초록을 통한 비디오 의미 파악의 정확도가 각 비디오 오에 따라 다를 수 있음을 확인할 수 있었다. 실험 결과를 자세히 분석해 본 결과, 순차적 영상 초록이 배열 모형 초록 보다 더 많은 정보를 주는 경우도 있었다. 순차적 영상 초록을 보여 주고 그 다음으로 배열 모형 기반 영상 초록을 보여 준 “최초의 비행기”에 관한 표본 비디오의 경우

한 피조사자의 처음 정확도가 0.50이었는데 배열 모형 기반 영상 초록을 본 이후에 정확도가 0.25로 감소하였다.

반대로 배열 모형의 구조에 맞는 비디오의 경우 배열 모형 영상 초록을 통해서 의미 파악이 훨씬 용이했음을 보여주는 사례도 있었다. 예를 들어서, 배열 모형 영상 초록을 보여 주고 그 다음으로 순차적 영상 초록을 보여 준 “지진 발생시 주의사항”에 관한 표본 비디오에서 한 피조사자의 처음 요약문 정확도가 0.70이었는데 두 번째 순차적 영상 초록을 본 이후에 정확도가 증가하는 대신 0.30으로 하락하는 경우도 생겨났다. 이는 반복해서 보여 주었음에도 불구하고 순차적 영상 초록에서 의미 파악이 어려웠다는 것을 보여 준다.

앞에서 언급한 것처럼 순차적 영상 초록을 먼저 본 후 배열 모형 기반 영상 초록을 보면 요약문의 정확도가 현저하게 증가하는 것을 확인할 수 있었다. 또한 각 비디오에 따라 배열 모형 초록이 더 많은 정보를 주거나 반대로 순차적 초록이 더 많은 의미를 전달하는 경우가 있음을 확인할 수 있었다. 즉, 비디오 내용이 사건의 순서 중심일 경우엔 순차적 초록이 더 많은 정보를 주며, 지식(내용)중심의 비디오의 경우는 배열모형 초록이 더 많은 정보를 제공하는 것으로 나타났다.

이러한 실험 결과를 바탕으로 디지털 도서관 환경에서 영상 초록의 활용 방안들을 제안하면 다음과 같다.

첫째, 앞의 실험 결과에서 나타난 두 유형의 비디오 특성들을 감안하여 본 연구에서는 초기값으로 순차적 영상 초록을 제시하고 반복 보기 버튼을 클릭하면 배열 모형 기반 영상 초록이 함께 출력되는 비디오 브라우징 인터페이스를 제안한다. 이와 같이 배열 모형 영상 초록을 순차적 영상 초록의 대안으로 주장하기 보다는 이 두 유형의 영상 초록을 모두 수용하는 방안을 제시하는 이유는 대부분의 이용자들이 영상 초록을 한번만 보는 것이 아니라 반복 보기를 하고 있는 상황에서 비디오에 대한 좀 더 정확한 의미 추출을 위한 좋은 방안으로 생각하기 때문이다. 더 나아가 배열 모형 영상 초록을 단독으로 볼 때의 요약문 정확도(0.42)에 비해서 순차적 영상 초록을 먼저 보고 배열 모형 기반 영상 초록을 보았을 때 요약문 정확도(0.51)가 매우 큰 차이로 증가하기 때문이다.

둘째, 텍스트 기반의 데이터베이스에서 텍스트 질의가 들어오면 문헌의 전체 원문을 분석하기 보다는 메타데이터에 기초한 색인 파일과 매칭시키는 것과 같은 원리로 제안된 배열 모형 영상 초록은 영상 질의의 매칭 자료로 활용한다. 즉, 영상 질의로 이미지(프레임)가 하나 또는 복수개가 사용될 수 있는데 이러한 이미지들을 제안된 배열 모형에 따라 구성한다면 더 정확한 검색 결과를 기대할 수 있다. 예를 들어서 영상 질의로 사용되는 두개의 이미지들이 배열 모형의 사건과 배경에 각각 해당된다면 구조화된 질의에서 각 질의 이미지를 사건과 배경에 할당되는 영상 초록의 키프레임들과 매칭시킬 때 사건(또는 배경)으로 구분된 키프레임들만을 분석하여 동일한 또는 유사한 이미지를 확인해 봄으로써 탐색 시간을 줄임과 동시에 탐색의 정확도도 높일 수 있을 것이다.

IV. 결론과 후속 연구

최근 이용자의 정보 요구는 높아가고 있지만 정보 현장의 멀티미디어정보의 검색 시스템은 이에 따라 가지 못하고 있는 것이 현실이다. 전통적으로 비디오 자료는 구조화된 메타데이터와 함께 수작업으로 만든 해제에 의존하여 검색된다. 즉, 비디오는 수작업으로 샷단위로 구분되고 각 샷별로 텍스트 해제가 훈련된 전문 사서에 의해서 완성된다. 이와 같이 수작업으로 작성한 해제와 구조화된 메타데이터의 결합이 이용자 검색의 기초가 되어 오고 있다. 그러나 이러한 텍스트 기반 검색은 특히 비디오의 내용 분석 작업이 수작업으로 수행될 때 급속도로 증가하고 있는 비디오 자료를 처리하는데 어려움이 있을 뿐만 아니라 비디오 자료의 구체적인 내용을 제시하지 못하는 문제가 생길 수 있다.

본 연구에서는 영상 초록의 키프레임 배열 모형을 영상소통 이론, FRBR 모형, 용어열 색인 이론 등을 이용하여 설계하고 이 모형에 따라서 영상 초록을 구성한 후 순차적 영상 초록과 요약문과 색인어 추출의 정확도 측면에서 어떤 차이를 보이는지 실험을 통해서 조사해 보았다. 실험 결과, 배열 모형 초록이 순차적 초록 보다 더 정확한 요약문을 추출하게 하는 것으로 나타났다. 또한 영상 초록의 반복 보기의 효과를 측정한 결과 순차적 영상 초록을 먼저 보고 배열 모형 영상 초록을 반복해서 본 경우에 요약문의 정확도가 크게 증가하는 것으로 나타났다.

이러한 실험 결과에 기초하여 디지털 도서관 환경에서 영상 초록의 두 가지 활용 방안 즉, OPAC 환경에서 영상 초록의 초기값으로 순차적 영상 초록을 제시하고 반복 보기를 원하는 경우 배열 모형 기반 영상 초록이 함께 출력되는 비디오 브라우징 인터페이스를 구성하는 것과 제안된 배열 모형 영상 초록을 영상 질의의 구조화된 매칭 자료로 활용하는 것에 대해서 제안하였다.

본 연구는 영상 초록의 키프레임 순서가 영상자료의 의미 파악에 어떤 영향을 미치는지에 대한 기초 연구로 많은 후속 연구들이 필요하다고 생각한다. 특히 본 연구를 통해서 내린 결론은 영상 초록의 키프레임 배열 모형도 비디오 장르에 따라서 다른 모형이 제시되어야 한다는 것이다. 영상소통 구조에 기반하여 제안한 본 연구의 키프레임 배열 모형은 다른 비디오 장르 특히 드라마 장르에 적용하면 더 좋은 효과가 나올 것으로 기대한다. 또한 다큐멘터리 자료에 더 적합한 배열 모형을 완성시키기 위해서 많은 다큐멘터리 비디오 자료에 대한 내용 분석과 함께 현재 단순히 사진, 행위자, 배경으로 되어 있는 단순한 구조에서 좀 더 정교한 배열 모형으로 다듬어야 하며 각 카테고리 키프레임(들)을 할당하는 좀 더 자세한 방법들에 대한 후속 연구도 필요하다고 본다. 끝으로 영상 초록이 비디오의 내용 전달을 위한 효율적인 도구가 되기 위해서는 무엇보다도 전체 클립에서 의미 있는 키프레임들을 추출하는 과정이 가장 중요하다고 생각한다. 따라서 공학자들에 의해서 연구된 이미지 프로세싱에 의해서 키프레임들을 추출하여 순차적 영상 초록을 구성하고 다른 한편 적합성 개념에 기초하여 세밀하게 설계된 키프레임 추출 알고리즘에 의해서 선정된 키프레임

16 한국도서관·정보학회지(제38권 제4호)

들을 정교한 배열 모형에 따라 재구성한 후 이 두 방법들이 비디오 의미 추출에서 어떤 차이를 보이는지에 대한 연구가 필요하다고 생각한다.

〈참고문헌은 각주로 대신함〉

〈부록〉 표준 색인어와 요약문

다음은 6개의 비디오 샘플 자료 중 한 비디오에 대한 표준 색인어 집단과 요약문이다.

■ 색인어 그룹

- 1) NASA Connect
- 2) 라이트 형제(Wright Brothers)
- 3) 비행 시뮬레이터(Flight Simulator)
- 4) 축(Axis)
- 5) 최초 비행기(First Flying Machine)*(0.5)
- 6) 글라이더(Glider)
- 7) Kitty Hawk(미국의 디젤엔진 추진의 초대형 항공모함을 의미하나 키티호크라는 이름은 1903년 12월 라이트형제가 세계 최초의 동력비행기를 타고 비행에 성공한 노스캐롤라이나 주의 작은 마을 이름에서 따온 것이다.)
- 8) 공학 기법(Engineering Method)
- 9) 비행기가 위아래로 움직임(Pitch)
- 10) 비행기 회전(Roll)
- 11) 비행기가 좌우로 이동(Yaw)
- 12) Wind Tunnel(풍동(風洞), 항공기의 모형이나 부품을 시험하는 통모양의 장치)
- 13) 엘리베이터 제어(Elevator Control)
- 14) 비행기 날개 휨(Wing Warp)
- 15) 비행기 방향타(Rudder)
- 16) 프로펠러(Propellor)
- 17) 날개(Wing)
- 18) 왼쪽(Lift)
- 19) 항력(Drag)
- 20) 라이트 경험(Right Experience)

■ 요약문

라이트 형제가 비행기(초기 글라이더)를 설계하고 개발하는 과정을 보여 주고 있다. 나레이터

(제니퍼)가 버지니아 주에 있는 “라이트 경험(Wright Experience)” 연구소의 책임자(Ken)를 만나서 라이트 형제의 글라이더에 대한 설명을 듣는데 라이트 형제는 글라이더의 문제점들을 하나 하나 해결해 가는 방법을 택했다고 한다. 또한 책임자는 제니퍼에게 이 연구소에 만든 라이트 글라이더의 모형을 보여 준다. Ken은 이 연구소에서 만든 모형과 라이트 형제의 1902년도 글라이더와의 가장 큰 차이는 1902년도 글라이더는 비행기가 위아래로 움직이고(pitch), 회전하고(roll) 그리고 좌우로 이동하는(yaw) 기능을 갖고 있지 못했다고 설명한다. 그리고 이 세 가지 기능에 대해서 좀 더 자세히 설명한다. 장소를 옮겨서 이 두 사람은 비행 시뮬레이션 사무실로 이동한다. 사무실에 있는 연구원이 제니퍼에게 비행기 시뮬레이터가 “wind tunnel”에 기초하여 만든 것이라고 설명한다. 제니퍼와 연구원은 시뮬레이터를 통해서 비행을 가상 경험해 보고 그 사이 연구원은 제니퍼에게 사용법을 설명한다. 다시 Ken은 제니퍼에게 1903년도 Kitty Hall이라고 불리는 비행기 모형을 가상 경험해 보지 않겠느냐고 제안하면서 비디오는 끝난다.