

영상처리와 기계학습을 통한 물체 인식

천정훈*

*고려대학교 컴퓨터정보통신대학원 컴퓨터정보통신공학과
e-mail : rousseau1@paran.com

Object detection with Video processing and machine learning

Jung-Hoon Chun*

*Dept. of Computer and Communication Engineering,
Graduate School of Computer & Information Technology, Korea University

요 약

본 논문은 전자 제어 시스템의 응용분야로 많은 연구가 이루어 지고 있는 영상 데이터를 통한 물체 인식을 위한 방법으로서 영상처리와 기계학습을 활용하는 알고리즘을 제안하고 실제 구현과 실험을 통해 물체 인식 특성을 파악하는 것에 대한 것이다. 물체의 정확한 인식을 위해서는 영상 데이터로부터 물체의 특성을 파악할 수 있는 알고리즘과 함께 파악된 특성을 학습을 통해 데이터를 확보하는 두 가지 과정이 요구된다. 본 논문에서는 이러한 방법을 구체화하기 위해 사람의 얼굴을 인식하는 알고리즘을 예로 들어 영상데이터로부터 물체를 인식하는 알고리즘을 구현하고 실험을 통해 알고리즘의 동작 특성을 확인한다.

1. 서론

최근 다양한 전자 제어 시스템들은 영상 처리 기술의 발달에 따라 다양한 센서를 영상데이터로 대체하는 시스템들의 개발에 많은 연구가 이루어 지고 있다. 자동차 제어 장치의 첨단화된 예를 보면, 차량에 장착된 카메라를 통해 영상 데이터의 처리를 통한 자동 주차 시스템은 이에 대한 좋은 예이며, 기존의 거리 측정 센서를 대체하여 보다 효과적인 어플리케이션의 개발이 이루어 지고 있다. 자동 주행을 위한 Lane Detection 시스템, 운전자의 눈을 감지하여 졸음을 방지하는 시스템 등은 영상 처리를 통한 새로운 시스템의 개발이 가능케 된 또 다른 예이다. 이러한 영상을 통한 물체 인식의 예들은 자동차 분야의 제어 시스템 뿐만 아니라 영상 데이터로부터 움직임을 감지하여 침입자를 경고하는 보안 시스템이나 공장 자동화에 활용되는 산업용 로봇의 물체 감지 시스템 등 다양한 분야에 활용되고 있다.

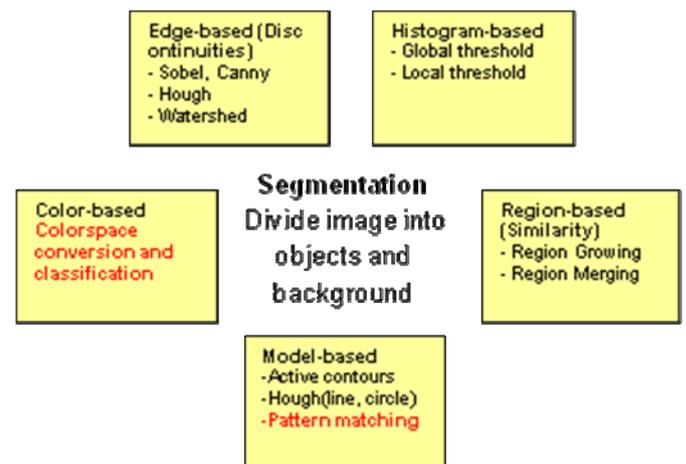
여기서 특히 고려해야 할 사항은 전자 제어 시스템의 실시간성이다. 자동차의 자동 주행 장치 등에 있어서 잘못된 물체의 인식은 치명적 사고의 원인이 될 수 있을 뿐 아니라 마이크로 프로세서 내에서 실시간의 제어가 이루어 져야 한다. 이에 따라 이러한 물체 인식 알고리즘은 간결하면서도 강건해야 한다.

본 논문에서는 물체 인식의 한 예로 영상처리 분야에 많은 연구가 이루어지고 있는 얼굴 인식 알고리즘을 구현하고 이의 적합성을 실험을 통해 확인해 본다.

2. 관련 연구

2-1. 물체 인식을 위한 영상 처리

물체 인식을 위한 영상 처리 알고리즘으로서 다양한 이미지 분할 방법(Image Segmentation Method)들이 제안되어 있다. 아래의 그림 1은 이러한 이미지 분할 방법들에 대하여 보여주고 있다. 이미지 분할 방법은 카메라로부터 전달되는 이미지 데이터를 통해 이미지 내부에 필요한 영역을 추출해 내기 위해 사용된다. 본 논문에서 얼굴 인식을 위해 활용하고자 하는 이미지 추출 방법은 그림 1에 있는 색기반 분할(Color-based Segmentation)과 모델 기반 분할(Model-based Segmentation) 두 가지 방법을 활용하고자 한다.



(그림 1) 다양한 이미지 분할 방법

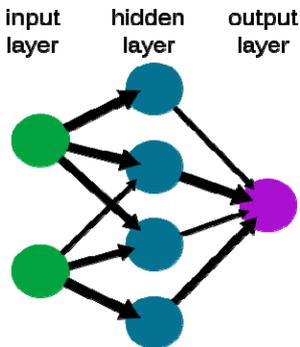
2-2. 패턴 매칭

모델 기반 이미지 분할 방법에서 가장 많이 활용되는 방법으로 물체에 대한 이미지의 패턴을 파악하고 새로운 이미지로부터 동일한 패턴의 이미지를 추출해 내는 패턴 매칭 방법이 있다. 패턴을 매칭하기 위한 기계학습 방법 중 이미지 처리를 위해서는 뉴럴 네트워크(Neural Network) 방법이 가장 많이 활용된다. 뉴럴 네트워크는 1960 년대에 인간처럼 판단하는 컴퓨터를 만들기 위해서 인간의 두뇌 구조를 모방하여 만들어 진 이론이다. 인간의 두뇌 구조를 보면 뉴런(Neuron)이라는 작은 신경조직이 거의 무한대로 연결된 그물망 구조를 가지고 있으며 뉴런은 전기적인 신호를 받아서 연결된 다른 뉴런으로 새로운 신호를 재 출력해준다. 즉 아래의 그림 2 와 같이 일반적인 뉴럴 네트워크는 어떤 값을 집어 넣을지 또는 어떤 값을 얻을지 결정하는 입력 값과 출력 값을 지정하는 입출력 층(input/output layer)과 신호를 받아서 전달하는 중간의 숨겨진 층(Hidden Layer) 등 세 개의 층(layer)로 구성된다. 중간의 숨겨진 층은 특히 다수의 숨겨진 층(Layer)들로 구성될 수 있으며 그 층들이 많아지면 계산은 복잡해 지나 정확도는 증가한다. 실시간 제어 시스템을 위한 물체 인식에서는 계산의 실시간성과 인식의 정확도를 위해 숨겨진 층의 개수를 정하는 것은 매우 중요해 질 수 있다.

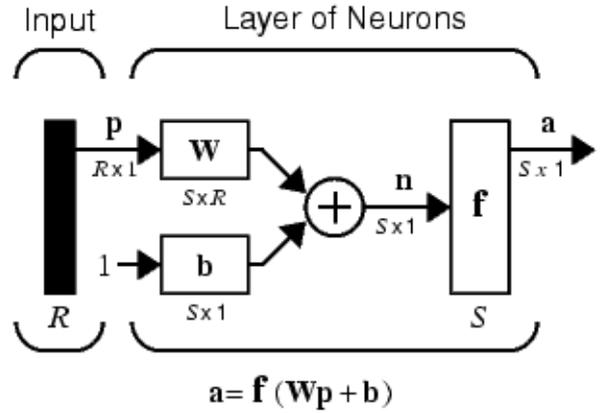
뉴럴 네트워크의 원리는 그림 3 에서와 같은 계산을 통해 입력에 대한 뉴럴 네트워크의 결과를 목표와 비교하여 오차를 최소화하기 위한 보정 값(Weight: 각 neuron 들의 연결 강도)

$$W = \begin{bmatrix} w_{1,1} & w_{1,2} & \dots & w_{1,R} \\ w_{2,1} & w_{2,2} & \dots & w_{2,R} \\ \dots & \dots & \dots & \dots \\ w_{S,1} & w_{S,2} & \dots & w_{S,R} \end{bmatrix}$$

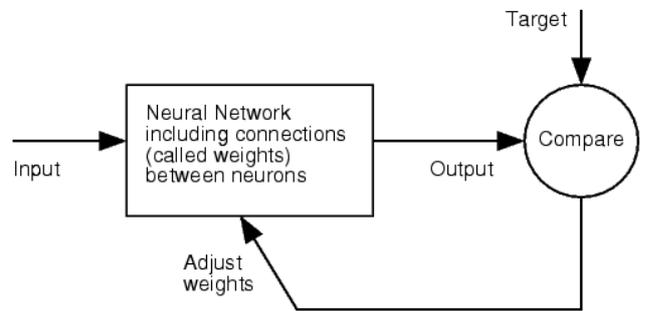
을 반복 수정하여 얻어 내는 것을 목표로 한다. 여기서 반복의 횟수는 학습을 의미한다. 그림 4 는 보정 값을 최적화하는 알고리즘의 순서도를 보여주고 있다.



(그림 2) 간결한 뉴럴 네트워크 구조



(그림 3) 뉴럴 네트워크의 계산



(그림 4) 뉴럴 네트워크의 동작 원리

3. 본론

3-1. 학습을 통한 얼굴 인식 알고리즘의 구현

본 연구에서는 얼굴 인식 알고리즘의 구현은 아래 그림 5 와 같은 순서로 진행하였다.

먼저 첫 번째 단계(Learning with existing face images)에서는 기 확보된 얼굴 이미지 데이터와 얼굴이 아닌 이미지 데이터로부터 뉴럴 네트워크를 통해 학습한다. 이때, 가보 변환(Gabor Transform)을 통하여 물체의 특성을 파악할 수 있는 형태로 이미지를 변환하고 변환된 이미지 데이터를 Input 으로 하여 학습한다.

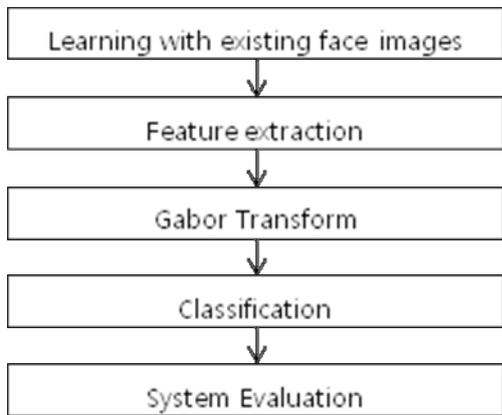
두 번째 단계(Feature extraction)에서는 새로운 이미지에서 사람의 얼굴과 유사한 패턴의 이미지를 추출한다. 이 때 추출을 위한 알고리즘으로 이미지의 색기반 분할(Color-based Segmentation)을 통해 얼굴 피부색과 유사한 이미지를 추출하고, 가로/세로의 비율에 범위를 주어 피부색과 유사한 분할된 이미지에서 얼굴일 가능성이 있는 이미지만을 재 추출한다. 이는 차후 수행하게 될 학습된 데이터와의 비교 시 대상 이미지의 수를 줄여 실시간성을 부여하기 위함이다.

세 번째 단계(Gabor Transform)에서는 추출된 이미지들을 학습된 네트워크의 인풋 데이터로 변환하는 과정이다. 얼굴 이미지의 특성을 표현하는 수단으로 가보 변환을 수행하였다.

네 번째 단계(Classification)에서는 학습 결과에 추

출 및 변환된 이미지 데이터를 입력하여 사람의 얼굴 인지 판별하는 과정을 거친다.

마지막 다섯 번째 단계(System Evaluation)에서는 최종적으로 판별된 결과를 평가한다.



(그림 5) 알고리즘의 동작 순서

3-2 뉴럴 네트워크를 이용한 얼굴 이미지 학습

얼굴 인식을 위해 학습할 이미지로는 그림 6 과 같이 다양한 얼굴 이미지뿐만 아니라 얼굴이 아닌 이미지를 함께 학습하도록 하였다. 이러한 이미지 데이터는 그림 6 과 같이 뉴럴 네트워크의 인풋 데이터로 활용되기 위해 가보 변환된다.

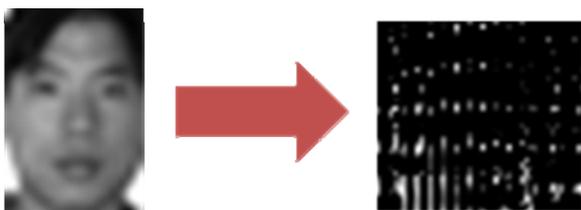


a. 학습될 얼굴 이미지

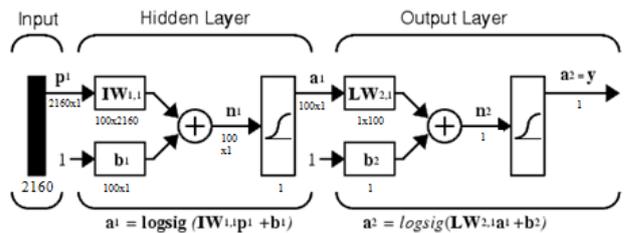


b. 학습될 얼굴이 아닌 이미지

(그림 5) 학습할 이미지 데이터



(그림 6) 이미지 데이터의 가보(Garbor) 변환



(그림 7) 구성 네트워크

학습을 위해 구성된 네트워크는 그림 7 과 같다. 활용된 인풋 데이터는 45x48 사이즈의 이미지 데이터가 2160x1 의 사이즈로 벡터화 되었으며, 아웃풋으로는 얼굴일 경우에 1 의 값을 얼굴이 아닐 경우 0 을 출력하게끔 구성되었다. 아웃풋의 범위를 0 과 1 사이로 변환하기 위하여 Log-sigmoid 전달 함수(transfer function)를 취하였다. 이는 boolean 으로 아웃풋을 출력해야 하는 경우, 즉 얼굴인지 아닌지를 판단하는 등의 학습 알고리즘에 효과적인 전달함수이다.

3-3 학습된 네트워크를 이용한 얼굴 판별 실험

그림 8 은 얼굴 판별을 위해 실험에 활용된 원본 이미지이다. 그림 8 의 이미지로부터 얼굴 대상으로 추출된 이미지는 그림 9 와 같다.

각 대상 이미지로부터 네트워크의 출력 값은

- 0.0345
- 0.0942
- 0.5335
- 0.1069
- 0.0973
- 0.0565
- 0.9289
- 0.8824

으로 판별 기준을 0.8 로 하여 0.8 보다 큰 값을 취할 경우,

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

의 결과가 나타난다. 이는 얼굴인지 아닌지를 판별하여 인풋 데이터들을 구별(classification)하는 부분이다. 즉, 원본 이미지에 각 얼굴로 인식된 이미지의 중심에 표시를 하게 되면 그림 10 과 같은 결과를 얻게 된다.



(그림 8) 시험에 사용된 원본 이미지



(그림 9) 얼굴 대상 추출 이미지



(그림 10) 얼굴 인식 결과

참고문헌

- [1] S Jayaraman 외, “Digital Image Processing”, McGraw-Hill, 2009
- [2] Thomas J. Anastasio, “Tutorial on Neural Systems Modeling”, Sinauer Associates, Inc., 2010
- [3] Erik Cuevas Jimenez 외, “Intelligent active vision systems for robots”, Cuvillier Verlag, Göttingen, 2007
- [4] Rafael C. Gonzalez, “Digital Image Processing Using MATLAB”, McGraw-Hill, 2011

4. 결론 및 향후 방향

본 논문에서는 최근 다양한 분야에서 활용되는 영상 데이터를 통한 물체 인식의 예로 얼굴인식 알고리즘을 설계, 구현하고 실험을 통해 확인하는 작업을 수행하여 보았다. 실제 구현된 알고리즘의 실시간 특성을 확보하기 위하여 학습을 위한 뉴럴 네트워크의 구성은 최대한 간결하게 하였으며, 이미지의 선 처리를 위해 이미지 추출 시 색기반 분할 및, 가로 세로 비율 들을 추가적인 추출 기법 등을 적용하여 보았다.

실험을 통해 알고리즘의 동작을 확인해 본 결과 정면의 얼굴 이미지에 대해서는 잘 동작하고 있는 것을 확인되었으나 실시간성을 확보하기 위하여 추가된 얼굴 비율을 이용한 얼굴 대상 이미지 추출에 의해 정면이 아닌 얼굴 이미지를 추출하는 부분에서 당연히 얼굴 이미지를 놓치는 상황을 가져왔다.

차후 이러한 현상의 개선과 알고리즘의 Robustness 확보를 위해 Cross Correlation 을 통한 pattern matching 방식의 활용 여부를 재 검토해 볼 필요가 있을 것이다.

또한, 실제 실시간 제어 시스템에의 적용을 통해 구현 알고리즘의 특정 마이크로 프로세서 내에서의 실시간 특성을 확인하는 실험을 수행할 예정이다.