[UCI]I804:24011-200000234045

*2007*년 *2*월
석사학위논문

비디오 내 움직임 객체의 의미적 표현을 통한 궤적 검색

조선대학교 대학원

컴퓨터학과

최 창

*2007*년 *2*월 석사학위논문

비디오 내 움직임 객체의 의미적 표현을 통한 궤적 검색

최 창

목 차

목 차 ····· i
표 목 차 ····· iii
그 림 목 차 ····· iv
ABSTRACT ····· vi

Ⅰ. 서론 ····· 1
A. 연구배경 및 목적 ····· 1
B. 연구 내용 및 구성 ····· 2

Ⅱ. 움직임 객체에 대한 시공간 관계 모델링 ····· 3
A. 시공간관계 ····· 3
1. 시간관계 ····· 3
2. 위상관계 ····· 4
3. 방향관계 ····· 8
B. 객체 움직임에 대한 계층 관계 모델링 ····· 10

Ⅲ. 움직임 객체에 대한 모션 동사의 적용 ····· 12
A. 움직임 객체에 대한 모션 동사의 정의 ····· 12
B. 움직임 객체 모델링의 검증 ····· 14
1. 제안된 모델링에서 유사도 측정 ····· 14
2. 비교평가 ····· 16

C. 움직임 객체에 대한 모션 동사의 확장 ····· 17
1. WordNet ····· 17
2. WordNet상의 유사도 측정을 통한 모션 동사의 확장 ····· 18
a. WordNet상에서 유사도 측정방법 ····· 18
b. 움직임 객체에 대한 모션 동사의 확장 ····· 20

Ⅳ. 의미기반 궤적 검색 시스템 ····· 23
A. 시스템 구성도 ····· 23
B. 비디오 전처리과정을 통한 움직임 객체의 궤적 추출 ····· 24
1. 비디오 내 움직임 객체 추출 ····· 24
2. 비디오 내 움직임 객체의 궤적 추출 ····· 26
C. 움직임 객체의 궤적과 제안한 모델링간의 유사도 측정 ····· 27
1. 의미적 움직임 객체 판별을 위한 규칙 ····· 28
2. TSR을 이용한 유사도 측정 ····· 32
D. 실험 및 성능평가 ····· 36

Ⅴ. 결론 및 제언 ····· 40

참고문헌 ····· 41

표 목 차

[표 2-1] Chang의 거리 유사도 관계 ····· 6
[표 2-2] [그림 2-12]에서 방향성에 따른 위상관계의 분류 ····· 10
[표 3-1] Levin의 동사의 분류 중 'verbs of Inherently directed motion' ···· 12
[표 3-2] Lisa Pearl의 'GO primitive and a particle' ····· 13
[표 3-3] [그림 2-13]에 대한 모션 동사의 매칭 ····· 14
[표 3-4] [그림 2-14]에서 첫 번째 단계에 대한 거리 ····· 15
[표 3-5] [그림 2-14]에서 1, 2 단계에 대한 유사도 ····· 16
[표 3-6] WordNet상의 엣지(edge)기반 유사성 측정 ····· 16
[표 3-7] WordNet의 "index.verb" ····· 18
[표 3-8] gloss기반 측정방법에서 관계 쌍 ····· 20
[표 3-9] 엣지(edge)기반의 유사도 측정과 모션동사의 확장 ····· 21
[표 3-10] 제안된 모델링에서 모션 동사의 유의어와 반의어 ····· 22
[표 4-1] 궤적추출 알고리즘 ····· 26
[표 4-2] 'Sample 5'에 대한 궤적 정보 ····· 27
[표 4-3] '*Include*', '*same*', '*is_include_by*'에 대한 규칙 ····· 29
[표 4-4] '*leave*'에 대한 규칙 ····· 30
[표 4-5] '*arrive*'에 대한 규칙 ····· 30
[표 4-6] '*come_up*' 대한 규칙 ····· 31
[표 4-7] '*depart*' 대한 규칙 ····· 31
[표 4-8] '*enter*'와 '*exit*'에 대한 규칙 ····· 32
[표 4-9] [그림 4-9]에 대한 유사도 측정 ····· 34
[표 4-10] [그림 4-10]과 [그림 4-11]에 대한 유사도 측정 ····· 36
[표 4-11] 의미 기반 궤적 검색의 Precision rate ····· 38

그 림 목 차

[그림 2-1] Allen이 제안한 X와 Y 사이의 시간관계 ····· 3
[그림 2-2] Egenhofer의 위상관계 ····· 4
[그림 2-3] W Ren의 위상관계 ····· 5
[그림 2-4] Chang의 위치 연산자와 거리관계 ····· 5
[그림 2-5] 좌표에 따른 위상관계 ····· 6
[그림 2-6] 좌표에 따른 모션의 기본 요소 ····· 7
[그림 2-7] 확장된 요소에 따른 위상관계 표현 ····· 7
[그림 2-8] I와 J에 대한 위상관계 표현 ····· 8
[그림 2-9] John Z의 방향관계 ····· 9
[그림 2-10] 9DLT scheme ····· 9
[그림 2-11] 방향성에 따른 모션 ····· 9
[그림 2-12] 객체 움직임에 대한 계층 관계 모델링 ····· 11
[그림 3-1] '*enter*'와 '*come_up*'간의 거리 유사도 ····· 15
[그림 3-2] '*walk*'에 대한 WordNet의 구조 ····· 19
[그림 4-1] 비디오 내 움직임 객체의 의미기반 궤적 검색 시스템의 구성도 ··· 23
[그림 4-2] 이미지에 대한 컬러 세그멘테이션(Color Segmantation) 결과 ····· 25
[그림 4-3] 객체 추출(Object Detection) 결과 ····· 26
[그림 4-4] 비디오 내 움직임 객체의 궤적과 제안한 모델 유사도 측정 방법 · 28
[그림 4-5] IP와 NMP에 따른 최초 움직임 객체의 위치 ····· 29
[그림 4-6] 'Sample 7'의 궤적과 의미적 궤적간의 비교 ····· 33
[그림 4-7] [그림 4-6] 대한 TSR ····· 33
[그림 4-8] 실제 비디오 내의 움직임 객체의 궤적 ····· 34
[그림 4-9] [그림 4-8]에서의 궤적 ····· 35
[그림 4-10] [그림 4-9]에 대한 TSR ····· 35
[그림 4-11] '*go_through*', '*turn_left*', '*return*'에 대한 T-TSR ····· 36
[그림 4-12] 제안한 규칙만 적용했을 경우의 문제점 ····· 37

[그림 4-13] TSR만 적용했을 경우의 문제점 ····· 37
[그림 4-14] 제안한 규칙과 TSR 모두를 고려한 결과 ····· 38

ABSTRACT

Trajectory Retrieval through Semantic Representation of Moving Objects in Video

Chang Choi
Advisor : Prof. Pankoo Kim, Ph.D
Department of Computer Science
Graduate School of Chosun University

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 (color, sound, form) of Audio, Video, Image. Especially in the video research field, varieties of compound multimedia features have been exploited.

There are two distinct research directions that related to the indexing and retrieving of video data. One research direction is by using the visual features of video objects such as color, shape and texture, and auditory features such as audio and music. The other one is by exploiting the spatio-temporal feature of objects to represent, index and retrieve in the video data. In the former area, there has been a lot of research on some specific video data, such as soccer and billiard video. but it is impossible to generalize the visual and auditory features, for every video differs with each other; In the latter one, however, temporal relation, topological relation and directional relation of the moving object in the video are able to acquire and use

for making a definition of semantics for moving object

Therefore, in this paper, It will specify how to give definitions to the motions of general moving objects by using the spatio-temporal relation, and match between the definition of moving objects and motion verbs in the vocabulary.

To achieve this, first the spatio-temporal relation models using topological relation and directional relation have been made, then a matching of the proposed models with the directional motion verbs proposed by Levin's verbs of Inherently directed motion is applied, finally the synonym and antonym are inserted by using WordNet.

For the similarity measuring between proposed modeling and trajectory of moving object in the video, first extracted the objects, and then compare the trajectories of moving objects with the proposed modeling. Because of the different features with each proposed modeling, finally the rules that have been generated will be applied to the similarity measurement by TSR(Tangent Space Representation).

Ⅰ. 서론

A. 연구배경 및 목적

최근 멀티미디어 데이터의 증가에 따라 효율적이고 정확한 멀티미디어 데이터의 검색의 필요성이 대두되고 있으며, 단지 텍스트매칭(Text Matching) 검색방법을 벗어나 오디오, 비디오, 이미지 등의 멀티미디어 데이터의 특징을 이용한 검색방법들이 연구되고 있다. 특히 비디오의 경우는 컬러, 음성, 형태 등 복합적이고 다양한 특징들을 이용한 검색이 활발히 연구되고 있다.

비디오 데이터의 색인(Indexing)과 검색(Retrieving)과 관련하여 크게 2 가지의 연구방향으로 나눌 수 있다[1]. 첫째는 비디오 내 객체의 컬러(Color), 형태(Shape), 질감(Texture) 등의 시각적인 특징과 오디오(Audio)와 음악(Music) 등의 청각적인 특징을 이용하는 것이고[2,3], 둘째는 비디오 내 객체의 시공간적 특징을 이용하여 비디오의 특정 부분을 검색, 색인하는 연구이다[4,5,6,7,8,9]. 전자의 경우는 축구, 당구 등 특정 비디오에 대한 연구들이 활발하게 이루어지고 있으나 비디오마다 시각적인 특징과 청각적인 특징이 서로 다르므로 이를 일반화시키기에는 다소 무리가 있다. 후자의 경우는 비디오 내 움직임 객체의 시간관계(Temporal relation), 위상관계(Topological relation), 방향관계(Directional relation) 등을 이용하여 일반적인 객체의 움직임에 대한 정보를 얻을 수 있고, 이를 통하여 특정 비디오에 대한 규칙을 설정하여 의미적인 객체의 움직임을 정의하려는 연구가 시도되고 있다.

위 두 접근법을 바탕으로 본 논문에서는 움직임 객체의 시공간적 특징을 이용하여 일반적인 객체의 움직임을 정의하고, 이들의 움직임을 모션 동사를 이용하여 어휘적으로 매칭시켜 의미적인 객체의 움직임을 정의해보고자 한다. 즉, 비디오 내 움직임 객체의 궤적(Trajectory)과 정의한 객체의 움직임 사이의 비교를 통해 비디오 내 움직임 객체의 의미적 검색에 활용해 보고자 한다.

B. 연구내용 및 구성

비디오 데이터의 의미적 이해를 위한 가장 중요한 요소는 객체의 움직임 정보라고 할 수 있다. 객체의 움직임 정보는 이를 구성하고 있는 프레임(Frame)들의 일련의 시간적인 관계를 통해 공간상의 위치변화이다. 다시 말해, 시간의 흐름에 따라 객체의 위상과 방향이 변하며, 이를 통해 움직임 정보가 구성된다. 여기서 움직이는 객체의 연속적인 변화를 모션(Motion)이라고 하며, 이 움직이는 객체가 이동한 자취를 궤적(trajectory)이라 한다.

객체의 위상 관계는 비디오에서 객체의 움직임 정의를 위한 기본적인 요소이며, 방향 관계는 비디오의 특성상 2차원에서 이 위상의 움직이는 방향에 대한 관계를 나타내고 있다. 따라서 본 논문에서는 시간에 따른 위상관계와 방향관계를 이용하여 시공간 관계 모델을 생성하고, 제안한 시공간 관계 모델을 Levin의 방향성 모션 동사들(verbs of Inherently directed motion)[10]을 이용하여 어휘적으로 매핑하였으며, 이를 WordNet[11]을 이용하여 동의어와 반의어 관계를 추가하였다. 또한 제안한 객체 움직임에 대한 관계를 모델링하고, 이를 이용하여 의미적 인식을 위한 각 모델간의 특징에 따른 규칙을 생성하였으며, 의미 기반 검색에 활용하였다. 특히, 비디오 내 움직임 객체의 궤적과 정의한 모델링간의 유사도 측정을 위해 각 모델간의 특징에 따른 규칙과 TSR(Tangent Space Representation)[12,13]을 이용하였다.

본 논문의 구성은 다음과 같다. 제 1장 서론에 이어 제 2장은 움직임 객체의 시간에 따른 위상관계와 공간관계를 통하여 객체 움직임을 위한 시공간 관계 모델링을 제안하며, 제 3장에서는 2장에 정의한 모델링을 Levin이 제안한 모션동사들과 어휘적으로 매칭하고 WordNet을 이용, 이를 확장하여 움직임 객체에 대한 모션 동사의 적용한다. 그리고 제 4장에서는 비디오 내 움직임 객체의 추출과 이 추출한 객체의 궤적 추출방법에 대해 기술하며, 비디오 내 움직임 객체의 궤적과 제안한 모델링간의 유사도 측정을 통해 정의한 규칙과 TSR을 이용하여 의미 기반으로 검색한다. 마지막으로 5장에서 향후연구 방향에 대해 서술하며 본 연구의 결론을 맺는다.

Ⅱ. 움직임 객체에 대한 시공간 관계 모델링

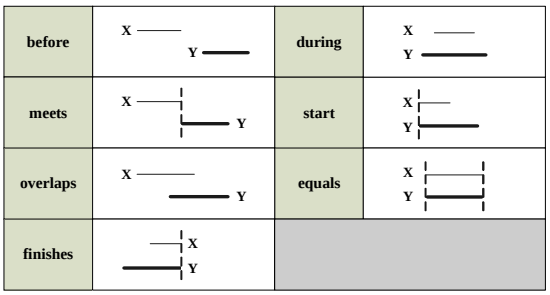
비디오에서 의미를 가지는 최소의 단위가 신(Scene)이며, 여기서 가장 중요한 요소 중 하나가 객체(Object)이다. 이는 객체에 따라 의미가 달라질 수 있으며, 이 객체의 움직임 또한 의미 변화의 중요한 요인이 되기 때문이다. 이에 본 장에서는 기존의 시공간관계를 이용한 연구들을 살펴보고, 이를 통하여 비디오에서 객체 움직임 분석을 위한 시공간 관계 모델링을 제안하고자 한다.

A. 시공간관계

시공간관계는 시간의 흐름에 따라 객체의 위상관계와 방향관계로 객체를 표현하는 것을 말한다. 본 논문에서는 Allen의 시간관계를 이용하고, 움직이는 객체의 위상관계와 방향관계는 관련된 연구들을 바탕으로 새로운 시공간 관계를 제시하였다.

1. 시간관계

시공간관계에서 시간위상관계연산자는 대상이 되는 객체들의 시간 값 사이의 위상 관계에 대한 연산으로 대부분 Allen[14]이 제안한 시간에 관한 연산자를 기초로 하여 많은 연구가 진행되고 있다. [그림 2-1]은 Allen이 제안한 시간관계를 도식화한 것으로 기본적인 관계를 표현한 것이다.



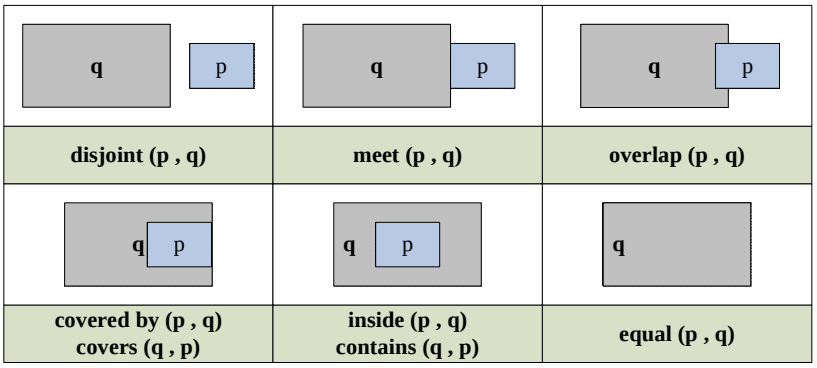
[그림 2-1] Allen이 제안한 X와 Y 사이의 시간관계

여기서 같은 비디오에서 시간관계에 따라 움직이는 객체는 위와 같은 관계를 이용할 수 있다. 그러나 본 논문에서는 각각 다른 비디오 혹은 같은 비디오의 신에서 움직이는 객체에 대한 궤적간의 유사도를 측정해야하므로, 위의 관계에서 '*equals*' 관계만 이용하기로 한다. 이러한 시간적 특징을 이용하여 움직이는 객체의 의미적 표현을 위해 우선 비디오의 위상관계를 통한 움직임을 분류하고, 방향관계를 적용하여 좀 더 세부적인 움직임에 대해 분석하였다.

2. 위상관계

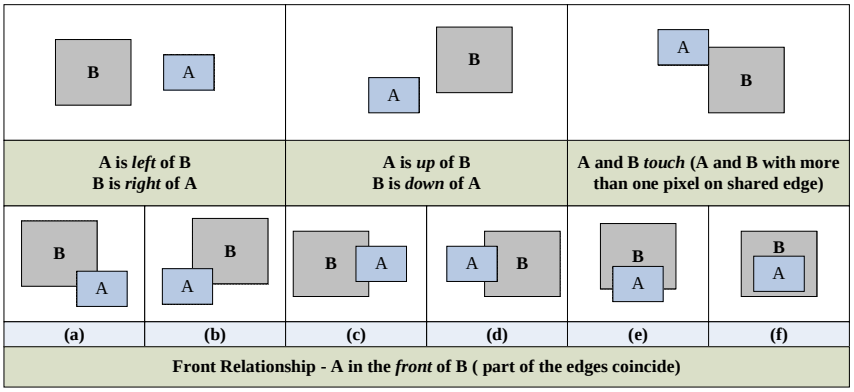
위상관계는 시간의 흐름에 따른 움직이는 객체의 위상관계를 나타내는 것으로, 기본에 되는 위상관계는 움직이는 두 객체 사이의 관계이다. 이를 시간에 따라 나누어 보면 한 객체가 움직이고 있을 때, 한 객체는 멈추어 있다고 말할 수 있다. 이러한 일련의 과정들의 반복을 통해 움직이는 두 객체 나아가 움직이는 다수의 객체 사이의 관계를 표현할 수 있다. 따라서 본 논문에서는 가장 기본이 되는 고정된 객체와 움직이는 객체 사이의 관계에 대해 표현하고, 이를 확장하도록 하겠다.

시공간관계를 통한 객체의 위상관계에 관한 많은 연구가 진행되었다. Egenhofer와 Franzosa[4]는 [그림 2-2]와 같이 공간 관계의 표현을 위해 8개의 위상관계 - '*equal*', '*inside*', '*cover*', '*overlap*', '*touch*', '*disjoint*', '*covered-by*', '*contains*' - 를 제안하였고, S.Y. Lee과 F.J. Hsu[5]는 2차원 상에서의 표현을 다양하게 확장하기 위하여 2D C-string을 제안하였다.



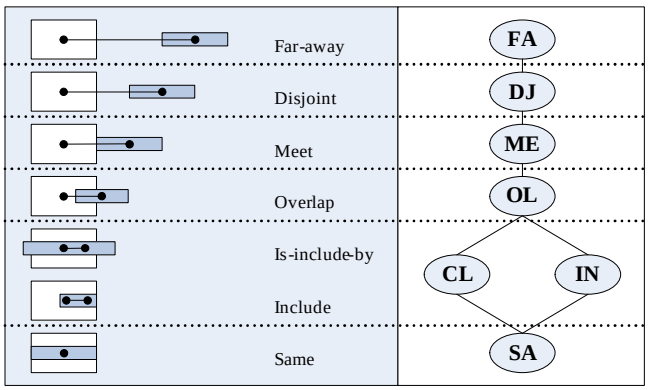
[그림 2-2] Egenhofer의 위상관계

W. Ren[8]은 [그림 2-3]과 같이 특별한 6가지의 위상관계 - '*Left*', '*Right*', '*Up*', '*Down*', '*Touch*', '*Front*' - 를 통하여 공간관계와 결합하였다.



[그림 2-3] W. Ren의 위상관계

Chang[6][7]에 의해 제안된 SMR 스키마(Scheme)와 SRR 스키마는 움직이지 않는 객체를 기준으로 움직이는 객체와의 공간적 관계를 기술하고 있다. SMR 스키마는 15가지의 위상 연산자를 생성하였으며, SRR 스키마는 SMR 스키마를 바탕으로 고정된 두 객체의 거리관계를 이용하여 7가지의 위상관계를 제안하였다. 다음 [그림 2-5]는 Chang[7]에 의해 제안된 SRR 스키마로 위상간의 거리관계 표현과 위치 연산자를 나타낸 것이다.



[그림 2-4] Chang의 위치 연산자와 거리관계

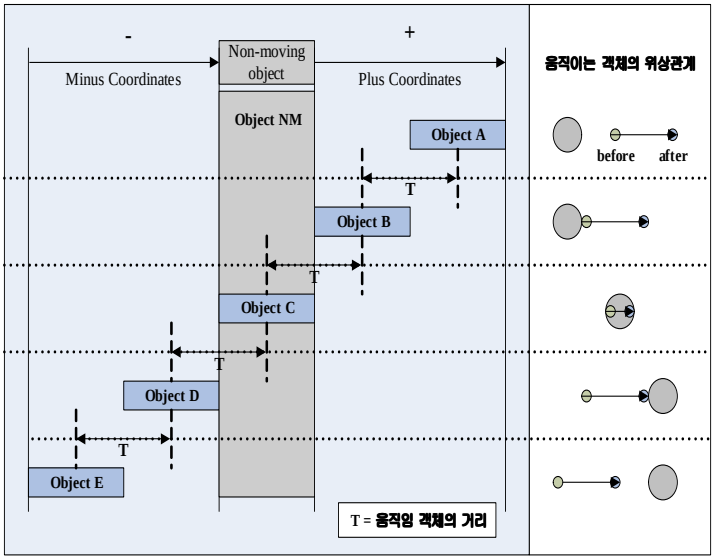
위의 위상관계모델은 움직임 객체의 위치에 따라 가장 인접한 위상관계를 표현하고 있다. 이는 방향성을 제외한 위상 관계적 측면에서 바라본 거리이다.

[표 2-1]은 [그림 2-4]의 위상관계에 따른 거리 유사도 관계이다.

	Far-away (FA)	Disjoint (DJ)	Meet (ME)	Overlap (OL)	Is-include-by (CL)	Same (SA)	Include (IN)
FA	0	1	2	3	4	5	4
DJ	1	0	1	2	3	4	3
ME	2	1	0	1	2	3	2
OL	3	2	1	0	1	2	1
CL	4	3	2	1	0	1	2
IN	5	4	3	2	1	0	1
SA	4	3	2	1	2	1	0

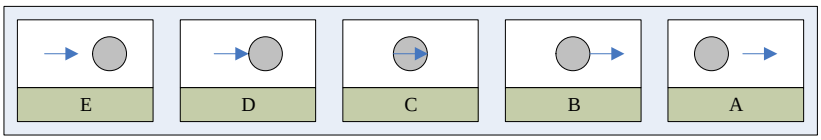
[표 2-1] Chang의 거리 유사도 관계

본 논문에서는 위의 관계를 기반으로 구분이 모호한 '*FA*'와 '*DJ*'는 같다고 가정하고 좌표에 따라 세부적인 위상 관계를 추가하였으며, 방향에 따른 관계를 설정하였다. [그림 2-5]에서는 좌표에 따라 움직이는 두 객체 사이의 관계를 표현하고 있다.



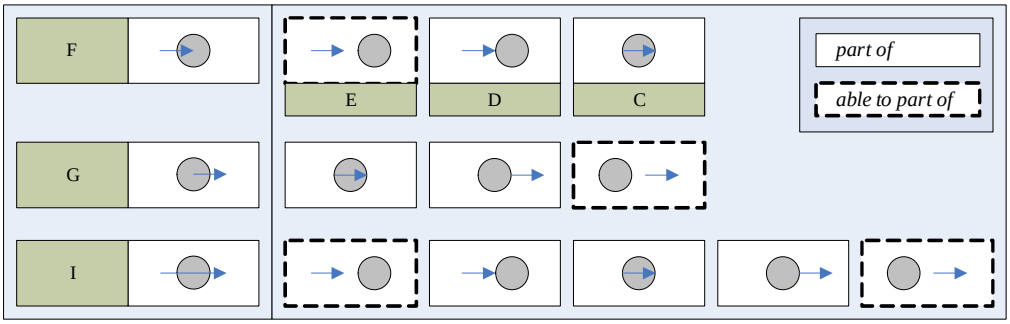
[그림 2-5] 좌표에 따른 위상관계

[그림 2-5]에서 좌표는 음의 좌표(Minus Coordinates)와 양의 좌표(Plus Coordinates)로 분류하여 움직이는 객체를 표현하고 있으며, 각각의 움직이는 객체 사이의 거리(T)가 일정하다고 가정했을 때 , 이를 통해서 좌표의 변화와 거리를 수치적으로 표현할 수 있다. 예를 들면, 고정된 객체 NM(Object NM)과 객체 C(Object C)처럼 같은 위상관계에서 객체 C가 객체 A로 움직였다고 가정했을 때, 좌표 값은 증가한다. 이와 같이 고정된 객체와 객체 A부터 객체 E까지의 위상관계를 표현하면 [그림 2-6]과 같다. 즉, [그림 2-5]에서 객체 E와 객체 NM의 관계가 [그림 2-6]에서 E로 표현되고 있으며, 객체 A와 객체 NM의 관계는 A로 표현할 수 있다.



[그림 2-6] 좌표에 따른 모션의 기본 요소

위의 5가지의 위상관계로는 모든 위상관계를 표현할 수 없으나 기본요소에 대한 위상관계들의 결합을 통하여 새로운 관계를 생성할 수 있고, 좀 더 다양한 관계를 표현할 수 있다. 다음 [그림 2-7]에서는 기본요소의 결합을 통한 확장된 요소의 위상관계 표현을 보여주고 있다.



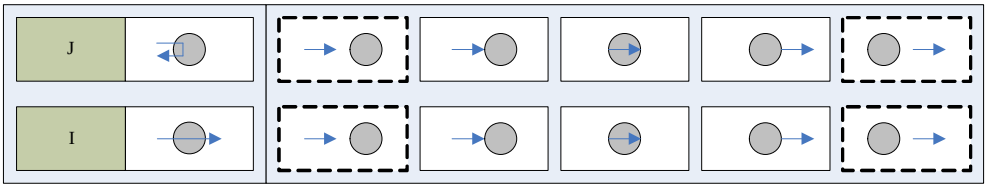
[그림 2-7] 확장된 요소에 따른 위상관계 표현

위의 그림에서 F는 기본위상관계에 따라 E와 D 그리고 C로 나누어 볼 수 있다.

여기서 E는 움직이는 객체의 위상에 따라 충분조건이 된다. 다시 말해, E의 존재여부와 관계없이 F를 정의할 수 있고, 움직임 객체의 상태에 따라 구성될 수도 있는 조건이 되는 것이다. 같은 방법으로 G와 I도 표현이 가능하다.

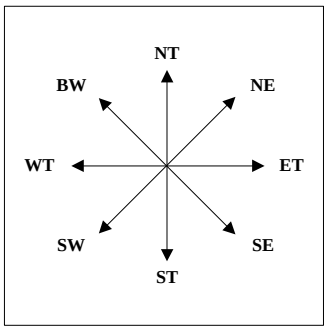
3. 방향관계

앞 절에서는 객체의 위상관계에 대해 살펴보았다. 그러나 이 위상관계들만을 고려한다면 [그림 2-8]과 같이 서로 같은 위상을 가지나 방향에 따라 서로 다른 의미를 가지는 움직임들을 구별하지 못하는 문제점이 있으므로, 객체간의 방향관계 또한 고려되어야 한다.

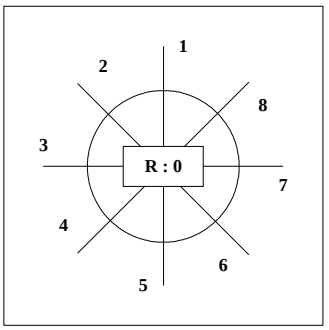


[그림 2-8] I와 J에 대한 위상관계 표현

기존의 비디오 데이터에서의 방향 표현에 대한 연구 중, John Z[16]는 [그림 2-9]와 같이 8가지의 방향 - *North(NT), Northwest(NW), Northeast(NE), West (WT), Southwest(SW), East(ET), Southeast(SE) and Southwest(SW)* - 에 따른 움직임 객체의 궤적을 표현하였고. Chang과 Jiang[15]은 [그림 2-10]과 같이 0부터 9까지의 9가지의 방향을 표현하고 객체들 간의 방향에 대한 공간적 관계를 9DLT scheme를 통해 제안하였다.



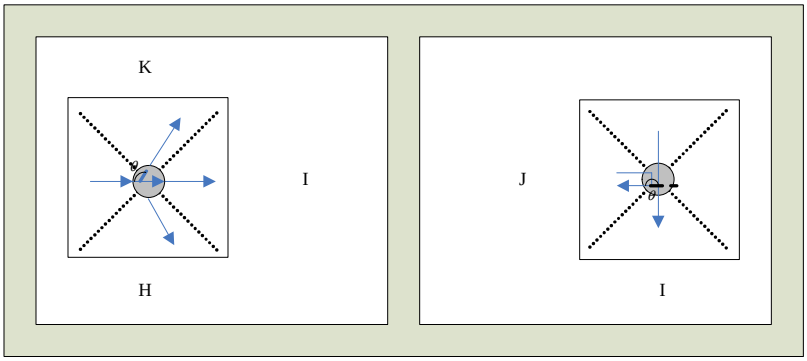
[그림 2-9] John Z의 방향관계



[그림 2-10] 9DLT scheme

이들은 기본적으로 8방향 관계에 따라 방향관계를 표현한 것으로, 본 논문에서는 이를 이용하여 객체의 방향을 표현하였다. 이에 따른 규칙과 유사도 측정방법은 4장에서 언급한다.

[그림 2-11]은 4가지 방향에 따른 움직임 변화 즉, H, I, J, K에 대한 변화를 나타내고 있다.



[그림 2-11] 방향성에 따른 모션

[그림 2-12]에서 θ이 북서(Northwest)방향과 남서(Southwest)방향으로 변화하였을 때, I가 아닌 K와 H로 표현됨을 보여주고 있다. 또한 I가 θ에 따라 J로 표현됨을 보여준다. 이를 통해, 모두 같은 위상관계를 가지고 있지만 방향에 따라 이들이

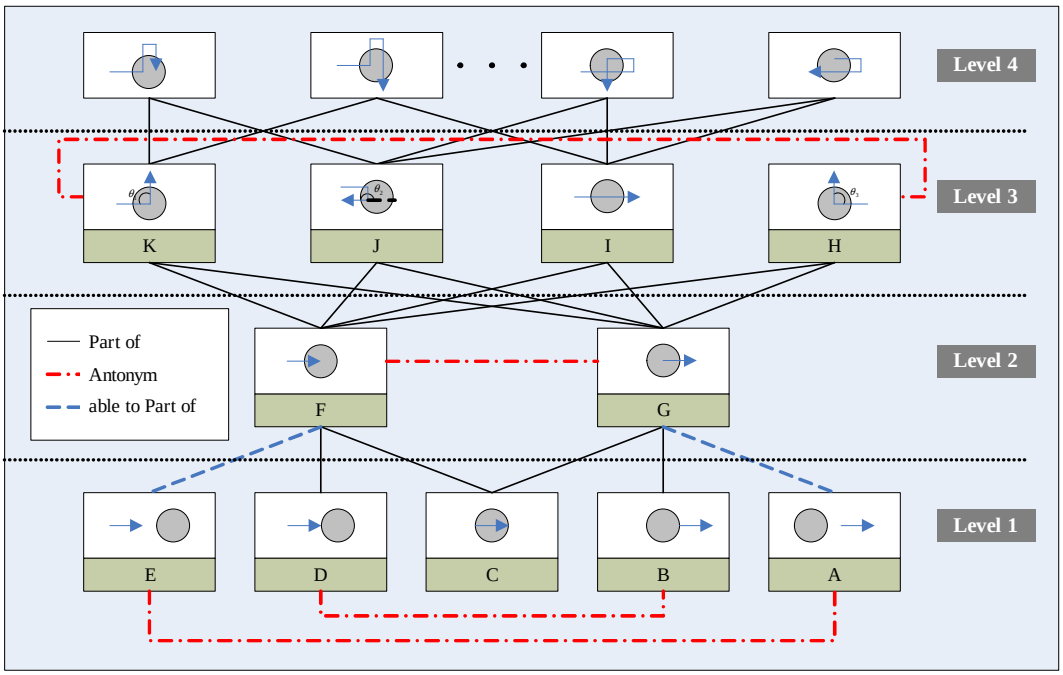
모두 다른 의미를 표현하고 있음을 보여주고 있다. [표 2-2]는 [그림 2-11]에서 θ값에 따른 [그림 2-12]에 대한 분류이다. 여기서 θ는 +180도 ~ -180도 사이의 값을 라디안(Radian) 값으로 표현하였다.

θ(Radian)	Motion verb
-2.356 ~ -0.785	K
-0.786 ~ 0.785	I
0.786 ~ 2.355	H
-3.141 ~ -2.355 2.356 ~ 3.142	J

[표 2-2] [그림 2-11]에서 방향성에 따른 위상관계의 분류

B. 객체 움직임에 대한 계층 관계 모델링

[그림 2-12]는 앞에서 정의한 새로운 위상관계와 방향관계를 바탕으로 의미적 표현을 위한 시공간 관계 모델링을 나타내고 있다. 다시 말하면, 비디오의 저차원적 특징을 모션 동사와 결합시키기 위한 의미적 레벨(Semantic Level)을 표현하고 있다. [그림 2-12]에서 우리는 크게 3가지 단계로 구분하였다. 첫 번째 단계(Level 1)는 모션의 기본 요소의 표현이고, 두 번째 단계(Level 2)는 모션의 확장된 요소의 표현이다. 여기서 고정된 객체와 움직이는 객체사이의 위상관계만을 고려하므로 2가지 외에 다른 경우는 성립이 되지 않는다. 마지막으로 세 번째 단계(Level 3)는 두 번째 단계의 결합으로 표현되는 모션의 표현이다. 이보다 상위 단계(Level 4)에서는 세 번째 단계의 결합으로 표현이 가능하나 의미부여에 어려움이 있으므로 본 논문에서는 세 번째 단계까지만 다루기로 한다.



[그림 2-12] 객체 움직임에 대한 계층 관계 모델링

[그림 2-12]에서 각 관계에 대해 살펴보자면 F와 G는 I의 하위 클래스가 된다. 그리고 이는 'Part of' 관계를 가지고 있다. 또한 E는 F를 구성하는데 꼭 필요한 요소는 아니지만, 부분 요소가 될 수 있는 자질을 가지고 있다. 따라서 이는 'able to part of'로 표현하였다. 또한 F와 G는 서로 반대의 의미를 가지고 있기 때문에 반의어(antonym) 관계로 나타낼 수 있다. 이와 같이, 비록 각 모션 동사는 공간적 관계는 다르다고 할지라도, 의미적 관계는 같기 때문에 계층적 의미 관계로 표현이 가능하다.

Ⅲ. 움직임 객체에 대한 모션 동사의 적용

2장에서는 시공간상에서 시간의 흐름에 따라 움직이는 객체에 대한 위상관계와 방향관계를 정의하였다. 3장에서는 제안한 모델을 모션동사(Motion Verb)를 이용하여 매칭하였으며, 링귀스틱 온톨로지의 일종인 WordNet을 이용하여 모션동사를 확장하였다. 또한, 이를 이용하여 제안한 모델의 의미적 표현에 대해 평가하였다.

A. 움직임 객체에 대한 모션 동사의 정의

2장에서 설명한 위상 관계에서 움직임 동사의 의미 및 개념과 움직임 객체들 사이의 시공간적 관계 표현의 매칭을 위해 인지적인 측면이 필요했다. 이를 위해 본 논문에서는 Levin[10]이 제안한 동사의 분류 중에서 방향성 모션 동사들 - verbs of Inherently directed motion - 로 분류된 동사들을 이용하여 제안한 모델과 매칭하였다. Levin의 동사 분류는 정확하게 각 클래스에 대한 구문을 기술하고 있다. 뿐만 아니라 각 클래스에 할당된 의미적 성분의 요점만을 표현하고 있다. 클래스는 몇몇 의미를 포함하고 있는 구문적 프레임쌍이 발생할 동사의 가능성에 기반하고 있다. 즉, 특이한 상황에서 발생하는 동사의 가능성에 기반을 두어 분류하고 있다

[표 3-1]은 Levin이 분류한 방향성이 있는 모션 동사들을 나타내고 있다.

1) verbs of Inherently directed motion *Class members*: advance, arrive, ascend, climb, come, cross, depart, descend, enter, escape, exit, fall, flee, go. leave, plunge, recede, return, rise, tumble

[표 3-1] Levin의 동사의 분류 중 'verbs of Inherently directed motion'

이 클래스 내 동사들은 정확한 방향은 아니더라도 움직임의 방향 표현을 포함한 의미를 가지고 있다. 그리고 이 동사들 중 움직임의 방법을 표현한 동사는 없으나 이 클래스 내 동사들은 모든 면에서 동일한 움직임은 아니다. 또한 이들은 움직임

의 동사에 따라 전치사구가 경로를 나타내거나 방향을 나타내기도 하고, 혹은 둘 다 표현하기도 한다. 즉, 시작, 끝, 경로를 어떻게 표현할 수 있느냐에 따라 다르다.

이러한 Levin이 분류한 방향성이 있는 모션 동사를 이용하여 Lisa Pearl[17]은 [표 3-2]와 같이 "Go +전치사" 형태로 변환하였으며, Tonia Bleam[18]은 문장에서 Levin이 분류한 방향성이 있는 모션 동사를 이용하여 의미적인 분석을 시도하였다.

Levin 51.1.a: Verbs of inherently directed motion	**GO primitive and a particle**
ascend, climb, rise	go up
come	go to
cross	go over
depart, escape, flee, leave	go away
descend, fall, plunge	go down
go	go
recede, return	go back
tumble, walk, run	go manner of motion

[표 3-2] Lisa Pearl의 'GO primitive and a particle'

이와 같이 Levin의 방향성이 있는 모션 동사의 분류를 통하여 많은 연구가 진행되고 있으며, WordNet과 더불어 어휘 분류에 관한 대표적인 연구라고 할 수 있다.

본 논문에서도 이를 이용하여 위상관계와 가장 비슷한 동사를 사용하였으나 움직임 객체의 고저에 대한 의미가 포함된 동사들(clime, rise 등)은 제외하였다. 또한 [그림 2-12]에서 C는 움직임 객체와 고정된 객체사이의 크기에 따라 기존의 비슷한 연구들에서와 같이 '*same*', '*include*', '*is_include_by*'로 구분하였고, 이들의 대표단어로 '*include*'로 설정하였다. [표 3-3]은 [그림 2-12]의 각 모델에 대한 모션 동사의 매칭을 보여주고 있다.

제안한 모델	제안한 모델에 대한 모션 동사의 정의
A	depart
B	leave
C	include
D	arrive
E	come_up
F	enter
G	exit
H	turn right
I	go through
J	return
K	turn left

[표 3-3] [그림 2-12]에 대한 모션 동사의 매칭

B. 움직임 객체 모델링의 검증

본 절에서는 WordNet을 이용하여 제안한 위상관계의 의미적 표현에 대해 평가하였다. WordNet은 인간의 어휘지식을 모방하여 다의성과 동의관계를 이용하였으며, 의미를 최대한 정확히 표현하고 있을 뿐만 아니라 개념간의 관계 표현 등을 통해 개념을 계층적으로 표현하고 있다.

1. 제안된 모델링에서 유사도 측정

[그림 2-12]에서 첫 번째 단계에 있는 5가지의 위상관계는 가장 기본이 되는 위상관계이며, 이는 수평적인 관계로 구성되어 있으며, 식(3. 1)을 통하여 거리를 측정하였다.

$$dis_{AB} = \sum_{k=1}^{n} |T_n| \qquad \text{식 (3. 1)}$$

[표 3-4]는 각각의 움직이는 객체 사이의 값(T)이 모두 1이라고 가정했을 때, 식(3. 1)을 통한 수치 값을 나타내고 있다.

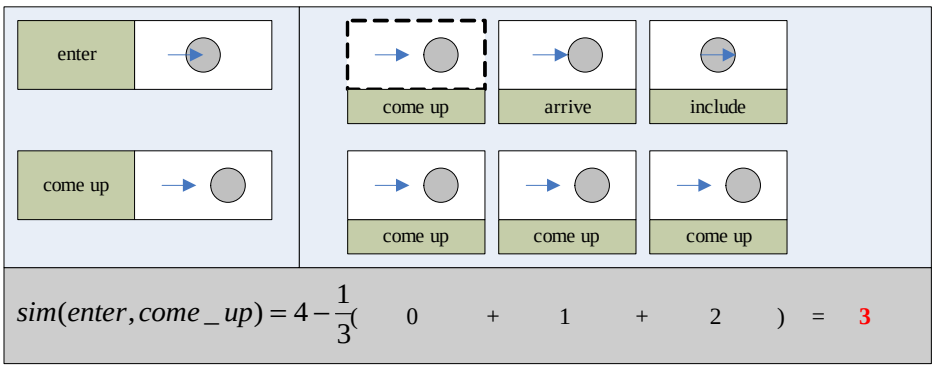
	come_up	arrive	include	leave	depart
come_up	0	1	2	3	4
arrive	1	0	1	2	3
include	2	1	0	1	2
leave	3	2	1	0	1
depart	4	3	2	1	0

[표 3-4] [그림 2-12]에서 첫 번째 단계에 대한 거리

이와 같이 기본 요소를 이용하여 [그림 2-7]과 같은 확장된 요소의 거리 즉 [그림 2-12]의 두 번째 단계는 다음과 같이 측정할 수 있다. [표 3-5]에서 기본요소들 사이의 최대 거리($s_{\max}$)는 4이고, 측정된 유사도는 기본 요소와 같이 4를 넘지 않는다. $m1$과 $m2$는 비교대상인 확장된 요소들을 말한다. 또한 n은 위상관계를 기본요소로 변환할 때의 개수를 말한다. 즉, $m1$의 수와 $m2$의 수에 최소공배수(Least Common Multiple)가 n이 된다. 여기서 같은 위상관계의 유사도 값은 '0'이 된다. 이를 수식으로 표현하면 식(3. 2)와 같다.

$$Sim(m1,m2) = S_{\max} - \frac{1}{n}\sum_{k=1}^{n} distance[m1_k, m2_k] \qquad \text{식 (3. 2)}$$

[그림 3-1]은 식(3.2)를 이용하여, '*enter*'와 '*come_up*'에 대한 거리 유사도 측정 예이다.



[그림 3-1] '*enter*'와 '*come_up*'간의 거리 유사도

위의 그림에서 *m*1의 수가 3이고, *m*2의 수가 1이므로, *n*은 3이 되고, '*enter*'와 '*come_up*'간의 거리 유사도는 3이다. 위와 같은 방법을 이용하여 객체의 유연한 확장이 이루어질 수 있고, 이 객체들 사이의 거리 유사도 또한 측정이 가능하다. 이와 같은 방법으로 [그림 2-12]의 1, 2 단계에 대한 거리 유사도를 측정하면 다음 [표 3-5]와 같다. 여기서 세 번째 단계(Level 3)에 또한 같은 방법으로 측정이 가능하다.

	come_up	arrive	include	leave	depart	enter	exit
come_up	4	3	2	1	0	3	1
arrive	3	4	3	2	1	3.33	2
include	2	3	4	3	2	3	3
leave	1	2	3	4	3	2	3.33
depart	0	1	2	3	4	1	3
enter	3	3.33	3	2	1	4	2
exit	1	2	3	3.33	3	2	4

[표 3-5] [그림 2-12]에서 1, 2 단계에 대한 유사도

2. 비교평가

제안한 위상관계의 비교 평가를 위해 WordNet을 이용하였다. [표 3-6]은 WordNet 내에서 각 단어들의 거리 즉, 엣지(edge)기반으로 유사도를 측정한 결과이다.

	come_up	arrive	include	leave	depart	enter	exit
come_up	1	0.25	0.25	0.25	0.2	0.25	0.2
arrive	0.25	1	0.33	0.33	0.25	0.33	0.25
include	0.25	0.33	1	0.33	0.25	0.33	0.25
leave	0.25	0.33	0.33	1	1	0.33	1
depart	0.2	0.25	0.25	1	1	0.25	0.5
enter	0.25	0.33	0.33	0.33	0.25	1	0.25
exit	0.2	0.25	0.25	1	0.5	0.25	1

[표 3-6] WordNet상의 엣지(edge)기반 유사성 측정

두 위상 관계에 대한 평가 척도로 본 논문에서는 상관계수를 이용한다. 상관계수란 두 변량 사이의 상관관계의 정도를 나타내는 수치(계수)로 0부터 1까지의 값을 가진다. 즉, 1에 가까우면 가까울수록 상관관계의 정도가 크다는 것을 의미한다. [표 3-5]의 제안한 모델과 [표 3-6]의 WordNet과의 상관계수는 0.663이라는 값을 얻었다. 여기서 상관계수가 높지 못한 이유는 WordNet이 인지적 측면을 많이 고려했다고는 하나 완전하지 않고 또한 본 논문에서 의미표현을 위해 사용한 동사와 WordNet내의 동사가 정확하게 매치되지 않는다는 점을 들 수 있다.

C. 움직임 객체에 대한 모션 동사의 확장

본 절에서는 정의한 모션 동사의 제한성을 극복하기 위해 WordNet상에서 유사도 측정을 통해 어휘를 확장하고자 한다. 먼저 [표 3-3]에 제안된 어휘에 대하여 동의어와 반의어를 이용하여 모션 동사를 확장하였다. 이를 통하여 객체 움직임에 대해 다양한 모션동사를 이용한 질의 처리가 가능하였다.

1. WordNet

링귀스틱 온톨로지(Linguistic Ontology)의 일종인 WordNet[11]은 인간의 어휘지식에 대한 심리언어학 연구의 성과를 토대로 1985년도 중반부터 프린스톤 대학 인지 과학연구실이 구축해온 영어어휘 데이터베이스(Database)이다. WordNet은 인간의 어휘지식을 모방하여 다의성과 동의 관계를 이용하였으며, 의미를 최대한 정확히 표현하고 있을 뿐만 아니라 개념간의 관계 표현 등을 통해 개념을 계층적으로 표현하고 있다. 또한 각 단어들은 synset이라는 동의어의 집합을 기준으로 노드들 사이의 관계를 이용하여 정의하고 있다. 즉 같은 단어라 하더라도 의미에 따라 synset이 다르며, 다른 synset이라고 하더라도 같은 단어로 표현할 수 있다.

지금까지의 가장 최신 버전인 WordNet 2.1의 경우, 동사에 대해 13,679개의 synset으로 구성되어 있고, 그 중에서 38번 그룹은 모션동사(motion verb)에 대한 분류를 나타내고 있으며 총 1,397개의 synset으로 구성되어 있다.

[표 3-7]은 WordNet의 "index.verb"의 내용으로 '*abandon*'은 의미에 따라 총 다

섯 개의 synset(02207318 00605740 02207028 02058122 00606404)으로 의미가 나뉘어 짐을 볼 수 있다.

```
aah v 1 1 @ 1 0 00856756
abacinate v 1 1 @ 1 0 02148661
abandon v 5 4 @ ~ $ + 5 5 02207318 00605740 02207028 02058122 00606404
abase v 1 3 @ ~ + 1 0 01782644
abash v 1 3 @ ~ + 1 0 01774948
abate v 2 2 @ + 2 0 00242091 00241861
abbreviate v 2 4 @ ~ $ + 2 1 00240551 00240702
abdicate v 1 2 @ + 1 0 02356275
abduce v 1 1 @ 1 0 01005783
abduct v 2 5 ! @ ~ + ; 2 0 01458315 01436738
aberrate v 2 3 @ + ; 2 0 02636750 02636443
abet v 1 2 @ + 1 1 02524939
abhor v 1 2 @ + 1 1 01757422
abide v 2 5 @ ~ $ + ; 2 2 02611851 00660036
                . . .
```

[표 3-7] WordNet의 "index.verb"

2. *WordNet*상의 유사도 측정을 통한 모션 동사의 확장

a. *WordNet*상에서 유사도 측정방법

WordNet을 이용한 유사성 측정 방법은 크게 정보량을 이용한 노드기반 방법[25]과 최단 거리를 이용한 에지기반 방법[26,27] 그리고 gloss기반 방법[28]으로 나눌 수 있는데, 각각에 대한 자세한 내용은 다음과 같다.

노드기반 유사성 측정은 노드의 확률을 기반으로 한 정보량(Entropy) 측정에 의한 접근방식이다. 먼저 정보량의 측정을 위한 노드의 확률을 구하면 식(3. 3)과 같다[25].

$$P(c) = \frac{freq(c)}{N} \qquad \text{식 (3. 3)}$$

여기서 N은 개념의 총 수를 의미하고 $P(c)$는 개념 c와 마주칠 확률로 계층적 구조의 경우 $freq(c)$는 개념 c에 포함된 모든 하위 개념들의 합을 의미한다. 만약 c_1 IS_A c_2 관계라면 $p(c_1) \le p(c_2)$가 된다. 그러므로 WordNet에서 유일한 Top Node ϕ의 확률은 1이 된다.

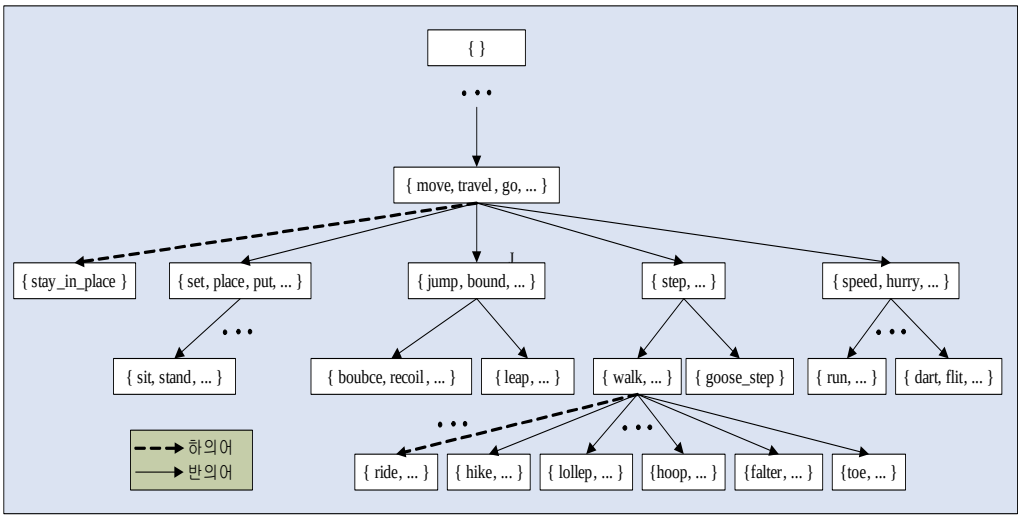
개념 c의 정보량은 $-\log P(c)$로 확률이 증가하면 정보량은 감소하므로 더 추상적인 상위 개념은 낮은 정보량을 가진다. 즉, ϕ의 정보량은 0이 된다.

$$H(c) = -\log P(c) \qquad \text{식 (3. 4)}$$

측정하고자 하는 두 개념이 공유한 정보는 WordNet에서 두 개념들을 포함하고 있는 개념의 정보량으로 표현할 수 있다. 따라서 두 개념간의 노드 기반 유사성 측정은 다음과 같이 나타낼 수 있다.

$$\text{Sim}\ (c_1, c_2) = \underset{c \in S(c_1,\ c_2)}{MAX} [-\log P(c)] \qquad \text{식 (3. 5)}$$

여기서 $S(c_1, c_2)$는 c_1과 c_2 두 개념을 포함하고 있는 상위 개념 c를 의미한다. 두 개념이 공유하고 있는 상위 개념이 많을수록 두 개념은 더 유사하다고 할 수 있다. 따라서 개념 간 유사성은 c_1과 c_2를 포함하는 상위 개념 중 가장 큰 H를 가지는 개념의 정보량이다. 예를 들어 [그림 3-2]에서 *'walk'*와 *'move'* 그리고 *'walk'*와 *'hike'*간의 유사성을 측정하면 *'walk'*와 'hike'가 공유한 개념은 *'walk'*와 *'move'*가 공유한 개념보다 더 많으므로 더 유사하다고 할 수 있다.



[그림 3-2] '*walk*'에 대한 WordNet의 구조

그러므로 *sim(walk, move)*는 $H(move)$이고, *sim(walk, hike)*은 $H(walk)$이다.

다음은 최단거리를 이용한 에지기반의 방법으로 간단하게 식(3. 6)과 같이 표현할 수 있다.

$$\mathrm{Sim}\ (c_1,c_2) = 2d_{\max} - \mathit{min}[\ len(c1,c2)\ ] \qquad \text{식 (3. 6)}$$

여기서 $d_{\max}$는 WordNet의 최대 깊이 값을 의미하며 $len(c_1,c_2)$은 최단거리 측정을 위한 함수이다. 에지 기반 측정 방법에는 이외에도 Wu Z. 방법과 Leacock C. 방법 등이 있다[26,27].

마지막으로 gloss기반의 측정방법은 유사성을 측정하고자 하는 두 단어의 gloss를 이용하는 것이다[28]. 이는 WSD(Word Sense Disambiguation)에서 많이 쓰이던 방법으로 gloss에서 공통된 단어를 찾는 것이다. 이때 사용하는 gloss는 측정하고자 하는 단어의 gloss 뿐만 아니라 하위어(hyponym), 상위어(hypernym) 등의 관계쌍(related_pairs)을 이용한다. 여기서 공통된 단어가 많으면 많을수록 두 단어 사이의 유사성은 커진다.

Related-pairs={(gloss, gloss), **(hype, hype),(hypo, hypo),(hype, gloss),(gloss, hype)**}

[표 3-8] gloss기반 측정방법에서 관계 쌍

이를 식으로 표현하면 다음과 같다.

$$sim(A,B) = \sum_{\alpha \in related-pairs, \beta \in related-pairs} score(\alpha(A) + \beta(B)) \qquad \text{식 (3. 7)}$$

위의 3가지 유사성 측정방법은 각기 장단점을 가지고 있으며, 활용에 따라 다른 성능을 보이고 있다.

b. 움직임 객체에 대한 모션 동사의 확장

본 논문에서는 WordNet에서 움직임 객체 표현을 위해 동사만을 사용하므로 동사의 특징에 대해 살펴보도록 한다. WordNet에서 동사는 계층 구조의 깊이가 4를

넘는 경우가 거의 없으며, 하나의 층에 여러 단어가 분포되어 있음을 볼 수 있다. 예를 들어 '*march*'와 '*strut*'가 포함된 계층을 L0층이라 하고, 위층을 L+1, 아래층을 L-1이라고 가정할 때, L+1에는 '*walk*'가 있고, L0에는 '*march*', '*strut*', '*traipse*', '*amble*'가 있으며, L+2에는 '*move*'가 있다. 이와 같이 계층의 차가 적은 동사의 경우는 유사도 측정 시 에지 기반의 측정방법이 효율적임을 알 수 있다.

에지 기반 측정방법을 이용한 '*exit*'와 '*go_out*'간 유사도 측정 결과는 [표 3-9]과 같다. 여기서 '*exit*'와 '*go_out*'가 각각 1번째 의미(#1)일 때 두 개념간 최단거리 1이 되므로 유사도 측정 결과는 1이 된다.

```
go_out#v#1 exit#v#1:
HyperTree: *Root*#v#1 move#v#3 exit#v#1
HyperTree: *Root*#v#1 move#v#3 exit#v#1
LCS: exit#v#1   Path length: 1.

                        . . .

go_out#v#1 exit#v#3:
HyperTree: *Root*#v#1 move#v#3 exit#v#1
HyperTree: *Root*#v#1 change#v#1 change_state#v#1 die#v#1
LCS: *Root*#v#1   Path length: 6.

                        . . .
go_out#v#1   exit#v#1   1
```

[표 3-9] 엣지(edge)기반의 유사도 측정과 모션동사의 확장

[표 3-10]에서는 [표 3-3]에 대한 어휘들과 그 어휘들에 대한 동의어(Synonym)와 반의어(Antonym)를 나타내고 있다.

제안한 모델링에서의 모션 동사		유의어(synonym)		반의어(antonym)	
Word	**Synset**	**Word**	**Synset**	**Word**	**Synset**
depart	01831361	go, go_away	01831361	come come_up	01831864
come_up	01831864	come	01831864	depart go go_away	01831361
arrive	01987643	come, get	01987643	leave go_forth go_away	01991207
leave	01991207	go_forth go_away	01991207	arrive come get	01987643
enter	01998301	come_in get_into get_in go_in move_into go_into	01998301	exit go_out get_out leave	01997376
exit	01997376	go_out get_out leave	01997376	enter come_in get_into get_in go_in move_into go_into	01998301
return	01986635	go_back get_back come_back	01986635		
go_through	02031827	pass go_across	02031827		

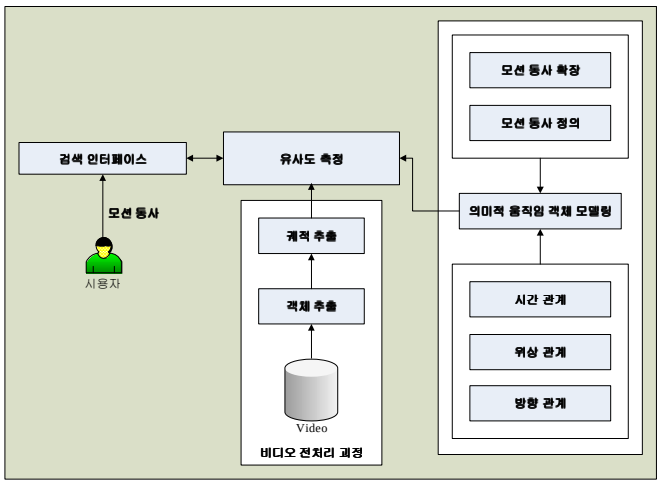
[표 3-10] 제안된 모델링에서 모션 동사의 유의어와 반의어

Ⅳ. 의미기반 궤적 검색 시스템

2장에서는 객체 움직임을 위한 시공간 관계 모델링을 제안하고, 3장에서는 이를 모션동사와 매칭하고 확장하였다. 본 장에서는 2, 3장에서 제안한 모델을 이용한 의미기반 궤적 검색에 대해 다루어 보고자 한다. 이를 위해 비디오에서 움직임 객체와 궤적(Trajectory)을 추출하고 객체의 움직임 비교를 위해 추출된 궤적 간 유사도를 측정하였다. 또한 움직임 객체의 의미적 인식을 위한 규칙을 설정하고, 이 규칙과 TSR을 이용하여 의미 기반 검색이 되도록 하였다.

A. 시스템 구성도

본 논문에서 제안하는 비디오 내 움직임 객체의 의미기반 궤적 검색 시스템의 전체 시스템 구성도는 [그림 4-1]과 같다. 이는 크게 3부분으로 나눌 수 있는데, 첫 번째는 2장과 3장에 기술한 의미적 움직임 관계 모델링이고, 두 번째는 비디오 전처리로 움직임 객체를 추출하고 이 객체의 궤적을 구하는 부분이다. 마지막은 생성된 규칙과 TSR을 이용한 유사도 측정으로 의미 기반 검색 부분이다.



[그림 4-1] 비디오 내 움직임 객체의 의미기반 궤적 검색 시스템의 구성도

B. 비디오 전처리과정을 통한 움직임 객체의 궤적 추출

1. 비디오 내 움직임 객체 추출

배경의 움직임이 없는 고정카메라 환경에서 객체 추출(Object Detection)의 전처리 과정으로 세그멘테이션(Object segmentation)을 하는데 이는 부분 영역 정보를 바탕으로 간소화하는 작업으로 블록기반과 영역기반으로 나눌 수 있다.

블록 기반 방법은 비디오의 특정 프레임을 작은 블록으로 나누어 처리하는 방법으로 이와 관련된 연구들을 살펴보면 다음과 같다. Wang[19]은 피부색 통계, 형태 제약, 얼굴의 광도 DCT((Discrete Cosine Transform) 계수의 분포에 따라 자동적으로 얼굴을 찾아내는 알고리즘을 제안하였고, De Queiroz[20]는 DCT 계수에서 halftone(사진의 명암에 따른 선명도 차이), Continuous-tone(농도가 연속적으로 변화하는 사진의 상), text 등이 포함되어 있는 특정 영역에서 JPEG 이미지를 세그멘테이션하는 알고리즘을 제안하였다. 또한 MPEG 비디오에서 객체를 추출하는 알고리즘을 사용하여 클러스터링(Clustering)과 세그멘테이션 방법의 제안하였고, Ji와 Park[21]은 참/거짓 모션 블록 분류(True/False Motion Block Classification)와 DCT 계수의 유사도를 기반으로 하여 동적 영역의 세그멘테이션 방법을 제안하였다. 마지막으로 Fatih Porikli[22]는 장면을 시공간적인 데이터 구조와 컬러정보를 구조적으로 표현하고 각 픽셀들에 대한 특징 벡터(Feature Vector)를 추출한다. 여기서 추출된 특징 벡터들을 기반으로 3차원 볼륨 확장(Volume growing)을 한 다음 각 볼륨들을 계층적으로 병합하여 여러 가지 수준의 객체를 추출한다. 병합과정에서 볼륨의 둘레, 궤적, 움직임 정보 등을 이용한다. 피부색을 포함하는지 움직임 정도는 어느 정도인지 등 기존의 방법들에 비해 사용자의 개입이 적어지고 장면 내 추적의 과정에서 발생하는 에러들을 줄여 보다 정확하고 세밀한 객체 분할의 결과를 보여준다. 이러한 블록 기반 방법들은 MPEG을 주로 사용하고 단순하여 원격진료 등에서 사용되나 사용자의 요구를 충분히 만족시키지 못하는 단점을 가지고 있다.

이에 반해 영역 기반 방법은 K-means, LBG, LVQ(Learning Vector Quantization)등의 벡터 양자화(Vector Quantization) 방법을 사용하며, 계속적으로

학습을 통하여 영역의 중심을 찾아내고 중심과의 거리를 이용하여 영역을 찾아내는 방법이다. Hariadi MoCHAMAD[23]는 픽셀(Pixel)의 좌표와 YUV 컬러 성분으로 5차원 특징 백터를 구성하고 이를 LVQ를 이용하여 학습시키는 방법을 사용하였다. 이처럼 비디오 전처리에서 가장 중요한 부분인 비디오에서의 특정 객체의 세그멘테이션 그리고 움직임 예측(Motion Estimation)이라고 할 수 있다. [그림 4-2](b)는 [그림 4-2](a)의 색상 값을 바탕으로 세그멘테이션 결과를 나타낸 것이다.

(a) Original Image　　(b) 컬러 세그멘테이션 결과

[그림 4-2] 이미지에 대한 컬러 세그멘테이션(Color Segmantation) 결과

이러한 세그멘테이션 과정을 거쳐 객체의 실제 이동 궤적을 추출하기 위해서는 객체 분리를 통해 객체의 위치를 추출하여야 한다. 이때 사용하는 알고리즘이 반자동 영역[2]으로 초기 영상에서 사용자가 분리하고자 하는 객체 영역의 경계를 그리면 해당 영역의 객체가 배경 영상으로부터 분리되고, 분리된 객체 영역 정보를 이용하여 연속된 프레임 상에서 자동으로 객체 분리를 행한다.

본 논문에서는 배경과 움직이는 객체를 분리하여 이들의 컬러 값의 차이를 이용하는 *Andrew Kirillov*[24]의 객체 추출(Object Detection) 알고리즘을 이용하였다. [그림 4-3]은 객체 추출의 결과를 나타내고 있다.

[그림 4-3] 객체 추출(Object Detection) 결과

2. 비디오 내 움직임 객체의 궤적 추출

객체 추출 과정을 통하여 움직이는 객체의 궤적을 추출할 수 있으며, 이에 대한 알고리즘은 [표 4-1]과 같다.

```
1   g.DrawRectangle(pen, rc);
2
3       rem_ps.X = rc.Left;         // 객체의 왼쪽 좌표
4       rem_ps.Y = rc.Bottom;     //객체의 하단 좌표
5       Pen pen_center = new Pen(Color.Red, 1);         객체번호표시를 위한 사각형
6
7       g.DrawImage(numbersBitmaps[n], rc.Left, rc.Top, 7, 9);     //객체 번호표시
8
9       if(rem_psnum == n)
10      {
11       g.DrawLine(pen_center,       rem_cps.X,       rem_cps.Y,       (rc.Right+rem_ps.X)/2
12      ,(rc.Top + rem_ps.Y)/2); //객체가 움직일 때, 움직이는 객체의 중앙값에 라인을 표시
13      }
14
15      rem_cps.X=(rc.Right + rem_ps.X)/2 ;  // 움직이는 객체의 X축 중간 값
16      rem_cps.Y=(rc.Top    + rem_ps.Y)/2;  // 움직이는 객체의 Y축 중간 값

                            . . .
```

[표 4-1] 궤적추출 알고리즘

[표 4-1]에서는 실시간으로 움직이는 객체를 MBR(Minimum boundary rectangle)로 추출하고 이에 대한 중앙값의 연결을 통해 시간에 따라 움직이는 객체의 궤적 좌표를 얻을 수 있었다. '*Sample 5*'에 대한 궤적 좌표는 [표 4-2]와 같다.

객체 번호	좌표(X,Y)
0	(8,72) … (16.72) … (32,72) … (80,72) (80,72) (88,80) (88,80) … (88,88) …

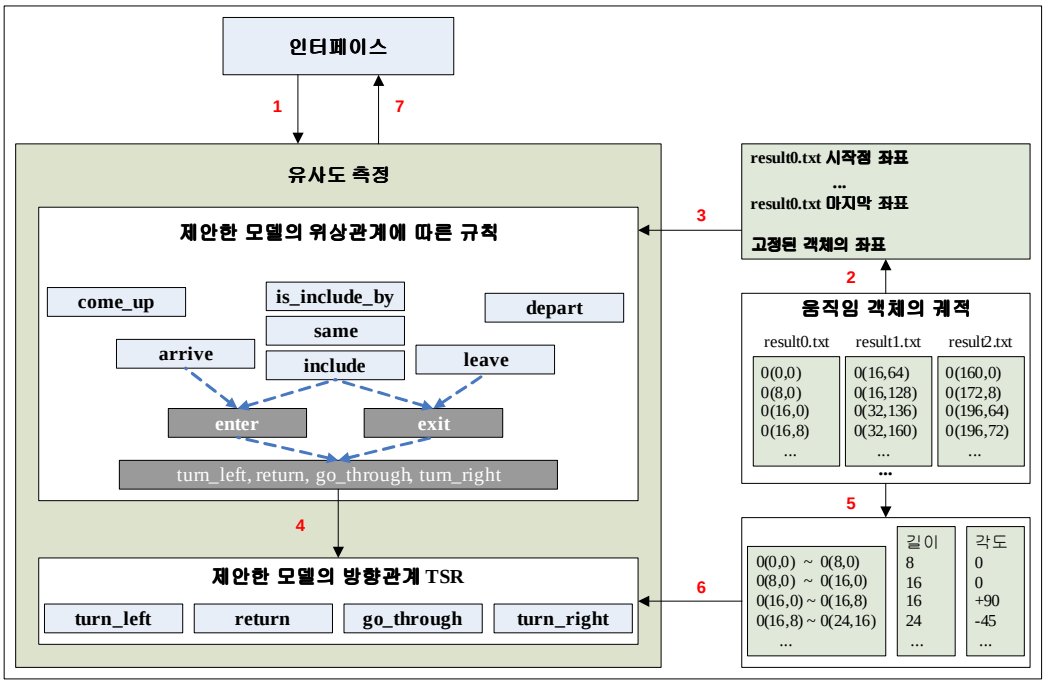
[표 4-2] 'Sample 5'에 대한 궤적 정보

C. 움직임 객체의 궤적과 제안한 모델링간의 유사도 측정

본 절에서는 앞 절에서 서술한 방법에 의해 추출된 비디오 내 움직임 객체의 궤적과 제안한 모델과 유사도 측정 방법에 대해 기술하겠다. 유사도 측정 순서를 살펴보면 다음과 같다.

1. 먼저 인터페이스에서 질의어를 받는다.
2. 비디오 내 움직임 객체에 대한 궤적좌표를 파일로부터 읽는다.
3. 움직임 객체에 대한 궤적좌표와 고정된 객체의 좌표를 제안된 모델의 위상관계에 따른 규칙과 비교하여 분류한다.
4. 만약 turn_left, return, go_through, turn_right의 경우는 방향관계를 고려해야한다.
5. 움직임 객체의 궤적 좌표를 바탕으로 각 좌표간의 길이와 각도로 변환한다.
6. 이 길이와 각도를 이용하여 TSR로 표현하고 4가지 위상관계와 면적을 구한다.
7. 면적의 합이 가장 작은 값을 찾아서 그에 맞는 모션동사를 인터페이스로 전달한다.

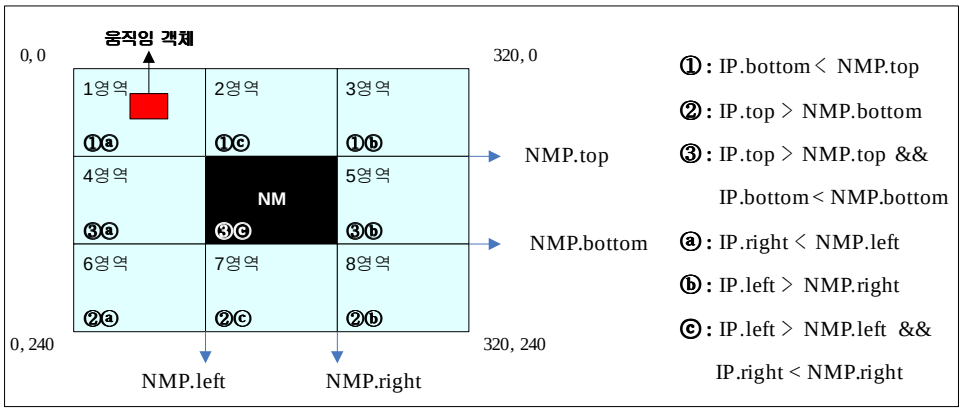
이와 같은 방법으로 유사도를 측정할 때, 의미적 움직임 객체 판별을 위한 규칙과 TSR을 이용하며 [그림 4-4]와 같다.



[그림 4-4] 비디오 내 움직임 객체의 궤적과 제안한 모델 유사도 측정 방법

1. 의미적 움직임 객체 판별을 위한 규칙

의미적 움직임 객체 판별을 위해 다음과 같은 알고리즘을 제안하고자한다. 고정된 객체의 좌표(Non-Moving object's Pixel)를 NMP라 하고 움직임 객체의 최초 궤적 좌표(Initial trajectory Pixel)를 IP, 최종 궤적 좌표(Final trajectory Pixel)를 FP라 할 때, IP와 FP의 관계는 IP, IP1, IP2, …, FP가 된다. 여기서 각 좌표를 IP로 예를 들면, IP의 왼쪽좌표는 IP.left, 오른쪽 좌표는 IP.right, 상단좌표는 IP.top, 하단좌표는 IP.bottom로 명명한다. 여기서 IP와 NMP에 따라 최초 움직임 객체의 위치는 [그림 4-4]와 같으며, x축 좌표는 1영역에서 3영역으로 y축 좌표는 1영역에서 6영역으로 증가한다.



[그림 4-5] IP와 NMP에 따른 최초 움직임 객체의 위치

[그림 4-5]에서 움직임 객체가 존재할 수 있는 영역을 총 9개(NM영역과 8영역)로 나누었다. 여기서 움직임 객체가 1영역에 존재하고 있는데, 이 같은 조건은 ①ⓐ의 조건을 만족할 때 성립한다. 이와 같은 최초 위치에 따라 [그림 2-12]의 각 모델들의 성립조건이 달라진다. 이를 자세히 살펴보면 다음과 같다.

먼저 [그림 2-12]에서 '*Include*'에 대해 살펴보면 '*Include*'는 움직이는 객체의 크기를 통해 3가지로 분류될 수 있으며 각 관계는 다음과 같다.

모션동사	조건	
	IP와 NMP	FP와 NMP
include	IP.left >= NMP.left IP.right <= NMP.right IP.top >= NMP.top IP.bottom. <= NMP.bottom	FP.left >= NMP.left FP.right <= NMP.right FP.top >= NMP.top FP.bottom. <= NMP.bottom
same	IP.left = NMP.left IP.right = NMP.right IP.top = NMP.top IP.bottom. = NMP.bottom	FP.left = NMP.left FP.right = NMP.right FP.top = NMP.top FP.bottom = NMP.bottom
is_include_by	IP.left <= NMP.left IP.right >= NMP.right IP.top <= NMP.top IP.bottom. >= NMP.bottom	FP.left <= NMP.left FP.right >= NMP.right FP.top <= NMP.top FP.bottom >= NMP.bottom

[표 4-3] '*Include*', '*same*', '*is_include_by*'에 대한 규칙

[그림 2-12]에서 '*same*'은 [그림 4-5]의 모든 NMP가 IP, FP와 같아야하며 '*is_Include_by*'는 모든 IP와 FP가 전 영역에 걸쳐서 존재해야한다. 또한 '*Include*'는 '*same*'이 아니면서 IP와 FP가 NM영역에 존재해야한다.

다음으로 '*leave*'에 대하여 살펴보면 영역에 관계없이 IP중 하나는 NMP와 같아야 하며 IP1 ~ FP는 NM영역에 존재하지 않아야 한다. 그리고 FP는 NMP를 기준으로 IP보다 멀어져야한다. 이와 같은 조건을 [표 4-4]와 같다.

모션동사	조건	
	공통조건	세부조건
leave	IP1,left ~ FP.left < NMP.left IP1,right ~ FP.right > NMP.right IP1,top ~ FP.top < NMP.top IP1,bottom ~ FP.bottom < NMP.bottom IP.right >= NMP.left IP.left >= NMP.right	IP.bottom = NMP.top FP.bottom < NMP.top
		IP.top = NMP.bottom FP.top > NMP.bottom
	IP1,left ~ FP.left < NMP.left IP1,right ~ FP.right > NMP.right IP1,top ~ FP.top < NMP.top IP1,bottom ~ FP.bottom < NMP.bottom IP.top =< NMP.bottom IP.bottom >= NMP.top	IP.left = NMP.right FP.left < NMP.right
		IP.right = NMP.left FP.right > NMP.left

[표 4-4] '*leave*'에 대한 규칙

'*arrive*'는 '*leave*'와 반대로 FP를 제외하고 IP ~ IP(n-1)까지의 모든 좌표가 NM영역에 존재하지 않아야 하며 FP중 하나는 NMP와 같아야 한다. 이는 [표 4-5]와 같다.

모션동사	조건	
	공통조건	세부조건
arrive	IP,left ~ IP(n-1).left < NMP.left IP,right ~ IP(n-1).right > NMP.right IP,top ~ IP(n-1).top < NMP.top IP,bottom ~ IP(n-1).bottom < NMP.bottom IP.right > NMP.left IP.left > NMP.right	FP.bottom = NMP.top IP.bottom < NMP.top
		FP.top = NMP.bottom IP.top > NMP.bottom
	IP,left ~ IP(n-1).left < NMP.left IP,right ~ IP(n-1).right > NMP.right IP,top ~ IP(n-1).top < NMP.top IP,bottom ~ IP(n-1).bottom < NMP.bottom IP.top < NMP.bottom IP.bottom > NMP.top	FP.left = NMP.right IP.left < NMP.right
		FP.right = NMP.left IP.right > NMP.left

[표 4-5] '*arrive*'에 대한 규칙

다음으로 '*come_up*'과 '*depart*'에 대하여 살펴보면, IP가 존재하는 영역에 따라서 조건이 달라진다. 예를 들어 '*come_up*'의 경우 IP가 1영역에 있을 때, FP는 1, 2, 4영역으로 제한한다. 또한 IP가 2, 4 ,5, 7 영역에 존재할 때, FP가 자신의 영역을 벗어나지 못하도록 제한한다. 이는 이 영역을 벗어날 경우는 의미적인 관계를 찾을 수 없기 때문이다. [표 4-6]은 '*come_up*'에 대한 규칙이다.

모션동사	조건		
	IP 영역	FP영역	세부조건
come_up	1	1 or 2 or 4	IP.top < FP.top IP.left < FP.left
	2	2	IP.top < FP.top
	3	2 or 3 or 5	IP.top < FP.top IP.left > FP.left
	4	4	IP.left < FP.left
	5	5	IP.left > FP.left
	6	4 or 6 or 7	IP.top > FP.top IP.left < FP.left
	7	7	IP.top > FP.top
	8	5 or 7 or 8	IP.top > FP.top IP.left > FP.left

[표 4-6] '*come_up*' 대한 규칙

'*depart*'는 '*come_up*'과 반대로 IP가 2, 4, 5, 7영역에 존재할 때, FP를 NM영역을 제외한 인접영역으로 제한하며, IP가 1, 3, 6, 8영역에 존재할 때, FP를 자신의 영역으로 제한한다. [표 4-7]은 '*depart*'에 대한 규칙이다.

모션동사	조건		
	IP 영역	FP영역	세부조건
depart	1	1	IP.top > FP.top IP.left > FP.left
	2	1 or 2 or 3	IP.top > FP.top
	3	3	IP.top > FP.top IP.left < FP.left
	4	1 or 4 or 6	IP.left > FP.left
	5	3 or 5 or 8	IP.left < FP.left
	6	6	IP.top < FP.top IP.left > FP.left
	7	6 or 7 or 8	IP.top < FP.top
	8	8	IP.top < FP.top IP.left < FP.left

[표 4-7] '*depart* 대한 규칙

'*enter*'와 '*exit*'는 2장에서 제안한 방법에 의해 조합으로 표현할 수 있으며, '*enter*'에 대한 규칙은 '*come_up*'과 '*arrive*' 그리고 '*include*'의 결합으로 표현할 수 있고, 여기서 IP는 '*come_up*'과 '*arrive*'의 규칙으로 FP는 '*include*'의 규칙으로 표현할 수 있다. [표 4-8]은 이 관계를 간결하게 표현한 것이다. '*enter*'는 [그림 4-4]에서 IP는 NM영역이 아닐 때 FP는 NM영역이여야 하며, 반대로 '*exit*'는 IP영역은 NM영역이여야 하고 FP는 NM영역이 아닐 때 성립한다.

모션동사	조건	
	IP와 NMP	FP와 NMP
enter	IP.bottom < NMP.top or IP.top > NMP.bottom IP.right < NMP.left or IP.left > NMP.right	FP.top > NMP.top FP.bottom < NMP.bottom FP.left > NMP.left FP.right < NMP.right
exit	IP.top > NMP.top IP.bottom < NMP.bottom IP.left > NMP.left IP.right < NMP.right	FP.bottom < NMP.top or FP.top > NMP.bottom FP.right < NMP.left or FP.left > NMP.right

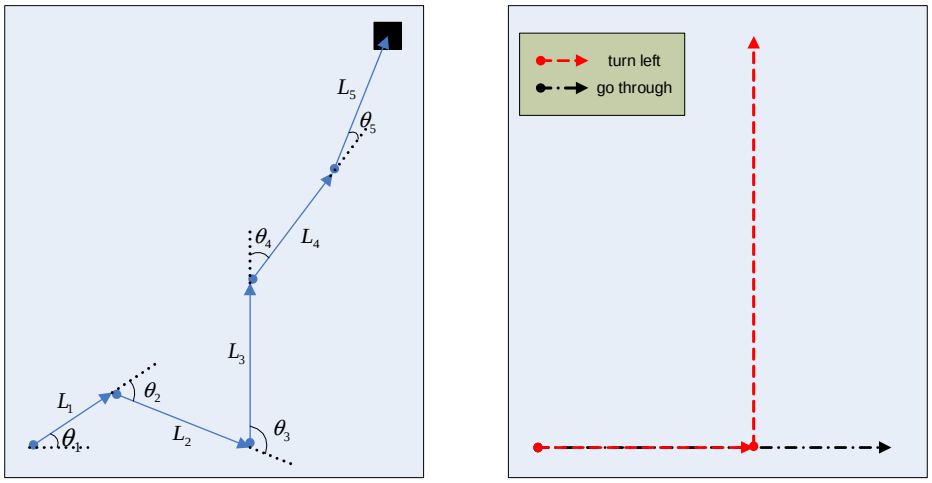
[표 4-8] '*enter*'와 '*exit*'에 대한 규칙

마지막으로 '*turn_left*', '*turn_right*', '*go_through*', '*return*' IP와 FP는 NM영역이 아니고, IP1 ~ IP(n-1)이 NM영역이야 한다. 여기서 이 4 가지 경우는 위와 같은 관계로 구별할 수 없기 때문에 TSR을 이용하여 유사도를 측정하였다.

2. *TSR*을 이용한 유사도 측정

TSR(Tangent Space Representation)[12,13]은 주로 2, 3차원 객체의 형태(shape)에 대한 유사도를 측정할 때 많이 사용된 방법이나 본 논문에 맞게 수정하여 적용하였다. TSR을 사용함으로써 객체의 방향과 이동거리를 고려한 유사성 측정이 가능하고 움직이는 객체의 궤적의 크기에 상관없이 정형화를 시킬 수 있으며, 면적을 통하여 유사도를 구할 수 있다.

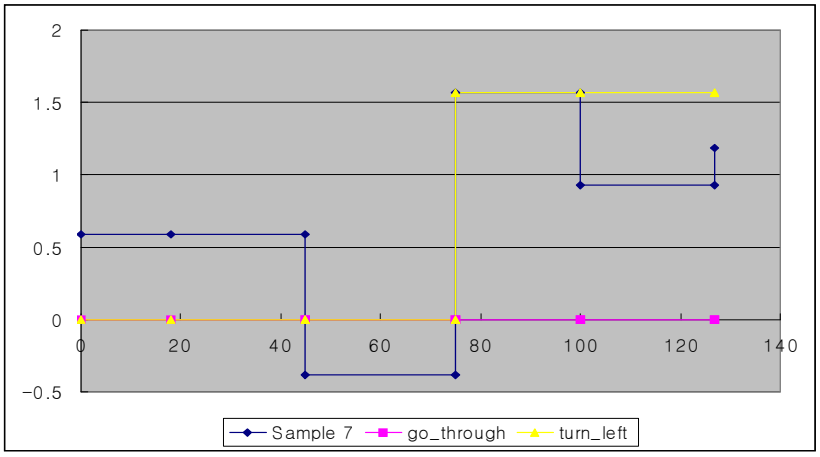
형태는 선으로 구성되어있고, 이 선은 좌표간의 연결이라고 할 수 있기 때문에, 궤적을 TSR로 표현할 수 있다. [그림 4-6]은'*Sample 7*'에 대한 궤적을 나타내고 있다.



(a) 'Sample 7'의 궤적 (b) (a)에 대한 'go_through'와 'turn_left' 비교

[그림 4-6] 'Sample 7'의 궤적과 의미적 궤적간의 비교

[그림 4-6](a)의 궤적에 대해서 TSR로 표현하면 [그림 4-7]에서 '*sample 7*'과 같다. 여기서 x축은 움직임 객체의 길이의 합(L_{sum}=L_1 ~L_5)을 나타내고 y축은 각도 변화의 합(θ_{sum}=θ_1 ~θ_5)를 나타내고 있다.



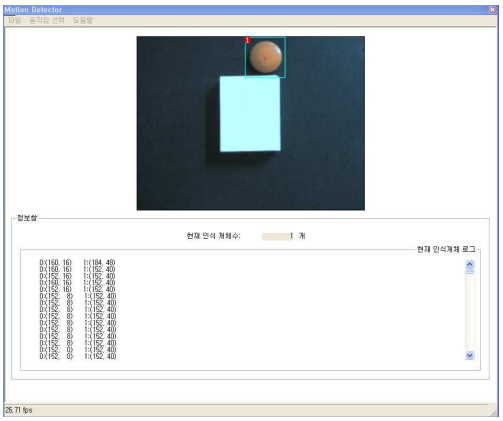
[그림 4-7] [그림 4-6] 대한 TSR

여기서 θ값은 라디안(Radian)값으로 변환하여 TSR로 표현한다. 예를 들어, θ_1이 33.6901도 이면 θ_1에 대한 라디안은 $33.6901\pi/180$이 되어 약 0.588값을 나타낸다. 그리고 [그림 4-9]에서 '*sample 7*'과 '*turn_left*' 그리고 '*sample 7*'과 '*go_through*'와의 유사도(면적의 차의 절대 값)를 계산하면 다음 [그림 4-6](a)와 (b)간의 유사도는 [표 4-9]와 같다.

	sample 7과 '*go_through*'	**sample 7과 'turn_left'**
면적 총계 (Similarity)	102.098	**57.163**

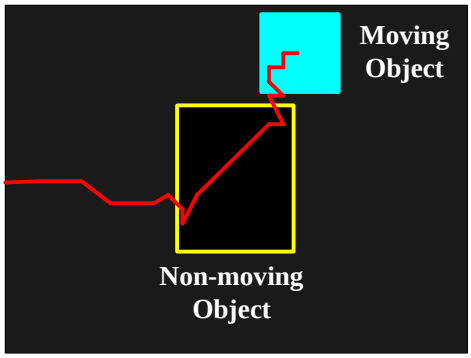
[표 4-9] [그림 4-6]에 대한 유사도 측정

[그림 4-8]은 *Andrew Kirillov*의 객체 추출(Object Detection) 알고리즘을 이용하여 실제 비디오에서 움직이는 객체와 궤적 추출이다. 여기서 파란색 MBR의 동그란 객체는 움직이는 객체로, 이 움직이는 객체에 대한 실시간 궤적은 아래 좌표와 같다.



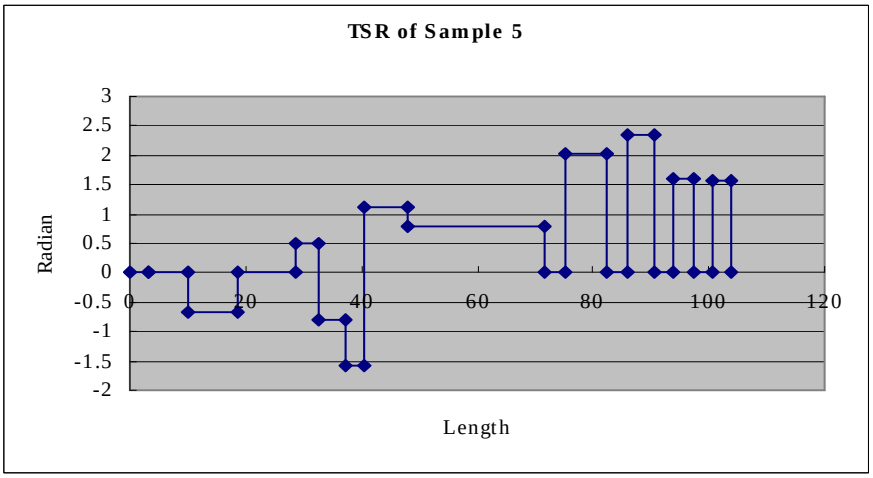
[그림 4-8] 실제 비디오 내의 움직임 객체의 궤적

[그림 4-9]는 [그림 4-8]에서 움직이는 객체와 고정된 객체 그리고 움직이는 객체의 궤적을 나타내고 있다. 여기서 궤적은 위의 좌표들의 연결선으로 다음과 같이 표현된다.



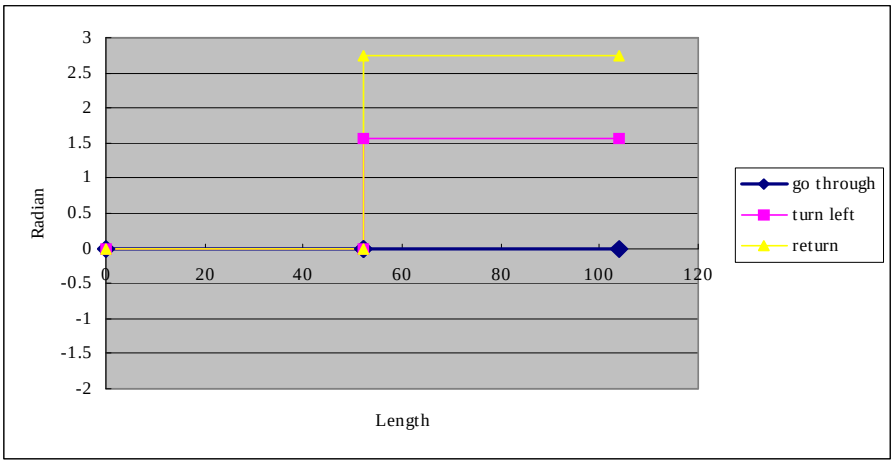
[그림 4-9] [그림 4-8]에서의 궤적

여기서 위의 궤적 정보를 이용하여 TSR로 변형한 그림은 [그림 4-10]이다.



[그림 4-10] [그림 4-9]에 대한 TSR

[그림 4-10]에서 Y축의 값의 '0'을 기준으로 하여 면적을 구하면 51.26031의 값을 얻었다. 이를 이용하여 [그림 2-12]의 움직임 객체의 시공간적 모델링과의 면적의 차에 의해 유사도 측정(Similarity Measure)이 가능하다. [그림 4-11]은 제안한 모델에서 *'go_through'*, *'turn_left'*, *'return'*에 대한 TSR이다.



[그림 4-11] *'go_through'*, *'turn_left'*, *'return'*에 대한 T-TSR

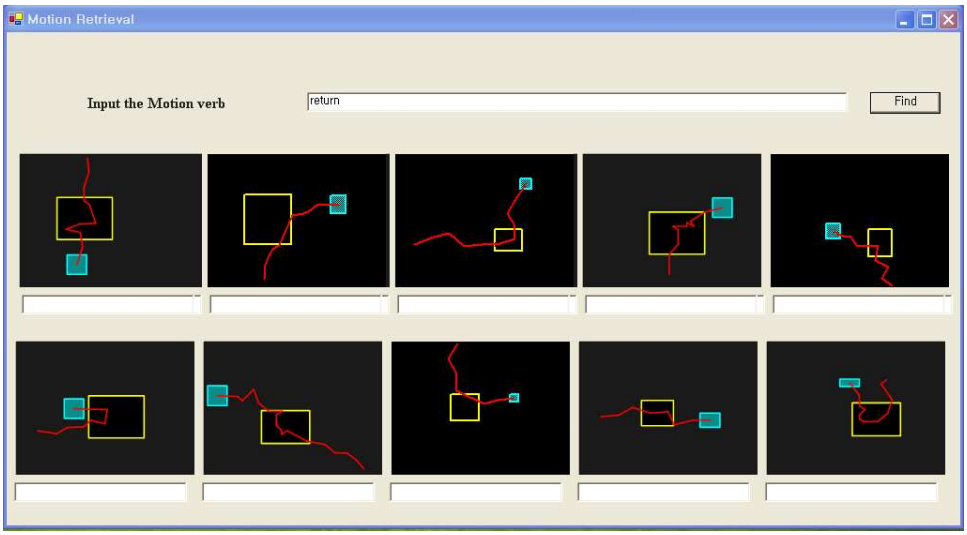
따라서 [그림 4-10]과 [그림 4-11]의 그래프간의 면적의 차를 통해 [표 4-10]과 같은 결과를 얻었다.

	go_through	**turn_left**	**return**
면적 총계 (Similarity)	51.26031	**30.4637**	97.0059

[표 4-10] [그림 4-10]과 [그림 4-11]에 대한 유사도 측정

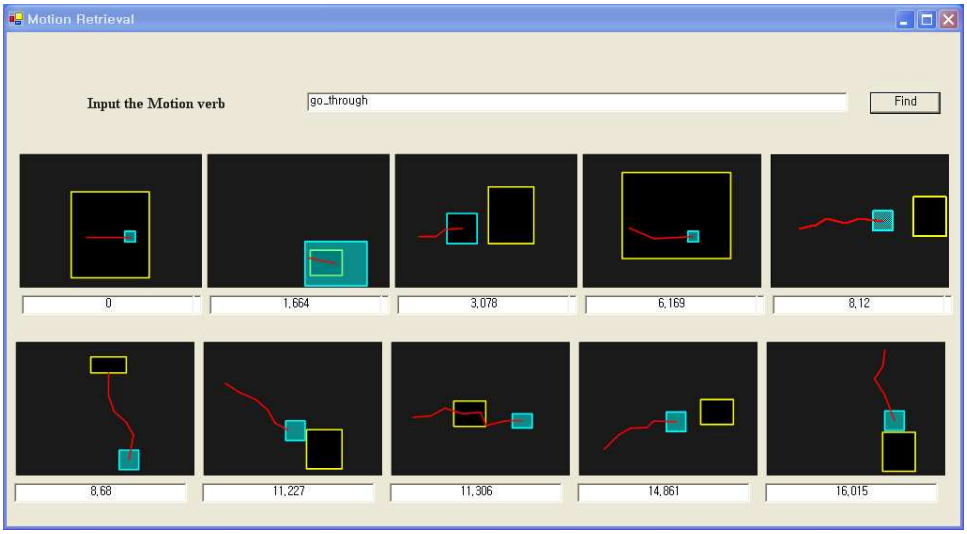
D. 실험 및 성능평가

의미 기반 궤적 검색 실험을 위해 총 94개의 궤적에 대해 실험을 하였으며, 정의한 모션동사를 이용하여 궤적을 검색했다. 먼저 제안된 규칙만 사용했을 경우는 방향관계를 측정하지 못하므로 [그림 4-12]와 같이 *'turn_left'*, *'go_through'*, *'return'*, 'turn_right'가 서로 구별되지 못하는 문제점을 가지고 있다.



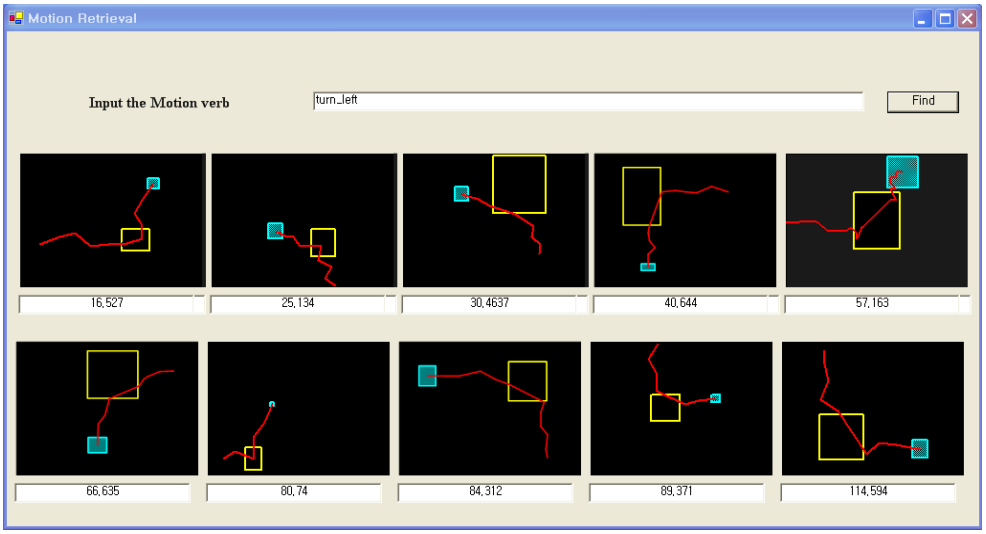
[그림 4-12] 제안한 규칙만 적용했을 경우의 문제점

또한 TSR만 적용했을 경우는 방향을 가지고 있는 모델에 대해서는 유사도 측정을 통해 구별을 하였으나 [그림 4-13]과 같이 '*turn_left*', '*return*', '*turn_right*'를 제외한 나머지 모델에 관해서는 구별하지 못하였다.



[그림 4-13] TSR만 적용했을 경우의 문제점

이러한 문제점을 해결하기 위해서 제안한 규칙과 TSR 모두를 고려하였을 때 실험 결과는 [그림 4-14]와 같다.



[그림 4-14] 제안한 규칙과 TSR 모두를 고려한 결과

본 논문에서는 비디오 내 움직임 객체의 궤적과 제안한 모델링간의 유사도 측정을 통하여 의미 기반 검색의 성능평가를 하였다. 이를 위해 검색된 궤적 중에서 질의와 관련된 궤적에 대한 비율인 Precision rate를 사용하였다. 다음 [표 4-11]은 의미 기반 궤적 검색의 Precision rate를 나타내고 있다.

적용대상	Precision rate
제안한 규칙	71.428 %
TSR	35.786 %
제안한 규칙 + TSR	**92.857 %**

[표 4-11] 의미 기반 궤적 검색의 Precision rate

실험 결과, 복잡한 궤적이나 움직임 객체가 여러 개가 있는 샘플은 객체 추출부

분에서 문제점을 보였다. 각각의 궤적들은 총 (320, 240) pixel로 구성이 되는데 객체의 움직임을 추출하는 과정에서 8 pixel 단위로 객체를 추출하기 때문에 이 보다 작은 객체는 인식하지 못하였다. 또한 컬러 값에 의존한 결과 빛에 대한 민감성 때문에 움직이는 객체의 그림자와 기타 움직이지 않는 부분에 대해서도 추출이 되는 등 문제점을 가지고 있다. 이는 향후 객체 추출 알고리즘 보완을 통해 해결할 수 있을 것이다.

또한 본 논문에서는 의미 기반 궤적 검색을 위해 움직임 객체에 대해 일반적인 모션 동사로만 매치하였으나, 향후 특정 비디오에 대해 적용 가능한 모션동사의 확장과 변형을 통해 좀 더 의미적인 검색이 가능할 것으로 사료된다.

V. 결론 및 제언

최근 멀티미디어 데이터의 증가에 따라 내용기반 검색뿐만 아니라 의미 및 개념 기반 검색의 필요성이 대두되고 있다. 이에 본 논문에서는 비디오 내 가장 중요한 의미적 정보인 객체의 움직임에 대한 의미적 표현에 관해 연구했다.

움직임 객체의 시간관계(Temporal relation), 위상관계(Topological relation), 방향관계(Directional relation)를 이용하여 객체의 움직임에 대한 정보를 얻고, 움직임 객체의 시공간 관계 모델링을 제안하였다. 이렇게 제안한 시공간 관계 모델들에 모션 동사를 매핑함으로써 의미적인 요소를 부여할 수 있었다. 이와 같이 제안한 의미적인 움직임 객체의 시공간 관계 모델과 비디오에서 움직임 객체의 궤적과의 비교를 통해 의미적인 움직임 객체의 궤적을 검색하는데 활용해 보았다. 이를 위해 각 시공간 관계 모델에 대한 규칙을 설정하고, 특히 방향성을 갖는 경우 TSR을 이용하여 유사도를 측정하였다.

본 논문에서 제안한 방법을 이용하여 기본적임 움직임 객체에 대한 의미는 얻을 수 있으나 객체의 정확한 의미 표현 및 분석에는 아직 어려움이 있다. 향후 시공간 관계 모델에서 축구, 당구 등 세부 도메인을 고려한 온톨로지(Ontology)를 구축하여 좀 더 다양한 시공간 관계를 설정하고, 어휘적 상·하위 관계와 동의어 관계 그리고 반어의 관계를 더 추가하여 좀 더 의미적인 시공간 관계 모델을 구축할 필요성이 있다. 또한 유사도 측정방법에서 제안한 규칙 세부적 수정보완과 TSR을 궤적 검색에 최적화 알고리즘도 필요하다. 이를 이용하여 향후 비디오에 대한 자동 나레이션 생성 및 검색 등에 활용할 수 있을 것으로 기대된다.

참고문헌

[1] Zaher A. Aghbari, "Studies on Modeling and Querying Video Databases", degree of Doctorate of Philosophy, Kyushu University, 2001

[2] Chang,S.-F., Hsu, W., Kennedy, L., Xie, L. , Yanagawa, A. , Zavesky, E., and Zhang, D.-Q., , "Video Search and High-Level Feature Extraction", TRECVID-2005.

[3] A. Ekin, A. Tekalp, and R. Mehrotra, "Automatic soccer video analysis and summarization", IEEE Trans. Image Processing, 12(7), pp. 796-807, 2003.

[4] M.Egenhofer and R.Franzosa, "Point-set topological spatial relations", International Journal of Geographical Information Systems, vol. 5, issue 2, pp.161-174, 1991.

[5] S.Y. Lee and F.J. Hsu "Spatial reasoning and similarity retrieval of images using 2D C-String knowledge representation", Pattern Recognition, vol. 25, no. 3, pp. 305-318, 1992.

[6] J.W. Chang, Y.j. Kim and K.J. Chang, "A Spatial Match Representation Scheme for Indexing and Querying in Iconic Image Databases", ACM International Conference on Information and Knowledge Management, pp.169-176, Nov. 1997.

[7] Jae-Woo Chang and Yeon-Jung Kim, "Spatial-Match Iconic Image Retrieval with Ranking in Multimedia Databases", Proceedings of Advances in Web-Age Information Management: Second International Conference, Jul. 2001.

[8] W. Ren, M Singh and S Singh, "Image Retrieval using Spatial Context", 9th International Workshop on Systems, Signals and Image Processing, November, 2002

[9] Pei-Yi Chen, Arbee L.P. Chen, "Video Retrieval Based on Video Motion Tracks of Moving Objects", Proceedings of SPIE Volume 5307, pp.550-558, 2003.

[10] Beth Levin, "English Verb Classes and Alternations : *A preliminary Investigation*", The University of Chicago Press, 1993.

[11] http://wordnet.princeton.edu/

[12] Sunkyoung Baek, Myunggwon Hwang, Miyoung Cho, Chang Choi, and Pankoo Kim, "Object Retrieval by Query with Sensibility based on the Kansei-Vocabulary Scale," Computer Vision in Human-Computer Interaction, The Proceeings of the ECCV2006 Workshop on HCI, LNCS 3979, pp.109-119, 2006.

[13] Myunggwon Hwang, Sunkyoung Baek, Hyunjang Kong, Juhyun Shin, Wonpil Kim, Soohyung Kim and Pankoo Kim, "Adaptive-Tangent Space Representation for Image Retrieval based on Kansei", LNAI 4293, pp.828-837, MICAI 2006.

[14] J. F. Allen, "Maintaining Knowledge About Temporal Intervals", Communications of the ACM, vol. 26, no. 11, pp.832-843, November 1983.

[15] C,C, Chang and J,H Jiang, " A fast Spatial Match Retrieval Using a Superimposed Coding Techniques.", International Symposium on ADT(Nara, Japan), pp. 71-78, 1994.

[16] John Z. Li, M. Tamer Ozsu, Duane Szafron, "Modeling of Moving Objects in a Video Data-base", In Proceedings of the International Conference on Multimedia Computing and Sys-tems, pp.336-343, 1997.

[17] Lisa Pearl, "Home to the castle and Other Locative Oddities", For graduate syntax class spring 2003.

[18] LT. Bleam, M. Palmer, and K. Vijay-Shanker. "Motion verbs and semantic features in TAG". In TAG+4 Workshop, 1998.

[19] H. Wang and S.F.Chang, "Automatic face region detection in MPEG video sequences", Electronic Imaging and Multimedia Systems, SPIE, .Photonics China, 1996.

[20] R. DeQueiroz, Z. Fan, and T. Tran, "Optimizing Block Thresholding Segmentation for Multilayer Compression of Compound Images", IEEE Transactions on Image Processing, 2000.

[21] S. Ji and H. W. Park, "Image Segmentation Of Color Image Based On Region Coherency", IEEE Transactions on Image Processing, 1998.

[22] Fatih Porikli, Huifang Sun, "Compressed Domain Video Object", IEEE Transactions on Circuits and Systems for Video Technology, may 2005.

[23] Hariadi MOCHAMAD Hui Chien LOY Takafumi AOKI, "Semi-Automatic Video Object Segmentation Using LVQ with Color and Spatial Features", IEICE TRANSACTIONS on Information and Systems Vol.E88-D No.7 pp.1553-1560, Jul. 2005.

[24] ihttp://www.codeproject.com/cs/media/Motion_Detection.asp

[25] Resnik P. "Using information content to evaluate semantic similarity", In Proceedings of the 14th International Joint Conference on Artificial Intelligence, pp.448-453. 1995.

[26] Wu Z. and Palmer M. "Verb Semantics and Lexical Selection", In Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics. Las Cruces, New Mexico, 1994.

[27] Leacock C. and Chodorow M, "Combining local context and WordNet similarity for word sense identification" In Fellbaum, pp.265-283, 1998.

[28] Banerjee S. and Pedersen T, "An Adapted Lesk Algorithm for Word Sense Disambiguation Using WordNet", In Proceeding of the Fourth International Conference on Computational Linguistics and Intelligent Text Processing (CICLING-02). Mexico City, 2002.