

BMA와 동적 GTM을 이용한 움직이는 객체의 얼굴 영역 검출에 관한 연구

장혜경, 김영호, 김대일, 홍종선, 강대성
동아대학교 전자공학과

A study on the face detection of moving object using BMA and dynamic GTM

Hye-Kyoung Jang, Young Ho Kim, Deail Kim, Jong-Sun Hon, Dae-Seong Kang
Department of Electronic Eng., Dong-A Univ.
E-mail: naoriori@hotmail.com

요약

본 논문에서는 video stream내의 움직이는 객체 정보를 추정하고 동적 GTM(genetic tree-map) 알고리즘을 사용하여 얼굴 영역 검출 기법을 제안한다. 기존의 일반적인 객체 추정 기법은 클러스터(cluster)과정을 통하여 영상 정보를 분할하고 그 중 움직이는 객체 부분을 복원함으로써 추정하였다. 제안하는 기법은 BMA(block matching algorithm)[1] 알고리즘을 사용하여 video stream에서 움직이는 객체 정보를 얻고 클러스터 알고리즘으로 PCA(principal component analysis)를 사용한다. PCA 기법은 입력 데이터에 관해 통계적 특성을 이용하여 주성분을 찾는다. 주축과 영역분할 알고리즘을 사용하여 데이터를 분할하고, 분할된 객체 정보를 사용하여 특정 객체만을 추정하는 것이 가능하다. 이렇게 추정된 객체를 얼굴영역의 feature에 대하여 신경망 학습인 동적 GTM 알고리즘을 사용하여 생성된 동적 GTM 맵의 정보에 따라 객체의 얼굴영역만을 추출해 낼 수 있다[2-6].

I. 서론

컴퓨터의 성능이 향상됨에 따라 계산 연구 분야뿐만 아니라, 여러 가지 다양한 분야에서 컴퓨터를 사용하고 있으며, 사용자의 폭도 상당히 넓어지게 되었다. 따라서 보다 친숙한 사용자 인터페이스를 위하여 화면에 직접 사람을 등장시켜 대화를 통하여 설명하고 작업을 수행하는 face-to-face 방법에 관한 연구도 활발히 진행되고 있다. 이와 같은 사용자 인터페이스를 위하여 얼굴에 관한 연구는 중요한 과제가 된다. 또한, 사용자 인터페이스뿐만 아니라 얼굴에 관한 연구는 여러 가지 다른 분야에서 요구되고 있다. 예를 들면, 얼굴 인식은 신분 증명, 보안 시스템, 인사 자료에 사용될 수 있으며, 치안 분야에서는

범인을 색출하기 위하여 얼굴 각 부분의 그림이나 사진을 보여 주면서 유사한 형태를 조합하여 몽타주를 작성하는 시스템에 사용될 수 있다. 얼굴 인식을 위한 시스템에서 얼굴의 모든 정보가 사용되어지는 것보다는 특징이 추출되어 그 특징으로 인식을 하는 것이 더 효과적이다.

얼굴 인식 성능은 객체의 분할 기술과 얼굴영역 검출 기술이 좌우한다. 객체 분할 기술에 있어서 움직임 추정은 tracking을 할 대상 object에 따라서, image sequence에 따라서, camera의 움직임 및 배경의 변화 유무에 따라서 적용되는 알고리즘이 상이하다. motion estimation은 크게 image region base에 의한 방법, feature point base에 의한 방법, line segment base에 의한 방법 등으로 나뉜다. 가장 많이 사용하는 방법으로 BMA, Optical Flow Methods(OFD)[7][8], Active Contour Model(SNAKE) 등의 방법이 많이 연구되어 왔다. 이들 각각마다 장단점을 가지며, 여러 곳에서 개선된 기법이 제안되고 있다. 이 중 BMA의 경우 직관적이며 이해가 쉬워 많이 사용되고 있다. 본 논문에서는 객체를 효율적이고 정확히 검출해 내기 위하여 BMA 기법을 사용하여 연구를 하였다. BMA로 획득한 영상 정보에 PCA의 통계적 특성을 적용하여 객체 정보를 추출하여 움직이는 객체와 움직이지 않는 배경 영역을 검출해낸다. 그중 움직이는 영역을 복원시켜서 객체만을 검출해 내었다. 얼굴영역 검출 알고리즘으로 사용한 동적 GTM 알고리즘의 가장 큰 특징은 기존의 신경망(neural network)에서 고려되지 못한 데이터의 특징(feature)에 대한 중요도를 유전 알고리즘(genetic algorithm)으로 구성하며 특징의 우선순위에 따라 트리 구조를 도입한 GTM을 적용한 것이다. 또한 단 1회의 학습만을 요구하게 되므로 학습시간을 줄일 수 있는 장점이 있다.

제안하는 얼굴 영역 추출 기법은 BMA를 이용하여 도출한 영상의 동적 정보로써 영상내의 객체 추출을 실시하며, 추출된 영상의 얼굴 영역을 신경망 학습인 동적 GTM을 통하여 검출해 낸다.

II. 본론

1. BMA(block matching algorithm)

영상 신호에서 운동이 발생 하였을 경우 개개의 화소가 각각 다른 방향으로 운동하기 보다는 유사한 방향으로 움직인다. 따라서 입력 영상을 작은 block으로 나눈 다음 이전 frame에 적당한 크기의 search area를 두고 search area내에서 가장 유사한 block을 찾아서 이동벡터를 구할 수 있다. 이와 같이 BMA는 현재 영상의 일정 영역과 가장 유사한 영역을 이전 영상에서 찾는다. matching 정도를 나타내는 척도로 제안된 방법은 NCCF(Normalized Cross-Correlation Function), MSE(Mean Square Error), MAD(Mean of the Absolute Frame Difference), NMAD(Number of bits necessary to binary code the Absolute Frame Difference)가 있다. 본 연구에서는 block 비교 시 기준으로 아래의 수식(MAD)을 사용한다.

$$MAD(i,j) = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} |f_k(x+m,y+n) - f_{k-1}(i+m,j+n)|$$

BMA는 matching 계산 시간이 많이 걸리기 때문에 이것을 극복하기위한 많은 알고리즘이 개발되고 있다[9]. 본 논문에서는 search 간격을 줄여 가면서 반복하여 검색함으로써 이동 벡터를 찾는 방법을 사용하였다.

2. PCA(principal component analysis)

주성분 해석 기법은 입력 공간내의 변동량이 많은 방향으로 주축을 발견하여 데이터의 중복성을 제거하는 기법으로 주성분 분석 기법이라고도 한다[10]. n차원의 정상적 확률 신호(stationary random signal)에서 입력 신호 $x(t) = [x_1, x_2, \dots, x_n]^T$ 의 자기상관 행렬 (autocorrelation matrix)은 $R_x = E[xx^T]$ 의 정규화된 고유치(normalized eigenvalues)인 $\{e_i\}$ ($\|e_i\| = 1, i = 1, 2, \dots, m$)에 의하여 확률신호 $x(t)$ 는 1차원 sub-space로 투영(projection) 될 수 있다. 따라서 $x(t)$ 의 주 성분(principal component)은 $a(t) = e_1^T x(t)$ 와 같이 R_x 의 정규화된 고유치 중 가장 큰 고유치 (e_1)의 고유성분(eigen-component)으로 표현할 수 있으며, m 개의 first eigen-component를 사용하면 n 차원의 확률신호를 m 차원으로 축소할 수 있다. 따라서 주성분 해석을 이용한 고차원 데이터의 차원 감소는 영상 분석에 있어 유용하다. 즉, $n \times n$ 의 2차원 영상을 n 개의 1차원 데이터 요소로 축소함으로써 상당한 양의 처리 시간과 메모리 공간을 줄일 수 있다. 또한 주성분 분석에 의하여 계산된 주축은

입력 벡터의 최대 고유치의 고유 벡터의 방향과 일치하며 입력 벡터의 변동량이 가장 큰 방향이다. 다음으로 주축과 수직을 이루면서 노드의 평균을 지나는 지점을 기준으로 영역을 2개로 분할한다. 각 영역의 평균 오차를 계산하여 평균 오차가 가장 큰 영역을 선별한 후 선택된 영역에 영역 분할을 재 적용한 다음, 객체 정보만을 가진 그룹을 영상 복원한다.

3. 동적 GTM(dynamic genetic tree-map)

대부분의 신경망은 추출된 특징을 우선순위나 중요도에 관계없이 뉴런의 구성요소로 사용하거나 실험을 통하여 얻어진 가중치(weight vector)를 사용하며, 얻어진 특징은 모두가 동일하게 훈련과 인식에 사용된다. 기존의 신경망(neural networks)에서 고려되지 못한 데이터의 특징(feature)에 대한 중요도를 유전 알고리즘(genetic algorithm)으로 구성하고, 특징의 우선순위에 따라 트리 구조를 도입한 GTM(genetic tree-map)을 적용한다. 데이터의 유사성에 따라서 신경망의 뉴런이 동적으로 분리되고 병합될 수 있도록 동적인 기능을 갖는 동적 GTM으로 확장한다[11].

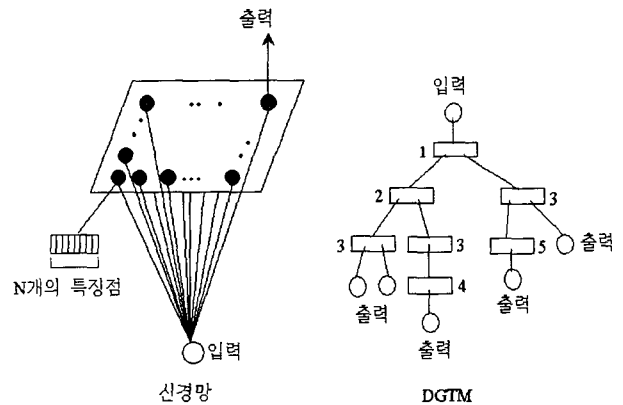


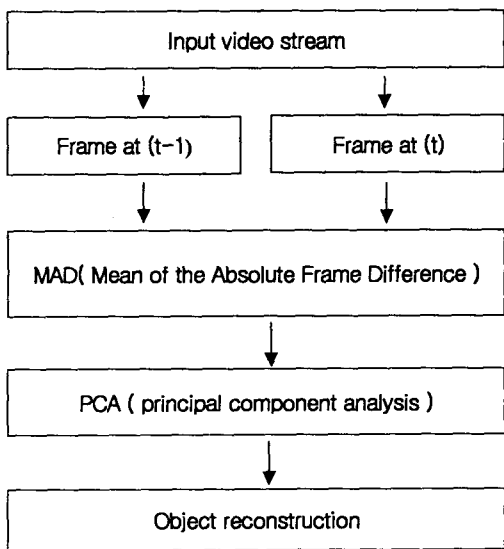
그림 1. 신경망과 동적 GTM 구조

그림 1에서는 기존 신경망과 동적 GTM에서 사용되는 특징에 대한 비교이다. 좌측 그림의 기존 신경망에서 각 노드는 N개 특징 모두를 갖고 있으며, 대체적으로 입력된 특징과의 최소 거리 방식이 자주 사용된다. 우측 그림은 동적 GTM에 사용되는 구조로서 각 레벨에서는 동일한 순서의 특징이 사용되지 않으며, 맵의 좌우에 표시된 순번에 따르는 특징이 구분에 사용되므로, 훈련에서 분석되었던 입력 데이터의 특징에 따라 분류에 적절한 특징으로 효율적인 인식이 가능하다. 첫 번째 우선순위의 특징을 통해 분류된 데이터들은 두 번째 우선순위의 특징에 따라 분류된다. 우선순위는 트리 구조의 레벨을 나타내며, 맵에 존재하지 않는 값이나 잘못된 출력이 발생하는 부분은 동적으로 맵을 확장할 수 있다. 또한 새로운

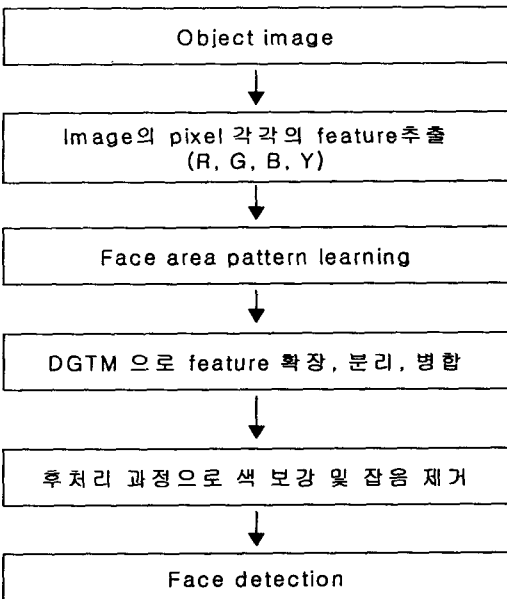
특징이 필요한 경우 마지막 노드에 추가할 수 있다.

III. 전체 블록 다이어그램

BMA 알고리즘을 사용하여 영상 내 객체의 동적 정보에 대하여 PCA를 이용하여 영역 분할하여 영상을 복원시키면 각각의 객체를 분류해 낼 수 있다. 추출된 객체의 얼굴 영역을 신경망 학습인 동적 GTM 알고리즘을 사용하여, 얼굴영역의 픽셀정보인 R, G, B, Y 성분을 학습시킨 동적 GTM 맵을 구성한 후 검출하였다. 그림 2는 전체 구성에 대한 블록 다이어그램이다.



(a) 객체 분할 구성도



(b) 얼굴 영역 검출 구성도

그림 2. 전체 블록 다이어그램

IV. 실험결과 및 고찰

본 논문에서는 실험에서 사용한 영상크기는 320X240이며 BMA 정보와 동적 GTM 알고리즘으로 시스템을 구하고 C 프로그래밍으로 구현하였다.

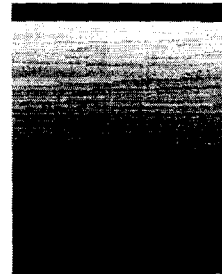


그림 3. 생성된 동적 GTM 맵

그림 3은 동적 GTM 알고리즘을 사용하여 생성시킨 학습 후의 동적 GTM 맵이며, 120명의 얼굴특징 정보(R, G, B, Y)를 학습시켜서 맵을 구성하였다. 그림 4는 실험에 사용할 입력 영상 video stream이며 각각 10장의 순차적인 data를 사용하였다

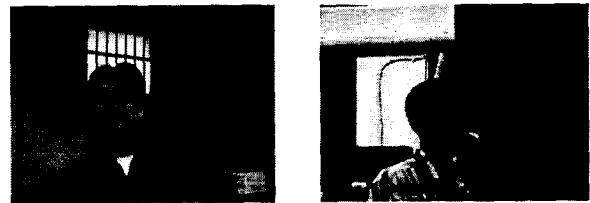


그림 4. 입력 video frames

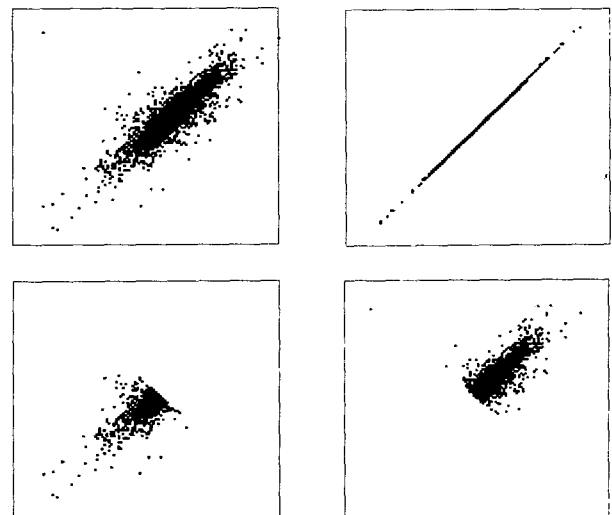


그림 5. PCA 영역분할 결과

그림 5는 BMA 정보에 관하여 PCA를 사용하여 객체와 배경정보의 영역을 분할한 결과이며 입력 벡터 영상, 주축 영상, 분류된 그룹 영상의 순으로 배열되어 있다.



그림 6. 입력영상에 대한 객체 추출 결과

그림 6은 입력 video frame에 대해 제안하는 기법으로 추정된 결과 영상이며 객체의 움직임이 빠를수록 추출된 객체의 윤곽선에서 약간의 오차가 있음을 확인하였다.

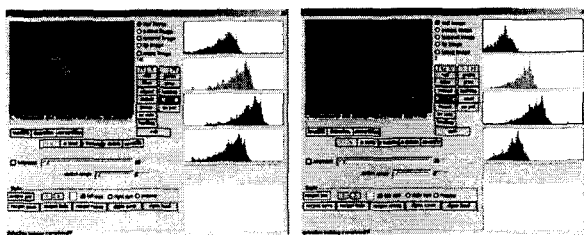


그림 7. 추출된 객체의 얼굴 영역 검출 결과

그림 7은 PCA로 분류된 data중 객체 정보만을 재복원하여 객체를 추출한 후, 동적 GTM 맵의 정보를 사용하여 얼굴 정보만을 추출해 내었다. 동적 GTM 맵의 정보를 사용하여 그림 7과 같이 얼굴 정보만을 추출해 내었으며, 이는 질의 영상에 얼굴영역의 픽셀 정보가 동적 GTM 맵 영역내 존재하기 때문에 얼굴 영역으로 인식한다. 우측 4개의 히스토그램은 질의 영상의 R, G, B, Y에 대한 성분을 히스토그램으로 표현하였다.

V. 결론

전체 영상에 대해서 직접 PCA 알고리즘을 적용하는 것은 많은 계산 속도와 메모리를 요구하게 된다. 이러한 문제를 해결하기 위하여 BMA 알고리즘을 이용하여 영상의 움직임 정보를 얻은 후 각각의 정보에 대하여 PCA를 적용하는 방법을 제안한다. PCA를 직접 사용하는 것보다 BMA를 사용하여 획득한 영상 정보를 PCA 알고리즘으로 분할하여 배경을 제외한 객체 부분만 복원함으로써 성능, 메모리 효율 면에서 좋음을 알 수 있었다.

동적 GTM 맵은 단지 색상 정보만을 사용하므로 배경 내의 유사 영역도 얼굴 영역으로 인식 할 수 있었으나 객체 정보만을 사용함으로써 오인식을 막을 수 있었다. 동적 GTM 알고리즘은 단 1회의 학습만으로 최적의 맵을 생성하므로 학습시간이 빠르며 생성된 동적 GTM 맵은 얼굴 정보를 가장 잘 나타낼 수 있는 임의의 특징인 R, G, B, Y 성분을 학습시킨 결과이므로 칼라 영상에서 얼굴영역만을 검출하는 성능이 우수함을 알 수 있었다.

VI. 참고 문헌

- [1] Ji-Hong Kim and Woo-Jin Song, "A Hierarchical Block Matching Algorithm Using Selective Elimination of Candidate Motion Vectors", IEICE Trans. Information and Systems, Vol. E82-D, No. 5, pp.985-992 1999.
- [2] Eli Saber and A. Murat Tekalp, "Face Detection Feature Extraction Using Color, Shape and Symmetry-Based Cost Functions", Proceedings of ICPR '96, pp. 654-658, 1996.
- [3] S. Gutta, J. Huang, B. Takacs and H. Wechsler, "Face Recognition Using Ensembles of Networks", Proceedings of ICPR '96, pp. 50-54, 1996.
- [4] Roberto Brunelli and Tomaso Poggio, "Face Recognition: Features versus Templates", IEEE Transactions on pattern analysis and machine intelligence, Vol. 15, No. 10, pp. 1042-1052, October 1993.
- [5] Srinivas Gutta and Harry Wechsler, "Face Recognition Using Hybrid Classifiers", Pattern Recognition, Vol. 30, No. 4, pp. 539-553, 1997.
- [6] Laurenz Wiskott, "Phantom Face for Face Analysis", Pattern Recognitions, Vol. 30, No. 6, pp. 837-846, 1997.
- [7] Lucas, B and Kanade.T (1981) An Iterative Image Registration Technique with an Application to Stereo Vision, Proc. of 7th International Joint Conference on Artificial Intelligence (IJCAI), pp 674-679
- [8] Carman Neustaedter, "An Evaluation of Optical Flow using Lucas and Kanade's Algorithm". Neustaedter 2002.
- [9] Ji-Hong Kim, Woo-Jin Song, "A Hierarchical Block Matching Algorithm Using Selective Elimination of Candidate Motion Vectors", IEICE Trans. Information and Systems, Vol.. E82-D, No. 5, pp. 985-992, 1999.
- [10] Oja, E., "A simplified neuron model as a principal component analyzer," Journal of Mathematical Biology 15, pp. 267-273, 1982.
- [11] 하성욱, 권기향, 강대성, "유전 목 지도의 동적확장", 정보과학회논문지:소프트웨어 및 응용 제29권 제6호 pp. 386-395 2002.6.