

□ 論 文 □

# 복합형GLVQ 神經網을 利用한 車種分類 模型開發

The Development of a Model for Vehicle Type Classification  
with a Hybrid GLVQ Neural Network

**趙 亨 基**

(亞洲大學校 交通研究所 先任研究員)

**吳 榮 泰**

(亞洲大學校 交通工學科 教授)

目 次

I. 서론	3. GLVQ
II. 이론적 고찰	4. GLVQ 복합모형
1. 신경망 이론	IV. 차종인식 실험
2. 차종구분 검지기의 국내외사례	1. 실험환경 및 방법
III. 차종구분 알고리즘 개발	2. 자료의 수집 및 처리
1. 통계량 추출방법(1)	3. 실험결과 및 비교분석
2. 통계량 추출방법(2)	V. 결론 및 향후과제
	참고문헌

ABSTRACT

Until recently, the inductive loop detectors(ILD) have been used to collect a traffic information in a part of traffic managment and control. The ILD is able to collect a various traffic data such as a occupancy time and non-occupancy time, traffic volume , etc. The occupancy time of these is very important information for traffic control algorithm, which is required a high accuracy. This accuracy may be improved by classifying a vehicle type with ILD. To classify a vehicle type based on a Analog-Digital Converted data collected form ILD, this study used a typical and a modified statistic method and General Learning Vector Quantization unsupervised neural network model and a hybrid model of GLVQ and statistic method. As a result, the hybrid model of GLVQ neural network model is superior to the other methods.

## I. 서 론

도로상에서 야기되는 각종 교통상태를 판단하고 이에 따른 적합한 신호제어를 수행하기 위해서는 목적인 교통정보를 수집할 수 있는 다양한 검지 센서들이 도로에 매설되거나 도로변에 설치되어야 한다. 교통정보를 수집하기 위하여 이용되는 검지 목적의 센서들은 유도식루프 검지기와 마이크로 웨이브 및 초음파 센서 등 매우 다양한 센서들이 사용되기는 하지만 국내는 물론 국외에서도 아직 ILD(Inductive Loop Detector : 이하 ILD라 함)가 주를 이루고 있는 실정이다. 비록 최근에 이러한 ILD의 관리체계 및 보수 유지에 관한 비용의 증대 및 수집자료의 질적, 양적 다양성을 위해 기존의 검지체계를 대체할 만한 다양한 검지센서에 관한 관심이 고조되고 있으며, 더욱이 최근 교통부문에서 도시부 및 고속도로에서의 버스전용차선 단속체계 도입의 필요성이 급속도로 부각되고 있어 기존 ILD의 대체연구가 더욱 가속화될 전망이다.

버스전용차선의 단속을 위해서는 단속차선에 설치된 검지기로 부터 차종에 관한 정보가 수집될 수 있어야 한다. 그러나 ILD를 통하여 수집될 수 있는 기초정보는 교통량(Volume), 점유시간(Occupancy Time) 및 비점유시간(Non-Occupancy Time)이며, 차종 등과 같은 기초정보는 단일 ILD를 기반으로 수집이 불가능한 것으로 인식되어 왔으며, 최근까지는 차종에 대한 정보를 수집하기 위하여 외국의 장비를 도입하여 교통량조사 목적으로 설치운영하고 있다.

국내의 경우 고속도로관리체계(Freeway Traffic Management System : FTMS)에서 ILD를 진행방향으로 2개 이상으로 설치하여 차종을 단지 2종 혹은 3종으로 구분하고, 이를 기반으로 지점속도(Spot Speed)를 예측할 수 있다

[14]. 그러나 이러한 방법은 정보의 질적 및 양적 다양성에 비하여 비경제적인 방법으로 그동안 도시부 관제 시스템에서는 지양되어 왔었다.

이에 본 연구는 도로를 통행하는 각 차종을 구분하기 위하여 기존의 단일 ILD 환경하에서 차량이 도로에 설치된 ILD를 통과시 각 차종별 발생하는 인덕턴스(inductance)의 변화를 기초자료로 수집하여 보다 효과적으로 인식할 수 있는 차종인식 알고리즘 개발에 그 목적이 있다.

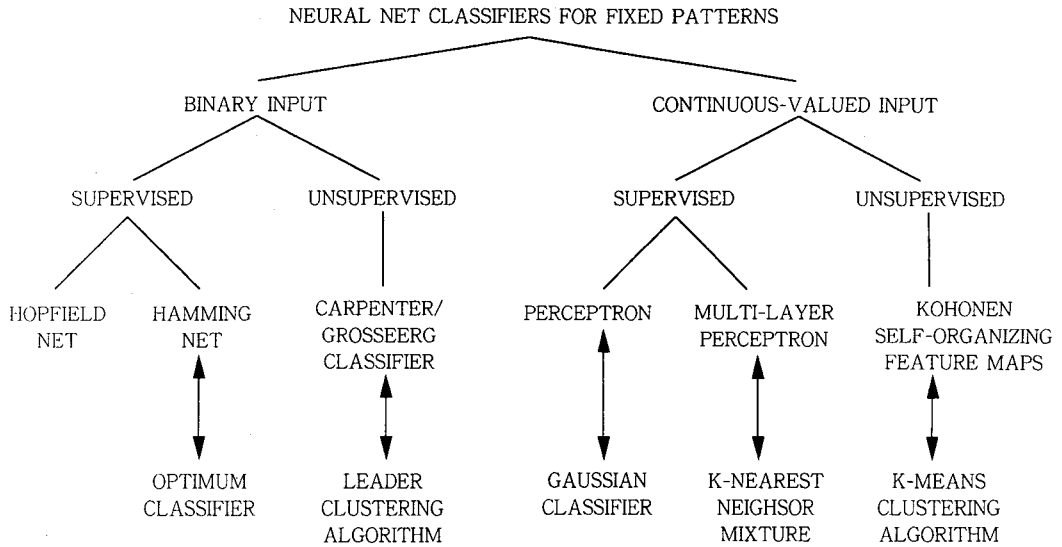
본 연구에서 차종을 인식하기 위하여 적용한 방법은 1993년 Nikhil R. Pal, James C. Bezdek, Eric C.-K. Tsao가 제안한 비교사학습신경망모형인 GLVQ(General Learning Vector Quantization) 알고리즘과 GLVQ에서 구분하지 못한 차종을 통계량을 이용하여 재분류하는 통계량-GLVQ복합모형을 적용하였으며, 특히 통계량-GLVQ복합모형은 본 연구에서 제안하는 알고리즘으로 이들 각 적용 모형별 비교분석 결과를 제시하였다.

## II. 이론적 고찰

### 1. 신경망 이론

신경망 이론은 단순퍼셉트론(single perceptron)을 기반으로 1960년대 초반에 출범한 이론으로 최근 다양한 모형들이 제안되고 있다. 이러한 신경망모형들은 학습방법에 따라 구분되며, 크게 교사학습방법(Supervisor Learning Method)과 비교사학습(Unsupervisor Learning Method)로 구분된다. 신경망 모형의 학습방법에 의한 구분은 다음 <그림 1>과 같다.

#### 1) 교사학습모형

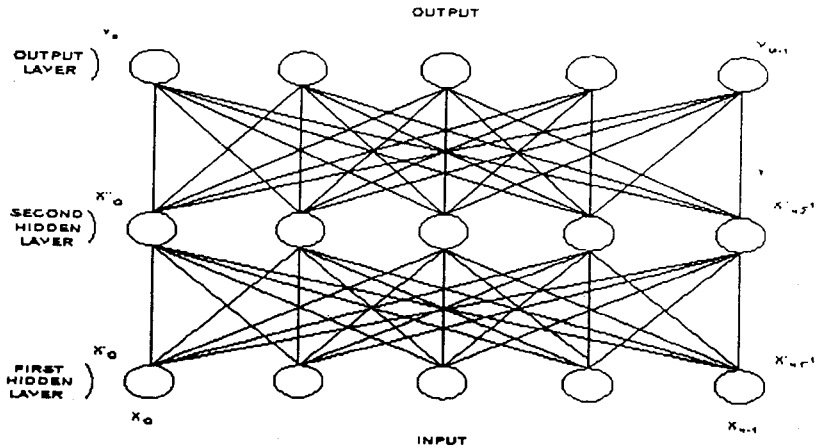


〈그림 1〉 신경망 모형의 구분(Lipmann : 1987)

교사학습방법은 초기 단순퍼셉트론에 기초한 모형으로 1980년대에 다층퍼셉트론(Multi-Perceptron)구조의 역전파신경망모형(Back Propagation Neural Network Model : BPN)이 정립되어 최근 예측 및 제어분야에 이르기까지 다양하게 응용되고 있다. 이 BPN 모형은 기존의 단순퍼셉트론이 가지고 있는 단점인 선형분리 문제를 근본적으로 해결할 수 있는 모형으로 〈그림 2〉와 같이 입력층과 출력층 사이에 은닉

층(Hidden Layer)을 최소한 하나 이상을 포함하는 구조를 가지고 있으며, 출력층에서 발생하는 오차(Error)를 출력층에서 부터 입력층에 이르기까지의 연결행렬들을 학습률(Learning Rate)의 비율로 매 반복과정(iteration)마다 조정함으로써 입력된 패턴을 학습하며, 모형의 특성상 학습효과가 우수하여, 우수한 인식효과를 기대할 수 있는 것이 특징이다.

그러나 이러한 BPN 이 가지고 있는 단점은



〈그림 2〉 일반화된 역전파신경망 모형의 구조(Architecture)

다음과 같다[1,20].

첫째, 지역최소(local minimum)에 빠질 수 있다.

BPN에서는 보통 최급강하법(Steepest Descent method)의 학습방법을 사용하는데 이는 현 위치에서 기울기가 적은 방향으로 이동하므로 학습결과 오차가 0인 전역최소(Global Minimum)가 아닌 지역최소에 도달할 수 있다.

둘째, 학습알고리즘에서 시그모이드변환함수(Sigmoid Function)의 입력치가 포화상태에 도달시 학습효과가 없게 된다.

학습율이 0과 1 사이에서 결정되며, 매 반복과정에서 발생하는 오차는 매우 작은 값이 발생되어 학습에 의한 오차보정율이 소숫점이하의 매우 작은 값이 되어 더 이상 오차보정이 안되는 경우가 발생되어 반복과정이 종료되면 그 결과가 최적인지 결정하기 어렵다.

셋째, 최적 인식을 위한 은닉층의 수를 결정하기 어렵다.

최근까지 은닉층의 수를 시행착오에 의한 임의의 탐색적 방법을 통하여 결정하고 있다.

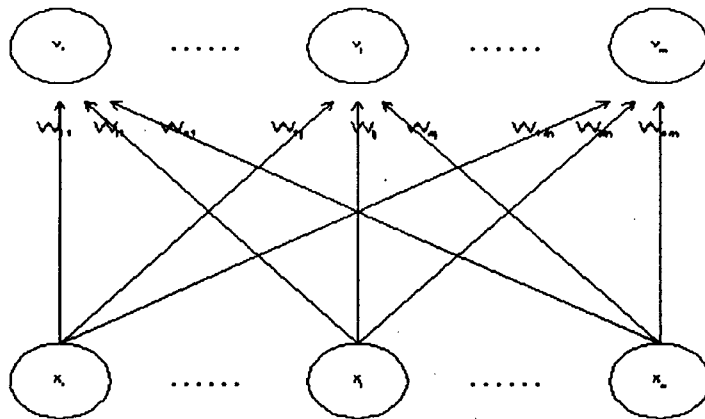
넷째, 각 은닉층에 구성되는 최적 뉴런(neuron)의 수를 결정하기 어렵다.

다섯째, 학습하는 데 있어서 다층구조에 의한 학습시간이 많이 소요된다.

여섯째, 다층구조에 의한 학습결과 현장적용에 있어서 실시간이 불가능한 경우가 발생할 수 있다. 매우 복잡한 망구조를 갖는 BPN의 경우 계산량이 많아 현장에서의 현장적용이 불가피한 경우가 발생할 수 있다. 최근 고성능 디지털신호처리(Digital Signal Process :DSP)의 개발 활성화로 실시간성 구현이 가능해지고 있으나 부대비용의 증대로 아직까지 개선의 여지가 남아있다.

2) 비교사학습모형

비교사학습방법에 의한 신경망은 오랜기간동안 꾸준히 연구된 모형들로서 Mcqueen 에 의하여 제안된 통계적 모형인 K-Means 알고리즘에 기초한 모형들이 주를 이루고 있다. 이러한 신경망모형은 크게 Kohonen Network과 Carpenter & Grossberg의 ART-I(Adaptive Resonance Theory- I), ART- II 모형이 있다. Kohonen Network은 자기조직화(Self Organization Feature Maps)모형과 학습벡터양자화(Learning Vector Quantization :LVQ)모형으로 구분된다. LVQ 망구조는 <그림 3>과 같다. 특히 이 방법은 입력패턴과 클러스터의 중심값간의 거리(Euclidian Distance)를 최소화하는 알고리즘으로 K-Means



<그림 3> LVQ 모형의 망구조

알고리즘과 대응된다는 점이 특징이며, 모형적 용시 사전에 입력 그룹에 대한 클러스터의 수를 정의해야 한다는 점이 단점으로 될 수 있으나 망구조가 간단하여 학습결과 현장적용에 있어서 실시간이 용이하다는 것이 장점이 될 수 있다.

ART-I 모형과 ART-II 모형은 1976년에 제안된 전형적인 비교사학습모형으로 망구조적으로 입력층(Input Layer)과 클러스터층(Cluster Layer) 그리고 인터페이스층(Interface Layer)으로 구성되며, 알고리즘이 수행시 패턴이 입력이 되면 기존의 클러스터들에 대하여 귀속여부를 평가하여 결정한다. 귀속여부결과 해당 클러스터가 존재하지 않으면 새로운 크로스터를 생성하여 귀속시키며, 귀속될 수 있는 클러스터가 있는 경우에는 해당 클러스터에 귀속시켜 학습을 진행하도록 되어 있다. ART-I 모형과 ART-II 모형의 차이점은 ART-I 모형은 2진화된 자료(binary data)만을 처리하는 반면에 ART-II 모형은 임의의 정량화된 연속값(continuous data)을 수행할 수 있다는 점이 다르다. Kohonen Network 과 ART 모형들의 차이는 Kohonen 모형이 클러스터의 수를 사전에 설정하여야 하는 반면에 ART 모형들은 Cluster의 수를 학습수행중 자율적으로 결정한다는 점이 다르다.

최근에 발표된 것으로 Pal.et.al.(1993)의 GLVQ (Generalized Learning Vector Quantization)알고리즘이 있다. 이 방법은 Kohonen의 LVQ알고리즘에서 초기 연결강도(initial weight)에 따라 클러스터의 중심값이 크게 영향을 받아 오분류(misclassification)되는 패턴이 발생하는 경우가 많다는 단점을 개선시킨 방법으로 모든 출력 노드에 대하여 학습을 함으로써 학습률, 반복횟수, 초기 연결강도의 값에 크게 영향을 받지 않고 일정한 클러스터의 중심값이 계산되기 때문에 각 클러스터별로 오분류 되는 패턴이 적게 나타난다는 장점이 있다. 실제 클러스터링 데이터로 많이 쓰이는 Fisher의 IRIS데이터를 가지고 실험을 하였을 경우 GLVQ알고리즘이 패턴 분류에 있어서 가장 정확도가 높은 것으로 나타났다.[19]

### 3) 교통분야의 적용사례

신경망이론의 교통분야에서의 적용된 연구사례는 최근 1989년부터 연구사례가 발표된바 있으며, 이때의 적용연구는 차량번호판 인식이었다. 교통부에서 주로 적용된 분야는 운전자행태분석에 주로 적용되었으며, 그의 모수 추정 및 항공 및 교통제어 등 짧은 시기에 다양한 분야에 적용된 바 있다. 각 분야별 연구사례가 다

<표 1> 교통부문에서의 신경망 적용사례 (MARK Dougherty : 1995)(계속)

분 야	적용연대	적용횟수	적용비율
운전자행태분석	1991 ~ 1994	12	23.0 %
모수추정(OD 추정)	1992 ~ 1993	7	13.4 %
차량감지 및 구분	1992 ~ 1993	6	11.5 %
도로의 보수유지	1992 ~ 1994	5	9.61 %
교통패턴 분석	1992 ~ 1993	5	9.61 %
비행기조작/운영	1991	4	7.70 %

분 야	적용연대	적용횟수	적용비율
교통량 예측	1992 ~ 1994	4	7.70 %
교통정책 및 경제	1991 ~ 1994	2	3.84 %
항공수송	1989 ~ 1991	2	3.84 %
해운수송	1991	2	3.84 %
submarine Vehicle	1990	1	1.90 %
지하철 운영	1994	1	1.90 %
교통제어	1991	1	1.90 %
총 계	1989 ~ 1994	52	100 %

또한 교통부문에 적용된 신경망 이론중 각 모형별 적용사례 건수는 다음 <표 2>와 같다.

<표2> 교통부문에 적용된 신경망 모형별 적용사례 (MARK Dougherty : 1995)

적 용 모 형	적 용 횟 수	비 고
BackPropagation	36	교통량단기수요예측 교통패턴 분석
Learning Vector Quantization	7	지하철 운영 적용(Fuzzy 결합) 해운수송적용(화상처리기반) 교통정책 및 경제 운전자 행태분석
Adaptive Resonance Theory	4	운전자 행태분석
Self Organization Feature Maps	2	비행기 조작/운영 운전자 행태분석
ADALINE	1	-
Hopfield Network	1	도로보수유지
Basis Function	1	-
Counterpropagation	1	모수추정(OD 추정)
Boltzman Machine	1	비행기 조작/운영

음 <표 1>에 제시되어 있다.

본 연구와 관련하여 상기 <표1>과 <표2>에서 차량감지 및 구분에 대한 연구는 1990년 화상처리기반의 차량번호판 인식에 관한 연구를 초석으로 한 연구로서 비디오 화상검지기에서의 차

량검출 및 차종 구분에 관한 연구사례에 준한 연구이다.

## 2. 차종구분 검지기의 국내외 사례

ILD 기반의 차종 구분에 관한 연구사례는 1980년부터 최근에 이르기까지 연구사례가 매우 적으며, ILD 기반의 상용화된 제품이 주류를 이루고 있다.

1) 국내 연구사례

국내의 경우 1994년 한국도로공사의 FTMS 관련사업의 일환으로 고속도로에 각 차선당 2개의 원형검지기를 설치하여 대/중/소에 대한

차종을 구분한 바 있다. 이는 본래의 목적상 차종을 구분하기 위함이 아니고 단지 통행차량의 길이를 수집하기 위함이었다.

1995년 단일 ILD기반의 차종분류 알고리즘개발에 관한 연구논문이 발표된 바 있다(1995 : 조한선). 이 연구는 ILD로부터 총 6종의 차종에 대하여 총 899개의 개별차량에 대하여 수집된 인덱턴스의 변화를 계측하여 통계량 추출방법을 이용하여 통행차량을 6종으로 구분하였으며, 적용한 방법 통계량추출방법과 이를 기반으로

<단계 1> 기준자료 설정단계

- |   |
|---|
| <ol style="list-style-type: none"> <li>총 6개 차종별 수집된 원형자료(Original Data) 그룹화</li> <li>각 차종별 샘플링 수 및 인덱턴스 값의 정규화 값을 이용한 각 그룹별 통계량 추출</li> </ol> |
|---|

<단계 2> 1차 차종구분 단계

- |  |
|--|
| <ol style="list-style-type: none"> <li>총 899개 자료를 대상으로 샘플링 수 및 인덱턴스 값의 정규화</li> <li>각 그룹별 최 근접 그룹결정</li> </ol> $\min \left\{ \sum_{i=1}^{N_{tot}}  (Back(i, j) - Veh_N(i))  \right\}$ <p>(단 Back(i, j)는 각 그룹별 추출된 통계량값이며, Veh(i)는 대상차량의 정규화된 자료이다)</p> |
|--|

<단계 3> 재 그룹형성 단계

- |   |
|---|
| <ol style="list-style-type: none"> <li>단계 2의 결과를 토대로 총 3개 그룹으로 재구성<br/>그룹(승용차,봉고)와 그룹(버스,11톤,25톤트럭) 그리고 그룹( 트레일러)으로 구성</li> <li>원형의 자료들을 샘플링 수에 대해서만 정규화</li> </ol> |
|---|

<단계 4> 2차 차종구분 단계

- |  |
|--|
| <ol style="list-style-type: none"> <li>&lt;단계 2&gt;의 구분방식에 의거하여 승용차,봉고,버스를 구분</li> <li>11톤트럭과 25톤트럭은 그룹화 재 구성</li> </ol> |
|--|

<단계 5> 3차 차종구분 단계

- |  |
|--|
| <ol style="list-style-type: none"> <li>원형자료를 다시 샘플링수와 인덱턴스값에 대하여 정규화</li> <li>11톤 트럭과 25톤 트럭 구분</li> </ol> |
|--|

<단계 6> 최종 차종구분 단계

- |  |
|--|
| <ol style="list-style-type: none"> <li>&lt;단계 5&gt;의 25톤자료들을 샘플링수에 대해서만 정규화</li> <li>25톤과 봉고 재 분류</li> </ol> |
|--|

논리적 알고리즘을 적용하여 인식결과 인식을 약94%를 제시한 바 있다. 이 연구에서 차종을 구분하기 위하여 적용한 방법은 다음과 같다.

이 연구에서 적용한 통계량 추출방법은 본 연구에서 적용하려는 통계량 방법(1)과 동일한 방법이지만 본 연구에서는 각 차종별 변화되는 인덱턴스에 대해서 정규화하지 않는다는 점에서 차이가 있다. 더욱이 이 연구(조한선)는 연구 목적으로 수집된 차종 이외의 5톤,8톤,15톤 트럭류 등에 대한 관련자료가 없어 실제 적용시 연구에서 의도한 인식을 기대할 수 없는 경우가 있다.

외국의 경우 본 연구와 관련하여 이미 오래 전인 1980년에 연구사례가 있었다. 이 연구는 ILD기반의 연구사례로서 두개의 ILD를 이용하였다. J. J. Reijmers는 이 연구에서 ILD를 통행하는 차량에 의해 발생하는 인덱턴스의 변화에 대한 특징을 추출하기 위하여 인덱턴스 변화의 최대값과 최소값 그리고 차량의 길이 그리고 샘플링수 및 샘플링된 인덱턴스의 총합산값을 이용하여 통행차량의 파형에 대한 특징을 규정하였다. 이렇게 수집된 파형의 특징을 기반으로 차종을 인식하기 위한 방법으로는 각 특성치들을 기초로 논리적 알고리즘(logical algorithm)을 모형으로 설정하여 총 5개 차종을 대상으로 실험하였다. 실험결과 약 6.3 %의 평균

2) 외국 사례

<변수의 정의>

- 1. N = 샘플링된 자료의 총수/차량
- 2.  $\sum S_i$  = 샘플링된 인덱턴스값의 총합/차량
- 3.  $t_1 - t_2$  = 차량의 속도를 결정하는 사건간격
- 4. Boundary : A(128), B(256)

<단계 1> 특징수출

$$1. N^* = \frac{100}{C_1 - C_0} \times N$$

<단계 2> 1,2종 그룹과 3,4,5종 그룹으로 대분류 단계

- 1. if  $N^* \leq 22$  then
  - 1.1  $A = \text{Max. } \frac{N}{\sum S_i}$
  - 1.2 if  $A \leq \text{Boundary}$  then
    - 대상차종은 Delivery Van
  - Else
    - 대상차종은 Passenger Car
  - End if
  - 1.3 차종구분과정 종결
  - end if
- 2. 다음 단계 계속



<단계 3> 3.4.5종 재분류 단계

```

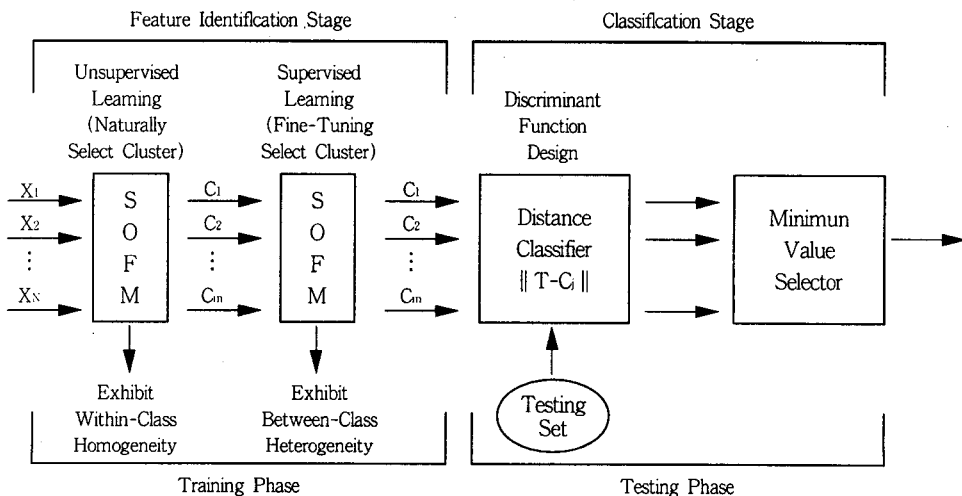
1. if N* ≥ 32 then
    1.1 B = Min.  $\frac{N}{\sum S_i}$ 
    1.2 if B = Boundary then
        대상차종은 Trailer
    Else
        대상차종은 Semi-Trailer
    End if
Else
    대상차종은 Truck
    차종구분 과정 종결
    End if
2. 차종구분단계 종결
    
```

오류율이 나타났다고 제시한 바 있다.[15] 이 연구에서 차종을 구분하기 위한 수행과정은 다음과 같다.

이 연구는 차종의 그룹을 재구성하여 적용하였다는 점에서 국내의 연구사례(1995 : 조한선)과 같다고 할 수 있으나 각 그룹별 구분을 위한 경계치 값은 도로를 통행하는 차량들의 특성에

따라 다르므로 국내의 경우에 적용시 그에 맞는 모수추정이 필수적으로 수반된다.

신경망이론을 적용한 차종구분관련 연구사례는 Radar에 의한 자료수집장치를 기반으로 승용차와 픽업차량 그리고 벤(Delivery van), 불도저(bulldozer)등 총 4개 차종을 인공신경망모델중 비교사학습모델인 LVQ(Learning Vector



<그림 4> LVQ를 이용한 차종인식 흐름도

Quantization)방법을 적용하였으며, 그 실험결과 인식률은 약 95%수준이었다.[16] 이 연구에 의한 연구수행과정은 다음 <그림 4>와 같다.

이 연구에서 적용한 신경망 모형인 LVQ는 본 연구에서 적용하는 GLVQ 이전에 개발된 모형으로 GLVQ와의 차이점은 학습방법이 다르다는 점에서 차이가 있다. GLVQ와 LVQ간의 수행도 차이는 이미 전절에서 언급한 바와 같으며, GLVQ의 학습방법은 3장 3절에서 설명되어진다.

### Ⅲ. 차종구분 알고리즘 개발

#### 3.1 방법론

본 연구에서는 효과적인 차종구분을 위한 모형을 구축하기 위하여 통계량추출방법 2가지 방법과 GLVQ 신경망모형 그리고 GLVQ신경망과 통계량 추출방법의 복합모형을 적용하였다.

통계량추출방법(1)은 전통적 통계량 산출방식으로 각 추출된 자료들에 대하여 기본적으로 정규화 1차처리 단계를 통하여 각 차종별 대표되는 평균값을 산출하여 대상차종을 식별하는 방식이며, 통계량추출방법(2)는 통계량추출방법(1)과 같으나 정규화 1차처리단계 이후 임의의 외곽점간의 거리제곱에 의한 평균값을 산출하는 2차처리 과정단계를 수행하는 것이 다르다. 또한 GLVQ모형은 신경망모형으로 수집된 모든 자료에 대하여 학습단계를 수행하여 학습된 결과자료를 기초로 대상차종을 인식하는 비교사 학습모형이다. 그리고 GLVQ복합모형은 GLVQ 신경망모형과, 통계량추출방법을 결합한 모형으로 GLVQ모형에서의 인식불능인 자료들을 통계량추출방법으로 효과적으로 보완할 수 있도록 한 모형이다.

#### 1. 통계량추출방법(1)

통계량 추출 방법에 의한 방법은 고전적인 클러스터링방법으로 널리 알려진 k-means 알고리즘과 유사한 관계로, 본 연구의 알고리즘 특성상 k-means 알고리즘의 특수형태라 할 수 있다.

k-means 알고리즘은 MacQueen 에 의하여 제안된 알고리즘으로서, 우선 패턴을 k개의 클러스터로 나눈 후 클러스터에 포함되어 있는 패턴들의 평균으로 클러스터의 중심값을 계산하고 이 중심값과 각 패턴과의 거리를 계산한 후 가장 거리가 가까운 클러스터에 패턴을 포함시키는 방법으로 그 조건은 (식 1)과 같다.[6]

$$x_i \in G \text{ iff } \|x_i - z_j\|^2 < \|x_i - z_k\|^2$$

여기서,  $1 \leq i \leq N, 1 \leq k \leq c, j \neq k$

N : 패턴 수

c : 클러스터 수

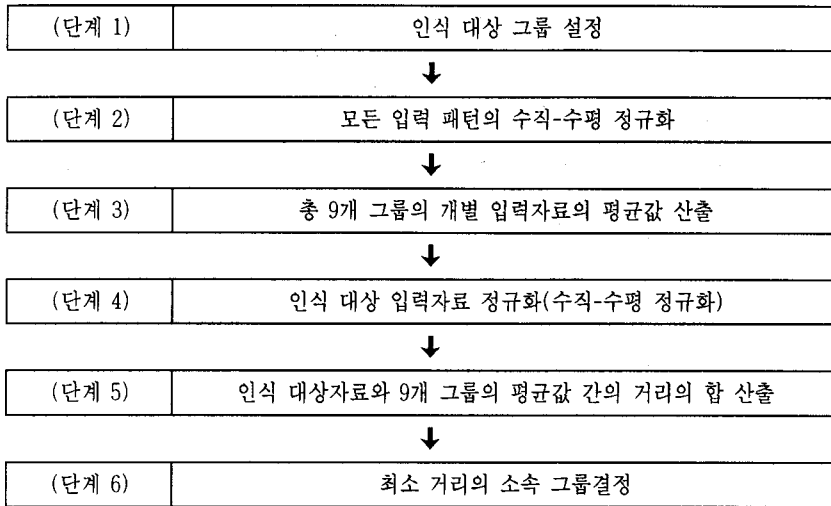
z : 클러스터 중심값 (식1)

그리고 이 계산은 각 클러스터의 중심값이 더 이상 변하지 않을 때까지 반복한다. 초기 k개의 클러스터의 중심값을 주는 방법에는 주어진 패턴에서 처음 k개의 패턴을 추출하여 중심값으로 하는 방법과 임의로 k개를 추출하여 중심값으로 하는 방법이 있는데 본 논문에서는 임의로 k개를 추출하여 클러스터의 초기 중심값으로 하였다.

통계량 추출방법은 각 대상 입력패턴을 구성하는 자료들간의 산출되는 평균값 통계량과 인식 대상의 입력값간의 거리를 기초로 하고 있다. 이 방법은 k-means에 의한 방법과 비교할 때 근본적인 접근방식은 같지만 본 통계량 추출방법이 사전에 각 그룹에 소속된 패턴의 입

력자료들의 통계량을 필요로하기 때문에 대상 입력패턴에 대한 차종정보을 알고리즘 수행전

에 알고 있어야 한다는 점이 다르다. 통계량 추출에 의한 인식방법의 알고리즘의



〈그림 5〉 통계량 추출방법(1)의 흐름도

흐름도는 다음과 같다.

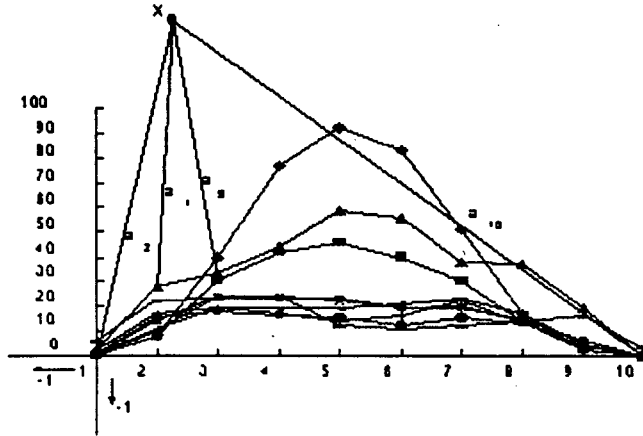
통계량 추출방법에 의한 인식과정에 있어서 평균치간의 거리는 일반적으로 비교사학습에 의한 전형적인 신경망 클러스터링 모델과 같이 유클리디안 거리(Euclidian Distance)를 의미한다.

〈그림 4〉의 흐름도에서 단계1에서는 수집된 모든 차종을 각각의 독립된 그룹으로 설정한다. 단계2에서는 수집된 총 1129개의 자료들에 대하여 샘플링수와 인덱터스의 값에 대하여 정규화를 수행하며, 본 연구에서는 샘플링의 수에 대해서만 정규화를 수행하였다. 단계 3에서는 단계1에서 규정한 각 그룹에 대하여 정규화과정에 의하여 변환된 자료에 기초하여 각 그룹별 대표되는 패턴요소(Pattern Element)의 평균치를 산출한다. 단계 4는 실제 인식실험을 수행하는 단계로서 인식대상의 차량에 대하여 샘플링의 수에 대한 정규화를 수행한다. 단계 5에서는

단계 1에서 설정한 각 그룹 중 가장 근접한 그룹을 결정하기 위하여 각 그룹별 대표치 패턴 요소와 정규화된 대상차량의 자료간 오차제곱의 합(Sum of Square Error : SSE)를 산출한다. 단계 6에서는 단계 5에서 각 그룹별 결정된 SSE들 중 가장 작은 그룹을 인식된 차종으로 한다.

## 2. 통계량추출방법(2)

통계량 추출방법(2)은 모든 처리과정이 방법 (1)과 동일하지만 차종인식에 사용되는 최종 자료군이 파형의 임의 외과 점과의 거리로 재변환되는 2차처리과정에 의하여 변환된다는 점에서 다르다. 제 2차 처리과정은 전처리과정에 의하여 정규화된 기초자료를 토대로 외곽점(extra-point)를 설정하여 설정된 지점과 각 입력값간 거리를 최종 자료로 활용한다.



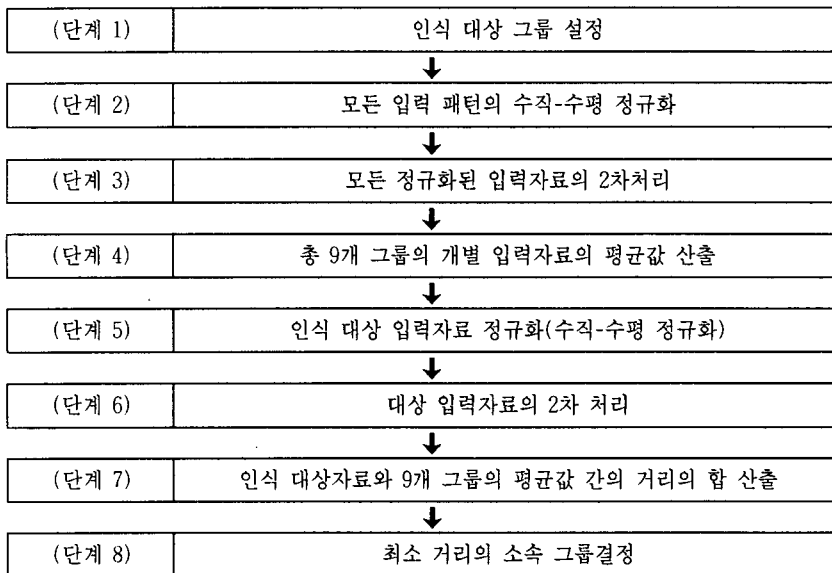
〈그림 6〉 2차처리에 관한 관계도(예 : 버스의 경우)

〈그림 6〉은 외관점  $x$ 와 버스의 특성자료의 경우에 대한 2차처리과정의 도면이다.  $a_1, a_2, \dots, a_{10}$ 은 각각  $x$ (기준점)으로부터 각 Data간 거리이며,  $\sqrt{(x_1-x_2)^2 + (y_1-y_2)^2}$ 의 기본식에 준하여 산출된다. 본 2차과정은 각 차종별 미묘한 파형의 특성을 증폭하기 위하여  $x$ (기준점)로부터의 각 거리의 제곱을 산출하도록 하였다.

$x$ 에 대한 기준점은 탐색적 방법에 의하여 산출하였으며, 본 연구에서는  $(x,y) \rightarrow (-500, 135)$ 을

$x$ 의 기준점으로 선택하여 적용하였다.

따라서 각 차종별 대표되는 자료는 〈그림 6〉에서와 같이 산출되는  $a_1^2, a_2^2, \dots, a_{10}^2$ 들의 평균치들을 산출하였으며, 인식기준은 통계량추출방법(1)과 동일하게 SSE(Sum of Square Error)를 산출하여 최소가 되는 그룹을 해당 차종으로 인식하게 된다. 이에 대한 알고리즘의 흐름도는 다음과 같다.



〈그림 7〉 통계량 추출방법(2)의 흐름도

### 3. GLVQ(Generalized Learning Vector Quantization)

Kohonen은 자기조직형태(Self Organization Feather Map : SOFM)의 개념을 입력형태의 특성들과 [6][7][8][9] 학습벡터 양자화(Learning Vector Quantization, LVQ)라 하는 알고리즘에 기초를 둔 연결정도를 조정하는 알고리즘을 제안하였다.[10][11][12][13]

실제로 LVQ 알고리즘은 클러스터링을 목적으로 한 알고리즘이 아니라 데이터의 프로토타입(prototype)을 설계하기 위하여 데이터 중에 가장 가까운 프로토타입(nearest prototype)들을 그룹화하는 알고리즘이다[14].

이 알고리즘은 근본적으로 입력치들과 중심치들간의 평균오차제곱(Mean Square Error)을 최소화하는 것으로 학습방법은 평균오차제곱에 대한 목적함수를 최소화를 최적화 시키므로서 양질의 분류가 가능하도록 하는 것이다. 여기서 오차의 제곱은 유클리디안 거리와 동일한 개념으로 취급될 수 있으며, LVQ의 학습규칙은 (식 2)와 같다.

$$v_{i,t} = v_{i,t-1} + \alpha_i(x_k - v_{i,t-1}) \quad (\text{식 2})$$

여기서,  $v_{i,t}$  는 반복 t시점에서 승자노드와 연결된 연결강도

$x_k$ 는 k번째 패턴벡터. ( $1 \leq k \leq N$ )

그러나 LVQ는 다음과 같은 몇몇 가지의 중요한 문제점들을 가지고 있다.

첫째, LVQ의 결과가 연결강도(weight)의 초기값에 매우 민감한 영향을 받는다. 즉, 클러스터의 중심에 대한 초기치가 입력 데이터의 볼록체 외부(out of convex hull)에 있는 경우에는 승자노드에 대해서만 학습을 하므로 승자가 아

닌 노드의 중심값에 대해서는 반복이 계속되더라도 초기의 연결강도값을 그대로 유지하게 된다.

둘째, 초기 학습률이 0에 가깝게 주어졌다면 반복이 계속되면서  $\alpha$ 는 감소하게 되어 0에 수렴하게 되므로 연결강도의 변화량이 허용오차안에 들어오지 않았는 데도 불구하고 (식 2)에서 두번째 항인  $\alpha_i(x_k - v_{i,t-1})$  값이 0에 가깝게 되어 연결강도의 조정이 이루어 지지 않는다. 따라서 학습이 끝나게 되기 때문에 오분류되는 패턴이 증가하게 된다.

이러한 문제점들을 극복하기 위한 방법으로 GLVQ알고리즘이 Pal, Beezdek, Tsao에 의하여 제안되었다[13]. 이 방법은 LVQ에서의 학습방법을 개선한 알고리즘으로 LVQ의 승자그룹학습(winner node learning)에 비승자노드의 학습(non winner node)을 부가하여 각각의 모든 입력값들에 대하여 학습을 수행하는 방법으로 승자노드와 입력값들간에 완전한 조화가 이루어질 경우에는 GLVQ는 LVQ와 같아지며, 승자그룹과 부조화가 증대되면 비승자노드의 학습과정에서 입력값들간의 충돌이 점차적으로 증대된다.

GLVQ는 손실함수(loss function)  $L_x$  (식 3)를 최소화 시키는 학습방법을 적용하였다[14].

$$L_x = \sum_{r=1}^c g_r \|x - v_r\|^2 \quad (\text{식3})$$

$$g_r = \begin{cases} 1 & , \text{if } r=i \\ \frac{1}{\sum_{j=1}^c g_j \|x - v_j\|^2} & , \text{if } r \neq i \end{cases}$$

여기서 승자노드일 경우에는  $L_x$ 를  $v_i$ (winner class center)로 미분하고, 승자노드가 아닌 경우에는  $v_j$ (non-winner class center)로 미분하여

(식 3)와 (식 4)로써 학습을 한다.

$$v_{i,t} = v_{i,t-1} + \alpha_t (x_k - v_{i,t-1}) \frac{D^2 - D + \|x_k - v_{i,t-1}\|^2}{D^2} \quad (\text{식4})$$

$$v_{r,t} = v_{r,t-1} + \alpha_t (x_k - v_{r,t-1}) \frac{\|x_k - v_{r,t-1}\|^2}{D^2} \quad \text{if } r \neq i \quad (\text{식5})$$

여기서,  $D = \sum_{r=1}^c \|x_k - v_r\|^2$ ,  $k=1, 2, \dots, n$ ;  $r=1, 2, \dots, c$  이고,

$\alpha_t$ 는 time  $t$ 에서의 학습률(learning rate)이다.

각 단계별 GLVQ 클러스터링 알고리즘은 다음과 같다.

GLVQ 클러스터링 알고리즘.

[단계 1] 패턴벡터  $X = \{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^p$

$c, T, \epsilon > 0$  결정

여기서,  $c$  : 클러스터의 수.

$T$  : 최대 반복횟수.

$\epsilon$  :  $\|v_t - v_{t-1}\|$ 의 허용오차

[단계 2] 초기 연결강도 결정.  $V_0 = (v_{1,0}, \dots, v_{c,0}) \in \mathbb{R}^p$

초기 학습률 결정.  $\alpha_0 \in (1, 0)$

[단계 3] For  $t=1, 2, \dots, T$ . 반복횟수에 따라 학습률 계산.  $\alpha_t = \alpha_0(1-t/T)$

For  $k=1, 2, \dots, n$ :

a. 승자노드의 발견.

$$\|x_k - v_{it-1}\| = \min \{ \|x_k - v_{jt-1}\| \}$$

$$1 \leq j \leq c$$

b. 승지노드 및 승자가 아닌 노드의 연결강도 수정

$$v_{i,t} = v_{i,t-1} + \alpha_t (x_k - v_{i,t-1}) \frac{D^2 - D + \|x_k - v_{i,t-1}\|^2}{D^2}$$

승자노드인 경우

$$v_{r,t} = v_{r,t-1} + \alpha_t (x_k - v_{r,t-1}) \frac{\|x_k - v_{r,t-1}\|^2}{D^2} \quad (r \neq i)$$

승자노드가 아닌 경우

Next k.

[단계 4] 에러 계산. 
$$E_t = \sum_{k=1}^n \sum_{r=1}^c |v_{rk,t} - v_{rk,t-1}|$$

[단계 5] if  $E_t \leq \epsilon$  then STOP; Else Next t.

$$U_{LVQ_{ik}} = \begin{cases} 1 : \|x_k - v_i\| \leq \|x_k - v_j\|, 1 \leq j \leq c, j \neq i \\ 0 : \text{otherwise} \end{cases}$$

$1 \leq i \leq c \text{ and } 1 \leq k \leq n.$

$U_{LVQ} = [U_{LVQ_{ik}}] : c \times n$  행렬로써  $x_k$  패턴이 클러스터  $c$ 에 포함되어 있는 상태를 나타내는 행렬.

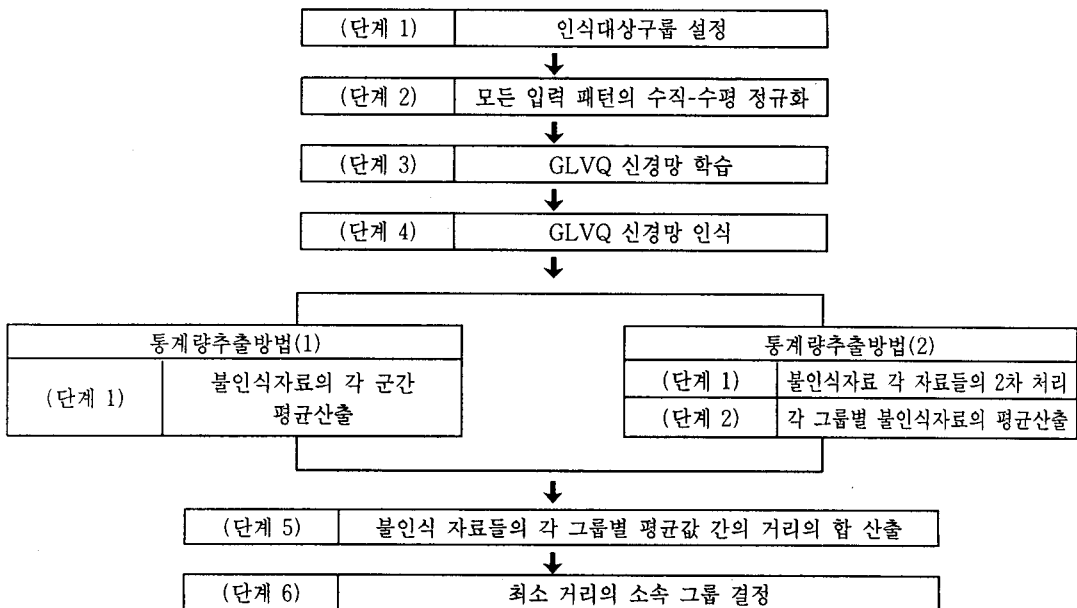
실제로 Nikhil R. Pal은 Anderson의 IRIS데이터를 가지고 반복횟수와 학습률을 변화시켜 가면서 LVQ와 GLVQ로 클러스터링 한 결과를 비교하였는데 GLVQ는 150개의 IRIS데이터중 에러의 갯수가 17개로 일정한 반면에 LVQ인 경우에는 100개에서 17개의 에러를 나타내고 있다.[13]

추출방법을 결합한 모형으로, 1차적으로 GLVQ에 의하여 학습하고 학습결과 인식이 불능인 자료들에 대하여 통계량 추출방법을 적용한 모형이다. 따라서 본 인식모형의 모체는 GLVQ가 되며, GLVQ에서 학습후 학습에 따른 인식결과에 따라 전절에 제시한 두가지의 통계량추출방법을 상호 결합하였다. 본 GLVQ와 통계량추출방법과의 결합모형의 수행흐름도는 다음 <그림 8>과 같다.

#### 4. GLVQ-통계량 추출방법의 결합모형

본 결합모형은 GLVQ신경망 모형과 통계량

<그림 8>에서 <단계 1>~<단계4>는 본 연구의



<그림 8> GLVQ와 통계량추출방법의 결합모형

전절 3의 GLVQ적용방법과 같으며, GLVQ의 인식결과를 토대로 인식율이 작은 차종들에 대하여 2차 인식과정을 수행한다. 본 연구에서는 GLVQ의 인식결과 트럭류의 경우 통계량 추출방법에 해당하는 2차인식과정을 수행한다. 이러한 2차 인식과정은 본 연구의 전절 1과 2에서 제시한 통계량추출방법을 연계하며, GLVQ와 연계되는 통계량 추출방법은 이들 두 통계량추출방법중 GLVQ의 낮은 인식율을 보완할 수 있는 모형을 연계한다.

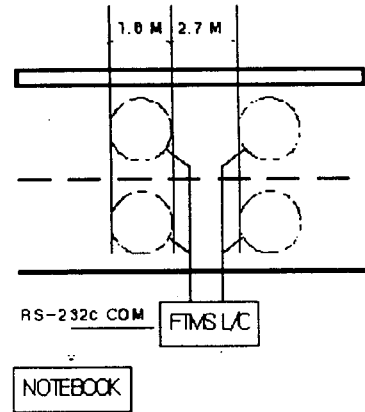
#### IV. 차종인식실험

##### 1. 실험환경 및 방법

본 연구에서 구축된 모형들은 현장에서 수집된 자료들을 기초로 각 모형별 인식률을 기준으로 수행도를 비교·분석 하였다.

본 연구에서는 차종을 인식하기 위하여 도로를 통행하는 각종 차량들을 9개 차종으로 규정하였으며, 각 규정된 차종은 승용차, 봉고, 버스, 2톤트럭, 5톤트럭, 8톤트럭, 11톤트럭, 15톤트럭, 트레일러로서 이들 각 차종에 대하여 1~9까지의 순번호를 설정하였다. 이들에 대한 자료는 고속도로 관리 시스템(Freeway Traffic Management System, FTMS)구축의 일환으로 설치된 총 320Km 구간중 중부고속도로의 구간중 판교-구리선에 설치된 직경 1.8M 크기의 원형 검지기로 부터 자료를 수집하였다. 자료수집 현장에서의 검지기 구성형태는 다음 <그림 9>과 같다.

이러한 원형 검지기로 부터 수집되는 자료는 통행차량이 검지기를 점유하는 동안 루프코일에서 발생하는 자장의 변화를 A/D 변환장치에 의하여 변환된 정보로서 목표한 차종(9종)에 관한 자료를 기초자료로 수집하였다.



<그림 9> 도로에 설치된 ILD

본 연구에서 수집된 자료들은 차량이 ILD에 점유된 시간동안 계측된 자료로서 점유시간은 통행속도에 직접적으로 영향을 받는다. 따라서 통행속도의 변화에 따라 매 20mmSec 단위마다 계측되는 샘플링수가 변화하게 된다. 이러한 영향은 각 차종에 대하여 샘플링된 자료들을 표준규격으로 변환하지 않고 원시 자료들을 그대로 클러스터링에 활용하는데 있어서 많은 문제들이 야기할 수 있다. 따라서 본 연구에서는 각 차종들에 대한 원시자료들을 수직성분(자장의 변화)과 수평성분(샘플링 순서) 각각에 대하여 2개 유형으로 정규화하였다.

이러한 파형의 정규화는 다양한 점유시간에 의한 파형 변화를 표준화함으로써 일률적이며, 효과적인 특징 추출을 위한 전처리과정이 될 수 있다. 검지기로 부터 수집된 파형을 정규화하기 위해서는 수집된 파형과 표준 규격에 대하여 다음의 사전정보를 알고 있어야 한다.

- ▷ 원시자료의 최대 샘플링수
- ▷ 표준규격의 수평성분의 크기

##### 1) 수평성분의 정규화

위에서 최대 및 최소값은 점유시간동안 계측



된 변환 자속변환 자료들중 최대값 과 최소값 을 의미하며, 대상자료의 표본수는 점유시간동안 계측한 횟 수를 의미한다. 표준규격의 가로 와 세로는 각각 계측된 주파수 자료와 계측순 위이며, 이들은 1 대 1 대응된다.

이상의 관련 정보가 주어졌을 때 (식 7)에 의하여 정규화가 가능하다.

$$X(i) = (X_{\min}-1) + \frac{i \times (X_{\max}-X_{\min})}{N} \quad (\text{식7})$$

X(i) = 정규화된 수평성분의 값.

X<sub>max</sub> = 표준규격의 수평성분 최대값.

X<sub>min</sub> = 표준 규격의 수평성분 최소값

N = 원시자료의 샘플링 수

i = 원시자료의 샘플링 순서

## 2) 수직성분의 정규화

수직성분의 정규화는 자료변환시 원시 자료의 특성이 왜곡될 수 있는 요소가 다분히 있는 관계로 적용시 매우 신중하여야 한다. 자료의 왜곡은 수직성분의 스케일이 동일 규격화됨으로 인하여 전체 파형의 수직적 증가 혹은 감소된 결과를 초래한다. 이때 닳은 꼴의 파형은 동일 패턴으로 변환되기 때문에 해당 패턴이 가지고 있는 고유한 특성이 왜곡될 수 있는 단점이 있다. 이러한 왜곡은 패턴의 자료 특성상(동일 패턴의 경우 닳은 꼴이지만 그 편차가 상대적으로 큰 경우) 의도적인 왜곡을 필요로 할 수도 있으나 근본적인 문제점은 해결될 수 없다.

반면에 샘플링시 수평영역에서 근접된 영역의 작은 정도의 미묘한 변화는 수직성분의 정규화 과정을 통하여 상쇄시킬 수 있다는 장점이 있다. 특히 어느정도의 비등속도 통행에 따른 자속성 변화의 파피는 이러한 수직성분의 정규화 과정을 통하여 극복될 수 있다. 이러한 수직성분의 정규화는 수직성분의 표준화 규격

을 원시자료보다 적은 스케일로 변환시 효과적이라 할 수 있다. 원시자료의 수직성분보다 표준화된 규격의 수직성분이 큰 경우에는 비록 표준화를 수행하여도 왜곡된 자료의 특성이 해당 패턴의 특성을 좌우할 수 있게 된다. 또한 수직성분의 증감에 따른 미묘한 변화까지 증폭하게 되어 결국 이들 각 요소적인 영향이 왜곡된 자료의 형태로 변환하게 된다. 수직성분에 대하여 정규화하는 기본식은 (식 8)와 같다.

$$Y(i) = (Y_{\min}-1) + \frac{(W_{ave}-W_{\min})(Y_{\max}-Y_{\min})}{(Y_{\max}-Y_{\min})} \quad (\text{식8})$$

Y(i) = 정규화된 수직성분의 값

Y<sub>max</sub> = 표준화 규격의 수직성분 최소값

Y<sub>min</sub> = 표준화 규격의 수직성분 최대값

W<sub>max</sub> = 원시자료의 수직성분 최대값

W<sub>min</sub> = 원시자료의 수직성분 최소값

W<sub>avg</sub> = 원시자료 수직성분의 자료

이러한 정규화 과정에서 수평과 수직성분 개별적인 정규화 과정에서는 큰 문제가 없으나 수평 및 수직 성분을 동시에 정규화를 하여야 할 경우 정규화된 하나의 수평성분에 복수개의 수직성분이 있을 수 있으며, 혹은 하나의 정규화된 수직성분에 복수개의 수평성분이 있을 수 있게 된다. 이러한 문제는 근본적으로 수평 및 수직성분의 자료의 형태(정수형 혹은 실수형)에 따라 발생될 수 있는 사항이다. 수직성분의 자료특성을 대상으로 할 경우에는 수평성분의 자료를 정수화하여 자료처리를 수행하고, 반대로 수평성분의 자료특성을 대상으로 할 경우에는 수직성분을 정수화하여 처리하면 위 문제를 극복할 수 있다.

한편으로 복수개의 문제는 각 성분에 대하여 정규화된 자료의 평균값을 산출하여 수평성분과 수직성분이 1 대 1 대응이 되도록 하였으며,

본 연구에서 적용한 각 모형별 알고리즘은 각각 DOS 환경하의 FORTRAN과 WINDOWS 환경의 Visual Basic 3.0으로 구현하였다.

## 2. 자료의 수집 및 처리

### 1) 자료의 수집

본 연구에서 수집된 자료는 전절에서 규정한 총 9종에 대하여 총 1129개에 대한 기초자료들을 수집하였으며, 이들자료에 대한 수집현황은 다음의 <표 4>과 같다. 수집된 총 1129개의 자료들중 승용차와 봉고의 자료들에 대하여 각각 2개의 씩, 총 4개의 기초자료가 불량자료로 수집되었다. 이러한 불량자료는 현장에서 자료수집시 노트북과의 통신시 발생된 장애로 판단되며, 실제 기초자료상 모두 0의 값을 갖는 자료이다.

<표-4> 자료 수집 현황

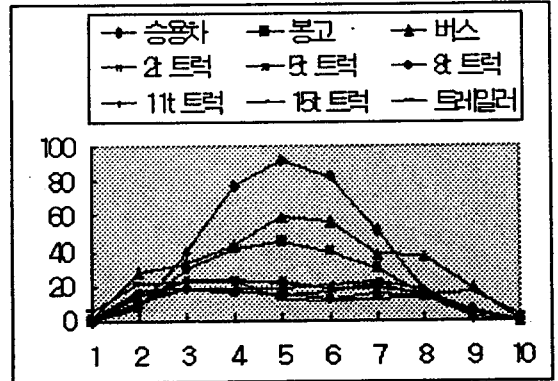
구 분	자료의 수	비 고
승 용 차	250 set	불량 2 set
봉 고	349 set	불량 2 set
버 스	46 set	
2톤 트럭	205 set	
5톤 트럭	83 set	
8톤 트럭	34 set	
11톤 트럭	35 set	
15톤 트럭	109 set	
트레일러	18 set	

(주) 불량은 현장장비로부터 통신오류로 발생된 자료임.

### 2) 자료의 처리

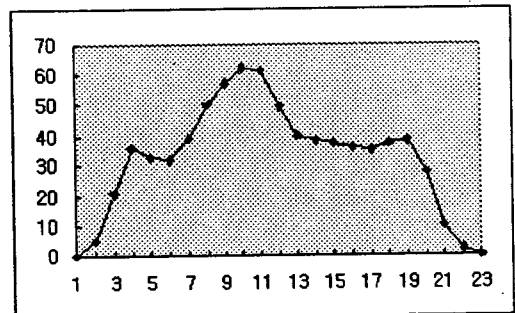
이들 자료들은 2차원의 표현시 수직성분은 인덱터스의 변화이며 수평성분은 샘플링 순서

이다. 본 연구에서의 자료처리는 전절의 식(7)과 식(8)에 의하여 수평성분을 10으로 설정하여 처리하였으며, 정규화된 각 차종들에 대한 파형의 형태는 (그림 9)과 같다.



<그림 10> 9개 차종의 샘플링된 파형

이 수집된 자료들은 이상적인 상태(점유하는 동안 등속도로 통행하는 경우)에서는 그 해당 차종에 대한 고유의 자속변화에 대한 특징이 명확하여 차종을 식별하는데 매우 용이하게 접근할 수 있다. 그러나 실제 상황에서는 (그림 10)와 같이 점유시간중 부분적 지체에 의한 다양한 파형변화가 있을 수 있으며, 파형 변화가 일정하다 하더라도 차량 점유시간의 다양한 변화로 이상적인 상태에서 차량이 갖는 고유의 파형을 파괴하는 요인이 된다.



<그림 11> 비등속에 의한 파형

이러한 현상은 평상시 등속도로 통행되는 고속도로 보다는 통행속도의 완급이 빈번한 도시가로 상에서 자주발생될 수 있는 요인이 된다. 이러한 고유 파형의 파괴는 다음과 같은 상황에서 발생된다.

첫째, 고속도로의 경우 검지기가 설치된 지역에서 사고에 의한 급진적 정체시 혹은 급정거의 경우에 발생될 수 있다.

둘째, 도시가로망의 경우 검지기가 정지선 부근과 상류부(링크의 길이가 300M 기준시 약 250M 부근에 설치된경우)[4]에 설치된 경우 상류부 혼잡 및 정지선에서의 출발손실시간에 따른 속도의 완급변화에 따라 파형의 변화가 발생될 수 있다.

이러한 현상은 도로를 통행하는 차량들의 비등속도통행에 의하여 나타날 수 있는 결과라 할 수 있다. 비등속도 통행에 의한 파형의 변화

는 어느정도의 수준까지는 인식분야에 있어서 효과적으로 활용될 수 있는 소위 전처리과정(pre-processing)들의 다양한 방법들을 토대로 극복할 수는 있으나 근본적인 해결방법은 될 수 없는 요소이다.

본 연구에서는 이러한 비등속도 통행으로 인한 부분적 증가분의 지체요인을 배제한 정상적인 등속도로 고속도로를 통행하는 차량들에 준하여 수집된 자료들을 이용하였다. 또한 하나의 보완장치로서 비등속도 통행에 의한 변화를 어느정도 완화시킬 수 있으며 자료특성의 동질성을 유지시킬 수 있는 방법으로 자속변화 및 점유시간의 변화에 대하여 각각을 정규화(Normalization)하는 전처리 과정을 이용하였다. 자료의 1차처리에 따른 각 9개 차종에 대한 정규화결과에 대한 파형은 다음 <그림 12>과 같다.

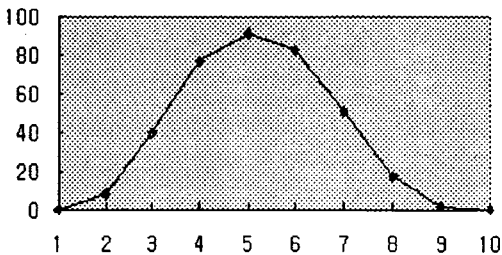


그림 11-a 승용차 (1종)

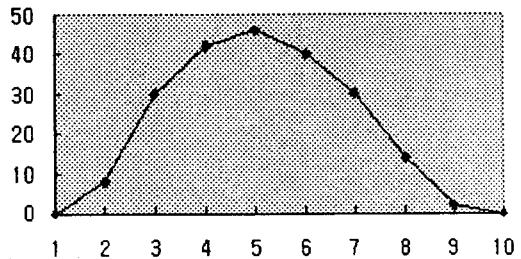


그림 11-b 봉고 (2종)

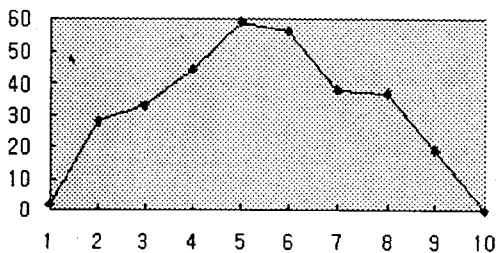


그림 11-c 버스 (3종)

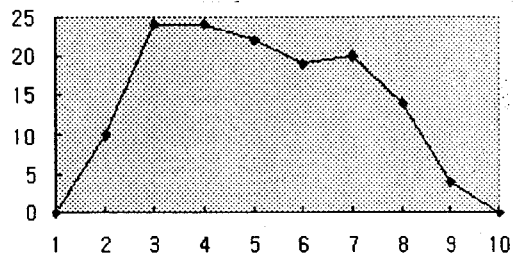


그림 11-d 2t 트럭(4종)

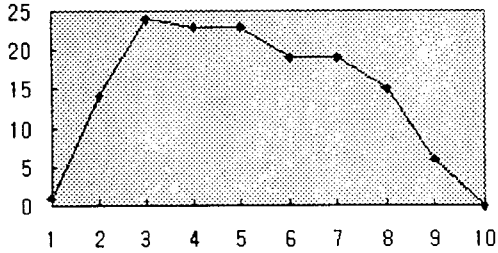


그림 11-e 5t 트럭(5종)

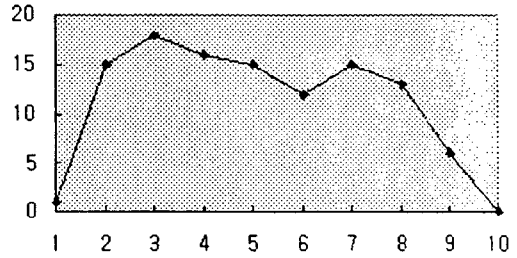


그림 11-f 8t 트럭(6종)

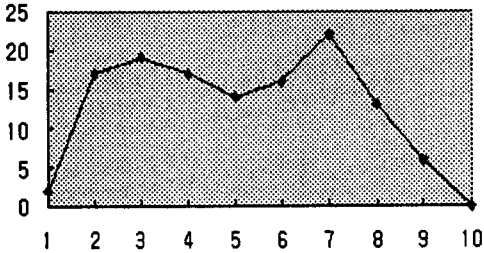


그림 11-g 11t 트럭(7종)

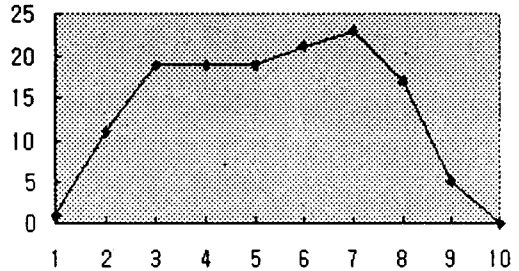


그림 11-h 15t 트럭(8종)

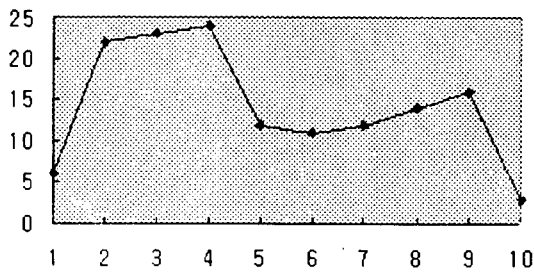


그림 11-i 트레일러(9종)

〈그림 12〉 9개 차종의 정규화된 자료

3. 실험결과 및 비교분석

총 9개 차종에 대한 총 1125개 패턴에 대한

여 통계량 추출방법과 비교사 학습모델인 GLVQ알고리즘을 적용한 차종별 인식 실험결과가 다음 <표 5>와 같다.

〈표-5〉 인식실험결과

구 분	통계량추출방법(1)					
	총자료수	총인식차량수	정인식 차량수	오인식차량수인	식불능수	인식률
1종 (승용차)	251	240	236	4	15	93.3 %
2종 (봉고)	347	299	285	14	62	77.9 %
3종 (버스)	46	60	45	15	1	94.0 %
4종 (2톤트럭)	203	177	126	51	77	49.7 %
5종 (5톤트럭)	83	123	58	65	25	45.3 %
6종 (8톤트럭)	34	78	25	53	9	49.7 %
7종 (11톤트럭)	35	43	27	16	8	64.3 %
8종 (15톤트럭)	108	87	82	5	26	69.1 %
9종 (트레일러)	18	18	18	0	0	100 %
총 계	1125	1125	902	223	223	71.4 %

구 분	통계량추출방법(2)					
	총자료수	총인식차량수	정인식 차량수	오인식차량수	인식불능수	인식률
1종 (승용차)	251	244	241	3	10	95.1 %
2종 (봉고)	347	302	292	10	55	80.8 %
3종 (버스)	46	58	45	13	1	94.3 %
4종 (2톤트럭)	203	171	126	45	77	50.7 %
5종 (5톤트럭)	83	123	58	65	25	45.3 %
6종 (8톤트럭)	34	78	25	53	9	49.0 %
7종 (11톤트럭)	35	43	27	16	8	64.3 %
8종 (15톤트럭)	108	88	83	5	25	70.1 %
9종 (트레일러)	18	18	18	0	0	100 %
총 계	1125	1125	915	210	210	72.1 %

구 분	GLVQ 신경망 모형					
	총자료수	총인식차량수	정인식 차량수	오인식차량수	인식불능수	인식률
1종 (승용차)	251	250	247	3	4	97.0 %
2종 (봉고)	347	352	327	25	20	81.2 %
3종 (버스)	46	46	46	0	0	100 %
4종 (2톤트럭)	203	477	-	-	-	-
5종 (5톤트럭)	83					
6종 (8톤트럭)	34					
7종 (11톤트럭)	35					
8종 (15톤트럭)	108					
9종 (트레일러)	18					
총 계	1125	1125	-	-	-	-

(-) 는 트럭류(4종~9종)들이 하나의 cluster에 속함으로 인하여 각 항목의 산출결과가 의미가 없음을 의미한다.

일반적으로 오인식율이라 함은 대상차종이 아님에도 불구하고 대상차종으로 잘못 인식하는 경우를 의미하는 것으로서 실제 이러한 오인식에 의한 분석을 수행하기 위해서는 수집된 각 차종별 자료의 수가 어느정도 균형을 유지할 수 있어야 효과적인 분석이 될 수 있다. 그러나 본연구에서는 차종별 자료의 수가 최대 317개에서 최소 18개까지 그 편차가 매우 심하

여 오인식에 의한 분석은 오히려 잘못된 해석을 유발할 수 있다.

따라서 본 연구에서는 인식율에 대한 분석상의 오류를 해결하기 위하여 상기의 예와 같이 각 Cluster내에 포함된 각 차종별 비율을 토대로 기대치 값으로 환산하여 인식율을 제시하였다.

이를 위한 관계식은 다음과 같다.

$$\text{인식율} = \frac{\text{정인식비율}(i) \times 100}{\text{정인식비율}(i) \times 100 + \text{오인식비율}(j) \times 100 + \text{인식불능비율}(i) \times 100} \quad (\text{식9})$$

정인식비율(i) = 특정 cluster(i) 에 속한 차종 i / 차종 i의 총 대수

오인식비율(j) = 특정 cluster (i)에 속한 차종 j / 차종 j의 총 대수

인식불능비율(i) = 특정 cluster(j)에 속한 차종 i / 차종 i의 총 대수

상기 <표 5>에 제시된 바와 같이 통계량 추출방법 1.2와 신경망 GLVQ 모두 1종(승용차)에서 부터 3종(버스)에 이르기까지 그 인식율이 81% ~ 100%수준에 이르고 있으며, 이들 승합차량군(1종~3종)에 대해서는 신경망 모형이

매우 우수한 것으로 나타났다. 그의 4종부터 9종은 모두 하나의 cluster에 편중되어 이들에 대한 인식은 전혀 수행되지 못한 결과를 얻었다.

본 연구의 적용모형에 따른 차종인식 및 오인식된 결과가 다음 <표 6>과 같다.

<표 6> 적용 모형별 인식 및 오인식된 결과

· 통계량 추출방법 (1)의 인식 및 오인식결과 (계속)

(차량대수)

군 집 \ 차 종	1 종	2 종	3 종	4 종	5 종	6 종	7 종	8 종	9 종	총 계
C (1)	236 (94.0%)	4 (1.1%)	-	-	-	-	-	-	-	240
C (2)	14 (5.57%)	285 (82.1%)	-	-	-	-	-	-	-	299
C (3)	1 (0.03%)	14 (4.0%)	45 (97.8%)	-	-	-	-	-	-	60
C (4)	-	39 (11.3%)	-	126 (62.1%)	9 (10.9%)	-	-	3 (2.7%)	-	177
C (5)	-	-	-	51 (25.1%)	58 (69.9%)	5 (14.8%)	3 (8.6%)	6 (5.5%)	-	123

· 통계량 추출방법 (1)의 인식 및 오인식결과

군 집 \ 차 종	1 종	2 종	3 종	4 종	5 종	6 종	7 종	8 종	9 종	총 계
C (6)	-	3 (0.8%)	-	25 (12.3%)	14 (16.8%)	25 (73.5%)	4 (11.4%)	7 (6.5%)	-	78
C (7)	-	2 (0.5%)	-	-	1 (1.2%)	3 (8.8%)	27 (77.1%)	10 (9.4%)	-	43
C (8)	-	-	1 (2.2%)	1 (0.5%)	1 (1.2%)	1 (2.9%)	1 (2.9%)	82 (75.9%)	-	87
C (9)	-	-	-	-	-	-	-	-	18 (100%)	18
총 계	251	347	46	203	83	34	35	108	18	1125

· 통계량 추출방법 (2)의 인식 및 오인식결과

(차량대수)

군 집 \ 차 종	1 종	2 종	3 종	4 종	5 종	6 종	7 종	8 종	9 종	총 계
C (1)	241 (96%)	3 (0.9%)	-	-	-	-	-	-	-	244
C (2)	10 (4%)	292 (84.1%)	-	-	-	-	-	-	-	302
C (3)	-	13 (3.7%)	45 (97.8%)	-	-	-	-	-	-	58
C (4)	-	34 (9.8%)	-	126 (62.1%)	9 (10.9%)	-	-	2 (1.8%)	-	171
C (5)	-	-	-	51 (25.1%)	58 (69.9%)	5 (14.8%)	3 (8.6%)	6 (5.5%)	-	123
C (6)	-	3 (0.9%)	-	25 (12.3%)	14 (16.8%)	25 (71.5%)	4 (11.4%)	7 (6.5%)	-	78
C (7)	-	2 (0.6%)	-	-	1 (1.2%)	3 (8.8%)	27 (77.1%)	10 (9.3%)	-	43
C (8)	-	-	1 (2.2%)	1 (0.5%)	1 (1.2%)	1 (2.9%)	1 (2.9%)	83 (76.9%)	-	88
C (9)	-	-	-	-	-	-	-	-	18 (100%)	18
총 계	251	347	46	203	83	34	35	108	18	1125

· GLVQ의 인식 및 오인식결과

(차량대수)

군 집 \ 차 종	1 종	2 종	3 종	4 종	5 종	6 종	7 종	8 종	9 종	총 계
C (1)	56 (22%)	3 (1%)	-	-	-	-	-	-	-	59
C (2)	78 (29%)	-	-	-	-	-	-	-	-	78
C (3)	-	17 (5%)	-	187 (92%)	78 (94%)	34 (100%)	35 (100%)	108 (100%)	18 (100%)	477
C (4)	-	-	14 (30%)	-	-	-	-	-	-	14
C (5)	2 (1%)	181 (52%)	-	16 (8%)	5 (6%)	-	-	-	-	204
C (6)	50 (22%)	-	-	-	-	-	-	-	-	50
C (7)	63 (25%)	-	-	-	-	-	-	-	-	63
C (8)	2 (1%)	146 (42%)	-	-	-	-	-	-	-	148
C (9)	-	-	32 (70%)	-	-	-	-	-	-	32
총 계	251	347	46	203	83	34	35	108	18	1125

· GLVQ-통계량 추출(1) 결합모형의 인식 및 오인식결과

(차량대수)

군 집 \ 차 종	1 종	2 종	3 종	4 종	5 종	6 종	7 종	8 종	9 종	총 계
C (1)	247 (98%)	3 (1%)	-	-	-	-	-	-	-	250
C (2)	4 (2%)	327 (94%)	-	16 (7.9%)	5 (6%)	-	-	-	-	352
C (3)	-	-	46 (100%)	-	-	-	-	-	-	46
C (4)	-	17 (5%)	-	129 (63.51%)	3 (3.6%)	1 (2.9%)	-	-	-	150
C (5)	-	-	-	40 (19.7%)	63 (75.9%)	4 (11.8%)	3 (8.6%)	1 (1%)	-	111
C (6)	-	-	-	17 (8.4%)	11 (13.3%)	25 (73.6%)	4 (11.4%)	7 (6.5%)	-	64
C (7)	-	-	-	-	1 (1.2%)	3 (8.8%)	27 (77.1%)	10 (9.2%)	-	41
C (8)	-	-	-	1 (0.5%)	-	1 (2.9%)	1 (2.9%)	90 (83.3%)	-	93
C (9)	-	-	-	-	-	-	-	-	18 (100%)	18
총 계	251	347	46	203	83	34	35	108	18	1125

(주) 통계량 추출방법의 경우 C(i)의 i는 군집화된 그룹으로 차종과 1:1 대응됨.  
 신경망 GLVQ의 경우 C(i)는 학습률 0.4로 최대 반복횟수 2500으로 학습된 군집임.  
 - : 해당 군집내 해당차종이 없음.



본 연구에서 적용하는 통계량 추출방법(1,2)와 GLVQ 방법에 의한 차종인식결과를 인식된 차량과 오인식된 차량으로 구분하여 나타낸 것으로 본 논문에서는 GLVQ의 경우는 초기 학습률을 0.4, 최대반복횟수를 2500회로 하여 산출된 결과이다.

〈표 6〉에서 가로방향의 차종은 실제 1종차종부터 9종차종의 구분을 의미하며, 세로방향의 C(i)는 각 차종을 인식하기 위한 군집으로 Cluster 된 군집이다. 따라서 통계량 추출방법 1과2의 경우 C(1)은 인식결과 1종 차량은 C(1)에 주로 속하게 되며, 2종차량은 C(2)에 속하게 되어 이들이 각각 1대1 대응된다. 또한 〈표 6〉내에 “( )”의 값은 가로의 각 차종의 총수에 대한 C(i)에 속한 차량의 비율을 의미한다. 따라서 가로의 차종과 세로의 C(i)가 일치하지 않는 경우는 해당차종에 대하여 오인식된 결과이다.

GLVQ는 알고리즘의 특성상 차종별 해당하는 군집(Cluster)이 학습과정이 진행됨에 따라 수시로 전이되므로 통계량 추출방법과 같이 가로의 차종에 대하여 세로의 C(i)가 일치하지 않는다. 따라서 〈표 6〉에서 GLVQ의 경우는 1종에 해당하는 군집은 C(1),C(2),C(6),C(7)이고, 2종에 해당하는 군집은 C(5),C(8)이며, 2종에 해당하는 군집은 C(4),C(9)이다.

GLVQ의 학습결과로 나타난 바와 같이 신경망 모형에 의한 방법은 9개 차종에 대하여 승합차종(승용차,봉고,버스)의 경우 효과적으로 인식이 가능하지만 그의 트럭군(승합차량군을 제외한 6개 차종)에 대해서는 각 클러스터별 속성들이 혼합되어 인식효과가 매우 낮게 나타났다. 이는 정규화 과정에서 산출된 데이터가 각 차종의 특징을 충분히 반영하지 못하였기 때문이다. 특히 트럭군(2t, 5t, 11t, 15t, 8t, trailer)은 파형의 유사성으로 인하여 그 특징을 구분할 수 없기 때문에 거의 하나의 클러스터에 혼합되어

있음을 알 수 있다.

그러나 통계량 추출방법(1,2)의 경우는 GLVQ와는 반대로 승합차량의 경우 인식률이 GLVQ보다는 낮게 나타나고 오히려 GLVQ에서 인식하지 못하는 트럭류(4종~9종)에 대해서는 보다 낮은 결과를 보여주고 있다. 〈표 6〉의 통계량추출방법의 인식 및 오인식결과에 있어서 특이할 만한 사항은 4종과 5종으로 이는 실제 2.5톤과 5톤 트럭을 구분하기 위함이었으나 실험결과로 볼 때 인식율이 크게 낮게 나타났다. 만약 4종과 5종을 통합하여 하나의 차종으로 인식할 경우에는 트럭류의 경우 본 연구의 결과보다 양호한 인식결과를 기대할 수 있을 것으로 판단된다.

따라서 GLVQ 신경망 모형에서 인식하지 못하는 트럭류(4종 ~ 9종)는 통계량추출방법과 결합된 모형에서 최소한 통계량 추출방법에서의 인식수준을 유지할 수 있으므로, GLVQ와 결합되는 통계량추출방법(1)과(2)는 승합차량의 경우(1종~3종)에 대해서 인식율이 약간의 차이가 있으나 이들의 결과가 GLVQ의 결과보다 우수하지 못하고, 트럭류에 대해서는 큰 차이가 없으므로 통계량 추출방법 1)를 GLVQ와 결합시킴으로서 GLVQ의 인식율이 낮은 트럭류에 대한 단점을 보완할 수 있다.

결국 통계량과 결합된 GLVQ모형은 1차적으로 차종을 인식하여 인식결과가 군집 3(표6의 C(3)를 의미함)이 결정되면 2차적으로 대상차종의 자료를 트럭류 4종~9종에 한하여 통계량 추출방법(1)로서 재인식한다.

인식 GLVQ와 통계량추출방법에 의한 결합모형의 실험결과가 다음 〈표7〉에 제시되어 있다.

GLVQ와 통계량추출방법(1)과의 결합모형은 GLVQ에서의 인식불능의 자료군을 통계량추출방법(1)이 보완하는 알고리즘으로 전체적으로

〈표 7〉 GLVQ와 통계량추출방법의 결합모형

구 분	GLVQ-통계량추출방법					
	총자료수	총인식차량수	정인식 차량수	오인식차량수	인식불능수	인식률
1종 (승용차)	251	250	247	3	4	97.0 %
2종 (봉고)	347	352	327	25	20	81.2 %
3종 (버스)	46	46	46	0	0	100 %
4종 (2톤트럭)	203	150	129	21	74	56.9 %
5종 (5톤트럭)	83	111	63	48	20	53.8 %
6종 (8톤트럭)	34	64	25	39	9	52.8 %
7종 (11톤트럭)	35	41	27	14	8	64.7 %
8종 (15톤트럭)	108	93	90	3	18	78.4 %
9종 (트레일러)	18	18	18	0	0	100 %
총 계	1125	1125	972	153	153	76.1 %

약 76.1%의 인식수준을 나타냈다.

인식실험결과 통계량 추출방법(1)의 인식율은 71.4% 이었으며, 통계량추출방법(2)는 72.1%로서 방법(1)보다 약 0.7% 향상된 인식율을 보이고 있다. 그러나 이러한 인식율의 향상은 전 차종에 대한 인식률상승이 아닌 1종~3종까지의 인식율이 상승된 효과이며, 그의 트럭류에 대해서는 미소한 상승이 있었을 뿐이다.

〈표 5〉의 통계량 추출방법(1)의 경우와 〈표 7〉의 각 4종~9종에 대해서 인식율이 달라진 이유는 〈표 7〉의 경우 GLVQ에 의하여 승합차량군(1종~3종)과 4종~9종까지의 트럭군이 인식결과 상호 오인식율이 달라지기 때문이다.

한편으로 GLVQ모형은 트럭류를 전체적으로 4종-9종까지 통합하여 인식할 수는 있으나 이들을 세분화하여 인식하지 못하는 관계로 그 인식률이 31.8%수준이었으며, GLVQ모형과 통계량추출방법과의 복합모형은 GLVQ모형에서 인식하지 못한 트럭류를 세분화하여 인식함으로써 약 74.3%의 인식율을 나타내어 본 연구에서 적용한 4모형중 가장 우수했다.

## V. 결론 및 향후과제

본 연구에서는 도로를 통행하는 차종들을 실시간으로 인식하기 위하여 4개의 적용가능한 방법들을 비교분석하여 그 결과를 제시하였다. 이상의 실험결과를 토대로한 결론은 다음과 같다.

첫째, 본 연구에서 적용된 두 통계량 추출방법은 경쟁적이며 좌우 대칭적 패턴들을 대상으로 효과적인 인식이 어렵다. 따라서 통계량 추출방법은 자료의 전처리과정에서 특징점 추출을 위한 새로운 처리과정이 필수적으로 요구된다.

둘째, 통계량 추출방법(2)는 통계량 추출방법(1)과 비교시 약 0.7%의 인식율 증가 효과가 있었으며, 자료의 특성을 증폭하는데 있어서 방법(1)에서의 단점을 어느정도 극복할 수 있었으나 트럭류(4종~8종)까지의 자료군에 대하여 2차전처리과정에서 명확한 특징점을 추출하는데 여전히 문제점을 안고 있다.

셋째, 신경망 모델인 GLVQ 알고리즘에 의한

방법은 차종별 불균형적인 자료의 수와 알고리즘에서의 정지규칙(stopping rule)의 약화로 인식률이 통계량추출방법 보다는 매우 낮은 결과를 얻으나 1종~3종 그리고 트럭류로서 총 4종으로 구분시 그 인식율은 매우 우수하였다.

넷째, GLVQ신경망 모형과 통계량추출방법(1)과의 결합 모형은 본 연구의 실험결과 트럭류에 대하여 GLVQ가 가지고 있는 인식불능에 대한 단점을 보완하는 기능을 하였으며, 그 인식율은 74.3% 수준이 었다.

다섯째, 차종인식 모형에 있어서 차종의 분류 기준에서 트럭류의 경우 4종과 5종이 매우 유사하여 이에 대한 통합여부와 차량의 적재량에 의한 구분보다는 차량의 전장길이 및 외관형태에 의하여 구분함이 타당할 것으로 판단된다.

여섯째, GLVQ 방법과 GLVQ와 통계량추출방법과의 결합모형은 적용결과 승합차량군과 트럭군의 대분류 효과와 승합차량군들중 승용차,봉고,버스의 인식률이 최소 81% 에서 최대 100% 수준에 이르고 있어 향후 국내 도시가로망과 고속도로에서의 버스전용차선의 운영관리를 위한 무인자동 감시체계구축의 가능성을 보여주고 있다. 특히 버스전용차선에 대하여 오직 버스에 국한할 경우에는 인식율 100%로 버스전용차선의 단속체계를 구현할 수 있을 것으로 기대된다.

이상의 결론으로 향후 다양한 신경망 모델의 응용으로 우수한 차종인식 모델의 구축이 가능하리라 판단되며, 이를 위해서는 전처리과정에 해당되는 특징점 추출에 관한 연구가 필수적으로 요구된다. 또한 GLVQ신경망 모형의 의 단점을 보완하기 위해서는 GLVQ의 학습방법의 개선 및 최근 널리 응용되고 있는 역전파신경망모형(Backpropagation Neural Network)과의 복합모형(Hybrid Model)의 구축에 대한 검토가 이루어질 필요가 있다.

이와 같은 알고리즘을 개발함으로써 다음과 같은 기대효과를 거둘수 있다.

첫째, 기존의 ILD검지기 수집정보의 질적 향상 및 양적 다양화를 도모할 수 있으며, 둘째, 도시가로망의 용량산정에 있어서 기존의 보정계수 산정방식에서 실측에 의한 용량산정을 가능토록 하며, 셋째, 유료도로 및 고속도로 톨게이트 자동 요금징수 체계에서의 운영효율을 향상시키며, 넷째, 기존의 차종식별 교통자료 수집을 위한 고가의 투자비용을 최소화할 수 있으며, 다섯째, 현 신신호시스템에서 운영되고 있는 대기행렬길이 예측알고리즘과 관련하여 보다 신뢰성이 높은 대기행렬 예측 결과 등 정보의 질적 향상을 기대할 수 있는 기초 정보를 제공하고, 여섯째, 향후 순수 국내기술에 의한 이동식 자동 교통정보 수집장치의 개발을 활성화할 수 있다.

## 참고문헌

- [1] 박병규, "신경망 이론에 의한 링크통행시간 예측모형의 개발", 한양대학교 석사학위논문, 1994
- [2] 道路交通安全協會, 서울특별시 交通信號制御시스템 研究用役施行 2次年度 結果報告書, 1993.
- [3] 道路交通安全協會, 서울특별시 交通信號制御시스템 研究用役施行 3次年度 結果報告書, 1993.
- [4] 韓國道路公社, 高速道路交通管制시스템 (FTMS) 構築을 위한 소프트웨어開發 및 시스템 管理 用役 中間報告書(1次), 1993.12.
- [5] 민준영, "腎經網 클러스터링의 性能評價", 成均館大學校 博士學位 論文, 1994.
- [6] 조한선, "단일루프검지기를 이용한 차종분류 알고리즘 개발", 아주대학교 碩士學位

- 論文. 1995, 2.
- [7] M.A.Ismail and M.S.Kamel, "Multidimensional Data Clustering Utilizing Hybrid Search Strategies", *Pattern Recognition*, Vol. 22, No. 1, 1989, pp.75-89.
- [8] James L. McClelland, Daid E. Rumelhart, *Exporations In Parallel Distributed Processing*, 3rd Ed.,1988.
- [9] Laurence Fausett, *Fundamentals of Neural Networks*,Prentice-Hall,1994
- [10] T. Kohonen, *Self-Organization and Associative Memory*, Berlin, Germany: Springer-Verlag, 3rd Ed., 1989.
- [11] T. Huntberger and P. Aimarangsee, *Parallel Self-Organization Feature Maps for Unsupervised Pattern Recognition*, *Int. J. General Systems*, Vol.16, 1989, PP.35 -372.
- [12] Helge Ritter, Thomas Martinets, Klaus Schulten, "Neural Computation and Self-Organizing Maps", Addison-Wesley Publishing Co. Inc., 1992.
- [13] John Hertz, Anders Krogh, Richard G. Palmer, "Introduction to the Theory Neural Computation", Addison-Wesley Publishing Co. Inc., 1991.
- [14] Nikhil R. Pal, James C. Bezdek, Eric C.-K. Tsao, "Generalized Clustering Networks and Kohonen's Self-Organizing Scheme," *IEEE Transactions on Neural Networks*, Vol.4, No.4, July 1993.
- [15] J. J. Reijmers, "On-Line Vehicle Classification". *IEEE Transaction on Vehicular Technology*, Vol VT-29, No.2, May 1980.
- [16] Clayton Stewart, Yi-Chuan Lu and Victor Larson, "A Neural Clustering Approach for High Resolution Radar Trget Classification", *Pattern Recognition*, Vol 4, No.4, April 1994, PP. 503-513.
- [17] Mark Dougherty, "A Review of Neural Networks Applied to Transpot", *Transpn. Res. C*, Vol.3, No.4,1995, pp.247-260
- [18] Richard P. Lippmann, "An Introduction to Computing with Neural Nets", *IEEE ASSP magazine*, APRIL 1987.
- [19] Kang, B.H., D.S. Hwang and J. H. Yoo, "Speeding Up Generalized Learning Vector Quantization", *Proc. of ICONIP (International Conference on Neural Information Processing)*, Oct, 1994.
- [20] Smith, Murry, "Neural Network for Statistical Modeling", Van Nostrand Reinhold, 1993.