

Gabor 웨이브렛과 FCM 군집화 알고리즘에 기반한 동적 연결모형에 의한 얼굴표정에서 특징점 추출

Feature-Point Extraction by Dynamic Linking Model bas
Wavelets and Fuzzy C-Means Clustering Algorithm

신 영 숙*
(Young-Suk Shin)

요약 본 논문은 Gabor 웨이브렛 변환을 이용하여 무표정을 포함한 표정영상에서 얼굴의 주요 요소들의 경계선을 추출한 후, FCM 군집화 알고리즘을 적용하여 무표정 영상에서 차원의 대표적인 특징점을 추출한다. 무표정 영상의 특징점들은 표정영상의 특징점들을 추출하기 위한 템플릿으로 사용되어지며, 표정영상의 특징점 추출은 무표정 영상의 특징점과 동적 연결모형을 이용하여 개략적인 정합과 정밀한 정합 과정의 두단계로 이루어진다. 본 논문에서는 Gabor 웨이브렛과 FCM 군집화 알고리즘을 기반으로 동적 연결모형을 이용하여 표정영상에서 특징점을 자동으로 추출할 수 있음을 제시한다. 본 연구결과는 자동 특징추출을 이용한 차원모형기반 얼굴 표정인식[1]에서 얼굴표정의 특징점을 자동으로 추출하는 데 적용되었다.

주제어 Gabor 웨이브렛 변환, FCM(Fuzzy C-Means) 군집화 알고리즘, 얼굴표정의 특징점 추출, 동적 연결모형

Abstract This paper extracts the edge of main components of face with Gabor wavelets transformation in facial expression images. FCM(Fuzzy C-Means) clustering algorithm then extracts the representative feature points of low dimensionality from the edge extracted in neutral face. The feature-points of the neutral face is used as a template to extract the feature-points of facial expression images. To match point to point feature points on an expression face against each feature point on a neutral face, it consists of two steps using a dynamic linking model, which are called the coarse mapping and the fine mapping. This paper presents an automatic extraction of feature-points by dynamic linking model based on Gabor wavelets and fuzzy C-means(FCM) algorithm. The result of this study was applied to extract features automatically in facial expression recognition based on dimension[1].

1. 서론

최근 얼굴 표정인식을 위한 연구가 여러 가지 방법으

로 수행되어져 왔다. 표정인식에 대한 기존 연구들을 살펴보면 크게 두 가지 범주의 연구들로 나눌 수 있다.

첫 번째 범주의 연구들로는 동영상에서의 표정인식 연구들이다. Suma 등[2], Mase[3], Yacoob과 Davis[4]등의 연구들에서 살펴볼 수 있다. 이 방법들은 구체적인 특징요소를 추출하는 것이 아니라 움직임 영역을 추출함으로써 직접적으로 얼굴의 움직임을 표현한다. 두 번째 범주의 연구들은 정지영상에서 이루어진다. Turk와 Pentland [5]의 주성분 분석(Principal Component Analysis)은 학습 얼굴패턴의 차원을 감소시켜 얼굴공간으로 변환 후 다시 역변환하는 과정을 통하여 이루어지

* 조선대학교 전자정보통신공학부
Division of Electronics and Information Communication
Engineering, Chosun University
연구세부분야: 전문가시스템, 컴퓨터 비전, 패턴인식, HCI
주소 : 501-759 광주 광역시 동구 서석동 375번지.
조선대학교 전자정보통신공학부
전화 : 연구실: (062)230-7011
Cellular: 019-654-8473
FAX : (062)233-6896
E-mail : ysshin86@hanmail.net

는데, 높은 성능을 보이는 반면, 처리시간이 오래 걸리는 단점이 있다. Padgett과 Cottrell[6] 역시 주성분 분석을 적용하여 얼굴 표정을 분류하였는데, 이들은 눈과 입 영역만을 특징영역으로 적용하였다. 최근에는 34개의 표시된 점들에 대해 Gabor 웨이브렛 변환 후 추출된 특징 정보를 이용하여 비교적 높은 성능을 보여준다[7]. 그러나, 얼굴의 특징점들의 위치를 사전에 미리 정의하고 그 위치 상에서의 특징정보를 추출하기 때문에 현실적으로 적용하기에는 일일이 수작업으로 특징점들의 위치를 기록해야 한다는 한계점이 존재한다.

본 연구에서는 표정영상에서 얼굴의 특징점을 자동으로 추출하기 위하여 Gabor 웨이브렛과 FCM군집화 알고리즘을 기반으로 동적 연결모형을 이용하여 특징추출을 수행하였다. 이는 다차원 표정공간에서 저차원 특징 벡터로의 표상을 가능하게 한다. Pilowsky, Thornton과 Stokes[8]은 얼굴표정으로부터 특정 내적상태를 인식하는 데 필수적인 요소로 얼굴표정 특징요소들간의 거리라고 밝히고 있다. 이러한 얼굴표정 특징요소들간의 거리는 동적 연결모형을 이용한 무표정으로부터 표정영상으로의 연결을 통한 얼굴표정 특징점 간의 거리측정을 통해 실현될 수 있다.

논문의 전개는 다음과 같다. 먼저 모든 표정영상에 이 차원 Gabor 웨이브렛 계수 히스토그램을 이용하여 얼굴의 주요 요소들의 경계선을 추출하였다. 경계선 정보들은 FCM 군집화 알고리즘을 통해 감소된 차원의 특징벡터로 표상된다. 이러한 저차원의 특징벡터들은 동적 연결모형을 통해 무표정에서 표정영상으로 대응점간의 정합을 빠르게 수행될 수 있도록 하였다.

2. 특징 추출을 위한 영상 전처리

연구에 사용된 모든 표정 영상들은 거의 정면 영상만을 대상으로 하며, 영상처리에 사용되기 전에 두 눈 안쪽사이의 거리를 일정한 화소값으로 보정하므로써 다양한 얼굴 크기를 정규화 하였다. 또한 왼쪽 눈의 안쪽 끝 점의 좌표 위치를 일정 위치로 고정시켰다.

2.1 Gabor 웨이브렛 변환

얼굴의 특징점을 추출하기 위하여 얼굴의 텍스처 정보를 이용하며, Gabor 웨이브렛 변환을 사용하여 표정 영상들을 전처리 한다. 각각의 영상은 한쌍의 Gabor 커널(kernel)과 컨벌루션(convolution)된다. Gabor 웨이브렛은 다중해상도와 다중방향으로 이루어진 필터들로 하나의 영상에 대해 여러개의 표상을 생성한다. 이차원

Gabor 웨이브렛 함수는 Daugman[9]에 의하여 제시된 형식과 동일하다.

$$\psi_{\vec{k}}(\vec{x}) = \frac{k^2}{\sigma^2} \exp\left(-\frac{k^2 x^2}{2 \sigma^2}\right) [\exp(i \vec{k} \cdot \vec{x}) - \exp(-\frac{\sigma^2}{2})] \quad (2.1.1)$$

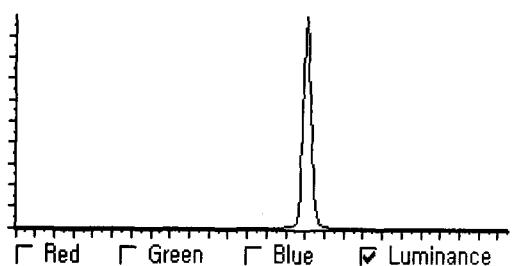
이 차원 웨이브벡터 $\vec{k} = (k, \theta)$ 는 필터의 공간 주파수 k 와 방향 θ 를 통제하며, 벡터 $\vec{x}(x, y)$ 는 영상을 이루는 각각의 화소를 나타낸다. σ 는 k 에 관련하여 Gaussian 원도우의 폭을 조절한다. (그림1)은 영상에 적용된 Gabor필터 쌍이다. 이 필터들은 90도 위상관계를 이루고 있다. 필터쌍들이 위상차이에 의하여 영상에 미치는 영향을 적게 하기 위하여 각각의 필터를 영상에 적용한 후, 생성된 2개의 컨벌루션된 영상 v_1, v_2 는 다음과 같이 처리된다.



(그림 1) 적용된 Gabor필터 쌍

$$S = \sqrt{v_1^2 + v_2^2} \quad (2.1.2)$$

(그림2)는 명도레벨에서 Gabor 웨이브렛 계수값들의 히스토그램을 나타낸 것이다. 이것은 히스토그램의 평균값을 이용하여 일정한 주파수와 방향에 대하여 얼굴의 주요 요소성분들의 경계선을 추출할 수 있음을 제시한다. (그림3)은 Gabor 웨이브렛 계수 히스토그램의 평균값을 적용하여 얼굴의 주요 요소들의 경계선이 추출된 결과이다.



(그림 2) Gabor웨이브렛 계수 히스토그램



(그림 3) 얼굴의 주요 요소
성분들의 경계선

3. 특징점 추출

이차원 Gabor 웨이브렛 계수 히스토그램의 평균값을 이용하여 추출된 점들은 얼굴의 주요 요소 성분들의 경계선 정보들이다. 이 영상에 있는 추출된 점들의 이웃하고 있는 화소들의 계수들을 살펴보면 높은 상관 관계를 보이며 중복된 값들로 나타난다. 이것은 영상을 이루는 화소들이 서로 상관성을 갖지 않으며 중복되지 않는 독립적인 계수를 갖는 대표적인 화소들로 다시 추출될 수 있는 가능성을 보인다. 이러한 독립적인 계수들을 갖는 화소들만을 추출하기 위해서는 화소들간의 거리 간격이 필요하며, 또한 얼굴의 주요 요소 성분들을 대표할 수 있는 기하학적인 특징점을 보존할 수 있어야 한다. 이러한 목적을 달성하기 위해서는 적절히 군집화되어 군집안에서 대표성 벡터들을 추출하는 것이 필요하다.

FCM 알고리즘은 표정영상의 주요 얼굴 요소(눈썹, 눈, 코, 입)들의 경계선들로부터, 대표적인 특징점을 추출하기 위한 템플릿으로 사용되어질 무표정 영상에 적용되었다.

3.1 무표정영상에서 대표적인 특징점 추출

FCM 군집화 알고리즘은 벡터 정량화를 기반으로 한 영상코드들을 압축하는 데 우수한 효과를 보인 사례들이 있다[10][11]. 따라서 본 연구에서는 얼굴의 주요 요소들의 경계선 정보들을 기반으로 얼굴의 기하학적인 특징들을 보유하면서 화소들간의 중복된 계수값들을 제거하고 대표성 벡터들만을 특징점으로 추출하기 위하여 FCM 군집화 알고리즘을 적용하였다.

FCM 군집화 알고리즘은 각 데이터의 점과 각 군집의

중심과의 거리를 염두에 두고 유사도 측정을 기반으로 한 목적함수의 최적화 방식이다. 이것은 n 차원의 유클리디안 공간상에서 N 개의 특징 벡터들로 이루어지는 집합 X 를 기반으로 한다.

$$X = \{x_1, x_2, \dots, x_N\},$$

사용된 목적함수는 다음과 같다.

$$J(U, c_1, \dots, c_c) = \sum_{i=1}^c J_i = \sum_{i=1}^c \sum_{j=1}^N u_{ij}^m d_{ij}^{-2} \quad (3.1.1)$$

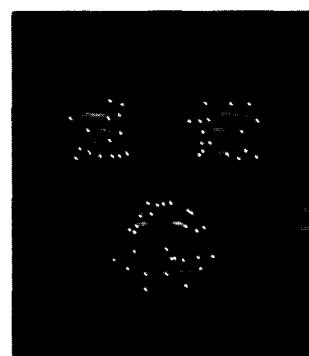
$$\sum_{i=1}^c u_{ij} = 1, \forall j = 1, \dots, N \quad (3.1.2)$$

c_i 는 각 군집 i 의 중심을 의미하며, $d_{ij} = \|c_i - X_j\|$ 는 각 데이터의 점과 각 군집의 중심과의 유클리디안 거리를 나타낸다. 소속함수 u_{ij} 는 0과 1사이의 실수값으로 각 군집에서 데이터의 소속정도를 평가한다. 식 (3.1.1)이 최소값에 도달하기 위한 필요 조건으로 식 (3.1.3)과 식 (3.1.4)이 있다.

$$c_i = \frac{\sum_{j=1}^N u_{ij}^m X_j}{\sum_{j=1}^N u_{ij}^m}, \quad (3.1.3)$$

$$u_{ij} = \frac{1}{\sum_{k=1}^c \frac{d_{ik}}{d_{kj}}^{2/(m-1)}}. \quad (3.1.4)$$

특징점 추출은 다음 단계를 이용하여 이루어졌다[12].



(그림 4) FCM을 적용하여
추출된 특징점($c=60$)

- 단계1: 데이터가 각 군집에 소속될 정도는 식 (3.1.2)을 만족하는 0과 1사이의 임의의 값으로 초기화한다.
- 단계2: 각 군집의 중심 c_i 은 식(3.1.3)을 사용하여 계산된다.
- 단계3: 목적함수는 (3.1.1)에 의하여 계산되며, 바로 전 분류결과와 현재의 분류결과의 에러차이를 비교하여 반복여부를 판단한다.
- 단계4: 식 (3.1.4)를 이용하여 데이터가 각 군집에 소속될 정도를 갱신하며, 단계2로 돌아간다.

(그림 4)은 무표정영상에서 FCM 군집화 알고리즘을 적용하여 군집수 60을 적용할 경우, 얼굴의 주요 요소들의 기하학적인 특징점들의 추출을 보인다. 본 연구에서 사용된 군집수는 인간의 시각으로 얼굴의 주요 요소들과 동일한 위상관계를 반영하는 범위에서 실험을 통하여 결정하였다.

3.2 동적 연결모형을 이용한 표정영상에서 특징점 추출

무표정 영상에서 대표적인 특징점들이 추출된 후에, 추출된 특징점들은 표정이 있는 영상의 표정변화의 정도를 측정하는 기준으로 사용된다. 따라서 무표정 영상의 추출된 특징점들은 사전에 추출된 표정영상들의 경계선 정보를 나타내는 점들로부터 특징점을 추출하는 템플릿으로 활용되었다. 표정영상의 특징점들은 동적 연결모형을 이용하여 추출되었다. 연구에 사용된 동적 연결모형은 Malsburg[13]와 Bienenstock[14]에서 사용되는 동적 연결구조와 유사하다. 이 알고리즘은 입력영상에 쪽워진 직사각형의 그래프에 의한 동적 연결구조를 갖는다. 동적 연결구조의 교차점의 특징 벡터들간의 비교를 통하여 얼굴간의 유사성을 판단하게 되며, 에지(edge)들의 거리를 비교함으로써 왜곡정도를 고려하게 된다.

무표정 영상과 표정 영상간의 동적인 연결을 위하여 두 개의 다른 영역이 고려되었다. 즉, 무표정 영역을 N 영역으로, 표정 영역을 E 영역으로 표기한다. 무표정 영역과 표정 영역은 (식 2.1.1)에서 정의된 Gabor 웨이브렛 변환을 통해 얻어진 젯(jet)들로 특징 벡터를 이룬다. 두 영역간의 동적인 연결을 위해 본 연구에서는 2개의 주파수, $k = \left\{ \frac{\pi}{4}, -\frac{\pi}{8} \right\}$ 와 30도 간격의 7개의 방향으로 이루어지는 14개의 Gabor 필터를 적용하였다.

무표정 영상에서 표정영상으로의 특징점 연결은 무표정 영상의 사전 선택된 점들과 표정영상의 연결되는 점

들과는 대략적으로 동일한 위상관계를 갖는다는 제한 아래 수행되어진다. 이는 영상처리에 사용되기 앞서 영상 정규화과정을 통해 얻어질 수 있다. 표정영상과 대략적으로 동일한 위상 관계를 보이는 무표정 영상의 특징점들은 다음의 두단계 과정을 통해 표정영상에서의 특징점으로 대응된다.

[1단계] 표정영상에서 얼굴의 주요요소들의 경계선 정보들을 기반으로 얼굴의 기하학적인 특징들을 보유하면서 대표성 벡터들만을 특징점으로 추출할 수 있는 "개략적인 정합"과정이 존재한다. 표정영상들의 추출된 경계선 데이터들은 1차적으로 무표정영상과 동일한 군집수로 분류된다. 분류된 각 군집의 중심을 이루는 데이터들의 좌표점과 무표정 영상의 대응점간의 유클리디안 거리를 측정하여 가장 작은 거리를 갖는 군집의 중심이 표정영상으로의 1차적인 대응점이 된다. (식 3.2.1)은 N 영역과 E 영역간의 유클리디안 거리 계산을 나타낸다.

\vec{x}_i^N 은 무표정 영상에서 템플릿으로 사용될 특징점들

\vec{x}_j^E 을 나타내며, \vec{x}_j 은 표정영상에서 경계선 데이터들이 무표정 영상과 동일한 군집수로 분류된 후 각 군집의 중심을 이루는 데이터들을 나타낸다.

$$\Delta_{ij}^{NE} = \vec{x}_i^N - \vec{x}_j^E \quad (3.2.1)$$

[2단계] 1단계에서 추출된 대표성 벡터를 포함하고 있는 각 군집 안의 점에서 Gabor 젯(jet)을 이용한 "정밀한 정합"과정을 수행한다. 1차적으로 무표정영상에서 표정영상의 대략적인 위치 결정 후, 다음으로는 무표정 영상의 각 특징점에서의 Gabor 웨이브렛 변환 후의 방향과 주파수로 이루어지는 Gabor 젯들과 표정 영상에서 1차적으로 선택된 군집안을 이루는 데이터들의 Gabor 젯간의 유사도 비교를 통하여 최종적인 대응점을 결정한다. (식 3.2.2)의 S는 무표정영상과 표정영상의

대응점간의 유사도 측정함수이다. \vec{j}_i^N 은 무표정 영상에서 템플릿으로 사용된 특징점들에서 Gabor 웨이브렛 변환 후의 방향과 주파수로 이루어지는 Gabor 젯들을

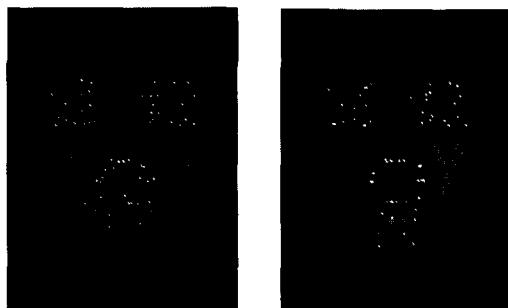
나타낸다. \vec{j}_i^E 은 표정 영상에서 [1단계]과정으로 선택된 군집의 데이터들에서 Gabor 웨이브렛 변환 후의 방향과 주파수로 이루어지는 Gabor 젯들을 나타낸다.

$$S(\vec{J}_i^N, \vec{J}_i^E) = \frac{\vec{J}_i^N \cdot \vec{J}_i^E}{\|\vec{J}_i^N\| \|\vec{J}_i^E\|} \quad (3.2.2)$$

따라서 무표정 영상의 선택된 점들에서 표정 영상의 점들로의 동적 연결모형은 최적화문제로 표현될 수 있으며, 목적함수 H 는 제안된 점들간의 연결 정도를 평가하게 된다. 따라서 식(3.2.3)과 같은 최소비용함수로 표현될 수 있다.

$$H = \Delta_{ij}^{NE} - \sum S(\vec{J}_i^N, \vec{J}_j^E) \quad (3.2.3)$$

- (1) 목적함수 H 는 최소값에 도달해야 한다.
- (2) 표정영상에서 찾아지는 대응점들은 무표정 영상의 대응점과 대략적으로 동일한 위상관계를 형성하도록 하는 최대거리를 초과하지 않아야 한다.



(그림 5) 동적 연결모형이 적용된 표정영상에서 특징점 추출

(그림 5)은 동적 연결모형을 적용하여 표정영상에서 특징점이 추출된 결과를 보인다. 추출된 특징점들은 표정을 인식하기 위하여 신경망의 입력층의 노드로 사용되었다. 입력노드는 표정영상에서 추출된 특징점들의 좌표값에서 무표정의 각각 대응되는 특징점들의 좌표값의 차이로 이루어진다. 따라서 입력노드는 60개의 특징점 각각에 대해 x, y 좌표값의 차이 값으로 이루어진 120개의 입력노드로 형성된다. [표1]은 이 값들이 신경망 분류기의 입력정보로 사용되기 위해 0과 1사이의 값으로 정규화된 결과를 보인다. 본 연구결과는 신경망 분류기의 입력정보로 이용되어 쾌-불쾌차원과 각성-수면 차원으로 이루어지는 차원모형 기반 얼굴 표정인식에 적용되었다[1].

<표 1> 추출된 특징점들이 표정인식을 위해 신경망의 입력정보로 사용(정규화이후)

입 력 데 이 터	0.37	0.45	0.35	0.50	0.28	0.43	0.33	0.55	0.47	0.61	0.60	0.50
	0.44	0.48	0.56	0.43	0.60	0.55	0.60	0.64	0.65	0.45	0.72	0.45
	0.53	0.48	0.53	0.57	0.58	0.52	0.40	0.57	0.51	0.39	0.42	0.57
	0.60	0.64	0.51	0.55	0.58	0.55	0.56	0.41	0.37	0.41	0.44	0.32
	0.51	0.64	0.47	0.50	0.56	0.55	0.44	0.61	0.44	0.57	0.49	0.36
	0.53	0.66	0.47	0.52	0.44	0.45	0.42	0.57	0.42	0.59	0.44	0.43
	0.56	0.57	0.51	0.55	0.56	0.57	0.56	0.59	0.53	0.55	0.42	0.57
	0.60	0.43	0.60	0.52	0.67	0.50	0.37	0.45	0.19	0.89	0.00	0.57
	0.14	0.84	0.05	0.70	0.60	1.00	0.16	0.16	0.16	0.14	0.37	0.02
	0.40	0.05	0.21	0.91	0.40	0.00	0.56	1.00	0.91	0.11	1.00	0.66

4. 결론 및 논의

본 논문은 Gabor 웨이브렛 변환과 FCM 군집화 알고리즘을 적용하여 무표정 영상에서 저차원의 대표적인 특징점을 추출할 수 있었으며, 이 특징점을 템플릿으로 동적 연결모형을 이용하여 표정영상에서 특징점들을 자동으로 추출할 수 있었다. 본 연구를 이용한 특징점 추출은 얼굴표정 특징요소들간의 움직임 거리를 측정 가능하게 함으로써 얼굴표정의 내적상태를 인식 가능하게 한다[1].

본 연구가 보다 범용적인 자동특징추출 시스템이 되기 위해서는 앞으로 다양한 얼굴크기와 기울어짐이 있는 변환영상에서도 특징 추출이 가능한 시스템으로 확장되어야 할 것이다. 제안된 모형을 개선하기 위해서는 기울어진 영상에서 기울어진 각도를 자동으로 감지한 후 이를 보정하도록 하는 연구가 필요하다. 영상들이 동일한 위상관계를 형성한다는 제한점은 정지영상의 표정인식에서 대두되는 문제로서, 동영상의 적용시 완화될 수 있을 것으로 판단된다.

참고문헌

- [1] 신영숙(2001). 자동 특징추출을 이용한 차원모형 기반 얼굴 표정인식. 연세대학교 본대학원 인지과학 협동과 박사학위논문.
- [2] Suma, M., Sugie, N., Fujimura, K.(1978). A Preliminary note on pattern recognition of emotional expression. In proceedings of the 4th International Joint Conference on Pattern Recognition, 408-410.
- [3] Mase, K.(1991). Recognition of facial expression

- from optical flow. *IEICE Transactions*, E 74, 10, 3474-3483.
- [4] Yacoob, Y. & Davis, L. (1996). Recognizing human facial expressions from long image sequences using optical flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18, 636-642.
 - [5] Turk, M. A. & Pentland, A. P.(1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1), 71-86.
 - [6] Padgett, C. & Cottrell, G.(1995). Identifying emotion in static face images. In *proceedings of the 2nd Joint Symposium on Neural Computation*, 5, 91-101.
 - [7] Lyons, M. & Akamatsu, S. (1998). Coding facial expressions with Gabor wavelets. *Proceeding of the third international conference on automatic face and gesture recognition*, 200-205.
 - [8] Pilowsky, I., Thornton, M., & Stokes, B. B.(1986). Towards the quantification of facial expressions with the use of a mathematic model of the face. In H. D. Ellis, M. A. Jeves FRSE, F. Newcombe & A. Young(Eds.), *Aspects of face processing*. Dordrecht: Martinus Nijhoff Publishers.
 - [9] Daugman, J.(1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters, *J. Opt. Soc. Amer.*, vol. 2, 1160-1169.
 - [10] Karayiannis, N. B.(1994). Maximum entropy clustering algorithms, Systems, Neural Nets, and Computing Technical Report No. 94-06, University of Houston June.
 - [11] Karayiannis, N. B. and Pai, P.(1995). Fuzzy vector quantization algorithms and their application in image compression, *IEEE Transactions on Image Processing*, in press.
 - [12] Bezdek, J. C.(1973). Fuzzy mathematics in pattern classification. Ph.D. thesis. Applied Math. Center, Cornell University, Ithaca.
 - [13] Malsburg, C.v.d.(1981). The correlation theory of brain function. *Internal Report 81-2, Dept. of Neurobiology*. MPI Biophysical Chemistry.
 - [14] Bienenstock, E. & Malsburg, C.v.d.(1987). A neural network for invariant pattern recognition. *Europhys. Lett* 4, 121-126.