

## (19)대한민국특허청(KR)

### (12) 공개특허공보(A)

(51) 。 Int. Cl. (11) 공개번호 10-2006-0094006  
*H04N 5/44* (2006.01) (43) 공개일자 2006년08월28일

(21) 출원번호 10-2005-0046074

(22) 출원일자 2005년05월31일

(30) 우선권주장 1020050014905 2005년02월23일 대한민국(KR)

(71) 출원인 인천대학교 산학협력단  
 인천 남구 도화동 177번지 인천대학교내

(72) 발명자 강현철  
 인천 연수구 동춘동 938번지 건영아파트 119동 303호  
 이진성  
 인천 남구 도화동 177번지

(74) 대리인 유미특허법인

심사청구 : 있음

#### (54) 신경망 기법을 이용한 뉴스 비디오의 앵커 화면 추출 장치및 그 방법

##### 요약

신경망 기법을 이용한 뉴스 비디오의 앵커 화면 추출 장치에서, 뉴스 비디오를 앵커 화면의 구조적 특성을 갖는 앵커 화면과 그렇지 못한 비 앵커 화면으로 분류한다. 그리고 외부로부터 뉴스 영상 데이터가 입력되면 장면 전환을 검출하고, 신경망 기법을 토대로 검출된 장면 전환에 해당하는 입력 패턴과 앵커 화면에 해당하는 다수의 분류 패턴을 비교하여, 상기 입력 패턴으로부터 앵커 화면을 검출한다. 상기 분류 패턴은 앵커의 수, 뉴스 아이콘의 유무, 그리고 뉴스 아이콘의 위치 중 적어도 하나를 포함하는 앵커 구성 변수에 따라 다수의 모델로 모델링되어 있다.

따라서 신경망 기법을 이용하여 뉴스 비디오에서 앵커 화면을 정확하고 효율적으로 추출할 수 있다.

##### 대표도

도 2

##### 색인어

신경망, 뉴스 비디오, 앵커화면, 장면전환

##### 명세서

##### 도면의 간단한 설명

도 1은 본 발명의 실시 예에 따른 뉴스 비디오의 공간적 구조를 나타낸 도이다.

도 2는 본 발명의 실시 예에 따른 앵커 화면 추출 장치의 구조도이다.

도 3은 본 발명의 실시 예에 따른 앵커 화면이 갖는 공간적 구성을 4가지 신경망 입력 모델로 모델링 한 도이다.

도 5a 및 도 5b는 본 발명의 실시 예에 따른 앵커 화면을 추출하기 위해, 비 앵커 화면이 갖는 공간적 구성을 신경망의 예리 입력 모델로 모델링 한 도이다.

도 6은 본 발명의 실시 예에 따른 앵커 화면 추출 방법의 흐름도이다.

도 7은 본 발명의 실시 예에 따른 장면 전환 추출 과정을 나타낸 도이다.

도 8은 도 7에 도시된 과정에 따라 장면 전환을 검출한 결과를 나타낸 예시도이다.

도 9는 본 발명의 실시 예에 따른 신경망 모델의 ADALINE 네트워크의 구조도이다.

도 10은 본 발명의 실시 예에 따른 신경망 기법을 이용하여 앵커 화면을 추출한 결과를 나타낸 도이다.

도 11은 본 발명의 실시 예에 대한 유효성을 검증하기 위한 프로그램의 화면 구조도이다.

도 12는 본 발명의 실시 예에 따른 앵커 화면 추출 장치를 적용할 수 있는 디지털 뉴스 비디오 아카이브 시스템의 구조 예이다.

## 발명의 상세한 설명

### 발명의 목적

#### 발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 영상 추출 장치 및 그 방법에 관한 것으로, 더욱 상세하게 말하자면, 디지털 뉴스 비디오 아카이브(archive) 시스템 구현을 위한 뉴스 비디오의 앵커 화면 추출 장치 및 그 방법에 관한 것이다.

다양한 지식 자산들 가운데 영상물의 경우 아날로그 방식의 제작과 보관/배포가 일반적인 상태에서, 최근 들어 급속하게 디지털화가 이루어지고 있는 상황이다. 즉 방송의 디지털화, DVD에 의한 VHS 테이프의 대체, 디지털캠코더의 보급, 보안 감시 시장의 디지털화 등을 그 사례로 들 수 있다.

디지털 영상물은 자기디스크, 자기테이프, 광디스크 등에 보관될 수 있으며, 다양한 저장매체에 저장된 디지털 영상물의 효율적 활용을 위해서는 방대한 분량의 영상물에서 원하는 장면에 대한 손쉽고 빠른 검색이 필수적이다. 또한 디지털 방송으로의 전환과 함께 디지털 영상물에 대한 검색 수요가 필수적으로 생겨날 수밖에 없다. 향후 디지털 방송이 일반화되고, 방송 산업 외에도 여타 기업/기관/행정부처에서의 디지털 이미지 및 영상물 형태의 지식자산이 급속도로 증가하고 있으므로, 이러한 디지털 영상물에 대한 효율적 저장 및 관리/검색 솔루션이 필수적으로 요구되고 있다.

이러한 요구에 따라 디지털화된 다양하고 방대한 뉴스 정보를 쉽게 데이터베이스화하고 검색할 수 있도록 하기 위한 디지털 뉴스 비디오 아카이브 시스템이 개발되고 있다.

디지털 뉴스 비디오 아카이브 시스템에서는 뉴스 비디오를 사건 단위로 저장 및 검색을 하여야 하며, 이를 위해서는 사건의 시작 부분인 앵커 화면을 검출할 수 있어야 한다. 앵커 화면(anchorperson frame)은 뉴스 비디오에서 앵커가 사건을 소개하는 구간이다. 앵커 화면은 사건의 시작을 의미하기 때문에 뉴스를 구조적으로 분할, 저장하기 위하여 중요할 뿐 아니라, 다음에 보도할 사건을 압축하여 표현하는 뉴스 아이콘과 자막을 담고 있기 때문에, 뉴스 비디오 아카이브를 구현하기 위해서는 앵커 화면의 추출이 필수적인 선결과제가 되어야 한다. 또한 주제어나 색인어를 추출하여 추후의 검색어로 사용하기 위해서도, 앵커 화면의 추출은 중요한 의미를 갖는다.

앵커 화면을 추출하기 위해서는 뉴스 비디오를 분석하여 내용을 기반으로 뉴스 비디오를 분할하여야 하는데, 이를 위하여 가장 많이 사용하는 방법은 내용의 변환이 일어나는 곳, 즉 장면 전환(scene change)을 검출하고, 이를 경계로 장면 단위로 뉴스 비디오를 분할하는 방법이다. 이러한 장면 전환은 연속된 화면에서 밝기나 색상의 변화량이 큰 곳에서 발생하기 때문에, 현재 프레임과 이전 프레임의 밝기나 색상 값의 차이나, 밝기나 색상의 히스토그램 차이 등과 같은 화면의 변화량을 하나의 수치로 표현할 수 있는 단위로 계산하여 그 변화량이 큰 지점을 장면 전환으로 검출할 수 있다.

종래에 장면 전환 검출을 이용하여 앵커 화면을 검출하는 방법으로는 장면 전환 중에서 앵커가 출현하는 화면의 특정한 배경색과 같은 색을 갖는 화면을 앵커 화면으로 간주하거나, 앵커의 얼굴색을 계산하여 얼굴색이 나타나는 화면을 앵커 화면으로 간주하거나, 또는 앵커의 얼굴을 인식하여 앵커 화면을 검출하는 방법 등이 있다. 그러나 이 방법들은 배경색이나 얼굴색과 유사한 색상이 많은 화면을 오 검출하는 오류가 많이 발생한다.

또 다른 방법으로는 앵커의 움직임이 작다는 가정하에 움직임 양이 작은 화면을 앵커 화면으로 간주하는 방법이 있으나, 움직임 양만으로는 앵커 화면을 정확하게 추출하기가 어렵다.

### 발명이 이루고자 하는 기술적 과제

그러므로 본 발명이 이루고자 하는 기술적 과제는 종래의 문제점을 해결하기 위한 것으로, 신경망 기법을 이용하여 뉴스 비디오에서 앵커 화면을 정확하게 추출하고자 하는데 있다.

특히, 본 발명이 이루고자 하는 기술적 과제는 뉴스 비디오에서 앵커 화면의 구조적 특성을 고려하여, 상기 앵커 화면을 정확하게 추출하고자 하는데 있다.

### 발명의 구성 및 작용

이러한 기술적 과제를 달성하기 위하여, 본 발명의 특징에 따른 앵커 화면 추출 장치는 뉴스 비디오에 해당하는 뉴스 영상 데이터로부터 앵커 화면을 추출하는 장치에서, 외부로부터 인가되는 상기 뉴스 영상 데이터를 입력받는 영상 입력부; 상기 뉴스 영상 데이터로부터 장면 전환(scene change)을 검출하고, 검출된 장면 전환에 해당하는 입력 패턴을 출력하는 장면 전환 검출부; 및 앵커 화면에 해당하는 다수의 분류 패턴이 저장되어 있는 분류 패턴 저장 모듈, 그리고 신경망 기법을 토대로 상기 입력 패턴과 상기 저장되어 있는 분류 패턴을 비교하여, 상기 입력 패턴으로부터 앵커 화면을 검출하는 패턴 비교 모듈을 포함하는 앵커 화면 검출부를 포함한다.

이러한 특징을 가지는 본 발명에서는, 상기 신경망 기법을 이용하여 앵커 화면을 추출하기 위하여, 상기 분류 패턴은 앵커의 수, 뉴스 아이콘의 유무, 그리고 뉴스 아이콘의 위치 중 적어도 하나를 포함하는 앵커 화면 구성 변수를 토대로 하여 모델링된 다수의 분류 패턴으로 이루어진다. 그리고 상기 분류 패턴은 앵커 화면과 비앵커 화면 중 하나로 분류되며, 상기 앵커 화면은 적어도 4가지 이상의 모델로 모델링되고, 상기 비앵커 화면은 적어도 8가지 이상의 모델로 모델링될 수 있다.

또한 본 발명의 다른 특징에 따른 앵커 화면 추출 방법은, 뉴스 비디오에 해당하는 뉴스 영상 데이터로부터 앵커 화면을 추출하는 장치의 앵커 화면 추출 방법에서, 외부로부터 상기 뉴스 영상 데이터--상기 뉴스 영상 데이터는 I-프레임을 포함하는 MPEG(motion picture expert group) 비트 스트림으로 이루어짐--가 입력되면, 상기 장치가 뉴스 영상 데이터로부터 I 프레임에 해당하는 화면만을 대상으로 장면 전환을 검출하고, 검출된 장면 전환에 해당하는 입력 패턴을 출력하는 단계; 및 상기 장치가 신경망 기법을 토대로 앵커 화면에 해당하는 다수의 분류 패턴과 상기 입력 패턴을 비교하여, 상기 입력 패턴으로부터 앵커 화면을 검출하는 단계를 포함하며, 상기 분류 패턴은 앵커의 수, 뉴스 아이콘의 유무, 그리고 뉴스 아이콘의 위치 중 적어도 하나를 포함하는 앵커 화면 구성 변수를 토대로 하여 모델링된 다수의 분류 패턴으로 이루어진다.

이하, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 본 발명을 용이하게 실시할 수 있는 가장 바람직한 실시 예를 첨부된 도면을 참조로 하여 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시 예에 한정되지 않는다.

뉴스는 시간 순으로 앵커가 말하는 부분과 뒤이어 나오는 비 앵커 구간인 기자가 설명하는 화면으로 규칙적으로 구성되어 있으며, 도 1에 전형적인 뉴스 비디오가 도시되어 있다. 첨부한 도 1에 도시되어 있듯이, 뉴스 비디오는 앵커 구간, 자막, 뉴스 아이콘 등으로 구성된다. 이러한 뉴스 비디오를 효율적으로 검색하기 위하여 뉴스를 사건별로 구조적으로 분할하여 저장하여야 한다. 이를 위해서는 뉴스 비디오에서 앵커 구간을 정확하게 검출하여야 한다.

따라서 본 발명의 실시 예에서는 넓은 범위의 디지털 비디오 아카이브 시스템에 속하는 디지털 뉴스 비디오 아카이브 시스템을 구현하는 데 필수적인 과제인 앵커 구간 즉, 앵커 화면을 신경망 기법을 이용하여 추출한다.

도 2는 본 발명의 실시 예에 따른 신경망 기법을 이용한 뉴스 비디오의 앵커 화면 추출 장치(이하, 설명의 편의상 "앵커 화면 추출 장치"라고 명명함)의 구조를 나타낸 도이다. 도 3은 도 2에 도시된 장면 전환 검출부 및 앵커 화면 추출부의 상세 구조가 도시되어 있다.

첨부한 도 2에 도시되어 있듯이, 본 발명의 실시 예에 따른 앵커 화면 추출 장치는 영상 입력부(10), 버퍼(20), 장면 전환 검출부(30), 앵커화면 추출부(40), 부호화부(50), 및 출력부(60)를 포함한다.

영상 입력부(10)는 외부로부터 인가되는 뉴스 영상 데이터를 제공받아 버퍼(20)에 버퍼링하며, 여기서 뉴스 영상 데이터는 카메라 등에 의하여 촬영되어 전기적으로 처리할 수 있도록 영상 처리된 데이터이다. 예를 들어, 뉴스 영상 데이터는 MPEG(Moving Picture Experts Group) 표준에 따라 부호화된 비트 스트림 형태로 처리된 데이터이다.

장면 전환 검출부(30)는 뉴스 영상 데이터로부터 장면 전환(scene change)을 검출하고, 이를 경계로 장면 단위로 뉴스 영상 데이터를 분할한다.

본 발명의 실시 예에 따른 장면 전환 검출부(30)는 도 3에 도시되어 있듯이, 뉴스 영상 데이터로부터 장면 전환 검출을 위한 DC 영상을 구하는 DC 영상 생성 모듈(31), DC 영상에 대하여 이진화를 수행하는 이진화 모듈(32), 이진화된 DC 영상을 정규화(normalization)하는 정규화 모듈(33)을 포함한다.

이러한 구조로 이루어지는 장면 전환 검출부(30)는 영상 데이터의 밝기 성분뿐만 아니라, 색상 성분의 조합에 의한 히스토그램 누적차를 이용하여 장면 전환을 검출한다. 그리고 검출된 장면 전환을 정규화하여 앵커 화면 검출을 위한 입력 패턴으로 출력한다.

한편, 앵커 화면 추출부(40)는 검출된 장면 전환에서 앵커 화면을 추출하며, 부호화부(50)는 이와 같이 추출된 앵커 화면을 부호화하여 출력부(60)를 통하여 출력하거나 저장한다.

본 발명의 실시 예에 따른 앵커 화면 추출부(40)는 신경망 기법을 이용하여 앵커 화면을 추출하기 위하여, 도 3에서와 같이, 앵커 화면 추출을 위한 신경망 기법에 따른 분류 패턴에 해당하는 데이터가 저장되어 있는 패턴 저장 모듈(41), 저장되어 있는 분류 패턴과 장면 전환 추출부(30)에 의하여 출력되는 입력 패턴을 비교하여 상기 입력 패턴이 어떠한 분류 패턴에 해당하는지를 판별함으로써, 앵커 화면을 추출하는 패턴 비교 모듈(42)을 포함한다.

신경망은 패턴 분류 분야에서 광범위하게 사용되는 기법으로, 미지의 입력 패턴을 입력 패턴과 가장 유사한 소정의 미리 정해진 패턴 클래스로 할당하는 방법으로, 입력 패턴을 분류한다. 본 발명의 실시 예에서는 신경망을 이용하여 앵커 화면을 검출하기 위하여, 앵커 화면을 다수의 앵커 모델로 모델링하였다.

다음에는 본 발명의 실시 예에 따른 앵커 화면에 대한 모델링에 따라 다수의 분류 패턴을 생성하는 방법에 대하여 설명한다.

앵커 화면 추출을 위해서는 먼저, 뉴스 영상 데이터에서 앵커가 존재하는 화면의 공간적 구조 특성을 분석하여 앵커의 얼굴 영역에 대한 사전 지식을 설정한 후, 이를 바탕으로 앵커의 얼굴을 검출하여 정확히 앵커로 인식된 화면을 키 프레임으로 설정하여야 한다.

구체적으로 앵커 화면에서는 앵커의 얼굴이 일정한 크기를 가지고 나타나고, 얼굴의 움직임은 아주 미세하다는 점, 비디오 전반에 걸쳐 나타나는 배경이 유사하다는 점의 공간적 특징을 이용할 수 있다. 따라서 이러한 공간적인 특징을 신경망의 패턴 클래스로 모델링 할 수 있다.

앵커 화면에는 보도할 뉴스를 압축하여 표현하는 뉴스 아이콘이 존재할 수도 있고 존재하지 않을 수도 있다. 하지만 뉴스 아이콘의 존재 여부와 상관없이 앵커의 얼굴 영역은 항상 일정하다는 사실을 알 수 있다. 따라서 이러한 경우를 모두 신경망 패턴 클래스로 설정하기 위하여, 앵커의 수와 뉴스 아이콘의 유무, 그리고 뉴스 아이콘의 위치를 토대로 하는 앵커 화면

의 공간적 구성에 따라 앵커 화면을 4가지 모델로 모델링하였다. 이하에서는 설명의 편의를 위하여 앵커 화면의 공간적 구성의 특징을 나타내는 앵커의 수와 뉴스 아이콘의 유무, 그리고 뉴스 아이콘의 위치를 "앵커 화면 구성 변수"라고 명명한다.

도 4는 본 발명의 실시 예에 따른 앵커 화면 구성 변수에 따라 앵커 화면을 4가지 모델로 모델링한 것을 나타낸 예시도이다.

첨부한 도 4에 도시되어 있듯이, 앵커 화면 구성 변수에 따라 앵커 화면을 단일 앵커이며 뉴스 아이콘이 없는 제1 모델(M1), 단일 앵커이며 앵커의 오른쪽 위(또는 아래) 부분에 뉴스 아이콘이 있는 제2 모델(M2), 단일 앵커이며 앵커의 왼쪽 위(또는 아래) 부분에 뉴스 아이콘이 있는 제3 모델(M3), 그리고 복수의 앵커가 나타나는 제4 모델(M4)로 모델링한다.

한편 비 앵커 화면 즉, 앵커가 나타나지 않는 화면도 화면의 밝기의 분포에 따라 모델링할 수 있으며, 도 5a 및 도 5b는 본 발명의 실시 예에 따라 비 앵커 화면을 모델링한 것을 나타낸 도이다. 비 앵커 화면을 도 5a 및 도 5b에 도시되어 있듯이, 화면 밝기의 분포에 따라 8가지의 예로 모델링할 수 있다.

위에 기술된 바와 같이, 본 발명의 실시 예에서는 패턴 클래스를 크게 앵커 화면과 비 앵커 화면의 두 가지 클래스로 설정하고, 각각의 클래스를 다시 4가지와 8가지의 서브클래스로 구성하여 신경망을 학습시킨다. 그리고 추후에 입력되는 뉴스 영상 데이터에 따른 입력 패턴과 가장 유사한 패턴 클래스를 찾음으로써, 입력 패턴이 앵커 화면인지 또는 비 앵커 화면인지를 판단하여 앵커 화면을 추출한다.

다음에는 위에 기술된 구조를 토대로 하여 본 발명의 실시 예에 따른 앵커 화면 추출 장치의 동작에 대하여 보다 구체적으로 설명한다.

도 6은 본 발명의 실시 예에 따른 앵커 화면 추출 방법의 흐름도이다.

외부의 방송 시스템으로부터 전송되거나 또는 카메라 등에 의하여 촬영되어 별도의 기록 매체(예: 비디오테이프, DVD 등)에 저장된 뉴스 영상 데이터가 영상 입력부(10)를 통하여 입력되면, 상기 뉴스 영상 데이터는 버퍼(20)에 버퍼링된다(S100). 이러한 뉴스 영상 데이터는 MPEG 비트 스트림으로 구성되어 있는 데이터이다.

일반적으로 MPEG에서는 모든 프레임을 개별 정화상으로 압축하는 것이 아니라, 인접 프레임 사이에 유사점이 많다는 점을 이용하여 동작 보상을 하는데 있어 예측과 보간을 이용한다. 그러나 한편 임의의 접근과 같은 VCR 식 제어가 가능해야 한다는 등의 여러 이유로 인해 MPEG 영상 데이터에는 자신이 가지고 있는 정보만으로도 복원될 수 있는 프레임이 규칙적으로 삽입되어야 한다.

그러므로 MPEG 영상 데이터는 정화상으로 압축된 프레임 I(Intra-coded) 프레임, 예측만을 한 P(Predictive-coded) 프레임, 양방향 보간을 한 B(Bidirectional-coded) 프레임이 일정한 패턴으로 섞여 있다. 여기서 I-프레임은 데이터 스트림의 어느 위치에도 올 수 있으며, 데이터의 임의의 접근을 위해 사용되며, 다른 이미지들의 참조 없이 부호화된다. I-프레임은 매크로 블록내에서 지정된 8×8 블록으로 나누어진 다음, 블록 단위로 부호화된다. 그리고 I-프레임으로 시작하는 연속적인 영상들의 집합을 GOP(Group Of Picture)라고 한다.

GOP에서 적어도 하나의 I 프레임이 한번은 부호화되기 때문에, 본 발명의 실시 예에 따른 장면 전환 검출부(30)는 이러한 MPEG의 뉴스 영상 데이터로부터 I 프레임에 해당하는 화면만을 대상으로 장면 전환을 검출한다. 즉, MPEG 비트 스트림을 완전하게 복호화 하지 않고 필요한 정보가 있는 비트 스트림만 복호화하는 부분 복호화를 수행하여, 복호화에 필요한 시간을 단축한다. 도 7에 본 발명의 실시 예에 따른 장면 전환 검출의 예가 도시되어 있다.

구체적으로 도 7을 토대로 하여 설명하면, 장면 전환 검출부(30)의 DC 영상 생성 모듈(31)은 뉴스 영상 데이터에서 I 프레임만을 대상으로 하여, I 프레임에서 소정 블록(8×8 블록)의 화소값을 대표하는 DC 계수만을 이용하여 부분적으로 복호화하여 DC 영상(40×30)을 구한다(S110~S120).

다음, 장면 전환 검출부(30)의 이진화 모듈(32)은 구해진 DC 영상에 대하여 이진화 과정을 수행한다(S130). 특히 상기에서 구해진 DC 영상의 밝기 값의 평균값을 임계치로 사용하며, 상기 DC 영상을 구성하는 화소의 밝기값과 임계치를 비교하여 이진화를 수행한다. 예를 들어, 소정 화소의 밝기값이 임계치보다 크면 흰색에 해당하는 값(예: +1)을 부여하고, 소정 화소의 밝기값이 임계치보다 작으면 검정색에 해당하는 값(예: -1)을 부여한다. 여기서 상기 임계치는 상기 평균값에 한정되지 않으며 주어진 환경에 따라 그 값이 다르게 설정될 수 있다.

그리고 정규화 모듈(33)이 이와 같이 이진화된 영상을 정규화하여 도 7에 예시된 바와 같은 정규화 패턴 즉, 입력 패턴을 생성한다(S140). 여기서는 이진화 영상을 10×10으로 정규화 하였으나, 이에 한정되지는 않는다. 도 8에 입력되는 뉴스 영상 데이터로부터 검출된 장면 전환의 예가 도시되어 있다.

위에 기술된 바와 같이 장면 전환이 검출되면, 검출된 장면 전환에 해당하는 입력 패턴을 신경망 기법을 이용하여 앵커 화면과 비앵커 화면의 두 가지의 클래스로 분류시킨다.

본 발명의 실시 예에서는 ADALINE 신경망을 이용하여 입력 패턴을 두 개의 클래스로 각각 분류하였으나, 반드시 이에 한정되지는 않는다. 도 9에 본 발명의 실시 예에 사용된 ADALINE 신경망의 구조가 도시되어 있다. 일반적으로 ADALINE 신경망은 도 9에서와 같이, 입력 레이어 및 출력 레이어로 이루어지며, 입력 레이어는 네트워크에 대한 각 입력을 위한 입력노드(input node)와 바이어스 노드(bias node)를 포함한다. 그리고 출력 레이어는 네트워크를 위한 출력을 산출하는 ADALINE 노드만을 포함한다. 입력 레이어에 있는 각 노드는 하나의 링크로 ADALINE 노드에 연결된다.

이러한 ADALINE 신경망으로 입력되는 입력 패턴은 미리 결정된 범위(일반적으로 -1.0~1.0이 사용된다)로 정규화된 성분들을 가져야 하기 때문에, 위에 기술된 바와 같이 검출된 장면 전환에 대한 영상 데이터가 정규화되었다. 이러한 정규화는 입력 패턴의 한 성분이 우세하여 네트워크의 동작을 교란하는 것을 방지한다.

ADALINE 신경망에서 요구되는 출력들은 항상 두 분류 클래스에 해당하는 두가지의 값들 중 하나가 되어야만 한다. 따라서, 본 발명의 실시 예에서는 입력 패턴에 대한 출력값이 앵커 화면 또는 비앵커 화면 중 하나가 되어야 한다. 여기서 ADALINE 신경망의 입력 벡터의 개수는 100으로 하고, 정규화 입력 패턴에서 화소 값이 0에 해당하면 입력패턴 -1, 화소 값이 255에 해당하는 +1의 입력패턴을 가지도록 하였다.

위에 기술된 바와 같은 신경망 네트워크를 기본으로 하여, 앵커 화면 추출부(40)의 패턴 비교 모듈(41)은 장면 전환 추출부(30)로부터 제공되는 입력 패턴과 저장 모듈(42)에 저장되어 있는 분류 패턴을 비교하여 입력 패턴에 대한 분류 클래스를 결정한다(S150~S160). 예를 들어, ADALINE 신경망을 토대로 입력 노드를 통하여 입력되는 입력 패턴의 값과 그것에 대응하는 링크의 값의 곱을 합한 다음에, 문턱값 함수가 가중된 합에 적용되어 출력값이 두 분류 클래스 중 하나의 값에 해당되도록 강제한다. 문턱값 함수는 어떤 양의 합(positive sum)을 분류 클래스 A(앵커 화면)와 연관된 값으로 변환하고, 어떤 음의 합(negative sum)은 분류 클래스 B(비앵커 화면)과 관련된 값으로 변환한다.

이러한 ADALINE 신경망은 훈련 과정 동안에 링크들과 관련된 가중 값들(weight values)을 수정함으로써 학습한다. 훈련은 보통 델타규칙으로서 알려진 학습규칙을 가지고 링크 값들을 조정한다.

델타규칙은 링크값들을 정정하기 위해서 ADALINE에 의해서 산출된 에러(하나의 패턴을 잘못된 카테고리로 정렬하는 것)를 사용하며 그래서 입력패턴이 제시되는 다음 번에는 정확한 답이 산출되도록 한다.

본 발명의 실시 예에 따른 델타규칙에 따라 산출되는 가중값은 다음 수학적 식 1에 따라 산출될 수 있다.

$$W_i(k+1) = W_i(k) + \alpha e_i(k)P(k)$$

여기서  $W_i$ 는 가중값을 나타내며,  $\alpha$ 는 학습율,  $e_i$ 는 에러 그리고  $P$ 는 입력값(입력패턴의 소정 화소값)을 나타낸다.

도 10에 위에 기술된 바와 같이 동작하는 본 발명의 실시 예에 따른 앵커 화면 추출 장치의 유효성을 검증하기 위한 프로그램의 화면 구조가 예시되어 있으며, 도 11에는 본 발명의 실시 예에 따른 신경망 기법을 이용하여 추출된 앵커 화면 결과가 예시되어 있다. 첨부한 도 10 및 도 11에 따르면 뉴스 비디오에서 정확하게 앵커 화면이 검출되는 것을 알 수 있다.

이러한 실시 예에 따른 앵커 화면 추출 장치를 도 12에 도시된 바와 같은 구조로 구현될 수 있는 디지털 뉴스 비디오 아카이브 시스템에 적용되어 사용될 수 있다.

비록, 본 발명이 가장 실제적이며 바람직한 실시 예를 참조하여 설명되었지만, 본 발명은 상기 개시된 실시 예에 한정되지 않으며, 후술되는 특허청구범위 내에 속하는 다양한 변형 및 등가물들도 포함한다.

## 발명의 효과

이상에서와 같이 본 발명의 실시 예에 따르면 뉴스 비디오의 효율적인 저장과 검색을 위한 디지털 뉴스 비디오 아카이브 시스템 구현을 위하여, 신경망 기법을 이용하여 앵커 화면을 정확하게 검출할 수 있다.

이러한 본 발명은 영상의 검색, 편집 분야 등에 응용될 수 있으며, 또한 디지털 뉴스 아카이브의 구현에 필수적인 기술이 되어 디지털 방송국의 초기 구축에 기여할 뿐만 아니라, 디지털 신문이나 영화 등의 다른 디지털 영상 아카이브에 응용될 수 있을 것으로 기대된다.

## (57) 청구의 범위

### 청구항 1.

뉴스 비디오에 해당하는 뉴스 영상 데이터로부터 앵커 화면을 추출하는 장치에서,

외부로부터 인가되는 상기 뉴스 영상 데이터를 입력받는 영상 입력부;

상기 뉴스 영상 데이터로부터 장면 전환(scene change)을 검출하고, 검출된 장면 전환에 해당하는 입력 패턴을 출력하는 장면 전환 검출부; 및

앵커 화면에 해당하는 다수의 분류 패턴이 저장되어 있는 분류 패턴 저장 모듈, 그리고 신경망 기법을 토대로 상기 입력 패턴과 상기 저장되어 있는 분류 패턴을 비교하여, 상기 입력 패턴으로부터 앵커 화면을 검출하는 패턴 비교 모듈을 포함하는 앵커 화면 검출부

를 포함하는 앵커 화면 추출 장치.

### 청구항 2.

제1항에 있어서

상기 분류 패턴은 앵커의 수, 뉴스 아이콘의 유무, 그리고 뉴스 아이콘의 위치 중 적어도 하나를 포함하는 앵커 화면 구성 변수를 토대로 하여 모델링된 다수의 분류 패턴으로 이루어지는 앵커 화면 추출 장치.

### 청구항 3.

제1항 또는 제2항에 있어서

상기 분류 패턴은 앵커 화면과 비앵커 화면 중 하나로 분류되고, 상기 앵커 화면에 해당하는 분류 패턴은 단일 앵커이며 뉴스 아이콘이 없는 제1 모델, 단일 앵커이며 앵커의 오른쪽 그리고 위 또는 아래 부분에 뉴스 아이콘이 있는 제2 모델, 단일 앵커이며 앵커의 왼쪽 그리고 위 또는 아래 부분에 뉴스 아이콘이 있는 제3 모델, 및 복수의 앵커가 나타나는 제4 모델 중 하나로 모델링되어 있는 앵커 화면 추출 장치.

### 청구항 4.

제3항에 있어서

상기 장면 전환 검출부는 상기 뉴스 영상 데이터로부터 장면 전환 검출을 위한 DC 영상을 구하는 DC 영상 생성 모듈, DC 영상에 대하여 이진화를 수행하는 이진화 모듈, 및 이진화된 DC 영상을 정규화(normalization)하여 입력 패턴을 생성하는 정규화 모듈을 포함하는 앵커 화면 추출 장치./

## 청구항 5.

제4항에 있어서

상기 뉴스 영상 데이터는 I-프레임을 포함하는 MPEG 비트 스트림으로 이루어지며, 상기 장면 전환 검출부는 상기 MPEG의 뉴스 영상 데이터로부터 I 프레임에 해당하는 화면만을 대상으로 장면 전환을 검출하는 앵커 화면 추출 장치.

## 청구항 6.

제3항에 있어서

상기 앵커 화면에 해당하는 분류 패턴은 제1 내지 제4 모델 중 하나로 모델링되어 있고, 상기 비앵커 화면에 해당하는 분류 패턴은 제1 내지 제8 에러 모델 중 하나로 모델링되어 있는 앵커 화면 추출 장치.

## 청구항 7.

뉴스 비디오에 해당하는 뉴스 영상 데이터로부터 앵커 화면을 추출하는 장치의 앵커 화면 추출 방법에서,

외부로부터 상기 뉴스 영상 데이터--상기 뉴스 영상 데이터는 I-프레임을 포함하는 MPEG 비트 스트림으로 이루어짐--가 입력되면, 상기 장치가 뉴스 영상 데이터로부터 I 프레임에 해당하는 화면만을 대상으로 장면 전환을 검출하고, 검출된 장면 전환에 해당하는 입력 패턴을 출력하는 단계; 및

상기 장치가 신경망 기법을 토대로 앵커 화면에 해당하는 다수의 분류 패턴과 상기 입력 패턴을 비교하여, 상기 입력 패턴으로부터 앵커 화면을 검출하는 단계

를 포함하며,

상기 분류 패턴은 앵커의 수, 뉴스 아이콘의 유무, 그리고 뉴스 아이콘의 위치 중 적어도 하나를 포함하는 앵커 화면 구성 변수를 토대로 하여 모델링된 다수의 분류 패턴으로 이루어지는 앵커 화면 추출 방법.

## 청구항 8.

제7항에 있어서

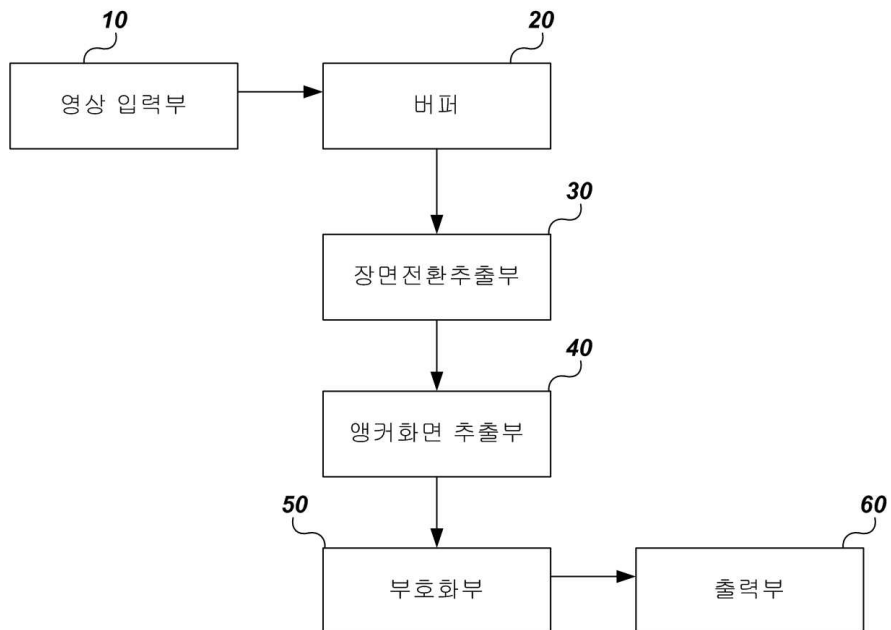
상기 분류 패턴은 상기 앵커 화면 구성 변수에 따라 모델링된, 단일 앵커이며 뉴스 아이콘이 없는 제1 모델, 단일 앵커이며 앵커의 오른 쪽 그리고 윗 또는 아래 부분에 뉴스 아이콘이 있는 제2 모델, 단일 앵커이며 앵커의 왼 쪽 그리고 윗 또는 아래 부분에 뉴스 아이콘이 있는 제3 모델, 및 복수의 앵커가 나타나는 제4 모델 중 어느 하나에 해당하는 분류 패턴으로 이루어지는 앵커 화면 추출 방법.

도면

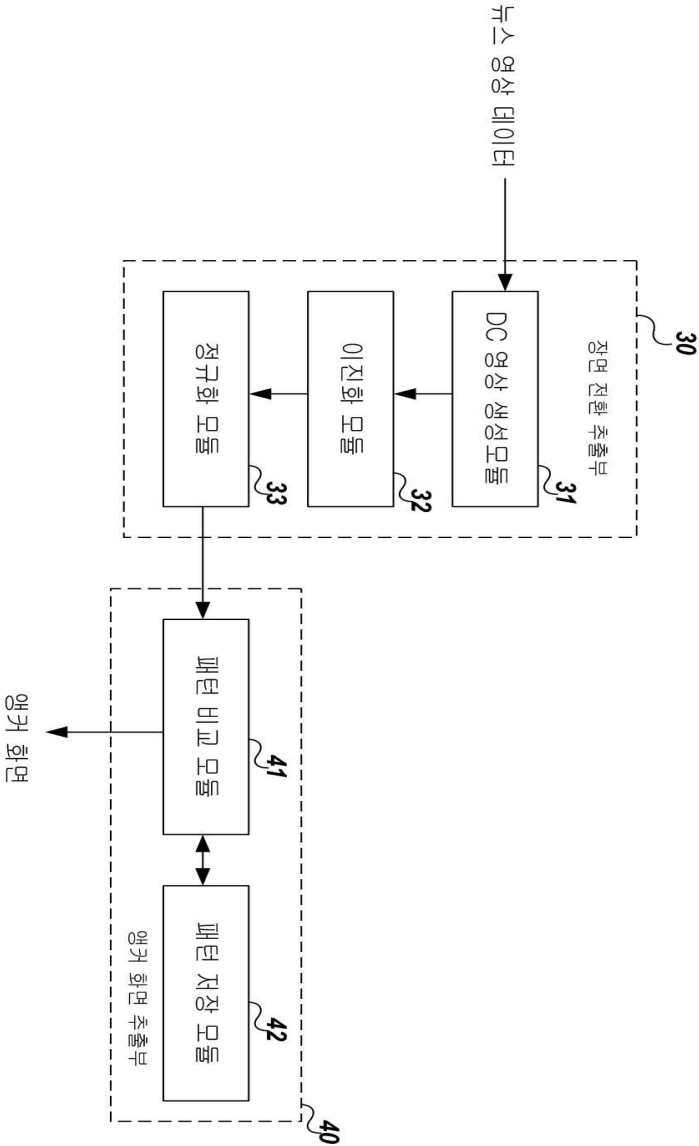
도면1



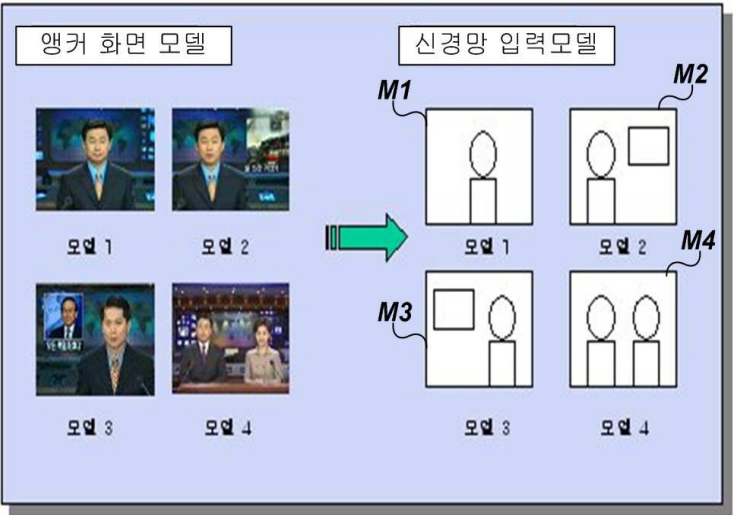
도면2



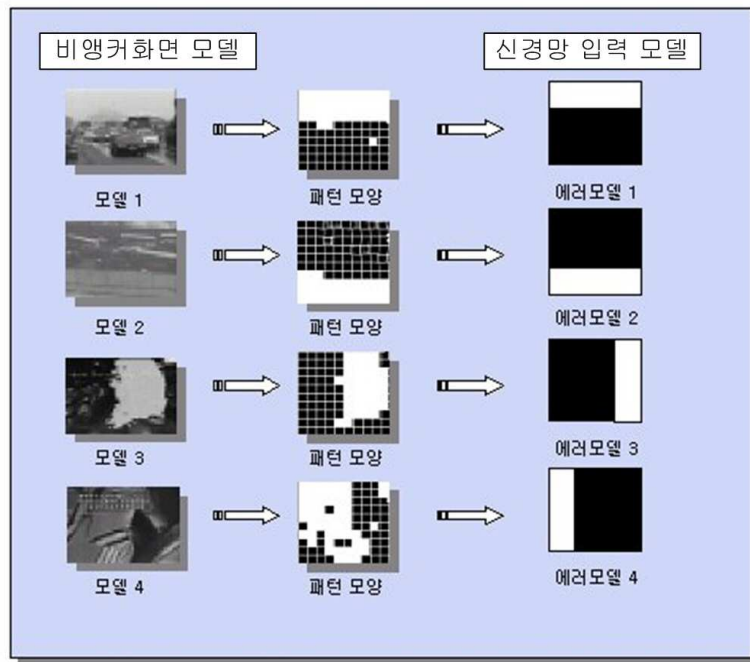
도면3



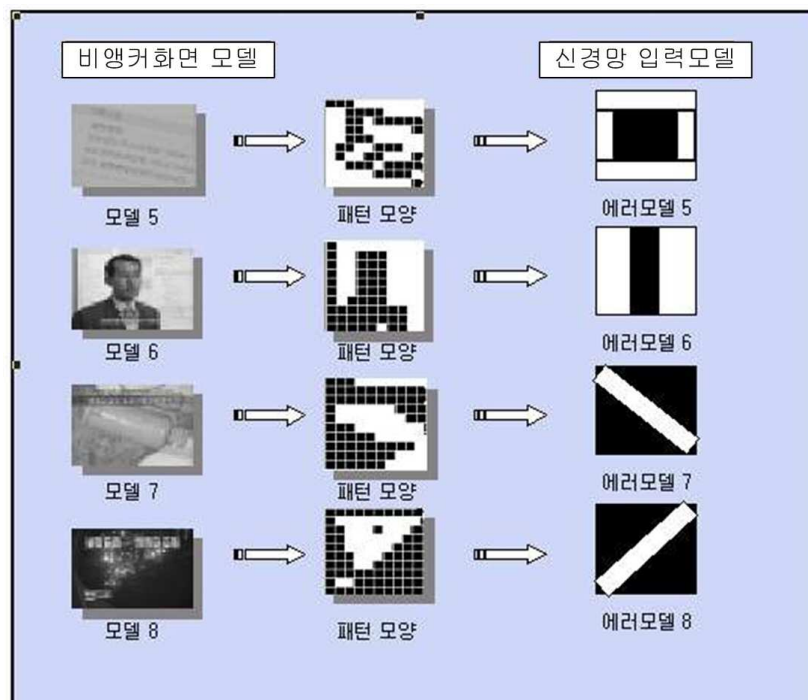
도면4



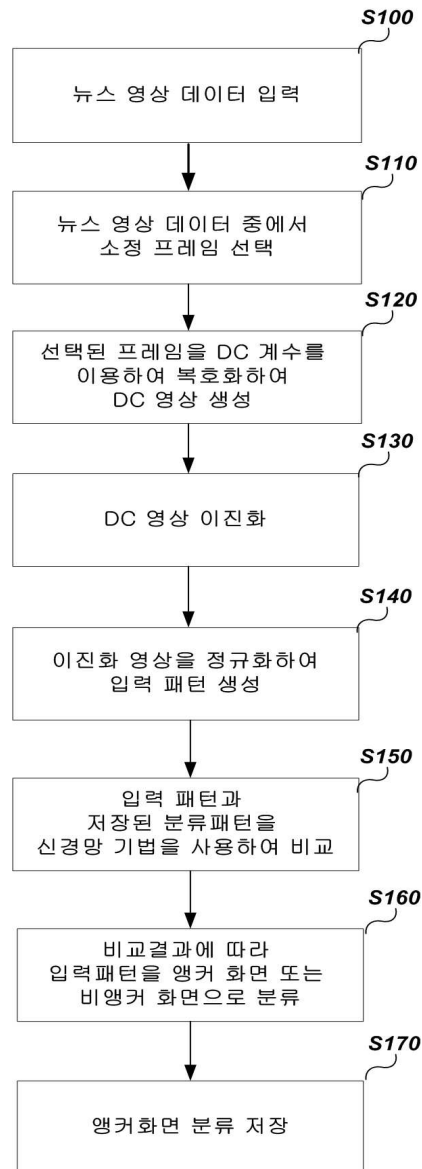
도면5a



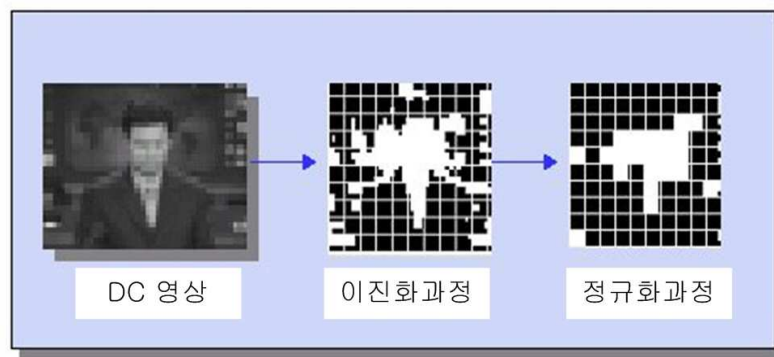
도면5b



도면6



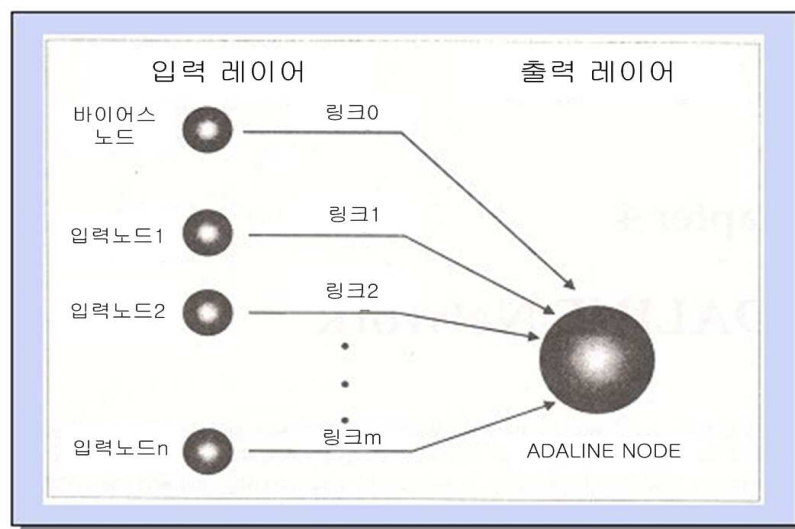
도면7



도면8



도면9



도면10



도면11



도면12

