

□ 기술해설 □

컴퓨터 비전의 처리 기술 및 응용†

승실대학교 최 형 일*
 김 명 원**

| ● 목 | 차 ● |
|------------------------|-------------------|
| 1. 서 론 | 3.1 산업용 시각 검사 시스템 |
| 2. 컴퓨터 비전의 처리 기술 | 3.2 항공 사진 분석 시스템 |
| 2.1 스케치 형성을 위한 처리 기술 | 3.3 문서 영상 분석 시스템 |
| 2.2 기호 묘사 형성을 위한 처리 기술 | 3.4 이동 물체 분석 시스템 |
| 3. 컴퓨터 비전의 응용 분야 | 4. 결 론 |

1. 서 론

“몸이 천냥이면 눈은 구백냥.” 이는 시각 정보 처리의 중요성을 강조한 문구이다. 인간은 주어진 환경에 적응하면서 방대한 양의 정보를 습득하고 학습하는 과정을 반복한다. 이러한 정보 획득 및 학습 활동의 많은 부분이 시각 활동을 통하여 구현된다. 예를 들어, 책을 읽고 물체를 인식하는 작업은 눈을 통하여 입력된 데이터가 두뇌의 시각 피질에서 분석되는 시각 활동에 의하여 구현된다. 인간의 시각 활동은 왼쪽 눈과 오른쪽 눈에 입력된 데이터를 비교하여 물체까지의 거리 정보를 추론하기도 하며, 연속적으로 입력되는 데이터를 분석하여 물체의 동작에 대한 정보를 추론한다. 이러한 시각 활동은 매우 신속히 진행되어 우리 인간은 눈을 통하여 데이터가 시각 피질에 전달되고 전달된 데이터가 분석되는 총체적인 처리 과정을 전혀 의식하지 못한다.

컴퓨터 비전이란 투시된 영상들로부터 장면에

대한 유용한 정보를 추출하는 작업을 의미한다. 즉, 여러 화소들의 배열인 영상으로부터 물리적인 대상을 명확하고 의미있게 해석하는 과정을 의미한다. 일반적으로 실세계는 불변의 성질(크기, 색상, 형태, 무게, ...)을 가진 물체들로 구성되어 있지만 이들에 대한 영상은 불명확하거나 오인된 변형물들을 포함하게 된다. 따라서 대부분의 컴퓨터 비전 시스템들은 입력 영상에서 추출된 정보와 실세계에 대한 가설 및 지식을 이용하여 불확실한 입력 영상을 해석하고 있다.

그러나, 여러 유형의 문제 영역에 적용 가능한 비전 시스템을 구축하는 작업은 매우 어려운 일이다. 하나의 장면은 많은 제약 조건하에서 영상으로 형성되기 때문에 영상 자체만으로는 그 장면을 회복하기에 충분한 정보가 제공되지 않으며, 여러 문제 영역에 대한 엄청난 양의 사전 지식을 시스템이 이해할 수 있도록 표현하는 작업은 매우 어려운 일이며, 영상 데이터는 그 양이 방대하여 많은 처리 시간을 요구하기 때문이다. 따라서, 대부분의 컴퓨터 비전 시스템은 특정한 문제 영역에 적용되도록 구성되는 것이 일반적인 예이다. 본 고에서는 주어진 장면에 대한 영상을 이해하기 위하여 필요한 처리 기술들을 설명하고, 이러한 기술들을 활용한 응용 분야를 살펴

†본 연구는 94년도 학술진흥재단 자유공모과제(E0830)에서 지원받았음을 밝힙니다.

*중신회원

**정회원

본다.

2. 컴퓨터 비전의 처리 기술

컴퓨터가 장면을 보고 이해할 수 있는 능력을 갖추기 위하여는 여러 단계의 처리 절차를 거쳐야 한다. 본 절에서는 각 단계에서 수행되어야 할 기본 작업들과 이러한 작업들을 수행하기 위하여 사용되는 기본 기술들에 대하여 설명한다. 장면을 보고 이해하는 과정은 입력된 영상을 사전에 미리 알고 있는 실세계의 모델로 변환하는 과정으로 해석될 수 있다. 입력 영상은 일반적으로 화소들의 배열로서 표현되고 실세계의 모델은 일반적으로 개념적이며 추상적인 기호들의 집합으로 표현된다. 예를 들어, 자연 장면에서 집, 도로, 하늘, 숲 등을 인식하는 작업은 화소들의 배열로 표현되는 영상을 다분히 추상적인 개념인 실세계의 모델(집, 도로, 하늘, 숲)로 변환하는 과정을 의미한다. 그러나, 입력 영상에 대한 표현과 실세계의 모델에 대한 표현과는 엄청난 차이가 있고, 실세계의 모델 역시 표현하고자 하는 대상에 따라서 매우 상이한 형태로 표현될 수 있다. 따라서, 입력 영상에서 실세계의 모델로의 변환은 단순한 공식에 의하여 정의될 수 없다. 그림 1은 [1]에서 기술된 두 단계의 변환 과정을 나타낸다.

그림 1의 첫 단계의 변환은 주로 영상 형성 과정에 대한 사전 지식을 활용하여 수행된다. 예를 들어, 화소값의 밝기 변화를 이용하여 에지를 검출하는 작업(edge detection), 좌측 영상과 우측 영상을 비교하여 거리 정보를 추론하는 작업(spatial stereo matching), t_1 시간에서의 영상과 t_2 시간에서의 영상을 비교하여 동작 정보를 추론하는 작업(temporal stereo matching), 광원과 표면과의 기하학적인 관계에 의하여 유도되는 명암 변화를 분석하여 표면의 특성을 추론하는 작업(shape from shading) 등을 생각할 수 있다. 두 번째 단계의 변환은 주로 응용 분야에 대한 사전 지식을 활용하여 수행하여 스케치 형태의 표현을 기호적인 표현으로 표현된다. 예를 들어, 에지점들을 분석하여 특정한 물체의 경계선을 추출하고 추출된 경계선을 그 물체를 대표하는

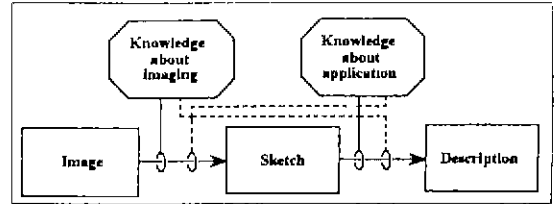


그림 1 입력 영상에서 실세계 모델로의 변환

기호로서 표현하는 작업, 공간적 스테레오 정합에서 추출된 거리 정보들을 분석하여 특정한 물체의 표면을 재구성하거나 또는 특정한 물체를 확인하는 작업, 시간적 스테레오 정합에서 추론된 동작 정보를 이용하여 특정한 물체가 현재의 장면에서 어떠한 속도 및 방향으로 움직이고 있는가를 확인하거나 또는 어떠한 운동성을 갖는 특정한 물체를 확인하는 작업 등을 생각할 수 있다.

그러나, 그림 1의 각 단계의 변환 과정은 비전 시스템을 적용하고자 하는 응용 분야의 특성 및 그 목적에 따라 매우 상이한 형태를 취할 수 있다. 예를 들어, 비전 시스템을 문자 인식(offline character recognition)에 적용하고자 하는 경우에는 그림 1의 첫번째 단계의 변환에서는 거리 정보 및 동작 정보를 추론할 필요가 없으며 두번째 단계의 작업에서는 매우 정교한 모델과의 정합(matching) 및 분류 과정(classification)이 요구된다.

2.1 스케치 형성을 위한 처리 기술

영상 분할 과정(image segmentation)이라고 불리는 스케치 형성 과정은 매우 중요한 처리 단계로서, 이 단계에서의 작업이 원활히 수행되는 경우 다음 단계의 작업이 매우 쉬워진다. 그러나, 스케치 형성 작업은 그 자체로서 완벽을 기할 수 없는 작업이다. 완벽한 스케치 형성이 이루어졌다는 사실은 이미 인식이 완료됐다는 것을 의미하기도 한다. 따라서, 최근에는 스케치 형성 과정을 독립적으로 생각하지 않고 다음 단계의 작업인 기호 묘사 형성 작업과 함께 해결하려는 시도도 이루어지고 있다. 본 고에서는

스케치 형성을 위하여 사용되는 주요한 기법들의 최근 동향을 기술한다.

2.1.1 에지 검출 기술

물체와 배경간의 경계에서는 밝기값의 급격한 변화가 생긴다는 전제하에서 대부분의 에지 검출기는 이러한 변화가 발생하는 화소들을 찾고자 한다. 전형적인 검출기로서 소벨 Sobel 에지 검출기, Marr & Hildreth 에지 검출기, 여러 방향의 마스크를 이용한 Kirsch 에지 검출기 등을 생각할 수 있다[1]. 이러한 대부분의 검출기들은 에지 유무를 판단하기 위하여 임계 방법을 사용한다. 즉, 각 화소의 에지 크기(edgeness)가 특정한 임계값 이상인 경우에 그 화소를 에지로 선언한다. 따라서, 임계값을 결정하는 방법이 검출기의 성능을 좌우하는 중요한 요인으로 작용한다. 영상의 전체 영역에 대하여 하나의 임계값을 사용할 것인가 또는 영상의 지역적인 특성을 고려하여 적응적으로(adaptively) 임계값을 결정할 것인가 하는 문제를 생각하여야 한다. 임계값 결정을 위하여 주어진 문제 영역의 특성을 활용할 것인가 또는 영상에 나타난 정보만을 활용할 것인가 하는 문제 역시 주요한 고려 사항이다. 임계값을 적응적으로 결정하기 위하여 퍼지 이론(fuzzy theory)을 도입하는 방법들에 대한 연구가 활발히 진행되고 있다. 예를 들어, 참고 문헌 [2]에서는 주어진 영상에서 퍼지 척도를 정의하고 정의된 척도에 의하여 임계값을 결정하는 방법을 소개하였다. 퍼지 척도를 정의하기 위하여 각 물체에 대한 히스토그램을 사전에 분석하는 방법, 즉 문제 영역에 관한 사전 지식을 퍼지 척도에 도입하는 방법을 사용하였다. Marr & Hildreth 검출기의 성능을 결정하는 σ 값을 영상내 지역적인 명암 분포에 따라 다르게 적용하므로써 미세한 에지 성분을 보호하고 동시에 불필요한 에지 및 잡음은 제거하는 적응적 에지 검출기에 대한 연구도 진행되고 있다.

임계값 결정 과정에 문제 영역에 관한 사전 지식을 도입하고자 하는 시도 역시 매우 활발히 진행되고 있다. 이러한 시도는 특히 문서 영상의 분석을 위하여 활발히 이루어지고 있다. 예를 들어, 참고문헌 [3]에서는 문서 영상이 갖는 특성

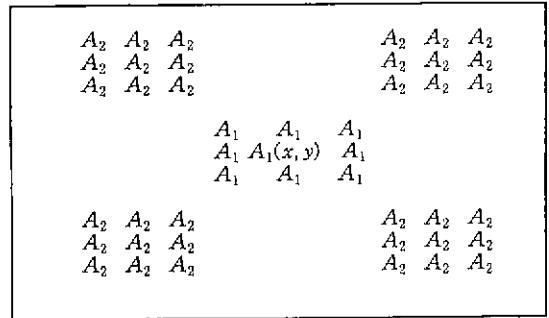


그림 2 문제 영역의 사전 지식을 이용한 적응 임계 방법

— 일반적으로 문자열 부분과 배경 부분이 번갈아 발생하는 특성—에 착안하여 적응적으로 임계값을 결정하는 방법을 제안하였다. 예를 들어, 그림 2의 A_1 영역이 문자열 영역이라면 A_2 영역은 배경 영역일 가능성이 높으며, 역으로 A_1 영역이 배경 영역이라면 A_2 영역은 문자열 영역일 가능성이 높다. 따라서, 두 영역간의 명암값을 지역적으로 임계값을 적응적으로 결정하는 방법을 제안하였다.

2.1.2 공간적 스테레오 정합 기술

일반적으로 스테레오 정합으로 알려진 공간적 스테레오 정합은 2개의 초점(좌측 및 우측)에서 촬영된 영상을 비교하여 3차원적인 거리 정보를 추론하고, 추론된 정보를 이용하여 물체의 표면을 재구성하는 것을 목표로 한다. 좌측 및 우측 영상간의 정합은 영역간의 정합을 시도하는 방법 [4]과 에지와 같은 특징점간의 정합을 시도하는 방법 [5]들로 분류될 수 있으며, 정합의 효율을 위하여 에피폴라 제약(epipolar constraints), 연속성 제약(smoothness constraints) 등과 같은 제약 조건들을 활용한다. 공간적 스테레오 정합에 관한 가장 고전적인 참고 문헌으로는 [6]을 생각할 수 있다.

좌측 및 우측 영상간의 정합을 시도하는 문제를 효율적으로 수행하기 위하여 정합 문제를 표면 재구성 문제와 동시에 해결하고자 하는 노력이 활발히 진행되고 있다. 예를 들어, 참고문헌 [7]에서는 스테레오 정합에 의하여 추론된 거리 정보들을 이용하여 표면 재구성을 구현하고, 동

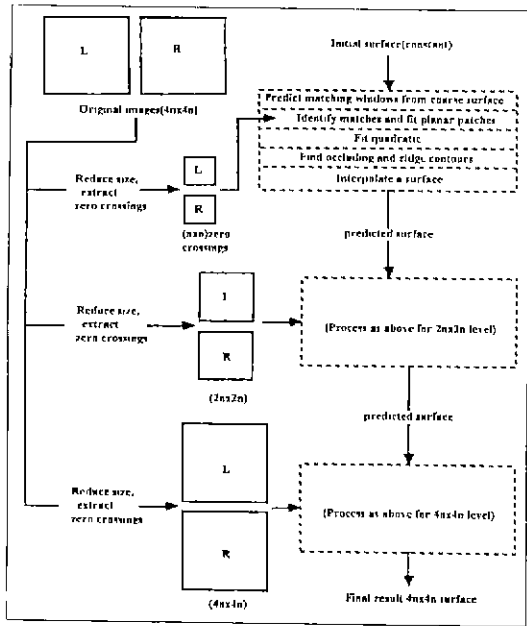


그림 3 정합과 표면 재구성의 통합

시에 재구성된 표면에 관한 정보를 이용하여 스테레오 정합을 재조정하는 방법을 계층적으로 반복하므로써 좀 더 정확한 정합을 이루고자 한다. 즉, 정합 문제를 독립적으로 취급하지 않고 물체의 표면과 연계하여 다루고자 하는 시도이다. 그림 3은 이 방법의 기본 구조를 보여준다. 영상의 해상도를 단계적으로 높혀가면서, 각 단계에서 정합을 시도하고 동시에 표면 재구성을 구현한다. 재구성된 표면은 다음 단계의 정합 과정에서 탐색 영역을 제한하기 위하여 사용된다.

공간적 스테레오 정합을 좀 더 정확하고 효율적으로 수행하기 위하여 지식 기반 스테레오 정합에 대한 연구 역시 활발히 진행되고 있다. 즉, 문제 영역에 대한 사전 지식을 정합 과정에 도입하는 방법이다. 참고문헌 [8]에서는 찾고자 하는 물체의 카메라까지의 거리를 사전에 미리 정의하여 정의된 거리내에 있는 정합점만을 탐색하는 방법을 시도하고 있으며, 참고문헌 [9]에서는 지식 베이스에 정의된 물체의 모델을 이용하여 정합을 시도하는 방법을 추구하고 있다. 예를들어, 지식 베이스에 정의된 물체의 특징들을 좌측 및 우측 영상에서 추출한 후, 추출된

특징들간의 정합을 구현하여 물체의 위치 및 거리 정보를 추론한다.

2.1.3 시간적 스테레오 정합 기술

인간은 항상 변화는 실세계에 대한 동영상을 분석하여 주위 환경 및 이동 물체에 대한 정보를 추출한다. 특히, 망막의 영상 수용체(photoreceptor)와 연결된 시신경에서는 영상의 시간적인 변화를 탐지하여 변화가 있는 곳에 주위를 집중시키는 역할을 수행한다고 알려져 있다. 물체의 동작 특성을 추출하기 위하여 약간의 시차(Δt)를 두고 연속적으로 입력되는 영상들을 비교 및 분석하는 대표적인 방법은 광류 측정에 의한 방법 [10]과 차영상 분석에 의한 방법 [11]이 있다. 이 방법들은 모두 두 시점에서 촬영된 영상간의 정합, 즉 시간적 스테레오 정합, 을 필요로 한다. 광류 분석에 의한 방법은 각 점 또는 영역의 속도 벡터를 추출하기 위하여 시간적 스테레오 정합을 필요로 하고, 차영상 분석에 의한 방법은 t_1 시점과 t_2 시점간의 차영상의 동작 영역과 t_2 시점과 t_3 시점간의 차영상의 동작 영역간의 시간적 스테레오 정합을 필요로 한다.

시간적 스테레오 정합을 위하여 사용되는 기본 기법들은 공간적 스테레오 정합을 위하여 사용되는 기본 기법들과 매우 유사하다. 물체의 동작 변화는 매우 완만히 진행되며 물체는 어느 정도의 크기를 갖는다는 전제하에서 많은 알고리즘들이 시공간 연속성 제약(spatio-temporal smoothness constraints)을 활용하고 있다. 특히, 물체의 이동 방향 및 속도에 관한 예측 정보를 활용하여 정합의 효율성을 높이고자 하는 노력이 활발히 진행되고 있다. 예를 들어, 참고 문헌 [12]에서는 최소 자승 적응 예측기(Least Square Adaptive Predictor)를 사용하여 정합 대상이 되는 이동 물체의 위치를 예측하고자 하였다. 두 시점간의 시차를 T 라 할 때 $n+1$ 시점에서의 예측 정보 $D(n+1)^*$ 는 n 시점에서의 물체의 위치 $D(n)$ 과 $n-1$ 시점에서의 물체의 위치 $D(n-1)$ 에 의하여 다음과 같이 산출된다.

$$D(n+1)^* = D(n) + \frac{D(n) - D(n-1)}{T} \cdot T$$

$$= 2 \cdot D(n) - D(n-1)$$

예측 정보를 추출하기 위하여 칼만-필터(Kalman filter)의 예측 기능을 활용하는 방안도 연구되고 있다. 예를들어, 참고문헌 [13]에서는 칼만-필터의 파라미터들을 이용하여 예측을 수행하고, 예측된 정보를 이용하여 두 시점간의 정합을 시도한 후 정합된 결과에 의하여 칼만-필터의 파라미터들을 수정하고, 수정된 파라미터들을 이용하여 다음 시점의 예측을 수행하는 작업을 반복한다.

2.2 기호 묘사 형성을 위한 처리 기술

스케치 표현을 기호 묘사 형태로 변환하는 과정은 실세계의 모델이 어떠한 형태로 정의되는가와 밀접한 관계가 있으며, 모델의 갯수가 복수개인 경우에는 인식을 위한 분류 작업(classification)을 수행하기에 적합한 형태가 무엇인가 하는 점도 고려되어야 한다. 본 고에서는 기호 묘사 기법들을 2차원 기호 묘사 기법과 3차원 기호 묘사 기법으로 구분하여 설명한다.

2.2.1 2차원 기호 묘사 형성을 위한 기술

2차원 기호 묘사는 주로 분할된 영역 및 에지들을 체계적으로 표현하기 위하여 사용된다. 에지들로 구성된 경계선을 표현하기 위하여 주로 사용되는 기호 묘사는 체인 코드(chain code), 푸리에 묘사(Fourier descriptor), 직선 및 곡선을 나타내는 파라미터들의 집합 등을 생각할 수 있고, 분할된 영역을 나타내기 위하여 주로 사용되는 기호 묘사는 4진 트리, 최소 포함 사각형 및 최소 포함 타원, 저널-모멘트(Zernike moments) 등을 생각할 수 있다. 이러한 기호 묘사들을 표현하는 방법 역시 다양한 형태로 나타난다. 벡터 형태로의 표현, 규칙에 의한 표현, 그래프에 의한 표현, 논리 모델에 의한 표현 등을 생각할 수 있다.

체인 코드는 가장 많이 사용되는 표현 기법 중의 하나로서 인접 경계 화소간의 방향 벡터를 코드화한다[14]. 하나의 방향 벡터는 2 또는 3-비트의 코드 워드(code word)로 표현되며, 일련의 코드 워드들과 시작 화소의 좌표값으로 하나

의 경계선을 나타낸다. 체인 코드의 미분치는 경계가 회전하더라도 불변이므로 회전된 물체간의 정합에 매우 유익하다.

푸리에 묘사 역시 가장 보편적인 표현 기법 중의 하나로서 사인형의 주기 함수(sinusoidal periodic functions)들을 사용하여 주어진 경계를 표현한다[15]. 주어진 경계를 푸리에 변환(Fourier transformation)하여 구해지는 계수들의 적절한 조합은 경계의 회전, 확대 및 축소, 위치 이동 등에 불변인 성질을 나타내므로 정합 과정에 효율적으로 사용될 수 있다.

4진 트리는 이전 영상을 계층적으로 표현하는 방법으로서 트리의 내부 노드는 4개의 자노드를 갖는다[16]. 트리의 단말 노드는 black 또는 white의 특성을 갖고 내부 노드는 gray의 특성을 갖는다. gray 노드는 흑화소와 백화소가 혼합된 영역을 나타내고, black 노드는 흑화소들로 구성된 영역 그리고 white 노드는 백화소들로 구성된 영역을 나타낸다. 트리의 상위 레벨에 위치한 노드들은 영상의 개략적인 형태를 나타내고, 트리의 하위 레벨에 위치한 노드들은 영상의 세세한 특성을 반영하므로, 이러한 계층적인 표현법은 계층적인 정합 과정에 효율적으로 사용될 수 있다.

이상과 같은 기호 묘사들을 표현하는 기법 역시 매우 다양한 형태를 취한다. 예를 들어, 항공 영상의 분석을 위한 SIGMA[17]에서는 각 영역의 특성을 프레임의 슬롯값으로 표현하였으며, 역시 항공 영상의 분석을 위한 SPAM[18]에서는 영역의 특성이 규칙의 전제부에 내장되도록 하였고, 문서 영상의 분석을 위한 ABLIS[19]에서는 흑판(blackboard)을 나타내는 그래프의 각 노드에 영역 및 경계의 특성이 등록되도록 하였다.

2.2.2 3차원 기호 묘사 형성을 위한 기술

3차원 기호 묘사는 주로 스케치 형태로 추출된 거리 정보 또는 속도 벡터들을 이용하여 물체의 표면을 나타내기 위하여 사용된다. 이를 위하여 일반적으로 사용되는 기법들은 일반화된 실린더(generalized cylinder), 표면 조각(surface patch), 확장된 가우시안 영상(extended gaussian image) 등을 생각할 수 있다.

입체 표현의 가장 일반적인 유형이 일반화된 실린더인데 이름에서 암시하듯이 일반화된 실린더는 실린더의 정의를 확장하여 생성되는 3차원 물체를 나타낸다. 실린더는 그 중심축(spine)을 따라 둥근 원반(disc)이 이동하여 만들어지는 입체로서 둥근 원반은 중심축에 항상 수직을 유지한다. 이러한 실린더를 일반화하기 위하여 다음과 같은 변화를 가한다. 중심축은 곡선이 될 수 있고, 횡단면(cross-section)의 형태는 둥근 원반이 아닌 임의의 형태를 취할 수 있고, 횡단면의 크기는 중심축을 따라 정의되는 특정한 이동 규칙에 의하여 변화될 수 있다. 일반화된 실린더를 사용하여 3차원 물체를 나타낸 예로서 참고 문헌 [20]을 들 수 있다. 즉, 평면의 빛을 투사하는 광원에 의하여 촬영된 그림 4-a와 같은 줄무늬 영상에서 유사한 방향성을 갖는 직선 성분들의 그룹을 형성한 후, 각 그룹에 속한 직선 성분들의 거리 정보들을 종합하여 그림 4-b와 같은 일반화된 실린더를 형성하였다.

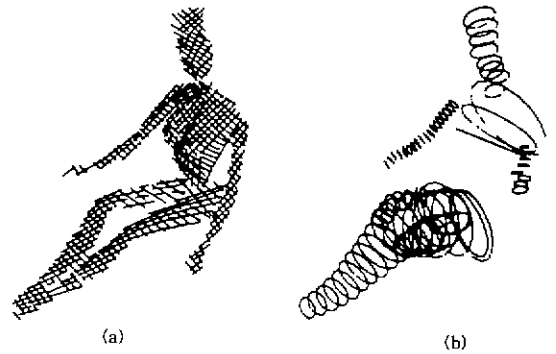


그림 4 일반화된 실린더에 의한 3차원 물체 표현

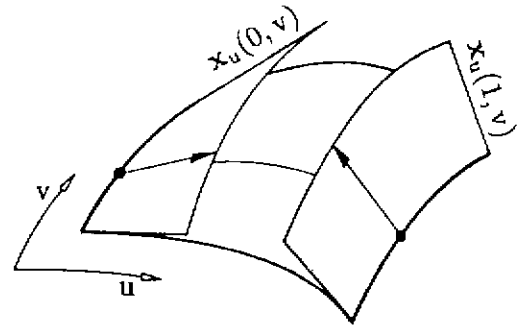


그림 5 쿤의 표면 조각에 의한 물체 표면 표현

표면 조각들의 집합에 의하여 물체의 표면을 표현하는 대표적인 방법은 쿤의 표면 조각(Coon's surface patch)을 사용하는 방법이다. 일반적으로 쿤의 표면 조각은 매개 변수 영역을 4각형으로 제한하여 사용한다. 따라서, 하나의 표면 조각은 4개의 경계 곡선들(bounding curves)과 표면의 경계 조건을 나타내는 혼합 함수(blending function)에 의하여 정의된다. 경계 곡선을 1차 함수로 정의하고 경계 조건을 표면이 경계에서 경계 곡선과 일치하는 조건으로 제한하면 쿤의 표면 조각은 그림 5과 같은 매우 단순한 양선형 표면으로 나타난다. 그림 5에서 경계 곡선 $x(t) = (x_x(t), x_y(t), x_z(t))$ 는 $0 \leq t \leq 1$ 에서 정의되는 3차원 공간상의 직선을 나타내고, x_0 와 x_1 은 경계 곡선의 양 끝점의 3차원 좌표를 나타낸다.

$$\text{경계 곡선} : x(t) = (1-t) \cdot x_0 + t \cdot x_1$$

$$\text{양선형 표면} : S(u, v)$$

$$= |1-u \ u| \cdot \begin{vmatrix} x(0, 0) & x(0, 1) \\ x(1, 0) & x(1, 1) \end{vmatrix} \cdot \begin{vmatrix} 1-v \\ v \end{vmatrix}$$

매끄러운 표면은 모든 점에서 탄젠트 평면(tangent plane)을 갖는다. 탄젠트 평면의 방위는 그 점에서의 표면 방위를 표현하는데 사용될 수

있다. 즉, 탄젠트 평면에 수직인 표면 법선(surface normal)을 이용하여 그 점에서의 표면 방위를 표현할 수 있다. 확장된 가우시안 영상은 이러한 표면 법선들의 분포를 나타내는 영상으로서 가우시안 구(gaussian sphere)라 불리는 단위구상에서 정의된다. 표면상의 한 점 x_{surface} 에 대응되는 가우시안 구의 한 점 x_{gaussian} 를 찾기 위하여 x_{surface} 의 표면 법선을 극각(polar angle) ζ 와 방위각(azimuth) η 로 표현한다. 구해진 ζ 와 η 를 갖는 가우시안 구상의 점이 x_{gaussian} 가 된다. 그림 6은 이러한 변환 과정을 보여준다. 확장된 가우시안 영상은 공간적 스테레오 정합을 통하여 구해진 거리 정보로부터 쉽게 계산될 수 있으며, 물체의 표면에 관한 정보를 함축적으로 나타낼 수 있다.

3. 컴퓨터 비전의 응용 분야

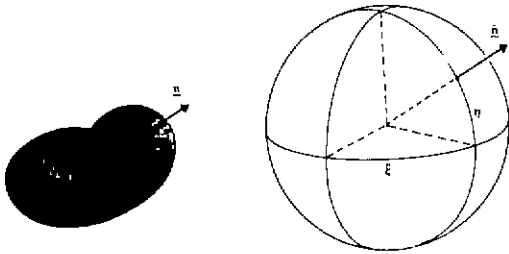


그림 6 표면 법선에 대응되는 가우시안 영상의 점

본 절에서는 컴퓨터 비전 시스템의 응용 분야에 대하여 기술한다. 컴퓨터가 인간처럼 보고 이해할 수 있다면 그 응용성은 대단히 클 것이다. 인간처럼 다양한 조건 및 여러 유형의 문제 영역에 적용될 수 있는 비전 시스템을 구축하는 작업은 아마도 불가능한 작업일 것이다. 그러나 특정한 문제 영역에 적용 가능한 비전 시스템의 구축을 위한 노력은 여러 분야에서 진행되고 있으며, 특히 산업용 시각 검사 시스템의 상품화는 활발히 추진하고 있는 실정이다.

3.1 산업용 시각 검사 시스템

시각 검사 시스템은 제품의 품질에 대한 시각적인 검사를 자동화하기 위하여 여러 유형의 제조 산업에 적용되고 있다. 특히 PCB(Printed Circuit Board)의 불량 유무를 확인하는 검사 시스템에 대한 연구는 매우 오랜 기간 동안 진행되어 왔으며, 최근에는 금속성 제품의 표면

검사, 제품의 분류 및 조립을 위한 시각 검사 등에 대한 연구가 활발히 진행되고 있다. 그림 7은 제품의 불량 유무를 판단하는 시각 검사 시스템의 기본 구조를 보여준다[21].

대부분의 산업용 비전 시스템은 실내에서 동작하므로, 조명 장치, 카메라의 위치 및 물체의 위치 등과 같은 주위 환경을 적절히 조절하는 작업이 가능하다. 환경을 제한하므로서 영상에 나타날 물체의 위치 및 형태에 대한 예측이 가능하고 영상 처리 작업 역시 비교적 쉽게 구현될 수 있다. 대부분의 비전 시스템이 그러하듯이 조명 장치는 강건한 시스템의 구축을 위하여 매우 중요한 역할을 담당한다. 조명 방법은 대상 물체의 후면에서 조명하는 후면 조명(back lighting) 그리고 대상 물체의 정면에서 조명하는 정면 조명(front lighting) 방법이 있다. 후면 조명은 물체와 배경을 분리하는 작업에 적합하고 특히 반투명 물체에 대한 검사 작업에 효율적이며, 정면 조명은 제품 표면의 이상 유무를 검사 하는데 주로 사용된다.

물체의 특성을 표현하는 모델은 주로 2차원적인 형태 분석에 의하여 구현된다. 예를 들어 표면의 텍스처(texture), 직선 및 원으로 구성된 도형의 형태, 분할된 영역 특히 blob의 형태를 분석하고, 분석된 데이터들을 벡터(vector) 또는 그래프 형태로 표현한다. 예를 들어, 참고 문헌 [22]에서는 광섬유 연결자(optical fiber connector)의 표면 검사를 위하여 에지 검출기를 사용하여 선조각(line segments)들을 추출한 후, 추

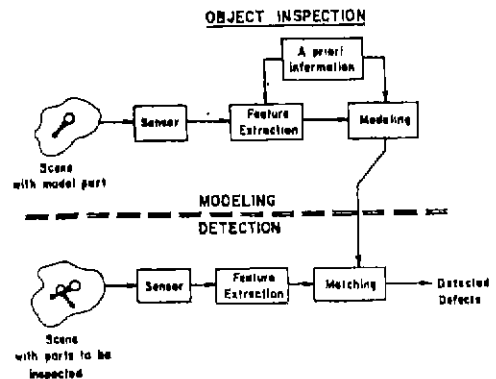
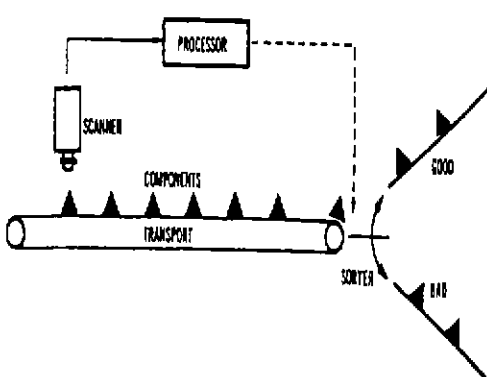


그림 7 시각 검사 시스템의 기본 구조

출된 선조각들을 모델과 비교하는 과정을 통하여 결합 유무를 판정한다.

3.2 항공 사진 분석 시스템

항공 사진을 분석하여 지형을 분석하고 토지의 사용도를 평가하는 작업은 매우 오랜 기간 동안 연구되어온 분야이다. 이러한 연구는 국방 과학 분야뿐 아니라, 항만 건설, 도시 계획, 국토 개발 및 지도 제작을 위하여 응용될 수 있는 분야이다. 대표적인 항공 사진 분석 시스템으로서 OPS5 도구를 이용하여 구현된 SPAM[18]과 프레임 형태의 지식을 이용하여 분석 작업을 수행하는 SIGMA[17]를 생각할 수 있다. 특히 SIGMA 시스템에서는 가설을 생성하고 검증하는 과정을 통하여 영상 해석 작업을 수행한다. 그림 8는 물체간의 공간적인 관계 그리고 이미 예시화된 물체를 이용하여 인접 물체에 대한 가설을 생성하고 검증하는 과정을 보여준다. 예를 들어, 물체간의 공간적인 관계 $f()$ 와 이미 예시화된 물체 S에 의하여 가설 $f(S)$ 가 생성되었다. 유사한 방법으로 물체간의 공간적인 관계 $h()$ 와 이미 예시화된 물체 U에 의하여 새로운 가설 $h(U)$ 가 생성되었다. 이제, 가설 $f(S)$ 와 가설 $h(U)$ 가 통합될 수 있는 가를 확인하여, 만약 이들이 통합 가능하면 통합된 가설 $f(S) \cup h(U)$ 를 생성한다. 생성된 통합 가설은 이제 물체 S와의 공간적인 관계 $f(S)$ 를 만족하고 동시에 물체 U와의 공간적인 관계 $h(U)$ 를 만족하는 물체 T를 추출하기

위하여 검증된다. 통합 가설 $f(S) \cap h(U)$ 는 다음과 같이 제약 조건들의 집합으로 표현된다.

- (AND (EQUAL OBJECT-TYPE ROAD)
- (AND (LESSP TOTAL-LENGTH 100)
- (GREATERP TOTAL-LENGTH 50))
- (AND (LESSP AVERAGE-WIDTH 15)
- (GREATERP AVERAGE-WIDTH 10))
- (AND (LESSP AVERAGE-DIRECTION 50)
- (GREATERP AVERAGE-DIRECTION 30)))

3.3 문서 영상 분석 시스템

문서 영상 분석을 위한 연구 역시 매우 오랜 기간 동안 연구되어온 분야로서 특히 광학 문자 인식 분야는 이미 어느 정도 만족할만한 성과를 거두고 있다[23]. 문자 인식을 위한 문서 영상은 주로 스캐너(scanner)를 통하여 입력되며, 입력된 영상에서 문자 부분만을 추출하기 위하여 구조 분석 작업이 선행된다. 예를 들어, 문서 영상에서 문자열 부분을 추출하기 위한 히스토그램 분석 작업, 입력된 문서가 기울어진 경우 이를 바로 잡기 위한 문자열의 방향성 분석 작업, 개별 문자 추출을 위한 글자 간격 분석 작업 등이 수행된다. 추출된 개별 문자를 인식하는 방법은 통계적인 방법, 구조 분석 방법, 구문 분석 방법 등으로 분류할 수 있으며, 각 인식 방법에 따라 사용되는 특징의 성격이 결정된다.

문서의 구조를 분석하는 대표적인 시스템으로

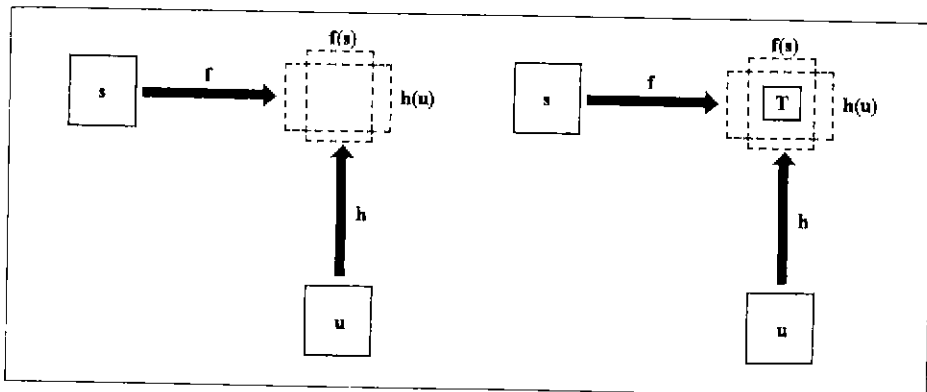


그림 8 가설 생성 및 검증 과정

서 우편물에 대한 영상을 분석하여 수취인 주소 영역을 추출하는 ABLIS[19] 시스템을 들 수 있다. 이 시스템은 기본적으로 흑판, 지식원 그리고 제어 모듈로 구성되는 흑판 모델(blackboard model)을 이용하여 분석 작업을 수행한다. 그림 9는 ABLIS 시스템의 기본 구성도를 보여준다. 흑판은 전역적 데이터베이스로서 현재의 영상 분석 상황을 묘사해 준다. 초기 처리 단계에서 얻어진 각 블록의 특성에 관한 정보 및 각 블록이 수취인 주소 블록이 될 수 있는 가능성을 나타내는 신뢰도 등이 흑판위에서 점진적으로 갱신된다. 각 지식원은 프로시듀어나 규칙의 집합 또는 논리 연산자로서 표현되고 선행 조건이라 불리는 자신이 활성화될 수 있는 조건을 갖는다. 제어 모듈은 현재의 흑판의 상황을 분석하여 적절한 지식원을 선택하고, 선택된 지식원을 흑판위의 각 블록에 적용하여 그들의 신뢰도를 갱신하고,

갱신된 블록이 수취인 주소 블록으로서의 조건을 만족하는가를 확인한다.

3.4 이동 물체 분석 시스템

움직이는 물체의 탐지 및 추적 그리고 인식에 대한 연구는 주로 의용 산업 분야와 국방 산업 분야에서 활발히 진행되어 왔으며 최근에는 보안 감시를 목적으로하는 시스템들이 소개되고 있다. 예를 들어, 보안이 요구되는 특정한 지역을 감시하는 카메라의 시계(field of view)에 움직이는 물체가 포착되면, 이에 대한 분석 및 추적 작업을 수행하는 시스템을 생각할 수 있다. 이러한 분야의 연구는 기본적으로 동영상의 분석을 요구한다. 즉, 연속적으로 촬영된 영상들을 비교하여 움직이는 물체를 확인하는 작업, 확인된 물체에 대한 분석 작업, 그리고 추적 작업을 수행하기

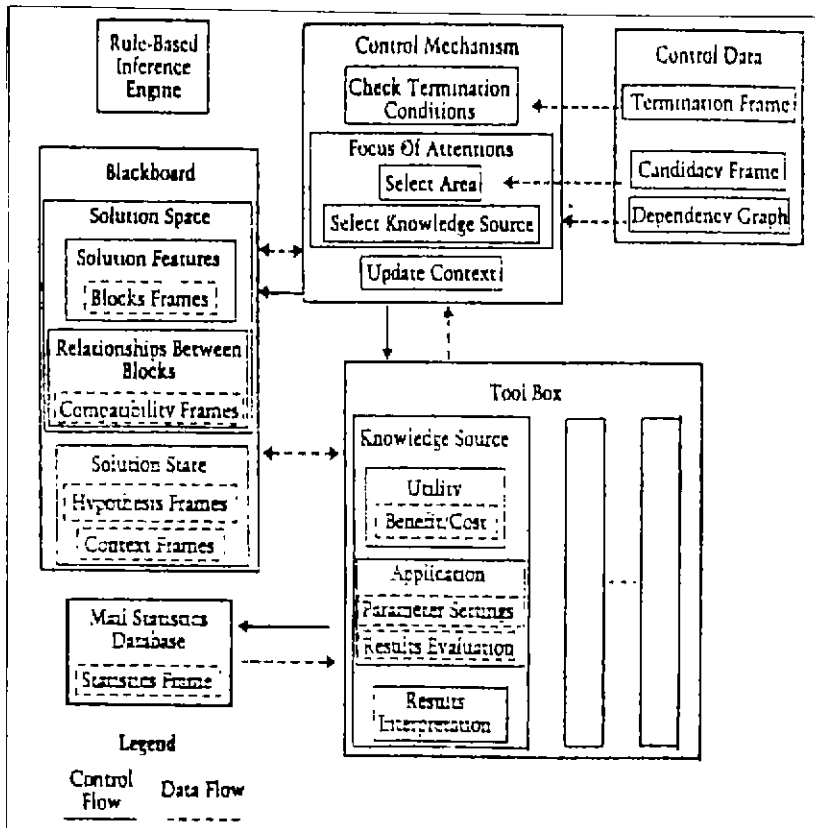


그림 9 ABLIS의 기본 구성도

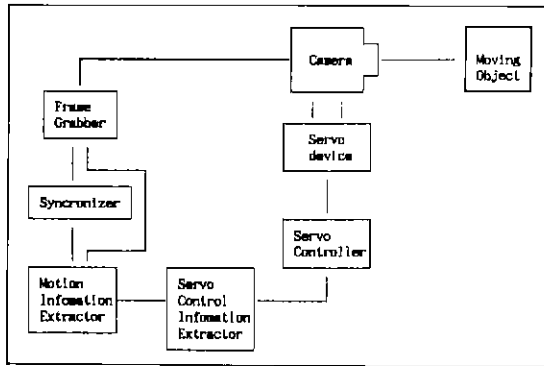


그림 10 실시간 이동 물체 추적 시스템

위하여 서보 장치(servo device)를 구동하는 작업 등이 요구되며, 특히 이러한 일련의 작업들을 실시간에 수행하기 위하여는 병렬 처리 알고리즘과 같은 다양한 영상 처리 작업을 필요로 한다.

그림 10은 참고문헌 [24]에서 소개한 실시간 이동 물체 추적 시스템의 개요도이다. 동기화기(synchronizer)는 이동 물체의 진입을 탐지하고 물체의 이동 속도에 적합한 영상 촬영 간격 Δt 를 결정한다. 동작 정보 추출기(motion information extractor)는 이동 물체의 속도 및 위치 정보를 추출하여 이를 서보 제어기(servo controller)에 전달한다. 서보 제어기는 전달받은 동작 정보를 좌우각(pan angle)과 상하각(tilt angle)로 변환하여 서보 장치(servo device)에 전달한다. 서보 장치를 전달받은 각도에 의하여 카메라의 시계(field of view)를 상하 및 좌우로 조정하여 추적중인 물체가 항상 영상의 중앙에 위치하도록 한다.

4. 결 론

본 고에서는 주어진 장면에 대한 영상을 이해하기 위하여 필요한 기본적인 처리 기술들을 설명하고, 이러한 기술들을 활용한 응용 분야를 살펴보았다. “보고 이해하는 범용 시스템의 구현” 이것이 컴퓨터 비전의 궁극적인 목표일 것이나, 이러한 목표는 현재의 과학 기술로는 물론이고 가까운 장래에서 실현되기 어려울 것이다. 그러나 특정한 분야에 특수한 목적을 구현

하기 위한 비전 시스템은 속속 출현되고 있는 상황이며, 특히 제조 산업 분야에서 생산 공정을 자동화하기 위하여 비전 시스템을 구축하고 있는 실정이다.

실제 현장에서 적용될 수 있는 강인한 비전 시스템을 구축하기 위하여는 패턴 인식, 퍼지 이론, 신경 회로망 이론 등과 같은 관련 분야의 연구가 접목되어야 하며, 특히 영상 분석의 초기 단계인 영상 분할(segmentation) 기법에 대한 좀더 심도있는 연구가 수행되어야 한다고 여겨진다. 영상 분할 단계의 중요성은 아무리 강조해도 지나치지 않다. 이 단계의 작업이 성공적으로 수행되면 다음 단계의 작업이 매우 쉬워진다. 대부분의 비전 시스템이 실세계에서 적용되지 못하는 이유는 분할 단계의 작업이 원활히 수행되지 못하기 때문이다.

참고문헌

- [1] B. K. P. Horn, Robot Vision, The MIT Press, 1986.
- [2] N. Yan and J. Keller, “Conditional Fuzzy Measures and Image Segmentation,” Proceeding of NAFIPS-92, pp. 32-36.
- [3] C. Wang and S. Srihari, “A Framework for Object Recognition in a Visually Complex Environment and its Application to Locating Address Blocks on Mail Pieces,” Int. Journal of Computer Vision, Vol. 2, 1988, pp. 125-151.
- [4] G. Medioni and R. Nevatia, “Segment-based Stereo Matching,” Computer Vision, Graphics, and Image Processing, 1985, pp. 2-18.
- [5] W. L. Grimson, “Computational Experiment with a Feature-based Stereo Algorithm,” IEEE. PAMI-7, 1985, pp. 17-33.
- [6] W. E. L. Grimson, From Images to Surfaces, The MIT Press, 1981.
- [7] W. Hoff and N. Ahuja, “Extracting Surfaces from Stereo Images: An Integrated Approach,” CH2465-3, 1987, pp. 284-294.
- [8] Y. Shirai, Three-dimensional Computer Vision, Springer-Verlag, 1987, pp. 122-140.
- [9] A. C. Kak et al., “Knowledge-based Stereo and Structured Light for 3D Robot Vision,” in Te-

chniques for 3D Machine Perception, edited by Rosenfeld, North-Holland, 1986, pp. 185-218.

[10] K. Kanatani, "Structure and Motion from Optical Flow under Perspective Projection," CVGIP, Vol. 38, 1987, pp. 122-146.

[11] R. Jain, "Extraction of Motion Information from Peripheral Process," IEEE PAMI-3, No. 5, 1979, pp. 489-503.

[12] L. Chen and S. Chang, "A Video Tracking System with Adaptive Predictor," Pattern Recognition, Vol. 25, No. 10, 1992, pp. 1171-1180.

[13] K. Karmann, "Time Recursive Motion Estimation using Dynamic Models for Motion Prediction," 10-th Int. Conf. on Pattern Recognition, Vol. 1, 1990, pp. 268-270.

[14] H. Freeman, "Computer Processing of Line Drawing Images," Computer Surveys, Vol. 6, 1974, pp. 57-98.

[15] E. Persoon and K.S. Fu, "Shape Discrimination using Fourier Descriptors," IEEE T. on Sys. Man and Cybern., Vol. 7, 1977, pp. 269-281.

[16] H. Samet, "Neighbor Finding Techniques for Images represented by Quadrees," CVGIP, Vol. 18, 1982, pp. 37-57.

[17] V. Hwang, "Hypothesis Integration in Image Understanding System," CVGIP, Vol. 36, 1986, pp. 321-371.

[18] D. M. McKeown et al., "Rule-based Interpretation of Aerial Imagery," Proc. of IEEE Workshop on Principles of Knowledgebased Systems, 1984.

[19] S. Srihari and C. Wang, "Recognizing Address Blocks on Mail Pieces," AI magazine winter 1987, pp. 25-40.

[20] G. J. Agin and T. O. Binford, "Computer Description of Curved Objects," IEEE T. on Computers, Vol. 25, 1976, pp. 439-449.

[21] R. Chin, "Algorithms and Techniques for Automated Visual Inspection," Handbook of Pat-

tern Recognition and Image Processing, edited by T.Y Young and K. S. Fu, Academic Press, 1986, pp. 588-612.

[22] B. Shahraray et al., "Defect Detection, Classification and Quantification in Optical Fiber Connectors," Proc. of Machine Vision Applications, 1990, pp. 15-22.

[23] Y. Y. Tang et al., "Document Analysis and Understanding: A Brief Survey," Proc. 1st Conf. on Document Analysis and Recognition, 1991, pp. 17-31.

[24] H. I. Choi, G. Y. Kim and B. K. Kim, "Real-time Tracking of Moving Object," Proc. of 3rd PRICAI, 1994, pp. 843-849.

김 명 원



1972 서울대학교 공과대학 응용수학(학사)
 1981 Univ. of Massachusetts (Amherst), computer science(석사)
 1986 Univ. of Texas(Austin), computer science(박사)
 1975 ~ 1978 한국과학기술연구원 연구원
 1982 ~ 1985 Univ. of Texas (Austin), Institute for Computing and Computer Application, 연구원
 1985~1987 AT & T Bell Labs. 연구원
 1987~1994 한국전자통신연구소, 기초기술연구부, 책임연구 책임연구원
 1994~현재 숭실대학교 인공지능학과 부교수
 관심분야 : 인공지능, Neural Network, Theory of Computation, Software Engineering

최 형 일



1979 연세대학교 전자공학과 공학사
 1982 미시간대학교 전산공학과 공학석사
 1987 미시간대학교 전산공학과 공학박사
 현재 숭실대학교 전자계산학과 부교수 재직중
 관심 분야 : 컴퓨터비전, 퍼지 및 신경망이론, 패턴인식, 지식기반시스템 등