

논문 2011-48SP-4-9

# Gabor 특징과 웨이브렛 영역의 BDIP와 BVLC 특징을 이용한 질감 특징 기반 언어 인식

( Texture Feature-Based Language Identification Using Gabor Feature  
and Wavelet-Domain BDIP and BVLC Features )

장 익 훈\*, 이 우 신\*\*, 김 남 철\*\*\*

( Ick Hoon Jang, Woo Shin Lee, and Nam Chul Kim )

## 요 약

본 논문에서는 Gabor 특징과 웨이브렛 영역의 BDIP와 BVLC 특징을 이용한 질감 특징 기반 언어 인식 방법을 제안한다. 제안된 방법에서는 먼저 시험 영상에 Gabor 변환과 웨이브렛 변환을 적용한다. 웨이브렛 영역의 상세 대역에는 Donoho의 연역치화를 적용하여 잡음을 제거한다. 이어서 Gabor 영상에는 크기 연산자를 적용하고 웨이브렛 부대역에는 BDIP와 BVLC 연산자를 적용한다. 그런 다음 Gabor 크기 영상과 BDIP, BVLC 부대역에 대하여 통계치를 계산하여 그 결과들을 벡터화하고 융합하여 특징 벡터로 사용한다. 분류 단계에서는 얼굴 인식에 주로 사용되는 WPCA를 분류기로 하여 시험 특징 벡터와 가장 유사한 학습 특징 벡터를 찾는다. 실험 결과 제안된 방법은 실험 문서 영상 DB에 대하여 비교적 낮은 특징 벡터 차원으로 매우 우수한 언어 인식 성능을 보여준다.

## Abstract

In this paper, we propose a texture feature-based language identification using Gabor feature and wavelet-domain BDIP (block difference of inverse probabilities) and BVLC (block variance of local correlation coefficients) features. In the proposed method, Gabor and wavelet transforms are first applied to a test image. The wavelet subbands are next denoised by Donoho's soft-thresholding. The magnitude operator is then applied to the Gabor image and the BDIP and BVLC operators to the wavelet subbands. Moments for Gabor magnitude image and each subband of BDIP and BVLC are computed and fused into a feature vector. In classification, the WPCA (whitened principal component analysis) classifier, which is usually adopted in the face identification, searches the training feature vector most similar to the test feature vector. Experimental results show that the proposed method yields excellent language identification with rather low feature dimension for a document image DB.

**Keywords :** Language identification, texture feature, BDIP, BVLC, Gabor transform, wavelet transform

## I. 서 론

최근 컴퓨터와 그 관련 기술이 비약적으로 발전함에

따라 종이 문서들을 문서 영상으로 스캔 한 다음 문자들을 OCR(optical character recognition) 시스템으로 인식하여 문서 파일로 저장, 관리하는 방법이 널리 사용되고 있다. 대부분의 OCR 시스템은 여러 언어의 OCR 엔진으로 구성되어 있는데, 어떤 언어의 OCR 엔진을 사용할 것인가는 주로 사용자가 선택하게 된다. 그런데 언어의 종류가 다양해지고 문서의 분량이 방대해질 경우 해당 언어의 OCR 엔진을 선택하는 데 많은 시간과 노력이 필요하게 된다. 따라서 문서 언어의 종류를 컴

\* 정회원, 경운대학교 디지털전자공학과  
(Department of Digital Electronic Engineering,  
Kyungwoon University)

\*\* 학생회원, \*\*\* 정회원-교신저자, 경북대학교 전자공학부

(School of Electronics Engineering, Kyungpook  
National University)

접수일자: 2010년11월6일, 수정완료일: 2011년5월4일

퓨터에서 자동으로 인식하게 되면 해당 언어의 OCR 엔진을 자동으로 선택하여 문자를 인식하고 그 결과를 효율적으로 저장, 관리 할 수 있게 된다<sup>[1]</sup>. 세계가 글로벌화 되면서 여러 언어로 된 문서를 접하게 되는 경우가 점점 증가하고 있어서 문서 영상에서 자동으로 언어를 인식하여 해당 언어의 OCR 엔진으로 문자를 인식할 수 있도록 하는 방법의 필요성이 증대되고 있다.

문서 영상에서의 문자 인식의 전단계로 사용하기 위한 언어 인식에 관한 기존의 방법들은 언어 별로 문자가 서로 다른 구조를 갖는 점에 근거한 문자 화소의 통계적 특징(statistical feature)을 이용하는 방법<sup>[2~4]</sup>과 문자 영상으로부터 추출한 질감 특징(texture feature)을 이용하는 방법<sup>[5~9]</sup>이 주류를 이루어왔다.

Hochberg 등<sup>[2]</sup>은 각 언어 별로 자주 사용되는 단어의 형태를 찾아내어 클러스터링(clustering) 시킨 다음 클러스터별 대표 틀(template)을 생성하고 인식하고자 하는 언어와 가장 유사한 클러스터를 찾아내어 해당 언어를 인식된 언어로 하였다. 이 방법에서는 13개국 언어를 인식하였다. Spitz<sup>[3]</sup>는 문자의 윗방향 요면(upward concavity)의 수직적 분포, 이진화된 문자 셀(character cell)의 정규화된 화소 분포, 그리고 단어 형태의 발생 빈도를 이용하여 6개국 언어를 인식하였다. Shijian과 Tan<sup>[4]</sup>은 문서 영상을 문자 또는 단어의 형태와 발생 빈도를 특징지우는 문서 벡터(document vector)로 만든 다음 가장 유사한 벡터를 찾는 방식으로 언어를 인식하는 방법을 제안하였다. 이 방법은 6개국 언어에 대한 실험에서 잡음과 왜곡이 있는 문서 영상에서도 우수한 성능을 보이는 것으로 알려졌다.

질감 특징을 이용한 언어 인식 방법에서는 각 언어의 문자가 시각적으로 다르게 나타난다는 사실에 근거하여 문서 영상으로부터 문자 블록을 얻은 다음 이를 질감 패턴으로 보고 그 특징을 추출하여 질감 분류(texture classification) 방식으로 언어를 인식한다<sup>[5~9]</sup>. 그러므로 질감 특징을 이용한 방법에서 가장 중요한 것은 각 언어에 해당하는 문자를 표현할 수 있는 특징을 추출하는 것이라고 할 수 있다. Pearke와 Tan<sup>[5]</sup>은 기존의 질감 분류에서 많이 사용되는 Gabor 필터를 이용하여 문서 영상에서 질감 특징을 추출하였다. 그리고 실험에서 Gabor 특징에 의한 언어 인식이 GLCM(gray-level co-occurrence matrix)<sup>[10]</sup>에 의한 것보다 성능이 우수함을 보였다. Tan<sup>[6]</sup>은 Gabor 필터를 이용하여 회전(rotation)에도 변하지 않는 질감 특징을 추출하였고,

Chan과 Coghill<sup>[7]</sup> 또한 Gabor 필터를 이용하였다. Busch 등<sup>[8]</sup>은 WLC(wavelet log co-occurrence)를 이용하여 질감 특징을 추출하였다. 본 저자들<sup>[9]</sup>은 [11~12]에서 제안된 웨이블릿 영역의 BDIP(block difference of inverse probabilities) 연산자와 BVLC(block variance of local correlation coefficients) 연산자를 이용하여 질감 특징을 추출한 다음 이들을 정규화된 크기 연산자를 이용하여 추출된 질감 특징과 융합(fusion)하여 언어 인식을 위한 질감 특징으로 사용하였다. [5]에서는 7개국 언어, [6]에서는 6개국 언어, [7]에서는 16개국 언어, [8]에서는 8개국 언어, [9]에서는 10개국 언어를 각각 인식하였다. 언어 인식에 대한 전반적인 개요는 [1]을 참조하면 된다.

언어 인식을 위한 질감 특징 추출로 가장 많이 사용된 것으로 알려진 Gabor 필터는 가우시안(Gaussian) 함수를 복소 지수 함수로 변조한 함수를 모 함수(mother function)로 하여 이를 스케일링(scaling)하고 회전하여 얻어진다<sup>[13]</sup>. Gabor 필터는 방향성과 주파수 선택성을 가지며 영상에서 국부적인 특정 주파수 성분과 방향성 정보를 효과적으로 표현할 수 있다.

한편 [9]에서 언어 인식을 위한 질감 특징 추출 연산자로 사용된 BDIP 연산자는 국부 밝기 변화를 국부 최대치로 정규화한 비선형 경사(nonlinear gradient) 연산자로서 국부적인 밝기 변화를 잘 표현하는 것으로 알려져 있다. 그리고 BVLC 연산자는 국부 영역에서 방향에 따른 국부 상관을 국부 분산으로 정규화하여 얻어지는 국부 상관 계수들의 최대 차이를 얻는 연산자로서 질감의 평활성(smoothness)을 잘 표현하는 것으로 알려져 있다. 이들 연산자는 언어 인식 뿐만 아니라 질감 분류<sup>[14]</sup>와 영상 검색<sup>[11~12]</sup>, 얼굴 인식<sup>[15~16]</sup>, ROI determination<sup>[17]</sup>에도 적용되어 우수한 성능을 보여주었다.

질감 특징은 여러 가지 특징을 융합하여 사용할 수 있다. 이러한 특징 융합에 의한 질감 분류 방법은 사용되는 특징 벡터의 차원이 커지면 계산량이 증가하고 경우에 따라서는 성능이 오히려 감소할 수 있다. 따라서 질감 분류 방법에 사용되는 방법을 문자 영상에 적용할 때는 특징 벡터의 차원을 고려하여 가능하면 낮은 차원의 특징으로 높은 인식률을 보이도록 하여야 한다. 그러기 위해서는 문자 영상의 특성을 잘 표현하면서 융합되었을 때 특징 벡터가 상호 보완하면서 시너지 효과를 얻을 수 있고 차원은 낮은 질감 특징들을 선택하여야 한다.

본 논문에서는 Gabor 필터 그리고 BDIP와 BVLC 연산자를 이용하여 문서 영상에서 질감 특징을 추출하고 이들을 이용하여 언어를 인식하는 방법을 제안한다. 제안된 방법에서는 먼저 시험 영상에 질감 분류에 우수한 성능을 보이는 것으로 알려진 Manjunath와 Ma<sup>[13]</sup>가 제안한 Gabor 필터를 적용하여 국부적인 특정 주파수 성분과 방향성 정보와 관련된 질감 특징을 추출한다. 그리고 시험 영상에 웨이브렛 변환을 적용하여 평활(smooth) 부대역과 상세(detail) 부대역을 얻고 각 상세 부대역에는 Donoho의 연역치화(soft-thresholding)<sup>[18]</sup>를 적용하여 잡음을 제거한다. 이어서 각 부대역에 BDIP와 BVLC 연산자를 적용하여 국부적인 밝기 변화와 질감의 평활성에 관련된 질감 특징을 추출한다. 그런 다음 추출된 각 부대역별 특징의 평균과 표준 편차를 구하고 이들을 벡터화한 다음 벡터 융합으로 얻어지는 벡터를 특징 벡터로 사용한다. 분류기로는 얼굴 인식에 주로 사용되는 것으로 알려진 WPCA(whitened principal component analysis)<sup>[16]</sup>를 사용한다.

## II. 질감 특징을 이용한 언어 인식과 질감 특징

### 1. 질감 특징을 이용한 언어 인식

그림 1은 기존의 질감 특징을 이용한 문서 영상에서의 언어 인식의 블록도를 보여준다. 편의상 학습(training)을 위한 영상 DB에는  $K$ 개의 언어를 가지면서 각 언어별로는  $J$ 개의 영상이 포함된 것으로 하자. 먼저 학습 단계에서는 영상 DB에 저장되어 있는 각 영상  $I_{k,j}$ ,  $k=1, \dots, K$ ,  $j=1, \dots, J$ 로부터 질감 특징을 추출한 다음 이들로 부터 특징 벡터  $\mathbf{f}_{k,j}$ 를 생성한다. 기존의 학습에서는 주로 특징 벡터  $\mathbf{f}_{k,j}$ 를 평균하여 각 언어에 대한 대표 특징 벡터 집합  $\{\bar{\mathbf{f}}_k | k=1, \dots, K\}$ 를 얻는다. 그리고 필요하다면 특징 벡터에 대한 공분산 행렬 집합  $\{\mathbf{C}_k | k=1, \dots, K\}$ 를 계산하고 통계치를 특징 DB에 저장한다. 시험(test) 단계에서는 언어 인식을 위한 시험 영상  $I$ 가 들어오면 이로부터 특징 벡터  $\mathbf{f}$ 를 추출하고 분류기에서 다음과 같이 DB에 저장되어 있는 특징 벡터  $\{\bar{\mathbf{f}}_k | k=1, \dots, K\}$  중에서  $\mathbf{f}$ 와 가장 유사한 벡터  $\bar{\mathbf{f}}_c$ 를 찾아 그 인덱스(index)  $c$ 를 출력한다.

$$c = \arg \min_{k \in \{1, \dots, K\}} d(\mathbf{f}, \bar{\mathbf{f}}_k) \quad (1)$$

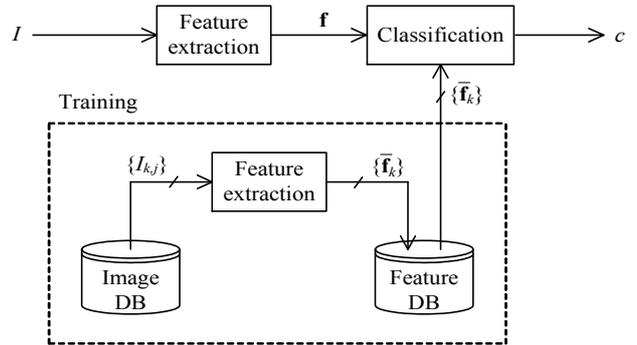


그림 1. 기존의 질감 특징을 이용한 언어 인식의 블록도

Fig. 1. Block diagram of typical texture feature-based language identification.

여기서  $d(\cdot)$ 는 두 벡터간의 거리(distance)를 나타낸다. 이때  $c$ 는 인식된 언어의 인덱스가 된다. 그리고 계산된 공분산 행렬  $\mathbf{C}_k$ 는 (1)식의 거리  $d(\cdot)$ 의 측정에 사용될 수가 있다.

### 2. GLCM과 Haralick 특징

밝기 값이 각각  $i$ 와  $j$ 이면서 거리가  $r$ 이고 각도가  $\theta$ 인 임의의 두 화소 간의 결합 확률 질량 함수(joint probability mass function)를  $P(i, j; r, \theta)$ 라 하자. 그러면 GLCM은 결합 확률 질량 함수의 추정치  $\hat{P}(i, j; r, \theta)$ 를 일컬으며, GLCP(GLC probability)라고도 한다<sup>[10]</sup>. 평균 GLCM  $\hat{P}(i, j)$ 는 다음과 같이 추정된 결합 확률 질량 함수  $\hat{P}(i, j; r, \theta)$ 를 모든 가능한  $(r, \theta)$ 에 대하여 평균하여 얻어진다.

$$\hat{P}(i, j) = \langle \hat{P}(i, j; r, \theta) \rangle_{(r, \theta)} \quad (2)$$

여기서  $\langle \cdot \rangle$ 은 평균 연산자를 나타낸다. Haralick 등<sup>[10]</sup>은 GLCM으로부터 얻게 되는 entropy, variance, moment, correlation 등의 28가지 질감 특징을 제안하였다. 그리고 Holmes 등<sup>[19]</sup>은 (2)식의 평균 GLCM을 이용하여 몇 가지 Haralick의 특징을 추출하였다.

### 3. Gabor 특징

2차원 Gabor 필터에서는 다음과 같이 표현되는 가우시안 함수를 복소 지수 함수로 변조한 함수를 모 함수로 사용한다<sup>[13]</sup>.

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right) + 2\pi j Wx\right] \quad (3)$$

여기서  $\sigma_x$ 와  $\sigma_y$ 는 가우시안 함수의 표준편차를 나타내며 모 함수의 주파수 대역을 정하는 상수로 사용된다. 그리고  $W$ 는 변조 주파수를 나타내며 모 함수의 중심 주파수를 정하는 상수로 사용된다. Gabor 필터는 다음과 같이 (3)식의 모 함수  $g(x,y)$ 를 스케일링하고 회전하여 얻어진다.

$$g_{m,n}(x,y) = a^{-2m}g(x',y') \quad (4)$$

여기서  $a(a > 1)$ 는 스케일링을 위한 상수,  $m(m = 0, \dots, M-1)$ 은 스케일 인자, 그리고  $n(n = 0, \dots, N-1)$ 은 회전 인자를 각각 나타낸다. 그리고  $x' = a^{-m}(x \cos \theta_n + y \sin \theta_n)$ ,  $y' = a^{-m}(-x \sin \theta_n + y \cos \theta_n)$ ,  $\theta_n = n\pi/N$ 이다. (3)식과 (4)식에서 보면 Gabor 필터  $g_{m,n}(x,y)$ 의 주파수 대역은 가우시안 함수의 표준편차  $\sigma_x$ 와  $\sigma_y$ , 그리고 스케일 인자  $m$ 에 의하여 결정되고, 중심 주파수는 변조 주파수  $W$ 에 의하여 결정되며, 방향성은 회전 인자  $n$ 에 의하여 결정된다는 것을 알 수 있다. 영상  $I$ 에 대한 Gabor 변환은 (3)식의 Gabor 필터  $g_{m,n}(x,y)$ 를  $I$ 에 적용하여 얻어지며, Gabor 특징은 Gabor 변환 영상  $GI = \{G^{(m,n)}, m = 0, \dots, M-1, n = 0, \dots, N-1\}$ 에 크기 연산자(magnitude operator)를 적용하여 얻은 영상  $G$ 로부터 추출된다.

#### 4. 공간 영역의 BDIP와 BVLC

공간 영역의 BDIP를 표현하기 위하여 영상  $I$ 의 화소  $p = (x,y)$ 에서의 밝기 값을  $I_p$ 로 하자. 그러면 영상  $I$ 의 화소  $p$ 에서의 BDIP  $D_p$ 는 다음과 같이 정의된다<sup>[11]</sup>.

$$D_p = \frac{\langle \widehat{I}_p - I_{p+q} \rangle_{q \in R}}{\widehat{I}_p} \quad (5)$$

여기서  $\langle \cdot \rangle_{q \in R}$ 은 국부 창(window)  $R$  내에서의 평균 연산자를 나타내며,  $\widehat{I}_p$ 는 다음과 같이 표현되는 화소  $p$ 를 중심으로 하는 국부 창  $R$  내에 있는 화소의 최대값을 나타낸다.

$$\widehat{I}_p = \max_{q \in R} I_{p+q} \quad (6)$$

(5)식에서  $D_p$ 의 분자는 창  $R$  내에서의 밝기 경사의 평

균을 나타낸다. 그러므로 BDIP는 국부 영역에서 최대값으로 정규화된 밝기 경사의 평균을 나타낸다는 것을 알 수 있다.  $\widehat{I}_p$ 는  $D_p$ 의 분모를 안정화하기 위하여  $\widehat{I}_p = \max(\widehat{I}_p, \delta_D)$ 로 클리핑(clipping)된다.

영상  $I$ 의 화소  $p$ 에서의 BVLC  $V_p$ 는 다음과 같이 정의된다<sup>[11]</sup>.

$$V_p = \max_{d \in O} \rho_p(d) - \min_{d \in O} \rho_p(d) \quad (7)$$

여기서  $\rho_p(d)$ 는 다음과 같이 정의되는 영상  $I$ 의 화소  $p$ 에서의 방향  $d$ 에 따른 국부 상관 계수를 나타낸다.

$$\rho_p(d) = \frac{\langle I_{p+d+q} I_{p+q} \rangle_{q \in R} - \mu_{I_{p+d}} \mu_{I_p}}{\sigma_{I_{p+d}} \sigma_{I_p}}, \quad d \in O \quad (8)$$

여기서  $\mu_{I_p}$ 와  $\sigma_{I_p}$ 는 화소  $p$ 에서의 국부 평균과 표준 편차를 각각 나타낸다. 그리고  $\mu_{I_{p+d}}$ 와  $\sigma_{I_{p+d}}$ 는 화소  $p+d$ 에서의 국부 평균과 표준 편차를 각각 나타낸다. 또한  $O$ 는 방향 집합을 나타내며  $O = \{(-a,0), (a,0), (0,-a), (0,a)\}$ 로 선택될 수 있다. (7)식과 (8)식에서 보면 BVLC는 국부 영역 내에서 국부 상관 계수의 최대 편차(deviation)를 나타낸다는 것을 알 수 있다.  $\sigma_{I_{p+d}}$ 와  $\sigma_{I_p}$ 도 상관 계수  $\rho_p(d)$ 의 분모를 안정화하기 위하여  $\delta_V$ 를 역치로 하여 클리핑된다.

#### 5. 웨이브렛 영역의 BDIP와 BVLC

영상  $I$ 에 대한 웨이브렛 변환에서는 그림 2와 같이  $I$ 에 1차원 저대역 필터(LPF)와 고대역 필터(HPF)를 수평, 수직 방향으로 분리적으로 적용하여 평활 부대역  $LL$ 과 3개의 상세 부대역  $LH, HL, HH$ 로 분해한다. 이들 부대역들은 다운 샘플링(downsampling)되어 첫 번째 스케일에서의 4개의 부대역  $\{W^{(1,LL)}, W^{(1,LH)}, W^{(1,HL)}, W^{(1,HH)}\}$ 를 얻는다.  $W^{(1,LL)}$ 은 다시 같은 방법으로 두 번째 스케일에서의 4개의 부대역  $\{W^{(2,LL)}, W^{(2,LH)}, W^{(2,HL)}, W^{(2,HH)}\}$ 로 분해된다. 이러한 과정을  $L$ 번째 스케일까지 적용하여 웨이브렛 영상  $WI = \{W^{(L,LL)}, \{W^{(l,b)}\}, l = 1, \dots, L, b \in \{LH, HL, HH\}\}$ 를 얻는다.

웨이브렛 변환된  $l$ 번째 스케일에서의  $b$  부대역의 화

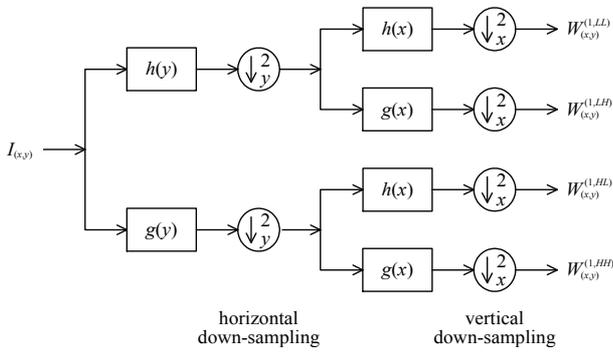


그림 2. 첫 번째 스케일까지 수행되는 웨이브렛 변환의 블록도, 여기서  $h$ 와  $g$ 는 각각 저대역 필터와 고대역 필터를 나타낸다

Fig. 2. Block diagram of wavelet transform up to the first scale, where  $h$  and  $g$  denote low-pass and high-pass filters, respectively.

소  $p$ 에서의 BDIP  $WD_p^{(l,b)}$ 는 다음과 같이 정의된다<sup>[11]</sup>.

$$WD_p^{(l,b)} = \frac{\langle \widehat{W}_p^{(l,b)} - W_{p+q}^{(l,b)} \rangle_{q \in R}}{\widehat{W}_p^{(l,LL)}}, \quad (9)$$

$l = 1, \dots, L, \quad b \in \{LL, LH, HL, HH\}$

여기서  $\widehat{W}_p^{(l,b)}$ 는  $l$ 번째 스케일  $b$  부대역에서 화소  $p$ 를 중심으로 하는 창  $R$ 내에서의 최대값을 나타낸다. 그리고  $W^{(l,b)}$ 의 화소  $p$ 에서의 BVLC  $WV_p^{(l,b)}$ 는 다음과 같이 정의된다<sup>[11]</sup>.

$$WV_p^{(l,b)} = \max_{d \in O} \rho_p^{(l,b)}(d) - \min_{d \in O} \rho_p^{(l,b)}(d), \quad (10)$$

$l = 1, \dots, L, \quad b \in \{LL, LH, HL, HH\}$

여기서  $\rho_p^{(l,b)}(d)$ 는 다음과 같이 정의되는  $W^{(l,b)}$ 의 화소  $p$ 에서의 방향  $d$ 에 따른 국부 상관 계수를 나타낸다.

$$\rho_p^{(l,b)}(d) = \frac{\langle W_{p+d+q}^{(l,b)} W_{p+q}^{(l,b)} \rangle_{q \in R} - \mu_{W_{p+d}^{(l,b)}} \mu_{W_p^{(l,b)}}}{\sigma_{W_{p+d}^{(l,b)}} \sigma_{W_p^{(l,b)}}}, \quad (11)$$

$d \in O$

여기서  $\mu_{W_p^{(l,b)}}$ 와  $\sigma_{W_p^{(l,b)}}$ 는  $l$ 번째 스케일  $b$  부대역에서 화소  $p$ 에 대한 국부 평균과 표준 편차를 각각 나타내고,  $\mu_{W_{p+d}^{(l,b)}}$ 와  $\sigma_{W_{p+d}^{(l,b)}}$ 는 화소  $p+d$ 에 대한 국부 평균과 표준 편차를 각각 나타낸다.

### III. 제안된 질감 특징을 이용한 언어 인식

그림 3은 제안된 질감 특징을 이용한 문서 영상에서의 언어 인식의 블록도를 보여준다. 언어 인식을 위한 시험 영상  $I$ 가 들어오면 먼저 Gabor 변환을 수행하여 Gabor 영상  $GI$ 를 얻고,  $GI$ 에 크기 연산자를 적용하여 얻은 영상  $G$ 로부터 통계치를 계산하여 그 결과를 특징 벡터  $\mathbf{f}_G$ 로 만든다. 다음으로 영상  $I$ 에 대한 웨이브렛 변환을 수행하여 웨이브렛 영상  $WI$ 를 얻고,  $WI$ 의 각 상세 부대역에 연역치화를 적용하여 잡음이 제거된 웨이브렛 영상  $\tilde{W}I$ 를 얻는다. 이어서  $\tilde{W}I$ 에 BDIP와 BVLC 연산자를 적용하여  $WD$ 와  $WV$  영상을 얻고, 이들로부터 통계치를 계산하여 그 결과를 각각 특징 벡터  $\mathbf{f}_{WD}$ 와  $\mathbf{f}_{WV}$ 로 만든다. 이들 특징 벡터  $\mathbf{f}_G$ ,  $\mathbf{f}_{WD}$ ,  $\mathbf{f}_{WV}$ 들은 벡터 융합을 거쳐 최종 특징 벡터  $\mathbf{f}$ 를 생성한다. 분류기에서는 특징 DB에 저장되어 있는 대표 특징 벡터  $\{\bar{\mathbf{f}}_k | k = 1, \dots, K\}$  중에서  $\mathbf{f}$ 와 가장 유사한 벡터  $\bar{\mathbf{f}}_c$ 를 찾아 그 인덱스  $c$ 를 출력한다.

#### 1. 상세 부대역에서의 연역치화

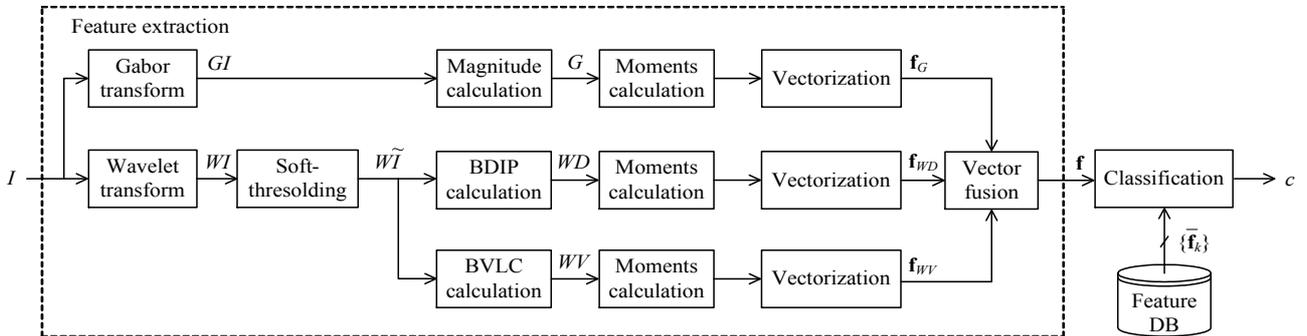


그림 3. 제안된 질감 특징 기반 언어 인식의 블록도

Fig. 3. Block diagram of the proposed texture feature-based language identification.

상세 부대역  $W^{(l,b)}$  의 화소  $p$  에서의 Donoho의 연역 치화는 다음과 같이 표현된다.

$$\widetilde{W}_p^{(l,b)} = \begin{cases} \text{sgn}(W_p^{(l,b)}) (|W_p^{(l,b)}| - \delta_l), & |W_p^{(l,b)}| > \delta_l \\ 0, & \text{otherwise} \end{cases} \quad l = 1, \dots, L, b \in \{LH, HL, HH\} \quad (12)$$

여기서  $\text{sgn}(\cdot)$  은 부호를 나타내고,  $\delta_l$  은  $l$  번째 스케일에서의 역치를 나타낸다. 편의상 상세 부대역에서 연역치화된 웨이브렛 영상을  $\widetilde{W} = \{W^{(L,LL)}, \{\widetilde{W}^{(l,b)}\}, l = 1, \dots, L, b \in \{LH, HL, HH\}\}$  로 표시하기로 하자. (12)식의 역치  $\delta_l$  은 영상  $I$  의 문자줄과 문자줄 사이의 빈 줄에서 추정된 잡음의 표준 편차  $\hat{\sigma}_n$  에 비례하도록, 즉  $\delta_l = \alpha_l \hat{\sigma}_n$  로 정한다.

## 2. 특징 벡터의 생성

특징 벡터  $\mathbf{f}_G$  를 생성하기 위하여 먼저  $m$  번째 스케일  $n$  번째 회전 방향에서의 Gabor 크기 영상  $|G^{(m,n)}|$  으로부터 전역 평균  $\mu_{G^{(m,n)}}$  과 전역 표준 편차  $\sigma_{G^{(m,n)}}$  을 계산하고 그 결과를 다음과 같이 벡터화하여  $\mathbf{f}_{G^{(m,n)}}$  를 얻는다.

$$\mathbf{f}_{G^{(m,n)}} = [\mu_{G^{(m,n)}}, \sigma_{G^{(m,n)}}] \quad (13)$$

이어서 다음과 같이 모든  $m$  과  $n$  에 대하여 얻어지는 벡터를 모아서  $\mathbf{f}_G$  를 생성한다.

$$\mathbf{f}_G = [\mathbf{f}_{G^{(m,n)}} | m = 0, \dots, M-1, n = 0, \dots, N-1] \quad (14)$$

다음으로 특징 벡터  $\mathbf{f}_{WD}$  를 생성하기 위해서는 먼저 웨이브렛 영역의 BDIP 영상  $WD$  의  $l$  번째 스케일  $b$  부대역  $WD^{(l,b)}$  로부터 전역 평균  $\mu_{WD^{(l,b)}}$  와 전역 표준 편차  $\sigma_{WD^{(l,b)}}$  를 계산하고 그 결과를 다음과 같이 벡터화하여  $\mathbf{f}_{WD^{(l,b)}}$  를 얻는다.

$$\mathbf{f}_{WD^{(l,b)}} = [\mu_{WD^{(l,b)}}, \sigma_{WD^{(l,b)}}] \quad (15)$$

그리고는 다음과 같이 모든  $l$  과  $b$  에 대하여 얻어지는 벡터를 모아서  $\mathbf{f}_{WD}$  를 생성한다.

$$\mathbf{f}_{WD} = [\mathbf{f}_{WD^{(l,b)}} | l = 1, \dots, L, b \in \{LL, LH, HL, HH\}] \quad (16)$$

같은 방법으로 특징 벡터  $\mathbf{f}_{WV}$  도  $WV$  영상의  $l$  번째

스케일  $b$  부대역  $WV^{(l,b)}$  로부터 전역 평균  $\mu_{WV^{(l,b)}}$  와 전역 표준편차  $\sigma_{WV^{(l,b)}}$  를 계산하여 그 결과를  $\mathbf{f}_{WV^{(l,b)}}$  로 벡터화하고 모든  $l$  과  $b$  에 대하여 얻어지는  $\mathbf{f}_{WV^{(l,b)}}$  벡터를 모아서 생성된다. 이때 BVLC 계산은 다음의 8 방향에 대하여 수행한다.

$$O = \{(u,v) | u,v \in \{-a, 0, a\}, (u,v) \neq (0,0)\} \quad (17)$$

## 3. 벡터 융합

특징 벡터  $\mathbf{f}_G, \mathbf{f}_{WD}, \mathbf{f}_{WV}$  는 그 값의 분포 범위가 서로 다를 수 있으므로 이들을 융합하되 그 값의 분포 범위가 같도록 각각 정규화(normalization)하면 효율적인 언어 인식이 되도록 할 수 있다. 이를 위하여 먼저 특징 벡터  $\mathbf{f}_G, \mathbf{f}_{WD}, \mathbf{f}_{WV}$  를 다음과 같이 모아서 융합된 특징 벡터  $\tilde{\mathbf{f}}$  를 얻는다.

$$\tilde{\mathbf{f}} = [\mathbf{f}_G, \mathbf{f}_{WD}, \mathbf{f}_{WV}] \quad (18)$$

그런 다음  $\tilde{\mathbf{f}}$  는 다음과 같이 특징 벡터  $\mathbf{f}_G, \mathbf{f}_{WD}, \mathbf{f}_{WV}$  의 표준 편차 벡터  $\sigma_G, \sigma_{WD}, \sigma_{WV}$  로 정규화 된다.

$$\mathbf{f} = \Sigma^{-1} \tilde{\mathbf{f}} \quad (19)$$

여기서  $\Sigma = \text{diag}[\sigma_G, \sigma_{WD}, \sigma_{WV}]$  이다. 이때  $\sigma_G$  는 학습 단계에서  $k$  번째 언어에 대한 Gabor 특징 벡터  $\mathbf{f}_{G_k}$  의 표준 편차 벡터  $\sigma_{G_k}$  를  $k$  에 대하여 평균하여 얻은 벡터를 나타낸다. 즉  $\sigma_G = \langle \sigma_{G_k} \rangle_k$  가 된다. 그리고  $\sigma_{WD}$  와  $\sigma_{WV}$  도  $WD, WV$  특징 벡터  $\mathbf{f}_{WD_k}, \mathbf{f}_{WV_k}$  에 대하여 같은 방식으로 얻어지는 양들을 나타낸다.

## 4. WPCA를 이용한 분류기

DB 내의  $k$  번째 언어에 해당하는 대표 특징 벡터로서  $J \times 1$  차원을 갖는  $\bar{\mathbf{f}}_k$  에 대한 WPCA는  $\bar{\mathbf{f}}_k$  에 대한 PCA 과정과 화이트닝(whitening) 과정으로 구성된다<sup>[6]</sup>. 편의상 대표 특징 벡터  $\bar{\mathbf{f}}_k$  의 공분산 행렬을  $\mathbf{C}_k$  로 표시하고, 내림차순으로 정렬된  $\mathbf{C}_k$  의 고유치와 고유 벡터를 각각  $\{\lambda_{k,i} | i = 1, \dots, U\}$  와  $\{\phi_{k,i} | i = 1, \dots, U\}$  로 표시하자. 그러면 대표 특징 벡터  $\bar{\mathbf{f}}_k$  에 대한 WPCA 는 다음과 같이 주어진다.

$$\bar{\mathbf{y}}_k = \tilde{\Lambda}_k^{-1/2} \Phi_k^T \bar{\mathbf{f}}_k \quad (20)$$

여기서  $\Phi_k$ 와  $\tilde{\Lambda}_k$ 는 각각 다음과 같이 고유 벡터로 이루어진 행렬과 고유치로 이루어진 대각선 행렬을 나타낸다.

$$\Phi_k = [\phi_{k,1}, \phi_{k,2}, \dots, \phi_{k,U}] \quad (21)$$

$$\tilde{\Lambda}_k = \text{diag}[\lambda_{k,1}, \lambda_{k,2}, \dots, \lambda_{k,V}, \delta, \dots, \delta] \quad (22)$$

이때 고유치는 안정화를 위하여 역치  $\delta$ 로 클리핑(clipping)되었다. 얼굴 인식에 사용되는 WPCA에서는 공분산 행렬을 모든 클래스에서 얻고 역치 보다 큰 고유치에 해당하는 고유 벡터만을 사용한다. 반면에 본 논문에서 사용하는 WPCA에서는 공분산 행렬을 각 클래스, 즉 각 언어에서 얻고 해당 언어의 모든 고유벡터를 사용한다.

본 논문에서 사용하는 WPCA를 이용한 분류기에서 시험 영상  $I$ 의 특징 벡터  $\mathbf{f}$ 와  $k$ 번째 언어에 해당하는 대표 특징 벡터  $\bar{\mathbf{f}}_k$ 의 거리는 (24)식으로 주어지는  $\mathbf{f}$ 에 대한 WPCA  $\mathbf{y}_k^{ts}$ 와 (20)식의  $\bar{\mathbf{y}}_k$  사이의 cosine 거리로서 다음과 같이 표현된다.

$$d(\mathbf{f}, \bar{\mathbf{f}}_k) = -\frac{\mathbf{y}_k^{ts} \cdot \bar{\mathbf{y}}_k}{\|\mathbf{y}_k^{ts}\| \|\bar{\mathbf{y}}_k\|} \quad (23)$$

여기서  $\mathbf{y}_k^{ts}$ 는 다음과 같이 주어진다.

$$\mathbf{y}_k^{ts} = \tilde{\Lambda}_k^{-1/2} \Phi_k^T \mathbf{f} \quad (24)$$

그리고  $\cdot$ 는 벡터의 내적(inner product)을 나타내고,  $\|\cdot\|$ 는 벡터의 크기를 나타낸다.

### IV. 실험 결과 및 검토

본 논문에서 제안된 언어 인식 방법의 성능을 평가하기 위하여 영어, 그리스어, 러시아어, 히브리어, 페르시아어, 힌디어, 태국어, 중국어, 일본어, 한국어

아어, 힌디어, 태국어, 중국어, 일본어, 한국어의 10개국 언어에 대하여 문서 영상 DB를 구축하고 시뮬레이션을 수행하였다. 문서 영상 DB를 구축하기 위하여 먼저 각 언어별로 문서를 스캔하여 모 영상(mother image)을 획득하였다. 이때 각 모 영상에는 두 가지의 폰트를 가지며 각 폰트 별로 정상적인 영상, 1.5도와 3.0도 만큼 회전된 영상, 그리고 0.8:1로 스케일링 된 영상의 총 8 종류의 영상을 가지도록 하였다. 각 모 영상은 128×128 크기의 영상으로 50장씩 나누어져 각 언어 별로 총 400장의 영상을 갖도록 하였다. 이때 각 언어별 400장의 영상 중에서 절반인 200장은 시험 영상으로, 나머지 절반인 200장은 학습 영상으로 사용하되 실험 조건에 따라서 학습 영상의 수를 16~200장으로 변화를 주었다. 그림 4는 실험에 사용된 10개국 언어 영상의 샘플을 보여준다.

실험에서는 제안된 언어 인식 방법을 Haralick의 특징, Gabor 특징, WLC 특징<sup>[8]</sup>, 그리고 웨이브렛 영역의 BDIP와 BVLC 특징에 의한 방법과 성능 비교를 하였다. 실험에서 사용된 Haralick 특징은 [8]에서 사용된 것으로서 거리가  $r=1,2$ 이고 각도는  $\theta=0, 45, 90, 135^\circ$ 인 GLCM을 이용한 energy, entropy, inertia, contrast, local homogeneity, cluster shade, cluster performance, information measure of correlation의 8가지이다. Gabor 필터를 위한 (3)식과 (4)식의 상수는  $\sigma_x=0.9809$ ,  $\sigma_y=1.2873$ ,  $W=0.4$ ,  $a=2.8284$ ,  $M=3$ ,  $N=4$ 로 하였다. 그리고 웨이브렛 변환은 Haar 필터를 사용하여 다운 샘플링을 하면서 첫 번째 스케일까지 수행하였다( $L=1$ ). 웨이브렛 변환을 수행할 때 영상 경계에서는 기 대칭(odd symmetry)을 사용하였다. 또한 웨이브렛 영역에서의 Donoho의 연역치화의 역치  $\delta_1$ 을 위한 비례상수  $\alpha_1$ 는 1.75로 하였고, BDIP와 BVLC 계산에서 사용되는 국부 창의 크기는 3×3으로 하였다. 따라서 (17)식에서  $a=1$ 이 된다. BDIP와

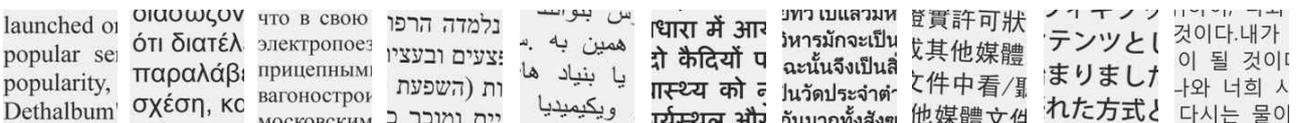


그림 4. 실험에 사용된 10개국 언어 영상의 샘플: 왼쪽부터 오른쪽 순서로 영어, 그리스어, 러시아어, 히브리어, 페르시아어, 힌디어, 태국어, 중국어, 일본어, 한국어

Fig. 4. Sample subimages of ten languages used in experiments: English, Greek, Russian, Hebrew, Persian, Hindi, Thai, Chinese, Japanese, Korean from left to right.

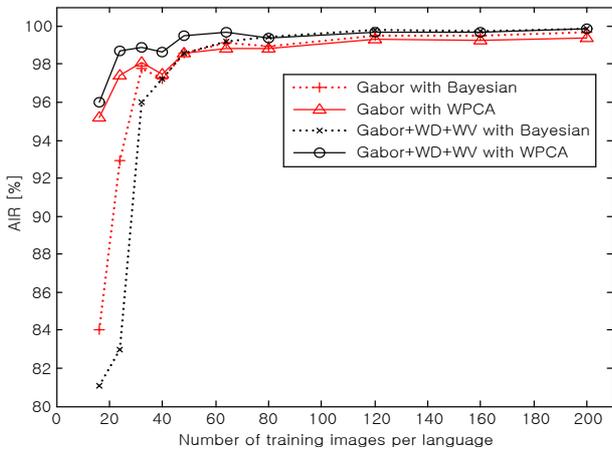


그림 5. 각 언어별 학습 영상의 수에 따른 Gabor 특징에 의한 방법과 제안된 방법의 Bayesian과 WPCA 분류기별 AIR 성능

Fig. 5. AIR performances of Gabor feature-based method and the proposed method with Bayesian and WPCA classifiers according to the number of training images per language.

BVLC의 안정화를 위한 역치  $\delta_D$ 와  $\delta_V$ 는 각각 실험적으로  $\delta_D = 2$ 와  $\delta_V^2 = 10^{-3}$ 으로 정하였다. WPCA에서의 고유치에 대한 역치  $\delta$ 는 모든 학습 영상에 대한 고유치를 내림차순으로 정렬한 다음 하위  $\varepsilon$ %를 역치화할 수 있는 값으로 결정하였다. 언어 인식 성능의 평가 척도로는 시험 영상의 수에 대한 정확하게 인식된 영상의 수의 비를 나타내는 평균 인식률(AIR: averaged identification rate)을 사용하였다.

그림 5는 각 언어별 학습 영상의 수에 따른 Gabor 특징에 의한 방법과 제안된 방법의 AIR 성능을 [9]에서 사용된 안정화된 Bayesian 분류기에 의한 경우와 본 논문에서 사용하는 WPCA 분류기에 의한 경우에 따라 보여준다. WPCA 분류기에서 고유치에 대한 역치  $\delta$ 를 위한  $\varepsilon$ 는 학습 영상의 수에 따라 0.1~1.0으로 선택되었다. 그림 5에서 WD와 WV는 각각 웨이블릿 영역의 BDIP와 BVLC 특징을 나타내며, Gabor+WD+WV는 이들이 Gabor 특징에 융합된 본 논문에서 제안된 특징을 나타낸다. 그림 5에서 보면 WPCA 분류기를 사용한 경우는 시험 영상의 수가 80 이하로 적을 때 안정화된 Bayesian 분류기를 사용한 경우에 비하여 성능이 우수하다는 것을 알 수 있다. 또한 제안된 Gabor+WD+WV 특징은 모든 시험 영상의 수에서 96.00%~99.85%의 AIR 성능을 보여 95.20%~99.35%의 성능을 보이는 Gabor 특징에 비해 0.4%~1.3% 정도 성능이 향상되어 결과적으로 에러율이 16.7%~76.9% 감소됨을 알 수 있

다. 이들 결과로부터 본 논문에서 사용된 특징들의 융합과 WPCA 분류기가 성능 향상을 가져온다는 것을 알 수 있다. 또한 실험 결과 Haralick 특징은 안정화된 Bayesian 분류기를 사용하는 경우에는 70.75%~82.45%의 AIR 성능을 보이고 WPCA를 사용하는 경우에는 62.10%~67.70%의 성능을 보이는 것으로 나타나서, 제안된 방법이 Haralick 특징에 의한 방법보다 성능이 훨씬 우수하다는 것을 알 수 있었다. 이외에도 제안된 방법의 연역치화에 의한 잡음 제거는 학습 영상의 수에 따라 최대 0.6%의 성능 향상을 가져온다는 것을 실험을 통하여 확인할 수 있었다.

표 1은 Gabor 특징에 WD와 WV 특징을 융합했을 때의 성능 향상 효과의 정도를 알아보기 위하여, 그림 5의 결과에서 특징 융합에 의하여 가장 큰 성능 향상을 보인 각 언어별 학습 영상의 수가 40인 경우의 각 특징에 따른 AIR 성능을 보여준다. 표 1에서 보면 Haralick 특징은 62.10%, Gabor 특징은 97.50%, WLC 특징은 93.70%, 그리고 융합된 WD+WV 특징은 86.20%의 AIR

표 1. 각 언어 별 학습 영상의 수가 40인 경우의 각 특징별 AIR 성능

Table 1. AIR performances according to features in case of 40 training images.

Types of texture feature	Haralick	Gabor	WLC	WD+WV	Gabor+WD+WV
AIR [%]	62.10	97.50	93.70	86.20	98.65
Feature dimension	8	24	96	16	40

표 2. 각 언어 별 학습 영상의 수가 40인 경우의 제안된 방법에 대한 분류 행렬, 여기서 AL은 실제 언어, IL은 인식된 언어를 나타낸다

Table 2. Confusion matrix for the proposed method in case of 40 training images per language, where AL (IL) denotes actual (identified) language.

AL \ IL	Eng	Gre	Rus	Heb	Per	Hin	Tha	Chn	Jap	Kor
Eng	99.5	1.5	4.0	0	0	0	0	0	0	0
Gre	0	97.0	1.0	0	0	0	0	0	0	0
Rus	0.5	0.5	95.0	0	0	0	0	0	0	0
Heb	0	0	0	100	0	0	0	0	0	0
Per	0	0	0	0	100	0	0	0	0	0
Hin	0	0	0	0	0	100	0	0	0	0
Tha	0	1.0	0	0	0	0	100	0	0	0
Chn	0	0	0	0	0	0	0	97.5	2.5	0
Jap	0	0	0	0	0	0	0	1.0	97.5	0
Kor	0	0	0	0	0	0	0	1.5	0	100
Error [%]	0.5	3.0	5.0	0	0	0	0	2.5	2.5	0

성능을 각각 보인다는 것을 알 수 있다. 반면에 제안된 Gabor+WD+WV 특징은 98.65%의 AIR 성능을 보여 WD+WV 특징이 Gabor 특징에 융합됨으로써 1.15%의 성능 향상으로 에러율이 절반 가까이 감소됨을 알 수 있다.

표 2는 각 언어별 학습 영상의 수가 40인 경우에 제안된 방법의 분류 행렬(confusion matrix)을 보여준다. 표 2에서 보면 제안된 방법이 영어, 그리스어, 러시아어, 중국어, 일본어에 대하여 0.5%~5.0%의 에러를 보이고, 히브리어, 페르시아어, 힌디어, 태국어는 에러가 없다는 것을 알 수 있다. 또한 가장 우수한 성능을 보이는 학습 영상의 수가 200인 경우에는 페르시아어와 중국어에 대하여 각각 1.0%와 0.5%의 아주 낮은 에러를 보이고, 나머지 8개국 언어에 대하여는 에러가 없다는 것을 실험을 통하여 알 수 있었다. 이상의 결과들로부터 제안된 방법은 특징 벡터의 차원이 비교적 낮은 40이면서도 우수한 언어 인식 성능을 보인다는 것을 알 수 있다.

## V. 결 론

본 논문에서는 Gabor 특징과 웨이브렛 영역의 BDIP와 BVLC 특징을 이용한 문서 영상에서의 질감 특징 기반 언어 인식 방법을 제안하였다. 제안된 방법에서는 Gabor 특징에 국부적인 밝기 변화와 질감의 평활성에 관련된 질감 특징을 추출하는 웨이브렛 영역의 BDIP와 BVLC를 융합한 것을 질감 특징으로 사용하였고 분류기로는 WPCA를 사용하였다. 실험 결과 제안된 방법은 실험 문서 영상 DB에 대하여 96.00%~99.85%의 AIR 성능을 보여 Gabor 특징에 의한 방법보다 에러율이 16.7%~76.9% 감소됨을 보여주었다. 그리고 기존의 Haralick 특징에 의한 방법보다는 월등한 성능 향상을 보였다. 또한 WPCA 분류기를 사용함으로써 각 언어별 시험 영상의 수가 80이하로 적을 때 기존의 안정화된 Bayesian 분류기를 사용하는 경우에 비하여 성능이 우수함을 보였다.

## 참 고 문 헌

- [1] D. Ghosh, T. Dube, and A. P. Shivaprasad, "Script recognition - a review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, Jan. 2010.
- [2] J. Hochberg, L. Kerns, P. Kelly, and T. Thomas, "Automatic script identification from document images using cluster-based templates," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 2, pp. 176-181, Feb. 1997.
- [3] A. L. Spitz, "Determination of the script and language content of document images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 3, pp. 235-245, Mar. 1997.
- [4] L. Shijian and C. L. Tan, "Script and language identification in noisy and degraded document images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 1, pp. 14-24, Jan. 2008.
- [5] G. S. Pearke and T. N. Tan, "Script and language identification from document images," in *Proc. IEEE Workshop on Document Image Analysis 97*, San Juan, Puerto Rico, Jun. 1997, pp. 10-17.
- [6] T. N. Tan, "Rotation invariant texture features and their use in automatic script identification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 7, pp. 743-756, Jul. 1998.
- [7] W. Chan and G. Coghil, "Text analysis using local energy," *Pattern Recognit.*, vol. 34, no. 12, pp. 2523-2532, Dec. 2001.
- [8] A. Busch, W. W. Boles, and S. Sridharan, "Texture for script identification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 11, pp. 1720-1732, Nov. 2005.
- [9] W. S. Lee, N. C. Kim, and I. H. Jang, "Texture feature-based language identification using wavelet-domain BDIP, BVLC, and NRMA features," in *Proc. IEEE International Workshop on Machine Learning for Signal Processing 2010*, Kittilä, Finland, Aug./Sep. 2010, pp. 444-449.
- [10] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 6, pp. 610-621, Nov. 1973.
- [11] Y. D. Chun, S. Y. Seo, and N. C. Kim, "Image retrieval using BDIP and BVLC moments," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 9, pp. 951-957, Sep. 2003.
- [12] Y. D. Chun, N. C. Kim, I. H. Jang, "Content-based image retrieval using multiresolution color and texture features," *IEEE Trans. Multimedia*, vol. 10, no. 6, pp. 1073-1084, Oct. 2008.
- [13] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Mach. Intell.*,

- vol. 18, no. 8, pp. 837-842, Aug. 1996.
- [14] H. J. So, M. H. Kim, and N. C. Kim, "Texture classification using wavelet-domain BDIP and BVLC features," in *Proc. 17th European Signal Processing Conf.*, Glasgow, Scotland, Aug. 2009, pp. 1117-1120.
- [15] H. J. So, M. H. Kim, Y. S. Chung, and N. C. Kim, "Face detection using sketch operators and vertical symmetry," *FAQS-2006, Lecture Notes in Artificial Intelligence*, vol. 4027, pp. 541-551, Jun. 2006.
- [16] Y. A. Ju, H. J. So, N. C. Kim, and M. H. Kim, "Face recognition using local statistics of gradients and correlations," in *Proc. 18th European Signal Processing Conf.*, Aalborg, Denmark, Aug. 2010, pp. 1169-1173.
- [17] T. D. Nguyen, S. H. Kim, and N. C. Kim, "An automatic body ROI determination for 3D visualization of a fetal ultrasound volume," *KES-2005, Lecture Notes in Artificial Intelligence*, vol. 3682, pp. 145-153, Sep. 2005.
- [18] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inform. Theory*, vol. 41, no. 3, pp. 613-627, May 1995.
- [19] Q. A. Holmes, D. R. Neusch, and R. A. Shuchman, "Textual features for image classification," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-22, no. 2, pp. 113-120, Mar. 1984.

— 저 자 소 개 —



**장 익 훈**(정회원)  
 1986년 경북대학교 전자공학과  
 학사 졸업.  
 1988년 경북대학교 전자공학과  
 석사 졸업.  
 1998년 경북대학교 전자공학과  
 박사 졸업.

1988년 2월~1994년 2월 국방과학연구소 연구원  
 2005년 1월~2006년 12월 한국정보처리학회  
 편집위원  
 1998년~현재 경운대학교 디지털전자공학과  
 부교수  
 <주관심분야 : 영상처리, 패턴인식, 컴퓨터비전,  
 멀티미디어>



**이 우 신**(학생회원)  
 2009년 경북대학교 전자공학과  
 학사 졸업.  
 2009년~현재 경북대학교 전자공  
 학과 석사과정.  
 <주관심분야 : 영상처리, 패턴인  
 식>



**김 남 철**(정회원)-교신저자  
 1978년 서울대학교 전자공학과  
 학사 졸업.  
 1980년 한국과학기술원 전기및  
 전자공학과 석사 졸업.  
 1984년 한국과학기술원 전기및  
 전자공학과 박사 졸업.

1984년~현재 경북대학교 IT 대학 전자공학부  
 교수  
 2002년 9월~2004년 8월 경북대학교 전자전기  
 컴퓨터 학부장  
 2005년 9월~2008년 8월 경북대학교  
 정보전산원장  
 <주관심분야 : 영상처리, 패턴인식, 초음파 신호  
 처리, 컴퓨터 비전, 영상압축>