

## 차종 시퀀스 패턴을 이용한 구간통행시간 계측

### Measurement of Travel Time Using Sequence Pattern of Vehicles

임 중 선\*      최 경 현\*\*      오 규 삼\*\*\*      박 종 현\*\*\*\*  
(Joong-Seon Lim)   (Gyung-Hyun Choi)   (Kyu-Sam Oh)   (Jong-Hun Park)

#### 요 약

교본 연구는, 구간속도 감지를 위한 기존의 방법인 프로브차량 방식과 차량 번호판 인식 방식의 문제점을 보완할 수 있는 대안으로써, 도로 구간 시.종점에서의 차량 시퀀스 패턴을 이용하여 구간속도 감지가 가능토록 하는 알고리즘을 개발, 제시하였다.

본 알고리즘은 구간 시.종점에서의 차량들을 '차종 순차(Precedence)패턴을 순서대로 나열한 일정한 길이의 시퀀스 그룹'으로 인식하고, 종점에서의 특정 시퀀스에 대응하는, 시점에서의 시퀀스를 탐색하여 가장 유사도가 높은 시퀀스를 동일 그룹으로 간주하여 해당 구간의 통행 시간을 산출하였다. 유사도 비용의 정의에 따라 세 가지의 모델을 제시하였으며, 차량 유출입에 의한 이상치를 제거하고 가공함으로써 정보제공 주기에 가장 적합한 구간 대표 통행시간을 산출할 수 있도록 하였다. 컴퓨터 모의 실험을 통해 구간길이와 통과차량 수를 증가시키면서 차종별, 시.종점의 시퀀스 길이별로 반복 시뮬레이션 한 결과, 평균 최대 오차율 3.46% 로서 현장 적용성에서 뛰어난 가능성을 보였다.

#### Abstract

In this paper, we propose the regional travel time measurement algorithm using the sequence pattern matching to the type of vehicles between the origin of the region and the end of the region, that could be able to overcome the limit of conventional method such as Probe Car Method or AVI Method by License Plate Recognition.

This algorithm recognizes the vehicles as a sequence group with a definite length, and measures the regional travel time by searching the sequence of the origin which is the most highly similar to the sequence of the end. According to the assumption of similarity cost function, there are proposed three types of algorithm, and it will be able to estimate the average travel time that is the most adequate to the information providing period by eliminating the abnormal value caused by inflow and outflow of vehicles. In the result of computer simulation by the length of region, the number of passing cars, the length of sequence, and the average maximum error rate are measured within 3.46%, which means that this algorithm is verified for its superior performance.

**Key words:** Travel time, sequence, pattern matching, similarity cost, vehicle identification

\* 주저자 : 한양대학교 산업공학과 박사과정  
\*\* 공저자 : 한양대학교 산업공학과 교수  
\*\*\* 공저자 : 한양대학교 산업공학과 석사  
\*\*\*\* 공저자 : 영산대학교 교통시스템학과 교수  
† 논문접수일 : 2008년 7월 11일  
† 논문심사일 : 2008년 8월 21일  
† 게재확정일 : 2008년 8월 22일

## I. 서 론

### 1. 연구의 배경 및 목적

교통관리시스템의 최하부 요소인 교통정보 수집 장치는, 루프검지기 및 영상 검지기를 중심으로 집중적인 투자가 이루어져 왔으며, 최근 들어 레이더 방식의 초단파 검지기에 대한 도입이 일부 진행되고 있는 중이다. 이러한 수집 기술들은 기본적으로 지점검지방식으로부터 출발한 것으로서, 지점 자료로부터 구간의 통행 시간을 추정 및 예측하기 위한 술한 연구에도 불구하고, 검지율이라는 문제점 뿐만 아니라 지점에 국한된 자료이기 때문에 해당 링크를 대변하지 못한다는 단점으로 인해 자료의 활용에 다소의 문제점을 안고 있는 실정이다.

최근의 검지기술은, 도시의 연담화 및 광역화와 맞물려 지점검지로부터 공간검지로, 검지율의 일관성 유지를 위해 단일검지체계에서 다양한 매체로부터의 데이터융합으로 발전해 가고 있다. 이러한 검지기술의 발전에 따라, AVI와 같은 도로변 시설물 설치 방식과 함께 도로상을 주행하는 일반 차량들을 프로브차량(Probe Car)으로 활용하기 위한 다양한 방안들이 개발되고 있으며, 그 수단으로는 전통적인 비콘(Beacon) 방식, 경찰청에서 추진중인 광역교통정보수집시스템(UTIS) 방식, 한국도로공사의 통행료 전자 지불 단말기(RF-ETCS)를 활용하는 방식 등이 시험설치 중에 있으며, 아울러 RFID Tag를 부착한 전자번호판 방식 등도 연구 중에 있으나, 대부분의 방식이 막대한 설비 투자를 필요로 하는 경제성의 문제와 개별 차량의 위치를 추적하는 방식에 따른 결측치 증가 및 정체시 정보 수집의 지연이 발생하는 등의 문제를 안고 있다.

본 연구에서는 개별 차량 인식에 따르는 이러한 문제 요인들을 개선하고자, 구간 시.종점에서의 차량들을 하나의 그룹으로 인식하는 새로운 인식 알고리즘을 제시함으로써 향후 구간 검지 관련 연구에 활용하도록 하고자 한다.

### 2. 연구의 범위 및 방법

본 논문은 다음과 같이 구성되어 있다. 2장에서

는 현재 국내에서 운영중인 수집 체계별 장단점을 토대로, 개별 차량 검지가 아닌 차종 시퀀스에 의한 구간검지라는 연구의 착안점을 제시하였다. 3장에서는 본 논문에서 제안하는 시퀀스 패턴 매칭에 의한 시점(Origin) 추정 알고리즘을 설명한다. 시점과 종점에서의 시퀀스의 유사성을 정량적으로 판단하기 위한 3가지 유형의 유사성 비용함수(Similarity Cost Function)를 가정하였으며, 최소 비용의 매칭 시퀀스에 의한 시점 추정을 통해 구간속도 추정 과정을 보여준다. 4장에서는 다양한 경우수별 시물레이션 수행함으로써 최소의 오차를 가지는 유사성 비용함수를 선정하고, 차종 분류의 정도, 시.종점 비교 시퀀스의 최적 길이 등, 시퀀스 패턴 매칭의 확률적 의미를 해석하였다.

## II. 교통정보 수집체계 검토

### 1. 기존 수집 체계 현황

통행시간 계측 또는 예측에 필요한 교통정보는 현장의 검지체계를 통해 수집, 변환된 후 센터로의 전송, 가공 등의 여러 단계를 거치게 되므로, 수집 체계에 대한 검토는 통행시간 연구의 가장 기본적인 부분이라 할 수 있다.

현재 국내의 교통정보 수집체계는 지점검지 방식과 구간검지 방식으로 대별할 수 있으며 각 방식별 검지기 유형은 <표 1>과 같다.

### 2. 검지방식별 일반적 장단점

#### 1) 지점검지 방식

지점검지방식은 지점검지기 자료를 구간 전체의 교통정보로 변환, 보정하여 운전자에게 제공하기 때문에 시시각각 변하는 도로 상의 모든 지점의 교통 특성을 반영할 수 없음에 따라, 지점을 통과하는 차량에 대한 교통정보는 비교적 상세히 제공할 수 있으나 구간 전체에 대해서는 부정확한 정보를 제공할 수도 있다는 한계를 안고 있다 [1].

〈표 1〉 감지 방식별 기존 검지기의 유형  
 <Table 1> Type of sensors by sensing method

형태	검지기	유형	수집 정보	
지점 검지 방식	매 설 형	루프	8각, 32각, 원형	교통량, 점유율, 속도, 점유시간, 차두시간
		지자기	능동식, 수동식	교통량, 점유율, 속도, 점유시간, 차두시간
		WIM	BP식, LC식, MS식	주행차량 무게, 속도, 교통량
	비 매 설 형	초단파	Doppler식, RADAR식	교통량, 점유율, 속도
		적외선	905nm LASER	교통량, 속도, 점유율, 차종, 차두거리
		초음파	Overhead식	교통량, 점유율, 속도, 점유시간
		영상	Loop-Emul.식, Tracking식	교통량, 점유율, 속도, 점유시간, 차종, 차두거리
구간 검지 방식	AVI	AVI Camera	ANPR식	교통량, 점유율, 구간속도
		ETC OBU	RF식	교통량, 점유율, 구간속도
	AVL	Beacon+CRF식		교통량, 점유율, 구간속도
		AVLS OBU+RSE식		교통량, 점유율, 구간속도
		Cell Phone Tracking		교통량, 점유율, 구간속도
		GPS		교통량, 점유율, 구간속도

\* 약어 (Abbreviation)

WIM : Weighing In Motion  
 LC : Load Call  
 AVI : Automatic Vehicle Identification  
 ETC : Electronic Toll Collection  
 CRF : Car Radio Frequency  
 RSE : Road Side Equipment

BP : Bending Plate  
 MS : Multiple Sensor  
 ANPR : Automatic Number Plate Recognition  
 OBU : On Board Unit  
 AVL : Automatic Vehicle Location  
 GPS : Global Positioning System

## 2) 구간검지-AVI 방식

자동 차량번호판 인식(ANPR) 방식에 의한 구간 검지는, 전체 차량 통행에 대해 대규모의 표본 추출이 가능하고, 차량 인식용의 별도 OBU를 필요로 하지 않으며, 번호판 인식의 자동화로 데이터 처리 시간을 획기적으로 감소시키고, 항시 재검토가 가능한 영구적인 기록을 제공하는 등의 장점이 있는 반면, 번호판 인식의 정확도가 날씨 및 대기 환경 변화에 매우 민감하며, 고가의 장비를 사용하는 관계로 경제성이 떨어진다는 단점이 있다.

즉, AVI 카메라와 같은 장비를 사용하여 구간속도 수집 및 통행시간 측정을 하고자 한다면, 측정하고자 하는 구간의 시점과 종점에 고가의 장비가 설치되어야 하며, 도로 네트워크에 적용할 경우에는 네트워크 내의 모든 링크에 장비가 설치되어야 하는 비효율성을 안게 된다 [2].

차량번호판 매칭을 통해 개별 차량을 인식하는 방식에 비해, 차량 정보가 사전에 입력돼 있는 차량용 단말기(OBU)의 정보를 인식하여 이를 프로브차량으로 활용하는 방식이 있으며, 예로서 최근 급속히 보급이 확대되고 있는 전자요금 징수(ETC)용 단말기를 이용하여 교통정보를 수집하는 방안이 검토 중에 있다. 이는 본래 교통정보 수집용으로 개발된 것이 아니므로 단말기의 보급 정도에 따라 정보 수집율이 좌우될 수 있다는 단점이 있다. 그러나, ETC 단말기는 필연적으로 향후 자연 발생적인 증가가 예상되는 만큼, 타 검지체계에 상호 보완 시스템으로 구성할 경우 기존의 인프라를 활용할 수 있다는 큰 장점을 안고 있다.

## 3) 구간검지-AVL 방식

AVL 방식은 프로브차량을 이용한 수집체계의

여러 유형 중 가장 일반적으로 이용되는 기술이라 할 수 있다. 프로브차량 방식은 ITS 시스템에서 실시간으로 데이터를 수집하기 위해 디자인 되었으며, 주요 적용 분야는 운행 시간 수집보다는 실시간 교통 상황 모니터링, 유고 검지, 경로 안내 등에 주로 사용되고 있다 [1, 2].

AVL 방식은 일단 기반 설비가 갖추어지면 그 이후부터 값 싸고 쉽게 지속적으로 정보를 수집할 수 있다. 즉, 정보의 단위당 비용이 저렴하며 정보 수집의 지속성을 보장한다는 장점이 있다.

반면에 AVL 방식은 설비 측면에서 초기 설치에 많은 비용이 필요하며 일단 통신 설비를 설치하고 나면 수집 대상 지역의 크기와 위치를 쉽게 조정하기가 어렵다는 단점이 있으며, 정보 수집 측면에서도 높은 수행비용으로 인해 소규모 정보 수집에 부적절하다. 또한 충분히 많은 프로브 차량이 확보되지 않을 경우 수집정보의 신뢰도가 현격히 떨어지며, 정체시에는 검지 지연이 발생 함에 따라 정체시 정보의 신뢰도가 저하 된다는 단점을 안고 있다.

### III. 시퀀스 패턴 매칭에 의한 구간속도 검지

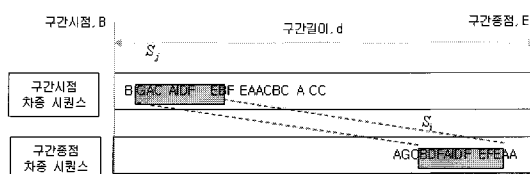
본 장에서는 기존의 개별 차량을 인식하는 방식이 아닌, 구간 시.종점에서의 차량들을 하나의 그룹으로 인식하여 구간 속도를 검지하는 새로운 인식 방안으로서 차종 시퀀스 패턴 매칭 알고리즘을 설명한다.

#### 1. 시퀀스 패턴 매칭 개요

차로 내 차종 순차를 시간 순으로 나열한 전체 시퀀스중 일정 길이의 부분 시퀀스로 잘라낸 부분을 차종시퀀스로 정의한다. 일정 길이의 구간에서 종점의 차종 시퀀스내 순차 패턴과 가장 유사도(Similarity)가 높은 시점의 차종 시퀀스를 탐색함으로써 시퀀스내 개별 차량의 시점을 추정할 수 있다.

<그림 1>은 이 과정을 개념적으로 보여준다.

이 문제는 구간 시점에서의 교통흐름 정보(시점이 j번째로 도착한 차량의 차종, 도착시간, 순간속도



<그림 1> 종점과 시점에서 검지된 시퀀스의 예  
<Fig. 1> Example of sensed sequence in a region

도)와 종점에서의 교통흐름 정보(종점을 i번째로 도착한 차량의 차종, 도착시간, 순간속도)가 주어졌을 때, 구간 종점에서의 T시간 동안의 교통흐름 정보영역(차종시퀀스  $S_j$ )과 최대한 일치하는 구간 시점에서의 교통흐름 정보영역(차종시퀀스  $S_i$ )을 찾는 것이다.

여기서 시퀀스 패턴의 유사성을 탐색하는 과정은, 인간 게놈 프로젝트(Human Genome Project)에서 염색체를 부분부분 절단한 후 운반체에 클로닝(Cloning)하고, 원본과의 대칭(A-G, T-C)을 통해 염색체의 서열을 밝히고(Sequencing), 유전자들에 대한 정보를 얻는(Identification) 작업 과정과 매우 유사한 부분이 있다 [3].

#### 2. 수행 알고리즘

이 알고리즘의 핵심사항은 차종 순차를 이용한 패턴매칭과 이것을 통한 시점 추정이다. 수행 알고리즘은 <표 2>와 같다.

알고리즘을 적용함에 있어서, 처음에는 시점에서 종점까지 차종 순차가 유지될 확률이 가장 큰 연속류 단일 차로로 대상으로 하였으며, 이어서 연속류 복수 차로와 교차로 구간으로 대상 범위를 확대 적용함으로써, 알고리즘의 강건함을 검증하였다.

실제 주어진 시간 창(Time Window)내에서 패턴 매칭 탐색이 이루어 지는 과정은 다음과 같다. 특정한 길이의 종점 시퀀스 i내 m번째 차량에 대해, 시점의 첫 번째 차량부터 시작하는 동일한 길이의 시점 시퀀스 1부터 시점의 마지막 차량 n이 포함된 시퀀스 n-k+1까지 비교하여, 다음 단원에서 정의된 유사도 비용을 계산한다. 이 과정이 끝나면 종점의 다음 시퀀스 i+1에 대해, 시점의 시간 창을 바꿔 위

<표 2> 차종 순차 패턴매칭에 의한 시점 추정 알고리즘  
 <Table 2> Estimation algorithm of origin by pattern matching of vehicle precedence

Step	Process
1	구간의 최소주행시간 ( $t_{min}$ )과 최대주행시간( $t_{max}$ )을 구한다.
2	구간 종점 도착 시간을 기준으로 Time Window - $t_{min} < W < (T + t_{max})$ 사이에 구간 시점을 통과한 차종 시퀀스를 대상으로
3	구간 종점에서 T(5분)시간 동안 통과한 차종 Sequence $S_i$ 과 최적 매칭되는 구간 시점에서의 차종 Sequence $S_j$ 을 찾는다.
4	매칭되는 차량 각각에 대하여 구간 소요시간을 계산한다.
5	구간 소요시간이 비정상적인 out layer를 제거한다.
6	구간 평균 소요시간 및 평균속도를 계산한다.

과정을 반복 수행하게 된다. 그러면 모든 시퀀스에 대한 유사도 비용이 계산되고 그 중 가장 낮은 유사도 비용으로 매칭 되는 종점 시퀀스와 시점 시퀀스를 선택한다.

이 방법은 일종의 전체 검색(Full Enumeration)이라 할 수 있으며, 비용함수는 당연한 최적화의 목표를 신속히 달성할 수 있도록 되도록이면 단순화 시킴이 좋은 만큼, 실제 적용 시에는 종료 조건을 적절히 조정할 수 있다 [4].

### 3. 유사도 측정 및 최적화 방법

#### 1) 유사도 측정 방법

주어진 시퀀스들의 유사한 정도를 표현하기 위해서는 유사성의 정량적인 표현 방법이 필요하다. 여기서는 진화계산 (Evolutionary Computation)에서 흔히 사용하는 비가능 해에 벌금을 부과하는 방식으로, 매칭 유사도 비용(Similarity Cost) 함수 개념을 사용하였다. 즉, 낮은 비용은 큰 유사성(High Similarity)과 같은 개념으로 본다.

실제 비용함수를 사용함에 있어서는, 어떻게 비용함수를 설계하고 이를 비가능 해에 적용하는가(각 제약조건에 어떻게 가중치를 두는가 등)가 가장 중요한 논점이 된다 [5].

본 논문에서의 제약조건은 차량 및 시퀀스의 순서차(Precedence Gap)만 고려하였으며, 차두간격(Headway)을 고려한 순서차의 크기는 고려하지 않았다.

#### 2) 비용함수 모델

본 논문에서는 세 가지 유형의 비용함수를 가정하였으며, 차종 순차의 변화(Match, Change, Insert, Remove)에 적절한 가중치를 주어 최적의 비용함수를 도출하였다. 후에 이에 대한 확률적 의미를 해석할 때, ‘시퀀스 패턴의 길이를 어느 정도로 하는 것이 좋은가’ 하는 점과 ‘차종 분류를 어느 정도 세밀하게 해야 하는가’가 중요 이슈가 될 수 있다. 비용함수를 나타내기 위한 표기는 다음과 같이 정의한다.

$i$  = 대상 종점 시퀀스의 순서

$j$  = 대상 시점 시퀀스의 순서

$k$  = 대상 종점 시퀀스 내 대상 차량의 순서

$l$  = 대상 시점 시퀀스 내 대상 차량의 순서

$D_k$  =  $k$  번째 차량의 차종

$D_l$  =  $l$  번째 차량의 차종

$C_{kl}^{ij}$  =  $i$  시퀀스내  $k$  차량과,  $j$  시퀀스내  $l$  차량의 유사도 비용

$C^{ij} = \sum C_{kl}^{ij}$  =  $i$  번째 시퀀스와  $j$  번째 시퀀스의 전체 유사도 비용

$\alpha, \beta$  = 시퀀스 순서에 대한 가중치

실제 고려된 비용함수들은 아래와 같다. 여기서  $M$ 은 아주 큰 벌칙비용 (Big-M)을 의미한다.

첫번째는, 비교 시퀀스 내에서 동일 차종의 차량 순서차를 이용한 것으로서, 시퀀스내 차량의 추월에 의한 순서 변화만 고려한 것이다. 식 (1)은 직관

적이며 단순하긴 하나, 차량 유출입 및 차로 변경 등에 따른 차종 순차의 차이를 충분히 반영하지는 못하고 있다.

$$C_{kl}^{ij} = \begin{cases} |k-l|, & \text{if } D_k = D_l \\ M, & \text{otherwise} \end{cases} \quad (1)$$

두번째는, 위의 차량 순서차에다가 시퀀스 순서에 대한 가중치를 준 것이다. 즉, 식(1)의 차량 순서 차에 의한 비용에 시퀀스 자체의 순서 차에 대한 벌칙을 부과한 것이다. 가중치는 지수 함수로 주어 시퀀스 순서차가 적을 경우 서서히 유사도 비용이 증가되고, 시퀀스 순서차가 많이 날 경우 유사도 비용이 급격히 증가되도록 하였다.

$$C_{kl}^{ij} = \begin{cases} (|k-l+1|)e^{\alpha|i-j|}, & \text{if } D_k = D_l \\ M, & \text{otherwise} \end{cases} \quad (2)$$

세번째는, 식(2)와 같이 차량 순서차에다가 시퀀스 순서차에 대한 가중치를 주되, 가중치를 제공근 함수로 주어 시퀀스 순서차가 적게 나는 경우에는 유사도 비용을 증가시키고 순서차가 많이 나는 경우에는 상대적으로 낮은 비율로 유사도 비용을 증가시키도록 하였다.

$$C_{kl}^{ij} = \begin{cases} (|k-l+1|)\sqrt{\beta(|i-j|+1)}, & \text{if } D_k = D_l \\ M, & \text{otherwise} \end{cases} \quad (3)$$

### 3) 유사도 비용 계산

비교된 시점 시퀀스의 전체 유사도 비용은 비교 시퀀스 내의 모든 차량에 대한 유사도 비용의 합으로 나타낸다.

즉,  $C^{ij} = \sum C_{kl}^{ij}$  에서 각  $i$ 에 대해

$$ij^* = \arg \text{Min}_j C^{ij} = \text{Argmin}_j C^{ij},$$

$$i^* = \arg \text{Min}_i C^{ij^*} = \text{Argmin}_i C^{ij^*}$$

로 정의하면 유사도 비용은

$$C^{i^*j^*} = \text{Min}_i \left[ \text{Min}_j C^{ij} \right] \quad (4)$$

가 되며, 이때 종점에서의  $i^*$ 번째 시퀀스와 시점

에서의  $j^*$ 번째 시퀀스가 가장 유사도가 높게 된다.

### 4) 이상치 제거 및 구간 통행시간 추정

비용함수 방법에서 비가능해에 가중치에 의한 벌칙을 부과하더라도 이 벌칙이 적절하지 못하면 비가능해가 가능해 보다 더 좋은 평가값을 가질 수 있다. 위 식(4)와 같은 최소화 문제에서는 목적함수 값이 항상 1보다 작도록 정의하고, 순서의 바뀔 때 따른 단위비용을 1로 줄으로써 제약 위반 시는 모두 1보다 큰 값을 갖도록 하여, 가능해가 비가능해에 의해 지배 받지 않도록 하였다.

이에 따라 유사도 비용이 가장 낮은 시퀀스, 다시 말해 유사도가 가장 높게 매치된 시점 시퀀스로부터 구간 통행 속도를 추정하게 된다. 매칭 시퀀스에 대한 각 차량의 통행 속도는 다음과 같다.

$$v_{kl} = \frac{\Delta t}{d} = \frac{t_{ek} - t_{bl}}{d} \quad (5)$$

여기서,  $v_{kl}$  = 추정 속도

$t_{ek}$  =  $k$ 번째 차량의 종점 통과시간

$t_{bl}$  =  $l$ 번째 차량의 시점 통과시간

$d$  = 종점과 시점간의 거리

위 식(5)를 통해 추정된 각 차량별 속도  $v_i$ 의 평균통행시간을 차량대수  $N$ 으로 나눈 조화평균이 매칭된 시종점 간의 구간 평균속도  $\bar{v}_o$ 가 된다 [6].

$$\bar{v}_o = \frac{N}{\sum_{i=1}^N \frac{1}{v_i}}$$

## IV. 시뮬레이션 및 평가

본 장에서는 III장에서 제시된 비용함수들에 대해, 설계속도별, 구간길이별, 통과 차량수별, 차종별, 시.종점의 시퀀스 패턴길이별로 조건을 설정하고 주어진 조건 별로 반복 시뮬레이션하여, 이에 대한 결과를 정리하였다.

### 1. 시뮬레이션 환경

설계속도가 각각 60km/h, 80km/h인 다차로 연속류 구간 도로를 설정한다. 설계속도 60km/h인 도로에서 구간 거리는 1km와 2km의 두 가지로, 설계속도 80km/h인 도로에서 구간 거리는 3km와 5km의 두 가지로 가정하였으며, 각각의 경우에 대해 검지기를 통과하는 차량 수는 50대, 100대, 150대, 200대로 가정하였다.

차종은 자동차관리법에서 정한 5종(승용, 승합중, 승합대, 화물-대, 화물-특수)에 컬러분류 4종(흑, 백, 적, 기타)을 결합하여 분류한다는 가정하에 A부터 J까지 10종류, A부터 O까지 15종류의 두 가지로 구분하였다.

시점을 통과하는 차량 간의 차두간격(시/대)은 지수분포를 따랐으며, 차량이 종점을 지나가는 시각은 시점 통과 시각에 종점까지의 주행시간을 합쳐 구하는데, 이때 차량의 주행속도는 60km/h ±10km/h, 80km/h±20km/h라 가정하고 정규분포에 따라 차량의 95%를 이 속도 범위에서 생성하였다.

전체 차량 시퀀스중 시.종점 비교를 위한 패턴 시퀀스의 적정한 길이를 정함에 있어서, 총 차량수 별로 종점에서의 패턴을 5,7,9,10,15로 정하고, 각 경우 별로 시점에서의 패턴 시퀀스의 길이를 다양하게 설정하였다.

전체적인 시뮬레이션 환경 설정은 <표 3> 및 <표 4>와 같다. 실험은 총 16가지의 유형에 대해, 비교 패턴 길이별 및 차종 분류 정도별로 총 312문제에 대해 수행하였다. 프로그램 환경은 Visual C++, .NET으로 구현하였으며, Pentium4-3GHz, RAM 1G의 PC를 사용하였다.

실험의 정확도를 측정하기 위한 오차율(%)은 다음과 같은 식으로 계산하였다.

$$\text{오차율 (\%)} = \frac{V_c}{V_r} \times 100$$

여기서  $V_c$  = 추정된 주행속도

$V_r$  = 구간내 전차량 실제 평균주행속도

<표 3> 시뮬레이션 환경별 문제 유형  
<Table 3> Types of problem by condition of simulation

설계 속도	60km/h							
	1km				2km			
구간 길이								
차량수	50	100	150	200	50	100	150	200
차종	10, 15							
시,종점 시퀀스 패턴 길이	<표 4 참조>							
문제 유형	P1	P2	P3	P4	P5	P6	P7	P8
설계 속도	80km/h							
	3km				4km			
구간 길이								
차량수	50	100	150	200	50	100	150	200
차종	10, 15							
시,종점 시퀀스 패턴 길이	<표 4 참조>							
문제 유형	P9	P10	P11	P12	P13	P14	P15	P16

<표 4> 총 차량수별 및 종점 패턴 길이별, 시점에서의 비교 패턴 길이

<Table 4> Length of pattern at origin vs end by total # of vehicles

총 차량수 (문제 유형)	종점에서의 비교 패턴 길이 (차량 수)				
	5	7	9	10	15
50 (P1,P5,P9,P13))	5,7,9	7,9,12	9,11,14		
100 (P2,P6,P10,P14)	5,7,9	7,9,12		10,12,15	
150 (P3,P7,P11,P15)	5,7,9	7,9,12		10,12,15	
200 (P4,P8,P12,P16)	5,7,9	7,9,12		10,12,15	15,17,20

### 2. 유사도 비용함수 모델별 비교 및 평가

#### 1) 오차율 비교

<표 5>, <표 6>은 앞 장에서 제시한 비용함수 식 (1), (2), (3) 세 모델에 대한 실험 결과이다. 평균 최소 오차율과 평균 최대 오차율을 볼 때, 식(3)의 최대 오차율이 7.67%로서 가장 근사한 값을보여 주었다. 이는 식(2)의 최대오차율 10.50%에 비해 매우 안정된 값이다.

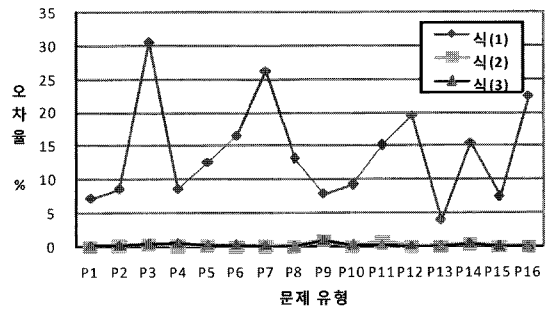
<표 5> 비용함수별 평균 최소 오차율  
<Table 5> The average minimum error rate

문제 유형	평균 최소 오차율(%)		
	식(1)	식(2)	식(3)
P1	2.47	4.05	4.19
P2	8.11	1.94	1.17
P3	17.55	2.14	2.33
P4	13.83	2.09	1.79
P5	6.19	3.50	2.28
P6	20.90	4.32	2.02
P7	18.56	2.83	2.43
P8	13.53	1.77	1.80
P9	23.80	4.50	4.34
P10	21.24	6.82	4.48
P11	14.70	9.78	7.58
P12	12.13	1.33	0.76
P13	8.02	3.38	3.54
P14	17.92	3.60	3.88
P15	29.83	2.41	1.66
P16	14.40	2.85	3.14
평균	15.20	3.46	2.88

식(1)의 오차율이 큰 이유는 시퀀스의 순서차를 고려하지 않아, 순서차는 크지만 동일 차종의 차량 시퀀스가 반복되어 발생하는 경우에 대한 비용을 계산할 수 없기 때문이다.

2) 안정성 비교

각 모델의 안정성을 알아보기 위해 평균 최소 오차율과 평균 최대오차율의 차이를 분석하였으며, 그 결과는 <그림 2>와 같다.



<그림 2> 평균 최소오차율과 평균 최대오차율의 차이  
<Fig. 2> Difference between the average maximum error rate and the average minimum error rate

식(1)은 문제 유형에 따라 높은 오차율 차이를 보이고 있어 안정적이지 못한 반면, 식(2)와 식(3)은 최대와 최소의 차이가 1% 범위 내에서 안정적인 결과를 보여준다. 식(3)은 시퀀스 순서차의 제공근을 이용하여 순서차가 적은 경우부터 시작하여 급격히 유사도 비용을 증가 시킴으로써 패턴의 순서차가 적은 경우의 가중치를 충분히 반영한 결과라고 할 수 있다.

3) 차종 분류 정도에 따른 오차율 비교

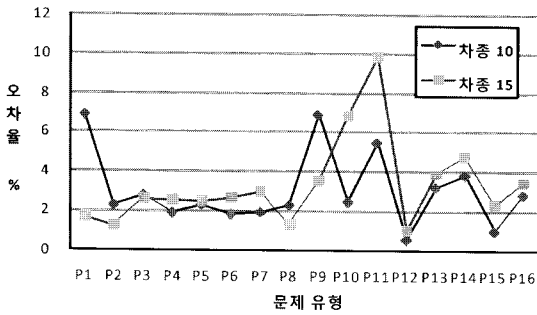
비용함수 설계시 고려한 바와 같이, 차종 분류의 세분화 정도가 본 알고리즘의 현장 적용성 및 비용함수의 확률적 해석에 영향을 미칠 것으로 보아, 가장 안정적인 결과를 보이고 있는 식(3)에 대해 차종 분류 정도에 따른 오차율을 분석하였으며, 그 결과는 <그림 3>과 같다.

차종이 세분화 되어 분포가 분산될수록 적합화

<표 6> 비용함수별 평균 최대 오차율  
<Table 6> The average maximum error rate

문제 유형	평균 최소 오차율(%)		
	식(1)	식(2)	식(3)
P1	9.61	4.06	4.29
P2	16.67	2.00	1.32
P3	48.11	2.42	2.71
P4	22.41	2.13	2.23
P5	18.75	3.67	2.39
P6	37.37	4.32	2.25
P7	44.79	2.91	2.47
P8	26.70	1.79	1.80
P9	31.66	5.35	5.21
P10	30.50	6.82	4.66
P11	29.87	10.50	7.67
P12	31.66	1.34	0.79
P13	11.90	3.38	3.54
P14	33.37	3.95	4.29
P15	37.27	2.42	1.66
P16	36.88	2.85	3.14
평균	29.22	3.61	3.06





<그림 3> 차종분류 정도에 대한 식(3)의 평균최대오차율  
<Fig. 3> Average maximum error rate of model (3) by the number of types of vehicles.

률이 증가할 것으로 예상하였으나, 실제 차종의 분류 정도는 오차율의 차이에 크게 영향을 미치지 않았다. 즉, 차종 분류 정도와 식(3)의 오차율 간에는 큰 상관 관계가 없는 것으로 분석 되었으며, 이는 현재의 차종분류(5종) 정도에 3~4가지의 컬러 정보를 조합하는 정도로 본 알고리즘의 적용이 가능함을 보여준다.

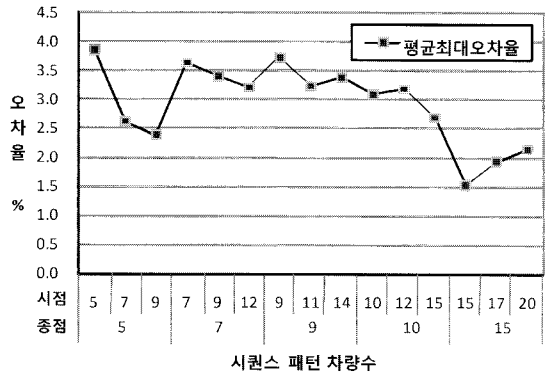
4) 패턴 길이별 오차율 비교

확률적 해석을 통해 비용함수를 최적화함에 있어서 시퀀스 패턴의 적절한 길이를 찾는 것이 또 하나의 중요한 과제이다. 이를 위해 식(3)에 대한, 종점 시퀀스의 길이에 따른 시점 시퀀스의 길이별 최대오차율을 분석하였으며, 그 결과는 <그림 4>와 같다.

종점 비교 시퀀스의 차량 수가 15대, 시점 비교 시퀀스의 차량 수가 15대일 때 가장 정확한 결과를 얻을 수 있음을 알 수 있다.

종점 비교 패턴의 차량 수가 많아진다면 더 좋은 결과를 얻을 수도 있을 것이나, 전체 차량 수가 적은 경우에는 시퀀스 패턴 비교의 의미가 없는 경우도 있을 수 있고, 반대로 전체 차량 수가 매우 많은 경우에는 계산량이 많아져서 실시간성이 떨어질 수 있으므로, 15대 정도를 최적 패턴 길이로 볼 수 있다.

설계 속도별, 구간 길이별, 총 차량 수별, 각 경우에 대한 최적의 시.종점 비교 시퀀스 패턴의 길이(차량 수)는 <표 7>과 같다.



<그림 4> 시.종점 비교 시퀀스 길이별 식(3)의 평균 최대오차율

<Fig. 4> Average maximum error rate of model (3) by the length of sequence

<표 7> 유형별 최적의 비교 시퀀스 길이  
<Table 7> Optimal length of sequence by cases

설계 속도		60km/h			
		1km		2km	
구간 길이					
총 차량 수(대)		150 이하	150 초과	150 이하	150 초과
제한최적패턴 길이(차량수)	종점	10	15	7	15
	시점	15	17	9	17
설계 속도		80 km/h			
		3km		5km	
구간 길이					
총 차량 수(대)		150 이하	150 초과	150 이하	150 초과
제한최적패턴 길이(차량수)	종점	10	10	10	15
	시점	15	10	12	15

5) 전체 검색의 실시간성

전체 검색의 경우 문제의 해를 구하는데 걸리는 계산시간은, 설계속도 60km/h, 총 차량 200대, 종점 비교 패턴 길이 15, 시점 비교 패턴 길이 20인 문제에서 최대 0.156초였다. 평균 계산시간이 0.042초이므로 전체검색을 이용한 유사도 계산에서, 계산시간이 큰 제약 조건이 되지는 않는다.

그러나, 보다 빠르고 효율적인 계산을 위하여, 유사도 비용이 임의의 경계값 이하로 나오면 그 시점에서 계산을 종료하고, 그때의 시퀀스 패턴을 시.종

접간 일치된 것으로 간주하여 통행 시간을 추정하는 방법도 충분한 정확도를 보장할 수 있다고 본다.

## V. 결론 및 추후 연구 과제

본 논문에서는 특정 구간 내 통행 차량들을 각 차량별 차종의 순서를 요소로 하는 일정 길이의 시퀀스 그룹의 연속으로 인식하고, 종점 시퀀스의 차종 순차 패턴에 가장 잘 매치되는 시점 시퀀스를 찾아 낸 후, 두 시퀀스 간의 구간 속도를 검지하는 알고리즘을 구현하고 검증하였다.

본 알고리즘의 컴퓨터 모의시험 결과 평균 최대 오차율 3.46% 로서 현장 적용 가능성을 보임에 따라, 구간속도 검지를 위한 기존의 방법인 AVI나 AVL과 같은 차량의 개별 인식 방식들에 비해, 보다 저렴한 비용의 구간 소통 정보 수집 방안이 될 수 있다고 본다.

본 연구로부터 도출된 결론은 다음과 같다.

첫째, 주어진 시퀀스들의 유사한 정도를 정량적으로 표시하기 위한 유사성 비용 함수는 차종 순서차 및 시퀀스 순서차의 제곱근 함수가 가장 적절하다.

둘째, 시퀀스 패턴의 길이는 15가 가장 적당한 것으로 나타났다.

셋째, 차종의 분류를 보다 더 세밀하게 하여도 오차율은 크게 변화가 없는 것으로 나타났다.

본 연구를 토대로 앞으로 추가적 연구가 필요한 분야는 다음과 같다.

첫째, 본 연구에서는 시뮬레이션 결과를 대상으로 모형의 성능을 검증하였으나, 향후 고속도로 및 도심 구간내 지점검지기 자료를 바탕으로 실험이 이루어 질 예정이며, 이에 따른 기존 구간통행속도 추정 알고리즘과의 비교분석이 수반되어야 할 것이다. 궁극적으로 차량의 개별 인식에 비해, 어떤 흐름의 덩어리(즉, 시퀀스)에 대한 패턴 매칭 방식이 어느 정도 장점을 갖는지가 정량화된다면, 현장 적용 가능성이 보다 더 탄실하게 뒷받침될 수 있을

것이다.

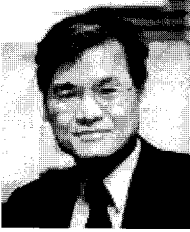
둘째, 실제로 차종을 인식하는 과정과 센서의 유형이 물리적으로 정의되어야 할 것이다. 본 연구에서는 저가형 카메라와 컬러 센서를 결합한 복합 센서를 가정하였다. 또한 시퀀스 내 패턴을 결정하는 요소로서, 차종 외에 시간조건, 차선조건, 진행조건 등을 동시에 고려하면 적합확률이 훨씬 올라갈 것이다.

셋째, 유사도 비용함수의 제약조건에서 제외시켰던 ‘시퀀스내 순서 차의 크기’, 즉 차두간격을 고려한 차간 간격을 비용함수에 포함하고 그 결과를 실제 데이터와 비교해 봄으로써, 알고리즘의 성능 향상을 꾀할 수 있을 것이다. 아울러 향후 복합센서(Sensor Fusion)에 대한 병행연구가 이루어져 본 연구가 구간통행시간 계측 또는 예측 모형 개발의 새로운 계기가 될 것으로 기대해 본다.

## 참고문헌

- [1] T. Turner, W. Eisele, R. Benz, and D. Holdener, *Travel Time data Collection Handbook: FHWA-PL-98-035*, TTI & Texas A&M University System, 1998.
- [2] 임중선, 최경현, *첨단 교통정보 시스템에서의 Probe Car 운영방안에 대한 연구*, 한양대학교, 2004. 12.
- [3] L. Rowen, G. Mahairas, and L. Hood, “Sequencing the human genome,” *Science*, vol. 278, pp. 605-607, Nov. 1997.
- [4] E. H. L. Aarts and J. H. J. Korst, *Simulated Annealing and Boltzmann Machines*, John Wiley & Sons, New York, 1989.
- [5] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, Springer-Verlag, pp. 107-120, 1995.
- [6] A. D. May, *Traffic Flow Fundamentals*, Prentice Hills Inc., 1980.

저자소개



임 중 선 (Lim, Joong-Sun)

1983년 2월 : 서강대학교 물리학과 (학사)  
1999년 2월 : 서강대학교 정보처리학과 (석사)  
2004년 8월 : 한양대학교 산업공학과 (박사 수료)  
1987년 2월~2005년 7월 : (주)LG CNS 수석  
2005년 8월~2006년 12월 : DB정보통신(주) 상무이사  
2007년 2월~2008년 4월 : (주)세인시스템 전무이사  
2008년 5월~현재 : (주)비오비네트웍스 대표이사



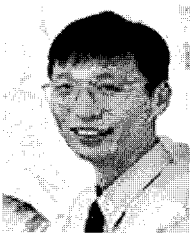
최 경 현 (Choi, Gyung-Hyun)

1980년 : 서강대학교 수학 학사  
1982년 : 서강대학교 대학원 수학 석사  
1989년 : Va. Tech. IEOR 석사  
1994년 : Va. Tech ISE 박사  
1997년~현재 : 한양대학교 시스템응용공학부 교수



오 규 삼 (Oh, Kyu-Sam)

2006년 2월 : 한양대학교 산업공학 학사  
2008년 8월 : 한양대학교 대학원 산업공학 석사  
2008년 8월~현재 : 삼성 SDS



박 중 헌 (Park, Jong-Hun)

1997년 11월~1999년 12월 : (주) LG교통정보 대표이사  
1999년 12월~2003년 9월 : (주) 로티스 대표이사, 연구소장  
2003년 9월~2004년 3월 : 서울시청 교통정보반 과장  
2004년 3월~현재 : 영산대학교 교통시스템학과 교수