

Repurposing, Surveillance, Fast Browsing Applications Using Motion Activity

박 철 수^{*}, 김 해 광

현대전자 정보통신연구소 기반기술실 MPEG-7 팀

{ cspark , cerise }@hei.co.kr

요약

본 논문에서는 효율적인 디지털 비디오 서비스를 위하여 콘텐츠의 움직임에 대한 크기, 시공간적 분포 그리고 방향성에 대한 통계적인 특징을 효율적으로 표현하는 움직임 정도 기술자에 대하여 논하고 움직임 정도가 중요한 특징으로 사용되는 비디오 검색, 리퍼포징, 원격 감시 그리고 브라우징 등의 서비스분야에 대하여 기술한다. 제안 기술은 멀티미디어 콘텐츠 색인과 구조에 대한 국제 표준을 개발중인 MPEG-7에서 실험모델로 채택되어 있다.

1. 서론

표현 미디어와 전달 미디어 그리고 이들을 운용하는 시스템 성능의 비약적인 발전으로 복수의 단일 미디어로 구성된 멀티미디어 데이터의 자유로운 생성, 빠른 검색 그리고 편리한 사용 및 재사용에 대한 욕구가 점점 증가하고 있다. 표현 미디어의 전자화에 의하여 자유로운 멀티미디어 데이터 생성에 대한 욕구가 충족되어 감에 따라, 방대한 양의 멀티미디어 데이터가 개인과 공용시스템에 산재하게 되었다. 이러한 멀티미디어 데이터의 증가에 비례하여 데이터를 재사용하기 위하여 검색에 소요되는 시간과 비용 또한 크게 증가하게 되었다. 따라서 멀티미디어 데이터를 빠르게 검색하여 편리하게 사용할 수 있는 디지털 신호 처리기법에 대한 필요성이 증가하게 되었다[1].

이미 널리 이용되고 있는 문자기반검색과 멀티미디어 데이터 기반검색의 주된 차이점은 검색에 사용되는 특징추출의 난해도와 멀티미디어 데이터 자체가 갖는 표현의 다형성을 들 수 있다. 현재 영상관련 멀티미디어 데이터 검색에서 가장 많이 연구되는 분야는 비교적 데이터 획득이 쉬운 정지영상이다. 정지영상 검색은 주로 내용기반기법이고 영상의 색상, 질감 그리고 모양 등에 대해서 회전, 크기 그리고 이동 등의 변화에 일관된 특징을 보이는 특징 정보 색인하여 검색에 활용하고있다. 비디오

검색은 정지영상에 비해 데이터 획득이 쉽지 않고, 대용량의 데이터를 저장하고 처리해야 하는 관계로 응용이 제한적이였다. 그러나, 디스크, 테이프 그리고 CD-ROM 등의 저장 미디어와 IMT2000, 인터넷 등의 전달 미디어의 급속한 발전으로 인하여 데이터 획득이 쉬워지면서 새로운 디지털 비디오 서비스 서비스가 출현하고 있으며, 보다 나은 서비스를 위하여 비디오 시퀀스 내의 콘텐츠를 간단하게 구별하고 효과적으로 브라우징할 수 있는 색인과 구조화 기술에 대한 필요성이 증가하고 있다. 기존의 비디오 색인과 구조화기법은 주로 시퀀스내의 색상과 움직임의 변화량에 사용하여 유사한 신호적 특성을 갖는 세그먼트로 구조화 수행하고 해당 세그먼트를 대표하는 대표영상의 특징정보로 색인 하였다. 이러한 구조화는 비디오를 보다 체계적으로 색인할 수 있어 많은 프레임으로 구성된 비디오의 경우 보다 효율적인 비디오 서비스를 지원할 수 있다.

멀티미디어 데이터 서비스 폭 넓은 확대를 위해서는 효율적이고 체계적인 멀티미디어 데이터 색인과 구조화기술이 필요하다. 따라서, MPEG-1,-2,-4 표준을 제정한 ISO/JTC1 MPEG 그룹은 멀티미디어 콘텐츠의 색인과 구조에 대한 멀티미디어 메타데이터 표준인 MPEG-7을 제안 중에 있다. 본 논문에서 제안하는 기술(P515)은 성능평가를 수행하는 핵심실험 기술 단계를 거쳐 MPEG-7 실험모델에 채택된 기술로서 사용자에게 감각적이며 지각적인 의미를

전달하는 콘텐츠 움직임 특성을 효과적으로 반영할 수 있는 움직임 정도기술자에 대하여 기술한다 [2]. 제안하는 움직임 기술자는 압축된 디지털 비디오 신호에 포함된 움직임 벡터를 별다른 처리 없이 사용할 수 있어 특징 추출에 소요되는 복잡도가 아주 작아 동영상 복호장치를 사용하는 서비스 시스템에서는 실시간 처리가 가능하며 제한된 성능을 갖는 이동통신 환경뿐 만 아니라 콘텐츠 기반 비디오 조작에 관련된 서비스분야에서도 적용 가능하다. 본 논문의 제 2장에서 제안하는 움직임 정도 색인 기술에 대하여 기술하고, 제 3장에서 움직임 정도가 중요한 특징으로 사용되는 몇 개의 응용분야를 언급한다. 제 4장에서는 결론과 향후 연구방향을 제시한다.

2. 움직임 정도 기술자

움직임 정도는 사용자에게 비디오 시퀀스 내의 콘텐츠 변화 정도에 대한 감각적이고 지각적인 정보를 전달하며, 콘텐츠를 구별하는 중요한 특징이 된다. 본 논문에서는 움직임 정도를 크게 크기와 방향으로 표현하며 색상, 모양 그리고 질감 등의 특징을 이용한 일반적인 검색에 비해 콘텐츠의 전체적인 변화량에 기반을 두기 때문에 보다 추상적이고 의미론적인 단계의 특징 기술이 가능하고 검색의 초기 단계에 적용하여 검색 대상의 범위를 축소 시켜 검색 속도를 크게 향상 시킬 수 있다. 제안하는 움직임 정도 기술자는 움직임 크기 기술자와 움직임 방향 기술자로 구성되며 콘텐츠의 전체적 움직임 정도 특성을 신축적으로 기술하기 위하여 움직임의 크기와 방향에 대한 시/공간적 평균, 시/공간적 분포의 통계적 특징을 나타내는 센트럴 모멘트 그리고 자세한 방향특징을 표현하기 위하여 히스토그램을 사용한다.

움직임 크기 기술자는 세크먼트 내의 한 프레임이나 전체에 대한 움직임 데이터에 대하여 움직임 크기의 특징을 표현하는 기술자이다. 일반적으로 움직임 정보를 표시하는 움직임 벡터는 크기와 방향 성분으로 구성되며 2 차원의 $MV=(MV_x, MV_y)$ 로 표현할 수 있다. 여기서 MV_x 는 수평방향의 움직임 성분이고 MV_y 는 수직방향의 움직임 성분이다. 2 차원 공간상의 움직임 거리를 나타내는 크기 I 와 움직임 방향을 나타내는 방향 ϕ 는 아래와 같이 표현된다.

$$I_k = \sqrt{(MV_{yk} \times MV_{yk} + MV_{xk} \times MV_{xk})} \quad (1)$$

$$\phi_k = \tan^{-1}(MV_{yk} / MV_{xk}) \quad (2)$$

식 (1)과 (2)에서 (MV_{xk}, MV_{yk}) 는 한 프레임의 임의의 크기를 갖는 M 개의 영역으로 분할 하였을 때, k 번째 영역에 대한 움직임 벡터이다. 한 프레임내의 콘텐츠의 평균적인 이동거리, 즉 변화정도를 나타내는 식 (3)의 λ_i 은 시퀀스 내의 i 번째 프레임의 움직임 크기 I 에 대한 평균으로 아래와 같이 표현된다.

$$\lambda_i = \frac{\sum_{j=0}^M I_j}{M} \quad (3)$$

ω_1 는 비디오 세그먼트 내의 T 개의 프레임들에 대한 움직임 크기의 평균 λ 에 대한 평균이며 식 (4)과 같이 표현된다.

$$\omega_1 = \frac{\sum_{i=0}^T \lambda_i}{T} \quad (4)$$

식 (4)에서 T 는 움직임 정보의 추출방법이나 목표하는 응용분야 특성에 따라 비디오 세그먼트 내의 전체 프레임 수와 동일 하지 않을 수 있다.

움직임 크기에 대한 센트럴 모멘트 ω_q 는 ω_1 에서 λ 에 대한 시간적 분포와 왜곡정도에 대한 특징을 나타내며 q 차원의 모멘트를 갖을 수 있으며, 표현하려는 움직임 정도의 정밀도에 따라 가변적인 개수를 사용할 수 있는데 실시간의 처리를 요하는 서비스나 비디오 검색의 초기단계에 움직임 정도를 특징정보로 사용할 경우 정확한 검색이 요하지 않으므로 ω_q 를 사용하지 않아도 된다.

$$\omega_p = \sqrt{\frac{\sum_{i=0}^T (\lambda_i - \omega_1)^q}{T}}, \quad q \geq 2 \quad (5)$$

그리고, 비디오 시퀀스의 i 번째 프레임에 대한 움직임 크기의 무작위 정도를 나타내는 σ_{λ_i} 와 세그먼트에 대한 평균적인 무작위 정도인 σ_{ω_1} 는 식 (6)와 식 (7)으로 표현된다

$$\sigma_{\lambda_i} = \sqrt{\frac{\sum_{k=0}^M (I_k - \lambda_i)^2}{M}} \quad (6)$$

$$\sigma_{\omega_1} = \frac{\sum_{i=0}^T \sigma_{\lambda_i}}{T} \quad (7)$$

식 (6)과 (7)에서 표현하는 무작위 정도(난잡도)는 프레임 내에 각기 다른 움직임 특성을 갖는 복수개의 물체가 포함되는 비디오 특성에서 비추어 볼 때 프레임 내의 움직임에 대한 인간의 시각 및 지각적 특성을 잘 반영할 수 있는 특징 정보이다. 움직임 크기의 표준편차에 센트럴 모멘트는 표준편차 평균 $\sigma_{\omega 1}$ 에서 각 프레임의 표준편차 $\sigma_{\omega i}$ 대한 공간적 분포 및 왜곡 정도에 대한 특징을 나타내며 q 차원의 모멘트를 갖을 수 있으며 보다 정확한 움직임 정도 패턴을 구별할 수 있는 특징이 필요할 경우에 사용될 수 있다. 아래의 식 (8)을 사용하여 얻을 수 있다.

$$\sigma_{\omega p} = \sqrt{\sum_{i=0}^T (\sigma_{\omega i} - \omega_{\omega 1})^q / T}, q \geq 2 \quad (8)$$

상기한 비디오 세그먼트 내의 콘텐츠의 움직임 크기에 대한 특징정보를 기술하는 움직임 크기 기술자는 다음의 표 1과 같이 구성된다.

움직임 크기 기술자			
평균	시간적 분포	편차 평균	공간적 분포
ω_1	$\omega_2 \dots \omega_q$	$\sigma_{\omega 1}$	$\sigma_{\omega 2} \dots \sigma_{\omega q}$

표 2. 움직임 크기 기술자

크기 기술자는 움직임 크기에 대한 기본적인 특징을 표현하는 움직임 정도 기술자의 기본 구성 요소이고 신축적인 구성할 수 있도록 q 의 모멘트로 구성하였고 평균적 특성과 시/공간적 움직임 분포에 대한 특징을 기술하고 있다.

움직임 방향 기술자는 비디오 세그먼트 내의 프레임이나 시퀀스에 대한 움직임 방향 특성을 표현하는 기술자로 방향에 대한 평균, 시/공간적 분포의 통계적 특징을 나타내는 센트럴 모멘트 그리고 지배적인 방향으로 구성되며 지배적 방향을 제외한 특징추출 방법과 과정은 움직임 크기 기술자와 동일하다. 세그먼트 내의 콘텐츠가 일관된 움직임 방향을 갖는 혹은 기대치가 존재하는 물체로 구성되어 있을 시에 유용하게 사용될 수 있는 특징 기술자이다. 지배적 움직임 방향을 나타내는 β_k 은 움직임 방향에 대한 빈도수를 크기에 따라 정렬 다음 상위 k 개를 지배적 움직임 방향을 나타내는 특징으로 사용한다. 움직임 방향 기술자는 표 3과 같이 표현된다.

움직임 방향 기술자				
평균	시간적 분포	편차 평균	공간적 분포	지배적 방향
β_1	$\theta_2 \dots \theta_p$	$\sigma_{\beta 1}$	$\sigma_{\beta 2} \dots \sigma_{\beta p}$	$\beta_1 \dots \beta_k$

표 3. 움직임 방향 기술자

제안하는 움직임 정도 기술자는 비디오 세그먼트의 콘텐츠를 검색하고 식별하는데 있어 몇 개의 대표영상의 특징 정보만을 사용하는 기존의 비디오 움직임 색인기법으로는 표현하기 어려운 비디오 전체, 대표영상 사이, 시간상 특정 구간에 대한 신호적 특성들과 시공간적 분포 그리고 변화 정도와 패턴등에 대한 지각적 특징을 기술할 수 있어 이러한 움직임 정도가 중요한 특징이 되는 디지털 비디오 서비스 응용들에 활용될 수 있다. 또한, 움직임 기술자의 특징정보는 응용분야와 움직임 표현의 정밀도를 고려하여 선택적으로 사용할 수 있으며, 기존의 동영상 압축 기법인 MPEG-1,-2,-4 와 H.263 를 사용하여 압축된 비디오 스트림의 움직임 정보를 특별한 부가 처리 없이 직접 사용할 수 있어 특징정보 추출에 소요되는 복잡도가 아주 작아 실시간 처리가 필요한 응용분야에도 적용 가능하다.

3. 움직임 정도 기술자를 활용한 디지털 비디오 응용 서비스

제안하는 움직임 기술자는 인터넷과 IMT2000과 같은 비디오 서비스가 지원되는 이동통신 환경, 화상회의, 디지털 방송 그리고 VOD와 같은 제한된 대역폭과 시스템 리소스를 효율적으로 활용하는데 움직임 정도가 중요하게 사용될 수 있는 다양한 디지털 비디오 서비스 분야에 사용될 수 있다. 그 중 대표적인 활용 일례로서 비디오 브라우징, 원격감시, 검색 그리고 리퍼포징을 제시한다.

3-1. 브라우징

브라우징은 사용자가 멀티미디어 데이터 베이스를 통하여 데이터들 사이에 존재하는 다양한 형태의 링크를 따라 이동하는 정보검색 활동을 칭한다. 비디오 같은 대용량의 미디어 데이터에 대하여 빠르고 효율적인 브라우징을 수행하기 위하여 지원이 되어야 할 필수적인 기능 요소 중에 하나는 사용자에게 사용중인 비디오에 대한 비디오의 구조, 요약정보, 대표영상 그리고 특징정보를 비주얼 형태로 유저 인터페이스를 통하여 제공하는 것이고

이를 위하여 사용자가 볼 수 있거나 들을 수 있는 수단을 사용할 수 있다. 또한 움직임 정도에 따라 몇 개의 범주로 분류하여 동일한 범주나 상이한 범주의 비디오 세그먼트를 빠르게 브라우징 할 수 있는 기능이 필요하다.

3.2 원격감시

원격감시는 제한된 공간상에 발생하는 사건들을 기록하고 감시해야 하는 통상적으로 발생 가능한 움직임 특성의 기대치가 존재하는 응용이다. 따라서, 통상적인 기대치와 다른 상이한 범주에 속한 움직임 특성을 갖는 비디오 세그먼트는 사건에 대한 중요한 정보를 담고 있기 때문에 움직임 정도를 사용할 경우 정도에 따라 기록해야 할 데이터량을 제어할 수 있으며 중요사건에 대한 효율적인 검색도 지원할 수 있다.

3.3. 비디오 검색

비디오 검색은 브라우징과 함께 필요한 정보를 찾는 탐색 활동의 다른 형태로서 통상 자연언어, 구조화된 컴퓨터 언어등을 사용하여 질의를 표현하거나 비슷한 예제를 부가정보로 제시하는 형태로 진행된다. 비디오 검색은 시간적 순서를 갖는 복수의 프레임들로 구성된 비디오 세그먼트 단위로 진행되며 통상 다양한 형태의 복수의 물체가 세그먼트 내에 존재하게 된다. 이러한 다양한 색상과 모양 그리고 질감의 영역이나 물체가 존재하게 된다. 따라서, 특정한 색상이나 형태의 모양 그리고 구성의 질감에 의한 검색 보다는 움직임 정도와 같은 통상적인 움직임 정도에 대한 정보를 검색의 초기단계에 사용하면 검색의 범위를 크게 줄일 수 있어 보다 빠른 검색이 가능하다. 또한 특정한 색상이나 모양 그리고 질감에 대한 특징을 사용할 수 없을 시에 콘텐츠의 변화에 기반을 둔 움직임 정보를 사용하면 보다 효율적인 검색을 할 수 있다. 일례로, 뉴스 프로그램의 데이터 베이스에서 앵커가 등장하는 장면을 검색하고자 할 경우, 매일 다른 의상과 다른 스튜디오와 다른 사람에 의해 진행되는 경우가 보통이므로 색상, 모양 그리고 질감을 사용하는 것보다 "움직임 정도가 거의 없는 장면을 검색"한 후 보다 구체적인 특징정보를 사용하는 것이 바람직하다. 또한, 비디오 내의 일관된 방향을 갖는 물체가 존재할 경우 방향에 대한 정보는 매우 유용하게 사용될 수 있다.

3-4. 비디오 리퍼포징

비디오 리퍼포징은 서버상의 부호화된 스트림을 사용하는 사용자 시스템 성능과 사용자 요구에 따라 적절하게 재가공하여 서비스하는 응용을 지칭한다. 특히, 이동통신 단말과 같이 제한된 대역폭과 성능을 갖는 시스템을 위하여 디지털 비디오 서비스를 수행할 경우, 움직임 정도는 대역폭을 보다 효율적으로 사용하기 위하여 전송 프레임 수, 즉 비트율을 제어하는데 다음과 같이 사용될 수 있다. 가령, 전송해야 할 GOP(Group Of Picture)에 대한 움직임 정도가 작을 경우 콘텐츠에 대한 중요한 변화가 발생하기 않은 것으로 간주하여 데이터를 전송하지 않거나 전송할 프레임 수를 줄이고, 반대로 움직임 정도가 크면 의미 있는 사건들이 발생한 것으로 간주하여 보다 많은 수의 프레임을 전송하여 사용자에게 보다 정확한 정보를 제공할 수 있다.

4. 실험 결과

본 논문에서 제안한 움직임 정도 기술자(이하 P515 라 칭함)는 MPEG-7 의 핵심 실험기술로 채택되어 타 기관에서 제안한 3 가지 기법들의 성능과 비교 평가되었고 움직임 정도와 방향성에 대한 특징 기술 기법이 실험모델에 채택되었다 [2][3][4][5]. 움직임 정도 기술자에 대한 핵심실험 기술단계에서 제안된 기법들의 객관적인 비교평가를 위하여 MPEG-7 테스트 데이터 셋 내의 28 개 비디오에 대하여 614 개의 샷을 추출하여 데이터베이스를 생성하였고 각 샷에 대하여 인간의 지각적인 평가에 의하여 움직임 정도가 매우 낮은 1 단계에서 아주 큰 5 단계까지의 Ground Truth 를 마련하였다 [8]. Ground Truth 데이터 셋의 분포는 아래의 표 4 와 같다.

움직임 정도	샷의 수
1	114
2	198
3	180
4	103
5	28

표 4. Ground Truth 데이터 셋의 분포

Ground Truth 데이터 셋에 대하여 제안한 기법들의 움직임 정도 특징들을 선형적인 정규화를 통하여 Ground Truth 와 동일하게 5 단계로 분류

하였고, 움직임 정도 기술의 정확성을 평가하기 위하여 Ground Truth Value(GTV)와 각 기법들의 움직임 정도의 일치한 횟수(히트율)를 평가하였고 GTV와 유사도를 측정하기 위하여 GTV에 대한 차의 합을 측정하였고 그에 대한 결과는 아래와 같다.

	A1	A2	A3	A4	A5	Total
P002	0.94	0.28	0.12	0.10	0.14	0.324
P075	0.00	0.25	0.50	0.00	0.00	0.229
P515	0.64	0.48	0.35	0.46	0.39	0.467
P627	0.70	0.70	0.37	0.20	0.10	0.419

표 5. 제안된 기법들의 GTV에 대한 히트율

	GTV에 대한 SAD
P 002	574
P 075	838
P 515	391
P 627	394

표 6. 제안된 기법들의 GTV에 대한 SAD

표 5와 표 6의 실험 결과에서 나타난 바와 같이 제안한 움직임 정도 기술자는 모든 단계의 움직임 정도에서 GTV를 가장 정확하고 근접하게 표현할 수 있음을 알 수 있다.

5. 결론

본 논문에서는 움직임 특징을 통계적인 방법을 사용하여 신속적으로 기술할 수 있는 움직임 정도 기술자를 제안하였다. 제안 기술이 인간이 지각적으로 느끼는 움직임 정도를 비교적 정확하고 근접하게 표현할 수 있음을 실험을 통하여 검증하였다. 아울러, 특징 추출 및 검색에 소요되는 복잡도가 작아 실시간 처리가 필요한 디지털 비디오 서비스분야에 효과적으로 사용될 수 있음을 보였다. 향후 연구 방향은 지속적인 MPEG-7 표준화 활동을 통하여 움직임 정도 기술자에 대한 보완 기술을 개발하여 표준에 반영하고, 개발된 기술들을 이용하여 디지털 비디오 서비스를 위한 시스템 개발 및 환경구축 하는 데 있다.

참고 문헌

1. MPEG Requirements Group, "MPEG-7: Context and Objectives", Doc. ISO/MPEG N2460, MPEG Atlantic City Meeting, October 1998
2. Cheol-Soo Park and Hae-Kwang Kim, "Temporally Aggregated Motion Descriptor", Doc. ISO/MPEG

P515, MPEG Seoul Meeting, March 1999

3. Ajay Divakaran, et. al., - A Bit Allocation Based Descriptor for MPEG-4/2/1 Compressed Video Sequences - ISO/IEC/SC29/WG11 MPEG98/P002 - Jan 1999

4. Xinding Sun, B. S. Manjunath - Motion Quantized α -Histogram as Video Unit Descriptor - ISO/IEC/SC29/WG11 MPEG99/P075 - Feb 1999

5. D. Manoranjan and V. V. Vinod - Video Segment Activity : A Visual Descriptor proposal for MPEG7 - ISO/IEC/SC29/WG11 MPEG99/P627 - Feb 1999

6. S. Jeannin and M. Bober, "Description of Core Experiments for MPEG-7 Motion/Shape", Doc. ISO/MPEG M2690, MPEG Seoul Meeting, March 1999

7. Philippe Joly and Hae-Kwang Kim, "Efficient Automatic Analysis of Camera Work and Microsegmentation of Video using Spatio-Temporal Images", Signal Processing : Image Communication, Vol. 8, No. 4, May 1996.

8. MPEG Requirements Group, "Distribution of MPEG-7 Content Set", Doc. ISO/MPEG N2468, MPEG Dublin Meeting, October 1998