

# 실내 환경에서의 현재 및 미래 위치 질의를 위한 색인과 동적 히스토그램 기법

김미경, 박현주

한밭대학교 정보통신공학과

e-mail : [mkkim922@naver.com](mailto:mkkim922@naver.com)

## An Index for Querying Current and Future Location and Dynamic Histogram Method in Indoor Environment

Mikyung Kim, Hyun-Ju Park

Dept of Information Communication Engineering, Hanbat National University

### 요약

이동객체 색인에 관한 기존 연구는 실외 환경의 빠른 속도로 이동하는 많은 이동객체들을 대상으로 과거 궤적과 현재 위치 및 미래 위치를 검색하는 색인 기법들이었다. 그러나 쇼핑몰이나 대형 박물관과 같은 건물의 실내나 놀이동산과 같은 한정된 영역의 이동객체들은 실외 환경에서처럼 빠른 이동속도를 가지고 있지 않다. 또한 이동객체들의 이동하는 영역이 한정되어 있을 뿐만 아니라 최대 수용인원 등이 정해져 있다.

본 논문에서는 이러한 특성을 갖는 실내 환경에 적합한 색인 기법으로 현재 그리고 미래 위치 질의 처리를 위한 공간 색인과 시간 색인을 분리한 두 개의 색인 구조를 갖는 새로운 색인 기법과 가까운 미래 시간의 특정 영역에 있는 이동객체들의 수에 대한 질의응답을 위한 시공간 히스토그램을 구성하고, 적용하는 방법을 제안한다.

### 1. 서론

최근 GPS(Global Positioning System) 기술 및 무선통신 기술의 발달로 위치기반 서비스에 대한 요구가 증가되고 있다. 그러나 GPS 기술을 이용한 위성기반 기술은 실외에서는 위치기반 서비스를 지원하기 위해 수 미터의 정확성을 보장하는 위치를 제공하고 있지만 실내에서는 그 기술을 사용할 수 없다. 그러므로 실내 위치 결정을 위한 일반적인 기술은 초음파나 라디오 주파수를 이용하거나 또는 이들의 기술을 조합한 것이다. 이들 기술로 생성한 위치 정보는 매우 정확도가 높다. 예를 들면 BAT 시스템[1]의 정확도는 보통 실제 위치의 9cm 이내로 계산된다. 이러한 실내위치정보 유용성은 추적 정보 제공과 같은 새로운 부류의 응용의 원인이 되었다. 하지만 사용자가 물리적인 공간 주변을 이동하기 때문에 사용자들에 따라 다르게 적용되는 위치인식 응용들이 필요하다.

위와 같은 기술들을 이용하여 위치 기반 서비스를

제공하려면 실외 환경뿐만 아니라 실내 환경 둘 다 이동객체의 위치를 저장하고 검색을 하기 위한 이동객체 데이터베이스가 필요하다. 또한 이동객체의 위치 정보는 시간이 지남에 따라 지속적으로 변하므로 이러한 동적인 데이터를 효율적으로 처리할 수 있는 색인에 대한 연구가 필요하다.

본 논문에서 제안한 색인 구조는 정적인 공간 차원의 색인과 동적인 시간 차원의 색인을 분리하여 구성한다. 즉, 실내의 공간차원을 중복되지 않게 분할하고 각 분할 공간에 대해 시간 차원의 색인을 유지하여 이동객체의 현재 및 가까운 미래위치에 대한 질의응답을 위한 색인기법을 제안한다. 그리고 각 셀에 있는 이동객체들의 수에 대한 질의응답을 할 수 있도록 질의 선택도 추정에 사용하는 히스토그램 기법을 사용하여 고정된 크기로 분할한 각 셀을 하나의 버킷으로 하는 히스토그램을 구성하여 이동객체들의 수를 요약한다.

## 2. 문제 정의와 관련연구

### 2.1 데이터 모델

시공간 데이터 모델에서는 시간이 지남에 따라 지속적으로 변하는 이동객체의 위치갱신에 대한 오버헤드를 줄이기 위해 이동객체의 정보를 시간함수로 표현한다. 이동객체  $O_i$ 의 위치 정보는  $(x_i, y_i, t_i)$ 가 되며, 이동객체 데이터베이스에 저장되는 이동객체  $O_i$ 의 정보로는 (객체 식별자, 현재 위치, 이동속도, 이동 방향, 이동시간) 등으로 표현할 수 있다. 이러한 이동객체 데이터베이스에 저장된 객체에 대한 정보의 명시적인 갱신은 객체의 이동속도나 방향이 변하거나 임계 시간이 지난 경우에 이동객체의 보고에 의해서 이루어진다.

### 2.2 질의 형식

이동객체에 대한 질의는 크게 이동객체의 미래 위치에 대한 질의와 이동객체의 위치이력에 대한 질의 두 가지로 분류할 수 있다. 그 중 미래 위치 질의는 이동객체의 현재 위치와 이동속도, 이동방향을 저장하고[2] 계산을 통하여 미래 위치를 예측하여 응답할 수 있다. 그리고 위치 이력[3]에 대한 질문은 다시 좌표 기반 질의와 궤적 기반질의로 분류할 수 있는데 좌표 기반 질의는 주어진 시간 구간과 주어진 영역 내에 있는 이동객체들을 포함하는 질의이고, 궤적 기반질의는 위상과 속도 같은 궤적 정보를 포함하는 질의이다.

그 중 본 논문은 이동객체들의 현재 및 미래 위치 질의 응답을 위한 질의처리와 특정 시간에 특정 영역의 객체들의 수에 대한 질의응답에 초점을 두었다.

### 2.3 이동 객체를 위한 색인과 질의 처리 관련 연구

이동객체 데이터베이스 사용자들로부터 전송되어 오는 많은 위치 정보를 효율적으로 저장, 검색, 갱신 등 관리를 할 수 있는 색인 연구에는 과거 궤적 검색을 위해 이동객체의 위치를 선분으로 모델링한 시공간 궤적인 3DR-트리[2]와 궤적 보존에 초점을 맞춘 TB-트리[3]가 있으며, 궤적 색인으로 공간적인 특성에 초점을 둔 SETI[4] 등이 있다.

또한 현재 위치 데이터를 점(Point)으로 나타내는 R-트리와 해쉬 기반 색인이 있으며, 현재 및 가까운 미래 위치를 시간에 대한 선형 함수로 표현한 TPR-트리[5]가 있다. TPR-트리는 이동객체를 시간에 대한 선형 함수로 표현한 색인 구조로 이동객체의 위치를 단순한 좌표로 저장하지 않고, 이동 방향과 이

동속도를 나타내는 벡터를 저장하여 그 변화가 특정 임계 값 이하인 경우 위치 관련 정보를 갱신하지 않아 갱신 횟수를 줄여 효율성을 높인 색인 구조이다.

### 2.4 히스토그램과 요약 질의 처리

히스토그램은 객체들을 커버하는 경계 값(bounding value)과 그 경계 값에 속하는 객체들의 수로 구성되어 있는 버킷들의 집합이다. 예를 들면, 공간 버킷은 공간 경계 사각형과 그 공간 경계 사각형 내에 있는 객체들의 수로 구성된다. 히스토그램은 한 버킷 내의 객체들이 균일하게 분포되어 있다고 가정한다.

히스토그램을 이용한 객체들의 수 평가 방법은 한 질의와 오버래핑하는 모든 버킷들을 선택하고, 해당 버킷에 대해 그 질의들을 만족하는 객체들의 수를 평가한 후 각 버킷에 대한 객체들의 평가된 수의 합을 구하게 된다. 히스토그램은 원래 객체가 아닌 히스토그램 구성과 관련된 샘플데이터를 이용하여 갱신하게 되며, 샘플 데이터가 갖는 정보는 식별자와 최종 갱신시간, 그 시간의 공간 위치, 속도 등의 정보들로 구성된다[6]. 히스토그램 구성은 공간 객체들의 공간 편향을 최소화하기 위해 최근에 Acharya[7] 등이 제안한 MinSkew 알고리즘을 이용하여 구성할 수 있다.

## 3. 실내의 이동객체들의 질의 처리를 위한 색인 및 히스토그램 구성

실내 환경에서 이동객체의 특징은 이동속도가 느리다는 것과 객체의 이동 공간의 시작 지점과 끝나는 지점이 같으며, 객체의 이동공간과 이동객체의 수가 한정되어 있다는 것이다.

### 3.1 실내 환경에서 가능한 질의

실내 환경에서 가능한 질의로는 놀이동산이나 대형 쇼핑몰에서 깜짝 이벤트를 위해 “현재(10분 후에) 가장 많은 이동객체들이 있는(있을 것 같은) 곳은?”과 같은 질의와 미아와 같은 특정 이동객체의 위치를 묻는 “미아의 현재 위치와 5분 이내에 도달할 수 있는 위치는?”과 같은 질의를 생각할 수 있다. 이러한 질의 처리를 위해서는 시공간 데이터베이스에서 이동객체의 미래 위치뿐만 아니라 특정 영역 내의 이동객체의 수를 알아낼 수 있어야 한다.

본 논문에서는 현재 및 가까운 미래 위치를 알아내기 위해 R-트리를 기반으로 하는 TPR-트리를 이

용하고, 특정 영역 내의 이동 객체의 수를 알아내기 위해 시공간 히스토그램을 이용한다.

### 3.2 실내 환경에 적합한 시공간 색인기법

위치 데이터 집합들에 대한 공간차원과 시간 차원들의 특성에는 중요한 차이점들이 있다. 즉, 공간차원들의 경계들은 상수이거나 시간차원이 변화함에 따라 연속적으로 변하는 위치 데이터에 비해 공간차원의 범위는 변화하지 않는다. 그러므로 색인구조는 공간차원을 정적으로 분할할 수 있으며, 시간 차원을 각 분할된 공간 내에서 단지 1-D(시간) 차원의 선들로 색인구조를 갖게 할 수 있다.

본 논문에서는 앞에서 예시한 것과 같은 질의들을 처리할 수 있는 효율적인 색인을 위해 공간을 오버래핑되지 않도록 고정된 크기의 셀들로 분할하고, 각 셀들 내에 있는 이동객체들에 대한 TPR-트리 색인을 구성하여 공간차원과 시간 차원을 분리하는 색인구조를 제안한다.

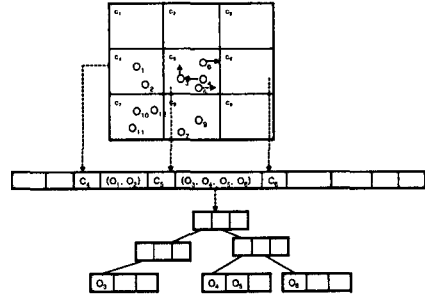
또한 가까운 미래에 대한 시공간 질의 처리를 위한 시간 매개변수 경계 사각형(Time Parameterized Bounding Rectangle)의 최대 크기 값은 셀의 크기와 같게 하고 이 값보다 커지는 경우 즉, 위치가 최대 경계 사각형 값을 벗어나는 이동객체들은 현재 셀을 벗어난 것이므로 현재의 셀에서 제거하고, 새로운 셀의 TPR-트리로 편입시킨다. 이 때 해당 셀의 TPR-트리 노드에 오버플로우가 발생하거나 언더플로우가 발생하게 되면, 해당 셀들에 대해 TPR-트리를 재구성하도록 한다.

앞의 예에서처럼 이동객체의 수가 가장 많은 영역을 묻는 질의응답은 해당 셀들의 식별자와 그 셀에 속한 이동객체의 수가 된다.

각 셀의 색인정보는 (셀 식별자, tptr, hptr)로 구성된다. 셀 식별자는 공간을 같은 크기로 나눈 셀들을 구별하기 위한 식별자이고, tptr은 이동객체의 위치 정보에 대한 TPR-트리 포인터이다. 그리고 hptr은 각 셀의 요약정보를 구성하고 있는 히스토그램에 대한 포인터이다. 각 셀에 있는 이동객체의 위치 정보는 셀 내의 이동객체 식별자와 그 객체의 현재 위치, 이동속도, 이동방향, 진입 가능한 셀 식별자, 또한 최근 지나온 셀의 식별자 등의 정보들을 데이터 파일의 형태로 유지한다. 이때 데이터 페이지는 같은 공간 셀에 있는 객체 정보만 포함하고, 각 데이터 페이지는 그 셀에 포함된 객체가 없을 때까지 유지한다.

또한 공간차원에 대한 그리드 셀의 색인과 공간

차원에 대한 TPR-트리 색인은 메모리에 저장하고, 셀 공간정보와 이동객체의 위치 정보, 셀에 대한 히스토그램은 디스크에 저장을 하여 유지한다.



[그림 1] 공간색인과 시간 색인을 분리한 색인구조

### 3.3 제안 색인에서 영역질의 처리

앞에서 언급한 것처럼 영역 검색을 위한 시공간 질의로는 다음과 같은 것들이 있다. “특정인의 현재 위치와 10분 후의 위치” 또는 “10분 후 가장 인파가 많은 곳(셀)” 등이 있을 수 있는데, 이러한 시공간 영역질의 검색은 본 논문에서 제안한 색인 구조에서 다음과 같은 단계를 통하여 수행될 수 있다.

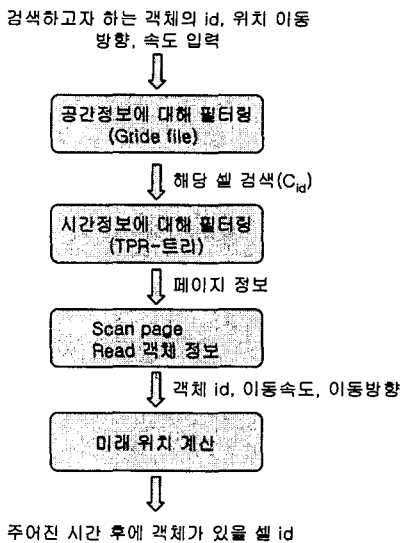
특정 객체의 미래 위치 검색을 위한 질의가 입력되면, 먼저 그 객체가 어느 셀에 해당되는지를 알아내기 위해 공간정보에 대해 필터링을 하게 된다. 그러면, 그 객체가 있는 해당 셀이 검색이 되고, 셀의 식별자  $C_{id}$ 를 반환한다. 그런 다음 반환 받은 셀의 id에 해당하는 디렉토리에서 이동객체 정보를 시간정보에 대해 작성한 색인 저장 주소를 읽게 되고, 시간정보에 대해 필터링을 하게 된다. 이 단계에서 해당 셀에 있는 객체의 노드를 검색하게 된다. 그림 2는 이와 같은 단계를 도식화 한 것이다.

또한 후자와 같은 질의는 우리가 구성한 공간 경계 사각형과 속도 경계 사각형 그리고 각 경계 사각형들에 속하는 객체들의 수로 구성된 공간 히스토그램의 요약정보를 이용하여 응답을 할 수 있다.

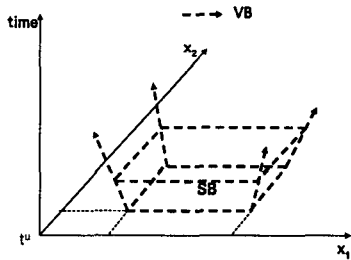
### 3.4 TPR-트리를 위한 시공간 히스토그램 구성

시공간 질의는 공간 사각형(Spatial Rectangle)과 시간으로 구성된다. 이 때 시간은 현재 시간보다 크거나 같은 시간이다. 시공간 히스토그램을 구성하는 시공간 버킷은 그림3에서 보는 것처럼 공간 경계 사각형(Spatial Bounding Rectangle), 속도 경계 사각형(Velocity Bounding Rectangle) 그리고 각 경계

사각형들에 속하는 객체들의 수로 구성된다.



[그림 2] 영역질의 처리 단계



[그림 3] 시공간 버킷

시공간 히스토그램은 공간 히스토그램과 시간 히스토그램으로 구성되는데, 본 논문에서는 최용진이 제안한 시공간 히스토그램을 구성하기 위해 constructSTH[6] 알고리즘을 이용하였으며, MinSkew[7] 알고리즘을 이용하여 하나의 셀이 한 버킷이 되도록 하는 공간 히스토그램과 속도 히스토그램을 작성한다.

#### 4. 결론

본 논문에서는 공간차원과 시간차원을 분리하여 실내 환경에 적용 가능한 색인기법을 제안하였다. 영역 질의 처리를 위한 경계 사각형은 셀의 크기와 같

게 하였다. 즉, 시간이 지남에 따라 확장되는 이동객체에 대한 매개변수 경계 사각형의 최대 크기를 셀의 크기와 같게 하여, 이동객체의 위치가 이 경계 사각형을 벗어나는 경우 현재 셀에서 이 이동객체를 제거하고 새로 진입한 이웃 셀에 삽입이 되도록 처리하였다. 또한 셀에 새로운 이동객체가 삽입되거나 제거될 때 오버플로우나 언더플로우가 발생하면 TPR-트리를 재구성하도록 하였다. 이러한 방법은 TPR-트리의 재구성으로 인한 오버헤드가 발생할 수 있지만 실내 환경은 실외 환경과 달리 이동객체의 이동속도가 느릴 뿐만 아니라 이동객체의 수 또한 실외 환경처럼 많지 않으므로 큰 문제가 되지 않는다.

#### 참고문헌

- [1] Harter, A., Hopper, A., Steggles, P., Ward, A., Webster, P. "The Anatomy of a Context-Aware Application." In *Mobile Computing and Networking* (Seattle, WA, USA, August 1999), pp. 59-68.
- [2] Yannis, Theodoridis, "Spatio-Temporal Indexing for Large Multimedia Applications", In Proc. Of the 3rd IEEE Conf. on Multimedia Computing and Systems, pp. 441-448, June 1996.
- [3] Pfoser, D., Jensen, C. S., AND Theodoridis, Y., "Novel Approaches in Query Processing for Moving Object Trajectories.", In Proceedings of the 26st VLDB Conf. (Cairo, Egypt, September 2000), pp. 395-406.
- [4] V. Prasad Chakka, Adam C. Everspaugh, Jignesh M. Patel, "Indexing Large Trajectory Data Sets with SETI", Proceedings of the 2003 CIDR Conference, 2003.
- [5] S. Saltenis, C. S. Jensen, S. T. Leutenegger, and M. A. Lopez, "Indexing the Positions of Continuously Moving Objects" , In Proc. ACM SIGMOD on Management of data, pp. 331 - 342, 2000.
- [6] 최용진, "시공간 질의를 위한 선택률 추정과 비용 모델." KAIST 박사학위 논문, 2003.
- [7] S. Archarya, V. Poosala, and S. Ramaswamy. "Selectivity Estimation in Spatial Databases." In Proceedings of the ACM SIGMOD Conference on Management of Data, pp. 13-24, 1999.