

GIS와 국가인구통계자료 통합에 의한 입지분석용 정밀인구통계지도 구축 방법

이용익¹, 홍성언^{2*}

¹지능형국토정보기술혁신사업단

²청주대학교 지척학과

Method on Constructing Precision Population-statistical Map Integrating GIS and National Census Data for Location Analysis

Yong-ik Lee¹ and Sung-eon Hong^{2*}

¹Korean Land Spatialization Group

²Dept. of Land Management, Cheongju University

요 약 본 연구에서는 다양한 입지분석에서 이용되고 있는 인구추정의 정확도와 신뢰성의 향상을 기하고자, GIS와 국가인구통계자료를 이용하여 정밀한 인구통계지도를 구축할 수 있는 방법을 제시하였다. 제시한 방법은 다음과 같다. 주거지, 상업 및 업무지의 토지이용이 인구와 상관성이 높다는 것을 분석·도출하여, 세부 토지이용(비오톱) 유형별로 다중회귀분석을 실시하였다. 그런 후 인구거주 밀도 별로 가중치를 부여하여 인구를 동별 토지이용 유형별로 재분배하는 방법으로 정밀한 인구통계통계 지도를 구축하였다. 본 연구의 방법은 그간 다양한 입지분석에서 이용되었던 인구추정 방법보다 정확도와 신뢰성의 향상을 가져올 것으로 기대된다.

Abstract The objective of the present study lies in providing the method to construct the precision population-statistical map for statistical demographics making full advantage of GIS and the national census data in an attempt to improve accuracy and reliability of population estimation applicable for a variety of location analysis. More specifically, it adopts the multiple regression analysis by segmented land use type(biotope) taking into account that the land use diversified as residence, commercial and office areas has the close connectivity and interdependence with population. Based on the analyzed result above, the study finalizes the modeling to construct demographic map with higher precision by prioritizing the population density by weight value and then re-distributing the population according to jurisdictional dong's and types of use for the land. The study consequently is expected to be conducive to improving precision and reliability rather than the existing method for population estimation widely acceptable thus far.

Key Words : National census data, location analysis, Population estimation, Precision population-statistical map, GIS

1. 서론

2007년 4월 1일 현재 통계법 제8조(또는 제9조)에 의거 승인받은 통계는 총 744종으로서 지정통계 92종, 일반통계 652종이다. 이중 정부기관에서 작성하는 통계는

750종(통계청은 54종)이며, 지정기관에서 작성하고 있는 통계는 174종이다. 이와 같이 국가통계정보는 다양한 통계정보를 제공하고 있는 반면에 실제 사용에 있어서는 그 사용빈도가 적다[9].

이러한 문제를 해결하기 위하여 현재 통계청에서는 국

*교신저자 : 홍성언(hongsu2005@cju.ac.kr)

접수일 : 09년 10월 06일

수정일 09년 10월 30일

게재확정일 09년 11월 12일

가통계정보를 GIS 기술 요소 중 시각화 기술을 활용하여 통계지리정보시스템, 블루슈머GIS, 지도로 보는 통계(2005 센서스 전수조사 제공), 즐겨 찾는 통계지도, 움직이는 통계지도, 통계네비게이터(2007년 현재, 대전광역시만 제공) 등과 같은 다양한 통합 서비스를 시행하고 있다.

지리정보시스템(GIS)과 국가통계정보 등 속성정보의 통합에 관한 연구는 GIS가 국내에 도입된 이후로 지속적인 관심의 대상이었다. 1999년 국토연구원의 조사결과에 따르면 국가통계정보와 GIS를 접목시키면 국가통계정보의 활용도가 향상될 것이라는 응답이 80.5%로 나타나 응답자의 대부분이 GIS와 국가통계정보를 접목시키면 활용도가 향상될 것으로 기대하는 것을 알 수 있다[2]. 구자문[1]은 도시환경 분석을 위하여 미국 TIGER (typologically integrated geographic encoding and referencing) 시스템을 예로, 통계자료와 수치지도와의 통합방안을 제안하였다. 김영표와 한선희[3]는 행정기관에서 사용하는 정책자료의 80% 이상이 지리적 정보인 것을 감안하여 통계정보의 원활한 활용을 위하여 국가통계정보와 지리정보체계기법을 접목시켜 여러 가지 유형의 통계지도 개발의 필요성을 주장하였다.

다양한 입지분석에서 가장 많이 이용되는 변수는 인구 통계에 관한 사항이다. 이러한 자료 역시 대부분 정부통계자료를 바탕으로 하고 있는데, 우리나라의 경우 정부통계가 읍·면·동 행정단위 이하로 제공되지 않아 세밀한 골목단위의 정밀 상권분석에서는 인구추정 기법을 적용하고 있다[10,6]. 문제는 이러한 인구의 추정에 있어 일반적으로 특정 입지분석의 주제에 적절하게 통계수집단위를 정밀하게 세분화하는 방법을 적용하기 보다는 단순한 면적 비례에 의하여 통계치를 배분하는 방법을 취하고 있어 추정의 정확도가 낮아질 수 있다는 것이다[8].

본 연구에서는 다양한 입지분석에서 이용되고 있는 인

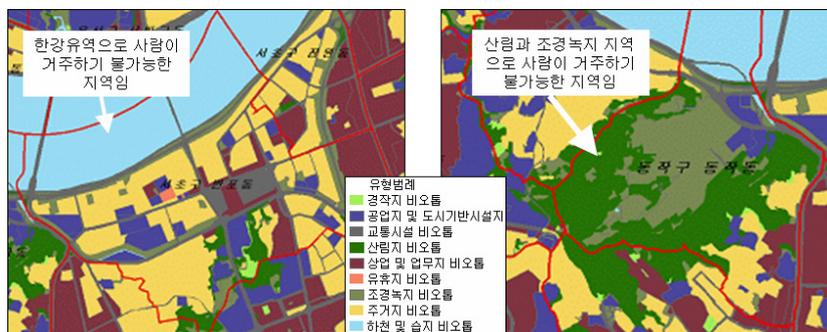
구추정의 정확도와 신뢰성의 향상을 위해 GIS와 국가인구통계자료를 이용하여 정밀한 인구통계지도를 구축할 수 있는 방법을 제시하고자 한다.

주요 연구내용과 방법으로는 먼저 인구통계자료를 현재의 구별·동별의 행정단위 보다 정밀하게 활용하기 위하여 공간데이터로 도시생태현황도와 행정구역도를, 속성데이터로 인구통계 자료를 이용하여 정밀인구통계지도 제작방법을 제시하였다. 정밀인구통계지도의 제작에 있어서는 두 종류의 수치지도(도시생태현황도/행정구역도)를 GIS의 공간분석 기법을 이용하여 하나의 수치지도로 제작하고, 속성자료는 인구와 토지이용의 상관분석과 회귀분석을 통하여 토지이용에 따른 거주지역의 가중치를 결정하여 거주가능 면적을 구하고 이를 바탕으로 인구를 재분배하였다.

2. 관련연구

입지분석을 위해서 공간데이터의 활용도 중요하지만 인구통계 데이터 역시 매우 중요한 요소이다. 일반적으로 입지분석에서 이용되는 인구데이터는 통계청에서 5년마다 전수조사되어 제공되는 데이터로, 제공의 최소 공간적 단위가 읍·면·동이다. 이러한 이유로 입지분석을 함에 있어 특정 공간(일반적으로 반경에 의한 원형으로 구획됨)의 포함 인구수 계산은 포함 면적의 비율로 인구수를 추정하는 단순 면적비례방법(simple area ratio operation)을 이용하고 있다[4,5,7,8].

다시 말해, 선행 연구에서는 통계청에서 제공하는 통계자료를 이용하는데 결정 입지 또는 상권 등의 분석에 있어 단일 행정구역 일부만이 포함될 경우 전체 인구수에 포함 혹은 인구수 배제, 또는 면적비례법칙을 이용하여 인구수를 추정하고 있다[4,7,8]. 이러한 방법들은 결과



(a)하천이 면적에 포함될 경우

(b)산림지가 면적에 포함될 경우

[그림 1] 단순면적비례에 의한 인구산정의 한계성

적으로 인구나 통계자료가 행정구역에 골고루 분포하고 있어야지만 의미가 있는 방법이다. 만약, 포함된 행정구역의 토지이용이 나대지 또는 임야, 하천 등 이라면 정확한 인구수의 추정이 이루어지지 못하는 한계성이 있다. 이렇게 현실성이 결여될 수 있는 문제점을 좀 더 구체적으로 기술하여 보면 다음과 같다.

도시생태현황도에서 제공하는 비오톱유형은 실제 도시의 토지이용과 생태현황에 대해서 서울시에서 조사한 자료를 근거로 하여 제작한 지도이다. 여러 문제점 중에서 그림 1에서 (a)의 경우, 북쪽지역의 대부분은 한강이 포함되어서 인구가 절대 존재하지 않는 지역임에도 불구하고 면적비례법칙을 이용하면 포함되는 면적의 비율만큼 인구가 추정되게 된다. 이와 유사한 경우로 (b)에서 동작구는 산림지역이 행정동면적의 대부분을 차지한다. 이처럼 사람이 거주하기 어려운 지역인 산림지도 면적비례법칙을 이용하면 사람이 거주하는 것으로 인구를 산정하게 된다.

따라서 사람이 실제 거주가 가능한 지역을 대상으로 인구를 계산하고, 이를 세분화하여 재분배 할 수 있는 새로운 방법의 모색이 필요하다.

3. 정밀인구통계지도 구축 방법

3.1 연구지역의 선정과 이용 자료

본 연구에서는 GIS와 국가인구통계자료를 통합하여 정밀인구통계지도를 구축할 수 있는 방법을 제시하고자 연구지역으로 서울시를 선정하였다. 서울시에는 현재 GIS관련된 다양한 공간데이터들이 구축되어 있다. 따라서 연구에서는 연구의 방법론을 제시하기 위한 필요데이터의 수집상황을 고려하여 서울시를 선정하였다.

연구지역인 서울시를 대상으로 정밀인구통계지를 구축하여 보고자 필요 공간데이터와 속성데이터를 수집하였다. 연구에서 이용한 공간데이터는 표 1과 같다. 서울시 행정동에 의해서 도시생태현황도를 구별하고, 정밀인

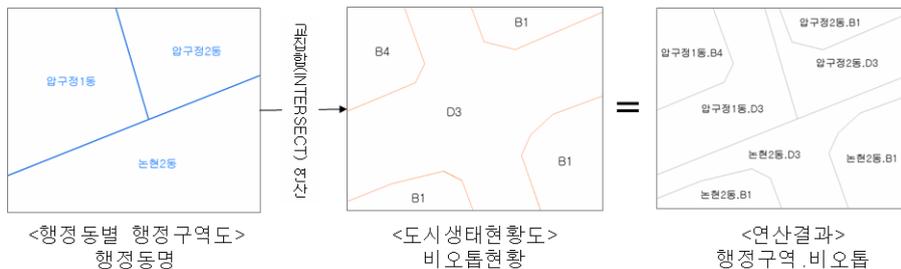
구통계지도를 구축하기 위해 도시생태현황도와 행정구역도를 기본도로 삼았다. 도시생태현황도는 서울시에서 제작한 주제도로써 서울시의 생태현황을 9개의 비오톱에 대하여 세분류로 63개로 분류한 매우 정밀한 서울시 토지 현황을 표현한 지도이다. 9개 비오톱은 토지이용현황과 매우 유사하게 나누어져 있기 때문에 토지이용에 따른 분류로 활용할 수 있다. 연구에서 정밀인구통계지도와 상권의 시설물 속성으로 이용할 테이블 형태의 자료를 정리하였다. 각 테이블은 공간 데이터를 편집한 후 인구통계 자료는 정밀인구통계지도의 속성자료로 이용하고 APT현황 자료는 아파트레이어의 속성자료로 이용된다.

【표 1】 연구에 이용된 공간 및 비공간 데이터

공간데이터		
데이터명(축척)	타입	내용
행정구역도(1:5,000)	polygon	행정동별 경계
도시생태현황도(1:5,000)	polygon	비오톱유형별 경계
아파트레이어	point	아파트
비공간 데이터		
테이블명	속성명	범주
인구통계 (행정동별)	인구수	인구학적 요인
	가구수	인구학적 요인
	가구당자동차비율	경제학적 요인
	월평균소득	경제학적 요인
	월평균식비 지출	경제학적 요인
APT현황	APT세대수	인구학적 요인

3.2 정밀인구통계 지도 구축 방법

상기에서 통계자료와 GIS기술을 통합하려는 노력과 인구통계가 입지분석에서 가지는 중요성에 대해서 살펴보고, 그에 대한 문제점을 살펴보았다. 이러한 문제점을 해결하기 위해 본 연구에서 제시하고자 하는 정밀인구통계지도 구축 방법에 대하여 개략적으로 기술하여 보면 다음과 같다. 먼저 인구를 정밀하게 배분하기 위해서는 동단위의 인구를 기반으로 하여 토지이용에 따라 정밀하게 배분하고자 한다. 이때 기준은 서울시 도시생태현황도



【그림 2】 정밀 인구 산정을 위한 공간데이터 연산

에서 제공하는 토지이용현황 자료를 이용하였다. 그 이유는 인구는 토지이용에 따라 거주가 가능하기도 하고 불가능하기도 때문이다.

정밀인구통계지도를 제작하기 위해서 GIS의 다양한 기법을 이용하였고, 공간연산에 앞서 행정동별 행정구역도와 인구통계자료를 행정동명을 이용하여 속성을 부여하였다. 그림 2와 같이, GIS 공간 연산 기법을 이용하여 정밀인구통계지도의 기반이 되는 공간데이터들을 교집합 연산을 하였다. 그 결과, 행정동 단위의 비오톱유형이 분류된 지도를 얻을 수 있었다.

속성부분으로는 행정동 단위로 각 비오톱 대분류별 면적을 구하여 합산하고, 행정동별 인구수와 각 비오톱 대분류 면적합을 상관분석을 실시하였다[표 2]. 상관분석에 이용된 다중회기 분석 모델은 일반적인 모델을 이용하였고, 상관분석시 독립변수는 면적(Bio_A12 등)으로, 종속변수는 인구수로 하여 상관성을 확인하였다. 다중상관분석 결과, 주거지, 상업 및 업무지, 유휴지 비오톱이 통계적으로 유의한 수준에서 인구와 양의 상관관계를 보였다. 하지만 유휴지는 의미상 인구가 거주하지 않는 지역이므로 제외하였다.

주거가 가능한 비오톱 유형으로 표 2와 같이 A와 B가 선정되었고, 그에 대해서 세분류를 살펴보면 표 3과 같다. 8개의 분석인자 중에서 인구수와 관련이 있는 인자를 찾기 위하여 다중회귀분석을 실시하였고, Bio_B1, Bio_B2, 그리고 Bio_B3은 유의수준이 기준을 넘어 인구

수를 결정하는 인자로서 기각하고, 나머지 5개 분석인자들을 이용하여 행정동의 사람이 거주하는 면적인 POP_Area를 구하였다.

[표 2] 동별 인구수와 비오톱 대분류 면적의 다중상관분석 결과

비오톱 대분류	상관 계수	유의 수준
A (주거지 비오톱)	0.41697	0.0001
B (상업 및 업무지 비오톱)	0.21683	0.0001
C (공업지 및 도시개발지 비오톱)	0.03478	0.4277
D (교통시설 비오톱)	0.08208	0.0609
E (조경녹지 비오톱)	0.07699	0.0789
F (하천 및 습지 비오톱)	0.05977	0.1727
G (경작지 비오톱)	-0.01282	0.7702
H (산림지 비오톱)	0.08459	0.0534
I (유휴지 비오톱)	0.11830	0.0068

POP_Area를 구하는데 있어서 2가지 방법을 비교해 보았다. 인구 거주 밀도 가중치를 고려하지 않은 모델과 가중치를 고려한 모델을 이용하였다. 가중치를 부여하기 위하여 다중회귀분석에 얻은 5개의 회귀계수를 가중치로 부여한 후 합산하였다).

서울시 522개 행정동에 대하여 두 방법의 POP_Area와 인구수와의 상관분석을 시행한 결과, 표 4와 같이 가중치를 부여한 것이 더 높은 상관성을 보이는 것을 확인할 수 있었다.

[표 3] 비오톱유형에 따른 분석인자

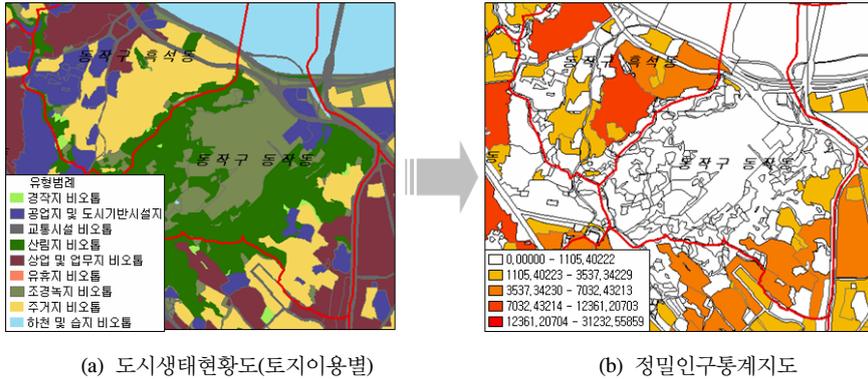
대분류	코드	비오톱유형	분석인자
A	A1	블투수포장비율이 높은(70%이상) 단독주택지	Bio_A12
	A2	블투수포장비율이 낮은(70%미만) 단독주택지	
	A3	4층 이하의 블투수포장비율이 높은 공동주택지	Bio_A34
	A4	4층 이하의 블투수포장비율이 낮은 공동주택지	
	A5	5-10층의 블투수포장비율이 높은 공동주택지	Bio_A56
	A6	5-10층의 블투수포장비율이 낮은 공동주택지	
	A7	11층 이상의 블투수포장비율이 높은 공동주택지	Bio_A78
	A8	11층 이상의 블투수포장비율이 낮은 공동주택지	
B	B1	5층 이하의 상업 및 업무지	Bio_B1
	B2	6-10층의 상업 및 업무지	Bio_B2
	B3	11층 이상의 상업 및 업무지	Bio_B3
	B4	블투수포장비율이 70%이상인 혼합지	Bio_B45 (주거와 상업용이 30%이상 혼재)
	B5	블투수포장비율이 70%미만인 혼합지	

1) POP_Area를 구하는 식은 아래의 수식과 같다.

$$POP_AREA = Bio_A12.Area * 0.02575 + Bio_A34.Area * 0.01495 + Bio_A56.Area * 0.02855 + Bio_A78.Area * 0.06038 + Bio_B45.Area * 0.03445$$

[표 4] 서울시의 아파트와 Bio_APT, 인구나 POP_Area의 상관분석 결과

독립변수	종속변수	밀도 가중치 없음	밀도 가중치 부여	증감
Bio_APT (공동주택지 총합)	APT (아파트)	0.81469	0.87508	+0.06
		<.0001	<.0001	
POP_Area (주거지역 총합)	POP (인구)	0.66333	0.77779	+0.11
		<.0001	<.0001	



[그림 3] 토지이용현황을 이용한 정밀인구통계지도

이러한 분석결과를 기초로 서울시의 25개 구에 대하여 POP_Area와 인구의 상관관계를 분석하였다. 왜냐하면, 구별마다 각기 다른 지역적 특색과 도시개발현황의 차이로 인하여 비오톱 유형에 따른 개발 밀도와 주거인구의 상관관계가 다를 수 있기 때문이다. 상관성을 확인하기 위하여 인구에 대해서는 POP_Area를, 아파트 가구수에 대해서는 Bio_APT를 관련지어 상관분석을 실시한 결과, 강남구의 Bio_APT를 제외하고는 모두 유의한 수준(0.05)에서 상관성이 향상된 것을 확인할 수 있었다. 이 결과를 전제로 하여 인구수는 주거지역의 개발 밀도에 영향을 받기 때문에 밀도 가중치를 이용하여 행정동별 내의 도시생태현황에 따라 인구를 재분배하였다.

다음의 수식은 동별 행정구역도와 도시생태현황도를 교집합연산을 한 결과물의 각 피처에 인구를 산정하는 수식이다. 다른 통계치는 인구로부터 시작하여 가구수를 구하고 구해진 가구수로 차량대수, 총수입의 합, 총식비지출의 합을 재계산하였다.

$$sub.district.pop = all.district.pop \times \frac{sub_pop_area}{pop_area}$$

(sub.district.pop: 새로 생성된 동별비오톱별면적의 인구, all.district.pop: 행정동 전체인구, sub_pop_area: 가중치를 적용한 동별비오톱면적, pop_area: 행정동 전체면적)

그림 3은 토지이용별 자료를 이용하여 동별 인구를 재

분배한 후 정밀인구통계지도를 제작한 모습이다. 그림과 같이, 주거지와 상업 및 업무지 비오톱에 대하여 인구를 재분배하였다. 그 이유는 인구나 비오톱 유형간의 통계분석에 의해 결정되어졌다. 간단히 설명하면, 도시생태현황도와 행정동별 구역도를 교집합을 실시하여 행정동 단위로 실제 비오톱 유형에 의해서 거주 가능한 지역과 불가능한 지역을 우선적으로 상관분석으로 통해 결정하였고, 주거 가능 지역에 대해서 다시 비오톱 세부 유형에 따라 가중치를 결정 하였다. 인구를 재분배함에 따라 현실과 유사하게 표현되는 것을 확인할 수 있다. 즉, 그림에서 산림지, 녹지지역, 공업지역 등은 인구수가 가장 적게 분배되었고, 주거지, 상업지 등에 인구수가 가장 많이 분배되어 있음을 확인할 수 있다(그림 (a),(b)의 범례참조).

인구수는 행정구역의 면적과 상관성이 높은 것을 확인할 수 있었다. 즉, 주거지역이 넓을수록 많은 수의 인구가 존재한다는 것이다. 정밀인구통계지도는 공간적 현상에 따른 인구분포를 섬세하게 확인할 수 있는 새로운 형태의 주제도로써 다양한 입지분석 분야에서 유용하게 사용될 것이다.

4. 결론

본 연구에서는 다양한 입지분석에서 이용되고 있는 인

구추정의 정확도와 신뢰성의 향상을 위해 GIS와 국가인구통계자료를 이용하여 정밀한 인구통계지도를 구축할 수 있는 방법을 제시하고자 하였다. 연구내용과 결과는 다음과 같다.

GIS 공간 연산 기법을 이용하여 정밀인구통계지도의 기반이 되는 공간데이터들을 교집합연산을 하여 행정동 단위의 비오톱유형이 분류된 지도를 구축하였다. 그리고 현실을 반영할 수 있는 인구수 계산을 위해, 행정동 단위로 각 비오톱 대분류별 면적을 구하여 합산하고, 행정동별 인구수와 각 비오톱 대분류 면적합을 상관분석을 실시한 결과, 주거지와 상업 및 업무지 비오톱이 인구와 상관성이 높은 것으로 분석되었다. 이를 기초로 세부 비오톱 유형별로 다중회귀분석을 실시하여 가중치를 부여한 후 인구를 동별 비오톱 유형별로 재분배하였다. 가중치를 부여하는 방법을 택한 이유는 상관분석 결과, 가중치를 부여하는 것이 높은 상관성(0.778)을 가지고 있는 것으로 나타났기 때문이었다.

본 연구의 정밀인구통계지도 구축 방법은 그간 다양한 입지분석에서 이용되었던 인구분석방법에서 보다 정확도를 향상시킬 수 있어 다양한 입지분석에 이용된다면 입지분석의 정확도 향상은 물론 신뢰도 향상에도 기여할 수 있을 것으로 기대된다.

참고문헌

[1] 구자문, “도시환경분석을 위한 센서스와 수치지도의 통합방안에 관한 연구”, 한국지리정보학회지, 1(1), pp. 39-51, 1998.

[2] 국토연구원, GIS를 활용한 국가통계정보의 활용도 제고방안 연구, 1999.

[3] 김영표, 한선희, “통계정보의 활용도 제고를 위한 통계지도 제작방안”, 국토연구, 제29권, 국토연구원, pp. 81-96, 2000.

[4] 심재현, 이성호, “대형할인점의 입지선정을 위한 의사결정에 관한 연구”, 대한토목학회논문집, 제28권 제5D호, pp. 705-712, 2008.

[5] 이상규, “대형할인점의 매출액 결정에 있어서 입지요인의 영향에 관한 연구”, 석사학위논문, 인하대학교 대학원, 2004.

[6] 이은영, 서원석, 이승곤, “GIS 분석을 통한 효과적인 Bran Chain 입지선정방안-스타벅스 코리아 사례를 중심으로-”, 관광연구저널 22(2), 한국관광연구학회, pp. 143-157, 2008.

[7] 이용익, 홍성언, 김정엽, 박수홍, “공간연관규칙을 이용한 대형할인점의 입지분석”, 대한지리학회지, 제41

권 제3호, pp. 319-330, 2006.

[8] 홍성언, 박수홍, “GIS와 AHP 의사결정 방법을 이용한 도시 근린공원의 입지 분석”, 대한지리학회지, 제38권 제5호, pp. 849-860, 2003.

[9] 통계청 홈페이지, <http://www.nso.go.kr>

[10] OPENmate 홈페이지, <http://www.openmate.co.kr>

이용익(Yong-ik Lee)

[정회원]



- 2005년 2월 : 인하대학교 지리정보공학과 (공학사)
- 2007년 8월 : 인하대학교 지리정보공학과 (공학석사)
- 2007년 9월 ~ 현재 : 지능형국토정보기술혁신사업단 연구원

<관심분야>

공간정보 서비스 모델, 공간분석(상권분석), 공간데이터 마이닝

홍성언(Sung-eon Hong)

[정회원]



- 2002년 2월 : 청주대학교 지적학과 (행정학석사)
- 2005년 8월 : 인하대학교 지리정보공학과 (공학박사)
- 2006년 3월 ~ 현재 : 청주대학교 지적학과 교수

<관심분야>

지적측량, GIS, LIS, SMCDM