

MPEG-7 기술자를 이용한 알파벳 인식

진 성 호, 강 호 경, 정 광 서, 노 용 만
한국 정보 통신 대학원 대학교 영상/비디오 시스템 연구실
전화 : 042-866-9251 / 핸드폰 : 011-9759-7698

Alphabet Recognition by Using MPEG-7 Descriptors

Sung Ho Jin, Ho Kyung Kang, Kwang Suh Jung, Yong Man Ro
Image/Video System Lab, Information and Communications University
E-mail : wh966@icu.ac.kr

Abstract

This paper has been studied for alphabet recognition by using MPEG-7 descriptors. As an application of MPEG-7, the algorithms in this paper have employed Homogeneous Texture descriptor and Edge Histogram descriptor. For confident matching a voting system is used with a tree structure,

따라서 homogeneous texture 와 edge histogram 기술자를 이용한 문자 인식 알고리즘을 구현하고 이를 실험해 결과를 구해 보고자 한다.

I. 서론

최근의 인터넷과 멀티미디어 표준화 기술의 급속한 발달 및 보급으로 디지털 형태의 비주얼 멀티미디어 데이터 베이스는 기하급수적으로 증가하고 있어 이러한 방대한 양의 비주얼 데이터의 효율적인 관리, 검색 및 인식에 대한 필요성에서 ISO/IEC JT11/SC/WG11에서는 MPEG-7을 통한 멀티미디어 검색 및 표현 기술들을 표준화하고 있다. 이러한 경향과 더불어 현재 표준화되고 있는 MPEG-7을 이용한 응용분야 개발의 중요성이 대두되고 있다.

본 논문은 그 중요성이 점차 강조되고 있는 MPEG-7의 기술자로 사용중인 homogeneous texture 와 edge histogram을 이용하여 alphabet에 대한 문자 인식을 수행한다. 이러한 시도는 MPEG-7의 응용폭을 확장시키는 역할을 담당할 것이다.

II. MPEG-7 기술자

2.1 Homogeneous texture 기술자

Homogeneous texture 기술자는 2차원 영상인 원의 이미지를 radon 변환을 통해서 1차원의 sinogram 얻어 1차원 DFT를 수행한 sinogram을 직교 좌표계에 배치한 것이다. 직교 좌표계에서 Radon 변환된 데이터는 저주파와 중간 주파수 영역에서 보다 조밀한 샘플 데이터를 가지게 되며, 고주파 영역에서는 적은 샘플 데이터를 가지게 된다. 이것은 휴먼 비주얼 특성과 관련된 저주파 정보의 강조 및 고주파 정보의 중요도 감소로 설명된다. 그림 1에서 주파수 이미지를 180도를 기준으로 방사 방향으로 각각 5개, 각 방향으로 6개씩 총 30개 채널로 나눠 휴먼 비주얼 특성을 이용하는 것을 보여주고 있다. 이렇게 나눈 각 채널의 energy 와 분산 정보를 추출하여 homogeneous texture의 특 벡터로 이용한다.[1]

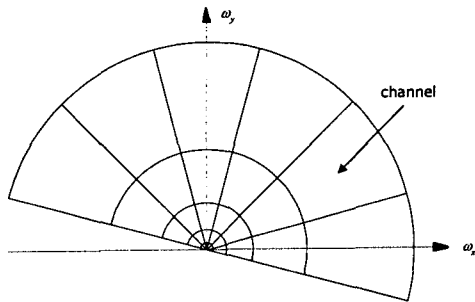


그림 1 휴먼 비주얼 필터에 의한 주파수 공간 분할

2.2 Edge histogram 기술자

Edge histogram은 원래의 이미지를 16개의 loc block으로 나눠 각각의 block에 대해 그림 2에서 보여주는 것처럼 수직, 수평, 45도, 135도, 무 방향성의 5 edge 정보를 추출하여 local edge histogram을 구한

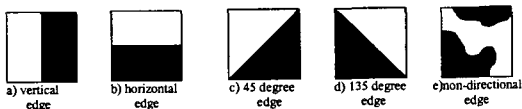


그림 3 5가지 형태의 에지들

그리고 이 정보들을 이용하여 global 과 semi-global histogram을 구해서 16 x 5개의 local edge, 5개의 global edge 과 13 x 5개의 semi-global edge bin들을 이용하여 이미지를 검색한다. semi-global histogram을 구하기 위한 13개의 sub-image 집합을 그림 3에서 보여준다.[2,3]

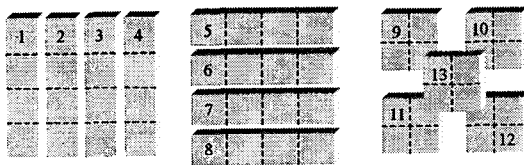


그림 4 semi-global histogram위한 sub-image들의 집합

III. 인식 알고리즘 구현

인식 알고리즘은 homogeneous texture를 위주로 하여 edge histogram을 사용한 알고리즘과 edge histogram을 위주로 하여 homogeneous texture를 사

용한 알고리즘 두 가지 형태로 구현하였다.

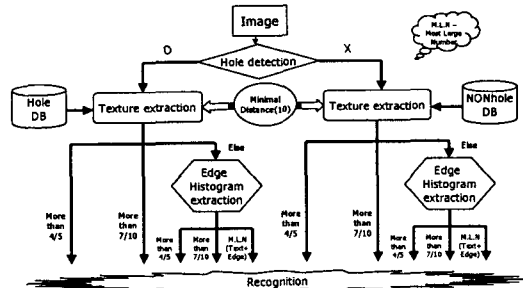


그림 5 homogeneous texture를 위주로 한 문자 인식 알고리즘

3.1 Homogeneous texture 위주 알고리즘

(1) 전처리 과정

Alphabet의 특성을 고려해서 전처리 과정으로 hole detection을 구현한다. 대·소문자 52개의 alphabet을 hole의 유·무를 통해서 hole이 있는 경우 15가지와 hole이 없는 경우 37가지 두 개의 집합으로 나누어서 각각의 데이터 베이스를 구축하게 하였다.

(2) Homogeneous texture 특징 추출 및 검색

전처리 과정을 통해서 분리된 이미지에 대해 homogeneous texture 특징을 추출한다. 이 알고리즘은 MPEG-7에서 구현된 알고리즘에서 각 채널별 완화 효과를 주기 위해 사용된 gabor filter를 시간 단축을 목적으로 사용하지 않았다. 입력된 이미지의 특징을 추출하여 데이터 베이스에 있는 이미지들과 비교하여 검색한다. 검색 방법은 트리 구조를 이용하여 3단계로 나뉜다. 먼저 그림 5에서 보이는 것처럼 가장 거리가 가까운 이미지 5개를 추출하여 5개중 4개 이상이 같은 문자를 나타내는 이미지이면 인식으로 인정한다. 여기서 검색되지 않는다면 상위 10개의 이미지를 추출하여 7개 이상이 같은 문자를 나타내면 인식으로 인정한다. 나머지의 경우 신뢰성을 위하여 edge histogram 통해서 인식을 시도한다.

(3) Edge histogram 특징 추출 및 검색

Edge histogram은 1, global 과 semi-global의 개수는 150개 bin들을 사용한다. 검색 방법은 homogeneous texture의 검색 방법과 유사하다. 위 입력된 이미지의 특징을 데이터 베이스에 있는 이미지와 비교하여 가장 거리가 가까운 상위 5개의 이미지를 검색하여 4개 이상이 같은 문자를 나타내면 검색으로

MPEG-7 기술자를 이용한 알파벳 인식

인정한다. 검색되지 않을 경우 상위 10개의 이미지를 검색하여 7개 이상이면 검색으로 인정한다. 나머지의 경우는 앞서 검색한 homogeneous texture의 상위 5개의 문자 이미지와 edge histogram의 통해 검색한 상위 5개의 문자 이미지들 중 가장 많은 수의 이미지를 가진 문자를 인식으로 인정한다.

3.2 Edge histogram 위주 알고리즘

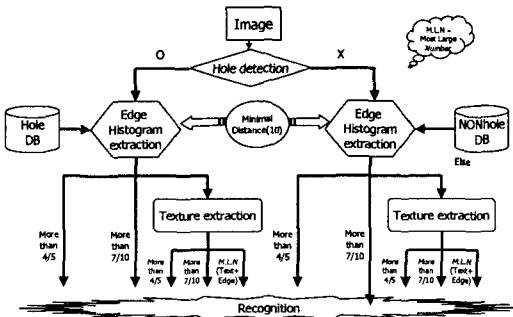


그림 6 Edge histogram을 위주로 한 문자 인식 알고리즘

그림 6에서 보여주는 것처럼 homogeneous texture 위주의 알고리즘과 edge histogram 위주의 알고리즘의 차이는 두 개의 위치만 변화시킨 것이다. 나머지 진처리 과정이나 검색 과정은 동일하다.

IV. 실험

실험에 사용된 alphabet 데이터 베이스는 서로 다른 10개 폰트, 520자로 구성되어 있다. 128 x 128 크기 이미지와 64 x 64 크기 이미지 두 가지 형태로 실험을 하였다. hole의 유·무를 통해서 각각 150개와 370개로 구성된 데이터 베이스로 나눌 수 있다.

4.1 128 x 128 이미지를 사용한 경우

(1) hole이 존재하는 경우

각 문자마다 10개의 폰트를 지니고 있고 이 문자 이미지들을 각각 입력 이미지로 하여 실험한 결과를 그림 7에서 보여주고 있다. 전체 150개 중에서 homogeneous texture 위주 알고리즘으로 검색했을 때 146개를 인식하고 edge histogram 위주 알고리즘의 경우 143개를 인식하였다.

Character	Texture 위주	Edge histogram 위주
A	1.0	9
B	1.0	9
C	1.0	9
D	1.0	1.0
E	1.0	1.0
F	1.0	1.0
G	1.0	1.0
H	1.0	1.0
I	1.0	1.0
J	1.0	8
K	1.0	1.0
L	1.0	1.0
M	1.0	8
N	1.0	1.0
O	1.0	1.0
P	1.0	1.0
Q	1.0	8
R	1.0	1.0
S	1.0	8
T	1.0	1.0
U	1.0	1.0
V	1.0	1.0
W	1.0	1.0
X	1.0	1.0
Y	1.0	1.0
Z	1.0	1.0
합계	146	143
인식률	97.33333333	95.33333333

그림 7 hole이 존재하는 이미지 인식 결과

(2) hole이 존재하지 않는 경우

전체 370개의 이미지 중 homogeneous texture 위주의 인식 알고리즘 결과는 361개이고 edge histogram 위주의 알고리즘 결과는 357개이다. 이 결과는 그림 8에서 볼 수 있다.

Character	Texture 위주	Edge histogram 위주
C	1.0	1.0
E	1.0	1.0
G	9	9
F	1.0	1.0
I	1.0	1.0
H	9	9
N	1.0	1.0
O	1.0	1.0
J	1.0	8
K	1.0	1.0
L	1.0	1.0
M	1.0	1.0
P	1.0	1.0
Q	1.0	1.0
R	1.0	1.0
S	1.0	1.0
T	1.0	1.0
U	1.0	1.0
V	1.0	1.0
W	1.0	1.0
X	1.0	1.0
Y	1.0	1.0
Z	1.0	1.0
합계	361	357
인식률	97.56756757	96.48648648

그림 8 hole이 존재하지 않는 이미지 인식 결과

(3) 전체 결과

Edge histogram 위주 알고리즘의 검색 결과보다 homogeneous texture 위주 알고리즘의 검색 결과가 약간 우월함을 보였다.

	Texture 위주	Edge histogram 위주
hole 유	142	129
hole 무	361	357
합계	503	486
인식률	98.79076923	93.48163948

그림 9 128 x 128 크기 이미지 인식 결과

4.2 64 x 64 이미지를 사용한 경우

(1) hole이 존재하는 경우

64 x 64 크기 이미지의 경우 hole 존재하는 경우는

인식결과에서 두 알고리즘 사이 큰 차이를 보이고 있다. 그림 10에서 이 차이를 보여주고 있다.

Character	Texture 위주	Edge histogram 위주
A	10	7
B	7	6
b	10	8
D	10	6
d	10	10
e	9	8
g	9	7
O	10	10
o	10	10
P	10	9
p	10	8
q	10	10
R	9	10
r	9	10
합 계	142	129
인식률	94.86486687	85

그림 10 hole이 존재하는 이미지 인식 결과

(2) hole이 존재하지 않는 경우

hole이 존재하지 않는 경우는 128 x 128 크기의 이미지와 비슷한 결과를 그림 11에서 보여주고 있다.

Character	Texture 위주	Edge histogram 위주
C	10	10
c	10	10
E	9	9
F	10	10
f	9	9
G	10	10
H	10	10
h	10	10
I	10	10
i	7	6
J	10	10
j	10	10
K	10	10
k	10	10
L	10	10
l	10	10
M	7	7
m	10	10
N	10	9
n	10	10
r	10	10
S	10	10
s	10	10
T	10	10
t	10	9
U	10	10
u	10	10
V	10	10
v	10	10
W	10	10
w	10	10
X	10	10
x	10	10
Y	10	9
y	10	10
Z	10	10
z	9	9
합 계	361	357
인식률	97.50756757	96.41848649

그림 11 hole이 존재하지 않는 이미지 인식 결과

(3) 전체 결과

64 x 64 크기 이미지의 인식 결과는 전체적으로 128 x 128 크기 이미지에 비해서 약간 떨어지지만 homogeneous texture 위주 알고리즘의 경우 비슷한 인식결과를 그림 12에서 보여주고 있다.

	Texture 위주	Edge histogram 위주
hole이 있음	142	129
hole이 없음	361	357
합 계	503	486
인식률	96.73276923	93.46155846

그림 12 64 x 64 크기 이미지 인식 결과

4.3 실험조건

다음과 같은 경우 같은 인식으로 인정하였다.

(1) 대·소문자가 유사한 경우

C·c, O·o, P·p, S·s, U·u, V·v, W·w, X·x, Z·z 이와 같은 경우 이미지의 정규화 과정을 통해서 같은 크기의 형태로 되어 같은 문자로 인정하였다.

(2) 특수한 경우

I·l 의 경우 같은 문자로 인정하였다.

V. 결론

현재 표준화되고 있는 MPEG-7의 응용분야를 연구하는 것은 표준화 못지 않게 중요한 분야로 대두되고 있다.

본 논문에서 그 한 분야로 alphabet 문자 인식을 구현하였고 실험 결과를 통해서 알 수 있듯이 충분한 가능성을 엿 볼 수 있다. 문제점인 이미지 정규화를 통해서 나타나는 대·소문자의 모호함을 해결하는 방법이 필요하며 앞으로 실제 인쇄물에 대한 문자 인식 실험을 실행 예정이다.

본 논문에서 보여주고 있는 alphabet 인식 실험은 MPEG-7의 응용분야의 한 부분을 보여주고 있고 문자 인식뿐만 아니라 이런 인식 방법을 통해서 이미지 검색을 실제로 또 다른 응용분야에 적용시킬 수 있을 것으로 사료된다.

참고문헌(또는 Reference)

- [1] Yong Man Ro, Ho Kyung Kang , "Texture feature extraction based on HVS for MPEG-7 homogeneous texture descriptor", SPIE 2000, Vol 4210, p72-81
- [2] MPEG-7 Visual part of eXperimentation Model Version 8.0, ISO/MPEG, w3673, La Baule, Oct. 2000.
- [3] Visual Working Draft 4.0, ISO/MPEG, w3522, Beijing, July. 2000.