

論文91-28B-12-7

신경 회로망과 Log-Polar Sampling 기법을 사용한 항공기 영상의 연상 인식

(Neural-Network and Log-Polar Sampling Based Associative
Pattern Recognizer for Aircraft Images)

金 鐘 岳*, 金 麟 哲**, 秦 成 一***

(Jong Oh Kim, In Cheol Kim, and Sung Il Chien)

要 約

본 논문에서는 신경 회로망을 이용한 항공기 인식용 영상 패턴 인식기 설계를 목표로 하였다. 2 차원 영상의 회전 및 크기 변화에 무관한 왜곡 불변 피처를 얻기 위해서 현재 물체 영상의 트래킹 및 인식 분야에서 최근 연구가 활발히 진행되고 있는 생물학적 시각 모델인 Log-polar Sampling과 Fourier 변환 기법을 도입하였다. 특히 영상 인식 결과와 더불어 영상 자체도 복원시킴으로써 인식 신뢰도가 복원 영상에 시각화 할 수 있도록 하였다. 다층 퍼셉트론 모델 신경 회로망의 학습에는 error backpropagation 알고리즘을 사용하였다.

Abstract

In this paper, we aimed to develop associative pattern recognizer based on neural network for aircraft identification. For obtaining invariant feature space description of an object regardless of its scale change and rotation, Log-polar sampling technique recently developed partly due to its similarity to the human visual system was introduced with Fourier transform post-processing. In addition to the recognition results, image recall was associatively performed and also used for the visualization of the recognition reliability. The multilayer perceptron model was learned by backpropagation algorithm.

I. 서 론

최근 패턴 인식 분야에서 새로운 인식기로 매우 유용하게 사용되고 있는 신경 회로망⁽¹⁾은 내재한 대규모 병렬 처리 기능(massive parallel processing)에

의해 하드 웨어 구현시 실시간 처리가 용이하고 적응 학습기능(adaptive learning capability)과 결함 극복 능력(fault tolerance capability)을 가지고 있어 널리 연구되어 오고 있다. 그리고 이러한 신경 회로망은 원하는 목적에 적용하는데 비교적 개발 시간이 짧은 장점을 가지고 있고, 장차 VLSI나 광 컴퓨터 기술등에 대한 연구가 완료되어 하드웨어적 구성이 실현되면 인식 과정에서 실시간 처리가 요구되는 항공기 인식 등에 매우 유용함을 알 수 있다. 본 연구에서는 이러한 특성을 가지는 신경 회로망 기법을

*準會員, ***正會員, 慶北大學校 電子工學科
(Dept. of Elec. Eng., Kyungpook National Univ.)

**正會員, CAS(株) 研究所
(Computer Aided System Engineering)

接受日字: 1991年 6月 14日

항공기 영상 인식 문제에 적용시켰다.

항공기 영상을 기술하기 위해 우선 왜곡 불변 피쳐 스페이스(invariant feature space)에 대한 연구가 선행되어야 한다. 즉 물체의 2차원 영상내에서 이동, 회전에 무관하며 아울러 zooming등에 의한 물체의 크기 변화에 불변인 피쳐의 도입이 필요하다. 이러한 대표적 피쳐로는 moment,^[2] Fourier descriptor,^[3] (L, ϕ) feature space^[4]등을 들 수 있다. 그러나 이러한 feature set은 모두 일반적으로 많은 계산량을 요하며 완전한 병렬 처리가 불가능하다. 따라서 본 연구에서는 완전한 병렬 처리 구조가 가능하며, 영상의 회전 및 크기 변화를 이동으로 변환시키는 Log-polar sampling^[5,6,7,8] 기법을 신경 회로망 전단으로 도입한다. 따라서 Log-polar sampling된 영상을 Fourier 변환시켜 절대값을 취하면 원하는 원 영상의 회전, 크기 변화에 무관한 왜곡 불변 피쳐 (invariant feature)를 얻을 수 있으며, 이러한 Fourier 변환 값을 적절히 sample하여 dimension을 축소시키면 신경 회로망에 사용할 feature space를 구성시킬 수 있다.

본 논문의 BP 알고리즘^[9]으로 학습된 다층 Perceptron 신경 회로망 인식기의 특징 중 하나는 신경 회로망의 출력으로서 인식 결과뿐만 아니라 정점(vertex) 좌표축으로 묘사된 표준 영상을 출력한다는 점이다. 즉 출력 영상을 구성하기 위해 vertex영상으로부터 좌표축을 적절히 정규화하여 인식 결과와 함께 학습시킨다. 이는 영상의 복원 뿐만 아니라 입력 영상의 인식이 이루어지더라도 그 신뢰성이 약하면 복원 영상이 왜곡되게 되어 인식 신뢰도를 시각적으로 측정할 수 있게 하였다.

II. 항공기 영상 인식 시스템

1. 항공기 영상 인식 시스템의 개요

본 연구에서는 인간의 시각 시스템(Human Visual System)과 비슷한 구조를 가지며, 영상의 회전과 크기 변화를 이동으로 변환함으로써 목표 추적^[6] 물체 인식^[5,8]에 사용되는 Log-polar sampling과 물체의 이동에 불변인 Fourier 변환 알고리즘을 적용하여 항공기 영상에 대해 회전과 크기 변화에 불변인 왜곡 불변 피쳐를 추출하였다. 또한 추출된 피쳐를 저주파 중심으로 다시 sampling 함으로써 MLP(multi-layer perceptron)의 입력 데이터량을 상당히 감소시켰을 뿐만 아니라, 신경 회로망의 학습 시간도 크게 감소시켰다. 그러나, Log-polar sampling은 물체의 중심점 좌표가 어긋날 경우 sampling된 영상에 왜곡이 발생함으로 인해 sampling전 중심점 좌표를 일치시

키는 과정이 선행되어야 한다. 본 실험에서는 사용된 항공기 영상을 256×256영상 평면의 중심으로 이동시킨 뒤 Log-polar sampling을 하여 전술한 문제점을 해결하였다.

입력 데이터에 대해 연상된 신경 회로망의 출력 데이터는 영상 복원을 위해 기준 영상으로 부터 구한 정규화된 정점 좌표값과 인식을 위한 부호 벡터로 나뉘어 진다. 전술한 항공기 영상의 인식 및 복원 시스템의 구성은 그림1과 같다.

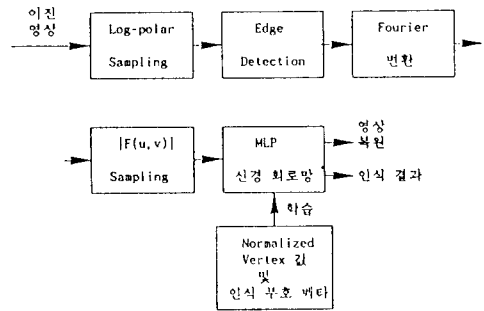


그림 1. 왜곡 불변 특성을 이용한 항공기 영상의 복원 및 인식 과정

Fig. 1. Schematic diagram of the proposed image recognition and recovery system.

2. 항공기 인식을 위한 데이터 베이스

본 실험에서 사용된 항공기 영상은 미국의 Purdue 및 Carnegie Mellon 대학에서 개발한 model-based CAD 소프트웨어^[10]를 이용하여 제작하였다. 제작된 영상은 이진 영상이며, 그림2와 같이, 256×256화소 크기를 가지는 영상 평면 내에서 화소 크기가 약 120×120이고 회전 각도가 0°인 항공기 영상을 기준 영상으로 사용하였다. 학습이 끝난 뒤 영상의 크기 변화와 회전 변화에 대한 회로망의 성능을 테스트하기 위해 256×256인 영상 평면내에서 크기가 각각 80×80, 100×100, 120×120, 140×140, 160×160이고, 회전 각이 0°, 120°, 240°인 영상을 제작, 실험용 영상으로 활용하였다. 그림3은 실험용 DC10 항공기 영상의 한 예를 나타낸다.

3. 왜곡 불변 피쳐 스페이스(Invariant Feature Space)

Log-polar sampling은 직각 좌표계의 영상을 Log-polar 좌표계의 영상으로 변환한다. 즉 직각 좌표상

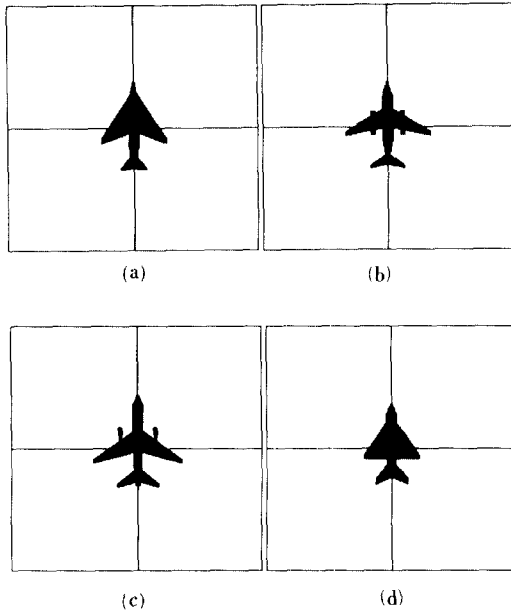


그림 2. 기준 항공기 영상
 Fig. 2. Four reference images.
 (a) DC10, (b) MiG21, (c) Phantom,
 (d) B737.

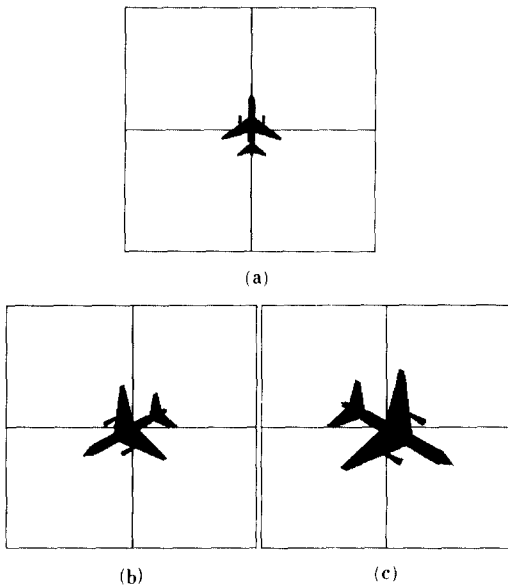


그림 3. 실험용 DC10 항공기 영상
 (a) 크기:80×80, 회전각:0°
 (b) 크기:120×120, 회전각 120°
 (c) 크기:160×160, 회전각:240°
 Fig. 3. DC10 aircraft images for experiment.
 (a) scale:80×80, rotation:0°,
 (b) scale:120×120, rotation:120°,
 (c) scale:160×160, rotation:240°.

의 한 점 (x, y) 가 식(1)과 같은 방법으로 Log-polar 평면상의 한 점 (u, v) 로 변환된다.

$$(u, v) = (\log(r), \theta) \quad (1)$$

여기서 $r = (x^2 + y^2)^{1/2}$ 이고, $\theta = \tan^{-1}(y/x)$ 이다.

본 실험에서는 화소의 크기가 256×256 인 직각 좌표 영상의 $(128, 128)$ 화소점을 중심으로 하여 5.6° 씩 떨어진 각 방사선에서 1.079925^n ($n=0, 1, 2, \dots, 63$) 간격으로 비선형 샘플링 함으로써 256×256 화소 크기의 영상에서 64×64 화소 크기의 Log-polar sampling된 영상을 얻었다. 여기서 5.6° 는 360° 를 64로 나누어 64×64 Log-polar 평면에서 직각 좌표계의 회전 변환인 θ 를 v 축으로의 이동으로 나타내기 위한 것이며, 1.079925 는 $1.079925^{63} = 127$ 이 되어 각 θ 에 대하여 64개를 샘플링함으로써 크기 변화를 u 축으로의 이동으로 나타내기 위한 것이다. 그림 4에서는 전술한 변화 과정을 나타내고 있다. 따라서, 이 변환은 그림 5의 sampling된 DC10 영상에서 알 수 있듯이 직각 좌표계에서 영상의 크기 변화를 Log-polar 좌표계에서 수직축으로의 이동으로, 회전 변화를 수직축으로의 순환 이동(cyclic shift)으로 바꾸어 준다. 그러나, Log-polar sampling은 직각 좌표상에서의 영상의 크기 변화가 Log-polar 좌표상에서 영상의 직접적인 이동으로 나타나지 않는다. 즉, 영상내의 내부 화소가 모두 1인 경우 영상이 확대되면 Log-polar sampling된 영상은 u 축으로의 이동으로 나타나지만 그로 인하여 생긴 빈 공간은 원 영상의 중심으로부터 샘플링된 화소로 채워져 결국 이동이 아닌 확장(stretching) 효과를 나타내게 된다. 따라서 확장 효과가 발생하더라도 변환된 영상의 경계 부분만을 고려한다면 이동이 발생했음을 알 수 있으므로 본 실험에서는 변환된 영상에서 경계 부분만을

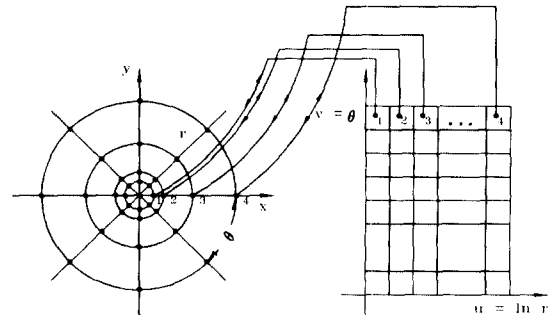


그림 4. Log-polar 샘플링 구조
 Fig. 4. Log-polar sampling structure.

추출함으로써 전술한 문제점을 해결하였다.

경계 추출 알고리즘에는 Sobel operator, Laplacian operator 등 여러 가지가 있으나, 본 실험에서는 간단한 raster scan 방식을 사용하여 경계를 추출하는 것으로 충분하였다. 즉, 변환된 영상이 이진 영상이므로 영상의 좌측 상단에서부터 우측으로 scan 하면서 각 화소의 이진 값이 변화하는 부분만을 고려하여, 영상이 변화하면 '1'로 그렇지 않으면 '0'으로 고정하여 경계를 추출하였다. 그림5의 하단에는 DC 10에 대해 Log-polar 좌표계로 변환된 영상으로부터 경계를 추출한 영상을 나타내었다. 그림5의 (a), (b), (c)에서 샘플링된 화소로 인한 확장 효과가 u축으로의 직접적인 이동으로 나타남을 알 수 있다.

Log-polar 좌표계로 변환된 영상은 크기 변화와 회전 변화를 이동으로 나타낸다. 따라서 Fourier 변환을 Log-polar sampling된 영상에 적용하여 크기 성분인 $|F(u, v)|$ 만을 고려하면, 변환된 영상은 이동에 불변한 성질을 가진다. 그림6에서는 DC10 항공기 영상중 크기가 120×120 인 기준 항공기 영상에 대하여 Fourier 변환의 크기 값인 $|F(u, v)|$ 를 3차원 그래프로 나타내었다. 그림에서 중심 부근에 가장 높이 솟아 있는 부분이 DC값을 나타내며, 이 DC값을 중심으로 magnitude값이 감소함으로써 Fourier 변환시 영상 정보가 DC값을 중심으로한 저주파 성분에 집중됨을 알 수 있다.

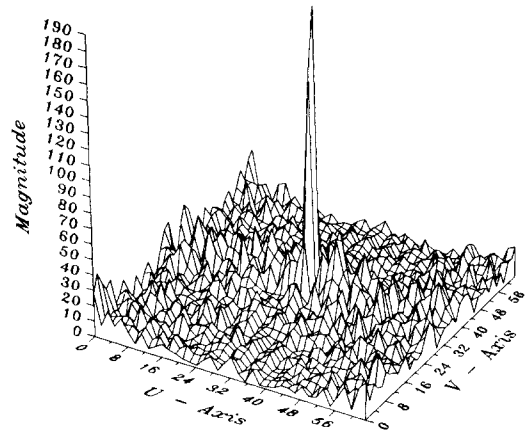


그림 6. DC10 기준 항공기 영상의 $|F(u, v)|$ 에 대한 3차원 graph

Fig. 6. 3-D graph of $|F(u, v)|$ of DC10 reference image.

III. MLP의 학습 알고리즘

본 실험에서는 항공기 영상의 인식과 복원을 위한 MLP의 학습 알고리즘으로, backpropagation(BP) 학습 알고리즘을 사용하였으며 또한 BP 알고리즘¹⁾의 단점인 느린 학습시간을 개선하기 위해 Fahlman¹¹⁾이 제안한 수정된 BP 알고리즘을 사용하였다. 이에 사용된 에러 기준 함수는 식(2)와 같다.

$$E = \sum_{p=1}^M \sum_{j=1}^N (t_{pj} - o_{pj})^2 \quad (2)$$

여기서 t 는 원하는 출력값(target value)이고 o 는 실제 sigmoid를 통한 출력 뉴런의 출력값을 나타내며, M 은 입력 패턴수를 그리고, N 은 출력 뉴런 수를 나타낸다. BP 알고리즘은 여러 경우에 광범위하게 사용되고 있고 참고문헌¹⁾에서 상술되어 있으므로 여기서는 자세한 언급은 피하고 Fahlman이 제시한 수정형 학습 알고리즘을 다음과 같이 요약 정리하여 수록한다.

(1) 먼저 아래층의 i 번째 뉴런과 윗층의 j 번째 뉴런 사이의 연결 강도(synaptic strength) 값인 w_{ji} 와 각 뉴런의 문턱값 θ_j 를 랜덤하게 초기화한다.

(2) 출력층에서 각각의 뉴런에 대한 실제 출력값을 계산한다.

$$o_j = 1 / (1 + \text{EXP} - [\sum_i w_{ji} o_i + \theta_j]) \quad (3)$$

(3) 출력 에러 정정 과정에서 출력층과 은닉층 각 뉴런에서의 에러값 δ_i 를 다음과 같이 계산한다.

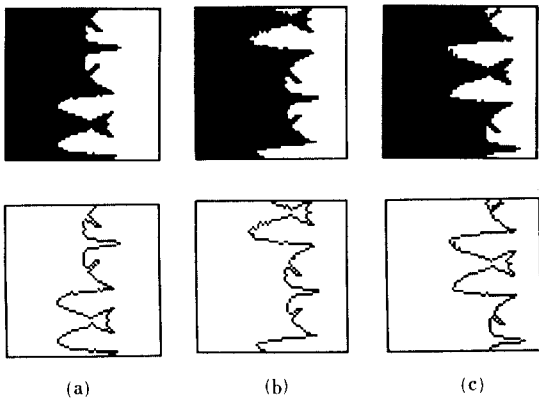


그림 5. DC10 실험용 항공기 영상에 대한 Log-polar sampling과 경계선 추출

Fig. 5. Log-polar sampling and edge extraction.

(a) scale: 80×80 , rotation: 0° ,

(b) scale: 120×120 , rotation: 120° ,

(c) scale: 160×160 , rotation: 240° .

$$\delta_1 = [\tanh^{-1}(t_j - o_j)] (o_j(1 - o_j) + 0.1) \text{ (출력층)} \quad (4)$$

$$\delta_j = o_j(1 - o_j) \sum_k \delta_k w_{kj} \text{ (은닉층)} \quad (5)$$

(4) 각 시냅스 값을 조정한다.

$$\Delta W_{ji}(t) = \epsilon \delta_j o_i + \alpha \Delta W_{ji}(t-1) \quad (6)$$

여기서 ϵ 는 학습률(learning rate)을 나타내며 α 는 관성항(momentum)을 나타낸다.

IV. 신경 회로망 인식기 Simulation

1. |F(u, v)|sampling과 신경회로망 학습을 위한 입·출력 패턴 벡터

Log-polar sampling중 경계 부분만을 추출하여 Fourier 변환할 경우 변환된 영상의 에너지는 |F(u, v)|의 DC값을 중심으로 저주파수 부근에 집중된다. 또한 Fourier 변환 특성상 영상을 변환할 경우 변환된 영상은 원점을 중심으로 대칭인 구조를 가지므로 우리는 DC값을 중심으로 반경이 10인 반원을 그려 정보가 집중되어 있는 152개의 피쳐를 일정하게 샘플링함으로써 전 영상을 모두 MLP의 입력 데이터로 사용하지 않고서도 안정된 영상복원과 인식이 이루어지도록 하였다. 그림7에서는 그림6의 Fourier 평면으로부터 샘플링된 152개의 입력 데이터를 나타낸다. 그림7과 그림6을 비교해보면 그림6의 DC값이 다른 샘플링 값에 비해 상당히 큼을 알 수 있다. 신경회로망의 입력 데이터는 0과 1사이의 값으로 고정되어야 하므로 DC값을 포함하여 정규화될 경우 상

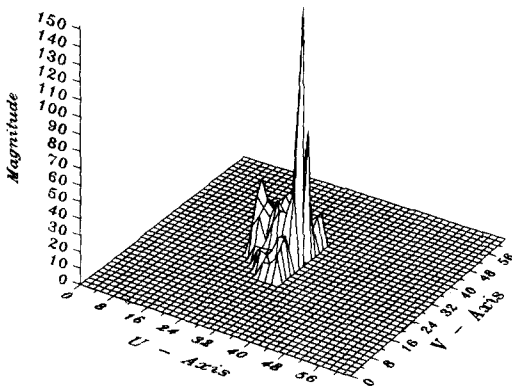


그림 7. DC10 기준 영상에 대해 추출된 152개의 입력 데이터

Fig. 7. Sampled input data of DC10 reference image.

대적으로 추출된 나머지 값들이 작아져 신경망의 학습에 큰 효과를 기대할 수 없다. 또한 항공기 상호간의 유사성으로 인해 너무 작은 값으로 정규화될 경우 데이터상에 별다른 차이가 없어 영상의 복원에 다소 어려움이 따를 것으로 예상되므로 본 실험에서는 DC값을 제외시켰다.

출력 벡터는 영상 복원과 인식 부분으로 나뉘어지며, 이중 영상복원 부분은 신경 회로망에 저장할 표준 영상에 대한 정보이다. 그러나 영상 정보는 원래 데이터량이 방대하므로 본 연구에서는 항공기 영상의 경계선(boundary)을 polygon으로 근사(approximation)시켜서 그림2에서 보인 120×120 기준 항공기로부터 polygon의 vertex X와 Y 정보를 추출하였다. 그림8은 기준 영상으로 부터 vertex를 추출한 출력 영상을 나타낸다. 그러나 신경 회로망의 출력 뉴런수는 고정되어야 하므로 우리는 그 수를 X축과 Y축에 대하여 각각 최대 vertex수(B737=38)인 38개로 정하였다.

또한 항공기의 인식을 위해 4bit 즉 4개의 출력 뉴런을 사용하였다. 이러한 목적의 부호 벡터는 1개의 "1" bit와 3개의 "0" bit로 이루어져 있으며 "1" bit의 위치로서 항공기를 인식하도록 하였다. 이 부호 벡터는 각 기종에 대해 표1처럼 정의되었다. 결론적으로 본 연구에서는 152개의 입력 node와 38×2+4=80개의 출력 node를 정의하였다.

표 1. 항공기 인식을 위한 출력 인식 부호 벡터
Table 1. Aircraft recognition code vector.

Aircraft	DC10	MiG21	Phantom	B737
인식 부호 벡터	0001	0010	0100	1000

2. 신경 회로망 학습 및 최적 신경망 구조

이론상으로는 각 기종마다 한개의 영상으로 학습을 완료할 수 있으나, 입력 영상의 크기 변화나 회전에 따른 Log-polar sampling이나, Fourier 변환에 변이(variation)가 존재하며, 특히 항공기의 회전에 Log-polar가 민감하므로 인해, 학습 영상으로 각 항공기마다 120×120의 크기를 갖고 0°, 120°, 240° 회전한 항공기를 사용하였다. 또한 학습후 신경 회로망의 성능을 시험하기 위해서 전술한 세 각도와 크기가 80×80에서 160×160까지 변화시킨 총 60개의 실험용 영상을 사용하였다.

본 논문에서는 은닉층의 경우 1개와 2개, 은닉층의 뉴런 수는 각 은닉층에 대하여 3, 5, 10, 15인 회로망

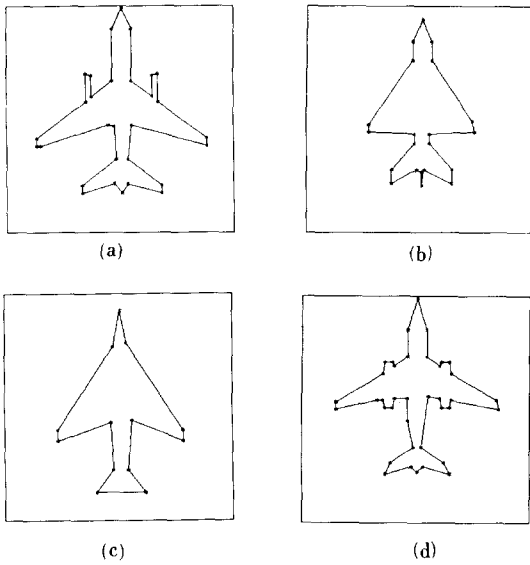


그림 8. Vertex를 추출한 MLP 출력 영상
Fig. 8. Recovered vertex images.

(a) DC10, (b) MiG21,
(c) Phantom, (d) B737.

을 설계하여 은닉층의 최적화 작업을 수행하였다.

이 결과를 표2에 나타내었다. 여기서 평균 패턴 에러란 실험용 항공기 영상에 대한 패턴 에러(Pss)를 전 패턴에 대해 더한 후 총 패턴 수로 나눈 값을 말하며, 또한 평균 최대 화소차는 기준 항공기 영상의 정점 좌표값과 비교해서 화소의 자가 가장 심한 좌표 값을 모든 패턴에 대하여 더한 뒤, 이 값을 전체 패턴 수로 나눈 값을 의미한다. 또한 표2에서 3-3은 각 은닉층에 3개씩의 뉴런을 가짐을 의미하며 Pc는 인식율을 나타내는데 이 정보는 전술한 바와 같이 출력 인식 부호 벡터의 내용으로 검증된다.

표 2. 각 은닉층과 은닉층 뉴런수에 대한 실험 결과

Table 2. Network performance depending on hidden layer variation.

	은닉뉴런수	평균 패턴 에러	평균 최대 화소차	Pc(%)
은닉층이 1층인 회로망	3	9.1×10^{-2}	5.2	100
	5	2.3×10^{-3}	1.5	100
	10	3.4×10^{-3}	1.9	100
	15	5.3×10^{-3}	2.1	100
은닉층이 2층인 회로망	3-3	1.8×10^{-4}	0.2	100
	5-5	4.3×10^{-4}	0.5	100
	10-10	7.2×10^{-3}	1.8	100
	15-15	8.6×10^{-3}	2.4	100

우리는 먼저 은닉층이 1층인 회로망을 설계하여 영상 복원과 인식 결과를 알아보았다. 표2에서 알 수 있는 바와 같이 한개의 은닉층(hidden layer)을 가지는 회로망의 경우 뉴런 수가 5개인 회로망이 가장 좋은 성능을 나타낸다. 그러나, 평균 최대 화소차가 1.5로써 한 영상당 1화소 이상의 차가 나는 경우가 있다. 따라서 은닉층이 1층인 회로망으로서는 만족할 만한 영상 복원을 할 수 없으므로 보다 나은 영상 복원을 위해 은닉층이 2층인 회로망을 설계하였다. 은닉층이 2층인 회로망은 각 층마다 3개의 뉴런을 가질 경우 가장 좋은 영상 인식 및 복원을 수행할 수 있음이 입증되었다. 표2에서 알 수 있는 바와 같이 은닉층이 2층인 회로망은 은닉층이 1층인 회로망에 비해 상당히 좋은 영상 복원 결과를 보인다.

본 실험에서 사용된 다층 구조 신경 회로망은 학습시 사용자가 지정해 주는 제한 파라메타(parameter)들, 즉 초기 연결강도 값의 범위, 학습율, 관성항의 값에 따라 회로망의 성능이 달라진다. 본 실험에서는 학습시 연결 강도를 -0.07 과 0.07 사이의 값으로 초기화하고 학습율은 0.1 로 관성항은 0.8 로 하여 coprocessor가 장착된 20Mhz 80386machine에서 총 12개의 패턴을 학습시켰다. 대부분의 경우 Xenix 상에서 3시간 이내에 학습이 완료 되었다.

3. 영상의 인식과 복원

본 실험에서는 기준 영상으로부터 구한 정점 좌표값으로 입력을 연상하도록 학습시켰다. 즉 각 영상으로부터 구한 $|F(u, v)|$ 를 신경망의 입력으로 사용하면 학습된 연결 강도값에 의해 출력값으로 정점 좌표값과 인식 부호 벡터를 얻게 된다.

표3에서는 은닉층이 1개일 경우 가장 우수한 성능을 보이는 은닉층의 뉴런수가 5개인 회로망의 실험용 영상에 대한 패턴 에러를 나타내었다. 표 3에서 패턴 에러가 3.0×10^{-3} 이하일 경우 비교적 안정된 영상으로 복원되며, 2.0×10^{-4} 이하로 패턴 에러가 떨어질 경우 영상 복원이 거의 완벽하게 이루어진다.

표3에서 신경 회로망의 입·출력 데이터로 사용한 크기가 120×120 인 영상은 전 기종의 항공기에 대해 회전 변화량에 관계없이 모두 완벽한 영상 복원이 이루어 졌다. 또한 DC10과 B737의 경우 패턴 에러가 적어 거의 완벽한 영상 복원과 인식 기능을 수행함을 알 수 있다. 그러나, MiG21과 Phantom의 경우 회전과 크기 변화에 따라 영상 복원시 상당한 에러가 발생하며, 특히 화소 크기가 80×80 인 영상과 100

표 3. 은닉층이 한층이고 은닉층의 뉴런수가 5개인 회로망의 실험용 영상에 대한 패턴 에러

Table 3. Pattern errors with a hidden layer and five neurons.

Aircraft Type	Rotation Angle	Image Size				
		80×80	100×100	120×120	140×140	160×160
DC10	0°	1.01×10^{-3}	6.86×10^{-4}	1.80×10^{-4}	3.00×10^{-4}	4.62×10^{-4}
	120°	3.21×10^{-3}	5.21×10^{-4}	9.01×10^{-4}	2.52×10^{-4}	2.87×10^{-4}
	240°	6.82×10^{-4}	4.77×10^{-4}	3.85×10^{-4}	4.23×10^{-4}	1.02×10^{-3}
MiG21	0°	1.22×10^{-2}	2.60×10^{-3}	1.91×10^{-3}	2.56×10^{-3}	5.20×10^{-3}
	120°	3.91×10^{-3}	1.18×10^{-3}	3.27×10^{-3}	3.22×10^{-3}	9.86×10^{-3}
	240°	2.53×10^{-3}	2.18×10^{-3}	3.27×10^{-3}	5.56×10^{-3}	5.58×10^{-3}
Phantom	0°	1.54×10^{-3}	5.55×10^{-4}	4.39×10^{-4}	9.04×10^{-4}	4.24×10^{-3}
	120°	1.19×10^{-3}	1.26×10^{-3}	9.06×10^{-4}	1.52×10^{-4}	6.15×10^{-4}
	240°	2.02×10^{-4}	6.12×10^{-3}	6.51×10^{-4}	6.06×10^{-4}	4.01×10^{-4}
B737	0°	2.96×10^{-3}	3.56×10^{-4}	4.85×10^{-4}	5.53×10^{-4}	1.50×10^{-4}
	120°	1.81×10^{-3}	3.52×10^{-4}	1.95×10^{-4}	2.50×10^{-4}	2.09×10^{-4}
	240°	1.35×10^{-4}	1.31×10^{-4}	6.19×10^{-4}	4.25×10^{-4}	4.51×10^{-4}

×100인 영상을 복원시 다소 문제가 발생하였다. 그러나, 영상을 인식한다는 측면에서 볼때 모두 완벽한 영상 인식이 수행됨으로써 인식에는 문제가 없었다.

그림9, 10에서는 MiG21과 Phantom중 패턴 에러가 가장 큰 경우의 항공기 영상과 복원 영상을 나타낸다. 그림 9(b)의 MiG21의 복원 영상을 그림8과 비교해 보면 꼬리날개 부근에서 왜곡 현상이 주로 나타나며 이는 인접한 좌표에서 정점값이 겹쳐 있으므로 인한 좌표간의 상호 간섭에 의한 것으로 추정된다. 그림10(b)의 Phantom은 그림8과 비교할 때 주로 우측 날개 부근에서 왜곡 현상이 발생하나, MiG21에 비하면 안정된 영상으로 복원됨을 알 수 있다. 그 이유는 MiG21과는 달리 vertex가 겹쳐 있지 않아서 좌표간에 별다른 간섭 현상이 발생하지 않기 때문이다.

실제로 영상 복원을 위해서는 회로망에서 출력된 정규화된 정점 좌표값에 기준 영상의 화소 크기인 120을 곱해 실제 항공기 영상의 정점 좌표값을 구해야 하므로 표3, 4의 정규화된 정점 좌표값에 대한 패턴 에러값에서 RMSE(root mean square error)를 구해 실제 정적 좌표값에서의 에러로 환산할 수 있다. 예를 들어 표4에서 크기가 100×100이고, 240° 회전한 MiG21의 경우 패턴 에러는 1.79×10^{-4} 이지만 실제 좌표상의 에러는 $(1.79 \times 10^{-4}/80)^{1/2} \times 120 \approx 0.18$ 로써 각 정점에서의 실제 좌표값의 에러는 한 좌표당 평균 0.18화소차가 난다.

표4와 같이 은닉층을 2층으로 늘리고, 은닉층의 뉴런수를 각 은닉층에 대해 3개로 줄일 경우 전술한

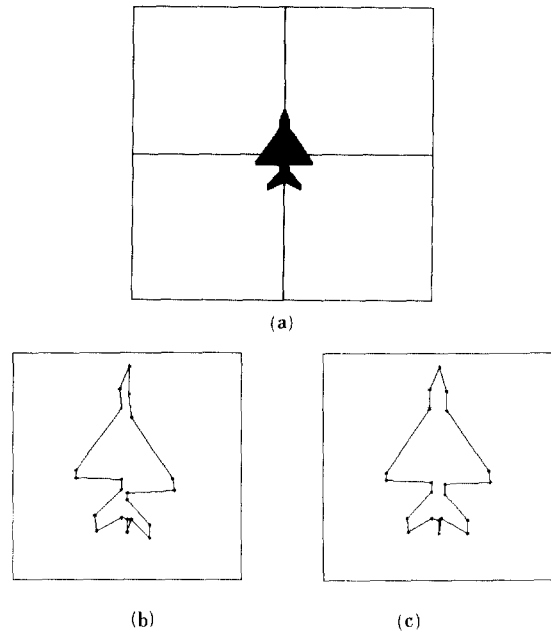


그림 9. 크기가 100×100이고, 0° 회전한 MiG21 영상
 (a) 원 영상
 (b) 은닉층이 한 층이고 은닉층의 뉴런수가 5개인 회로망의 복원 영상
 (c) 은닉층이 두층이고 은닉층의 뉴런수가 3-3인 회로망의 복원 영상

Fig. 9. Image recoveries of MiG21 with scale of 100×100and 0°.
 (a) original image,
 (b) recovered image with a hidden layer and five neurons,
 (c) recovered image with two hidden layers and 3-3 neurons.

MiG21의 패턴 에러가 2.60×10^{-2} 에서 1.32×10^{-3} 으로 감소하면서 그림9, 10(c)와 같이 영상 복원 또한 거의 완벽하게 이루어지며, 이는 그림8의 기준 영상에 대한 vertex 영상과 별다른 차이를 발견할 수 없었다. 따라서, 은닉층이 2층이고 각 은닉층에 대해 뉴런수가 3개일 경우는 패턴 에러가 최대인 경우도 거의 완벽한 vertex 영상으로 복원됨을 알 수 있다.

V. 결 론

본 논문에서는 항공기 영상에 대해 인간의 시각 인식 시스템과 유사한 구조를 가지며 영상의 크기 변화 및 회전을 이동으로 변환시키는 Log-polar sampling과 이동 불변 특성을 갖는 Fourier 변환을 이용하여 영상의 크기 변화와 회전에 무관한 왜곡 불변 특성을 추출하였다. 또한 이를 다시 저주파 성분을 중심으로 샘플링한 값을 입력으로 하여, 영상 복

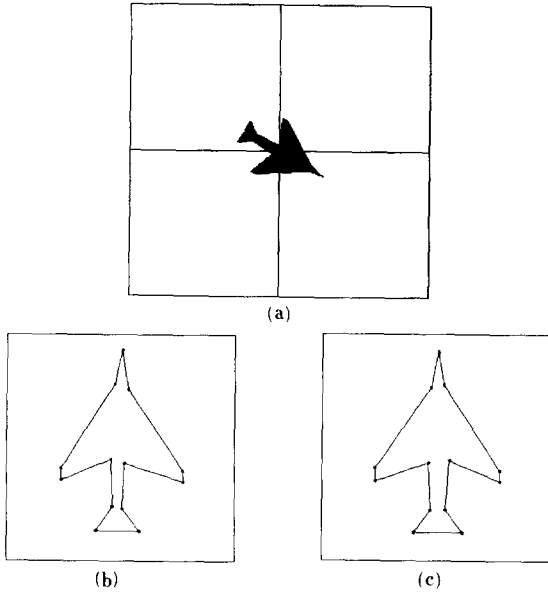


그림 10. 크기가 100×100 이고, 240° 회전한 Phantom 영상

- (a) 원 영상
- (b) 은닉층이 한층이고 은닉층의 뉴런수가 5개인 회로망의 복원 영상
- (c) 은닉층이 두층이고 은닉층의 뉴런수가 3-3인 회로망의 복원 영상

Fig. 10. Image recoveries of Phantom with scale of 100×100 and 240° .

- (a) original image,
- (b) recovered image with a hidden layer and five neurons,
- (c) recovered image with two hidden layers and 3-3 neurons.

표 4. 은닉층이 2층이고 각 은닉층에 대해 은닉층의 뉴런수가 3개인 회로망의 실험용 영상에 대한 패턴 에러

Table 4. Pattern errors with two hidden layers and 3-3 neurons.

Aircraft Type	Rotation Angle	Image Size				
		80×80	100×100	120×120	140×140	160×160
DC10	0°	2.77×10^{-4}	1.01×10^{-4}	1.16×10^{-4}	3.87×10^{-4}	2.90×10^{-4}
	120°	1.10×10^{-4}	1.73×10^{-4}	1.19×10^{-4}	9.80×10^{-5}	9.48×10^{-5}
	240°	3.84×10^{-4}	1.55×10^{-4}	1.05×10^{-4}	1.55×10^{-4}	2.96×10^{-5}
MiG21	0°	7.91×10^{-4}	1.32×10^{-4}	9.61×10^{-5}	1.71×10^{-4}	1.02×10^{-4}
	120°	2.46×10^{-4}	5.11×10^{-4}	1.08×10^{-4}	9.71×10^{-5}	9.84×10^{-5}
	240°	1.37×10^{-3}	1.79×10^{-4}	1.08×10^{-4}	3.55×10^{-4}	9.70×10^{-5}
Phantom	0°	6.02×10^{-4}	6.03×10^{-4}	6.07×10^{-4}	6.02×10^{-4}	6.00×10^{-4}
	120°	7.29×10^{-4}	6.53×10^{-4}	6.04×10^{-4}	6.00×10^{-4}	6.82×10^{-4}
	240°	6.05×10^{-4}	1.73×10^{-4}	5.99×10^{-4}	6.46×10^{-4}	6.06×10^{-4}
B737	0°	1.26×10^{-4}	6.24×10^{-5}	5.46×10^{-5}	5.61×10^{-5}	5.42×10^{-5}
	120°	8.75×10^{-5}	5.88×10^{-5}	5.01×10^{-5}	6.07×10^{-5}	5.20×10^{-5}
	240°	1.05×10^{-4}	5.19×10^{-5}	6.91×10^{-5}	5.01×10^{-5}	5.10×10^{-5}

복원을 위한 기준 영상의 정규화된 정점 좌표값과 영상 인식을 위한 부호 벡터를 출력으로 하는 MLP를 설계하였다. 따라서 기존의 신경 회로망이 주로 패턴 인식 결과만을 출력하는 패턴 인식기로 사용해 왔음에 비해 본 실험에서는 인식 결과와 더불어 정점 좌표축으로 묘사된 기준 영상을 출력함으로써 인식 신뢰도를 시각적으로 측정할 수 있게 하였다.

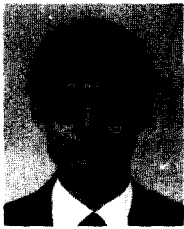
지금까지는 일반적으로 MLP의 입출력으로 디지털화된 값을 사용하였으나, 본 실험에서는 입력과 출력 모두를 아날로그 값으로 연상하는 MLP를 설계하여 영상 인식과 복원을 수행하였다. 본 실험에서 사용된 MLP의 학습 알고리즘으로는 backpropagation 알고리즘을 사용하였으며, 시뮬레이션 결과, 영상의 크기 변화와 회전에 관계없이 MLP를 이용해 안정된 영상 복원과 인식 기능을 수행할 수 있었다. 또한, 은닉층의 수와 은닉층의 뉴런 수를 최적화함으로써 보다 향상된 기능을 발휘하는 MLP를 설계할 수 있었다.

參 考 文 獻

- [1] Darpa, *Darpa Neural Network Study*, AFCEA International Press, 1987.
- [2] M. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. on Information Theory*, vol. IT-8, pp. 179-187, 1962.
- [3] T.P. Wallace and P.A. Wintz, "An efficient three-dimensional aircraft recognition algorithm using normalized fourier descriptors," *Computer Vision, Graphics, and Image Processing*, vol. 13, pp. 99-126, 1980.
- [4] D. Casasent and S.I. Chien, "Efficient (ϕ, s) string code symbolic feature generation," *Optics Communications*, vol. 67, no. 2, pp. 103-106, 1988.
- [5] R.A. Messner and H.H. Szu, "An image processing architecture for real time generation of scale and rotation invariant patterns," *Computer Vision, Graphics, and Image Processing*, vol. 31, pp. 50-66, 1985.
- [6] J.G. Bailey and R.A. Messner, "Docking target design and spacecraft tracking system stability," *SPIE Proc.* vol. 1192, Philadelphia, pp. 820-831, Nov. 1989.
- [7] C.F.R. Weiman and R.D. Juday "Tracking algorithms using log-polar mapped image coordinates," *SPIE Proc.*, vol. 1192, Philadelphia, pp. 843-853, Nov. 1989.

- [8] H. Wechsler and G.L. Zimmerman, "2-D Invariant object recognition using distributed associative memory," *IEEE Trans. on Patt., Anal., and Mach. Intell.*, vol. PAMI-10, pp. 811-821, 1988.
- [9] D.E. Rumelhart, G.E. Hinton, and R.J. Williams, "Learning internal representation by error propagation," *Parallel Distributed Processing*, vol. 1, D.E. Rumelhart and J.L. McClelland (Eds.), Cambridge, MA: MIT Press, pp. 318-32, 1987.
- [10] D. Casasent and S. Liebowitz, "Model-based Knowledge-based Optical Processors," *Applied Optics*, vol. 26, no. 10, 1987.
- [11] S.E. Fahlman, "Faster-learning variations on backpropagation: an empirical study," *Proc. Connectionist Models Summer School*, Carnegie Mellon University, pp. 38-51, 1988.

 著 者 紹 介



金 鐘 吾 (準會員)

1967年 12月 13日生. 1990年 경북대학교 전자공학과 졸업. 1990 年 3月~현재 경북대학교 대학원 전자공학과 재학중임. 주관심 분야는 신경회로망, 패턴인식 등 임.

秦 成 一 (正會員) 第27卷 第11號 參照

현재 경북대학교 전자공학과 조교수

金 麟 哲 (正會員) 第28卷 B編 第6號 參照

현재 CAS(주) 연구소 연구원