

퍼셉트론을 이용한 Neuro-FAX 방식에 관한 연구

正會員 金 海 洙* 正會員 李 根 泳**

A Study on the Neuro-FAX algorithm Using the Perceptron Network

Hae Soo Kim*, Keun Young Lee** *Regular Members*

요 약

본 논문에서는 보다 효율적인 전송과 깨끗한 화질로 복원 문서를 구성 할 수 있도록 신경회로망을 이용한 Neuro-FAX 방식을 제안하였다. 이 방식은 문자부분과 도형부분을 따로 처리하는 것으로 문자 부분에 대해서는 각 문자에 신경회로망을 적용시켜 문자를 인식한 후 인식된 정보만을 전송하고, 도형 부분은 기존의 방식으로 그대로 전송하는 방식으로 전송 효율도 높이고, 짧은 전송시간으로 인하여, 전송선로의 혼잡을 줄일 수 있고, 수신측에서 각 문자 패턴을 가지고 있게 하여 복원된 문서를 기존의 팩시밀리 문서보다 깨끗하게 처리하여 고품질을 얻을 수 있도록 하였다.

ABSTRACT

In this paper, we proposed a Neuro-FAX algorithm having high compression rate and good reconstruction capability in spite of noise and fonts. This algorithm processes the character part and the image part separately. In the character part, we recognized each characters in document using neural networks, and transmitted the information recognized. And we transmitted the image part as it is by the conventional method. With character set in receiving terminal, it can produce nice document of noise free characters and different font.

I. 서 론

CCITT에서는 팩시밀리상의 데이터를 보다 효율적으로 전송할 수 있도록 최적의 방식을 계속 권고하고 있으며 현재의 권고안인 Group 4는 2차원 부호화를 이용하여 고해상도로 이루어진 A4 용지 (7.7 line/mm와 line당 2048화소)를 1분내외에 전송이

가능하도록 권고하고 있으나 보다 효율적으로 문자를 전송하기 위하여 문자 부분과 도형 부분을 따로 전송하는 방식 역시 권고안을 통해 제시하고 있다^{[1][2][3]}.

본 논문에서는 CCITT에서 권고하고 있는 문자와 도형을 따로 전송하는 방식에 부합되면서 압축율과 복원문서의 화질도 개선될 수 있는 방식을 제안한다. 즉, 문자 부분은 각 문자를 추출하여 가능한한 모든 문자를 인식하고 인식된 문자의 정보만을 전송하여 전송 효율도 높이고 수신단에서는 보유하고 있는 문

*安養專門大學 電子通信科
Dept. of Electronic Communications Anyang Jr. College
**成均館大學校 電子工學科
Dept. of Electronic Engineering Sung Kyun Kwan Univ.
論文番號 : 93-2 (接受1992. 4. 29)

자 패턴을 이용하여 문서를 복원함으로써 문서의 디지털 발생하는 에러의 영향을 최소화시키면서 희망하는 글자체로 복원이 가능하여 고화질의 문서를 얻을 수 있도록 하였으며, 인식을 위해서 사용된 방식은 인간의 두뇌가 처리하는 방식을 모방하여 높은 인식력과 빠른 처리 능력등의 장점을 지닌 신경회로망을 이용하였으며 신경회로망의 많은 모델중 퍼셉트론이 많은 패턴의 인식에 가장 유리하여 이를 이용하였다. 도형 부분은 팩시밀리의 특성상 그대로를 복원하여야 하므로 기존의 방식인 MMR 부호화를 적용하여 전송하였다. 그리고 훼손이나 학습의 부족으로 인식이 불가능한 문자에 대해서도 도형으로 처리하여 복원시 발생할 수 있는 문제점을 제거하였다. 이상과 같이 도형과 문자 부분을 분리하여 처리하면서 문자 부분을 신경회로망을 이용하여 인식하여 처리하는 방식을 Neuro-FAX라고 부르기로 한다.

Neuro-FAX 방식을 문서에 적용해 본 결과 순수하게 문자만으로 구성된 문서에 대해서는 기존의 팩시밀리 전송 방식에 비해 약 10:1 정도의 높은 압축율을, 도형이 80%이상 포함된 문서에 대해서도 3:1이상의 압축율을 얻을 수 있으면서 기존의 팩시밀리 문서보다 깨끗하게 처리하여 고화질을 얻을 수 있고, 짧은 전송시간으로 인하여 전송선로의 혼잡을 줄일 수 있고, 전화비등의 경제적인 면까지 해결할 수 있는 장점을 가지며, 데이터베이스 구성시에도 문자정보만으로 구성하므로 보다 효율적인 운영이 가능하다.

그리고 팩시밀리에서의 전송을 최소로 하기 위해서는 문자뿐 아니라 숫자, 특수문자 모두를 인식하여야 최적의 부호화를 이룰 수 있으므로 본 논문에서는

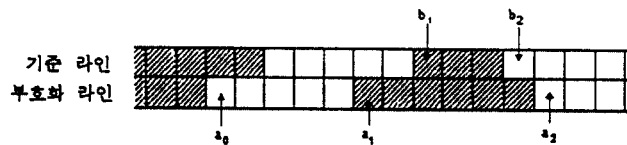
ASCII 부호에 포함된 대부분의 문자, 즉 키보드에서 발생할 수 있는 모든 표준문자를 인식하도록 하여 가능한 모든 문자를 인식할 수 있는 Neuro-FAX방식이 되도록 중점을 두었으며, 인식이 불가능한 훼손된 문자와 도형등에 대해서도 별도로 전송하는 방식을 취하여 완전한 화상복원을 이루도록 하였다.

II. 기존의 FAX 방식

1. 이 장에서는 CCITT에서 권고하여 현재 사용하고 있는 디지털 전송방식으로 높은 압축율과 효율을 보이고 있는 Group 3, Group 4 방식인 개선된 Huffman 부호화와 개선된 READ 부호화 방식, 그리고 MMR 부호화에 대한 전송방식에 대해 알아본다.

1. 개선된 Huffman 부호화(MHC: Modified Huffman Code)^{1) | 2) | 4)}

개선된 Huffman 부호화는 화상 데이터에 있어서는 0 데이터가 한번 발생하면 다음 데이터 역시 0이 발생할 확률이 크므로 이를 이용한 연속길이 부호화와 빈도가 낮은 데이터로 인하여 전체 부호길이가 길어지는 단점을 없애고져, 발생빈도에 따라서 부호를 설정하는 방식인 Huffman 부호화를 결합한 형태로 먼저 연속길이 부호화 방식을 이용하여 데이터의 연속길이를 구한 후 이렇게 구한 연속길이들을 Huffman 부호화하는 1차원 부호화 방식으로 CCITT에서는 팩시밀리에서 발생하는 데이터들의 발생확률을 모두 구하여 가장 타당성있는 Huffman 부호표를 제시하여 이의 사용을 권고하고 있다. 이 부호표는 실험을 통하여서도 효율성과 타당성을 확



- a_0 : 현재 줄의 출발점이나 기준점
- a_1 : a_0 오른쪽의 변화점, 즉 데이터가 a_0 와 다른 점
- a_2 : a_1 오른쪽의 변화점, 즉 데이터가 a_1 과 다른 점
- b_1 : 이전 줄에서 a_0 의 오른쪽에서 a_0 와 다른 데이터를 가진 점
- b_2 : b_1 의 오른쪽의 변화점, 즉 데이터가 b_1 과 다른 점

그림 1. 화소의 변화
Fig. 1. Changing picture elements

인할 수 있었다.

2. 개선된 READ 부호화(MRC : Modified Relative Element Address Designate Code)^{[1][2][4]}

개선된 READ 부호화는 2차원 부호화로 이전의 줄을 기준으로 하여 현재 줄과의 차이만을 부호화하는 방식으로 1차원 부호화 방식보다 훨씬 높은 압축율을 보이고 있다.

부호화 방식은 화상 데이터의 특성상 현재 줄은 이전 줄과의 차이가 작기 때문에 이전 줄을 기준으로 하여 차이나는 정도만을 전송하는 방식을 취하고 있으며, 이때 나타나는 현상은 크게 3가지로 구분될 수 있다.

첫번째로 통과방식(pass mode)로 이전의 줄이 현재 줄에서는 나타나지 않는 형태로 이전 줄에서 데이터가 끝난 형태이며, 그림 1에서 b_2 가 a_1 의 오른쪽에 나타난다.

두번째로 수직방식(vertical mode)으로 이전 줄과 현재 줄에서의 변화가 적은 형태로 차이값만큼만 전송하며 보통 3이하의 변화가 이에 포함되며, a_1 과 b_1 의 차에 따라 $V(0)$, $V_R(1)$, $V_R(2)$, $V_R(3)$, $V_L(1)$, $V_L(2)$, $V_L(3)$ 로 구분되며, 여기서 첨자 R, L은 a_1 이 b_1 의 오른쪽에 존재하면 R, 왼쪽에서 존재하면 L로 표시한다.

세번째는 수평방식(horizontal mode)으로 이전 줄과 현재 줄의 변화가 큰 경우로 이때는 이전 줄과의 차이를 전송하는 것보다는 현재 줄의 변화량을 그대로 전송하는 것이 효율적이므로 수평방식이라고 불리는 부호코드 M과 a_0a_1 의 거리와 a_1a_2 의 거리를 전달하는 방식이다.

이상의 3가지 방식에 따라 각 줄을 전송하며, 중간에서 발생할 수 있는 데이터의 손실을 위하여 지정된 줄마다 1차원 부호화를 취하여 데이터의 손실이나 에러로 인한 영향이 지정된 줄 이내로 줄이는 방식을 취한다. 팩시밀리상의 이진 데이터는 특성상 수직방식이 가장 많이 발생하므로 전체적으로 상당히 효율적인 방식이다.

Group 3 권고안은 개선된 Huffman 부호화와 개선된 READ 부호화를 결합하여 사용하는 형태로 중간에서 발생할 수 있는 에러와 데이터 훼손의 영향을 줄이기 위하여 지정된 줄 간격마다 1차원 부호화인 개선된 Huffman 부호화를 취하고 나머지에 대해서는 2차원 부호화인 개선된 READ 부호화를 취하는 방식으로 발생한 에러의 영향이 지정된 줄 간격이내

에 한정되도록 하고 있다. 보통 이용하고 있는 줄의 간격은 3줄이다.

3. MMR 부호화(Modified Modified READ 부호화)^{[2][4]}

MMR 부호화 방식은 Group 4에서 권고안으로 규정하고 있으며, Group 3에서 규정한 팩시밀리 부호화 방식과 원리적으로 동일한 부호화 방식이지만, Group 3에서는 1차원 부호화와 2차원 부호화가 결합된 형태인 반면 Group 4의 MMR 부호화 방식은 순수한 2차원 부호화 방식으로 Group 3의 권고안 방식보다 효율적이다. 부호화 방식은 부호화할 문서의 최초의 줄은 가상적인 백으로 구성된 줄이라고 가정하여 1차원 부호화를 제거하였다. 그리고 Group 3의 권고안에서는 지정된 줄, 보통 3줄마다 1차원 부호화 방식을 취한 반면 Group 4에서는 모두 2차원 부호화 방식인 READ 부호화만을 취한다는 점과 매줄이 끝날 때마다 줄바꿈 정보를 전송하지 않는다는 점으로 줄바꿈 정보뒤에 1차원 부호화인지 2차원 부호화인지를 알리는 비트정보 또한 전송하지 않는다는 점이다. 이런 점으로 인하여 Group 3의 권고안에 비해 전송 정보량과 전송시간면에서 많은 장점을 가지고 있다.

전송에 사용하는 통과방식과 수직방식에 사용하는 부호표는 표 1의 2차원 READ 부호표와 수평방식에

표 1. 개선된 READ 부호화표

Table 1. Modified READ code table

방식	부호화될 요소		표시법	부 호
통과	b_1, b_2		P	0001
수평	a_0a_1, a_1a_2		H	$001+M(a_0a_1)$ $+M(a_1a_2)$
수직	b_1 바로 밑 a_1	$a_1b_1=0$	$V(0)$	1
	b_1 오른쪽 a_1	$a_1b_1=1$	$V_R(1)$	011
		$a_1b_1=2$	$V_R(2)$	000011
		$a_1b_1=3$	$V_R(3)$	0000011
	b_1 왼쪽 a_1	$a_1b_1=1$	$V_L(1)$	010
		$a_1b_1=2$	$V_L(2)$	000010
$a_1b_1=3$		$V_L(3)$	0000010	
줄바꿈(end of line)				00000000001
다음줄 1차원 부호화				EOL+1
다음줄 2차원 부호화				EOL+0

사용하는 부호표는 개선된 Huffman 부호표와 동일하다.

Ⅲ. 신경회로망 [6][7][10][11].

복잡한 구조를 가지는 인간의 두뇌가 처리하는 방식을 부분적으로 모방한 신경회로망은 아직 인간의 능력을 완전하게 모방하지는 못하지만, 병렬로 구성이 되어 정보를 기억하는 방식, 정보를 처리하는 방식등에서 많은 장점을 보여주고 있으며, 현재 많은 분야에서 활용되고 있다.

첫째, 연상기억기능으로 일부 훼손된 입력패턴으로 부터 원래의 정보를 찾아낼 수 있다. 둘째, 일반화기능으로 학습된 여러 패턴으로 부터 일반적으로 나타나는 특징을 추출하여 인식할 수 있다. 셋째, 최근접 패턴 인출기능으로 학습시킨 패턴과 대응하는 패턴이 존재하지 않을 때 가장 비슷한 패턴을 찾아낼 수 있다. 넷째, 결합허용기능으로 회로망의 일부가 훼손되었다하더라도 전체 시스템의 성능이 급격히 저하되지 않고 완만하게 저하된다는 특징을 가지고 있다.

신경회로망은 적용분야에 따라 적합한 모델이 많이 개발되어 있으며 본 논문에서는 단층 퍼셉트론 모델을 이용하였다. 단층 퍼셉트론의 기능으로 XOR와 같은 논리함수를 처리할 수 없어 침체기를 초래하기도 했지만 패턴분류에 있어서는 만족한 결과를 보이고 있어 이를 이용하였으며, 연결강도와 임계치는 고정되거나 여러가지 알고리즘을 사용하여 변경할 수 있는데 최초의 학습 알고리즘은 페셉트론 수렴정리 (Perceptron convergence theorem)로 다음과 같이 나타낼 수 있다.

$$W_i(t+1) = W_i(t) + \eta [d(t) - y(t)] x_i(t)$$

여기서 $d(t)$, $y(t)$ 는 각각 기대되는 출력과 실제 출력을 의미하며 η 는 적응율을 제어하는 변수이다. 페셉트론의 결과가 정확하다면 $[d(t) - y(t)]$ 는 0의 값을 가지며 이때 연결강도의 변화는 없다. Rosenblatt는 두 유형이 선형적으로 분리가능하면, 즉 두 패턴 유형이 하나의 초평면의 양쪽에 놓이게 된다면, 그 알고리즘은 수렴하고 이것에 의해 구해진 초평면은 패턴의 유형을 정확하게 분류할 수 있음을 증명하였다.

Ⅳ. 제안한 Neuro-FAX 방식

본 장에서는 신경회로망을 이용하여 기존의 방식처럼 전송하여야 할 문서를 그대로 전송하는 것이 아니라 문서속에 포함된 각각의 문자들은 인식하여 인식된 문자의 부호만을 전송하고, 도형 부분은 기존의 방식을 이용하여 전송함으로써 전송효율도 높이고 복원화상도 깨끗한 Neuro-FAX 방식을 제안한다.

1. 문자 추출 방식

문자의 추출과정은 먼저 이진영상을 입력받은 후 입력된 화상 데이터에서 각각의 문자를 분리시키는 과정이다.

1-1. 이진영상 입력

이진영상의 입력은 CCD-TV 카메라를 이용하는 방법과 image scanner를 이용하는 두가지 방식이 있는데 CCD-TV 카메라를 이용하는 방법은 짧은 시간에 데이터를 추출할 수 있고 자연화상, 즉 화소당 밝기가 여러 단계로 구분되는 화상데이터의 입력에는 효율적인 반면 이진화상의 입력에 대해서는 깨끗한 화상의 입력이 어렵다. 이는 주위의 밝기와 대상체의 위치등의 주위 환경에 많은 영향을 받기 때문이다. 이에 반해 image scanner를 이용하는 방법은 데이터 입력시 시간이 소용되는 반면 깨끗한 화상을 얻을 수 있고, 주위의 밝기나 입력 대상체의 위치등 주위 환경에 전혀 영향을 받지 않는 장점을 가지고 있다. 본 연구에서는 이진영상에 중점을 두고 있고, 기존의 팩시밀리와 같은 방식인 줄단위의 화상 입력 방식에 중점을 두기 위하여 image scanner를 이용하여 입력 데이터를 취득하였다.

1-2. 문자 추출

신경회로망에 각 문자들에 대한 학습과 인식을 시키기 위하여 입력받은 화상데이터에서 각각의 문자를 분리시켜야 한다. 분리 과정으로 입력 문서에서 먼저 발생할 수 있는 잡음들을 제거한 후 줄과 줄사이를 구분하였는데 이는 전체 한 열의 데이터가 모두 0이면 이 열은 줄과 줄의 사이라고 간주하여 이를 구분의 기준으로 하였다. 이런 방식을 이용하여 각 줄의 분리가 가능하였으며 이런 과정을 거쳐 얻은 분리된 각 줄에 대하여 이번에는 행을 검사하여 한 행의 전체가 0이면 문자와 사이로 간주하여 이를 문자와 문자의 경계로 하여 분리하였다.

이런 과정에서 한 문자가 둘로 나누어지는 분할 오

차와 두 문자가 하나로 연결되는 병합 오차, 두가지 형태의 오차가 발생할 수 있다. 이의 해결책으로 병합 오차는 소문자 j가 세로로 28크기로 가장 큰 문자이므로 이보다 크게 분리된 문자는 두 문자가 붙어서 입력된 것으로 간주하여 문자의 가운데를 기준으로 하여 가장 적은 흑 화소로 구성된 열을 나누어 분리 추출하였다. 한 문자가 둘로 나뉘어 추출되는 분할 오차의 경우는 추출된 문자의 크기가 최소크기 이하인 경우는 한 문자가 나누어진 경우로 판단하여 가장 가깝게 인접해있는 문자에 결합하여 처리하였다. 이런 과정을 거쳐 28×28 크기로 각 문자를 추출하였다. 또한 줄과 줄 사이가 30이상의 크기인 경우에는 도형 정보로 간주하였으며, 도형정보속에 포함되어 있을 수도 있는 문자를 추출하기 위하여 도형정보에 대해서는 30×30 크기의 마스크를 적용하여 문자를 분리하였다. 200dpi로 문서를 입력시 세 문자이상 이 연결되어 발생하는 경우는 발생하지 않았으나 여러 문자가 함께 연결되어 발생하는 오차는 Lee 교수의 제안방법^[14]의 방식을 참조하였으며, 분리가 불가능한 경우에는 도형문자로 처리하므로 알고리즘상 큰 문제는 발생하지 않는다.

2. 학습과정

분리 추출한 각 문자를 신경회로망에 학습시킨 후 학습된 회로망을 통해 인식이 가능하며, 여기서 학습이라함은 신경회로망의 각 연결강도를 스스로 결정하는 것으로, 방식으로는 모든 문자를 학습시키는 것이 가장 인식율도 높고 안정되겠지만, 학습시간과 신경회로망의 복잡도를 고려하여 가장 대표되는 문자종류와 그의 발생할 수 있는 예외적인 몇몇 문자만 학습시켜 학습시간과 회로망 자체의 복잡도를 낮추었으며, 학습에 사용된 문자도 그대로 학습시간과 회로망 자체의 복잡도를 낮추었으며, 학습에 사용된 문자도 그대로 학습시키면 28×28의 크기이므로 784개의 입력을 가져야 하기 때문에 좀더 효율적으로 처리하기 위하여 음영 부호화 방식을 이용하여 학습하고 학습된 신경회로망을 이용하여 인식하였다. 학습에서는 ASCII 부호중 키보드에서 발생하는 모든 영문자, 숫자 및 특수기호를 대상으로 하였다.

음영 부호화 방법은 Burr에 의해 제안되었으며^[16], 많은 활용을 보이고 있는 특징점 추출 방법^[13]에 비해 전처리과정없이 바로 적용이 가능한 장점을 가지고 있으며 입력량의 많은 감축을 취할 수 있는 점이 부각된 좋은 방법이나 입력항들을 실수로 처리하여

보다 정확한 결과를 위하여 16개정도의 많은 바 마스크를 사용하였기 때문에 학습시 오랜시간과 오차를 줄이기 위해서는 대형 컴퓨터가 동원되어야 한다는 문제점을 가지고 있었으나, 본 연구에서는 실수가 아닌 정수를 취하여 이런 문제를 해결함으로써 인하여 바 마스크의 수도 8개만 가지고 데이터의 특성을 분류할 수 있었고, 정수값을 이용하였기 때문에 입력항면에 있어서는 큰 효과를 얻지는 않았으나 학습에 소요되는 시간과 오차에 대해서도 훨씬 여유를 두면서 전처리 과정을 없애는 효과를 얻을 수 있었으며, 대형컴퓨터인 경우에는 빠른 속도로 인하여 입력항을 실수로 처리 가능하여 입력항면에서도 많은 효과를 얻을 수 있으리라 기대된다. 또한 음영 부호화로 얻은 결과를 퍼셉트론에 적용함으로써 인하여 새로운 문자항이 추가되더라도 쉽게 대응할 수 있고, 연결강도도 정수로 처리하여 학습시간의 단축과 하드웨어 구성상의 용이점을 얻을 수 있었다.

음영 부호화 처리방식은 간단하여 설치된 바 마스크에 비추어지는 그림자에 의해 그림자가 있으면 1, 없으면 0이 되는 방법으로 바 마스크를 몇개 구성하느냐에 따라 4-바, 6-바, 8-바 마스크로 구분된다. 그림 2에 각 바 마스크의 형태를 나타내었다.

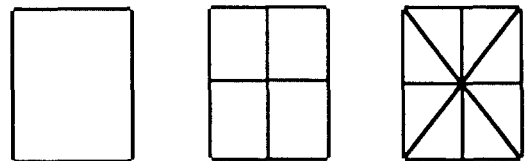


그림 2 바 마스크
Fig. 2. Bar Mask

이상의 과정을 통해 얻은 결과를 이용하여 신경회로망의 연결강도를 조절하였으며, 학습 방법은 지도 학습으로 입력 데이터로는 바 마스크를 통해 얻은 결과가 이용되며 입력 데이터와 원하는 출력값이 함께 학습쌍(training pair)으로 이용되었다.

패턴 분류를 위해 사용되는 모델은 여러가지가 있으나 다른 모델들은 패턴의 수가 적은 경우에는 뛰어난 성능을 나타내고 있으나 본 논문의 경우에는 98개의 패턴을 분류해야하므로 이렇게 많은 경우에는 분류가 불가능하였다. 예로 가장 효율적이라고 알려진 Back Propagation Network의 경우에도 한계를 가지고 있어 급기야는 학습에 사용된 패턴의 분류도 불

가능하였다. 이는 다층 구조로 인하여 하나의 연결강도가 여러 패턴의 결과에 영향을 미치기 때문이다.

단층 퍼셉트론에서는 많은 연결강도중에서 원하는 하나의 결과항과의 연결강도만이 영향을 미치기 때문에 이 연결강도만 조절하면 되므로 다른 신경회로망에 비해 학습이 비교적 간단하며 본 논문에서는 1, -1, 0 세 가지 정수값만 연결강도로 이용하였기 때문에 학습이 간단하고, 새로운 학습쌍에 대해서도 전체 연결강도가 영향을 받는 것이 아니라 해당하는 부분의 연결강도만 영향을 받기 때문에 학습시간이나 방식이 간단하다.

3. 인 식

이상의 방법과 같이 퍼셉트론 신경회로망의 각 연결강도를 조절하여 학습시킨 후 인식을 실시하였으며, 여기서 인식 과정은 입력 데이터가 입력되었을 때 각각의 입력항이 해당되는 연결강도에 곱하여지고 각 곱해진 결과가 출력층에서 임계치와 비교하여 결과가 임계치보다 크면 그 결과로 인식한 것이다.

$$y(t) = \sum W_i x(i)$$

으로 나타나낼 수 있다.

이와 같은 학습을 거친 신경회로망에 대해 인식을 실시하면 학습시킨 문자는 거의 100% 인식이 가능하며, 학습시키지 않은 입력문자에 대해서도 95% 이상 인식이 가능하였다. 그리고 원래의 퍼셉트론 회로망은 연결강도가 실수값으로 처리되기 때문에 학습시간이 많이 걸리는 문제가 있었으나 본 방식은 정수값, 그중에서도 -1, 0, 1 세가지 값만 가지고 연결강도를 조절하였기 때문에 학습시간도 줄일 수 있고, 회로망 자체도 간단하며, 집적화에도 유리하다. 그리고 팩시밀리의 부호화 데이터를 최소로 줄이기 위해서는 가능한 문자인식이 완벽하여야 하며, 이를 위하여 ASCII 부호에서 나타나는 거의 대부분의 문자, 즉 키-보드에서 발생하는 모든 문자를 학습에 이용하였으며, 인식되지 못한 문자에 대해서는 계속 학습시켜 100% 인식을 목표로 하였으며, 문자의 훼손으로 인하여 인식이 불가능한 경우에 대해서는 도형정보와 함께 처리하였다.

4. 위치정보의 전송방식

이상의 과정을 거쳐 각 문자는 인식이 이루어지며,

인식된 정보는 가장 최소의 정보로 나타낼 수 있도록 아스키(ASCII) 문자(8비트 단위)로 부호화되며, 복원하기 위해서는 인식된 문자를 어느 위치에 복구해야 할 지를 결정하여야 한다. 이를 위하여 먼저 문자를 추출할 때 추출한 위치를 기억하고 있어야 하며, 문서를 복원하기 전에 각 문자의 위치를 먼저 알아야 정확한 복원화상을 얻을 수 있다.

일반적인 A4용지를 팩시밀리로 전송하기 위하여 디지털화하면 고해상도인 경우 가로로 약 2432화소, 세로로 2233줄 정도의 크기이므로 이를 모두 나타내기 위해서는 가로, 세로 모두 12비트씩을 할당하여야 정확한 위치를 나타낼 수 있다. 그러므로 인식이 이루어지면 1문자당 8비트로 부호화하여 문자 모두를 나타낼 수 있는데 비해, 위치정보는 각 문자당 가로, 세로에 12비트씩 24비트가 필요하여 전송데이터의 75%를 위치 정보가 차지하기 때문에 이의 효율적인 부호화가 반드시 필요하다. 이의 해결 방안으로 CSM에서의 방식^[15], PMS에서의 방식^[15] 등 많은 연구가 이루어지고 있으나 본 논문에서는 인쇄체 문자로 구성된 화상 데이터의 경우에는 문자가 워드 프로세서로 작성되었던 지, 타자기로 작성되었던 지 상관없이 문자의 크기는 거의 동일하기 때문에 한 문자가 발생하면 다음 문자의 위치는 예측가능하므로 이를 이용하여 보다 효율적으로 부호화하기 위하여 두 가지 방안을 제시하였다.

4-1. 발생확률에 의한 방법

발생하는 수평, 수직값의 차이에 대한 모든 정보를 취득한 후 이 정보에서 다음에 발생할 문자를 예측한 후 예측오차에 대한 확률을 구하고 구한 확률을 이용하여 Huffman 부호화를 적용하는 방법으로, 이 방법은 전송해야 하는 데이터 수는 많은 압축을 이룰 수 있으나 발생확률이 모든 문서에서 동일하지 않으므로 일반성의 문제가 있으나 문서상 큰 변화를 보이지는 않으므로 적용에 큰 문제는 없다고 생각된다.

4-2. 부호화에 의한 방법

일반성을 고려한 방법으로 먼저 각 줄의 첫 문자의 위치는 그대로 전송한다. 이때 수직좌표에는 10비트, 수평좌표에는 7비트를 할당하며 여기서 수평에 적은 비트를 할당한 것은 문서상의 특성에 의한 것이다.

다음 문자의 수직성분은 인쇄체인 경우 거의 일정한 간격을 보이므로 이를 고려하여 미리 예상하고 있는 값과의 차이분만 전송한다. 인쇄체 문서의 특성상

기의 대부분이 3이하의 예측오차내에 포함되어 예측 오차 3까지는 표 2의 부호표를 이용하고, 차이값이 3 이상이면 그 이상이라는 부호 111 다음에 부호 1 비트(+이면 1, -이면 0), 다음에 크기를 전송하며 크기는 5비트를 할당한다.

수평좌표 역시 인쇄체 문서의 특성상 비슷한 특성을 보이므로 표 2의 할당표에 의해 전송하며, 3이상 일 때는 7 비트를 할당한다. 그리고 수평좌표가 감소할 때는 다음 줄로의 줄바꿈을 의미하므로 줄바꿈 부호 111100000(9비트)를 전송하는 방식으로 위치정보를 전달한다.

이 방법은 문서 데이터의 전체적인 통계에 의한 방법이므로 일반성을 가지고 있다는 점이 부각되며, 앞 방법에 비해서 압축면에서는 조금 미흡하다.

표 2. 위치정보부호 할당표
Table 2. Code Table for Position Information

차이값	부호
0	000
-1	001
1	010
-2	011
2	100
-3	101
3	110
크이상	111

5. 도형부분의 처리

팩시밀리에서 사용하는 문서들에는 문자뿐 아니라 싸인이나 기호, 그림등의 도형이 포함될 수 있으나 도형을 모두 인식한다는 것은 현재의 인식능력으로는 역부족이며 또 인식을 한다고 하더라도 이를 모두 패턴으로 처리하는 것은 팩시밀리의 장점을 부각시키지 못하기 때문에 Neuro-FAX 방식에서는 도형부분에 대해서는 기존의 방식대로 부호화하여 팩시밀리의 장점을 그대로 지니도록 하였다.

본 연구에서는 28×28 크기를 넘어선 형태와 인식이 불가능한 문자에 대해서는 문자가 아니라 도형으로 인식하여 처리하였으며, 도형의 처리는 도형의 다양한 형태를 고려하여 2진 데이터를 전송하는 데 있어 가장 효율적으로 알려져 있고, 실제 적용되고 있는 Group 4 권고안인 MMR 부호화를 적용하여 처

리하였다. 이렇게 처리함으로 인하여 도형부분에 대한 팩시밀리의 장점을 활용하면서 인식이 불가능한 문자도 도형부분과 같이 처리하여 해결하였다.

V. 실험 및 고찰

본 논문의 타당성과 효율성을 입증하기 위하여 실험에 사용한 문서로 그림 3은 순수하게 문자만으로 구성된 형태이고 그림 4은 도형이 포함되어 있는 문서이고, 그림 5는 CCITT에서 제시한 표준 이진 문서로 축소하여 제시하였다. 그리고 그림 3과 그림 4의 문서는 명조체로 워드 프로세서로 작성된 문서이고, 그림 5의 문서는 타자기에 작성된 문서이다.

Drive carefully and slowly when pedestrians are about, particularly in crowded shopping streets, when you see a bus stopped, or near a parked mobile shop. Watch out for pedestrians coming from behind parked or stopped vehicles, or from other places where you might not be able to see them.

Three out of four pedestrians killed or seriously injured are either under fifteen or over sixty. The young and the elderly may not judge speeds very well, and may step into the road when you do not expect them. Give them, and infirm, or blind, or disabled people, plenty of time to cross the road.

Drive slowly near schools, and look out for children getting on or off school buses. Stop when signalled to do so by a school crossing patrol showing a "Stop Children" sign.

Be careful near a parked ice-cream van-children are more interested in ice-cream than in traffic.

When coming to a zebra crossing, be ready to slow down or stop to let people cross. You must give way once they have stepped on to a crossing. Signal to other drivers that you mean to slow down or stop. Give yourself more time to slow down or stop on wet or icy roads.

Never overtake just before a zebra crossing.

From The Highway Code

그림 3. 실험 문서 1

Fig. 3. Test Document 1

이들 문서들을 Image scanner를 이용하여 200DPI로 입력받았으며, 먼저 포함되어 있는 잡음이나 훼손을 제거하였다. 정보의 최소단위는 영문자 i, j의 소문자 꼭지점이 최소의 경우 4개의 화소로 구성되어 있으므로 2개의 화소이하로 구성된 1의 집단은 잡음으로 간주하여 제거하였고, 2개의 화소로 구성되어



THE SLEREXE COMPANY LIMITED

SAPORS LANE - BOOLE - DORSET - BH 25 8 ER

TELEPHONE BOOLE (945 13) 51617 - TELEX 123456

Our Ref. 350/PJC/EAC

18th January, 1972.

Dr. P.N. Cundall,
Mining Surveys Ltd.,
Holroyd Road,
Reading,
Berks.

Dear Pete,

Permit me to introduce you to the facility of facsimile transmission.

In facsimile a photocell is caused to perform a raster scan over the subject copy. The variations of print density on the document cause the photocell to generate an analogous electrical video signal. This signal is used to modulate a carrier, which is transmitted to a remote destination over a radio or cable communications link.

At the remote terminal, demodulation reconstructs the video signal, which is used to modulate the density of print produced by a printing device. This device is scanning in a raster scan synchronised with that at the transmitting terminal. As a result, a facsimile copy of the subject document is produced.

Probably you have uses for this facility in your organisation.

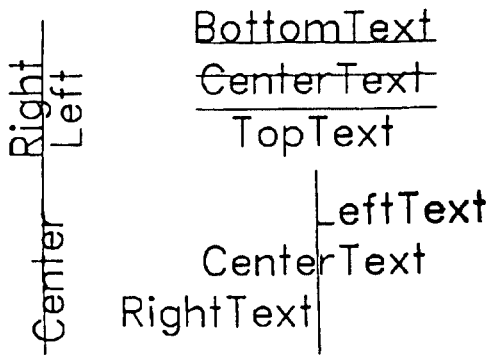
Yours sincerely,

P.J. CROSS
Group Leader - Facsimile Research

Registered in England: No. 2008
Registered Office: 10 Vicars Lane, Hford, Essex.

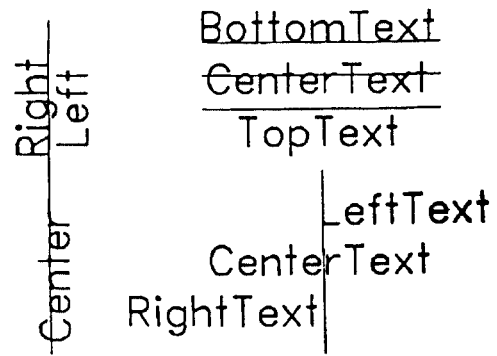
그림 5. 실험 문서 3

Fig. 5. Test Document 3



Several options for text alignment are supported. Notice that the CENTER TEXT line (vertical orientation at left) is centered according to the total font height. This includes blank header space above the characters. All alignments include this header space, as well as a brief margin following the end of a string (or a single character).

그림 4. 실험 문서 2
Fig. 4. Test Document 2



Several options for text alignment are supported. Notice that the CENTER TEXT line (vertical orientation at left) is centered according to the total font height. This includes blank header space above the characters. All alignments include this header space, as well as a brief margin following the end of a string (or a single character).

그림 7. 복원 문서 2
Fig. 7. Reconstructed Document 2

Drive carefully and slowly when pedestrians are about, particularly in crowded shopping streets, when you see a bus stopped, or near a parked mobile shop. Watch out for pedestrians coming from behind parked or stopped vehicles, or from other places where you might not be able to see them.

Three out of four pedestrians killed or seriously injured are either under fifteen or over sixty. The young and the elderly may not judge speeds very well, and may step into the road when you do not expect them. Give them, and infirm, or blind, or disabled people, plenty of time to cross the road.

Drive slowly near schools, and look out for children getting on or off school buses. Stop when signalled to do so by a school crossing patrol showing a "Stop Children" sign.

Be careful near a parked ice-cream van-children are more interested in ice-cream than in traffic.

When coming to a zebra crossing, be ready to slow down or stop to let people cross. You must give way once they have stepped on to a crossing. Signal to other drivers that you mean to slow down or stop. Give yourself more time to slow down or stop on wet or icy roads.

Never overtake just before a zebra crossing.

From The Highway Code

그림 6. 복원 문서 1
Fig. 6. Reconstructed Document 1

있고 1의 집단에 쌓여있는 0의 집단은 문자의 훼손으로 간주하여 처리하였다.

이런 과정을 거쳐 처리된 문서데이터에서 각 문자를 추출하였으며, 추출 과정으로 먼저 가로로 전체가 0로 구성된 수평성분이 연속인 곳은 줄과 줄사이로 간주하여 이를 기준으로 각 줄을 추출하고, 추출된 줄내에서 수직으로 전체가 0으로 구성된 곳은 문자와 문자 사이로 간주하여 분리하였다. 이때 문자와 문자가 입력시 붙어서 입력되는 병합 오차와 한 문자가 둘로 나누어져 입력되는 분할 오차 역시 무시할 수 없으며 이는 분리된 문자의 크기를 조사하여 크기가 최대 문자의 크기보다 큰 경우는 두 문자가 붙은 경우로 이때는 그 가운데를 기준으로 하여 근처의 가장 적은 혹은 화소로 구성된 세로 줄을 분리시켰고, 분할 오차의 경우는 발생한 문자의 크기를 조사하여 최소 크기 이하의 분할 오차로 간주하여 가장 가까이 위치한 문자로 포함하여 처리하였다. 이렇게 분리된 문자들을 중심이동하여 28크기로 정규화시켰다.

분리된 문자를 학습시키기 위하여 입력 데이터와 원하는 결과값으로 구성되는 학습쌍을 바 마스크를 이용하는 음영 부호화 방식을 이용하여 구성 하였다. 음영 방식으로 8개의 바 마스크를 이용하였으며, 이는 각 성분의 정보를 최대한 추출해내기 위해서였고 각 마스크당 28개의 데이터가 필요하지만 데이터



THE SLEREXE COMPANY LIMITED

SAPORS LANE - BOOLE - DORSET - BH 25 8 ER

TELEPHONE BOOLE (945 13) 51617 - TELEX 123456

Our Ref. 350/PJC/EAC

18th January, 1972.

Dr. P.N. Cundall,
Mining Surveys Ltd.,
Holroyd Road,
Reading,
Berks.

Dear Pete,

Permit me to introduce you to the facility of facsimile transmission.

In facsimile a photocell is caused to perform a raster scan over the subject copy. The variations of print density on the document cause the photocell to generate an analogous electrical video signal. This signal is used to modulate a carrier, which is transmitted to a remote destination over a radio or cable communications link.

At the remote terminal, demodulation reconstructs the video signal, which is used to modulate the density of print produced by a printing device. This device is scanning in a raster scan synchronised with that at the transmitting terminal. As a result, a facsimile copy of the subject document is produced.

Probably you have used for this facility in your organisation.

Yours sincerely,

P.J. CROSS
Group Leader - Facsimile Research

Registered in England: No. 2038
Registered Office: 60 Vicars Lane, Hford, Essex.

그림 8. 복원 문서 3

Fig. 8. Reconstructed Document 3

의 발생분포에 의해 처음과 마지막 부분의 2개의 화소씩은 거의 대부분 데이터가 발생되지 않고 많은 정보량을 포함하고 있지 않아서 제외시키고, 24개의 데이터만 가지고도 결과를 얻을 수 있었으므로 한 문자당 192개의 입력데이터가 필요하였다. 음영 부호화를 이용하는 방법은 세선화와 특징점 추출등의 전처리가 필요치 않고 단지 화소가 존재하는 지의 여부만으로 입력 데이터를 구할 수 있으므로 시간이나 소요되는 기억장치의 감소등의 장점을 가질 수 있어 효율적이면서도 데이터의 분포, 특성, 크기, 위치정보등을 쉽게 얻을 수 있었다.

이렇게 구한 입력패턴을 이용하여 학습을 실시하였으며, 학습으로 얻은 신경회로망의 가중치를 이용하여 학습에 사용된 문자는 물론 학습시키지 않은 문자에 대해서도 거의 100%의 인식이 가능하도록 하였다.

인식된 문자를 원 데이터로 복원하기 위해서는 인식된 문자가 위치할 장소를 알고 있어야 하므로 이 위치정보를 먼저 전송한 후 인식문자를 전송하여야 한다. 위치정보의 전송은 발생확률에 의한 방법에 따라 발생한 확률을 구한 후 이를 이용하여 Huffman 부호표를 만들어 전송한 결과와 본 논문에서 제안하는 부호표에 의해 부호화하는 방법, 두 가지를 적용하였다. 부호표에 의한 방법이 더 많은 정보를 요구하고 있지만 기존의 다른 방법에 비해서는 상당히 압축된 결과이며, 일반성 또한 가지고 있어 많은 곳에서의 활용이 기대된다.

이상의 방식에 의해 문자만으로 구성된 이진 데이터의 전송이 가능하며, 결과의 문자체 종류는 수신단에서 선택 가능하며 사용자의 취미에 맞게 선택할 수 있다.

그리고 문자만으로 구성된 문서가 아니라 도형이 포함된 문서에 대해서는 문자와 도형부분을 따로 처리하였으며, 본 연구에서는 28크기 이상과 인식이 불가능한 문자에 대해서는 도형으로 간주하여 처리하였고 이들 도형부분은 Group 4방식인 MMR 부호화 방식을 이용하여 전송하고, 문자부분은 제안한 방식으로 신경회로망을 이용하여 인식한 후 부호화하였다.

이상의 방법을 실험 문서에 적용하였을 때 총 소요된 정보량을 표 3에 나타내었으며, 비교를 위하여 CCITT에서 권고한 Group 3방식과 Group 4방식으로 전송한 결과들 역시 제시하였다.

이 결과는 IBM PC-486을 이용하여 시뮬레이션한 결과이며, 소요된 시간은 문서에 포함된 문자의 수에 따라 차이를 나타내지만 평균적으로 1문자를 추출하는 데 소요되는 시간은 약 0.05초이며, 인식에 소요되는 시간은 병렬로 처리되어야하나 PC로 처리하였을 때 문자당 0.09초였다.

이상의 결과에서 알 수 있듯이 권고안 Group 3에 비해서는 최대 10.187:1, 최소 3.022:1의 압축율을 얻을 수 있었고, Group 4에 비해서는 최대 7.489:1 최소 1.951:1의 압축율을 얻을 수 있어 제안하는 방식의 효율성을 입증할 수 있었다.

그리고 CCITT에서 표준 문서로 제시한 그림 5의 실험 문서에 본 논문에서 제시한 방식을 적용하여 보다 객관적인 결과를 제시하였으며, 타이프로 구성된 문서에 있어서도 적용 가능하다는 점을 보였다.

이상의 실험 결과를 통하여 볼 때 Neuro-FAX 방식을 이용하며 기존의 팩시밀리 방식보다 평균 5.67:1 정도의 데이터를 줄일 수 있어, 전송시간이나 저장용량을 줄일 수 있어 기존의 방식에 비해 많은 장점

표 3. 총 소요 정보량

Table 3. Total Data Bits

		구 분			총소요데이터(bit) / 압축율(bpp)		
		위치	문자	도형	총	G3(MRC)	G4(MMR)
문서1	발생확률	5643	7864		13507 / 0.011	137593 / 0.116	101157 / 0.086
	부호화	6936	7864		14800 / 0.013		
문서2	발생확률	1562	2280	23647	27489 / 0.030	83063 / 0.091	53637 / 0.058
	부호화	2337	2280	23647	28264 / 0.031		
문서3	발생확률	11995	6296	17427	35718 / 0.015	157434 / 0.068	108027 / 0.047
	부호화	14068	6296	17427	37791 / 0.016		

을 가짐을 보였으며, 인식에 사용된 문자가 영문자를 중심으로 하고 있어 다른 문자, 즉 한글이나 일어등은 인식이 되지않아 도형으로 처리되고 있으나 보다 폭 넓은 문자의 인식만 이루어진다면 더욱 효과적인 방식으로 발전 가능하리라 예상된다.

VI. 결 론

본 연구에서는 점차 확고하게 사무자동화 기기로서 자리를 잡고 있는 팩시밀리에서의 데이터 전송 방식을 개선하여 인쇄체 문자에 대해서만 정보 손실 (information loss) 방식을 이용하고 나머지 부분은 기존의 팩시밀리 방식을 이용하여 기존 팩시밀리가 가지는 장점을 최대한 활용하면서 빠른 전송과 고품질의 복원문서를 얻을 수 있는 Neuro-Fax 방식을 제안하였다. 신경회로망의 놀라운 인식능력은 인간두뇌의 조직과 방식을 이용한 것으로 최근에 와서 많은 연구와 적용이 이루어지고 있으며 기억의 분산저장이나 병렬처리방식은 많은 장점을 가지고 있어 패턴 인식분야를 중심으로한 많은 분야에서 응용과 연구가 이루어지고 있다. 본 연구에서는 CCITT에서 제시하고 있는 권고안중에서 문자와 도형부분을 따로 처리하는 방식에 부합되면서 각 문자를 신경회로망의 인식능력을 이용하여 인식한 후 인식된 문자정보만을 전송하는 방식으로 수신측에서는 미리 가지고 있는 각 문자 패턴으로 문서를 복원함으로써 복원문서의 일그러짐이나 잡음, 훼손을 방지할 수 있고, 전송시간의 감소로 인하여 선로상의 혼잡을 줄일 수 있으며, 전화비등의 경제적인 몇가지 해결할 수 있는 많은 장점을 얻고자 하였으며, 이를 최대한 이용하기 위하여 팩시밀리 문서상에서 발생할 수 있는 모든 영문자, 숫자, 특수기호를 신경회로망에 학습시켰으며 이들의 거의 완전한 학습으로 인하여 더욱 효율적인 압축과 복원이 가능하였다. 그리고 문자가 아닌 기호나 도형, 혹은 손으로 작성한 문자, 문자의 훼손이나 학습의 미흡으로 인하여 인식이 불가능한 부분에 대해서는 기존의 팩시밀리 전송방식을 이용하여 복원시켜 기존 방식의 장점을 최대한 활용하였다. 실험결과 기존의 팩시밀리 전송방식인 CCITT 권고안 Group 3과 Group 4 방식에 비해 순수하게 문자만으로 구성되었을 때에는 최대 10:1 이상의 높은 압축율과 깨끗한 복원 화상을 얻을 수 있었고, 많은 도형이 포함되어 있는 문서에 있어서도 최대 3.022:1 정도의 압축율로 화상을 복원시킬 수 있어 제안한 방식

의 효율성을 입증할 수 있었고, 송신시 문자정보만을 전송하기 때문에 처리시간면에서도 많은 단축을 얻을 수 있고, 회로가 저항과 합산기만으로 구성되기 때문에 하드웨어 역시 간단하여 앞으로의 가능성을 보여주고 있다.

본 논문에서는 영문자를 중심으로한 ASCII 코드에서 발생하는 특수문자에 대해서 거의 완벽한 인식을 이루었으나, 아직 한글이나 다른 문자에 대한 인식은 실시하지 않아서 현재의 알고리즘으로는 다른 문자는 인식을 할 수 없어 도형으로 처리되기 때문에 다른 문자가 많이 포함된 문서에 있어서는 기존의 방식에 비해서 압축율면에서는 만족할 만큼 개선되지는 않았으나 현재의 결과로 볼 때 다른 문자에 대해서도 전처리 과정이나 명확한 특징점에 대한 연구가 완료되면 인식이 가능하여 제안한 Neuro-FAX 방식을 그대로 적용할 수 있어 지속적인 연구가 요구된다.

참 고 문 헌

1. CCITT Revised Recommendation T.4, "Standardization of Group 3 Facsimile apparatus for Document Transmission," 1984.
2. CCITT Recommendation T.6, "Facsimile Coding Schemes and Coding Control Functions for Group 4 Facsimile Apparatus," 1984.
3. CCITT Recommendation T.5, "General Aspects of Group 4 Facsimile Apparatus," 1984.
4. Yasuhiko Yasuda, "Advances in FAX," Proc. IEEE, Vol. 73, No. 4, April, 1985.
5. Rafeal C. Gonzalez & Paul Wintz, "Digital Image Processing," Addison-Wesley, 1977.
6. R. Hecht-Nielsen, "Neurocomputing : Picking the human brain," IEEE Spectrum, pp. 36-41, March, 1988.
7. Philip D. Wasserman, "Neural Computing-Theory and Practice," Van Nostrand Reinhold, 1989.
8. Yoh-Han Pao, "Adaptive Pattern Recognition and Neural Networks," Addison-Wesley Pub. Co., 1989.
9. Marily McCord Nelson & W. T. Illingworth, "A Practical guide to Neural Nets," Texas Instruments, 1991.

10. David E. Rumelhart & James L. McClelland, "Parallel Distributed Processing," The MIT Press, 1987.
11. 甘利俊一, "神經回路網の 數理-腦の情報處理樣式," 産業圖書, 1978.
12. J. S. Denker & W. R. Garner, "Neural Network Recognizer for Hand-Written ZIP Code Digits," AT & T Bell Lab., 1987.
13. Y. Le Cun & L. D. Jackel, "Hand-Written Digit Recognition : Applications of Neural Network Chips and Automatic Learning," IEEE Comm. Magazine, Nov., 1989.
14. Keun Young Lee & Murray J. J. Holt, "Recursive Segmentation of Binary Character Patterns for Improved Data Compression in Digital Facsimile," IERA Fifth International Conference and Tutorial Day on Digital Processing of Signals in Communications, Sep., 1988.
15. Murray J. J. Holt, "Data Compression Schemes for Binary Digital Images," Loughborough University of Technology, August, 1987.
16. D. J. Burr, "Experiments on Neural Net Recognition of Spoken and Written Text," IEEE Trans. Speech and Signal Processing, vol. 36, No. 7, pp. 1162-1168, Jul., 1988.



金 海 洙(Hae Soo Kim) 正會員
 1960年 11月 20日生
 1986年 2月 : 성균관대학교 전자공학과 졸업(공학사)
 1988年 2月 : 성균관대학교 대학원 전자공학과 졸업(공학석사)
 1992年 8月 : 성균관대학교 대학원 전자공학과(공학박사)

1991年 3月 ~ 현재 : 안양전문대학 전자통신과 전임강사

李 根 泳(Keun Young Lee) 正會員
 1947년 12월 30일생
 1973년 : 전남대학교 전기공학과 졸업
 1975년 : 한양대학교 대학원 전자공학과 석사학위 취득
 1978년 : 한양대학교 전자공학과 박사학위 취득
 1979년 3월 ~ 1980년 2월 : Denmark 공과대학(연구)
 1987년 9월 ~ 1988년 8월 : 영국 Loughborough대학(연구)
 1977년 3월 ~ 1981년 8월 : 광운공대 조교수
 1981년 9월 ~ 현재 : 성균관대학교 전자공학과 교수
 ※ 관심분야 : 스윙칭 이론, 마이크로 프로세서, 영상처리, 신경회로망등.