

# 특성정보를 이용한 영상기반 인터넷 동영상 검색 시스템의 구현

손정식\*, 이원석  
연세대학교 컴퓨터과학과

## An Implementation of a Image-based Internet Video Search System using Feature Information

JeongSig Sohn\*, WonSuk Lee  
Dept. of Computer Science Yonsei University

### 요 약

본 논문에서는 인터넷 동영상 정보들을 보다 효율적이고 정확하게 검색할 수 있는 시스템을 구현하기 위해, 특성 정보에 기반한 영상기반 검색과 예제를 통한 질의를 사용한 시스템을 제안한다. 기존의 검색 시스템과는 달리 본 시스템에서는 사용자가 질의로 제출한 동영상의 특성정보를 추출하여 데이터마이닝하여 프로파일로 만들고, 인터넷에서 검색된 동영상의 특성정보와 비교하는 기법을 사용한다. 또한 질의 동영상 프로파일들을 통합하고, 특징 영역 세그먼트들의 중요도를 저장해 둬으로써 사용자의 질의를 하나의 척도로 처리 하며, 작성된 프로파일과 비교할 동영상의 특성정보를 영역별로 가중치를 주어 보정함으로써 검색을 원하는 동영상의 특징을 강조한 질의가 가능하도록 한다. 끝으로 가중치를 주는 방법의 변화에 따른 시스템이 출력하는 검색 결과의 변화를 관찰하여 보정 방법의 성능을 평가한다.

### 1. 서론

인터넷 망을 이용한 동영상 서비스가 쉬워지게 되고 인터넷의 각 분야에서 서비스되고 있는 동영상의 수가 증가함에 따라 과거 문서 검색 시스템과 유사하게 필요한 동영상을 찾을 수 있도록 도와줄 수 있는 검색 시스템의 필요성이 대두 되었다. 그러한 필요성에 따라서 동영상 검색 엔진이 등장했으나 그 중 상당수의 사이트들은 동영상 링크를 가리키는 문자나 화일명을 검색어로 사용한다. 이러한 시스템으로도 효과를 거둘 수는 있지만, 실제 내용을 보지 않고 그 내용을 짐작하는 점에서 한계를 가지게 된다. 이러한 한계를 극복하기 위해 영상기반 검색 방법이 대두되었고 많은 방법들이 연구되었다.

본 논문에서는 데이터마이닝을 응용하여, 영상기반 검색을 지원하고 예제를 통한 질의 방법을 도입, 검색자로부터 표본 동영상을 받아 프로파일화 하고

그 프로파일들을 조직화하여 좀더 정확하고 빠른 검색을 돕도록 하며, 검색하고자 하는 동영상의 특징들을 부각시켜 비교 판정할 수 있는 방법들 고안하여 새로운 시스템의 구현을 시도한다.

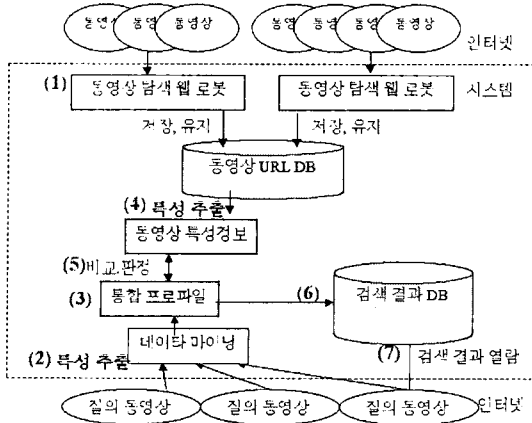
### 2. 관련연구

일반적인 동영상 검색 시스템은 크게 인터넷으로부터 검색대상이 되는 동영상 링크를 검색하여 그 동영상으로부터 특성정보(feature)를 추출하는 부분, 추출된 특성정보를 색인화 시키는 부분, 사용자의 검색명령을 받고 직접 색인과 비교하여 결과를 판정하는 부분으로 나눌 수 있다. 각 부분의 구현을 위해 인터넷의 문서를 자율적으로 수집하는 웹로봇의 연구, 수집된 문서, 콘텐츠를 이해하고 그 색인을 검출해 내고 구조화 하는 연구, 이를 위한 콘텐츠의 특성 정보와 같은 내용 정보를 추출하는 연구[5]가 뒷받침

되고 있다.

기존에 연구된 검색 시스템들을 분류하면 크게 링크, 화인이름과 같은 간접 정보를 비교 검색하는 시스템과 콘텐츠 내용 기반 비교 검색 시스템으로 분류 할 수 있다. 간접 정보를 사용하는 시스템은 대부분 문자 질의 방식을 사용하고[8], 내용 기반 정보를 사용하는 시스템은 문자 질의를 사용하는 것과 [1][2], 예제 질의를 사용하는 것으로 나눌 수 있다 [4][6].

### 3. 인터넷 동영상 검색 시스템



[그림 1] 인터넷 동영상 검색 시스템

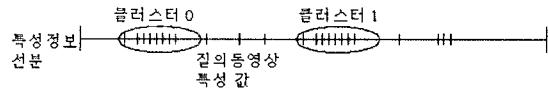
#### 3.1 인터넷 동영상 검색시스템

본 시스템의 작업의 흐름은 그림 1에서와 같다. 먼저 동영상 탐색 웹 로봇이 동영상의 URL을 수집하여 동영상 URL DB를 구축한다. 그리고 사용자들로부터 질의 동영상 집합을 받아 특성정보를 추출, 클러스터링 하여 통합 프로파일로 제작한다. 그 다음 URL DB에서 검색할 동영상의 링크를 받아 접속하여 특성 추출한다. 추출된 특성과 통합 프로파일을 비교하여 어느 사용자의 질의에 가까운지를 비교 판정한다.

질의 동영상의 특성정보를 추출하여 그것을 프로파일(profile)로 제작하고 그것을 인터넷 동영상의 특성정보와 비교하는 부분에 이 논문의 주안을 두었고, 위 그림에서 굵은 선으로 표기되어있다. 다음 절에 이를 구현하는 이론에 대하여 설명한다.

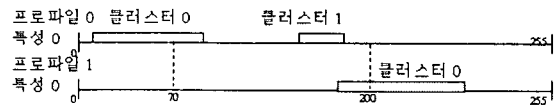
#### 3.2 프로파일 제작

동영상의 특성정보는 색상 원소, 색상 수, 색상 변화량, 질감, 무늬의 종류로 구성되어있다. 질의 동영상을 검색자가 하나만 제시한다면 중요하지 않은 특징들이 의도하지 않은 판정 결과를 가져 오는 경우가 생긴다. 그래서 검색자가 원하는 특징들을 갖춘 질의 동영상을 다수 모아서 제시된 각 동영상들이 모두 갖춘 특징 만이 강조된 프로파일의 필요성이 대두 되었다. 이를 위해 제시된 모든 동영상들을 각각 트랜잭션(transaction)으로 설정하고 그 특성정보들을 클러스터링 하는 데이터마이닝 방법을 사용한다. 데이터마이닝의 기법 중 클러스터링(clustering)을 사용하면 그림 2와 같이 여러 동영상들 사이의 유사한 특징들로 판단되는 것들을 하나의 묶음(cluster)으로 추출하는 것이 가능하다.



[그림 2] 특성 값 군집에 의한 클러스터 생성

이 클러스터들은 앞서 설명한 동영상의 각 특성 정보 값들이 가질 수 있는 값의 범위(domain) 내에서의 한 영역을 나타낸다. 또한 이 영역을 구성하는 질의 동영상의 비율을 지지도(support)라고 한다. 이 영역은 제시된 동영상들이 모여져 강조된 특징이 특성으로 추출될 때 나타나는 값들의 영역으로 정의되고, 다른 동영상의 특성을 추출하여 값을 비교했을 때 그 값이 범위 내에 존재하면 그 동영상은 그 클러스터의 지지도 만큼 그 특징을 가진다 라고 판단할 수 있다.

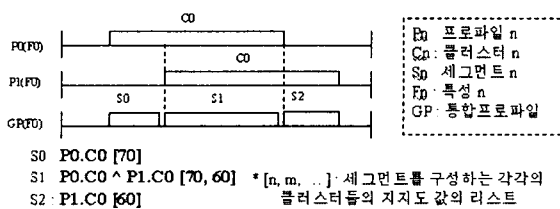


[그림 3] 특성값 선분상에 표시된 클러스터

그림 3은 프로파일0과 1의 특성0에 대한 클러스터가 위와 같이 존재한다고 하고, 이 프로파일에 대해 임의 동영상의 한 프레임을 비교하는 것을 나타낸 것이다. 여기서 동영상의 특성0에 대한 값이 200이라면 프로파일1의 클러스터0에 속하게 되고 클러

스터의 지지도만큼 프로파일과 유사하다고 할 수 있다. 모든 특성에 대하여 각 프로파일 별로 유사도를 구한 평균을 해당 프레임의 유사도라 하고, 짧은 동영상의 경우 모든 프레임의 평균을 그 동영상과 해당 프로파일 간의 유사도라고 할 수 있고, 긴 동영상의 경우 시간에 따른 유사도의 변화를 구할 수 있다. 이러한 클러스터들의 정보를 모아둔 것을 본 논문에서는 프로파일(profile)이라고 한다.

### 3.3 통합 프로파일 구성 및 비교



[그림 4] 프로파일 통합에 의해 세그먼트가 생성되는 과정

그림 3과 같이 통합 프로파일을 제작해 두면 비교 판정 시 통합 프로파일을 나타내는 선분 위에서 비교 대상 동영상의 해당 특성 값이 놓여졌을 때, 그 특성 값이 어느 세그먼트에 속하는지를 한번의 탐색으로 알 수 있고, 즉시 그 세그먼트를 구성하는 클러스터와 그것의 지지도, 그리고 그것이 어느 프로파일에 속하는지를 판단 가능하다. 이를 토대로 유사도식을 구하면 다음과 같다.

$$\text{유사도} = (\text{클러스터의 지지도}) / \text{총 클러스터 수}$$

### 3.4 가중치에 의한 프로파일 보정

표본 동영상으로 프로파일을 만들거나 만들어진 프로파일을 비교하는 과정에서 특징을 강조한다든지, 노이즈를 제거할 수 있는 보정을 가할 필요성이 있다.

프로파일을 제작하면서 유사한 클러스터를 묶어 블럭을 만들게 하고, 블럭 내에 존재하는 클러스터들의 영역 대비 개수를 구하여 그 값이 미리 설정한 최소 값을 넘어서는 블럭을 동적 블럭으로, 그렇지 않은 블럭을 정적 블럭으로 분류 하는 것이 가능하다. 이렇게 프로파일을 정적, 동적으로 구분하고, 정적인 특징을 강조하여 동영상상을 비교할 때는 정적 가중치를, 반대의 경우 동적 가중치를 부여함으로써 화면

내에서 중요한 요소만을 선별하여 절의하는 것이 가능하다.

또한 동영상 화면을 물리적으로 등분하여 그 특징에 해당하는 물체가 존재하는 영역과 그렇지 않은 영역으로 선별하여 가중치를 주는 것도 가능하다. 이 경우도 위와 유사하게 가중치를 준 영역으로부터 추출된 특성정보가 통합 프로파일과 비교 되어 유사도가 계산 될 때 주어진 가중치가 가해져서 유사도의 변화를 크게 하여 판정을 수월하게 한다.

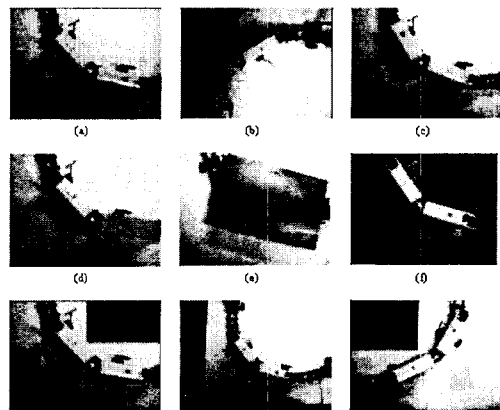
### 3.5 인터넷 동영상 검색 시스템의 구성

본 시스템은 작업량이 많은, 동영상 URL을 수집하는 부분과 동영상의 특성정보를 추출하는 부분이 서로 구분되어 동작할 수 있도록 전체를 다음 세가지 부분으로 분할되어 동작한다.

- URL 수집 부분
- 동영상 특성정보 추출 부분
- 통합 프로파일 구축 및 판정 부분

그리고 특성 추출 부분에서 작업 효율을 높이기 위해서 동시에 여러 개의 추출 프로그램을 실행시킬 수 있도록 하고, 그것을 관리하는 데몬을 둔다. URL DB는 시스템 내부에서 그 접근이 많으므로 검색자가 자신의 검색결과를 살펴볼 수 있는 DB는 독자적으로 둔다.

### 3.6 모의실험 및 결과



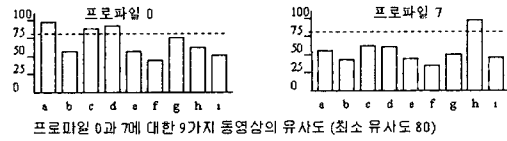
- (a) 열차가 정상적으로 돈다.
- (b) 90도 회전된 레일을 돈다.
- (c) 레일을 거꾸로 돈다.
- (d) 열차에 주황색 증이를 붙이고 돈다.
- (e) 열차에 파란색 증이를 붙이고 돈다.
- (f) 바닥 색깔이 다르다.
- (g) 바닥에 다른 물체들이 놓여있다.
- (h) 50% 줌 아웃(zoom out)한 영상.
- (i) 바닥에 다른 물체가 있고 회전된 레일.

[그림 5] 실험에 사용될 동영상 집합

본 실험에서 사용한 동영상 그림 5에서와 같이 균일한 색을 가진 바닥 위에 설치된 레일을 등속으로 주행하는 실험용 열차를 위에서 바닥을 바라보는 시각에서 촬영한 것을 사용한다.

■ 기본 프로파일에 대한 각 동영상의 유사도

- 시스템의 주 기능에 대한 시연

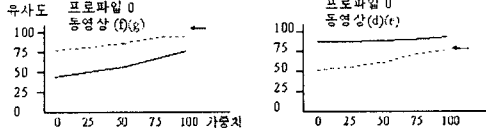


[그림 6] 기본 실험 결과

그림 6은 동영상 a,h로 만들어진 프로파일0,7과 모든 동영상과의 유사도를 나타낸 결과이다. 최소 유사도를 80으로 설정하면 각각 3개,1개의 검색결과가 나온다. 그리고 프로파일0에 가중치를 주어서 유사도의 변화를 얻은 결과를 그림 7,8에서 볼 수 있다.

■ 동적.정적 블럭 가중치 변화에 따른 유사도 변화

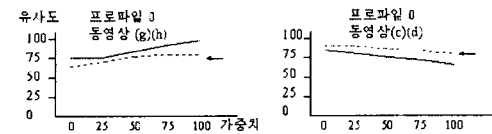
정적 동적 블럭 중 하나만 높은 동영상에 특정 가중치를 변화  
가중치 변화에 의한 유사도의 증가로 보였 효과를 검증



[그림 7] 동적.정적 영역 가중치 실험 결과

■ 영역별 가중치 변화에 따른 유사도 변화

특정 영역만 높은 동영상에 영역 가중치를 변화  
가중치 변화에 의한 유사도의 증가로 보였 효과를 검증



(영상에서 열차가 있는 영역만 가중치를 줌)

[그림 8] 영역별 가중치 실험 결과

[표 1] 영상기반 검색 시스템과 본 시스템의 비교

	분류 대상	영상 추출 방식	비교 방식	특징
[Kao]	MPEG-형식	자막 추출인식	문자 기준	자막만 사용
[Yu]	동영상	신경망 이용 키워드	문자 기준	
[Shih-Fu]	동영상	객체 움직임만 추출	그림 기준	객체의 행동 도시
[Yeh]	정지영상	특성정보 추출	정지영상 기준	정지영상만 지원
본시스템	동영상	특성정보 추출	동영상 기준	

5. 결론

본 논문에서는 검색자의 질의 동영상의 특성정보를 추출, 그것을 프로파일로 만들고 인터넷 동영상의 특성정보와 비교함으로써 동영상을 검색할 수 있는 영상기반 예제 중심 검색 시스템을 구축하였다. 그리고 프로파일을 통합하여 비교 효율을 높이고, 질의 동영상의 특성에 따른 가중치를 부여할 수 있는 옵션을 사용함으로써 질의 동영상의 판정을 보정할 수 있음을 실험으로 검증하였다.

향후 연구과제로 님은 형태의 동영상을 검색하는 것이 힘든 문제를 해결하고, 검색자가 사용하는 인터페이스의 향상을 통해서 검색 옵션을 쉽게 줄 수 있도록 하는 것이 필요하다. 그리고 동영상의 정적. 동적 영역을 구분하는데 더 정확한 정적. 동적 블럭을 찾아 낼 수 있는 방법론의 연구가 필요하다.

[참고문헌]

- [1] Carlson, Brian, "Taking on visual recognition's tough on-line test: Web still & video image content search & retrieval", Advanced Imaging vol. 12, no. 4, 3pp, 1997
- [2] Kao, O.; Joubert, G.R., "A content based Internet search engine for analysis and archival of MPEG-1 compressed newsfeeds", Multimedia and Expo, 2000. ICME 2000., pp.1155 -1158 vol.2
- [3] Yu, Hong-heather; Wolf, Wayne, "Hierarchical, multi-resolution method for dictionary-driven content-based image retrieval", IEEE INT CONF IMAGE PROCESS, vol. 2, pp. 823-826,
- [4] Swain, M.J., "Searching for multimedia on the World Wide Web", Multimedia Computing and Systems, 1999. IEEE International Conference on, pp. 32 - 37 vol.1
- [5] Shih-Fu Chang; Chen, W.; Meng, H.J.; Sundaram, H.; Di Zhong, "A fully automated content-based video search engine supporting spatiotemporal queries", Circuits and Systems for Video Technology, IEEE Transactions on , Volume: 8, pp. 602 -615
- [6] Makhoul, J.; etc, "Speech and language technologies for audio indexing and retrieval", Proceedings of the IEEE , Volume: 88, pp. 1338 -1353
- [7] Yeh, C.H.; Kuo, C.J., "Index-based fast search algorithm of image database on Internet", Multimedia and Expo, 2000., pp.1195 -1198 vol.2
- [8] Martin, P.; Eklund, P.W. , "Knowledge retrieval and the World Wide Web", IEEE Intelligent Systems, pp. 18 -25