

신경망을 이용한 자막 크기에 무관한 연결 객체 기반의 자막 추출

Connected Component-based Regardless of Caption Size Caption Extraction with Neural Network

정제희, 윤태복, 김동문, 이지형

성균관 대학교 정보통신 공학부 전자전기컴퓨터 전공

E-mail: gulingi@skku.edu, tbyoon@skku.edu, skyscraper@naver.com, jhlee@ece.skku.ac.kr

요 약

영상에 나타나는 자막은 영상과 관계가 있는 정보를 포함한다. 이러한 자막의 정보를 이용하기 위해 영상으로부터 자막을 추출하는 연구는 근래에 들어 활발히 진행되고 있다. 기존의 연구는 일정한 높이의 자막이나 획의 두께를 가진 자막만을 추출하였다. 본 논문에서는 일정 크기 이상의 크기에 무관한 자막을 추출하는 방법을 제안한다. 먼저, 자막 추출을 위해서 영상에 포함된 픽셀들의 연결 객체를 생성하였다. 그리고 연결 객체 중에서 자막의 형태적인 특징의 패턴을 분석하고, 패턴을 이용하여 자막을 추출하였다. 실험에 사용된 영상은 다큐멘터리, 쇼 프로그램과 같은 대중 방송에서 획득하였으며, 실험 결과는 다양한 크기의 자막을 포함한 영상을 사용하여 실험하였고, 자막 추출의 결과는 찾아진 연결객체 중에 자막의 비율과 자막 중에 찾아진 자막의 비율로 분석하였다. 제안한 방법에 의해 다양한 크기의 자막을 추출할 수 있었다.

Key Words : 자막 추출, 신경망, 자막, 연결 객체

1. 서 론

지식을 전달하는 역할을 하는 콘텐츠에는 문서, 영상, 음악, 동영상 등의 다양한 종류가 존재한다. 특히, 영상 콘텐츠는 동영상을 이루는 요소이면서 자체적으로도 시각적으로 의미를 전달하기에 효율적인 콘텐츠라 할 수 있다. 하지만 영상 콘텐츠는 다른 콘텐츠들에 비해 대부분 용량이 거대한 편이기에 영상 콘텐츠를 효율적으로 저장하고 검색하는 연구는 활발히 진행되고 있다.

영상에 포함된 자막은 영상의 내용과 밀접한 관계를 가지고 있기 때문에 자막의 정보를 이용한 영상의 저장 및 검색을 위한 연구는 예전부터 진행되어 왔다[1],[2].

자막 추출에는 자막의 다양한 크기, 밝기, 색상, 위치, 글꼴 등의 여러 가지 어려움이 존재한다. 특히, 자막의 크기가 일정 수준을 벗어나면 제안한 방법이 정상적인 작동을 하지 않는 경우가 존재하였다. 이러한 단점을 극복하기 위해 다중 해상도 조절 방법을 사용하여 다양한 크기의 자막을 추출하고자 하였으나, 조절된 해상도에 따라 같은 연산을 반복적으로 수

행하여야만 했다[1].

본 논문에서는 자막을 이루는 경계선을 기반으로 연결 객체를 생성하여, 신경망을 통해 자막과 비 자막의 패턴을 분석한다. 분석된 패턴을 이용하여 연결 객체를 분류하는 자막의 크기에 무관하게 자막을 추출하는 방법을 제안하고자 한다.

본 논문은 2장에서 기존의 논문들에 대한 분석을 할 것이다. 3장에서 제안하는 방법을 서술하고, 4장에서 실험 결과를 정리할 것이며, 마지막으로 5장에서 결론을 내린다.

2. 관련 연구

자막을 추출의 기존의 연구[1],[2]들은 크게 3가지 단계로 자막을 추출하였다. 각 단계는 다음과 같다. 영상의 자막 존재 여부를 판단하는 자막 탐색 단계와 영상의 자막이 존재하는 지역을 결정하는 자막 위치 찾기 단계, 자막이 있다고 판단되는 지역에서 자막을 찾아내는 자막 분리 단계로 이루어져 있다.

영상에 포함된 자막은 위치, 크기, 색상, 글꼴 등의 다양한 특성을 갖기 때문에 자막 추출

에는 어려움이 따른다. 이러한 자막의 특성을 모두 고려하여 자막을 추출하는 것은 상당히 어렵다고 볼 수 있기 때문에 대부분의 자막 추출 방법은 자막에 대해 일정한 가정을 전제로 자막 추출 방법을 제안하였다.

다양한 크기의 자막을 검색하기 위해서 다중 해상도 기법[1]을 사용하였고, 이러한 방법은 해상도가 조절된 영상에서 같은 단계를 반복해야만 했다. [2]에서는 획 기반 자막 추출 방법을 제안하였으나, 획의 간격으로 정한 범위를 벗어난 자막의 획은 자막으로 인식하지 못하였다.

3. 자막 크기에 무관한 자막 추출 방법

3.1 연결 객체

자막을 이루는 픽셀들은 하나의 픽셀만으로 자막을 표현할 수 없으며, 자막은 연결 객체로 구성되어 있다고 할 수 있다. 이러한 조건을 전제로 자막을 이루는 모든 픽셀들을 하나의 연결 객체로 구성한 후, 연결 객체의 패턴을 분석하여 자막 연결 객체와 비 자막 연결 객체로 분류하여 자막의 여부를 판단하였다.

3.2 자막의 특성

자막은 다양한 특성을 갖기 때문에 어떠한 상황에서도 만족하는 자막의 특성을 결정하기는 매우 어려운 일이다. 그래서 본 논문에서는 추출 대상 자막을 정하기 위해 다음과 같이 가정한다. 첫째, 자막은 일정한 색상으로 이루어지며, 배경보다는 밝은 색상을 갖는다. 둘째, 자막을 이루는 획의 에지는 반드시 반대방향으로 대응하는 에지가 존재한다.

3.3 신경망 학습

신경망 학습은 연결 객체의 학습을 통해 이루어졌다. 연결 객체 데이터는 총 8,800개이며, 자막은 5,900개, 비 자막은 1,900개로 구성되어 있다. 학습 데이터는 자막 5,000개와 비 자막 2,500개로 구성되어 총 7,500개이며, 테스트 데이터는 자막 900개와 비 자막 400개로 총 1,300개로 구성하였다. 입력 특징은 총 10로 구성하였고, 자세한 서술은 3.4.1절에서 언급하겠

다.

3.4 제안하는 방법

제안하는 방법은 색상 영상을 입력으로 하여, 영상에서 자막과 그 이외의 부분으로 이진화된 영상을 출력하는 것을 목적으로 한다.

3.4.1 자막 찾기 단계

입력으로 들어온 컬러 영상은 자막 추출 과정의 복잡성을 낮추기 위해 256단계의 밝기를 가진 명암도 영상 G로 변환한다.

자막은 배경과 대비를 이루기 때문에 자막과 배경의 경계지점은 반드시 에지가 존재한다. 에지란 방향성을 갖는 픽셀로서 밝기의 변화의 정도를 표현하여 준다. 이러한 에지를 연결한 연결 객체는 자막의 획의 경계라고 할 수 있다.

명암도 영상 G로부터 에지를 추출은 로버츠 에지 추출 방법[3]을 사용한 (1)의 식을 이용하여 각 방향(좌상, 상, 우상, 좌, 우, 좌하, 하, 우하)의 에지를 갖는 방향 에지 영상 E_{tl} , E_t , E_{tr} , E_l , E_r , E_b , E_{bl} , E_{br} 을 만들어 낸다. 이때 각 영상들은 임계치를 적용하여 일정 수준이하의 에지는 제거하여, 충분한 대비를 갖지 않는 에지를 제거하였다.

방향 에지 영상은 하나의 에지 영상은 한 방향의 에지들의 정보만이 표현하기 때문에, 모든 방향의 에지를 포함하는 영상이 필요하다. 그러기 위해 각 방향 에지 영상에서 하나라도 에지가 존재한다면 에지로 판단하여 모든 방향 에지 영상 E_{all} 을 생성한다.

$$\begin{aligned}
 E_{tl}(x,y) &= G(x,y) - G(x-1,y+1) \\
 E_t(x,y) &= G(x,y) - G(x,y+1) \\
 E_{tr}(x,y) &= G(x,y) - G(x+1,y+1) \\
 E_l(x,y) &= G(x,y) - G(x-1,y) \\
 E_r(x,y) &= G(x,y) - G(x+1,y) \\
 E_{bl}(x,y) &= G(x,y) - G(x-1,y-1) \\
 E_b(x,y) &= G(x,y) - G(x,y-1) \\
 E_{br}(x,y) &= G(x,y) - G(x+1,y-1)
 \end{aligned} \tag{1}$$

모든 방향 에지 영상의 에지들은 자막의 외부를 구성한다고 할 수 있다. 연결 객체의 패턴을 분석하기 위해 모든 방향 에지를 4-이웃 연결 객체 방법으로 조건1)을 만족하고, 각 라

$$\begin{aligned}
 1) \quad CC(x,y) &= \begin{cases} CC(x,y-1) = 0, & E_{all}(x,y-1) = 0 \\ CC(x,y-1) = CC(x,y), & E_{all}(x,y-1) \neq 0 \end{cases} \\
 CC(x,y) &= \begin{cases} CC(x-1,y) = 0, & E_{all}(x-1,y) = 0 \\ CC(x-1,y) = CC(x,y), & E_{all}(x-1,y) \neq 0 \end{cases}
 \end{aligned}$$

벨은 고유한 값을 갖는 1차 연결 객체 영상 CC_{first} 를 생성한다.

1차 연결 객체 영상은 자막이 아닌 배경의 에지까지도 라벨로 형성되는 경우가 존재한다. 이러한 문제점을 개선하기 위해 같은 연결 객체 내부에서 반대로 대응되는 에지가 존재하지 않는 에지는 제거한다. 또한, 자막의 색상은 배경보다 밝기 때문에 서로 대응되는 에지의 중간점에 존재하는 밝기가 에지보다 어두운 경우엔 자막을 이루는 에지가 아니라 판단하고 제거한다.

이렇게 자막의 에지가 아니라 판단된 에지들을 제거하면, 배경 연결 객체의 비율이 줄어들게 된다. 남겨진 에지들을 2차 라벨링을 수행하여 2차 연결 객체 영상 CC_{second} 를 생성한다.

연결 객체 패턴 분석을 위해서 2차 연결 객체 영상의 10가지 특징을 추출하고, 신경망 학습을 수행한다. 학습에 앞서 각 연결 객체의 각 에지는 하나의 픽셀에서도 다양한 방향이 존재할 수 있기 때문에, 한 픽셀에 다양한 방향의 에지가 존재하는 것을 제거하기 위해 그림 1에서 사용되는 마스크를 이용하여 각 방향 에지 영상에서 중첩되는 에지를 제거한다.

그림 1의 3 * 3의 마스크에서 중앙 점을 현재 검색하고자하는 픽셀에 위치하고, 주변의 에지 존재 여부를 판단하면, 에지의 방향성을 정제할 수 있다. 각 방향 에지 영상에 마스크를 사용한다. 예를 들어 E_{tr} 은 mE_{tr} 를 이용한다. 이때, X의 좌표와 대응되는 부분에 에지가 없고, O의 좌표와 대응되는 부분에 에지가 존재한다면 그 점은 그 방향을 갖는 성분이라고 할 수 있다.

방향 에지 정제 마스크를 통해 생성한 정제된 방향 에지 영상 (RE_{tl} , RE_t , RE_{tr} , RE_l , RE_r , RE_b , RE_{bl} , RE_b , RE_{br})에서 존재하는 에지들을 이용해 연결 객체의 패턴을 분석하기 위해 신경망에 쓰일 데이터를 추출한다. 10개의 입력 값은 다음과 같으며, 각 입력 값은 연결 객체마다 계산한다.

1. RE_{tl} 의 픽셀의 수 / E_{all} 의 픽셀의 수
2. RE_t 의 픽셀의 수 / E_{all} 의 픽셀의 수
3. RE_{tr} 의 픽셀의 수 / E_{all} 의 픽셀의 수
4. RE_l 의 픽셀의 수 / E_{all} 의 픽셀의 수
5. RE_r 의 픽셀의 수 / E_{all} 의 픽셀의 수
6. RE_{bl} 의 픽셀의 수 / E_{all} 의 픽셀의 수
7. RE_b 의 픽셀의 수 / E_{all} 의 픽셀의 수
8. RE_{br} 의 픽셀의 수 / E_{all} 의 픽셀의 수
9. 각 수평으로 대응되는 에지의 중앙 픽셀의 밝기들의 표준 편차
10. 각 수직으로 대응되는 에지의 중앙 픽셀의 밝기들의 표준 편차

mE_{tl}	<table border="1"><tr><td>X</td><td>X</td><td>X</td></tr><tr><td>X</td><td>O</td><td>O</td></tr><tr><td>O</td><td></td><td></td></tr></table>	X	X	X	X	O	O	O			<table border="1"><tr><td>X</td><td>X</td><td></td></tr><tr><td>X</td><td>O</td><td>O</td></tr><tr><td></td><td>O</td><td></td></tr></table>	X	X		X	O	O		O		<table border="1"><tr><td>X</td><td>X</td><td>O</td></tr><tr><td>X</td><td>O</td><td></td></tr><tr><td>X</td><td>O</td><td></td></tr></table>	X	X	O	X	O		X	O	
X	X	X																												
X	O	O																												
O																														
X	X																													
X	O	O																												
	O																													
X	X	O																												
X	O																													
X	O																													
mE_t	<table border="1"><tr><td>X</td><td>X</td><td>X</td></tr><tr><td>X</td><td>O</td><td>O</td></tr><tr><td>O</td><td></td><td></td></tr></table>	X	X	X	X	O	O	O			<table border="1"><tr><td>X</td><td>X</td><td>X</td></tr><tr><td>O</td><td>O</td><td>O</td></tr><tr><td></td><td></td><td></td></tr></table>	X	X	X	O	O	O				<table border="1"><tr><td>X</td><td>X</td><td>X</td></tr><tr><td>O</td><td>O</td><td>X</td></tr><tr><td></td><td></td><td>O</td></tr></table>	X	X	X	O	O	X			O
X	X	X																												
X	O	O																												
O																														
X	X	X																												
O	O	O																												
X	X	X																												
O	O	X																												
		O																												
mE_{tr}	<table border="1"><tr><td>X</td><td>X</td><td>X</td></tr><tr><td>O</td><td>O</td><td>X</td></tr><tr><td></td><td></td><td>O</td></tr></table>	X	X	X	O	O	X			O	<table border="1"><tr><td></td><td>X</td><td>X</td></tr><tr><td>O</td><td>O</td><td>X</td></tr><tr><td></td><td>O</td><td></td></tr></table>		X	X	O	O	X		O		<table border="1"><tr><td>O</td><td>X</td><td>X</td></tr><tr><td></td><td>O</td><td>X</td></tr><tr><td></td><td>O</td><td>X</td></tr></table>	O	X	X		O	X		O	X
X	X	X																												
O	O	X																												
		O																												
	X	X																												
O	O	X																												
	O																													
O	X	X																												
	O	X																												
	O	X																												
mE_l	<table border="1"><tr><td>X</td><td>X</td><td>O</td></tr><tr><td>X</td><td>O</td><td></td></tr><tr><td>X</td><td>O</td><td></td></tr></table>	X	X	O	X	O		X	O		<table border="1"><tr><td>X</td><td>O</td><td></td></tr><tr><td>X</td><td>O</td><td></td></tr><tr><td>X</td><td>O</td><td></td></tr></table>	X	O		X	O		X	O		<table border="1"><tr><td>X</td><td>O</td><td></td></tr><tr><td>X</td><td>O</td><td></td></tr><tr><td>X</td><td>X</td><td>O</td></tr></table>	X	O		X	O		X	X	O
X	X	O																												
X	O																													
X	O																													
X	O																													
X	O																													
X	O																													
X	O																													
X	O																													
X	X	O																												
mE_r	<table border="1"><tr><td>O</td><td>X</td><td>X</td></tr><tr><td></td><td>O</td><td>X</td></tr><tr><td></td><td>O</td><td>X</td></tr></table>	O	X	X		O	X		O	X	<table border="1"><tr><td></td><td>O</td><td>X</td></tr><tr><td></td><td>O</td><td>X</td></tr><tr><td></td><td>O</td><td>X</td></tr></table>		O	X		O	X		O	X	<table border="1"><tr><td></td><td>O</td><td>X</td></tr><tr><td></td><td>O</td><td>X</td></tr><tr><td>O</td><td>X</td><td>X</td></tr></table>		O	X		O	X	O	X	X
O	X	X																												
	O	X																												
	O	X																												
	O	X																												
	O	X																												
	O	X																												
	O	X																												
	O	X																												
O	X	X																												
mE_{bl}	<table border="1"><tr><td>X</td><td>O</td><td></td></tr><tr><td>X</td><td>O</td><td></td></tr><tr><td>X</td><td>X</td><td>O</td></tr></table>	X	O		X	O		X	X	O	<table border="1"><tr><td></td><td>O</td><td></td></tr><tr><td>X</td><td>O</td><td>O</td></tr><tr><td>X</td><td>X</td><td></td></tr></table>		O		X	O	O	X	X		<table border="1"><tr><td>O</td><td></td><td></td></tr><tr><td>X</td><td>O</td><td>O</td></tr><tr><td>X</td><td>X</td><td>X</td></tr></table>	O			X	O	O	X	X	X
X	O																													
X	O																													
X	X	O																												
	O																													
X	O	O																												
X	X																													
O																														
X	O	O																												
X	X	X																												
mE_b	<table border="1"><tr><td>O</td><td></td><td></td></tr><tr><td>X</td><td>O</td><td>O</td></tr><tr><td>X</td><td>X</td><td>X</td></tr></table>	O			X	O	O	X	X	X	<table border="1"><tr><td></td><td></td><td></td></tr><tr><td>O</td><td>O</td><td>O</td></tr><tr><td>X</td><td>X</td><td>X</td></tr></table>				O	O	O	X	X	X	<table border="1"><tr><td></td><td></td><td>O</td></tr><tr><td>O</td><td>O</td><td>X</td></tr><tr><td>X</td><td>X</td><td>X</td></tr></table>			O	O	O	X	X	X	X
O																														
X	O	O																												
X	X	X																												
O	O	O																												
X	X	X																												
		O																												
O	O	X																												
X	X	X																												
mE_{br}	<table border="1"><tr><td></td><td></td><td>O</td></tr><tr><td>O</td><td>O</td><td>X</td></tr><tr><td>X</td><td>X</td><td>X</td></tr></table>			O	O	O	X	X	X	X	<table border="1"><tr><td></td><td>O</td><td></td></tr><tr><td>O</td><td>O</td><td>X</td></tr><tr><td></td><td>X</td><td>X</td></tr></table>		O		O	O	X		X	X	<table border="1"><tr><td></td><td>O</td><td>X</td></tr><tr><td></td><td>O</td><td>X</td></tr><tr><td>O</td><td>X</td><td>X</td></tr></table>		O	X		O	X	O	X	X
		O																												
O	O	X																												
X	X	X																												
	O																													
O	O	X																												
	X	X																												
	O	X																												
	O	X																												
O	X	X																												

그림 1. 방향 에지 정제 마스크

위의 10가지 특징 값은 폰트의 종류에 영향을 받으나 동일한 폰트에 경우엔 큰 차이가 없다고 볼 수 있다.

신경망 학습으로 분석한 자막의 연결 객체의 패턴을 사용하여 연결 객체를 자막과 배경으로 분류하고 최종적으로 자막으로 판단한 연결 객체만을 남긴다.

3.4.2 자막 위치 찾기 단계

신경망 학습을 통해 분석된 자막 연결 객체의 패턴은 자막과 배경의 이진화를 수행할 것이다. 효율적인 이진화 작업을 위해선 자막이 존재하는 지역을 분리하고, 이진화를 수행하기 위한 최소한의 범위를 지정한다.

이진화를 수행할 범위를 정하기 위해서 패턴 분석에 의해 남겨진 연결 객체들로 범위를 지정하는 과정을 수행한다. 그러나 한글의 특성상 한 글자는 분리된 여러 획들로 구성이 되기 때문에 이러한 획들을 하나로 모아서 하나의 범위로 지정한다면 좀 더 많은 정보를 이용한 이진화가 가능하다. 그러기 위해 [1]에서 제안한 다중 수평, 수직 프로젝션 방법을 사용한다.

3.4.3 자막 분리 단계

수평, 수직 프로젝션 방법에 의해 자막의 분리를 수행하는 지역이 축소되었다. 자막은 배경보다는 밝다는 가정에 따라, 자막과 배경을 이루는 경계인 자막을 이루는 획의 에지는 자막과 배경의 밝기의 임계값이라고 할 수 있다. 그러므로 획 기반 이진화 방법[2]을 사용하여, 자막이 존재하는 위치로부터 연결 객체의 평균적인 밝기를 구하면 적절한 임계치를 구할 수 있다. 구해진 임계치를 이용하여 자막 위치로부터 자막과 배경을 이진화를 수행한다.

4. 실험 결과

제안한 방법에서는 크기에 무관하게 자막을 추출하는 자막 추출 방법을 제안하였다. 실험 환경은 펜티엄 4 CPU에 1기가 RAM을 사용한 컴퓨터의 VC++ 6.0 환경에서 실험하였다.

실험 데이터는 대중 방송에서 얻어진 영상을 사용하였으며, 총 31장의 영상을 사용하였다.

영상은 뉴스 및 다큐멘터리의 영상이 10개, 쇼 영상이 21개의 영상으로 총 603의 글자가 자막으로 존재하였다. 그 중에서 찾아진 자막의 글자는 총 516개였다. 글자는 15pixel에서 45pixel의 높이로 다양한 크기로 존재하였으며, 색상은 비 자막보단 밝았다.

표 1은 대중 방송으로부터 얻은 영상인 실험 데이터를 통해 얻은 자막 추출의 결과다. Precision이란 전체 찾아진 연결 객체 중에서 찾아진 자막의 비율이며, Recall이란 찾아진 자막 중에서 옳게 찾아진 자막의 비율이다.

실험 결과에서 Precision의 경우엔 자막과 유사한 비 자막 연결 객체가 패턴 분석 이후에

	Precision(%)	Recall(%)
쇼	70%	89%
뉴스	69%	83%

표 1 자막 추출 결과

도 많이 남아있음을 보여준다. 뉴스 프로에 비해 쇼 프로가 비교적 좋은 성능을 보였다. 그 이후는 뉴스 프로그램의 경우엔 지정된 범위 안에 많은 수의 글자가 존재하기 위해 작게 위치하는 경우가 존재했기 때문이었다. 이러한 결과를 보았을 때 크기가 일정 수준 이하의 자막을 추출하기 위해선 알고리즘의 개선이 필요하다고 할 수 있다.

대부분의 쇼 프로에서는 자막이 눈에 띄도록 하기 위해 영상에서 자막은 자막 테두리를 갖는 경우가 존재하였다. 이러한 경우에는 제안하는 방법이 크기에 무관한 특성 값으로 패턴을 분

석하였기 때문에 자막이 테두리보다 밝은 경우에는 제대로 자막을 추출하였다

5. 결 론

본 논문에서는 영상에서 크기에 무관하게 자막을 추출하는 연결 객체 기반의 자막 추출 방법을 제안하였다. 기존의 자막 추출 방법은 다중 해상도 방법을 사용하여, 다양한 크기의 자막을 추출하였다. 하지만 다양한 해상도에서 반복적인 작업을 수행하는 단점이 존재하였다. 제안하는 방법은 다중 해상도 방법이 아니어도 다양한 크기의 자막이 추출 가능하였다.

실험 결과의 Precision은 70%, Recall은 85% 정도로 나타났으며 자막과 배경의 밝기가 유사한 자막의 경우엔 오작동이 발생 하였고, 20pixel이하의 자막은 에지의 추출이 제대로 수행되지 않아서 정상적인 자막 추출이 수행되지 않는 단점을 보였다.

차후에 수행할 연구는 작은 글자의 추출도 가능하며, 합쳐진 연결 객체의 분리가 가능한 에지 추출 방법을 연구할 것이다. 또한, 신경망 학습을 통한 자막 연결 객체 구분 방법은 100%를 보장할 수 없기 때문에 잘 못 판별된 연결 객체를 제거하거나 복원하는 연구를 할 것이다.

감사의 글 : 본 연구는 21세기 프론티어 연구 개발 사업의 일환으로 추천되고 있는 정보통신부의 유비쿼터스컴퓨팅 및 네트워크원천기반기술 개발사업의 지원을 받았습니다.

참 고 문 헌

- [1] R. Lyu, "A Comprehensive Method for Multilingual Video Text Detection, Localization and Extraction," IEEE Transaction on Circuits and Systems for Video Technology, Vol. 15, NO. 2, FEB. 2005
- [2] 정종면, "모폴로지 연산을 이용한 강건한 자막 추출 알고리즘", 신호처리합동 학술대회 논문지, 제 19권 1호, 2006
- [3] R. C. Gonzalez, Digital Image Processing, 2nd edition, Prentice Hall, New Jersey, 2001