

# 개선된 퍼지연상기억장치에 기반한 장면전환 검출 (Shot Transition Detection based on Improved Fuzzy Association Memory)

이 동 하 \* 고 일 주 \* 김 계 영 \*\* 최 형 일 \*\*\*  
(Dong Ha Lee) (Il Ju Ko) (Gye Young Kim) (Hyung Il Choi)

**요약** 학습과 추론을 위하여 유용한 방법으로 퍼지연상기억장치가 있다. 본 논문에서는 보다 효과적으로 추론결과를 유도하기 위하여 퍼지연상기억장치를 학습하는 단계에서 오류 역전파를 통하여 노드들 사이의 연결가중치를 재조정하는 방법과 퍼지규칙들을 간결화하는 방법을 제안한다. 제안된 방법은 비디오 데이터의 장면전환을 검출하는 분야에 적용하여 성능평가를 수행한다.

장면전환 검출을 위한 처리과정은 특징추출, 학습, 그리고 추론 등의 세 단계로 구성된다. 학습과 추론을 위하여 사용한 영상특징은 상관관계, 명도차, 비유사도이며, 퍼지소속함수는 훈련 데이터에 대한 히스토그램 분석을 통하여 생성된다. 성능평가를 위하여 다양한 종류의 비디오 데이터를 사용하였으며, 제안하는 방법과 인위적 규칙에 의한 방법 그리고 퍼지연상기억장치를 이용한 방법의 장단점을 분석한 결과를 제시한다.

**키워드** : 장면전환, 퍼지연상기억장치, 오류역전파 알고리즘

**Abstract** Fuzzy associative memory is very useful for machine learning and inference. To derive an inference result effectively, this paper proposes the method re-adjusting the connection weights between the nodes of fuzzy associative memory using error back propagation, and simplifying the fuzzy rules.

The proposed technique is applied to the field of shot transition detection. The shot transition detection system is consisted of three phases, feature extraction, learning, and inference. The used image features for learning and inference are correlation, intensity difference, and dissimilarity. And the fuzzy membership function is generated by analyzing the histogram of each feature. The proposed system is tested on 500 shots, and the experimental and comparative results are presented.

**Key words** : Shot Transition, Fuzzy Associative Memory, Back Propagation Algorithm

## 1. 서 론

생각을 전달하는 방법은 인간의 역사와 함께 발전되어, 최근에 이르러서는 정보통신 기술의 발전으로 전자매체를 통하여 멀티미디어 정보전달이 활발히 사용되고 있다. 특히, 비디오 자료는 영상과 음성 그리고 문자정보 등의 복합적인 정보를 표현하는 것으로 멀티미디어의 특징을 가장 잘 나타내고 있으며 현재까지의 개발된 정보전달

방법들 중에서 가장 유용한 것으로 고려되고 있다. 그러나, 비디오 자료는 자료 양이 방대할 뿐 아니라 형태 또한 매우 다양하여 효율적인 관리가 용이하지 않아, 보다 효과적으로 비디오 자료를 관리하고 검색할 수 방법인 비디오 분할, 색인, 검색에 대한 연구가 요구된다.

비디오 자료를 구성하는 단위는 프레임(frame), 장면(shot), 그리고 에피소드(episode)라 할 수 있다. 프레임은 비디오 자료를 구성하는 최소의 단위로 필름 한 장에 해당되며, 장면은 끊기지 않고 촬영된 일련의 프레임들로 시공간적으로 연속적인 행동을 보여주는 일련의 영상들이다. 마지막으로 에피소드는 여러 개의 장면이 결합되어 하나의 줄거리를 형성하는 의미적인 단위이다. 비디오 자료를 분할, 색인, 검색하는데 있어서 기본 단위는 장면이므로 분할을 위하여 장면의 경계를 탐지하는 방법은 필수적이다. 이 경계를 장면전환이라 하며,

\* 비 회 원 : 송실대학교 컴퓨터학부

winrive@vision.ssu.ac.kr

andy1026@vision.ssu.ac.kr

\*\* 중신회원 : 송실대학교 컴퓨터학부 교수

gykim@computing.soongsil.ac.kr

\*\*\* 중신회원 : 송실대학교 미디어학부 교수

hic@computing.soongsil.ac.kr

논문접수 : 2001년 10월 23일

심사완료 : 2002년 6월 18일

장면전환은 비디오의 편집으로 인하여 발생하는 것으로 급진적인 전환과 점진적인 전환으로 분류할 수 있다. 급진적인 전환인 컷(cut)은 두 장면 사이에 급격한 변화가 나타나는 것으로 뉴스나 스포츠 중계와 같은 비디오 자료에서 많이 나타난다. 점진적인 전환은 페이드 인(fade in), 페이드 아웃(fade out), 디졸브(dissolve)로 세분된다. 페이드 인은 고정영상(constant image)으로부터 점진적으로 장면이 선명해지는 전환을 의미하고, 페이드 아웃은 그 반대이다. 디졸브는 페이드 인과 페이드 아웃이 교차해서 발생하는 장면전환을 의미한다. 이들 점진적인 장면전환은 광고, 영화, 또는 뮤직비디오 등의 예술적인 비디오에서 많이 나타난다.

장면전환을 자동으로 탐지하는 방법에는 히스토그램에 근거한 방법[1][4][5], 에지를 이용하는 방법[3][7], 규칙에 기반한 방법[6][8] 등이 있다. 히스토그램에 근거한 방법은 인접한 두 영상 사이의 명암값 분포를 나타내는 히스토그램의 차이를 측정하여 장면전환을 검출한다. 이 방법은 급격한 전환을 검출하는데는 효과적이지만, 페이드(fade)나 디졸브(dissolve)와 같이 점진적인 전환은 검출할 수 없는 문제점이 있다[2]. 에지를 이용하는 방법은 현재영상과 이전영상 각각에 대하여 에지를 검출한 후, 현재영상에서 새로 나타나는 에지와 사라지는 에지의 변화량을 측정하여 장면을 검출한다[3][7]. 이 방법은 페이드나 디졸브와 같은 점진적인 전환은 히스토그램에 근거한 방법보다는 잘 검출하지만, 영상 특징에 따라 에지 추출 및 비교에 오류가 발생하므로 전체적인 성능이 나쁜 단점이 있다[8]. 보다 효율적으로 장면전환을 검출하기 위하여 참고문헌 [6]과 [8] 등에서는 여러 가지 특징값을 사용하는 방법이 제안하였다. 즉, 특징값들에 의하여 장면전환이 발생할 수 있는 모든 경우에 대하여 규칙을 생성한 후, 측정된 특징값이 정의된 규칙을 만족하는가를 판단하여 장면전환을 추론한다. 그러나, 이 방법은 규칙이 적용될 수 있는 범위 내에서만 장면전환을 검출할 있어 다양한 입력영상을 수용하지 못하는 문제가 있다[8]. 이러한 문제들을 해결하고 보다 정확히 장면전환을 검출하기 위하여 본 논문에서는 퍼지연상기억장치(Fuzzy Associative Memory, 이하 FAM이라 함)를 개선한 방법을 제안하고, 실험결과를 보인다. 본 논문에서 제안하는 방법인 개선된 퍼지연상기억장치(Improved Fuzzy Associative Memory, 이하 IFAM이라 함)를 이용한 방법은 특징값들과 장면전환 사이의 규칙을 생성 및 검증을 통하여 장면전환을 추론하는 방법이다. 그림 1은 본 논문이 제안하고 있는 장면전환 추론시스템의 전체적인 개요를 나타낸 것이다.

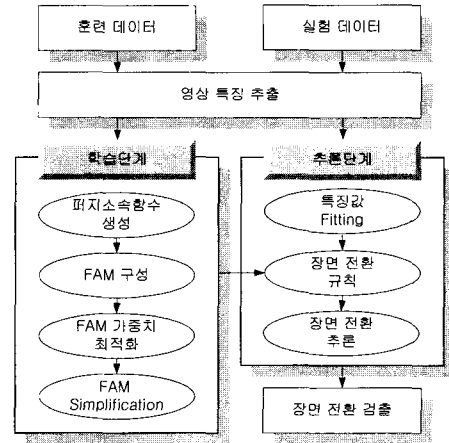


그림 1 시스템 전체 구조

본 논문의 구성은 다음과 같다. 제1절에서는 본 논문의 기본 방향에 관하여 설명하였고, 제2절에서는 본 논문에서 제안하는 IFAM에 관하여 설명한다. 제3절에서는 장면전환을 검출하기 위하여 사용하는 척도들과 이들을 입력되는 비디오 데이터로부터 추출하는 방법 그리고 IFAM의 입력으로 사용하기 위하여 퍼지화하는 방법이 대하여 기술한다. 제4절에서 실험결과를 보인 후, 마지막 제5절에서는 결론 및 향후연구에 관하여 논술한다.

## 2. IFAM의 학습과 추론

FAM은 연상기억장치(Associative Memory)와 퍼지이론(Fuzzy Theory)을 결합한 것으로 입력 퍼지집합과 출력 퍼지집합 사이의 관계를 정의 할 수 있는 것으로 기본적인 구조는 그림 2와 같다. 즉, FAM은 입력층(input layer), 퍼지화층(fuzzification layer), 조건부층(antecedent layer), 결론부층(consequent layer) 등 4개의 층과 연결강도를 나타내는 가중치로 구성된다.

그림 2와 같은 구조를 가지는 FAM을 사용하여 결론을 추론하기 위해서는 다음과 같은 두 가지 문제를 해결하여야 한다. 즉, 조건부층과 결론부층을 연결하는 가중치를 결정하는 문제와 입력자료를 퍼지자료로 변환하는 퍼지화 문제를 해결하여야 한다. 본 논문에서는 가중치를 결정하는 방법에 관하여 먼저 설명한 후, 입력자료를 퍼지자료로 변환하는 퍼지화에 관해서는 다음 절에서 기술한다.

가중치를 결정하는 대표적인 방법으로는 조건부와 결론부 모두가 활성화될 때, 이들을 연결하는 가중치를 증가시키는 Hebbian 학습을 통한 방법이 있다[9].

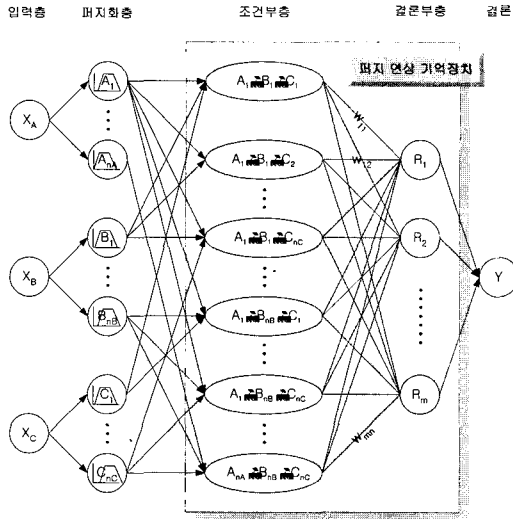


그림 2 FAM의 기본적인 구조

Hebbian 학습을 통하여 생성된 가중치는 지역최소를 포함하고 있어 추론된 결과에 오류를 내포할 수 있는 문제점이 있다[11]. 이 문제를 해결하기 위하여 본 논문에서는 오류 역전파(back propagation) 알고리즘을 사용하여 가중치를 재조정하는 방법을 제안한다. 제안하는 방법의 학습절차는 그림 3과 같이 Hebbian 학습을 통하여 초기 연결강도를 결정한 다음 오류 역전파 알고리즘을 이용하여 연결강도를 재조정하는 단계로 구성된다.

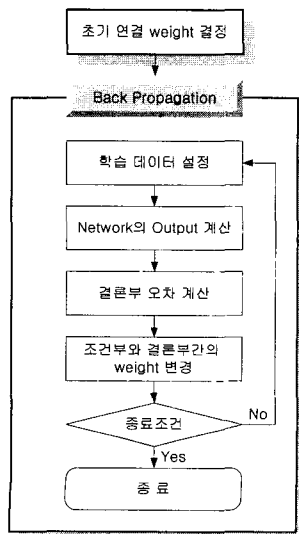


그림 3 IFAM에서의 연결강도학습

Hebbian 학습을 통하여 초기연결 가중치를 결정하는 과정은 식 1과 같다. 식 1에서  $y_j$ 가 가지는 값은 입력 패턴의 결론이  $j$ 번째이면 1이고, 그렇지 않으면 0이다. 초기연결강도를 결정하는 식 1은 모든 입력패턴에 대하여 적용된다.

$$w_{i,j}(t) = \frac{\lambda \cdot w_{i,j}(t-1) + \Delta w_{i,j}(t)}{\lambda + 1} \quad (1)$$

$$\Delta w_{ji} = \mu_{A_i}(x_i) \otimes y_j$$

where  $\lambda$  : 학습 회수  
 $\otimes$  : 퍼지 최소 연산

식 1만을 사용하여 결정된 가중치를 사용하여 입력된 패턴을 추론한 경우, 지역적 최소(local minima)로 인하여 추론결과가 부정확할 수 있다. 이를 해결하여 보다 정확한 추론결과를 유도하는 가중치를 얻기 위하여 본 논문에서는 오류역전파 알고리즘을 사용하여 FAM의 연결강도를 다음 같은 과정을 통하여 재조정한다.

FAM은 입력층의 각 노드에 입력패턴이 주어지면, 퍼지화층과 조건부층에 있는 노드들에 의하여 처리되어 최종적으로 출력층에서 추론결과를 출력한다. 이때, 기대값을 알 수 있으면 결론노드에 의하여 출력된 값과의 차를 알 수 있으므로, 본 논문에서는 이 차를 최소화하도록 연결강도를 재조정한다. 따라서, 본 논문에서 제안하는 방법은 감시학습(Supervised Learning)의 일종으로 학습과정에서 입력패턴과 추론결과를 함께 제공하여야 한다. 연결강도를 조정하기 위해 기본적으로 식 2와 같이 델타규칙을 사용한다. 즉, 식 2를 사용하여  $p$ 번째 제시된 입력/출력에 대하여 델타규칙에 따라 노드  $i$ 에서 노드  $j$ 로의 연결 강도를 조정한다.

$$w_{ji}^p = w_{ji}^p + \Delta w_{ji}^p \quad (2)$$

$$\Delta w_{ji}^p = \eta (t_j^p - o_j^p) i_i^p = \eta \delta_j^p i_i^p$$

식 2에서  $t_j^p$ 는  $p$ 번째 목표 출력 패턴의  $j$ 성분,  $o_j^p$ 는  $p$ 번째 입력패턴으로부터 FAM이 계산한 출력의  $j$ 성분,  $i_i^p$ 는  $p$ 번째 입력패턴의  $i$ 성분,  $\delta_j^p$ 는 목표출력과 실제출력의 오차,  $\Delta w_{ji}^p$ 는 입력층  $i$ 노드로부터 출력층  $j$ 노드에의 연결 강도의 변화량이다.

식 2를 사용하여 가중치를 조정할 경우, 해당 결론노드와 연결된 가중치는 모두 동일한  $\Delta w_{ji}^p$ 가 더하여진다. 이 경우 가중치 조정을 위한 처리가 수렴하지 않고 발산하는 경우가 발생할 수 있다. 이 문제점을 해결하기 위한 가중치를 가지는  $w_{ji}^p$ 에 대하여서는 상대적으로 더 적은 값을, 큰 가중치를 가지는  $w_{ji}^p$ 에 대해서는 보다 더 큰 값이 더하여야 하므로 본 논문에서는 시그모이드(sigmoid) 함수를 사용했다[10]. 비선형적으로 가중치를 변경하기 위하여 사용하는 시그모이드 함수는 그림 4와

같으며, 이 함수를 본 논문에서는 두 가지의 단계 즉, 결론부층의 해당 노드로 입력되는 가중치를 반영하기 위한 경우와 출력을 반영하기 위한 경우에 사용한다.

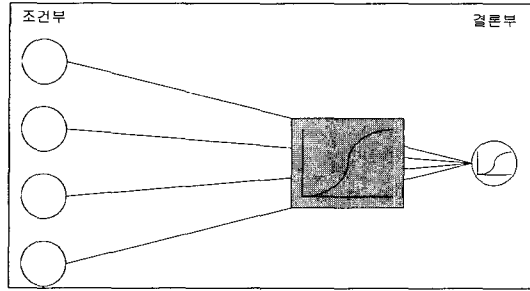


그림 4 가중치를 고려한 조건부와 결론부의 연결

본 논문에서 결론부층의 해당 노드로 입력되는 가중치를 반영하기 위한 시그모이드 함수는 식 3과 같다. 식 3에서 a와 c는 조건부층의 노드에서 일정한 값 이상의 가중치를 갖는 노드의 개수에 따라 달라진다.

$$g(w) = \frac{1}{1 + e^{-a(w-c)}} \quad (3)$$

식 3의 함수를 사용하여 반영율을 고려할 경우 해당 노드의 출력은 식 4와 같으며, 이 출력 또한 비선형성을 고려하면 식 5와 같다. 즉, 비선형성을 고려하여 결론부층의 어떤 노드가 추론하는 결과는 식 5와 같다.

$$R_j^p = \sum_{i=1}^m (\mu_{A_i}(x_i) \cdot w_{ji} \cdot g(w_{ji})) \quad (4)$$

$$O_j^p = f(R_j^p) = \frac{1}{1 + e^{-R_j^p}} \quad (5)$$

식 4와 식 5를 통하여 연결강도의 변화량은 식 6과 같이 정리할 수 있다. 식 6에서  $l^p$ 는 해당 결론노드인 경우는 1의 값을 그렇지 않는 경우는 0의 값을 가진다.

$$\frac{\partial O_j^p}{\partial R_j^p} = O_j^p(1 - O_j^p) \quad (6)$$

$$\delta_j^p = O_j^p(1 - O_j^p)(T_j^p - O_j^p)$$

$$\Delta w_{ji}^p = \eta \delta_j^p w_{ji}$$

따라서, 최종적으로 연결강도를 변경하는 방법은 식 7과 같다.

$$w_{ji}^p(t+1) = w_{ji}^p(t) + g(w_{ji}) \cdot \Delta w_{ji}^p \quad (7)$$

오류 역전파를 통하여 가중치를 조정하는 과정은 모든 훈련 데이터에 대해 반복적으로 수행한 후, 평가함수를 통하여 종료 또는 계속을 판단한다. 평가함수는 식 8과 같이 추론된 출력값과 목표 출력값 사이의 차를 제곱한 총합 E(t)의 변화도를 사전에 설정된 임계값  $\theta_E$

과 비교함에 의하여 이루어진다. 즉, 식 8에서 변화도인 E(t)-E(t-1)가  $\theta_E$  보다 적은 경우가 일정 수 이상 일 때 종료한다.

$$E(t-1) - E(t) < \theta_E \quad (8)$$

$$E(t) = \sum_{p=1}^n E^p(t)$$

$$E^p(t) = \frac{1}{2} \sum_{j=1}^m (T_j^p - O_j^p)^2$$

지금까지는 FAM을 구성하는 방법과 보다 정확한 추론을 위하여 조건부와 결론부를 연결하는 가중치행렬의 값을 조정하는 방법에 관하여 기술하였다. 이때, 가중치행렬의 원소들 각각은 하나의 퍼지규칙을 의미하므로 가중치 행렬의 크기가  $n \times m$  일 때, 생성되는 퍼지규칙의 수 역시  $n \times m$ 개이다. 생성된 퍼지규칙들 중에는 불필요하거나 중복된 규칙을 포함하고 있다. 추론에 소요되는 계산시간을 단축하고 보다 정확한 추론을 위하여 불필요하거나 중복된 규칙은 제거하여야 하는데 그 방법은 다음과 같다.

학습 과정을 거쳐 각각의 조건부는 연결 강도들을 가지고 모든 결론 클래스에 연결된다. 이때, i번째 조건부 노드  $A_i$ 에서 가장 높은 연결 강도를 갖는 간선의 강도를  $w^{\max}$ 라고 하면 이 조건부 노드가 결론에 대한 기여도는 식 9와 같다.

$$Peak_i = \frac{w^{\max}(A_i)}{\sum_{j=1}^m w_{ji}} \quad (9)$$

$$w^{\max}(A_i) = \max_{i \leq j \leq m} \{w_{ji}\}$$

Peak<sub>i</sub>가 낮고  $w^{\max}$ 가 낮은 조건부 노드는 결론을 유도하는데 공헌도 또한 낮음을 의미한다. 따라서, 식 10과 같이 임계값을 사용하여 해당하는 노드를 삭제함으로써 추론에 요소 되는 계산시간을 단축할 수 있다.

$$\text{Delete}(A_i), \text{ if } ((Peak_i < \theta_{Peak}) \&\& (w^{\max}(A_i) < \theta_{weight})) \quad (10)$$

식 10에서  $\theta_{Peak}$ 는 조건부의 기여도에 대한 임계치를 의미하며,  $\theta_{weight}$ 는 조건부가 가지는 최대연결강도에 대한 임계치를 나타낸다.

IFAM의 학습과 가지치기가 완료된 후에는 새로운 입력패턴에 대한 결론을 추론하는 그 과정은 식 11과 같다.

$$y^* = \max_{1 \leq j \leq m} \left\{ \sum_{i=1}^n \mu_{A_i}(x_i) \cdot w_{ji} \right\} \quad (11)$$

### 3. 장면전화 검출을 위한 특징추출과 퍼지화

본 논문에서 제안하는 IFAM을 사용하여 장면전환을 검출하기 위하여 입력영상으로부터 추출하는 특징은 상

상관관계(Correlation), 명도차(Difference of intensity) 그리고 비유사도(Dissimilarity) 이다. 이 특징들은 HIS(Hue, Intensity, Saturation) 칼라공간에서 추출되며 그 정의는 다음과 같다.

**상관관계**

상관관계는 인접하는 두 영상이 전체적으로 다른 정도를 나타내는 척도로 식 12와 같다. 식 12에서  $F_c$ 은 현재시점의 영상과 이전시점의 영상 사이의 색상과 명도의 상관관계로 구성되며, 값의 범위는 0에서 1사이이다. 즉,  $F_c$ 이 0에 근접할수록 두 영상의 차이가 큰 것이고, 반대로 1에 접근할수록 동일한 장면일 가능성이 높다.

$$F_c = \alpha \cdot H_c + \beta \cdot I_c \quad (12)$$

$$H_c = \frac{\sigma_{12}^H}{\sigma_1^H \cdot \sigma_2^H}, \quad I_c = \frac{\sigma_{12}^I}{\sigma_1^I \cdot \sigma_2^I}$$

where  $\begin{cases} H_c : \text{인접한 영상사이의 색상 상관관계} \\ I_c : \text{인접한 영상사이의 명도 상관관계} \\ \alpha, \beta : \text{가중치 } (0 < \alpha \leq \beta < 1, \alpha + \beta = 1) \\ \sigma : \text{표준편차} \end{cases}$

**명도차**

명도차는 페이드 장면전환의 특성을 나타내는 척도로 식 13과 같다. 식 14에서  $F_d$ 는 현재시점의 영상과 이전시점의 영상 사이의 명도차를 절대 명암값의 차로 나눈 것으로 정의되며, 값의 범위는 -1에서 1사이이다. 즉,  $F_d$ 가 -1에 근접할수록 페이드 아웃일 확률이 높으며, 반대로 1에 근접할수록 페이드 인일 확률이 높다.

$$F_d = \frac{1}{2} \times \left( \frac{I_{diff}}{I_{Adiff}} + 1 \right) \quad (13)$$

$$I_{diff} = \frac{\sum_{i=1}^M \sum_{j=1}^N (I(i, j, t) - I(i, j, t-1))}{I_{max} \cdot M \cdot N}$$

$$I_{Adiff} = \frac{\sum_{i=1}^M \sum_{j=1}^N |I(i, j, t) - I(i, j, t-1)|}{I_{max} \cdot M \cdot N}$$

where,  $I_{max}$  = 최대명도값,  $M$  = 영상의 세로축 크기  $N$  = 영상의 가로축 크기

**비유사도**

비유사도는 디졸브 장면전환의 특성을 나타내는 척도로 식 14와 같다. 식 14에서  $F_{Diff}$ 는 현재시점  $t$ 의 영상과 이전시점  $t-1$ 의 영상 사이의 색상의 비율과 채도의 비율을 합한 값이며, 값의 범위는 0에서 1이다.  $F_{Diff}$ 의 값이 1에 근접할수록 컷이나 디졸브일 확률이 높으며, 반대로 0에 근접할수록 디졸브일 확률이 낮다. 그리고  $\alpha, \beta$ 는 색상, 채도에 대한 가중치로 사용되는데 본 논문에서는 색상값에 대한 가중치보다 채도에 보다 높은 가중치를 부여한다.

$$F_c = \alpha \cdot \left| \frac{H_{diff}}{H_{Adiff}} \right| + \beta \cdot \left| \frac{S_{diff}}{S_{Adiff}} \right| \quad (14)$$

$$H_{Adiff} = \frac{\sum_{i=1}^M \sum_{j=1}^N |H(i, j, t) - H(i, j, t-1)|}{H_{max} \cdot M \cdot N}$$

$$H_{diff} = \frac{\sum_{i=1}^M \sum_{j=1}^N (H(i, j, t) - H(i, j, t-1))}{H_{max} \cdot M \cdot N}$$

$$S_{Adiff} = \frac{\sum_{i=1}^M \sum_{j=1}^N |S(i, j, t) - S(i, j, t-1)|}{S_{max} \cdot M \cdot N}$$

$$S_{diff} = \frac{\sum_{i=1}^M \sum_{j=1}^N (S(i, j, t) - S(i, j, t-1))}{S_{max} \cdot M \cdot N}$$

where  $\begin{cases} H_{max} = \text{hue level}, S_{max} = \text{saturation level} \\ 0 < \alpha, \beta < 1, \alpha + \beta = 1 \\ 0 \leq F_{diff} \leq 1 \end{cases}$

**퍼지화**

IFAM의 학습과 추론을 위하여 상기된 영상특징들은 추출된 다음 퍼지화되어야 하는데, 이를 위하여 소속함수를 정의하여야 한다. 특징값의 분포를 반영한 퍼지소속함수를 생성하기 위해서 본 논문에서는 히스토그램 분석을 통하여 소속함수를 생성한다[9]. 즉, 히스토그램 분석을 통하여 사다리꼴 형태의 퍼지소속함수를 생성한다. 소속함수를 생성하기 위해 먼저 학습자료로부터 각 특징들이 가지는 값을 추출하여, 특징별 히스토그램을 생성한 다음, 히스토그램 평활화(smoothing)를 수행하여 크기가 작은 지역 최대값(local maxima)과 지역 최소값(local minima)을 삭제한다. 평활화된 히스토그램으로부터 사다리꼴 형태의 퍼지소속함수를 생성하는 과정은 식 15과 같다.

$$\theta_{height} = \max\{\mu_T(c - \sigma_T), \mu_T(c + \sigma_T)\} \quad (15)$$

- [ Step 1 ]  $cl = cr = c$
- [ Step 2 ] while  $(\mu_T(cl) > \theta_{height}), cl--$
- [ Step 3 ] while  $(\mu_T(cr) > \theta_{height}), cr++$

식 15에서  $\mu_T$ 는 퍼지소속함수의 퍼지화 값을 의미하며, 사다리꼴의 중심  $c$ 로부터 좌우로 표준편차  $\sigma_T$ 만큼 떨어진 곳의 삼각 소속함수의 값 중에서 높은 값으로  $\theta_{height}$ 를 구한다. 알고리즘 순서에 따라 중심값  $c$ 로 부

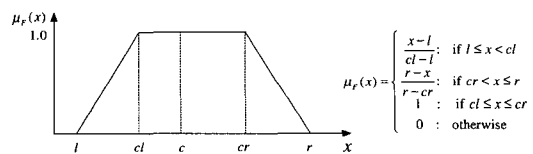
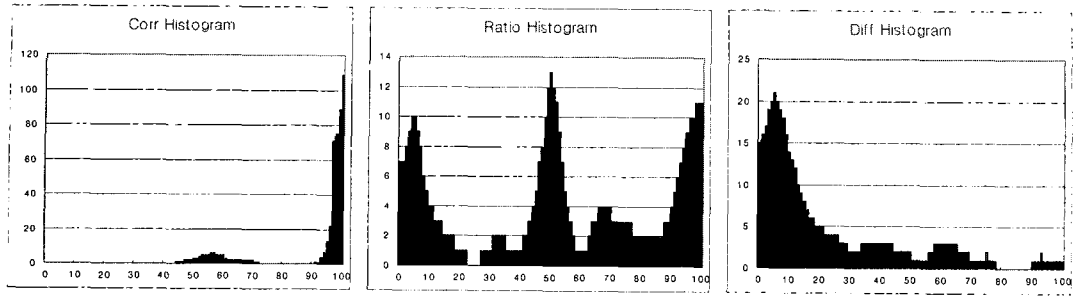


그림 5 사다리꼴 퍼지 소속 함수



(a) 상관관계에 대한 히스토그램 (b) 명도차에 대한 히스토그램 (c) 비유사도에 대한 히스토그램

그림 6 평활화된 히스토그램

터  $\theta_{height}$  보다 큰 조건을 만족하는 가장 좌측값(left most value)과 가장 우측값(right most value)으로부터  $cl$ 과  $cr$ 을 결정한다. 본 논문에서 사용하는 퍼지소속함수를 그림으로 도시하면 그림 5와 같다.

4. 실험 및 결과

이 절에서는 본 논문에서 제안한 IFAM을 사용하여 장면전환추론 방법을 다양한 비디오 데이터에 적용하여 실험한 결과에 대해 기술한다. 실험을 위하여 사용한 컴퓨터는 800MHz의 속도를 갖는 Pentium III PC이고, 구현을 위하여 사용한 프로그래밍 언어는 Microsoft의 Visual C++이다. 실험을 위하여 스테이(stay), 컷, 페이드 인, 페이드 아웃, 디졸브의 특성을 가지는 비디오 데이터 각각에 대하여 100개씩 총 500개의 훈련데이터를 사용했다. 이 실험자료는 뉴스, 드라마, 뮤직비디오, 광고 등에 사용된 다양한 비디오 데이터로 부터 취득하였다.

IFAM을 구성하기 위한 첫 번째 단계는 퍼지소속함수를 생성하는 것이다. 그림 6에서는 퍼지소속함수생성을 위하여 각 특징값의 분포를 평활화한 히스토그램을 보여주고 있다.

평활화된 히스토그램으로부터 생성된 퍼지소속함수는 그림 7과 같다. 그림 7에서 보는 바와 같이 히스토그램 분석을 통하여 얻어진 퍼지소속함수는 상관관계에 대해서는 2개, 명도차와 비유사도에 대해서는 각각 5개가 생성되었다. 따라서 이 퍼지소속함수로 부터 생성된 퍼지연상기억장치는  $2 \times 5 \times 5$ 개의 조건부층을 가진다.

그림 8에서는 오류 역전파를 반복함에 따라 IFAM이 가지는 오차제곱의 총합과 규칙수의 변화를 보여주고 있다. 즉, 오류 역전파를 통하여 가중치를 재조정함에 따라 조건부층의 노드로부터 결론부층의 노드에 이르는 연결강도가 점차 안정화되어 오차가 점차적으로 감소하는 것을 알 수 있다. 생성된 규칙의 수는 일정한 수에 도달할 때까지 급격히 감소하다가 변화가 없음을 알 수 있다.

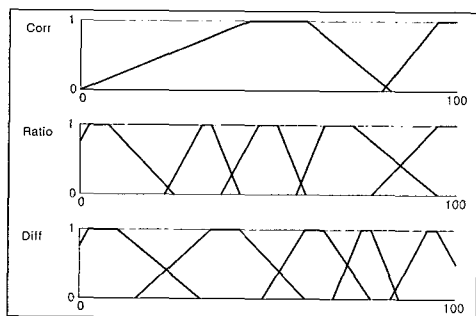
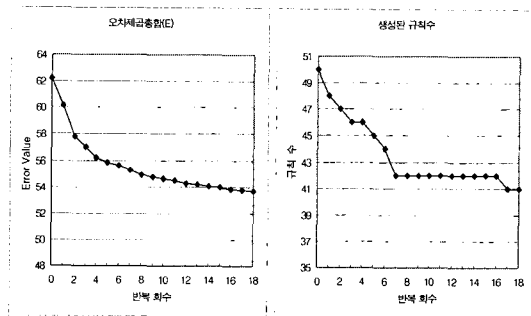


그림 7 생성된 퍼지소속함수



(a) 오차 제곱 총합의 변화 (b) 규칙수의 변화

그림 8 오류 역전파 알고리즘 수행에 따른 변환

본 논문에서 제안하는 방법의 성능을 평가하기 위하여 인위적 규칙[8]에 의한 방법과 Hebbian 학습에 의한 방법과 비교 분석하였다. 표 1은 상기의 세 가지 방법 각각에 대한 장면전환검출의 정확도를 나타낸 것이다. 표 1에서  $N_c$ 는 정확하게 검출한 장면 전환의 개수

표 1 장면 전환 검출의 정확도

Method	Cut			Fade in			Fade out			Dissolve		
	N <sub>c</sub>	N <sub>m</sub>	N <sub>f</sub>	N <sub>c</sub>	N <sub>m</sub>	N <sub>f</sub>	N <sub>c</sub>	N <sub>m</sub>	N <sub>f</sub>	N <sub>c</sub>	N <sub>m</sub>	N <sub>f</sub>
인위적인 규칙	31	0	12	7	1	1	8	0	0	5	3	7
Hebbian	30	1	7	8	0	1	8	0	0	7	1	1
제안한 방법	30	1	7	8	0	0	8	0	0	8	0	0

를, N<sub>f</sub>는 잘못 검출한 장면 전환의 개수를, N<sub>m</sub>은 검출하지 못한 장면 전환의 개수를 각각 의미한다.

표 1에 의하여 세 방법 모두가 잘못 검출한 컷의 개수가 많은데, 이것은 영상의 명도가 0에 가까운 두 인접 영상간의 상관관계를 계산함에 있어서, 영상간의 작은 차이가 정수화 될 때 커다란 비율로 확대되어 계산오차를 크게 하기 때문이다. 인위적인 규칙을 이용한 방법은 동작물체의 움직임이 있거나 카메라의 동작이 있는 스테이 구간을 디졸브로 오검출하는 경우가 많았다. 페이드 인과 페이드 아웃의 경우는 세 방법 모두 높은 정확도를 보여 주었고, 디졸브인 경우는 인위적인 규칙을 이용한 추론 방법은 디졸브 구간을 스테이로 오검출하는 경우가 많았다.

전체적인 성능평가를 위하여 본 논문에서는 식 16의  $R_{precision}$  과  $R_{recall}$ 의 두 가지 척도를 사용했다[7].  $R_{precision}$ 은 영상에서 검출된 장면전환 중에서 정확하게 검출된 장면전환의 비율을 의미하고,  $R_{recall}$ 은 영상에 존재하는 장면전환 중에서 정확하게 검출된 장면전환의 비율을 의미한다. 표 2에서는 식 16을 이용하여 각 장면전환에 성능을 평가한 결과를 보인다.

$$R_{precision} = \frac{N_c}{N_c + N_f} \quad (16)$$

$$R_{recall} = \frac{N_c}{N_c + N_m}$$

표 2 성능평가

(a)  $R_{precision}$

Method	Cut	Fade in	Fade out	Dissolve
인위적인 규칙	0.720	0.875	1.000	0.416
Hebbian	0.810	0.889	1.000	0.857
제안한 방법	0.810	1.000	1.000	1.000

(b)  $R_{recall}$

Method	Cut	Fade in	Fade out	Dissolve
인위적인 규칙	1.000	0.875	1.000	0.625
Hebbian	0.968	1.000	1.000	0.875
제안한 방법	0.968	1.000	1.000	1.000

본 논문에서 제안하는 방법은 오류 역전과 알고리즘을 통해 규칙을 학습하는 단계는 많은 시간을 필요로 하지만, 생성된 규칙을 이용하여 장면전환을 검출하는 과정은 다른 두 가지 방법과 차이가 없다. 표 2를 통해 본 논문에서 제안하는 방법이 장면전환을 검출하는 성능은 가장 우수함을 알 수 있다.

### 5. 결론 및 향후연구

본 논문에서는 FAM의 성능을 향상시키기 위하여 학습과정에서 오류 역전과를 통하여 가중치를 재조정과 노드 삭제를 통하여 추론에 소요되는 계산시간을 단축할 수 있는 IFAM을 제안하고, 이를 비디오 데이터에서 장면전환을 검출하는 분야에 적용하는 방법에 대하여 설명한 후, 실험결과를 제시한 다음 기존의 방법과 비교 분석하였다. 그 결과 기존의 주요 방법인 인위적인 규칙을 이용한 추론방법과 Hebbian 학습을 통하여 생성된 규칙을 이용한 추론 방법 보다 본 논문에서 제안한 방법의 정확도면에서 가장 우수한 성능을 가는 것을 알 수 있었다. 향후연구로는 퍼지화층까지 오류 역전과를 통하여 가중치를 재조정하는 방법이라 고려된다.

### 참고 문헌

- [1] A. Nagasaka and Y. Tanaka, "Automatic Video Indexing and Full-Video Search for Object Appearance," the Second Working Conference on Visual Database Systems, pp. 119-133, 1991.
- [2] HongJiang Zang, Atreyi Kankanhalli and Stephen W. Smoliar, "Automatic Partitioning of Full-Motion Video," Multimedia Systems, Vol. 1, pp. 10-28, 1993.
- [3] Ramin Zabih, Justin Miller and Kevin Mai, "A Feature-Based Algorithm for Detecting and Classifying Scene Breaks," Proc. of ACM Multimedia, pp. 189-200, 1995.
- [4] S. J. Dennis, R. Kasturi, U. Gargi and S. Antani, "An Evaluation of Color Histogram Based Methods in Video Indexing," Research Progress Report CSE-96-053 for the contract MDA 904-95-C 2263, 1995.

- [5] Yi Wu and David Suter, "A Comparison of Methods for Scene Change Detection in Noisy Image Sequence," First International Conference on Visual Information Systems, pp. 459-468, 1996.
- [6] 문철호, "복합 규칙을 이용한 장면전환 검출", 숭실대학교 석사학위청구논문, 1998.
- [7] Rainer Lienhart, "Comparison of Automatic Shot Boundary Detection Algorithms," Conf. of Storage and Retrieval for Image and Video Databases VII, pp 290-301, 1999
- [8] 장석우, "카메라의 동작을 보정한 장면전환 검출", 숭실대학교 박사학위청구논문, 2000
- [9] 장대식, 최형일, "퍼지연상기억장치에 기반한 퍼지추론 시스템", 한국정보과학회논문지(B), pp. 1693-1705, 1995
- [10] Earl Gose, Richard JohnsonBaugh and Steve Jost, "Pattern Recognition and Image Analysis," Prentice Hall Press, 1996
- [11] Yurichiro Anazai, "Pattern Recognition and Machine Learning," pp. 325-327, Academic Press, Inc., 1992.



이 동 하

1999년 8월 숭실대학교 컴퓨터학부 졸업(공학사). 2001년 8월 숭실대학교 대학원 컴퓨터학과 졸업(공학석사). 관심분야는 퍼지시스템, 신경망, 영상 처리/분석 등



고 일 주

1992년 2월 숭실대학교 전자계산학과 졸업(공학사). 1994년 2월 숭실대학교 대학원 컴퓨터학과 졸업(공학석사). 1997년 8월 숭실대학교 대학원 컴퓨터학과 졸업(공학박사). 관심분야는 퍼지시스템, 시각 인식, 패턴 인식, 감성 에이전트 등

김 계 영

정보과학회논문지 : 소프트웨어 및 응용  
제 29 권 제 5 호 참조

최 형 일

정보과학회논문지 : 소프트웨어 및 응용  
제 29 권 제 5 호 참조