

내용 기반 영상 검색을 이용한 실시간 몽타주 시스템 설계

Real-time Montage System Design using Contents Based Image Retrieval

주저자 : 최현석 (Hyeon Seok Choi)

중앙대학교 첨단영상대학원 영상공학과 예술공학전공

공동저자 : 배성준 (Seong Joon Bae)

중앙대학교 첨단영상대학원 영상공학과

공동저자 : 김태용 (Tae Yong Kim)

중앙대학교 첨단영상대학원 영상공학과 교수

공동저자 : 최종수 (Jong Soo Choi)

중앙대학교 첨단영상대학원 영상공학과 교수

“본 논문은 교육인적자원부, 산업자원부, 노동부의 출연자금으로 수행한 최우수실험실지원사업의 연구 결과입니다.”

1. 서 론

- 1-1 연구의 배경
- 1-2 연구의 목적 및 필요성
- 1-3 연구의 방법 및 범위

2. 영상 검색 기술과 몽타주 이론

- 2-1 뉴미디어의 개념
- 2-2 영상 검색 기술
- 2-3 몽타주 이론

3. 관련 작품연구

- 3-1 한 단어영화
- 3-2 소프트 시네마

4. 실시간 몽타주 시스템

- 4-1 내용 기반 영상 검색
- 4-2 실시간 움직임 감지
- 4-3 실시간 몽타주 시스템 고찰

5. 결 론

참고문헌

(要約)

본 논문에서는 내용 기반 영상 검색을 이용하여 사용자가 원하는 영상을 쉽게 찾아내고, 이를 자동 재구성함으로써, 독창적인 영상 언어라 일컬어지는 몽타주를 사용자 중심의 관점에서 구현하고자 한다.

본 논문에서 제안하는 실시간 몽타주 시스템은 이산 푸리에 변환(Discrete Fourier Transform)을 이용해 사용자가 선택한 영상의 특징을 찾고, 유클리디안 거리(Euclidean Distance)를 이용해 데이터베이스에 있는 영상과 유사도를 비교함으로써, 빠르고 효과적으로 사용자가 원하는 영상을 검색할 수 있다. 또한 카메라 트래킹(Camera Tracking)에 의해 실시간으로 사용자의 움직임 영상을 취득하고, 취득된 영상을 검색된 사용자의 영상과 함께 자동 재구성함으로써, 손쉽게 사용자의 의도에 맞춘 영상 재구성을 하게 된다.

본 시스템은 사용자를 즐겁게 참여 시킬 수 있는 뉴미디어 영상 디자인 툴(엔터테인먼트)이다. 일반적으로 영상을 시청하는 소극적 영상의 수용자에서 벗어나 영상 재생산의 적극적 주체가 되는, 사용자 중심의 새로운 영화(미디어기반 엔터테인먼트)의 토대가 될 것으로 기대된다.

(Abstract)

In this paper , we introduce "Contents Based Image Retrieval" which helps a user find the images he or she needs more easily and reconfigures the images automatically. With this system, we try to realize the language of (motion) picture, that is, the Montage from the viewpoint of the user.

The Real-time Montage System introduced in this paper uses 'Discrete Fourier Transform'. Through this, the user can find the feature of the image selected and compare the analogousness with the image in the database. This kind of system leads to the user's speedy and effective retrieving. Also, we can acquire the movement image of the user by Camera Tracking in Real-time. The movement image acquired is to be reconfigured automatically with the image of the user. In this way, we can get an easy and speedy image reconfiguration which sets to the user's intention.

This system is a New Media Design tool(entertainment) which induces a user enjoy participating in it. In this system, Thus, the user is not just a passive consumer of one-way image channels but an active subject of image reproduction in this system. It is expected to be a foundation for a new style of user-centered movie (media based entertainment).

(Keyword)

montage, contents based image retrieval, new media

1. 서론

1-1. 연구의 배경

영화, 방송 기술의 발달과 디스플레이 장치의 보급으로 현대 사회에서 영상은 매우 보편적인 정보 전달 매체가 되었다. 또한, 인터넷 등 네트워크기술의 발달과 디지털 카메라, 캠코더, 컴퓨터 등 하드웨어의 발달과 보급으로 사용자는 언제 어디서나 손쉽게 영상에 접근할 수 있게 되었다. 그러나 우리가 일상생활에서 쉽게 접할 수 있는 기존 미디어는 정보 전달이 일방적이고 획일적이기 때문에, 다양한 가치관과 정보에 대한 욕구에 적합하지 않다. 이를 반증하듯 인터넷 기반의 개인화 된 미디어(1인 미디어)는 놀라울 만한 속도로 확산되며 또 한 번의 커뮤니케이션 혁명 시대를 열고 있다. 미니홈피, 블로그, 브이로그(Vlog), 포드캐스팅(Podcasting) 등에 탑재된 디지털 영상들은 폭발적인 양과 초고속도로 퍼져나가고 공유되고 있다. 그와 더불어 즐기고, 참여하려는 사용자들의 관심은 더욱 높아지고 있다.

이에 따라 미디어 아트 분야에서는 상승된 욕구를 충족 시켜줄 새로운 시도가 이어지고 있다. 그리고 보다 체계적인 연구 그룹인 레프 마노비치(Lev Manovich)¹⁾의 소프트 시네마(SOFT CINEMA), 제프리 쇼(Jeffrey Shaw)²⁾의 아이 시네마(iCinema)³⁾, MIT 미디어랩(MIT Media lab)의 미디어 패브릭스(Media Fabrics, 구 Interactive Cinema)⁴⁾ 등에서 컴퓨터 비전, 가상현실, AI(artificial intelligence) 등 첨단 기술과 융합된 새로운 미디어 기반 연구가 활발히 진행되고 있다.

그러나 한층 더 높아진 욕구와 다양한 연구에도 불구하고 일반인들은 움직임(motion picture)방식의 영상 제작에 많은 어려움을 겪고 있다. 이는 기존 미디어(영화, 방송) 제작 방식이 보다 많은 시간과 경험을 요구하기 때문이다. 우선 원하는 영상을 얻기 위해서는 그곳으로 가야만 한다. 또 촬영을 위해선 노출, 구도, 쇼트 구성 등보다 풍부한 경험이 필요하다. 보다 좋은 질의 영상을 만들려면 전문적인 장비(수동 조절 가능한 카메라, 교환렌즈, 조명기구 등)와 스태프도 필요하다. 그리고 소설이나 드라마 같은 플롯중심 이야기 구성은 보다 많은 연출을 요구하게 된다. 프로덕션 디자인, 스토리보드, 보다 미장센⁵⁾과 배우도 필요하다. 애니메이션의 경우는 보다 더 노동집약적 노력과 전문적 경험이 필요하다. 따라서 보다 손쉽고 효율적인, 새로운 미디어에 대응할 수 있는, 구상이나 설계가 절실히 필요하다.

1) 모스크바 태생인 마노비치는 미술, 건축, 컴퓨터공학을 전공하고 뉴욕으로 이주한 후 뉴욕대에서 인지과학 석사, 로체스터대에서 시각과 문화학 박사학위를 받음. 마셜 맥루한 이후 가장 광범위하면서도 예리하게 미디어 역사를 다룬 것으로 평가됨.
<http://www.manovich.net>

2) 호주 멜버른 태생인 쇼는 건축과 조각을 전공함. 1990년에는 그의 작품 The Legible City로 the Prix ArsElectronica로 부터 Distinction상을 수상하기도 함. ZKM Center for Art and Media Karlsruhe의 시각 미디어연구소장.
<http://www.jeffrey-shaw.net>

3) <http://www.icinema.unsw.edu.au/>

4) <http://mf.media.mit.edu/>

5) mise-en-scène 촬영상 안의 요소들 사이의 관계, 카메라의 움직임, 구도, 조명 등과 같은 편집단계 이전의 과정.

1-2. 연구의 목표 및 필요성

뉴미디어는 사용자 개개인에게 맞춤화되고, 최적화된 환경을 제공하는 것을 가장 중요시한다. 영상을 만든 사람이 보는 사람에게 일방적으로 내용을 전달하는 것이 아니라, 개인의 특성과 상황에 알맞은 영상을 제공받을 수 있고, 이러한 영상을 사용자가 의도하는 방향으로 쉽게 재가공하여 시각 메시지로 구성하며 즐길 수 있어야 한다.

본 논문에서는 내용 기반 영상 검색 기술을 이용하여 사용자가 원하는 영상을 쉽게 찾아내고, 이를 자동 재구성함으로써, 독창적인 영상 언어라고 일컬어지는 몽타주를 사용자 중심의 관점에서 구현하는데 연구의 초점이 있다. 이를 위해 제안된 실시간 몽타주 시스템은 사용자가 아주 손쉽게 놀이하듯 시각 메시지를 구성하고, 긴장감(suspense)을 유발 시킬 뿐만 아니라 상호작용의 재미까지 느낄 수 있는 뉴미디어 기반 영상 디자인 툴 또는 엔터테인먼트이다. 이는 C세대⁶⁾의 유희적 욕구를 충족시켜 줄 수 있다.

누구나 영상의 수용자에서 벗어나 적극적 주체로서, 자신만의 영상 언어를 효과적으로 구성 할 수 있는 기술적 예술적 토대를 마련함으로써, 앞으로 새로운 영화(미디어기반 엔터테인먼트)에 한발 더 다가가기라 기대된다.

1-3. 연구의 방법 및 범위

본 논문에서는 이러한 뉴미디어처럼 상호작용이 가능하고 영상 재구성이 용이한 실시간 몽타주 시스템에 대해 기술하고자 한다. 본 논문은 2장에서는 먼저 뉴미디어의 기본 적인 개념을 정의한다. 그리고 데이터베이스 서사를 지원해 줄 수 있는 영상 검색 기술에 대한 이론적 고찰을 한다. 이어서 치밀한 인과율에 얽매이지 않는, 독창적인 몽타주 구성 방식의 뉴미디어적인 특징을 본 시스템과 연관 지어 살펴본다. 3장에서는 관련 작품인 한단어영화 (onewordmovie)와 소프트 시네마 (SOFT CINEMA)의 장단점을 고찰한다. 4장에서는 본 시스템의 내용 기반 영상 검색에 사용된 보다 효율적인 고속 푸리에 변환(Fast Fourier Transform)과 유클리디안 거리(euclidean distance)를 설명하고, 카메라 트래킹 기술은 무엇이며 사용자 움직임 입력 방법에는 어떠한 알고리즘이 사용되었는지 알아본다. 이어 에이치시아이(HCI)2006 '에이치시아이♥키즈(HCI♥KIDS)'에 전시된 아이 플립북(I.flipbook, Ver.2.0) 중심으로 본 논문에서 제안한 실시간 몽타주 시스템을 고찰한다. 마지막 5장에서는 결론 및 향후계획을 도출한다.

2. 영상 검색 기술과 몽타주 이론

1장에서 본 연구의 바탕이 되는 뉴미디어의 개념을 정의하고 2장에서 내용 기반 영상 검색과 자동화(뉴미디어 원리³⁾와의 관계를 밝히고 검색기술에 관한 이론적 고찰을 한다. 이어 3장에서 러시아 몽타주 이론을 살피고 뉴미디어의 가변성(뉴미디어 원리⁴⁾)과의 연관관계를 설명한다.

6) LG경제연구원, 'C-Generation'은 콘텐츠 세대를 뜻하는 것으로 사진·음악·동영상 등과 같은 콘텐츠를 자신이 직접 디지털 기기로 만든 파일을 인터넷상에 올리고 이를 다른 사람과 공유하려는 세대.

2-1. 뉴미디어의 개념

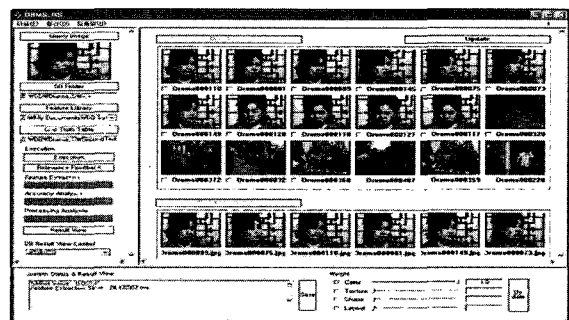
개인화 된 미디어(1인 미디어)는 틀과 같은 성격을 지니고 있다. 이는 자동화와 가변성에 기인한다. 다시 말해 온전히 인간에 의해 제작되기보다는 부분적으로 컴퓨터에 의해 자동적으로 수행된다. 자동화는 웹 디자이너에 의해 제작된 템플릿들을 사용하여, 데이터베이스로부터, 안내에 따라 몇 번 클릭하면 수초 안에 생성되는 블로그와 콘텐츠가 좋은 예이다. 그리고 동일한 방법으로 생성된 다른 사람의 블로그는 나의 그것과는 똑같지 않다. 이를 (가변적) 맞춤형이라고 말할 수 있다. 물론 모듈성과 수적재현이 기본이 되었기에 가능한 일이다. 레프 마노비치는 이러한 특성들을 컴퓨터 화의 과정을 거치고 있는 '뉴미디어의 원리'로 설명한다. 수적재현(뉴미디어 원리1)과 모듈성(뉴미디어 원리2)이라는 뉴미디어의 '물리적' 기초 원리에서 시작하여, 보다 심도 있고 그 영향력이 광범위한 '자동화(뉴미디어 원리3)'와 '가변성(뉴미디어 원리4)'의 원리로 옮겨간다. 마지막으로 문화적 '부호 변환(뉴미디어 원리5)'을 통해서 미디어 컴퓨터 화의 가장 중요한 결과를 기술하고 있다. 그리고 그는 존재하는 모든 미디어를 컴퓨터에서 처리할 수 있도록 숫자화 된 자료로 전환하는 것, 그래서 모든 문화가 컴퓨터를 매체로 하여 생산, 배포, 의사소통하는 형태로 바뀌는 것으로서 뉴미디어를 정의한다. 뉴미디어라고 하면 단순히 디지털을 떠올리기 쉽기에 대한 다양한 정의들도 많으나 철저한 탐구를 바탕으로 한 그의 정의가 보다 독특하고, 설득적이라 평가 할 수 있다.

2-2. 영상 검색 기술

우선 '뉴미디어의 원리'중 본 논문과 관련성이 보다 높은 '자동화(뉴미디어 원리3)'를 중심으로 접근해본다. 앞에서 간략하게 설명했던 미디어 제작의 자동화와 함께, 점점 더 자동화되고 있는 또 다른 영역은 '미디어 접근' 부분이다. 인터넷은 거대한 하나의 미디어 데이터베이스이다. 현재 정보의 과잉상태에서 이미지와 같은 뉴미디어 객체를 어떻게 제작하느냐 보단 어떻게 찾느냐 하는 것이 더 중요해 졌다. 어떤 특별한 이미지를 새로 제작한다면 이는 이미 존재하지만 찾기 어렵기 때문에 확률이 높다. 현대 사회는 사진 발명 이후 150년 이상 지나며 전례 없는 양의 미디어 자료들이 축적되었다. 이에 따라 미디어 혁명의 다음 단계, 즉 뉴미디어의 출현은 새로운 것들을 제작하는 것만큼이나 이미 존재하는 미디어 객체에 접근하고 다시 사용하는 것에 대해관심을 가지고 있는 미디어 사회의 두 번째 단계에 부합되는 것이다.⁸⁾ 이제 공학적인 측면에서 영상 검색 기술을 살펴본다. 이러한 방대한 정보들 중 사용자가 원하는 정보를 쉽고, 빠르게 그리고 정확하게 얻을 수 있는 방법을 연구하는 것이 정보 검색(IR: Information Retrieval)이다.⁹⁾ 일반적으로 멀티미디어 정보검색 방법은 세 가지로 분류될 수 있다. 첫 번째는 원시적인 관찰에 의한 검색이다. 이 방법은 순차적인 검색 (재생, 빨리감기, 되감기)만을 제공하고 있어 정보를 검색하기 위해

서는 많은 시간을 소비해야 한다. 두 번째는 텍스트 기반(Text Based) 영상 검색 방법으로 모든 멀티미디어 검색 대상에 사람이 직접 색인을 부여하고 탐색자는 그 색인에 부합되는 탐색어를 사용하여 원하는 정보를 검색하는 방법이다. 이 방법은 대용량의 멀티미디어 데이터에 대한 수작업에 의한 색인 작업이 시간과 비용면에서 너무나 방대하고, 멀티미디어 데이터의 성격상 데이터의 속성들을 소수의 키워드로 표현하기가 어렵다. 그리고 데이터 내용에 대해 색인자와 탐색자가 주관적 관점의 텍스트로 표현하게 됨으로써 객관성이 결여되어 검색의 불일치를 발생시켜 검색의 효율성이 크게 떨어진다. 세 번째는 내용 기반 영상 검색(CBIR : Contents Based Image Retrieval) 방법이다. 멀티미디어 데이터에서 그 내용을 대표할 수 있는 특징을 추출하고, 이를 기반으로 색인과 검색을 수행한다. 관리자에 의해 수동으로 삽입되는 주석 입력 방식이 아닌 영상 내 특징을 자동으로 추출하여 색인 과정에 사용함으로써 데이터베이스 구축에 필요한 시간 및 인력의 소모를 줄인다는 장점이 있다.¹⁰⁾

검색을 위한 대표적인 시각특징정보로는 색상(color)과 질감(texture), 그리고 모양(shape)이 있다. 색상은 내용 기반 영상 검색 기법에 있어 가장 중요한 요소로서 모든 인지 색상은 RGB 색상의 조합으로 만들어 낼 수 있다. 일반적으로 색상은 효율적인 특징 벡터로의 사용을 위해 양자화 하여 사용한다. 질감은 영상 내의 시각 요소가 조밀하고 균등하게 나열 되어 있을 경우의 시각적 무늬를 말한다. 질감 특징은 크게 구조적인 분석과 통계적인 분석 기법에 의해 추출되는데, 통계적 분석 기법이 영상의 화소 단위 분석이라면 구조적인 분석은 좀 더 큰 매크로 영역에서의 질감 분석이라고 할 수 있다. 모양에 대한 특징 정보를 얻는 것은 다른 특징에 비해 매우 어렵다. 이미지에서 분할하고자 하는 대상 물체가 얼굴인지, 꽃인지, 비행기 인지에 따라 각기 모양을 얻어내는 방법이 달라져야 하기 때문이다. 객체의 모양 특징은 영상에서의 에지 검출 후에 얻어지며, 해당 객체의 면적, 편심률, 원형성, 곡률 등의 특징 정보를 검색에 이용한다.



< 그림 1 > 이미지 질의 방식의 예

이미지 질의 방식은 폭넓게 사용되고 있는 검색 방법이다. 다양한 이미지 목록으로부터 질의 영상을 선택하여 데이터베이스의 영상과 특징을 비교한다(그림 1 참조). 이러한 방법은 사용자의 편의라는 측면에서 매우 유용하다.¹¹⁾

7) 박해미, 뉴미디어의 언어, 독립영화, 69, 2004

8) 레프마노비치, 서정신역, 뉴미디어의 언어, 79-80, 생각의 나무, 2004

9) Glenn Becker. Information in Images, chap 6: Content-based Query of Image Databases. Thomson Technology Labs, 1997. Online book.

10) 노진구, 내용기반 영상정보 검색기술에 관한 이론적 고찰, 한국도서관/정보학회지 제31권 제1호 2000.03

11) 김낙우 최중수, 내용 기반 영상 검색 기술의 현황 및 전망, 한국멀티미디어학회지 제8권 제2호 2004. 06

2-3. 몽타주¹²⁾ 이론

1920년대부터 베르토프 등 야심에 찬 젊은 러시아 감독들은 “있는 그대로의 사물을 촬영한 장면에 편집기법을 가미하여 전혀 다른 생동감을 연출할 수 있다”는 러시아 스타일의 몽타주(montage)를 시도하기 시작했다. **A + B = AB (연결의 몽타주):** 콜레쇼프, 푸토프킨은 실험을 통해 영화의 예술적 오리지널리티를 발견했다. 표정 없는 얼굴 + 테이블 위의 스프 접시, 표정 없는 얼굴 + 죽은 여성이 누워있는 영상, 표정 없는 얼굴 + 장난감을 가지고 노는 어린 소녀, 이 세 가지 몽타주를 관객들에게 보였을 때 그 결과는 놀라운 것이었다. 배우의 표정은 모두 같은 것이었음에도 불구하고 관객들은 배우의 연기를 보고 열광했고 배고픈 눈빛, 슬픈 모습, 기쁜 미소에 감탄했다. 어떤 쇼트와 어떤 쇼트가 연결되느냐에 따라 전혀 다른 의미, 새로운 이미지, 새로운 의미를 도출할 수 있다는 것을 증명해 낸 것이다. **A + B = C (충돌의 몽타주):** 에이젠슈타인은 보다 예술적인 독창성과 철학적 체계를 부여했다는 평가를 받고 있다. 女(여자) + 子(남자) = 好(좋아하다), 犬(개) + 口(입) = 口犬(짖다). 서로 상충되는 요소들의 결합에 의해 일반적인 논리적 사고방식과는 달리 원시적인(감각적인) 사고 과정인 ‘연상적 사고(image thinking)’에 의해 새로운 의미, 개념으로 전환된다(일반적인 개념이 아닌 새로운 개념으로의 전환). 서사를 이어가거나 리듬을 만들어 내는 선에서 멈추지 않고 쇼트들을 연결할 때마다 일종의 충돌이 일어나며 앞 쇼트와 이어지는 쇼트가 충돌할 때 거기서 발생하는 충격은 관객의 마음에 전혀 예기치 않은 갈등과 긴장감(suspense)으로 자리 잡고 있어서 그 갈등은 하나의 새로운 관념(인식)에 대한 지각을 낳는다는 것이다.¹³⁾ **몽타주, 간격 이론:** 베르토프는 영화에 대한 철학적, 근본적인 사유를 통해 진정한 의미의 새로운 예술형태를 모색하였다.¹⁴⁾ 영화 눈(Kino-Eye)이 있는 그대로의 현실을 포착해낸다면, 몽타주는 포착한 현실들을 가지고 새로운 현실(Kino-Pravda)을 열어보여야 한다. 에이젠슈타인과 달리 쇼트 자체가 어떤 의미나 충격을 전달해 줄 수 있으며 쇼트내의 간격, 빈틈, 지체 등은 몽타주에 의해 완전히 조직화되지 않는 사실의 흔적을 가지고 있는 것이므로 물질적 재료의 진정성을 가지고 있는 것이다. 바로 이 쇼트 이미지 내부의 균열과 간격에서 현실의 의미가 도출되어 나올 수 있다. 그리고 간격의 두 번째 측면인 병치란 쇼트와 쇼트의 연결인 몽타주를 의미한다. 편집(몽타주) 없는 영화는 있을 수 없기 때문에 모든 영화에는 간격이 있다. 하지만 대부분의 영화에서 간격 또는 차이는 애써 서사적 재구성의 맥락 속으로 통합시킨다. 하지만 그는 이 간격을 일부러 극대화한다. 이는 두 연속적인 이미지들 간의 거리두기를 표시하는 것이 아니라, 반대로 멀리 있는 그리고 우리의 인간적 지각의 관점에서 볼 때 같은 정도(공통된 단위)를 가질 수 없는 두 이미지들의 상호관계를 표시한다는 것이다.¹⁵⁾

12) 몽타주는 프랑스어의 '조립한다'에서 나온 말로 본래는 건축용어. 이미 1915년 독일의 다다이스트 사진가 존 하트필드가 사진의 단편을 붙여 모아 포토몽타주(공간몽타주)라고 호칭했고, 영화에서는 1920년대 초기 프랑스의 이론가, 특히 무식나크에 의하여 사용됨(시간몽타주).

13) 남명자, 세르게이 에이젠슈타인의 영화편집이론에 대한 고찰, 인문논총 7, 1997

몽타주 이론은 영화를 기술론에서 예술론으로 승화 시켰을 뿐만 아니라 연극(또는 소설)의 플롯중심의 극적 구성방식에서 벗어나 가장 영화(영상)적인 언어를 만들어 냈다고 평가 받았으나 더 이상 크게 발전하지 못했다.¹⁶⁾ 이후 실험적인 기록 영화, 전위 영화, 미디어 아트(영상), 그리고 보다 대중적인 뮤직 비디오나 CF 등 현대적 영상 또는 미디어에서 그 명맥을 찾아 볼 수 있는데, 이는 몽타주 방식의 표현이 보다 진보적인 영상 언어임을 입증해 주는 역설적 예이며, 더 나아가 뉴미디어에 보다 적합한 영상 구성 방식이라 말할 수 있다. 그렇다면 몽타주의 어떤 특성이 뉴미디어와 더욱 밀접한 연관성을 띠는 걸까? 이를 가변성에서 찾아볼 수 있다. 앞서 살펴본 다양한 몽타주 이론은 분절된 이미지(객체)¹⁷⁾들을 재편집하여 그 구성 방식에 따라 상이한 내용, 의미, 결과를 만들어낼 수 있다는 점에서 가변성이라는 뉴미디어의 특성을 가진다. 즉, 같은 DB에서 영상을 어떻게 구성하느냐에 따라 내용을 강화시키거나 상반된 의미 또는 제3의 결과를 지어낼 수 있다는 점이 데이터베이스 서사를 가능하게 하며 다양한 판본을 생성해준다. 이 시점에서 앞의 2-1에서 언급한 ‘뉴미디어의 원리’중 ‘가변성(뉴미디어 원리4)을 보다 자세히 살펴볼 필요가 있다. 뉴미디어 객체는 하나로 고정된 것이 아니라, 잠재적으로는 서로 다른 무한한 판본으로 존재할 수 있다. 기존 미디어는 텍스트, 시각적 요소 그리고 청각적 요소들을 특별한 구성, 순서로 한번 저장되면 영원히 지속되었다. 복사본들은 모두 동일하다. 이와 대조적인 특성을 가지고 있는 뉴미디어는 가변성을 띤다. 이러한 가변성의 원리는 미디어 기술의 변화가 역사적으로 사회 변화와 어떻게 연결되는지를 보여 준다. 옛 미디어가 산업사회의 논리에 상응하는 것이라면, 뉴미디어는 순응보다는 개성을 존중하는 후기 산업사회의 논리에 부합한다. 모든 시민은 다수의(그러나 무한하지는 않은) 선택으로부터 자신의 맞춤 생활방식을 구축하고 자신의 이데올로기를 선택할 수 있다. 이런 식으로 뉴미디어 기술은 독특한 개인들로 이루어진 이상사회라는 유토피아를 가장 완벽하게 실현한다.¹⁸⁾ 앞에서 살펴본 몽타주 이론은 그 각각이 하나로 수렴하려는 선형적인 행진이 아닌 서로 다른 고유의 미적 변수를 지닌 다양한 언어(형식)로 많은 가능성을 내포하고 있다. 그들의 다양한 시도가 독창적인 결과를 가져왔다면, 본 논문에서 제안하는 시스템은 사용자가 단순히 몽타주 방식으로 시각 메시지를 구성해볼 뿐만 아니라, 보다 손쉽게 다양한 시도를 탐구해 나갈으로써, 열려있는 가능성을 독창적인 자신만의 영상 언어(몽타주)로 발전시켜 나갈 수 있는 편리하고 효율적인 환경 제공을 목표로 한다.

14) 자료필름을 해체하고 배치하며 거기서 이미지를 추출했으며 관객이 능동적이고 주체적으로 영화를 수용 할 수 있도록 해야 한다고 주장함. 그리고 영화에서 의미를 보여주는 방식이 반드시 선형적이거나 통합체적일 필요는 없으며, 뒤섞일 수 있고 유사한 것들로 대체할 수 있다고 생각함.

15) 이지영, 이미지의 물질성과 내재성에 대한 연구 - 지가 베르토프(Dziga Vertov)의 영화 이론을 중심으로 -, 174, 2004

16) 스탈린 시대로 접어들면서 실험적인 영화들은 형식주의라는 비난을 받음. 미국 또한 배척하여 현재 대표적인 상업영화(할리우드)들은 플롯중심의 고전적 이야기 구성방식을 주로 사용함.

17) 분절된 이미지(객체)는 ‘모듈성(뉴미디어 원리2)’을 됴.

18) 레프마노비치, 서정신역, 뉴미디어의 언어, 86-87, 생각의나무, 2004

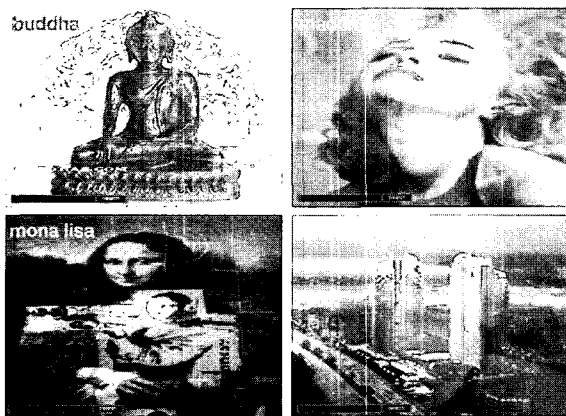
3. 관련 작품연구

3-1. 한 단어 영화 (onewordmovie)

한 단어 영화는 텍스트 기반 영상 검색 엔진을 이용하고 있다.¹⁹⁾ 사용자가 텍스트로 질의하는 영상을 순차적으로 보여 줌으로써, 마치 단어에 해당되는 영상들이 이어지는 영화를 보는 듯한 느낌을 준다. 한 단어 영화는 인터넷에서 가장 대중적으로 쓰이는 검색엔진(Google)의 이미지 검색을 이용해 데이터베이스를 구성하였고, 사용자가 텍스트로 질의를 하게 되면 DB에서 그에 맞는 영상을 검색해서, 이에 해당하는 영상들을 미리 결정된 우선순위에 의해 순차적으로 보여준다. <그림 2-1>은 온라인 플랫폼의 사용자 화면이다.



< 그림 2-1 > '한 단어 영화' 사용자 화면 (온라인 플랫폼)



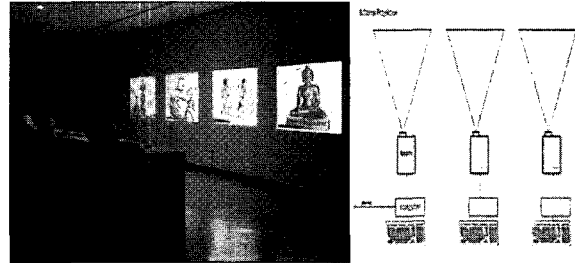
< 그림 2-2 > 각각 4개 단어의 상영 모습

<그림 2-2>는 각각 4개 단어의 상영 모습을 모은 것이다. 기획 브로글(Brogie), 프로그램 집머만(Zimmermann)이며 한국, 네덜란드, 독일, 싱가포르 및 일본 등 세계 각지에서 전시 되 호평을 받은 바 있다. 온라인 플랫폼은 일본 넷아츠(netarts) 2005²⁰⁾에서 대상을 받았다. <그림 2-3 >는 2004 미디어 시티 서울(3rd Seoul International Media Art Biennale)과 전주 국제 영화제(The Jiff Festival)에 전시 상영된 모습 및 시스템

19) <http://www.onewordmovie.ch/>

20) The Project netarts.org is an online-museum by the Machida City Museum of Graphic Arts, Tokyo. <http://www.netarts.org/>

구성도이다. <그림 2-4>는 공공장소에서 휴대폰 SMS 서비스를 이용한 상영 모습이다. 아무런 가이드도 필요 없는 직관적인 인터페이스로, 미디어 아트 경험이 없는 사람들도 열성적으로 참여하여 자신의 고유 키워드를 바탕으로 개인적 트레일러를 가질 수 있지만, 한 단어 이상의 의도는 전혀 반영할 수 없다. 그리고 근본적으로, 텍스트 기반 검색은 영상을 색인한 사람의 주관성을 배제시킬 수 없다.



< 그림 2-3 > 직: 전시 상영 모습 우: 시스템 구성도



< 그림 2-4 > 공공장소(public space) 상영 모습

3-2. 소프트 시네마 (soft cinema)

소프트 시네마는 데이터베이스 서사방식의 뉴미디어 영화이다.²¹⁾ 맞춤 생활방식의 후기산업사회에서 사용자는 반드시 기존 영화의 형식을 따라갈 필요가 더 이상 없다는 관점에서 출발한다. 보다 간략하게 말하자면, 소프트웨어 같은 영화, 영화 같은 소프트웨어이다.

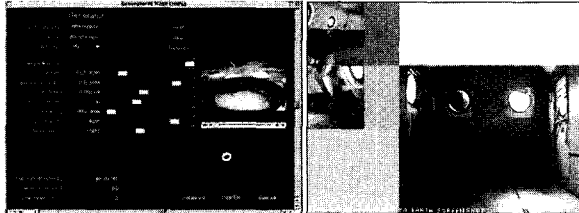
id	word	time	score	url	type	status	category	description	keywords
1	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
2	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
3	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
4	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
5	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
6	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
7	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
8	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
9	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
10	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
11	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
12	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
13	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
14	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
15	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
16	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
17	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
18	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
19	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
20	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
21	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
22	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
23	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
24	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
25	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
26	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
27	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
28	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
29	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
30	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
31	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
32	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
33	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
34	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
35	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
36	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
37	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
38	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
39	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
40	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
41	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
42	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
43	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
44	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
45	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
46	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
47	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
48	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
49	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple
50	apple	0.0000000	0.0000000	http://www.apple.com	apple	no	medium	apple	apple

< 그림 3-1 > 비디오 데이터베이스 : keywords view

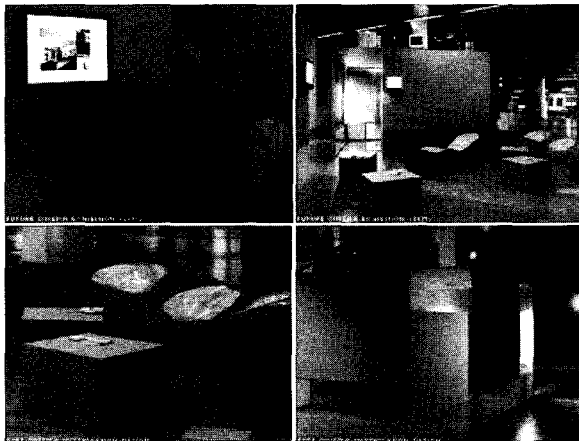
우선 비디오클립의 속성(activity, average grays, distance, geolocation, sub motion, cam motion, discription 등)을 분류한다. 그리고 이를 수치와 텍스트로 규칙을 부여한 DB로 구축한다(그림 3-1 참조). 사용자는 작성기(Generator)에서 이

21) <http://www.softcinema.net/index.htm>

러한 비디오 클립의 속성 값을 조절, 선택해 영상을 구성한다(그림 3-2 좌 참조). 소프트 시네마는 2002 - 2003년 독일 제트케이엠(ZKM)의 미래 영화(Future Cinema) 전시에서 상영되었다(그림 3-3 참조). 또한 2004년 세네프(SeNef)에 초청받아 삼성 미디어 라운지 전시에서 벽걸이용 TV 한 대당 1편씩 총 3편(Texas, Mission to Earth, Absences)이 상영되었다. 이 3편의 영화는 2005년 엠아이티 출판사(MIT Press)에서 DVD로 발매되었다(그림 3-2 우 참조).



< 그림 3-2 > 좌: 작성기(Generator) 우: 상영 화면(DVD, 2005)



< 그림 3-3 > 미래 영화(Future Cinema) 전시 상영 모습

소프트 시네마는 '뉴미디어의 언어' 저자 레프 마노비치의 연구 작품이다. 이는 알고리즘을 이용한 소프트웨어 같은 형태로 몇몇 선택을 통해 영화가 자동으로 구성된다. 그리고 같은 DB에서 사용자의 의도에 따라 가변 판본을 생성할 수 있는 데이터베이스 서상방식을 사용한다. 또한 상영 포맷도 멀티스크린 포맷으로 보다 많은 내용과 정보의 공간 몽타주 구성이 가능하다(그림 3-2 우 참조). 그러나 세네프(SeNef) 전시 상영에서 사용자 인터페이스는 제공되지 않았다. 이는 소프트웨어 사용에 가이드가 필요할 뿐만 아니라 보다 많은 경험이 요구되기 때문으로 분석된다.²²⁾ 제트케이엠(ZKM) 전시의 경우도 사용자 인터페이스가 제공되지 않은 것으로 추정된다.²³⁾ 물론 DVD에도 소프트웨어는 제공되지 않았다. 이는 실제 전시 상영 시 소프트 시네마가 가지고 있는 장점을 발휘하지 못하고 결국 기존 영화처럼 관람만 해야 한다는 이야기다. 그리고 DB의 경우 멀티미디어 데이터의 성격상 속성들을 소수의 키워드로 다 표현하기 어렵다는 단점을 가지고 있다. 그밖에 DB 구축을 위한 수작업 색인, 필요한 비디오 클립 제작 등 많은 시간과 노력이 필요하다.

22) DB를 구축한자(제작자)가 아니라면 이에 대한 파악이 어려움.
23) 사진 및 자료를 분석한 결과.

4. 실시간 몽타주 시스템

본 논문에서 제안하는 시스템은 앞서 살펴본 관련 작품의 장단점 연구를 기초로 하여, 기획 단계부터 쉽고 직관적인 인터페이스, 사용자의 의도를 반영할 수 있는 시각 메시지(몽타주) 구성 방법 그리고 관람, 감상 시 상호작용하는 재미를 가장 우선시 하여 설계하였다.

이 시스템은 크게 3부분으로 구성된다. 사용자가 선택한 영상을 이용해 유사한 영상을 검색하는 내용 기반 영상 검색 부분과 사용자의 영상 삽입을 위한 실시간 움직임 감시 부분, 그리고 결과 영상을 보여주는 부분으로 구성되어 있다. 본 장에서는 내용 기반 영상 검색 부분과 실시간 움직임 감시 부분을 따로 나누어 설명하고, 그 결과 영상을 보여주는 부분은 실시간 몽타주 시스템 고찰에서 함께 기술한다.

4.1. 내용 기반 영상 검색

앞서 2-2에서 살펴본 것처럼 디지털 형태로 변환되고 저장되는 이미지, 즉 영상 자료들에 대한 검색 방식은 문서를 검색하는 방식과는 접근 방법이 크게 다르다. 텍스트에 기반하여 이미지를 표현하고 검색하고자 했던 과거의 이미지 검색 방식은 사용자에게 직관적이지 못하고, 영상을 텍스트로 표현한다는 것 자체가 주관성의 개입여지가 매우 많기 때문에 검색에는 적합하지 않다.²⁴⁾ 그렇기 때문에 이미지의 특징을 직접적으로 이용해서 검색하는 내용 기반 영상 검색(CBIR : Contents Based Image Retrieval) 기술에 대한 연구가 활발히 진행되고 있다. 이미지의 특징을 직접적으로 표현하는 방식은 크게 색상(Color), 형태(Shape), 질감(Texture)으로 나눌 수 있다.²⁵⁾ 이 중 색상과 형태를 이용한 방식은 영상의 종류(그림, 사진 등)에 따라 매우 다른 결과를 보여주는 경우가 있기 때문에, 본 시스템에서는 여러 영상에 적용이 용이하고 구현이 간단한 질감 기반 검색 방식을 이용하였다. 질감 기반 검색 방식에도 여러 가지가 있지만, 효과적인 질감 기반 검색 기법으로 검색 기법으로 널리 알려진 주파수 영역으로의 변환을 통한 비교 방식을 사용하였다.

사용자가 선택한 영상과 데이터베이스에 있는 영상들을 주파수 영역으로 변화시키기 위해서는 일반적으로 식 (1)과 같이 2차원 푸리에 변환(Fourier Transform)을 이용하게 된다.

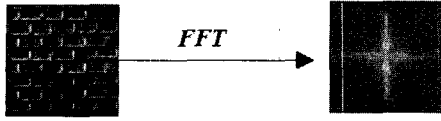
$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-2\pi i (ux + uy)} dx dy \quad (1)$$

본 시스템에서 사용하는 영상은 디지털화된 영상이기 때문에 식 (2)와 같이 푸리에 변환의 이산적 형태인 이산 푸리에 변환(Discrete Fourier Transform)을 사용해야 하지만 계산의 효율성을 위해서 이산 푸리에 변환의 반복 연산을 효율적으로 개선한 고속 푸리에 변환(Fast Fourier Transform)을 이용해 영상을 주파수 영역으로 변환한다.

24) Kjersti Aas and Line Eikvil. A survey on: Content-based access to image and video databases. Report 915, Norwegian Computing Center, March 1997.

25) Atsuo Yoshitaka and Tadao Ichikawa. A survey on content-based retrieval for multimedia databases. IEEE Transactions on Knowledge and Data Engineering, 11(1):8193, January/February 1999.

$$F(u, v) = \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) e^{-2\pi i \left(\frac{ux}{N} + \frac{vy}{N} \right)} \quad (2)$$



< 그림 4 > Fast Fourier Transform

<그림4>는 원 영상을 FFT를 이용해 주파수 영역으로 변환한 영상이다. 이와 같이 주파수 영역으로 변환시킨 영상들을 다시 $N \times N$ 크기의 영상으로 변환시킨 후, 식 (3)을 이용해 $b \times b$ 크기의 영역으로 나누고, 각 영역의 명암 값의 합(B_n)을 구한다.

$$B_n = \sqrt{\sum_{n=0}^{N^2-1} \sum_{u=nb}^{nb+b-1} \sum_{v=nb}^{nb+b-1} F(u, v)} \quad (3)$$

각 영역의 명암 값의 합을 다른 영상과의 비교를 정확히 하기 위해 식 (4)를 이용하여 정규화(Normalization)하면, 이로부터 영상 전체에서 해당 영역이 차지하는 비중(C_n)을 구할 수 있다.

$$A_n = \frac{N^2-1}{b} B_n, \quad C_n = \frac{B_n}{A_n} \quad (4)$$

이와 같은 과정을 사용자가 선택한 기준 영상과 데이터베이스에 있는 모든 영상에 적용한 후, 기준 영상과 데이터베이스의 영상간의 유사도 비교를 한다. 유사도는 식 (5)와 같이 유클리디안 거리(euclidean distance)를 이용해 빠르고 효과적으로 비교하게 된다. 비교 결과 유사한 영상일수록 유클리디안 거리가 작게 되고, 이 순서대로 영상 순서를 결정하게 된다.

$$D_n = \sum_{n=0}^{N^2-1} C_n^q - C_n^d \quad (5)$$

4-2. 실시간 움직임 감지

컴퓨터 비전이란 여러 매체(Voltage, Ampere, Frequency, 시간, 온도, 위치 데이터)중에서 카메라를 이용하여 2차원의 위치 데이터를 입력 받아 검사, 제어하는 방법으로서 비 접촉 방식이며, 인간의 육안 검사와 비슷한 역할을 하는 장치를 말한다. 이러한 컴퓨터 비전(Computer Vision) 분야에서 가장 활발하게 연구되고 있는 분야 중 하나가 영상 트래킹(tracking) 부분이다. 물체 인식, 움직임 추출, 영상 검색 등 다양한 분야에서 연구, 활용하고 있다.²⁶⁾ 일반 센서의 물리적 한계로 인한 대안으로 사용된 모션 트래킹은 센서와 유사한 결과를 얻을 수 있지만 회로를 이용한 센서나, 무선 센

26) J.K. Aggarwal and Q. Cai. Human motion analysis: a review. Computer Vision and Image Understanding, 73(3):295304, 1999.

서 그리고 CATV 카메라, CCD카메라, 아주 저렴한 가격의 웹 캠(PC CAMERA)으로 카메라 렌즈의 넓은 화각을 이용하여 보다 확장된 공간 영역을 동시에 통제할 수 있다는 점을 장점으로 들 수 있다.

본 논문에서는 사용자의 영상의 삽입 시기를 결정하는 방법으로 차분영상 기반 영상 트래킹을 이용한다. 본 시스템에서는 사용자가 움직일 때마다 사용자의 영상을 삽입하게 되는데, 움직임을 파악하기 위해서는 카메라로 입력되는 영상의 변화를 감지해야 한다. 식 (6)과 같이 현재 영상과 이전 영상 각 화소의 명암 값을 비교해서 그 차이가 일정 한계치 이상이면 해당 화소에 변화가 있는 것으로 감지하도록 하였다.

$$\text{if } |f(x, y) - F(x, y)| > 30, \quad P(x, y) = 1$$

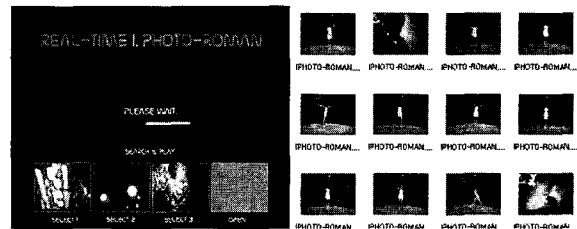
$$\text{else } P(x, y) = 0 \quad (6)$$

식 (7)을 이용해서 한계치를 넘는 픽셀의 수(T)가 사용자가 설정한 수 보다 많게 되면, 사용자가 움직인 것으로 추정해 이 영상을 실시간으로 저장하고, 저장된 영상을 화면에 보여 준다.

$$T = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} P(x, y) \quad (7)$$

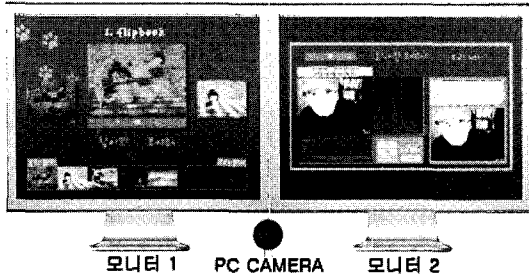
4-3. 실시간 몽타주 시스템 고찰

본 논문에서 제안하는 실시간 몽타주 시스템은 현재까지 아이 포토로망(I. PHOTO-ROMAN)과 아이 플립북(I. flipbook) 두 개의 버전이 있다. 연구의 진행 사항을 살피기 위해 먼저 첫 번째 버전을 간단히 살펴본 후 두 번째 버전을 보다 자세히 고찰해본다.



< 그림 5 > 아이 포토로망(I. PHOTO-ROMAN, Ver1.0) 검색화면과 실행 결과, 중앙대학교 공연 영상제 전시 상영, 2005

아이 포토로망(I. PHOTO-ROMAN)은 불특정 다수의 사진을 DB로 사용하고 하나의 이미지만 질의 할 수 있다. 사용자 임의의 이미지를 불러오거나 화면에 제공된 샘플을 선택하여 검색을 수행한다. 검색 순위에 따라 영상을 순차적으로 보여줌으로써 자동 구성된 이미지들은 시각적 패턴, 리듬을 형성한다. 여기에 사용자의 영상을 실시간으로 삽입함으로써 사용자는 검색된 영상과 사용자의 영상이 마치 하나의 시퀀스로 연결된 것처럼 보이게 되며, 영상구성에 의한 의미를 창출함으로써 몽타주 효과를 느낄 수 있게 된다(그림 5참조). 아이 플립북 (I. flipbook, Ver2.0)은 '에이치시아이♥키즈(HCI♥KIDS)' 전시를 위한 타겟 지향형 시스템으로 설계되었다. <그림 6-1>과 같이 모니터 2대와 PC카메라 1개로 구성되어 있다.

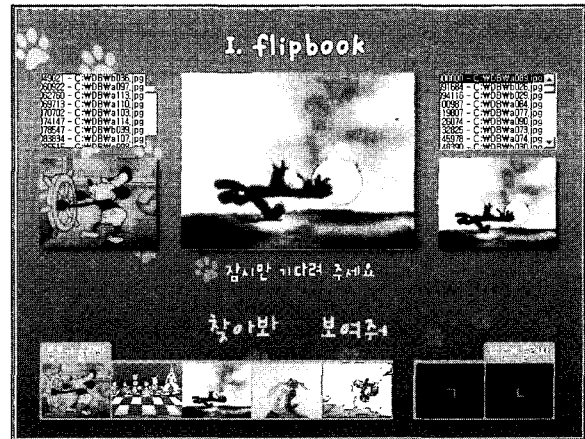


< 그림 6-1 > 아이 플립북(I. flipbook, Ver2.0) 시스템 구성

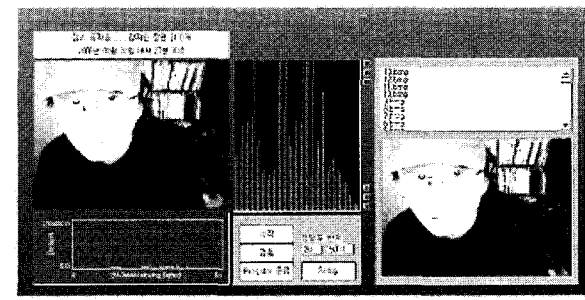
<그림 6-2>는 시스템 모니터 1번, 검색화면 이다. 사용자가 검색을 하고자 하는 기준 영상을 □□ □□ 선택 또는 사용자가 다른 임의의 이미지를 직접 불러오고, 검색을 원하는 데이터베이스를 선택할 수 있다. 선택된 영상과 유사한 영상을 데이터베이스에서 자동으로 검색해서 유사한 영상일수록 우선순위를 주게 된다. 질의 영상과 데이터베이스의 영상은 256 x 256 크기로 일정하게 변화시켰고, 비교를 위해 분할하는 영역의 크기는 16 x 16으로 하였다. 또한 사용자의 움직임 을 실시간으로 감시하는 부분에서는 변화된 화소의 양이 전체의 20%가 되면 움직임이 있는 것으로 판단하도록 하였고, 사용자가 언제든지 수치를 조절할 수 있도록 하여 사용자 움직임 모습의 입력 정도를 변경할 수 있도록 하였다. <그림 6-3>는 모니터 2번, 사용자의 움직임을 감시하는 화면 이다.

검색이 끝나고 실시간 몽타주 시스템을 실행시키면 <그림 6-4>와 같이 애니메이션으로 구성된 5개 카테고리 500개의 데이터베이스 영상 중 검색된 영상을 우선순위에 따라 몽타주 방식으로 □□ □□을 병치하여 보여준다. 이 때 사용자가 움직이게 되면 사용자의 움직임을 실시간으로 검색된 영상 사이에 삽입하게 된다. <그림 6-5>는 시스템 실행 시 순차적으로(약 0.3초 간격) 보여주는 영상을 나열한 것이다. 가장 처음 나오는 영상이 질의 □□에 의해 검색된 영상이다. 이후 영상은 질의 □□에 의해 검색된 영상이다. □□ □□은 약 4초 마다 병치되며 보여준다. 이를 이용하면 사용자 의도에 따라 몽타주 방식의 짧은 영상메시지를 구성할 수 있다. 생쥐 미키의 휘파람부는 모습과 고양이 펠릭스의 뛰는 모습이 계속 부딪치면 아이들은 아주 자연스럽게 긴장감과 호기심을 가지게 된다. 이는 가장 단순한 차원의 의미생성이라 말할 수 있다. <그림 6-5>를 보면 사용자의 움직임이 있을 때마다 검색된 영상 사이에 사용자의 모습이 삽입 된 것을 알 수 있다. 이는 아이들에게 또 다른 재미를 줄 뿐만 아니라 만화 중간 중간에 나오는 자신의 모습을 보고 자신이 만화 속에 나온다는 착각을 일으키기도 했다. 이는 단순화된 플로토타입 모델에 가장 걸맞은 몽타주 효과의 한 예라 할 수 있다. 아이들은 비교적 짧은 시간 안에 의도를 가지고 필요한 영상을 찾아내고 손쉽게 영상을 재구성 할 수 있다. 그리고 자신의 얼굴 표정, 제스처 등을 가지고 이를 즉각적으로 발전시켜 나갔다. 축약된 형태로나마 영상제작 전 과정을 경험했을 뿐만 아니라 이 과정 자체를 즐겼다. 사용자의 시간과 노력을 줄여 줌으로써 보다 유희적으로 접근하고, 더욱 감각적으로

반응하며 독특한 개성을 발휘할 수 있도록 도움을 주었다.



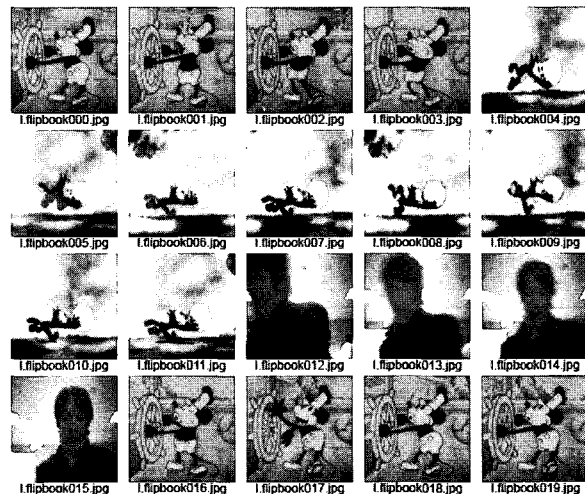
< 그림 6-2 > 내용 기반 영상 검색 화면



< 그림 6-3 > 실시간 움직임 감시화면



< 그림 6-4 > DB를 구성하고 있는 영상들



< 그림 6-5 > 결과 동영상을 순차적으로 영상을 나열한 것

본 시스템은, 앞에서 살펴본 두 가지 버전처럼, 그 목적에 따라 인터페이스를 달리하고 기능을 추가시키거나 빼서 용도에 맞는 새로운 버전으로 손쉽게 업그레이드 할 수 있는 모듈 또는 틀 적 성격을 띤다. 이는 사용자의 경우도 마찬가지다. 개개인의 차이, 반복 경험의 정도에 따라 영상을 선택하

거나 임의의 이미지를 불러와 질의 할 수 있으며 DB를 교체하여 사용자 개인이 소유하고 있는 다른 영상들을 언제든지 스스로 활용할 수 있다. 자신의 모습까지 원하는 시점에 삽입시킴으로써 상호작용을 극대화한 자신만의 작품, 개성 넘치는 영상 메시지를 손쉽게 구성할 수 있다. 즉 또 다른 의미의 1인 미디어(미디어기반 엔터테인먼트)로써 무한한 가변 판본을 만들어 낼 수 있다.



< 그림 6-6 > '에이치시아이♥키즈' 아이 플립북(Ver2.0) 전시 모습

5. 결론

본 논문에서는, 내용기반 영상검색과 실시간 움직임 감시를 통해, 기존의 영상들을 사용자가 의도하는 방향으로 쉽게 시각 메시지로 구성하며 즐길 수 있는 실시간 몽타주 시스템에 대해 기술하였다. '에이치시아이♥키즈(HCI♥KIDS)'에 확인된 바와 같이 비교적 짧은 시간 안에 아이들이 의도를 가지고 손쉽게 영상을 재구성 하고 자신의 얼굴 표정, 제스처 등을 가지고 영상 구성을 즉각적으로 발전시켜 나가며 이 과정 자체를 즐겼다는 점이 의미하는 바가 크다.

그러나 보다 다층적이고 복잡한 내용구성은 아직 어렵다. □□□□ 그리고 사용자 움직임(제스처, 표정)만으로 구성할 수 있는 영상메시지는 짧을 수밖에 없다. 만약 찾아낸 양이 많다고 계속 보여주게 되면 초반에 느꼈던 영상에 대한 긴장감(suspense)은 곧 사라질 것이다. 사용자의 의도를 정확히 반영할 수 있는 인터페이스, 물리적 요소, 특히 새로운 시점을 생성²⁷⁾하고 내용을 증강시킬 수 있는 가상 객체 삽입²⁸⁾ 등의 확장이 절실히 필요하다. 그리고 1-2분미만의 검색 속도는 느리지 않지만 더 많은 데이터베이스를 검색할 경우 상대적으로 더 빠른 속도와 보다 광범위한 데이터베이스를 이용하기 위해 네트워크 기술이 필요하다. 인터넷은 거대한 하나의 미디어 데이터베이스이다. 물론 서로 떨어져 있는 제 3자와 공동창작도 고려해 본다. 향후 이와 같은 기능 및 기술을 구현 할 수 있는 연구를 지속하여 프로튜어²⁹⁾의 빠른 사전작업, 아이디어 중심의 독창적 저작과 실험 행위 (live

27) MPI(Multiple Perspective Interaction) 여러 시점에서 취득한 2차원 영상을 이용해서 3차원 복원 없이 사용자가 원하는 새로운 시점의 2차원 장면을 생성하는 기법

28) AR(Augmentd Reality) 실세계에 3차원 가상물체를 겹쳐 보여주어 현실을 증강시켜주는 기술

29) LG경제연구원, Proteur : Professional + Amateur

performance)를 도와줄 수 있는 환경을 제공할 계획이다.

1920년대 러시아 혁명기 예술가(영화인)들의 다양한 영화적 시도가 몽타주 이론을 통해 독창성과 가능성을 열어놓았다면, 디지털 혁명이라 불리 우는 오늘날, 사용자가 기존의 틀이나 외부적 환경 요인에 얽매이지 않고 보다 편리하게 그리고 자유롭게 다양한 영상 구성을 할 수 있도록 함으로써 열려있는 가능성을 자신만의 영상 언어로 발전시켜 나갈 수 있는 단초를 제공하려 노력했다. 본 논문에서 제안하는 시스템은 다수의 선택으로부터 자신의 맞춤 생활방식을 구축하고 자신의 이데올로기를 선택하며 독특한 개성을 존중하는 후기산업사회의 욕구에 대응하는 뉴미디어 영상 디자인 틀이다. 이는 누구나 영상의 수용자에서 벗어나 적극적 주체로서 자신만의 영상 언어를 효과적으로 구성 할 수 있는 기술적 예술적 토대를 마련함으로써 앞으로 새로운 영화(미디어기반 엔터테인먼트)에 한발 더 다가가기라 기대된다.

참고문헌

- Glenn Becker. Information in Images, chap6: Content-based Query of Image Databases. Thomson Technology Labs, 1997. Online book.
<http://www.thomtech.com:80/mmedia/tmr97/tmr97.html>
- Kjersti Aas and Line Eikvil. A survey on: Content-based access to image and video databases. Report 915, Norwegian Computing Center, March 1997.
<http://www.nr.no/home/kjersti/video.html>
- Atsuo Yoshitaka and Tadao Ichikawa. A survey on content-based retrieval for multimedia databases. IEEE Transactions on Knowledge and Data Engineering, 11(1):8193, January/February 1999.
- J.K. Aggarwal and Q. Cai. Human motion analysis: a review. Computer Vision and Image Understanding, 73(3):295304, 1999.
- 김낙우 최종수, 내용 기반 영상 검색 기술의 현황 및 전망, 한국멀티미디어학회지 제8권 제2호 2004. 06
- 이지영, 이미지의 물질성과 내재성에 대한 연구: 지가 베르토프(Dziga Vertov)의 영화 이론을 중심으로, 2004
- 이지현, 영상에 있어서 몽타주 이론에 관한 고찰, 디지털디자인학연구 제8권, 2003
- 정현두, 소비에트 영화의 태동과 지가 베르토프, 376, 노어노문학 제14권 제2호, 2002.11
- 남명자, 세르게이 에이젠슈타인의 영화편집 이론에 대한 고찰, 인문논총 제7집, 1997
- 지가 베르토프, 김영란역, 키노아이 : 영화의 혁명가 지가 베르토프, 이매진, 2006. 01
- 랄프 슈넬, 강호진역, 미디어 미학 : 시청각 지각형식들의 역사와 이론에 대하여, 이론과실천, 2005. 10
- 레프마노비치, 서정신역, 뉴미디어의 언어, 생각의나무, 2004
- 이원곤, 디지털화 영상과 가상공간 : '새로운 예술'의 전개를 위한 지평, 연세대학교 출판부, 2004.02
- 교육, 디지털 스토리텔링, 황금가지, 2003.10
- S. 채트먼, 한용환역, 이야기와 담론, 푸른사상, 2003.09
- 박성수, 디지털 영화의 미학, 문화과학사, 2001.03
- 김용수, 영화에서의 몽타주 이론, 열화당, 1996.09