

논문 2014-51-6-11

Signed Local Directional Pattern을 이용한 강력한 얼굴 표정인식

(Robust Facial Expression Recognition Based on Signed Local Directional Pattern)

류 병 용*, 김 재 면*, 안 기 옥*, 송 기 훈*, 채 옥 삼**

(Byungyong Ryu, Jaemyun Kim, Kiok Ahn, Gihun Song, and Oksam Chae[Ⓢ])

요 약

본 논문에서는 얼굴 표정인식을 위한 새로운 지역 미세 패턴 기술 방법인 Signed Local Directional Pattern(SLDP)을 제안한다. SLDP는 얼굴 영상의 텍스처 정보를 표현하기 위해 에지 정보를 이용한다. 이는 기존의 방법들에 비해 뛰어난 구별 성능과 효율적인 코드 생성을 가능하게 한다. SLDP는 마스크 범위 이웃 화소들을 이용하여 에지 반응 값을 계산하고 이들 중 부호를 고려하여 에지 반응 값이 큰 에지 방향 정보를 가지고 만들어진다. 이는 기존 LDP에서 구별하지 못하던 비슷한 에지 구조에 밝기 값이 반대인 지역 패턴을 구별할 수 있다. 본 논문에서는 얼굴 표정인식을 위해 얼굴 영상을 여러 영역으로 분할하고 각 영역으로부터 SLDP코드의 분포를 계산한다. 각 분포는 얼굴의 지역적인 특징을 나타내고 이들 특징을 연결해서 얼굴 전체를 나타내는 얼굴 특징 벡터를 생성한다. 본 논문에서는 생성된 얼굴 특징 벡터와 SVM(Support Vector Machine)을 이용해서 Cohn-Kanade 데이터베이스와 JAFFE 데이터베이스에서 얼굴 표정인식을 수행했다. SLDP는 표정인식에서 기존 방법들보다 뛰어난 결과를 보여주었다.

Abstract

In this paper, we proposed a new local micro pattern, Signed Local Directional Pattern(SLDP). SLDP uses information of edges to represent the face's texture. This can produce a more discriminating and efficient code than other state-of-the-art methods. Each micro pattern of SLDP is encoded by sign and its major directions in which maximum edge responses exist—which allows it to distinguish among similar edge patterns that have different intensity transitions. In this paper, we divide the face image into several regions, each of which is used to calculate the distributions of the SLDP codes. Each distribution represents features of the region and these features are concatenated into a feature vector. We carried out facial expression recognition with feature vectors and SVM(Support Vector Machine) on Cohn-Kanade and JAFFE databases. SLDP shows better classification accuracy than other existing methods.

Keywords : facial expression recognition, signed local directional pattern, support vector machine

* 정회원, ** 평생회원, 경희대학교 컴퓨터공학과
(Computer Engineering, Kyung Hee University)

Ⓢ Corresponding Author(E-mail: oschae@khu.ac.kr)

※ 본 연구는 중소기업청의 기술혁신개발사업의 일환
으로 수행하였음. [SA113709, 사용자 표정인식 기반
의 3차원 아바타 생성 기술 개발]

접수일자: 2014년03월03일, 수정일자: 2014년05월06일
수정완료: 2014년05월28일

I. 서 론

최근 사람의 감정을 인식하고 응용하는 분야가 많아
지고 있다. 감정을 인식하고 분석하는 응용 분야는 제
품을 차별화 하고 사용자의 관심과 구매를 유도 할 수
있는 잠재력이 있기 때문에 이에 관한 연구가 활발하게

진행되고 있다^[3-4]. 최근에는 감정을 인식하는 방법 중 하나로 얼굴 표정 인식이 많이 연구되고 있다^[1-2].

얼굴 표정인식은 크게 기하학적 특징 기반 방법(Geometric feature based method)과 외형 특징 기반 방법(Appearance feature based method) 2가지가 있다^[6]. 기하학적 특징 방법은 얼굴의 구성요소(눈, 코, 입, 입술, 눈썹 등)들의 위치와 이들 사이의 관계를 이용하여 얼굴의 형태와 표정의 변화를 표현하는 방법이다. 이러한 방법은 얼굴 특징을 소수의 인자로 간단하게 표현이 가능하고 위치, 크기, 방위 변화에 쉽게 대응이 가능하지만, 얼굴의 구성요소를 정확하게 검출하는 것이 매우 어려운 문제점이 있다^[5].

외형 특징 기반 방법(Appearance feature based method)은 전체적인 또는 지역적인 얼굴 영역을 대상으로 얼굴 영상의 특징을 표현한다. 이때 사용되는 영상의 크기에 따라서 전역특징기반 방법(Holistic image based approach)^[7-11]과 지역특징기반 방법(Local image based approach)로 구분된다. 전역특징기반 방법은 얼굴 영상을 구성하는 각 화소를 하나의 특징 요소(feature element)로 보고 영상 전체를 하나의 특징벡터로 표현한다. 이러한 방법은 별도의 특징추출과정을 거치지 않고 영상 전체를 대상으로 하기 때문에 처리방법이 간단하지만 얼굴 영상의 표정변화는 얼굴 영상의 일부분에 나타난다는 특징을 수용하지 못하기 때문에 표정인식에서 성능이 떨어지는 문제점이 있다. 반면 지역특징기반 방법은 얼굴 영상의 표정 변화가 지역적으로 나타내는 특징을 수용할 수 있도록 변화가 일어날 수 있는 영역에 지역적인 윈도우를 설정하고 윈도우 내의 밝기 분포를 바탕으로 특징을 추출하여 전체적인 얼굴의 형태를 나타내는 특징 벡터를 구성한다. 지역특징기반 방법은 표정 변화의 특징을 수용했기 때문에 전역특징 기반 방법보다 표정인식에서 안정적이고 좋은 얼굴 영상 분석 결과를 보여준다. 따라서 최근 지역특징기반 방법의 얼굴 표정인식 연구가 가장 활발하게 진행 되는 중이다.

지역특징기반 얼굴 표정인식의 일반적인 방법은 다음과 같은 단계를 거쳐 수행 된다. 첫 번째 얼굴 영상을 입력 받는다. 입력 받은 영상은 얼굴 영역만 포함하고 있다고 가정한다. 입력 받은 얼굴 영상을 지역적으로 분석하기 위해서 균일한 블록으로 분할한다. 분할된 각 블록내의 화소들은 3x3, 5x5, 7x7 같은 마스크 범위 이

웃화소를 사용하여 패턴을 부호화 한다. 각 블록마다 블록에 포함된 화소 숫자만큼 패턴이 계산되고 계산된 패턴을 하나의 히스토그램에 누적한다. 이와 같은 과정을 거치면 블록 숫자만큼 히스토그램이 생성되고 각 히스토그램은 하나의 지역 패턴을 표현하게 된다. 지역특징기반 방법은 이렇게 만들어진 지역 패턴을 연결해 얼굴 영상 전체를 표현하는 특징 벡터를 만든다. 얼굴 표정 인식 시스템은 이와 같이 생성된 특징 벡터와 분류 방법을 이용하여 얼굴 표정인식을 수행한다.

위와 같은 지역특징기반 얼굴 표정인식 단계에서 가장 중요한 부분은 각 화소들의 패턴 부호화 방법이다. 패턴 부호화 방법에 따라 얼굴 표정변화에 따른 지역적인 얼굴 특징을 올바르게 표현 가능하고 얼굴 표정인식 성능에 결정적인 영향을 줄 수 있다. 따라서 강력한 패턴 부호화 방법에 초점을 맞춘 많은 연구들이 수행되었다. 이러한 연구 중 가장 널리 사용되고 대표적인 알고리즘은 LBP(Local Binary Pattern)이다^[12-14]. LBP는 얼굴 영상이 조명이나 촬영 환경에 따라서 전체적인 밝기 값이 변한다는 문제점을 극복하기 위해 이웃화소의 밝기 값 관계를 부호화 한다. 이러한 LBP부호화 방법은 전체적인 밝기 값 변화에는 강력하고 미세한 부분까지 패턴으로 표현이 가능하기 때문에 얼굴인식과 texture 패턴 인식에 좋은 성능을 보여준다. 하지만 임의의 잡음이 포함되는 경우 LBP는 코드의 일관성을 유지할 수 없는 문제점이 있다^[15-16].

LBP의 잡음에 취약한 문제점을 해결하기 위해서 Gabor filter^[11]가 적용된 영상에 LBP를 적용하는 LGBPHS(Local Gabor Binary Pattern Histogram Sequence)^[24] 방법이 제안 되었다. LBP의 잡음 취약성의 문제를 다소 해결 했지만 과도한 계산량과 세밀한 패턴의 표현력이 떨어지는 점이 문제가 되고 있다. 최근 좋은 결과를 보여주는 LDP(Local Directional Pattern)^[25]방법은 이웃의 모든 화소를 사용해서 에지 반응 방향 정보를 부호화했기 때문에 LBP에 비해 잡음에 강력한 특징을 가진다. 또한 에지 반응 방향 정보는 표정 변화로 생기는 주름, 얼굴 특징의 변화를 밝기 값 관계 보다 정확하게 표현 가능하기 때문에 LDP는 LBP보다 표정인식에서 상대적으로 좋은 결과를 제공한다. 하지만^[25]에서 제안한 LDP는 부호를 고려하지 않고 에지 반응 정보를 부호화했기 때문에 일부 패턴을 정확하게 표현하지 못하는 문제점이 있다.

본 논문에서는 위와 같은 부호를 고려하지 않은 LDP의 문제점을 극복하기 위해 에지 반응 정보를 이용할 때 부호를 고려하는 새로운 방법으로 Signed Local Directional Pattern(이하, SLDP)을 제안하고 얼굴 표정 인식 성능을 분석하였다. SLDP를 이용한 얼굴 표정 인식의 과정은 지역특징기반 얼굴 표정인식과 같이 얼굴 영상을 균일한 블록으로 분할한 뒤 각 블록내의 화소들을 제안된 SLDP를 이용해서 패턴 부호화 한다. 각 블록의 히스토그램을 구하고 이를 연결해서 얼굴 전체를 표현하는 히스토그램을 만들어 얼굴 전체를 나타내는 얼굴 특징 벡터를 만든다. 이렇게 만들어진 특징 벡터와 SVM을 이용해서 얼굴 표정인식을 수행했다. 이때 정확한 얼굴 표정인식 비교 검증을 위해서 널리 사용되는 Cohn Kanade (CK) 얼굴 표정 데이터베이스^[17]와 Japanese female facial expression (JAFFE) 데이터베이스^[11]를 사용했다. 이를 통해 제안된 SLDP가 LBP와 LDP보다 강력한 얼굴 특징 추출이 가능하고 뛰어난 얼굴 표정인식 성능을 보여주는 방법임을 확인하였다.

본 논문의 구성은 다음과 같다. II장에서는 이전 연구인 LBP와 LDP의 패턴 생성 방법을 간략하게 설명한다. III장에서는 제안된 SLDP를 이용한 특징 벡터 생성 방법을 설명하고 IV장에서는 SLDP를 이용한 얼굴표정 인식에 방법에 대해 소개한다. V장에서는 얼굴 표정인식 실험을 수행하고 VI에서는 결론으로 끝을 맺는다.

II. 이전 연구

1. Local Binary Pattern(LBP)

가. 기본적인 LBP 개념

기본적인 LBP^[15]는 그림 1과 같이 3x3크기 윈도우의 8개의 이웃 화소들을 중앙 화소를 임계값으로 이진화하여 얻어지는 비트 패턴을 순차적으로 나열함으로써 얻

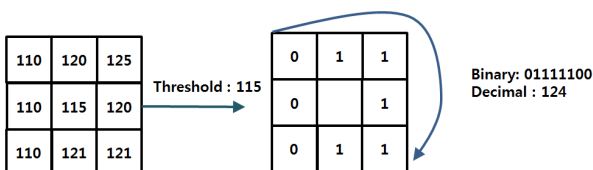


그림 1. 기본 LBP 연산자
Fig. 1. Basic LBP operator.

어진다.

나. 다양한 크기의 LBP 코드

기본적인 LBP 코드 생성 방법의 문제점은 3x3 이웃 화소가 수용하기 어려운 크기가 큰 구조를 가진 특징은 정확하게 부호화 할 수 없는 점이다. 크기가 다른 특징에 대응하기 위해서 기본적인 LBP코드는 크기가 다른 이웃 화소들을 사용 할 수 있는 방법으로 확장 되었다^[15]. 일반화된 LBP는 수식 (1) 과 같다. 여기서 R 은 이웃 화소를 형성하는 원의 반지름을 나타내고 P 는 사용할 이웃 화소의 숫자를 나타낸다. g_c 는 임계치로 사용될 중심 화소의 화소값이고 g_p 는 이웃 화소의 화소값을 나타낸다. 그림 2는 여러 가지 P 와 R 에 대한 여러 가지 이웃 화소 예를 보여준다. 그림 2에서 (8, 1), (16, 2), (24, 3)은 (P, R)값을 나타낸다. P 와 R 값에 따라서 정의된 이웃 화소들은 정확하게 입력 영상의 화소위에 위치하지 않을 수도 있다. 이런 경우 양선형 보간법 (bilinear interpolation)을 사용해서 이웃 화소 값을 계산한다.

$$LBP(P, R) = \sum_{p=0}^{P-1} s(g_p - g_c)2^p, \quad (1)$$

$$s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

다. 균일국부이진패턴 LBP(Uniform LBP)

LBP 코드(8, 1)는 256가지 패턴을 표현 할 수 있다.^[15]에서는 이러한 패턴 중 특정패턴이 다른 패턴보다 많은 정보를 가진 다는 것을 보여주었다. LBP코드를 비트로 표현 했을 때 순환을 고려 0에서 1로 또는 1에서 0으로 변하는 숫자가 2이하인 경우를 균일국부이진패턴 (uniform patterns)이라고 정의한다. 예를 들어 00000000(0변이), 01000000(1변이)는 균일국부이진패턴이고 11010100(6변이), 00110110(4변이)는 균일국부이진패턴이 아니다. 이들 균일국부이진패턴은 다른 코드들에 비해서 발생 빈도가 높고 많기 때문에 많은 정보를 나타낼 수 있다. 한 실험에서 얼굴 인식 및 표정인식에 널리 쓰이는 FERET 데이터베이스^[26]를 이용해서 균일국부이진패턴 발생 빈도를 조사 했다. 이때 (8, 1)의 경우 LBP 코드의 90.6% (8, 2)의 경우 LBP코드의 85.2%가 균일국부이진패턴 이었다^[20]. (8, 1)의 경우 256가지

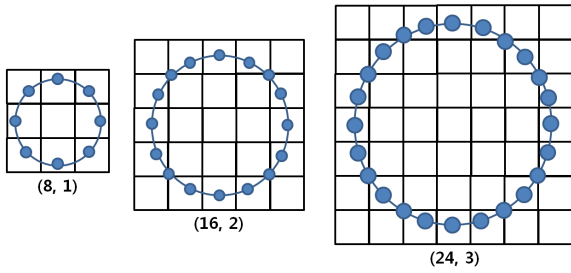


그림 2. P와 R값에 따른 이웃 화소 예 (8, 1), (16, 2), (24, 3)

Fig. 2. Examples of neighborhoods according to P and R (8, 1), (16, 2), (24, 3).

패턴 중 58가지가 균일국부이진패턴으로 나타나고 이를 얼굴 표정 영상을 분석을 위해 사용할 때 균일국부이진 패턴은 58가지로 각각 따로 표현하고 나머지 코드는 동일한 코드 하나로 표현해서 사용한다. 이렇게 만들어진 균일국부이진패턴은 59가지 패턴을 가지게 되고 이는 256가지 패턴에 비해 훨씬 적은 숫자로 비슷한 효율을 나타내는 패턴 표현이 가능하다.

라. LBP의 장·단점

이와 같이 생성된 LBP코드는 단조로운 밝기 변화에 강력 특성을 가진다. 또한 LBP코드는 코드 생성의 단순함과 강력한 성능으로 texture 분석, 얼굴 인식, 얼굴 표정인식 등 다양한 분야에 널리 사용된다. 하지만 LBP는 코드 생성시 1비트 값 생성에 오직 이웃화소 하나와 중심 화소만 사용하기 때문에 잡음에 매우 취약하고 단조롭지 않은 밝기 변화에 코드가 많이 변하게 된다.^[15-16]

2. Local Directional Pattern(LDP)

가. LDP의 등장 배경

에지 반응 값은 주변 이웃화소를 모두 사용하여 계산되는 특징 때문에 화소의 밝기 값 보다 밝기 변화와 잡음에 덜 민감하고 에지 반응 값을 이용해서 지역 미세 패턴을 기술하면 좀 더 안정적이고 많은 정보를 코드에 담을 수 있는 이점이 있다. 이와 같은 사실을 기반으로 에지 정보를 이용하여 지역 미세 패턴을 기술하는 LDP^[25]가 제안되었다. LDP는 LBP의 단점을 극복하고 지역 미세 패턴을 좀 더 정확하게 기술할 수 있는 장점을 가진다.^[19]

나. LDP 코드 생성 방법

LDP는 각 화소에 대해서 m x m 지역 패턴 윈도우를 설정하고 여기에 8 방향 m x m Kirsch 에지 반응 마스크^[18]를 이용하여 8방향 에지 반응 값을 구한다. 이때 m의 변화에 따라서 표현 가능한 패턴에 다양한 변화가 가능하다. 그림 3은 8개의 3 x 3 Kirsch 에지 반응 마스크의 예이다. 각 8 방향 에지 반응 중 동쪽부터 반 시계 방향으로 각 방향을 나타내는 비트를 설정한다. 이때 동쪽 방향(m₀)이 LSB(Least Significant Bit)가 되고 남서 방향(m₇)이 MSB(Most Significant Bit)가 되도록 설정한다. 8 방향의 모든 에지 반응 값이 유의미한 값을 가지는 것은 아니다. 따라서 계산된 8 방향 에지 반응 값을 절대값 크기에 따라 정렬하고 상위 k개를 선정 해당 방향을 나타내는 비트를 1로 설정하고 나머지 방향을 나타내는 비트는 0으로 설정한다. 그림 4는 m과 k가 각각 3일 경우, 주어진 3 x 3 지역패턴에 대한 LDP 생성 과정을 보여준다. m₄, m₁, m₀의 방향의 에지 반응 크기가 1, 2, 3위로 상위 3개의 반응 크기를 나타내고 이 방향을 나타내는 비트 4, 1, 0이 1로 설정되어 LDP 코드 00010011(19)가 생성된다.

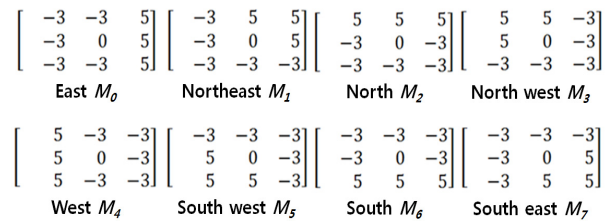


그림 3. 3x3 8방향 Kirsch 에지 반응 마스크
Fig. 3. 3x3 Kirsch edge masks in eight directions.

다. 잡음에 강력한 LDP

에지 반응 값은 화소 밝기 값보다 안정적인 값을 가지기 때문에 LDP는 일부 노이즈가 발생하거나 단조롭지 않은 밝기 변화에도 같은 패턴을 유지한다. 그림 5는 영상 원본과 가우시안 백색 잡음을 적용한 영상에서의 LDP코드와 LBP코드의 차이를 보여준다. LBP는 가장 오른쪽 비트가 1에서 0으로 변했지만 LDP는 같은 코드를 보여준다^[25]. 이는 에지 반응 값을 계산하기 위해서 이웃 화소의 값을 모두 사용하기 때문에 잡음을 강력한 모습을 보여준다.

162	180	175
193	20	27
190	20	25

Mask index	M_7	M_6	M_5	M_4	M_3	M_2	M_1	M_0
Mask Value	-2340	-1036	308	1444	1364	1220	140	-1100
LDP Rank	1	6	7	2	3	4	8	5
LDP bit	1	0	0	1	1	0	0	0
LDP decimal	152							

그림 4. LDP 코드 생성 과정
Fig. 4. The process to generate LDP code.

라. 얼굴 표정인식을 위한 LDP 코드 최적화

LBP의 균일국부이진패턴과 비슷한 개념으로 LDP는 에지반응 크기가 큰 상위 k개의 방향만 부호화 하는 방법으로 코드의 길이를 줄여주고 효율성을 높여 줄 수 있다. 이때 k의 숫자에 따라서 LDP코드의 길이가 달라진다. 예를 들면, 8개의 방향에서 크기가 큰 1개의 방향만(k=1) 부호화 하면 8가지 코드가 생성되고 8개의 방향에서 크기가 큰 2개의 방향만(k=2) 부호화 하면 28가지(8개에서 2개를 뽑는 조합을 이용 계산 가능)의 코드가 생성된다. 따라서 얼굴 표정 인식을 수행하는데 가장 적합한 k값을 찾기 위한 실험이 [25]에서 수행되었다. 여러 가지 k값을 이용해서 실험한 결과 k가 3일 때 최적의 결과를 보여주었다. 또한 LDP를 얼굴 표정인식에 적용 했을 때 LBP에 비해서 안정적이고 뛰어난 결과를 보여주었다.

85	32	26	LBP 10000011 LDP 00010011
53	50	10	
60	38	45	

81	29	32	LBP 10000010 LDP 00010011
38	58	15	
65	43	47	

그림 5. LDP와 LBP의 안정성 비교
Fig. 5. Stability of LDP vs. LBP.

III. Signed Local Directional Pattern

1. LDP의 문제점

LDP^[25]방법은 에지 반응 값을 이용해서 LBP^[15]보다 안정적인 결과를 보여주지만 결정적인 문제점이 있다. LDP는 8방향 Kirsch 에지 반응 마스크^[18]를 사용해서 에지 반응 크기를 계산한다. 계산된 반응 크기는 음수와 양수가 모두 존재하게 된다. 하지만 LDP에서는 이들 부호를 무시하고 에지 반응 절대값 크기 상위 k개를 선정했다. 하지만 부호를 무시할 경우 특징이 전혀 다른 특정 패턴에서 같은 코드를 생성하는 문제점이 발생한다.

그림 6은 위에서 전술한 LDP의 문제점이 나타나는 예를 보여준다. 그림 6에서 서로 다른 반대되는 특징을 가지는 두 3x3 지역 픽셀 값을 이용해서 k=3인 LDP 코드 값을 계산했다. 이때 3x3 지역 픽셀 값은 서로 다른 에지 특징을 가지는 것을 눈으로 확인할 수 있다. 상단의 3x3 지역 픽셀은 동남쪽 방향으로 밝은 화소 값에서 어두운 화소 값으로 변하는 에지를 나타내고 하단 3x3 지역 픽셀은 동남쪽 방향으로 어두운 화소 값에서 밝은 화소 값으로 변하는 에지를 나타낸다. 에지의 방향을 부호화 할 때 두 에지는 반대 방향을 가지기 때문에 반드시 구별해야 한다. 하지만 LDP코드는 같은 값을 가진다. 이는 LDP를 이용해서 지역 패턴을 기술할 때 그림 6과 같은 에지 패턴을 구별하지 못한다는 것을 의미한다. 이러한 문제가 생기는 원인은 LDP에서 부호를 무시하고 절대값이 큰 순서대로 3개(k=3)의 방향만 1로 설정했기 때문에 발생한다. 에지 반응 값의 부호는 에지의 방향을 나타내는 중요한 정보이기 때문에 코드를 생성할 때 이를 반드시 고려해야 한다. 하지만 LDP에서는 이를 무시했기 때문에 이와 같은 문제점이 발생하게 된다. 얼굴 표정인식에서는 주름, 눈, 눈썹, 눈 꼬리 등 얼굴 특징에서 발생하는 에지 정보가 매우 중요하기 때문에 특정 에지 패턴을 구분하지 못하는 단점은 치명적인 문제점을 나타낼 수 있다.

162	180	175
193	20	27
190	20	25

Mask index	M_7	M_6	M_5	M_4	M_3	M_2	M_1	M_0
Mask Value	-2340	-1036	308	1444	1364	1220	140	-1100
LDP Rank	1	6	7	2	3	4	8	5
LDP Code	1	0	0	1	1	0	0	0
SLDP Rank	8	6	4	1	2	3	5	7
SLDP Code	0	0	0	1	1	1	0	0

93	75	80
62	235	228
65	235	230

Mask index	M_7	M_6	M_5	M_4	M_3	M_2	M_1	M_0
Mask Value	2340	1036	-308	-1444	-1364	-1220	-140	1100
LDP Rank	1	6	7	2	3	4	8	5
LDP Code	1	0	0	1	1	0	0	0
SLDP Rank	1	3	5	8	7	6	4	2
SLDP Code	1	1	0	0	0	0	0	1

그림 6. 다른 방향의 에지에서 같은 코드를 생성하는 LDP와 다른 코드를 생성하는 SLDP 비교
Fig. 6. LDP generates the same codes in edges of different directions but SLDP generates the different codes.

2. Signed Local Directional Pattern 생성 방법

위에서 전술한 LDP의 문제점은 부호를 무시했기 때문에 발생한다. 따라서 본 논문에서는 위와 같은 문제를 해결하기 위한 방법으로 부호를 고려한 LDP(Signed

Local Directional Pattern)코드 생성방법을 제안한다. SLDP는 LDP와 같이 각 화소에 $m \times m$ 의 지역 패턴 윈도우를 설정하고 여기에 8 방향 $m \times m$ Kirsch 에지 반응 마스크^[18]를 적용한다. Kirsch 에지 마스크를 적용해서 얻어진 8개의 에지 반응 값 중 큰 순서대로 k 개를 선정한다. 이때 LDP에서는 에지 반응 값의 절대값이 큰 순서대로 k 개를 선정했지만 본 논문에서 제안한 SLDP방법은 부호를 고려해서 k 개를 선정한다. 선정된 k 개의 에지 반응을 이용해 SLDP 코드를 만드는 방법은 수식 (2)와 같다. 여기서 m_k 는 Kirsch 에지 마스크로 계산된 8 방향의 에지 반응 값이고 m_k 는 에지 반응 값을 순서대로 나열 했을 때 k 번째 에지 반응 값이다. 이는 LDP와 같이 선정된 방향에 해당하는 비트를 1으로 설정해 주고 나머지는 0으로 설정해 코드를 생성하는 것과 같다. 그림 6은 LDP에서 구별하지 못했던 서로 다른 특징을 가진 두 3×3 지역 픽셀 값을 제안된 SLDP 방법과 LDP방법을 이용해서 지역 미세 패턴 코드를 생성하는 과정을 비교해서 보여준다. LDP와 다르게 SLDP에서는 에지 반응 값의 순위를 부호를 고려해서 정하고 에지 반응 값이 큰 방향 3개에 해당하는 비트를 1로 설정해준다. LDP에서는 부호를 고려하지 않았기 때문에 같은 비트 패턴을 생성 했지만 SLDP에서는 부호를 고려했기 때문에 다른 패턴을 생성 했다. 이는 SLDP는 LDP와 다르게 에지의 방향 패턴을 올바르게 부호화 할 수 있다는 것을 보여준다.

$$SLDP_k = \sum_{i=0}^7 s(m_i) \times 2^i, \tag{2}$$

$$s(a) = \begin{cases} 1 & a \geq m_k \\ 0 & a < m_k \end{cases}$$

SLDP는 LDP의 장점인 잡음과 단조롭지 않은 밝기 변화에 강력한 특징을 그대로 가지면서 LDP의 단점인 반대 방향 에지를 구별하지 못하는 문제점을 해결했다. SLDP는 LDP가 구별 하지 못하는 에지를 구별 가능하기 때문에 LDP보다 뛰어난 에지 구별능력을 제공할 수 있기 때문에 얼굴 표정인식에서 사용되는 주요 특징 (주름, 눈, 눈썹, 입, 등)을 정확하게 기술 할 수 있다.

3. SLDP 특징벡터 생성 방법

각 화소에서 SLDP코드 계산 한 뒤 입력된 $M \times N$ 크기를 가진 얼굴 영상은 SLDP 히스토그램을 이용해

서 표현이 가능하다. 수식 (3)은 $M \times N$ 크기를 가진 입력 영상을 히스토그램에 누적하는 방법을 나타낸다. 여기서 m, n 은 각 화소의 위치를 나타내고 히스토그램의 빈(bin)의 크기는 SLDP코드의 k 값에 따라 8개의 방향에서 k 개를 선택하는 조합의 숫자로 결정된다. 예를 들어 $k=3$ 인 경우 히스토그램 빈(bin) 크기는 56이 된다. 계산된 히스토그램은 영상 전체 기술하는 특징벡터로 사용할 수 있다.

$$H(i) = \sum_{m=1}^M \sum_{n=1}^N f(SLDP_k(m, n), i), \tag{3}$$

$$f(a, b) = \begin{cases} 1, & a = b \\ 0, & a \neq b \end{cases}$$

SLDP코드는 입력 영상의 에지, 점, 모퉁이 같은 세부적인 정보를 가지고 있지만 영상 전체에서 계산된 SLDP코드를 하나의 히스토그램에 누적하면 위치정보를 모두 상실하기 때문에 세부적인 정보가 사라지고 오직 SLDP코드 발생 확률만 보여주게 된다. 얼굴 표정 변화는 얼굴의 전역적인 변화보다는 지역전역 변화로 나타나기 때문에 표정인식을 위해서는 얼굴 영상을 지역적으로 분석하는 것이 필요하다^[20~21]. 따라서 그림 7과 같이 입력 영상을 g 개의 영역으로 분할해서 R_0, R_1, \dots, R_{g-1} 의 영역을 정의하고 각 영역에서 히스토그램 H_g 를 계산한다. 최종적으로 얼굴 전체의 특징 벡터를 표현하기 위해 지역적으로 계산된 히스토그램을 모두 연결해서 하나의 큰 히스토그램을 만들어준다. 이와 같이 생성된 히스토그램은 분할된 영역 수 g 와 SLDP코드 길이의 곱으로 연결된 히스토그램 크기가 정해진다.

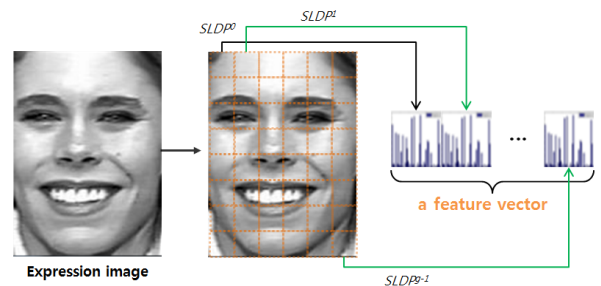


그림 7. 얼굴 표정 영상은 작은 영역으로 나눈 해당 영역의 히스토그램을 계산한다. 계산된 히스토그램을 연결해서 얼굴특징 벡터를 생성한다.

Fig. 7. Expression image is divided into small regions from which SLDP histograms are calculated and concatenated into a face feature vectors.

연결된 히스토그램은 그 자체로 표정 영상을 나타내는 특징벡터가 된다.

IV. SLDP를 이용한 얼굴 표정인식

얼굴 표정인식은 SLDP를 이용해 생성된 얼굴 특징 벡터를 Template matching, linear discriminant analysis, linear programming, SVM(Support Vector Machine)과 같은 기계학습(machine learning) 방법으로 학습 및 분류하는 방법으로 수행 할 수 있다. 기존 LBP를 이용한 얼굴 표정인식에서 언급한 기계학습 방법들의 비교 평가가 수행 되었다^[9]. 여기에서 SVM이 가장 좋은 결과를 보여주었기 때문에 본 논문에서 제안된 SLDP방법의 얼굴 표정인식은 SVM을 이용해서 수행 한다.

1. Support Vector Machine(SVM)

SVM 방법은 통계적으로 잘 만들어진 학습 이론이다. SVM은 컴퓨터 비전 여러 분야에 다양한 분류 작업에 있어 성공적으로 사용되고 있다^[22]. SVM은 목시적인 데이터 사상(mapping)을 더 높은 차원의 특징 벡터 공간으로 수행한다. 그리고 데이터 구분을 위해서 최대 여유분을 가진 선형 구별 초평면(Hyper-Plane)을 찾는다.

$$T = \{(s_i, l_i), i = 1, 2, \dots, L\}, \quad (4)$$

수식 4와 같은 학습 데이터가 있다. 여기서 $s_i \in \mathbb{R}^p$ 이고 $l_i \in \{-1, 1\}$, 이다. 이와 같은 학습 데이터가 주어질 때 새로운 테스트 데이터 x 는 수식 5를 이용해서 분류된다. 여기서 α_i 는 dual optimization problem의 Lagrange multipliers이고 b 는 편향 값 또는 임계 치 매 개변수이고 K 는 커널 함수이다. x_i 중 $\alpha_i > 0$ 인 학습 표본을 support vector라고 부른다. 이 support vector를 이용해서 분류를 위한 초평면(Hyper-plane)의 여유분을 최대화 한다. 커널 함수는 주로 Linear, Polynomial, Radial Basis Function(RBF) 커널 함수가 사용된다.

$$f(x) = \text{sign} \left(\sum_{i=1}^L \alpha_i l_i K(x_i, x) + b \right), \quad (5)$$

원래 SVM은 2가지 분류(-1, 1)를 가진 데이터를 구별하는 방법이다. 하지만 표정인식 같은 경우 6가지 또는 7가지 표정을 가지고 있다. 이를 SVM을 이용해서 구별하기 위해서는 1가지 표정을 하나의 분류로 나머지 표정 모두를 다른 하나의 분류로 구분해서 2가지로 구분 하여 적용 하는 방법 또는 여러 개의 2가지 구별 방법으로 나눠서 사용하는 방법이 있다. 본 논문에서 1 가지 표정을 하나의 분류로 나머지 표정 모두를 다른 하나의 분류로 놓는 방법을 사용했다.

V. 실험 및 결과

1. 실험 환경

대부분의 얼굴 표정인식 시스템은 화남(Anger), 역겨움(Disgust), 두려움(Fear), 즐거움(Joy), 슬픔(Sad), 놀람(Surprise)과 같은 일반적인 6가지 감정[9] 표현을 인식한다. 이 6가지 표정은 무표정을 포함해서 7가지 표정으로 확장될 수 있다. 본 논문에서는 6가지 표정인식과 7가지 표정인식 모두 수행 가능한 시스템을 개발 얼굴 표정인식 실험을 수행 하였다. 이때 2가지 널리 알려진 Cohn-Kanade 얼굴 표정 영상 데이터베이스^[17]와 Japanese Female Facial Expression(JAFFE) 데이터베이스^[11]를 이용했다.

Cohn-Kanade(C-K)데이터 베이스는 18세에서 30세까지의 100명의 대학생으로 구성되어 있다. 학생들 중 65%가 여성이고 15%가 아프리카 인종 미국인이고 3%가 아시아인 또는 라틴계열이다. 학생들에게 무표정부터 화남, 역겨움, 두려움, 즐거움, 슬픔, 놀람 같은 표정을 짓도록 시키고 표정 변화를 영상에 담아 데이터베이스를 구축하였다. 무표정으로부터 일반적인 6가지 표정으로 변하는 영상은 640x480 또는 640x690 해상도와 그레이 스케일 컬러를 가진 연속된 디지털 영상으로 구성되어 있다. 본 논문에서 개발한 평가 시스템은 408 가지 연속된 표정 변화 영상을 97명의 학생으로부터 선택했다. 연속된 표정 변화 영상들은 6가지 일반적인 표정 중 하나를 나타낸다. 6가지 일반적인 연속된 표정 변화 영상 중 가장 뚜렷하게 표정은 나타내는 영상 3개를 뽑아 1,224장의 영상을 만들었다. 여기에 모든 연속된 표정 변화 영상의 첫 영상(무표정) 408장을 더해서 총 1,632장의 7가지 일반적인 표정을 나타내는 영상을 만들었다.

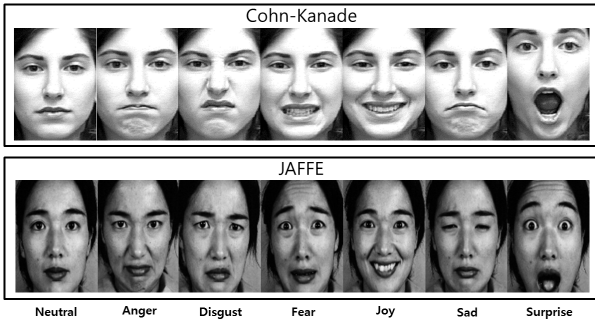


그림 8. Cohn-Kanade와 JAFFE 데이터베이스의 7가지 일반적인 표정을 나타내는 표정 영상 표본
 Fig. 8. Sample expression images of each prototypic expression from Cohn-Kanade and JAFFE databases.

JAFFE 데이터베이스는 10명의 일본 여성 영상 213장으로 구성되어 있다. 각각의 영상은 256x256 해상도를 가지고 있고 각 영상은 다른 표정을 나타내고 있다. 얼굴은 정면을 바라보고 여성들의 머리는 얼굴 표정이 드러나도록 얼굴 바깥으로 단정하게 넘겼다. 얼굴에 균일한 조도를 만들기 위해서 텅스텐 라이트를 사용해서 얼굴 영상을 촬영했다. 실제 사진을 찍은 여성들의 실명은 노출되지 않았고 약자를 이용해서 표시하였다. KA, KL, KM, KR, MK, NA, NM, TM, UY, YM은 해당 여성들의 약자다. Cohn-Kanade와 JAFFE 데이터베이스의 7가지 표정에 해당하는 표정 영상 표본은 그림 8에서 확인할 수 있다.

데이터베이스에서 선택된 표정영상을 바로 얼굴 표정인식에 사용 할 수 없다. 외형기반 얼굴 표정인식을 성공적으로 수행하기 위해서는 해당 표정영상에서 얼굴 표정 인식에 필요한 얼굴 영역만 잘라내서 사용해야 하기 때문이다. 얼굴 영역을 찾아내기 위해서 눈의 위치를 사용한다. Cohn-Kanade 데이터베이스는 정확한 눈 위치 정보를 제공하기 때문에 제공된 눈 위치 정보를 사용해서 얼굴 영역을 찾아낸다. JAFFE 데이터베이스는 눈 위치 정보를 제공하지 않기 때문에 기존의 좋은 성능을 제공하는 눈 검출 알고리즘^[23]을 이용한다. 눈 위치를 이용한 얼굴 영역 선택의 원리는 매우 간단하다. 눈과 눈 사이의 거리를 D로 놓는다. 눈 바깥쪽 왼쪽 오른쪽 영역은 각각 0.5D로 설정한다. 얼굴 영역의 세로 길이는 2.7D로 놓고 눈 아래쪽은 2D 위쪽은 0.7D로 놓는다. 계산된 얼굴 영역을 잘라내고 해상도는 150x110으로 변경한다. 본 실험에서는 모든 얼굴 영상

을 6 x 7 블록(가로 블록 숫자, 세로 블록숫자)으로 분할 한 다음 얼굴 표정 인식을 수행하였다.

SLDP는 밝기와 잡음에 강력한 특징을 가지고 있기 때문에 데이터베이스의 영상을 잘라내고 해상도 변경 후에 별도의 처리를 하지 않는다. 본 논문의 실험에서는 얼굴 표정인식을 위해서 10 분할 상호 검증(10-fold cross validation)을 수행했다. 10 분할 상호 검증이란 Cohn-Kanade 데이터베이스와 JAFFE 데이터베이스로부터 얻어진 각 데이터 셋을 10개의 집단으로 임의 분할하고 9개의 그룹을 학습데이터로 나머지 1개의 그룹을 검증 데이터로 사용해서 얼굴 표정 인식을 수행하는 것을 말한다. 위와 같은 과정을 총 10번 반복하면 데이터 셋의 모든 데이터에 대해서 표정 인식을 수행 할 수 있고 이렇게 상호 검증을 수행한다. 상호 검증을 위해서 10개의 집단으로 임의 분할했기 때문에 결과가 매번 달라질 수 있다. 이와 같은 문제점을 완화하기 위해서 상호검증을 총 10번 반복해서 수행하고 평균을 구했다. 즉 본 실험에서는 각 데이터 셋을 10개의 집단으로 임의 분할하고 각 집단의 인식률을 계산하여 상호검증을 수행하고 이와 같은 상호검증을 10번 반복한 뒤 평균을 구해서 제안된 방법의 결과를 검증했다.

2. 표정인식을 위한 최적의 SLDP 결정

SLDP는 LDP와 같이 선정한 방향 비트수 k에 따라 표정인식 성능이 달라 질 수 있다. 방향 비트수가 증가하면 세밀한 패턴 까지도 표현이 가능하지만 코드의 길이가 길어지고 예지 반응 값이 작은 방향 비트는 잡음에 취약하게 된다. SLDP는 LDP보다 강력한 구별 능력을 보여주기 때문에 LDP와 다른 특징을 보여 줄 수 있다. 따라서 본 실험에서는 SLDP의 선정 방향 비트수 k에 따른 최적의 SLDP 코드를 결정한다. 실험을 위해 실험 환경에서 정해진 블록 숫자 6x7(가로축 분할 및 세로축 분할)에 k값을 1부터 4까지 다르게 해서 실험을 수행했다. 실험은 Cohn Kanade 데이터베이스와 SVM을 이용해서 7가지 표정인식과 6가지 표정인식을 수행했다. 이때 SVM의 커널 함수는 Linear, Polynomial, RBF함수를 사용했다.

표 1과 표 2는 6x7로 분할된 블록 숫자에서 k값에 따른 6가지 표정인식과 7가지 표정인식 성능을 Cohn-Kanade 데이터베이스에서 측정한 결과다. 여기에서 6가지 표정인식과 7가지 표정인식 모두 k값이 2일

표 1. Cohn-Kanade 데이터베이스에서 k값에 따른 6 가지 얼굴 표정인식 성능

Table 1. The performance of 6-class face expression recognition for different k on Cohn Kanade database.

6-class expressions(%)			
k number	Linear	Polynomial	RBF
	Kernels(%)	Kernels(%)	Kernels(%)
k=1	98.17 ± 0.60	99.19 ± 0.58	99.40 ± 0.52
k=2	98.77 ± 0.49	98.53 ± 0.82	99.41 ± 0.31
k=3	98.93 ± 0.32	99.13 ± 0.68	99.32 ± 0.54
k=4	98.79 ± 0.67	99.17 ± 0.50	99.31 ± 0.45

표 2. Cohn-Kanade 데이터베이스에서 k값에 따른 7 가지 얼굴 표정인식 성능

Table 2. The performance of 7-class face expression recognition for different number of regions.

7-class expressions(%)			
k number	Linear	Polynomial	RBF
	Kernels(%)	Kernels(%)	Kernels(%)
k=1	92.78 ± 0.93	95.51 ± 0.78	96.21 ± 0.32
k=2	93.87 ± 0.55	95.44 ± 0.52	96.55 ± 0.69
k=3	93.79 ± 0.41	95.72 ± 0.43	96.20 ± 0.67
k=4	93.73 ± 0.65	95.29 ± 0.61	96.51 ± 0.68

때 가장 좋은 인식률을 보여준다. LDP의 경우 k값이 3 일 때 가장 좋은 결과를 보여준 것과 달리 SLDP는 k값이 2일 때 가장 좋은 결과를 보여준다. 이는 부호를 고려한 SLDP의 구별 능력이 LDP보다 뛰어나기 때문에 LDP보다 적은 수의 k값으로 충분한 예지 방향 정보를 부호화 할 수 있다고 판단된다. 또한 k값이 작을수록 코드의 길이가 줄어들기 때문에 SLDP는 LDP보다 효율적인 코드로 지역 패턴 표현이 가능하다. 여기에서 얻어진 결과를 바탕으로 본 논문에서는 최적의 SLDP 특징 벡터 생성을 위한 최적의 k값은 2로 결정하고 결정된 값을 바탕으로 비교 평가를 수행한다.

3. Cohn-Kanade DB에서 표정인식 비교 평가

실험으로 얻어진 최적의 k값 2인 SLDP를 이용해서 다른 방법들과 비교 평가를 Cohn-Kanade 데이터베이스에서 수행 했다. 6가지 표정인식과 7가지 표정인식을 수행하기위해 Cohn Kanade 데이터베이스에서 추출한 얼굴 영상 1224장과 1632장을 사용해서 표정인식을 수행했다. SVM은 잘 만들어진 기계 학습 기술이고 패턴 인식에서 뛰어난 분류 성능을 보여준다. 따라서 본 논

표 3. Cohn-Kanade 데이터베이스에서 SVM을 이용한 6가지 얼굴 표정인식 방법 성능 비교

Table 3. The performance of 6-class expression recognition with different methods using SVM on Cohn Kanade database.

6-class expressions(%)			
Feature descriptor	Linear	Polynomial	RBF
	Kernels(%)	Kernels(%)	Kernels(%)
LBP ^[5]	91.5 ± 3.1	91.5 ± 3.1	92.6 ± 2.9
Gabor ^[27]	89.4 ± 3.0	89.4 ± 3.0	89.8 ± 3.1
LDP ^[25]	94.9 ± 1.2	94.9 ± 1.2	96.4 ± 0.9
SLDP	98.8 ± 0.5	98.5 ± 0.8	99.4 ± 0.3

표 4. Cohn-Kanade 데이터베이스에서 SVM을 이용한 7가지 얼굴 표정인식 방법 성능 비교

Table 4. The performance of 7-class expression recognition with different methods using SVM on JAFFE database.

7-class expressions(%)			
Feature descriptor	Linear	Polynomial	RBF
	Kernels(%)	Kernels(%)	Kernels(%)
LBP ^[5]	88.1 ± 3.8	88.1 ± 3.8	88.9 ± 3.5
Gabor ^[27]	86.6 ± 4.1	86.6 ± 4.1	86.8 ± 3.6
LDP ^[25]	92.8 ± 1.7	92.8 ± 1.7	93.4 ± 1.5
SLDP	93.9 ± 0.5	95.4 ± 0.5	96.5 ± 0.7

표 5. Cohn-Kanade 데이터베이스에서 RBF커널 SVM을 이용한 6표정 인식(%) 혼동 행렬(Confusion matrix)

Table 5. Confusion matrix of 6-class expression recognition (%) using SVM with RBF kernel function on Cohn Kanade database.

	Anger	Disgust	Fear	Joy	Sad	Surprise
Anger	99.0	0.0	0.1	0.0	0.9	0.0
Disgust	0.2	99.7	0.1	0.0	0.0	0.0
Fear	0.2	0.0	99.5	0.3	0.0	0.0
Joy	0.1	0.0	0.1	99.8	0.0	0.0
Sad	0.5	0.0	0.0	0.0	99.5	0.0
Surprise	0.8	0.0	0.0	0.0	0.0	99.2

문 비교 평가에서는 Linear, Polynomial, RBF 커널 함수를 사용한 SVM을 이용해서 표정인식을 수행 했다. Cohn Kanade 데이터베이스에서 6가지 및 7가지 얼굴 인식 성능 비교 결과는 표 3과 4에서 볼 수 있다. 본 논문에서 제안한 SLDP는 LDP보다 k값이 작기 때문에 코드의 길이가 LDP보다 작다. 하지만 LDP보다 6가지 표정과 7가지 표정에서 훨씬 더 안정적이고 강력한 결과를 보여준다.

표 6. Cohn-Kanade 데이터베이스에서 RBF커널 SVM을 이용한 7표정 인식(%) 혼동 행렬(Confusion matrix)

Table 6. Confusion matrix of 7-class expression recognition (%) using SVM with RBF kernel function on Cohn Kanade database.

	Anger	Disgust	Fear	Joy	Sad	Surprise	Neutral
Anger	96.9	0.2	0.0	0.0	0.0	0.0	2.9
Disgust	0.5	98.3	0.0	0.0	0.0	0.0	1.2
Fear	0.2	0.1	97.5	0.0	0.0	0.0	2.2
Joy	0.0	0.0	0.1	99.9	0.0	0.0	0.0
Sad	0.5	0.0	0.0	0.0	98.2	0.0	1.3
Surprise	0.0	0.0	0.0	0.0	0.0	99.5	0.5
Neutral	5.2	0.8	1.4	0.2	1.4	0.2	90.8

지금까지는 모든 표정의 평균 인식 성능을 가지고 성능을 논의 했다. 좀 더 자세하게 각 표정별 성능을 알아보기 위해서 6가지 표정인식과 6가지 표정인식의 각 혼동 행렬(Confusion matrix)을 표 5와 6에 제시 했다. 6가지 표정에서 높은 인식률은 가지던 화남(Anger), 혐오(Disgust), 슬픔(Sad)이 7가지 표정인식에서는 무표정(Neutral)로 잘못 분류되었기 때문에 7가지 표정인식에서 다소 낮은 성능을 보였다. 하지만 흥분(Surprise)과 즐거움(Joy)은 무표정(Neutral)으로 잘못 분류되는 경우가 극히 적었다.

4. JAFFE DB에서 표정인식 비교 평가

Cohn-Kanade 데이터베이스에서 수행했던 것과 같은 방식으로 JAFFE 데이터베이스에서도 3가지 커널 함수를 이용한 SVM을 이용해서 얼굴 표정 인식을 수행했다. 6가지 표정인식과 7가지 표정인식을 수행하기 위해 JAFFE 데이터베이스에서 추출한 7가지 얼굴 표정 영상 213장과 무표정을 제외한 183장을 사용해서 표정인식을 수행했다. JAFFE 데이터베이스에서 6가지 및 7가지 얼굴인식 성능 비교 결과는 표 7과 8에서 볼 수 있다. 본 논문에서 제안한 SLDP는 적은 코드길이에도 불구하고 6가지 표정 인식과 7가지 표정인식에서 기존 방법들 보다 좋은 결과를 보여준다. 특히 7가지 표정에서 기존 방법들 보다 매우 뛰어난 모습을 보여준다. 하지만 전체적으로 Cohn-Kanade 데이터베이스 결과에 비해 전반적으로 나쁜 결과를 보여준다.

Cohn Kanade 데이터베이스 수행했던 것과 같이 좀 더 자세한 표정별 결과를 분석하기 위해서 JAFFE 데이터베이스 6가지 표정인식과 6가지 표정인식 결과의

표 7. JAFFE 데이터베이스에서 SVM을 이용한 6가지 얼굴 표정인식 방법 성능 비교

Table 7. The performance of 6-class expression recognition with different methods using SVM on Cohn Kanade database.

6-class expressions(%)			
Feature descriptor	Linear Kernels(%)	Polynomial Kernels(%)	RBF Kernels(%)
LBP ^[5]	86.7 ± 4.1	86.7 ± 4.1	87.5 ± 5.1
Gabor ^[27]	85.1 ± 5.0	85.1 ± 5.0	85.8 ± 4.1
LDP ^[25]	89.9 ± 5.2	89.9 ± 5.2	90.1 ± 4.9
SLDP	91.0 ± 1.3	89.5 ± 2.6	91.2 ± 1.6

표 8. JAFFE 데이터베이스에서 SVM을 이용한 7가지 얼굴 표정인식 방법 성능 비교

Table 8. using SVM on JAFFE database.

7-class expressions(%)			
Feature descriptor	Linear Kernels(%)	Polynomial Kernels(%)	RBF Kernels(%)
LBP ^[5]	80.7 ± 5.5	80.7 ± 5.5	81.9 ± 5.2
Gabor ^[27]	79.7 ± 4.2	79.7 ± 4.2	80.8 ± 3.7
LDP ^[25]	84.9 ± 4.7	84.9 ± 4.7	85.4 ± 4.0
SLDP	90.3 ± 0.7	88.8 ± 1.9	90.6 ± 1.4

표 9. JAFFE 데이터베이스에서 RBF커널 SVM을 이용한 제안된 방법의 6표정 인식(%) 혼동 행렬(Confusion matrix)

Table 9. Confusion matrix of 6-class expression recognition (%) of the proposed method using SVM with RBF kernel function on JAFFE database.

	Anger	Disgust	Fear	Joy	Sad	Surprise
Anger	90.3	2.7	0.0	0.0	7.0	0.0
Disgust	2.1	94.5	0.0	0.0	3.4	0.0
Fear	0.0	7.8	87.8	0.0	3.4	0.9
Joy	0.0	0.0	0.0	99.4	0.6	0.0
Sad	1.0	0.0	11.9	3.2	83.9	0.0
Surprise	0.0	0.0	3.7	4.7	0.0	91.7

각 혼동 행렬(Confusion matrix)을 표 9과 10에 제시 했다. Cohn-Kanade 데이터베이스 결과와 다르게 화남(Anger)과 슬픔(Sad)의 구별과 슬픔(Sad)과 두려움(Fear)의 구별이 잘 이루어지지 않았다. 또한 혐오(Disgust)와 두려움(Fear)도 Cohn-Kanade 데이터베이스 보다 잘 구별되지 않았다. 이는 그림 9와 같이 JAFFE의 몇몇 표정 영상이 다른 표정 영상과 매우 유사한 영상을 가지기 때문이다. 이와 같은 이유로

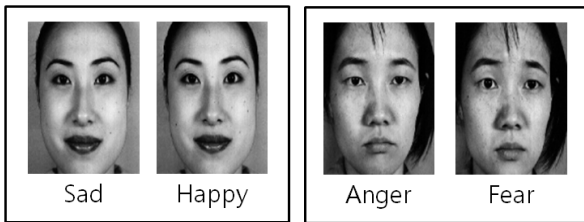


그림 9. JAFFE 데이터베이스의 다른 표정의 비슷한 영상 표본

Fig. 9. The samples of similar images which are labeled with different expressions in JAFFE.

표 10. JAFFE 데이터베이스에서 RBF커널 SVM을 이용한 제안된 방법의 7표정 인식(%) 혼동 행렬 (Confusion matrix)

Table 10. Confusion matrix of 7-class expression recognition (%) of the proposed method using SVM with RBF kernel function on JAFFE database.

	Anger	Disgust	Fear	Joy	Sad	Surprise	Neutral
Anger	90.0	1.3	0.0	0.0	8.7	0.0	0.0
Disgust	1.4	95.2	1.4	0.0	2.1	0.0	0.0
Fear	0.0	8.8	83.1	0.0	6.3	1.3	0.6
Joy	0.0	0.0	0.0	98.1	1.9	0.0	0.0
Sad	0.6	0.0	12.3	3.2	81.9	0.0	1.9
Surprise	0.0	0.0	3.3	4.7	0.0	89.3	2.7
Neutral	0.0	0.0	0.0	0.0	0.0	2.7	97.3

JAFFE의 표정인식이 전반적으로 Cohn-Kanade데이터베이스 보다 나쁜 결과가 나오게 된다. 하지만 Cohn-Kanade데이터베이스와 다르게 7가지 표정인식에서는 각 표정이 무표정으로 잘못 판단되는 경우가 매우 적었으며 이로 인해 JAFFE 데이터베이스 결과는 6가지 표정인식과 7가지 표정 인식의 차이가 매우 적었다.

VI. 결 론

본 논문에서는 표정 인식을 위한 새로운 지역 미세 패턴으로 SLDP를 제안했다. SLDP는 얼굴 표정인식을 위한 얼굴 영상의 지역 정보를 정확하게 기본 방법들 보다 효율적인 코드로 표현이 가능하다. SLDP는 지역 패턴 코드를 생성하기 위해 주 예지의 방향 정보를 사용했기 때문에 밝기 정보를 사용한 것 보다 안정적이고 잡음에 강력하다. 또한 지역 패턴 생성 시 부호를 고려하기 때문에 기존 LDP에서 구별이 어려웠던 밝기 값이 반대인 비슷한 모양의 예지를 구별할 수 있다. 그

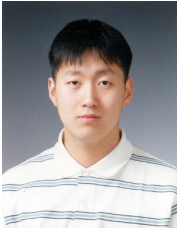
결과 기존 방법들 보다 뛰어난 얼굴 표정인식 성능을 보여준다.

REFERENCES

- [1] Y.L. Tian et al., Real World Real-time Automatic Recognition of Facial Expressions, Proc. IEEE Workshop Performance Evaluation of Tracking and Surveillance, 2003.
- [2] C. Shan, S. Gong, and P.W. McOwan, "Robust Facial Expression Recognition using Local Binary Patterns," Proc. IEEE Int. Conf. Image Process., pp. 914-917, Genoa, Italy, September 2005.
- [3] M.C. Hwang et al., "Person Identification System for Future Digital TV with Intelligence," IEEE Trans. Consum. Electron., vol. 53, no. 1, pp. 218-226, 2007.
- [4] P. Corcoran et al., "Biometric Access Control for Digital Media Streams in Home Networks," IEEE Trans. Consum. Electron., vol. 53, no. 3, pp. 917-925, 2007.
- [5] C. Shan, S. Gong, and P.W. McOwan, "Facial Expression Recognition based on Local Binary Patterns: A Comprehensive Study," Image Vision Comput., vol. 27, no. 6, pp. 803-816, May 2009.
- [6] Y. Tian, T. Kanade, and J.F. Cohn, "Facial Expression Analysis," Handbook of Face Recognition, Springer, pp. 247-275, Oct. 2003.
- [7] M.A. Turk and A.P. Pentland, "Face Recognition Using Eigenfaces," Proc. Comput. Vision Pattern Recog., pp. 586-591, 1991.
- [8] C. Padgett and G. Cottrell, "Representing Face Images for Emotion Classification," Advances in Neural Information Processing Systems, M. Mozer, M. Jordan, and T. Petsche, eds., vol. 9, Cambridge, Mass.: MIT Press, 1997.
- [9] M.S. Bartlett, J.R. Movellan, and T.J. Sejnowski, "Face Recognition by Independent Component Analysis," IEEE Trans. Neural Networks, vol. 13, no. 6, pp. 1450-1464, 2002.
- [10] C.C. Fa and F.Y. Shin, "Recognizing Facial Action Units using Independent Component Analysis and Support Vector Machine," Pattern Recog., vol. 39, no. 9, pp. 1795-1798, 2006.
- [11] M.J. Lyons, J. Budynek, and S. Akamatsu, "Automatic Classification of Single Facial Images," IEEE Trans. Pattern Anal. Mach. Intell., vol. 21, no. 12, pp. 1357-1362, 1999.

- [12] T. Ojala and M. Pietikainen, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971-987, 2002.
- [13] X. Feng, M. Pietikainen, and A. Hadid, "Facial Expression Recognition with Local Binary Patterns and Linear Programming," *Pattern Recog. Image Anal.*, vol. 15, no. 2, pp. 546-548, 2005.
- [14] G. Zhao and M. Pietikainen, "Dynamic Texture Recognition using Local Binary Patterns with An Application to Facial Expressions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 6, pp. 915-928, 2007.
- [15] H. Zhou, R. Wang, and C. Wang, "A Novel Extended Local Binary Pattern Operator for Texture Analysis," *Inf. Science*, vol. 178, no. 22, pp. 4314-4325, 2008.
- [16] T. Jabid, M.H. Kabir, and O.S. Chae, "Local Directional Pattern (LDP) for Face Recognition," *IEEE Int. Conf. Consum. Electron.*, pp. 329-330, Las Vegas, U.S.A, Jan. 2010.
- [17] T. Kanade, J. Cohn, and Y. Tian, "Comprehensive Database for Facial Expression Analysis," *IEEE Int. Conf. Autom. Face Gesture Recog.*, pp. 46-53, Grenoble, France, Mar. 2000.
- [18] W.K. Pratt, *Digital Image Processing*, Wiley, New York, pp. 489, 1978.
- [19] T. Jabid, M.H. Kabir, and O.S. Chae, "Local Directional Pattern (LDP): A Robust Image Descriptor for Object Recognition," *IEEE Int. Conf. Adv. Video and Signal-Based Surveillance*, pp. 482-487, Boston, U.S.A, Sep. 2010.
- [20] T. Ahonen, A. Hadid, and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037-2041, 2006.
- [21] S. Gundimada and V.K. Asari, "Facial Recognition Using Multisensor Images Based on Localized Kernel Eigen Spaces," *IEEE Trans. Image Process.*, vol. 18, no. 6, pp. 1314-1325, 2009.
- [22] C. Cortes and V. Vapnik, "Support Vector Networks," *Machine Learning*, vol. 20, no. 3, pp. 273-297, 1995.
- [23] Z. Niu et al., "2D Cascaded AdaBoost for Eye Localization," *Proc. IEEE Int. Conf. Pattern Recog.*, pp. 1216-1219, Hong Kong, Aug. 2006.
- [24] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang, "Local Gabor Binary Pattern Histogram Sequence (LGBPHS): A Novel Non-Statistical Model for Face Representation and Recognition," *Proc. IEEE Int'l Conf. Computer Vision*, pp. I: 786-791, Beijing, China, Oct. 2005.
- [25] T. Jabid, M. H. Kabir, and O. Chae, "Robust facial expression recognition based on local directional pattern," *ETRI Journal*, vol. 32, pp. 784 - 794, 2010.
- [26] P. Phillips, H. Moon, S. Rizvi, and P. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 10, pp. 1090 - 1104, Oct. 2000.
- [27] M.S. Bartlett et al., "Recognizing Facial Expression: Machine Learning and Application to Spontaneous Behavior," *IEEE Conf. Computer Vision and Pattern Recog.*, pp. 568-573, San Diego, CA. U.S.A, June 2005.

저 자 소 개



류 병 용(정회원)
2010년 경희대학교 컴퓨터공학과
학사 졸업
2010년~현재 경희대학교 컴퓨터
공학과 석박통합과정
<주관심분야 : 멀티미디어 데이
터처리 및 통신, Robot Vision, 영
상처리, Signal Processing 등>



송 기 훈(정회원)
2014년 경희대학교 컴퓨터 공학과
학사 졸업
2014년~현재 경희대학교 컴퓨터
공학과 석박통합과정
<주관심분야 : 멀티미디어 데이
터처리 및 통신, Robot Vision, 영
상처리, Signal Processing 등>



김 재 면(정회원)
2010년 경희대학교 컴퓨터공학과
학사 졸업
2010년~현재 경희대학교 컴퓨터
공학과 석박통합과정
<주관심분야 : 영상처리, 컴퓨터
비전, 패턴인식, 신호처리>



채 옥 삼(평생회원)
1982년 오클라호마 주립대학 전기
및 컴퓨터공학 석사 졸업.
1986년 오클라호마 주립대학 전기
및 컴퓨터공학 박사 졸업.
1986년~1988년 Texas Instrument
Image Processing Lab.
선임연구원
1988년~현재 경희대학교 컴퓨터공학과 교수
<주관심분야 : 멀티미디어 데이터처리 및 통신,
Robot Vision, 영상처리, Signal Processing 등>



안 기 옥(정회원)
2004년 경희대학교 전자계산
공학과 석사 졸업
2004년~2006년 (주)엠지시스템
주임 연구원
2006년~2012년 미디어코러스
2012년~현재 이타기술 선임연구
원 및 경희대학교 겸임
교수

<주관심분야 : 멀티미디어 데이터처리 및 통신,
Robot Vision, 영상처리, Signal Processing 등>