

감성 기반의 자동 텍스타일 영상 분류 시스템

김영래[○] 신윤희 김은이
건국대학교 신기술융합학과

kisssgood@konkuk.ac.kr, ninharsa@konkuk.ac.kr, eykim@konkuk.ac.kr

Automatic Textile-Image Classification System using Human Emotion

Youngrae Kim[○] Yunhee Shin Eun Yi Kim

Dept. of advanced technology fusion, school of Internet and multimedia Eng. Konkuk Univ., Korea

요 약

본 논문에서는 감성을 기반으로 텍스타일 영상을 자동으로 분류할 수 있는 시스템을 제안한다. 이 때, 사용된 감성 그룹은 고바야시의 10가지 감성 키워드 - {romantic, clear, natural, casual, elegant, chic, dynamic, classic, dandy, modern} - 를 이용한다. 제안된 시스템은 특징 추출과 분류로 구성된다. 특징 추출 단계에서는 텍스타일을 구성하는 대표 컬러를 추출하기 위해서 양자화 기법을 이용하고, 패턴정보를 표현하기 위해서는 웨이블릿 변환 후의 통계적인 정보를 이용한다. 신경망 기반의 분류기는 추출된 특징들을 입력으로 받아 입력 텍스타일 영상을 분류한다. 제안된 감성인식 방법의 효율성을 증명하기 위해서 220장의 텍스타일 영상에서 실험한 결과 제안된 방법은 99%의 정확도를 보였다. 이러한 실험 결과는 제안된 방법이 다양한 텍스타일 영상에 대해 일반화되어 사용될 수 있음을 보여주었다.

1. 서 론

인간의 감성에 대한 예측은 다양한 비즈니스와 과학 및 공학 응용에서 매우 중요하다. 특히 감성지향 제품군으로 분류되는 패션, 섬유, 인테리어 산업에서의 감성은 매우 중요하다^[8]. 이러한 중요성 때문에 감성에 따른 텍스타일 영상의 분류가 최근 많은 주목을 받고 있다. 감성에 따른 텍스타일 영상의 분류는 멀티미디어 검색 시스템과 결합하여 사용자에게 임의의 텍스타일 영상에 포함된 감성을 기준으로 영상을 색인하고 사용자의 질의에 대해 데이터를 검색하여 서비스 할 수 있게 한다. 또한 사용자가 바라는 감성의 친을 자동으로 선택할 수 있게 하기 때문에 사용자의 요구대로 텍스타일의 기획과 생산을 가능하게 한다.

텍스타일 영상에서의 자동 감성 분류를 위해서는 영상에 포함된 컬러 및 패턴 등의 물리적인 특징을 이용하여 감성 정보를 표현 해야만 한다. 따라서 최근 이러한 시각 정보를 이용한 감성 연구가 이루어지고 있다.^[9] 이러한 연구 중 대표적인 몇 가지를 설명하면 다음과 같다. 고바야시는 컬러가 인간의 감성에 미치는 영향에 대해 조사하였으며 이를 분석함으로써 특정 컬러로 구성된 영상을 특정 감성 그룹으로 맵핑할 수 있는 "Color image scale"을 개발하였다^[1]. 그러나 고바야시의 연구는 이러한 요소들 간의 상호관계를 조사하였지만, 영상으로부터 이러한 요소들을 자동으로 추출하고, 추출된 정보를 자동으로 분류하는 시스템 개발은 이루어지지 않았다. 반면, 소옌은 주파수성분과 밝기 및 색조 정보를 이용하여 감

성을 인식하는 자동 시스템을 개발하였다^[2]. 주파수 정보 및 컬러 정보가 감성을 인식하기 위해 유용한 정보지만 시스템 평가를 위해 13개의 임의의 텍스타일에 대해서만 감성을 평가하였기 때문에 실제 문제에 적용하기에는 역시 한계를 가지고 있다. 이 외에, 자동으로 추출하는 연구가 진행되어 왔지만, 대부분의 연구에서는 단일 컬러 및 패턴을 가지는 인위적인 영상을 사용하기 때문에, 실제 응용 분야에 적용하는데 문제가 있었다.

그러므로, 우리의 이전 연구에서는 고바야시가 정의한 감성 키워드를 기반으로 컬러 정보와 감성간의 상관관계를 조사하여, 컬러 정보를 이용한 퍼지 기반의 자동감성인식 시스템을 개발하였다^[2]. 이후에 패턴과 감성간의 상관관계를 조사하여 이를 바탕으로 패턴정보를 자동으로 추출하여 인식할 수 있는 시스템을 개발하였다^[4]. 이러한 연구들은 컬러 및 패턴의 특징 정보와 감성간의 상관관계를 조사하여 특정 패턴 및 컬러가 특정 감성과 높은 상관관계를 가지고 있다는 것을 증명하였다. 그러나 이들 단일 특징을 이용하여 수행된 결과 특정 감성만을 정확히 인식하거나 특정 도메인의 텍스타일 영상에서만 적용될 수 있다는 한계를 가지고 있다. 이를 극복하기 위해 복합 특징을 이용한 연구가 수행되어야 한다.

따라서 본 논문에서는 고바야시의 감성 그룹을 기반으로 텍스타일 영상을 자동으로 분류할 수 있는 시스템을 제안한다. 제안된 시스템은 특징 추출과 분류로 구성된다. 특징 추출 단계에서는 텍스타일 영상을 구성하는 대표 컬러를 추출하기 위해서 양자화 기법을 이용하고, 패턴정보를 표현하기 위해서는 웨이블릿 변환 후의 통계적

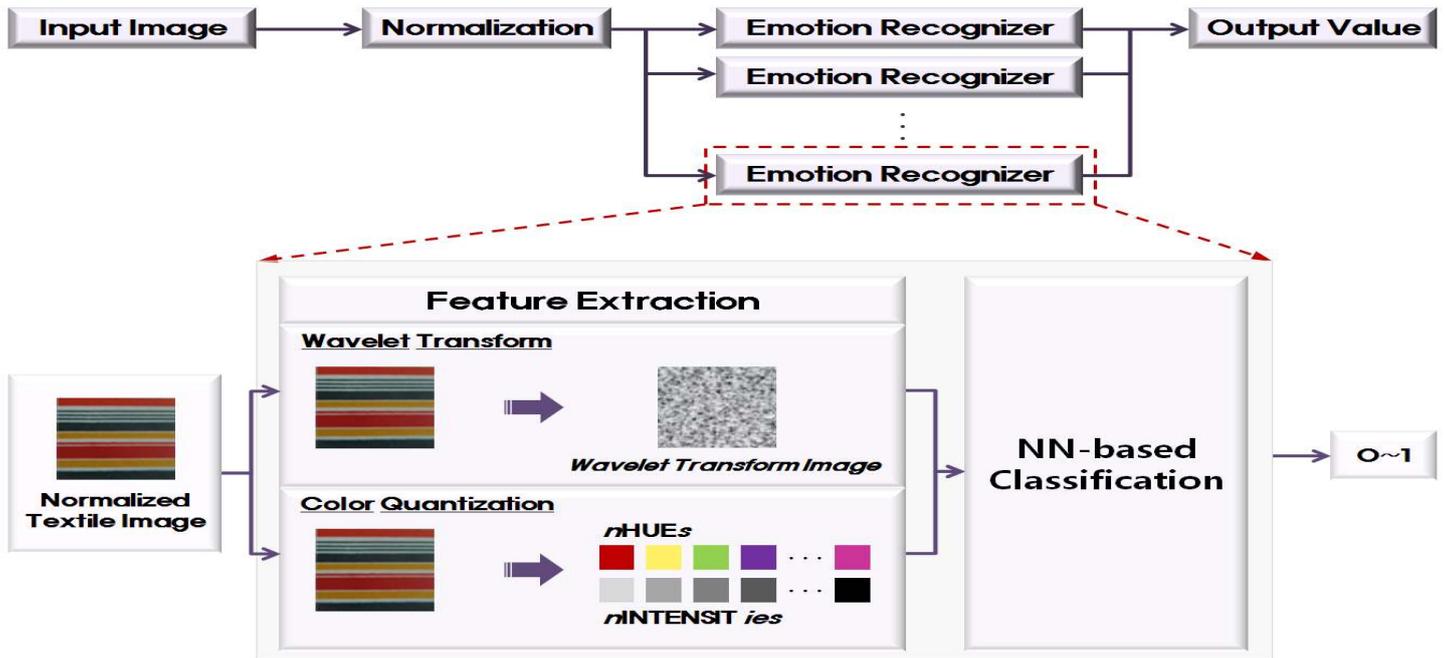


그림 1. 신경망 기반의 개별 감성인식 시스템

인 정보를 이용한다. 신경망 기반의 분류기는 추출된 특징들을 입력으로 받아 입력 텍스타일 영상을 분류한다. 제안된 감성인식 방법의 효율성을 증명하기 위해서 220장의 텍스타일 영상에서 실험한 결과 제안된 방법은 99%의 정확도를 보였다.

본 논문의 구성은 다음과 같다. 2장에서는 제안된 시스템을 소개하고, 실험 결과가 3장에서 설명되며, 4장에서 결론으로 논문을 마무리 한다.

2. 제안된 시스템

본 논문에서 입력된 텍스타일 영상은 고바야시의 10가지 감성 키워드에 따라 분류된다. 여기서 사용된 감성 키워드는 다음과 같다. { romantic, clear, natural, casual, elegant, chic, dynamic, classic, dandy, modern } 텍스타일 영상이 이들 감성 그룹과 정확하게 맵핑되기 위하여 컬러와 패턴 정보를 사용한다. 그림 1은 10개의 감성에 대응된 10개의 개별 텍스타일 영상 분류기를 보여준다. 각 감성 인식기는 특징 추출 부분과 분류기로 구성된다. 텍스타일 영상에서 대표컬러는 양자화 방법을 이용하여 추출되고, 패턴은 웨이블릿 변환 후의 통계적인 분석에 의해 표현된다. 신경망 기반의 분류기는 추출된 특징들을 입력으로 받아 해당 감성을 출력하고 이에 따라 입력 텍스타일 영상을 분류한다.

2.1 컬러 정보추출

본 논문은 고바야시의 이론에 기반하여 입력 텍스타일 영상에서의 대표 컬러 정보를 추출한다^[1]. 감성과 컬러를 맵핑하기 위한 공간인 Color Scale Image는 soft-hard, warm-cool, clear-grayish의 세 축으로 표현된다. 3 가지 기본 감성축은 HSI 컬러 공간에서의 세 가

지 축 색조(hue), 채도(saturation), 명도(intensity)에 각각 대응한다. 따라서 텍스타일 영상을 구성하는 대표 컬러를 추출하기 위해서 먼저 HSI 컬러공간으로 변환한다.

본 논문에서는 명도와 색조 두 정보만을 감성분류를 위한 특징으로 사용한다. 제안된 방법은 양자화기법을 적용하여 I 와 I 채널에서 몇 개의 대표색을 추출한다. 컬러 대푯값은 입력 이미지의 컬러 분포에 대해 사용 빈도수가 가장 높은 n 개로부터 구해지고 이후 다른 모든 컬러들은 가장 가까운 대표색으로 할당된다. 컬러 양자화 결과는 그림 3에서 보인다.

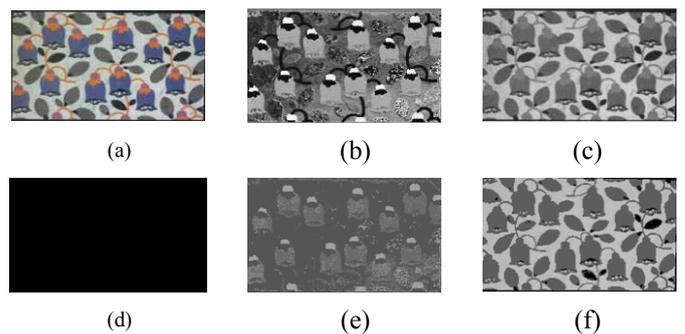


그림 2. 컬러 양자화 결과 : (a) 원본 영상, (b) 색조 변환, (c) 명도 변환, (d) 채도 변환, (e) 양자화된 색조(H) (Q=5), (f) 양자화된 명도(I) (Q=5)

2.2 패턴정보추출

본 논문에서는 패턴 정보의 추출을 위해 웨이블릿 변환을 사용한다. 웨이블릿 변환은 입력영상을 각각 수직, 수평적 특징들을 나타내는 계수들로 구성된 4개의 sub-block LL, LH, HL, HH로 나눈다. LL은 원 영상을 변환 레벨만큼 압축한 것으로 모든 웨이블릿의

sub-block 중에서 가장 중요한 정보를 포함하고 있으며 나머지 sub-block은 수직, 수평, 대각선 방향에 대한 에지 정보를 나타내고 있다 [6].

제안된 방법은 LL레벨에서 웨이블릿 변환을 계속적으로 수행하여 4개씩의 sub-block을 얻을 수 있다. 이는 P번 동안 반복적으로 처리되어 3P+1의 sub-block을 생성한다. 다음의 식을 이용하여 각 sub-block의 3가지 대푯값을 구한다.

$$M(I) = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} I(i, j) \quad (1)$$

$$\mu_2(I) = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (I(i, j) - M(I))^2 \quad (2)$$

$$\mu_3(I) = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (I(i, j) - M(I))^3 \quad (3)$$

영상 N×N이 주어졌을 때, 수식(1)은 평균을 나타내고, 수식 (2)와 (3)은 각각 2차와 3차 momentum을 나타낸다. 이러한 파라미터들은 블록에서 추출되어지고, 3(3P+1)의 파라미터의 합계는 P차 웨이블릿 변환 후에 생성된다.

2.3 분류

본 논문에서는 다층 퍼셉트론(multilayer perceptron, MLP) 기반의 분류기를 사용한다. 제안된 신경망은 입력층, 은닉층, 출력층으로 구성되어 있으며 인접한 층들은 모두 연결되어 있다.

분류기는 3(3P+1)+2Q의 입력을 받는다. 이때 P는 웨이블릿 변환의 레벨이며, Q는 양자화된 컬러들의 개수이다. 다양한 실험을 통해 패턴특징과 컬러특징의 균형을 맞추는 것이 중요하다는 사실을 얻어냈다. 즉, (P, Q)의 최적값은 (3P+1)≈2Q의 하나일 것이다. 따라서 (P, Q)의 값을 (3, 5)로 선택하였다. 본 논문에서는 신경망 가중치의 학습을 위해 텍스타일 영상, I와 설문조사를 통해 그 영상에 색인된 감성 값, d의 쌍으로 구성된 (I, d)를 학습 패턴으로 사용한다.

신경망의 출력 값은 0에서 1사이의 실수로 정규화되며, 0.5보다 크면 해당 감성이라고 판단하고 그렇지 않은 경우에는 해당 감성과 반대 감성을 가진다고 본다.

3. 실험

제안된 방법은 Window XP 환경에서 MS Visual C++를 사용하여 구현하였다. 신경망의 학습 시 중요한 파라미터인 오차율, momentum, 그리고 iteration은 각각 0.001, 0.5, 5000으로 설정된다. 실험에서 사용된 220장의 텍스타일 영상은 패턴북1)으로부터 수집되었다. 제안된 방법의 성능을 평가하기 위해서 220장에 대한 ground truth가

필요하다. 이는 지난 연구의 설문조사 과정을 통하여 각 텍스타일 영상들을 10가지 감성그룹으로 분류할 수 있다 [4]. 표 1은 이러한 데이터 샘플을 보여주며, 총 220장의 실험 데이터 중 120장은 신경망의 학습을 위해 사용되었고, 100장은 테스트에 사용되었다.

표 1. 실험 데이터

	Count	Samples		
Training Data	120			
Test Data	100			

표 2는 제안된 시스템의 성능을 보여주며, 제안된 시스템은 100%의 정확도와 99%의 재현율을 보인다.

표 2. 인위적인 도메인에서의 성능

Emotion type	# of data	# of detected data	False alarm	False dismissal	Recall (%)	Precision (%)
romantic	20	20	0	0	100	100
clear	20	18	1	1	90	100
natural	20	20	0	0	100	100
casual	20	20	0	0	100	100
elegant	20	20	0	0	100	100
chic	20	20	0	0	100	100
dynamic	20	20	0	0	100	100
classic	20	20	0	0	100	100
dandy	20	20	0	0	100	100
modern	20	20	0	0	100	100
Average	200	198	1	1	99	100

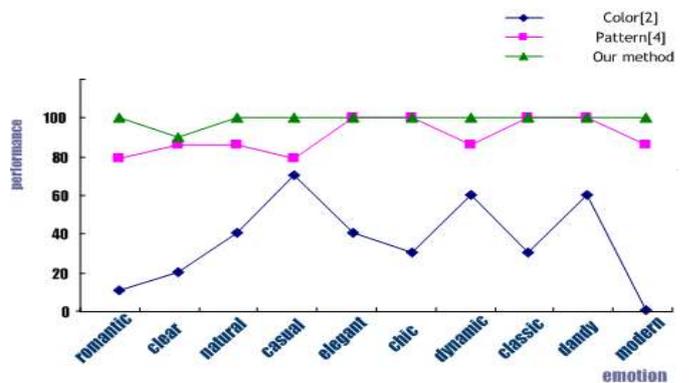


그림 3. 정확도 비교 결과

제안된 시스템의 효율성을 증명하기 위하여 기존의 방법들과 비교되었다. [2]는 컬러 정보 기반의 시스템이며 [4]는 패턴 정보 기반의 시스템으로 단일 특징과 복합 특징을 사용했

1) Meller, Susan, "Textile designs : 200 years of European and American patterns for printed fabrics organized by motif, style, color, layout", Harry N. Abrams, 1991.

을 때의 차이를 알아보기 위하여 선택되었다.

이들의 실험 결과는 그림 3에서 보인다. 실험에서 컬러 정보 기반의 시스템은 36%, 패턴 정보 기반의 시스템은 90%의 정확도를 보였으며, 제안된 시스템은 99%의 정확도로 이들 중 가장 좋은 성능을 보였다. 즉, 이를 통해 복합 특징을 사용하는 것이 단일 특징을 사용했을 때보다 더욱 좋은 성능을 보이는 것을 알 수 있다.

제안된 시스템이 실제 의류, 패션, 디자인 관련 산업에 적용될 수 있는지 알아보기 위해 학습에 사용되지 않은 다른 데이터를 사용하여 실험을 실시했다. 이들 데이터는 총 170장으로 실제 의류 제작에서 사용되는 천들과 벽지, 그리고 카펫 등의 천들을 동대문 광장시장에서 수집하였으며, 학습 데이터보다 복잡한 패턴과 컬러로 조합되어있다.

표 3은 패션 데이터에 대한 결과로 제안된 시스템은 100%의 정확도와 99.74%의 재현율을 보였다.

표 3. 다양한 실제 데이터에서의 성능

Emotion type	# of data	# of detected data	False alarm	False dismissal	Recall (%)	Precision (%)
romantic	37	37	0	0	100	100
clear	37	36	1	0	97	100
natural	34	34	0	0	100	100
casual	36	36	0	0	100	100
elegant	33	33	0	0	100	100
chic	34	34	0	0	100	100
dynamic	36	36	0	0	100	100
classic	34	34	0	0	100	100
dandy	38	38	0	0	100	100
modern	35	35	0	0	100	100
Average	354	353	1	0	99.7	100

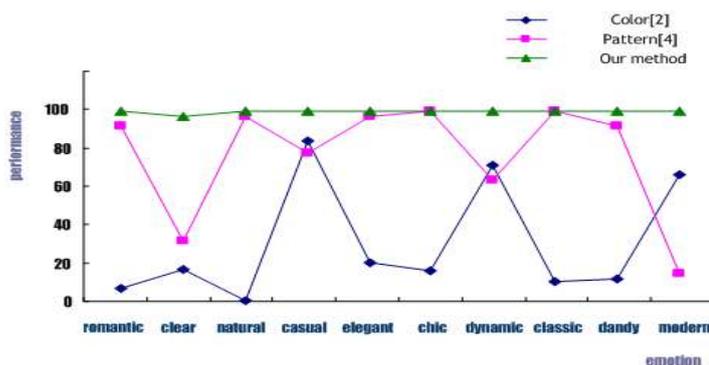


그림 4. 다양한 실제 데이터에서의 성능 비교

그림 4는 다양한 실제 데이터에 대해 제안된 방법과 기존의 연구들의 정확도를 보여준다. [2]의 방법은 30%, [4]의 방법은 62%, 제안된 방법은 99%의 정확도를 보였다. 그림에서 알 수 있듯이 제안된 시스템은 영상의 복잡도가 높음에도 불구하고 여전히 좋은 성능을 보였지만 앞선 실험의 결과와 비교했을 때 기존의 연구들은 급속하게 정확도가 하락했다. 이것은 단일 특징을 사용한 방법이 학습 데이터에 대한 영상은 정확하게 인식하지만, 학습에 사용되지 않았던 데이터에 대해서는 정확하게 인식

할 수 없는 반면 복합 특징 기반의 제안된 방법은 도메인에 상관없이 정확하게 인식할 수 있다. 이러한 실험을 통해 제안된 시스템이 단일 특징을 사용한 다른 시스템보다 더 효율적이라는 것을 보여준다. 따라서 제안된 방법은 다양한 실제적 데이터에 대해 일반적으로 적용될 수 있음을 보여준다.

4. 결론

본 논문에서는 텍스타일 영상에 포함된 컬러와 패턴의 복합 특징 정보를 사용하여 일반화된 감성 인식 방법을 제안하였다. 제안된 방법은 특징추출 부분과 분류 부분으로 구성되었다. 텍스타일 영상으로부터 컬러와 패턴 정보를 기술하기 위하여 컬러 양자화와 웨이블릿 변환을 이용했다. 그리고 신경망을 사용하여 인식된 감성을 기반으로 텍스타일 영상을 분류한다.

제안된 감성인식 방법의 효율성을 증명하기 위한 실험에서 99%의 정확도를 보였다. 이러한 실험 결과를 통해 제안된 방법이 e-Business에 기반을 두는 쇼핑몰의 패션 업체나 섬유산업, 그리고 인테리어 관련 분야, 디자인 관련 분야 등의 많은 응용분야에 적용이 가능함을 보여준다.

참고문헌

- [1] Shigenobu Kobayashi, "COLOR IMAGE SCALE", Kodansha, 1991.
- [2] Eun Yi Kim, Soo-jeong Kim, Hyun-jin Koo, Karpjoo Jeong, Jee-in Kim, "Emotion-based Textile Indexing using Colors and Texture," Lecture Notes in Computer Sciences, vol. 3613, pp. 1077-1080, 2005.
- [3] T. Soen, T. Shimada, and M. Akita, "Objective Evaluation of Color Design ", Color Res. Appl., Vol. 12, pp. 187-194, 1987.
- [4] Na Yeon Kim, Yunhee Shin, Soo-jeong Kim, Jee-in Kim, Karpjoo Jeong, Hyunjin Koo, Eun Yi Kim, "gEmotion Recognition System using Neural networks in Textile images", Jorunal of KIISE: Software and Applications, Vol. 34, No. 9, 2007.
- [5] Soo-jeong Kim, "An Emotion Recognition Framework for Textile Images", Thesis for the Degree of Doctor of Philosophy in Konkuk graduate school, 2007.
- [6] Huiping Li, David Doermann, and Omid Kia, "Automatic Text Detection and Tracking in Digital Video", IEEE Transactions on Image Processing, Vol. 9, No. 1, January 2000.
- [7] 류형주, 채송아, 김수정, 김은이, 김지인, 정갑주, 구현진, "퍼지시스템을 이용한 텍스타일 인텍싱", 한국정보과학회 가을 학술발표논문집 Vol.31, No.2, 2004.
- [8] <http://www.multitherapy.com>.
- [9] 권혜영, 금근하, 김하얀, 이경미, 박우창, 이은옥, "텍스타일 영상의 내용과 메타데이터의 결합을 통한 검색", 한국정보과학회 가을 학술발표논문집 Vol.34, No.2, 2007