

강인한 특징 변수 선별과 신경망을 이용한 장면 전환점 검출 기법

홍승범[†], 홍교영^{**}

요 약

본 논문은 여러 가지 장면 검출 방식들 중 강인한 특징 변수들의 선별과 신경망을 이용하여 향상된 장면 전환점 검출 기법을 제안한다. 기존의 장면 전환점 검출 방식에서는 인접한 프레임 간에 단일 특징과 고정된 임계값을 주로 사용하였다. 하지만, 비디오 시퀀스 내의 장면 전환점에서는 인접한 프레임 간의 내용(content)인 컬러, 모양, 배경 혹은 질감 등이 동시에 변화한다. 따라서 단일 특징보다는 상호 보완 관계를 갖는 강인한 특징을 이용하여 장면 전환점을 효율적으로 검출한다. 본 논문에서 강인한 특징 변수들을 선택하기 위해, 데이터 마이닝 기법 중 대표적인 CART(classification and regression tree)를 이용하고, 다차원 변수에 따른 임계값을 선정하기 위해 역전파 신경망(backpropagation neural net)을 이용한다. 제안한 방식과 대표적인 특징 추출인 PCA(principal component analysis)기법을 비교하여 특징 변수의 추출 성능을 평가한다. 실험 결과에 따라 제안된 방식이 PCA 기법과 비교하여 우수한 성능이 나타남을 확인한다.

Robust Feature Selection and Shot Change Detection Method Using the Neural Networks

Seung Bum Hong[†], Gyo Young Hong^{**}

ABSTRACT

In this paper, we propose an enhancement shot change detection method using the neural net and the robust feature selection out of multiple features. The previous shot change detection methods usually used single feature and fixed threshold between consecutive frames. However, contents such as color, shape, background, and texture change simultaneously at shot change points in a video sequence. Therefore, in this paper, we detect the shot changes effectively using robust features, which are supplementary each other, rather than using single feature. In this paper, we use the typical CART (classification and regression tree) of data mining method to select the robust features, and the backpropagation neural net to determine the threshold of the each selected features. And to evaluation the performance of the robust feature selection, we compare the proposed method to the PCA(principal component analysis) method of the typical feature selection. According to the experimental result, it was revealed that the performance of our method had better than the PCA method.

Key words: Shot Change Detection(장면 전환 검출), Feature Selection(특징 변수 추출), CART (Classification And Regression Tree), Backpropagation Neural Net(역전파 신경망)

* 교신저자(Corresponding Author) : 홍승범, 주소 : 충청남도 서산시 해미면 대곡리(356-706), 전화 : 041)660-1531, FAX : 041)660-1412, E-mail : sbhong@hanseo.ac.kr
접수일 : 2003년 9월 5일, 완료일 : 2003년 12월 15일

[†] 한서대학교 항공전자시뮬레이션학과 전임강사

^{**} 한서대학교 항공전자시뮬레이션학과 조교수
(E-mail : kiathgy@hanseo.ac.kr)

1. 서 론

멀티미디어의 방대한 데이터양을 저장, 브라우징(browsing), 인덱싱(indexing), 그리고 검색(retrieval)을 위해 많은 연구가 진행되어왔다. 특히 내용 기반 검색 기법은 영상내의 내용(contents)인 컬러, 모양, 그리고 질감 등의 특징을 추출하여 데이터베이스

를 구성하거나 질의 영상의 검색을 위한 속성들(attributions)로 사용하게 되므로 영상 데이터베이스의 인덱싱 및 검색 분야에서 활발히 연구되어오고 있다[1,2].

내용량의 비디오 시퀀스에서 효율적으로 검색, 관리하기 위해 전체 비디오 시퀀스를 샷(shot) 혹은 장면(scene) 단위로 비디오 분할한 후 각 단위를 대표 프레임(key frame)을 추출하게 된다. 샷은 시간 혹은 공간에서 연속적인 동작 특성을 가지고 있으며, 한 대의 카메라로 연속적으로 녹화된 프레임 시퀀스로 구성된다. 그리고 장면은 의미적으로 연관성이 많은 샷들의 집합들로 정의된다[3-5,10]. 본 논문에서는 샷과 장면은 프레임간의 컬러, 배경, 그리고 객체의 유사성이 높게 나타나게 되므로 구분 없이 접근한다.

기존의 장면 전환점을 검출한 방식을 살펴보면, Yusoff[3]는 급격한 장면 변화를 검사하기 위해 공간 도메인과 주파수 도메인 등을 대표하는 다섯 가지의 검출 기법을 선정하여 [4]에서 ROC(receiver operating characteristic) 곡선을 이용하였다[4,6]. 하지만 임계값 결정을 위한 비디오 시퀀스별로 ROC 곡선을 구하기 때문에 많은 연산이 필요한 문제점을 가지고 있다. Lupatini[5]는 다양한 비디오 시퀀스 종류(영화, 뉴스, 다큐멘터리, 그리고 광고)들에 대하여 히스토그램 기반, 움직임 기반, 윤곽선 기반 알고리즘과 같은 여러 가지 장면 전환점 검출 기법을 적용하여 각 종류별 뛰어난 성능을 발휘하는 방식을 추출하였지만, 검출 기법 간 통합하지는 않았다.

본 논문에서는 여러 가지 장면 검출 방식들 중 성능이 뛰어난 방식들을 선별하는 특징 변수 추출(feature selection) 기법과 다차원 변수로 패턴을 구분하는 분류 기법(classification method)을 결합하여 향상된 장면 전환점 검출 기법을 제안한다.

대표적인 특징 변수 추출 방식인 PCA(principal component analysis) 기법은 데이터 분포 중 넓은 분포를 갖는 주축을 선택하여 강인한 특징 변수를 선별하는 방식이다. 그러나 데이터 분포 내에 고립된 데이터(outlier data)가 존재할 경우 주축을 결정에 많은 영향을 주는 문제점이 있다[6,7]. 그래서 데이터 분포의 통계적 접근보다 여러 변수로 패턴을 분류하는 기법인 CART(classification and regression tree)를 이용하여 강인한 특징 변수들을 선별한다. 또한 여러 가지의 특징 변수로 효과적인 장면 전환점의

임계값을 결정하기 위해 역전파 신경망(backpropagation neural net)을 사용하고, 제안한 CART 기법을 PCA 기법과 비교하여 장면 전환점 검출 성능을 평가한다.

본 논문의 구성은 2장에서 기존의 장면 전환점 검출 기법들을 소개하고, 3장에서는 강인한 특징 변수를 선택하는 CART 기법과 PCA 기법을 설명한다. 그리고 4장에서는 역전파 신경망을 설명하고, 5장과 6장에서는 장면 전환점 검출을 수행한 후 성능을 평가하고 결론을 맺는다.

2. 장면 전환점 검출을 위한 기존의 특징 변수

본 논문에서는 인접한 프레임간의 유사도를 검사하기 위해 차이 기반 기법에 기반을 둔 9가지 그리고 상관 기반 기법에 기반을 둔 4가지 등으로 모두 13가지의 기존 특징 변수를 추출하며 이 변수들은 각각의 기법 이름에 해당하는 축약 명칭들을 특징 변수 이름으로 대응하여 사용한다.

2.1 차이 기반 기법

차이 기반 기법은 인접한 두 프레임(f_1 , f_2)사이의 비유사도 측정하여 기준 임계값을 초과할 때 서로 다른 장면으로 판정하는 방식이다[3,5]. 비유사도의 측정 단위에 따라 픽셀, 블록, 히스토그램 기법으로 나누어진다. 이 기법은 구현이 간단하고 연산량이 적다는 특징을 가지고 있어 널리 사용된다. 비유사도 측정 함수 $D(f_1, f_2)$ 를 수식 (1)과 같이 정의한다.

$$D(f_1, f_2) = \frac{1}{S_f} \sum_i |F_{f_1}(i) - F_{f_2}(i)| \quad (1)$$

여기서, f_1 , f_2 는 인접한 프레임을 의미하며, S_f 는 정규화 계수(normalized coefficient)이고, i 는 픽셀, 블록, 히스토그램을 나타내는 변수이다. 그리고 $F(i)$ 는 각 픽셀, 블록, 그리고 히스토그램을 나타내는 함수이다.

히스토그램 기법은 각 프레임별로 전역(global) 혹은 국부(local) 영역에서 컬러(HSI)별 히스토그램을 획득하여 인접한 프레임간 히스토그램 분포 차들의 평균을 구하여 비유사성을 검사한다. 이 방식은 카메라 혹은 객체의 움직임에 민감하지 않으며, 장면 전환점을 찾는 방식 중 가장 뛰어난 성능을 발휘하지

만, 유사한 히스토그램 분포를 갖는 프레임 경우 동일한 장면으로 판정하는 단점을 가진다. 따라서 이러한 단점을 해결하기 위해 전체 프레임을 16개의 블록으로 나누어 국부 히스토그램 차이 기법을 이용하여 전역 히스토그램의 단점을 해결하였다. 따라서 이 기법에 의해 아래의 다섯 가지 특징 변수를 정의한다.

① *LGH*[luminance global histogram] :

전체 영상의 밝기 정보에 대한 히스토그램 변화량을 검출

② *HGH*[hue global histogram with 64bins] :

히스토그램 분포를 64bin(bin)으로 하는 전체 영상의 색상(hue)정보에 대한 히스토그램 변화량을 검출

③ *LLH*[luminance local histogram] :

국부 영상의 밝기 정보에 대한 히스토그램 변화량을 검출

④ *HGH16*[hue global histogram with 16 bins] :

히스토그램 분포를 16bin으로 하는 전체 영상의 색상정보에 대한 히스토그램 변화량을 검출

⑤ *SGH8*[saturation global histogram with 8 bins] :

히스토그램 분포를 8bin으로 하는 전체 영상의 채도(saturation)정보에 대한 히스토그램 변화량을 검출

픽셀 기법은 인접한 프레임에서 대응하는 픽셀간 차들의 평균을 구하여 비유사성을 검사한다. 이 방식은 뉴스와 같이 배경이 고정된 경우 뛰어난 성능을 발휘하고 히스토그램 기법의 유사한 히스토그램 분포의 문제점을 해결하는 방법이다. 따라서 이 기법에 의해 아래의 두 가지 특징 변수를 정의한다.

① *LPD*[luminance pixel difference] :

인접한 프레임간의 밝기 정보의 변화량을 검출

② *CPD*[RGB pixel difference] :

인접한 프레임간의 RGB 정보의 변화량을 검출
블록 기반 기법은 프레임을 동일한 블록 크기로 나누어 평균과 편차를 구하여 비유사성을 검사한다. 이 방식은 픽셀 기반과 유사한 방식으로 잡음과 플래시(flash)에 의한 영향을 적게 받는 특징을 가진다. 이 기법에 의해 아래의 두 가지 특징 변수를 정의한다.

① *LLR*[likelihood ratio] :

블록의 평균과 분산을 이용하여 변화량을 검출

② *LBM*[luminance block mean] :

각 블록별 평균의 차를 이용한 변화량을 검출

2.2 상관 기반 기법

상관 기반 기법은 인접한 프레임($f1, f2$)에서 칼라 히스토그램을 구한 후 프레임내의 칼라 주성분을 추출하여 유사도를 파악하는 방식이다[10,11]. 상관 측정 함수 $C(f1, f2)$ 는 수식 (2)와 같이 정의한다.

$$C(f1, f2) = \quad (2)$$

$$\frac{1}{C_f} \sum_{x,y,z} \min\{H_{f1}(x, y, z), H_{f2}(x, y, z)\}$$

여기서, $f1, f2$ 는 인접한 프레임을 의미하고, C_f 는 정규화 계수, x, y, z 는 임의의 칼라 성분을 나타낸다. 마지막으로 $H_{f1}()$ 과 $H_{f2}()$ 는 프레임 $f1$ 과 $f2$ 의 칼라 히스토그램이다.

칼라 히스토그램 상관은 전역과 국부에 따라 유사도를 구할 수 있다. 전역 칼라 상관 기법은 히스토그램기법과 동일한 단점을 가지고 있고, 이를 국부 칼라 히스토그램 상관(color histogram correlation) 기법을 통하여 해결하였다. 상관 기반 기법에 의해 아래의 4 가지 특징 변수를 정의한다.

① *LGC*[luminance global correlation] :

전체 영상의 히스토그램 분포를 128bin으로 하는 밝기 정보에 대한 상관도를 검사

② *HSGC*[hue saturation global correlation] :

전체 영상의 색상 히스토그램 분포를 16bin과 색채 히스토그램 분포를 8bin으로 하는 상관도 검사

③ *LISGC*[HSI global correlation] :

①과 ②의 히스토그램 분포를 이용하여 상관도 검사

④ *LLC*[luminance local correlation] :

전체 영상을 16블록으로 나눈 후 각 블록마다 히스토그램 분포를 128bin으로 하는 밝기 정보에 대한 상관도 검사

지금까지 언급한 기존의 장면 전환점 기법들은 인접한 프레임사이의 변화량을 찾게 되고, 임의로 정의된 임계값에 의해 유사도가 높은 경우 동일한 장면으로 판정한다. 따라서 장면 전환점 검출 기법은 장면 전환점을 결정하기 위한 임계값 설정이 중요한 의미를 가진다. 임계값은 비디오 시퀀스의 종류와 비디오 촬영 시 배경과 편집 방법, 그리고 다양한 카메라의 움직임 등 여러 가지의 이유로 고정된 값으로 결정되지 않는다. 또한 특징 변수별 검출 성능에 따라 임계값이 다르게 결정된다.

위에서 언급된 특징 변수들의 성능을 확인하기 위

해 두 편의 영화와 다큐멘터리를 이용한다. 영화로는 “미션 임파서블(the Mission Impossible)”과 “메트릭스(the Matrix)”를 선택하고, 다큐멘터리는 “낙동강은 있다”라는 비디오 공모전 작품을 선택하였다. 영화는 초당 10프레임마다 캡처하였다. 장면 전환점을 *sc* (shot change) 그리고 장면을 *nsc* (no-shot change)로 표기하며 전체 시퀀스는 표 1에 나타내었다.

본 논문에서 성능을 평가하기 위해 소환 비율 (*recall ratio*)과 정확 비율 (*precision ratio*) 지수를 사용한다.

$$recall = \frac{S_c}{S_c + S_M} \times 100[\%] \quad (3)$$

$$precision = \frac{S_c}{S_c + S_F} \times 100[\%] \quad (4)$$

여기서, S_c 는 정확한 장면 전환된 프레임의 수, S_M 은 장면 전환점을 미검출된(missed detection) 프레임의 수이고, S_F 는 장면 전환점으로 오검출된(false alarm) 프레임의 수이다.

표 1의 실험 데이터에 앞에서 언급한 13가지 특징 변수를 사용하여 장면 전환점을 검출하고, 각 임계값은 소환 비율과 정확 비율이 각각 85% 이상이 되도록 설정하였다. 그 결과 매트릭스는 *HSGC*, 미션 임파서블은 *LLC*, 그리고 다큐멘터리는 *LHSGC* 특징 변수가 우수하게 나타났다. 이 결과를 실험 데이터 중 매트릭스에 동시에 적용하여 그림 1에 나타내었다.

그림 1에서 *HSGC*와 *LHSGC*의 임계값은 0.15로 선정하고 *LLC*는 0.25로 선정하였다. 실제 장면 전환 지점은 **A**와 **B**이고, **C**, **D** 그리고 **E**는 오검출된 지점이다. 그림에서 볼 수 있듯이 *LLC*는 **A**와 **B**는 정확하게 검출하지만, **E**점은 오검출 된다. 그리고 *LHSGC*는 **C**와 **D**점을 오검출 되고, *HSGC*는 **B**점은 미검출 되고, **C**와 **D**는 오검출된다. 이와 같이 각 특징 변수들이 독립적으로 적용할 경우 오검출과 미검출이 발생하게 되지만, *LLC*, *LHSGC*와 *HSGC* 특징 변수를 동시에 적용하면, 미검출 지점은 없어지고, 오검출 지점은 제거된다.

이러한 예에서 알 수 있듯이 임의의 비디오 시퀀스에 여러 가지의 특징 변수를 사용하면 각 특징 변수의 검출 특성 때문에 상호 보완성이 발생한다. 별도로 비디오 종류에 적합한 특징 변수를 찾을 필요 없이 여러 가지의 특징 변수를 이용하므로 검출 성능

표 1. 실험 데이터

종류	전체 프레임수	sc 수	nsc 수
매트릭스	2167	48	2119
미션 임파서블	2429	30	2399
다큐멘터리	2716	37	2679

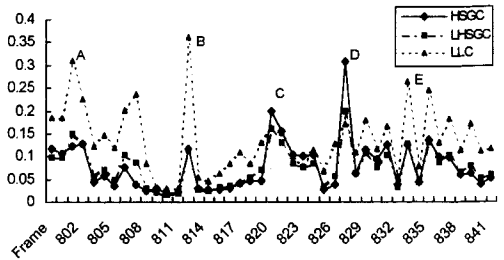


그림 1. 특징 변수에 따른 실험 결과.

을 높일 수 있다.

3. 강인한 특징 변수의 선별 기법

특징 변수(feature selection) 선별은 모집단을 구성하는 다양한 패턴을 구분하는 여러 가지 특징 변수 중 분류 능력이 뛰어난 특징 변수를 찾는 방식이다[6].

특징 변수의 선별하기 위해 기존의 특징 변수 선별 방식인 PCA 기법과 본 논문에서 제안한 CART 기법을 설명한다.

3.1 PCA 기법

PCA 기법은 다차원의 변수를 정보의 손실을 최소화하면서 낮은 차원의 변수로 선형 변환하는 방법이다[6,7]. 즉, 다차원인 변수들을 축소, 요약하여 차원의 단순화와 더불어 일반적으로 서로 상관되어 있는 변수들 상호간의 복잡한 구조를 단순화하는데 그 목적을 두고 있다.

전체 학습 샘플이 n 개, d 차원의 특징 변수 x_i 집합을 가지며, 특징 프레임 행렬 $X = [x_1, x_2, \dots, x_n]$ 는 각 특징 변수들을 포함한다. 이 행렬을 PCA 기법을 수행하는 것은 d 차원 특징 벡터를 갖는 각 프레임 i 를 k 차원의 정렬된 특징 공간($k \ll d$)상으로 투영하는 것이다. 따라서 여러 방면의 다른 방향으로 투영 (projection)하는 것을 의미하기 때문에 수식 (5)과 같이 임의의 새로운 단위 벡터(v)에 투영된 특징

벡터(x_i)로 나타낼 수 있다.

$$v(x_i) = v^T(x_i - \mu) \tag{5}$$

여기서, μ 는 수식 (6)과 같이 특징 벡터별 전체 샘플에 대한 평균을 정의한다.

$$\mu = \frac{1}{n} \sum_i x_i \tag{6}$$

각 특징 벡터 중 $v(x_i)$ 중 가장 큰 분산을 찾기 위해, 수식 (7)과 같이 각 특징에 대한 분산을 찾게 된다. 여기서, 전체 분산이 최대가 되기 위한 고유 벡터 (unit eigen vector)인 $var(v)$ 를 얻는다.

$$\begin{aligned} var(v) &= \frac{1}{n-1} \sum_i v(x_i)v(x_i)^T \\ &= v^T \Sigma v \end{aligned} \tag{7}$$

여기서, Σ 는 수식 (7)과 같다.

$$\Sigma = \frac{1}{n-1} \sum_i (x_i - \mu)(x_i - \mu)^T \tag{8}$$

수식 (8)에 얻은 결과는 단위 고유 벡터이므로 각 성분은 서로 직교(orthogonal)하며, 분산이 큰 순서대로 나열된다.

따라서 PCA 기법은 데이터 분포가 넓게 분포하는 변수를 찾게 되므로, 데이터 중 고립된 데이터에 많은 영향을 받는 문제점을 가지고 있다.

위의 방식에 따라, 표 1의 실험 데이터로 PCA 기법을 수행하였다. 그림 2는 각 실험데이터 시퀀스별 고유 벡터 차원과 주성분에 의한 성분의 분산 비율을 나타내었다. 그림 2에서 주성분 값이 1인 경우, 세 가지 시퀀스 분산은 약 70%, 80% 그리고 88%로 주성분 축에 의해 전체 데이터의 선형 매핑 정도를 알 수 있다. 본 논문에서는 주성분에 의해 전체 데이터 분포의 98%를 포함하는 변수들을 선택하였다.

그림 2에서 수직 점선은 전체 누적 분산의 98%지점을 나타낸 것으로 이 때의 특징 변수를 추출하였다. 따라서 전체 시퀀스에서 총 15개의 특징 변수를 얻게 되고 투표 기법(voting method)을 이용하여 상위 6개 변수만을 선택하였다. 각 선별된 변수로는 픽셀 기반 기법의 LPD, 히스토그램 기법의 LGH, HGH, 그리고 LLH 특징 변수가 선출되었고, 마지막으로 상관에 의한 LLC와 LHS GC 특징 변수가 선별되었다.

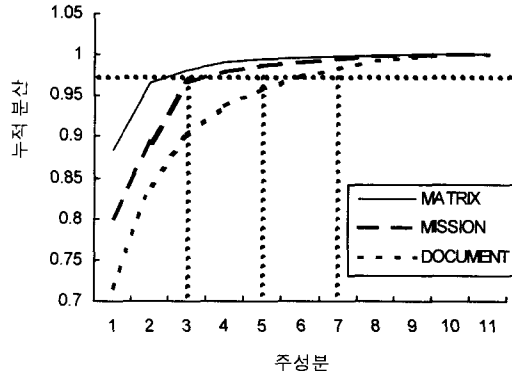


그림 2. 비디오 시퀀스별 주성분 대 누적 분산

3.2 CART 기법

CART는 대표적인 데이터 마이닝(data mining) 기법으로 이진 나무 구조(binary tree structure)를 이용하여 데이터 내의 숨겨진 패턴, 규칙, 관계 등의 정보를 탐색하는 방법이다[7]. 즉, CART는 학습 데이터를 다양한 속성들을 통하여 클래스들을 구분하는 분류 규칙뿐만 아니라 새로운 데이터에 대한 예측(prediction)하는 분류 기법이다[8].

CART에서 상위노드에서 하위노드로 분류 규칙은 다양성 혹은 불순물을 줄여나가는 방식이다. 즉, 상위 노드에서 10개의 클래스가 존재하는 것을 하위 노드들에서 불순물을 절반으로 감소시키는 특징 변수 혹은 속성을 선택하고 임계값을 결정하게 된다. 분류의 종료는 각 노드에 불순물이 존재하지 않을 때까지 계속 수행하지만 무한히 반복할 수 없으므로 종료 규칙(stopping rule)을 만족할 때까지만 수행한다[8].

전체 데이터를 분류할 때 유일한 한 가지의 변수만으로 최적의 분류 결과가 생성되는 경우는 극히 드물다. 그러므로 각 특징 변수들에 대한 이진 분류를 수행하였을 때 전체 불순물 함유율을 계산할 수 있다. 이것을 변수 중요도라고 하며 수식 (9)와 같다. 수식 (9)는 임의의 노드 t 에서 x_m 번째 변수로 최적 임계값($\overline{S_m}$)이 결정되었을 때, 변수 중요도($M(t)$)이다.

$$M(x_m) = \sum_{t \in T} \Delta I(\overline{S_m}, t) \tag{9}$$

여기서, $\overline{S_m}$ 는 대체 변수(surrogate variable)에 의한 분할점, t 는 임의의 노드이고, T 는 최적 노드 혹은

교차 타당에 의한 생성된 최종 분류 나무이다. 그리고 $\Delta K(t)$ 는 노드 t 의 분할점 \overline{S}_m 에 따른 불순물 정도를 나타내며 수식 (10)과 같다.

$$\Delta K(\overline{S}_m, t) = i(t) - i_L(t)p_L(t) - i_R(t)p_R(t) \quad (10)$$

여기서, $i(t)$ 는 노드 t 의 불순물 함유율이며, $i_L(t)$ 과 $p_L(t)$ 는 왼쪽 노드의 불순물 함유율과 클래스 분포확률이다. 그리고 $i_R(t)$ 과 $p_R(t)$ 는 오른쪽 노드의 불순물 함유율과 클래스 분포확률이다.

수식 (10)에서 임의의 노드 t 에서 하위 노드로 분할을 수행한 후 불순물 제거 정도를 측정하기 위한 것이다. 따라서 \overline{S}_m 에 의해서 불순물 제거정도를 누적하여 각 변수별로 중요도를 연산하게 된다.

예로 표 1의 비디오 시퀀스 중 매트릭스로 변수 중요도를 찾아보면, 첫 번째로 모든 변수를 통하여 데이터가 최종 노드에서 불순물이 영이 될 때까지 이진 분류 나무를 생성한다. 두 번째로 각 최종 노드의 소요 비용을 구한다. 여기서, 소요 비용은 데이터를 분류할 때 각 최종 노드에 함유된 불순물을 계산한 것으로 그림 3과 같다. 마지막으로 최소 비용이 소요되는 최종 노드에서 변수 중요도를 구하게 된다. 따라서 그림 3에서 최소 비용의 최종 노드는 2일 때이고, 수식 (9)에 의해 변수 중요도를 구한다. 그 결과는 표 2와 같다.

여기서, 그림 3에서 C. V. Rel. Cost (+)와 (-)는 교차 검증(cross-validation)을 수행하였을 때 발생하는 최대(+) 그리고 최소(-) 예측 비용을 의미하고, Resub. Rel. Cost는 학습 데이터에 대한 실제 비용을 의미한다.

CART에서 특징 변수의 추출은 그림 3과 표 2에서 얻었던 방식과 동일하게 표 1의 비디오 시퀀스에 대하여 변수 중요도를 구하였다. 각 시퀀스별 변수 중요도는 표 2와 동일하지 않기 때문에 각 시퀀스에 얻는 변수 중요도를 누적하여 상위 6개만을 선택하였다. 그 결과 블록 기법의 LLR, 히스토그램 기법의 LLH와 HGH, 그리고 상관 기법에 의한 LHSGC, HSGC와 LLC 등 이다.

PCA 기법과 CART 기법에 의한 특징 추출의 결과를 비교해 보면, PCA 기법은 학습 데이터 분포 상에서 주축인 변수를 선택하는 방식인 반면 CART의

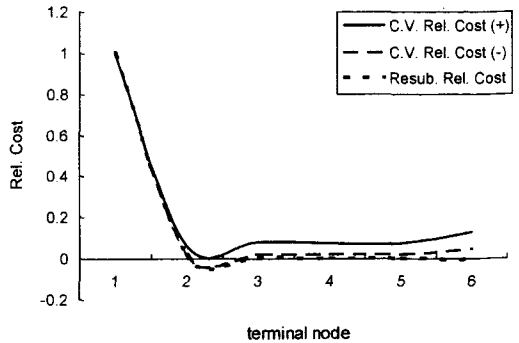


그림 3. matrix 시퀀스의 교차 검증 및 학습 관련 비용

표 2. 변수 중요도

변 수	HSGC	LLC	LHSGC	LLH	HGH	LGH
중요도	100	99.25	98.2	97.94	95.66	95.48

경우는 학습 데이터의 분류에 관련이 높은 변수만을 선택하므로 접근상의 차이점을 가지고 있다.

4. 역전파 신경망을 이용한 장면 전환점 검출

3장에서 추출된 다차원 특징 변수로 장면 전환점을 찾기 위해, 지도 학습 기법(supervisor learning method)의 대표적인 방식인 신경망을 이용한다.

신경망은 인간의 신경 구조를 모델링 한 것으로 단순한 기능을 하는 유닛이 복잡하게 연결된 구조로 되어 있다. 신경망의 구성은 셀(cells), 노드(nodes), 혹은 유닛(units)으로 불리는 많은 수의 뉴런으로 이루어지고, 각 유닛들은 특정 구조에 따라 가중치를 갖는 연결선으로 연결되어 있다. 그리고 각 유닛의 역할은 다른 여러 유닛들로부터의 입력을 연결선의 가중치를 고려하여 더한 후 그 결과를 적당한 활성화 함수(activation function)로 처리한 후 연결된 다른 유닛으로 출력한다[6]. 그림 4는 입력층, 은닉층, 그리고, 출력층으로 구성되는 다층 신경망의 구조이다.

그림 4에서 X는 입력층을 나타내고, Y는 출력층을 나타낸다. 그리고 Z는 은닉층으로 입력층에 따른 가중치 v 에 따른 결과를 출력층으로 전달하게 된다. 따라서 각 입력 데이터에 따라 가중치 값을 훈련하므로 최적의 분류 결과를 생성하는 방식이다.

본 논문에서는 지도 학습 방법 중 델타 학습 법칙의 일종인 오류 역 전파 신경망을 사용하였으며, 실

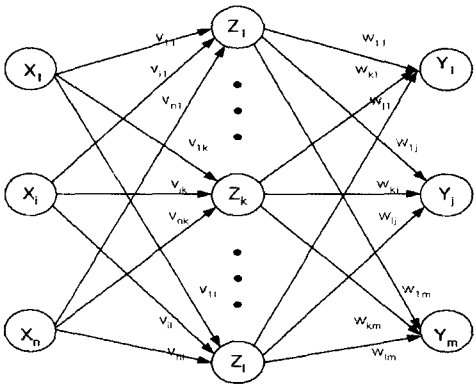


그림 4. 다층망 신경망의 구조

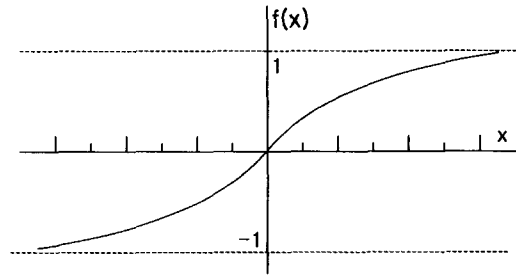


그림 5. tan sigmoid

제 문제에 있어서 가장 널리 사용되는 학습법이다. 델타 학습 법칙의 기본은 현재 주어진 연결 가중치로 생성되는 오차값을 구하여 이를 감소시키는 방향으로 연결 가중치의 값을 조정하는 것이다. 하지만 간단한 XOR 문제도 해결하지 못하는 단점을 가진다. 오류 역 전파 규칙은 이러한 문제를 해결하기 위한 방법의 일종으로 다층의 신경망을 학습시키는데 적합하다[6,7,9].

신경망을 역 전파 방식으로 학습시키는 것은 크게 다음의 세 단계로 나눌 수 있다. 첫 번째는 신경망의 순방향으로 입력층의 값을 가중치를 통해서 출력층의 값을 계산한다. 두 번째는 올바른 출력값과 신경망을 통해서 얻은 출력값의 에러를 계산하고, 오류 역 전파 알고리즘을 통해서 출력층에서 발생한 오류 값을 최하단의 층까지 전파시켜 각 층의 가중치들의 에러를 계산한다. 세 번째는 이렇게 얻어진 값을 가지고 각 층의 가중치의 값을 갱신한다. 초기 조건에 만족할 때까지 위의 세 단계를 반복 수행하게 된다.

그리고 은닉층은 하나의 신경세포가 낼 수 있는 출력은 한정되어 있으므로 특정 값을 넘어갈 수 없다. 여러 개의 입력층의 값이 은닉층으로 입력하게 되면 신경망 모델은 활성화함수의 임계값에 의해 출력 값을 얻게 된다. 본 논문에서는 아래와 같이 두 가지 활성화함수를 이용한다[6].

① tan-sigmoid 활성화 함수[13]

$$f(x) = \frac{1 - \exp(-\alpha x)}{1 + \exp(-\alpha x)}, \sigma > 0 \quad (11)$$

그림 5는 은닉층에 사용되는 활성화함수로 정규화(normalized)된 임계값을 갖는다.

② linear 활성화 함수

그림 6은 출력층의 활성화 함수로 출력의 임계값을 결정한다.

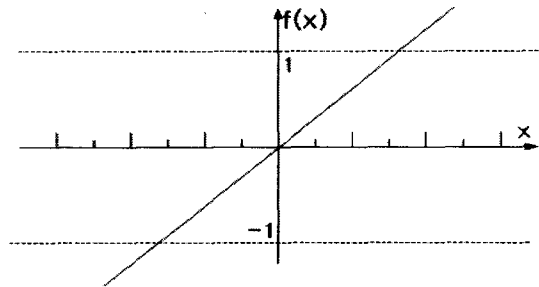


그림 6. linear 활성화 함수

5. 실험 및 검토

본 논문에서 제안된 특징 추출 기법을 통한 검출 성능을 검증하기 위하여 두 편의 영화와 한 편의 다큐멘터리를 선정하였다(표 1). 특징 변수로는 2장에서 설명한 13가지의 특징 변수인 LGH, HGH, LLH, HGH16, SGH8, LPD, CPD, LLR, LBM, LGC, HSGC, LHS GC, 그리고 LLC를 이용하였다. 특징 변수를 추출위한 소프트웨어는 Visual C++6.0으로 작성하였으며, 특징 추출을 위한 CART는 salford systems사의 CART[8]를 사용하였고, 이 결과 추출된 특징 변수는 모두 6개로 HGH, LLH, LLR, HSGC, LHS GC, 그리고 LLC 등이다. 그리고 PCA 기법은 R[12]을 이용하였으며, 선별된 특징 변수로는 LGH, HGH, LLH, LPD, CPD, LHS GC, 그리고 LLC 등으로 6개이다. 그리고 각 선별된 특징 변수의 성능을 검증하기 위한 역전파 신경망은 Matlab에서 구현하였다. 그리고 역전파 신경망으로 임계값을 선정하기 위해 가중치를 학습하게 된다. 그래서 “매트릭스”는

학습 데이터(learning data)로 선정하고, 나머지 데이터는 시험 데이터(test data)로 선정하였다.

신경망의 성능은 입력 변수, 은닉층의 뉴런과 계층 수 그리고 학습 데이터의 패턴 분포 및 시험 데이터의 패턴 분포에 따라 그 성능이 많이 좌우된다. 여기서 은닉층의 뉴런과 계층 수에 따라 훈련 시간이 많이 소요되는 단점을 가지고 있다. 따라서 다양한 실험 중 최고의 성능을 발휘하는 지점을 선별하여 표 3에 나타내었다.

표 3에서 역전파 신경망에 사용된 초기 변수를 나타내었다. 은닉층은 1개 층으로 선정하였고, 은닉층의 뉴런 개수는 13가지 특징 변수는 60개로 선정하였고, 특징 추출한 경우는 30개로 선정하였다. 은닉층의 활성화 함수는 -1에서 1사이로 한정하기 위해 tansigmoid 함수를 사용하고, 출력 활성화 함수는 linear 활성화 함수를 이용한다. 그리고 최종 목적 에러율은 10^{-3} 으로 선정하여 학습 데이터를 10^5 번 훈련하여 내부 가중치를 결정하였다.

표 4의 결과를 정리해 보면, 13가지 특징 변수를 사용하여 장면 전환점 검출 결과 소환 비율은 93.7%이고, 정확비율은 92.5%이다. 또한, 특징 추출을 이용한 결과를 비교해 볼 때, PCA의 기법에서 소환과 정확 비율은 각각 93.2%, 91%이고, CART는 각각 95.4%, 91.9%로 좋은 성능을 발휘하였다. 이런 결과를 통하여 CART 기법은 PCA 기법에 비하여 소환 비율과 정확 비율보다 좋은 성능을 발휘하였다. 그리고 전체 특징 변수를 사용하였을 때와 비교하여 소환 비율은 보다 뛰어난 성능을 발휘하였다. 이와 같은 결과로 다양한 시퀀스에서 우수한 성능을 발휘하는

표 3. 신경망의 파라미터

특징 변수 (개)	13	6
Number of Hidden layers	1	1
Hidden layer activation function	tansigmoid	tansigmoid
Output layer activation function	linear	linear
Number of hidden layer neurons	50	30
Learning rate	0.01	0.01
sum-squared error goal	0.001	0.001
number if epoches	100,000	100,000

표 4. 시험 데이터 결과

실험 영상	NN (13)		NN (PCA)		NN (CART)	
	소환	정확	소환	정확	소환	정확
매트릭스	87.3	88.9	82.8	87.3	88.9	87.3
미션 임파서블	93.8	93.8	96.8	90.9	100	90.9
다큐멘터리	100	94.9	100	94.9	97.4	97.4
평균	93.7	92.5	93.2	91.0	95.4	91.9

특징 변수를 선택한다면, 변수간의 상호 보완성을 이용하여 뛰어난 검출 성능을 가짐을 확인하였다.

본 논문에서 제안한 방식에 따른 오검출과 미검출 프레임을 살펴보면, 오검출의 경우는 갑작스러운 밝기 변화(섬광)와 잡음 등이며, 미검출의 경우는 유사한 히스토그램 분포 때문이다.

6. 결 론

본 논문에서는 장면 전환점 검출 기법의 특성에 따라 발생하는 검출 기법간의 상호 보완성을 이용하기 위해 여러 가지 장면 전환점 검출 기법을 동시에 적용하였다.

비디오 시퀀스에서 다양한 특징을 찾기 위해 차이 기반 방식과 상관 기반 방식을 이용하여 13가지의 특징 변수를 선정하였고, 특징 변수 중 강인한 특징 변수 추출을 위해 CART를 이용하여 6가지의 특징 변수를 선택하였다. 그리고 효과적인 임계값을 선정하기 위해 역전파 신경망을 이용하여 장면 전환점을 찾았다. 아울러 제안한 특징 변수 추출 방식을 기존 PCA 기법과 비교 검토하였다.

실험 결과 소환 비율과 정확 비율은 전체 특징 변수일 경우 평균 93.7%와 92.5%로 나타났고, PCA 기법은 평균 93.18%와 91%이고, CART 기법은 각각 평균 95.4%와 91.9%로 나타났다. 따라서 기존의 PCA 기법에 비교하여 CART 기법은 우수한 검출 성능을 가짐을 확인하였고, 전체 특징 변수와 비교하여 유사한 성능을 발휘함을 확인하였다. 향후의 비디오 요약과 색인 및 검색에 적용하므로 뛰어난 색인과 검색 방식을 제공할 것이다.

참 고 문 헌

[1] 김태희, 이흥희, 정동석, "MPEG 압축 영역에서

내용 곡선을 이용한 Video 요약 기법”, 한국통신학회논문지, 제27권 10호, pp.1021-1028, 2002.

[2] 이경화, 강대성, “PCA에 기반한 압축영역에서의 MPEG Video 검색 기법”, 전자공학회논문지, 제40권 SP편 제 1호, pp.28-33, 2003.

[3] Y. Yusoff, W. Christmas, and J. Kitter, “A Study on Automatic Shot Change Detection”, Proc. 3rd European Conference on Multimedia Applications, Services and Tech.(ECMAST), LNCS 1425, Springer, pp. 177-189, May 1998.

[4] Y. Yusoff, K. Kitter, and W. Christmas, “Combining Multiple Experts for Classifying Shot Changes in Video Sequences”, IEEE International Conference on Multimedia Computing and Systems, Vol. 2, pp. 700-704, 1999.

[5] G. Lupatini, C. Saraceno, and R. Leonardi, “Scene break detection : a comparison”, Proceedings of 8th International Workshop on Continuous-Media Databases and Application, pp. 34-41, 1998.

[6] D. A. Forsyth, and J. Ponce, Computer Vision A Modern Approach, Prentice Hall, New Jersey, 2003.

[7] R. D. Duda, P. E. Hart, and D. G. Stork, Pattern Classification, Wiley-Interscience, New York, pp. 394-411, 2001.

[8] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and Regression Trees, Chapman & Hall/CRC, New York, 1998

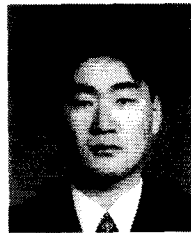
[9] R. K. M. Rao, K. R. Ramakrishnan, N. Balakrishnan, and S. H. Srinivasan, “Neural Net Based Scene Change Detection for Video Classification”, IEEE 3rd Workshop on Multimedia Signal Processing, pp. 247-252, 1999.

[10] T. Lin and H.J. Zhang, “Automatic Video Scene Extraction by Shot Grouping”, Proc. 15th International Conference on Pattern Recognition '00, pp39-42, 2000.

[11] Y. Zhuang, Y. Rui, T. S. Huang, and S. Mehrotra, “Adaptive Key Frame Extraction Using Unsupervised Clustering”, IEEE International Conference on Image Processing, Vol.1, pp866-870, 1998

[12] <http://www.r-project.org/>

[13] <http://www.mathworks.com/>



홍 승 범

1995년 2월 한국항공대학교 항공통신정보공학과(공학사)
 1997년 2월 한국항공대학교 항공통신정보공학과(공학석사)
 2003년 8월 한국항공대학교 항공통신정보공학과(공학박사)
 2001년 ~ 2003년 12월 이레스페이스 선임연구원
 2004년 3월 ~ 현재 한서대학교 항공전자시뮬레이션학과 전임강사
 관심분야 : 컴퓨터 비전 및 영상 처리, 항공기 시뮬레이터



홍 교 영

1990년 2월 한국항공대학교 항공전자공학과(공학사)
 1992년 2월 한국항공대학교 항공전자공학과(공학석사)
 1999년 8월 한국항공대학교 항공전자공학과(공학박사)
 1993년 ~ 2002년 8월 (주)대한항공 한국항공기술연구원 선임연구원
 2002년 8월 ~ 현재 한서대학교 항공전자시뮬레이션학과 조교수
 관심분야 : 항공전자, 항법, 항공기 시스템