

Temporal texture를 이용한 비디오 검색의 성능분석

홍지수^U 김영복 김도년 조동섭
세종대학교 전산학과, 이화여자대학교 컴퓨터학과
hongjs@gce.sejong.ac.kr, yungbkim@sejong.ac.kr,
(dnkim, dscho)@mm.ewha.ac.kr

Study on Performance Analysis of Video Retrieval Using Temporal Texture

Ji-Su Hong^U Do-Nyun Kim Yung-Bok Kim Dong-Sub Cho
Dept. of Computer Science, Sejong University
Dept. of Computer Science & Engineering, Ewha Womans University

요 약

모든 물체의 표면은 독특한 성질을 보유하고 있으므로, 비디오 검색에 있어 텍스처는 형상이나 색과 더불어 중요한 변수로 사용될 수 있다. 비디오 검색에 있어서 중요한 것은 어떤 영상의 특징을 올바르게 추출하고 잘 분류하여 표현하는 것이다. Temporal texture는 무한한 시공간적 범위의 복잡하고, 추상적인 움직임 패턴도 특징화시킬 수 있으므로, temporal texture 패턴을 얼마나 잘 이용할 수 있는냐는 비디오 검색의 성능에 많은 영향을 끼칠 수 있다. 본 논문은 temporal texture의 서로 다른 특징을 가진 세 가지의 모델을 선정하여 비교한다. 특히, 특징 추출의 분류가 정확하게 이루어지느냐에 초점을 맞추어서 분석하였다. 분류의 성능은 두 가지 변수 즉, 어떤 성질의 모델이며 비디오 데이터인가에 따라 달라지게 된다. 이들 모델링이 분류하기까지 걸리는 시간의 차이는 무시할 수 있을 정도의 시간차이므로, 정확도를 위주로 성능을 분석했다.

1. 서론

영상의 검색에 있어서 형상, 색과 더불어 텍스처 특성 또한 중요한 역할을 할 수 있다. 기존의 텍스처 분석의 경우, 인간의 시각적이며 직감적 특징이랄 수 있는 텍스처의 굵기, 거침의 정도, 대비, 방향성, 규칙성과 같은 저수준의 패턴 특성을 먼저 추출한다. 그리고 나서 이렇게 추출된 특성들을 기반으로 분류·분할의 단계를 거쳐서야 원하는 속성을 가진 영역으로 사용될 수 있다. 동영상은 위와 같이 공간적 특성만을 추출하는 방법만으로 표현하기엔 나타내야할 정보가 많기 때문에, 기존의 정지 영상의 특성 이외에도 시간적 특성과 주기성을 고려한 방법이 요구된다.

동영상에서 추출한 움직임 특징은 비디오 데이터 베이스에서의 내용 기반 검색에 널리 쓰일 수 있으며, 감시 시스템이나 인간의 제스처 해석기와 같은 움직임 기반의 인지 함수 등에 중요한 단서로 작용하고 있다[1]. 자연 세계에서 빈번하게 나타나는 움직임의 하나인 temporal texture의 특징을 추출하는 것은 시공간적으로 무한한 범위를 가진 영상이나 복잡한 영상에 있어 더욱 유리하다. 물 위의 잔 물결이나 새 떼들의 움직임, 또는 차가 움직이는 동안 보이는 경관 등도 temporal texture 모델

링을 이용하여 추출할 수 있는 움직임의 예이다.

본 논문에서는 주요한 temporal texture 모델링을 비교하고, 그 모델들을 바탕으로 영상의 특징 추출 및 분류 과정을 구현했을 시, 그 성능을 분석하고 평가한다. 본 논문의 구성은 다음과 같다. 제2장에서는 주요한 temporal texture 모델들의 기본 구조를 설명하고, 이에 따라 각 모델들을 상호 비교한다. 3장에서는 각 모델들이 temporal texture를 보유한 영상들을 정확히 분류하는 것을 중점으로 성능을 분석하였다.

2. 관련연구

Temporal texture를 이용하여 형상을 추출한 연구는 그리 많지 않다. 최근 3가지 주요 temporal texture 모델들에서 나타나는 특징을 정리하자면, 다음과 같다.

첫 번째, Polana의 모델은 시각적 흐름 필드로부터 유도된 시공간적 특성에서 나타나는 비균일성을 이용하여 영상에 나타나는 움직임 패턴들을 분류한다. 그리고 걷는 사람의 특성이나 날으는 새들 등과 같은 움직임의 주기적 성질에 근거하여 시간적으로 반복적인 움직임을 탐색하고 인식할 수 있는 저수준 계산 접근방법을 제시하여, 복잡하고 다양한 움직임 패턴을 포함하는 실세계 이

미지열에 대한 기법을 제시하였다.

이의 특징은 움직임 패턴의 인식에 통계적 기법을 적용하여, 추상화된 객체 모델을 사용하거나 궤도 표현 등을 사용하지 않고 움직임의 저수준적 특성만을 이용하여 움직임 인식을 가능하게 했다는 것이다. 이 모델은 조명 상태와 색의 변화에 그다지 영향을 받지 않지만, 불안정한 temporal texture로부터 정확히 시각적 흐름을 예측하기 힘들다는 난제를 안고 있다[2].

두 번째, Szummer가 개발한 모델은 인식 뿐 아니라, 텍스처 합성에도 사용 가능한, STAR(spatio-temporal autoregressive)모델을 이용하여 temporal texture를 모델링하였다. 이는 시계열에 적합할 뿐 아니라 텍스처 이미지 같은 공간 데이터에 대해서도 적합하다는 장점을 가진다. 이 모델은 이미지 내의 지역적 밝기에 의존한 모델이어서, 노이즈에 매우 약하다. 또한 선형성을 바탕으로 하여, 형성된 모델 pixel 간의 관계가 비선형적일 경우나 주기적 움직임, 팽창/수축, 회전, 나선형 움직임을 보일 때에는 분석이 불가능하다. 그리고 어떤 형태의 가속도도 모델링하지 못하므로, 영상의 속도 벡터는 어느 지역이든 비슷해야 한다[3].

세 번째, Otsuka의 모델은 이미지 시퀀스의 여러 프레임으로부터 얻은 시공간적 영역에서의 움직임 궤도의 표면에 초점을 두었다. 그래서 그 표면의 지역적 영역을 tangent 평면을 분포로 나타내어 시공간적인 특징을 추출하는 모델이다. 이 모델은 노이즈에 강하고 temporal texture에 복잡한 motion과 변형의 조합도 모델링할 수 있는 장점을 가진다[4].

Polana의 모델과 달리, 후자 두 모델들은 저수준 움직임 특성만을 이용한 모델링이 아닌 객체 모델과 움직임 궤도의 표현을 이용하여 움직임 특성들을 추출하는 특징을 가지고 있다. 이들 모델들 자체가 가지는 특성에 따라 그 결과는 상이하다.

3. Temporal texture 모델의 성능 분석

각 temporal texture 모델들의 특성 분류에 대한 성능을 평가하고자, 여러 텍스처 샘플들이 실험을 위해 사용되었다. 방향성이 있는 텍스처와 비방향성 텍스처 모두를 표현하는 이미지열들의 집합이 디지털화되어 이용되었다. 또한, 균일한 팽창을 표현하는 이미지열과 회전을 나타내는 이미지열도 실험에 쓰여져 다양한 이미지 입력에 모델링의 결과가 어떻게 나오는지 주목하였다.

실험에 쓰여진 입력들은 그림 1과 같이 흐르는 강의 모습과 수증기, 끓는 물의 모습 등의 텍스처를 포함한 것들로, Hi-8 비디오프로 촬영했으며, 170×115 크기의 120 프레임으로 이뤄진 4초 분량의 그레이 영상들이다.

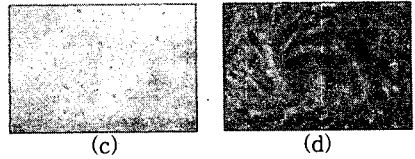
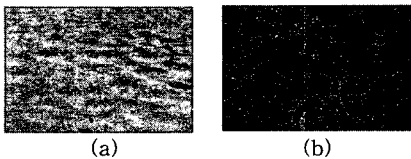


그림 1. Temporal texture 이미지열의 예 : (a) 강물, (b) 끓는 물, (c) 수증기 (d) 변기로 흘러 들어가는 물

Temporal texture의 샘플을 분류하는 실험은 영상 내의 강조된 부분, 엔트로피, 속도, 방향성, 밀도, 교합 등의 특징을 평균화하여 이를 바탕으로 분류하게 된다.

표 1은 입력 영상들로 실험한 각 모델들의 분류 정확도를 나타낸 것이다. 여기서 모델명은 각 모델의 개발자의 이름을 딴 것이다. 분류 성공률은 다음의 식을 기반으로 구해졌다. 여기서 P는 분류 정확도, D는 추출한 특징의 수, F는 추출에 실패한 특징의 수이다.

$$\text{Precision}(P) = F/D + F$$

식1. 분류 정확도 도출식

| 영상 모델명 | 흔들리는 나뭇잎 | 흐르는 강물 | 끓는 물 | 수증기 |
|---------|----------|--------|------|------|
| Polana | 100 | 100 | 99.4 | 99.1 |
| Szummer | 98.7 | 98.6 | 98.4 | 98.8 |
| Otsuka | 98.3 | 98.4 | 98.2 | 98.3 |

표 1. 영상에 따른 각 모델의 분류 성공률 (%)

표 1에서 볼 수 있듯이 세 모델 모두 자연 세계의 막연한 현상인 temporal texture의 특징을 성공적으로 추출하였다. 그러나, 이 테스트 영상들은 급격한 움직임 없이 거의 균일한 속도를 가지며, 카메라의 움직임 또한 없이 촬영된 것이다. 하지만, 관찰자인 카메라는 움직임을 갖고, 이에 따라 영상에 담기는 대상 또한 움직이게 된다. 위의 실험은 그러한 영상의 움직임을 배제한 조건에서 수행되었다.

이에 다음의 조건을 더해서, 이들 모델들을 테스트했다. 첫째, 카메라의 움직임에 의해서 회전하는 영상을 테스트해 보고, 둘째, 카메라가 근접함에 따라 영상이 팽창되어 나타난 테스트 영상으로 실험했을 경우와 마지막으로, 움직임 자체에 급격한 변화가 있는 경우로, 예를 들자면, 그림 1의 (d)와 같은 경우이다. 그 그림은 방향과 속도가 급격하게 변하는 움직임을 포함한다. 다음 실험에서는 위 세 조건에 temporal texture를 보유한 영상으로 각 모델들의 성능을 분석해보았다.

Polana의 모델은 영상의 팽창, 축소, 회전 등에도 모델링이 가능하였지만, 변형이 급격하게 이루어질 경우는 제대로 모델링이 되지 않았다. 즉, 영상이 일정 수치로

균일하게 팽창하거나 회전할 경우에는 모델링이 가능하지만 그렇지 않을 경우는 특징을 정확히 추출해낼 수 없었다.

Szummer의 모델은 일정 방향으로 비슷한 속도를 가지고 변화하는 temporal texture에 강한 모델링 효과를 나타내었다. 그러나, 선형성을 바탕으로 한 Szummer의 STAR 모델은 배수구에 물이 빠져나가는 영상으로 테스트했을 때에, 물결의 중심을 따라 가장자리와 중심의 속도가 달라 그 특징을 제대로 추출하지 못 했다. 또한, 카메라의 원근 조절을 통한 대상의 팽창이나 축소 또한 모델링하지 못 하여, 극히 제한적인 영역의 temporal texture만을 모델링할 수 있었다.

Otsuka의 모델은 주요 움직임의 윤곽들로부터 변형의 속도를 추정하고 그 속도에 의해 결정된 tangent 평면들의 분포로부터 공간적 특징, 즉, 윤곽의 방향성 또한 얻을 수 있다. 그래서, 물체의 회전이나 확대와 같이 움직임에 큰 변형이 있을 때에도 모델링이 가능했다. 또한, 기상 레이더 이미지열과 같이 열화가 심한 영상에도 강한 성능을 보였다.

표 2는 이들 모델들이 영상의 특징에 따라 나타내는 결과를 비교한 것이다.

| 영상 모델명 | 팽창/축소, 회전, 나선형 움직임 | 가속도 |
|---------|-----------------------|---------|
| Polana | 균일할 시, 분석 가능 | 추출이 어려움 |
| Szummer | 일정 속도의 단방향 움직임만 분석 가능 | |
| Otsuka | 분석 가능 | |

표 2. 영상의 특징에 따른 각 모델의 특징 분석 결과

4. 결론

Temporal texture는 무한한 시공간적 범위의 복잡하고, 추상적인 움직임 패턴이다. 이를 특징화시킬 수 있고, temporal texture 패턴을 얼마나 잘 이용할 수 있는냐는 비디오 검색의 성능에 많은 영향을 끼칠 수 있다. Temporal texture의 특징에 따라 분류하고 추출할 수 있는 모델들의 특징들은 서로 매우 상이하셔서 입력 영상에 따라 그리고, 각 모델들의 특징에 따라 그 결과가 많이 달라진다.

본 논문은 서로 다른 특징을 가진 세 가지의 모델을 선정하여 비교하였다. temporal texture의 특징을 각 모델 별로 추출하고, 영상들로부터 특징들을 추출 분류하는 실험을 하였다. 여러 텍스처 샘플들이 실험을 위해 사용되었으며, 방향성과 비방향성 텍스처 모두를 표현하는 이미지열들의 집합이 디지털화되어 이용되었다. 또한, 일정 방향으로의 움직임 이외에 회전이나 팽창/축소와 같은 변형을 포함하는 영상이 실험에 쓰여져 다양한 움직임을 가지는 이미지 입력에 모델링의 결과가 어떻게 나오는지에 주목하였다.

분류 실험은 영상 내의 강조, 속도 벡터, 방향성, 밀도와 같은 특징들을 평균화 한 다음, 그를 바탕으로 분

류가 이루어졌다.

성능 분석의 결과, 세 모델 모두 영상 내에 급격한 움직임, 즉, 가속도나 회전과 같은 변형이 있을 경우에는 움직임을 정확히 분류하고 추출하는 비율이 많이 저하되었다. 영상의 변형에 취약한 temporal texture 모델로는 향후 디지털 라이브러리에서의 영상 검색에서의 효율적 대응이 어려울 것이므로 이 문제에 대한 집중적 해결이 요망된다.

5. 참고 문헌

- [1] C.Cedras and M.Shah, "Motion-based recognition: a survey", Image and Vision Computing, 13(2) : 129- 155, 1995.
- [2] Ramprasad B.Polana, "Temporal Texture and Activity Recognition", University of Rochester, Ph.D Thesis, 1994.
- [3] Marcin Olof Szummer, M.O, "Temporal texture modeling", MIT Media Lab, Technical Report, 1995.
- [4] K. Otsuka, T. Horikoshi, and S. Suzuki, "Feature Extraction of Temporal Texture based on Spatiotemporal Motion Trajectory", IEEE Proc. ICPR'98, Vol. 2 , pp. 1047-1051 , 1998.
- [5] Fang Liu, "Modeling Spatial and Temporal Textures", MIT Press, Ph.D Thesis, 1997.
- [6] 김희승, "영상인식", 생능, pp.175-199, 1993.
- [7] Mihran Tuceryan, Anil K. Jain, "Texture Analysis", Handbook of Pattern Recognition and Computer Vision, pp.235-276, 1983.