

길찾기 과정의 도로명주소 체계 연계를 위한 선형 객체 매칭 방법 Line Matching Method for Linking Wayfinding Process with the Road Name Address System

방윤식* · 유기윤**
Bang, Yoon Sik · Yu, Ki Yun

要 旨

지난 2012년부터 도로명주소가 본격 시행 및 활용되고 있지만, 아직도 상당 부분에서는 기존의 지번주소가 많이 통용되고 있다. 이는 일반인들의 공간 인식체계와 도로명주소의 공간 구조화 방식의 간극으로 인한 문제이다. 따라서 도로명주소 기반의 공간 인식 체계가 자리잡기 위해서는 생활 속에서 활용되는 각종 공간정보들이 도로명에 의하여 주소정보를 부여받을 수 있어야 한다. 본 연구는 공간 인식 과정이 가장 중요하게 나타나는 길찾기 과정에서의 도로명주소 체계의 연계를 목적으로, 이를 위하여 필요한 공간데이터의 기하학적 매칭 방법론을 설계 및 구현하였다. 도로명주소 기본도의 도로구간 레이어와 보행자용 도로 네트워크에 대하여, 개별 도로 객체를 중심으로 네트워크 이웃을 생성하였다. 그 다음, 생성된 이웃 집합 간의 기하학적 유사도 비교를 통하여, 네트워크 데이터의 각 객체에 매칭되는 도로구간을 탐색하였다. 매칭 성능은 F0.5 값을 기준으로 0.936의 결과를 얻었으며, 유사도 값을 기준으로 10% 수동 검사를 수행한 결과 이 값을 0.978까지 향상시킬 수 있었다. 이렇게 생성된 매칭 대응관계를 이용하여, 보행자용 도로 네트워크 데이터에 도로명 정보를 부여하였다. 이러한 방법론을 통하여, 도로명주소를 기반으로 한 길찾기 서비스 제공 및 공간 인식체계 정착에 도움을 줄 수 있다.

핵심용어 : 도로명주소, 길찾기 서비스, 보행자용 도로 네트워크, 기하학적 매칭

Abstract

The road name address system has been in effect in Korea since 2012. However, the existing address system is still being used in many fields because of the difference between the spatial awareness of people and the road name address system. For the spatial awareness based on the road name address system, various spatial datasets in daily life should be referenced by the road names. The goal of this paper is to link the road name address system with the wayfinding process, which is closely related to the spatial awareness. To achieve our goal, we designed and implemented a geometric matching method for spatial data sets. This method generates network neighborhoods from road objects in the 'road name address map' and the 'pedestrian network data'. Then it computes the geometric similarities between the neighborhoods to identify corresponding road name for each object in the network data. The performance by F0.5 was assessed at 0.936 and it was improved to 0.978 by the manual check for 10% of the test data selected by the similarity. By help of our method, the road name address system can be utilized in the wayfinding services, and further in the spatial awareness of people.

Keywords : Road Name Address, Wayfinding Service, Pedestrian Network Data, Geometric Matching

1. 서 론

도로명주소 체계는 지난 100여년간 지속되어 온 지번주소체계의 문제점을 해소하고, 경로 안내의 용이성 강화, 국제표준 채택 등의 효과를 달성하기 위하여 도

입된 도로 기반의 주소체계이다. 1997년 시범사업을 시작으로, 2011년 전국 고지 및 고시, 2013년까지 병행 사용 기간을 거쳐 2014년부터 전면 시행되었다(Road Name Address, 2016).

이에 따라 공공기관이나 관련 업무에 대해서는 도로

Received: 2016.12.12, accepted: 2016.12.21

* 정희원 · 서울대학교 건설환경공학부 박사과정(Member, Doctoral Student, Department of Civil and Environmental Engineering, Seoul National University, bangys1004@snu.ac.kr)

** 교신저자 · 정희원 · 서울대학교 건설환경공학부 교수(Corresponding Author, Member, Professor, Department of Civil and Environmental Engineering, Seoul National University, kiyun@snu.ac.kr)

명주소 활용 비율이 상당히 높아졌으나, 아직 일상생활에서는 도로명주소가 익숙하게 활용되지 않고 있다. 고시 후 5년여 시간이 지났음에도 택배, 배달, 택시 등의 생활 서비스 이용 시 기존 지번주소의 활용 빈도가 여전히 높은 상황이다.

이러한 문제의 근본적인 원인은, 일반인들의 공간 인식체계와 도로명주소의 공간 구조화 방식의 간극이라고 볼 수 있다. 도로 네트워크 데이터, 주소 데이터, 구역 데이터 등 공간정보 관련 활동에 활용되는 기반 데이터들이 기존의 지번주소 체계를 기반으로 제작된 것들이 대부분이다.

도로명주소 기반의 공간 인식체계가 자리잡기 위해서는 공간 탐색이 가장 적극적으로 이루어지는 행위, 즉 ‘길찾기’ 과정에서의 도로명주소 활용이 매우 중요하다. 도로명주소 체계 활용의 대표적인 장점으로 길찾기의 편의성이 홍보되고 있음에도 불구하고, 정작 수많은 길찾기 또는 길안내 서비스에서 도로명주소는 그리 많이 활용되지 않고 있다.

최근에는 보행 친화적 환경 실현이 도시 설계 및 재개발에서 큰 비중을 차지하고 있다. 이러한 흐름을 고려할 때, 차량용 내비게이션 서비스보다는 보행자를 위한 길찾기 및 길안내 서비스에 대하여 도로명주소 활용 개선이 더욱 중점적으로 이루어질 필요가 있다 (Kim, 2015).

도로명주소 활용 확대를 위한 연구로서, 행정자치부에서는 도로명주소 활용을 위한 주소체계 및 시설물 개선방안에 대한 연구를 수행한 바 있다(CMI Consortium, 2006). Roh(2009)는 도로명주소 기반 지오코딩의 위치 정확도를 평가하고 개선방안을 제안하였다. 또한 Lee et al.(2015)은 도로명주소 기본도의 구축현황을 파악하고 위치정확도 오류를 분석하였다. Seok and Lee(2016)는 도로명주소 기반 3차원 지오코딩 및 역지오코딩 기법을 제안하였다.

도로명주소 체계와 타 공간정보와의 연계를 위한 연구로, Kang and Jang(2012)은 도로명주소와 연계한 국가기초구역 데이터 모델을 제안하였다. Bang et al.(2012)은 수치지도 도로 레이어와 도로명주소 기본도 사이의 매칭 및 연계 방안을 제시하였다. 또한 Yang(2013)은 도로명주소 기본도를 기초로 부동산개발안내도를 구축하는 방안을 제시하였으며, Yeom et al.(2014)은 도로명주소 기본도와 수치지도의 건물 레이어 간 매칭 및 갱신 방법을 연구하였다.

이에 본 연구에서는 도로명주소 기본도를 활용하여, 길찾기 서비스에 사용되는 보행자용 네트워크 데이터와 도로명주소 체계의 연계를 목적으로 하였다. 이를

위하여 필요한 공간데이터의 기하학적 매칭 방법론을 설계 및 구현하였다.

2. 현황 및 필요성

2.1 도로 네트워크 데이터

보행자를 위한 길찾기 서비스는 일반적으로 보행자가 이동 가능한 모든 경로를 표현하는 지리적 네트워크 데이터를 기반으로 이루어진다. 이러한 네트워크 데이터는 차량용 내비게이션 서비스의 발전과 함께 구축되어 왔으며, 도로의 선형을 나타내는 링크(link)와 각 링크들의 시작/끝점을 나타내는 노드(node)로 구성되어 있다. 최근에는 보행자용 길찾기를 제공하기 위하여 인도, 보행자 전용길, 횡단시설 등을 더욱 상세하게 표현한 보행자용 네트워크 데이터(pedestrian network data, PND)가 구축 활용되고 있다.

보행자용 네트워크 데이터의 링크 객체는 보통 “도로 유형”, “길이”, “폭”, “통행제한”, “시작점 노드”, “끝점 노드” 등의 속성정보들을 포함하고 있다. 그 외에 특정 용도에 따라 필요한 추가적인 속성 (시설물 유형, 방향, 경사도, 포장재질 등)들이 구축 및 활용되고 있다.

그러나 도로명에는 속성정보에 포함되어 있지 않은 경우가 많다. 이는 보행자 네트워크 데이터가 기존 차량용 네트워크 데이터를 가공하여 구축되었는데, 이 데이터들이 도로명주소 체계와 무관하게 작성되었기 때문이다. 지번주소가 등록된 공간데이터의 경우 도로명주소 안내시스템(www.juso.go.kr)에서 제공하는 매칭테이블을 이용하여 도로명주소 값을 부여할 수 있지만, 단일 주소값이 없는 선형 공간객체에 해당 도로명을 부여하는 것은 불가능하다.

길안내 과정에서의 도로명주소 체계 연계를 위해서는 공식 도로명 및 관련 정보가 서비스 기반 데이터에 포함되어야 한다. 현재 차량용 내비게이션 서비스의 경우, 도로명주소 제도가 전면 시행되면서 그에 필요한 장소 DB 및 도로 네트워크 데이터가 도로명주소 정보를 포함하도록 개량되어 왔다.

그러나 보행자 길찾기 서비스의 경우에는 배경 네트워크 데이터 자체가 충분히 구축되어 있지 않으며, 구축된 데이터 역시 도로명주소 정보와의 연계성이 부족한 상태이다.

2.2 도로명주소 체계

2007년 제정된 <도로명주소법>의 정의에 따르면 도로명주소란 도로명, 기초번호, 건물번호, 상세주소에 의

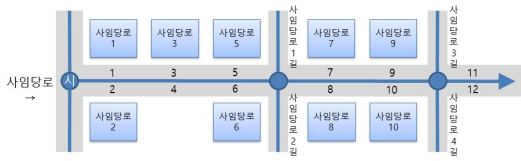


Figure 1. Basic rules for assignment of the road name address

하여 건물의 주소를 표기하는 방식이다(Ministry of the Interior, 2013). 이에 따라 주소 부여의 기준이 되는 각 도로의 시작점에서 끝점을 연결한 가상의 선형인 도로 구간이 설정되었으며, 각 도로구간마다 도로명이 부여되었다. 도로구간은 도로의 규모에 따라 대로-로-길 순서의 위계를 갖도록 설정되었다. 건물번호는 도로에 연결된 건물들에 대하여 순서대로 부여된다. 건물번호는 도로구간의 시점에서부터 종점 방향으로 순서대로 좌측에는 홀수, 우측에는 짝수 번호를 부여하는 것을 원칙으로 한다. 도로명주소 부여의 기본 원칙을 간단히 나타내면 Fig. 1과 같다(Road Name Address, 2016).

지번주소는 법정동 내 각 필지에 대하여 유일하게 지정된 번호를 이용하여 주소를 나타낸다. 지번주소 제정 초기에는 이 번호가 실제 위치에 따라 체계적으로 부여되었지만, 오랜 시간 동안 필지의 통합과 분할이 계속되면서 지번의 공간적 연속성이 약화되었다. 이에 따라 지번주소의 위치찾기 기능이 거의 상실되었으며, 이에 대한 대안으로 도로명주소가 도입되었다.

지번주소는 여러 단계에 걸쳐 국토를 평면적으로 분할함으로써 각 영역에 대한 주소를 부여하는 반면, 도로명주소는 장소에 연결된 도로를 기준으로 건물에 대한 주소를 부여하는 방식이다. 보다 근본적으로, 두 주소체계는 공간의 구조화 방식에 가장 큰 차이가 있다. 지번주소 체계에서는 각 장소가 여러 단계에 걸친 공간 분할의 결과로 특정된다. 반면 도로명주소 체계에서는 각 장소가 해당 위치로의 접근 과정에 의하여 특정된다. 이러한 관점에서 볼 때, 도로명주소는 길찾기 과정에 보다 적합한 위치참조 방식이라고 할 수 있다. 따라서 길찾기에 사용되는 기반 데이터 역시 도로명주소 체계하에서 활용될 수 있도록 주소값이 부여되어야 한다.

2.3 필요한 기술적 과제

길찾기에 사용되는 기반 데이터 중 장소정보 데이터, 즉 POI(point of interest) 데이터는 각 객체가 점 형태로 표현되어 있다. 각 점은 단일한 좌표값을 갖기 때문에 지번주소 체계하의 위치 탐색, 즉 역 지오코딩(reverse geocoding) 과정을 통하여 지번주소 값을 찾

을 수 있으며, 매칭 테이블을 이용하여 도로명주소 값도 쉽게 찾을 수 있다. 그러나 배경 네트워크 데이터의 경우, 각 링크 객체들은 선 형태로서 단일 좌표값으로 표현되지 않기 때문에 지번주소 체계하에서는 주소값을 부여할 수 없다. 대신, 도로명주소 지도에 존재하는 도로구간 데이터와의 매칭을 통하여 직접적으로 ‘도로명’ 값을 부여하는 과정이 필요하다. 이를 위하여, 선형 공간데이터 간 기하학적 매칭 방법을 개발하여 적용하고자 한다.

기하학적 매칭을 통하여 배경 네트워크 데이터의 링크 데이터와 도로명주소 지도의 도로구간 데이터 사이의 대응 관계를 찾아낸 다음에는, 도로명주소 정보를 이용하여 네트워크 데이터를 갱신한다.

3. 데이터 매칭 및 갱신 방법론

본 장에서는 위에서 제시한 기술적 과제를 해결하기 위한 방법론을 연구 개발하였다.

3.1 보행자용 네트워크 데이터와 도로명주소 데이터의 기하학적 매칭

선형 데이터 사이의 기하학적 매칭을 위한 접근법에는 여러 가지가 있다. 그 중에서도 본 연구에서와 같이 각 선형 객체들이 평면상에서 서로 연결되어 있는 네트워크 구조를 형성하고 있는 경우에 대한 여러 방법론들이 개발되어 왔다. 대표적인 방법으로는 버퍼 그로잉(buffer growing) 기법(Walter and Fritsch, 1999)이 있다.

그러나 이 방법은 두 데이터셋이 일정한 오차 범위 내의 위치 차이를 가질 때 적합하다. 본 연구의 대상 데이터셋들은 표현 방식 및 세밀도의 측면에서 서로 큰 차이가 있기 때문에, 공간객체 간 위치오차에 영향을 덜 받는 매칭 기법을 적용할 필요가 있다.

본 연구에서는 개별 공간객체 간 위치오차로 인한 영향을 최소화하기 위하여, 좀 더 넓은 범위의 데이터셋을 기하학적으로 비교하는 방법을 제안하였다. 먼저, 보행자용 네트워크 데이터의 한 객체(기준 객체)를 중심으로, 미리 정해진 탐색 범위(R_C) 내에 있는 도로명주소 도로구간(road section) 객체들을 매칭 후보로 선정한다. 그 다음, 기준 객체로부터 거리가 일정 기준(D_N)보다 가까운 이웃 객체들을 추출하고, 마찬가지로 각 매칭 후보들을 중심으로 각자의 이웃 객체들을 추출한 다음, 이웃 객체들끼리의 기하학적 유사도를 측정하였다. 복수의 매칭 후보들 중 측정된 기하학적 유사도 값이 가장 높은 도로구간 객체를 매칭 대상으로 선택하였다. 이와 같은 매칭 프로세스의 대략적인 흐름도는 Fig. 2와 같다.

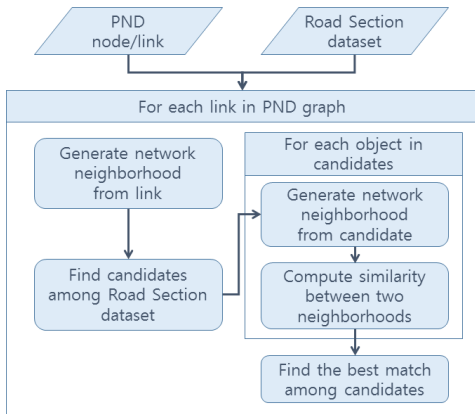


Figure 2. Flowchart of matching process between PND and the road section dataset

3.1.1 네트워크 이웃 생성

어떤 객체의 이웃 객체들은 중심 객체로부터의 네트워크 거리 값이 일정 기준(D_N) 이하인 객체들을 추출함으로써 생성할 수 있다. 여기에서 가장 중요한 점은, 이웃 객체 추출 시 유클리드 거리가 아닌 네트워크 거리를 기준으로 가까운 객체를 추출한다는 점이다. 두 지점 사이의 네트워크 거리는 한 점에서 다른 점까지 네트워크를 따라 이동하는 데 지난 거리로 정의된다. 이 과정에서, Zhan(1998)의 연구를 참조하여 네트워크 데이터 내에서 최단경로 탐색 알고리즘을 적용하였다. 이렇게 추출된 이웃 객체들의 집합은 네트워크 이웃(network neighborhood)라고 정의하였다. 네트워크 이웃 생성 과정은 아래 Fig. 3과 같다.

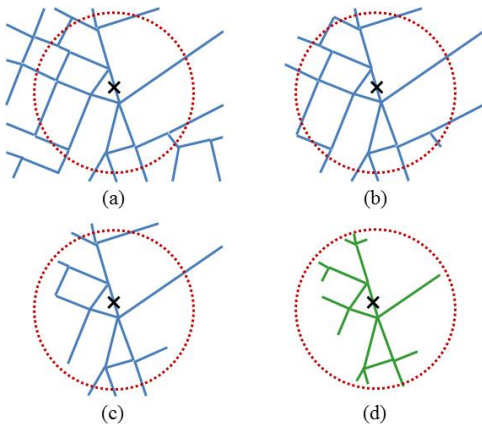


Figure 3. Establishing the network neighborhood of a road segment object within the distance D_N which is depicted by a circle drawn in dotted line

위 Fig. 3의 프로세스를 단계별로 설명하면 다음과 같다.

- 기준 객체의 중심점을 찾고, 그 중심점으로부터 최단 유클리드 거리 값이 D_N 이하인 주변객체 추출
- 추출된 각 주변객체들의 양 끝점으로부터 기준 객체 중심점까지의 최단 경로 계산
- 각 주변객체에 대하여, 계산된 두 최단 경로 중 하나 이상이 D_N 이하인 객체 추출
- 최종 추출된 이웃 객체들에 대하여, 네트워크 거리 D_N 초과하는 부분 잘라내기

3.1.2 기하학적 유사도 측정

기준 객체와 그의 매칭 후보들을 중심으로 생성된 네트워크 이웃 사이의 기하학적 유사도를 측정하기 위하여, 본 연구에서는 하우스도르프 거리(Hausdorff distance) 개념을 활용하였다. 하우스도르프 거리는 두 공간객체 사이의 거리를 계산하는 방법으로서, Eq. (1)에 의하여 정의된다(Min et al., 2007).

$$HD(A,B) := \max\{h(A,B), h(B,A)\} \quad (1)$$

이 식에서 $h(A,B)$ 는 방향 하우스도르프 거리(directed Hausdorff distance)인데, B 를 중심으로 하여 A 를 모두 포함하는 버퍼 영역의 반경의 최솟값으로서, 다음 Eq. (2)에 의하여 정의된다.

$$h(A,B) = \max_{p_A \in A} \{ \min_{p_B \in B} \{ d(p_A, p_B) \} \} \quad (2)$$

where, p_A and p_B denote the arbitrary points consisting A and B , respectively, and $d(p_A, p_B)$ denotes the shortest path between p_A and p_B

즉, 하우스도르프 거리는 두 객체 사이의 방향 하우스도르프 거리 값 중 더 큰 값으로 정의된다. 그러나 본 연구에서는, 비교 대상이 되는 보행자용 네트워크 데이터와 도로명주소 도로구간 데이터 사이의 세밀도 차이가 있음을 가정하였으므로, 비교 객체 중 일부분이라도 기하학적 유사성이 높으면 매칭 가능성이 높다고 볼 수 있다. 따라서 본래 하우스도르프의 정의와 다르게 두 객체의 방향 하우스도르프 거리 값 중 더 ‘작은’ 값을 두 객체 사이의 거리 값으로 택하였다.

이 거리 값이 작을수록 유사도가 높으므로, 최종 유사도 값은 이 거리 값의 역수로 정의하였다. 보행자용 네트워크 데이터의 링크 a 의 네트워크 이웃을 A , 도로

명주소 도로구간 객체 b 의 네트워크 이웃을 B 라 하면, 둘 사이의 유사도 $S(A,B)$ 는 다음 Eq. (3)과 같이 정의된다.

$$S(A,B) := \frac{1}{HD_{\min}(A,B)} \quad (3)$$

$$= \frac{1}{\min\{h(A,B), h(B,A)\}}$$

3.1.3 최적 매칭 결과 선택

앞 단계에서 보행자용 네트워크 데이터의 각 링크 객체를 기준으로, 주변에 위치한 도로명주소 도로구간 데이터의 객체들을 추출하여 후보 객체로 선택하였다. 이 단계에서는 후보 객체들과의 네트워크 이웃 유사도 값을 기준으로, 유사도 값이 가장 높은 후보 객체를 매칭 결과로 선택하였다(Fig. 4). 이 때, 최대 유사도가 후보 탐색 범위(R_C)의 역수보다 작으면 매칭 결과가 존재하지 않는 것으로 판정하였다. 즉, 후보 객체들에 대하여 Eq. (3)으로부터 계산한 하우스도르프 거리(HD_{\min}) 값이 R_C 보다 큰 경우는 형태적 유사성이 낮은 것으로 간주하여 후보에서 제외하였다.

또한 복수의 매칭 후보 객체들이 서로 비슷한 수준의 유사도 값을 갖는 경우도 발생할 수 있는데, 이는 교차로와 같이 둘 이상의 도로명주소 도로구간 객체가 만나는 위치에서 주로 발생하게 된다. 이러한 복수 매칭 계

이스를 찾아내기 위한 판정 기준은 다음과 같다. 매칭 후보 선택 범위를 R_C , 도로명주소 도로구간 데이터의 최대 위치오차를 E_{RS} 라 하면, 후보 객체 b_i 의 복수 매칭 여부는 다음 Eq. (4)에 의하여 판정한다.

$$b_i \in RS_{\max}, \text{ if } S(A, B_i) > \frac{1}{R_C} \text{ and}$$

$$\frac{S(A, B_{\max}) - S(A, B_i)}{S(A, B_{\max})} < \frac{E_{RS}}{R_C} \quad (4)$$

where, RS_{\max} denotes the set of road sections of maximum similarity among candidates, B_i denotes the network neighbor of b_i , and B_{\max} denotes the one with the highest similarity

위의 과정을 거쳐 추출한 객체 집합 RS_{\max} 가 둘 이상의 원소를 갖는 경우, 다시 말해 복수 매칭이 발생한 경우에는, 유사도 비교 대상을 ‘네트워크 이웃’에서 개별 객체 단위로 축소, $S(a, b_i)$ 값을 계산하여 이 값이 가장 큰 후보를 최종 매칭 대상으로 결정하였다. 즉, 개별 객체의 형상이 조금이나마 더 유사한 도로구간 객체를 찾도록 하였다.

3.2 보행자용 네트워크 데이터의 갱신

기하학적 매칭 프로세스를 통하여 네트워크 데이터의 각 링크에 대응되는 도로명주소 지도 도로구간 객체

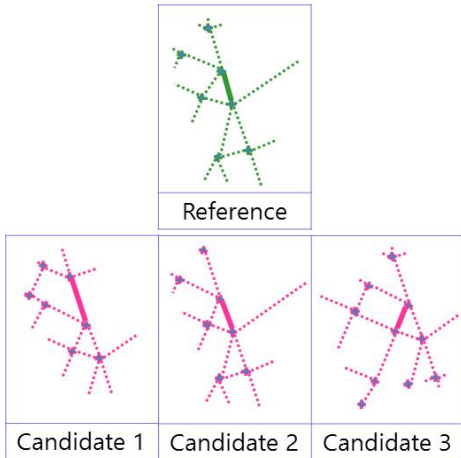


Figure 4. Finding the best match of a PND object from three candidates by comparing network neighborhoods. The candidate 2 shows the highest similarity value and therefore is the best match

Table 1. Attributes which can be updated from the road section dataset to the PND

| Attribute (Eng.) | Attribute (Kor.) | Description | Type |
|------------------|------------------|-------------------------|------|
| RN | 도로명_한글 | road name in Korean | Char |
| ENG_RN | 도로명_영문 | road name in English | Char |
| RN_CD | 도로관리번호 | road section identifier | Char |
| SIG_CD | 시군구코드 | city identifier | Char |
| ROAD_BT | 도로폭 | road width | Num |
| ROA_CLS_SE | 도로위계기능구분 | road class | Char |
| RDS_DPN_SE | 도로구간종속구분 | status of dependence | Char |
| PAR_RDS_NO | 분기도로구간일련번호 | dependent road number | Num |
| RDS_MAN_NO | 도로구간일련번호 | road number | Num |

가 탐색되면, 해당 도로구간 객체의 도로명주소 정보를 해당 링크의 속성정보에 추가하여야 한다. 갱신될 수 있는 도로명주소 정보의 목록은 Table 1과 같다.

이 단계에서는 각 네트워크 데이터의 링크 객체 속성 테이블이 도로명주소 정보를 포함하도록 데이터베이스 스키마를 확장하고, 매칭 관계를 이용하여 확장된 속성 테이블의 레코드에 속성정보를 입력한다.

4. 실험 및 평가

4.1 대상 데이터

본 연구에서는 보행자 길안내 서비스를 위하여 제작된 보행자용 네트워크 데이터를 매칭 대상 데이터로 선정하였다. 공간적 범위는 서울특별시 양천구 신월동-신정동 일대 1개 도엽(1:1000 축척의 수치지도 기준)으로 설정하였다. 총 2903개의 링크로 이루어져 있으며, ESRI Shapefile 형식(.shp)으로 제작되어 있다(Fig. 5(a)).

프로세스의 효율성을 위하여, 보행자용 네트워크 데이터 링크 객체 중 도로명주소 도로구간 데이터와의 매칭이 명확하게 불가능한 아파트단지·학교 등 단지 내 도로 객체들을 미리 걸러내었다. 그 결과 374개의 링크가 제외되어 나머지 2529개의 링크에 매칭 프로세스를 적용하였다.

도로명주소 안내도 중, 도로명주소 관련 정보를 담고 있는 핵심 레이아웃 도로구간(TL_SPRD_MANAGE)을 활용하였다. 서울특별시 영역의 데이터를 도로명주소 안내시스템에서 제공받아 활용하였다(Fig. 5(b)).



Figure 5. Used data: (a) a tile of the PND set, (b) road section data from the road name address map

4.2 매칭 방법론 적용

매칭 후 성능 평가를 위하여, 전처리된 네트워크 데이터 링크 중 임의로 150개를 선택하여 수동으로 대응되는 도로명주소 도로구간 객체를 찾아 참조 데이터를 미리 구축하였다. 참조 데이터 중 118개 (78.7%) 링크가 대응쌍을 명백하게 찾을 수 있는 ‘매칭 가능’ 객체였다. 나머지 32개는 대응쌍을 찾을 수 없는 ‘매칭 불가’ 객체로서, 앞서 필터링되지 않은 단지 내 도로, 공원 내 도로, 기타 소로 등이 이에 해당되었다.

매칭 프로세스에 적용할 주요 계수들은 다음과 같이 산정하였다. 매칭 후보 탐색 범위(R_C)는 보통 보행자 길이 도로 가장자리에 그려진다는 점을 감안하여, 실험 지역 내 가장 넓은 도로 폭 값(54m, 국회대로)의 1/2 수준, 즉 27m로 설정하였다. 따라서 매칭 실패 여부를 결정하는 임계값은 $1/R_C = 0.037m^{-1}$ 로 결정되었다. 네트워크 이웃 추출 임계 거리(D_N)는 보행자용 네트워크 데이터에서 가장 긴 링크의 길이가 523.7m임을 감안하여, 그 절반인 261.9m로 설정하였다. 또한 도로명주소 도로구간 데이터의 최대 위치오차(E_{RS})는, 도로명주소 지도의 원도 격인 1:1000 축척 수치지도의 규정 최대오

Table 2. The matching table between PND and road section data (Part)

| PND link ID | Road section object ID | Similarity (m^{-1}) |
|-------------|------------------------|-------------------------|
| 0 | 40859 | 0.8592 |
| 1 | 50257 | 0.7701 |
| 2 | 50197 | 0.8223 |
| ... | ... | ... |

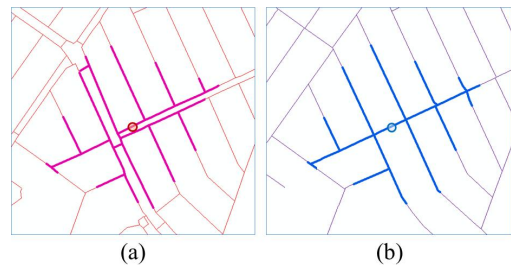


Figure 6. An example of matching by comparing network neighborhoods: (a) A PND link object and its network neighborhood drawn by thick lines, (b) One of the candidates in the road section data and its network neighborhood drawn by thick lines.

Similarity between them was $0.259m^{-1}$

차 값인 0.7m로 가정하였다(National Geographic Information Institute, 2015).

그 다음, 매칭 대상 두 데이터셋에 대하여, 네트워크 이웃 기반 매칭 기법을 적용하였다(Fig. 6). 실험은 Matlab 8.0 기반으로 구현하였으며, 그 결과는 객체 ID 가 기록된 매칭 테이블(Table 2)로 정리하였다.

4.3 매칭 결과 평가

매칭 프로세스 수행 결과, 보행자용 네트워크 데이터의 링크 2529개 중 1863개 (73.7%)에 대하여 대응되는 도로구간 매칭쌍 객체를 찾을 수 있었다. 앞서 구축한 참조 데이터를 활용하여, 매칭 결과를 평가하였다. 참조 데이터 내에서 ‘매칭 가능’으로 분류된 118개 링크 중 111개에 대하여 대응쌍을 찾아냈으며, 그 중 109개가 참조 데이터의 대응쌍과 일치한 결과를 나타내었다. 반면 ‘매칭 불가’로 분류된 32개 링크 중 27개에 대해서 대응쌍을 찾지 못하였으며 나머지 5개는 ‘매칭 불가’임에도 대응쌍이 탐색되는 결과를 얻었다(Table 3).

즉, 매칭이 이루어진 116개 링크 중 옳게 매칭된 링크는 109개로, 94.0%의 정밀도(precision)를 나타내었다. 또한 ‘매칭 가능’한 118개 링크 중 옳게 매칭된 링크는 109개이므로 92.4%의 재현율(recall)을 나타내었다. 이 결과를 이진 분류 기법의 평가에 사용되는 지표인 F-score를 통하여 평가하였다. 일반적인 F-score는 다음 Eq. (5)와 같이 계산된다(Van Rijsbergen, 1979).

$$F_{\beta} = (1 + \beta^2) \frac{P \cdot R}{(\beta^2 \cdot P) + R} \quad (5)$$

where, P and R denote precision and recall, respectively, and β denotes the importance ratio between P and R

본 연구의 결과를 길안내에 적용할 때, 도로명이 누락되는 경우보다 틀린 도로명을 안내하는 경우의 손실 비용이 높으므로, 재현율보다 정밀도에 더 높은 중요도를

Table 3. Performance of matching results

| Reference Experiment | | Matched | Unmatched | Total |
|-------------------------|-----|---------|-----------|-------|
| | | Correct | 109 | |
| Incorrect | 2 | | | |
| Total | 111 | | | |
| Unmatched | | 7 | 27 | 34 |
| Total | | 118 | 32 | 150 |

부여하였다. 중요도 차이를 2배라고 가정하면 $\beta = 0.5$ 이므로, F-score는 다음 Eq. (6)과 같이 계산된다.

$$F_{0.5} = (1 + 0.5^2) \frac{P \cdot R}{(0.5^2 \cdot P) + R} = 0.936 \quad (6)$$

한편, 매칭이 이루어진 116개 링크 중 5개는 ‘매칭 불가’임에도 매칭이 이루어졌으며, 2개는 매칭은 되었으나 틀린 도로구간 객체로 매칭되는 오류가 나타났다. 실제로 이 매칭 결과를 이용하여 보행자 길안내 서비스에 데이터를 적용한다면 사용자에게 틀린 도로명을 안내하게 되는 것이다. 따라서 실제 서비스에 적용하기 전 수작업에 의한 오류 보정 과정이 필요하다. 본 연구에서는 기하학적 유사도를 기준으로 매칭 여부를 판단하기 때문에, 이 유사도 값을 이용하여 오류 검사를 수행할 객체를 검출하였다. 수동 검사 비율을 10%로 가정하고, 매칭이 이루어진 116개 링크 중 매칭된 최대 유사도 값이 작은 순서대로 12개를 추출하여 오류 여부를 검사하였다. 그 결과 오매칭된 2개 중 2개 모두 검출되었으며, ‘매칭 불가’인 5개 중 3개를 검출할 수 있었다. 반대로 매칭에 실패한 34개 링크 중 최대 유사도 값이 큰 순서대로 4개를 추출하여 검사한 결과 2개는 ‘매칭 가능’으로 검출되었다. 따라서 본 연구에서 활용한 유사도 기준으로 10%의 수동 검사 비율을 적용하면, 정밀도는 98.3%까지, 재현율은 95.8%까지 향상시킬 수 있었다. $F_{0.5}$ 값은 0.936에서 0.978까지 향상시킬 수 있었다(Table 4).

4.4 보행자용 네트워크 데이터 갱신

위의 매칭 결과를 이용하여, 보행자용 네트워크 데이터의 각 링크에 도로명 관련 속성정보가 포함되도록 갱신하였다. 갱신 결과의 일부를 표시하면 Fig. 7과 같다.

Table 4. Performance of matching results after 10% manual check

| Reference Experiment | | Matched | Unmatched | Total |
|-------------------------|-----------|---------|-----------|-------|
| Matched | Correct | 113 | 2 | 115 |
| | Incorrect | 0 | | |
| | Total | 113 | | |
| Unmatched | | 5 | 30 | 35 |
| Total | | 118 | 32 | 150 |

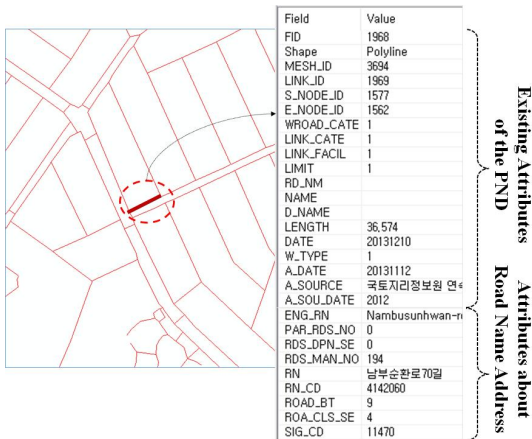


Figure 7. An example of the updated PND link with attributes about the road name address using the matching results

5. 결론

도로명주소 체계는 도로의 이름과 구간 내 위치를 기반으로 위치를 표시하는 새로운 주소체계로서, 지난 2012년부터 본격적으로 시행되었다. 그러나 실생활에서 활용되는 기존 공간정보들과 충분히 연계되지 않아 도로명주소 제도가 확실히 자리 잡지 못하고 있다. 이에 대한 해결 방안으로서, 본 연구에서는 길찾기 서비스에서 사용되는 배경 네트워크 데이터와 도로명주소 도로구간 데이터 사이의 기하학적 매칭 방법론을 제시하였다. 이를 위하여, 비교 대상 공간객체를 중심으로 네트워크 이웃을 생성하고, 그 사이의 기하학적 유사도를 측정함으로써 최적 매칭 결과를 도출하였다. 참조 데이터를 기준으로 매칭 성능을 평가한 결과, $F_{0.5}$ score를 기준으로 약 93.6%의 정확도를 나타내었다. 또한 본 연구에서 적용한 유사도 기준으로 10% 객체를 수동으로 검사하면 97.8%까지 향상시킬 수 있었다. 이러한 방법을 적용하여 네트워크 데이터에 도로명 정보를 등록함으로써, 공간 인식에 가장 중요한 길찾기 과정에서 도로명주소가 연계될 수 있도록 하였다. 본 연구의 기법을 확장하면, 다양한 형태의 공간정보들을 도로명주소와 연계할 수 있는 일반적인 주소 탐색 기법을 제안할 수 있을 것이다. 이러한 방법론을 통하여, 길찾기를 위한 기반 데이터 구축 비용 절감 및 서비스 효율성 증대, 그리고 도로명주소 제도의 확고한 정착 등의 효과를 기대할 수 있다.

감사의 글

본 연구는 "국토교통부 국토공간정보연구사업 국토공간정보의 빅데이터 관리, 분석 및 서비스 플랫폼 기술개발(16NSIP-B081011-03)과제"의 연구비 지원에 의해 연구되었음.

References

- Bang, Y. S., Ga, C. O. and Yu, K. Y., 2012, Matching and attribute conflating method for linking the digital map with the road name address system – focused on the road centerline layer, Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography, Vol. 30, No. 4, pp. 379–388.
- CMI Consortium, 2006, A study on maintenance of the address system and facilities for application of the road name address – part I, Research report, Ministry of the Interior, Republic of Korea, pp. 47–71.
- Kang, Y. O. and Jang, Y. J., 2012, A study on the integrated topological data model based on linear features of new address and national basic districts, Journal of the Korean Cartographic Association, Vol. 12, No. 3, pp. 63–79.
- Kim, K. M., 2015, Establishing walk-friendly wayfinding system, World and City, Vol. 10, pp. 24–31.
- Lee, J. S., Kim, J. H., Kim, M. G. and Yun, H. D., 2015, Error and accuracy analysis about road name address for reliability improvement and efficient utilization, Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology, Vol. 5, No. 2, pp. 223–230.
- Min, D., Zhilin, L. and Xiaoyong, C., 2007, Extended Hausdorff distance for spatial objects in GIS, International Journal of Geographical Information Science, Vol. 21, No. 4, pp. 459–475.
- Ministry of the Interior, 2013, Road name address act, The National Law Information Center, <http://www.law.go.kr/lsInfoP.do?lsiSeq=136735>
- National Geographic Information Institute, 2015, Regulations for generating digital map, The National Law Information Center, <http://www.law.go.kr/LSW/lsInfoP.do?lsiSeq=108484>
- Road Name Address, 2016, Introduction road name address, Ministry of the Interior, <http://www.juso>

- go.kr/CommonPageLink.do?link=/street/GuideBook
10. Roh, H. J., 2009, Assessment and improvement method of position accuracy of geocoding application for new address system in Korea, *Journal of the Korean Cartographic Association*, Vol. 9, No. 1, pp. 63–72.
 11. Seok, S. M. and Lee, J. Y., 2016, Development of geocoding and reverse geocoding method implemented for street-based addresses in Korea, *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, Vol. 34, No. 1, pp. 33–42.
 12. Van Rijsbergen, C. J., 1979, *Information retrieval*, Butterworths, UK, pp. 112–128.
 13. Walter, V. and Fritsch, D., 1999, Matching spatial data sets: a statistical approach, *International Journal for Geographical Information Science*, Vol. 13, No. 5, pp. 445–473.
 14. Yang, S. C., 2013, A study on construction of real estate development map using the road name address map, *Journal of the Korean Society for Geospatial Information System*, Vol. 21, No. 3, pp. 47–54.
 15. Yeom, J. H., Huh, Y. and Lee, J. B., 2014, Building matching analysis and new building update for the integrated use of the digital map and the road name address map, *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, Vol. 32, No. 5, pp. 459–467.
 16. Zhan, F. B., 1998, Shortest path algorithms: an evaluation using real road networks, *Transportation Science*, Vol. 32, No. 1, pp. 65–73.