

비디오 자료의 의미 추출을 위한 영상 초록의 효용성에 관한
실험적 연구*

An Experimental Study on the Effectiveness of Storyboard Surrogates in
the Meanings Extraction of Digital Videos

김 현 희 (Hyun-Hee Kim)**

*이 논문은 2006년도 정부재원(교육인적자원부 학술연구조성사업비)으로 한국학술진흥재단의 지원을 받아 연구되었음(KRF-2006-321-H00031).

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

초 록

본 연구에서는 12개의 표본 비디오 집단과 14명의 피조사자들을 이용하여 영상 초록 및 전체 클립 보기를 통한 색인어 및 요약문 추출의 정확도를 측정해 보았다. 측정 결과 첫째, 비디오 유형에 따라 정확도가 차이가 있는 것으로 나타났으며 이는 이미지에 주로 의존하여 정보를 표출하는 비디오의 경우 텍스트 초록만으로 의미 파악을 하기에는 한계가 있으며 텍스트 초록이 영상 초록과 함께 사용되었을 때 시너지 효과를 낼 수 있음을 보여주고 있다. 둘째, 영상 초록의 색인어 및 요약문 정확도가 전체 클립의 정확도 보다 떨어지지만 절반치에 근접한 것으로 나타나 영상 초록이 비디오 의미 추출에 효율적으로 활용될 수 있음을 확인하였다. 또한 영상 초록의 색인어 정확도(0.45)가 요약문 정확도(0.40) 보다 더 높게 나타나 영상 초록을 통해서 색인어 추출 작업을 더 효율적으로 할 수 있음을 확인할 수 있었다. 이러한 실험 결과에 기초하여 영상 초록이 색인어 또는 요약문 추출 작업에 활용될 수 있을 뿐만 아니라, 디지털 도서관 환경에서 텍스트 초록과 같은 다른 메타데이터 요소들과 함께 사용된다면 이용자의 적합성 판정을 좀 더 용이하게 할 것이며, 더 나아가 영상 질의의 매칭 자료로도 이용될 수 있음을 제안하였다. 끝으로 영상 초록의 품질을 높이기 위한 키프레임 추출 알고리즘 및 키프레임 배열 모형 설계 등 후속 연구에 대해서 제안하였다.

ABSTRACT

This study is designed to assess whether storyboard surrogates are useful enough to be utilized for indexing sources as well as for metadata elements using 12 sample videos and 14 participants. Study shows that first, the match rates of index terms and summaries are significantly different according to video types, which means storyboard surrogates are especially useful for the type of videos of conveying their meanings mainly through images. Second, participants could assign subject keywords and summaries to digital video, sacrificing a little loss of full video clips' match rates. Moreover, the match rate of index terms (0.45) is higher than that of summaries (0.40). This means storyboard surrogates could be more useful for indexing videos rather than summarizing them. The study suggests that 1) storyboard surrogates can be used as sources for indexing and abstracting digital videos; 2) using storyboard surrogates along with other metadata elements (e.g., text-based abstracts) can be more useful for users' relevance judgement; and 3) storyboard surrogates can be utilized as match sources of image-based queries. Finally, in order to improve storyboard surrogates quality, this study proposes future studies: constructing key frame extraction algorithms and designing key frame arrangement models.

키워드: 디지털 비디오, 영상 초록, 스토리보드, 주제 색인, 요약, 텍스트 초록, 영상 질의, 키프레임, 메타데이터, 통보적 초록, 지시적 초록, 비디오 클립, 디지털 도서관, Open Video Digital Library, storyboard surrogates

1. 서 론

1.1 연구 필요성과 목적

최근 방송국, 대학 등의 정보를 생산, 보유하는 기관의 정보 관련 부서들의 환경이 급변하고 있다. 방송국 정보시스템의 경우 디지털 방송의 출현으로 방송 정보 환경에 커다란 변혁을 초래하고 있으며, 대학의 경우, 디지털 도서관의 등장으로 텍스트를 포함한 전자책, 영상강의 자료 등 멀티미디어 정보가 디지털 콘텐츠로 구축되고 있다. 특히 이들 기관의 정보 시스템에 저장된 멀티미디어 콘텐츠는 영상 및 음성 정보 뿐만 아니라 문자 정보를 비롯한 여러 의미 정보를 내포하고 있다. 또한 시각적 관계, 시간적 관계, 공간적 관계, 비정형적이고 대용량적 관계 등의 특징 정보들을 내포하고 있으므로 기존의 텍스트 기반의 질의, 메타데이터(초록, 표제 등)만으로 이용자들이 적합한 자료를 찾는 것이 어려워지고 있다.

비디오 자료에 접근하기 위해서 다양한 비디오 대체물 즉, 메타데이터 요소들(예, 표제, 키워드, 초록 등)이 사용되고 있다. 그 중 비디오의 간명한 표현을 보여 주면서 주요 메시지를 나타내는 비디오 초록은 전체 비디오 자료를 보지 않고도 적합한 자료인지를 판단하게 해 주는 핵심 메타데이터 요소로 알려져 있다(Mu and Marchionini 2003). 이제까지 방송국 자료실이나 대학 도서관에서 영상 콘텐츠의 초록을 구성하는 방법을 살펴보면 수작업으로 비디오 클립을 장면 단위로 구분하고 각 장면별로 텍스트 초록을 부여하거나 또는 이러한 구분 없이 전체 비디오 클립에 대한 텍스트 초록을 부여하고 있는 것이 일반적이다. 그러나 이러한 내용 분석 방식은 특히 모든 과정이 수작업으로 수행될 때 급속도로 증가하고 있는 비디오 자료를 처리하는데 어려움이 있을 뿐만 아니라 영상 자료의 구체적인 내용을 제시하지 못하는 제한점을 가지고 있다.

한편 공학 분야 연구자들은 이미지 프로세싱에 초점을 맞춰서 비디오 콘텐츠의 영상 초록을 작성하는 연구를 진행하고 있다. 즉, 이미지의 색, 질감, 객체를 확인하거나 다이어로그, 서술과 같은 음성 정보를 인식하여 비디오 자료들을 샷별로 분할하고 분할된 콘텐츠에서 키프레임을 추출하고 추출된 키프레임들을 순서대로 나열하여 영상 초록을 구축하고 있어 앞에서 언급한 텍스트 초록의 제한점들을 보완하고 있다. 현 검색 환경에서 비디오 초록은 텍스트(텍스트 초록)이거나 이미지(영상 초록) 또는 이 두 개의 결합이 될 수 있다. 특히 영상 초록은 영상 질의를 사용하여 영상 자료를 검색할 때 영상 질의의 매칭 자료로 활용되어 검색 효율성을 높일 수 있다. 이는 전통적인 텍스트 기반 검색에서 검색어가 텍스트 자료 전체와 비교되는

것 대신 표제, 초록 등 자료의 일부와 비교되는 원리와 유사하다.

본 연구는 비디오 자료의 의미 추출을 위한 영상 초록의 효용성을 조사, 평가한 후 영상 초록이 색인어 및 요약문 추출 작업에 활용될 수 있을 뿐만 아니라, 디지털 도서관의 OPAC 시스템의 메타데이터의 한 요소로 활용되어 적합성 판정을 위한 이용자의 비디오 브라우징 및 영상 질의가 이용되는 비디오 검색 환경에 이용될 수 있음을 제안하여 영상 자료의 색인과 검색을 좀 더 용이하게 하는 것을 목적으로 한다.

1.2 연구 방법

비디오 자료의 의미 추출을 위한 영상 초록의 효용성을 조사하기 위해서 5개의 연구 가설과 2개의 연구 문제를 설정하고 이를 검증하기 위해서 표본 비디오 12개를 이용하여 영상 초록과 전체 클립을 볼 수 있는 웹사이트를 구축하였다. 본 연구에서는 이미지의 의미전달의 가능성을 조사하는데 초점을 맞추었기 때문에 자막이 있는 키프레임들을 영상 초록에서 되도록 제외시켰다. 구성된 영상 초록과 전체 클립 보기를 통한 색인어 및 요약문 추출의 정확도를 측정하기 위해서 피조사자 14명을 M대학의 문헌정보학 전공 또는 복수 전공 학생 중에서 임의로 추출하였다. 선정된 피조사자 14명을 컴퓨터실에서 작업을 하도록 하였다. 즉, 피조사자에게 설문지를 배포하고 각 영상 초록에 대해서 색인어를 4개 이상 할당하도록 하고, 비디오 요약문은 두 문장 이상 기술하도록 하였다. 그런 다음 사운드가 제거된 전체 클립을 보게 한 후 영상 초록과 똑같이 색인어와 요약문을 기술하도록 하였고 이외에 영상 초록의 기능과 영상 초록과 전체 클립의 장단점에 대한 추가적인 질문을 포함시켰다. 총 실험시간은 실험에 대한 안내를 포함하여 약 3시간 30분이 소요되었다.

색인어와 요약문의 정확도 측정은 표본 비디오에 대한 표준 색인어와 요약문을 구성한 후 이를 피조사자가 작성한 색인어 그리고 요약문과 대조한 후 어느 정도 일치하느냐에 따라 0~1 사이의 값을 주어서 측정하였다. 측정자의 주관적인 판단이 개입될 가능성이 높기 때문에 두 명이 정확도를 측정하고 그 평균값을 활용하였다. 이러한 측정값을 기초로 하여 일원배치 분산분석과 사후검정, t-검증 등의 통계기법을 활용하여 가설 검증을 하였다.

1.3 용어 정의

- 1) 비디오 샷: 동영상의 최소단위이다. 하나의 비디오 시퀀스는 비디오 클립, 장면, 샷, 프

레이프 등의 요소로 구성된다. 하나의 비디오 클립은 하나 이상의 비디오 장면으로 구성되고, 비디오의 각 장면은 하나의 주제를 갖는 연속된 일련의 비디오 샷으로 구성된다. 즉, 비디오의 각 장면은 비슷한 의미를 내포하고 있는 특정 대상을 연속 촬영한 일련의 샷들의 그룹으로 구성되는 비디오의 단위이다. 샷은 비디오 분할의 기본 단위로 사용된다.

2)컷: 장면 전환이란 용어로도 표현되며, 샷과 샷의 경계를 의미함과 동시에 장면과 장면의 경계이기도 하다.

3) 비디오 클립: 동영상의 단위로 하나의 비디오 시퀀스는 비디오 클립, 장면, 샷, 프레임 등의 요소로 구성되며, 하나의 비디오 클립은 하나 이상의 비디오 장면으로 구성된다.

4) 영상 초록: 비디오의 전체 내용을 이해할 수 있도록 비디오를 샷 단위로 분할하고 각 샷을 대표하는 키프레임들을 나열한 것을 지칭한다. 스토리보드라 지칭하기도 한다.

2. 이론적 배경

2.1 선행 연구

비디오의 영상 초록과 영상 질의에 대한 선행 연구를 살펴보면 다음과 같다. Christel 등(1997)은 여러 유형의 비디오 대체물을 확인하였는데 이는 키프레임과 영상 초록(스토리보드)이다. 그들은 키프레임은 비디오를 나타내는 개별적인 이미지이며 텍스트 문헌에 할당된 키워드와 유사한 반면, 영상 초록은 비디오로부터 추출하여 순서대로 나열한 일련의 키프레임들로 구성된다고 하였다.

먼저 영상(또는 이미지) 및 텍스트 초록의 특성 및 차이 그리고 이들을 동시에 사용했을 때의 시너지 효과를 보고하는 연구들이다. Hughes 등(2003)은 텍스트가 비디오가 무엇에 관한 것인지(what the video was about)에 대한 정보를 전달하는 것인데 반해, 이미지는 비디오가 무엇인지(what the video was like)에 관한 정보를 전달하는 것으로 특히 이미지 정보는 적합성 판단을 확실하게 할 때 자주 이용되고 있다고 제안하였다. Wildemuth 등(2002)은 텍스트는 적합성을 결정하는 과정에서 도움을 줄 수 있는 반면 비텍스트 대체물(surrogates)은 텍스트 대체물을 효과적으로 확대하는데 사용된다고 주장한다. Robertson 등(1998)은 웹페이지에서 작은 이미지(thumbnails)의 이용에 대한 연구에서 작은 이미지와 텍스트의 결합은 탐색

에서의 착오를 줄인다고 보고하였다. Cristel와 Warmack (2001)은 뉴스 비디오의 내비게이션을 위해서 텍스트 초록과 함께 사용된 영상 초록이 텍스트 없이 사용된 경우 보다 작업 시간이 더 짧았고 이용자의 선호도는 더 높았다고 보고한다. Li 등(2000)은 비디오의 장르에 따라서 샷이미지의 유용성이 달라질 수 있음을 제안하였는데, 샷이미지가 뉴스, 여행 및 스포츠에 가장 유용하게 이용되는 반면 서술(narration)을 통해서 주로 정보를 전달하는 강의 및 회의 자료는 유용성이 가장 낮았다고 보고하였다.

Robertson 등(1998)은 이미지와 텍스트의 결합은 탐색에서의 착오를 줄인다고 보고하였다. 국내 연구로 정보 검색을 다루었지만 텍스트와 이미지 정보의 특성을 제안한 연구로 고영승과 주경수 (2005)의 연구가 있다. 이들은 상위-레벨 메타데이터 요소들(텍스트 기반 검색)을 이용하면 원하는 영상을 빠르게 검색할 수 있고, 하위-레벨 메타데이터 요소들(이미지 형태를 중심으로 한 내용 기반 검색)을 이용하면 확실한 결정을 내리기 어려운 영상을 판정하는데 도움을 받을 수 있으며, 이 두 가지 요소들을 이용하면 이것들의 장점들을 모두 취할 수 있다고 제안하였다. 이러한 연구들은 텍스트와 영상 메타데이터를 동시에 사용하였을 때 비디오 자료의 내용 파악이 훨씬 용이하다는 점을 보여주고 있다.

다음은 비디오 색인 작업의 정보원으로 영상 초록을 사용했을 때의 효율성을 다루는 연구이다. Stachowicz (2002)는 6개의 표본 비디오를 활용하여 영상 초록을 사용하여 색인작업을 한 경우 전체 클립을 사용한 경우 보다 대략 82% 정도 시간을 절약할 수 있으면서 검색성은 전체 클립에 비해 단지 6% 정도 낮아진다고 주장하였다. 즉, 그녀는 비디오 주제 색인을 하기 위해서 비디오 전체를 체크하는 대신 자동적으로 추출한 영상 초록만을 체크하는 것이 검색 효율성을 거의 저하시키지 않으면서 신속한 색인 작업을 수행할 수 있음을 실험적으로 증명하였다. 이는 전통적으로 텍스트 문헌의 색인어를 추출할 때 문헌 전체를 분석하기 보다는 표제와 초록 등 원문 대체물을 분석하는 것과 같은 원리이다. Lyer와 Lewis (2006)는 세 개의 표본 비디오를 이용하여 영상 초록 보기와 전체 클립 보기를 통한 색인어 추출의 정확도를 측정하였다. 영상 초록의 정확도는 56%인데 반해서 전체 클립의 정확도는 89%로 나타났다. 또한 그들은 비디오의 영상 초록을 구성할 때 키프레임이 비디오에서 나타난 순서대로 배열하는 방법 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 않았다.

비디오 검색 환경에서 텍스트 질의 대신 영상 질의가 이용될 수 있다. Smeaton과 Browne

(2006)는 키프레임 이미지 매칭, 텍스트 탐색, 샷단위의 브라우징 및 객체 기반 검색이 가능한 비디오 검색 시스템을 소개한다. 그들은 비디오 검색에서 영상 질의는 이용자 정보 요구가 일단 정지 이미지로 표현되어야 하고 키프레임이 샷의 내용을 잘 나타냈을 때에만 성공적으로 적용될 수 있다고 제안한다(Browne and Smeaton 2005). IBM에서 개발한 QBIC 시스템은 색, 질감, 모양 등의 특징 조합을 이용한 영상 질의를 지원하는데 질의 방식으로는 팔레트에서의 검색, 예제 기반 검색, 모양 그려서 검색 등이 제공되고 있다(Niblack et al. 1997).

이와 같이 많은 연구들이 비디오 영상 초록이 지시적 또는 통보적 초록의 기능을 한다고 가정하고, 비디오의 영상 초록이 전통적인 문헌에서 초록이 하는 것처럼 똑같은 기능을 하고 있으며 또한 비디오의 키프레임이 키워드와 유사한 기능을 하고 있다고 보고하고 있다. 이러한 연구 결과를 검증하기 위해서 다양한 연구 방법 또는 큰 표본 집단을 이용한 후속 연구가 필요하다.

2.2 연구 문제와 가설

본 연구에서는 다음의 다섯 가지 가정에 대해서 검토하고자 한다. 첫째, 같은 장르에 속해 있는 비디오도 비디오의 유형에 따라서 영상 초록 보기를 통한 색인어 및 요약문 추출의 정확도가 차이가 나는지 알아보하고자 한다. 둘째, 영상 초록의 각 키프레임이 키워드의 역할을 하며 이들 키프레임들이 문맥적으로 연결되어 있지 않고 순차적으로 연결되어 있기 때문에 색인어 정확도가 요약문 정확도 보다 더 높게 나타날 것으로 예측해 본다.

셋째, 비디오 영상 초록과 사운드가 제거된 전체 비디오 클립이 비디오 주제 파악에 어떤 차이를 보이는지 알아보하고자 한다. 즉, 키프레임만을 보여주는 경우와 모든 프레임들을 보여주는 경우의 차이가 무엇인지 조사하고자 한다. 영상 초록의 키프레임이 키워드와 유사한 기능을 하고 있기 때문에 영상 초록과 사운드가 제거된 비디오 클립간의 색인어 추출의 정확도에는 크게 차이가 없을 것이라고 가정한다. 다만 텍스트 초록과 달리 영상 초록의 키프레임들이 문맥적으로 연결되어 있는 것이 아니기 때문에 전체 내용 파악(요약)의 정확성은 비디오 클립이 영상 초록 보다 더 높을 것으로 가정한다. 넷째, 영상 초록이 지시적 초록의 역할을 하는지 아니면 통보적 초록의 역할을 하는지 알아보하고자 한다. 끝으로, 영상 초록 보기와 전체 클립 보기의 장단점들이 무엇인지 조사하고자 한다. 이러한 다섯 가지 이슈에 기초하여 다음과 같은 다섯 개의 연구 가설과 두 개의 연구 문제를 구성하였다.

2.2.1 연구 가설

연구가설은 다음과 같이 다섯 개로 구성되어 있다.

- 1) 비디오 유형에 따라 요약문 정확도가 차이가 있을 것이다.
- 2) 비디오 유형에 따라 색인어 정확도가 차이가 있을 것이다.
- 3) 영상 초록의 색인어 정확도와 요약문 정확도는 차이가 있을 것이다.
- 4) 사운드가 제거된 전체 비디오 클립의 색인어 정확도와 영상 초록의 색인어 정확도는 차이가 없을 것이다.
- 5) 사운드가 제거된 전체 비디오 클립의 요약문 정확도와 영상 초록의 요약문 정확도는 차이가 있을 것이다.

2.2.2 연구 문제

두 개의 연구문제는 다음과 같다.

- 1) 영상 초록은 지시적 초록의 역할을 하는가? 아니면 통보적 초록의 역할을 하는가?
- 2) 비디오 자료의 의미 추출을 위한 영상 초록 보기를 전체 클립 보기와 비교하면 어떤 장점과 단점을 가지고 있는가?

3. 영상 초록의 효용성 조사와 분석

3.1 실험 설계

3.1.1 비디오 표본 선정

비디오 표본 자료의 장르에 따라서 연구 결과가 달라질 수 있기 때문에 영상으로 의미를 전달하는데 큰 문제가 없는 다큐멘터리 비디오를 선택하였고 재생시간에 따라 영상 초록의 기능이 달라질 수 있기 때문에 2분 17초에서 10분 사이에 있는 비디오를 선정하였다. 표본 자료는 노스캐로리나 대학의 문헌정보학과에서 구축한 Open Video Digital Library(OVDL, www.open-video.org)에 올려진 비디오 자료에서 총 12건을 선정하였다(Marchionini and Geisler 2002). 비디오 유형에 따른 색인어 및 요약문 추출의 정확도 차이를 알아보기 위해서 12개의 표본 비디오를 주로 이미지에 의존하여 내용을 전달하는 4개의 비디오, 서술을 통해서 의미를 전달하는 4개의 비디오 그리고 이미지와 서술을 모두 사용하고 있는 4개의 비디오로 구성하였다.

OVDL 비디오 자료의 영상 초록은 하이브리드 방식을 통해서 구성된 것이다. 첫 단계에서는 컬러 히스토그램과 같은 키프레임간의 차이를 이용하여 자동적으로 키프레임 후보 집단을 추출하고 추출된 키프레임 집단에서 최종적으로 사람들에 의해서 키프레임들이 선정된다. 키프레임수는 최대 36개까지 가능하게 구성하였다.¹⁾ OVDL에 있는 키프레임의 크기가 너무 작아 판독하는데 어려움이 있었다. 따라서 본 연구의 실험을 위해서 비디오 플레이어인 KMPlayer를 사용하여 OVDL 사이트의 영상 초록에 있는 키프레임대로 전체 클립에서 다시 추출하였다. 다만 본 연구는 이미지의 의미전달의 가능성을 조사하는데 초점을 맞추었기 때문에 OVDL 사이트에 있는 프레임이라도 자막이 있는 키프레임들은 제외시켰다.

3.1.2 실험 시스템

1) 초기 화면

실험을 위한 영상 초록 목록의 초기화면은 <그림 1>의 왼쪽 부분과 같다. 표제를 포함한 주제 파악에 증거가 되는 일체의 메타데이터 요소를 제외시키고 오직 비디오코드만 보여주었다. 전체 클립은 웹사이트의 주소를 달리하여 <그림 1>의 오른쪽에 나와 있는 목록처럼 구성하였다. 이와 같이 두 목록의 웹사이트 주소를 달리한 이유는 영상 초록과 전체 클립을 같은 화면에 모아 두면 피조사자들이 전체 클립을 영상 초록 보기 이전에 먼저 볼 가능성이 있기 때문이다.

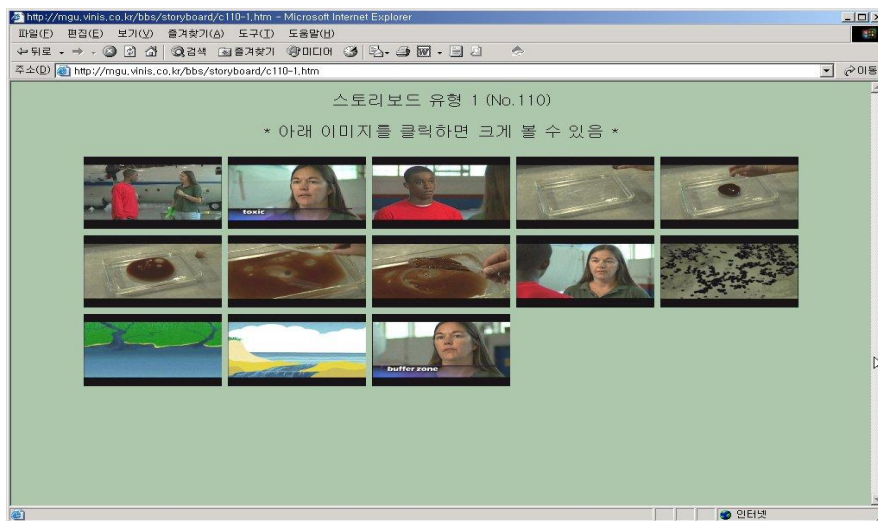
비디오코드	영상초록	비디오코드	재생시간	전체 클립
NO.101	유형 1	NO.101	4분 47초	링크
NO.102	유형 1	NO.102	2분 17초	링크
NO.103	유형 1	NO.103	4분 55초	링크
NO.104	유형 1	NO.104	2분 36초	링크
NO.105	유형 1	NO.105	3분 41초	링크
NO.108	유형 1	NO.108	4분 01초	링크
NO.109	유형 1	NO.109	4분 18초	링크
NO.110	유형 1	NO.110	3분 31초	링크
NO.112	유형 1	NO.112	2분 46초	링크
NO.113	유형 1	NO.113	2분 51초	링크
NO.114	유형 1	NO.114	4분 02초	링크
NO.115	유형 1	NO.115	10분 00초	링크

<그림 1> 영상 초록과 전체 클립의 초기화면

1)OVDL 프로젝트의 책임자인 Gary Marchionini 교수와 이메일을 통해서 얻은 정보이다.

2) 영상 초록

<그림 1>의 초기화면에서 비디오 110을 선택하면 <그림 2>와 같은 영상 초록이 출력된다. 이 비디오는 “기름 유출의 영향“이라는 표제를 갖고 있는 비디오로 기름 유출이 환경에 미치는 영향을 이야기하고, 기름 유출이 어떻게 추적되고 기름 유출을 해결하기 위해서 해야 할 방안에 대해서 이야기 한다. 이 영상 초록은 이미지와 서술을 함께 사용하여 정보를 전달하는 비디오 유형(유형 2)이다. 각 키프레임을 클릭하면 더 큰 크기(720 X 480)로 확대해서 볼 수 있다.



<그림 2> 영상 초록

3) 전체 클립

영상 초록에 대한 조사가 끝난 후 전체 클립 목록이 있는 사이트로 가서 각 비디오에 대한 전체 클립을 보도록 하였다(<그림 3> 참조). 각 비디오를 클릭하면 KMPlayer가 실행되어 전체 클립을 볼 수 있도록 실험 환경을 구현하였다.



<그림 3> 전체 클립의 한 화면

3.1.3 설문지와 피조사자

설문지는 크게 두 종류로 구성하였다. 첫 번째 설문지는 각 영상 초록에 대해서 색인어(키워드)를 4개 이상 할당하고, 비디오 요약문은 두 문장 이상 기술하도록 하였다. 두 번째 설문지는 사운드가 제거된 전체 클립을 보게 한 후 영상 초록과 똑같이 색인어와 요약문을 기술하도록 하였고 이외에 두 가지 추가적인 질문을 포함시켰다. 즉, 영상 초록의 기능이 통보적 초록인지, 지시적 초록인지 아니면 초록 기능을 전혀 하지 못하는지 등을 선택하는 질문이 있고 비디오 의미 추출을 위한 영상 초록 보기와 전체 클립 보기의 장단점들을 기술하도록 하는 질문이 포함되었다.

피조사자 14명은 M대학의 문헌정보학 전공 또는 복수 전공 학생 중에서 임의로 추출한 다음 컴퓨터실에서 각 피조사자에게 한 대의 컴퓨터를 사용하도록 하여 실험을 진행하였다. 먼저 실험에 대한 설명을 15분 정도 한 다음 첫 번째 설문지를 배포하고 영상 초록이 있는 웹사이트 주소를 알려 주어 12개의 표본 비디오에 대해서 4개 이상의 색인어 할당과 2 문장 이상의 요약문을 각각 적도록 하였다. 즉, 각 비디오에 대해서 14개의 결과값이 할당되어 총 사례수는 168개가 되었다. 각 비디오의 영상 초록에 대한 색인어와 요약문을 적는데 5분을 할당하여 총 60여분을 주었다.

영상 초록에 대한 설문지를 회수한 다음 두 번째 설문지를 배포하고 전체 클립이 있는 웹사이트 주소를 알려 주어 앞 단계에서 사용한 동일한 12개의 비디오에 대해서 색인어와 요약문을 적도록 하였다. 즉, 각 비디오 클립에 대해서 14개의 결과값이 할당되어 총 사례수

는 168개가 되었다. 12개 비디오의 전체 클립을 볼 수 있는 60 여분과 각 클립에 대한 색인어와 요약문을 적는데 5분을 할당하고 다른 두개의 질문에 10분을 할당하여 총 130분을 주었다. 총 실험시간은 실험에 대한 안내를 포함하여 약 3시간 30분이 소요되었다.

3.1.4 색인어와 요약문의 정확도 측정 방법

Open Video Digital Library 사이트에 올려진 색인어와 텍스트 초록을 표준자료로 이용하고자 하였으나 12개 중 3개 비디오에 대한 색인어 또는 텍스트 초록이 없었고, 나머지 9개 비디오에 대한 색인어와 초록도 대부분 색인어수가 너무 적거나 초록이 너무 짧아서 전체 클립을 반복해서 듣고 표준 색인어와 요약문을 재구성하였다. 비디오에 4~15개의 표준 색인어와 두 문장 이상의 표준 요약문을 할당한 다음 정확도를 측정하였다. 색인어의 정확도 측정은 피조사자가 기술한 색인어가 표준 색인어 그룹에 있으면 0.25점을 주었고 제목과 같은 색인어를 할당한 경우에는 가중치를 주어서 0.50을 주었는데 이는 초록 영상에서 열거한 하나의 사물이나 인물 등을 제목과 동일한 개념을 나타내는 색인어와 동일시할 수 없기 때문이다. 예를 들어서, 앞의 “기름 유출의 영향“이라는 비디오에서 특정 피조사자가 제목과 동일한 색인어(0.50)와 1개의 표준 색인어(0.25)를 할당하고 2개의 비표준어를 할당했다면 정확도는 0.75가 된다. 요약문의 경우 표준 초록과 내용이 어느 정도 일치하느냐에 따라 0~1 사이의 값을 주었다. 측정자의 주관적인 판단이 개입될 가능성이 높기 때문에 두 명이 정확도를 측정하고 그 평균값을 활용하였다.

3.2 데이터 분석과 가설 검증

3.2.1 비디오 유형에 따른 정확도 차이 분석

1)통계 분석: 일원배치 분산분석과 사후 검증

표본 비디오 그룹은 세 가지 유형, 즉 의미 전달을 위해서 이미지에 주로 의존하는 유형 1, 이미지와 서술에 의존하는 유형 2 그리고 서술에 주로 의존하는 유형 3으로 구분되어 있다. 이러한 세 가지 유형의 정확도 평균값의 차이를 분석하기 위해서 일원배치 분산분석과 사후 검증을 이용하였다. 이미지에 의존하여 정보를 표출하는 비디오의 색인어 및 요약문 정확도 평균값이 가장 높을 것이며 서술에 의해서 정보를 전달하는 비디오는 평균값이 가장 낮을 것이며 의미를 전달하는데 이미지와 서술을 모두 사용하는 비디오는 중간 정도의 평균값을 갖는 것으로

가정하였다. 분석 결과, 세 가지 유형의 비디오그룹 간에 영상 초록과 전체 클립을 통한 색인어 및 요약문 추출의 정확도가 통계적으로 유의미한 차이를 보였다. 비디오 유형에 따른 항목별 값의 차이를 좀 더 구체적으로 알기 위해서 사후 검증을 수행하였다. <표 1>의 첫 부분인 기술 통계값을 살펴보면 서술에 의해서 의미 전달을 하는 비디오 유형 3의 경우 영상 초록과 전체 클립에서 요약문 정확도(0.28, 0.50)가 색인어 정확도(0.39, 0.57)에 비해서 상대적으로 더 낮은 평균값을 나타내고 있다.

<표 1> 일원배치 분산분석과 사후검증 결과

기술 통계값 (비디오 유형에 따른 항목별 평균)				
비디오 유형(수)	색인어 (초록)	요약 (초록)	색인어 (클립)	요약 (클립)
유형1(이미지)(4개)	0.51	0.50	0.77	0.76
유형2(이미지+서술) (4개)	0.46	0.42	0.60	0.61
유형3(서술)(4개)	0.39	0.28	0.57	0.50
일원배치 분산분석 결과				
유형	F값	유의확률		
색인어(초록)	3.61	0.03*		
요약(초록)	16.06	0.00**		
색인어(클립)	6.13	0.00**		
요약(클립)	6.89	0.00**		
사후 검증 (다중비교) 결과				
대 응	유의확률	대 응	유의확률	
유형1[색인어(초록)] vs. 유형2[색인어(초록)]	0.63	유형1[색인어(클립)] vs. 유형2[색인어(클립)]	0.03*	
유형1[색인어(초록)] vs. 유형3[색인어(초록)]	0.02*	유형1[색인어(클립)] vs. 유형3[색인어(클립)]	0.00**	
유형2[색인어(초록)] vs. 유형3[색인어(초록)]	0.35	유형2[색인어(클립)] vs. 유형3[색인어(클립)]	0.95	
유형1[요약(초록)] vs. 유형2[요약(초록)]	0.13	유형1[요약(클립)] vs. 유형2[요약(클립)]	0.12	
유형1[요약(초록)] vs. 유형3[요약(초록)]	0.00**	유형1[요약(클립)] vs. 유형3[요약(클립)]	0.00**	
유형2[요약(초록)] vs. 유형3[요약(초록)]	0.01*	유형2[요약(클립)] vs. 유형3[요약(클립)]	0.36	

**는 0.01 수준에서 유의미하고 *는 0.05 수준에서 유의미함

2)가설 검증: 연구가설 1과 2

“가설1: 비디오 유형에 따라 색인어 정확도가 차이가 있을 것이다”와 “가설2: 비디오 유형

에 따라 요약문 정확도가 차이가 있을 것이다” 라는 가설들을 검증하기 위해서 일원배치 분산 분석을 한 결과, 가설1과 가설2가 유의확률수준 0.05에서 통계적으로 유의미한 결과가 나와서 가설이 검증되었다(<표 1> 참조).

<표 1>의 맨 마지막 부분의 “사후 검증(다중 비교)”의 결과를 살펴보면 “유형1[색인어(초록)] vs. 유형3[색인어(초록)]”, “유형1[요약(초록)] vs. 유형3[요약(초록)]”, “유형2[요약(초록)] vs. 유형3[요약(초록)]”, “유형1[색인어(클립)] vs. 유형2[색인어(클립)]”, “유형1[색인어(클립)] vs. 유형3[색인어(클립)]” 및 “유형1[요약(클립)] vs. 유형3[요약(클립)]”이 통계적으로 유의미한 차이를 보였다. 이러한 결과는 비디오의 의미 추출을 위한 영상 초록의 효용성이 비디오 장르에 따라 달라지기도 하지만 동일한 장르안에서도 의미 전달을 위해서 이미지를 어느 정도 사용했는지에 따라서도 달라진다는 것을 알 수 있다.

3.2.2 영상 초록의 색인어 및 요약문 정확도 차이 분석

1)통계 분석: T-검증

영상 초록의 색인어 및 요약문의 정확도 차이를 검증하기 위해서 t-검증을 하였다. 분석 결과, 색인어 정확도가 요약문 정확도 보다 더 높았으며 이 차이는 통계적으로도 유의미한 것으로 나타났다(<표 2> 참조).

<표 2> T-검증 (1)

기술 통계값 (항목별 평균)		
색인어 (초록)	요약 (초록)	
0.45	0.40	
t-검증 결과		
대 응	t값	유의 확률
색인어(초록) vs. 요약(초록)	3.04	0.00**

2)가설 검증: 연구 가설 3

“가설 3: 영상 초록의 색인어 정확도와 요약문 정확도에는 차이가 있을 것이다”라는 가설을 검증하기 위해서 t-검증을 한 결과, 색인어의 정확도(0.45)가 요약문의 정확도(0.40) 보다 더 높게 나타났으며 이러한 차이가 유의확률수준 0.05에서 통계적으로 유의미한 결과가 나와서 가설이 검증되었다(<표 2> 참조). 이는 영상 초록의 각 키프레임이 색인어의 역할을 하며, 초록의 키프레임들이 문맥적으로 연결되어 있지 않고 있기 때문에 정확하게 전체 내용을 파악하는데 어려움이 예상된다는 사실을 다시 한번 확인시켜주는 결과이다.

3.2.3 전체 클립과 영상 초록의 색인어 및 요약문 정확도 차이 분석

1) 통계 분석: T-검증

사운드가 제거된 전체 클립과 영상 초록을 통한 색인어 및 요약문 추출의 정확도에 차이가 있는지를 검토하기 위해서 t-검증을 하였다. 분석 결과, 전체 클립을 통해서 파악한 색인어 및 요약문의 정확도가 영상 초록의 정확도 보다 더 높은 것으로 나타났고 통계적으로도 유의미한 차이를 보였다(<표 3> 참조).

<표 3> T-검증 (2)

기술 통계값 (항목별 평균)						
색인어(초록)		색인어 (클립)		요약(초록)		요약 (클립)
.45		.64		.40		.62
t-검증 결과						
대응	T값	유의확률	대응	T값	유의확률	
색인어(초록) vs. 색인어(클립)	-6.15	0.00**	요약(초록) vs. 요약(클립)	-6.94	0.00**	

2) 가설 검증: 연구 가설 4과 5

“가설 4: 사운드가 제거된 전체 비디오 클립의 색인어 정확도와 영상 초록의 색인어 정확도는 차이가 없을 것이다”와 “가설 5: 사운드가 제거된 전체 비디오 클립의 요약문 정확도와 영상 초록의 요약문 정확도는 차이가 있을 것이다”라는 가설들을 검증하기 위해서 t-검증을 한 결과, 먼저 가설4는 전체 클립과 영상 초록의 색인어 정확도가 유의확률수준 0.05에서 통계적으로 유의미한 결과가 나와서 차이가 있는 것으로 나타나 가설이 검증되지 못했다(<표 3> 참조). 이는 비록 사운드가 제거되었지만 전체 클립을 보고 의미 파악이 좀 더 명확해 졌을 것이며 이에 따라 더 정확한 색인어 추출이 가능했다고 예상해 볼 수 있다. 가설5는 전체 클립과 영상 초록의 요약문 정확도가 유의확률수준 0.05에서 통계적으로 유의미한 결과가 나와서 차이가 있는 것으로 나타나 가설이 검증되었다(<표 3> 참조).

3.3 연구문제 분석

3.3.1 영상 초록의 기능

영상 초록이 통보적 초록의 역할을 하는지 아니면 지시적 초록의 역할을 하는지 파악하기

위해서 피조사자에게 질문을 한 결과 9명(64.3%)이 지시적 초록의 역할을 한다고 답변하였으며 그 다음으로 4명(28.6%)이 통보적 초록의 역할을 한다고 답변하였다. 1명(7.0%)은 영상 초록이 통보적 초록은 물론 지시적 초록의 역할을 하기에 미흡하다고 답변하였다(<표 4> 참조).

<표 4> 영상 초록의 기능

초록 사례	통보적초록	지시적초록	통보적 초록(또는 지시적초록)으로 미흡
사례수(명)	4	9	1
비율(%)	28.6%	64.3%	7.0%

3.3.2 영상 초록과 전체 클립의 장단점

영상 초록이 전체 클립과 비교하여 어떤 장단점이 있는지 피조사자들에게 조사한 결과, “시간을 절약할 수 있다”, “정확하지는 않지만 대충 내용을 파악할 수 있고, 자신이 찾는 주제와 부합한지 여부를 알 수 있다”, “여러 개의 영상이 있을 때 한 가지 영상을 선택하는데 도움을 줄 수 있다” 그리고 “전체적인 내용에 접근하기에 앞서서 이해력을 증진시켜주고, 내용을 파악하는데 있어 좀 더 쉽게 파악할 수 있다”를 영상 초록의 장점으로 가장 많이 언급하였다. 이러한 분석 결과는 앞에서 조사한 것 처럼 영상 초록이 통보적 초록의 역할을 하기에는 미흡하다는 것을 나타낸다고 생각한다(<표 5> 참조).

<표 5> 장단점

	영상초록	전체클립
장점	<ul style="list-style-type: none"> -전체적인 내용 중에서 꼭 필요한 부분(요약된 부분)을 간단히 알 수 있고, 기억에 남는다. -정확하지는 않지만 대충 내용을 파악할 수 있고, 자신이 찾는 주제와 부합하는지 여부를 알 수 있다. -영상을 보기 전에 전체적인 흐름을 미리 알 수 있고, 핵심 프레임을 알 수도 있어서 여러 가지 영상을 봐야할 때 대신할 수 있었다. -전체적인 내용에 접근하기에 앞서서 이해력을 증진시켜주고, 내용을 파악하는데 있어 좀 더 쉽게 파악할 수 있다. 	<ul style="list-style-type: none"> -자신이 원하는 부분을 직접 찾아볼 수 있고, 내용 파악이 쉽다. -처음부터 끝까지 전체적인 내용과 흐름을 알 수 있다. -내용을 정확하게 알 수 있고, 전체적인 내용 파악이 쉽다. -영상초록으로 이해할 수 없었던 부분을 모두 이해할 수 있었다. -영상에 대한 정보가 더 풍부해진다. -영상의 전반적인 요점과 앞뒤 맥락을 쉽게 파악할 수 있다. -영상의 소리가 없어도 내용의 흐름을 볼 수 있어서 의미 파악이 된다.

	<ul style="list-style-type: none"> -시간을 절약할 수 있다. -여러 개의 영상이 있을 때 한 가지 영상을 선택하는데 도움을 줄 수 있다. -영상을 보기 전에 화면의 품질이나 구성을 알 수 있다. -영상의 내용을 파악할 수 있기 때문에 빠른 자료 검색에 도움이 된다. -주요사건이 무엇인지 파악 가능하여 주제를 알 수 있다. 	
단점	<ul style="list-style-type: none"> -캡처의 앞뒤 순서가 완벽하게 이어지지 않을 때도 그것을 이어지는 것으로 혼동하고 순서를 통해 내용을 판단할 경우 내용 파악이 잘 안될 수 있다. - 중요 장면을 선정하는 것에 대한 기준이 모호한 것 같다. - 각자 중요하게 생각하고, 기억에 남는 부분이 달라 요약이 효과적이지 않게 되면 내용에 대한 추측이 힘들다. -정확한 내용 파악이 힘들고, 주제가 어려운 경우 전혀 감을 못 잡을 수도 있다. -영상의 유형에 따라 영상 초록의 의미파악이 어려울 수도 있다. -프레임간의 연계성이 부족하다. 영상의 부분, 부분을 잘라서 만든 것이기 때문에 이용자의 요구와 맞는 자료를 찾기가 힘들다. -텍스트 초록과 함께 쓰이지 않으면 이용자 입장에서 불편할 것 같다. -배경지식이 없이는 내용을 파악하지 못할 수 있다. 	<ul style="list-style-type: none"> -봐야 할 영상이 많거나 영상이 길 경우 보기가 번거롭고 불편하다. -불필요한 부분을 봐야 하는 경우가 있고, 중요한 부분이 한 눈에 들어오지 않는다. -시간이 너무 많이 걸려서 이용자에게 부담을 준다. -소리를 들을 수 없어 알려주고자 하는 것을 파악하기 힘들었다. -불필요한 클립도 보여줌으로서 핵심을 파악하기 힘들고, 내용을 이해하는데 혼란을 줄 수가 있다. -화면이 순식간에 지나가기 때문에 깊은 인상을 남기기 어려워서 동영상은 두 번 이상 봐야한다. -실제 온라인상에서 쓰이게 되면 용량이 커지고 검색속도가 느려질 것 같다. -원하는 정보를 빠르게 발췌하기가 쉽지 않다. -동영상 플레이어와 같은 도구가 필요하다.

이에 반해서 영상 초록의 단점으로는 “캡처의 앞뒤 순서가 완벽하게 이어지지 않을 때도 그것을 이어지는 것으로 혼동하고 순서를 통해 내용을 판단할 경우 내용 파악이 잘 안될 수 있다”, “정확한 내용 파악이 힘들고, 주제가 어려운 경우 전혀 감을 못 잡을 수도 있다”, “영상의 유형에 따라 영상 초록의 의미파악이 어려울 수도 있다”, “텍스트 초록과 함께 쓰이지 않으면 이용자 입장에서 불편할 것 같다”, “배경지식이 없이는 내용을 파악하지 못할 수 있다” 그리고 “중요장면을 선정하는 것에 대한 기준이 모호한 것 같다”를 영상 초록의 단점으로 가장 많이 꼽았다.

전반적으로 전체 클립의 장단점은 영상 초록의 장단점을 역으로 생각하는 경향을 나타냈

다. 즉, 전체 클립의 장점으로 “자신이 원하는 부분을 직접 찾아볼 수 있고, 내용 파악이 쉽다”, “내용을 정확하게 알 수 있고, 전체적인 내용 파악이 쉽다”, “영상 초록으로 이해할 수 없었던 부분을 모두 이해할 수 있었다” 그리고 “영상의 전반적인 요점과 앞뒤 맥락을 쉽게 파악할 수 있다”를 가장 많이 언급하였다. 한편 전체 클립의 단점으로는 “봐야 할 영상이 많거나 영상이 길 경우 보기가 번거롭고 불편하다”, “불필요한 부분을 봐야하는 경우가 있고, 중요한 부분이 한 눈에 들어오지 않는다”, “시간이 너무 많이 걸려서 이용자에게 부담을 준다”, “불필요한 클립도 보여줌으로써 핵심을 파악하기 힘들고, 내용을 이해하는데 혼란을 줄 수가 있다”, “화면이 순식간에 지나가기 때문에 깊은 인상을 남기기 어려워서 동영상을 두 번 이상 봐야한다”, “실제 온라인상에서 쓰이게 되면 용량이 커지고 검색속도가 느려질 것 같다” 그리고 “원하는 정보를 빠르게 발췌하기가 쉽지 않다”를 가장 많이 언급하였다.

4. 논의와 후속 연구

연구 가설 및 문제의 검증 결과 다음과 같은 논점을 이끌어 내었다. 첫째, 영상 초록 보기를 통한 색인어 및 요약문 추출의 정확도가 다른 선행 연구에 비교하여 다소 낮은 값으로 나타난 것은 영상 초록의 키프레임 구성시 자막이 들어 있는 키프레임들을 제외시킨 점과 피조사자에게 비디오 코드 이외의 다른 메타데이터 요소를 제시하지 않았기 때문으로 생각된다. 둘째, 같은 장르에 속해 있는 비디오라도 의미 전달을 위해서 이미지를 어느 정도 사용했는지에 따라서 색인어 및 요약문 추출의 정확도가 달라진다는 것을 알 수 있다. 특히 이미지에 의존하여 내용을 전달하는 비디오 유형1의 정확도(색인어: 0.51, 요약문: 0.50)와 서술에 의존한 비디오 유형3의 정확도(색인어: 0.39, 요약문: 0.28)간에 많은 차이가 있었다. 이러한 연구 결과는 비디오 유형3의 경우 영상 초록만으로 의미 파악 또는 적합성 판정을 하는데 한계를 드러내기도 하지만 반대로 이미지에 주로 의존하여 정보를 표출하는 비디오 유형1의 경우 텍스트 초록만으로 의미 파악을 하기에는 한계가 있으며 텍스트 초록이 영상 초록과 함께 사용되었을 때 시너지 효과를 낼 수 있음을 보여주고 있다. 결론적으로 영상 초록은 디지털 도서관의 OPAC 시스템 환경에서 제목, 색인어, 텍스트 초록 등 다른 메타데이터 요소들과 함께 사용되었을 때 비디오 장르나 유형에 관계없이 멀티미디어 자료의 검색과 브라우징의 효율성을 증진시킬 수 있을 것으로 생각된다.

셋째, 영상 초록 보기를 통한 색인어와 요약문 추출의 정확도(색인어 정확도:0.45, 요약문 정확도:0.40)가 전체 클립 보기를 통한 색인어와 요약문 추출의 정확도(색인어 정확도:0.64,

요약문 정확도:0.62) 보다 떨어지지만 영상 초록의 정확도가 절반치(0.50)에 근접한 것으로 나타났다. 무엇보다도 의미 추출에 소요되는 시간을 대략 82% 정도 감소시킬 수 있다는 점이 영상 초록의 최대의 강점으로 보인다(Stachowicz, 2002). 또한 영상 초록의 색인어 정확도(0.45)가 요약문 정확도(0.40) 보다 더 높게 나타났고 두 값의 차이도 통계적으로 유의미한 것으로 나타나 영상 초록을 통해서 요약문을 추출하는 것 보다 색인어를 추출하는 것이 더 효율적임을 확인할 수 있었다.

본 연구의 분석 결과에 기초하여 앞으로 연구되어야 할 사항들에 대해서 기술하면 다음과 같다. 첫째, 영상 초록을 통해서 추출한 색인어의 정확도가 요약문의 정확도 보다 더 높은 것으로 나타났다. 요약문의 정확도가 상대적으로 낮은 것은 앞에서 언급한 것처럼 초록의 키프레임들이 문맥적으로 연결되어 있지 않기 때문인 것으로 생각된다. 키프레임 순서가 영상 초록의 이해에 어떠한 영향을 미치는지를 판단하기 위해 영상 초록에 관한 키프레임 순서와 요약 개념 순서 사이의 상관관계를 측정된 결과 요약문 가운데 61%에 포함된 아이디어는 키프레임 순서와 직접적으로 상호관련이 있는 것으로 나타났다(Lyer and Lewis, 2006). 이는 영상 초록에서 키프레임 순서의 중요성을 강조하는 것이다. 이러한 가정에 기초하여 영상물의 소통 구조 이론과 용어간의 구문적 및 어의적 관계를 고려하는 온톨로지 이론을 기초로 하여 영상 초록의 키프레임들을 문맥적으로 연결시키는 키프레임 배열 모형을 설계하여 설계된 모형에 따라서 키프레임들을 재구성한 영상 초록을 사용하거나 또는 배열모형에 기반한 영상 초록을 순차적으로 구성한 영상 초록과 함께 사용한다면 비디오의 의미 추출을 훨씬 더 정확하게 수행할 수 있다고 생각한다.

둘째, 영상 초록이 비디오의 내용 전달을 위한 효율적인 도구가 되기 위해서는 전체 클립에서 의미 있는 키프레임 추출 작업을 가능하게 하는 키프레임 추출 알고리즘에 대한 연구가 무엇보다도 중요하다고 생각된다. 공학 분야 연구자들은 이미지 프로세싱에 초점을 맞춰서 즉, 이미지의 색, 질감, 객체, 음성 정보 등을 인식하여 비디오 자료들을 분할하고 이를 기초로 하여 영상 초록을 구축하고 있다. 그러나 이러한 이미지 프로세싱 방법에만 의존하여 의미 있는 키프레임들을 추출하는 데는 한계가 있다고 생각한다. 따라서 공학 분야에서 발전하고 있는 이미지 프로세싱 연구 결과와 함께 사람들이 비디오 자료의 주제를 판단할 때 많이 사용하고 있는 객체의 종류 또는 상징물에 대한 조사와 비디오자료의 적합성 문제를 체계적으로 고찰한 후 영상소통구조 이론 및 편집기술 등을 활용하여 영상 초록 구성을 위한 키프레임 추출 알고리즘의 설계 또한 지속적으로 연구되어야 할 분야라고 생각한다.

끝으로, 현재 사용되고 있는 영상 초록은 정지된 키프레임들을 나열하고 있는데 이러한 이미지들이 요약된 비디오 사운드와 결합되어 사용된다면 의미 파악이 훨씬 쉬워질 것으로 판단

된다. 따라서 전체 클립에서 사용하고 있는 전체 사운드에서 핵심적인 사운드만 이끌어 내는 이론과 방법도 장기적으로 연구되어야 할 분야라고 생각한다.

5. 결 론

최근 ‘멀티미디어 정보학’이라는 용어도 생겨나고 있다. 이러한 추세가 계속된다면 문헌정보학은 연구와 교육에서 문헌정보 뿐만 아니라 멀티미디어 정보도 상당한 비중으로 다루어야 할 것임이 분명해 보인다. 문헌정보학의 연구 분야인 정보 조직, 정보 검색 등에서 영상 정보의 구조화, 검색 방법 그리고 더 나아가서는 영상 정보의 처리 까지 다루어야 하는 것은 문헌정보학 분야에 있는 연구자와 실무자가 피할 수 없는 일이 된 것 같다. 멀티미디어 정보의 생성 과정, 저장, 검색 및 보존에 대한 이론적인 측면과 기술적인 측면을 심도 있게 이해하기 위해서는 영상 콘텐츠가 분할되어 처리되는 물리적인 색인 및 초록 방법에 대한 이해와 함께 전통적인 문헌정보학의 정보검색 이론체계를 멀티미디어 정보에 접목시키려는 노력이 필요해 보인다.

본 연구에서는 비디오 자료의 의미 추출을 위한 영상 초록의 효용성을 조사하기 위해서 표본 비디오 12개와 피조사자 14명을 이용하여 영상 초록과 전체 클립의 색인어 및 요약문의 정확도를 측정해 보았다. 표본 비디오 자료에서 자막이 있는 키프레임들은 되도록 제외시키고 피조사자들에게 표제를 포함한 메타데이터 요소들을 전혀 제시하지 않고 영상 초록만을 보여 주었기 때문에 색인어 및 요약문 추출의 정확도가 다른 연구에 비해서 낮을 것으로 예측된다. 측정 결과 첫째, 비디오 유형에 따라 정확도가 차이가 있는 것으로 나타났다. 특히 이미지에 의존하여 내용을 전달하는 비디오 유형1의 정확도(색인어: 0.51, 요약문: 0.50)와 서술에 의존한 비디오 유형3의 정확도(색인어: 0.39, 요약문: 0.28)간에 큰 차이가 있었다. 이는 이미지에 주로 의존하여 정보를 표출하는 비디오의 경우 텍스트 초록만으로 의미 파악을 하기에는 한계가 있으며 텍스트 초록이 영상 초록과 함께 사용되었을 때 시너지 효과를 낼 수 있음을 보여 주고 있다.

둘째, 영상 초록의 색인어 정확도(0.45)와 요약문 정확도(0.40)가 전체 클립의 정확도 보다 떨어지지만 절반치(0.50)에 근접한 것으로 나타나 영상 초록이 비디오 의미 추출에 효율적으로 활용될 수 있음을 확인하였다. 또한 영상 초록의 색인어 정확도가 요약문 정확도 보다 더 높게 나타나 영상 초록을 통해서 색인어 추출 작업을 더 효율적으로 할 수 있음을 확인할 수

있었다. 이러한 연구 결과에 기초하여 본 연구는 영상 초록이 색인어 및 요약문 추출 작업에 활용될 수 있을 뿐만 아니라, 디지털 도서관의 OPAC 시스템의 메타데이터의 한 요소로 활용되어 적합성 판정을 위한 이용자의 비디오 브라우징 및 영상 질의가 이용되는 비디오 검색 환경에 이용할 수 있음을 제안하고, 영상 초록의 품질을 높이기 위한 후속 연구들에 대해서도 제안하였다.

6. 참고 자료

- 고영승, 주경수. 2005. MPEG-7 기반의 의료영상 검색시스템 개발. 『멀티미디어 학회논문지』, 8(8): 1032-1041.
- 김재우. 2002. 영상 기호의 의미 작용에 관한 연구. 『동서언론』. 6:41-67.
- 김무규. 2006. 영화 영상에 관한 기호학적 이해의 문제점과 대안: '사이공간' 개념을 중심으로. 『언론과 학연구』, 6(3); 65-98.
- 존, 버거. 1991. 『영상커뮤니케이션과 사회』. 강명구 역. 서울: 나남출판.
- 황민선. 2005. 출판물과 출판물을 원작으로 하는 영상물의 매체 특성 비교 연구. 『한국출판학연구』 48: 371-392.
- Browne, P. and Smeaton, A. F. 2005. "Video retrieval using dialogue, keyframe similarity and video objects." *ICIP 2005 - International Conference on Image Processing*, Genova, Italy, 11-14.
- Choi, Y. and Rasmussen, E.M. 2002. "User's relevance criteria in image retrieval in American history." *Information Processing and Management*, 38(5):695-726.
- Christel, M.G. et al. 2002. "Collages as dynamic summaries for news video." In *Proceedings of the 10th ACM International Conference on Multimedia* (pp. 561-569). New York: ACM Press.
- Christel, M. G. and Warmack, A. 2001. "The effect of text in storyboards for video navigation." In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol.3 (pp. 1409-1412). New York: IEEE Press
- Hughes, A. et al. 2003. "Text or pictures? An eye-tracking study of how people view digital video surrogates." In *Proceedings of CIVR 2003* (pp.271-280).
- Li, F. et al. 2000. Browsing digital video. *Proc. CHI '00*, pp.169-176.
- Lyer, H. and Lewis, C. D. 2007. "Prioritization strategies for video storyboard keyframes." *JASIST* 58(5): 629-644.
- Marchionini, G. and Geisler, G. 2002. "The Open video digital library." *D-Lib Magazine*. 8(12). [cited 2006.4.25].

- Markkula, M. and Sormunen, E. 1998. "Searching for photos - journalistic practices in pictorial IR." In: Eakins, J.P., et al. (Eds), *The Challenge of Image Retrieval*. Newcastle upon Tyne, 5-6 Feb 1998. British Computer Society (BCS), Electronic Workshops in Computing.
- Mu, X. and Marchionini, G. 2003. "Enriched video semantic metadata: Authorization, integration, and presentation." In *Proceedings of the Annual Meeting of the American Society for Information Science and Technology* (pp. 316-322). Silver Spring, MD: American Society for Information Science and Technology.
- Niblack, W. et al. 1997. "Updates to the QBIC system." *Proceedings- SPIE The International Society for Optical Engineering Storage and Retrieval for Image and Video Database VI* 1997; Issue 3312, 150-161.
- Robertson, G. et al. 1998. "Data mountain: Using spatial memory for document management." In *Proceedings of the 11th Annual ACM Symposium on User Interface Software and Technology* (pp. 153-162). New York: ACM Press.
- Smeaton, A.F. and Browne, P. 2006. "A usage study of retrieval modalities for video shot retrieval." *Information Processing and Management*, 42(5):1330-1344.
- Smeaton, A. F. 2004. "Indexing, browsing, and searching of digital video." In *Annual Review of Information Science and Technology* Edited by Blaise Cronin, Vol.38: 371-407.
- Stachowicz, C. 2002. *The effectiveness of storyboard surrogates in the subject indexing of digital video*. Master's Degree Thesis, Univ. of North Carolina, Chapel Hill.
- Wildemuth, B.M. et al. 2002. "Alternative surrogates for video objects in a digital library: Users' perspectives on their relative usability." In *Proceedings of the 6th European Conference on Digital Libraries* (pp. 493-507). New York: Springer.
- Yang, M. 2005. *An exploration of users' video relevance criteria*. Unpublished doctoral dissertation, University of North Carolina at Chapel Hill.
- Yang, M., Wildemuth, B., & Marchionini, G. 2004. "The Relative Effectiveness of Concept-Based versus Content-Based Video Retrieval." *Proceedings of the ACM Multimedia Conference*, 368-371.

<부 록: 표준 색인어와 요약문>

다음은 12개의 비디오 샘플 자료 중 두 개의 비디오에 대한 표준 색인어 집단과 요약문이다.

	비디오 1	비디오 2
색인어 그룹	1)자전거 2)자전거 안전 상태 점검 3)자 전거 타는 규칙 4)교통 5)운송수단 6)질서 7)정지 표시판	1)The Ocean Odyssey 2)조수(Tides) 3) 파도(Waves) 4)Oceans(해양, 바다) 5) Impact of Oil Spills(기름 유출의 영향) 6) 기름 유출을 해결하는 방안 7)독성물질
요약문	자전거 기계적 이상이 없는지 점검하는 방 법을 보여주고 좋은 상태로 유지하는 방법 도 보여준다. 소년과 소녀는 자전거 타기 알아야 할 정확한 자전거 규칙을 보여주고 안전하게 타는 법을 보여준다. 따라서 이 비디오의 주제는 크게 두 가지로 자전거 점 검과 자전거의 운행 규칙을 들 수 있다.	기름 유출이 환경에 미치는 영향을 이야기 하고, 기름 유출이 어떻게 추적되고 이런 기 름 유출을 해결하기 위해서 해야 할 방안에 대해서 이야기 한다.