

## 영상의 한국적 감성 형용사 추출 및 감성 콘텐츠 프레임 워크 설계

백성은\*, 신성윤\*, 이양원\*

### Extract korean sensitivity adjective from image and Design for sensitivity contents framework

Seong-Eun Baek \*, Seong-Yoon Shin \*, Yang-Won Rhee \*

#### 요 약

본 논문에서는 영상으로부터 한국적 감성 데이터를 추출하고 감성 콘텐츠 프레임워크(ISC) 구축을 통해 영상데이터를 다양한 멀티미디어 데이터와 함께 사용자에게 전달하는 시스템을 제안한다. 이를 위해 먼저 영상으로부터 객체를 분리해내고 각각의 특징벡터 모델을 구축한다. 특징벡터모델은 요소인자로서 벡터공간의 데이터로 입력되고, 정규화된 한국어 형용사 데이터와의 근접도를 비교하여 대표 감성을 표현하게 된다. 또한 감성 콘텐츠 데이터의 3층 구조를 이용하여 멀티미디어 데이터를 분류하고 대표 감성형용사와의 인자 사상을 통하여 다양한 영상-감성 콘텐츠를 제공하는 프레임워크를 설계한다. 이것은 다양한 영상의 추상데이터에 대응한 감성표현을 추출하여 영상에 내포되어 있는 정확한 의미를 관찰자에게 전달할 수 있다.

▶ Keyword : 감성공학, Kansei Data Analysis, 추상 영상, 감성 콘텐츠 프레임 워크, 감성 형용사

#### 1. 서 론

영상으로부터 감성을 추출하는 주제는 인지과학 분야와 감성공학 분야 등에서 연구되고 있다. 대표적인 연구 분야로는 감성기반 멀티미디어 분류 및 검색이 있는데, Takagi는 인간의 감성에 기반한 검색을 위해 감성 인자 공간을 구성하고 영상 공간과의 사상(mapping)관계를 정의하여 대화형 유전자 알고리즘(IGA)을 통하여 영상검색에 사용하였고[12], Colombo등은 Itten[7]의 컬러이론을 이용해 예술 영상을 감성에 기반하여 분석하였다[8].

이러한 형태는 전체 영상의 특징벡터를 사용하므로 정확한 의미의 감성 조합을 구성하기 어렵고, 대표적인 10개 미만의 감성만을 사용하여 실험하였기 때문에 일반화하기 어렵다. 이를 해결키 위해서는 다양한 감성지표를 도입함과 동시에 영상을 구별 가능한 객체 단위로 분류하는 선행 작업이 먼저 이루어져야한다. 또한 시각센서로부터 추출한 한 컷의 정지영상에는 많은 정서적 내용들을 담고 있는데, 이것은 시간적인 흐름에 따른 의미 변화를 테마로 규정하는 매체라기보다, 대표하고자 하는 촬영 시점의 정적이며 함축적 의미를 담고 있으므로, 실제 시각적으로 작은 정보를 통해 큰 의미를 전달하고자 하는 현상매체에 가깝고,

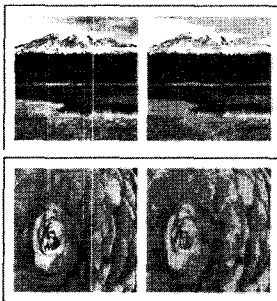
• 제1저자 : 백성은

\* 군산대학교 컴퓨터정보과학과

산출물을 보는 관찰자의 환경적 영향에 따라 다른 내용을 전달하기도 한다. 따라서 영상의 정확한 의미를 전달코자 할 때에, 내부 함축적인 의미관계는 전체 영상의 느낌과 더불어 영상에 존재하는 객체들의 느낌으로부터 도출해야 하며, 각 객체들의 데이터는 감성을 전달코자 하는 대상의 환경적 배경을 고려해야 한다. 본 논문에서는 감성 공간을 분석하기 위해 감성인자 공간(factor space) 분석법(9)을 참고하였고, 디자인과 관련된 연구를 토대로 구축된 I.R.I 형용사 이미지 스케일(10)을 이용해 한국적 감성 형용사를 추출하였다. 추출된 감성형용사는 주 제어 분류기에 의해 구성된 멀티미디어 콘텐츠와 연동하여 입력된 영상과 가장 유사한 멀티미디어 콘텐츠를 제공하게 된다.

## 2. 영상의 추상 모델링

영상으로부터 의미 정보를 추출하기 위해 먼저 컬러, 패턴, 외곽선과 같은 영상특징을 각각의 객체로부터 추출하는 과정이 필요하다. 먼저 영역분할을 통해 객체를 구분하고 각각의 객체 특징은 다시 특징벡터공간으로 바뀌어 대표 특징 벡터를 추출한다. 추출이 완료되면 대표 특징 벡터를 가지는 객체 클러스터 군들을 하나의 영상 모델로 추상화하게 된다. 이러한 과정을 통해 특징검출을 객체단위(object sequence)로 한정할 수 있고 영상이 담고 있는 의미를 객체 단위로 분리하게 된다.



〈그림 1〉 영상 분리

### 2.1 객체 분할

컬러 영상을 영역화(segmentation) 하는 방법에는 크게 영역 기반(region-based) 방법과 경계선 기반(contour-based) 방법이 있다. 대표적인 영역기반방법으로 Mean-shift 알고리즘과 K-means 알고리즘이 있는데, 이 중 K-mean 알고리즘은 구현하기가 쉽고 컬러공간을 구분하여 K의 개수와 평균값을 자동 적 응적으로 초기화할 수 있다.

K개의 평균값이 결정되면 아래의 식 (1). 에서

$$X \in C_i(t), \text{ iff } d(X, Y_j) < d(X, Y_i) \text{ for all } \forall i \neq j \dots (1)$$

와 같이 K개의 중심화소  $Y_i$ 를 대상으로 이들 화소값 X에 근접한 화소들을 모은 후에 이들 중심 화소들을 중심으로 거리를 계산하여 근접하게 군집된 화소들의 평균값을 구하여, 다음 반복처리의 중심값으로 결정하게 된다. 군집과 평균 과정은 더 이상 평균값이 변하지 않을 때까지 반복한다.

### 2.2 영상 특징

인간의 감성에 영향을 주는 보편적인 영상 요소로는 색상과 질감이라는 사실이 Soen 등에 의해 입증된바 있다.[3] 이와 함께 영상의 외곽선 정보는 객체지각을 지도하는 데이터로 주어진다. 본 논문에서는 추상영상을 구성하는 데에 있어 인간 시각에 가장 먼저 지각되는 1차 조건인 색상정보를 색채 지각의 보편적인 구성형태인 H.S.I. 컬러공간으로 변환하여 사용하도록 하였다. Hue, Saturation, Intensity 세가지의 데이터는 컬러공간 좌표로 특징벡터 모델의 구성요소가 된다. 그리고 질감정보 T는 엣지패턴의 각과 길이의 빈도수로 이루어져있으며, 외곽선 B의 S는 외곽 커널의 중심값이다.

$$V_n = \{C(H, S, I)_n, T(A_m, L_m)_n, B(S)_n\} \dots (2)$$

### 2.3 특징벡터공간 모델링

객체단위(Object Sequence)로 분리된 특징벡터는 구조적인 형태의 모델로 영상을 대표하게 된다.

$cf$  는 클러스터 군의 화소 빈도수 이고,  $M$ 은 클러스터 평균이다.

$$VO_n = cf_n(M - V_n)^2 \dots (3)$$

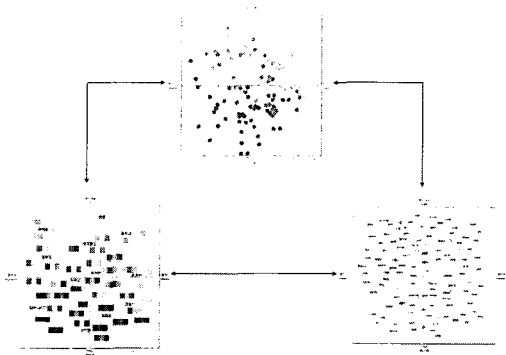
## 3. 감성 추출

### 3.1 감성 데이터

한 정서의 상태가 어떤 생리학적 변수들의 측정치에 의해 한가지로 지목될 수 있을 때, 그 생리적 측정 항목들을 우리는 정서 특정적 생리라고 정의 할 수 있다[5]. 이러한 생리적 특징은 매우 다양하고 연속적이므로 언어적 형태를 통하여 전달하는 것은 어려운 일이다. 그러나 인지추상화의 측면에서 감성

표현이라는 주제의 핵심은 임의의 대상 물체에 대한 구체적인 느낌을 서술하기보다, 대략적이고 모호한 공통적인 특성을 기술하고, 이 특성간의 조합을 통해 형성된 감성 집합의 형성에 있다. 또한 이러한 감성 표현들은 실제 경험적인 영향에서 오는 것일 가능성이 높고 이는 사회적 영향과 개인적 특성에 따라 다른 표현을 이끌어 낼 수 있다.

인간의 정서를 통계적으로 접근하여 감성공학의 기틀을 마련한 Kansei 공학[9]에서는 이와 같이 모호한 데이터들을 분류 가능한 인자 공간(factor space) 내에서의 집합으로 분류하고 대표 감성을 추출하였다. 공통감성거리를 측정하기 위한 요소인 인자분석[4]은 데이터 변동을 설명할 때 잠추어진 요인을 발견하기 위해 사용되는 다변량 해석 방법의 하나로 샘플과 감성표현으로 이루어진 인자공간을 분석하는 방법 중 하나이다. 인자분석을 위해 Osgood이 창안한 의미분별척도법(Semantic Differential Scale, SD method)[6]으로 구성된 각 샘플과 평가자의 평균데이터를 이용하여 인자분석을 하고 감성표현간의 관계를 인자공간으로 파악하는 것을 시도한다. <그림 2>는 한국적 형용사를 SD법을 사용하여 구성한 컬러-형용사 척도이다.



<그림 2> I.R.I Color Image Scale

감성 데이터 분석의 인자 분석 모델에서는 상품과 감성표현을 데이터 행렬로 구성하고 있는데, 여기에서 M x N 데이터 행렬의 표준화 데이터 Zmn 는

$$z_{mn} = \sum_{l=1}^L f_{ml} a_{nl} + u_{mn} d_n \dots\dots\dots (4)$$

라고 나타낸다. l 은 공통인자(common factor)라고 부르고, f<sub>ml</sub> 은 l 의 샘플 m 의 값으로 인자특점(factor score)으로 불린다. a<sub>nl</sub> 은 공통인자 l 과 감성표현 n 과의 관계를

나타내고 인자패턴(factor pattern) 혹은 인자부하량(factor loading)이라고 불린다. 인자부하행렬 A의 제n행을

$$a_n = (a_{n1}, \dots, a_{nL}) \dots\dots\dots (5)$$

이라 하면 이것은 인자공간에 있어서의 감성 표현 n의 좌표이다. 그리고 u<sub>mn</sub> 은 감성표현 n의 독자인자(unique factor)의 샘플 m의 값으로 독자인자 스코어라고 불려진다. 마지막으로 d<sub>n</sub> 은 감성표현 n의 독자인자 가중치고 d<sup>2</sup><sub>n</sub> 은 독자성이라고 불린다. 즉 데이터로부터 독자인자라는 오차항을 뺀 양을 샘플 m의 인자공간에 있어서의 좌표 벡터와, 감성표현 n의 인자공간에 있어서의 좌표벡터간의 내적으로 표현하는 것이다. 구성된 인자공간은 다차원 공간 좌표축 간에 존재하는 각각의 감성 공간을 교합하는 관계 함수를 필요로 하며, 이러한 관계함수는 동적이여야 하고 감성데이터에 독립적이고 다양한 감성 공간에 적용가능 해야만 한다. 이를 위해 n개의 차원 공간을 동일 차원으로 정규화하는 함수는 아래와 같다.

$$f(N_n) = \sum_{i=1}^n X_i (d^2_n - d^2_{n'}) \dots\dots\dots (6)$$

각 차원의 영역 외의 양을 인자의 가중치로 변환하면 다른 좌표축을 가지는 다차원 공간을 정규화 할 수 있다.

### 3.2 감성 형용사 맵

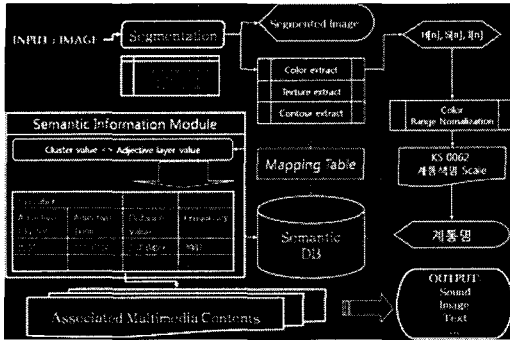
특징벡터 모델의 좌표공간과 감성 형용사 공간의 좌표는 서로 일치하므로, 벡터 모델의 중심영역(center point area)과 형용사좌표의 중심거리를 계산하여 가장 가까운 감성 형용사를 선택한다. 대상 중심간 유클리드 거리는 동일 차원의 다른 공간을 가지는 각각의 원소인 특징모델 va와 형용사모델 sa 차이를 이용하여 아래와 같이 구한다.

$$d(SO_n, VO_n) = \|sa_n - va_n\| \dots\dots\dots (7)$$

구해진 거리의 최소값으로 클러스터를 대표하는 감성표현을 구할 수 있고, 이 감성 표현들의 집합은 더 큰 감성 표현 형용사의 하위 클러스터로 존재하게 된다.

각각의 객체로부터 추출된 감성 형용사가 동일한 클러스터내에 존재할 경우 영상의 의미 가중치는 높아지고, 이는 일관성있는 감성 추론을 불러일으킨다. 따라서 동일한 군집 혹은 동일한 감성 형용사 발생 빈도수가 높아질 수록 최종 감성 콘텐츠의 선택 또한 정확률이 높아지게 된다.

#### 4. 영상-감성 콘텐츠 프레임 워크



〈그림 3〉 영상-감성 콘텐츠 디자인 플로우

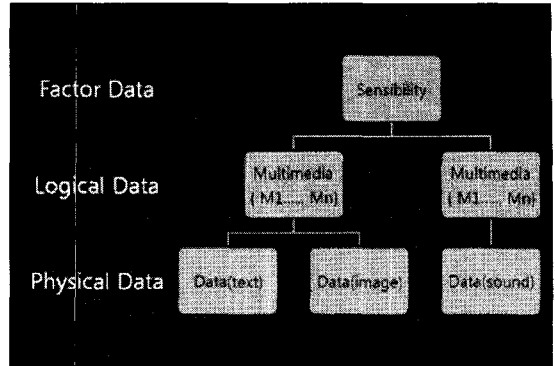
영상-감성 콘텐츠(Image-Sensitivity Contents, ISC)의 변환 과정은 〈그림 3〉과 같은 순서를 따라 작동하는 구조이다. 초기에 벡터모델과 객체구조에 대한 사전지식을 초기화한다. 그리고 영상이 들어오면 객체를 분리하여 특징을 추출하고, 인자공간 분석과 주제어분류기를 통해 정의된 각 형용사들의 의미 그룹(Semantic Database)과 사상하여 유사도가 가장 높은 형용사와 감성 콘텐츠를 출력한다.

#### 4.2 공간 사상 테이블

사상 테이블(Mapping Table)에서는 소속 함수에 의하여 감성 표현 과의 거리를 측정하고 유사 데이터의 양이 증가하면 공간 간의 연결 가중치를 증가시킨다.이 중 최대 연결치를 가지는 콘텐츠를 구하여 입력 영상의 감성 형용사 집합에 더한다.

#### 4.3 감성 데이터 구조

데이터를 이루는 멀티미디어의 구조는 〈그림 4〉와 같은 3단계의 계층구조로 설계한다. 물리적인 멀티미디어 데이터는 형용사 데이터의 하부 레이어인 논리 계층의 리스트에 연결(hyperlink, Anchor)되어 감성데이터에 따라 다른 콘텐츠 제공을 쉽게 전환할 수 있도록 하고, 각 계층의 논리적인 관계 구조가 바뀌더라도, 빠르고 유연하게 적용할 수 있도록 한다. 이는 공통적인 감성 공간뿐만 아니라, 개인적이거나 다른 감성 표현에 더 쉬운 반응을 보일 수 있는 사람들을 고려함으로써 좀더 정확한 콘텐츠의 제공에 기여할 수 있다.



〈그림 4〉 감성 데이터 계층 구조

#### 5. 실험 및 평가

제안된 감성영상 프레임워크의 실험을 위해서 자연 풍경을 담은 450여장의 이미지와 12개의 감성 공간 그룹에 할당된 “포근한, 아기자기한, 수수한, 스포티한, 기운찬...”등과 같은 60개의 감성 형용사를 사용하였다. 이는 자연 풍경을 묘사할 때 자주 사용하는 감성 형용사 중에서 대표적인 것을 선택한 것이다



〈그림 5〉 실험을 위한 영상의 예

#### 5.1 감성 표현의 정확성

본 실험의 검증을 위하여 무작위로 선발된 50명의 평가자에게 감성표현 맵을 제공하고, 실험영상을 보여준 뒤 이에 대한 가장 가까운 감성형용사를 적어도 한 개 이상 선택하도록 하였다. 그리고 같은 영상을 ISC 시스템에 적용한 결과를 비교하여 정확률을 산출하였다.

〈표 1〉 감성표현 정확도

대상	감성군 적합도	감성표현 정확도	평균 복잡도 증가율
객체 별	79.2%	87.0%	21%
전체 감성군	46.1%	53.5%	45%

표 1에서 보인 실험 결과에서 감성표현 정확도에 따라 감성군의 적합성 또한 근사하게 비례하여 상승함을 알 수 있다. 또한 영상의 특징 데이터양이 증가 할수록 감성표현의

복잡도가 증가하였으나, 객체 별로 분리하여 대표감성을 추출하였을 때에는 영향을 적게 받았다.

### 5.2 ISC 시스템의 감성 콘텐츠 유사성

감성 콘텐츠 지원 데이터의 신뢰도를 보장하기 위하여 감성 형용사와 감성 콘텐츠의 유사도를 구하였다. 앞의 실험과 더불어 동시에 동일한 평가집단을 대상으로 선택된 콘텐츠 중 적합한 데이터의 수와 임의로 분류한 제공한 콘텐츠의 유사도를 평가하는 실험을 진행하였다. 이 데이터를 기준으로 감성 콘텐츠 논리 계층 구조를 수정하는 작업을 수행하였다.

〈표 2〉 감성표현-콘텐츠 유사도

감성표현영커	콘텐츠 수	평균 유사도	피드백 횟수
60	180	80%	3

## 6. 결 론

영상의 추상화된 감성을 표현하는 것은 전체적인 영상 패턴의 구조를 파악하는 것이 아니라, 구별 가능한 영상 내부 객체 간의 관계와 각각의 감성을 파악하여 이들 간의 대표 감성간의 관계를 비교함으로써 판단하는 것이 바람직하다. 각각 추상화에 대한 개념은 경험적으로 누적될 수 있는 사회 구성원들을 통해 얻은 통계에 의해 가중치를 보완할 수 있고, 이는 현재 감성공학분야의 많은 연구를 통하여 증명되었다. 정확히, 감성 표현에 관련된 데이터가 실제로 개인별 특성에 의존적이고, 영상으로부터 추출한 객체 특성이 그 객체를 지각하는 대상의 감성을 모두 표현한다고 말하기 어렵다. 따라서 지역적인 공통성을 가지는 통계적 방법을 통해 감성 표현의 균집을 만들고 이에 대한 감성표현 추출의 정확률을 높이는 것이 본 연구의 목적이었으며, 이를 실험한 결과, 영상으로부터 한국적 형용사를 추출하는 부분에서 높은 정확률을 보여 좋은 성능을 보여주었다. 이렇게 표본집단과의 유사도가 높게 추출된 감성표현은 다양한 형태의 멀티미디어를 통해 확장하는 것이 가능한데, 감성공간변동에 적합한 설계 구조를 통해 동적으로 변화하는 인지 특성 변화에 대응할 수 있다.

정확한 영상의 감각 추상화는 지각장애인을 위해 주어진 영상매체를 다른 감각적 매체로 변환하는 것을 가능하게 하고, 감성표현과 관련된 멀티미디어 콘텐츠 검색을 지원할 수 있다. 또한 지능형 로봇의 인공지능에 포함될 자의적 감성 (self-emotion)과 같은 시각적 센서에 의한 감정 표현 메커니즘을 디자인하는 역할에 본 연구가 공헌하리라 생각한다.

## 참고문헌

- [1] J.-S. Um, K.-B. Eum, and J.-W. Lee, "A Study of the Emotional Evaluation Models of Color Patterns Based on the Adaptive Fuzzy System and the Neural Network," *Color Research and Application*, vol. 27, no. 3, pp. 208-216, 2002
- [2] 박정호, 한성배, 양선모, 김형범, 이순요, "다변량해석 기법을 활용한 감성 데이터베이스 구축에 관한 연구", *대한인간공학회*, vol 1. 1996
- [3] N. Kawamoto and T. Soen, "Objective evaluation of color design II," *Color Res. Appl.* Vol.18, pp.260-266, '1993
- [4] C. Spearman, General intelligence, objectively determined and measured. *American Journal of Psychology*, Vol. 15, pp. 201-293, 1904.
- [5] 우제린, "정서 특정적 생리의 탐색을 모색하는 감성공학의 패러다임과 실천방법," *한국감성과학회지*, Vol 4, No. 2, pp. 1-13. December 2001
- [6] C.E. Osgood, G.J.Suci and P.H. Tannenbaum, "The Measurement of Meaning", Univ. of Illinois Press, 1957
- [7] J. Itten, *Art of Color (Kunst der Farbe)*, Otto Maier Verlag, Ravensburg, Germany, 1961 (in German)
- [8] C.Colombo, A. Del Bimbo, and P. Pala, "Semantics in Visual Information Retrieval," *IEEE Multimedia*, vol. 6, no. 3, pp.38-53, 1999.
- [9] Y. Nakamori, "Kansei Data Analysis", sambuk, 2000
- [10] 이복신, "형용사를 이용한 색채 감성척도의 개발" *한국디자인학회*, vol 6, pp.17-18, 1998
- [11] Y. Rui and T.S.Huang, "Optimal learning in image retrieval," *IEEE Proc. of Conf. on Computer Vision and Pattern Recognition*, Vol.1, pp.236-243, 2001
- [12] H. Takagi, T. Noda, and S-B. Cho, "Psychological Space to Hold Impression among Media in Common for Media Database Retrieval System." In *Proc. IEEE Int. Conf. on System, Man, and Cybernetics*, pp.263-268, 1999