

컴퓨터 비전(Computer Vision)

박래홍

(서강대학교 전자공학과 조교수)

1. 서론

우리는 매일 우리가 받아들이는 외부 정보의 약 70~80%를 시각을 통하여 받아들이고, 의식적인 혹은 무의식적인 처리과정을 거쳐 외부 상황을 이해하고, 이에 따른 여러가지 반응을 행하게 된다. 백 번 듣는 것보다 한번 보는 것이 낫다는 말이 있듯이 인간은 아무리 복잡한 경치나 상황이라 할지라도 한번 보면 쉽게 이해하고 인식할 수 있지만 이것을 말이나 글로 표현하려면 그 장면에 맞는 많은 묘사와 서술이 불가피하게 되며 실사 복잡하고 길게 표현하더라도 불충분한 경우가 많게 된다. 이렇듯 사람의 눈 혹은 비전 시스템(vision system)을 통해 얻은 영상은 실로 방대한 양의, 3차원 세계에 대한 정보를 포함하고 있다.

1970년대 들어 반도체, 마이크로 프로세서, 컴퓨터 기술들이 발전함에 따라 디지털신호처리 기술은 급격히 성장하여 왔으며 인간이 행하던 일을 기계가 대신 해내려는 시도가 급기야는 방대한 데이터 양을 갖는 영상을 처리하는 영상처리^{1)~4)}, 컴퓨터 비전^{5)~9)}, 패턴인식^{10)~13)}, 영상이해^{9), 13)} 등 영상처리 영역에까지 미치게 되었다.

컴퓨터등과 같이 사람이 만든 기계가 사람이나 동물과 같이 센서(sensor)를 통하여 입력영상으로부

터 추출한 정보에 의해 주위 사물이나 환경을 인지하고 거기에 따르는 반응을 행하는 것을 다루는 분야를 컴퓨터 비전(computer vision) 혹은 기계 지각(machine perception)이라 한다. 컴퓨터 비전에서 다루는 것으로는 직선 혹은 곡선 성분 및 물체의 둘레 추출, 영역분할(region segmentation), 물체의 둘레에 의한 물체 인식(recognition), 분류(classification) 및 정합(matching), 시변영상에서의 움직이는 물체의 탐지 및 추적, 텍스처(texture) 해석, 그리고 화상 해석등을 들 수 있다. 이러한 컴퓨터 비전의 응용분야로는 로봇트 비전 및 제품 검사등과 같은 산업용, 문자인식, 지문인식 군사용, 추적기, 원격감지, 항해 및 정찰, 의학, 그리고 생물학등이 있다.

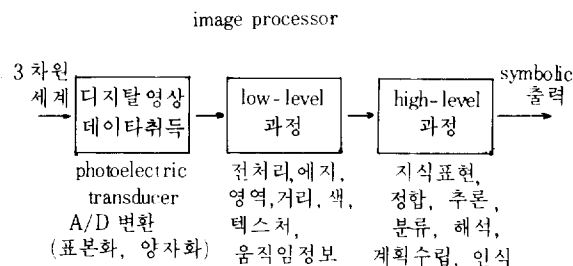


그림 1. 컴퓨터 비전 시스템의 개요

특집 : 회로 및 신호처리

컴퓨터 비전은 그림 1에서와 같이 영상 데이터 취득, low-level처리 과정 및 high-level 처리 과정등으로 크게 나눌 수 있다. 영상 데이터 취득 부분은 실제 3차원의 물체나 장면의 광신호 혹은 음향신호에 의한 영상을 전기적 신호로 바꾸어 주고 이를 디지털화시키는 스캐너(scanner)나 디지털라이저(digitizer)를 거쳐 디지털 컴퓨터가 처리하기 알맞는 디지털 영상신호로 바꾸어 주는 과정이다. Low-level 처리 과정은 잡음제거, 영상개선이나 복원등의 전처리과정을 거쳐서 에지(edge)검출, 영역분할, 색, 거리, 움직임, 텍스처 해석등을 통해 high-level에 필요한 데이터를 제공해 주는 영상 데이터 계산과정이라 할 수 있다. High-level과정은 low-level 정보로부터 얻은, 같은 특성을 갖는 화소들을 묶어(aggregation) 얻은 물체들의 전체 영상내에서의 상호 연관관계를 고려하여 3차원 영상정보를 표현, 정합, 인식, 추론 및 계획수립 과정등을 거쳐 3차원 영상정보를 이해하고 해석하여 필요한 출력을 내는 과정이라 볼 수 있다. 일반적으로 영상신호처리에서의 입력 및 출력은 영상인 반면, 컴퓨터 비전에서는 디지털 입력 영상을 받아 들여 low-level 및 high-level 처리과정을 거쳐 직선이나 곡선과 같은 특징이나 인식한 물체의 특징등을 가지고 표현한 실제 3차원 세계의 symbolic출력인 경우가 많다.

본 고에서는 서론에 이어 영상 데이터 취득 과정에 대해 간단히 알아보고, 디지털 신호처리 기술을 이용한 low-level 처리과정에 대해 주로 설명한 후, high-level 처리과정에 대해 간단히 언급하고 결론을 맺는다.

2. 영상 데이터 취득

사람의 눈은 파장이 380nm(자외선)~700nm(적외선)사이 에 있는 빛에 민감하며 사람의 눈을 통해 들어오는 아날로그(analog)입력 영상신호는 광원의 공간 및 주파수 특성뿐 아니라 물체의 표면 특성이 고려된 반사 혹은 투과 과정, 관찰시 보는 각도 및 빛 에너지를 탐지하는 센서의 특성에 따라 주어지게 된다.¹⁾ 사람 눈의 빛의 밝기에 대한 반응은 들어오는 빛의 밝기의 로그함수로 주어진다고 실험적으로 알려져 있다. ¹⁾⁻²⁾ 이렇게 빛 에너지가 전기

신호로 바뀌어진 후 아날로그 영상신호를 스캐너나 디지털라이저를 거쳐 디지털 영상신호로 바꾸어 주는 아래의 과정들이 필요하게 된다.

많은 정보량을 가지고 있는 화상을 디지털화시켜 표현, 저장 및 전송하기 위해서는 엄청난 양의 메모리가 필요하다. 먼저 아날로그 화상을 디지털 컴퓨터가 받아들여 처리가 가능하려면 유한개의 값으로 이산화(discretization)시켜야 한다. 즉 좌표평면상에서의 해상도와 관계가 있는 표본화(sampling)과정과 각 화소(pixel 혹은 picture element)에서의 빛의 밝기(intensity)를 유한개의 값으로 대응시키는 양자화(quantization) 과정을 거쳐 아날로그 신호를 디지털 신호로 바꾸어 표시하게 된다.³⁾⁻⁴⁾ 표본화 과정 및 양자화 과정 모두 균일(uniform)하게 혹은 불균일(nonuniform)하게 행할 수 있으며 표본화 과정에서 화소를 사각형, 삼각형, 혹은 육각형 화소등으로 표현할 수 있으나 계산상 편리하므로 보통 직사각형 혹은 정사각형 화소를 많이 사용한다.

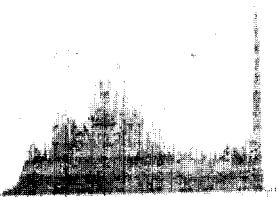
주어진 화상을 $N \times N$ 개의 화소로 표본화하고 각 화소의 밝기를 m 비트(bit) (2^m 개의 레벨)로 양자화했다고 하면 화면 하나를 디지털화하여 표현하기 위해서는 총 $N * N * m$ 비트가 필요하게 된다. 예로 $N=512$, $m=8$ 인 경우, 정지 화상 하나를 표현하기 위해서는 약 2M비트라는 엄청난 양의 메모리가 필요하게 되며, 또한 응용에 및 입력 화면에 따라 N 과 m 의 선택은 다르지만 우리 눈이 느끼기에 만족할만한 적절한 N 과 m 의 값은 상호 관련이 있으며, 물론 주어진 화상의 종류에 따라 다르지만 실험 결과 일반적으로 원래의 아날로그 화상을 이산화시켜 어느 정도 충실하게 표현하기 위해서는 N 은 128-256이상, m 은 6-7이상이 적합하다고 한다.⁵⁾

또한 앞에서 언급한 것처럼 한 정지화상을 표현하기 위한 이산화 과정뿐 아니라 시간에 따라 변화하는 시변영상(time varying imagery)인 경우 시간축에 대해서도 이산화가 필요하게 된다. 예로 NTS C(National Television Systems Committee) 텔레비전의 경우와 같이 초당 30프레임(frame)을 보낸다고 하면, $N=512$, $m=8$ 로 가정할 경우, 초당 약 60 M비트가 필요하게 된다. 이렇게 하여 공간축, 밝기축 및 시간축에서 이산화과정을 거쳐 표현된 디지털

탈 입력영상은 다음 장에서 설명할 low-level 처리과정을 거치게 된다. $N=256$, $m=8$ 인 경우의 디지털 영상 데이터의 예를 그림 2(a)에 보였으며 이 영상의 밝기 분포인 히스토그램(histogram)을 그림 2(b)에 보였고, 그림 2(c)에서는 그림 2(a)영상을 좌표축에서 64×64 로 재표본화하여, 각 화소의 빛의 밝기를 256레벨로 양자화한 디지털 영상을 3차원으로 그려 나타내었다.



(a) 교회 화상



(b) 히스토그램



(c) 3차원 표시

그림 2. 디지털 영상신호

3. Low-level 처리

입력화상의 화질에 따라, 시각적 효과를 좋게 하거나 혹은 컴퓨터의 효율적인 처리를 위해서, 잡음을 감소시키거나 화질 개선(image enhancement) 혹은 복원(image restoration) 등의 전처리 과정을 거쳐 컴퓨터 비전을 위한 low-level 처리과정을 거치게 된다. Low-level 처리 과정에서는 디지털 영상신호를

입력으로 디지털 신호처리기술을 사용하여 좌표 평면에서, 혹은 Fourier 변환등을 통해 주파수 평면에서 필요한 계산과정을 수행하여 물체의 둘레, 분할된 영역, 움직임 정보, 텍스처, 색 정보등과 같은 특징(feature)을 입력영상에서 추출해내고 분할된 각 영역을 표현하는 작업을 수행하게 된다. 이러한 특징추출 과정은 그 다음의 high-level 처리과정의 성공여부를 좌우하게 되는 중요한 단계이다.

본 장에서는 빛의 밝기 혹은 색의 급격한 변화라 할 수 있는 에지 추출과정, 비슷한 성질을 가지고 있는 화소들끼리 묶는 영역 분할과정, 화상의 규칙적인 배열 혹은 불규칙한 배열등을 고려하는 텍스처 해석과정 및 시변영상에서 움직이는 물체의 탐지 과정등에 대한 기초적인 이론을 설명하고, 이러한 에지, 분할된 영역, 텍스처 및 움직임 정보등의 특징을 고려하여 묶은 2차원 혹은 3차원 구조들의 정량적 표현방법에 대해 간단히 설명하려 한다.

3.1 에지 검출

에지란 빛의 밝기 혹은 색의 불연속적인 급격한 변화가 있는 점들을 의미하며 화상에서 모양정보(shape information)를 이해하는데 있어 중요한 요소가 된다. ^(4-8, 10-17) 이것은 물체의 표면 성질에 따라 생기거나, 물체에서 면의 교차에 의해 생기거나, 광원에 의한 그림자 때문에 생기거나, 혹은 표본화시 생기는 오차등에 의해 생긴다. 실제 화상에서 에지는 위에서 언급한 여러 복합적인 요인에 의해 생기게 되어 정확한 에지를 구하기가 어려운 경우가 많다.

에지를 구하는 것을 수학적으로 표현하면, 에지를 찾는 것은 빛의 밝기등의 급격한 변화 위치를 찾는 것이므로 미분을 하여 최대, 최소인 점을 찾거나 두 번 미분하여 영교차(zero crossing) 위치를 찾으면 된다. 이산적으로 표현된 디지털 영상에서는 에지 추출 연산자를 차 연산자(difference operator)에 근거하여 에지의 각 방향에 따라 여러개의 마스크(mask)나 템프리트(template), 혹은 라플라시안(Laplacian) 연산자등으로 표현할 수 있다.

대표적인 에지 검출 방법은 우선 여러 방향의 에지 마스크¹⁵⁾ 예로 수직, 수평, 대각선 및 반대각선 방향의 에지 추출용 마스크를 씌어 얻은 각 방향

특집 : 회로 및 신호처리

에지 값들의 제곱의 합의 제곱근값, 혹은 계산상 편하게 각 방향 에지 절대값의 합, 또는 절대값의 최대값등의 비선형 계산을 거친 후, 임계값보다 큰지 작은지에 따라 그 점이 에지인지 아닌지를 나타내는 2치 (bilevel) 에지 맵 (map) 을 얻게 된다. 임계값은 전체적인 (global) 임계값, 혹은 국부적인 특성을 고려하여 국부 적응 (locally adaptive) 임계값을 사용한다. 대표적인 에지 연산자로는 Roberts 연산자, Sobel 연산자, Prewitt 연산자, Kirsch 연산자등이 있으며, 간단한 1차원 차 연산자 및 수직, 수평방향의 Sobel 연산자와 $256 \times 256 \times 8$ 비트로 표시된 Girl 원래영상 및 이 영상에 Sobel 연산자를 취해 임계값 = 25로 하여 구한 에지를 그림 3에 보였다.

-1	0	1
----	---	---

-1
0
1

(a) 1차원 차 연산자

-1	0	1
-2	0	2
-1	0	1

-1	-2	-1
0	0	0
1	2	1

(b) 2차원 Sobel 연산자



(c) 원래영상



(d) 추출된 에지

그림 3. 에지 연산자 및 추출된 에지

이밖에도 각 방향 마스크로 최소 오차가 되게 근사시키는 Hueckel 연산자, 눈의 평활화 (smoothing) 특성을 고려하여 Gaussian 필터링 후 에지 검출을 행하는 방법¹⁶⁾ 이차미분의 영교차점을 찾는 방법¹⁷⁾, 마스크들 사이의 직교성을 이용하여 위에서 설명한 예처럼 크기에 따라 에지를 판별하는 것이 아니라 에지 subspace와 이루는 각에 따라 에지를 결정하는 Frei-Chen 연산자 및 방향성이 없는 라프라시안 연산자등이 있다. 또한 위에서 언급한 방법 이외에도 물체의 모델을 가정하거나 통계적인 특성을 고려한 에지 검출방법, 순차적으로 화소를 주사해

가며 비용 (cost) 을 정의하여 제한조건에 맞게 탐색을 하는 순차적인 (sequential) 에지 검출방법 및 응용에 따라 3차원 영상정보에 적용이 가능한 3차원 에지 검출 방법등이 있다.

이렇게 구한 에지는 화상에서 급격한 변화를 나타내는 점이기에 때문에 물체의 둘레 (boundary) 를 구하거나, 물체 둘레 모양에 따라 물체 인식 및 분류를 행하는데 이용되거나, 비슷한 성질을 가진 영역으로 나누는 영상 영역화에 이용되거나, 텍스처 해석에 응용되거나, 비슷한 성질을 갖는 부분과 둘레 부분을 나누어 따로 코딩 (coding) 하는 texture-contour 코딩과 같은 높은 영상 감축율을 갖는 소위 제 2세대 코딩¹⁸⁾ 등에 이용된다.

3.2 영역 분할

영역 분할이란 비슷한 특징 (예로 밝기, 색, 텍스처, 움직임등) 을 갖는 의미있는 화소들끼리 묶어서 전체 화상을 이러한 비슷한 성질을 갖는 영역들로 분할하고 각 영역을 labeling 하는 것을 말한다.^{4) - 8)} ^{19) - 21)} 영역을 나누는 것과 에지를 구하는 것은 dual 인 관계에 있으며, 100% 완전하게 만족할 만한 분할된 영역을 얻거나 에지를 잘 찾아내는 경우는 거의 없으므로 이 둘을 서로 결합하여 쓰기도 한다.

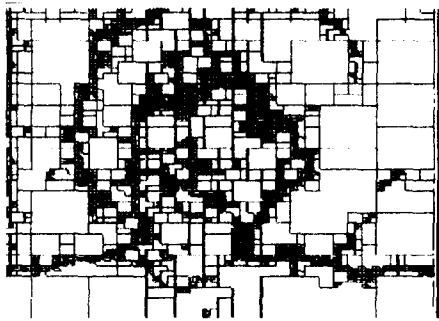
영역분할 방법은 비슷한 화소끼리 묶어 나가는 bottomup 방식인 merge 방법과 다른 성질을 가진 영역을 나누어 나가는 top-down 방식인 split 방법으로 크게 나눌 수 있으며 이러한 low-level 분할 방법에서는 화상의 모델 혹은 물체의 사전지식은 이용되지 않는다. 더 나아가, 이 둘의 결합형태인 split- and -merge 방법, 전술한 에지 정보에 근거를 둔 방법 및 3차원 물체의 모델, 사전지식 및 그림자 정보등을 분할에 이용하는 high-level 분할 방법²⁾ 등이 있다.

merge 방법은 예로 빛의 밝기의 균일성이나 텍스처 특징이나 영역 둘레의 길이 및 강도를 고려하여 전체적인 혹은 국부적인 임계값에 따라 합쳐줄 것인지의 여부를 결정한다. 그림 4에 256×256 크기의 8비트로 양자화된 원래 영상 및 빛의 밝기의 최대와 최소의 차가 40보다 작은 경우 합쳐서 영역분할된 결과를 보였다.

Split 방법은 균일하지 않은 영역을 계속 4 부분으



(a) 원래 화상



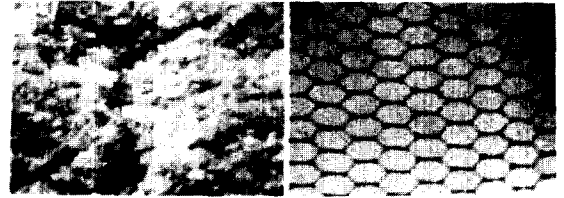
(b) 분할된 화상

그림 4. 영상 영역화

로 잘게 나누어 나가는 quad-tree 방법, 화상의 통계적인 특성을 나타내는 히스토그램의 분포에 따라 나눌 것인지 나누지 않을 것인지를 결정하여, 여러 빛의 밝기 영역을 가진 multimodal 히스토그램의 경우 임계값을 설정하여 각 영역이 unimodal 히스토그램 특성을 갖게끔 갈라나가는 thresholding 방법 등이 있다.^{20), 21)} 물론 임계값은 전체적인 특성에 따라 전체적인 고정 임계값을, 혹은 화상을 여러개의 부영상(subblock)으로 나누어 각 부영상마다 히스토그램 특성에 의해 나누어 주는 국부 적응 임계값을 사용할 수도 있다. 또한 예로 다차원 히스토그램으로 표현할 수 있는 색영상의 경우 각 특징의 완전한 피크(peak)가 없어 더 이상 분할할 수 없을 때까지 계속 분할하는 recursive 분할방법도 있다.

3.3 텍스처

2차원 영상 필드(field)의 해석 및 분류하는데 있어 중요한 특징이라 할 수 있는 텍스처^{5), 8), 22), 23)}는 앞에서 언급한 예지나 영상분할 방법으로는 쉽게 설명할 수 없는 물체의 구조 혹은 그들의 결합관계등을 잘 나타내주는 물체의 표면에 대한 조직구조를



(a) 바위

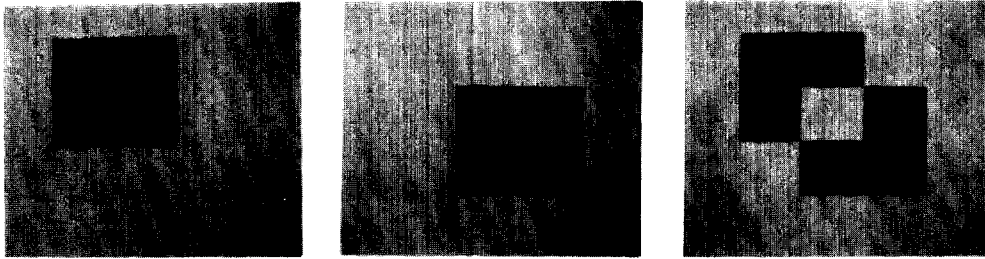
(b) 그물

그림 5. 텍스처

말하는데 여기에는 물체들 사이에 반복성이 없는 불규칙(random) 텍스처와 반복적인 형태를 가지는 규칙(regular) 텍스처가 있다. 불규칙한 텍스처의 예로는 잔디, 나무, 모래, 바다 파도등과 같은 자연이 있으며 규칙적인 텍스처로는 벽, 벽지, 타일, 직물등과 같은 인공적인 것들이 있다. 텍스처의 예로 그림 5에 불규칙적인 바위 및 규칙적인 그물 화상을 보였다.

텍스처 정보로부터 물체 표면에 대한 정보 및 표면의 방향성등의 정보를 얻을 수가 있는데 텍스처의 추상적인 개념(예로 얼만큼 거친지, 직선에 가까운지, 굵기가 굵은지, 불규칙한지, 방향성이 있는지등)을 정량화시키는 과정이 필요하게 된다. 이러한 과정은 텍스처의 특징을 정의하고 그 특징의 히스토그램 혹은 2차원 어레이(array)등을 고려하여 여러가지 정량적인 수치의 도입을 통해 텍스처의 해석 및 분류에 이용하게 된다.^{23), 22)} 텍스처 해석에서는 사람 눈의 특성상 1차 및 2차 통계적 특성이 중요한 역할을 하는 것으로 알려져 있다.

텍스처 영상을 해석하는 방법은 크게 두가지로 나눌 수 있는데 첫째는 통계적인(statistical) 해석방법이며 둘째는 구조적인(structural) 해석방법이다. 전자는 과거에 많이 쓰이던 방법으로 autocorrelation 방법, 여러가지 디지털 변환 방법 및 전력 밀도함수, 같은 빛의 밝기의 계속되는 화소 길이인 run-length, cooccurrence 행렬, Markov 모델, 좌표평면에서의 빛의 밝기 차의 분포등과 같은 영상 전체의 특성을 사용하여 텍스처의 특성을 정량적으로 표시하는 방법으로 불규칙 및 규칙 텍스처 해석에 사용된다. 후자는 최근에 활발히 연구되는 방법으로 텍스처를 이루는 기본 구조(primitive)의 추출과 이들의 공간상의 연결관계를 이용함으로써 텍스처를 해석하는 방법으로 주로 규칙적인 텍스처 해석에 쓰인다.



(a) 첫번째 프레임 (b) 두번째 프레임 (c) (a) 와 (b) 의 차화상

그림 6. 움직이는 물체를 포함하는 2 프레임 영상

텍스처 해석은 x선 사진, 조직사진등에서와 같이 의학적 판단이나 인공위성사진에서의 지형 분류, 영상분할, 지문인식등에 쓰인다. 또한 텍스처 해석에서 얻어진 특징들을 이용하여 텍스처 합성²⁴⁾을 행할 수가 있는데 이렇게 함으로써 영상신호 전송시 특징만을 보내주고 수신단에서는 텍스처 합성을 통해 송신단에서의 원래 텍스처 영상을 재현하여, 높은 영상정보 감축을 행할 수 있다.

3.4 움직임 해석

앞서 설명한 방법들은 움직임이 고려되지 않은 영상해석 방법인데, 이러한 정적인 화상의 시간적인 연속으로 볼 수 있는 움직이는 영상에 대한 연구가 영역 분할, occlusion 해석 및 3차원 정보의 추출등에서 최근 활발하게 진행되고 있다. ^{5), 25), 26)} 움직임 추정을 통해, 검출된 이동물체의 변위 및 속도등에 대한 정보를 얻을 수 있으며 이는 영역 분할 등에 응용될 수 있다. 이동물체의 변위추정은 프레임간의 이동 물체의 상관성을 이용하여 이루어 지는데 부영상 단위로 변위를 추정하는 block matching 방법과 화소 단위로 반복적으로 변위를 추정하는 pel recursive 방법이 있다. 후자는 화소 단위로 이동값을 추출하여 정확한 값을 얻을 수 반면에 계산량이 많고 복잡하다. 전자는 부영상 단위로 이동 변위를 추출하는 만큼 정확도는 떨어지지만 계산량이 많지 않아 실 시간 처리에 유리하다.

위에서 언급한 움직임 정보를 이용하여 물체의 3차원 정보를 알아 내거나 산업, 군사용의 추적기에 사용하거나 데이터 감축, 혹은 동적 영역 분할등에 이용할 수 있다. 동적 영역 분할 방법에는 각 영상에서 에지, 모서리, 무게중심, 둘레등과 같은 특

징을 고려하여 두 영상사이의 정합과정을 거쳐 같은 속도를 갖는 특징들끼리 묶어 영역분할을 하는 특징에 근거한 영역분할방법 및 그림 6에서와 같은 차영상 (difference imange), 상관관계 및 시간공간 경사법등을 이용하여 화소 단위로 움직임을 추정하여 같은 움직임을 갖는 화소들끼리 묶어주는 화소에 근거한 영역분할 방법으로 크게 나눌 수 있다. 이밖에도 구름 사진을 토대로 한 구름의 이동 추적, teleconferencing, 산업용 검사 및 의학등에 많이 응용되고 있다.

3.5 2차원 및 3차원 구조의 표현

앞에서 설명한 특징들에 의해 물체의 둘레, 영역을 구하고 혹은 3차원 물체를 배경 또는 서로 다른 물체와 분류를 행한 다음에는 2차원 및 3차원 물체에 대한 표현이 필요하게 된다. 이것들을 표현하는 일반적인 방법은 없지만, 2차원 물체의 둘레 혹은 영역 및 3차원 물체의 모양 정보를 정량화시키는 방법으로서, 여러가지 양이 정의되어 있다. 예를 들면 2차원 물체의 경우, 둘레를 이루는 변의 길이 분포, 각의 크기, 각의 변화량, 면적, 둘레, 꼭지점의 좌표, 지름 및 둘레의 규칙성, 주방향, 직선성, 곡선성 등 여러가지가 시도되어 왔지만 이것들은 표본화 크기, 물체의 위치, 혹은 방향등에 영향을 받으며, 이렇게 정량적으로 표현된 특징들로 물체를 묘사, 정합, 인식 및 분류를 행하게 된다. 이밖에도 2차원 물체의 둘레를 근사적인 직선으로 표현하기 위한, 그림 7에 예로 보인 (128×128로 이루어진 화상에서 115개의 화소로 이루어진 원을 12개의 점들을 직선으로 연결하여 구성하는), 다각형 근사화

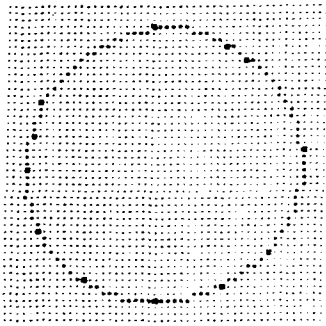


그림 7. 다각형 근사화

(polygonal approximation), 둘레를 chain 코드로 표시하는 방법, Fourier descriptor로 표현하는 방법과 2 차원 영역을 간단한 부분으로 분해하여 표시하는 방법, quad-tree를 이용하는 방법, 직선 혹은 곡선 성분을 추출하여 표현하는 Hough 변환 방법, 골격화상 (skeleton) 및 둘레까지의 거리로 표현되는 medial axis 변환, 그리고 크기, 위치, 회전에 영향을 받지않는 (invariant) 모멘트(moment)를 정의하여 정량화시키는 방법등이 있다.^{4), 5), 8)}

3 차원 물체의 표현방식으로는 3 차원의 기저함수를 사용하여 근사시키거나, 3 차원 물체를 이루는 2 차원 표면으로 나누는 표현하는 방법, 간단한 3 차원 기본구조로 분해하여 표현하는 방법, 단면이 축을 따라 이동하여 3 차원 물체를 이루었다고 표현하는 일반화된 실린더 모델 (generalized cylinder model), 건축, 공학등에서 사용하는 line drawing 방법등이 있다.

Low-level 처리과정을 거쳐 분할된 물체를 얻은 후, 이렇게 표현된 2 차원 혹은 3 차원 물체는, 즉 밝기 변화량의 기하학적 분포 및 구조를 다루는 기본 스케치 (primal sketch) 과정을 거쳐, 그 다음의 high-level 단계로 넘어가 지식표현, 정합, 인식등에 쓰이게 된다. 혹은 관측자를 중심으로 한 좌표계에서 보이는 면의 방향, 거리, 혹은 이러한 양들의 contour의 불연속적인 양들을 다루는 2½ sketch (본질적영상) 과정을 거쳐 3 차원 형상을 표시하게 된다.⁸⁾

4. High-level 처리

High-level 처리과정에서는 입력으로 들어온 화상에서 low-level 처리를 거쳐 추출하여 얻은 특징들과 물체들 사이의 연관관계를 컴퓨터가 이미 가지고 있는 3 차원 세계에 대한 지식베이스 (knowledge base)와 연관시켜 정합, 추론 혹은 update 과정을 거쳐 목적을 달성하여 symbolic 출력을 내게된다. High-level 처리과정에서는 참, 거짓으로 주어지는 명제논리 (propositional logic) 및 참, 거짓뿐 아니라 표현의 문법적 구조와 의미를 포함한 술어논리 (predicate logic) 혹은 의미 네트워크 (semantic network)를 사용하여 지식표현, 정합, 인식, 믿음 (belief), 추론 및 계획수립을 행하게 된다.^{5), 23)}

표현된 지식은 정보를 쉽게 꺼낼 수 있고 확장할 수 있으며 정합, 인식, 추론 및 계획수립등 컴퓨터가 처리하기에 적합하여야 한다. 컴퓨터 비전을 위한 지식표현은 명제논리 및 술어논리를 사용하거나, 규칙에 근거한 생성 시스템 (production system) 을 사용하거나, 물체를 마다 (node)로 표현하고 각 링크 (link)는 물체들 사이의 관계를 나타내는 것으로 표현하는, 그래프 형태를 가지며 인간 기억의 심리학적 모형으로 개발된 의미 네트워크를 사용하여 표현할 수 있다. 링크는 물체 사이의 관계의 종류에 따라 방향성이 정의될 수도 있다. 이렇게 표현된 지식을 이용하여 그 다음의 정합, 인식등의 과정을 수행하게 된다.²⁴⁾

정합이란 외부에서 들어온 3 차원의 시각입력과 이미 존재하는 물체에 대한 다른 표현을 비교하는 과정이며 이러한 과정을 거쳐 인식, 계획수립을 거쳐 반응으로 나타나게 된다. 이러한 정합을 위해서는 완전한 정합은 존재하지 않는 경우가 있으므로 정합의 정도를 나타내는 척도로 영역사이의 성질 혹은 관계의 정합도 및 정합된 영역수를 고려하여 정량적인 양을 정의하여 사용한다.²⁵⁾

추론은 주어진 다른 사실로부터 사실을 이끌어 내는 인간의 합리적 사고의 한 과정을 말하는데 이를 위해 술어 논리를 사용하여 표현된 지식을 modus ponens, modus tollens, 치환등과 같은 추론규칙, 정리증명 규칙을 사용하여 새로운 지식을 이끌어 낸다. 또는 생성 시스템과 같이 규칙 집합 및 이규칙

을 사용할 수 있는 조건들로 이루어진 일종의 규칙에 근거한 시스템들을 사용할 수 있다. Line drawing으로 표현된 3차원 물체의 line labeling도 넓은 의미의 추론이라 할 수 있다.⁹⁾

비전 시스템에서 목적수행을 위해 전체과정을 미리 결정하는 계획수립을 얼마나 유효적절하게 최소의 비용을 들이고 하느냐 하는 것이 문제가 된다. 각 행동(action)이 취해지기 위한 전제조건 및 후의 상태, 각 행동에 필요한 비용 및 제한조건에 따라 탐색(search)하여 최소의 비용으로 목적까지 도달하는 길을 설정하는 작업이 계획수립이다. 이 계획수립에 따라 해당되는 제어를 행하여 최종 symbolic 출력 결과를 얻게 된다.

5. 결 론

본 고에서는 컴퓨터 비전의 각 과정을, 특히 디지털 신호처리 기술을 사용하는 low-level 처리과정을 주로 소개하였다. 영상처리의 거의 대부분의 알고리즘이 그렇듯이 사람의 시각이 최후의 판단기준이 되므로 사람의 정확한 시각 모델 및 사람 눈의 특성을 고려한 효과적인 알고리즘의 개발이 필요하다. 크기, 방향, 필기체와 같은 서로 다른 표현 방식등을 고려한 처리과정의 개발 및 다양한 종류의 입력 영상에 적합하고 조명의 영향을 적게 받는 알고리즘의 개발이 요구된다. 또한 컴퓨터 비전을 위한 low-level 처리과정과 high-level 처리과정의 일의 분담 및 연결등이 효율적으로 이루어져야 할 것이며 움직이는 물체의 해석등과 같은 방대한 데이터양을 가지며 빠른 처리가 요구되는 응용예에서는 실시간 처리를 위한 고속 알고리즘 및 하드웨어의 개발이 필수 불가결하다. 최근의 추세로 보아 컴퓨터 비전의 연구 및 응용범위가 나날이 증대되어 가고 있으나 결국 인간의 눈의 한계 및 기계가 가질 수 없는 인간고유의 능력등이 컴퓨터 비전의 한계를 결정짓게 될 것이다.

참 고 문 헌

- 1) R.C.Gonzalez and P. Wintz, Digital Image Processing. Addison-Wesley, 1977.
- 2) W.K.Pratt, Digital Image Processing. John Wiley & Sons, 1978.
- 3) E.L.Hall, Computer Image Processing and Recognition. Academic Press, 1979.
- 4) A.Rosenfeld and A.C.Kak, Digital Picture Processing, second ed., vol. 1 & 2, Academic Press, 1982.
- 5) D.H.Ballard and C.M.Brown, Computer Vision, Prentice-Hall, 1982.
- 6) R.Nevatia, Machine Perception. Prentice-Hall, 1982.
- 7) D.Marr, Vision. W.H.Freeman and Co., 1982.
- 8) M.D.Levine, Vision in Man and Machine. McGraw-Hill, 1985.
- 9) A.Rosenfeld, "Image analysis : Problems, progress and prospects," Pattern Recognition, vol. 17, no. 1, pp. 3-12, 1984.
- 10) J.T.Tou and R.C.Gonzalez, Pattern Recognition Principles. Addison-Wesley, 1974.
- 11) R.C.Gonzalez and M.G.Thomason, Syntactic Pattern recognition An Introduction. Addison-Wesley, 1978.
- 12) K.S.Fu, Syntactic Pattern Recognition and Applications. Prentice-Hall, 1982.
- 13) R.O.Duda and P.Hart, Pattern Classification and Scene Analysis. Wiley Interscience, 1973.
- 14) L.S.Davis, "Edge detection by compass gradient masks," Comput. Graphics Image Processing, vol. 4, pp. 248-270, 1975.
- 15) G.S.Robinson, "Edge detection by compass gradient masks," Comput. Graphics Image Processing, vol. 6, pp. 492-501, 1977.
- 16) D.Marr and E. Hildreth, "Theory of edge detection," Proc. Roy. Soc. London, vol. 207, pp. 187-217, Mar. 1980.
- 17) R.M.Haralick, "Digital step edges from zero crossing of second directional derivatives," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-6, no. 1, pp. 58-68, Jan. 1984.
- 18) M.Kunt, A.Ikonomopoulos, and M.Kocher, "Second-generation image coding techniques," Proc. IEEE, vol. 73, pp. 549-574, Apr. 1985.
- 19) K.S.Fu and J.K.Mui, "A survey on image segmentation," Pattern Recognition, vol. 13, pp. 3-16, 1981.
- 20) J.Kittler and J. Illingworth, "Threshold selection based on a simple image statistic," Comput. Vision, Graphics, Image Processing, vol. 30, pp. 125-147, 1985.
- 21) J.S.Weszka, "A survey of threshold selection techniques,"

- Comput. Graphics Image Processing, vol. 7, pp. 259-265, 1978.
- 22) R.M.Haralick, "Statistical and structural approaches to texture," Proc. IEEE, vol. 67, no. 5, pp. 786-804, 1979.
 - 23) R.W.Conners and C.A.Harlow, "A theoretical comparison of texture algorithms," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-2, no. 3, pp. 204-222, May 1980.
 - 24) R.Chellappa and R.L.Kashyap, "Texture synthesis using 2-D noncausal autoregressive models," IEEE Trans. Acoust., Speech, Signal Proc., vol. ASSP-33, pp. 194-203, Feb. 1985.
 - 25) T.S.Huang(ed.), Image Sequence Processing and Dynamic Scene Analysis, Springer-Verlag, 1983.
 - 26) R.Jain, "Extraction of motion information from peripheral processes," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-3, no. 5, pp. 489-503, Sept. 1981.
 - 27) R.Nevatia and K.Price, "Locating structures in aerial images," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-4, no. 5, pp. 476-484, Sept. 1982.
 - 28) L.S.Davis, "shape matching using relaxation techniques," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-1, no. 1, pp. 60-72, Jan. 1979.