

다양한 실내 환경변수로부터 강인한 객체 검출

최미영*, 김계영**, 최형일***

Robust Object Detection from Indoor Environmental Factors

Mi-young Choi*, Gye-young Kim**, Hyung-il Choi***

요약

본 논문에서는 다양한 환경변수가 존재하는 실내에서 효율적인 객체를 검출하기 위한 방법을 제안한다. 일반적으로 실내 환경은 조명의 변화와 객체에 의해 발생된 그림자, 바닥면에 반사된 조명성분 등으로 인하여 정확한 객체 검출이 이루어지기 어려운 환경이다. 먼저 객체검출을 위한 배경영상을 생성한다. 영상 내에 객체가 존재하는 경우 이전에 생성된 배경영상과 현재 입력영상간의 유사도 비교를 통해 보정된 배경영상을 생성한다. 배경영상과 입력영상으로 생성한 평균영상과 보정된 배경영상을 이용하여 혼합영상을 생성한다. 마지막으로 혼합영상을 이용하여 입력된 영상으로부터 객체를 검출한다. 검출된 객체를 보완하기 위해 레이블링 과정을 통해 잡음 성분을 제거한 후 모폴로지 기법을 적용하여 객체영역 보완한다. 따라서 조명의 변화나 그림자와 같은 환경변수로부터 강인한 객체를 검출한다. 본 논문에서 제안한 시스템은 변형된 조명성분과 그림자 성분이 포함되어 있는 혼합영상을 사용하기 때문에 기존시스템보다 객체영역 검출이 더욱 효과적이다.

Abstract

In this paper, we propose a detection method of reduced computational complexity aimed at separating the moving objects from the background in a generic video sequence. In generally, indoor environments, it is difficult to accurately detect the object because environmental factors, such as lighting changes, shadows, reflections on the floor. First, the background image to detect an object is created. If an object exists in video, on a previously created background images for similarity comparison between the current input image and to detect objects through several operations to generate a mixture image. Mixed-use video and video inputs to detect objects. To complement the objects detected through the labeling process to remove noise components and then apply the technique of morphology complements the object area. Environment variable such as, lighting changes and shadows, to the strength of the object is detected. In this paper, we proposed that environmental factors, such as lighting changes, shadows, reflections on the floor, including the system uses mixture images. Therefore, the existing system more effectively than the object region is detected.

▶ Keyword : 객체 검출(object detection), 배경영상 모델링(background image modeling), 혼합영상(mixture image)

• 제1저자 : 최미영

• 투고일 : 2010. 01. 18, 심사일 : 2010. 01. 28, 게재확정일 : 2010. 02. 22.

* 송실대학교 BK21 전임연구원 ** 송실대학교 IT대학 컴퓨터학과 부교수 *** 송실대학교 IT대학 미디어학과 교수

※ 본 연구는 송실대학교 BK21 디지털영상산학공동사업단에서 지원을 받았음.

※ 본 연구는 송실대학교 교내연구비 지원으로 이루어졌음.

I. 서론

최근 객체 검출 및 인식에 관한 연구는 무인 감시, 보안 출입 관리, 공장 자동화, 지능형 교통 시스템 등을 목표로 컴퓨터 시각 분야 뿐만 아니라, 인간과 컴퓨터간의 상호작용 등 여러 응용 분야에서 활발히 진행되고 있으며 관심 또한 점차 증대되고 있다. 컴퓨터 비전 기술을 이용하여 다양한 환경에서 실시간 객체 검출 및 추적하는 일은 매우 어려운 일이다. 그럼에도 불구하고 컴퓨터 성능 향상과 영상처리 기법들의 발전으로 인하여 현재까지도 많은 연구가 활발히 진행되고 있으며, 특히 실내 환경에서 고정된 카메라를 사용한 객체 검출 및 추적에 관한 연구가 활발하게 진행되고 있다. 주로 실내 환경에서 고정된 카메라를 사용하여 객체를 검출하고 추적하는 가장 큰 목적은 영상 감시를 이용한 외부인의 침입을 조기에 발견하고 대처하기 위함이다. 검출하고자 하는 움직이는 객체는 사람에게 해당된다.

기존의 연구를 살펴보면, 객체를 검출하기 위한 방법으로 차영상[1][2][3][4][5] 및 배경영상[6][7][8][9]을 생성하여 객체 검출에 이용하기도 하며, 객체의 대상이 사람일 경우 컬러 성분 [10][11]을 사용하기도 한다. 그러나 위에서 언급한 방법들의 경우, 실내 외 환경변수에 의해 여러 가지 문제점이 발생된다.

본 논문에서는 실내 환경에서 발생할 수 있는 다양한 환경 변수에서 신속하고 정확하게 객체를 검출하기 위한 시스템을 제안한다. 제안된 방법은 배경영상과 입력영상에 의해 생성된 혼합영상을 이용하여 객체를 검출한다.

II. 관련 연구

객체를 검출하기 위한 기존의 다양한 연구 방법 중에서 몇몇 대표적인 방식들을 살펴보자.

먼저 차영상[1][2][3][4][5]을 이용하는 방법은 모션을 추출하는 가장 기본적인 방법으로, 두 인접한 프레임간의 각 픽셀의 차이를 이용하는 것이다. 픽셀간의 차이로 움직임을 검출하기 때문에 조명의 변화가 생기게 되면 같은 픽셀이라도 실제로는 다른 영상 값을 갖게 되며, 객체의 움직임이 느린 경우 정확한 객체를 검출하는 것이 매우 어렵다. 또한 이러한 오류를 막기 위한 후처리 과정은 연산 처리시간이 길어지게 된다.

컬러 성분을 이용한 방법 중 대표적으로 스킨 컬러[10][11]를 사용하여 객체에 해당하는 사람을 검출하는 경우도 조명의 변화에 따라 검출율의 현격한 차이를 나타낸다.

또한, 고정된 카메라를 이용하여 객체를 검출하는 경우 많이 사용되어지는 방법으로는 배경영상을 생성하여 객체를 검출하는 방식이다[6][7][8][9]. 대부분의 경우, 일반적으로 배경영상을 일정한 시간 간격에 의해 갱신하며 입력영상과의 차에 의해 객체를 검출한다. 그러나, 다양한 외부요인으로 인하여 정확한 객체 검출이 어려운 것이 사실이다.

그 예로 첫째, 실내 환경의 경우 객체의 등장으로 인하여 영상의 전체적인 밝기 값이 변화하며 이에 의해 배경영상과의 차이가 발생한다. 둘째, 객체가 움직일 때마다 객체에 의해 생성된 그림자 성분 때문에 객체 검출에 어려움이 발생한다. 셋째, 바닥에 의해 반사되어진 조명성분의 변화로 정확한 객체 검출이 어렵다는 단점이 있다.

III. 본론

본 논문은 실내 환경에서 발생하는 다양한 환경변수를 고려하여 정확한 객체를 검출하는 방법을 제안한다. 실시간으로 사람이나 움직이는 객체를 검출하는데 있어 정확도에 가장 영향을 미치는 요인은 조명 환경의 변화이다. 외부 조명환경이나 그림자에 의한 오검출로부터 사람 또는 객체를 정확하고 신속하게 검출하는 것을 목적으로 한다.

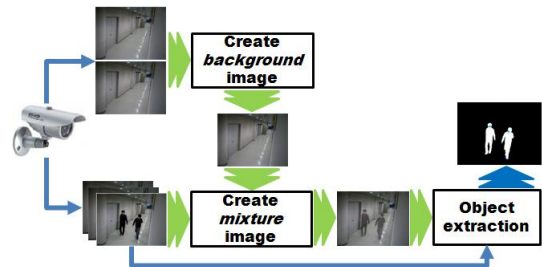


그림 1. 시스템 흐름도
Fig. 1. The system flow

본 논문에서 제안하는 시스템 흐름도는 그림 1과 같다. 먼저 CCD 카메라에 의해 획득된 객체가 포함되지 않은 영상으로부터 배경영상을 모델링한다. 이후 모델링된 기준 배경영상과 현재 입력영상 간의 유사도 비고를 통하여 생성된 혼합영상을 이용하여 객체를 검출한다. 검출된 객체의 수직 프로파일을 이용하여 관심영역을 지정한다. 이후 배경영상과 현재 입력영상과의 차영상에 의해 검출된 객체영역을 이용하여 해당 관심영역 내에 존재하는 객체영역을 보완함으로써 다양한 실내 환경변수의 영향을 제거한 강인한 객체를 검출한다.

3.1 배경영상 모델링

일반적으로 시스템에서 받아들이는 배경영상은 시간이 지남에 따라 다양한 조명의 변화가 발생하기 때문에 고정적인 배경영상을 사용하게 되면 조명성분의 영향으로 인해 배경영역이 객체로 추출되는 오류가 발생된다. 따라서 실시간으로 배경을 갱신하여야 하며 그 방법은 식 1에 의해 생성된다.

$$BG_{rgb}(x, y) = \frac{I_{rgb}(x, y) * w_1 + BC_{rgb}(x, y) * w_2}{2} \quad (1)$$

여기서, $BG_{rgb}(x, y)$ 는 갱신된 배경영상을 의미하며, $I_{rgb}(x, y)$ 는 현재 프레임의 $BC_{rgb}(x, y)$ 는 현재 배경영상을 나타내며, 각 영상에 대한 가중치인 w_1 , w_2 는 각각 0.4와 0.6을 적용하여 배경영상을 생성하였다. 그림 2는 배경영상 생성 과정을 나타내고 있다.

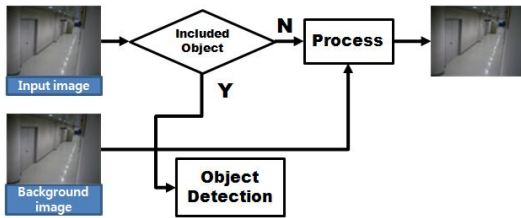


그림 2. 배경영상을 생성하는 알고리즘
Fig. 2. The algorithm that generates the background image

배경영상 생성 과정은 먼저 객체가 배경으로 인식되는 오류를 피하기 위해 입력된 영상에 객체가 존재하는지 여부를 판단한다. 입력영상과 배경영상의 차영상을 생성한 후 잡음제거 과정을 거쳐 움직임 벡터의 존재 유무에 따라 판단한다. 배경영상 생성은 일정한 시간 간격에 의해 계속 갱신된다.

3.2 혼합영상 생성 및 객체 검출

일반적으로 배경영상을 모델링하여 이용하는 방식들은 배경영상과 입력영상간의 차영상에 의해 객체를 검출한다. 그러나 이러한 방식들의 경우 객체에 의해 생성된 그림자 영역과 실내의 조명 변화에 의해 정확한 객체 검출이 이루어지지 않는다. 그림 3은 위에서 설명한 실내 환경 변수에 의해 발생하는 문제점들을 나타내고 있다. 그림 3 (A)에 속한 영상은 등장 인물의 수가 1인으로 구성된 영상이며, 모든 차영상은 동일한 임계값($th = 20$)을 적용하였다. 이때 사용된 배경영상은 객체가 등장하기 직전에 보정된 영상을 이용하여 검출을 수행하였다. 이와 같은 경우는 객체의 등장으로 인해 조명 변화가 적은 경우에 해당되며, 프레임 255, 305와 같이 미세한 잡음 성분이 검출되었다. 또한 프레임 547, 619, 665는 그림자가 객체에 포함되어 검출된 경우를 나타낸다. 그림 3의 (B)는 등장인물의 수가 2~4명으로 구성된 경우이다. 가장 커다란 문제점은 조명의 변화, 즉 등장인물의 수가 늘어남에 따라 객체에 의해 굴절 및 산란된 조명 성분에 의해 변질된 밝기성분이다. 이로 인해 Frame_2 291, Frame_4 205, 266, 588과 같이 배경에 해당하는 영역들과 그림자에 해당하는 영역이 객체영역으로 오검출 되는 문제점이 발생한다. 차영상 내에 존재하는 색상 성분 중 흰색은 배경과 입력영상간의 유사도가

Frame Number	Frame_1 #205	Frame_1 #255	Frame_1 #305	Frame_1 #440	Frame_1 #547	Frame_1 #619	Frame_1 #665
Input Image (A)							
Difference image							
Frame Number	Frame_2 #136	Frame_2 #219	Frame_2 #240	Frame_2 #329	Frame_4 #205	Frame_4 #266	Frame_4 #588
Input Image (B)							
Difference image							

그림 3. 배경영상과 입력영상에 의해 생성된 차영상
Fig. 3. Background image and input image generated by difference image

가장 낮은 부분에 해당되며, 흰색을 제외한 나머지 색상성분들은 유사도의 차이에 따라 부여된 색상들이며, 해당 성분들은 임계값의 설정에 따라 달라질 수 있다.

본 논문에서는 배경영상을 이용함에 따라 발생하는 문제점을 해결 및 보완하기 위해 혼합영상과 입력영상을 이용하여 정확한 객체 검출을 수행하였다.

3.2.1 혼합영상 생성

배경영상을 사용하여 객체를 검출하는데 있어 가장 큰 문제점은 현재 입력영상과의 화소 밝기 차에 의해 발생한다. 따라서 본 논문에서 제안하는 방법은 배경영상과 입력영상을 이용하여 혼합영상을 생성하여 객체검출 시 발생할 수 있는 문제점을 최소화 하였다. 혼합영상을 생성하는 방법은 다음과 같다.

Step 1. 먼저 배경영상과 입력영상을 이용하여 식 2와 같이 평균 영상을 생성한다. 생성된 평균 영상은 그림 4의 (a)이다.

$$Avr = \frac{BG_{rgb}(x, y) + I_{rgb}(x, y)}{2} \dots\dots\dots (2)$$

Step 2. 보정된 배경영상 Φ 은 식 3에 의해 생성하며 그 방법은 다음과 같다. 입력영상 $I_{rgb}(x, y)$ 과 배경영상 $BG_{rgb}(x, y)$ 의 단순 차영상을 적용하여 설정한 임계값($th = 50$)과 비교하여 새롭게 보정된 배경영상을 생성한다. 차영상한 결과 값이 임계값보다 크면 배경영상의 화소 값을 가져오고, 차영상한 결과 값이 임계값보다 작으면 입력영상의 화소 값을 가져와서 배경영상을 보정한다. 그림 4의 (b)는 조명의 변화를 포함한 새롭게 보정된 배경영상이다.

$$\Phi = \begin{cases} BG_{rgb}, & \{I_{rgb}(x, y) - BG_{rgb}(x, y)\} > th \\ I_{rgb}, & Otherwise \end{cases} \dots\dots\dots (3)$$

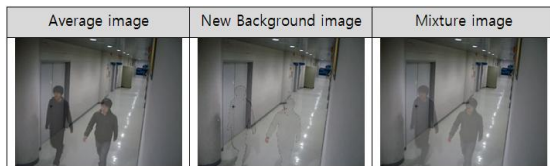


그림 4. 혼합영상 생성 결과 (a), (b), (c)
Fig. 4. The result of creating a mixture image. (a), (b), (c)

Step 3. 혼합영상(Mix)은 평균영상과 보정된 배경영상을 사용하여 다음의 식 4에 의해 생성된다. 그림 4의 (c)와 같이 변형된 조명성분과 그림자 성분이 포함되어 있는 혼합영상이 획득된다.

$$Mix = \frac{Avr_{rgb}(x, y) + \Phi_{rgb}(x, y)}{2} \dots\dots\dots (4)$$

3.2.2 객체 검출

객체 검출과정은 혼합영상과 현재영상과의 차영상을 이용하여 이루어진다. 차영상을 획득하는 방법은 식 (5)와 같이 표현 된다.

$$\delta I(x, y) = abs(I(x, y) - Mix(x, y)) \dots\dots\dots (5)$$

여기서, $\delta I(x, y)$ 는 현재영상과 혼합영상의 차영상을 나타낸다. 그림 5(a)는 배경영상과 입력영상을, 그림 5(b)는 혼합영상과 입력영상을 이용하여 검출된 결과를 나타내고 있다. 그림안쪽에 표시된 타원형의 내부를 살펴보면 배경영상의 경우 그림자영역이 존재하지 않으며, 입력영상의 경우 그림자영역이 존재한다. 그러나 혼합영상의 경우 입력영상의 그림자와 비슷한 밝기 값의 영역이 배경에 포함되어 있으며, 객체 부분 역시 그림자의 밝기 값과 비슷한 형태로 표현되어 있다. 결과영상에서 나타나듯이 혼합영상과 입력영상을 이용한 객체영역 검출이 더 효과적임을 알 수 있다.



배경영상(왼쪽), 입력영상(가운데), 차영상결과(오른쪽)
그림 5. 객체 검출 결과

Fig. 5. The Result of Object Detection.

혼합영상에 의해 검출된 영상은 레이블링에 의해 일정한 크기를 가지는 미소영역을 제거시킨다. 이후 모폴로지 기법인 채움 연산을 통해 객체 내부에 존재하는 공백 영역을 제거시켜 객체영역의 정확도를 높인다. 모폴로지 기법은 식 (6)에 의해 채움 연산을 수행한다.



그림 6. 결과영상
Fig. 6. Result image

$$g_o(x, y) = [(\delta I \ominus M) \oplus M](x, y) \dots\dots\dots (6)$$

$$g_c(x, y) = [(g_o \oplus M) \ominus M](x, y)$$

g_o 는 opening 연산을 의미하며, g_c 는 closing 연산을 나타낸다. δI 는 검출된 객체영상이며, M 은 3×3 크기의 회선 마스크를 나타낸다. 그림 6은 미소영역과 객체 내부의 채움 연산을 수행한 최종 결과 영상을 나타낸다.

IV. 실험결과

본 논문에서 제안한 객체 검출 시스템은 Windows XP환경에서 Visual C++ 6.0을 이용하여 구현하였으며, Intel(R) Core(TM)2 Quad CPU Q6600 2.40GHz와 2GB RAM의 하드웨어에서 실험하였다. 카메라는 CCD카메라를 사용하였으며, 실험 영상은 초당 24프레임, 영상 사이즈는 640×480을 사용하였다. 실험에 사용된 총 프레임 수는 5609개 이며, 배경영상을 생성하기 위한 프레임 수는 포함되지 않았다. 시스템 성능 평가는 1명 이상 5명 이하의 영상을 이용 하였다. 기존 시스템의 경우 총 검출율은 77.5%이며, 에러율은 22.5%이다. 제안된 시스템은 총 검출률 92%이며, 에러율은 8%이다.

표 1. 시스템 성능 평가
Table 1. Test of system performance.

	총 프레임 수		검출율(%)		에러율(%)	
	1인	2인 이상	1인	2인 이상	1인	2인 이상
기존 시스템	1665	3944	87	68	13	32
제안한 시스템	1665	3944	98	86	2	14

기존 시스템[7]은 생성된 배경영상과 입력영상에 의해 객체를 검출하는 방법이다. 이 시스템의 특징은 배경영상을 일정시간마다 갱신하여 객체를 검출하는 방식으로, 실내 환경에서 사람 또는 객체의 등장에 의해 굴절 및 변질된 조명성분으로 인하여 객체를 검출할 때 오검출이 발생한다. 표 1에 나타난 결과처럼 2인 이상의 경우 객체에 의해 굴절된 조명성분에 의한 검출 성능이 현저히 낮아지는 것을 볼 수 있다. 그러나 제안한 시스템의 경우는 변형된 조명성분과 그림자 성분이 포함되어 있는 혼합영상을 사용하기 때문에 2인 이상의 경우에서도 검출률이 높게 나타나는 것을 볼 수 있다. 따라서 기존 시스템보다 혼합영상과 입력영상을 이용한 객체영역 검출이 더욱 효과적임을 알 수 있다.

V. 결론

일반적으로, 고정된 배경영상과 입력영상간의 차영상을 적용하면 배경에 해당하는 영역들과 그림자에 해당하는 영역이 객체영역으로 오검출 되는 문제점이 발생하기 때문에, 본 논문에서는 실시간으로 입력되는 영상으로부터 배경영상을 모델링하고 일정한 시간 간격으로 배경영상을 갱신한다. 배경영상과 입력영상을 이용하여 평균영상을 생성하고, 조명의 변화를 포함한 새로운 보정된 배경영상을 생성한다. 평균영상과 보정된 배경영상을 이용하여 최종적으로 혼합영상을 생성한다. 혼합영상은 조명의 변화와 객체에 의해 발생된 그림자 성분, 바닥면에 반사된 조명성분이 모두 반영되어 생성되기 때문에 실험에서와 같이 정확하게 객체를 검출할 수 있었으며, 객체 검출 시 발생할 수 있는 문제점을 최소화하였다.

그림 6의 Frame_4 #205에서와 같이 줄무늬 형태의 복잡한

질감 영역에서 배경으로 인식되는 경우가 발생한다. 이러한 경우 모폴로지 연산에 의해 내부에 존재하는 공백을 채우지 못하는 결과를 초래한다. 향후 혼합영상 생성시 객체 내의 복잡한 질감 영역을 고려하는 처리가 고려되어야 한다.

참고문헌

[1] Dar-Shyang Lee, "Effective Gaussian mixture learning for video background subtraction," *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, VOL. 27, NO. 5, pp. 827-832, May 2005.

[2] A. M. McIvor, "Background subtraction techniques," *Proc. IEEE Conf. Systems, Man, and Cybernetics*, pp. 3099-3104, April 2004.

[3] 김현기 외 4인, "적응적 3프레임 차분 방법 기반 템플릿을 이용한 객체 추적," *퍼지및지능시스템학회 논문지*, 제 17권, 제 3호, 349-354쪽, 2007년 1월.

[4] 장희준, 고혜선, 최영우, 한영준, 한현수, "동영상에서 칼만 예측기와 블록 차영상을 이용한 얼굴영역 검출기법," *한국지능시스템학회 논문지*, 163-172쪽, 2005년 4월.

[5] 주홍식, 이병선, 이은주, "차영상과 컬러정보를 이용한 목표물 계수 알고리즘," *한국정보기술응용학회 논문지*, 83-83쪽, 2004년 6월.

[6] Teixeira, L. F., Cardoso, J. s., and Corte-Real, L. "Object segmentation using background modelling and cascaded change detection," *Journal of Multimedia*, 2(5), pp. 55-64, September 2007.

[7] 안성진 외 4인, "중요지역 보안을 위한 조명환경 적응형 실시간 영상 감시 시스템," *전자공학회 논문지*, 제 44권, 제 2호, 116-124쪽, 2007년 3월.

[8] O. Javed, K. Shafique, and M. Shah, "A hierarchical approach to robust background subtraction using color and gradient information," in *Proc. IEEE Workshop Motion Video Computing*, pp. 22-27, December 2002.

[9] 조현태, 장재니, 강남오, 백준기, "배경 생성 기법을 이용한 다중 카메라 객체 추적 시스템 구현," *2008년 대한전자공학회 하계학술대회*, 제 31권, 제 1호, 947-948쪽, 2008년 6월.

[10] 지영석, 한영준, 한현수, "불안정한 조명 환경에 강인한 참조 배경 영상의 갱신 기법," *한국컴퓨터정보학회 논문지*, 제 15권, 제 1호, 91-102쪽, 2010년 1월.

[11] 정중교, 박상성, 장동식, "피부색과 Haar-like feature를 이용한 실시간 얼굴검출," *한국컴퓨터정보학회 논문지*, 제 10권, 제 4호, 113-121쪽, 2005년 9월.

저자소개



최미영

2001 : 백석대학교 컴퓨터학과 공학사.
 2003 : 백석대학교 컴퓨터학과 공학석사.
 2008 : 숭실대학교 미디어학과 공학박사.
 2008~현재 : 숭실대학교 BK21 Post Doc.
 관심분야 : 컴퓨터비전, 영상처리, 패턴인식, HCI



김계영

1990 : 숭실대학교 전자계산학과 공학사.
 1992 : 숭실대학교 컴퓨터학과 공학석사.
 1996 : 숭실대학교 컴퓨터학과 공학박사.
 1996~1997 : 한국전자통신연구원 Post Doc.
 1997~2001 : 한국전력공사 전력연구원 선임.
 2001~ 현재 : 숭실대학교 컴퓨터학부 부교수.
 관심분야 : 컴퓨터비전, 형태인식, 생체인식, 증강현실, 영상 및 신호처리



최형일

1979 : 연세대학교 전자공학과 공학사.
 1982 : 미시간대학교 전산공학과 공학석사.
 1987 : 미시간대학교 전산공학과 공학박사.
 1987~현재 : 숭실대학교 IT대학 미디어학과 교수.
 2005~2006 : 한국정보과학회이사.
 2005~현재 : 한국디지털미디어학회 이사.
 2008~현재 : 숭실대학교 문화콘텐츠기술(CT) 연구소 소장.
 관심분야 : 컴퓨터비전, 퍼지 및 신경망이론, 패턴인식, 지식기반시스템