

모양 시퀀스 기술자를 이용한 효과적인 동작 표현 및 검색 방법

최 민 석[†]

요 약

동영상에서 객체의 움직임은 동영상의 내용을 분석하는데 중요한 요소로 작용한다. 따라서 움직임 정보를 이용하여 동영상 내용을 분석하고 검색하기 위한 많은 방법들이 제안되었다. 그러나 대부분의 방법들은 객체 자체의 동작 보다는 움직임의 방향이나 경로를 분석하는 쪽으로 치중되었다. 본 논문에서는 객체의 움직임에 의한 모양 변화를 이용하여 객체의 동작을 표현하고 비교하기 위한 모양 시퀀스 기술자(descriptor)를 제안한다. 객체의 움직임 정보는 입력된 이미지 시퀀스에서 객체 영역을 추출하여 연속된 2차원 모양 정보로 표현되고, 각각의 2차원 모양 정보는 모양 기술자를 이용하여 1차원 모양 특징 값으로 변환된다. 순서에 따라 배열된 모양 기술자들을 시간 축으로 주파수 변환한 후 저주파 영역의 계수를 취하여 모양 시퀀스 기술자를 얻게 된다. 실험을 통하여 제안된 방법이 객체의 동작 정보를 매우 효과적으로 표현 및 비교 가능하여 내용 기반 동영상 검색, 동작 인식 등의 인지적 관점의 움직임 분석 응용에 적용 가능성을 보였다.

키워드 : 모양 시퀀스 기술자, 모양 시퀀스, 동작 인식, 동작 검색

Efficient Representation and Matching of Object Movement using Shape Sequence Descriptor

Min-seok Choi[†]

ABSTRACT

Motion of object in a video clip often plays an important role in characterizing the content of the clip. A number of methods have been developed to analyze and retrieve video contents using motion information. However, most of these methods focused more on the analysis of direction or trajectory of motion but less on the analysis of the movement of an object itself. In this paper, we propose the shape sequence descriptor to describe and compare the movement based on the shape deformation caused by object motion along the time. A movement information is first represented a sequence of 2D shape of object extracted from input image sequence, and then 2D shape information is converted 1D shape feature using the shape descriptor. The shape sequence descriptor is obtained from the shape descriptor sequence by frequency transform along the time. Our experiment results show that the proposed method can be very simple and effective to describe the object movement and can be applicable to semantic applications such as content-based video retrieval and human movement recognition.

Keywords : Shape sequence descriptor, Shape sequence, Movement recognition, Movement retrieval

1. 서 론

인터넷 및 네트워크 환경의 고속화와 디지털 카메라 및 카메라 내장 휴대전화 등의 보급, 특히 UCC(User Created Contents) 또는 UGC(User Generated Contents)라 불리는 일반인들의 동영상 정보 콘텐츠가 새로운 문화 경향으로 자리잡으면서 디지털 형태의 동영상 정보의 급속한 증가를 가

져왔다. 이러한 디지털 정보의 홍수 속에서 원하는 정보를 쉽게 찾고 접근하고 관리할 수 있어야 정보로서의 가치를 가지게 될 것이다. 그러나, 기존의 사람에 의한 문자 색인 방식의 관리 방법으로는 기하급수적으로 증가하는 디지털 멀티미디어 정보를 효율적으로 관리하는데 한계를 가진다. 이러한 상황은 디지털 멀티미디어 정보의 보다 빠르고 편리한 검색 및 관리를 위하여 동영상 등의 디지털 데이터 자체의 내용 및 특성에 기반한 검색 방법의 필요성을 낳게 되었고 다양한 내용 기반 검색 방법들이 연구되고 있다.

동영상 분석 및 검색 응용에서 움직임 정보는 다른 정보

[†] 정 회 원 : 삼육대학교 경영정보학과 겸임교수
논문접수 : 2008년 3월 28일
심사완료 : 2008년 5월 8일

들 보다 중요한 의미를 가진다 [1,2]. 따라서 움직임 정보를 이용하여 동영상의 내용을 분석하고 검색하는 방법들이 많이 제안되었다. 움직임 정보 분석은 카메라의 움직임과 객체의 움직임 궤적 분석과 관련된 연구로부터 시작하여 최근에는 객체의 동작을 분석하여 장면의 내용을 분류하려는 방향으로 진행되고 있다 [3,4]. 이 장에서는 먼저 객체의 동작을 분석하기 위한 기존의 연구 흐름을 살펴 보고, 그 중 제안된 방법과 비교 대상이 되는 기존의 방법들을 간략히 설명하고자 한다.

1.1. 기존의 동작 분석 방법들에 대한 고찰

컴퓨터 비전 분야에서 동영상 등의 연속된 영상들에서 움직임을 분석하는 것은 중요한 연구 분야 중 하나이다. 특히 인간의 동작을 인식하기 위한 많은 연구들이 진행되었고, 현재도 진행 중이다. 이러한 연구는 그 접근 방법에 따라 다음과 같이 분류 할 수 있다[3,4].

- 모델 기반 접근 방법(Model-Based Approach)
- 모양 기반 접근 방법(Appearance-Based Approach)
- 움직임 기반 접근 방법(Motion-Based Approach)

모델 기반 접근 방법은 각 시간에 따른 사람의 자세를 3차원 모델을 이용하여 기술하고 모델 정합을 이용하여 동작을 인식하는 방법이다. 이러한 방법은 객체의 세밀한 분리가 필요하며, 각 신체 부분에 대한 모델 정렬 과정이 필요하다 [4-10].

모양 기반 접근 방법은 모델 기반 접근 방법과 달리 단지 움직임의 2차원 모양만을 이용하려는 시도이다 [11-15]. 동작은 객체의 연속된 2차원 모양으로 기술되고, 많은 방법들이 배경과 분리된 객체의 정규화된 이미지를 이용한다.

움직임 기반 방법은 신체의 정적인 자세를 참조하지 않고 움직임 자체만을 특성화 하려는 방법이다 [16-21]. 신체의 각 부분을 간단한 단일 덩어리 형태의 영역으로 구성하고 구조적 특징 대신 미리 정의된 영역의 움직임을 추적하여 동작을 분석한다.

1.2 모양 기반 접근 방법

모양 정보는 객체의 인식, 정합, 등록 등의 응용에서 매우 중요한 역할을 한다. 따라서 객체의 모양을 효과적으로 표현하고 중요한 모양 특징을 특징 값으로 정의하는 많은 연구들이 진행되어 왔다. 이러한 기존 연구들을 기반으로 동영상 분석에서도 객체의 시간에 따른 모양 변화를 효과적으로 분석하여 동영상이 전달하고자 하는 중요 정보를 얻어내는 것이 가능하다 [22]. 본 논문에서도 앞서 언급한 3가지 접근 방법 중 모양 기반 접근 방법을 이용한 객체의 동작 표현 및 인식 방법을 제안하고자 한다.

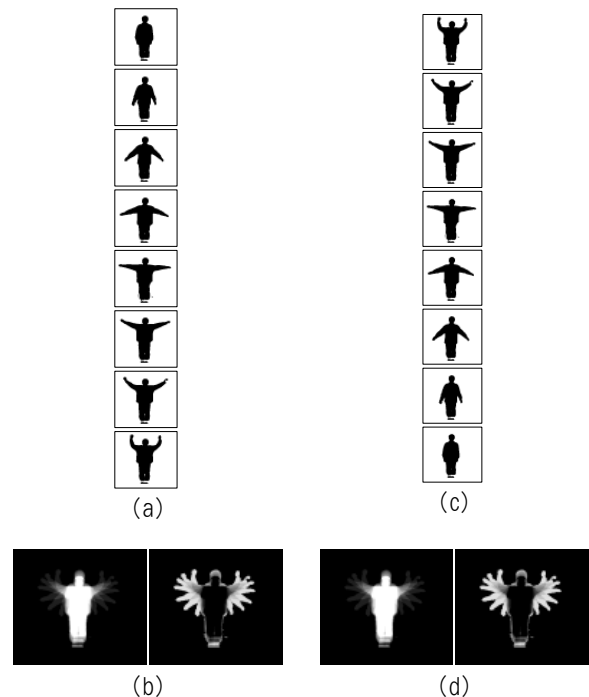
본 절에서는 제안된 방법의 객관적 평가를 위하여 실험 비교 대상이 된 기존의 모양 기반 접근 방법 중 MPEG-7에

제안된 모양 변화 기술자(shape variation descriptor)와 Bobick 과 Davis가 제안한 Temporal Template을 이용하는 방법에 대하여 간략히 살펴 보고자 한다.

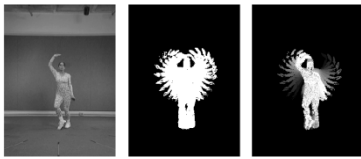
멀티미디어 콘텐츠 기술(description)에 대한 국제 표준인 MPEG-7에는 몇 가지 움직임 기술자가 제안되어 채택 되었다 [23,24]. 대부분의 방법은 움직임의 방향이나 형태, 궤적 등을 분석하는 것에 치중되었고, 객체의 동작 자체의 분석에 초점을 맞추지는 않았다. 그러나 동영상의 내용을 특성화 하는데 객체의 동작 정보가 중요한 역할을 하는 경우가 많기 때문에 객체의 동작 정보를 기술하기 위한 모양 변화 기술자가 제안되어 채택되었다 [25].

모양 변화 기술자의 목적은 동영상에서 비슷한 동작을 검색하고자 하는 것이다. 따라서 다양하고 복잡한 형태의 동작들을 의미상 비슷한 그룹으로 분류하기 위하여 3차원 동작 정보를 (그림 1)과 같은 간단한 2차원 영상형태의 모양 변화 누적도(shape variation map)를 이용하여 표현하였다. 이러한 모양 변화 누적도는 동작의 순서 정보를 보존하지 않고 있기 때문에 이를 이용한 모양 변화 기술자는 좀더 세분화된 개별 동작들을 구분하지 못하게 된다. (그림 1)의 예시와 같이 손을 올리거나 내리는 동작, 앉거나 일어서는 동작 등의 정확한 동작을 구분하기 위해서는 동작의 순서 정보가 반드시 필요하게 된다.

Bobick과 Davis는 인간의 동작을 표현하고 인식하기 위하여 2차원 temporal template을 제안 하였다 [4]. Temporal template은 (그림 2)와 같이 두 장의 이미지로 구성된다. 하나



(그림 1) 손을 올리고 내리는 두 가지 동작의 비교 (a) 손 올리기, (b) (a) 동작에서 추출된 모양 변화 누적도, (c) 손 내리기, (d) (c) 동작에서 추출된 모양 변화 누적도



(그림 2) MEI 와 MHI의 예. (b) MEI, (c) MHI

는 binary motion-energy image (MEI)라 부르며 연속된 이미지에서 움직임이 발생한 위치를 나타낸다. 다른 하나는 motion-history image (MHI)로 밝기가 움직임의 순서의 합수로 표현된 이미지이다.

인간의 동작을 MEI와 MHI로 표현한 후, Hu의 모멘트를 이용하여 두 영상의 특징을 기술하였다. Hu의 모멘트를 이용함으로써 실시간 적용을 위한 낮은 계산 복잡도를 얻을 수 있었지만, 모양 정보를 좀더 강인하게 표현하는 데는 부족함이 있다.

본 논문에서는 인간의 동작을 보다 효과적으로 표현하기 위한 모양 기반 접근 방법으로 모양 시퀀스 기술자를 제안한다. 모양 시퀀스 기술자는 객체의 동작 변화와 그 순서 정보를 모두 포함하여 정확한 동작 정보를 표현 할 수 있다. 2장에서는 제안된 모양 시퀀스 기술자에 대하여 설명하고, 3 장에서는 실험을 통하여 제안된 방법의 성능을 검증하고, 4 장에서는 결론을 맺는다.

2. 모양 시퀀스 기술자

2.1 모양 기술자 배열

객체의 모양 변화와 그 순서를 표현하기 위해 비디오의 각 프레임에서 객체의 모양 특징 값을 순서대로 추출한다. (그림 3)에서와 같이 각 프레임의 객체 모양 정보는 모양 기술자를 이용하여 1차원 모양 특징 값으로 변환되고, 이 특징 값들을 시간 순서에 따라 나열하여 시간에 따른 모양 변화를 표현하기 위한 2차원 특징 값을 만든다. 즉, 모양 기술자 배열은 순서 정보를 유지한 채 움직이는 객체의 모양 변화 정보를 표현하게 된다. 각 프레임에서 객체의 1차원 모양 특징 값을 추출하기 위해서는 MPEG-7 모양 기술자인 ART (Angular Radial Transform) 기술자[24,25]를 이용하였다.

두 동작 사이의 비교 시 대상 동작들의 프레임 수의 차이는 추출된 두 개의 모양 기술자 배열 사이의 유사도 계산을 어렵게 만든다. 따라서, 추출된 모양 기술자 배열을 시간 축으로 선형 보간(linear interpolation)을 적용하여 같은 길이로 정규화 하였다. 각 프레임 사이의 중간 모양을 예측하기 위해서는 3차원 보간이 필요하지만, 계산 복잡도를 고려하여 두 프레임 사이의 2차원 보간을 차선책으로 선택하였다. ART 모양 기술자의 특성에 따라 1차원 모양 기술자 특징 값 사이의 보간이 두 이미지 사이의 보간 후 모양 특징 값 추출과 같은 결과를 얻을 수 있음을 수식 (1)을 통하여 알 수 있다. 이러한 시간 정규화는 대상 응용에 따라 필요한 경우 선택적으로 적용할 수 있다.

$$\int_0^1 \int_0^{2\pi} V^*(\rho, \theta) [\alpha f_1(\rho, \theta) + \beta f_2(\rho, \theta)] d\rho d\theta$$

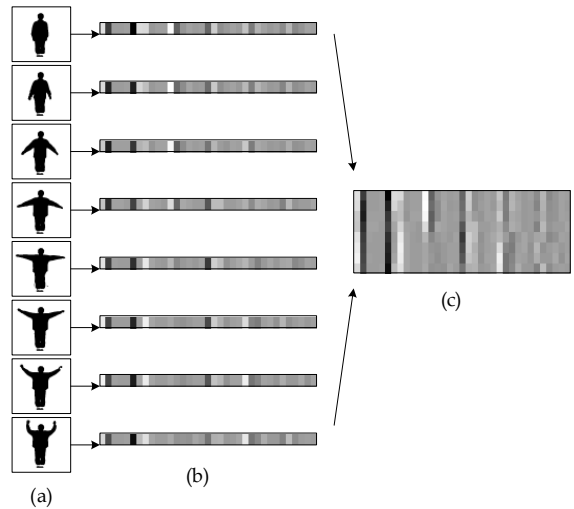
$$= \alpha \int_0^1 \int_0^{2\pi} V^*(\rho, \theta) f_1(\rho, \theta) d\rho d\theta + \beta \int_0^1 \int_0^{2\pi} V^*(\rho, \theta) f_2(\rho, \theta) d\rho d\theta$$

(1)

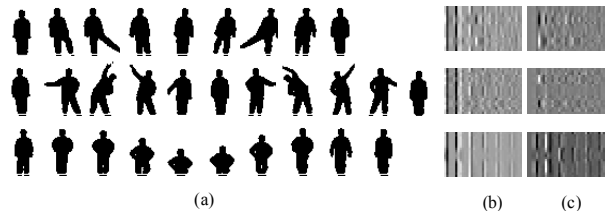
여기서 $V^*(\rho, \theta)$ 는 ART 기저함수, $f(\rho, \theta)$ 는 대상 이미지, α 와 β 는 보간 계수이다.

(그림 3)은 모양 시퀀스를 어떻게 생성하는지 보여주고 있다. 첫 번째 열은 동영상 클립의 연속된 프레임에서 객체를 분리하여 표시한 것이고, 두 번째 열은 첫 번째 열의 추출된 객체 영역에서 ART 기술자를 추출한 것이다. 마지막은 ART 기술자들을 순서대로 나열하여 2차원 배열을 만든 것이다.

MPEG-7에 정의된 ART 기술자는 회전 불변 특성을 얻기 위해 ART 계수의 절대값만을 사용하지만, 본 논문에서는 회전 운동을 포함한 모든 움직임을 기술하기 위하여 모양 기술자 배열을 생성할 때 복소수 형식의 ART 계수를 그대로 이용하였다. (그림 4)는 인간의 몇몇 동작에 대하여 추출된 모양 시퀀스의 예를 보여주고 있다.



(그림 3) 모양 기술자 시퀀스 생성 과정. (a) 입력 이미지, (b) 각 객체의 모양 특징값, (c) 모양 기술자 시퀀스



(그림 4) 모양 기술자 배열 예제. (a) 입력 동작, (b) 모양 기술자 배열 실수부, (c) 모양 기술자 배열 허수부

2.2 모양 시퀀스 기술자

비록 모양 기술자 배열이 3차원 움직임 정보를 2차원 배열로 압축하기는 하였지만, 움직임 인식 및 검색 시스템 구

성을 위해서 보다 효율적이고 축약된 특징 값이 필요하다.

일반적으로 신호 변환은 신호를 주파수 공간 같은 다른 특징 공간으로 배치시키며 신호의 공간 주파수 특성 분석을 위한 정보를 제공해 준다. 주파수 영역에서 낮은 주파수 성분은 신호의 계략적인 윤곽을 나타내며, 높은 주파수 성분은 신호의 세부적 내용을 나타낸다. 이러한 이유로 모양 기술자 배열을 구성하는 ART 계수의 각 열을 시간 축을 따라 주파수 공간으로 변환하면 비슷한 동작 집단의 공통된 특성과 개별화된 특성을 분리하여 얻을 수 있다. 그리하여 (그림 5)에서 보여지는 것과 같이 주파수 변환 후 저주파 영역의 계수들을 각 동작을 표현하는 특징 값으로 선택하여 모양 시퀀스 기술자로 사용한다.

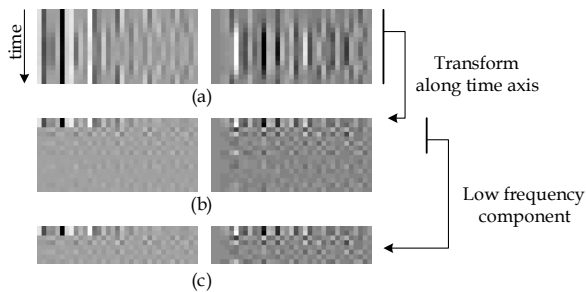
동영상으로부터 모양 시퀀스 기술자를 추출하는 과정을 정리하면 아래와 같다.

- a. 각 동영상 프레임에서 객체 영역을 추출.
- b. 각 객체 영역에서 ART 계수 추출.
- c. ART 계수들을 시간에 따라 배열 후 정규화 하여 모양 기술자 배열 생성.
- d. 생성된 모양 기술자 배열의 각 열을 시간 축으로 주파수 변환.
- e. 저주파 영역의 성분만 선택하여 모양 시퀀스 기술자 생성.

추출된 모양 시퀀스 기술자를 이용하여 동작 비교 시, 입력 동작 Q와 데이터 베이스 상의 동작 S 사이의 유사도 계산은 아래 수식 (2)와 같이 각 기술자 성분의 차이를 절대 값을 취하여 더하는 것으로 정의한다.

$$\text{Distance}(Q,S) = \sum_i \sum_j |Q(i,j) - S(i,j)| \quad (2)$$

여기서, $Q(i,j)$ 와 $S(i,j)$ 은 비교하는 각 기술자의 i 번째 행과 j 번째 열의 값을 의미한다.

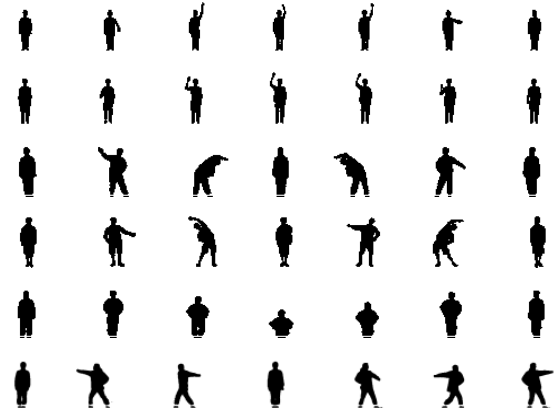


(그림 5) 모양 시퀀스 기술자 생성 과정. (a) 모양 기술자 배열, (b) 주파수 변환된 배열, (c) 모양 시퀀스 기술자

3. 실험 결과

3.1 실험 데이터 셋

실험을 위한 데이터는 5명의 사람들이 22가지 동작을 취



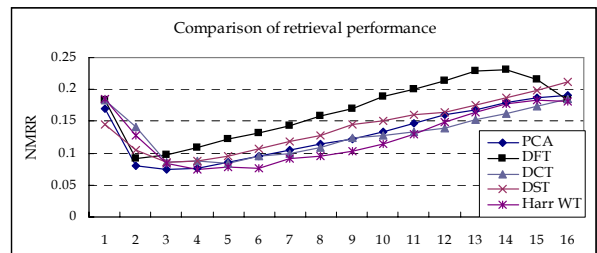
(그림 6) 데이터 셋에 포함된 동작 예제

하는 것을 촬영 후 객체를 분리한 110개의 동영상 클립을 이용하였다. 각 동영상의 프레임 수는 동작의 종류와 동작을 행하는 사람의 특성에 따라 20에서 80 프레임 사이로 변화한다. 몇 가지 동작은 모양의 변화는 비슷하지만 순서가 다른 것들도 포함되었다. (그림 6)은 데이터 셋에 포함된 동작 예제를 보여 준다.

3.2 특징 값 추출과정에서 사용되는 주파수 변환에 대한 고찰

이 절의 실험 목적은 모양 시퀀스 기술자 추출 과정에 사용되는 주파수 변환의 방법에 따른 성능 차이를 비교하고 분석하기 위한 것이다. 주파수 변환 방법으로 PCA, DFT, DCT, DST, Harr Wavelet을 이용하여 비교 하였으며, 또한 각 주파수 변환 방법에서 주파수 변환 후 사용한 저주파 영역 계수의 수를 변화시켜가며 검색 성능을 비교 하였다. 검색 효율에 대한 정량적 비교를 위하여 평가 척도로 MPEG-7 표준화 과정에서 검색 성능 평가 척도로 사용된 NMRR(Normalized Modified Retrieval Rank)을 사용하였다 [23,25]. NMRR은 0과 1 사이의 값을 가지며, 0에 가까울수록 높은 검색 성능을 나타낸다. (그림 7)은 각 경우에 대한 검색 성능을 종합하여 보여 주고 있다.

(그림 7)에서 보여지는 것과 같이 PCA 후 3개의 저주파 계수를 취한 경우 최고의 검색 성능을 얻을 수 있었다. 하지만, PCA의 기저 함수는 데이터 셋에 따라 변하기 때문에 Harr wavelet 변환 후 4개의 계수를 취하는 방법을 권장할



(그림 7) 주파수 변환의 종류와 사용된 계수의 수에 따른 검색 성능 비교. y축은 검색 효율(NMRR), x축은 사용된 계수의 수를 나타낸다.

수 있다. 또한 위 결과에서 앞서 언급한 주파수 변환의 특성대로 고주파 영역으로 갈수록 각 신호의 개별적 특징들이 많이 반영되어 검색 효율이 떨어짐을 알 수 있다.

3.3 검색 효율 비교

검색 효율 비교 실험에서는 객관적 평가를 위하여 temporal template에서 Hu의 모멘트를 추출하는 방법, temporal template에서 ART 기술자를 추출하는 방법, 모양 변화 기술자를 이용하는 방법, 그리고 제안된 모양 시퀀스 기술자를 이용하는 방법을 비교 하였다. 모양 시퀀스 기술자는 Harr wavelet 변환 후 4개의 계수를 선택하여 사용하였다. 각 방법에 따른 검색 효율 측정 결과는 아래 표 1에 정리되어 있다.

<표 1>에서 보여 지듯이 모양 시퀀스 기술자가 가장 좋은 성능을 보여 주었다. 비록 temporal template에 ART 기술자를 적용하였을 경우 Hu 모멘트를 이용하는 것보다 좋은 성능을 보여 주었지만, temporal template의 경우 동작의 미묘한 변화를 표현하는데 모양 시퀀스 기술자 보다 부족함을 보였다. 모양 변화 기술자의 경우 앞에서 언급하였듯이 동작의 순서 정보를 가지고 있지 못하기 때문에 모양 변화는 비슷하지만 순서가 다른 동작들을 구분하지 못하여 성능 저하를 가져왔다.

<표 1> 검색 효율 비교 (NMRR)

Method	NMRR
Hu 모멘트를 이용한 Temporal Template	0.353
ART 기술자를 이용한 Temporal Template	0.244
모양 변화 기술자	0.271
모양 시퀀스 기술자	0.076*

3.4 동작 인식 성능 비교

이 절의 실험에서는 동작 인식 응용의 적용을 위해 모양 시퀀스 기술자를 이용한 간단한 동작 인식 시스템을 구성하고 동작 인식을 수행 후 각 동작의 인식률을 비교하였다. 동작 인식 시스템은 복잡한 학습이나 통계적 모델을 적용하지 않고 단순히 22가지 동작 그룹에 속한 데이터 중 하나씩을 그룹의 대표 모델로 선택하고 나머지 데이터들을 질의하여 대표 모델들과 비교 하여 가장 유사도가 높은 모델의 동작을 인식 동작으로 판단 하였다. 실험 과정을 정리하면 아래와 같다.

- a. 각 동작 그룹에서 하나의 대표 모델을 선택.
- b. 그룹의 나머지 데이터들로 질의 후 대표 모델들과 비교하여 동작 인식률 계산.
- c. 대표 모델을 바꿔가며 앞의 (b) 과정을 반복.
- d. 모든 질의에 대한 평균 인식률 계산.

인식률은 수식 (3)과 같이 전체 질의 중 올바른 인식한 질의의 비율로 정의 하였다.

$$Recognition\ rate(\%) = \frac{correct\ response}{total\ query} \times 100 \quad (3)$$

<표 2>는 인식률 실험에 대한 결과를 정리하여 보여 주고 있다.

실험 결과는 인식률에서도 모양 시퀀스 기술자가 가장 좋은 성능을 나타냄을 보여 준다. 또한 복잡한 학습이나 모델 선정 과정 없이 간단한 거리 계산 만으로도 좋은 인식률을 보임을 알 수 있다.

<표 2> 각 방법에 대한 인식률 비교

Method	Recognition rate
Hu 모멘트를 이용한 Temporal Template	45.91%
ART 기술자를 이용한 Temporal Template	62.73%
모양 시퀀스 기술자	89.09%*

4. 결 론

본 논문에서는 객체의 동작을 인식하고 검색하기 위한 모양 시퀀스 기술자를 제안하였다. 제안된 방법은 객체의 모양 정보에 기반하여 동작을 표현하기 때문에 객체의 모양 및 그 동작을 모두 구분하는 것을 목적으로 한다. 동영상에서 각 객체의 연속된 모양 정보는 모양 기술자를 이용하여 축약된 특징 값으로 변환되고, 정규화 과정과 주파수 변환 과정을 거쳐 모양 시퀀스 기술자가 생성된다. 모양 시퀀스 기술자는 시간에 따른 객체의 움직이는 동작을 간단하면서도 효과적인 방법으로 표현할 수 있으며, 빠르고 정량적인 비교를 가능하게 한다. 실험 결과를 통하여 비정형 객체의 모양 변화에 의한 동작을 인식하거나 검색하는데 기대할 만한 성능을 보여 주었으며, 사람의 움직임에 기초한 동작 인식 및 인지적 관점의 내용 기반 검색 응용에 적용 가능성을 보였다.

참 고 문 헌

[1] S. F. Chang et al, "A Fully Automated Content-Based Video Search Engine Supporting Multi-Objects Spatio-Temporal Queries," *IEEE Transaction on Circuit and Systems for Video Technology*, Vol.8, No.5, pp.602-615, 1998.

[2] Y. P. Tan, S. R. Kulkarni and P. J. Ramadge, "Rapid estimation of camera motion from compressed video with application to video annotation," *IEEE Transaction on Circuit and Systems for Video Technology*, Vol.10, No.1, pp.133-146, 2000.

[3] J. Aggarwal and Q. Cai, "Human Motion Analysis: A review," *Computer Vision and Image Understanding*, Vol.73, No.3, pp.428-440, 1999.

- [4] Aron F. Bobick and James W. Davis, "The Recognition of Human Movement Using Temporal Templates," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.23, No.3, pp.257-267, 2001.
- [5] J. A. Webb and J. K. Aggarwal, "Visually interpreting the motion of objects in space," *IEEE Computer*, Vol.14, No.8, pp.40-46, 1981.
- [6] F. J. Perales and J. Torres, "A system for human motion matching between synthetic and real image based on a biomechanic graphical model," *Proceedings of IEEE Computer Society Workshop on Motion of Non-Rigid and Articulated Objects*, pp.83-88, November 1994.
- [7] J. Rehg and T. Kanade, "Model-Based Tracking of Self-Occluding Articulated Objects," *Proceedings of Int'l Conference on Computer Vision*, pp.612-617, August 1995.
- [8] L. Goncalves, E. DiBernardo, E. Ursella and P. Perona, "Monocular Tracking of the Human Arm in 3D," *Proceedings of Int'l Conference on Computer Vision*, pp.764-770, August 1995.
- [9] D. Hogg, "Model-Based Vision: A Paradigm to See a Walking Person," *Image and Vision Computing*, Vol.1, No.1, pp.5-20, 1983.
- [10] K. Rohr, "Towards Model-Based Recognition of Human Movements in Image Sequences," *CVGIP, Image Understanding*, Vol.59, No.1, pp.94-115, 1994.
- [11] K. Akita, "Image Sequence Analysis of Real World Human Motion," *Pattern Recognition*, Vol.17, No.1, pp.73-83, 1984.
- [12] T. Darrell and A. Pentland, "Space-Time Gestures," *Proceedings of Computer Vision and Pattern Recognition*, pp.335-340, 1993.
- [13] Y. Cui, D. Swets and J. Weng, "Learning-Based Hand Sign Recognition Using Shoslif-m," *Proceedings of Int'l Conference on Computer Vision*, pp.631-636, August 1995.
- [14] J. Yamato, J. Ohya and K. Ishii, "Recognizing Human Action in Time Sequential Images Using Hidden Markov Models," *Proceedings of Computer Vision and Pattern Recognition*, pp.379-385, 1992.
- [15] A. Wilson and A. Bobick, "Learning Visual Behavior for Gesture Analysis," *Proceedings of IEEE Int'l Symposium on Computer Vision*, November 1995.
- [16] J. Little and J. Boyd, "Describing Motion for Recognition," *Proceedings of IEEE Int'l Symposium on Computer Vision*, pp.235-240, November 1995.
- [17] R. Polana and R. Nelson, "Low Level Recognition of Human Motion," *Proceedings of IEEE Workshop Non-Rigid and Articulated Motion*, pp.77-82, 1994.
- [18] E. Shavit and A. Jepson, "Motion Understanding Using Phase Portraits," *Proceedings of IJCAI Workshop: Looking at People*, pp.101-108, 1993.
- [19] J. M. Siskind, "Grounding Language in Perception," *Artificial Intelligence Rev.*, Vol.8, pp.371-391, 1995.
- [20] Y. Yacoob and L. Davis, "Recognizing Human Facial Expressions Form Long Image Sequences Using Optical Flow," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.18, No.6, pp.636-642, 1996.
- [21] I. Essa and A. Pentland, "Coding, Analysis, Interpretation, and Recognition of Facial Expressions," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.19, No.7, pp.757-763, 1997.
- [22] Ashok Veeraraghavan, Amit K. Roy-Chowdhury and Rama Chellappa, "Matching Shape Sequences in Video with Applications in Human Movement Analysis," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.27, No.12, pp.1896-1909, 2005.
- [23] MPEG-7 Visual Group, "Text of ISO/IEC 15938-3/FDIS Information technology - Multimedia content description interface - Part 3 Visual", ISO/IEC JTC1/ SC29/WG11 N4358, Sydney, July 2001.
- [24] B. S. Manjunath, Philippe Salembier and Thomas Sikora, Introduction to MPEG-7: multimedia content description interface, West Sussex, England: John Wiley & Sons, 2002.
- [25] Min-seok Choi and Whoi-yul Kim, "The description and retrieval of a sequence of moving objects using Shape Variation Map," *Pattern Recognition Letters*, Vol.25, issue 12, pp.1369-1375, 2004.
- [26] MPEG-7 Visual Group, "Descriptor of Core Experiments for MPEG-7 Color/Texture Descriptors," ISO/IEC JTC1/ SC29/WG11 N2929, Melbourne, October 1999.



최민석

e-mail : mschoi@hanyang.ac.kr

1996년 한양대학교 전자공학과(학사)

1998년 한양대학교 대학원 전자공학과
(공학석사)

2004년 한양대학교 대학원 전자통신전파
공학과(공학박사)

2002년~현재 삼육대학교 경영정보학과 겸임교수

2007년~현재 삼육대학교 정보전력건설링연구소 책임연구원

2007년~현재 아이노티 커뮤니케이션 기술이사

관심분야 : Contents-base Image/Video Retrieval, Shape Analysis, Computer Vision