

컴퓨터 비전을 위한 인공지능 모델의 실현에 관하여

李 秀 東

蔚山大學校 電氣 및 電子工學科 副教授

I. 서 론

지난 4 반세기동안 컴퓨터비전에 관한 연구는 대학이 주가 된 소수의 A.I 연구소에 의하여 소규모로 그 연구가 진행되어 왔다. 임의의 복합이미지를 분석하고 해석하여 하나의 대상으로 인지할 수 있는 효율적인 컴퓨터비전시스템(즉 general-purpose system)의 개발은 그 방대한 정보량 또는 프로세싱시간의 제약등으로 현재로도 극히 어려운 일로 간주되고 있다. 따라서 주로 특정한 대상을 염두에 둔 컴퓨터비전시스템(즉 goal-oriented system)의 개발이 작금의 주 연구과제가 된다고 할 수 있겠다.

그러나 이 분야의 연구수행을 위해서는 인식심리학(cognitive psychology), 패턴인식, 이미지이해(image understanding or scene analysis), 컴퓨터 하드웨어 및 소프트웨어, 컴퓨터그래픽스, 전기 전자공학, 신경생리학(neuro-physiology) 등의 다양한 학문영역을 인공지능(artificial intelligence)이란 한 범주속에 흡수시켜 결과 시스템을 구성해야하므로, 그 목적을 향한 접근 방법 또한 다양해질 수 밖에 없다고 하겠다.

그러나 어떤 방법으로 목적하는 시스템을 구성하였던, 이 시스템모델이 가져야 하는 구조적 특성, 내부의 특성 또는 이러한 모델이 프로세싱과정을 통하여 자연발생적으로 나타내는 창발적 특성(emergent properties)은 인간의 이미지프로세싱의 과정과 흡사한 기능을 보일 때, 융통성있는 인공지능모델이 된다고 할 수 있다. 여기서 창발적 특성이라 함은 시스템 고유의 활동이 아니나, 각 국부 프로세스들간의 또 다른 조합(적당한 양의 피드백을 포함한)에 의해 예기하지 않았던 다른 종류의 활동을 시스템이 나타내는 것을 의미한다.

본 논문에서는 패턴인식과 이미지이해를 위해 컴퓨터비전시스템의 구성시 고려하여야 할 특징들과 또 이 특징들을 수용할 수 있는 시스템모델의 구조 및 제어등

에 관한 제언을 하고자 한다.

II. 시스템모델의 구조

제언 1 : "시스템모델은 이미지이해기법(즉 segmentation followed by interpretation)과 패턴인식기법(즉 direct recognition)을 동적으로 결합시킬 수 있는 구조를 가져야 한다."

이미지이해는 어떤 미지의 입력 이미지를 받아들여, 이 이미지를 의미를 가질 수 있는 최소의 부분패턴(primitive pattern)으로 분할(segmentation)하고, 이렇게 분할된 부분패턴을 심볼형태등으로 서술될 수 있게 변환시켜 description을 얻고, 이들 description 사이의 구조관계(syntactic relation)를 해석하여, 예를 들면 '구름이 떠도는 하늘아래 소가 풀을 뜯고 있다.'는 등의 이미지 해석이 가능하도록 한다.

그러나 이러한 이미지이해기법(즉 segmentation followed by interpretation)은 모든 이미지의 조각들을 심볼로 변환·처리하게되므로, 이미지 자체가 가지는 특질(어떤 경우에는 중요한 변수가 될 수도 있는)을 간과하게 되어, 전반적인 이미지이해에 있어 오류를 범할 수 있다는 약점이 따르게 된다.

Fishler(1979)는 이 점에 관해 많은 문제들은 꼭 심볼형태로 변환시킬 수 있다고 보장될 수 없다. 또한 우리가 가지는 시각세계(visual world)에 관한 충분한 이해의 부족, 심볼형태를 처리할 수 있는 규칙적인 기법의 부재등으로 인해 대개의 비전은 현상적인 실험과 유사한 방법으로 처리할 수 밖에 없다. 따라서 자연이미지(natural image)의 처리에 있어서 이미지자체정보의 도입은 시스템프로세싱의 전과정을 통하여 필수불가결하다고 할 수 밖에 없다. 그리고 제한하였다!¹⁾

따라서 이미지 자체 공간정보개념(즉 spatial properties)을 잃지 않은 상태에서 전반적인 시스템프로세싱이 이루어져야 하며, 여기에 패턴인식기법(즉 direct

recognition)의 도입이 필연적으로 요청된다. 즉 패턴 인식기법은 어떤 미지의 입력 이미지를 있는 그대로 인식하고자 하여, 이미지에 있어서 적은 정보량까지도 포함시키는 픽셀(pixel-picture elements) 레벨의 프로세싱을 가능하게 하기 때문이다.

그러나 부분패턴들의 무수한 조합으로 구성되는 방대한 수의 복합이미지를 각개로 인식하기에는 엄청난 메모리가 소요되고 또한 처리시간에 따른 병렬연산처리기법등이 요구되어, 현실적으로는 많은 제약을 받을 수 밖에 없다.

그러므로 복합이미지를 기본패턴으로 분할시켜 프로세싱을 수행하는 이미지해기법과 기본패턴의 공간정보를 그대로 유지시키며 프로세싱을 수행하는 패턴인식기법 양자를 동적으로 결합시켜 수용할 수 있는 시스템모델이 구성되어야 한다고 하겠으며, 이 양 기법은 서로간의 대화를 통해 프로세싱을 수행할 수 있어야 한다고 하겠다. 즉 feed-forward 방식이 아닌 feed-back의 구조를 가진 이중레벨(two-level)의 시스템구조가 요청되어 진다.

이 이중레벨의 시스템구조중에서 저차층 즉 이미지 해석부에서는 이미지를 의미를 가지는 부분패턴으로 분할시키는 역할을 수행하게 하고, 고차층 즉 패턴인식부에서는 저차층에서 분할한 부분패턴들을 순차적으로 인식하여, 이들간의 관계를 지위층으로 결론을 내리게 한다. 또한 고차층의 결론을 항상 저차층의 결론을 항상 저차층으로 피드백시켜서 저차층의 프로세스가 올바르게 진행되게끔 감시 또는 유도해 주며, 미흡하다고 느낄 시는 저차층의 프로세스 자체를 교정해줄 수 있어야 한다. 즉 시스템모델은 straight-forward방식이 아닌 동적(dynamic)구조를 가져야 한다.

Ⅲ. 시스템모델과 데이터베이스

지식의 표현은 “인공지능” 시스템의 구성에 있어서 가장 중요한 문제중의 하나이다. 일반적으로 데이터베이스 모델은 대상(object), 대상사이의 관계(relation), 제한 조건(constraints), 상태(states), 상태를 천이시키는 행동(action), 행동을 일으키게 하는 에이전트(agents), 그리고 시스템에 의해 이해되어야 할 목적 description등이 포함된다. 본 논문에서는 컴퓨터비전시스템을 위해 데이터베이스를 지식데이터베이스(knowledge-based database), 프레임데이터베이스(frame-based database), 결과데이터베이스(result database)로 분류하고 각기 이들 데이터베이스가 갖추어야 할 특질을 다음과 같이 살펴 보았다.

제언 2 : “지식데이터베이스는 주어진 이미지를 기본 패턴들로 분할시키는데 필요한 문제영역 주변환경이 되는 a priori knowledge와 문제를 해결하기위한 heuristic information으로 구성한다.

A priori knowledge는 제기된 문제의 이미지영역에서 도출가능한 기본패턴들과 이들 사이의 구조적 관계등을 포함하며, heuristic information은 이미지를 분할시키고자 할 때시작이 될 수 있는 실마리(clue)점들, 분할시 적용해야 할 행동(action)들, 그리고 이들 행동들중의 어느 하나를 선택할 수 있는 hypothesis generation scheme등을 포함시켜야 한다.

실마리점들은 이미지에서 여러 edge들이 만나는 점중에서 추출가능하며, 행동들은 일반적으로 제한을 가한 production system으로 서술할 수 있으며, 이미지의 분할시작점, 분할방향, 분할후의 이미지 상태의 변경, 구조변경, 예외처리등의 기법이 일련의 행동을 일으킬 수 있게끔 구성되어야 한다. 또한 실제 적용되고 있는 행동은 이미지로부터 의미있는 기본패턴을 추출해야 하나, 어느 경우에는 이 행동이 목적과는 위배되는 방향으로 진행될 수도 있다. 즉 위배(violation)가 발생되었을 시는 지금까지 취해왔던 행동은 즉시 중단시키고 다른 행동을 유발시키게 할 수 있는 hypothesis generation scheme이 요청되며 이러한 scheme에 의해 시스템모델은 추론 능력(inferential power)을 갖추게 된다. 이때 추론을 위한 지식이 hypothesis generation scheme에 포함되어야 함은 물론이다. 즉 이미지해석에 있어 heuristic information이 간여하여 수많은 조합이 가능한 프로세스를 제한시켜 빨리 결론에 도달할 수 있게 한다. 이때 heuristic information의 양이 많을수록 프로세스는 줄어들게 되므로 적절한 타협이 이루어져야 할 것이다.

그리고 a priori knowledge는 나중에 기술할 패턴인식에 관계하게 된다.

제언 3 : “프레임데이터베이스는 문제영역에서 추출가능한 이미지프레임(즉 image-specific data)들을 수용하며, 외부교육에 의해 학습형태로 구성한다.

프레임데이터베이스는 인간의 두뇌중 long-term memory에 해당되며, 이러한 데이터베이스는 외부의 교육(training)에 의해 점진적으로 학습받을 수 있는 구조를 가져야 하며(즉 learning machine), 차후의 이용시에는 융통성을 발휘해야 한다. 여기서 융통성이란 이 데이터베이스의 입력이 내부상태와 조금 다르더라도 지

능을 가지고(주변여건에 비추어) 대처하여 대상을 판별할 수 있는 능력을 의미한다.

실제적으로는 이미지 자체의 정보는 패턴인식과 밀접한 관계를 가지므로 패턴인식부 내부에 구축하는 것이 좋을 것이다. 그리고 이때 데이터베이스는 주의집중(attention)기능과 망각(forgetting)기능을 아울러 가져야 하며 이러한 추가기능들에 의하여 소위 말하는 군중속의 시끄러움중에서도 자기의 이름만을 경청하여 귀 기울일 수 있고, 사람의 얼굴중 코끝만을 바라볼 수도 있게 되는 것이다.

이미지의 교육은 문제영역에서 나타날 수 있는 가능한 이미지 부분패턴으로부터 프로토타입을 추출하고 그의 변형을 함께 교육함으로 이루어 지며, 이 방법은 차후의 이용시 시스템이 융통성을 발휘하게 하는 원인이 된다.

제언 4 : “결과데이터베이스는 시스템 프로세싱의 전 과정을 통하여 발생되어지는 모든 중간 결과들로 구성된다.

이미지해석부와 패턴인식부의 전 프로세싱 과정중에 발생되어 지는 중간 결과들은 직접 또는 간접적으로 다음의 프로세스에 영향을 끼치게 된다. 즉 이들 중간 결과들을 분석하고 다음 프로세스에 적용시킴으로써 현재 까지 진행되어 오고있는 프로세스의 당위성을 판단할 수 있으며 또한 다음 프로세스를 목적방향으로 유도시킬 수 있는 예측하에 프로세스를 구동시킬 수도 있게 된다.

그리고 이미지해석부에서 분할한 기본패턴들의 공간 정보(spatial information)도 결과 데이터베이스에 보관시켜, 시스템 프로세싱의 전과정을 통해 국부이미지 자체의 개념을 유지시켜가며 패턴인식을 할 수 있게 한다. 또한 결과데이터베이스에 발생되어진 데이터의 시퀀스를 관찰해 봄으로, 시스템모델의 동작특성(behavioural properties)을 해석할 수 있으며, 시스템모델이 나타낼 수 있는 가능한 창발적특성(emergent properties)들을 찾아낼 수 있게 된다.

이상의 세가지 범주의 데이터베이스는 이미지해석과 패턴인식을 원활히 수행하는데 직접 또는 간접으로 간여하여 각 프로세스가 올바르게 진행되게끔 하고, 시스템모델이 지능을 가지고 동작하게 하며 또한 이 지능이 점점 개발되어가는 방향으로 동작할 수 있게 한다. 즉 “인공지능”에 의한 이미지해석과 패턴인식이 이루어지게 됨으로 시스템모델은 융통성 있는 판단능력과 창발적특성을 발휘할 수 있게 된다.

IV. 시스템모델의 제어

제어 5 : “시스템 프로세스의 제어는 고차층의 결론을 저차층의 프로세스에 간여시킴으로 저차층의 프로세스를 유도, 수정, 또는 취소시킬 수 있게 해야 하며, 전체 프로세싱시퀀스를 결정지을 수 있게 해야 한다.”

인간의 이미지 프로세싱에 있어서 전체제어는 결코 국부의 프로세싱에 전권을 위임하지 않는다. 즉 중추부는 국부적인 프로세스의 진행과정을 끊임없이 감시해가며 때로는 프로세스의 조건들을 고침으로 그 목적을 변경시키기도 하며 또 어떤 경우에는 그 프로세스가 종료되기도 전에 중단시키기도 한다.

시스템의 제어에 있어 고려해야 할 사항은 프로세스 모듈들간에 얼마나 효과적으로 융통성있는 대화를 하게 할 수 있는가, 결론 도출시 그 결론은 얼마나 문제영역 주변의 heuristic information을 잘 활용하여 이끌어 내어지게 할 수 있는가, 또한 만들어진 결론을 얼마나 적절한 타이밍으로 보류하기도 하고, 다음 프로세스에 전달하게 할 수도 있는가 하는 점들이다.¹¹⁾

시스템제어는 시스템이 받아들인 입력 이미지로부터 실마리점을 추출하고, 추출된 실마리점과 지식데이터베이스내의 heuristic information을 토대로 적용시킬 프로세스를 찾는데서부터 시작이 된다. 그리고 이러한 hypothesis의 발생에는 default 값, 프로세스의 tendency, priority등과 앞단의 결론등이 관계된다. 만약 적용시킨 hypothesis가 전체 시스템의 동작에 위배를 초래할 경우에는 다음 hypothesis가 hypothesis generation scheme에 의해 자동적으로 발생되어져야 하고 전에 발생되었던 hypothesis는 폐기되어져야 한다.

그러나 지금까지의 시스템제어는 단지 국부적인 프로세스의 구동들만에 관계되어, 시스템 전체 동작제어에는 직접적인 영향을 끼칠 수 없다. 그러므로 국부 프로세스들의 조합에 의해 특정지워지는 전체동작의 제어를 위해서는 고차층의 결론이 저차층의 각 프로세스를 유도하여 올바른 시스템 동작특성을 나타내도록 하여야 한다. 여기에 결론피드백루프(decision feedback loop)를 도입하여 출력상태가 다시 입력상태로 변환되는 동적인 오토마타 특성을 보일 수 있게 한다.

즉 시스템의 동작은 인간의 예에서 관찰할 수 있는 바와 같이 프로세스가 진행됨에 따라 자체의 지능에 의거하여 자신의 동작을 교정해 나가는 능력을 가질 수 있도록 제어되어져야 하는 것이다.

V. 패턴인식기의 구조

제언 6 : “패턴인식기는 픽셀(pixel) 레벨에서 랜덤 프

로새성이 가능한 학습기(learning machine) 형태의 구조를 가져야 하며, 학습받은 내용은 차후 인식시 융통성있게 이용되도록 구성되어야 한다.”

패턴인식에서 학습(training)이란 문제영역에서의 기본패턴들을 시스템 내부에 미리 구성해 두는 것을 의미한다. 학습기(learning machine)의 구현을 위해서는 문제영역에서 관측한 이미지들로부터 한 크라스의 기본패턴을 대별할 수 있는 특성을 갖춘, 즉 일반화시킨, 프로토타이프 패턴을 추출하고 이들을 교육시켜 시스템의 내부가 이미지 프레임을 형성하도록 해야 한다. 시스템 구조 또한 이들 이미지 프레임을 수행할 수 있는 구조를 가져야 하며 한번 수용한 패턴은 외부에서 망각기능을 부여하기 전까지는 시스템 프로세싱의 전 과정을 통하여 유지되어야 한다.

한편 미지의 입력패턴을 인식할 시는 교육에 의해 구성된 내부상태와 외부상태를 끊임없이 비교해가며, 변형을 제거시켜 감으로 하나의 결론을 내리게 하는 방법을 채택함으로써 시스템에서 인식은 물론 preprocessing의 기능까지도 수행할 수 있게 한다.

이미지의 인식에 있어 prototype hypothesis에 따르면, 인간은 이미지를 주어진 그대로 기억한다기보다 이들 이미지의 변형된 프로토타이프를 random mapping에 의해 기억하고, 인식시에는 외부의 대상과 이렇게 구축한 프로토타이프를 비교해가며 자신이 만족할 때에야 비로소 결론을 내린다.^[1]

여기에 패턴인식기도 결론을 입력측에 피드백시켜서 하나의 대상을 만족할 때까지 재확인(confirm)인식하게 하는 구조를 가지게 해야 한다는 것을 알 수 있다. 또한 이러한 피드백루프는 패턴인식기 자체에도

창발적특성을 부여해 주어서, 글자의 인식만을 목표로 개발된 시스템이 단어의 인식은 물론 나아가서는 자연언어의 인식까지도 가능하게 하는 것이다.^[1]

VI. 결 론

이상에서 살펴본 바와 같이 컴퓨터 비전을 위한 “인공지능” 모델의 개발을 위해서는 시스템의 데이터구조, 시스템의 내부상태와 프로세스등이 동적으로 결합되어야 하며, 또한 동적인 제어가 이루어져야 한다고 하겠다.

이러한 모든 조건과 제언들을 만족하는 시스템 모델은 인간의 인식심리와 유사한 기능을 발휘할 것이며, 인간의 패턴인식기능을 조감해 볼 수 있는 하나의 가능성을 제시할 것이라고 생각한다.

參 考 文 獻

- [1] Fishler, M.A., “On the Representation of Natural Scenes,” *Computer Vision Systems*, Hanson, A.R. and Riseman, E.M. (Eds.), Academic Press, pp. 47-52, 1978.
- [2] Pylyshyn, Z.W., “Complexity and the Study of Artificial Intelligence,” *Philosophical Perspectives in Artificial Intelligence*, Ringle, M.(Eds.), The Harvest Press, 1979.
- [3] Klatzky, R.L., “Human Memory: Structures and Processes,” W.H. Freeman and Company, 1975.
- [4] Lee, S.D., “Dynamics of Feedback for Artificially Intelligent Pattern Recognition,” MSc. Dissertation, Brunel Univ., England, 1981. *

♣ 用 語 解 說 ♣

simplex / duplex modem

Modem은 그 종류에 따라 3가지의 다른 형태로 동작 될 수 있다 : (1) simplex에서는 정보가 한 방향으로만 전달되고 ; (2) half-duplex에서는 어느 시간에서는 정보가 한 방향으로만 전달되나, 시간에 따라 그 전달 방향이 바뀔 수 있고 ; (3) full-duplex에서는 정보가 두 방향으로 동시에 전달 될 수 있다.