

비디오에서 프로젝션을 이용한 문자 인식

백정욱* · 신성윤* · 이양원*

*군산대학교

Identification of Korea Traditional Color Harmony

Jeong-Uk Baek* · Seong-Yoon Shin* · Yang-Won Rhee*

*Kunsan National University

E-mail : s3397220@kunsan.ac.kr

요 약

비디오에서 우선 장면 전환 검출을 통해 생성된 키 프레임에 대해 프로젝션을 통하여 문자 인식을 수행하도록 한다. 텍스트의 자간 분리를 수직 프로젝션에 의해 분리한다. 자소는 초성, 중성, 종성으로 분리하고 6가지 유형으로 분리한다. 자소 패턴 분리는 수평 프로젝션을 통하여 6가지 유형에 맞도록 분리한다. 자소는 수평, 수직, 사선, 역사선 방향으로 분리한다. 자소의 인식은 4-방향 프로젝션과 위치정보를 이용하여 인식하도록 한다.

ABSTRACT

In Video, key frame generated from the scene change detection is to perform character recognition through the projections. The separation between the text are separated by a vertical projection. Phoneme is separated Cho-sung, Jung-sung, and Jong-sung and is divided 6 types. Phoneme pattern is separated to suitable 6 types through the horizontal projection. Phoneme are separated horizontal, vertical, diagonal, reverse-diagonal direction. Phoneme is recognized using the 4-direction projection and location information.

키워드

프로젝션, 자소 분리, 위치 정보

I. 서 론

비디오에서 한글 문자 인식은 자간 분리와 자소 분리, 그리고 자소와 문자의 인식 과정을 통하여 하나의 문자로 인식하게 된다.

본 논문에서는 수직 프로젝션에 의한 자간 분리와 초성, 중성, 종성에 따른 6가지 유형의 자소의 유형 분리를 수행한다. 그리고 수평 프로젝션에 의한 자소 패턴 분리를 수행하고 4방향 프로젝션에 의해 완전한 자소 분리를 수행한다.

II. 관련연구

한글 자막에서 동일한 유형의 문자는 초성, 중성, 종성이 모두 존재하는 영역이 거의 일정하지만 모음의 형태에 따라 자음과 모음이 존재하는

영역이 약간씩 차이가 있다. 빌렛영상에 포함된 문자인식을 위한 비전시스템 개발[1]과 혼용 문자 코드 집합을 위한 계층적 다중 문자 인식기[2] 등이 현대의 대표적인 문자 인식 시스템이다.

III. 문자 인식

문자 인식을 위해 우선 자간 분리를 수행한다. 자간 분리는 수직 프로젝션에 의한 히스토그램 차이로 분리한다.

다음으로 자소의 유형을 6형식으로 분리한다.

다음으로 자소의 패턴을 분리한다. 수평 프로젝션에 의해 다음 그림 1과 같이 분리한다.

위치조건(좌측)	두 선분 사이의 관계
$F < A$	선분 DE 는 새로운 패턴의 시작 선분
$D > B$	선분 AB 는 기존 패턴의 끝 선분
$F < B$	선분 AB 는 선분 DE 와 선분 DF 의 다음 선분으로 분리
$E > C$	C를 시작점으로 하는 선분은 선분 AB 에 병합
기타의 경우	선분 AB 와 선분 DE 는 같은 영역 선분

그림 1. 자소 패턴 분리

자소의 분리는 4-방향 벡터에 대한 프로젝션에 의하여 수행된다.

다음으로 자음과 모음의 인식과 문자 인식 단계로 나뉜다. 4-방향 프로젝션과 위치 정보를 이용하여 수행된다. 자음/모음의 인식 특징은 그림 2, 그림 3과 같다.

자음	프로젝션 특징	위치정보
ㄱ	$h1, d1$	$h1(x2,y2)=d1(x1,y1), h1(y1,y2)<d1(y2)$
ㄴ	$h1, v1$	$h1(x1,y1)=v1(x2,y2), v1(x1,x2)<h1(x2), v1(y1)<h1(x1,x2)$
ㄷ	$h1, h2, v1$	$h1(x1,y1)=v1(x1,y1), h2(x1,y1)=v1(x2,y2), (h1(x2),h2(x2))>v1(x1,x2), h1(y1,y2)<h2(y1,y2)$
ㄹ	$h1, h2, h3, v1, v2$	$h1(x2,y2)=v1(x1,y1), v1(x2,y2)=h2(x2,y2), v1(y1)<v2(y1)$ $h2(x1,y1)=v2(x1,y1), v2(x2,y2)=h3(x1,y1), v1(y2)<v2(y2)$ $h1(y1,y2)<h2(y1,y2)<h3(y1,y2), v1(x1,x2)>v2(x1,x2)$
ㅁ	$h1, h2, v1, v2$	$h1(x1,y1)=v1(x1,y1), h1(x2,y2)=v2(x1,y1), h2(x1,y1)=v1(x2,y2), h2(h2,y2)=v2(x2,y2), h1(y1,y2)<h2(y1,y2), v1(x1,x2)<v2(x1,x2)$
ㅂ	$h1, h2, v1, v2$	$(h1(x1),h2(x1))=v1(x1,x2), (h1(x2),h2(x2))=v2(x1,x2)$ $h2(x1,y1)=v1(x2,y2), h2(h2,y2)=v2(x2,y2), v1(x1,x2)<v2(x1,x2), h1(y1,y2)<h2(y2,y2)$
ㅅ	$d1, i1$	$d1(x2)<i1(x1)<i1(x2), d1(y1)<i1(h1)<(d1(y2),i1(y2))$
ㅇ	$h1, h2, v1, v2, d1, d2, i1, i2$	$h1(x1,y1)=d1(x1,y1), d1(x2,y2)=v1(x1,y1), v1(x2,y2)=i1(x1,y1), i1(x2,y2)=h2(x1,y1), h2(x2,y2)=d2(x2,y2), d2(x1,y1)=v2(x2,y2), v2(x1,y1)=i2(x2,y2), i2(x1,y1)=h1(x2,y2)$
ㅈ	$h1, d1, i1$	$h1(x2,y2)=d1(x1,y1), d1(x2)<i1(x1)<i1(x2), d1(y1)<i1(h1)<(d1(y2),i1(y2))$
ㅊ	$h1, h2, d1, i1$	$h1(y1,y2)<h2(y1,y2), h2(x2,y2)=d1(x1,y1), d1(x2)<i1(x1)<i1(x2), d1(y1)<i1(h1)<(d1(y2),i1(y2))$
ㅋ	$h1, h2, d1$	$h1(x2,y2)=d1(x1,y1), h1(y1,y2)<h2(y1,y2)<d1(y2)$ $h1(x1,y1)=v1(x1,y1), h3(x1,y1)=v1(x2,y2)$
ㅌ	$h1, h2, h3, v1$	$h1(y1,y2)<h2(y1,y2)<h3(y1,y2), h1(x1)=h2(x1)=h3(x1)=v1(x1,x2)$
ㅍ	$h1, h2, v1, v2$	$h1(y1,y2)<h2(y1,y2), v1(x1,x2)<v2(x1,x2), (h1(x1),h2(x1))<v1(x1,x2)<v2(x1,x2)<(h1(x2),h2(x2))$ $h1(y1,y2)<h2(y1,y2)<h3(y1,y2)$
ㅎ	$h1, h2, h3, h4, v1, v2, d1, d2, i1, i2$	$h3(x1,y1)=d1(x1,y1), d1(x2,y2)=v1(x1,y1), v1(x2,y2)=i1(x1,y1), i1(x2,y2)=h4(x1,y1), h4(x2,y2)=d2(x2,y2), d2(x1,y1)=v2(x2,y2), v2(x1,y1)=i2(x2,y2), i2(x1,y1)=h3(x2,y2)$

그림 2. 자음의 인식

모음	프로젝션 특징	위치정보
ㅏ	$h1,v1$	$v1(y1)<h1(y1)<v1(y2), v1(x1,x2)=h1(x1), h1(x2)>v1(x1,x2)$
ㅑ	$h1,h2,v1$	$v1(y1)<h1(y1)<h2(y1)<v1(y2), v1(x1,x2)=h1(x1)=h2(x1), (h1(x2),h2(x2))>v1(x1,x2)$
ㅓ	$h1,v1$	$v1(y1)<h1(y1)<v1(y2), v1(x1,x2)=h2(x1), h1(x1)<v1(x1,x2)$
ㅕ	$h1,h2,v1$	$v1(y1)<h1(y1)<h2(y1)<v1(y2), v1(x1,x2)=h1(x2)=h2(x2), (h1(x1),h2(x1))<v1(x1,x2)$
ㅗ	$h1,v1$	$h1(y1,y2)=v1(y2), v1(y1)<h1(y1,y2), h1(x1)<v1(x1,x2)<h1(x2)$
ㅛ	$h1,v1,v2$	$h1(y1,y2)=v1(y2)=v2(y2), (v1(y1),v2(y1))<h1(y1,y2), h1(x1)<v1(x1,x2)<v2(x1,x2)<h1(x2)$
ㅜ	$h1,v1$	$h1(y1,y2)=v1(y1), v1(y2)>h1(y1,y2), h1(x1)<v1(x1,x2)<h1(x2)$
ㅠ	$h1,v1,v2$	$h1(y1,y2)=v1(y1)=v2(y1), (v1(y2),v2(y2))<h1(y1,y2), h1(x1)<v1(x1,x2)<v2(x1,x2)<h1(x2)$
ㅡ	$h1$	$h1(x1)<h1(x2), h1(y1)=h1(y2)$
ㅣ	$v1$	$v1(x1)=v1(x2), v1(y1)<v1(y2)$
ㅞ	$h1,v1,v2$	$v1(x1,x2)=h1(x1), h1(x2)=v2(x1,x2), (v1(y1),v2(y1))<h1(y1,y2)<(v1(y2),v2(y2))$
ㅟ	$h1,h2,v1,v2$	$v1(x1,x2)=h1(x1)=h2(x1), h1(x2)=h2(x2)=v2(x1,x2), (v1(y1),v2(y1))<h1(y1,y2)<h2(y1,y2)<(v1(y2),v2(y2))$
ㅠ	$h1,v1,v2$	$h1(x1)<v1(x1,x2)<v2(x1,x2), h1(x2)=v1(x1,x2), (v1(y1),v2(y1))<h1(y1,y2)<v1(y2),v2(y2)$
ㅡ	$h1,h2,v1,v2$	$h1(x2)=h2(x2)=v1(x1,x2), (h1(x1),h2(x1))<v1(x1,x2)<v2(x1,x2), (v1(y1),v2(y1))<h1(y1,y2)<(v1(y2),v2(y2))$

그림 3. 모음의 인식

문자의 인식은 다음 그림 4와 같이 자소를 조합하여 얻어진다.

초성		중성		종성	
자소	코드	자소	코드	자소	코드
ㄱ	0	ㅏ	0	없음	0
ㄴ	1	ㅑ	1	ㄱ	1
ㄷ	2	ㅓ	2	ㄴ	2
ㄹ	3	ㅕ	3	ㄷ	3
ㅁ	4	ㅗ	4	ㄹ	4
ㅂ	5	ㅛ	5	ㅁ	5
ㅅ	6	ㅜ	6	ㅂ	6
ㅇ	7	ㅠ	7	ㅅ	7
ㅈ	8	ㅡ	8	ㅇ	8
ㅊ	9	ㅣ	9	ㅈ	9
ㅋ	10	ㅞ	10	ㅊ	10
ㅌ	11	ㅟ	11	ㅋ	11
ㅍ	12	ㅠ	12	ㅌ	12
ㅎ	13	ㅡ	13	ㅍ	13
				ㅎ	14

그림 4. 자소 조합에 의한 문자 인식

V. 실험

문자 인식 실험 결과 대부분의 텍스트는 인식이 가능하지만 오인식의 경우는 자막 텍스트에 연결성이 많이 존재하거나 배경 화면이 불분명한 경우로 나왔다. 전체적으로 인식률은 95% 이상인 것으로 나왔다.

V. 결론

비디오에서 장면 전환 검출을 통해 생성된 키 프레임에 대상으로 프로젝션을 통하여 문자 인식을 수행하도록 하였다. 텍스트의 자간 분리는 수직 프로젝션에 의해 수행하였다. 자소는 초성, 중성, 종성으로 분리하고 6가지 유형으로 분리하였다. 자소 패턴 분리는 수평 프로젝션을 통하여 6가지 유형에 맞도록 분리하였다. 자소는 수평, 수직, 사선, 역사선 방향으로 분리하였다. 자소의 인식은 4-방향 프로젝션과 위치정보를 이용하여 인식하도록 하였다. 본 논문에서 제시한 방법으로 95% 이상의 인식률을 보여주었다.

참고문헌

- [1] 박상국, “빌렛영상에 포함된 문자인식을 위한 비전시스템 개발,” 한국산업정보학회논문지, 제13권, 제1호, pp. 22-29, 2008.
- [2] 김도현, 박재현, 김철기, 차의영, “혼용 문자 코드 집합을 위한 계층적 다중 문자 인식기,” 한국해양정보통신학회논문지, 제11권, 제10호, pp. 1977-1985, 2007년 10월.