

□ 특 집 □

패턴인식의 개관

한국과학기술원 오 영 환*

● 목

차 ●

- I. 서 론
- II. 패턴인식의 개념
- III. 패턴인식의 문제점
- IV. 패턴인식 시스템

- 4.1 패턴 전처리 시스템
- 4.2 특징 추출
- 4.3 식별
- V 결 론

I. 서 론

정보화 시대에 살고 있는 우리들은, 방대한 정보의 범람 속에서 외부로부터 끊임 없이 정보를 받아들이고 이를 처리하여 대처해 가면서 생활을 영위한다. 외부로부터의 정보는 인간의 5각인 시각, 청각, 촉각, 미각, 후각이라 불리우는 감각기관들을 통해 받아들여지며, 이는 대뇌로 전달되어 처리된다. 이와 같이 외부 정보를 두뇌에서 처리하는 과정을 패턴인식(pattern recognition)이라 한다. 패턴은 사람이 인식하려는 정보를 물리적으로 표현한 대상으로 정의 할 수 있으며, 인식(recognition)은 대상물의 존재를 아는 인지(cognition)의 과정에서 한걸음 더 나아가 구체적인 의미나 정보내용을 판정하는 과정을 의미한다. 그러나, 보통 패턴인식이라 말할 때 컴퓨터를 중심으로 한 기계에 의한 자동인식을 가리키며, 대상으로 하는 감각기관도 시각과 청각을 지칭하는 경우가 일반적이다.

대상 연구분야는 문자인식, 음성인식, 화상인식으로 크게 나눌 수 있으며, 응용분야는 매우 다양하다. 문자인식의 경우도 비교적 인식 알고리즘이 간단한 바코드(bar code), OMR, 자기인크문자인식(MICR), 광학문자인식(OCR) 등과 같

은 특수문자인식을 비롯하여, 인쇄 또는 필기된 문자를 인식하는 off-line 문자인식, 펜컴퓨터로 대표되는 on-line 문자인식 등, 제약조건 등에 따라 구현의 난이도도 크게 다르다. 음성인식은 대상음성의 종류(단어나 절을 떼어 발생시키는 구분발성음성, 연속발성한 연속음성), 어휘수(수십 단어, 수십~수백단어, 수천단어, 무제한 어휘), 화자수(특정화자, 학습 가능한 다수화자, 불특정 다수화자), 환경(방음실, 일반사무실, 소음이 심한 현장 등) 등의 기술적요인의 조합에 따라 구현 가능성이 크게 좌우 된다. 한편, 화상인식의 인식대상은 크게 2차원 화상과 3차원 화상으로 나누어 볼 수 있으며, 보통 화상인식이라 부를 때는 전자를 말하고, 후자의 경우는 computer vision으로 호칭하는 것이 일반적이다. 2차원 화상으로는 문자를 제외한 평면화상이 모두 인식 대상이 될 수 있으며, 도면, 현미경, X-ray 화상, remote sensing data, 지문인식 등의 광범위한 응용분야를 가지고 있다. 3차원 화상은 입체정보를 필요로 하는 물체의 형상인식에 연구의 초점이 맞추어지며 응용분야로는 robot vision, 얼굴인식 등을 들 수 있으며 보통 인공지능분야로 분류한다.

이하, 일반적인 패턴인식 시스템에서 공통적인 기술적 요인에 대해 간단히 기술한다. II장과

*중신회원

III장에서는 패턴인식의 일반적 개념을, IV장에서는 패턴인식 시스템의 처리과정을 설명한다.

II. 패턴인식의 개념

패턴이란 사람이 인지할 수 있는 물리적으로 표현된 정보의 외형을 말하며, 인식이란 패턴을 구체적인 부류로 확정하는 일을 말한다. 물체를 보고, 단지 대상물의 존재를 아는 과정은 인지라 한다. 그러므로 우리들이 패턴인식이라 부르는 경우, 단순히 패턴의 성질이나 특징을 정량적으로 관측, 측정하는 것 뿐만 아니라, 패턴과는 차원이 다른 의미나 정보의 내용을 한정하는 것을 말한다. 외형으로서의 패턴에 내용으로서의 개념을 대응시키는 작업이 패턴인식이다.

패턴인식의 정의는 사람에 따라 조금씩 다를 수 있다. 가장 넓게 해석한다면, 외계의 정보 모델을 자기 안에 형성하는 과정이라고 할 수 있다. 그러나 이러한 넓은 의미의 패턴인식과정을, 현재의 기술수준으로 완전히 실현할 수 있는 단계에는 와 있지 못하다. 공학적으로는 외계의 관측과 해석에 의해 의미 자기내부에 형성된 모델에 따라, 주어진 패턴에 분류표시를 하여 그것이 속하는 카테고리의 명칭을 출력하는 과정이라고 할 수 있다.

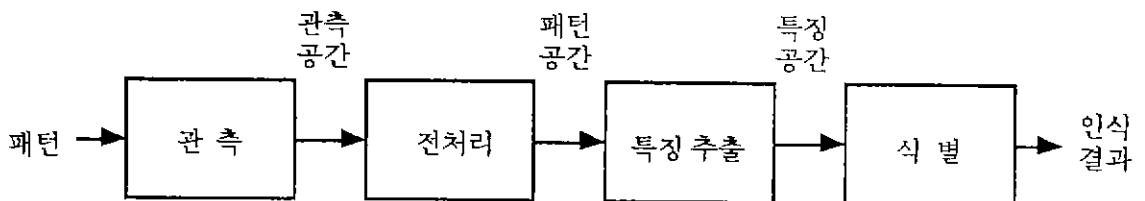
패턴에는 우리가 인식대상으로 하는 정보를 표현하여 전달하기 위하여 인간이 의도적으로 발생시킨 것과, 생물이나 자연현상, 사회현상 등과 같이 자연히 나타나는 경우가 있으며, 후자는 다시, 목적하는 정보를 얻기 위하여 의도적으로 관측하는 경우와 자연스럽게 관측되는 경우로 나눌 수 있다. 음성이나 문자는 언어적인 정보를 전달하기 위해서 의도적으로 발생시킨 대표적인 패턴이며, X-선 사진이나 멀티 스펙트럼 사진은

의도적으로 관측한 대표적인 패턴이다. 그러나, 풍경이나 구름 모양 등은 자연히 관측되는 패턴의 좋은 예이다.

패턴인식에 관한 연구는, 인간이나 동물의 패턴인식 능력 자체를 탐구하는 분야와, 특정한 응용을 위해 주어진 인식 작업이 가능한 인식기기를 설계하는 이론 및 기술개발분야의 두가지로 나눌 수 있다. 전자의 경우는 심리학, 생리학, 생물학 등의 지식을 많이 이용하는데 반해, 후자의 경우는 공학, 컴퓨터, 정보과학에서 다루는 분야이다. 본 고에서 다루는 분야는 인식기기의 설계이론 및 기술개발인 후자이며, 주로 공학적 접근에 의한 한정 인식기기의 실용화에 관한 기술이다.

일반적인 패턴인식 시스템은 (그림 1과) 같이 구성된다. 먼저, 카메라, 마이크 등을 이용하여 관측된 자료는 다소의 불필요한 정보를 가지고 있으므로, 컴퓨터에서 인식을 위해 필요한 자료만을 얻어내는 전처리를 하게 된다. 다음은 입력된 패턴을 인식하기 용이한 작은 단위로 분할하는 과정을 거친다. 관측된 패턴은 여러가지 다른 환경하에서 관측되므로, 같은 패턴이라도 스케일이 크게 다르게 되는데, 인식을 위해서 크기, 길이 등의 정규화를 하게 된다. 특징추출 단계에서는 인식에 중요한 실마리가 되는 특징을 얻어내고, 이 특징을 비교하여 패턴의 식별을 하게 된다.

(그림 1)내에서의 각 단계는 절대적인 것이 아니며, 전처리(preprocessing), 분할(segmentation), 정규화(normalization), 특징추출(feature extraction) 등의 순서를 일부 바꾼다든지, 일부를 생략할 수 있다. 특히, 분할은 어려운 문제를 많이 포함하기 때문에, 패턴을 완전히 인식하고 나서야 비로소 분할이 가능한 경우도 많다.



(그림 1) 패턴 인식 시스템의 기본 구성

III. 패턴인식의 문제점

본 장에서는 패턴인식의 문제점에 대하여 살펴본다. 일반적으로 패턴인식의 문제는 질적인 문제와 양적인 문제로 나눌 수 있다.

질적인 문제는 패턴인식의 본질적인 문제로, 해(solution)는 알고 있으나 해법이 불명인 경우를 말한다. 인간의 패턴인식기능에 의해 해의 존재는 알 수 있으나, 인간의 인식과정은 알기 어렵다. 인간은 뛰어난 패턴인식기능을 가지고 있으나, 그 과정은 무의식적이면서 직관(intuition)과 경험에 바탕을 두고 있는데 반해, 이 과정을 객관화하고 정량화시켜 알고리즘의 형태로 정식화하는 일반적인 방법이 아직 완성되어 있지 않기 때문이다. 따라서, 공학적 관점에서는 부분적으로 체계적인 접근을 하여, 일부기능의 실용화를 목표로 하고 있다.

이러한 인간의 패턴인식기능 형성과정에 대해서는 경험주의적인 관점과 구조주의적인 관점이 있다. 경험주의적인 관점은 인식기능을 개개인의 학습과 경험에 의해, 인류의 발생과 발달의 역사속에서 얻어지는 능력으로 본다. 이에 반하여 구조주의적인 관점은 인식능력을 인류발전의 유전적인 배경하에서 발전하여, 현존하는 사람이 선천적으로 지니게 되는 특수한 구조에 의해서 가능한 능력으로 본다.

이러한 질적인 어려움은 인식대상의 문자나 음성, 기하학적 도형이나 물체와 같이 인간이 정보 전달을 위해 의도적으로 발생시킨 패턴의 경우에 특히 강하게 나타난다.

그런데, 어떠한 문제가 수학적으로 유한한 과정에 의해 명확히 기술되면 그 문제는 일단 질적으로 해결된 것으로 볼 수 있으나, 공학적 측면에서 바라보면 양적으로 방대하여 기억용량면이나 인식소요시간 면에서 실용상의 구현이 불가능한 경우가 많다. 원리적으로나 수학적으로 가능하다는 사실 이외에도 비용을 포함해서 공학적으로 허용될 수 있는 범위의 기억용량과 시간내에 가능하지 않으면 공학적으로는 실용상의 의미가 없게 된다.

이러한 양적인 문제점은 필기문자나 불특정

다수의 화자가 발생한 음성인식과 같이 다양한 변형 패턴을 대상으로 하는 경우나, 명암화상 또는 색채화상 패턴과 같은 정보량이 많은 패턴을 대상으로 하는 경우에 특히 강하게 나타난다. 그러나, 과거에는 양적인 문제점으로 인하여 실용 불가능 했던 기술도 하드웨어의 발달에 따라 가능해진 부분도 있다. 앞으로는 양적인 문제에 구애받지 않고 오직 질적인 문제가 중요한 관점이 되는 시기가 올 것으로 기대할 수도 있다.

물론 위와 같은 두 가지의 어려움은 명확히 분리할 수 있는 것이 아니며 현실적으로는 혼재하는 경우가 대부분이다.

IV. 패턴인식 시스템

본 장에서는 (그림 1)에 도시된 패턴인식 시스템을 전처리, 특징추출, 식별로 나누어 설명한다. 특히, 패턴인식과정의 핵심이라고 할 수 있는 식별단계에서는 최근 많이 사용되고 있는 방법들의 기본개념을 주로 언급한다.

4.1 패턴 전처리 시스템

(그림 1)에서 전처리, 분할 및 정규화를 넓은 의미로 패턴 전처리 부분이라 한다. 관측된 그대로의 자료는 많은 양의 불필요한 부분을 포함하고 있으므로, 인식을 위해서는 시스템에서 필요로 하는 자료만을 추출할 필요가 있다.

먼저, 관측의 과정에서 입력되는 잡음을 제거한다. 잡음을 넓게 정의하면 “인식에 쓸모 없는 성분요소”라 할 수 있으나 여기서는 좁은 의미로 “대상 또는 관측계에서 발행하는 불규칙성의 성분요소”를 잡음으로 생각하기로 한다. 잡음과 신호는 기본적으로는 둘 사이의 통계적 성질의 차를 이용하여 분리한다.

음성과 같은 시간파형의 경우는 잡음과 신호의 주파수 대역이 서로 다르면, 여파기를 사용하여 분리할 수 있다. 전원의 험(hum)잡음에 대한 리플 필터(ripple filter), 고주파 잡음에 대한 저역 통과 필터 등을 예로 들 수 있다. 문자도형이나 화상과 같은 2차원 도형에서는 공간 주파수의

차를 이용한 잡음 제거가 푸리에(Fourier) 변화 또는 2차원 중첩적분(convolution)선형 필터에 의해서 이루어 진다.

전처리 단계에서 제거해야 하는 요소 중에는 잡음 이외에도 관측시의 왜곡이 있다. 관측계에서는 되도록 왜곡(distortion)이 없는 관측을 하는 것이 이상적이나, 실제로는 기술적 혹은 경제적 이유에 의해 다소의 왜곡이 있다. 화상관측에 사용하는 비점 주사기(flying spot scanner; FSS)를 예로 들면 비점 편향의 비선형성에 의한 위치의 왜곡, 장소에 따른 밝기의 변화에 따른 셰이딩(shading), 사진 필름농도의 비선형성에 의한 감마(gamma)특성 등이 있다.

이들 왜곡은 미리 특성을 측정해 놓고 디지털 화상으로 컴퓨터 안에 저장한 후 보정하는 방법도 있다. 이들 특성이 시간적으로 변화될 우려가 있는 경우에는 대상에 기준이 되는 표시를 해 놓고 그 표시를 이용하여 보정하면 된다.

패턴의 잡음 및 왜곡을 제거한 후에는, 인식하기 용이한 단위로 나누는 데, 이 과정을 분할이라고 한다. 패턴의 분할에는 다음과 같은 두 가지 경우가 있다. 첫째는, 분리해서 쓰여진 문자와 같이 각 패턴이 고립해 있어서 단순히 외곽만을 결정하면 되는 경우이며, 둘째는 필기체 영문자나 연속 음성과 같이 패턴이 연속되어 있고, 그로부터 인식단위가 되는 단일 패턴을 분리하는 경우이다.

전자의 경우는 문제가 비교적 간단하지만, 분할에 있어서의 위치 결정의 정밀도가 인식시에 문제시되는 일이 있다. 또한, 단위 패턴 사이가 분리되어 있어도 거리가 가까우면 분할이 어렵게 되는 경우가 있으며, 특히 한글이나 한자처럼 단위 패턴속에 분리된 부분 패턴을 포함하는 것이 보통이다.

분할된 인식 단위는 표준 패턴과의 비교를 용이하게 하기 위하여 정규화 과정을 거친다. 패턴은 생성된 때나 생성된 후에 여러가지 조작에 의해 다양한 변형이 생긴다. 이러한 패턴의 변형을 회복하여 기준적인 패턴으로 변환하는 과정이 정규화이다. 그러나, 복잡한 변형을 회복하는 데에 일반적으로 유효한 방법은 없으며, 정규화는 개별적인 인식방법에 포함된다.

정규화의 대상이 되는 것으로는 위치, 크기, 진폭 등이 있으며, 선형 변환의 범위에서 정규화하는 것이 보통이다. 위치의 정규화는 평행이동된 패턴을 기준이 되는 위치로 되돌리는 것에 해당한다. n 차원 벡터 \bar{x} 에 대한 패턴 $f(\bar{x})$ 가 주어질 때, 기준 패턴 $g(\bar{x})$ 에 대해 평행이동의 관계에 있다고 하자. 이때의 정규화는

$$f(\bar{x} + \bar{a}) = g(\bar{x})$$

를 만족하는 미지의 파라미터 \bar{a} 를 정하는 일이 된다. \bar{a} 의 값이 다소 변동하더라도 $g(\bar{x})$ 가 안정되는 경우가 바람직하다. 크기의 정규화는

$$f(A\bar{x}) = g(\bar{x})$$

만족하는 파라미터 A 를 정하는 것이다. 진폭의 정규화는

$$K(f(\bar{x}) - m) = g(\bar{x})$$

를 만족하는 K , m 을 결정하는 것이다. 여기서 K 는 비율, m 은 기준값 파라미터이다.

정규화는 해석적인 방법이므로 인쇄문자처럼 해석 모델이 잘 맞아 들어가는 대상에 대해서는 적합하지만, 변형이 복잡한 필기문자와 같은 대상에도 반드시 적합한 것은 아니다. 후자와 같은 대상에서는 변형의 범위를 대강 규제하여 상세한 변형은 인식 방법으로 흡수한다.

이상과 같은 전처리 과정은 인식의 준비과정이라고 할 수 있으므로, 되도록 간단한 작업으로, 인식에 비해 많은 시간을 요하지 않으며, 인식에 필요한 특성을 크게 변화시키지 않는 범위에서 수행된다. 그러나, 인식대상이 명확하게 알려지지 않은 상태에서는 잡음을 추출하거나, 인식단위의 분할을 하는 것 등이 불가능할 경우도 있다. 그러므로, 여기서 기술한 전처리의 내용이 인식과 특별히 구별되지 않고, 동시에 행해지는 경우도 많다.

4.2 특징 추출

특징 추출은 변형에 대하여 안정된 물리량을 추출하는 것을 말한다. 예를 들어, 음성의 주파수 스펙트럼은 시간에 대하여 일정한 특징이 된다.

특징 추출의 목적은 정보의 손실 없이 패턴의 차원을 줄여, 중복성(redundancy)을 없애고, 인식에 필요한 시간 및 기억 공간을 줄이는데 있다.

특징 추출은 대상으로 하는 패턴의 특성에 크게 의존하므로 패턴의 성격을 깊이 규명할 필요가 있다. 그러므로 패턴인식 중에서도 특징 추출에 대해서는 일반적인 이론을 세우기가 어려우며, 현재 부분적으로 진보되면서 축적되어 가는 단계에 있다.

특징 추출은 관측에 의해서 얻어진 관측값의 집합을 특징량으로 변환하는 과정이라고 할 수도 있다. 관측값과 특징은 일반적으로는 성질이 서로 다른 양으로, 필기 문자 인식을 예로 들면, 관측값은 그물눈 패턴(mesh pattern)인데 비해, 특징은 단점, 교차점, 굴곡, 각 등이 되는 경우이다.

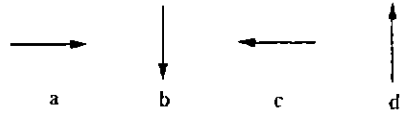
특징을 선택할 때는 급간(interclass, interset)의 상이점과 급내(intraclass, intraset)의 공통점을 잘 표현하도록 하는 것이 중요하다. 일반적으로 사용되는 특징 추출 방법에는 푸리에 변환(Fourier transform), 다변량 분석(multivariate analysis), K-L 변환(Karhunen-Loeve transform) 등이 있다.

4.3 식별

식별은 미지의 패턴이 주어졌을 때 그 패턴이 어느 범주에 속하는가를 결정하는 것이다. 식별의 방법을 크게 분류하면 결정 이론적 방법(deterministic approach)과 구조해석적 방법(syntactic approach)으로 나눌 수 있다.

결정 이론적 방법은 미지 패턴으로부터 추출한 특징에 대해서 식별 함수를 할당하는 일로 볼 수 있으며, 패턴 인식에 있어서 이론적으로 가장 상세히 연구되고 있는 분야이다. 결정 이론적 방법은 어떤 식별 함수를 계산해야 하므로, 각 특징량이 수치로 구해져야 한다.

한편, 필기 문자나 음성과 같이 매우 다양한 변형을 보이는 패턴에 있어서는 특징이 양적이라기 보다 질적인 것이기 때문에, 특징의 존재 유무나 출현하는 양상 등에 의해 인식을 행할 필요가 있다. 이러한 대상의 특징은 규칙에 따라



(그림 2) 4방향 단위 벡터

문법으로 규정되어 있으므로, 식별은 미지 패턴의 특징 계열이 이러한 문법에 적합한지의 여부로 결정된다. 이를 구조해석적 방법 또는 언어이론적 방법이라 부른다.

구조해석적 접근은 대상 패턴의 성질이나 사용하는 특징의 성격에 의해 크게 영향 받고 패턴 의존성이 큰 데 비해, 결정 이론적 접근은 일반적으로 해석이 가능하다.

본 장에서는 구조해석적 접근의 한 예와, 결정이론적 방법의 기본이라 할 수 있는 군집화와 비선형 정합에 대하여 간단히 언급하고, 현재 활발히 연구되고 있는 분야인 신경회로망과 은닉 마코프 모델(hidden Markov model)에 대하여 기술한다.

4.3.1 구조해석적 접근

구조해석적 방법(syntactic method), 또는 언어이론적 접근법으로 불리는 방법은, 질적 패턴 구조를 규정하는 법칙을 추정하여 패턴을 식별하려는 것이다.

간단한 예를 들면, 그림 2와 같은 4방향 단위 벡터 a, b, c, d로 패턴을 표현할 수 있다. 임의의 선도형이 벡터의 연쇄(chain)에 의해 벡터열의 형태로 쓸 수 있다.

즉

$$A = a^1 b^3 c^4 d^3$$

은 직사각형이고,

$$B = a^3 b^2 cdc^3 d^2 ad$$

는 직사각형이 아니다. 일반적으로

$$S = a^m b^n c^m d^m$$

은 직사각형을 나타내고, 이러한 직사각형의 표현은 문법구조에 의해 규정된다.

구조 문법에서는 위의 a, d, c, d를 프리미티브

(primitive)라 부르고, 문장 S를 생성 규칙(rewriting rule, production rule)에 따라 생성함으로써 프리미티브의 연쇄를 도출한다. 반대로 어떤 연쇄가 주어지면, 그 연쇄가 대상 문법에서 도출되었는지의 여부에 따라 그 언어에 속하는가를 판정한다. 패턴 인식에 있어서는 패턴의 특징을 프리미티브로 잡아, 각 카테고리를 규정하는 특징간의 관계를 생성 규칙으로 표현하면 패턴생성이 가능하게 되고, 패턴인식 과정은 주어진 패턴의 구문 해석으로 기술할 수 있다.

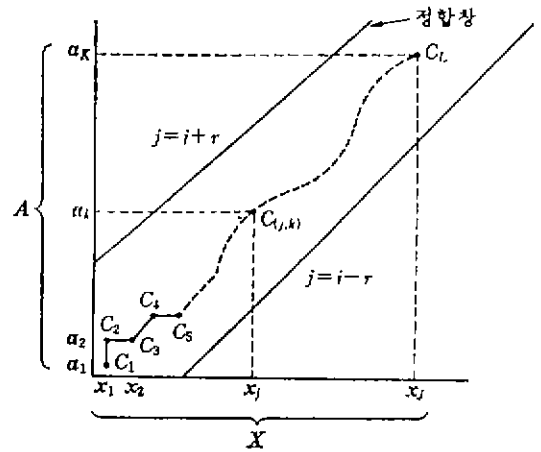
4.3.2 군집화(clustering)

일반적인 식별에 있어서 카테고리는 확정되어 있고, 그 식별 함수도 이미 알고 있는 경우가 많다. 그러나 경우에 따라서 카테고리가 확정되어 있지 않거나, 카테고리가 정해져 있어도 그 안에서의 패턴의 분포가 흩어져 있어 한 개의 식별 함수로는 기술이 곤란한 경우도 있다. 패턴 분포가 주어졌을 때, 같은 종류라고 생각할 수 있는 몇개의 서브클래스(subclass)로 분할하는 것을 군집화(clustering)라고 한다. 군집을 형성할 때에는 두 개의 패턴이 동일 군집에 속하는지 여부를 결정할 필요가 있으며, 패턴간의 거리를 정의해야 한다. 흔히 사용되는 방법에는 K-평균(K-means) 알고리즘과 Isodata 알고리즘[2] 등이 있다.

4.3.3 비선형 정합(nonlinear matching)

표준 패턴과 미지의 패턴을 비교할 때, 두 패턴의 길이가 서로 다를 경우에는 선형신축으로 두 패턴의 길이를 같게 할 수 있다. 그러나, 패턴 내의 특징의 변화가 선형이 아닐 경우에는 선형신축은 의미가 없게 된다. 특히, 단어별로 발성된 음성에서 단어 내의 각 음소의 길이는 발성할 때 마다 크게 변화한다. 그러므로, 표준 패턴과 미지의 패턴을 비교할 때, 음소의 길이의 차를 조절하여 최적의 정합을 해야 한다. 즉, 시간축 정규화가 필요하다. 음성 패턴과 같이 시계열로 입력되는 패턴을 표준 패턴과 조합하기 위한 이론으로 동적 계획법(dynamic programming : DP)을 이용한 정합법이 있다[17, 18].

음성 패턴을 일정한 주기로 분석하여 그 주



(그림 3) 시계열로 이루어진 음성 패턴의 시간축 정규화 정합

파수 성분을 특징 패턴 X_j 로 표현하면, 입력 패턴은 특징 벡터의 시계열 패턴 $X_1, X_2, \dots, X_j, \dots, X_J$ 로 나타낼 수 있다. 음성은 발성하는 법이나 기분에 따라 동일 발성자가 동일 단어를 발성하여도 특징 벡터나 시계열적 구조가 일정하지 않고 복잡하게 변동한다. 그러므로 시간축 방향의 변동을 비선형으로 신축시켜 표준 패턴과의 정합을 계산하는 방법이 소련의 Vintsyuk과 일본의 Sakoe 등에 의해 제안되었다.

입력 패턴을 $X=x_1, x_2, \dots, x_j, \dots, x_J$ 라고 하고 이것과 대조하는 표준 패턴을 $A=a_1, a_2, \dots, a_k, \dots, a_K$ 라 하자 두 시계열 패턴의 길이 J, K는 서로 다르고, 시계열의 대응도 선형으로 신축되어 있지 않으므로, 그림 3과 같이 시간 대응 계열 $C=c_1, c_2, \dots, c_L$ 을 생각하여 각 시간 대응점 C_l 에서 입력 패턴 X와 표준 패턴 A와의 거리 $d(C_l)=d(j, k)$ 를 계산한다. 거리의 정의로는 두 특징 벡터 간의 유클리드 거리나 최우 벡터 거리 등이 쓰이고 있다. 시간 대응 계열 C에서의 두 특징 벡터 패턴 계열의 가중 평균을 구하고, 그것을 최소로 하는 계열 $\{C_l\}$ 을 구하는 문제가 된다.

$$D(X, A) = \min_c \frac{\sum_{l=1}^L d(C_l)w_l}{\sum_{l=1}^L w_l} \quad (1)$$

여기서 가중값 w_l 은 시간 대응 계열 C를 구성하는 대응점 수에 따른 거리의 차를 정규화하기 위한 계수이다. 식 (1)의 분모가 계열 C에 의존하지

않으면 분자는 가산식의 최소화 문제가 되고 동적 프로그래밍 방법을 이용할 수 있다. 보통은 시간 대응 계열 C의 선택법에 제한을 두어, 계산의 효율화를 도모하며 극단적인 대응을 방지하고 있다.

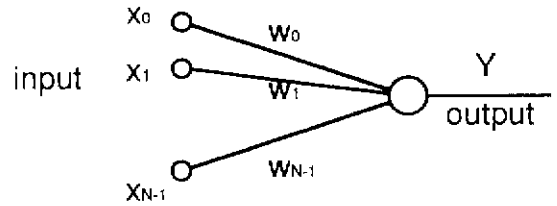
4.3.4 신경회로망

복잡한 인지 작업을 하는 인간의 뇌는 약 10억내지 100억개의 신경 세포(neuron)로 이루어져 있다. 각 신경 세포는 약 천개내지 10만개의 다른 신경세포와 신경연접(synapse)을 통해 연결되어 정보를 전달한다. 이를 응용한 신경회로망의 구조는 (그림 4)와 같이 표현할 수 있다. 하나의 노드(node)는 N개의 입력을 위한 연결선을 가지며, 각각의 연결선에 가해지는 입력은 연결선의 가중치(weight)가 곱해진 후 합산된 값이 어느 임계치 이상이면 출력이 되는데, 출력은 비선형 함수에 의하여 결정된다. 하나의 노드의 출력은 또다른 여러개의 노드의 입력이 된다. 실제로 인식의 대상이 되는 패턴은 여기서 가중치의 형태로 저장 된다고 볼 수 있다. 즉, 입력이 어느 일정한 형태를 가지면 출력이 일정한 값이 되는 것이다.

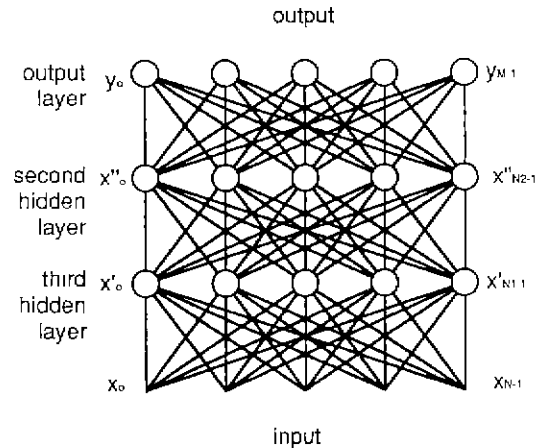
신경회로망에서의 학습은 패턴의 부류에 따라 가중치를 적당한 값으로 지정하는 과정이라고 할 수 있다. 즉, 원하는 패턴이 입력으로 주어질 때 지정된 노드가 일정한 출력값을 내도록 하는 것이다. 신경회로망은 학습방법에 따라 교사학습(supervised learning)과 비교사학습(unsupervised learning)으로 나눌 수 있다. 교사학습은 입력 패턴과 출력의 형태가 주어지면, 이에 따라 가중치가 조정되는 것이고, 비교사 학습은 다수의 입력 패턴만이 주어지면 신경회로망 자체에서 분류하여 서로 상이한 패턴 부류에 대하여 서로 다른 출력값을 내도록 하는 것이다.

초기에 제안된 신경회로망으로는 Hopfield network과 Hamming Net 등이 있으며, 이 후에 다양한 특성을 가지는 여러가지 모델에 제안되었고, 그 구조도 복잡화 되어 가고 있다.

여기서는 현재 가장 많이 응용되고 있는 모델의 하나인 다층 퍼셉트론(multi-layer perceptron)에 대하여 간단히 살펴보기로 한다. 세계의



(그림 4) 신경회로망의 기본 구조



(그림 5) 3층 퍼셉트론

층을 가지는 다층 퍼셉트론은 (그림 5)와 같이 구성된다. 그림에서 3층 퍼셉트론은 하나의 입력 노드열과 출력 노드열, 그리고 두 개의 은닉 노드열(hidden layer)로 구성되어 있다. 노드의 수는 응용 분야에 따라 임의로 정하며, 각각의 노드열의 사이에는 비선형 특성이 도입 된다. 일반적으로 다층 퍼셉트론의 층 수를 말할 때는 비선형이 도입된 수를 말하지만, 노드열의 수를 층 수로 말할 때도 있다.

그림 4의 3층 퍼셉트론에는 N개의 입력, M개의 출력 노드가 있다. 비선형 함수로는 어느 것이든 가능하지만, 주로 시그모이드(sigmoid)함수가 사용된다. 패턴 분류 과정에서는 가장 큰 값을 가지는 출력 노드에 해당하는 패턴 부류가 선택된다. 그림에서 첫째 은닉층, 둘째 은닉층 노드의 출력을 각각 $x'_0, x'_1, \dots, x'_{N1-1}$ 각 노드의 내부 임계치를 각각 $\theta'_0, \theta'_1, \dots, \theta'_{N1-1}$ 입력에서 첫번째 은닉층까지의 가중치를 w_{ij} , 첫번째 은닉층에서 두번째 은닉층까지의 가중치를 w''_{ij} , 두번째 은닉층에서 출력까지의 가중치를 w''_{ij} 이라 하면, 출력은 식 (2)와 같이

구해진다. 여기서, $S(\bullet)$ 는 비선형 시그모이드 함수이다. 다층 퍼셉트론의 학습에는 오류역전파(error back-propagation)알고리즘이 사용된다.

$$\begin{aligned} x'_i &= S\left(\sum_{j=0}^{N_1-1} w_{ij}x_j - \theta_i\right), \quad 1 < i < N_1 \\ x''_i &= S\left(\sum_{j=0}^{N_1-1} w'_{ij}x'_j - \theta'_i\right), \quad 1 < i < N_2 \\ y_i &= S\left(\sum_{j=0}^{N_2-1} w''_{ij}x''_j - \theta''_i\right), \quad 1 < i < M \end{aligned} \quad (2)$$

4.3.5 HMM

HMM은 상태(state)천이의 과정으로 실세계를 모델링하는 Markov chain을 기본으로 하고 있다. 먼저, 이해를 쉽게 하기 위하여 여러가지 색의 공이 들어 있는 항아리를 가정해 보자. N개의 항아리가 있고, 모든 공은 M가지 색 중의 하나라고 하자. 먼저 하나의 항아리를 선택한 다음 공을 하나 꺼낸다. 여기서, 항아리의 선택이나, 공의 선택은 모두 확률적으로 이루어진다. 다음에는 다시 다른 항아리를 선택하여 공을 꺼낸다. 이렇게 하여 모두 T개의 공을 순서대로 꺼내었다고 하자. 각각의 항아리에 들어 있는 공의 수나 색은 일정하므로, 우리는 순서대로 선택된 공을 관찰함으로써, 각각의 항아리들이 어떤 순서대로 선택되었는지 추측할 수 있을 것이다.

여기서, 선택된 공들은 어떤 모델의 출력이라고 할 수 있다. 예를 들면, 기호의 열로 변화된 음성이나, 획의 열로 변화된 필기체 문자 등이 이에 해당한다. '가'라고 하는 문자의 모델은 확률적으로 임의의 획의 열을 만들어 내는 것이다. 우리가 실세계에서 관찰할 수 있는 것은 단지 이 출력 뿐이다. 모든 항아리와 각각의 항아리가 선택될 확률은 어느 패턴 부류의 모델이라고 볼 수 있다. 즉, 하나의 패턴 부류에 대하여 하나의 모델이 대응된다. 우리가 출력을 관찰하므로써 가장 확률이 높은 항아리의 열을 찾아내는 것은 바로 패턴을 관찰하여 그 패턴의 부류를 찾아내는 것과 같다.

이제, 보다 정식적으로 HMM을 기술하도록 하자. 위의 예에서의 항아리는 패턴 모델의 상

태이며, 항아리의 선택은 상태의 천이에 해당한다. 선택된 공들은 출력 기호열에 해당한다.

위와 같은 모델을 사용하여 패턴을 분류하고, 모델을 학습시키는일은 다음의 세가지 문제로 귀착된다.

첫째는, 주어진 모델에 대하여 하나의 출력 기호열이 생성되는 확률을 구하는 것이다. 이것은 이미 만들어진 모델을 이용하여, 출력 기호열에 해당하는 패턴의 분류를 하는 것에 해당한다. 이 문제를 해결하는 가장 간단한 방법은 모든 가능한 상태 천이에 대하여 출력 확률을 계산하는 것이지만, 지수적으로 증가하는 계산량을 필요로 하므로 불가능하다. 실제로는 각 상태마다 단계적으로 확률을 구하는 forward-backward 방법이 사용된다.

두번째 문제는, 주어진 관측 기호열, 즉 모델의 출력에 적당한 상태 천이 열을 찾아내는 것이다. 이것은 시스템의 분석을 위해 필요하다. 이 문제의 해결에는, 전체적인 것을 고려하지 않고 각 시간에서 가장 적당한 상태를 선택해 가면서 경로를 찾는 Vierbi 알고리즘을 사용한다.

세번째 문제는, 하나의 모델에서 주어진 관측 기호열이 생성될 확률을 최대로 하도록 모델 파라미터를 조정하는 일이다. 이것은 시스템의 학습에 해당한다. 이 문제의 해결에는 Baum-Welch reestimation 식이 사용된다.

HMM에는 응용 분야에 따라, 또는 모델의 난이도에 따라 여러가지 형태의 모델이 사용된다.

또한, HMM은 출력되는 심벌의 type에 따라 이산형 HMM(Discrete HMM)과 연속형 HMM(Continuous HMM) 등으로 크게 나누어지는데, 이산 HMM의 경우 벡터 양자화 과정을 통해서 얻어진 codebook을 사용하여 입력 패턴을 해당 색인(index)으로 부호화하는 전처리 과정을 필요로 하나, 연속성 HMM의 경우는 입력 자료의 값 그 자체를 그대로 입력하여 사용한다. 전자의 경우, 구현이 간단하고 학습시간 및 학습할 모델의 파라미터들이 적다는 장점이 있으나, 벡터 양자화(vector quantization)시 정보가 유실되는 단점이 있다.

V. 결 론

본 고에서는 패턴 인식에 관하여, 기본적으로 공통적인 내용을 다루었다. 패턴 인식에서 대상으로 하는 분야는 문자, 음성, 화상 등 다양하므로, 각각의 특성을 반영하여 특색 있는 여러 이론이 제안되고 실용화 되어 왔다. 각 분야는 서로 상이한 세계를 다루고 있어, 고유한 접근 방법이 존재하지만, 주어진 표본 패턴군에서 어떠한 공통된 성질이나 함수를 찾아 내야 하는 기본적인 문제는 모든 분야에 적용된다.

현재는 하드웨어 기술의 급속한 발달로, 과거에 시간적, 기억 공간적, 경제적인 이유로 실용 불가능 했던 이론들이 실현되어가고 있다. 또한 정보과학의 진보와 더불어 자동 패턴 인식의 요구도 높아가고 있다. 그러나, 현재까지의 패턴 인식 기술이 현대 사회가 요구하고 있는 수준에 완전히 이르지 못하고 있으며, 아직도 많은 부분은 전시회의 진열장을 벗어나지 못하고 있다. 소비자는 인간에 가까운 능력을 가진 시스템을 기대하고 있는 반면, 실용화 시스템은 인간의 인식 능력과는 거리가 먼 시스템적 방법을 기반으로 하여 구현되어 왔기 때문이다.

우리가 처리 대상으로 하는 패턴은 단순히 코딩되어 있는 것이 아니라, 반드시 의미 정보와 함께 존재한다. 따라서, 외형상의 구분으로 일관되어 왔던 패턴 인식 기술에 환경과 의미에 관한 정보를 효율적으로 추출해 내는 이론의 확립이 요구되고 있다. 단순한 패턴 인식이 아닌, 패턴 이해(pattern understanding)의 방향으로 연구의 중심이 옮겨져 가는 추세도 이에 부응하는 방향으로 이해되어야 할 것이다.

참 고 문 헌

- 오영환, "패턴 인식론", 정익사, (1991).
- J. T. Tou and R. C. Gonzalez : Pattern Recognition Principles, Addison-Wesley, (1974).
- R. O. Duda and P. E. Hart : pattern Classification and Scene Analysis, John Wiley, (1973).
- K. Fukunaga : introduction to Statistical Pattern Recognition, Academic Press, (1972).
- G. S. Sebestyen : Decision Making Processes in Pattern Recognition, Macmillan, (1962).
- K. S. Fu : Sequential Methods in Pattern Recognition and Machine Learning, Academic Press, (1968).
- S. Watanabe(Ed.) : Frontiers of Pattern Recognition, Academic Press, (1972).
- J. R. Ullman : Pattern Recognition Technique, Butterworths Group, London, (1973).
- F. W. Smith : pattern Classifier Design by Linear Programming, IEEE Trans, Comput, C-17, (1968) pp. 367~372.
- J. K. Chang and S. J. Dwyer : New Multiclass Classification Method : modified Maximum Likelihood Decision Rule, Proc. First IJCP, (1973). pp. 334~339.
- P. K. Rajasekaran, M. D. Srinath : Unsupervised Learning in Nongaussian Pattern Recognition, Pattern Recognition, 4, (1972). pp. 401~416.
- K. Fukunaga, W. C. G. Koonts : Application of the Karhunen-Loeve Expansion To Feature Extraction and Ordering. IEEE Trans, Comput. C-19, (1970). pp. 311~318.
- J. Karhunen and E. Oja : "New Methods for Stochastic Approximation of Truncated Karhunen-Loeve Expansion," Proc. 6th, int. Conf. On Pattern Recognition, (1970). pp. 550~593.
- I. Gitman and M. D. Levine : "An Algorithm for Detecting Unimodal Fuzzy Sets and its Application as a Clustering Technique," IEEE Trans, Comput, C-19, (1970). pp. 585~593.
- J. A. Hartign : "Clustering Algorithms," John Wiley, (1975).
- N. V. Findler and J. V. Leeuwen : "A Family of Similarity Measures Between Two Strings," IEEE Trans. Pattern-Anal. & Math. Intell., PAMI-1, (1979). pp. 116~118.
- T. K. Vintsyuk : "Speech Discrimination by Dynamic Programming," Cybernetics, 4, (1968). pp. 81~88.
- H. Sakoe : "Two-Level Dp-Matching : A Dynamic Programming-Based Pattern matching Algorithm," IEEE Tran. Acoustics, Speech, and Signal Processing. ASSP.
- R. P. Lippmann, "An Introduction to Computing with Neural Network," IEEE ASSP Magazine,

- April, (1987). pp. 4~22.
20. L. R. Rabiner and B. H. Juang, "An Introduction to Hidden Markov Models," IEEE ASSP Magazine, January, (1986). pp. 4~15.

오 영 환



- 1972 서울대학교 공과대학(공학사)
1974 서울대학교 교육대학교(석사)
1980 Tokyo Institute of Technology 정보공학(박사)
1981 충북대학교 공과대학 조교수
1981 서울대학교 공과대학 강사

1983 Univ. of California, Davis 연구 교수
1985 한국과학기술원 전산학과 교수
관심분야: 음성인식, 음성합성, 패턴인식 등
