

비디오 내 이동 객체의 색인 정보를 이용한 궤적 유사도 측정 기법

(Similarity Measurement Method of Trajectory using Indexing Information of Moving Object in Video)

김정인*, 최창*, 김판구**

(Jeong In Kim, Chang Choi, Pan Koo Kim)

요약

멀티미디어 데이터의 사용이 증대됨에 따라, 이를 관리하고 검색하기 위한 다양한 연구 및 시스템이 개발되고 있다. 하지만 일반적인 검색 방법이 비디오 데이터 내 관련 태그정보나 제목을 통해 검색이 되기 때문에 많은 어려움이 있다. 따라서 본 논문에서는 비디오 검색을 위해 비디오 내 이동 객체의 정보를 이용한 궤적 정보를 통해 유사도 측정 기법에 대해 기술한다. 전체적인 과정은 CCTV 비디오 데이터를 그레이 스케일화 하여, 이동 객체를 추출한 후 라벨링 과정을 통해 궤적을 추출한다. 이를 통해 유사도 측정을 위한 TSR(Tangent Space Representation)과 DTW(Dynamic Time Warping) 알고리즘을 사용하여 두 알고리즘을 비교 분석한다.

■ 중심어 : | 궤적 유사도 | TSR | DTW |

Abstract

The recent proliferation of multimedia data necessitates the effectively and efficiently retrieving of multimedia data. These research not only focus on the retrieving methods of text matching but also on using the multimedia data features. Therefore, this paper is a similarity measurement method of trajectory using indexing information of moving object in video. for similarity measurement. This method consists of 2 steps. Firstly, Video data is processed indexing for trajectory extraction of moving objects using CCTV. Finally, we describe to compare DTW(Dynamic Time Warping) to TSR(Tangent Space Representation) algorithm.

■ keyword : | Trajectory Similarity | TSR | DTW |

I. 서 론

최근 컴퓨터의 발전과 응용 기술의 발달로 인해 멀티미디어 데이터 사용이 증대되었다. 이에 따라 멀티미디어 데이터를 효과적으로 저장하고 검색하기 위한 다양한 연구들이 진행되고 있다. 특히, 멀티미디어 데이터의 한 종류인 비디오 데이터에 대한 많은 연구가 진행되고 있으며 많은 시스템이 개발되었다[1].

비디오 데이터는 영상, 소리, 텍스트 등 다양한 정보로 이루어져 있기 때문에 검색이 매우 복잡하고 까다로우므로 이를 효과적으로 검색하기 위한 기법들이 필요하다[2,3]. 비디오 데이터 검색 기법으로는 비-파라미터(Non-Parametric) 기반의 움직임 분류를 통한 비디오 영상 검색 기법이 있다. 이는 비디오 압

축 스트림을 정규화시킨 움직임 벡터를 추출하여, 각각의 움직임 벡터를 양자화함으로써 평균, 방향, 분산 등을 고려하여 영상을 검색하는 기법이다[4]. 키 프레임(Key Frame) 추출 기법을 이용한 내용 기반 검색 기법은 기준 누적 히스토그램을 활용한 검색 방법이나, 매크로 블록 타입 정보를 이용한 추출 방법이 아닌 압축 영역 내에 히스토그램의 차이 정보를 이용하여 키 프레임을 추출하여 비디오 검색에 활용하였다[5]. 이처럼, 키 프레임을 이용한 많은 연구가 진행되었지만, 현재까지도 비디오 검색에 있어서 많은 어려움이 존재한다. 따라서 본 논문에서는 도로 CCTV 비디오 데이터를 프레임별로 분류하고 이를 그레이 스케일화 한 뒤, 이동 객체를 추출한다. 추출된 이동 객체는 라벨링 과정을 거쳐 궤적을 추출해 궤적 유사도 기법인 TSR(Tangent Space Representation)과 DTW(Dynamic

* 학생회원, 조선대학교 컴퓨터공학과

**정회원, 조선대학교 컴퓨터공학과

본 연구는 교육과학기술부와 한국연구재단의 지역혁신인력양성사업으로 수행된 연구결과임(2011-04-기-03-023).

접수번호 : J1_201200004

접수일자 : 2012년 08월 07일

심사완료일 : 2012년 09월 21일

교신저자 : 김판구 e-mail : pkkim@chosun.ac.kr

Time Warping) 과정을 거쳐 두 가지 유사도 기법의 비교 분석을 해보고자 한다.

본 논문의 구성은 다음과 같다. 2장에서는 비디오 데이터 내 객체와 관련된 연구에 대해 살펴보고, 3장에서는 도로 CCTV 비디오 데이터 내 이동 객체를 추출하여 색인화하는 방법에 관해 기술한다. 4장에서는 각 궤적의 유사도를 측정하기 위해 TSR(Tangent Space Representation)과 DTW(Dynamic Time Warping) 알고리즘을 사용하고 두 알고리즘을 비교 분석한다. 마지막으로, 5장에서는 결론 및 향후 연구 방향을 제시한다.

II. 관련 연구

객체의 움직임 정보는 비디오 데이터의 의미적 이해를 위한 가장 주요한 요소라고 할 수 있으며, 이 움직이는 객체가 이동한 자취를 궤적(trajectory)이라 한다. 앞서 언급한 이동 객체 궤적을 측정하는 기법에는 EU(Euclidean distance), DTW(Dynamic Time Warping distance), ERP(Edit distance with Real Penalty), LCSS(Longest Common Sub-Sequences), EDR(Edit Distance in Eeal Sequence), TSR(Tangent Space Representation) 등이 있으며 이를 토대로 다양한 연구들이 진행되고 있다.

본 논문에서는 이동 객체 궤적들에 대한 유사성을 측정하는 알고리즘으로 서로 다른 두 개의 커브를 비교하는데 많이 쓰이는 DTW(Dynamic Time Warping) 알고리즘과 2, 3차원 객체의 형태(Shape)에 대한 유사도를 측정할 때 많이 사용되는 TSR(Tangent Space Representation) 알고리즘을 사용해 비디오 데이터 내 객체의 궤적을 비교하고자 한다[6,7].

1. Dynamic Time Warping

DTW(Dynamic Time Warping)는 Sakoe와 Chiba에 의해 1978년 처음 소개되었다. DTW 기법은 음성인식, 생물정보 공학, 온라인 필기체 문자 인식 등에서도 널리 사용되고 있는 알고리즘으로, 두 Time Series 간의 거리(Distance)를 최소화하는 방향으로 Dynamic하게 매칭시켜 각 템플릿과의 누적 거리를 계산하여 최소가 되는 클래스로 인식한다.

그림 1을 살펴보면 DTW로 매칭을 하였을 때 단순 Euclidean distance로 매칭했을 때와 달리 부분적으로 웨곡되거나 shift 된 파형에 대해서도 적절한 매칭이 가능하다는 것을 알 수 있다.

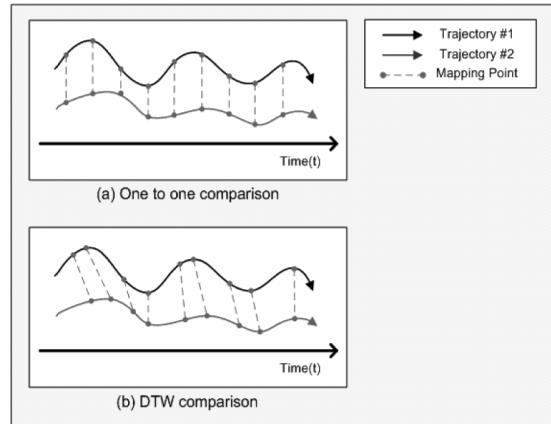


그림 1. 단순 매칭과 DTW 매칭[6]

2. Tangent Space Representation

TSR(Tangent Space Representation)은 주로 2, 3차원 객체의 형태(Shape)에 대한 유사도를 측정할 때 많이 사용된 방법이나 본 논문에서는 궤적 간 유사성 측정을 위해 적용하였다. 객체의 형태는 선으로 구성되어 있고, 이 선은 좌표 간의 연결이라고 할 수 있기 때문에, 궤적을 TSR로 표현할 수 있다. 이를 통해 객체의 방향과 이동거리를 고려한 유사성 측정이 가능하고 움직이는 객체의 궤적의 크기에 상관없이 정형화를 시킬 수 있으며, 면적을 통하여 유사도를 구할 수 있다.

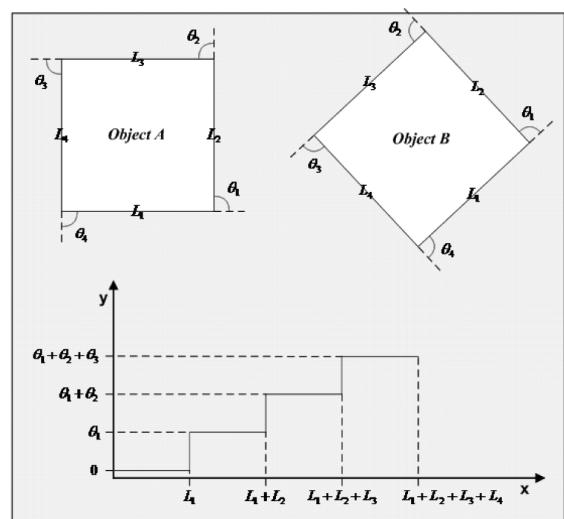


그림 2. 이동 객체 궤적의 TSR 변환[7]

그림 2는 비디오 데이터에서 추출한 이동 객체의 궤적을 TSR(Tangent Space Representation)로 변환한 것이며, 이를 통해 객체의 궤적 사이의 유사도를 측정할 수 있다.

III. 이동 객체의 색인화

비디오 내 움직임 객체들은 시간(t)의 흐름에 따라 2차원 평면 공간(x, y)에 움직임 정보가 표현된다. 즉, 시간에 따른 이동 객체의 위치가 x좌표와 y 좌표의 집합으로 나타나게 되며, 이를 통해 이동 객체의 궤적 정보로 색인된다. 이와 같이 이동 객체의 움직임 인식을 위해서는 색인화 과정이 필수적이며, 일반적으로 다음과 같은 5단계를 거치게 된다.

- 1 단계 : 입력 영상의 Grayscale로 변환
- 2 단계 : 배경 영상 제거 및 이미지 보정
- 3 단계 : 객체 추출 및 라벨링
- 4 단계 : 이동 객체의 궤적 추출
- 5 단계 : 객체 색인화

첫 번째 단계에서는 입력 영상을 프레임별로 분할한다. 분할한 프레임별 이미지는 RGB색상으로 각각 0~255 값으로 표현되며, 이를 그레이 스케일화하여 색상 차이를 256 단계로 변화시켜 처리를 용이하게 한다. 두 번째 단계에서는 프레임별 차(Difference) 영상을 통해 배경을 제거하는데, 고정된 영상만 가능하다. 일반적으로 픽셀 기반의 배경 분리(Pixel-based Background Subtraction) 알고리즘에 의해 배경과 이동 객체를 분리하는 방법이 사용된다. 이는 임계 값 설정 및 주변 픽셀과의 연관성을 고려가 필요하기 때문에, 본 논문에서는 이를 이용하고자 한다. 이와 같은 차 영상은 Median Filter와 Threshold 연산 등의 과정을 통해 노이즈를 제거하고, 이미지를 보정한다. 세 번째 단계에서는 비디오마다 이동 객체의 수가 다르므로 라벨링(Labeling)을 통해 각각의 이동 객체를 분류하고 저장한다.



그림 3. 이동 객체 라벨링(Labeling)[6]

그림 3은 이동 객체 라벨링에 대한 예로써 이동 객체를 MBR(Minimum Bounding Rectangle)로 추출함으로써 이동 객체를 분류하고 있으며, 가장 큰 사각형은 고정 객체 영역으로

향후 이동 객체와 고정 객체 사이의 의미적 인식을 위해 사용자가 직접 지정한 영역이다.

네 번째 단계에서는 추출된 이동 객체에 대해 그 움직임 궤적 정보를 x좌표와 y좌표로 추출하는 단계이며, 마지막으로 다섯 번째 단계는 이동 객체를 색인화 하는 단계로 그림 4와 같이 이동 객체의 움직임과 고정 객체를 인덱싱하여 나타내고 있으며, 비디오 정보에서 사각형은 설정한 고정 객체의 영역으로 최초 좌표, 최종 좌표, 높이, 넓이 값을 보이고 있으며, 움직임 정보에서는 동일한 고정 객체를 표현한다. 또한, 움직임 정보는 이동 객체에 대한 영역을 인덱싱하여 보여주고 있으며, 움직임 궤적을 표현하고 있다.

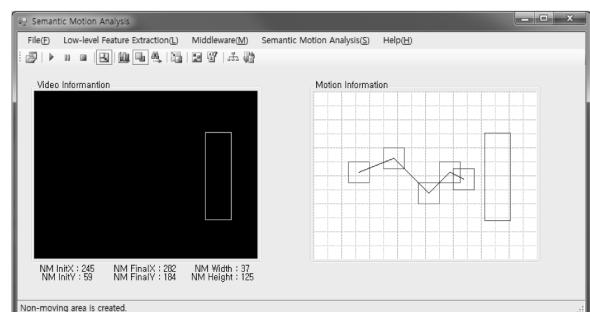


그림 4. 이동 객체 정보의 색인화

이와 같이 색인된 이동 객체의 색인 정보를 바탕으로 이동 객체의 궤적 유사도 측정을 통해 비슷한 움직임을 분류할 수 있다.

IV. 효율적인 궤적 유사도 측정 기법

본 논문에서는 TSR과 DTW 알고리즘을 이용하여 궤적 유사도 측정에 이용하고자 한다. TSR은 이동 객체의 움직임 정보를 움직인 거리와 각도를 이용하여 면적의 합을 통해 유사성 측정이 가능하다. 그러나 세밀한 처리를 수행하지 않기 때문에 빠른 수행시간을 보이는 장점을 가지고 있다. DTW는 서로 상이한 두 개의 커브를 비교하는데 많이 쓰이는 알고리즘으로 유사한 정도를 파악할 수 있는 알고리즘이다. 이는 시간이 변화하면서 각 t시간에 가장 가까운 거리의 커브상의 점을 매칭하여 비교함으로써 일반적인 비교방법보다 더욱 직관적인 방법이다. 따라서 단순 Euclidean distance로 매칭했을 때와 달리 부분적으로 왜곡되거나 shift 된 과정에 대해서도 적절한 매칭이 가능하다는 장점을 가지고 있다.

그림 3에서 보듯이 입력 영상은 도로 CCTV영상을 대상으로 하며, 고정객체와 이동 객체를 정의한다. 고정 객체로는 교차로, 건널목 등이라고 가정하며 MBR로 표현하여 정의한다. 이를 통해 이동 객체의 궤적에 대해 “교차로로 들어간다(enter)”, “교

차로를 통과하다(go through)", "왼쪽 방향으로 이동하려한다(turn left)", "오른쪽 방향으로 이동하려한다(turn right)", "유턴한다(return) 등의 의미적 정보를 추출할 수 있다. 또한, 같은 움직임을 나타내는 궤적들 사이의 유사성을 본 논문에서는 TSR과 DTW 알고리즘을 수정하여 고정 객체를 기준으로 이동 객체의 궤적 유사도 측정에 활용한다.

1. TSR을 이용한 이동 객체의 궤적 분류

그림 5는 샘플에 대한 이동 객체의 움직임 궤적을 TSR 그래프로 표현한 것이며, 궤적 분류를 위한 기준은 그림 7과 같다.

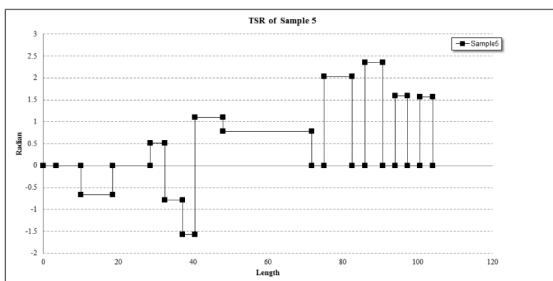


그림 5. TSR 그래프 표현 예[7]

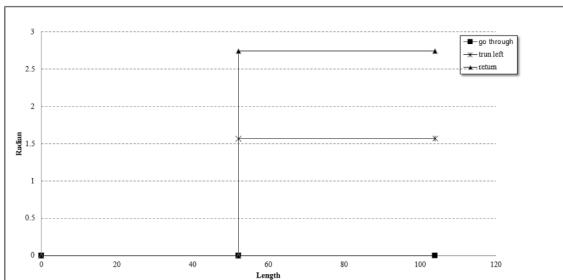


그림 6. 궤적 분류를 위한 기준 움직임의 TSR 그래프 표현[7]

위 그림 5와 그림 6의 면적의 합을 구하면 go_through와는 51.26, turn_left와는 30.46, return은 97.01의 값을 얻을 수 있으며, 가장 값이 작은 turn_left와 가장 유사함을 보이고 있다.

2. DTW를 이용한 이동 객체의 궤적 분류

DTW도 그림 6과 같이 비교 대상이 되는 기준 커브가 필요하며 이는 표 1과 같다. 또한, 기준 커브와 입력된 커브를 더욱 정확하게 비교하기 위하여 기준좌표로의 원점 이동이 필요하다.

표 1. 궤적 그룹핑을 위한 DTW의 기준 커브 함수[6]

$$= \pm 1/ax \quad (0.0000 \leq x \leq a) \quad (1)$$

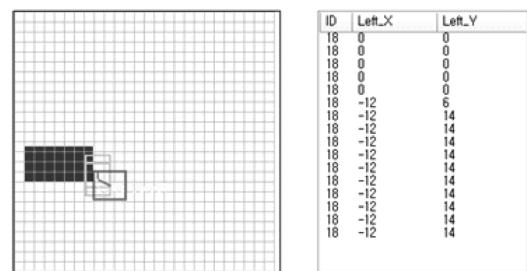
$$y = \pm 1/ax^2 \quad (-a.0000 \leq x \leq 0.0000) \quad (2)$$

$$y = \pm |a|^{1/2}(x) \quad (0.0000 \leq x \leq a) \quad (3)$$

$$y = \pm |\sqrt{a}|^{1/2}\sqrt{(x)} \quad (-a \leq x \leq 0.0000) \quad (4)$$

*a는 측정 영상의 화면 좌표계에 의존하는 상수이다.

위의 표 1의 기준 커브를 이용하여 DTW기반 유사도 측정이 가능하며, 그림 7은 DTW를 이용한 궤적 분류의 예이다. 그림 7에서 4방위를 기준으로 진행방향을 고려하며, 고정 객체 이동 객체의 오른쪽에 위치하고 있음을 알 수 있다. 결과적으로 이동 객체의 궤적은 표 1과 같이 X축과 Y축의 표준편차를 이용하여 각 성분 중 가장 작은 유사도를 가진 성분의 분석과 진행 방향을 알 수 있다.



Dynamic Time Warping 유사도 분석

결과보기		
선택 대상 1 비교=1679		
선택 대상 2 비교=2453		
선택 대상 3 비교=3412		
선택 대상 4 비교=3344		
Go_Through(수행성분) + Turn Right		

그림 7. DTW를 이용한 궤적 분류의 예[6]

V. 결 론

본 논문에서는 도로 CCTV 비디오 데이터를 이용하여 궤적을 추출하고, 추출한 궤적을 통해 궤적 간의 유사도 측정하기 위해서는 TSR 알고리즘과 DTW 알고리즘을 비교하였다. 두 기법은 비교한 결과 TSR 알고리즘은 궤적의 유사도를 측정하는데 있어 빠른 수행시간을 보이는 장점을 가지고 있으나 정밀한 곡선에 적용하는 데 문제가 있었다. 또한, 면적의 합을 구할 시 상쇄되는 면적의 합으로 인한 문제점이 있다. DTW 알고리

즘은 객체의 인식률이 TSR보다 높았지만, 객체의 수가 증가할 수록 유사도를 측정하는 계산량이 증가하여 수행시간이 오래 걸리는 단점을 지녔다. 앞서 언급한 두 유사도 측정 알고리즘을 접목하여 새로운 유사도 측정 알고리즘을 개발한다면 지금보다 더 나은 비디오 데이터 내 이동객체의 유사도를 측정하는 기법 도출이 가능할 것으로 기대된다.

참 고 문 헌

- [1] Yi Li, Linda and G. Shapiro, "Object and Concept Recognition for Content-Based Image Retrieval," in *Lecture Note in Computer Science*, Springer-Verlag, to Appear, 2004.
- [2] Chang,S.-F., Hsu,W.,Kennedy, L.,Xie,L., Yanagawa,A., Zavesky,E., and Zhang,D.-Q., "Video Search and High-LevelFeature Extraction", TRECVID-2005.
- [3] A.Ekin, A.Tekalp, and R.Mehrotra, "Automatic soccervideo analysis and summarization", *IEEE Trans. Image Processing*, 12(7), pp.796~807, 2003.
- [4] 김낙우, 최종수, "비-파라미터 기반의 움직임 분류를 통한 비디오 검색 기법", 전자공학회지 제43권 제2호 통권 제308호, pp. 1-11, 2006. 03.
- [5] 홍보현, 엄민영, 김명호, 최윤식, "고속 key frame 추출 기법을 이용한 내용 기반 비디오 검색 기법", 대한전기학회 학술대회 논문집, pp. 539-541, 2005.10.
- [6] 이홍렬, 김판구, "의미적 움직임 모델링 기반 비디오 내 이동 객체 그룹화", 학위논문(석사), 2008.
- [7] 최창, 김판구, "비디오 내 움직임 객체의 의미적 표현을 통한 궤적 검색", 학위논문(석사), 2007.

저 자 소 개



김정인(학생회원)

2011년 조선대학교 컴퓨터공학부 공학사 졸업.
2011년 ~ 현재 조선대학교 석·박사 연계 과정.

<주관심분야 : 소셜 네트워크, 시맨틱 웹과 온톨로지, 정보검색>



최창(학생회원)

2005년 조선대학교 컴퓨터공학과 공학사 졸업.
2007년 조선대학교 컴퓨터학과 이학석사 졸업.
2012년 조선대학교 컴퓨터공학과 공학박사 졸업.

<주관심분야 : 의미적정보처리, 시맨틱 웹, 멀티미디어 데이터 처리>



김판구(종신회원)

1988년 조선대학교 컴퓨터공학과 공학사 졸업.
1990년 서울대학교 컴퓨터공학과 공학석사 졸업.
1994년 서울대학교 컴퓨터공학과 공학박사 졸업.

1994년 ~ 현재 조선대학교 교수

<주관심분야 : 시맨틱 웹과 온톨로지, 정보검색, 멀티미디어 데이터의 의미적 표현, 감성정보처리>