

신경망을 이용한 감성인식 시스템

김나연^{1*} 신윤희¹ 김수정¹ 김지인¹ 정갑주¹ 구현진² 김은이¹

¹건국대학교

²한국원사직물시험 연구원

{yeon0830*, ninharsa, cryolite, jnkm, jeongk, eykim}@konkuk.ac.kr¹, hhkoo@fiti.re.kr²

Emotion Recognition System Using Neural Networks

Na-yeon Kim^{1*}, Yunhee Sin¹, Soo-jeong Kim¹, Jee-in Kim¹, Karpjoo Jeong¹, Hyun-jin Koo² and Eun-yi Kim¹

¹Dept. of Internet and Multimedia Eng., Konkuk Univ.

²FITI Testing and Research Institute

요 약

본 논문에서는 신경망을 이용하여 텍스타일 영상으로부터 인간의 감성을 인식할 수 있는 시스템을 제안한다. 자동감성인식 시스템의 구현을 위해 220장의 텍스타일 영상을 수집한 후, 일반인 20명을 대상으로 설문조사를 실시하였다. 이를 통해 감성과 패턴간의 상관관계, 즉 특정 패턴이 사람의 감성에 영향을 준다는 것을 발견하였다. 따라서 본 연구에서는 영상에 포함된 패턴의 인식을 위해 신경망을 이용하였으며, 이때 패턴 정보의 추출을 위해 두 가지 특징 추출 방법을 사용한다. 첫 번째는 auto-regressive method를 이용한 raw-pixel data extraction scheme (RDES)을 사용하는 것이고, 두 번째는 wavelet transformed data extraction scheme (WTDES)을 사용하는 것이다.

제안된 시스템의 유용성을 증명하기 위해서 실제 100장의 텍스타일 영상에서의 감성을 인식하는데 사용했다. 그 결과 RDES와 WTDES를 사용했을 때 각각 71%와 90%의 인식률로, WTDES를 사용했을 때가 RDES를 사용했을 때보다 더 좋은 성능을 보였다. 데이터 추출방법에 따라 다소 차이가 있었지만 전체적으로 약 81%의 정확도를 보였다. 이러한 실험 결과는 제안된 방법이 감성인식 기반으로 텍스타일 데이터를 검색 할 수 있는 시스템 및 다양한 산업 분야에 응용 가능함을 보여주었다.

1. 서 론

감성인식은 기계로 하여금 인간이 느끼는 감성을 자동으로 식별하게 하는 것이다. 감성인식은 최근 다양한 시스템에 결합되어 사용자의 얼굴표정과 제스처 및 음성으로부터 감성을 인식함을 통해 사용자에게 보다 편리하고 직관적인 인터페이스를 제공하기 위한 목적으로 사용되고 있다. 또한 감성인식은 멀티미디어 검색 시스템과 결합하여 사용자에게 내용기반 검색 및 조작을 제공할 수 있다. 즉, 임의의 영상 데이터에 포함된 감성을 자동으로 인식하여 감성을 기준으로 영상을 색인하고, 사용자의 질의에 대해 데이터를 검색하여 서비스 할 수 있게 한다. 특히 텍스타일 영상에 대한 감성인식기반의 검색 시스템은 사용자가 바라는 감성의 천을 자동으로 선택할 수 있게 하기 때문에 사용자의 요구대로 직물의 기획과 생산을 가능하게 한다. 따라서 인터넷 쇼핑몰의 패션업체나 섬유산업, 그리고 디자인 관련분야 등의 많은 응용분야에 적용될 수 있다. 이러한 중요성과 다양한 응용 분야 때문에 최근 많은 연구자들로부터 주목을 받고 있으며, 국내외적으로 연구가 수행되고 있다.

영상에 포함된 감성을 자동으로 인식하기 위해서는 먼저 감성에 영향을 주는 중요한 특징들의 추출이 필요하다. 영상에는 색상 정보 및 텍스처, 그리고 패턴 등의 정보가 있으며 이러한 정보들은 감성에 다양한 영향을 줄 수 있다.

본 논문에서는 임의의 입력된 텍스타일 영상에 포함된 감성을 자동으로 인식하기 위한 신경망 기반의 새로운 감성인식 시스템을 제안한다. 제안된 방법의 기본 가정은 영상에 나타나는 특정 패턴이 사람의 특정 감성에 많은 영향을 줄 수 있다는 것이다. 이러한 가정의 정당성을 증명하기 위해서 먼저 패턴과

감성과의 상관관계를 조사하였고, 그 결과를 통해 특정 패턴이 몇 개의 감성과 높은 상관관계를 가지고 있다는 것을 증명하였다. 이러한 사실을 바탕으로 본 논문에서는 자동 감성인식 시스템을 개발한다.

본 논문에서는 영상에 포함된 패턴의 인식을 위해 신경망을 이용한다. 신경망의 입력을 위해 두 가지 추출 방법을 사용하여 텍스타일 영상으로부터 특징을 추출한다. 하나는 auto-regressive method를 이용한 raw-pixel data extraction scheme (RDES)을 사용하는 것이고, 다른 하나는 wavelet transformed data extraction scheme (WTDES)을 이용하여 생성된 sub-block들의 통계적인 특징을 입력으로 사용하는 것이다. 전자는 입력 영상에서 63×63 윈도우 내에 있는 raw-pixel data 중 중요한 화소만을 신경망의 입력으로 사용하는 것이고, 후자는 64×64의 입력 영상에 6차 웨이블릿 변환을 적용하여 각 sub-block들로부터 3가지의 대푯값을 추출하여 사용하는 것이다. 각각의 특징 추출 방법을 통해 역전파 알고리즘(back-propagation algorithm: BP)을 이용하여 신경망을 학습한다.

특징 추출 방법에 따라서 성능의 차이가 있기는 하지만 평균적으로 81%의 인식률을 보여주었으며, 이는 제안된 인식 시스템이 산업 분야 및 감성 기반의 검색 서비스 시스템에 응용 가능함을 보여주었다.

2. 데이터 수집 및 분석

제안된 감성인식 시스템을 구현하기 위해서 먼저 텍스타일 영상에 포함된 패턴과 감성과의 관계에 대한 설문 조사를 실시하였다. 이를 위해 먼저 패턴북 으로부터 텍스타일 영상 데이

터를 수집하였다.

수집된 텍스타일 영상들은 영상에 포함된 패턴의 복잡도에 따라 크게 단일패턴과 복합패턴으로 분류된다. 단일패턴은 다시 형태에 따라서 '도형', '선', 그리고 '식물'로 분류되고, 각 단일 패턴은 다시 세부 항목으로 분류된다. 그 결과 총 9개의 그룹 ({네모, 세모, 원, 가로선, 세로선, 체크선, 곡선, 꽃, 잎})으로 텍스타일 영상은 분류된다. 복합패턴은 이러한 기본패턴 중 두개 이상의 결합으로 구성된다. 이렇게 분류된 데이터들에 대해서 10가지 감성 기준 ({romantic, clear, natural, casual, elegant, chic, dynamic, classic, dandy, modern})에 따라 설문조사를 실시하였다. 설문조사는 패턴에 따라 분류된 총 220장의 영상에 대해 일반인 20명을 대상으로 실시되었다. 이 중 120장은 신경망의 학습을 위해, 나머지 100장은 시스템의 성능을 실험하기 위한 테스트 데이터로 사용되었다.

이렇게 얻은 설문조사 결과를 토대로 9개의 단일패턴과 10가지 감성과의 상관관계를 조사하기 위해 그림 1(a)~(c)와 같이 히스토그램으로 표현했다. 각 히스토그램은 임의의 감성과 패턴과의 관계를 나타내고, 가로축은 감성 값을, 세로축은 각 감성 값에서의 설문 조사자들의 응답 빈도수를 나타낸다.

먼저, 그림 1(a)는 '꽃' 패턴과 감성 'modern'과의 관계를 보여주는 히스토그램이다. 그림에서 설문조사자의 응답이 대부분 -11에서 -18 사이에 분포되어 있는 것을 볼 수 있는데, 이는 '꽃'패턴이 주어진 감성 'modern'과 반대의 감성을 가지고 있다는 것을 뜻한다. 이와 달리, 그림 1(c)는 설문조사자의 응답이 대부분 6에서 15사이의 값을 가지므로 '페이즐리'가 'dynamic'한 감성을 가지는 패턴임을 나타낸다. 마지막으로 그림 1(b)는 0을 중심으로 분포되어 있는데, 이는 해당 패턴과 조사 감성과의 상관관계가 없음을 나타낸다.

이러한 히스토그램 분석 과정을 통해 각각의 텍스타일을 감성 값에 따라 분류하였으며, 이로부터 특정 패턴과 몇 가지 감성 그룹간의 상관관계의 유무를 조사하였다. 그림 1(d)는 패턴과 감성간의 상관관계를 트리구조로 도식화한 것이다.

그림 1(d)에서 보이듯이 각 도형마다 특정 감성이 매핑되어 있다. 예를 들어, '네모'와 '가로선' 패턴은 'dandy'를 주 감성으로 갖고 있고, '세모'와 '곡선' 그리고 '잎'의 패턴은 'dynamic'을 주 감성으로 갖고 있다.

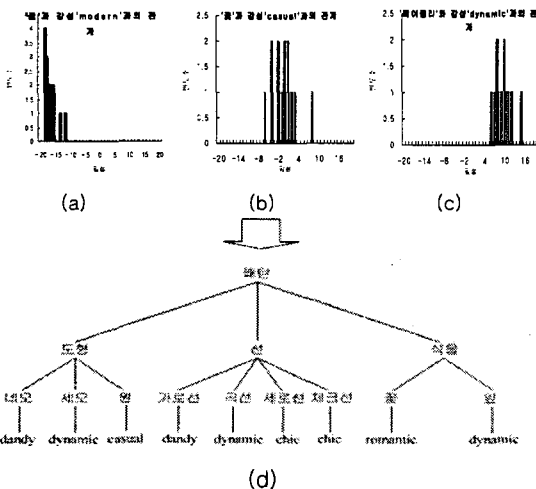


그림 1. 패턴과 감성간의 상관관계 (a)꽃 vs. modern (b)꽃 vs. casual (c)페이즐리 vs. dynamic (d)패턴과 감성간의 상관관계 그래프

3. 제안된 방법

설문 조사 분석을 통해 텍스타일 영상의 포함된 감성은 그 영상을 대표하는 패턴 정보에 많이 의존한다는 사실을 알 수 있었다. 이러한 패턴 정보를 인식하기 위해서 본 시스템에서는 신경망을 이용한다. 신경망의 입력으로는 크게 두 가지 방법을 사용하는데 하나는 RDES를 사용하는 것이고, 다른 하나는 WTDES를 사용하는 것이다.

3.1 특징 추출

3.1.1 Raw-pixel data extraction scheme (RDES)

Raw data는 말 그대로 가공되지 않은 데이터로써 원 영상 그대로를 사용함을 뜻하며 raw-pixel data는 텍스처를 표현하는 가장 쉽고 보편적인 방법으로 많이 사용되고 있다.

하지만 주어진 입력 윈도우 $M \times M$ 내에 모든 화소 값을 신경망의 입력으로 사용하는 것은 입력 노드의 개수를 지나치게 크게 만들기 때문에 신경망 학습 및 테스트의 속도를 느리게 만드는 주된 요인이 된다. 따라서 입력 노드 개수를 줄이기 위해 본 논문에서는 autoregressive method를 사용하여 M^2 의 노드 개수를 $4M-3$ 개로 줄여 사용하기로 한다.

3.1.2 Wavelet transformed data extraction scheme (WTDES)

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 예지 정보를 나타내고 있다.

본 논문에서는 64×64 의 입력 영상에 대해서 6차 웨이블릿 변환을 적용한다. 그 결과 생성된 총 24개의 sub-block에 대해서 각각 다음의 식을 이용하여 3가지 sub-block의 대표값을 추출한다.

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

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

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

영상이 $N \times N$ 이 주어졌을 때, 수식(1)은 평균값을 나타내고, 수식(2)와 (3)은 각각 2차와 3차 momentum을 나타낸다. 6차 웨이블릿 변환에서는 총 24개의 sub-block이 생성되기 때문에 64×64 영상에 대해 72개의 특징 값이 나오며 이들은 신경망의 입력으로 사용된다.

이 특징 추출 방법은 RDES와 비교했을 때, 보다 적은 입력 노드의 수를 사용하기 때문에 신경망 학습 및 테스트에 필요한 시간을 감소할 수 있다.

3.2 신경망 기반의 인식

그림 2는 10개의 감성에 대응된 10개의 개별 감성 인식 시스템을 보여준다. 각 감성 인식기는 입력된 영상에 해당 감성의 존재 여부에 따라 텍스타일 영상을 분류한다. 각 감성 인식기는 사용된 특징 추출 방법에 따라서 두 가지로 나누어지며 이들은 특징 추출 방법의 우수성을 비교하기 위해서 사용된다.

주어진 시스템은 먼저 입력된 영상을 64×64 의 크기로 정규

화 한다. 그 후 두 가지 특징 추출 방법을 사용하여 신경망의 입력될 특징 값을 추출한다.

첫 번째 RDES를 이용하여 신경망의 입력 값을 추출하는 경우에는 63×63의 윈도우 크기가 주어졌을 때, 총 249개의 입력노드와 1개의 출력노드로 구성된 신경망을 이용한다. 이 때, 은닉노드의 수는 실험을 통해 정해진다.

두 번째 WTDES일 경우에는 6차 웨이블릿 변환 시 생성된

24개의 sub-block으로부터 총 72개의 특징 값이 나오므로 72개의 입력노드와 1개의 출력노드로 구성된 신경망을 인식기로 이용하고, 이 경우에도 은닉노드의 수는 실험을 통해 정해진다.

신경망의 출력 값은 0에서 1 사이의 실수로 정규화 되며, 0.5보다 크면 해당 감성이 있다고 판단하고 그렇지 않은 경우에는 해당 감성이 없다고 본다.

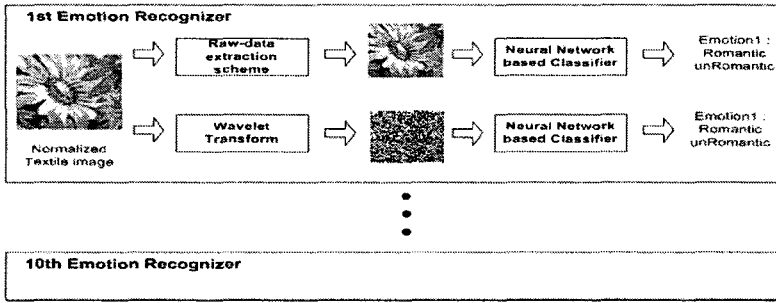


그림 2. 신경망 기반의 개별 감성 인식 시스템

4. 실험 및 결과

본 논문에서는 신경망 학습을 위해 오차율과 momentum, 그리고 iteration을 각각 0.001, 0.5, 5000으로 설정하고, 제안된 시스템의 평가를 위해 설문조사로부터 얻어진 100개의 데이터를 사용하였다.

표 1. 개별 감성 인식 시스템의 성능 분석 (%)

	raw-pixel data(27×27)		raw-pixel data(63×63)		wavelet transform	
	recall	precision	recall	precision	recall	precision
ROMANTIC	57	56	79	100	79	100
CLEAR	79	100	79	100	86	100
NATURAL	64	100	71	80	86	100
CASUAL	57	100	64	63	79	100
ELEGANT	50	0	64	50	100	100
CHIC	71	80	64	67	100	100
DYNAMIC	79	100	79	75	86	100
CLASSIC	71	100	71	71	100	100
DANDY	64	100	71	71	100	100
MODERN	79	83	64	63	86	78
AVERAGE	67	82	71	74	90	98

표 1은 제안된 감성인식 시스템의 전체적인 정확도를 보여주며, 성능 분석을 위해 precision과 recall, 이 두 가지 평가지표를 사용한다.

감성인식 시스템은 크게 두 가지 데이터 추출 방법에 의해서 구분 되는데, WTDES를 사용한 결과가 RDES를 사용한 결과보다 더 좋은 결과를 보여주는 것을 알 수 있었다. 그리고 RDES를 사용했을 경우에는 윈도우 크기가 커질수록 조금 더 좋은 결과를 보여주었지만, 그 차이가 그렇게 크지 않았다. 이러한 실험결과로부터 우리가 알 수 있는 것은 RDES를 사용한 경우 윈도우 크기가 커짐에 따라 정확도의 개선은 보이지만 그에 따른 계산시간이 늘어날 때 때문에 정확도와 속도 사이의 trade-off가 발생한다. 반면, WTDES를 사용하였을 경우에는 더 높은 인식률을 보일 뿐만 아니라 더 적은 입력 노드를 사용함으로써 계산 시간을 줄일 수 있어 속도 및 정확도면에서 더 우수한 성능을 보여준다. 이러한 성능 차이는 WTDES에서 각

sub-block내 입력 화소의 색상 정보만 이용한 것이 아니라 수직, 수평 방향의 에지 성분들을 입력으로 사용하고, 또한 6차 웨이블릿 변환을 사용함으로써 압축된 영상의 전체적인 특징을 사용하였기 때문이다.

5. 결론

본 논문에서는 패턴과 감성의 상관관계로부터 감성의 인식은 패턴 인식으로 인해 수행될 수 있음을 밝혔고, 영상에 포함된 패턴을 인식하기 위해서 신경망 기반의 분류기를 개발하였다.

개발된 시스템은 전체적으로 좋은 성능을 보였지만, 몇 가지 특정 감성을 인식하는데 어려움을 보여주었다. 이러한 문제점은 이산 코사인 변환(Discrete Cosine Transform: DCT)을 이용하여 주파수 성분을 패턴 정보로 표현하는데 사용한다면, 쉽게 해결 될 것으로 보인다. 따라서 현재 새로운 특징을 검출해 내는 방법을 개발 중이며, 또한 단일 특징 추출법만이 아니라 여러 가지 추출 방법을 통합하여 사용하도록 시스템을 확장하고 있다.

참고 문헌

- [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.
- [2] Shigenobu Kobayashi, "COLOR IMAGE SCALE", Kodansha, 1991.
- [3] T. Soen, T. Shimada, and M. Akita, "Objective Evaluation of Color Design II", Color Res. Appl., Vol. 12, pp. 187-194, 1987.
- [4] 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.
- [5] 류형주, 채송아, 김수정, 김은어, 김지연, 정갑주, 구현진, "퍼지시스템을 이용한 텍스타일 인덱싱", 한국정보과학회 가을 학술발표논문집 Vol.31, No.2, 2004.