

부동산 가격변동 핫스팟 탐색을 위한 공간통계기법*

손학기** · 박기호***

A Spatial Statistical Method for Exploring Hotspots of House Price Volatility*

Hakgi Sohn** · Keyho Park***

요약 : 투기가 발생할 가능성이 높은 지역은 일정지역 내의 대다수 경제주체가 적응적 소유자와 수요자일 때 형성된다. 이 지역의 가격변동은 타 지역에 비해서 가격상승 폭이 크고, 개별 부동산들의 주변 부동산들과 가격변동의 방향이 동질적인 특성을 가진 가격변동 핫스팟 패턴을 형성한다. 본 연구의 목적은 투기과열지역을 정량적으로 탐색하기 위한 가격변동 핫스팟 탐색법을 개발하는 것이다. 가격변동 핫스팟 탐색법은 크게 2단계로 구성된다. 첫째 단계는 정규모형의 공간스캔통계량을 이용하여 타 지역에 비해 높은 가격상승이 이루어진 공간클러스터를 탐색한다. 둘째 단계는 국지 Moran I를 이용하여 공간클러스터 내의 개별부동산들이 그 주변 부동산과의 가격변동 방향이 동질적인가, 즉 공간연관성을 가지는가를 평가한다. 개발된 방법을 공간적으로는 참여정부에서 부동산 문제의 중심으로 알려진 서울시 강남·서초·송파구에 적용하였고, 시간적으로는 참여정부 주요 부동산 대책의 하나인 10.29 대책을 전후로 한 2003년 8,9,10,11월의 가격변동 자료에 적용하였다. 10.29 대책발표 전인 8, 9월에는 개포동을 중심으로 가격변동 핫스팟이 발견되었고, 10월은 10.29 대책에 의해서 소강상태를 보이다가 11월에는 가격변동 콜드스팟이 발견되었다. 이 결과는 제안된 방법이 기존 단순 시각화만을 통해서 탐색할 수 없었던 투기과열지역을 정량적 방법을 통해서 시공간적으로 탐색할 수 있음을 보여준다.

주요어 : 가격변동 핫스팟, 투기과열지역, 공간클러스터, 공간스캔통계량, 공간연관

Abstract : The purpose of this paper is to develop a method for exploring hotspot patterns of house price volatility where there is a high fluctuation in price and homogeneity of direction of price volatility. These patterns are formed when the majority of householders in an area show an adaptive tendency in their decision making. This paper suggests a method that consists of two analytical parts. The first part uses spatial scan statistics to detect spatial clusters of houses with a positive range of price volatility. The second part utilizes local Moran's I to evaluate the homogeneity of direction of price volatility within each cluster. The method is applied to the areas of Gangnam-Gu, Seocho-Gu, and Songpa-Gu in Seoul from August to November of 2003; the Participatory Government of Korea designated these areas and this period as the most speculative. The results of the analysis show that the area around Gaepo-Dong was as a hotspot before the Government's anti-speculative 10.29 policy in 2003; the house prices in the same area stabilized in October, 2003 and the area was identified as a coldspot in December, 2003. This case study shows that the suggested method enables exploration of hotspot of house price volatility at micro spatial scales which had not been detected by visual analysis.

Key Words : hotspot of house price volatility, overheated speculation area, spatial cluster, spatial scan statistics, spatial association

* 본 연구는 학술진흥재단의 2005년 신진연구인력 연구장려금 지원사업에 의해서 지원을 받았습니다.

** 국토연구원 책임연구원(Associate Research Fellow, Korea Research Institute for Human and Settlements), hgsohn@krihs.re.kr

*** 서울대지리학과 교수(professor, Department of Geogrpahy, Seoul National University), khp@snu.ac.kr

1. 서론

투기과열지역은 부동산시장 활성화 정책 등으로 인하여 불확실성이 높아진 상태에서, 위협에도 불구하고 높은 기대이익을 추구하는 적응적 소유자와 수요자의 경제행위에 의해서 형성된다. 적응적 소유자는 불법행위에 대한 불이익의 위협에도 불구하고 주변 부동산과 담합 등의 전략적 제휴를 통해서 인위적으로 공급을 제한하거나 부동산 가격을 끌어 올림으로써 높은 기대이익을 추구하는 경제행위 경향이 있다. 적응적 수요자는 높은 기대이익을 추구하기 위해서 타인의 의사결정에 편승하고 타 수요자에 비해 먼저 매수하기 위해서 웃돈을 지불하는 경제행위 경향을 가지고 있다. 이 적응적 수요자가 일정 지역에 발생하면 공급을 초과하는 가수요가 형성되어 급격한 가격상승이 이루어진다. 이 적응적 소유자와 수요자의 경제행위가 공간상에 중첩되어 이루어지는 지역은 타 지역에 비해서 가격상승 폭이 크고, 개별 부동산의 가격변동의 방향이 동질적인 특성을 가진 가격변동 핫스팟 패턴을 형성한다(손학기, 2008).

이 가격변동 핫스팟은 불확실성의 수준에 따라 소유자와 수요자의 경제행위 경향에 의해서 동적으로 결정되기 때문에 우리는 그 위치, 크기, 모양을 사전에 파악할 수 없다. 이러한 특성에 의해서 행정구역 또는 고정된 공간단위를 기반으로 투기의 발생가능성이 높은 지역을 탐색하고자 하는 토지투기 지표(건교부, 2005) 및 각종 부동산정책(투기과열지역: 주택건설촉진법 32조의 5, 투기지역: 소득세법 제96조 1)은 다음과 같은 한계가 있다. 첫째, 가격급등이 발생한 핫스팟 지역이 행정구역 내에서 작은 비중을 차지할 경우에 실제 이 지역에 가격이 급등하더라도 행정구역의 평균이 크게 변하지 않기 때문에 이 지역 탐색 민감도가 떨어질 수 있다. 둘째, 핫스팟이 행정경계에 걸쳐 있을 경우 이 지역을 탐색하지 못한다. 셋째, 기존 방법이 통계적 검정을 포함하지 않고 있기 때문에 탐색 결과를 객관적으로 신뢰하기 어렵다. 이와 같은 한계에 의해서 정부는 선제적으로 대응할 수 없을 뿐만 아니라 사회적 문제로 대두된 이후에야 대응하게 된다. 하지만 그 중요

성에 비해 아직까지 가격변동 핫스팟을 탐색할 수 있는 방법은 체계적으로 연구된 바 없다.

이에 본 연구의 목적은 공간통계적 기법을 활용하여 가격변동 핫스팟 탐색법을 개발함으로써 근사적으로 투기과열지역일 가능성이 높은 지역을 찾고자 한다. 2장에서는 가격변동 핫스팟의 개념과 기존 연구에서 핫스팟을 탐색하는 방법을 검토한다. 3장에서는 공간통계를 이용하여 가격변동 핫스팟을 탐색하는 방법을 개발한다. 4장은 개발된 방법을 사례지역에 적용하여 경험적으로 알려진 투기과열지역과 비교한다. 5장은 연구를 요약하고 발전방향을 논의한다.

2. 가격변동 핫스팟에 대한 선행연구

1) 가격변동 핫스팟

아담 스미스 이래의 고전경제학이 전제하는 경제주체의 인간상은 합리적으로 선택·판단하고 효용의 극대화를 추구하는 경제적 인간이다. 하지만 현실에서 만나게 되는 인간상은 합리적 인간과 함께 휴리스틱(heuristic)과 바이어스(bias)를 가진 비합리적인 또는 준합리적인 특성을 가진 경제적 인간이다(이명희(역), 2007).

손학기(2008)는 부동산 시장의 가격이 합리적인 경제주체 뿐만 아니라 준합리적인 또는 적응적인 경제주체의 경제행위에 의해서 결정된다고 가정하고, 각각의 경우에 있어 부동산 소유자와 수요자의 경제행위 경향을 유형화하였다. 그리고 개별 부동산에서의 경제행위가 각 유형의 부동산 소유자와 수요자에 의해서 이루어질 때 발생할 수 있는 가격변동의 공간패턴을 그림 1과 같이 모형화 하였다. 시장 내에서 발생할 수 있는 가격변동의 다양한 공간패턴 중에서 그림 1의 (a), (b), (d) 내의 타원과 같이 일정지역이 타 지역에 비해서 가격상승 폭이 크고(공간클러스터), 그 내부의 개별 부동산의 가격변동의 방향이 주변 부동산의 가격변동방향과 동질적일 때(공간연관관계) 가격변동 핫스팟(hotspot of house price volatility)이라 정의하였다.

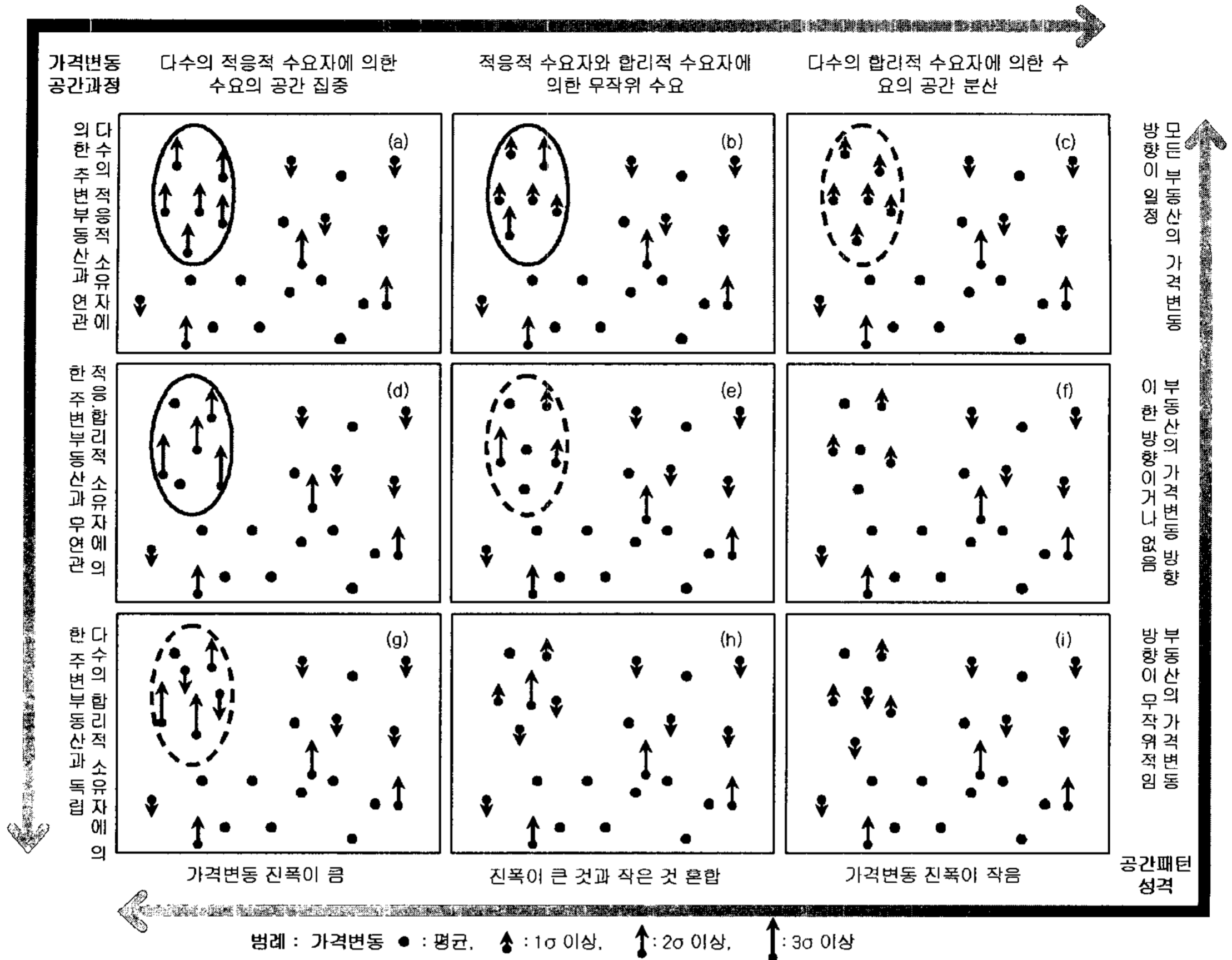


그림 1. 개별 부동산의 가격변동 공간과정에 의해서 발생할 수 있는 공간패턴 예시: (a), (b), (d)의 타원은 가격변동 핫스팟, (f), (h), (i)의 타원은 무작위 가격변동패턴; 그림에서 각 점은 부동산의 위치를 나타내고, 전체지역의 모든 부동산의 평균을●, 표준편차(σ)를 화살표(↑)로 표현하였고, 표준편차가 커지면 길이가 긴 화살표를 사용함(원자료: 손학기(2008), 그림3).

이 패턴은 시장에서 불확실성과 함께 가격상승에 대한 기대가 높을 때 적응적 소유자와 수요자의 경제행위가 공간적으로 중첩되는 곳에 형성될 가능성이 높다고 주장하였다. 반대로 일정지역이 타 지역에 비해서 가격 하락 폭이 크고, 개별 부동산의 가격변동 방향이 일정할 때 가격변동 콜드스팟(coldspot of house price volatility)으로 정의하였다. 이 패턴은 핫스팟과 화살표의 방향이 반대라는 점 이외에 다른 것은 동일하다. 마지막으로 이 두 가지 경우에 포함되지 않는 경우, 즉 그림 1의 (f), (h), (i)의 타원과 같이 일정지역의 가격변동 폭이 타 지역과 비슷하거나 가격변동 방향이 일정하지 않을 때를 가격변동 무작위 패턴(random pattern of house price volatility)으로 정의하였다. 이 패턴은

합리적 소유자와 수요자의 경제행위에 의해서 형성될 가능성이 높은 또는 우연에 의해서 발생할 가능성이 높다고 주장하였다. 그리고 가격변동 핫스팟을 탐색함으로써 투기과열지역을 근사적으로 탐색할 수 있다고 주장하였다. 가격변동 핫스팟을 통해서 탐색된 지역은 부동산 투기의 결과로 이례적인 가격변동이 발생한 지역으로 판단할 수 있고, 투기과열의 가능성이 높은 지역이라 할 수 있다.

2) 핫스팟 탐색 관련 연구에 대한 검토

가격변동 핫스팟을 관찰하고자 할 때 개별 부동산의 평균대비 가격변동 편차를 지도상에 표현하여 공간 클

러스터 패턴과 같은 특정 패턴을 시각적으로 탐색하는 것은 매우 자연스러운 탐색 방법이라 할 수 있다. 하지만 견고한 지도학적 원리를 적용해서 만든 패턴일 경우라도 적절한 해석을 얻을 수 있다고 확신할 수 없다 (Rheingans and Landreth 1995; Gahegan and O'Brien 1997; MacEachren and Kraak 1997). 즉 지도학적 원리에 의한 시각화를 통해서 생성된 공간 클러스터 패턴은 직감적으로 파악하는데 도움을 주지만 연구자의 주관이 개입될 여지가 있고, 공간클러스터가 존재하는데 탐색하지 못할 경우와 반대로 존재하지 않는데 탐색할 가능성도 존재한다. 또한 그 경계가 연구자마다 주관적으로 경계를 설정할 수 있어 반복적인 분석에서 동일한 결과를 얻을 수 없다. 하지만 Cleveland(1993)은 정량적으로 분석된 결과(통계적 분석 정보)와 함께 지도학적 원리에 의한 시각화를 결합함으로써 자료가 가진 패턴에 대한 적절한 해석이 가능하다고 주장하였다.

정량적 분석을 통한 핫스팟 탐색은 범죄의 집중발생 지역과 질병의 집중발생지역을 찾는 과정에서 주로 사용되고 있다. 범죄가 특정 장소에 집중되는 패턴을 갖는다는 사실은 Brantingham 과 Brantingham(1982)에 의해서 처음 알려졌고, 이 지역에 대한 순찰활동이 범죄를 억제할 수 있음이 실증적으로 밝혀졌다 (Sherman, 1995; Spelman, 1995). 이러한 범죄가 집중하는 지역은 흔히 범죄 핫스팟으로 불리며, 범죄사건이 집중되는 확인할 수 있는 경계를 갖는 장소 혹은 적은 규모의 지역으로 정의된다(Anselin *et al.*, 2000; Levine, 2006). 이 지역은 범죄가 무작위로 발생하는 타 지역에 비해서 범죄가 이례적으로 많이 발생하는 지역이라 할 수도 있다.

범죄 핫스팟 지역을 탐색하기 위한 시도는 범죄발생 위치를 지도에 핀으로 표시하는 핀맵(pin-map)에서 공간통계와 GIS를 활용한 범죄패턴 분석까지 다양하다 (Hirschfield *et al.*, 1995). Sherman과 Weisburd (1995)는 위치, 시간 간격, 사건의 수의 관점에서 핫스팟의 경계를 고정된 경계(예, 행정경계나 방형 그리드)를 이용하여 객관적으로 정의하였다. 고정된 핫스팟 경계는 이 지역의 시계열 자료를 활용한 추가적인 분석 및 보고서 작성 등의 장점이 있지만 실제 핫스팟이

고정된 핫스팟 경계를 가로지르거나 크기가 변하게 된다는 단점이 있다.(Anselin *et al.*, 2000). 이러한 문제를 해결하기 위해서 Block(1994)은 STAC(Spatial and Temporal Analysis of Crime) 소프트웨어에서 관찰된 범죄자료를 대상으로 특별히 고안된 클러스터링 알고리즘을 통해서 타원을 동적인 핫스팟으로 사용하였다. 한편, LISA(local indicator of spatial association)의 일종인 국지적 Moran I를 이용하여 한 지역의 강도 등의 범죄 행위가 타 지역과 공간적으로 밀접하게 연관되어 있는가를 판단하기도 한다(Anselin and Bao, 1997; Murray *et al.*, 2001).

질병의 집중발생지역 탐색 연구는 존 스노우에 의해서 시작되었다고 할 수 있다. 존 스노우는 질병이 발생하는 위치를 지도화함으로써 콜레라의 발생 원인이 상수도 시스템과 깊이 관련되어 있음을 밝혀내었다(Cliff and Haggett, 1988). 그 이후로 질병이 우연에 의해서 발생할 것 같지 않을 정도로 지리적으로 제한된 지역에 충분한 크기로 발생하는 패턴, 즉 질병의 공간클러스터 패턴(Knox, 1989) 또는 질병 핫스팟을 통계적으로 찾기 위한 여러 탐색법이 제안되어 왔다(Elliot *et al.*, 2000). 이 방법들이 질병에 대한 특별한 병리적 가정(aetiological hypothesis)이 없는 상태인 질병의 창궐 초기에 적용되면 질병 핫스팟의 발생과 그 공간범위를 탐색할 수 있는 능력을 제공할 수 있기 때문에 최근에는 보건관련 바이오테러리즘 활동(health-related bioterrorism activity) 감시에도 사용되고 있다(Lawson and Kleinman, 2005).

Sankoh and Becher(2002)는 기존에 질병클러스터 탐색에 사용된 방법들을 자료의 형태와 탐색법의 목적에 따라 분류하여 제시하였다. 표 1은 그 중에서 점형 자료의 경우에 사용할 수 있는 방법을 정리한 것이다. 이 표는 연구 지역 전체에서 공간클러스터의 존재 여부를 판단하는 전역적(global) 분석법, 기대 이상으로 많은 혹은 적은 사건들이 발생한 핫스팟 지역이 어디인가를 판단하는 국지적(local) 분석법, 그리고 특정 지역 주변에 기대 이상으로 많은 사건이 발생하였는가를 판단하는 집중(focused) 분석법으로 나누어진다. 이 중에서 가격변동 핫스팟 탐색은 핫스팟이 발생한 지역을 탐색하기 때문에 국지적 분석법에 속한다. 이들 중

표 1. 질병클러스터 탐색에 주로 사용하는 방법 (원자료: Sankoh and Becher(2002), 표 , 수정)

구분	방법
전역분석 (global)	Moran's I(1950), Horn's Ro(1966), Mantel & Bailer(1970), Ripley's K Function(1980), Esseen's test(1983), Whittemore et al.(1987), k-NN Cuzick & Edwards(1990), Alexander's NNA(1991), Alt & Vach(1991), Grimson's U(1993), Tango's Excess Events Test(1995), Anderson & Titteringham's (1997)
국지분석 (local)	Turnbull's CEEP(1990), Besag & Newell's R(1993), Grimson's MAX(1993), Spatial Scan Kulldorff & Nagarwalla(1995), Anderson-Titteringham(1997), Bithell's M with uniform kernel(1999), Isotonic spatial scan Kulldorff(1999)
집중분석 (focused)	Fixed Cut-off Lyon et al.(1981), Isotonic regression Stone(1988), Stone MLR(1988), Hills & Alexander's Z(1989), Focused k-NN Cuzick & Edwards(1990), Focused Besag & Newell(1991), Score Test Waller et al.(1992) Lawson(1993), Bithell's linear rank score(1995), Maximum X2 Lagazi et al. (1996), Cumulative X2 Lagazi et al.(1996), Diggle et al.(1999)

에서 국지적 공간클러스터 탐색법은 Openshaw et al.(1999)의 GAM(Geographic Analysis Machine)에서 시작하여 현재 Kulldorff(1997)의 공간스캔통계량까지 발전되어 왔다. 공간스캔통계량은 동적으로 공간클러스터를 탐색하면서도 우도비(likelihood ratio)를 통해서 기존 국지적 공간클러스터 탐색법의 문제점이었던 다중 테스트 문제(multiple test problem)을 해결한 방법이다. 이 방법은 현재 대부분의 질병 클러스터 탐색 연구에 이용되고 있다.

이상에서 검토한 범죄나 질병분야의 핫스팟 탐색법을 부동산 가격변동 핫스팟을 탐색하기 위한 정량적 분석법으로 직접적으로 적용할 수 없다. 그 이유는 부동산 가격변동 핫스팟이 타 지역에 비해서 높은 가격 상승이 일어난 공간클러스터인 동시에 공간클러스터 내의 개별 부동산의 가격변동이 그 주변 부동산과 가격변동 방향이 동질적인, 즉 공간적으로 연관된 지역이라는 2가지 특성을 가지고 있기 때문이다. 범죄 및 질병에서 널리 사용되는 방법들은 위에서 검토하였듯이 이들 중에서 1가지 특성만을 탐색하는 방법들이다. 또한 검토된 대부분의 제안된 탐색법들이 범죄 및 질병의 발생 건수자료와 같은 이벤트자료를 이용한 공간클러스터 탐색법이다. 최근에 와서야 Agarwal et al. (2006)에 의해서 가격변동과 같은 연속형 자료를 대상으로 하는 공간클러스터 탐색법이 제안된 상태이다.

3. 가격변동 핫스팟 탐색법 개발

1) 가격변동 핫스팟 탐색법의 개요

일반 재화시장에서 볼 수 있는 동일 상품의 동일 가격체계와 달리, 개별 부동산의 가격은 개별성과 위치 고정성에 의해서 각 부동산별로 다르게 형성된다. 심지어 동일 시점에 지어진 동일 아파트의 층별 가격도 다를 수 있다(강원철 등, 2004). 하지만 특정 기준 시점에 대비해 볼 때 각 부동산의 가격변동 차이는 크지 않을 것이다. 즉 비슷한 주택형(평형)일 경우 주변 아파트의 가격변동은 준공 일시나 시공사가 다를 경우라도 비슷한 정도로 이루어진다. 이것은 호재와 같은 가격 상승 요인이 영향권 내의 대부분의 부동산에 비슷한 수준으로 영향을 미치기 때문이다. 이때의 가격변동 상태는 고전경제학이 전제하는 합리적 경제주체의 경제행위에 의해서 우연하게 형성될 수 있는 또는 인의적인 의도가 배제된 상태에서 형성될 수 있는 가격변동 상태이고, 공간적으로는 공간 무작위 가격변동 패턴이라 할 수 있다. 반면에 가격변동 핫스팟은 불확실성 속에서도 높은 미래 기대이익을 바라는 적응적 경제주체의 경제행위 의해서 인위적으로 형성될 수 있는 가격변동 상태이고, 타 지역에 비해서 가격상승 폭이 크고 그 주변부동산과 가격변동 방향이 동질적인 지역

으로 정의된다.

가격변동 핫스팟을 탐색하기 위해서 본 연구에서는 모든 개별 부동산이 위치(x, y 좌표)와 기준 시점대비 분석시점의 가격 변동률을 속성으로 가지고 있다고 가정한다. 가격변동 핫스팟 탐색법은 그림 2와 같이 4단계로 구성하였다. 1단계는 가격변동 핫스팟이 될 가능성이 있는 모든 후보 지역을 개별 부동산의 위치를 중심으로 가변 원이나 타원 등으로 조합함으로써 생성하는 부분이다. 2단계는 타 지역에 비해서 가격상승 폭이 큰 공간클러스터를 탐색 부분이다. 3단계는 주변 부동산과 가격변동 방향이 동질적인가를 판단하는 공간 연관성 평가하는 부분이다. 4단계는 다수준(multi-level)으로 탐색된 가격변동 핫스팟을 중첩함으로써 단일 수준으로 탐색할 때 잃을 수 있는 가격변동 핫스팟의 크기 및 모양에 대한 상세함을 유지하는 부분이다. 2단계에서 공간클러스터가 탐색될 때에는 3단계로 진행되지만, 공간클러스터가 탐색되지 않을 시에는 다른 가격변동 핫스팟 후보군을 대상으로 1단계부터 다시 시작한다. 3단계의 공간연관관계 평가를 통과할 경우 4단계로 진행되지만, 공간연관관계를 통과하지 못할 때에는 다른 가격변동 핫스팟 후보군을 대상으로 1단계부

터 다시 시작한다. 이 탐색법에 있어서 주요 부분은 가격변동의 공간클러스터를 탐색하는 2단계와 주변 부동산과의 공간연관성을 평가하는 3단계이라 할 수 있다. 이어지는 절에서는 위에서 개념적으로 설명된 과정을 상세히 설명한다.

2) 모형의 단계별 설명

(1) 가격변동 핫스팟의 후보군 생성

① 탐색 도형의 모양 선정

가격변동 핫스팟 지역의 경계는 행정경계와 같이 고정된 공간범위를 갖는 것이 아니라, 특정 시점에서 적응적 경제주체의 경제행위에 의해서 동적으로 결정된다. 이 지역은 전체 연구지역의 평균 가격변동을 기준으로 동적으로 형성되기 때문에 사전에 그 위치, 크기, 그리고 모양을 알 수 없지만, 최소한 특정 중심점을 기준으로 일정 반경을 공유하는 아메바(ameba)와 같은 형태일 것이라 쉽게 생각할 수 있다.

본 연구가 불규칙적이고 동적으로 형성되는 가상의 경계를 정확히 찾는 것이 아니라 투기과열지역 등 특정 현상이 발생할 가능성이 높은 지역을 찾는 것이기

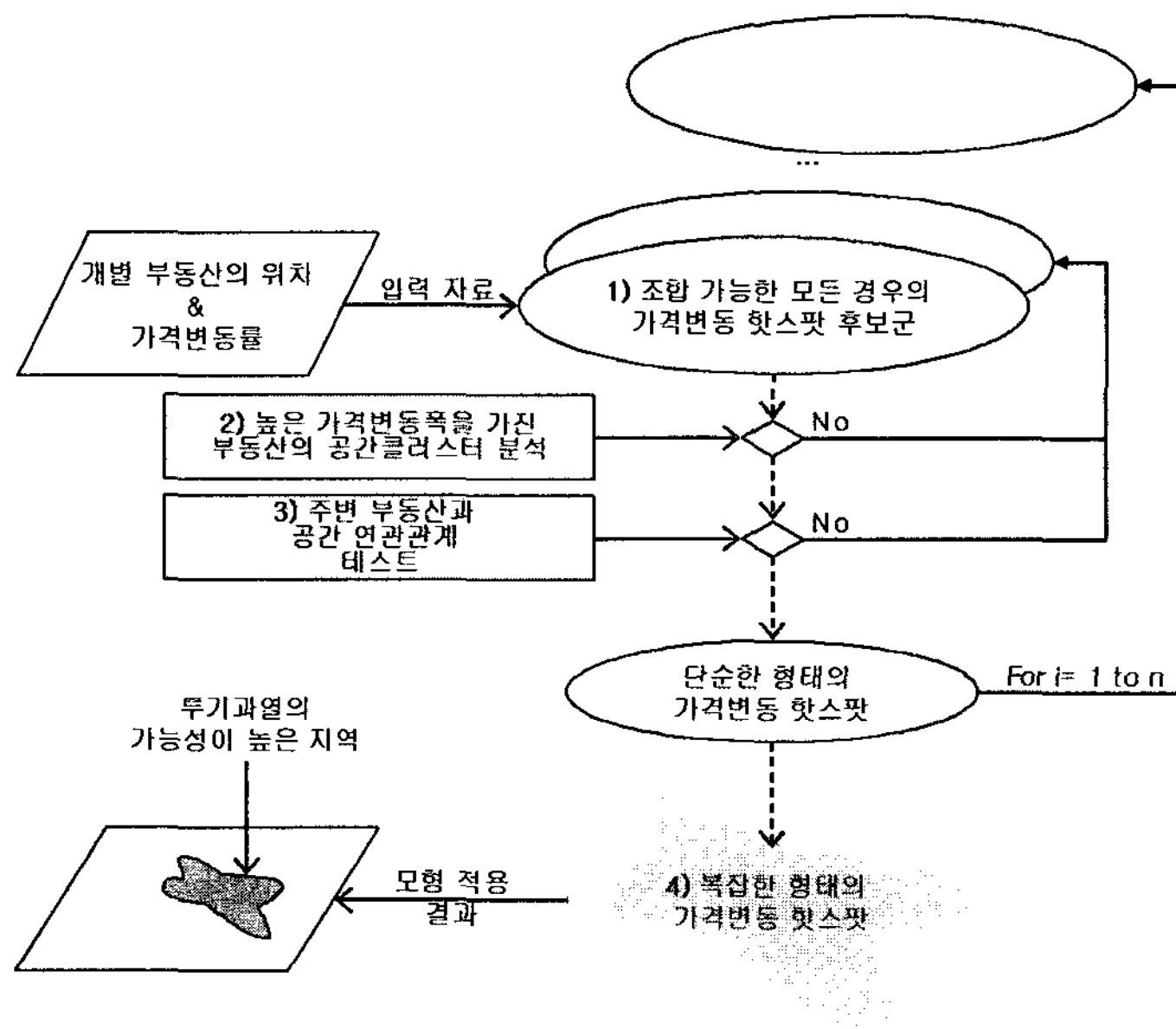


그림 2. 가격변동 핫스팟 탐색 모형

때문에 원, 타원, 직사각형 등의 비교적 단순하지만 그 지역을 잘 대표할 수 있는 도형을 이용하여 이 지역을 탐색할 수 있다. 원은 동일한 반경을 갖는 지역을 나타내기 때문에 공간통계 분야에서 널리 이용되어 왔고 (Openshaw *et al.*, 1987; Besag and Newell, 1991; Tunbull *et al.*, 1990; kullodorff, 1997), 타원은 원으로 탐색할 수 없었던 방향성이 있는 핫스팟을 탐색하기 위해서 사용되어져 왔으며(kullodorff, 2006), 직사각형은 원을 이용할 때보다 빠르게 공간클러스터를 탐색하기 위해서 사용되어 왔다(Neill and Moore, 2005).

본 연구에서는 원에 비해 방향성을 고려하여 탐색할 수 있는 타원을 탐색 도형으로 사용한다. 타원은 중심점 좌표와 이심률(eccentricity), 각(angle), 크기(size) 중 하나만 다를지라도 다른 타원으로 정의되기 때문에 원에 비해 다양한 조합으로 가격변동 핫스팟의 가능성을 평가할 수 있다. 타원의 중심점 좌표는 원과 같으며, 이심률은 단축반경에 대한 장축반경의 비로 표현된다. 이심률이 증가할수록 길고(long) 좁은(narrow) 타원을 의미한다. 즉 그 비율이 1인 경우 특별한 경우로 원을 나타내고, 2인 경우 단축에 비해서 장축이 2배인 타원이 된다. 이심률이 큰 타원을 이용하여 도로나 하천을 따라 형성된 핫스팟을 탐색할 수 있지만, 이심률이 너무 큰 경우는, 실질적 관계가 없음에도 불구하고 단지 통계적 조건을 만족함으로써 핫스팟으로 잘못 탐색할 가능성이 있기 때문에 주의하여야 한다. 그리고 방향성은 모든 방향(360°)에 대해서 고려할 수 있겠지만 계산의 편이를 위해서 15° 또는 30° 간격으로 주요 축을 설정할 수 있다. 또한 타원은 대칭성이 있기 때문에 1사분면에서의 축 설정으로 모든 가능한 타원을 설정할 수 있다. 따라서 15° 간격으로 타원의 주요 축을 설정하는 경우 0, 15, 30, 45, 60, 75, 90, -15, -30, -45, -60, -75를 주요 축으로 이용한다.

② 탐색 도형의 중심점과 크기 선정

우리는 가격변동 핫스팟의 크기와 위치를 사전에 알 수 없기 때문에 실제 가격변동 핫스팟이 될 수 있는 가능성이 높은 다양한 후보군을 만들어야 한다. 후보군을 만들기 위해서는 타원의 중심점과 타원의 크기가 결정되어야 한다. 탐색도형의 중심점은 임의의 모눈

(grid)의 교차점(Openshaw *et al.*, 1987), 자료의 위치(Besag and Newell, 1991; kullodorff, 1997), 무작위로 선정된 위치(Fortheringham and Zhan, 1996) 등이 사용되고 있다. 이 중심점의 결정은 절대적인 기준이 있는 것이 아니라 탐색하고자 하는 현상과 제일 근접한 결과를 도출해 낼 수 있고 보다 객관적인 방법이 선호되고 있다. 모눈과 무작위 방법은 그 상세도를 주관적으로 결정해야 하는 단점이 있기 때문에 자료의 위치를 핫스팟의 중심이라 가정한 '자료의 위치를 중심으로 하는 방법'이 많은 연구에서 사용하고 있다. 본 연구에서는 자료의 위치를 중심으로 타원의 중심을 설정한다.

탐색도형의 크기는 가격변동 핫스팟의 크기가 사전에 결정되어 있지 않기 때문에 작은 크기에서 큰 크기의 가격변동 핫스팟을 탐색할 수 있도록 가변 타원을 사용한다. 가변 타원은 타원의 중심점을 중심으로 일정거리 만큼의 장축 또는 단축의 길이를 증가시킴으로써 만들 수 있다. 가변 타원의 증가 폭을 결정하는 방법은 2가지가 있다. 하나는 연구자가 정한 고정된 거리를 사용하는 방법이다(Openshaw *et al.*, 1987). 다른 하나는 각 탐색의 중심점에서 다른 모든 탐색점과의 거리를 미리 계산하여 순위를 산정하고, 작은 순위에서 큰 순위로 이동하는데 그 때의 거리를 가변 타원의 증가 폭으로 사용하는 방법이다(Kullodorff, 1997). 본 연구에서는 다른 탐색점과의 거리를 이용하여 가변 타원을 만드는 방법을 사용한다.

(2) 가격상승 폭이 큰 개별 부동산의 공간집중성 평가

① 공간스캔통계량에 의한 가격상승의 공간클러스터 탐색

가격상승의 공간클러스터는 그림1의 (a), (b), (d)와 같이 전체지역 평균 대비 높은 가격상승이 발생한 개별 부동산의 공간적 집적 지역을 의미한다. 공간스캔통계량은 무수히 많은 가격변동 후보군 중에서 후보군 자신과 자신 이외의 지역을 비교할 때 가장 크게 차이가 나는 지역을 찾는 방법이다. 다시 말해 "R지역(ex, 가격상승 공간클러스터)에서 관찰된 부동산의 가격변동의 평균이 Rc지역(가격상승 공간클러스터 이외의 지

역)에서 관찰된 부동산의 가격변동의 평균과 통계적으로 다르다고 볼 수 있는가?”라는 통계적 차이 함수 (statistical discrepancy function)로 정의된다(Agarwal *et al.*, 2006). 이 함수는 우도비 테스트(likelihood ratio test)로 유도되고, 이 우도비 테스트는 Rc지역과 최대로 차이가 나는 R지역을 탐색할 수 있는 최대의 통계적 파워(most statistical power)를 가진 통계량임이 Kulldorff(1997)에 의해 증명되었다. 이 우도비 테스트는 H0의 우도 대비 H1의 우도의 비를 측정한다. 이 테스트의 결과는 H1의 강도에 대한 측정치이다. 이 숫자가 더 크면 클수록 H1은 참(true)이 될 가능성이 높아지고, R지역은 참의 공간클러스터를 대표할 가능성이 커진다.

통계적 공간클러스터 탐색을 위한 우도비 테스트는 다음을 가정한다.

- H0 : R지역의 가격변동 분포(m_r)
= Rc지역의 가격변동 분포(b_r)
- H1 : R지역의 가격변동 분포(m_r)
≠ Rc지역의 가격변동 분포(b_r)

공간스캔통계량은 자료의 분포에 따라 각기 다른 모형이 개발되어 있다. 질병에 걸린 사람(실험군)과 질병에 걸리지 않은 사람(대조군)으로 이루어진 자료는 Bernoulli 모형(Kulldorff *et al.*, 1995, Kulldorff, 1997), 일정 지역의 환자의 수에 관한 자료는 포아송 모형(kulldorff, 1997), 연령별 암환자에 관한 자료는 오디널 모형(Jung *et al.*, 2006), 생존시간에 관한 자료는 지수 모형(Huang *et al.*, 2006), 그리고 연속형 자

료에 대해서는 정규 모형(Kulldorff. *et al.*, 2006; Agarwal *et al.*, 2006)이 개발되어 있다. 본 연구에서 사용하는 자료는 연속형 자료이기 때문에 공간스캔통계량 중에서 정규 모형을 사용하여야 한다.

다음은 Agarwal *et al.*(2006)을 참조하여 정규 모형을 통계적 차이 함수(statistical discrepancy function) 형태로 정리하였다. 이모형은 다음의 정규 분포를 가정한다. 하나의 확률변수 y 는 평균(mean) μ 와 분산 $1/\tau^2$ 을 갖는 ($y \sim N(\mu, 1/\tau^2)$)정규 분포를 따른다고 가정한다. 그리고 R지역의 관찰 값을 독립된 $N(\mu_r, 1/\tau_r^2)$, ($i \in R$)라 가정하고, Rc지역의 관찰 값을 독립된 $N(\mu_{rc}, 1/\tau_{rc}^2)$ 로 가정한다. 이때 자료의 분산의 Φ 는, 분산으로 표준화한 대표값을 G 로, 공간클러스터 후보의 기대 값을 b_r 로, 그리고 공간클러스터 후보의 관찰 값을 m_r 이라 할 수 있고, 그림 3.은 각 기호가 의미하는 공간적 위치를 표시하고 있으며, 아래는 각각에 해당하는 식이다.

$$\Phi_{\text{분산}_R} = (\sum_{i \in R} \tau_i^2)^{-1}, \quad \Phi_{\text{분산}_{R^c}} = (\sum_{i \in R^c} \tau_i^2)^{-1} \text{이고,}$$

$$G_{\text{대표값}_R} = \frac{\sum_{i \in R} \tau_i^2 y_i}{\sum_{i \in R} \tau_i^2},$$

$$G_{\text{대표값}_{R^c}} = \frac{\sum_{i \in R+R^c} \tau_i^2 y_i}{\sum_{i \in R+R^c} \tau_i^2} \text{의 경우에}$$

$$b_{\text{기대값}_R} = \frac{1/\Phi}{(1/\Phi_R + 1/\Phi_{R^c})} = \frac{\sum_{i \in R} \tau_i^2}{\sum_{i \in R+R^c} \tau_i^2} \text{이고,}$$

$$m_{\text{기대값}_R} = \frac{\sum_{i \in R} \tau_i^2 y_i}{\sum_{i \in R+R^c} \tau_i^2} \text{이라 할 수 있다.}$$

위와 같이 기대값 b_r 과 관찰값 m_r 이 주어졌을 경우

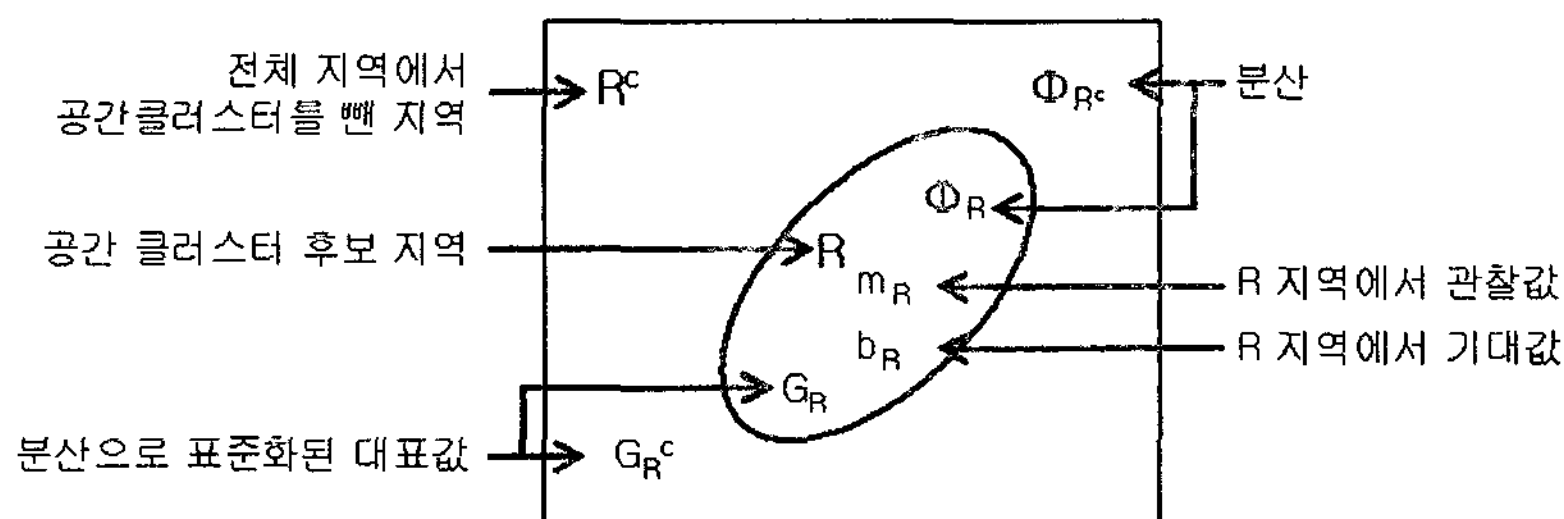


그림 3. 자료가 Normal 모형의 공간스캔통계량에서 사용되는 변수의 의미와 위치

의 통계적 차이 함수 $d_r(b_r, m_r)$ 는 다음과 같다.

$$\begin{aligned} d_r(b_r, m_r) \frac{\Phi}{G} &= m_R G \frac{m_R}{b_R} - \frac{b_R}{G} G \frac{m_R}{b_R} + (1-m_R)G \\ &\quad \frac{1-m_R}{1-b_R} - \frac{1-b_R}{G} G \frac{1-m_R}{1-b_R} \\ &= G \left(\frac{m_R^2}{b_R} + \frac{(1-m_R)^2}{1-b_R} \right) - 1 \\ &= G \frac{(m_R-b_R)^2}{b_R(1-b_R)} \\ \therefore d_r(b_r, m_r) &= c \frac{(m_R-b_R)^2}{b_R(1-b_R)}, \\ c(>0) & \text{ is Fixed constant} \end{aligned}$$

정규모형의 공간스캔통계량은 위의 통계적 차이 함수를 이용하여 모든 후보군 중에서 최대 통계량을 갖는 하나의 공간클러스터를 찾을 수 있다. 이 탐색된 공간클러스터를 대상으로 몬테카를로 테스트를 수행함으로써 이 찾아진 공간클러스터가 우연에 의해서 발생할 수 있는가에 대한 여부를 통계적으로 확인할 수 있다(Kulldorff, 1997).

한편, 대상지역에 공간클러스터가 1개 존재할 경우는 최대 공간스캔통계량을 가진 공간클러스터에 대해서 유의성 검정을 수행함으로써 분석이 종료된다. 하지만 대상지역에 정확하게 몇 개의 이례적 가격변동 집중 지역이 존재하는 지에 대해서 사전에 알 수 없기 때문에 반복적으로 위의 방법을 적용하는 것이 필요하다. Kulldorff(2006)는 반복 스캔(iterative scan)방법을 이용하여 복수의 공간클러스터를 탐색할 수 있다고 주장하였다. 이 방법은 각 분석에서 탐색된 최대 가능 공간클러스터에 해당하는 자료를 분석 자료 집합에서 제거하고, 남은 자료를 대상으로 다시 공간스캔통계량을 계산한다. 이를 정해진 횟수만큼 수행함으로써 복수의 공간클러스터를 탐색한다.

② 다수준에서 후보군 생성

공간스캔통계량은 탐색 윈도우의 최대 크기(max spatial cluster size)를 각기 다르게 제한하면 크기나 모양이 약간 다른 공간클러스터를 탐색할 수 있다. 그림 4는 가상의 가격변동 면(surface)에서 공간스캔통계

량을 사용하여 다양한 가격상승 공간클러스터가 탐색될 수 있음을 보여준다. 가격변동 평균선이 전체지역의 평균이라고 할 때, 평균선 위로 솟은 산(mountain)은 가격상승 공간클러스터의 후보가 될 것이고, 이 가격상승 공간클러스터 후보를 최대 크기의 공간탐색 윈도우가 제한될 때 다양하게 탐색될 수 있는 가격변동 클러스터를 보여주고 있다. (A)는 이상적인 가격변동면이 단봉의 산일 경우를 가정하고 있다. 탐색 윈도우(a)는 작은 크기의 탐색윈도우에 의해서 산의 정상 부분을 포함하는 가격상승 공간클러스터로 탐색하고, 탐색 윈도우(b)는 큰 크기의 탐색윈도우에 의해서 산의 대부분을 포함하는 가격상승 공간클러스터로 탐색한다. (B)에서 이상적인 가격변동면이 쌍봉의 형태일 경우를 가정하고 있다. 탐색 윈도우(c)는 작은 크기의 탐색윈도우에 의해서 산의 코어 부분 포함하는 2개의 가격상승 공간클러스터로 탐색하고, 탐색 윈도우(d)는 큰 크기의 탐색윈도우에 의해서 2개의 산 정상을 대부분을 포함하는 가격상승 공간클러스터로 탐색한다. 공간스캔통계량은 동일한 상황에서 단지 클러스터를 구성하는 개별 부동산의 개수가 증가할 경우 공간클러스터가 될 가능성이 더 높다고 판단하기 때문에, 그림 4에서 다른 조건이 같다면 면적이 더 큰 진한색의 타원이 가격변동 공간클러스터로 탐색되는 경향이 있다.

탐색 윈도우의 최대 크기는 공간클러스터에 대한 사전정보가 없다면 전체 면적의 절반 또는 자료 개수의 절반 이하로 설정하는 것이 일반적이다. 그러나 본 연구와 같이 탐색된 가격변동 공간클러스터를 대상으로 2차적인 분석이 필요한 경우는 2차 분석에서 많은 대안이 검토될 수 있도록 다양한 크기로 탐색윈도우를 설정할 필요가 있다. 예로 그림 4의 (A)에서 (a)와 (b)는 다른 최대 탐색 윈도우 설정에 의해서 공간클러스터로 탐색될 수 있다고 할 때, (b)가 더 높은 공간스캔통계량을 가지고 있기 때문에 최대 가능 공간클러스터로 탐색된다. 하지만 (b)는 공간클러스터이지만 그 내부에 있는 개별 부동산의 가격변동 방향이 그 주변 부동산과 동질적이라 할 수 없다. 따라서 (b)만을 공간클러스터로 탐색하는 경우 (a)와 같이 공간클러스터이면서 그 내부의 개별 부동산과 그 주변 부동산의 가격변동 방향이 동질적인 지역을 찾지 못할 것이다. 이와 같이 공

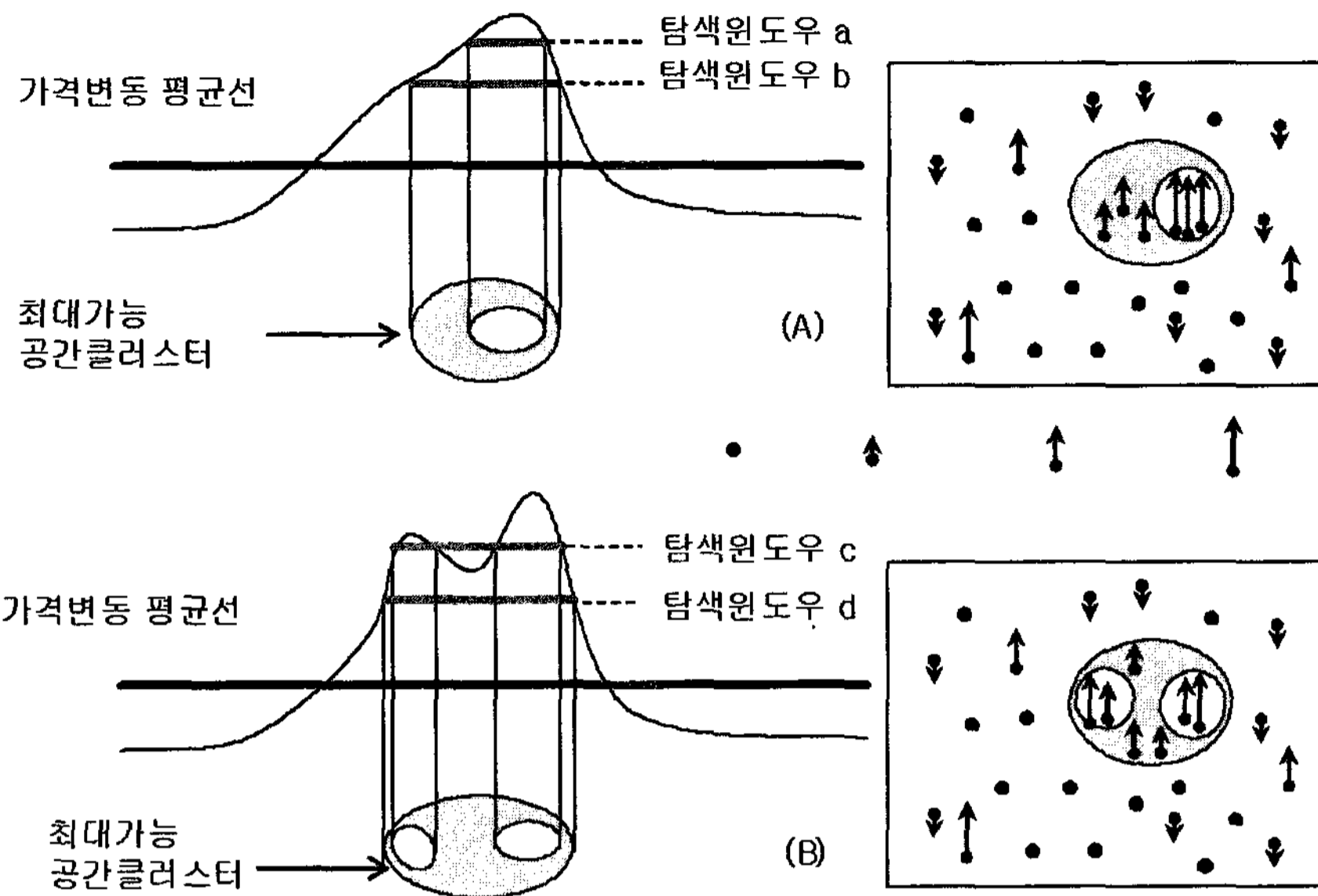


그림 4. 가변 윈도우의 최대 크기를 다르게 제한할 때 탐색되는 다양한 모양과 크기의 공간클러스터

간클러스터이면서 그 내부의 개별 부동산들의 가격변동 방향이 동질적인 지역을 찾기 위해서, 본 연구에서는 다른 크기를 갖는 복수의 최대 탐색윈도우 크기를 사용하여 반복적으로 공간클러스터를 탐색하는 다수준의 공간클러스터 탐색법을 사용한다.

한편, 다수준의 공간클러스터를 탐색하기 위해서는 몇 개의 수준으로 얼마 정도의 최대 크기로 할 것인가를 결정하여야 한다. 더 많은 수준으로 가격변동 공간클러스터를 탐색하면 복잡한 형태의 가격변동 핫스팟을 탐색할 가능성은 높아지지만 같은 자료를 반복적으로 분석함으로써 다중 테스트 문제(multiple test problem)가 발생한다. 따라서 무한히 많은 수준을 사용하는 것은 바람직하지 못하다. 본 연구에서는 개별 수준의 p-value의 누적 총합이 유의수준을 0.05를 넘지 않도록 제한한다. 즉, 본페로니 수정(Bonferroni adjustment)에 의해 p-value 0.01인 가격변동 공간클러스터 탐색을 5회할 수 있다. 그리고 최대 공간클러스터의 반경은 연구 지역에서 가능한 최대 공간클러스터 크기를 기준으로 일정 간격으로 위에서 결정된 다수준의 개수만큼 감소시킨다. 즉, 최대 가능확률이 경우의 최대 공간클러스터의 크기는 전체 면적의 절반이 될 때의 지름이기 때문이 이 길이를 5등분 하고, 각 등분

되는 길이를 각각의 최대 공간클러스터의 크기로 사용할 수 있다. 또한 가격상승 공간클러스터에 대한 사전 지식이 있는 경우는 연구자가 그 크기를 결정할 수 있다.

(3) 주변 부동산과의 공간연관성 평가

앞 절의 가격상승이 큰 부동산의 공간집중성 평가를 통해서 일정 지역이 가격상승 공간클러스터로 평가되었다 할지라도 이 지역의 부동산들이 주변 부동산과 가격변동 방향이 동질적이라고 할 수 없다. 이 탐색된 가격상승 공간클러스터 내부의 부동산들이 주변 부동산과 가격변동 방향이 일정한가를 판단하기 위해서 이들 부동산들 간의 공간연관관계를 평가하여야 한다.

이 공간연관관계는 두 가지 측면에서 검토할 수 있다. 하나는 자신이 전체평균대비 가격상승 값일 때 주변 부동산도 같이 가격상승 값을 갖는가를 평가하는 핫스팟(hot spot) 테스트이다. 그림 1에서 (a), (b), (c)의 타원과 같이 공간 클러스터 내의 대부분의 부동산들이 주변 부동산과의 가격변동 방향이 일치 할 때는 높은 공간연관성을 갖는다고 할 수 있다. 이들 중에서 (a) 타원 내의 개별 부동산들의 가격상승 폭이 가장 크기 때문에 핫스팟이 될 가능성이 가장 높다고 하겠다.

표 2. 국지적 공간자기상관지수의 특성(자료: Lee, 2001)

	Local Moran's I(i)	Local Geary's C(i)	Getis-Ord's Statistics G(i)	Local Lee's S(i)
Spatial Clusters	Effective	Not Effective	Effective	Effective
Hot spots vs. Cold spots	Neutra	lNot Effective	Effective	Neutral
Local homogeneity	Not Effective	Effective	Not Effective	Not Effective
Spatial outliers	Effective	Not Effective	Not Effective	Not Effective
Dependency on Reference Data	Strong	Strong	Not Strong	Not Strong

다른 하나는 자신이 높은 가격상승 값일 때 반대로 주위의 부동산 중에서 가격하락 값을 가지고 있는가를 평가하는 공간 이례관계(spatial outlier) 테스트이다. 그림 1 (a)의 타원은 공간클러스터 내부의 대부분의 부동산에서 자기가 높은 가격상승일 때 주변 부동산도 높은 가격상승인 공간연관관계(High-High)이고, (d)의 타원은 자기가 높은 가격상승 값을 가질 때 주변 부동산이 최소한 전체 평균 가격상승 값을 가지고 있거나 자기가 전체 가격변동의 평균값일 때 주변 부동산은 평균 이상이거나 자신과 같은 평균값 정도의 값을 가지는 관계이다. (g)의 타원은 자기가 높은 가격상승 값을 가질 때 주변 부동산이 높은 가격하락 값을 가지는 공간 이례관계(High-Low 또는 Low-High)관계이다.

한편, 공간연관관계를 측정하기 위해서는 먼저 개별 부동산 사이의 공간적 근린관계를 정의하고 이를 정량화하는 작업이 필요하다. 공간근린관계는 근린 가중치 행렬 (spatial proximity weight matrix) W 로 표현되고 W_{ij} 는 부동산 i 와 j 간의 근린 정도를 나타낸다. W 를 구성하는 가장 기본적인 형태는 이진 연결성 (binary connectivity)으로서, 두개의 폴리곤이 인접하면 1, 인접하지 않으면 0의 가중치를 부여하는 것이다(Wong and Lee, 2005). 하지만 본 연구의 자료는 점형자료이기 때문에 연결성을 바로 사용할 수 없다. 본 연구에서는 거리가 멀어질수록 부동산 소유자간의 전략적 제휴 강도가 약해진다고 가정함으로써 거리조락함수에 의한 영향력을 각 관찰 부동산 간의 W_{ij} 로 사용한 근린가중행렬을 사용한다. 거리조락함수는 역거리제곱 (inverse distance square)함수를 사용한다.

공간연관관계 평가는 대상지역 전체에 공간연관관

계가 존재하는지 여부를 평가하는 전역적 방법과 공간연관관계가 존재하는 일정 지역을 탐색해 주는 국지적 방법이 존재한다. 본 연구는 1차적으로 탐색된 가격상승 공간클러스터 내의 개별 부동산의 가격변동이 일정 방향을 유지하고 있는가를 평가하기 위해서 국지적 공간자기연관지수를 사용한다. 국지적 공간연관지수는 관찰지점과 그 인접한 이웃사이에 높은 유사성이 존재하는 경우 공간적인 연관관계가 높다고 판단한다. 대표적인 지수로는 Anselin(1995)가 제안한 국지 모란 $I_{(i)}$, Geary(1952)와 Boots(2002)가 제안한 Geary's $C_{(i)}$, Ord and Getis(1995)가 제안한 $G_{(i)}$, 그리고 Lee(2001)이 제안한 $S_{(i)}$ 가 있다. 이 지수들은 표 2와 같이 주변 부동산과의 공간연관관계를 다른 측면에서 측정하고 있다. 이들 중에서 국지 모란 I 가 공간 이례관계를 탐색하는데 더 많이 이용되고 있다. 즉, 국지 모란 I 를 기초로 만들어진 Moran Scatterplot을 이용할 경우에 우리는 공간 이례관계를 쉽게 구분할 수 있다(Anselin, 1995).

가격변동 핫스팟은 그림 1의 (a), (b), 그리고 (d)의 타원인 경우로 알려져 있다(손학기, 2008). 이것을 공간연관성 관점에서 해석해 보면, 가격변동 핫스팟 내부의 부동산의 가격변동은 공간 이례관계(High-Low)를 포함하지 않으면서 일정 수준 이상의 공간연관관계를 형성하여야 한다. 이것은 부동산의 개별성과 위치 고정성에 의해서 일정지역 내의 모든 부동산에서 실제 투기과열이 발생한 경우라도 타 지역보다 높은 가격상승이 이루어지지 않는 경우가 있기 때문이다. 또한 부동산 과열지역이라면 이 지역의 부동산 가격변동은 최소한 타 지역에 비해서 하락하지 않은 점도 고려되었

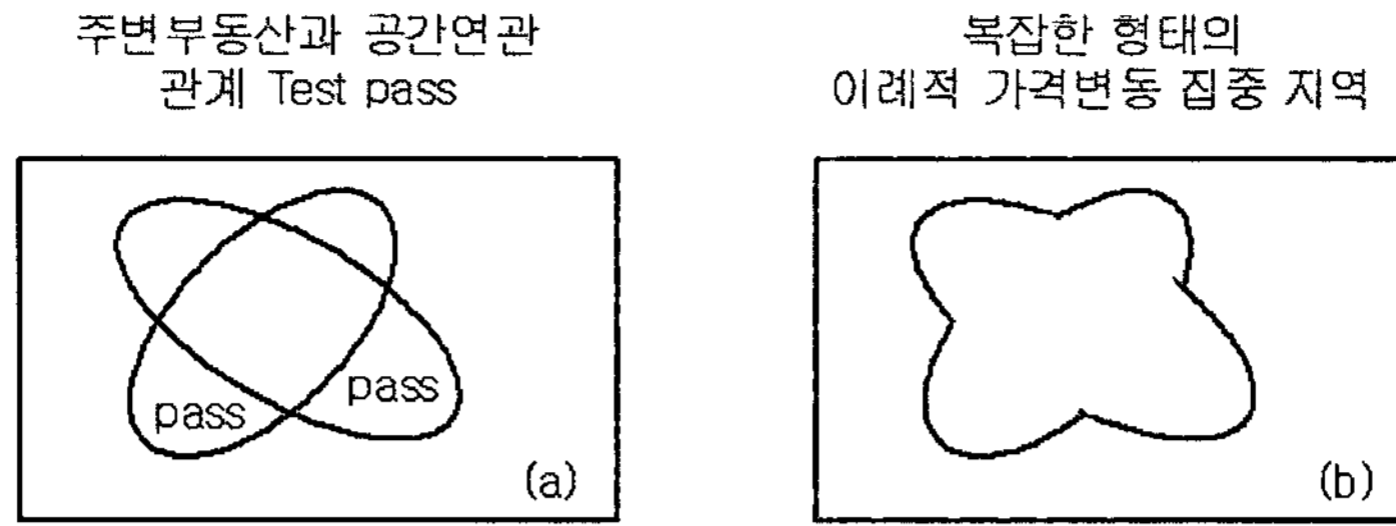


그림 5. 복잡한 형태의 이례적 가격변동 집중 지역 형성 과정

기 때문이다. 따라서 본 연구에서 가격변동 핫스팟의 공간연관관계는 가격상승 공간클러스터에 포함되는 모든 개별 부동산이 주변 부동산과 High-Low 또는 Low-High의 공간 이례관계를 가지지 않으면서, 부동산의 개별성과 위치고정성을 고려하여 1표준편차 이상의 공간연관관계를 가진 부동산이 가격변동 핫스팟에 70%이상 존재하는가를 평가한다. 즉, 30% 정도의 부동산은 각 부동산의 개별성에 의해서 주변 부동산과 특별한 공간연관관계를 보이지 않는다고 판단한다.

(4) 복잡한 형태의 가격변동 핫스팟 생성

각각 수준별로 가격상승 공간클러스터 평가와 주변 부동산과의 공간연관관계 테스트를 통과한 가격변동 핫스팟 후보는 동일한 가격변동 핫스팟이 탐색될 경우도 있고 공간적으로 최소한 일부분이 중첩되는 다른 크기와 모양의 가격변동 핫스팟들이 탐색될 수도 있다. 다수준(multi-level)으로 탐색된 가격변동 핫스팟을 하나로 중첩함으로써 단일 수준으로 탐색할 때 잃을 수 있는 가격변동 핫스팟의 크기 및 모양에 대한 상세함을 유지할 수 있다. 본 연구에서는 이것이 주로 탐색 윈도우의 단순화(원이나 타원)에 의해서 실제 가격변동 핫스팟이 불규칙한 모양임에도 불구하고 단순한 형태의 탐색윈도우로 탐색된다고 가정한다. 따라서 [그림 5]와 같이 공간적으로 중첩되는 가격변동 핫스팟을 합집합으로 결합하여 최종의 가격변동 핫스팟을 도출한다. 그럼에도 불구하고 이 경계는 실제 현상, 즉 투기과열지역의 경계를 정확하게 표현하고 있다고 볼 수 없다. 단지 단순한 도형을 이용하여 그 가능성이 가장 큰 지역을 나타낸 것이다.

4. 가격변동 핫스팟 탐색법 적용

1) 사례연구 지역 및 자료처리

연구 대상지역은 1997년 외환위기 이후 전국의 타 지역에 비해서 높은 가격상승률을 기록하였고, 소위 버블세븐지역에 해당하는 서초·강남·송파 3개구를 대상으로 하였다. 이 지역은 전국의 부동산가격상승률 및 물가상승률에 비해 높은 가격상승이 이루어져 2003년 초 주택투기지역(소득세법 제96조)지역으로 지정되었다. 특히 참여정부의 강력한 부동산 대책 중 하나인 10.29 부동산 대책은 이 지역의 집값상승을 막기 위한 대책이라고 할 수 있다. 따라서 10.29 부동산 대책을 전후한 2003년 8월부터 2003년 11월 사이 이 지역의 일부는 투기과열의 가능성이 높은 지역이라 할 수 있다. 우리는 이 시기 이 지역의 아파트를 대상으로 높은 미래 기대이익을 바라는 적응적 경제주체가 경제행위를 하였다고 쉽게 알 수 있다.

연구의 자료는 2003년 8월부터 2003년 11월까지의 국민은행에서 제공하는 개별 부동산 단지의 월별 아파트 시세자료를 사용하였다. 이 자료는 2007년 2월 기준 562개 아파트 단지에 1997개 주택형(평형별) 시세자료로 구성된다. 한 단지 내에 여러 종류의 주택형이 존재하기 때문에 각각을 주택형 당 시세로 변환하고 이를 평균하여 한 단지를 대표하는 시세로 사용하였다. 그림 6은 분석에서 사용된 서초·강남·송파구의 562개 아파트 단지의 위치를 점으로 표현하고 있다. 가격변동률 지수는 기준시점(t)의 한 아파트단지의 평

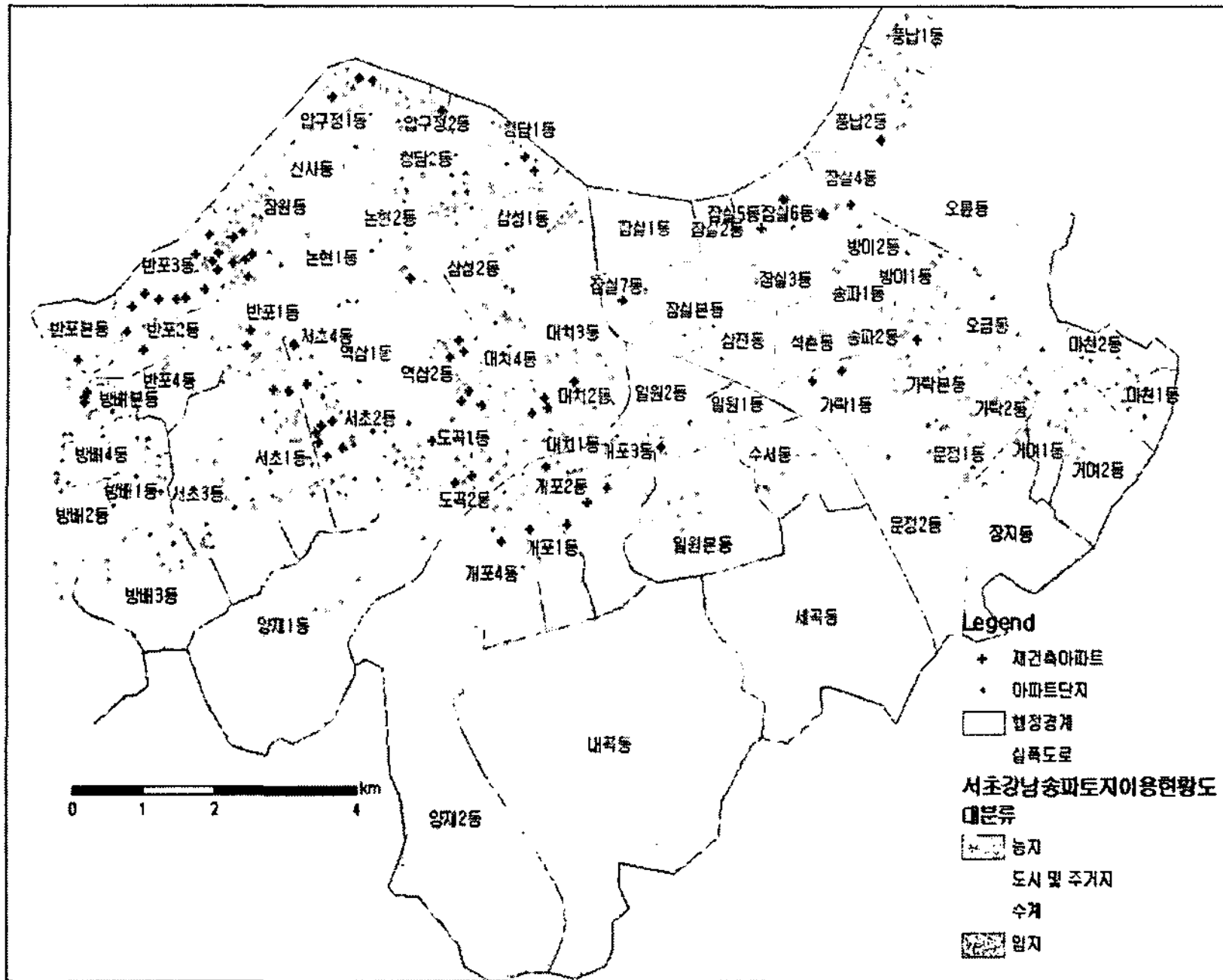


그림 6. 시세가 자료가 있는 서초·강남·송파구 아파트단지의 위치

균 시세에서 기준 직전 월(t-1)의 평균 시세를 뺀 값에 직전 월(t-1)의 평균 평당 시세로 나누고 100을 곱함으로써 산출하였다. 시세가 직전월 대비 5% 증가한 경우에 변동률지수는 5가 된다. 전분기 대비나 전년 대비 가격변동률 지수를 산정할 때는 대비되는 시점만 바뀌

고 다른 부분은 동일하다.

$$\begin{aligned} & \text{전월대비 가격변동률 지수} \\ & = \frac{(\text{기준 시점 가격}(t) - \text{직전 월 가격}(t-1))}{\text{직전 월 가격}(t-1)} \times 100 \end{aligned}$$

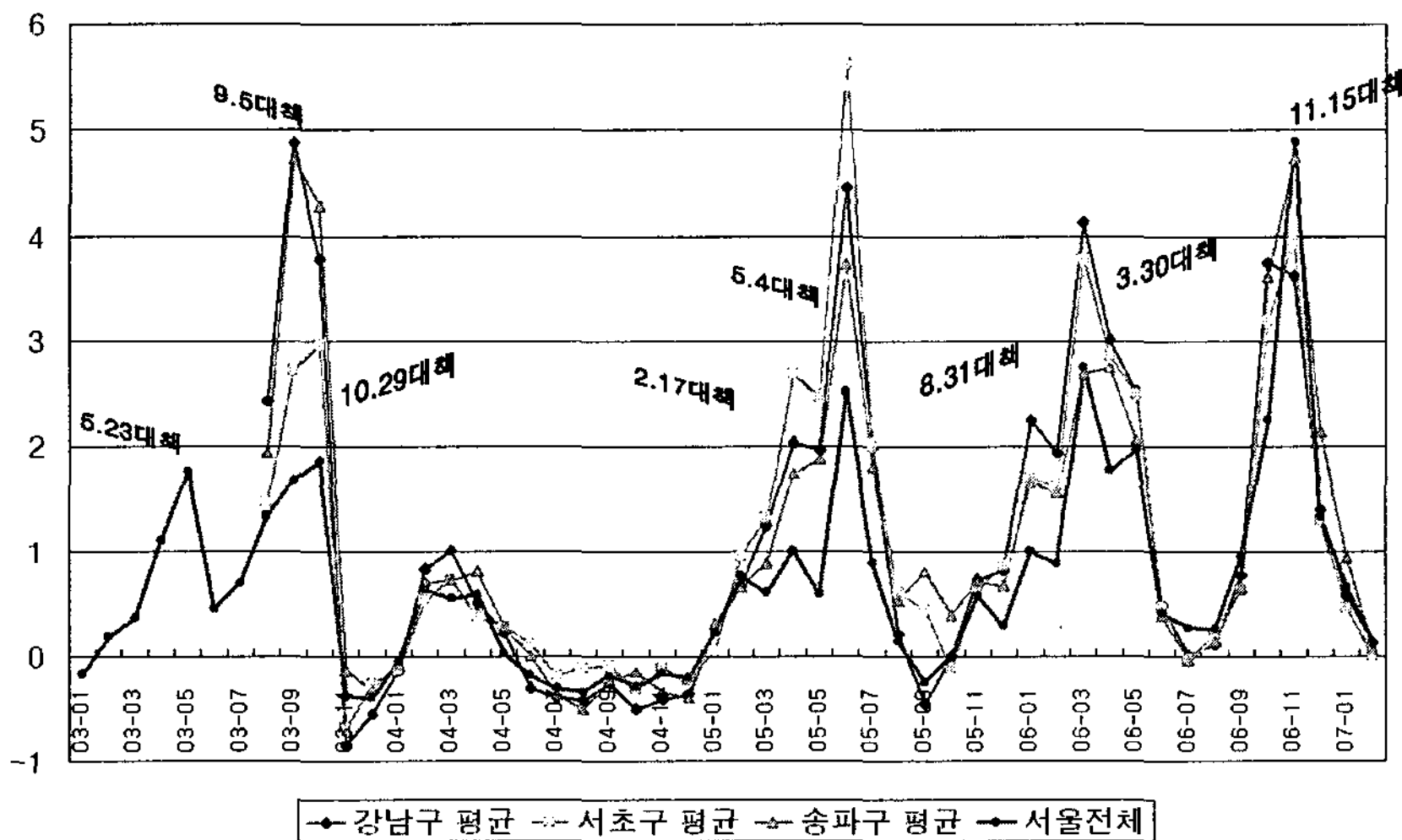


그림 7. 부동산 대책과 서초·강남·송파구의 전월대비 시세변동과의 관계

$$= \{ \text{기준 시점 가격}(t) / \text{직전 월 가격}(t-1) - 1 \} \times 100$$

이 지수를 이용하여 연구대상지역의 2003년 8월부터 2003년 12월까지 가격변동률을 그림 7과 같이 그래프를 작성하였다.

이 그래프에서 전월대비 가격변동률은 각 부동산 대책 발표 전까지는 상승하고, 발표 후에는 급락하는 것을 확인할 수 있었고, 재건축아파트가 일반아파트보다 가격 급등과 급락의 폭이 크다는 것을 알 수 있다. 이처럼 전월대비 가격변동률은 부동산 대책에 민감하게 반응하였다. 그러나 전년 대비나 전분기 대비 가격변동지수는 전월대비 가격변동률보다 부동산대책에 민감하게 반응하지 못하였다. 따라서 본 연구에서는 전월대비 가격변동률이 시시각각으로 변하는 가격변동 상황을 잘 표현할 수 있다고 판단하여 가격변동 자료로 사용하였다.

2) 부동산 가격변동 핫스팟 탐색

개발된 가격변동 핫스팟 탐색 모형을 서초·강남·송파 3개구 지역에 적용한다. 모형의 1단계인 가격변동 핫스팟 후보군 생성과 2단계인 가격상승의 공간클러스터 탐색은 SatScan을 이용한다. SatScan에서 탐색 윈도우로 이심률이 1, 1.2, 1.4, 1.6, 1.8이고, 각도가 0, 15, 30, 45, 60, 75, 90, -15, -30, -45, -60, -75인 타원을 사용하고, 타원의 중심점은 부동산의 위치로 사용한다. 복잡한 형태의 이례적 가격변동 핫스팟 지역을 탐색하기 위해서 장축의 길이가 각각 500, 1000, 1500, 2000인 4단계의 최대 공간탐색범위를 이용한다. 이 파라미터는 가격변동 핫스팟이 연구에 사용된 구경계를 벗어나지 않는다는 가정하에서 사용한다. 4단계의 가격상승 공간클러스터 탐색이 수행되기 때문에 각각의 공간클러스터 탐색의 유의수준을 p-value가 0.01 이하로 하여 총 p-value가 0.05 이하가 되도록 하였다.

3단계인 주변부동산과 공간연관성 평가와 4단계인 복잡한 형태의 가격변동 핫스팟 생성은 ArcGIS를 이용한다. 2단계에서 탐색된 가격상승 공간클러스터를 대상으로 주변 부동산과의 공간 연관관계 평가는 ArcGIS의 국지 Moran I를 기초로 작성된 Moran Scatterplot을 이용한다. 탐색된 가격상승 공간클러스터 내 개별 부

동산 중에서 공간 이례관계가 존재하지 않고, z-score에 의한 신뢰수준이 1 표준편차 이상의 공간연관관계를 가진 부동산이 이 지역 내에 70%이상 존재하면 유의한 공간 연관관계를 가지고 있다고 판단한다. 다수준으로 탐색된 가격변동 핫스팟이 공간상에 일부분이라도 중첩되는 경우 ArcGIS에서 Dissolve를 수행하여 탐색된 집합의 합집합을 구한다.

탐색된 가격변동 핫스팟은 분석시 사용된 자료를 지도화 함으로써 시각적으로 쉽게 평가할 수 있도록 하였다. 각 분석 시점에 대해서 [부록]과 같이 2가지 종류의 지도를 작성한다. [부록]의 좌측에는 탐색된 가격변동 핫스팟과 함께 각 부동산의 주변 부동산과의 공간 연관관계 정도를 국지 Moran I의 통계적 유의성을 표현하였다. 각 부동산의 공간유의성을 표준편차로 구분된 점형 심볼로 표현하였다. 그리고 바탕에는 IDW(Inverse Distance Weight)를 이용하여 분석시 사용된 자료를 보간하여 지도화 한다. 평균에 해당하는 값은 노란색 계열을 사용하고 평균이상은 붉은색, 평균이하는 푸른색 계열을 사용한다. 이때 토지이용이 도시나 주거지가 아닌 수계, 농지, 그리고 임지는 마스크(masking)함으로써 시각적인 왜곡을 줄인다. [부록]의 우측에는 각 부동산의 가격변동률을 표준편차로 단계구분하고 점형 심볼로 표현하였다. 붉은 색은 1표준편차 이상, 황색계열은 1표준편차이하이고 -1표준편차 이상, 그리고 청색계열은 -1 표준편차 이하인 지역을 표현하였다. 이때 지도의 바탕은 첫 번째 지도와 같은 방법으로 작성하였다.

3) 탐색 결과

본 연구에서는 가격변동 핫스팟 탐색법을 이용하여 부동산 투기과열의 가능성이 높은 시기와 지역을 대상으로 실증적으로 가격변동 핫스팟을 탐색하였다. 10.29 부동산 대책 2개월 전인 2003년 8월의 개별 부동산의 전월대비 가격변동의 전체 연구지역의 평균은 1.98이고 표준편차는 2.78이었다. [부록]의 2003년 8월의 우측 도면은 개별 부동산의 가격변동을 단계구분도이다. 우리는 이 지도를 통해서 개포동과 대치동 주변에 2 표준편차 이상을 가진 부동산이 공간적으로 집

중되어 있는 것을 시각적 확인할 수 있다. 하지만 이 시각적 탐색법을 통해서 투기과열의 가능성이 높은 지역의 경계를 확정하기 어렵고, 그 지역이 우연에 의해서 형성된 지역인지 아니면 실제로 투기과열의 가능성이 높은가에 대한 검증할 수 없다(Mac Eachren and Kraak 1997).

그러나 [부록]의 2003년 8월의 좌측 지도 또는 이 지도의 핫스팟을 확대한 그림 8에서는 이 시기 부동산 문제의 시작으로 알려진 은마아파트 등의 다수의 재건축 아파트를 포함하는 지역, 즉 가격변동 핫스팟을 확인할 수 있었다. 이 탐색된 가격변동 핫스팟의 통계적 특성은 전체 평균변동 1.98보다 3배 이상의 가격변동이 이루어진 지역이고, 공간이레관계가 없는 상태에서 70%이상의 부동산이 자신이 가격이 상승할 때 주변 부동산의 가격이 상승하는 공간연관관계(high-high)를 가지고 있었다. 구체적으로 이 지역에는 5개의 주상복합아파트와 14개의 재건축 아파트를 포함한 총 53개의 아파트 단지로 구성되었다. 이 지역의 대표적인 재건축아파트 단지는 개포주공1·2·3단지, 현대3차, 개포시영, 은마, 진달래1·2·3차가 있었다. 이 부동산들은 이 시기 부동산 투기의 대상이 된 것으로 널리 알려져 있다.

9.5 대책이 발표되고 10.29 대책이 발표되기 1개월 전인 9월은 전체 평균과 편차가 4.1이었다. 이 시기의 가격변동 핫스팟은 [부록]의 2003년 9월 좌측 도면과 이 도면을 확대한 그림 9에서 개포동, 대치동, 수서동 일대에서 3개를 확인하였다. 이 지역들은 평균보다 2~3배 이상이었으며, 공간이레지점이 없는 상태에서 주변 부동산과의 공간연관을 가진 부동산이 70%이상이었다. 이 시기에 가격변동 핫스팟에 포함된 아파트는 3

개의 재건축 아파트를 포함한 24개의 아파트였다. 3개의 재건축 아파트는 개포리키, 개포한신, 그리고 대우 아파트이었다.

10.29 대책이 발표된 2003년 10월의 전월대비 가격변동 평균은 3.6이고 표준편차는 4.2이었다. [부록]의 2003년 10월 우측 지도에서 비록 2표준편차 이상의 가격변동을 보인 부동산이 존재하였지만, 좌측 지도에서는 어떠한 가격변동 핫스팟이 탐색되지 않았다. 이것은 10.29 대책으로 미래 기대이익에 대한 기대가 줄어들게 됨으로써 수요자의 (가)수요가 전월에 비해서 줄어들거나 분산된 것으로 생각할 수 있다.

10.29 대책이 발표 1개월 후인 2003년 11월의 전월대비 가격변동 평균은 -0.62이고 표준편차는 3.09이었다. [부록]의 2003년 11월 우측 지도에서 다수의 -2표준편차 이하의 가격변동 부동산을 확인할 수 있고, 우측 도면에서 1개의 가격변동 콜드스팟을 확인할 수 있었다. 이 지역은 전체지역 평균적인 가격변동 -0.62보다 18배 이상의 가격하락이 발생하였다. 이 지역에 포함된 부동산은 총 7개로 개포주공1·2·3단지 등의 재건축 아파트를 포함하고 있었다. 이 시기 가격하락을 주도한 부동산은 재건축 아파트였다.

이상에서 10.29 대책을 전후하여 형성된 가격변동 핫스팟과 콜드스팟을 탐색하였다. 대책 전에는 가격변동 핫스팟이 탐색되었고, 대책 시점에는 소강상태를 보이다가 대책 1개월 후에는 가격변동 콜드스팟이 탐색되었다. 그리고 가격변동 핫스팟 탐색모형을 사용하여 참여정부의 주요 부동산 대책(10.29) 대책전후로 형성되었을 것이라 잘 알려진 투기과열의 가능성이 높은 지역과 부동산 침체 가능성이 높은 지역을 미시적이고 적시에 탐색할 수 있음을 확인하였다.

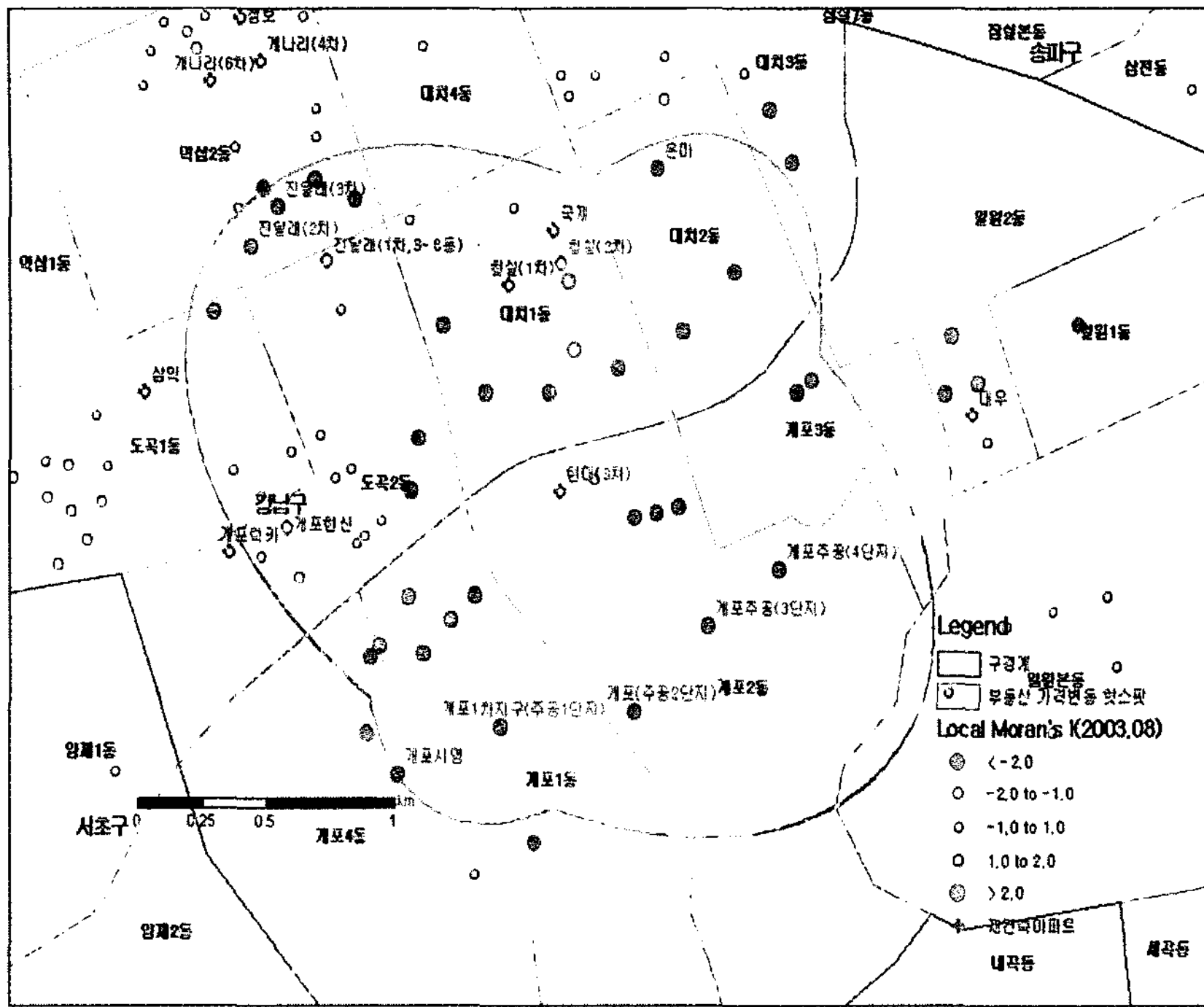


그림 8. 복잡한 형태의 부동산 가격변동 핫스팟 탐색 결과 (2003. 8)

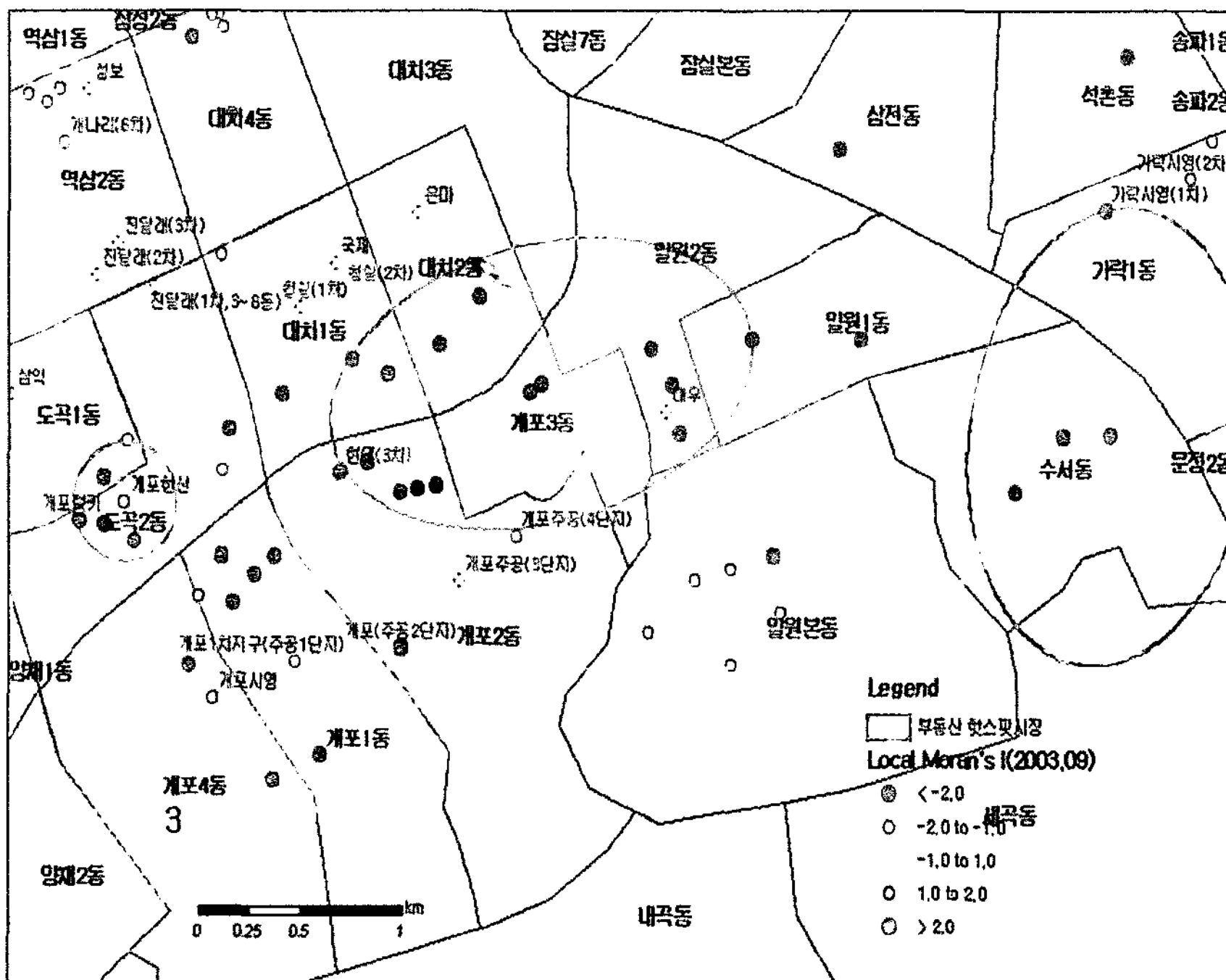


그림 9. 부동산 가격변동 핫스팟 탐색 결과 (2003. 9)

5. 결론

투기과열의 가능성이 높은 지역을 가격변동 핫스팟 패턴을 통해서 적시에 높은 공간해상도(개별 부동산 단위)로 탐색하는 것은 부동산 정책의 효율성에 직결된다. 하지만 가격변동 핫스팟은 행정경계와 같이 고정된 공간범위를 가지는 것이 아니라 그 위치, 크기, 그리고 모양을 사전에 알 수 없기 때문에 이 특성을 고려할 적절한 방법론의 부재로 그동안 체계적인 연구가 수행되지 못하였다. 최근 공간통계학 분야에서 공간클러스터와 공간연관관계를 탐색할 수 있는 방법이 제안됨에 따라 다양한 분야에서 특정 현상을 설명하기 위한 공간패턴 탐색 연구가 활발히 진행되고 있다.

이러한 배경에서 본 연구는 가격변동 핫스팟의 후보군을 대상으로 이 지역의 공간집중성은 공간스캔통계량의 정규모형으로 평가하고, 주변 부동산과의 공간연관관계는 국지 Moran I를 통해서 평가하는 모형을 개발하였다. 그리고 이 모형을 사례연구에 적용하여, 서울 강남·서초·송파 3개구 지역에서 2003년 8월부터 11월 사이에 투기과열지역 및 부동산 침체 가능성이 높은 지역을 탐색하였다.

연구지역은 2003년 초에 주택투기지역으로 지정되었고 참여정부의 부동산대책의 주요 대상지역으로 투기과열지역이라 할 수 있다. 특히 2003년 10.29 대책 전에는 지속적으로 가격이 상승할 것이라는 기대심리에 의해서 투기과열이 발생하고, 대책 이후는 이 기대가 무너짐으로써 부동산침체가 일어났으리라 쉽게 예상할 수 있다. 탐색결과 대책 발표 전인 8월과 9월은 대치동과 개포동 일대에 가격변동 핫스팟을 발견하였다. 대책 시점인 10월은 소강상태로 가격변동 핫스팟이나 콜드스팟을 발견할 수 없었다. 그러나 대책 시점 1개월 후에는 대치동 일대에서 가격변동 콜드스팟을 발견하였다. 이상의 결과는 제한된 방법이 가격변동 핫스팟의 경계를 근사적으로 획정할 수 있다는 것 외에도 연속된 시점간의 비교를 통해서 현재의 부동산시장의 상황을 모니터링 할 수 있는 방법임을 보여준다.

제안된 탐색법을 비유적으로 말하자면 인간의 뼈에 대한 이상을 검사하기 위해서 촬영하는 x-ray와 같은

방법이다. 이 방법을 통해서 시시각각으로 변화하는 부동산시장의 상황을 단층 촬영하듯 관찰할 수 있다. 하지만 이 방법은 현 부동산시장의 상황을 만든 직접적인 요인(예, 재건축 아파트의 가격상승과 주변지역으로의 가격상승 전파)을 설명하지 못하는 한계가 있다. 이것은 방법론의 한계라기보다 제안된 방법을 사용하여 문제가 될 가능성이 높은 시점과 지역으로 관찰범위를 좁힘으로써 위와 같은 가격변동의 직접적인 요인을 탐색할 수 있는 도구의 성격을 가지고 있기 때문이다. 이 탐색법의 향후 발전 방향은 공간클러스터링 평가와 공간연관관계 평가가 이원화됨으로써 탐색된 가격변동 핫스팟을 통계적으로 엄밀하지 못한 부분을 극복해야 한다. 즉, 공간연관관계를 고려할 수 있는 공간클러스터 탐색법이 개발되어야 한다. 응용측면의 발전 방향은 이 탐색법을 대상지역의 확대와 다양한 시점에 적용하여 방법론의 안정성을 검토할 필요가 있다.

文獻

- 강원철 · 김용창 · 서순탁 · 서후석 · 임재만, 2004, 부동산학 개론, 부동산114 주식회사, 서울.
- 건교부, 2005, 토지를 활용한 부동산시장 상시모니터링체계 구축방안 연구, 국토연구원.
- 손학기, 2008, “투기과열지역의 공간패턴 모형화,” 대한지리학회지, 43(1), 104-116.
- 이명희(역), 2007, 행동경제학, 지형, 서울(友野 典男, 2006, 行動經濟學 經濟は「感情」で動いてい, 光文社, japan).
- Agarwal, J.D., Phillips, M., and Venkatasubramanian, S., 2006, The hunting of the bump: on maximizing statistical discrepancy. *Proc. 17th Ann. ACM-SIAM Symp. on Disc. Alg.*, 1137-1146.
- Anselin, L., 1995, Local indicators of spatial association-LISA, *Geographical Analysis*, 27, 93-115.
- Anselin, L. and Bao, S., 1997, Exploratory Spatial Data Analysis Linking SpaceStat and ArcView, in Fischer, M. and Getis, A.(eds.), *Recent*

- Development in Spatial Analysis*, 35-59, Springer-Verlag, Berlin.
- Anselin, L., Cohen, J., Cook, D., Gorr, W., and Tita, G., 2000, *Spatial analyses of crime*, Criminal Justice 2000, National Institute of Justice, Washington, D.C.
- Besag, J. and Newell, J., 1991, The detection of clusters in rare diseases, *Journal of the Royal Statistical Society A*, 154, 143-155.
- Block, 1994, STAC hot spot areas: A statistical tool for law enforcement decisions, Chicago: Illinois Criminal Justice and Information Authority, Retrieved 28 April 2008 from the World Wide Web: <http://www.icjia.state.il.us/public/pdf/stac/Hotspot.pdf>.
- Brantingham, P.L. and Brantingham, P.J., 1982, Mobility, notoriety and crime: A study of crime patterns in urban nodal points, *Journal of Environmental Systems*, 11, 89-99.
- Boots, B.N., 2002, Local measures of spatial association, *Ecoscience*, 9, 168-176.
- Cleveland, W.S., 1993, *Visualizing data*, Summit, Hobart Press, New Jersey.
- Cliff, A.D. and Haggett, P., 1988, *Atlas of disease distributions: analytic approaches to epidemiological data*, Blackwell Publishers, Oxford.
- Elliott, P.J., Wakefield, C., Best, N.G., and Briggs, D.J., 2000, *Spatial Epidemiology: Methods and Applications*, Oxford University Press, Oxford.
- Fotheringham, A.S. and Zhan, F.B., 1996, A Comparison of Three Exploratory Methods for Cluster Detection in Spatial Point patterns, *Geographical Analysis*, 28, 200-218.
- Gahegan, M. and O'Brien, D., 1997, A strategy and architecture for the visualization of complex geographical datasets, *International Journal of Pattern Recognition and Artificial Intelligence*, 11, 239-261.
- Getis, A. and Ord, T.K., 1995, The analysis of spatial association by use of distance statistics, *Geographic Analysis*, 24, 188-206.
- Hirschfield, A., Brown, P., and Todd, P., 1995, GIS and the analysis of spatially referenced crime data: experiences in Merseyside, UK, *International Journal of Geographical Information Science*, 9(2), 191-220.
- Huang, L., Kulldorff, M., and Gregorio, D., 2007, A spatial scan statistic for survival data, *Biometrics*, 63(1), 109-118.
- Jung, I., Kulldorff, M., and Klassen, A., 2006, A spatial scan statistic for ordinal data, *Statistics in Medicine*, 26(7), 1594-1607.
- Knox, G., 1989, Detection of clusters, In Methodology of Enquiries into Disease Clustering, Elliott, P.(ed.), 17-22, *Small Area Health Statistics Unit*, London.
- Kulldorff, M., 1997, A spatial scan statistic, *Communications in Statistics-Theory and methods*, 26, 1481-1496.
- Kulldorff, M., 2006, SaTScan v7.0: SatScan user guide, Retrieved 22 January 2008 from the World Wide Web: <http://www.satscan.org>.
- Kulldorff, M. and Nagarwalla, N., 1995, Spatial disease clusters: detection and inference, *Statistics in Medicine*, 14, 799-810.
- Kulldorff, M., Huang, L., Pickle, L., and Duczmal, L., 2006, An elliptic spatial scan statistic, *Statistics in Medicine*, 25(22), 3929-3943.
- Lawson, A.B. and Kleinman, K., 2005, *Spatial and Syndromic Surveillance for Public Health*, John Wiley & Sons, Inc.
- Lee, S.L., 2001, *Spatial Association Measures for An ESDA-GIS Framework: Developments, Significance Tests, And Applications to Spatio-Temporal Income Dynamics of U.S. Labor Market Areas, 1969-1999*, Ph.D. Dissertation, Department of Geography, The Ohio State University.
- Levine, N., 2006, Crime Mapping and the Crimestat Program, *Geographical Analysis*, 38, 41-56.
- MacEachren, A.M. and Kraak, M.J., 1997, Exploratory cartographic visualization: Advancing the agenda, *Computers and Geosciences*, 23, 335-343.

- Murray, A.T., McGuffog, I., Western, J.S., and Mullins, P., 2001, Exploratory spatial data analysis techniques for examining urban crime, *British Journal of Criminology*, 41, 309-329.
- Neill, D.D. and Moore, A.W., 2005, Anomalous spatial cluster detection, *Proceedings of the KDD 2005 Workshop on Data Mining Methods for Anomaly Detection*, Retrieved 28 April 2008 from the World Wide Web: <http://www.cs.cmu.edu/~neill/papers/ADKDD-Neill.pdf>.
- Openshaw, S. and Albanides, S., 1999, Applying geocomputation to the analysis of spatial distributions, In Longley, P., Goodchild, M., Maguire, D., and Rhind, D.(2d ed.), *Geographical Information Systems*, John Wiley & Sons, Inc.
- Openshaw, S., Charlton, M., Wymer, C., and Craft, A., 1987, A Mark 1 geographical analysis machine for the automated analysis of point datasets, *International Journal of Geographical Information Systems*, 1, 335-358.
- Ord, J.K. and Getis, A., 1995, Local spatial autocorrelation statistics: Distributional issues and an application, *Geographical Analysis*, 27, 286-306.
- Rheingans, P. and Landreth, C., 1995, Perceptual principles for effective visualizations In Grinstein, G. and Levkowitz, H.(ed), *Perceptual issues in visualization*, Berlin: Springer-Verlag.
- Sankoh, O.A. and Becher, H., 2002, Disease cluster detection methods in epidemiology with application to data on childhood mortality in rural Burkina Faso, West Africa, *Epidemiology and Biostatistics in Medicine and Biology*, 33, 460-472.
- Sherman, L.W., 1995. Hot spots of crime and criminal careers of places, In Eck, J.E. and Weisburd, D.(ed), *Crime and place*, Criminal Justice Press, Monsey, NY, 35-52.
- Sherman, L.W. and Weisburd, D.A., 1995, General deterrent effects of police patrol in crime hot spots: A randomized, controlled trial, *Justice Quarterly*, 12, 625-648.
- Spelman, L.W., 1995, Criminal careers of public places, In Eck, J.E. and Weisburd, D.(eds.), *Crime and place*, Crime Prevention Studies, 4, Criminal Justice Press.
- Turnbull, B.W., Iwano, E.J., Burnett, W.S., Howe, H.L., and Clarke, L.C., 1990, Monitoring for clusters of disease: Application to leukemia incidence in upstate New York, *American Journal of Epidemiology*, 32, S136-S143.
- Wong, D.W.S. and Lee, J., 2005, *Statistical analysis of geographic information with ArcView GIS and ArcGIS*, John Wiley & Sons, Inc.
- 교신: 손학기, 431-712, 경기도 안양시 동안구 관양동 1591-6, 국토연구원 (이메일: hgsohn@krihs.re.kr, 전화, 031-380-0225)
- Correspondence: Hakgi Sohn, Korea Research Institute for Human Settlements, 1591-6 Gangyang-dong, Dongan-gu, Anyang-si, Gyeonggi-do, 431-712, Korea (e-mail: hgsohn@krihs.re.kr, Tel: 031-380-0225)
- 최초투고일 08. 04. 28
최종접수일 08. 06. 12

[부록]

