적응 학습 능력을 가진 SEJONG-NET

박혜영○ 이일병
연세대학교 컴퓨터과학과

Adaptive SEJONG-NET

Hyeyoung Park Yillbyung Lee
Dept. of Computer Science, Yonsei University
E-mail: hypark@csai.yonsei.ac.kr

요 약

SEJONG-NET은 시각 문자패턴의 인식 과정을 설명할 수 있는 적절한 패러다임을 제공하기 위해 척추동물의 시신경계 구조와 기능을 모방하여 만든 문자인식 모형이다. 초기에는 온라인 한글 인식을 위하여 설계되었으며, 이후 다양한 문자 집합이나 오프라인 한글 문자를 위한 모델들이 개발되었다. 현재까지 개발된 여러 SEJONG-NET 모델이 가지고 있는 문제점은 경직성이라고 할 수 있다. 즉, 설계 초기에 고려한 인식 대상 문자 집합과 문자 패턴에 대해서만 인식이 가능하고, 변형된 패턴을 기존의 패턴으로 근사화하여 해석하거나 새로운 패턴에 대하여 그것을 추가 학습하는 것이 불가능하다. 따라서 본 논문은 SEJONG-NET의 이러한 제약점을 해결하여 한글 인식 문제에 일반적으로 적용될 수 있도록 개선하는 것을 목적으로 한다. 이를 위해 상위층에서는 인간이 가지고 있는 문자에 대한 구조적인 지식을 표현하고 학습을 통해 추가적으로 습득할 수 있는 형태로 구현하였고, 하위층에서는 상위층에서 쓰이는 구조적인 지식을 표현하는데 적합한 특징을 추출해 낼 수 있도록 구현하였다. 특히 하위층에서는 인간의 초기 시각 피질에서 감지되는 특징들을 추출하도록 구현하여 사용되는 특징이 일반성을 가질 수 있도록 하였다. 이러한 방법을 기반으로 하여 본 논문에서는 변형된 패턴에 대한 적응 학습 능력을 가지며 인지과학적인 사실에 보다 충실하도록 개선된, 온라인 한글 인식을 위한 SEJONG-NET 모델을 제안한다.

I. 서 론

문자 인식 시스템을 만들때 가장 먼저 고려해야 할 것은 인식 대상이 되는 문자 집합과 그 특징이다. 대상이 되는 문자 집합의 특성에 따라 그에 적합한 특징 추출 방법이 적용되고 추출된 특징을 어떤 방식으로 인식할 것인지를 정하게 된다. 한글의 경우 가장 두드러지는 구조적인 특징은 간단한 직선 성분들이 정해진 규칙에 따라 2차원상에 배열됨으로써 하나의 문자를 이룬다는 것이다. 이러한 한글의 구조적인 특성에 부합되는 인식 방법으로 주로 사용된 것이 구문론적 인식 방법이다. 입력 문자의 각 획들을 기본획으로 근사화시켜서 인식하고, 인식된 획들의 위치 관계나 필기 순서 등을 이용하여 정의된 오토마타 등으로 문자를 인식한다. 인식하는

이러한 구문론적 방법을 이용한 인식의 경우에는 필기 한글에서 나타나는 다양한 패턴의 변화에 어떻게 대응하느냐가 인식 성능을 좌우한다. 필기 문자의 경우 획들은 간단한 직선 성분으로 표현되기보다는 영문자, 숫자 등에서 주로 나타나는 곡선적인 성분으로 변형되어 나타나기 때문에 변형된 획을 기본획으로 근사화하는 방법이 인식 성능에 영향을 미친다. 또한 흘려 쓴 문자의 경우 빈번히 발생하는 필기 순서의 변형이나 획 또는 자소간의 이어쓰기와 같은 변형에 대하여서는 오토마타의 설계시에 얼마나 많은 변형을 고려해 주느냐가 주안점이 된다. 그러나 설계자가 수없이 많은 변형을 모두 고려한다는 것은 불가능한 일이고, 따라서 이러한 방법을 사용한 경우에는 인식 가능한 문자 패턴에 제약이 가해지는 것이 불가피하다. 이러한 문제를 해결하기 위해 획 인식 과정에서 신경망을 사용하여 인식 성능을 높이려는 연구와 문자 조합 규칙을 적응적으로 학습하려는 대한 연구들이 진행되었다[3, 4, 5].

SEJONG-NET은 동물의 시신경계 구조를 바탕으로하여 하위층에서는 획인식을 위한 특징 추출을 하고, 상위층에서는 획,자소,문자를 인식함에 있어서 문자의 구조적인 지식을 이용하도록 설계된 모델로 구문론적인 방법과 마찬가지의 문제를 가지고 있다. 비록 한글 인식을 위한 초기의 SEJONG-NET모델 이후에도 그 기본 구조를 이용하여 다양한 문자 집합에 대한 인식 모델들이 개발되었지만[1], 획의 변형을 효과적으로 흡수하거나 필기 방식에 따른 다양한 변형에 대한 인식 능력을 갖고 있지는 못하다. 이러한 문제의 해결을 통해 SEJONG-NET은 한글 인식에 있어서 보다 좋은 성능을 보여줄 수 있을 뿐 아니라, 다양한 문자 인식에 일반적으로 적용될 수 있는 보다 효과적인 모델로 발전할 수 있을것이라 기대된다.

II. SEJONG-NET의 기본구조

SEJONG-NET은 시각 문자 패턴의 인식 과정을 설명할 수 있는 적절한 패러다임을 제공하기 위해 척추동물의 시신경계 구조와 기능을 모방하여 만든 문자 인식 모형이다. SEJONG-NET의 기본 특징은 계층적인 처리 구조를 가진다는 것이다. 이러한 처리 구조는, 동물의 시각 시스템은 입력된 정보가 최종적으로 특정한 일반 개념으로 분류되기까지 단계적인 층들을 거치면서 추상적인 개념으로 변환되는 구조를 가지고 있다는 사실에 기반을 둔 것이다[6]. 또 다른 중요한 특징은, SEJONG-NET이 다루는 정보 형태는 정적 시각 패턴이 아니라 시간에 따라 변화하는 동적 시각 패턴도 함께 다룬다는 점이다. 시공간적인 정보는 각각 시간 정보와 공간 정보를 다루는 분리된 두 처리 경로를 통해 병렬적으로 처리된다. 이것은 인간의 경우 형태(form), 움직임(motion), 색깔(color)을 처리하는 분리된 경로가 존재한다는 생리/해부학적인 실험결과를 바탕으로 하고 있다[6, 7].

〈그림 1〉과 〈그림 2〉에 SEJONG-NET의 기본 구조와 기능적인 처리 과정을 도시하였다. 마우스나 펜과 같은 온라인 입력 기기로 부터 들어오는 정보는 2차원 이미지로 표현되고, 이 패턴이 각 층을 거치면서 변환되고 추상화되어 최종적으로 하나의 문자 개념으로

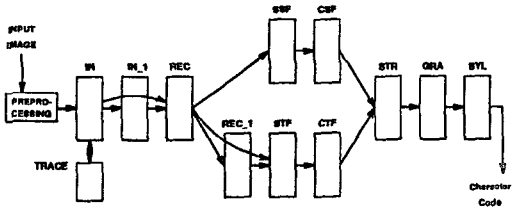

그림 1: SEJONG-NET의 기본 구조

분류된다. 입력된 이차원 이미지를 갖고 있는 IN층은 망막상에 맺힌 영상을 표현하는 것으로 자극의 유무에 따라 이진값으로 표현된다. 다음 단계인 REC층은 망막에 있는 감광세포(photoreceptor)를 모델링한 것으로, 최근에 들어온 자극에 가장 크게 반응하는 출력 패턴을 가진다. REC층으로부터의 출력은 두개의 분리된 경로를 거치면서 처리되는데, 공간적 정보를 다루는 경로와 시간적 정보를 다루는 경로가 그것이다. 각각의 경로는 두개의 층으로 구성되는데 비교적 간단한 특징을 찾아내는 층(SSF층/STF층)과 그들로부터의 출력 결과를 조합하여 보다 복잡한 특징을 추출해내는 층(CSF층/CTF층)이 존재한다. 두개의 경로를 거치면서 보다 추상화된 정보는 STR층에서 다시 합쳐진다. SEJONG-NET은 초기에 온라인 한글을 인식하기 위해 개발되었기 때문에 한글에 적합한 형태로 그 구조를 설계하였다. STR층과 GRA층, SYL층은 각각 한글을 구성하는 기본 성분이 되는 획과 자소, 음절을 인식하도록 설계되었다. 특히 한글은 비교적 간단한 획으로 구성되기 때문에 상위층들은 한글의 구조적 지식을 표현하는 간단한 오토마타로 표현되었다.

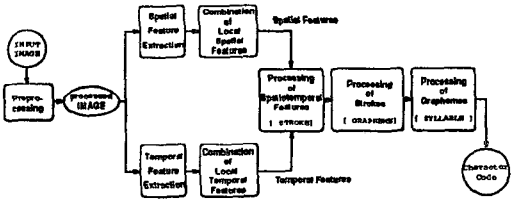

그림 2: SEJONG-NET의 기능적인 처리 과정

최초의 SEJONG-NET 모델이 온라인 한글 인식에 적용되어 긍정적인 결과를 보인 이래 영문자와 한자 등 다양한 문자 집합을 인식 대상으로 하는 SEJONG-NET 모델들이 개발되었다[1]. 또한 구조적인 지식을 이용하여 정적문자로부터 동적 정보를 추출하고 이를 인식에 사용함으로써, 오프라인 한글을 인식하려는 연구가 있었다[2]. 그러나 이러한 여러가지 모델들이 개발되었음에도 불구하고, 한글의 경우만 살펴 볼 때 변형된 패턴에 대한 인식 성능면에서는 많은 발전을 이루지 못하였다. 그 이유는 여러가지 변형된 모델들이 최초의 SEJONG-NET의 근본적인 문제에 접근하기 보다는 단순히 인식하고 싶은 문자 집합을 처리하기 위한 과정을 추가하는 방식으로 발전을 거듭해 왔기 때문이다. 그 결과 SEJONG-NET의 인식 대상 문자 집합의 수는 늘었지만, 각각의 문자 집합에서 나타나는 다양한 변형을 처리하는 능력을 갖추지 못했으며 각 층의 기능과 처리 과정은 일관성을 잃어버리게 되었다.

이러한 문제의 근본 원인은 SEJONG-NET이 처음 설계 당시에 다양한 문자 집합이나 변형을 충분히 고려하지 않았기 각 층이 추출해 내는 특징이나 처리 과정이 일반성을 갖고 있지 못하다는 데 있다. 따라서 현재의 SEJONG-NET이 당면한 가장 큰 과제는, 기본 구조를 재정립하고 각 층의 기능이 보다 일반성을 가질 수 있도록 구현함으로써 패턴의 변형에 대한 적응력을 갖출 수 있도록 하는 것이다.

III. 제안하는 SEJONG-NET

3.1 제안하는 SEJONG-NET

인식에 있어서 문자 패턴의 변형에 대한 문제는 크게 세가지로 나눌 수 있다. 첫째는 획 자체가 가지는 변형으로, 한글의 경우 직선적인 기본획이 필기 형태에 따라 곡선 성분이나 다른 기울기로 나타난다. 두번째는 획간의 연결 문제로, 한글의 경우 자소내의 흘림 및 자소간의 흘림 형태에 따라 수없이 많은 변형이 존재한다. 세번째로는 온라인의 경우 획순의 변형 문제이다. 여러 개의 기본획이 복잡하게 조합되는 한글의 경우 획순 변형 문제 역시 필기자에 따라 대단히 많은 변형을 가지고 있다.

이상과 같은 문제를 해결하기 위해 본 논문에서는 우선 SEJONG-NET이 바탕으로 하고 있는 척추동물의 시각 시스템의 처리과정에 대하여 재검토하였고, 이를 참조하여 다음과 같은 가정을 바탕으로 SEJONG-NET을 재구성하였다.

- 시각 패턴의 처리를 구성하는 각 계층에는 특정한 패턴에 대해 반응하는 세포들이 존재하고, 이 세포들은 입력 패턴으로부터 자신이 감지하고자 하는 특징들에 반응하여 능동적으로 발화하는 것으로 볼 수 있다[6]. 이와 같은 사실을 바탕으로 SEJONG-NET의 각 층에는 특정 패턴을 감지하는 감지기가 존재하여 하위층으로부터 들어오는 입력에 대하여 자신의 특징과 부합되는 것이 존재하는지 탐색한다. 이렇게 구현함으로써 각 층에서는 잡음이 포함된 입력 패턴에 대해서도 자신과 부합되는 부분을 감지해 내고, 그에 선택적 주의를 할 수 있으므로 패턴의 변형을 충분히 수용할 수 있을 것이다.
- 인간의 초기 시각 피질에서 감지되는 특징을 SEJONG-NET의 하위층에서 추출하도록 구현하였다. 인간의 초기 시각 피질에서 추출되는 특징들은 다양한 시각 패턴을 인식하는 기본 요소들로, 이 특징을 SEJONG-NET에서 사용함으로써 다양한 패턴 집합이나 그 변형에 영향을 받지 않고 문제 해결에 적용할 수 있을 것이다.
- 인간이 하나의 문자를 학습할 때에는, 우선 그 문자를 구성하는 기본 요소들을 설정하고 그것들이 조합되는 규칙을 기억하는 것으로 간주된다. 한글의 경우 직선적인 기본획이 문자의 기본 구성 요소가 되고, 그것들의 조합 규칙을 학습함으로써 하나의 문자를 학습한다. 이것은 어린이가 한글을 배우는 과정과 부합된다고 볼 수 있다. 이렇게 구성된 시스템에서 변형된 패턴은 그 기본 요소와 조합 방식을 학습함으로써 충분히 인식될 수 있을 것이다.

이상과 같은 가정을 바탕으로하여 재구성된 SEJONG-NET의 각 층의 기능은 다음과 같다.

3.2 IN 층, TRACE 층

SEJONG-NET은 마우스나 펜과 같은 입력 기기를 사용한다. 이러한 입력 기기의 경우, 기계적 속도와 인간의 필기 속도 차이에 의해 입력은 불연속점들의 열로 주어진다. 이러한 입력은 메우기(relaxation) 과정을 거쳐서 불연속적인 점들의 연결성을 유지시켜준다. 메우기을 거쳐 복원된 이미지는 51×51 크기로 정규화되어 이진값으로 저장된다. 또한, 입력된 점들의 순서 정보는 상위층의 처리를 용이하게 하기 위해서 TRACE층에 저장된다. IN층은 단순히 자극의 유무만을 표시하는 것으로, 처리가 가해지기 전의 순수한 이미지를 갖고 있다. SEJONG-NET은 시간에 따라 변화하는 시각 패턴을 다루는 것으로, 필기가 진행되는 동안 IN층은 계속 새로운 자극을 받아들인다. (그림 1)에 나타나 있는 IN_1층은 한 단위시간 전의 IN층의 이미지를 나타내는 것으로 자극의 변화를 탐지하기 위해 사용된다.

3.3 REC 층

REC층은 망막에 들어온 빛 자극을 가장 먼저 받아들이는 감광세포를 모델링한 것이다. 빛 자극에 대한 감광세포의 반응 패턴은 자극이 주어지는 순간에 강하게 발화하고, 자극이 지속되는 동안 안정된 값을 유지한다.

REC층의 출력은 이러한 특성을 나타낼 수 있도록 모델링하여 가장 최근에 입력된 자극이 가장 큰 값을 가지고 나머지는 안정된 값을 갖도록 하였다. IN층과 IN_1층의 정보를 사용하여 REC층의 값을 구하는 식은 다음과 같다.

$$REC(t) = W_1 \times IN_1(t) + W_2 \times (IN(t) - IN_1(t)) \quad (1)$$

여기서 W_1과 W_2의 값은 임의로 정해진 값으로, 자극이 주어질 당시의 값이 그 이후의 값보다 크기만 하면 된다. 이렇게 계산된 REC층의 출력은 시간 정보를 다루는 경로와 공간 정보를 다루는 경로로 분리되어 병렬적으로 처리된다. (그림 1)에 있는 REC_1층은 IN_1층과 마찬가지로 단위시간 전의 REC층의 값을 갖고 있다.

3.4 STF 층

인간의 경우에 시간적 정보를 다루는 경로는, 정확히 말해서 움직임에 대한 정보, 즉 움직이는 방향(direction selectivity), 속도(velocity selectivity)등을 다루는 부분이다. 그러나 문자 인식의 경우에는 복잡한 시간 정보는 크게 중요하지 않다고 여겨지므로, STF층에서는 보다 초기 단계에서 이루어지는 간단한 시간 정보만을 추출하도록 하였다.

그림 3: ON-OFF 신경절세포의 발화 패턴

망막에서 이루어지는 정보처리중 가장 마지막 단계를 담당하고 있는 것은 신경절세포(ganglion cell)이다. 신경절세포는 그 발화 패턴에 따라 몇가지로 나누어지는데 그중에 시간 정보를 처리한다고 여겨지는 것은 ON-OFF 신경절 세포로서, 그 발화 패턴은 (그림 3)과 같다[6].

SEJONG-NET에 들어오는 자극은 한번 ON되면 하나의 문자가 인식될 때까지 계속 그 자극이 주어지므로 STF층에서는 자극이 없다가 생길때 발화하는 패턴을 생성하면 된다. STF층의 계산식은 다음과 같다.

$$STF(t) = \begin{cases} 1 & \text{if } REC(t) - REC_1(t) > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

3.5 CTF 층

CTF층은 STF층으로부터 정보를 받아 그것을 조합하여 출력을 생성하고, 이 출력은 STR층에서 공간적인 정보와 합쳐진다. 여기서, CTF층의 역할을 정하기 전에 먼저 시간적 정보와 공간적 정보가 합쳐지는 것이 어떤 역할을 하는 것인지를 먼저 고려해야 한다. 인간의 시각 시스템이 형태, 움직임, 색깔에 관한 3가지 정보를 서로 다른 경로를 통해 처리한다는 것은 생리/해부학적으로 그를 뒷받침하는 많은 증거들이 존재한다[6, 7]. 그러나 분리된 경로들이 어떻게 서로 결합되며, 결합되는 부분에서는 어떤 기능을 하는 지에 대해서는 자세히 알려져 있지 않다. 다만 분리된 경로들의 결합이 선택적 주의를 하는데 중요한 역할을 한다는 가정이 지배적이다. 따라서 본 논문에서 CTF층은 하나의 획이 완성되었는지를 판단하여 하나의 획 정보가 모두 입력되면 STR층이 현재 들어온 획에 선택적 주의를 할 수 있도록 신호를 보내는 기능을 갖도록 하였다. 하나의 획이란 연속된 세포들의 열이라고 생각할 수 있다. 따라서 CTF층은 STF층에서 현재 발화된 세포와 이전에 발화된 세포의 위치를 비교하여 인접해 있지 않으면, 이전 세포까지를 하나의 획으로 간주하여 STR층으로 신호를 보낸다.

3.6 SSF 층

동물의 시각 시스템에서 공간적 정보를 다루는 부분에서 특히 초기 시각피질에 대해서는 Hubel과 Wiesel의 실험과 가정이 일반적으로 받아들여지고 있다. 가장 중요한 공간적인 특징은 기울기 성분으로, Hubel과 Wiesel은 고양이의 초기 시각피질에 기울기를 탐색하는 세포(Orientation Detector)가 있음을 밝히고 그 수용야의 특성에 따라 단순 세포과 복합세포로 분류하였다[6]. 단순세포와 복합세포의 수용야가 (그림 4)에 나타나있다. 단순세포는 수용야의 흥

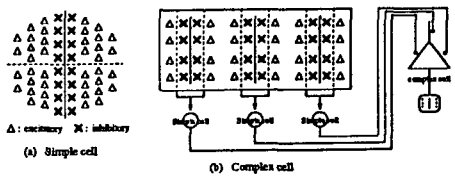

그림 4: (a) 단순세포의 수용야 (b) 복합세포의 수용야와 연결형태

분성 영역에만 자극이 주어질 때 발화함으로서 특정 기울기를 감지한다. 복합세포는 단순 세포로부터 정보를 받는다고 알려져 있고, 단순세포보다는 위치에 덜 민감하여 특정 기울기가 어느 정도 다른 위치에 나타나더라도 발화하는 위치 변형에 적응성을 갖고 있다[6].

SSF층은 이상의 사실을 바탕으로하여 기울기를 탐지하도록 구현하였다. 0°, 45°, 90°, 135°의 4개의 기울기 성분을 탐색하기 위해 4개의 필터를 사용한다. 각 필터는 각 기울기를 탐색하는 세포를 나타내는 것으로 그 수용야의 형태는 (그림 5)와 같다. 수용야의 형태

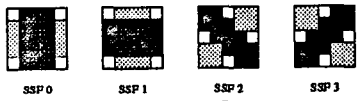

그림 5: SSF층 필터의 수용야

는 단순세포와 복합세포의 수용야를 합친 것으로, 이렇게 한 이유는 입력 패턴에서 직선 획이 울퉁불퉁한 형태로 나타나는 경우에도 올바르게 기울기를 추출할 수 있도록 하기 위해서이다. 또 크기를 5×5로 정한 것은 기울기 성분의 특징을 나타낼 수 있는 최소의 단위를 선택한 것이다. 필터의 가중치는 -1, 0, 1의 값중의 하나를 가지며 기본적으로 학습을 고려하지 않았다. 동물의 경우 기울기를 감지하는 세포의 학습이 가능하고, 발달 초기에 특정 기울기 성분만을 볼 수 있게 하면 다른 기울기 성분을 감지하는 세포가 존재하지 않게 된다는 실험 결과가 있다[7]. 그러나 일반적인 패턴을 다루게 하기 위해서는 특정 학습 패턴을 가지고 학습시켜 만든 가중치를 사용하는 것보다는 일반적인 동물의 시각 시스템에 나타나는 특징을 검출할 수 있도록 가중치를 고정시키는 것이 바람직하다고 생각하였다. 또한 4개의 기울기 성분만을 검출하도록 한 것은 모든 곡선 성분이 4가지 기울기 성분의 조합으로 표현 가능하다는 가정을 바탕으로 한다. 결국 SSF층은 각각의 기울기 성분에 대한 발화값을 표현하는 4개의 면으로 구성된다.

3.7 CSF 층

Hubel과 Wiesel은 단순세포와 복합세포의 다음에는 보다 복잡한 특징을 검출하는 초복합세포가 있을 것이라는 가정을 세웠고, 그 가정을 뒷받침하는 실험들이 현재까지 계속되고 있다[8, 9]. 초복합세포가 하는 기능중 중요한 것은 길이 감지이다. 이밖에 최근의 연구결과에 의하면 길이 감지 이외에 특정 기울기 성분의 곡률을 감지

하는 세포들이 존재한다는 것이 밝혀짐으로써 Hubel의 가정을 다시 한번 지지하였다[8]. 그러나 이러한 길이 감지 및 곡률 감지 세포와 초기 시각 피질에 있는 end-free 세포들과의 연결 관계에 대한 확실한 실험적 자료는 존재하지 않으며, 몇몇 가정이 제시되고 있다. 한 예로 하나의 길이를 감지하는 세포는 특정 기울기 성분을 감지하는 복합세포들로부터 입력을 받아들이는데, 가운데 있는 세포로부터는 흥분성 입력을 받아들이고, 양 끝에 있는 세포들로부터는 억제성 입력을 받아 들인다. 이렇게 하면 가운데 있는 세포들의 전체 수용야 길이에 따라 감지해내는 길이가 정해지게 되는 것이다.

이상과 같은 생리학적 실험 결과와 가정을 바탕으로하여, CSF층은 초복합세포와 같은 기능을 수행하도록 하였다. 즉, 길이를 감지하고 곡률을 감지하는 것이 CSF층의 주된 역할이다. CSF층은 SSF층의 4개의 면으로부터의 출력 결과를 입력으로 받는다. 우선 길이를 감지하기 위해서는 각 면에 대하여 하나의 활성화된 세포를 찾고, 그 세포를 중심으로 해당 기울기 방향으로 수용야를 늘려가면서 억제성 필터를 적용한다. 억제성 필터가 놓여진 곳에 활성화된 세포가 존재하지 않으면 그곳이 해당 기울기를 가진 하나의 개체가 끝난 지점으로 받아들이고, 그때의 길이가 정보로 얻어진다. 동시에 SSF층에서 하나의 독립적인 세포들로 존재하던 것들이 묶여져 하나의 개체로 추상화하는 역할을 수행한다. 각 기울기 성분에 대한 억제성 필터는 3×3 크기로 구현하였고, 그 수용야 형태는 (그림 6)과 같다.

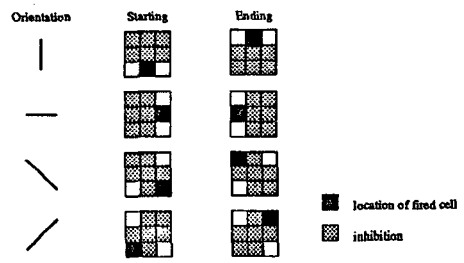

그림 6: 길이 감지를 위한 필터의 수용야

곡률을 감지할 때는 길이 감지의 과정에서 찾아진 개체들을 중심으로 수행한다. 각각의 개체에 대해 양 끝부분에 곡률을 감지할 수 있는 필터를 적용하여 필터링 값을 계산한다. 각 필터는 해당 기울기의 휘어짐과 관련된 기울기 성분을 담당하는 SSF층의 면에 놓여진다. 하나의 기울기 성분에 대해 휘어짐은 두 방향으로 생각할 수 있는데 ((그림 7) 참조) 전체 곡률은 curvature1에서 curvature2를 뺀 값으로 나타낸다. CSF층의 출력은 두가지로 나눌 수 있는데, 하

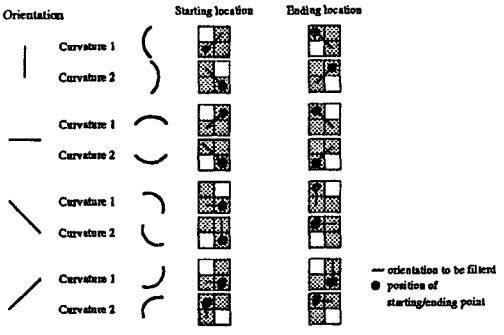

그림 7: 곡률 감지를 위한 필터의 수용야

나는 찾아진 개체들의 리스트이다. 해당 리스트는 각 개체의 기울기 성분과 크기, 시작점과 끝점, 중간점의 좌표, 곡률값 정보를 가진다.

3.8 STR 층

STR층은 CTF층으로부터 신호를 받아 동기화되어 CSF층으로부터 추출된 각 개체들과 그들간의 위치관계를 나타내는 그래프를 만든다. 이 그래프의 기본 구조는 개념그래프(conceptual graph)이론을 따른다[10]. 개념그래프는 개념노드(concept node)와 관계노드(relation node)의 두가지 노드와 방향성 있는 아크로 구성된다. 노드들간의 직접적인 연결은 서로 다른 두 종류의 노드들 사이에서만 가능하다. STR층의 경우, CSF층에서 추출된 각 개체들을 개념노드로 구성하고, 그들간의 위치 관계를 관계노드로 구성한다. 연결의 방향성은 개념노드로부터 관계노드로 향하도록 한다. 개념노드가 갖고 있는 정보는 CSF층의 개체가 갖고 있는 정보와 동일하다. 만약 해당 개체가 있는 위치에 다른 기울기 성분이 함께 존재한다면 겹쳐진 정도를 그것의 방향 성분이 발화한 값으로 함께 갖고 있는다. 예를 들어, CSF층의 기울기 0°인 개체와 연결된 STR층의 노드가 하나 있다면, 그것의 기울기 성분은 0°이고 이것을 나타내는 값이 1이 된다. 그런데 그 위치에서 45° 성분도 함께 탐지되었다면, 중첩된 부분의 크기를 계산해서 그 비의 값을 0° 기울기 성분 값으로 가진다. 이렇게 하는 이유는 필기에 따라 나타나는 기울기 정도의 변화를 흡수할 수 있도록 하기 위해서이다. 관계노드는 상대적인 위치를 나타내는 4가지 경우와 접촉된 부위를 나타내는 9가지 경우로 표현한다. 이때 두개의 개념노드 중 어떤 것을 중심으로 계산하느냐에 따라 값이 달라지므로, 두가지 경우에 대한 값을 모두 갖고 있도록 하였다. 이들 값은 두 개념노드의 처음,끝,중간 점들의 좌표를 이용하여 계산된다. 예를 들어 오른쪽 위치를 나타내는 값은 중심이되는 개념노드의 처음,중간, 끝의 x좌표가 상대방의 x좌표보다 모두 작으면 1, 모두 크면 -1, 그외에는 0으로 나타내었다. 또 접촉 부위를 나타내는 값은 처음과 처음의 경우는 두개의 시작점의 좌표사이의 거리를 계산하여 그 역수값을 취하였다.

STR층의 역할은 하나의 획에 선택적 주의를 하여 개념그래프를 만들어 내는 것이고, 이 개념그래프는 GRA층이 가지고 있는 자소에 대한 개념그래프와의 매칭을 통해 인식된다.

3.9 GRA 층

GRA층은 14개의 단자음과 10개의 단모음에 대한 기본 개념그래프를 가지고 있어서 이것이 하나의 자소를 탐지하는 세포 역할을 한다. 이 개념그래프의 구조는 STR층의 경우와 동일하다. 단, 각 노드가 갖고 있는 값들(기울기 성분 값, 곡률값, 위치관계 값)과 개념노드와 관계노드 사이의 연결 강도 및 각 개념노드가 하나의 자소를 형성하는 데 미치는 기여도를 값으로 갖고 있어서 각각 학습이 가능하다. 실제시에 기본 개념그래프들도 학습을 통해서 얻을 수 있는데, 각각의 자소에 대해 그 기본 패턴을 작성해 주면, 그것을 입력으로 받아 형성된 STR층의 개념그래프를 해당 자소의 기본 개념그래프로 등록된다.

인식시에 각 자소에 대한 개념그래프들은 STR층에서 생성된 개념그래프에서 자신과 부합되는 부분을 찾아낸다. 이때 각 개념노드는 휴지,대기,활성화의 3가지 상태를 가질 수 있다. 먼저 하나의 관계노드가 대기 상태로 놓여지고, 동시에 그것과 연결된 두개의 개념노드가 대기 상태에 놓여진다. 이 대기 상태에 놓여진 개념노드 중 하나가 자신과 부합되는 STR층의 개념노드를 탐색하여 발견되면 활성화 상태가 되고 그것과 연결된 관계노드를 모두 대기 상태로 놓는다. 대기 상태로 놓여진 관계노드는 자신과 연결된 개념노드 중 휴지 상태이 있는 것을 대기 상태로 전환한다. 이때 만약 연결된 두 개념노드가 모두 활성화 상태라면 자신과 대응하는 STR층의 관계노드와의 부합도를 계산하고 활성화 상태로 변환된다. 이렇게 하여 모든 관계노드가 활성화되거나, STR층의 노드를 모두 탐색하고 나면 관계노드의 활성화 값이 그 가중치와 곱해져서 개념노드에 전달된다. 값을 전달받은 개념노드는 자신의 부합도와 관계노드의 부합도를 합하여 출력값을 계산하고, 각 개념노드의 출력값이 가중합되어 하나의 자소 그래프의 출력값으로 정해진다. 출력값이 일정값이상인 경우에

해당 자소는 후보 자소 리스트에 첨가되어 SYL층으로 전달된다.

GRA층에서의 학습은 두가지 경우로 나누어 생각할 수 있다. 첫번째는 초기 학습의 경우와 새로운 패턴이 들어오는 경우로, 이때는 입력된 자소를 표현하는 STR층의 개념그래프를 새로운 기본 그래프로 추가한다. 두번째로 그 변형의 정도가 심하지 않아서 해당자소로 인식되기는 하였으나, 그 부합도가 낮은 경우에는 패턴을 표현하는 개념그래프를 기본 그래프가 포함할 수 있도록 수정함으로써 변형을 흡수할 수 있도록 하였다.

3.10 SYL 층

SYL층에서는 GRA층으로부터 추출된 후보 자소들을 조합하여 하나의 문자 후보를 생성한다. 이때 한글의 구조적인 지식이 사용된다. 우선 후보 자소 중 자음만을 선택하여 자음을 초성으로 하는 문자 후보를 생성한다. 그후에 중성, 종성의 순서로 조합을 맞춰 나가되, 그 자소간의 위치 관계를 조사하여 부적합한 것은 제외시킨다. 또 GRA층에서는 단모음과 단자음만을 인식하므로, 선택된 자음이 쌍자음이나 복자음을 형성할 수 있는 경우에는 필요한 자음을 더 탐색하기도 한다. 이것은 복모음의 경우도 마찬가지로 적용된다. 이렇게 하여 서로의 위치 관계가 맞는 자소열은 후보 문자로 등록되고 그 문자들은 다시 부합도를 계산한다. 부합도는 각각의 자소가 갖고 있는 부합도의 합에 자소의 개수로 나누어 계산한다. 그런데 이렇게 선택된 후보문자들 중에는 전체 입력중 일부만을 사용하여 만들어진 것들도 있고, 각 자소가 서로 같은 개념노드를 가지고 형성된 경우도 있다. 이런 경우를 제외시키기 위하여, 앞서 계산된 부합도에 3가지 값을 더하여 전체 부합도를 계산하는데, 위치부합도, 잡음 정도, 중첩 정도가 그것이다. 먼저 위치 부합도란 하나의 문자후보를 찾을 때 계산된 각 자소들간의 위치 관계에 대한 부합도를 말한다. 이 값은 초,중,종성에 따라 놓여져야할 위치가 다르므로 각각에 대한 계산식이 따로 존재해야 하는데, 이것은 경험적인 규칙으로 정하였다. 잡음 정도란 후보 문자가 전체 이미지의 일부분만을 표현하는 경우를 배제하기 위한 것으로, 후보 문자에 사용되지 않은 나머지 부분들의 크기를 계산한다. 마지막으로 중첩 정도란 두개의 자소가 같은 개념노드를 갖고 있는 경우를 고려한 것으로, 각 자소에 포함되는 개념노드를 조사하여 중첩되는 가지수 만큼의 값을 계산한다. 전체 부합도를 계산할 때는 자소들의 부합도, 위치관계 부합도, 잡음 정도, 중첩 정도의 4가지 값을 합해준다. 그 결과 부합도가 가장 큰 문자가 최종 인식 결과로 출력된다.

IV. 결과 및 토의

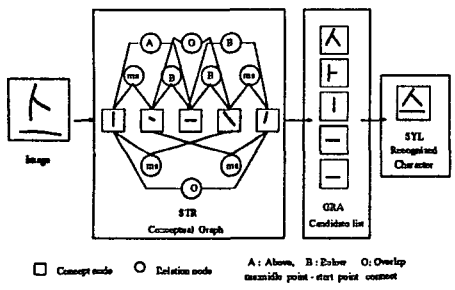

그림 8: 문자 '스'의 인식과정

(그림 8)에 한글 '스'에 대한 개념그래프와 GRA층에서 추출한 자소 후보 리스트, 그리고 SYL층의 출력이 나타나있다. 이 경우에 자음 'ㅅ'의 획들은 그 기울기가 4개중 하나로 정확히 분류되지 않아서 사선 성분 이외에도 수직/수평 성분도 함께 탐지되었다. 탐지된 5개의 개체들이 개념그래프의 개념노드가 되고 각각이 관계노드로 연결된다. (그림 8)에서 관계노드상의 값은 14개의 위치 관계를 나타내는 값 중 가장 큰 것을 표시한 것이다. 개념그래프가 형성되면 GRA층에서는 기본 개념그래프들과 비교하여 자소들을 탐지해 낸다. 이때는 자소의 위치와 무관하게 획들의 조합으로부터 가능한 모든 자소들을 탐지해낸다. (그림 8)에서는 모두 5개의 자소가 탐지되었다. SYL층에서는 탐지된 자소들로부터 가능한 모든 문자 조합을 찾는다. 탐지의 기준은 자소들의 위치 관계가 되므로, 초성과 중성의 위치 관계가 맞지 않는 것은 탐지 도중에 제외된다. 결국 조합 가능한 문자는 '스'가 되어 이것이 인식 결과가 된다. 이렇게 인식된 후에 GRA층의 개념그래프의 학습이 일어날 수 있다. 즉,'ㅅ'의 개념그래프는 초기 학습시에 두개의 사선 성분으로 구성된 형태를 갖고 있다. 그런데 이 경우에 두 사선 성분이 각각 수직/수평 성분과 중첩되어 있으므로 개념노드는 해당 기울기 성분도 학습하고, 이후부터는 수직/수평 성분이 주되게 발화하는 경우에도 'ㅅ'으로 인식할 수 있다.

본 논문에서 제안한 SEJONG-NET은 다음과 같은 특징을 갖고 있다. 첫째로 하위층에서 추출되는 특징을 동물의 시각 시스템에서 추출되는 특징들로 구현함으로써 다양한 패턴의 변화에 일관성있게 사용될 수 있다. 둘째로 이러한 특징 추출은 입력획을 하나의 기본획으로 근사화시킬 때 발생하는 오류를 줄이고, 상위층의 세포들이 자신과 부합되는 부분에 선택적으로 주의하여 변형된 패턴에서도 올바른 자소/문자를 탐지해 낼 수 있다. 셋째로 하나의 자소를 기본 특징과 조합 규칙으로 표현함으로써 변형된 자소 패턴에 대해서도 정의된 기본 특징과 조합 규칙을 사용하여 표현할 수 있고, 따라서 변형된 패턴에 대한 적응적인 학습이 가능하다. 넷째로 한글 자체내의 변형뿐 아니라 영문자, 한자, 숫자의 경우도 같은 원리로 적용될 수 있으므로 다양한 문자 집합으로의 확장이 가능하다. 이상과 같은 특징을 갖춘 SEJONG-NET은 기존 모델의 큰 문제점으로 간주된 경직성의 문제에 보다 궁극적인 해결책으로 여겨질 수 있다. 따라서 인지과학적인 실험 결과를 바탕으로 하여 적응적인 학습 능력을 갖춘 본 모형은 일반적인 문자 인식 문제에 적용될 수 있음을 보여주었다.

참고 문헌

[1] 이일병, "동적 문자인식 모형 사례연구," 한국정보과학회지, 제 9권 제 1호, pp.64-77, 1991.2.

[2] 이관용, 이일병,"한글의 구조적 지식을 이용한 동적 정보의 복원 및 SEJONG-NET에의 적용", 연세대학교 박사학위 논문, 1994.8

[3] 성태진, 방승양, "문자 조합 규칙 학습에 의한 온라인 한글 인식," 한국정보과학회 논문지, 제 20권 제 3호, pp.305-316, 93.3.

[4] 김백섭, "실시간 적응학습에 의한 온라인 한글 문자인식," 한국정보과학회 논문지, 제 21권 제 6호, pp.1048-1058, 1994.6.

[5] 정봉만, 권오석, 김태균, "온라인 입력 한글의 적응학습과 인식에 관한 연구," 한국정보과학회 논문지, 제 16권 제 5호, pp.487-497, 1989.9.

[6] Kandel, Schwartz, Jessell, PRICIPLES OF NEURAL SCIENCE, Elsevier, 3rd Ed.. 1991

[7] Audice G. Leventhal, The Neural Basis of Visual Function, Macmillan Press, 1991

[8] Versavel, Orban, Lagae, "Responses of Visual Cortical Neurons to Curved Stimulus and Chevrons," Vision Research, Vol.30, No.2, pp.235-248, 1990

[9] Dobbins, Zucker, Cynader, "Endstopped neurons in the visual cortex as a substrate for calculating curvature," Nature, Vol.329, No.1, pp.438-441, 1987

[10] John F.Sowa, CONCEPTUAL STRUCTURES: Information Processing in Mind and Machine, Addison-Wesley,1984