

상호활성화모형에 기반한 문자인식방법의 타당성 평가

김상기⁰, 박창수, 방송양, 김대진
포항공과대학교 컴퓨터공학과
{ihistory, bylkuse, sybang, dkim}@postech.ac.kr

Viability of Korean character recognition method based on IAM

Sang-Ki Kim⁰, Chang-Su Park, Sung-Yang Bang, Dae-Jin Kim
Department of Computer Science & Engineering, POSTECH

요약

여기서 우리는 한국어의 글자를 인식하기 위한 계산모형을 제시한다. 이 모형은 상호활성화모형에 기반하고 있으나, 입력처리부분이 개선되었다. 우리는 기존의 상호활성화모형에서 모든 입력이 규격화된 형태라는 가정을 배제하였다. 또한 우리는 성능향상을 위하여 다중해상도경로 학습을 결합시켰다. 우리는 제안된 모형을 구현하고, 그것을 실제 자료에 적용해 보았다. 우리는 모의실험결과를 통하여 제안된 모형이 실용적인 의미를 갖고 있음과 다중해상도 경로가 실제적으로 인식 성능 향상에 도움을 주고 있음을 확인할 수 있었다.

1. 서론

영어의 시각적 단어인식 모형으로서, 단어우월효과(Word Superiority Effect)를 설명하기 위하여 개발된 상호활성화모형(Interactive Activation Model)이 있다 [1]. 우리는 일찍이 이 모형을 기반으로 한글인식모형을 개발한 바 있다[3]. 이 모형은 한글의 글자우월효과(Character Superiority Effect)를 잘 설명할 뿐만 아니라, 실제적인 한글문자인식 시스템에 사용될 수 있는 가능성을 보여주었다. 이 논문에서는 이 모형을 실제 사용할 수 있는 문자인식 계산모형으로 만드는 것을 목적으로 한다.

이런 관점에서 이전의 모형은 규격화된 모양에서 특정위치의 선분의 존재 여부를 입력으로 한다는 점이 문제가 된다. 이를 해결하기 위하여 Neocognitron[4]을 사용하였다. 또한 사람의 시각 정보 처리에서 다중해상도 경로 학습이 있다[5]. 경로의 수나 해상도의 정도 등에 대해서 아직 통일된 견해는 없지만, 매우 설득력 있고 실용적인 가치가 있어 제안된 모델에 적용하였다.

2절에서는 제안된 모형의 기초가 될 한글용 상호활성화모형을 요약한다. 확장된 입력처리 부분에 관한 설명은 3절에서 한다. 4절에서는 제안된 모형을 써서 실행한 인쇄체 한글인식 초기실험에 대한 보고를 하고, 5절에서는 결론을 제시한다.

2. 제안된 방법

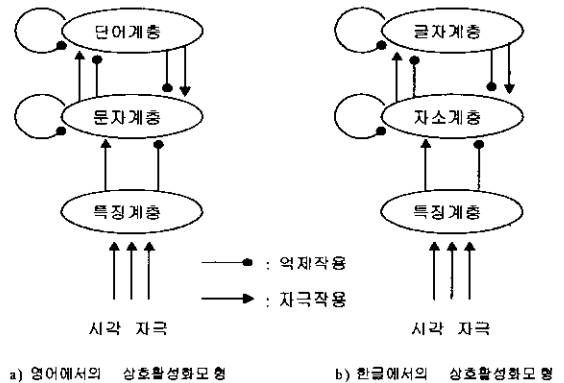


그림 1. 상호활성화모형의 기본 구조

영어에 적용된 상호활성화모형의 기본적인 구조는 그림 1의 a)와 같다. 특징계층은 입력되는 시각 자극을 감지하는 부분이고 문자계층은 단어의 특정 위치에서 각 알파벳의 활성화 정도를 나타내는 부분이며 단어계층은 이미 습득된 단어정보와 주어진 시각 자극에 의한 각 단어들의 활성화 정도를 나타내는 부분이다. 각 계층안에서는 여러 노드들이 있고, 각 계층의 노드들은 같은 계층의 다른 노드들 내지는 다른 계층의 노드들과 상호작용

을 한다. 노드사이의 상호작용에는 노드의 활성화 정도를 크게하는 자극작용과 반대로 활성화 정도를 감소시키는 억제작용이 있다.

이러한 상호작용이 반복적으로 이루어지면 활성화 값이 다른 노드보다 월등히 크게되는 노드가 나타나는데 그러한 노드에 해당하는 값으로 주어진 자극을 인지하게 된다. 그렇지만 영어와 한글은 그 구성이 다르기 때문에 영어의 문자계층과 단어계층을 한글의 자소계층과 글자계층으로 바꾸어 그림 1의 b)와 같이 하였다.

3. 입력처리의 확장

3.1 기본방향

앞서 설명한 상호활성화모형의 경우 입력을 규격화된 선분들의 유무를 이용하고 있으므로 실제적인 인식문제에 적용하기에 적합하지 않다. 이러한 문제를 해결하기 위해 신경망을 입력으로 이용한다.

또한 인식을 좀 더 효과적으로 수행하기 위해서 여러 다른 해상도에서 추출된 특징을 사용하기로 하였다. 저해상도의 특징은 입력의 전체적인 모형을 추출하는데 사용되고, 그 정보가 빠르게 전달된다. 반면, 고해상도의 특징은 자소를 정확하게 확인하기 위해서 입력에서 좀 더 자세한 정보를 추출하게 하는데 사용된다. 상호활성화모형의 입장에서 보면 저해상도 경로는 하향식처리에 도움을 주고, 고해상도 경로는 상향식처리에 도움을 준다고 할 수 있다.

제안된 모형에서 저해상도, 중해상도, 고해상도의 세 가지 경로를 사용한다. 다음에서 각각의 경로에 사용되는 특징과 구조에 대해서 설명한다.

3.2 저해상도 경로

저해상도 경로의 역할에 대해서는 여러가지로 생각해 볼 수 있으나 우선 글자의 대략적인 모양으로 인식의 대상이 되는 글자의 후보 범위를 축소시키는 것으로 하였다. 저해상도에서 추출된 특징을 이용하여 전체 글자를 미리 군집화해 봤다가, 입력이 들어오면 입력과 유사한 군집에 속한 글자들은 자극 신호를받게 된다.

저해상도 경로의 특징을 얻기위해서 입력영상을 겹침 없이 5x5의 윈도우로 나누고 해당 윈도우안에서 1의 값을 갖는 화소의 비율을 얻어낸다. 우리는 앞서 설명한 상호활성화모형의 글자계층에 포함되는 모든 글자를 자기조직화지도(SOM)을 이용해 군집화하였다.

3.3 중해상도 경로

중해상도 경로를 입력된 글자의 유형을 결정하는데

도움을 주기위해 사용하였다. 중해상도 경로의 특징을 추출하기 위해서 우리는 입력 영상에 겹침을 허용하면서 작은 크기의 윈도우를 씌우고, 각각의 윈도우로부터 하나의 특징을 추출하게 된다. 그리고 이 경로의 학습을 위해 다중계층퍼셉트론(MLP) 신경망을 사용한다.

3.4 고해상도 경로

고해상도 경로에서 얻어지는 정보는 글자의 각 위치에 해당하는 자소의 종류이다. 고해상도 경로는 실제적으로 글자의 각 유형에 대응하는 6개의 분리된 경로로 이루어지고, 이들은 병렬적으로 동작한다. 각각의 경로는 입력이 자신에 해당하는 유형이라고 가정하고 입력을 처리한다. 또한 한글에서 각 글자는 유형별로 2~4개의 자소로 이루어지므로, 각 유형의 경로내에서도 2~4개의 자소인식경로가 존재한다. 결국 고해상도 경로내에는 병렬적으로 작동하는 17개의 각기 다른 인식이 존재하는 것이다.

앞서 언급했듯이 우리는 각각의 자소 인식기에 Neocognitron형태의 피드포워드 신경망을 사용한다. 각 자소인식기는 두 개의 은닉계층으로 구성되고, 각 계층은 단순세포층과 복합세포층의 쌍으로 구성된다. 단순세포층의 노드는 주어진 수용영역안에서 미리 정해진 특징을 감지한다. 그리고 복합세포층의 노드는 그 전 계층의 수용영역안에서 하나의 노드라도 발화하면 자신도 발화함으로써 어느 정도의 왜곡을 흡수하게 된다.

3.5 모형의 동작

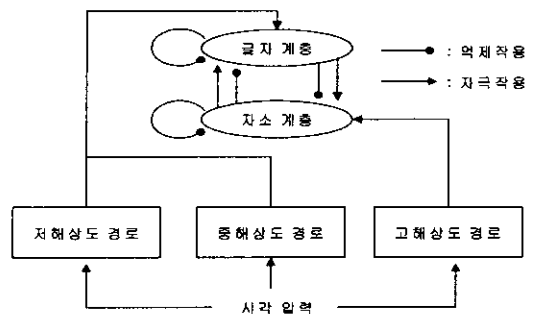


그림 2. 모형의 전체적인 구조

제안된 모형의 전체적인 구조가 그림 2에 있다. 앞서 설명한대로 세 개의 다른 경로로부터의 출력신호를 통합하기 위한 기본적인 구조는 상호활성화모형이다. 원래의 상호활성화모형에서 입력으로부터 특징을 감지하는 특징계층을 고해상도 경로가 대신하게 된다. 저해상도 경로와 중해상도 경로로부터의 출력은 하향식처리와 관련되기 때문에 문자계층으로 바로 입력된다.

4. 모의 실험

제안된 모델의 검증을 위해 여러가지의 실험을 해보았다. 여기서는 두가지 실험결과를 보이겠다. 첫 번째 실험은 한글 상용글자 2,350자를 인쇄체 폰트에 대해 모델을 적용한 실험이고 두 번째 실험은 상호활성화가 모델에 기여하는 정도를 알기위해 수행한 실험이다.

표 1. 모의실험1의 결과

| 사용된 경로 | 훈련자료 | 평가자료 |
|-------------|-------|-------|
| 고해상도 경로 | 97.8% | 93.4% |
| 저·고해상도 경로 | 98.6% | 95.1% |
| 중·고해상도 경로 | 98.9% | 98.0% |
| 저·중·고해상도 경로 | 99.2% | 98.3% |

첫 번째 실험에서 각 글자는 30x30크기의 이진 화소 영상으로 5%의 무작위 잡음을 가하였다. 저해상도 경로에서는 겹침이 없는 5x5 윈도우를 이용 6x6의 실수를 입력으로 하고, 중해상도 경로는 겹침이 있는 5x5 윈도우를 이용 10x10의 실수를 입력으로 한다. 실험결과는 표1과 같다. 표에서 보듯이 저해상도와 중해상도 경로에서의 출력은 시스템의 전체적인 성능을 향상시키는데 도움을 주고 있다. 또한 평가자료의 경우, 저해상도와 중해상도 경로의 역할이 더 크게 나타나는 것을 확인할 수 있다.

두 번째 실험에서는 중해상도와 고해상도 두 경로에 대해서 상호활성화 모델을 적용한 경우와 가중 합산(weighted sum)을 한 경우를 비교하였다. 그림 3에서 보는 바와 같이 상호활성화 모델을 적용한 경우가 월등히 높은 결과를 보이고 있다. 이는 상호활성화 모델이 인식 성능 향상에 큰 도움을 주는 것을 나타낸다.

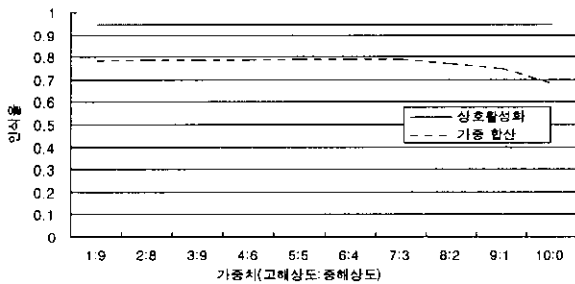


그림 3. 상호활성화 모델과 가중 합산 비교

5. 결론

여기서 제안된 모형은 인지모형에 기반한 실제적인 문자인식시스템의 개발을 위한 시도이다. 그러기 위해서

우리는 한글형 상호활성화모형에서 출발하였다. 그리고 규격화된 형태가 아닌 실제 영상을 입력으로 받아들이기 위해 입력부분을 개선하였다. 더 나아가 인간 시각 처리의 다중해상도 경로 학습의 잇점을 취하기 위하여 세가지 다른 해상도 경로를 결합하였다. 각 경로는 여러가지 신경망 구조로 구현되었다.

제안된 모형의 성능과 가능성을 검증하기 위하여 실제적인 한글자료를 이용하여 모의실험을 행하였다. 모의 실험을 통하여 우리는 시스템이 예상했던대로 동작함을 확인했다. 우리는 저·중해상도 경로가 고해상도 경로를 보충해주면서 시스템의 전체적인 성능을 향상시키는데 도움을 준다는 것을 확인하였다.

6. 참고문헌

- [1] McClelland J. L. & Rumelhart D.E., An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account of Basic Findings, Psychological Review, VOL.88, No. 5, pp. 375-407, 1981.
- [2] Glyn W. Humphreys & Vicki Bruce, Visual Cognition:Computational, Experimental, and Neuropsychological Perspectives, Lawrence Erlbaum Associates Ltd., Publishers, pp. 241-247, 1989.
- [3] Park C. S. & Bang S. Y., Modeling Character Superiority Effect in Korean Characters by Using IAM, Proceedings of Biologically Motivated Computer Vision, Vol. 1, pp. 316-325, 2000.
- [4] Kunihiko Fukushima, Sei Miyake & Takayuki Ito, Neocognitron:A Neural Network Model for a Mechanism of Visual Pattern Recognition, IEEE Transaction on systems, man and cybernetics, Vol. 13, No. 5, pp. 826-834, 1983.
- [5] Toshio Inui, A Model of Human Visual Memory:Data Compression with Multi-resolution, Scandinavian Conference on Image Analysis, Vol. 6, pp. 325-332, 1989.
- [6] Kathrayn T. Spoehr & Stephen W. Lehmkuhle, Visual Information Processing, W. H. Freeman and Company, 1982.
- [7] 김 재갑, 한글 글자 맥락에서의 자모 지각, 학위논문(박사), 서울대학교, 1994.