

이미지 개체 표현을 위한 논리적 프레임워크

최준호**, 김성석*, 최 창*, 김판구**

*조선대학교 전자계산학과

**조선대학교 컴퓨터공학과

e-mail : spica@chosun.ac.kr

A Logical Framework for Image Object Representation

Jun-Ho Choi**, Sung-Suk Kim*, Chang Choi*, Pan-Koo Kim**

*Dept. of Computer Science, Chosun University

**Dept. of Computer Engineering, Chosun University

요 약

의미 기반 영상 검색은 Color, Texture, Region 정보, Spatial Color Distribution 등의 저차원 특징 정보와 이미지 데이터에 의미를 부여하기 위해 주석 처리하는 것이 일반적이다. 그리고 부여된 키워드나 시소러스와 같은 어휘 사전을 이용하여 의미기반 정보검색을 수행하고 있지만, 기존의 키워드기반 텍스트 정보검색의 한계를 벗어나지 못하는 문제를 야기 시킨다.

따라서 본 논문에서는 WordNet 어휘 사전을 확장한 개념적 어휘 체계를 갖는 대형 Ontology를 기반으로 하여 이미지 데이터 내의 객체 인식과 추출된 객체간의 관계를 정의하여 이를 논리적으로 표현할 수 있는 방법을 제시하고자 한다.

1. 서론

최근 이미지나 비디오와 같은 시각 데이터에 내포된 의미적(Semantical) 요소를 표현하고, 그 내용을 검색하기 위한 개념 모델을 제시하는 연구가 다양하게 진행되고 있다. 최근까지 연구되어온 내용은 이미지 데이터의 경우 의미기반 정보검색을 위해서는 기존의 Color, Texture, Region 정보, Spatial Color Distribution 등의 저차원(Low level) 특징 정보와 이미지 데이터에 의미(Semantic)를 부여하기 위해 주석 처리하는 것이 일반적이다. 그리고 부여된 키워드를 Thesaurus와 같은 어휘 사전을 이용하여 의미기반 정보검색을 수행하고 있지만, 이런 경우 기존의 키워드기반 텍스트 정보검색의 한계를 벗어나지 못하는 문제가 발생할 수 있다[10].

따라서 본 논문에서는 WordNet 어휘 사전을 확장한 개념적 어휘 체계를 갖는 대형 Ontology를 기반으로 하여 이미지 데이터 내의 객체 인식과 추출된 객체간의 관계를 정의하여 이를 논리적으로 표현할 수 있는 방법을 제시하고자 한다.

2. 이미지 내 개체간의 관계표현

이미지 내 개체의 의미적 표현을 위해서는 이미지에 포함된 개체를 추출하는 과정이 선행되어야 한

다. 이 과정은 이미지의 저차원(Low-Level)특징을 추출하는 부분에 해당하는 것으로, Semantic Image Segmentation 과정을 거친다[2]. Semantic Image Segmentation 과정은 이미지의 일정 영역 내에 존재하는 Color, Texture 등의 값을 추출할 수 있도록 하는 Segmentation 단계와 이 과정을 통해 분할된 영역과 그 속성값(Color, Texture 등)을 기반으로 시각-공간 규칙을 적용하는 Region Merging 단계로 구분할 수 있다. Semantic Image Segmentation 과정을 통해 얻어진 특징 정보를 설계된 HFV(Heterogeneous Feature Vector) 구조로 표현한 후, 개념적 의미 부여를 위한 Semantic-Learning 알고리즘을 적용시켜 학습을 시키고 분류 알고리즘을 통해 유사 이미지를 분류한다.



(그림2.1) Quadtree를 이용한 영상 분할

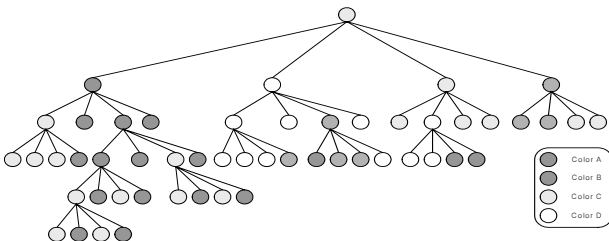
여기서 얻는 HFV는 각 이미지에서 의미 있는 영

역을 특징으로 갖는 Semantic Segmentation Tree로 구성된다. 이러한 의미 있는 영역 트리를 생성함으로써 인간의 의미적 인지에 해당하는 영역 기반 표현보다 High-Level의 성질을 추가한다. 그러나 인간의 Semantic-Understanding은 이러한 영역기반 표현보다 더 High-Level에 속한다. 따라서 더 가까운 의미적인 접근을 위해 Low-Level 과정을 거쳐 얻어진 정보를 이용하여 High-Level에서 의미정보를 포함하는 개체를 표현하는 방법을 모색하였다. 이러한 정보를 표현하기 위해, 우선 표현하고자하는 이미지의 개체정보에 대한 벡터(HFV) 구조를 알고, 그 구조를 토대로 개체를 High-Level에서 논리적으로 구조화시킨 후 구조화된 개체를 이용하여 개체간의 관계를 표현한다. 이러한 개체 간 관계가 정립되면 개체에 의미적인 요소를 부여할 수 있다.

3. 추출된 객체의 Description Scheme

3.1 Color Structure Description

Low-Level에서 추출하여 벡터정보(HFV)로 표현된 개체를 인간의 감성이 개입되는 High-Level에서 논리적인 표현으로 나타내고, 그 개체가 갖는 속성에 대한 구조를 논리적으로 정의해야 한다. 또한 이미지 내 개체를 어떻게 효율적으로 Logical 하게 표현할 것인가에 대해서 고려한다. 본 논문에서는 Color 정보를 이용하여 속성을 구조화 하였다. 예를 들면, 영역 분할 알고리즘을 이용하여 특정 영역의 컬러들의 집합을 다음 (그림 3.1)과 같이 표현할 수 있다.



11100000111000000001000000110000010000010100000010000
 ABAAABBBBBAABABBABBAABABBBDDDDDCDCBCDD
 AADDBBAACCCAA

(그림3.1) Color Structure Description의 예

3.2 Color Structure Description의 논리적 표현

제안된 Color Structure Description의 표현을 High-Level의 논리적 표현을 위해 여러 Description Language를 이용하여 Visual Semantic Ontology 모델링의 기본적인 구조를 설계하려고 한다. 다음은 (그림 3.2)의 Color Structure Description을 XML 형태로 표현한 예이다.

```
<DocID>Image URL</DocID>
<MultimediaAsset>
<URL>Image Name</URL>
<!-- General descriptors of the image -->
<HNumBins>64</HNumBins>
<Histogram>
<Bin Occurence="0.0">0</Bin>
<!-- ...중략... -->
</Histogram>
<NumDominantColors>3</NumDominantColors>
<DominantColor/>
<!-- Structure-based descriptors -->
<Structure>11100000111000000001000000110000010000010100000010000
000010000</Structure>
<Labels>ABAAABBBBBAABABBABBAABABBBDDDDDCDC
BCDDAADDDBBAACCCAA</Labels>
</MultimediaAsset>
```

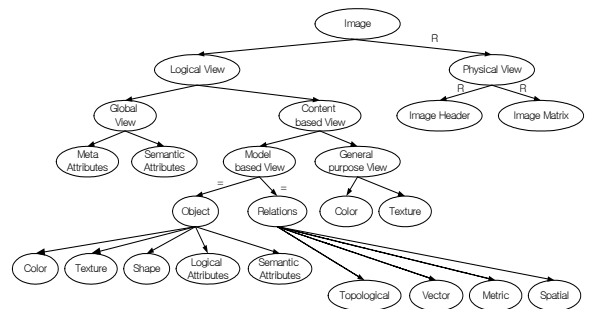
(그림3.2) Color Structure Description의 XML 표현

4. Visual Semantic Ontology의 모델링

본 장에서는 영상 내 객체 추출 단계를 거친 후, 각 객체간의 움직임 표현 등이 이루어진 다음에, 그 움직임의 내용 이해를 개념적 수준에서 이루어지도록 하기 위해 관련 Ontology를 구축하기 위한 방법을 기술한다.

4.1 Visual Semantic Ontology의 의미 계층 구조화

기존의 Ontology는 개념과 개념간의 관계 표현을 통해 개념을 계층적 구조적으로 표현하고 있다. 이러한 Ontology의 특징을 Visual Semantic Ontology에 적용시켜 이미지 데이터 내의 Visual 개체 및 추출 정보를 의미적으로 계층 구조화하고자 한다. 이를 위해 개체의 체계적 분류가 필요하며 개체와 개체간의 관계 정립이 요구된다. 다음 (그림4.1)은 이미지 데이터의 개념적 정보를 구조화한 예이다.



(그림 4.1) 이미지 데이터의 Semantic Schema

4.2 상위 레벨 구현을 위한 의미적 Ontology 모델

이미지 데이터는 크게 물리적 표현과 논리적 표현으로 나눌 수 있다. 이 중 이미지의 의미적인 요소는 논리적인 표현으로 기술이 가능하며, 이는 이미지 서술 방법과 이미지 개체들의 특징, 개체간의 관계를 포함하고 있다. 이미지의 Ontology 표현을 위해 기본 블록을 설정하고, 해당 블록 내의 픽셀 집

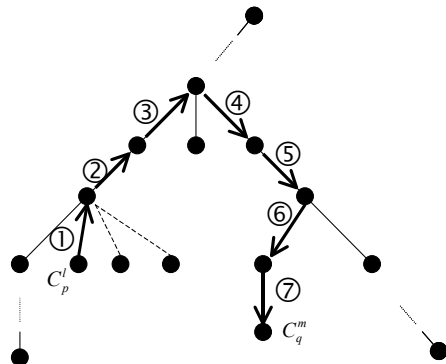
(표 4.1) 이미지 Ontology를 위한 표현 구조

개념	설명
	표현식
Image	픽셀 값의 n차원 배열 Image: $\{M_1(t_1), M_2(t_2), \dots\}, M_i(t_i) \subseteq R$
Geometric Structure	영역 R에 대한 위치(S) 집합, D(Domain), A(Attribute) geometry: $[S_1, \dots, S_m, a_1, \dots, a_n], S_i \subseteq R, a_i \in D(A_i)$
Image Processing Function	Method Ontology의 컴포넌트 process_image: image \rightarrow image
Geometric Image Classification Function	Image와 Geometric Structure 집합의 매핑 classify_image: image \rightarrow {geometry}
Functional Structure	이미지 각 영역R의 범위 V와 연결된 함수 속성 funct_struct: [image, range, attribute_function]
Application Domain Classes	개념 정의 클래스 app_class
Matching	Geometric Structure와 Function Structure의 인스턴스 geometric_matching = (geometry _i , app_class) functional_matching = (funct_struct _i , app_class)
Spatial Configuration	Geometric Structure와 Function Structure의 매칭 집합 정의 spat_conf: {(geometry _i , app_concept), i=1, n} spat_conf: {(funct_struct _i , app_concept), i=1, n}

합 R을 이용하여 블록 내 영역을 표현한다. 이를 위해 (표4.1)과 같은 표현식을 정의하였다. 또한, 상위 레벨의 이미지 검색을 위해 논리적 이미지 데이터에 대한 모델 설계가 요구되는데, 특히 개체의 의미적 요소를 검색하기 위해 논리적 표현 중 컬러 값이나 정립된 모델을 이용하여 Shape과 같은 다른 논리적 요소를 이용하여 이미지의 의미를 부여하는 방안도 고려하여야 한다.

4.3 관계 표현법을 이용해 개체간의 관계규칙 정의

이미지 내에는 크고 작은 여러 개의 Visual 개체가 존재하며 각각의 개체는 포함, 인접 등의 관계를 가질 수 있다. 본 연구에서는 이러한 Visual 개체간의 관계를 표현하기 위해 관계 규칙을 정의하고자 한다. 또한 이를 검색에 활용하기 위해 개체의 의미적 관련 질의어와 부합되도록 하는 개체 표현 방안을 도출한다.



(그림4.2) 개념간 유사도 측정

Ontology 상의 개념 기반 검색을 위해서는 운영 관리 구조를 마련하여 개념기반 정보 색인, 공유, 검색이 가능하도록 그 기초를 제공하도록 한다. Ontology를 적용해서 개념 기반 정보검색을 위해서는 개념 간 유사도 측정이 중요한데, 이를 위해 WordNet을 이용한다. WordNet의 경우 IS_A, HAS_PART, MEMBER_OF, HAS_SUBSTANCE 등과 같은 관계로 구성되어 있으며, 이들에 대해 각각 다른 가중치를 주는 문제, 또 경로 거리에 대한 문제 등을 고려하고 있다.

링크 Pi로 연결된 두 개념 C_i와 C_j⁻¹에 대해서 개념 거리 S_{P_i}^s(C_i, C_j⁻¹) = δφw_{i→(l-1)}^{P_i}로 고려하였으며, 여기서 l은 레벨을, δ은 Intra-level conceptual distance control function으로, 그리고, φ은 Inter-level conceptual distance control function을 정의하고 있으며, w_{i→(l-1)}^{P_i}은 Travel Cost이다. (그림4.2)는 WordNet 상에서의 개념 간 유사도를 측정하는 Weighted graph travel cost 방법을 표현한 예이다.

이와 같이 기존의 Ontology 공간상의 개념 간 유사성 측정을 위한 계산 도구를 마련하고 있지만, Visual Semantic Ontology상에서 다른 이미지간 유사도 측정 방법과 비교·분석할 필요가 있다.

5. Visual Semantic Ontology의 적용 방안

의미 기반 이미지 검색을 위해 구축된 Visual Semantic Ontology는 각각의 이미지에 대해 컬러, shape, texture 등과 같은 많은 의미적 요소를 포함하고 있다. 이러한 방대한 의미적 요소를 효율적으

로 관리하고 표현하기 위해 이미지 데이터베이스 스키마 설정 및 상세 설계가 필요하다. 즉, HFV에 의해 추출된 특징들은 실제 스트링 형태이기 때문에 Table 및 Relation에 기반을 둔 데이터베이스에 적용이 가능하다. 이를 위해서 데이터베이스에 저장될 이미지의 요소를 물리적인 요소와 논리적인 요소로 분류하여 HFV에서 해당되는 값을 추출한 후 이를 검색에 활용할 수 있다. 따라서 HFV로 표현된 내용을 실제 관계형(Relation) 데이터베이스에 적용시킬 때 효율적 적용에 대해 이미지 데이터의 물리적인 요소와 논리적인 요소를 Physical Image Representation과 Logical Image Representation의 영역으로 분류하였다.

6. 결론

본 논문에서는 이미지 데이터 내의 객체 인식과 추출된 객체간의 관계를 정의하여 이를 논리적으로 표현할 수 있는 방법을 제시하였다.

이미지 기반으로 구축된 본 프레임워크를 시간의 개념을 도입하여 비디오 영역으로 확장한다. 구축된 의미적 Ontology는 비디오 내 정보를 표현하는데도 유용하므로 적용이 가능하다. 이때 필요한 과정으로 3차원 Mesh 모델 생성, 각 Shot에 대한 시각적 속성값 및 Motion을 표현하는 Syntactic Content 생성, 각 개체의 시공간적 특성을 기술하기 위한 실시간적 3차원 Bounding Volume의 생성, 각 Volume Trajectory를 갖는 개체와 그 움직임의 공간적 속성 등을 Description Logic을 기반으로 표현 등을 들 수 있는데, 이러한 특성을 고려하여 비디오 내용 기술과 접목시키면 비디오 데이터의 Visual Semantic Ontology 구축이 가능할 것이다.

참고문헌

- [1] H.Denman, et al, "Content Based Analysis for Video from Snooker Broadcasts", Image and Video Retrieval, LNCS2383, July, 2002.
- [2] M.Pickering, et al, "Video Retrieval by Feature Learning in Key Frames", Image and Video Retrieval, LNCS2383, July, 2002.
- [3] R.Leonardi, et al, "Semantic Indexing of Multimedia Documents", IEEE Multimedia, 2002, April-June, Vol. 9, No. 2.
- [4] Yixin Chen, James Z. Wang, "A region-based fuzzy feature matching approach to content-based image retrieval," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 9, pp. 1252-1267, 2002.
- [5] A. B. Benitez and S.-F. Chang, "Multimedia Knowledge Integration, Summarization and Evaluation", Proceedings of the 2002 International Workshop on Multimedia Data Mining in conjunction with the International Conference on Knowledge Discovery & Data Mining (MDM/KDD-2002), Edmonton, Alberta, Canada, July 23-26, 2002; also Columbia University ADVENT Technical Report #003, 2002.
- [6] A. B. Benitez and S.-F. Chang, "Semantic Knowledge

- Construction from Annotated Image Collections", Proceedings of the 2002 International Conference On Multimedia & Expo (ICME-2002), Lausanne, Switzerland, Aug 26-29, 2002; also Columbia University ADVENT Technical Report #002, 2002.
- [7] Yixin Chen, James Z. Wang, Jia Li, "FIRM: Fuzzily Integrated Region Matching for content-based image retrieval," Proc. ACM Multimedia, pp. 543-545, Ottawa, ACM, September 2001. Also in e-proceedings.
- [8] J. Assfalg, et al, "Semantic Annotation of Sports Videos", IEEE Multimedia, 2002, April-June, Vol. 9, No. 2.
- [9] Peter L. Stanchev, "Object-Oriented Image Model", 16th International Conference on Computers and Their Applications (CATA-2001), March 2001.
- [10] PanKoo Kim, Youngchoon Park, Forouzan Golshani, and Sethuraman Panchanathan, "A Logical Framework for Visual Information Modeling and Management", Circuit Systems and Signal Processing, Sept. 2001.