

# 영역확장을 이용한 뉴스 비디오 자막 자동 추출 Automatic News Video Caption Segmentation Using Region Growing

이 상 호\*, 손 광 훈\*, 박 질 남\*\*

\*연세대학교 전파공학과

\*\*평화방송 TV기술부

Sang Ho Lee, Chull Nam Park, and Kwang Hoon Sohn

\*Dept. of Radio Communication & Engineering, Yonsei University

\*\*Pyonghwa Broadcasting Corp.

e-mail : cocook@nownuri.net

## 요 약

본 논문에서는 뉴스 비디오 자막의 구조적 문자 특징을 이용한 자동 자막추출방법을 기술하였다. 기존의 방법들은 경계 추출이나 이진화 과정 후에 화소값의 변화 profile에 나타나는 문자의 굵기, 간격 등의 문자 특징을 이용하는데 반해, 본 논문에서 제안한 방식은 뉴스 비디오 자막에서 문자와 배경을 구분하는 테두리선을 배경과 문자를 구분하는 경계로 하여, 적절한 이진화 과정과 영역 확장 기법을 이용하여 문자를 추출하였다. 제안한 방법은 문자에 대한 사전지식이 없어도 되며, 사용한 영역확장 기법은 기존의 방법에서의 경계 추출기법보다 간단하면서도 효과적인 문자추출 결과를 나타낸다.

본 논문에서는 경계검출이나 문자의 굵기, 간격등의 특성대신 뉴스 자막의 테두리선 특징을 이용하며, 문자영역의 추출은 1차미분을, 문자와 배경의 구분은 영역확장을 이용한다. 따라서, 문자에 대한 사전 정보를 알 필요가 없으며, 문자 내부와 외부의 테두리선만 있다면 명도차가 크지 않아도 된다.

본 논문에서는 II장에서 뉴스 자막의 특징, III장에서는 문자 추출에 사용된 기본적 기법인 1차 미분과 영역확장에 대하여 설명하며, IV장에서는 제안한 알고리즘과 실험결과를, 그리고 V장에서 결과와 함께 추후 연구과제에 대해 기술한다.

## II. 뉴스 자막의 특징

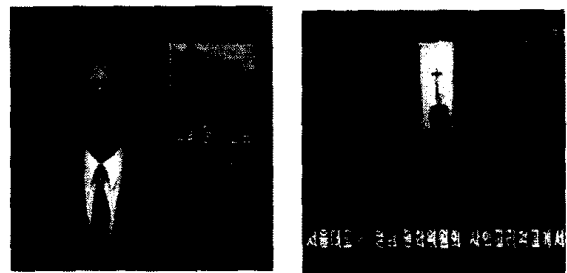
### I. 서 론

매일의 방대한 뉴스 자료를 수작업으로 색인화하는 것은 많은 시간과 노력이 필요하다. 이 경우 뉴스 내용을 가장 함축적으로 표현하고 있는 자막을 자동적으로 추출하여 색인화한다면 효과적인 뉴스 내용 데이터 베이스 시스템이 될 것이다. 이런 데이터 베이스는 빠른 뉴스 내용 검색을 필요로 하는 NOD(News on Demand) 서비스를 위해서는 필수적이며, 이를 위해 효과적인 뉴스 내용의 색인화를 위한 연구가 필요하다.

뉴스 비디오의 자막 추출에 관한 외국의 연구를 살펴보면, 고정된 문자 후보영역에서 명도값의 변화가 큰 곳 사이의 간격이, 미리 알고 있는 문자의 간격과 일치하면 하나의 문자로 간주하여 추출한다<sup>[1]</sup>.

국내의 한글문자의 추출 방법에 관한 기존 연구에서는 문자의 규칙적 반복 특성에서 나타나는 주파수 특성을 이용하거나<sup>[2]</sup>, 경계 검출을 통해 나타난 문자 획의 굵기나 문자 간격 등의 특성을 이용하여 문자 영역을 추출하는 방식도 제안되어있다<sup>[3]</sup>. 또한 가로 방향의 인접 화소간의 명도차이가 큰 곳의 수를 조사하여 문자 후보 영역을 추출한다<sup>[4]</sup>. 위의 기존의 방법들에서는 문자의 정보를 미리 알아야 하거나 또는 문자 내부와 외부의 명도값 차이가 커야만 한다.

뉴스 비디오에서 나타나는 자막은 그림1과 같이 크게 뉴스 아이콘 자막과 뉴스 내용 자막으로 나눌 수 있다.



(a) 뉴스 아이콘 자막

(b) 뉴스 내용 자막

그림 1: 뉴스 자막

이러한 뉴스 자막은 다음과 같은 특징을 가진다.

#### (1) 자막의 위치 특징

뉴스 프로그램에서 아이콘 자막이 존재하는 곳은 앵커 장면이 나타나는 영상에서 앵커의 오른쪽 일정 영역이며, 이 자막의 위치는 같은 프로그램에서는 항상 동일하다. 마찬가지로 뉴스내용자막 역시 일정한 영역에 표시된다. 따라서 이 고정된 위치 특성을 이용하면

문자의 후보 영역을 선정하기 위해서 한 프레임의 전체 영상을 검색할 필요가 없이 특정한 일부부분만을 검색하면 된다.

### 2. 자막의 문자 특징

일반적인 자막의 문자는 길이와 배경색상이 불규칙하지만 몇 가지 주목할 만한 공통적 특징을 가지고 있다.

첫 번째 특징은 문자내부는 대부분 단일색상으로 표현된다는 것으로, 단일한 색상 값을 갖는 영역을 조사하여 문자를 판단하는데 이용될 수 있다.

두 번째 특징은 문자의 가로 세로 크기가 비교적 균일하다는 것이다. 따라서 최종적으로 추출되는 문자영역은 직사각형의 영역이 된다.

세 번째 특징은 문자의 테두리가 문자내부에 비해서 상당히 낮은 명도값을 갖는다는 것이다. 이것은 문자를 두드러지게 하기 위한 목적으로 사용된 것으로 제안된 방법에서는 이 특징을 가장 크게 이용하였다. 일반적으로 테두리선 명도값은 0~50의 값을 가진다.

## III. 문자 추출에 이용된 두 가지 기본 기법

뉴스비디오 자막의 추출을 위해서 본 논문에서는 크게 다음과 같은 1차미분과 영역확장 기법을 사용한다.

### 1. 1차 미분

앞에서 말한 문자 특성에서 보면 문자는 테두리와 내부의 명도차가 크다. 따라서 문자에서의 주파수 특성은 고주파 부분이 되고 배경에서는 저주파가 된다. 급격한 명도차를 나타내는 부분을 미분하여 보면 그림 2와 같은 특성이 나타남을 알 수 있게 된다.



그림 2 : 경계 부분의 미분 특성

기존의 방법에서는 미분을 통해 얻은 경계선의 간격과 미리 알고 있는 문자획의 굵기를 비교하여 문자영역의 위치를 알아낸다. 그러나, 본 논문에서는 식(1)과 같은 3-point Lagrange interpolation에 해당하는 1차 미분 영상의 profile 특성을 이용하여 문자 영역을 추출한다.

$$f^*(x) = \sum_{j=0}^3 f_j \phi_j(x) \quad (1)$$

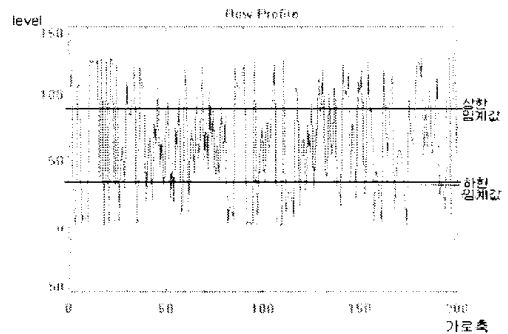
$$\Phi_j(x) = \frac{\prod_{i=0, i \neq j}^3 (x - x_i)}{\prod_{i=0, i \neq j}^3 (x_j - x_i)}$$

여기서,  $\Phi_j(x)$ 는 이차 미분값,  $x$ 는 화소의 좌표,  $f_j$ 는 명도값을 나타낸다.

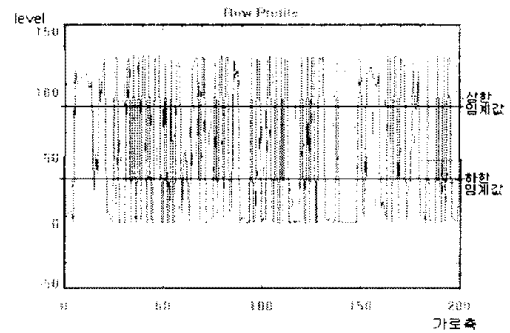
그림 3은 문자 후보 영역의 1차 미분 영상을 보여주고 있으며 그림 4는 문자 부분과 배경부분의 미분영상 profile을 나타낸다.



그림 3 : 1차 미분 영상



(a) 문자영역 내의 Profile



(b) 문자영역 밖의 Profile

그림 4 : 미분영상의 profile

위 그림 4에서 보듯이 명도차가 큰 곳에서는 불규칙한 미분값을 갖게 되고, 명도차가 작은 곳에서는 균일한 분포의 값을 갖게 된다. 즉, 문자부분과 배경부분은 그 명도차의 특성 차이로 인해 뚜렷하게 서로 다른 미분특성을 가진다. 문자영역 밖의 profile은 최대, 최소값이 규칙적으로 반복되는데 비해 문자영역 내의 profile은 불규칙적으로 반복됨을 알 수 있고 이를 이용하여 문자 영역을 추출할 수 있게 된다.

## 2. 영역확장(Region Growing)

제안한 방법에서는 명암도의 유사성에 근거한 영역확장 방식을 이용하여 영상의 자막을 분할한다. 영역확장은 seed point( $x_i$ )에서 시작하여 유사한 특성을 가진 이웃한 화소들( $x_j$ )이 한 영역으로 판단되도록 다음의 기준대로 계속 영역을 확장시키게 된다.

$$|x_i - x_j| \leq \text{Threshold} \rightarrow \text{확장}$$

$$|x_i - x_j| > \text{Threshold} \rightarrow \text{정지}$$

뉴스 자막의 경우 문자 내부의 명도가 균일하고 테두리와의 명도차가 크기 때문에 배경과 문자를 분리하는데 이 방식을 효과적으로 사용할 수 있다.

## IV. 뉴스자막 자동 분할 알고리즘

제안한 알고리즘의 개요는 그림 5와 같다. 즉, 흑백으로 전환된 입력영상은 자막의 종류에 따라 원 영상의 일정한 위치에 해당하는 작은 영상이다. 이 입력영상은 1차미분되고 문자영역이 추출된다. 추출된 문자영역영상은 배경과 문자내부가 문자테두리선에 의해 구분되므로 배경부분을 영역확장을 통해 하나의 영역으로 묶고 이 영역과 영역확장되지 않은 영역(문자내부영역)을 이진화를 통해 분리시키고 후처리를 한다.

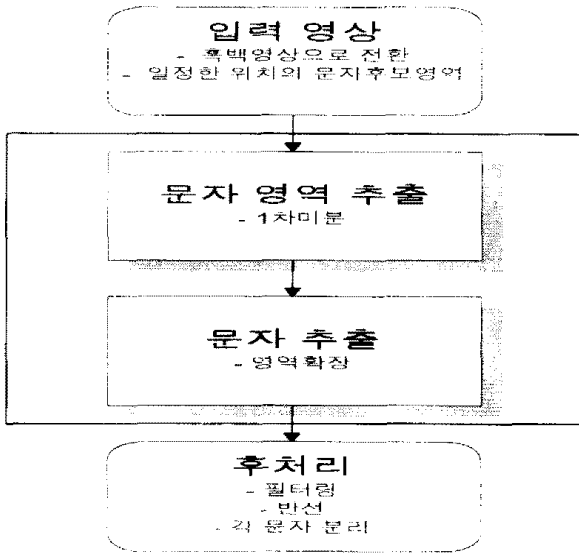


그림 5 : 뉴스 자막 추출 개요

### 1. 입력 영상

본 논문에서는 뉴스 자막 추출이 목적이므로 먼저 컬러 원 영상을 처리가 쉽도록 흑백영상으로 바꾼다. 그리고, 그림 6과 같이 뉴스 아이콘 자막과 뉴스 내용에 따라 일정한 위치의 영역만을 문자 후보 영역으로 한정시킨다.



(a) 뉴스 아이콘 자막



(b) 뉴스 내용 자막

그림 6 : 문자 후보 영역

### 2. 후보 영역 추출

문자 후보 영역을 1차 미분하고 다음의 범위에 속하는 미분값에 해당하는 가로부분만을 선택해서 그림 7과 같은 가로 방향의 문자 영역을 추출한다.

- 하한 임계값  $\leq$  "추출된 문자영역"  $\leq$  상한 임계값
- 상한 임계값 = (미분 영상의) 최대값 + 표준편차
- 하한 임계값 = (미분 영상의) 최소값 - 표준편차



그림 7 : 가로 방향으로 추출된 문자 영역

### 3. 문자 추출

먼저 그림 8과 같이 추출된 문자영역의 상하좌우에 각각 '문자 테두리선 굵기 +  $\alpha$  (실험값:5)'에 해당하는 만큼을 확장시킨다. 이것은 한글 문자의 일부 획의 손실을 방지하며, 문자가 검정 테두리선으로 완전히 싸이게 하는 효과를 준다. 그리고, 문자의 검정 테두리선을 더욱 뚜렷이 하기 위해 임계값(보통 50~100)이하는 모두 0으로 한다.



그림 8 : 확장된 문자 영역

확장된 문자영역 영상에 다음의 조건에 따라 문자 외부로부터 영역확장을 한다. 이 때, 문자 테두리선이 문자 내

부와 외부의 경계가 되므로 문자 내부로부터의 영역확장은 외부로부터의 영역확장과 같은 결과를 나타낸다. 따라서 이것은 문자의 내부를 찾기 위해 기존의 방법들처럼 문자의 정보를 이용할 필요가 없는 장점을 가진다.

· seed point

그림 8 영상의 가장 위 line과 아래 line의 각 화소 중에서 명도값이 50 이상인 화소.

· 상한 임계값 : seed point 화소의 명도값 + 50

· 하한 임계값 : seed point 화소의 명도값 - 50

이상과 같은 조건으로 영역확장된 영역은 0, 그렇지 않은 영역은 1로 이진화를 한다. 이때, 뉴스 아이콘 자막에서 '통'의 'o'과 같이 외부에서 영역확장 되지 않는 부분은 적절한 임계값(실험값:190)으로 이진화를 한다. 그림 9는 영역확장후 이진화를 한 영상을 보여준다.



그림 9 : 외부로부터 영역확장된 영상

#### 4. 후처리

영역확장 후 이진화된 영상은 뉴스 자막의 문자를 뚜렷이 보여준다. 그러나, 그림 9의 뉴스 아이콘 자막에서 '교'의 왼쪽 흰 부분과 같이 배경부분중 문자 내부와 같은 명도값을 가지는 부분은 배경으로 인식되지 못하고 남게 된다. 이것은 앞의 영역확장과 같은 방법으로 가장 윗 line에서 흰색 부분이 나타나는 곳을 seed point로 하여 흰색 부분이 한 영역으로 되도록 영역확장을 한 후 그 부분을 명도값 '0'으로 하여 제거한다.

또, 영역확장시에 테두리선의 불균일한 명도값으로 인해 문자 내부가 일부 침범될 우려가 있으며 이로 인해 획이 매끄럽지 못한 현상이 나타난다. 이를 해결하기 위해 후처리 과정으로 메디안 필터링을 하는데 문자의 세로획의 굵기가 가늘기 때문에 세로방향으로만 필터링을 한다. 그림 10은 후처리한 영상을 보여준다



그림 10 : 후처리된 영상

영상의 문자내부와 외부를 만진시킨 후, 문자와 문자사이에서 명도값 255만을 갖는 선의 반복을 조사하고, 문자의 평면적인 가로 크기를 비교하여 각 문자를 구분한다. 그림 11은 최종적으로 추출된 뉴스 자막의 문자를 보여준다.

## 교회 통계 연감

### 서울대학교 '정의 평화위원회' 사회교리학교에서

그림 11 : 추출된 뉴스 자막 문자

#### V. 결론

본 논문에서는 내용기반 뉴스 데이터 베이스를 위한 뉴스 자막 추출 방법에 대해 기술하였다.

뉴스 자막의 특징인 테두리선을 문자 외부와의 경계로 하여 문자 외부를 시작점으로 하는 영역확장을 하여 문자 테두리선에서 확장이 멈추도록 하였다. 그리고, 문자 외부와 내부를 이진화를 통해 구분하였다.

실험결과 문자에 대한 정보를 이용하지 않고도, 뉴스 아이콘 자막과 뉴스 내용 자막에 상관없이 1차미분에 의한 문자영역추출과 영역확장에 의한 문자 추출 결과 모두 우수함을 알 수 있다. 그러나, 일부 문자 테두리선이 명확하지 않은 문자는 추출이 어려운 단점이 있어서 이 부분에 대한 연구가 진행되어야 한다고 생각한다.

#### 참고문헌

- [1] S. Kannangara, E. Asbun, R. X. Browning, and E. J. Delp, "The Use of Nonlinear Filtering in Automatic Video Title Capture," Proceedings of the 1997 IEEE/EURASIP Workshop on Nonlinear Signal and Image Processing, pp. 8-10, September, 1997.
- [2] 진 병태, 정 세윤, 이 재연, 배 영래, "뉴스 비디오 자막 분할 방법," 대한 전자공학회 하계종합학술대회 논문집, 20권, 1호, pp. 467-470, 1997년 6월.
- [3] 진 병태, 윤 호섭, "신호처리 기법을 응용한 차량번호판 추출방법," 전자공학회논문지, 30권, 7호, pp. 92-100, 1993년 7월.
- [4] 박 영규, 김 성국, 유 원영, 김 준철, 이 준환, "MPEG II 뉴스영상에서의 문자영역 추출 및 문자인식," 제11회 신호처리합동학술대회논문집, 11권, 1호, pp. 117-120, 1998년 10월.