

3 차원 물체인식을 위한 Vision 시스템

閔 鉉 午*, 朴 來 弘**

西江大學校 電子工學科
大學院*, 助教授(工博)**

I. 序 論

인간의 살아가는 과정은 외계로부터의 많은 데이터에 대한 의식적인 혹은 무의식적인 반응이라 할 수 있다. 이에 따라 이러한 입력으로 들어온 수많은 데이터를 분석하고 표현함과 동시에 그것들로부터 정보를 끌어내는 일, 즉 입력을 이해하는 일이 필요하게 된다. 여기에서는 이러한 인간에게 들어온 입력중 특히 vision에 관해 다루고자 한다. 인간은 아무리 복잡한 장면이나 화상이라 할지라도 한번 보면 쉽게 이해할 수 있는 반면 현 단계의 vision 시스템으로는 그것을 기대하기 어렵다. 왜냐하면 sensing 장치를 통하여 얻은 화상 depth에 대한 정보를 가지고 있지 않기 때문에 주어진 2차원의 화상으로부터 3차원 물체에 대한 3차원 정보를 추출해내야 하기 때문이다. 이 과정을 위해서는 각 물체의 인식뿐 아니라 더 나아가 각 물체끼리 결합하여 이루는 모든 가능성에 대한 knowledge-base가 필요하게 되어 많은 메모리 및 계산을 요하게 된다. 더구나 화상에서 물체끼리 서로 중첩되어 있거나 잡음의 영향이 많은 때는 이러한 과정이 더욱 더 어렵게 된다. computer vision 시스템을 위한 연구가 자동항해, scene 해석 및 산업계에서는 로봇트와 연관지어 물품 검사, 분류, 용접등 많은 응용분야에 걸쳐 20여년 동안 계속되어 왔으나 아직까지는 초보단계에 지나지 않는다.

일반적인 computer vision 시스템은 주어진 화상으로부터 edge와 같은 정보를 추출하기 위한 preprocessing 과정, 물체의 3차원 특징을 추출하는 과정 및 이것들로부터 이미 알고 있는 물체와 match시켜 물체인식 및 물체사이의 관계를 얻어내어 화상을 이해하는 단계로 나누어 생각할 수 있다.

주어진 2차원 화상으로부터 3차원 정보를 얻어내기 위하여 밝기, 그림자, 꼭지점의 모양, texture등과 같은 화상의 여러가지 특징들을 constraint로 사용하

여 모든 가능성에 대하여 고려하여 unique하게 물체를 결정할 수 있다. 그러나 이 모든 가능성에 대한 계산량이 너무 방대하여 실시간으로 처리하기 위해서는 1초에 적어도 $10^{10} \sim 10^{13}$ 개의 명령을 수행할 수 있는 프로세서를 요하게 된다.^[1] 이에 반해 좀 더 체계적인 방법으로 물체인식과 matching 및 더 나아가 화상이해를 위해 각 물체 및 이러한 물체들로부터 이루어지는 보다 복잡한 물체들에 대한 knowledge-base를 가지고 들어온 입력 데이터와 비교하여 화상을 이해하는 model-based 방법이 있다. 물론 이 경우에서는 물체인식이 크기 축소, 보는 각도등에 따라 달라지게 되므로 모든 변화 가능성에 대해서도 고려를 해야 한다.

여기에서는 model-based 시스템의 예로써 stereo 화상으로부터 block 모델을 사용하여 도식지 사진으로부터 빌딩의 3차원 묘사에 응용된 mosaic 시스템과 3차원 물체를 cylinder와 cylinder의 양 끝을 나타내는 타원으로 이루어진 generalized cylinder로^[2] 분석하여 이것을 비행장에서 찍은 사진에 적용하여 비행기를 인식하고 더 나아가 주어진 비행기 모델 match에 적용된 acronym 시스템에 대해 간단히 소개하고자 한다.

II. Mosaic 시스템

3차원 mosaic 시스템은 다른 각도로부터 얻은 복잡한 도시의 항공 stereo 화상으로부터 3차원 정보를 추출하여 전체 도시의 3차원 모델을 점차적으로 구하는데 응용되었다.^[3] 여기서 도시의 항공사진은 카메라의 위치를 달리하여 같은 scene에 대해 보는 각도만 다르게 되어 있는 stereo 화상으로 이루어져 있다고 본다.

하나의 화상으로는, 예를 들어 겹친부분이 있거나 보는 각도에 따라, 부분적인 정보만 얻을 수 있기 때문에 도식지 항공사진과 같은 복잡한 화상을 3차원적으로 완전히 이해하기는 어려우며, 한 화상에서 얻을

수 없는 화상의 특징을 같은 scene 이면서 보는 각도를 달리한 stereo 화상에서는 추출하기 쉬운 경우가 있게 된다. Scene에 대한 부분적인 이해를 주어진 화상에서 얻은 후 이 화상으로부터의 예측과정 및 다른 각도에서 본 stereo 화상을 통해 3차원 물체에 대한 정보를 그때 그때 점차적으로 update 하는 것이 incremental mosaic vision 시스템의 아이디어라 할 수 있다.

Mosaic vision 시스템의 기본구조에 대해 살펴보면, 먼저 주어진 stereo 화상으로부터 화상의 특징을 얻어내는 stereo 화상 분석 과정이 있다. 즉 3×3 sobel operator를 사용하여 edge를 추출한 후 필요에 따라 line fitting이나 linking 과정을 거쳐 주어진 stereo 화상에서 edge를 얻는다. 이 edge로부터 5×5 의 window를 사용하여 이 window내에서 edge들이 만나 이루는 접합을 찾아 나가면서 접합의 모양에 따라 각각 L접합, T접합, Arrow접합 및 Fork접합등으로 분류한다.⁽⁴⁾ 그 다음 과정으로서 stereo 화상에 나타난 각 접합들을 서로 match 시키는 과정을 거쳐 여러가지의 match 가능성중 유일한 match를 선택한다. 여기에서 두 stereo 화상사이의 관계는 사진을 찍은 카메라의 위치등 두 화상을 얻은 조건을 미리 알고 있다고 가정하여 주어진 화상으로부터 이것의 stereo 화상은 어느 정도 예측 가능하게 된다. 이러한 예측과정을 거쳐 예를 들어 도시 빌딩으로서 주어지는 각 직선의 특성 및 접합을 고려하여 stereo 화상으로부터 완전한 접합 정보를 얻는 과정이 수반된다.

다음 단계로서 주어진 stereo 화상의 접합 정보로부터 3차원 물체를 구성하는 단계가 있는데, 먼저 구한 접합들로부터 면을 구성하고 다시 이 면들로부터 물체의 3차원 구조를 좌표 공간에서 구하게 된다. 이 과정에서는 각 단계에서 선, 접합 및 면에 대한 constraint가 적용되며, 더 나아가 빌딩의 지붕, 벽등을 나타내는 edge의 성질등을 고려한 knowledge-base에 의해 이것이 규칙으로 적용된다. 예를 들어 missing된 빌딩의 수직 edge를 knowkdge-base에 의해 삽입시키거나 불필요한 edge가 있을 때는 가까이 있는 두 edge를 합쳐주거나 혹은 제거시킨다. 다시 말해 우리가 관심있는 화상에 내포되어 있는 3차원 물체의 성질 또는 구조에 대한 모든 특징을 knowledge-base로 가지고 있어 이것들을 규칙으로 적용시켜 2차원 정보로부터 3차원 정보를 추출해 내게 된다.

결론적으로 말해 이 incremental mosaic vision 시스템을 적용하여 복잡한 도시의 항공 stereo 사진으

로부터 도시를 구성하는 물체에 대한 knowledge-base를 적용하여 3차원 도시 모델을 좌표 공간에서 얻을 수 있고⁽⁵⁾ 이 분석 방법을 확장한다면 여러 각도에서 얻은 같은 scene에 대한 많은 화상으로부터 점차적으로 update시켜 보다 정확한 3차원 물체구조를 얻는데 응용할 수 있다.

III. Acronym 시스템

Acronym vision 시스템은 자세한 3차원 기하학적 지식을 knowledge-base로 사용하여 물체의 인식, 또는 주어진 2차원 화상에서 3차원 정보를 얻음으로써 직접적으로 화상을 이해하는 시스템이다. 이 시스템에서는 3차원 물체를 기술하기 위하여 2차원 면이 어떠한 직선 혹은 곡선 축을 따라 형성된 generalized cylinder 개념을 쓰고 있다. 즉 acronym vision 시스템은 축과 단면을 이용하여 기하학적 3차원 물체를 수학적으로 쉽게 표현하여 각 물체 모델에 대한 constraint로 사용하고, 더 나아가 이러한 물체들로 이루어진 보다 더 큰 물체 또한 constraint를 사용하여 표현할 수 있다. 이 시스템을 이용하여 비행장에서 찍은 비행장 사진을 분석하여 주어진 비행기의 3차원 모델을 constraint로 사용하여 예측 및 matching 과정을 거쳐 비행기를 인식하고 그 기종을 찾아내어 비행장 사진을 이해할 수 있다.⁽⁶⁾

Acronym vision 시스템은 class와 더 나아가 subclass의 기하학적 모델에 대한 규칙을 constraint로서 가지고 있다. 위에서 언급한 예에서는 일반적인 비행기의 기하학적인 구조에 대한 일반적인 constraint가 있고 이 비행기를 분류하여 L-1011 모델 혹은 Boeing-747 모델에 대한 constraint가 있으며 더 나아가 Boeing-747 모델을 Boeing-747B와 Boeing-747 SP 모델까지로 세분하여 모든 비행기 모델에 대한 constraint를 knowledge-base로 하고 있다. 즉 비행기를 cylinder 형태의 동체, 대칭인 앞, 뒤의 날개, 방향타 및 뒷엔진등의 기하학적인 구조로 세분하여 각 구조의 두께, 길이, 반지름등의 크기의 범위가 위에서 언급한 각 비행기 모델에 대해 constraint로 주어지게 된다. 물론 비행장 사진을 찍은 카메라의 위치등이 또한 constraint에 포함되어 512×512 로 주어진 화상에 적용된다. 즉 acronym vision 시스템에 주어진 knowledge-base에는 비행기를 구성하는 각 구조를 고려한 각 비행기 모델에 대한 constraint, 카메라의 위치, 각등을 포함하는 항공사진 촬영시의 화상촬영 모델 및 화상의 크기 모델등이 포함되어 있다.

Acronym vision 시스템을 이용하여 항공사진을 이해하는 과정을 간단히 살펴보면 주어진 화상에서 edge를 찾은 후, 이것으로부터 3 차원 generalized cylinder의 2 차원 표현인 ribbon과 타원을 찾은 후에 비행기 구조에 대한 knowledge-base에 의해 주어진 규칙을 사용하여 예측과정을 거쳐 edge linking을 행한 후, 이 데이터로부터, 예를 들면 비행기의 동체, 날개등의 후보를 찾아 match시켜 나가면서 주어진 데이터에 맞는 비행기 모델을 찾는다. 예를 들어 비행기의 왼쪽 날개 부분만 가지고도 동체 및 오른쪽 날개등 나머지 구조의 위치를 예측한 후 실제의 화상에 나타나는 것과 비교하여 예측과정을 시험할 수 있는 기능을 갖는다.

결론적으로 말해 acronym vision 시스템은 앞에서 예를 들어 설명한 바와 같이 주어진 3 차원 물체의 기하학적인 구조에 대한 지식을 knowledge-base로 하여 주어진 화상에서 그 물체를 인식, 예측 및 추리 과정을 거쳐 화상을 이해하는 vision 시스템이다.

IV. 結 論

주어진 2 차원 화상에서 3 차원 정보를 끌어내고 물체를 인식하여 물체사이의 관계를 표현하여 화상을 이해하는데는 아직까지 많은 발전을 요한다. 현재의 vision 시스템은 제한된 경우의 간단한 화면만 다루고 있으며 그 속도 또한 실시간으로 처리하기에는 너무 늦다. 일반적인 실시간 처리가 가능한 vision 시스템의

개발을 위해서는 화상에 나타나는 물체와 그 물체들 사이의 관계에 대한 보다 많은 지식이 필요할 뿐 아니라 이를 처리할 수 있는 큰 저장능력 및 고속처리능력을 갖는 프로세서를 필요로 하게 된다. 이러한 프로세서가 가능하다면 좀 더 복잡한 3 차원 물체를 인간 눈과 같이 실시간에 쉽게 이해할 수 있을 것이다.

參 考 文 獻

- [1] Takeo Kanade and Raj Reddy, "Computer vision: the challenge of imperfect inputs," *IEEE Spectrum*, vol. 20, pp. 88-91, Nov., 1983.
- [2] Gerald J. Agin and Thomas O. Binford, "Computer description of curved objects," *IEEE Trans. Comput.*, vol. C-25, pp. 439-449, April, 1976.
- [3] Martin Herman, Takeo Kanade, and Shigeru Kuroe, "Incremental acquisition of a three-dimensional scene model from images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAMI-6, pp. 331-340, May, 1984.
- [4] Patrick Henry Winston, *Artificial Intelligence*, second ed., Addison-Wesley, 1984.
- [5] Rodney A. Brooks, "Model-based three-dimensional interpretations of two-dimensional images," *IEEE Trans. pattern Analysis and Machine Intelligence*, vol. PAMI-5, pp. 140-150, March, 1983. *

案 內 文

본 학회에서는 이달부터 종전의 회비납입방법(우체국의 소액환 및 대체구좌 이용)과 은행 지로를 통한 각종회비 납입을 병행하여 실시하고 있습니다.

은행의 지로를 이용하실 때에는 전은행 99번(지로)창구에서 본 학회의 지로번호(7510904)를 기입하시고 회비를 납입하여 주시기 바랍니다.