

# 의류 영상에서 컬러 영상 척도를 이용한 감성 인식

## Emotion Recognition Using The Color Image Scale in Clothing Images

이슬기, 우효정, 유성필, 김동우, 안재형  
충북대학교 전자정보대학 정보통신공학부

Seul-Gi Lee(lsg1271@hanmail.net), Hyo-Jeong Woo(redwoo@naver.com),  
Sung-Pil Ryu(cucus@chol.com), Dong-Woo Kim(dubssi@paran.com),  
Jae-Hyeong Ahn(jhahn@chungbuk.ac.kr)

### 요약

감성 인식은 기계가 작업을 자동적으로 판단하여 인간의 감성을 인식하도록 하는 것이다. 인간의 감성은 매우 주관적이므로 객관적인 측정이 불가능하다. 따라서 감성 인식의 목표는 가급적 많은 사람들이 공감하는 척도를 만드는 것이다. 영상에서의 감성 인식은 영상의 여러 가지 특징을 감성과 매칭하는 방법으로 구현된다. 본 논문에서는 고바야시의 영상 척도를 참고하여 의류 영상에서 색상 특징을 이용한 감성 인식 시스템을 제안한다. 제안 시스템은 영상 척도의 색상 정보를 데이터화하고, 입력 의류 영상에서 추출한 색상을 저장된 색상 데이터와 비교하여 가장 유사한 색상을 찾고, 이 색상과 연관된 감성을 출력한다. 제안된 시스템은 한 의류 영상에서 최대 3개까지 감성을 찾을 수 있다. 제안한 시스템의 성능을 평가하기 위하여 70명의 관찰자가 검사하였다. 그 결과에 따르면 제안한 시스템이 인식한 감성과 관찰자의 감성이 매우 유사하였다.

■ 중심어 : | 과학기술 | 영상처리 | 감성인식 | 의류 | 영상척도 |

### Abstract

Emotion recognition is defined as that machines automatically recognize human emotions. Because the human emotions is very subjective, it is impossible to measure objectively. Therefore, the goal of emotion recognition is to obtain a measure that is agreed by as many people as possible. Emotion recognition in a image is implemented as the method that matches human emotions to the various features of the image. In the paper, we propose an emotion recognition system using color features of clothing image based on the Kobayashi's image scale. The proposed system stores colors of image scale into a database. And extracted major colors from a input clothing image are compared with those in the database. The proposed system can obtain three emotions maximally. In order to evaluate the system performance 70 observers are tested. The test results shows that recognized emotions of the proposed system are very similar to the observers emotions.

■ keyword : | Science Technology | Image Processing | Emotion Recognition | Clothing | Image Scale |

## I. 서론

IT 기술이 발전할수록 기계가 인간처럼 말하고, 듣고, 스스로 판단하여 결정을 내리는 능력과 더불어 인간의 감성을 인식하기 위한 다양한 연구가 시도되고 있다 [1][2].

기계에 의한 감성 인식의 목표는 기계가 인식한 감성에 대하여 다수의 사람들이 동의하는 것이다. 인간의 감성은 매우 다양하여 회로에라 같은 기본적인 감정에서 함부터 말로 표현할 수 없는 미묘한 느낌까지 모두 포함된다. 게다가 똑같은 사물을 놓고도 사람들이 느끼는 감성은 저마다 다를 수 있기 때문에 모든 인간이 아닌 대다수를 만족시키는 방향으로 연구되고 있다[3].

인간의 감성은 오감 중에서 시각에 의해 가장 영향을 받는다. 더군다나 같은 내용을 갖고 있는 영상일지라도 그 영상의 색상, 질감, 형태와 같은 특징에 따라 다양한 감성으로 느껴질 수 있다. 영상의 특징 중 색상 특징은 잡음에 덜 민감하고 영상 내의 변화에 영향을 적게 받아 영상을 분류하는 적합한 기준이 된다[4]. 이와 마찬가지로 영상에서의 감성 인식에서도 색상 특징이 주로 이용된다. 본 논문에서는 생활의류 영상에서 색상 특징을 이용하여 감성을 인식하는 방법을 제안한다. 설문 조사를 통하여 제안 방법으로 감성을 인식한 결과가 실제 사람들이 느끼는 감성과 유사하다는 것을 입증했다. 본 논문의 구성은 다음과 같다. 2장에서는 관련 연구에 대하여 소개한다. 3장에서는 제안하는 방법의 알고리즘과 전체 시스템을 단계적으로 설명한다. 4장에서는 실험 결과를 기술하고, 5장에서 결론으로 논문을 마무리한다.

## II. 관련 연구

### 1. 고바야시의 영상 척도

고바야시는 오래전부터 색채에 따른 감성 변화에 대한 연구를 해오면서 사람들이 갖는 색상 조합에 대한 감성은 조금씩 다를 수 있지만 공통되는 부분이 많다는 것을 발견하였다. 고바야시는 공통되는 부분을 명확하게 정의하고자 노력했고, 1981년 연구의 결과로서 색상

조합에 따라 사람들이 느끼는 공통적인 감성을 규정하고 그에 따른 감성 모델을 제시하였다[5]. 고바야시의 영상 척도는 색상 조합에 따른 공통적인 감성을 분산 분석, 클러스터 분석, 요인 분석, SD법, 색채 투영법 등의 심리학적 연구에 기초를 두고 과학적으로 개발된 모델이다.

영상 척도는 [그림 1]과 같이 Warm-Cool, Soft-Hard 의 좌표축 상에 감성으로 표현된 색상 조합들이 배치되어 있다. 영상 척도에 있는 색상은 Munsell의 10색상이 채용되었고 유채색 12톤과 무채색 10색으로 표현된다.

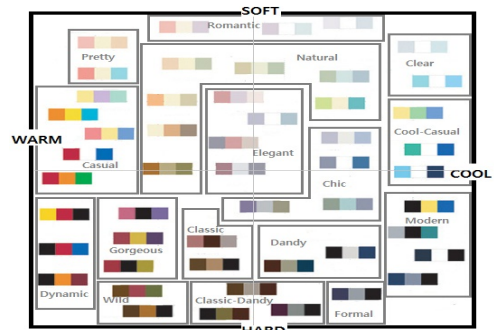


그림 1. 고바야시의 2차원 영상 척도 모델

고바야시가 규정한 감성은 총 180개로 인간의 언어에서 사용하는 형용사로 일대일로 대응된다. 이 감성 형용사는 다시 15개의 감성 그룹으로 분류된다[6].

### 2. K-means 클러스터링

K-means 클러스터링은 주어진 데이터를 K개의 클러스터로 묶는 알고리즘이다. 이 알고리즘은 임의의 점 (x,y)을 이웃하는 클러스터의 중심점들과 식 (1)과 같은 거리를 계산하여 그 중에서 가장 가까운 거리에 있는 클러스터에 배정하는 것이 핵심이다.

$$d(x,y;k) = |c(x,y) - \mu_k|^2 \quad k=1,2,\dots,K \quad (1)$$

여기서 c(x,y)는 점(x,y)에 있는 데이터 값이며,  $\mu_k$ 는 k 번째 클러스터의 중심값이다. K-means 알고리즘은 초기 중심점으로부터 클러스터의 경계선을 새로 도출하

고, 그 후 새로운 클러스터의 중심점을 재계산한다. 이러한 과정을 반복하되, 만약 클러스터의 경계가 바뀌지 않거나, 중심점이 변하지 않으면 반복을 중지한다[7].

### 3. 영상에서의 감성 인식에 대한 기존 연구

초기의 영상 감성 인식은 벽지와 옷감과 같은 일정한 패턴이 존재하는 텍스타일 영상에서 색상 특징을 감성에 매칭하는 방법이 연구되었다[8]. 미리 정의된 감성 색상과 텍스타일 영상의 색상을 히스토그램으로 비교하여 유사도를 계산하고, 유사도가 가장 높은 감성 색상이 해당 영상의 감성을 대표하도록 하였다. 이 방법은 단순하고 일정한 패턴이 존재하는 텍스타일 영상에서 좋은 결과를 얻을 수 있었다.

최근에 연구된 감성 인식 방법으로는 패턴 학습을 반영하여 업데이트된 감성을 매칭하는 기계 학습법이 있다[9]. 기계 학습의 종류에는 K-NN, neural network, SVM(support vector machine)이 대표적이다. 기계 학습의 입력 값은 영상의 색상이나 주파수 정보 같은 특징이 들어간다. 입력 값에 대한 처리를 미리 학습시켜 놓으면, 새로운 영상에 대하여 기존에 학습된 결과에 따라 새로운 영상에 대한 처리를 진행할 수 있다. 이러한 방법은 결과에 대한 정확도를 높일 수 있지만 미리 학습된 내용이 정확하지 않을 경우에는 오히려 정확도가 떨어질 수 있다는 것이 단점이었다.

### III. 제안 방법

본 논문에서는 의류 영상에서 색상 정보를 추출하여 감성을 인식하는 방법을 제안한다. 제안하는 방법은 [그림 2]와 같이 먼저 의류 영상을 색상 데이터에 따라 K-means 클러스터링으로 분할하고, 각 클러스터의 중심 색상을 추출한다. 중심 색상들은 고바야시의 영상 척도에서 사용한 색상들로 재정의 하고, 임의로 3개씩 조합하여 고바야시의 감성과 매칭한다. 가장 유사한 감성이 발견되면 해당 감성 그룹 일치 횟수를 증가시킨다. 모든 조합에 대해 감성 매칭이 완료되면 일치 횟수대로 감성 그룹을 정렬하고, 가장 많은 수치를 기록한

감성 그룹이 대표 감성이 된다. 이러한 과정을 K-means 클러스터링부터 반복하여 대표 감성이 바뀌지 않으면 반복을 중지한다.

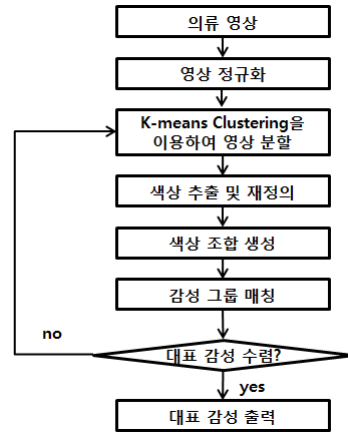


그림 2. 감성 인식 알고리즘

#### 1. 영상 정규화

생활 의류를 카메라로 촬영한 후에 영상을 처리하기에 앞서 영상의 크기를 정규화 하는 과정이 필요하다. 영상의 크기가 너무 크면 영상을 처리하는 과정에서 실행 시간이 너무 길어질 수 있고, 반면에 너무 작으면 의류의 형태나 색상을 식별하기가 어려워져 감성 인식 결과가 좋지 않거나 아예 불가능할 수 있다. [표 1]은 영상의 크기에 따른 평균 실행 시간과 인식된 감성의 평균 개수를 나타낸 것이다.

표 1. 영상의 크기에 따른 실행 시간과 감성 인식

영상 크기	평균 실행 시간	평균 감성 인식
100×100	0.633 [sec]	0.31 [개]
200×200	2.041 [sec]	1.33 [개]
300×300	7.000 [sec]	2.41 [개]
400×400	17.000 [sec]	2.53 [개]
500×500	32.000 [sec]	2.21 [개]

[표 1]에서 알 수 있듯이 영상의 크기가 500×500 화소보다 커질 경우 영상 분할 과정에서 실행 시간이 크게 증가하면서 오히려 감성 인식 효율은 떨어지는 것을 알 수 있다. 본 논문에서는 이러한 점을 감안하여 영상의 크기를 300×300화소로 정규화 하였다.

## 2. 영상 분할 및 색상 추출

정규화를 마친 영상에 대하여 K-means clustering으로 영상 분할을 수행한다. 이 때 최대 클러스터의 개수는 배경부분을 포함하여 8개로 정하였다. 고바야시의 감성은 총 180개이므로 배경부분을 제외한 클러스터가 7개이면 그 중 큰 순서대로 3개씩 선택해도  $r_3=210$ 개의 감성을 인식할 수 있기 때문이다.

영상 분할을 마치면 각 클러스터의 중심 색상을 추출한다. 중심 색상은 임의의 컬러로 이루어졌으므로 고바야시의 감성 그룹에서 사용한 색상으로 바꾸어야 한다. 이를 위하여 식 (2)와 같이 색상 거리를 정의하였다.

$$d(k,l) = |\mu_k - r_l|^2 \quad l=1,2,\dots,L \quad (2)$$

여기서  $r_l$ 은 고바야시 감성 그룹에서 사용한 색상 중  $l$  번째 색상이다. 재정의 색상은 이 색상 거리가 최소화 되는  $r_l$ 색상으로 선택한다.

## 3. 색상 조합 생성

K개의 클러스터의 재정의 색상이 중복될 경우에는 하나로 통합한다. 색상 조합은 서로 다른 3개의 재정의 색상을 조합하여 구성한다. 만약 전체 색상이 2개 이하일 경우에는 처리에서 제외하며 감성 인식 실패로 간주한다. 색상 조합은 최대  $r_{C_3}=35$ 개가 만들어질 수 있다.

## 4. 감성 그룹 매칭

감성 그룹 매칭 과정은 생성된 색상 조합이 고바야시의 어떤 감성 그룹에 매칭 되는 지 확인한다. 만약 찾았으면 해당 감성 그룹의 일치 횟수를 증가시키고, 어떤 그룹에도 색상 조합이 없을 경우 그 색상 조합은 폐기한다. 여러 개의 색상 조합이 같은 감성 그룹에 속하면 해당 감성 그룹의 일치 횟수가 계속 증가하게 된다.

감성 인식 결과의 정확도를 높이기 위해서 영상 분할부터 감성 그룹 매칭까지의 과정을 반복한다. 만약 감성 그룹의 일치 횟수 순위가 수렴되면 반복을 중지하며, 수렴이 되지 않더라도 10회에서 중지한다.

반복하는 과정마다 인식되는 감성이 다르기 때문에 일치 횟수가 높은 순서로 감성 그룹을 최대 3개까지 확

면에 출력한다. 만약 10회에서 강제로 중지했을 때는 대표 감성 하나만 출력한다.

## IV. 실험 결과

제안한 방법의 우수성을 확인하기 위하여 생활 여성복 102벌에 대하여 실험을 진행하였다. 실험에 사용된 영상은 여성복을 직접 디지털 카메라로 촬영하여 구하였다. 실험 환경은 Intel(R) Core(TM)2 Quad CPU에 3GB 메모리 Windows 7 Home Premium K 32비트의 PC이며, 시뮬레이션을 위한 소프트웨어는 Visual Studio 2010에서 MFC로 구현하였다. [그림 3]은 촬영한 의류 영상에 대한 감성 인식 시뮬레이션 화면이다.



그림 3. 의류 영상의 감성 인식 시뮬레이션 화면

여성복 영상 전체에 대한 시뮬레이션에서 첫 번째로 인식한 감성으로 romantic이 가장 많았으며, formal과 wild 항목은 전혀 인식되지 않았다. [그림 4]는 인식한 대표 감성을 일부 나타낸 것이다.



그림 4. 의류 영상에서 인식한 대표 감성

[그림 4]에 따르면 대표 감성이 인간이 인식하는 감성과 매우 유사하다는 것을 느낄 수 있다. 이 판단의 신뢰도를 측정하기 위하여 70명의 사람들에게 15개의 감성이 인식된 영상을 하나씩 보여주고 관측자가 느끼는 감성과 공감하는 지를 조사하여 [표 2]에 나타내었다.

표 2. 대표 감성에 대한 관측자들의 공감도

감성 그룹	공감	비공감	비율(%)
E1 romantic	53	17	75.71
E2 natural	66	4	94.29
E3 casual	54	16	77.14
E4 elegant	36	34	51.43
E5 pretty	53	17	75.71
E6 chic	47	23	67.14
E7 classic	34	36	48.57
E8 modern	45	25	64.29
E9 gorgeous	47	23	67.14
E10 dynamic	35	35	50.00
E11 formal	26	44	37.14
E12 cool-casual	52	18	74.29
E13 dandy	33	37	47.14
E14 wild	24	46	34.29
E15 clear	42	28	60
평균			61.62

[표 2]에 따르면 시물레이션에서 인식한 대표 감성에 대한 사람들의 공감도는 61%로 그다지 높지는 않았다. 이러한 이유는 한 의류에서 느낄 수 있는 감성이 여러 개 일뿐 아니라 개인 간의 의류 감성 선호도도 있기 때문으로 생각된다.

[그림 5]는 사람들이 의류 영상을 보고 느낀 감성의 비율과 시물레이션에서 인식한 대표 감성 및 제2감성, 제3감성을 모두 포함하여 계산한 비율을 2차원 평면에 나타낸 것이다.

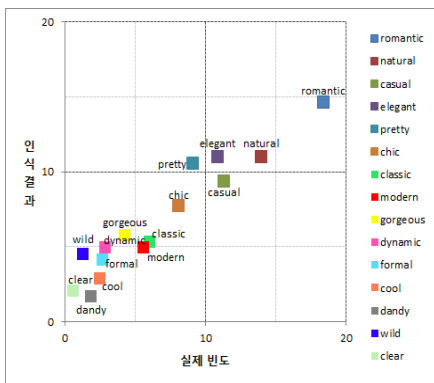


그림 5. 실제 빈도와 감성 인식 결과 비교

[그림 5]에 따르면 실제 사람들이 의류 영상에 대해 느끼는 감성의 인식 빈도와 시물레이션에서 찾은 인식 감성 인식 결과가 선형적인 관계라는 것을 보여준다. 이는 제안한 방법이 실제 사람들이 느끼는 감성과 유사하다는 것을 의미하는 것이다. 그리고 출현 빈도수가 높은 그래프의 제1사분면을 보면 romantic, natural, elegant와 같은 감성이 분포하고 있다는 것은 여성복에 대한 전체적인 감성이 고바야시의 감성 척도 모델에서 Warm-Soft쪽으로 치우치고 있음을 나타낸다.

### V. 결론

본 논문에서는 의류 영상에서 추출한 색상을 고바야시의 감성 척도에 매칭하여 영상에 포함된 감성을 인식하는 방법을 제안하였다. 제안한 방법을 적용한 시물레이션 인식 결과에 대해 실제 사람들이 느끼는 감성과의 공감도는 대표 감성인 경우에 61%에 달했으며, 한 영상에서 최대 3개까지 인식된 감성에 대해서는 실제 사람들이 느끼는 감성 비율과 매우 유사함을 보였다. 또한 여성복에서 느끼는 감성과 동일하게 고바야시의 2차원 감성 척도 모델에서 Warm 및 Soft쪽으로 감성이 치우침을 볼 수 있었다.

향후 실행 속도나 정확도에 대한 연구를 통해 보강하면 의류 업계나 쇼핑몰 등에서 소비자에게 감성적인 상품 추천 서비스를 제공하는 시스템도 구현할 수 있을 것이다.

### 참고 문헌

- [1] 안상민, 황민철, 김동근, 김종화, 박상인, “개인화 프로세스를 적용한 실시간 감성인식 기술”, 한국감성과학회 학술지, Vol.15, No.1, pp.133-140, 2012.
- [2] 김명규, 김정호, 차명훈, 채수환, “텍스트 문서 기반의 감성 인식 시스템”, 한국감성과학회 학술지, Vol.12, No.4, pp.433-442, 2009.
- [3] 조운호, 박규식, “성별 구분을 통한 음성 감성인식 성능 향상에 대한 연구”, 전자공학회논문지,

Vol.45, No.4, pp.107-114, 2008.

- [4] 김영래, 신윤희, 김은이, “텍스타일 영상에서의 감성 기반 검색 시스템”, 한국산업정보학회 학술대회논문집, Vol.2009, No.5, pp.82-87, 2009.
- [5] S. Kobayashi, “The Aim and Method of the Color 감성 척도”, Color Res.Appl, Vol.16, No.2, pp.93-107, 1981.
- [6] S. Kobayashi, COLOR 감성 척도, publishing of Kodansha, Tokyo, pp.15-150, 1991.
- [7] 김경환, 정진우, 최윤식, “K-MEANS CLUSTERING 기반 영상의 공간 해상도 축소 변환을 위한 효율적 움직임 벡터 재예측 방법”, 한국방송공학회 학술발표대회 논문집, Vol.2010, No.7, pp.567-569, 2011.
- [8] 이경미, 차은미, 박우창, 이은옥, “감성 색상을 이용한 텍스타일 영상 검색”, 덕성여자대학교 자연과학연구소 자연과학 논문집, Vol.14, pp.1-9, 2008.
- [9] 정민주, 기계학습을 이용한 영상에서의 실시간 감성 추출, 한신대학교 대학원, 석사학위논문, 2012.

저 자 소 개

**이 슬 기(Seul-Gi Lee)** 준회원



- 2012년 2월 : 충북대학교 정보통신공학과(공학사)
- 2014년 2월 : 충북대학교 정보통신공학과(공학석사)  
<관심분야> : 영상처리, 내용기반검색, 감성인식

**우 효 정(Hyo-Jeong Woo)** 준회원



- 2012년 2월 : 충북대학교 정보통신공학과(공학사)
- 2014년 2월 : 충북대학교 정보통신공학과(공학석사)  
<관심분야> : 영상처리, 컴퓨터비전, 감성인식

**류 성 필(Sung-Pil Ryu)** 정회원



- 2014년 2월 : 충북대학교 정보통신공학과(공학박사)
- 2004년 2월 ~ 2006년 3월 : LG 전자 단말연구소 주임연구원
- 2006년 9월 ~ 2008년 2월 : 충북대학교 정보통신공학과 누리초빙 전임강사
- 2009년 1월 ~ 2013년 1월 : HERO Tech 기술이사  
<관심분야> : 멀티미디어 정보처리, MPEG, H.264, 컴퓨터 비전

**김 동 우(Dong-Woo Kim)** 정회원



- 1997년 2월 : 충북대학교 정보통신공학과(공학사)
- 2002년 2월 : 충북대학교 정보통신공학과(공학석사)
- 2006년 2월 : 충북대학교 정보통신공학과(공학박사)
- 2013년 3월 ~ 현재 : 충북대학교 시간강사  
<관심분야> : 영상처리, 컴퓨터 비전, 네트워크프로그래밍

**안 재 형(Jae-Hyeong Ahn)** 정회원



- 1981년 : 충북대학교 전기공학과(공학사)
- 1983년 : 한국과학기술원 전기 및 전자공학과(공학석사)
- 1990년 : 한국과학기술원 전기 및 전자공학과(공학박사)
- 1987년 ~ 현재 : 충북대학교 정보통신공학부 교수  
<관심분야> : 영상통신, 영상 인식, 멀티미디어 정보처리, 웹 프로그래밍, 모바일 프로그래밍