

고속도로 교통자료 처리기법 통합평가 시스템 개발

An Evaluation System For Freeway Traffic Data Processing Techniques

오 동 옥* 오 철** 남 궁 성*** 전 세 길****
(Dong-Wook Oh) (Cheol Oh) (Sung Namkoong) (Se-Kil Jeon)

요 약

지능형 교통체계(ITS)는 검지시스템의 지원으로 도로 현장에서 전개되는 교통상황을 실시간 모니터링 한다. 또한 이렇게 수집된 교통자료는 다양한 처리 및 가공 기법을 적용하여 교통운영, 교통계획, 교통안전 등 교통분야에서 다양하게 추가 활용된다. 그러나 실시간으로 검지되는 교통자료에 대한 활용 및 신뢰성이 선진 외국에 비해 낮은 실정으로 신뢰성 있는 교통정보 제공 및 분석을 위하여 교통자료의 품질확보가 우선 되어야 한다. 품질확보를 위한 기존의 연구에서는 기본 자료처리 3단계인 이상치 검지 및 제거, 결측자료 보정, 평활화 각각에 대한 알고리즘 및 관련 파라미터 추정에 초점이 맞추어져 왔다. 그러나 각각의 자료처리 과정은 사용자의 궁극적인 자료 활용목적에 따라 다양한 상호연관성을 나타낸다. 따라서 사용자에 보다 신뢰성 높은 교통자료를 제공하기 위해서는 자료처리 기법을 통합평가할 수 있는 방법론 및 도구가 필요하다. 본 연구에서는 차량 검지기로부터 수집되는 교통자료의 통합평가 방법론을 연구하였고, 사용자의 자료 활용 목적에 부합하는 교통자료를 손쉽게 추출할 수 있는 도구를 개발하였다. 교통자료 품질확보를 위하여 자료처리 방법선택, 입력파라미터 선택, 자료처리 알고리즘 등록이 가능하도록 하였으며 품질확보를 위하여 반복적 자료처리가 가능하도록 하였다. 또한, 실제 경부선 2개 구간에서 수집된 교통자료를 본 연구에서 개발한 방법론 및 도구에 적용한 사례를 제시하였다.

Abstract

Real-time traffic data are readily obtainable by traffic surveillance systems of intelligent transportation systems (ITS). Such data greatly support further applications in the field of traffic operations, planning, and safety. However, traffic data should be appropriately processed to fully exploit the benefits of data collection capability. Rather than developing individual data processing techniques, which is major concern of existing studies, this study proposes a novel methodology for evaluating data processing techniques in an integrated manner. Also, a tool for implementing the proposed methodology is developed. Users can extract useful and more reliable traffic data based upon their ultimate purpose of data usage by the evaluation tool developed in this study. Actual freeway traffic data are, as an example, fed into the evaluation tool, and results are discussed.

Key words: Traffic data, filtering, imputation, smoothing, data quality

† 본 연구는 한국도로공사 도로교통연구원 교통정보 이력자료 데이터베이스 설계 및 구축 연구과제의 지원으로 수행하였습니다.

* 주저자 : 한양대학교 교통시스템공학과 석사과정

** 공저자 : 한양대학교 교통시스템공학과 교수

*** 공저자 : 한국도로공사 도로교통연구원 수석연구원

**** 공저자 : 한국도로공사 도로교통연구원 박사후연수자

† 논문접수일 : 2008년 6월 23일

† 논문심사일 : 2008년 6월 30일

† 게재확정일 : 2008년 8월 21일

I. 서론

지능형 교통체계(ITS: Intelligent Transportation Systems)는 기존의 교통에 정보통신, 전기 및 전자, 기계, 제어 등의 기술을 접목시켜 도로 및 교통체계의 안전 및 운영효율을 극대화하여 교통문제를 해결하는 종합적인 솔루션이다. 지능형 교통체계가 기존의 교통체계와 차별화될 수 있는 가장 핵심적인 주제 중의 하나는 실시간 교통모니터링이라고 할 수 있다. 이는 현재 도로 현장에서 전개되는 교통상황을 실시간으로 파악할 수 있는 기술 환경을 의미하는데, 다양한 검지체계의 구축으로 가능하다. 이러한 지능형 교통체계를 위해서는 도로상의 차량 흐름에 대한 자료수집이 필수적이며, 이를 위하여 ITS 구축시 도로상에 교통자료를 수집할 수 있는 많은 차량검지기가 설치된다. 도로상에 설치된 차량검지기는 교통량, 속도, 점유율 등의 교통자료를 수집하여 통신선을 통해 교통관리센터로 전송하고 교통관리센터에서는 수집된 자료를 적절한 알고리즘으로 가공하여 실시간 교통제어나 도로이용자에게 유용한 교통정보를 제공하고 있다. 그러나 현재 수집되고 있는 교통자료는 해외 우수사례에 비해 신뢰성이 낮은 실정이다. 그러므로 다양한 추가활용을 위하여 우선적으로 차량검지기 자료의 품질 확보가 절실하다. 이를 위해 본 연구에서는 다양한 교통자료 처리기법을 통합적으로 평가하여 신뢰성 높은 자료의 취득을 지원할 수 있는 자료처리 통합평가 시스템을 개발하였다.

II. 관련연구 및 현황

1. 국내외 자료처리 사례검토

한국도로공사 고속도로 관리시스템(FTMS)에서는 도로에 설치된 루프검지기를 이용하여 교통량, 속도, 점유율 및 차량길이를 검지하여 30초 주기로 집계되며, 전국 각 지역에서 수집된 차로별 원시데

이터를 가공하고 있다. 고속도로 관리시스템에서는 오류판단 과정만 거치며 결측보정과 평활화 알고리즘은 운영되지 않고 있다. 고속국도 우회도로 교통정보시스템(ARTIS)에서는 크게 원시자료 수집, 오류판단, 결측보정, 평활화를 모두 실시하고 있으며, 원시자료 중 오류자료와 누락자료는 문제처리 후 DB에 저장한다. 또한 누락이나 오류자료는 아니지만 비정상적인 교통패턴을 보여주는 자료는 문제된 자료와 함께 그 이력을 기록하여 별도 관리를 하고 있다. 결측처리는 해당지점의 데이터는 공간적으로 참고지점과 시간적으로 교통상태가 유사하다는 전제하에 시행하고 있다. 결측처리 방법으로는 우선 공간적 추세 활용법으로 해당지점과 비슷한 교통특성을 보이는 지점의 데이터를 활용하여 누락자료를 보정하고 매 지점마다 인위적으로 우선순위가 부여된 n 개의 참고지점을 선정하여 시행하고 있다. 시간적 추세활용법은 공간적 추세활용법이 불가능한 경우에 시행하며 이전자료 이용법과 이동평균법이 있다. 다음으로 공간적, 시간적 추세활용법이 불가능한 경우 이력자료 활용법을 시행하고 있다. 이력자료 활용법은 데이터베이스에 저장된 과거의 요일별, 시간별 이력자료를 통해 보정하는 방법이다. 평활화 과정은 자료의 왜곡현상이 심각하고 원시자료의 변화폭이 심하여 2004년 1월 이후로 평활화 계수 k 의 값을 1로 설정하여 평활화 과정을 미운영하고 있다[1].

국의 사례로 미국의 캘리포니아 PeMS[2]는 매 30초마다 고속도로의 루프검지기 자료를 수집한다. 자료 체크와 진단 시스템이 존재하여 웹페이지 내에서 PeMS 자료의 품질에 대한 평가 결과를 조회할 수 있다. 또한 자료에 대한 품질과 검지기 작동률 및 검지된 검지기 에러비율을 보여준다. 버지니아 ADMS[3] 교통관리 시스템에서는 오류판단과정에서 검지기 단위의 평가와 지점 단위의 평가가 동시에 이루어지며, 오류에 따른 별도의 관리가 이루어지고 있다. 결측보정 방법으로는 실시간 결측보정시 과거 이력자료를 사용하고 결측된 값은 과거 패턴에 근거하여 평균값을 사용한다.

2. 자료처리 알고리즘 연구

자료처리와 관련한 연구로는 크게 자료처리에 관련한 알고리즘 연구와 차량검지기 자료의 품질평가 연구로 구분할 수 있다. 국내 자료처리 알고리즘에 관한 연구에서는 ITS 시스템의 결측자료 보정에 관한 연구에서 한대철[4]은 교통량과 속도 자료의 보정에 관하여 각각 다른 결측보정 방법을 제안하였는데 교통량 자료는 평일, 토요일, 일요일에 대해서 유사지점의 데이터를 이용한 자기회귀 보정을 적용하고, 속도 자료는 직전 5주의 평일 자료를 모두 포함한 산술평균 보정방법을 제안하였다. 그리고 차량 검지자료 결측보정에 관한 연구에서 김정연[5]은 결측자료에 대해 인접지점 참조방식과 이력자료를 활용한 방법론을 적용한 보정처리를 적용하여 인접지점 참조방식에 비해 이력자료를 활용한 보정처리방법이 원데이터에 가까운 값을 가지게 된다는 결과를 제시하였다. 국외 자료처리 알고리즘에 관한 연구로는 오류판단에 관해 Al-deek[6]은 수집된 자료의 품질향상을 위해 루프검지기 자료에 엔트로피 개념을 적용한 자료 필터링 알고리즘을 제시하였고, Wall and Dailey[7]는 correction factor를 제안하여 검지된 교통량 자료를 보정하는 방법론을 제시한 바 있다. 오류 보정 및 결측치 생성에 관해서는 Turner[8]은 오류자료를 검지하는 방법론과 함께 이상치, 결측치, 자료의 정확도를 측정하는 기법을 제시하였다. Zhong[9]은 인공신경망을 이용한 결측치 생성모형을 개발하고 기존의 기법과의 비교평가를 수행하였다. 다양한 집계간격을 이용한 통계적인 자료 필터링 기법도 Schmoyer[10]에 의해 개발되었다. 이력자료를 활용한 예로서 Bertini와 El-Geneidy[11]는 대중교통 서비스의 효과평가를 위한 지표를 개발하여 제시하기도 하였다.

차량검지기 자료의 품질평가와 관련된 연구로는 차량검지기 성능평가 방법론 연구에서 장진환[12]은 교통제어 및 교통정보 제공을 위한 교통자료 정확도의 중요성을 평가하기 위한 방법론으로 평가항목 선정 및 방법을 제안하였으며, 평가방법으로 부등계수, 편기비율, 분산비율, 공분산비율, MAPE

(Mean Absolute Percent Error), 퍼센트오차, 상관계수, RMSE(Root Mean Square Error), APE(Absolute Percent Error)를 사용하여 분석한 결과 MAPE가 사용자의 이해와 성능평가에서 가장 적절하다고 평가하였고, 수집주기가 개별차량일 경우 퍼센트 오차가 적절하다고 평가하였다. Loop식 차량검지기 신뢰성 평가를 위한 모의시험 장치 개발연구에서 이정준[13]은 Loop식 차량검지장치의 Inductance Wave에 대한 충실한 재현이 가능한 특징을 갖는 새로운 평가장비의 시험 개발을 평가하였다.

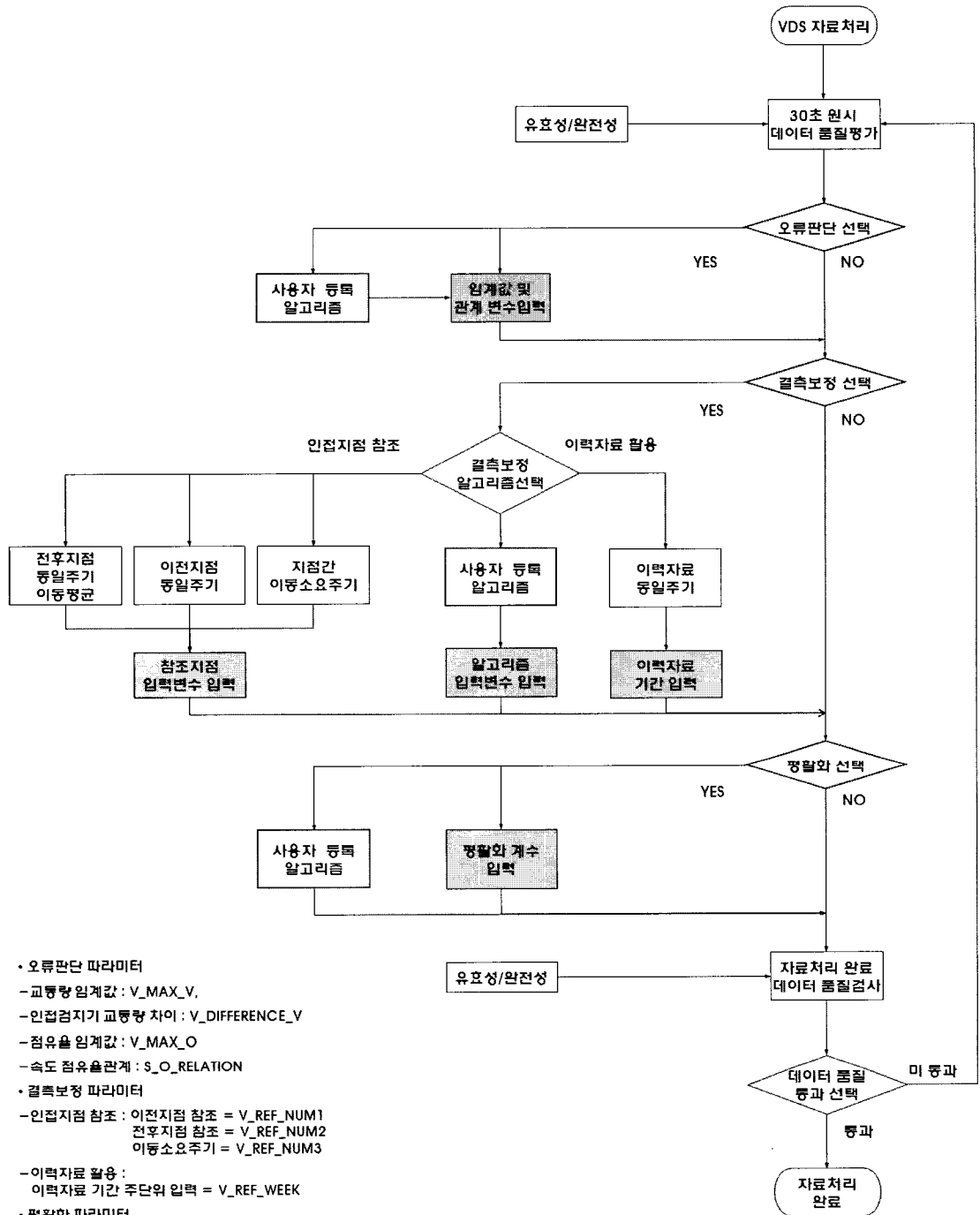
기존의 자료처리 관련 연구는 개별 자료처리 과정을 통한 자료품질의 향상을 평가하는데 중점을 두고 있다. 본 연구에서는 자료처리 단계별로 적용되는 기법을 통합적으로 평가하여 교통자료의 품질을 체계적으로 개선할 수 있는 분석방법론을 제안하였다. 이를 통해 적용 알고리즘의 업그레이드 및 신규개발이 용이하도록 하였다.

III. 자료처리 분석 방법론

교통자료의 품질을 향상시키기 위한 자료처리 분석 방법론은 <그림 1>과 같이 나타낼 수 있다. 첫 번째로 30초 간격으로 수집되는 원시데이터의 품질을 검사 후 오류판단 검사를 선택한다. 이 단계에서 오류판단 조건의 임계값검사와 관계검사 파라미터를 입력하게 된다. 두 번째로 결측보정을 여부를 선택하고 결측보정을 수행할 알고리즘과 그에 따른 파라미터를 입력한다. 자료처리의 마지막 단계로 평활화를 선택한 후 평활화 계수값을 입력한다. 자료처리 과정을 거치면서 입력하게 되는 파라미터를 정리하면 <표 1>과 같다.

오류판단, 결측보정, 평활화를 수행한 후 자료처리가 완료된 데이터에 대해서는 다시 품질검사를 실시하여 자료 품질의 개선 여부를 확인한다.

자료처리 과정에서 사용되는 알고리즘은 기존 연구결과 제안된 방법을 적용하였으며, 또한 새로 등록된 알고리즘이나 이미 등록된 알고리즘간의 순서 및 파라미터를 변화시켜 자료처리할 수 있도록 하였다.



- 오류판단 파라미터
- 교통량 임계값 : V_MAX_V,
- 인접검지기 교통량 차이 : V_DIFFERENCE_V
- 점유율 임계값 : V_MAX_O
- 속도 점유율관계 : S_O_RELATION
- 결측보정 파라미터
- 인접지점 참조 : 이전지점 참조 = V_REF_NUM1
전후지점 참조 = V_REF_NUM2
이동소요주기 = V_REF_NUM3
- 이력자료 활용 : 이력자료 기간 주단위 입력 = V_REF_WEEK
- 평활화 파라미터
- 평활화 계수 : V_K

〈그림 1〉 자료처리 분석 방법
 <Fig. 1> Proposed evaluation framework

<표 1> 자료처리 입력 파라미터
<Table 1> Input parameters

	입력파라미터	내용
오류 판단	V_MAX_V	교통량 임계값
	V_DIFFERENCE_V	쌍루프간 교통량 차이 임계값
	V_MAX_O	점유율 임계값
	S_O_RELATION	속도-점유율관계 파라미터
결측 보정	V_REF_NUM1	이전지점 참조 최대지점값
	V_REF_NUM2	전후지점 참조 최대지점값
	V_REF_NUM3	이동소요주기 참조 최대지점값
	V_REF_WEEK	이력자료기간 주단위값
평활화	V_K	평활화 계수값

1. 품질평가

품질평가 방법으로는 완전성과 유효성을 두 가지 기준으로 원시데이터와 자료처리 완료 후 두 번에 걸쳐 평가된다. 완전성은 30초 동안 루프별로 수집된 전체 원시자료 중에서 결측되지 않은 자료의 비율을 의미하며, 완전성이 높다는 것은 이용 가능한 자료의 비율이 높다는 것을 의미한다. 완전성의 평가방법은 식(1)과 같이 이루어진다.

$$\text{완전성 (\%)} = \frac{x}{X} \times 100 \quad (1)$$

여기서, $X = 30$ 초 동안 루프별 전체 원시자료
 $x = 30$ 초 동안 루프별 관측 원시자료 중 결측되지 않은 자료

유효성은 30초 동안 루프별 수집된 이용 가능한 자료 중에서 오류판단 기준에 의거하여 오류가 없는 자료의 비율을 말하며 평가방법은 식(2)와 같이 이루어진다.

$$\text{유효성 (\%)} = \frac{\bar{x}}{x} \times 100 \quad (2)$$

여기서, $\bar{x} = 30$ 초 동안 루프별 관측 원시자료 중 결측되지 않은 자료 중 오류가 없는 유효한 자료

<표 2> 오류판단 기준
<Table 2> Decision criteria for outlier detection

구분	항목	오류판단기준
임계값 검사	교통량	$V < 0$ or $V > 30 \rightarrow V, O$ (BAD)
		인접 검지기와의 차 $> 2 \rightarrow$ (BAD)
	점유율	$O < 0$ or $O > 100 \rightarrow$ (BAD)
관계 검사	속도-점유율비	$2.7V \leq \frac{S \times O}{12} \leq 18V$

$x = 30$ 초 동안 루프별 관측 원시자료 중 결측되지 않은 자료

2. 오류판단

오류판단 기준은 기본적으로 <표 2>와 같이 이루어진다. 오류판단 기준에서 오류판단 임계값검사와 관계검사의 입력변수를 파라미터화하여 사용자에 따라 다른 값을 적용할 수 있도록 하였다. 교통량 임계값에서는 교통량의 최대값과 인접 검지기와의 교통량 차이값을 파라미터로 하였고, 점유율에서는 점유율 값의 최대값으로 하였다. 관계검사에서는 속도-점유율비에서 차량길이 이용법을 적용하였고, 위 식에 들어가는 상수항 값을 파라미터화하여 변화시킬 수 있도록 하였다. 여기서 관계검사는 교통량, 속도, 점유율 값을 이용하여 산출한 차량길이의 최대값 및 최소값을 산정하여 오류판단을 실시한다.

3. 결측보정

결측보정 방법은 2가지를 적용하였으며, 유출입이 없는 폐쇄구간내 연속된 지점상의 교통 패턴이 유사하다는 전제조건을 가진다. 첫 번째 방법은 정상류 상태에서 결측발생시 인접지점을 참조하여 보정하는 인접지점 참조 방법이다. 두 번째 방법은 인접지점이 모두 결측되어 보정처리가 불가능할 경우 사전에 구축된 이력자료를 활용하는 방법이다. 인접지점 참조방법에는 전후지점 동일주기 이동평균방법, 이전지점 동일주기방법, 지점간 이동소요주기방법을 적용하였고, 여기서 참조할 지점

의 최대위치값을 파라미터로 입력하여 결측 발생 시 결측된 지점으로부터 가장 가까운 인접 검지기를 참조하며, 참조할 지점이 결측이 되었을 때는 최대위치부터 순차적으로 참조 지점을 찾아 결측을 보정한다. 이력자료 활용방법은 이력자료 동일 주기 적용방법을 적용하였으며, 여기서 사용할 이력자료 기간을 주단위로 파라미터화하여 설정한 이력자료기간의 값을 평균하여 결측보정을 실시하게 된다. 인접지점 참조방법에 사용되는 알고리즘은 식(3)~식(5)와 같이 나타낼 수 있다.

1) 전후지점 동일주기 이동평균 방법

결측 발생시 전후 인접 검지기 자료의 평균으로 보정처리하는 방법 식(3)으로 참조지점이 결측시 최대지점으로 설정된 위치까지 참조를 수행한다.

$$\left\{ \begin{aligned} \widehat{V}_{j,l,t,d} &= \frac{V_{i,l,t,d} + V_{k,l,t,d}}{2} \\ \widehat{O}_{j,l,t,d} &= \frac{O_{i,l,t,d} + O_{k,l,t,d}}{2} \\ \widehat{S}_{j,l,t,d} &= \frac{S_{i,l,t,d} + S_{k,l,t,d}}{2} \end{aligned} \right. \quad (3)$$

여기서,

- $\widehat{V}_{j,l,t,d}$: 지점 j , 차로 l , 시간 t , 날짜 d 의 교통량 추정치
- $\widehat{O}_{j,l,t,d}$: 지점 j , 차로 l , 시간 t , 날짜 d 의 점유율 추정치
- $\widehat{S}_{j,l,t,d}$: 지점 j , 차로 l , 시간 t , 날짜 d 의 속도 추정치
- $V_{i,l,t,d}$: 지점 j , 차로 l , 시간 t , 날짜 d 의 교통량 관측치
- $O_{i,l,t,d}$: 지점 j , 차로 l , 시간 t , 날짜 d 의 점유율 관측치
- $S_{i,l,t,d}$: 지점 j , 차로 l , 시간 t , 날짜 d 의 속도 관측치

2) 이전지점 동일주기방법

결측 발생시 상류부 인접검지기 자료의 동일주기로 보정처리 하는 방법 식(4)으로 위와 마찬가지로 참조지점이 결측시 최대지점으로 설정된 위치까지 참조를 수행한다.

$$\left\{ \begin{aligned} \widehat{V}_{j,l,t,d} &= V_{i,l,t,d} \\ \widehat{O}_{j,l,t,d} &= O_{i,l,t,d} \\ \widehat{S}_{j,l,t,d} &= S_{i,l,t,d} \end{aligned} \right. \quad (4)$$

3) 지점간 이동 소요주기 적용방법

속도자료와 지점간의 거리를 바탕으로 이전 상류지점에서의 해당주기를 찾아 결측자료를 보정하는 방법으로 식(6)에 의해 구해진 값을 이용하여 식(5)로 보정처리 된다.

$$\left\{ \begin{aligned} \widehat{V}_{j,l,t,d} &= V_{i,l,(t-n),d} \\ \widehat{O}_{j,l,t,d} &= O_{i,l,(t-n),d} \\ \widehat{S}_{j,l,t,d} &= S_{i,l,(t-n),d} \end{aligned} \right. \quad (5)$$

여기서,

$$n = \frac{L}{\left[S_{j,l,(t-1),d} \times \frac{1000}{3600} \right]} \times timeinterval \quad (6)$$

$n \geq 0$

L : 지점간 거리(m) (S의 단위가 km/h라 가정)

n : 소요시간, 수집주기 단위로 환산

$timeinterval$: 수집주기, 30초

4. 평활화

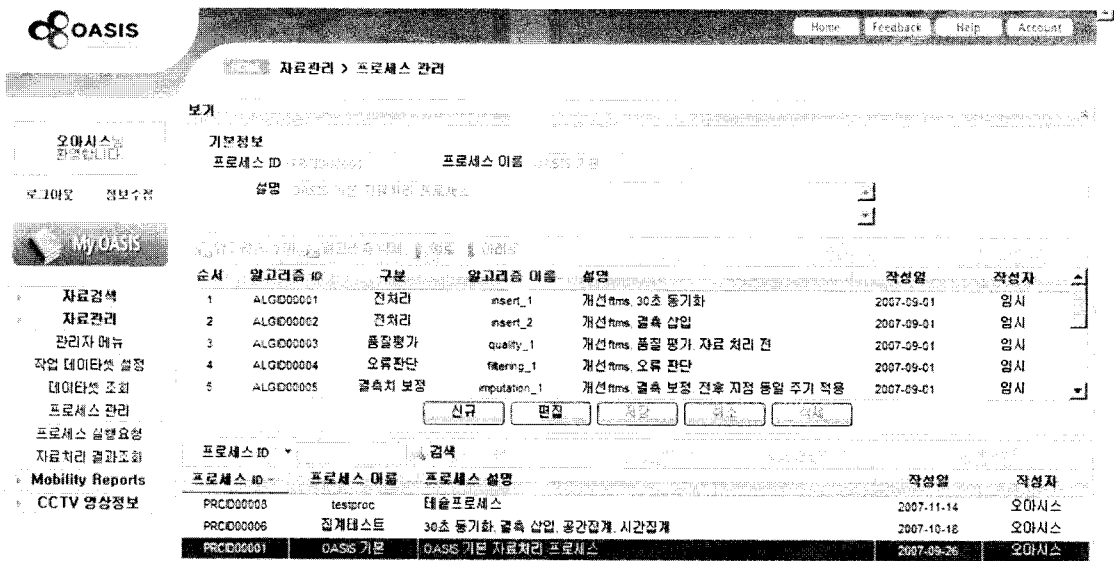
평활화는 보정이 수행된 주기 단위 시간 자료에 대해서 최근 수집된 자료에서부터 과거에 수집된 자료까지 새로운 자료에 대해 지수적으로 가중치를 부여하는 방법인 지수이동평활법을 적용하였으며, 여기서의 평활화 계수값 K 를 파라미터화하여 0~1사이값을 선택할 수 있도록 하였다.

IV. 교통자료 처리기법 분석도구 개발

1. 사용자 인터페이스

본 연구에서 제안한 자료처리를 위한 사용자 인터페이스는 자료처리된 데이터를 저장하기위한 Data set 공간 관리와 자료처리 프로세스 관리화면, 자료처리 실행요청, 자료처리 결과를 조회하는 화면으로 구성된다.

Data set 관리는 Data set을 생성할 때, 해당 Data set 시공간적 범위인 대상기간과 대상구간에 대해



<그림 2> 사용자 인터페이스중 프로세스 관리화면 예
 <Fig. 2> A snapshot of graphical user interface

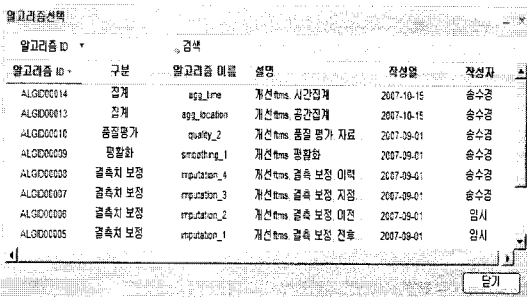
설정할 수 있다. 대상구간은 노선단위로 가능하며, 시작지점과 종료지점은 해당 노선에 설치된 검지기(LDS(Local Detection System) 아이디이다. 대상구간을 설정할 때는 먼저 노선을 설정하고, 시작지점과 종료지점에 해당하는 LDS 아이디를 설정한다.

대상구간을 설정할 때는 상행방향 기준으로 설정해야 한다. LDS는 상하행 지점자료를 모두 포함하기 때문에 상행기준으로만 설정하면 상하행 자료에 대해서 모두 자료처리가 이루어질 수 있도록 하였다.

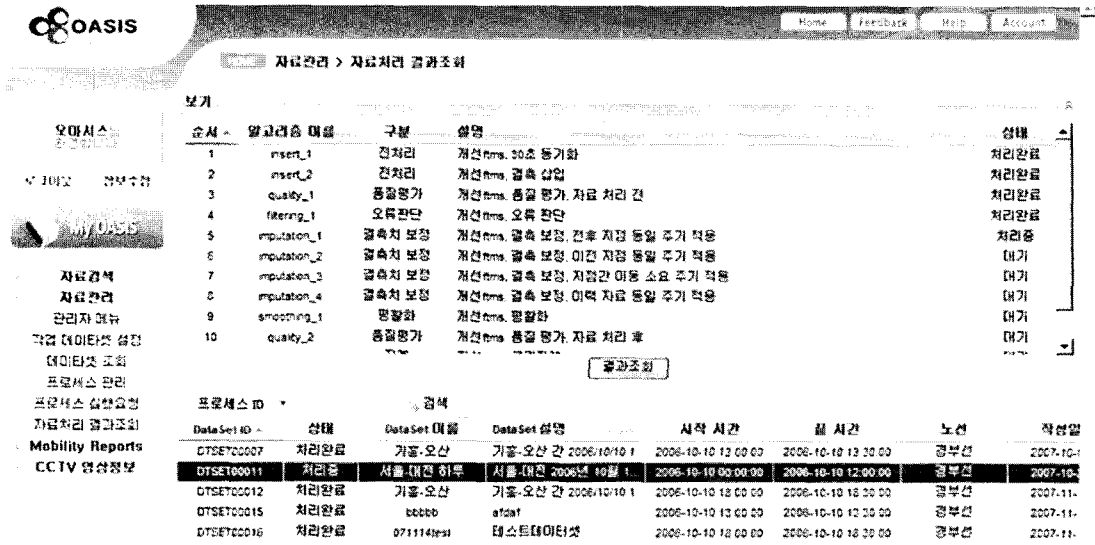
알고리즘을 조합하기 위해서 알고리즘 추가를 하면 <그림 3>과 같이 선택할 알고리즘을 순차적으로 선택하거나 알고리즘의 실행순서의 변경이 가능하다.

알고리즘은 오직 시스템 관리자만이 등록, 수정할 수 있으며, 일반 사용자들은 관리자가 생성해둔 알고리즘을 프로세스 생성 시에 사용하는 것만 가능하다. 프로세스의 정보는 데이터베이스의 SYS_DP_PROCESS 테이블에 저장되며, 알고리즘은 SYS_ALG 테이블에 저장된다.

프로세스 실행은 자료가공 프로세스를 실행하기 위해서 사용자는 먼저 하단의 프로세스 리스트로부터 사용할 프로세스를 선택한다. 프로세스를 선택하고 나면, 가공을 원하는 Data set을 선택한다. 프로세스와 Data set을 선택하고 나면, 프로세스를 구성하는 알고리즘들이 요구하는 파라미터들을 설정하여야 한다. 파라미터 값을 모두 설정하고 나서, 실행요청 버튼을 클릭하면 자료처리 프로세스 실행요청이 완료된다. 프로세스 실행요청이 이루어지면, SYS_DP_RSLT 테이블에 새 레코드가 추가된다. 자료처리 서비스는 SYS_DP_RSLT 테이블을



<그림 3> 알고리즘 추가 화면 예
 <Fig. 3> An example of algorithm selection



<그림 4> 자료처리 결과 예
 <Fig. 4> A snapshot of data processing results

주기적으로 조회하여, 요청된 프로세스가 있다면, 해당 프로세스를 자동적으로 실행한다. 자료처리 프로세스가 실행되면, 프로세스의 진행상황을 사용자가 모니터링할 수 있도록 SYS_DP_RSLT 테이블과 SYS_DP_STATUS 테이블에 계속해서 기록하게 된다. 프로세스 실행요청 시에 설정되는 파라미터는 SYS_DP_PARAM 테이블에 저장된다. 자료처리 실행요청 후, <그림 4>과 같이 자료처리 결과조회 메뉴를 통해서 사용자는 자료처리 과정을 모니터링할 수 있다.

<표 3> 시간적·공간적 범위
 <Table 3> Data set for evaluation

	시간범위	공간범위
Data set I	2007년10월13일 00시 00분 00초 ~ 23시 59분 00초	경부고속도로 북대구IC~동동JC
Data set II	2007년10월14일 00시 00분 00초 ~ 23시 59분 00초	경부고속도로 동동JC~동대구IC

V. 자료처리 품질평가 실험

품질 비교는 원시데이터와 자료처리과정을 모두 거친 데이터 간의 완전성과 유효성을 기준으로 평가하였다.

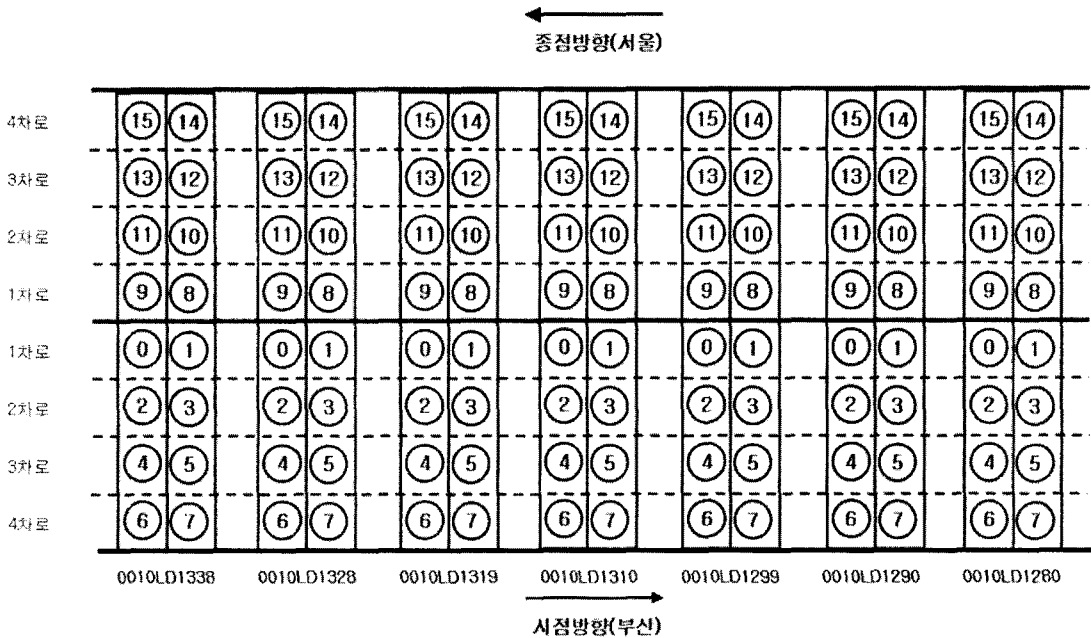
1. 적용지점 선정

1) 시간적·공간적 범위

자료처리 품질 비교 적용의 시간적·공간적 범위는 <표 3>과 같다.

2) 검지기 위치 및 루프 논리적 순서

품질 비교를 위한 Data set은 구간내 차로수 변화로 인한 결측보정 오류를 방지하기 위해 동일한 차로수를 갖는 두 개구간에서 추출되었다. Data set I은 구간내 7개의 검지기가 설치되어 있고, Data set II는 구간내 5개의 검지기가 설치되어 있다. 검지기의 위치 및 관계에로 Data set I은<그림 5>와 같으며 원안의 번호는 설치된 루프의 논리적 순서를 나타낸다. 또한 LDS_아이디의 앞숫자는 설치된 노선명을 나타내며, 뒷숫자는 검지기 이정을 나타낸다.



<그림 5> Data set I의 검지기 위치 및 관계
 <Fig. 5> Loop configuration for dataset I

3) 자료처리조건

자료처리를 위한 입력조건은 <표 4>와 같이 오류판단의 임계값으로 교통량 최대값 30대, 인접검지기 교통량 차이 2대, 점유율 최대값 100% 하였다. 관계검사는 속도-점유율관계를 사용하였고 결측보정은 인접지점 참조 방법을 적용하였다. Data set I의 입력파라미터는 참조지점 최대값 2를 적용하였고, Data set II의 입력파라미터는 참조지점 최대값 3을 적용하였다. 평활화는 평활화 계

<표 4> 자료처리 파라미터 입력 조건

<Table 4> Input parameters for data processing

입력파라미터	Data set I	Data set II
V_MAX_V	30	30
V_DIFFERENCE_V	2	2
O_MAX_O	100	100
V_REF_NUM1	2	3
V_REF_NUM2	2	3
V_REF_NUM3	2	3

수 1을 적용하였다.

2. 원시 데이터와 품질 평가

1) 원시데이터

원시데이터는 30초 간격으로 각 지점에 설치된 0~15번 루프를 통해 교통량, 속도, 점유율이 집계된다. 원시자료는 <표 5>와 같이 24시간동안 10월 13일은 7개 지점에서 총322,448개의 교통자료가 기록되고 결측데이터는 9,312개, 오류데이터는 56,992개, 10월 14일은 5개 지점에서 총230,320개의 교통자료가 기록되었고, 결측데이터는 46,144개, 오류데이터는 34,624개 기록되었다.

<표 5> 원시데이터

<Table 5> Raw data

	10월 13일	10월 14일
총데이터	322,448	230,320
결측데이터	9,312	46,144
오류데이터	56,992	34,624

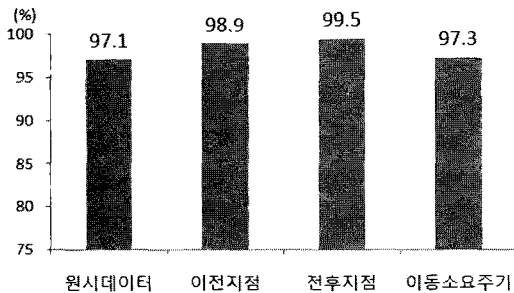
<표 6> 자료처리데이터
<Table 6> Data processing results

		10월13일	10월14일
이전지점 동일주기 방법	결측	3,514	15,686
	오류	13,286	21,354
전후지점 동일주기 이동평균 방법	결측	1,688	3,780
	오류	2,278	6,006
지점간 이동 소요주기 방법	결측	8,718	45,750
	오류	47,236	32,012

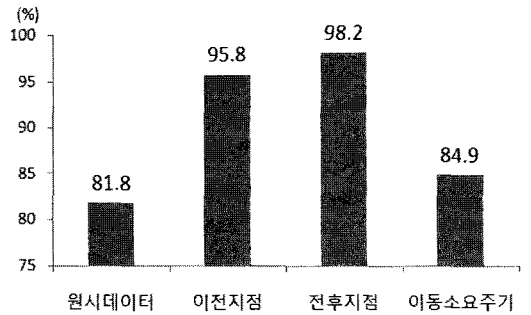
2) 자료처리 데이터

오류판단, 결측보정, 평활화의 자료처리과정을 모두 거친 데이터들을 결측보정 방법에 따라 보면 <표 6>과 같고, 여기서 결측은 참조 지점내 값이 없어서 보정을 수행하지 못한 데이터를 말하고, 오류는 오류판단 기준에 따라 오류자료로 판단된 데이터이다.

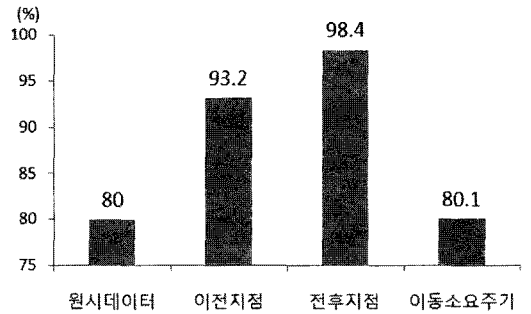
품질평가 결과는 결측보정 방법별로 구분하여 유효성과 완전성을 측정하였으며, 결과는 <그림 6>~<그림 9>와 같다. Data set I의 원시자료 품질은 완전성 97.1%, 유효성 81.8%으로 평가되었고 자료처리 후 품질은 전후지점 동일주기 이동평균 방법이 가장 높은 품질향상(완전성 2.4%, 유효성 16.4%)을 보인다. 또한 Data set II의 원시자료 품질은 완전성 80.0%, 유효성 81.2%으로 평가되었고 자료처리 후 품질은 전후지점 동일주기 이동평균 방법이 가장 높은 품질향상(완전성 18.4%, 유효성 15.6%)을 보이고 있다. 품질평가 결과 자료품질은 모두 향상되나 자료처리 방법 및 원시데이터의 상태에 따라 다른 품질을 보여준다.



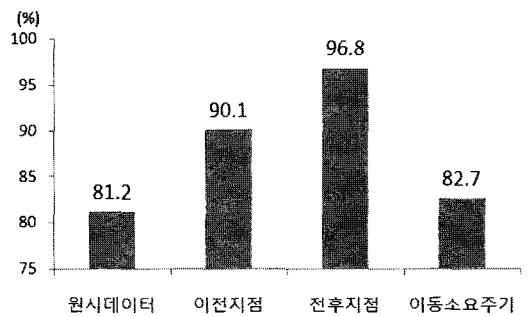
<그림 6> Data set I 완전성 비교
<Fig. 6> Completeness evaluation (data set I)



<그림 7> Data set I 유효성 비교
<Fig. 7> Validity evaluation (data set I)



<그림 8> Data set II 완전성 비교
<Fig. 8> Completeness evaluation (data set II)



<그림 9> Data set II 유효성 비교
<Fig. 9> Validity evaluation (data set II)

VI. 결론 및 향후연구과제

지능형교통체계는 도로 및 교통체계의 안전 및 운영효율을 극대화 시킬 수 있으며 이것을 지원하기 위하여 실시간 교통모니터링 중 하나인 차량검지기를 통한 교통자료의 운영과 활용이 필요하다.

또한 교통자료의 활용과 제공분야가 점차 늘어나는 추세에 따라 수집되는 교통자료를 운영 및 활용하기 위해서는 자료의 품질이 신뢰성을 가져야 하며 이러한 교통자료의 신뢰성을 향상시키기 위하여 본 연구에서는 교통자료 처리기법 통합분석 방법론을 개발 하였다. 도로에서 수집되는 원시데이터를 신뢰성 있는 자료로 가공하기 위한 새로운 자료처리 알고리즘을 등록하여 테스트 및 운영할 수 있으며, 기존에 사용 중인 알고리즘의 파라미터를 변화시키거나 알고리즘 선택을 통하여 사용자가 원하는 수준의 자료품질을 얻을 수 있도록 하였다. 또한 자료처리의 통합평가 방법 및 개선방향 제시가 가능한 도구를 개발이 가능하다. 이를 적용시켜본 결과 동일한 2일간의 원시데이터를 각각 다른 알고리즘으로 자료처리를 한 결과 소폭의 차이는 있지만 모두 원시 데이터 보다 향상된 자료 품질을 보이고 있다.

향후 연구 과제로는 자료의 품질을 향상 시킬 수 있는 신규 알고리즘 개발이 필요하며, 기존 운영 중인 알고리즘에 대해서는 품질향상을 위한 연구가 요구된다.

참고 문헌

- [1] 한국도로공사 도로교통기술원, 차량검지기 자료의 효율적 수집저장 및 관리체계 연구, 2006. 12.
- [2] <https://pems.eecs.berkeley.edu/>
- [3] <http://trafficdataarchive.ce.virginia.edu/>
- [4] 한대철, 최숙양, 문학룡, 김성현, "ITS 시스템의 결측자료 보정에 관한 연구," 한국ITS학회 제3회 추계학술대회 논문집, pp. 270~275, 2004. 11.
- [5] 김정연, 이영인, 백승걸, 남궁성, "차량 검지자료 결측 보정처리에 관한 연구(이력자료 활용 방안을 중심으로)," 대한교통학회지, 제24권, 제7호, pp. 27~40, 2006. 12.
- [6] H. M. Al-Dee and C. Chandra, "New algorithms for filtering and imputation of real-time and archived dual-loop detector data in I-4 data warehouse," *Transportation Research Record*, vol. 867, pp. 116-126, 2004.
- [7] Z. Wall and D. Dailey, "Algorithm for detecting and correcting errors in archived traffic data," *Transportation Research Record*, vol. 1855, pp. 183-190, 2003.
- [8] S. Turner, L. Albert, B. Gajewski, and W. Eisele, "Archived intelligent transportation system data quality: preliminary analyses of San Antonio transguide data," *Transportation Research Record*, vol. 1719, pp. 77-84, 2000.
- [9] M. Zhong, S. Sharma, and P. Lingras, "Genetically designed models for accurate imputation of missing traffic counts," *Transportation Research Record*, vol. 1879, pp. 71-79, 2004.
- [10] R. Schmoyer, P. Hu, and R. Goeltz, "Statistical data filtering aggregation to hour totals of intelligent transportation system 30-s and 5-min vehicle counts," *Transportation Research Record*, vol. 1769, pp. 79-86, 2001.
- [11] R. Bertini and A. El-Geneidy, "Generating transit performance measures with archived data," *Transportation Research Record*, vol. 1841, pp. 109-119, 2003.
- [12] 장진환, 백남철, 강원의, "차량검지기 성능평가 방법론 연구," 한국ITS학회 제3회 추계학술대회 논문집, pp.198~201, 2004. 11.
- [13] 이정준, 오영태, "Loop식 차량검지기의 신뢰성 평가를 위한 모의 시험장치 핵심부 개발," 대한교통학회 제51회 학술발표회 논문집, pp. 577~584, 2006. 2.

저자소개



오 동 옥 (Oh, Dong-Wook)

2007년 3월 ~ 현재 : 한양대학교 교통공학과석사과정재학중(교통공학전공)
2001년 3월 ~ 2007년 2월 : 한양대학교 공학학사(교통시스템공학전공)



오 철 (Oh, Cheol)

1999년 9월 ~ 2003년 12월 : University of California, Irvine 토목 환경 공학과
공학박사 (교통시스템공학 전공) .
1995년 8월 ~ 1997년 8월 : 한양대학교 대학원 교통공학과(교통공학 전공)
2006년 3월 ~ 현재 : 한양대학교 교통시스템공학과 조교수
2004년 4월 ~ 2006년 2월 : 한국교통연구원 첨단교통기술연구실 책임연구원
2004년 1월 ~ 2004년 3월 : Post Doctorate Research Associate, Institute of
Transportation Studies University of California, Irvine



남 궁 성 (Namkoong, Seong)

2004년 11월 ~ 2008년 현재 : 한국도로공사 도로교통연구원 수석연구원
2002년 9월 ~ 2004년 10월 : Smart travel Lab. Univ. of Virginia, Research Scientist
1996년 10월 ~ 2002년 8월 : 한국도로공사 도로연구소 교통연구그룹장
2007년 1월 ~ 2008년 12월 : 아주대학교 ITS대학원 강사
1991년 3월 ~ 1996년 8월 : 한양대학교 부설 산업과학연구소 연구원



전 세 길 (Jeon, Se-Gil)

2000년 2월 ~ 2004년 8월 : 단국대학교 전자컴퓨터공학과 공학박사(정보처리전공)
2006년 3월 ~ 현재 : 한국도로공사 도로교통연구원 박사후연수자
2005년 1월 ~ 2006년 2월 : 단국대학교 BMIC 박사후연수자
2000년 1월 ~ 2006년 2월 : 단국대학교, 서울여대, 국민대 강사
2000년 1월 ~ 2002년 8월 : (주)패스싸인 선임연구원