

연속적인 스카이라인 질의의 정적 유효 영역을 이용한 효율적인 처리

(Efficient Processing using Static Validity Circle for Continuous Skyline Queries)

이 종 혁 [†] 박 영 배 ^{**}
 (ZhongHe-Li) (Young-Bae Park)

요 약 시간이 변함에 따라 위치 좌표를 변경하는 모바일 환경에서 이동 객체는 자신의 위치를 기준으로 질의를 요청한다. 연속적인 스카이라인 질의 처리를 위한 효율적인 영역 결정 기법에서는 이동 객체의 속도와 방향과는 무관한 최적화된 스카이라인 영역(OSR: Optimal Skyline Region)을 미리 계산하여 질의에 답할 수 있다. 이에 따라 이동 객체의 위치를 중심으로 하고 가장 가까운 영역 변까지의 거리를 반경으로 하는 원(*Vcircle*: Validity Circle)을 유효 영역으로 결정하여 질의 발생 빈도를 감소하는 기법이 제안되고 있다. 그러나 원은 최초 질의가 발생한 시점의 이동 객체 위치에 따라 면적이 가변적이므로 질의 발생 빈도도 가변적이고, 객체가 최적화된 스카이라인 영역 내에서 이동하는 경우에 재질의가 빈번하게 발생하는 문제점이 발생한다. 예를 들어 사용자는 “현재 위치에서 가깝고 숙박료가 싸고, 해변과의 거리가 가까운 호텔을 검색하라”는 질의를 할 수 있다. 이 경우, 이동 객체와 대상 객체의 거리뿐만 아니라 대상 객체의 다중 속성을 고려해야하고, 스카이라인 질의 결과는 이동 객체의 현재 위치와 대상 객체의 거리에 따라 유효하지 않을 수 있으므로 이동 객체의 위치 변경에 따라 스카이라인을 재계산해야 하며, 새로운 결과를 요청하기 위해 연속적인 질의가 발생한다.

이 논문에서는 항상 불록 다각형을 형성하는 최적화된 스카이라인 영역의 특징을 이용하여 스카이라인 영역의 최대내부원(*IVcircle*: Interior Validity Circle)을 정적 유효 영역으로 결정하는 기법을 제안한다. 실험을 통하여 영역내의 평균 질의 발생 빈도를 기존의 *Vcircle*을 이용한 동적 유효 영역 결정 기법보다 평균 52.55%가 감소함을 보인다.

키워드 : 스카이라인 질의, 위치 의존 질의, 이동객체

Abstract Moving objects in a mobile environment to change their position based on the change of time require a query with their position as a basis. Efficient Regional Decision for Continuous Skyline Queries requires objectively pre-calculating the OSR(Optimal Skyline Region) regardless of the speed and direction of the moving objects. It proposes techniques to reduce the frequency of continuous queries by choosing a *VCircle*(Validity Circle) as safe location which has radius of the distance to the closest region with position on the moving object at center. But, a *VCircle*'s area varies based on the Moving object's position from first marked time of continuous query. Therefore, the frequency of its continuous query is variable and also when the object moves inside of OSR, query can re-occur frequently.

In this paper, we suggest a technique of selecting an *IVCircle*(Interior Validity Circle) in a Skyline Region as the static Safe Region using the characteristics of the OSR. An Interior *IVCircle* can be calculated in advance when the OSR is decided. Our experiment shows that the frequency of using *IVcircle* as safe region reduced than that of using *VCircle* as safe region by 52.55%.

Key words : Skyline Query, Location-dependent Query, Moving Object

† 학생회원 : 명지대학교 컴퓨터공학과
 mslzh@mju.ac.kr
 ** 종신회원 : 명지대학교 컴퓨터공학과 교수
 parkyb@mju.ac.kr
 논문접수 : 2006년 8월 4일
 심사완료 : 2006년 10월 30일

1. 서 론

이동 객체에 대한 대표적인 위치 의존 질의(Location Dependent Query)는 질의의 현재 위치와 가장 가까운 대상 객체를 검색하는 최 근접 질의(Nearest Neighbor

Query)이다. “현재 위치에서 가장 가까운 호텔을 검색하라”는 질의를 예로 들 수 있다. 위치정보를 변경하는 모바일 환경에서 이동 객체는 질의를 요청한 후 계속해서 위치를 변경하기 때문에 기존 질의 결과가 현재 위치에서는 유효하지 않을 수 있으므로, 위치 변경에 따른 연속적인 질의가 발생한다[1]. 연속적인 질의는 네트워크 부하의 증가와 전체 대상객체를 빈번히 재계산해야 하는 문제점이 발생한다. 이러한 문제점을 해결하기 위하여 질의 결과의 유효성 검증을 통한 효율적인 질의 결과의 갱신 기법들[2,3,7-9]이 연구되고 있다.

스카이라인 질의(Skyline Query)는 전체 대상 객체 집합에서 대상 객체의 여러 속성들을 다른 객체가 지배하지 않는 관심 있을만한 객체 집합을 검색한다[3]. “숙박료가 싸고, 해변과의 거리가 가까운 호텔을 검색하라”의 예처럼 스카이라인 질의는 대상 객체의 단일 속성이 아닌 다중 속성을 고려해야하며, 질의 결과는 단일 객체가 아닌 대상 객체 집합으로 구성된다.

이동 객체에 대한 질의 처리 기법[2,3,7-9]들은 단순히 동적 속성인 대상객체와의 거리만 고려하였고, 기존의 스카이라인 질의 처리 기법들[4-6,10]은 질의 위치와 상관없이 숙박료, 해변과의 거리 등 대상 객체의 정적 속성만을 고려하였다.

연속적인 스카이라인 질의에 위한 효율적인 영역 결정 기법[1]에서는 이동 객체의 속도와 방향과는 무관한 최적화된 스카이라인 영역(OSR: Optimal Skyline Region)을 미리 계산하여 질의를 처리한다. 또한 이동 객체의 위치를 중심으로 하고 가장 가까운 영역 변까지의 거리(MINDIST)를 반경으로 하는 원(Vcircle)을 유효 영역으로 결정하여 질의 발생 빈도를 감소하는 기법을 제안하고 있다. 그러나 원(Vcircle)은 최초 질의가 발생한 시점의 이동 객체 위치에 따라 면적이 가변적이므로 질의 발생 빈도도 가변적이고, 객체가 최적화된 스카이라인 영역(OSR) 내에서 이동하는 경우에도 재 질의가 빈번하게 발생하는 문제점이 있다.

연속적인 스카이라인 질의의 발생 빈도를 결정하는 요소는 유효 영역의 면적이다. 즉 질의 발생 빈도는 유효 영역의 면적에 반비례하므로 유효 영역의 면적이 최적화된 스카이라인 영역(OSR)을 얼마나 많이 차지하느냐에 따라 질의 발생 빈도가 결정된다.

이 논문에서는 항상 불록 다각형을 형성하는 최적화된 스카이라인 영역(OSR)의 특징을 이용하여 스카이라인 영역의 최대내부원(IVcircle)을 정적 유효 영역으로 결정하는 기법을 제안한다. 최대내부원(IVcircle)은 질의 발생 시점이 아닌 OSR를 결정할 때 선 처리(pre-processing)할 수 있고, 상대적으로 원(Vcircle)보다 OSR를 차지하는 비율이 높아 유효 영역으로 결정하는

것이 효율적임을 다양한 실험을 통하여 증명한다.

2. 관련연구

2.1 스카이라인 질의(Skyline Queries)

스카이라인이란 전체 객체 집합에서 다른 객체보다 나쁘지 않는 관심 있을 만한 객체 집합이다. 예를 들어 “숙박료가 싸고, 해변과의 거리가 가까운 호텔을 검색하라.” 스카이라인 질의에서 숙박료가 5만원이고, 해변과의 거리가 1km 인 호텔 a와 숙박료가 8만원이고 해변과의 거리가 5km인 호텔 b의 경우 호텔 a는 b를 지배한다($h_a < h_b$)[1].

그림 1은 호텔의 숙박료(y)와 해변과의 거리(x)에 대한 속성 값을 좌표 평면상에 표현하며, 사용자가 관심 있을 만한 호텔은 숙박료와 해변과의 거리 속성에 대해서 다른 호텔이 지배하지 않는 a, i, k로서 a, i, k를 잇는 선을 스카이라인이라 하고, 이 객체들의 집합이 스카이라인 질의의 결과이다.

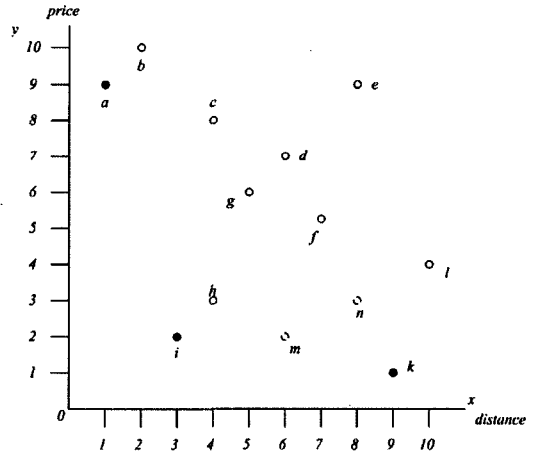


그림 1 호텔의 정적 속성 관계

2.2 이동 객체에 대한 스카이라인 질의

2.2.1 이동 객체에 대한 스카이라인 질의

이동 객체란 시간에 따라 위치정보가 계속 변경되는 공간 객체를 의미한다. 이동 객체는 질의를 요청한 후 계속해서 위치를 변경하기 때문에 처음 질의 결과가 현재 위치에서는 유효하지 않을 수 있으므로, 위치 변경에 따른 연속적인 질의가 발생한다. 연속적인 질의는 네트워크 부하의 증가와 전체 대상 객체에 대해 빈번한 재계산에 따른 비용 증가 등 문제점이 발생한다. 연속적인 질의로 인한 계산 횟수의 증가를 줄이기 위하여 초기에 계산된 결과의 유효성을 검증하여 결과를 갱신하는 기법들이 연구되고 있다[2,9].

기존의 대상 객체의 정적 속성에 대한 스카이라인 질의의 결과는 질의의 위치와 상관없이 항상 고정이다. 그러나 이동 객체에 대한 스카이라인 질의는 이동 객체의 위치 변경에 따른 동적 속성인 이동 객체와 대상 객체의 거리 때문에 질의의 위치에 따라 결과는 가변적이다. 질의 결과 집합의 객체들은 대상 객체의 정적인 속성에 대한 스카이라인 객체와 정적인 속성에 대해서 지배하는 객체보다 이동 객체와 가까운 객체들로 이루어진다.

이동 객체에 대한 스카이라인 질의의 결과(*Skyline_{total}*)는 질의의 위치와는 상관없이 정적 속성에 대한 스카이라인 객체(*Skyline_{static}*)를 항상 포함하고, 정적 속성에 대해 지배를 받을지라도 이동 객체와의 거리라는 동적 속성으로 다른 객체가 지배하지 않는 가변적인 객체(*Skyline_{dynamic}*)도 질의 결과에 포함된다. 즉 이동 객체에 대한 스카이라인 질의의 결과는 객체의 정적 속성에 대해 지배하는 객체보다 이동 객체와 가까운 객체들과 현재 위치에서 가장 가까운 객체를 포함한다.

예를 들어 호텔의 정적 속성 관계가 그림 2(a)와 같을 경우 “현재 위치(Q_1)에서 가장 가까운 숙박료가 싸고, 해변과의 거리가 가까운 호텔을 검색하라.”는 질의의 결과는 그림 2(b)와 같이 대상 객체의 정적속성(숙박료, 해변과의 거리)에 대한 스카이라인 H1 과 H4를 질의의 위치에 관계없이 포함하고, Q_1 은 동적 속성인 H3과 이동 객체와의 거리가 정적 속성에 대해 지배하는 H1과 H4와의 거리보다 상대적으로 가까움으로 H3도 스카이라인 질의 결과에 포함된다.

2.2.2 스카이라인 영역의 정의

스카이라인 영역(*SR*: *Skyline Region*)이란 질의의 위치에 따라 결과에 포함될 수 있는 객체가(*Skyline_{dynamic}*)

정적 속성에 대하여 p_i 를 지배하는 객체들보다 질의의 위치와 가까운 영역으로 대상 객체 p_i 가 스카이라인 질의 결과에 포함될 수 있는 영역이다[1].

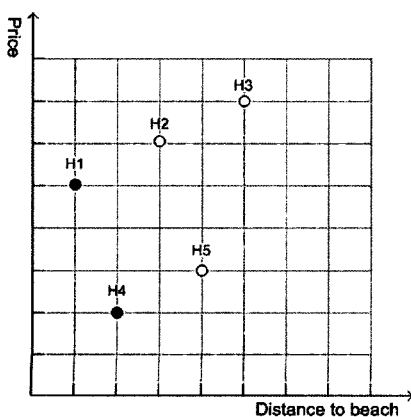
$$SR_i = \{qdist(q, p_i) < dist(q, p_j), \forall p_j < p_i\}$$

p_i 의 스카이라인 영역 SR_i 은 p_i 의 위치좌표와 정적인 속성에 대해서 p_i 를 지배하는 모든 p_j 의 위치좌표가 생성하는 수직 이등분선들이 형성하는 영역으로 계산할 수 있다.

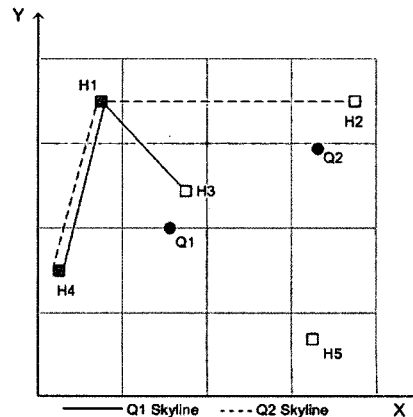
그림 3(a)와 같이 전체 객체 집합에서 호텔 H7를 정적 속성으로 지배하는 호텔들의 집합은 {H4}이며, 호텔의 위치 좌표가 그림 3(b)와 같을 경우, H7과 지배하는 호텔들과의 수직이등분선이 형성하는 스카이라인 영역은 그림 3(b)의 음영부분이다. 이동 객체가 이 영역 안에 포함되어 있을 경우 H7은 정적 속성에 대해서 지배하는 객체 H4보다 상대적으로 가깝기 때문에 항상 스카이라인 결과에 포함된다.

이동 객체의 위치 변경에 따라 현재 위치와 전체 객체의 거리를 계산하여 스카이라인을 재계산하는 것은 비효율적이다. 대상 객체의 정적인 속성을 이용하여 대상 객체의 스카이라인 영역을 미리 계산하고 현재 위치와 영역간의 관계를 판별함으로써 전체 객체와의 거리를 재계산하지 않고도 위치에 따른 스카이라인을 계산할 수 있다[1].

그림 4(a)는 그림 3의 데이터 집합에 대한 호텔 H2,H3의 스카이라인 영역을 나타내고 그림 4(b)는 호텔 H5,H6,H7의 스카이라인 영역을 나타낸다. 그림 4(c)는 모든 스카이라인 영역을 나타내고 있다. 그림 4(c)에서 Q_1 을 포함한 영역은 SR_3 , SR_6 이므로, 최적화된 스카이라인 영역 OSR_1 내에서 Q_1 의 스카이라인은 항상 정적

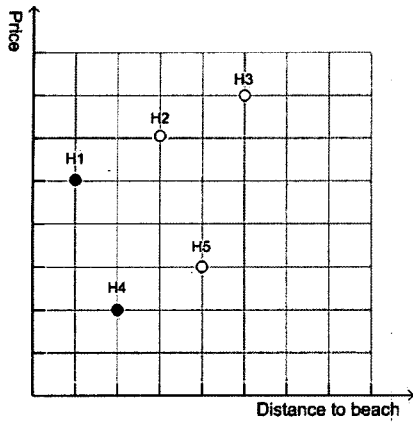


(a) 호텔의 정적 속성 관계

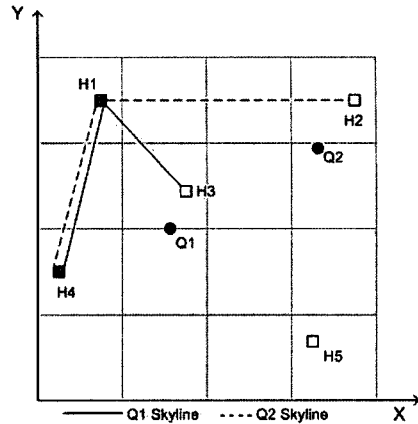


(b) 호텔의 공간 좌표

그림 2 질의 위치에 따른 스카이라인

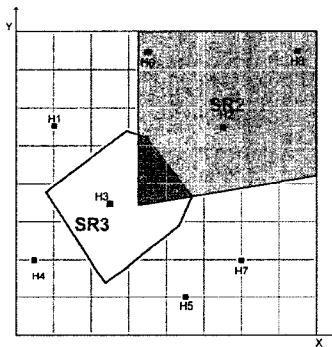


(a) 호텔 H7의 정적 속성 관계

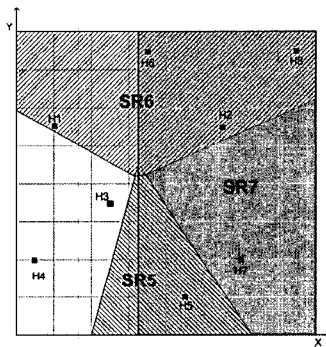


(b) H7의 스카이라인 영역

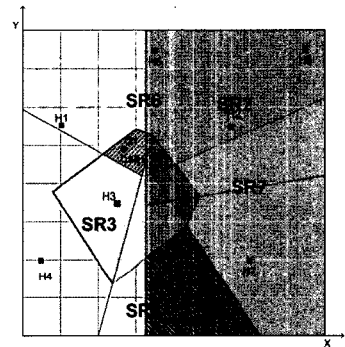
그림 3 H7의 정적 속성 관계에 의한 스카이라인 영역



(a) 호텔 H2,H3의 스카이라인 영역



(b) 호텔 H6,H7의 스카이라인 영역



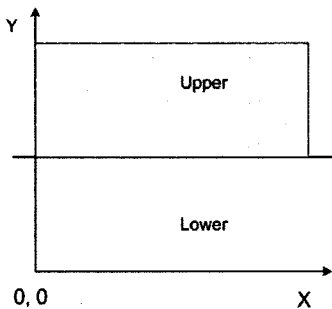
(c) 전체 스카이라인 영역

그림 4 호텔의 스카이라인 영역

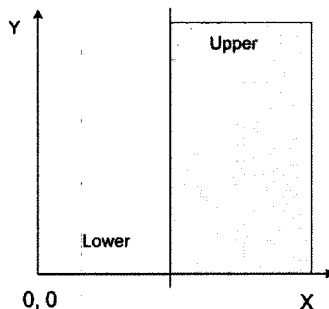
속성에 의한 {H1, H4, H8}과 동적 속성에 의한 {H3, H6}이다. 마찬가지로 Q_2 를 포함한 영역은 SR_5, SR_7 이므로 OSR_2 영역 내에서 Q_1 의 스카이라인은 항상 정적 속성에 의한 {H1, H4, H8}과 동적 속성에 의한 {H5, H7}이다.

2.2.3 스카이라인 영역 결정 방법

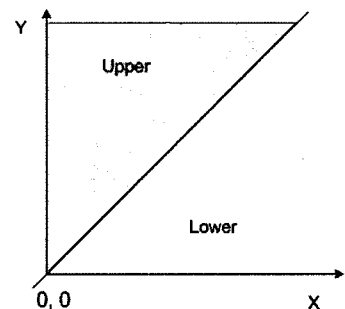
좌표 평면상에 선분에 의하여 만들어지는 영역은 부등식으로 나타낼 수 있다. 직선의 방정식을 $b \cdot y = a \cdot x + c$ 로 표현할 때, 직선의 윗부분 영역의 부등식은 $b \cdot y > a \cdot x + c$ 이고 아랫부분 영역은 $b \cdot y < a \cdot x + c$ 이다.



(a) $b \cdot y > c$
($a=0, b \cdot c > 0$)



(b) $a \cdot x > c$
($b=0, a \cdot c > 0$)



(c) $b \cdot y > a \cdot x + c$

그림 5 부등식 영역의 표현

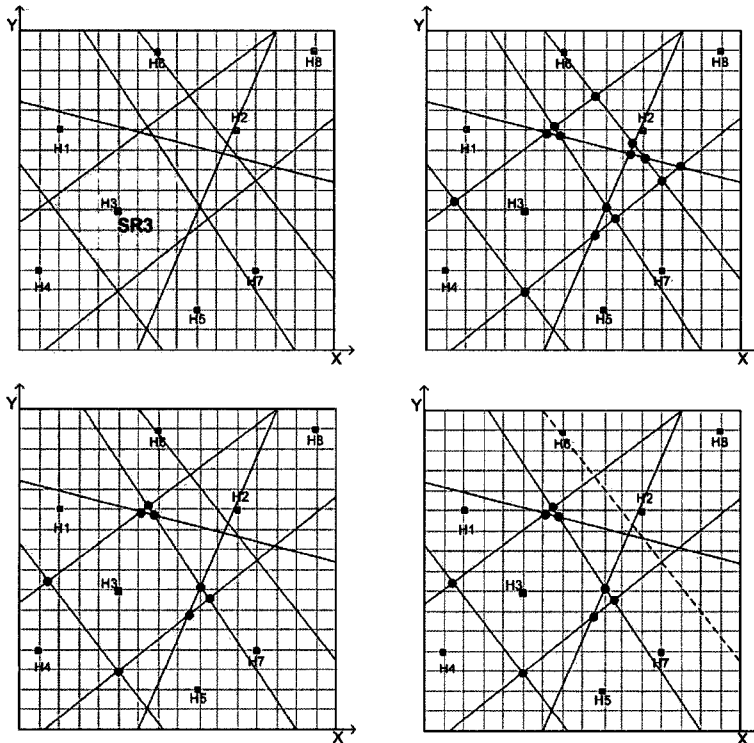


그림 6 스카이라인 영역 결정 4단계

대상 객체 p_i 의 스카이라인 영역(SR_i)의 정보는 영역을 형성하는 수직이등분선을 계산하고, 수직이등분선에 대한 p_i 의 위치를 판별하여 부등식 형태로 표현한다. 따라서 부등식 영역의 표현을 이용하여 스카이라인 영역을 결정하는 방법은 다음 4단계로 나눈다[1].

- [1단계] 정적 속성에 대하여 지배하는 객체들과 수직이등분선을 계산
- [2단계] SR_i 를 형성하는 수직이등분선들의 모든 교차점 계산
- [3단계] 연립부등식을 만족하는 교차점 계산(최적화된 영역의 꼭지점)
- [4단계] 꼭지점 2개를 지나지 않는 선분 제거

3. 다각형의 최대내부원을 이용한 정적 유효 영역 결정 기법

3.1 기존의 연속적인 스카이라인 질의 처리 과정

기존의 이동 객체(클라이언트 혹은 단말기)의 연속적인 스카이라인 질의는 다음과 같다. 그림 7(a)처럼 화살표 방향으로 영역에 진입한 이동 객체는 Q_1 지점에서 위치좌표(x_1, y_1)를 서버에 전송하면서 처음으로 질의를 요청한다. 서버는 이동 객체가 위치한 영역 OSR_1 을 판단하

고 Q_1 에서 가장 가까운 영역 변까지의 거리 $MINDIST(R_1)$ 을 계산하여 선 처리된 스카이라인 결과 집합과 함께 이동 객체에 전송한다. 이동 객체가 Q_2 까지 이동하면(그림 7(b)) 전 질의의 위치(Q_1)에서 Q_2 까지 이동한 거리인 S_1 을 계산하여 Q_1 질의에서 받은 스카이라인 결과의 유효성을 판단한다. 만약 $S_1 < R_1$ 이면 그 결과는 계속 유효하지만 $S_1 > R_1$ 이므로 결과가 유효하지 않기에 현재위치 Q_2 에서 새로운 질의가 발생한다. 마찬가지로 Q_3 로 이동한(그림 7(c)) 이동 객체는 전 질의의 위치 Q_2 에서 Q_3 까지 이동한 거리 S_2 와 전 유효 영역의 반경인 R_2 를 비교하여 $S_2 > R_2$ 이므로 재 질의를 하게 된다.

여기서 Q_1, Q_2, Q_3 은 모두 최적화된 스카이라인 영역 OSR_i 내에 있으므로 세 번의 질의 결과는 똑 같다. 그러나 이동 객체는 R_i 을 반경으로 하는 원을 유효 영역($Vcircle$)으로 간주하고 그 $Vcircle$ 이탈여부에 따라 재질의 여부를 결정한다. 이러한 현상은 이동 객체의 위치에 따라 최적화된 스카이라인 영역(OSR_i)의 유효 영역($Vcircle$)이 동적으로 결정되므로 최적화된 스카이라인 영역(OSR_i)을 효과적으로 활용하지 못했기 때문이다.

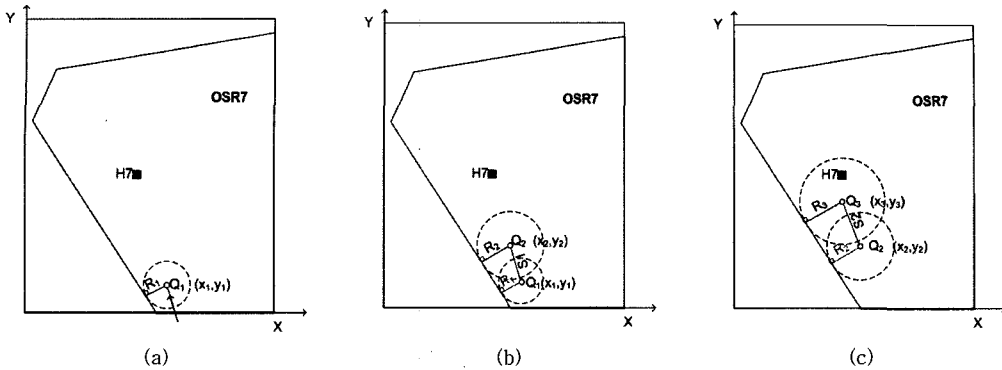


그림 7 Vcirde 동적 유효 영역

그러나 이 논문에서는 IVcirde을 유효 영역으로 이용하여 이동 객체가 최적화된 스카이라인 영역(OSR₇)을 벗어날 때까지는 최대한 서버에 재질의 하지 않도록 하는데 있다.

3.2 IVcirde을 이용한 연속적인 스카이라인 질의 처리 과정

3.2.1 내부원의 면적과 블록다각형의 면적의 관계

다각형의 최대내부원을 이용하여 최적화된 스카이라인 영역(OSR)의 정적 유효 영역을 결정하는 방법은 유효 영역의 면적을 최대한 크게 하여 영역 내에서의 불필요한 질의를 줄이기 위함이다. 그림 8의 그래프는 실험 결과로서, 정다각형의 각의 개수(x)가 많을수록 내접하는 원의 면적과 정다각형의 면적(OSR)의 비가 높다는 것을 알 수 있다. 대상 객체의 수가 많을수록 지배 관계가 복잡하기 때문에 OSR을 형성하는 다각형의 각의 개수는 상대적으로 많아진다. 또한 OSR은 대상 객체의 정적 속성에 의하여 생성된 수직이등분선으로 형성되므로 항상 블록 다각형이며, 블록 다각형의 특성은 각의 개수가 많을수록 원에 근접한다. 최대내부원(IVcirde)의 면적이 다각형에서 차지하는 비율은 다각형의 각의 개수에 정비례한다.

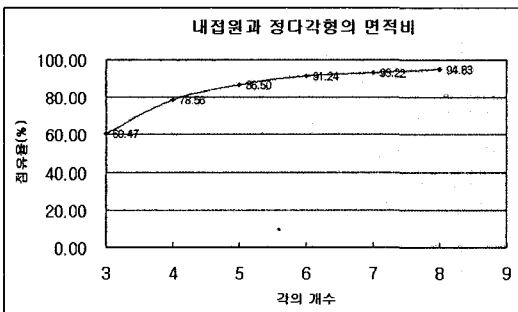


그림 8 내접원과 정다각형의 면적의 관계

3.2.2 IVcirde을 이용한 연속적인 스카이라인 질의 처리 과정

다각형의 최대내부원을 이용한 이동 객체의 연속적인 스카이라인 질의 처리 과정은 다음과 같이 5단계로 나눈다.

- [1단계] 이동 객체는 처음 질의 요청 시 위치좌표 (x_1, y_1) 를 서버에 전송
- [2단계] 서버는 이동 객체의 좌표값을 측정하여 속한 최적화된 스카이라인 영역을 판단
- [3단계] 해당 영역 정보에 의하여 이동 객체의 좌표값과 영역 변까지의 가장 가까운 거리 R 을 계산
- [4단계] 선 처리된 스카이라인 결과 집합과 최대내부원의 중심좌표 (x, y) , 내부원의 반경 R , 질의 위치에서 가장 가까운 변까지의 거리 R 을 이동 객체에 전송
- [5단계] 이동 객체에서는 일정 시간 간격으로 처음 질의의 위치에서 이동한 거리 S_1 를 계산 하여
 If $S_1 < R$
 처음 질의의 결과는 유효.

Else

현재 위치에서 내심 (x, y) 까지 거리 L_i 를 계산.

If $L_i < R$

처음 질의의 결과는 유효.

Else

영역 이탈로 판단하고 질의를 재요청.

IVcirde을 이용한 이동 객체의 연속적인 스카이라인 질의는 다음과 같다. 그림 9-step1에서처럼 이동 객체가 화살표 방향으로 OSR₇영역으로 진입한다. 이동 객체는 Q₁ 지점에서 위치좌표 (x_1, y_1) 를 서버에 전송하면서 처음으로 질의를 요청한다. 서버는 이동 객체가 위치한 영역 OSR₇을 판단하고 해당 영역 정보에 의해 Q₁에서 가장

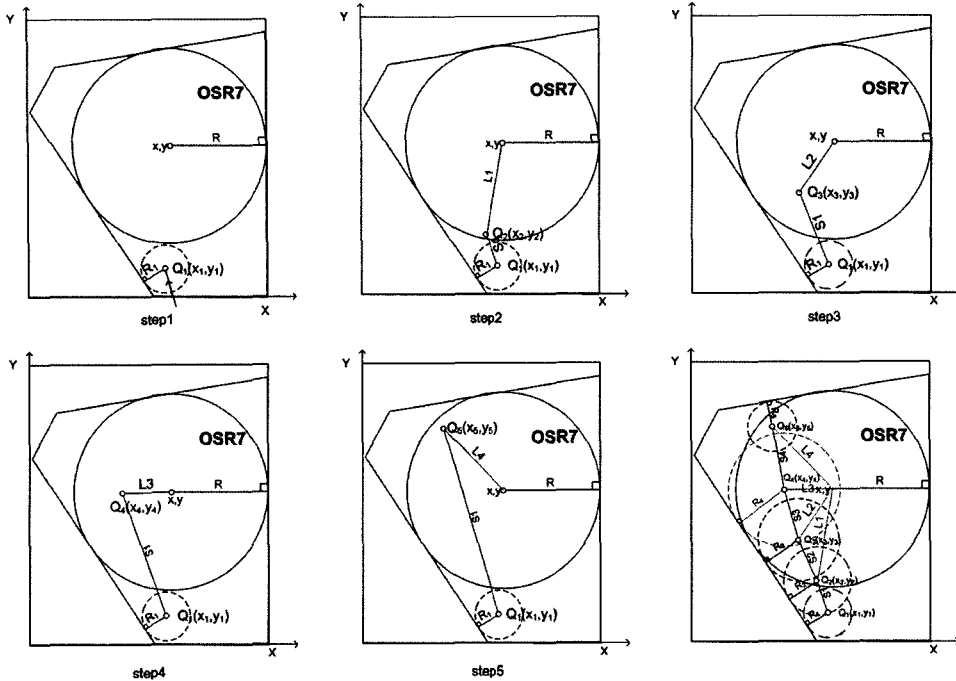


그림 9 최적화된 스카이라인 영역의 최대내부원

가까운 영역 변까지의 거리 R_1 과 선 처리된 스카이라인 결과 집합, IV_{airde} 의 중심좌표 (x, y) , IV_{airde} 의 반경 R 을 이동 객체에 전송한다.

Q_2 로 이동한(그림 9-step2) 이동 객체는 S_1 과 R_1 을 비교한다. $S_1 > R_1$ 이므로 Q_2 에서 내심까지의 거리 L_1 을 계산하여 IV_{airde} 의 반경 R 과 비교한다. $L_1 < R$ 이므로 IV_{airde} 을 이탈하지 않았기에 기존 질의의 결과는 계속 유효하다. 즉 재질이가 발생하지 않는다. 마찬가지로 Q_3 로 이동한(그림 9-step3) 이동 객체는 S_1 과 R_1 을 비교하여 $S_1 > R_1$ 이므로 Q_3 에서 내심까지의 거리 L_2 를 계산한다. $L_2 < R$ 이므로 재질이가 발생하지 않고 기존 질의의 결과는 계속 유효하다. 나머지 경우에도 마찬가지로 이동 객체가 이동한 거리가 $MINDIST$ 보다는 크지만 최대내부원(IV_{airde})을 이탈하지 않았으므로 기존 질의의 결과는 계속 유효하다.

그림 9에서처럼 이동 객체가 대상 객체 H7의 최적화된 스카이라인 영역 OSR_7 을 Q_1 에서 Q_5 의 궤적으로 관통할 때 V_{airde} 을 이용한 유효 영역 결정 기법은 5회의 영역 내 질의가 발생하지만 IV_{airde} 을 유효 영역을 결정하였을 때는 1회의 질의만 발생하였다. 이동 객체의 위치에 따라 동적으로 유효 영역이 결정되는 것이 아니라 블록 다각형의 특성을 이용한 최대내부원을 정적 유효

영역으로 결정한다. 유효 영역의 면적을 최대한 크게 하여 최적화된 스카이라인 영역(OSR)을 효과적으로 활용하여 영역 내에서의 불필요한 재질의 빈도를 감소할 수 있다.

3.3 블록 다각형의 최대내부원 추출 방법

블록 다각형의 최대내부원은 다각형의 최소 세 개 변에 내접하는 가장 큰 원이다. 내접원의 정의로 각의 이등분선이 교차되는 점이 내접원의 중심이 되고 내접원의 중심에서 변까지의 거리가 내접원의 반경이다.

블록 다각형의 최대내부원은 삼각형의 내접원을 찾는 방법을 이용하여 찾을 수 있다. 그림 10(a),(b),(c),(d)는 블록 다각형의 최대내부원을 추출하는 과정이다.

- 1단계: 다각형의 변을 이용한 삼각형 그리기.
- 2단계: 삼각형의 내접원 찾기.
- 3단계: 연립부등식을 이용하여 영역 밖의 내접원 제거.
- 4단계: 영역 내의 최대내부원 추출.

3.3.1 다각형의 변을 이용한 삼각형 그리기

다각형의 변(a,b,c,d)을 연장하여 임의의 3개변으로 삼각형(P1, P2, P3, P4)을 만든다. 블록 다각형의 각의 수가 3이면 형성되는 삼각형은 하나이며, 각의 수가 4이면 형성되는 삼각형은 4개이다. 블록 다각형의 각의 수가 n 일 때 형성되는 삼각형의 개수는 nC_3 으로 표시할 수 있다.

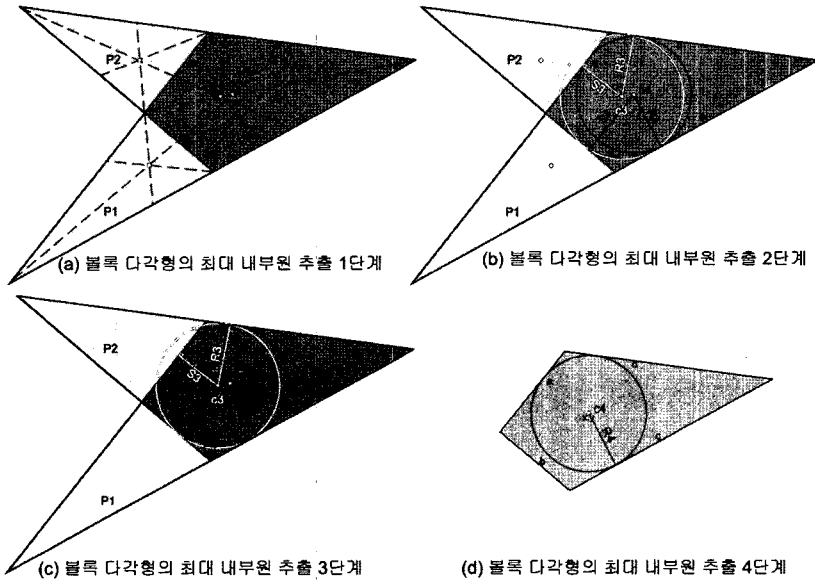


그림 10 볼록 다각형의 최대내부원 추출 순서

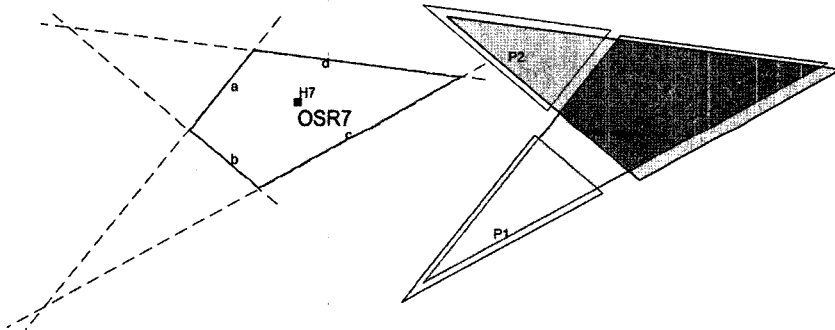


그림 11 다각형의 변으로 구성된 삼각형

그림 11에서처럼 대상 객체 H7의 최적화된 스카이라인 영역(OSR_r)의 변(a,b,c,d)의 연장선으로 네 개의 삼각형 P1, P2, P3, P4를 그릴 수 있다.

3.3.2 삼각형의 내접원 찾기

3.3.1에서 그린 모든 삼각형들에 대해 각의 이등분선을 그려 삼각형의 내심을 계산하고 내심에서 변까지의 거리를 반경으로 한 내접원을 그린다. 삼각형의 내심은 세 각의 이등분선의 교차점으로 찾을 수 있다.

최적화된 스카이라인 영역(OSR)을 결정할 때 영역을 구성하는 선분의 기울기와 꼭지점의 좌표들은 이미 구해져 있으므로 삼각형의 각의 이등분선은 각을 형성하는 두 선분의 기울기를 이용하여 다음과 같은 방법으로 찾을 수 있다.

그림 12에서 선분 L₁, L₃의 기울기 a₁, a₃는 선분이 지나는 두 개의 꼭지점을 이용하여 최적화된 스카이라

인 영역(OSR)을 결정할 때 이미 구해진 수치이다. 또한 선분 L₁은 직선의 방정식 $y = a_1 \cdot x + c_1$ 으로 선분 L₃는 $y = a_3 \cdot x + c_3$ 으로 표시할 수 있다.

선분 L₁과 L₃로 형성된 각의 이등분선 선분 L₂의 기울기 a₂는 $a_2 = \tan(\frac{\alpha + \beta}{2})$ 로 표시할 수 있다. 여기에서 α와 β는 각각 L₁과 L₃이 X축과 형성된 각이다.

$a_2 = \tan(\frac{\alpha + \beta}{2})$ 를 a₁과 a₃으로 표시하면

$$\tan(\frac{\alpha + \beta}{2}) = \sqrt{\frac{1 - \cos(\alpha + \beta)}{1 + \cos(\alpha + \beta)}}$$

$$\tan(\frac{\alpha + \beta}{2}) = \sqrt{\frac{1 - (\cos\alpha \cdot \sin\beta - \cos\beta \cdot \sin\alpha)}{1 + (\cos\alpha \cdot \sin\beta - \cos\beta \cdot \sin\alpha)}}$$

여기서 $\cos\alpha = \frac{S}{L_1}, \cos\beta = \frac{S}{L_3}, \sin\alpha = \frac{h_1}{L_1}, \sin\beta = \frac{h_3}{L_3}$

를 대입하면,

$$\tan\left(\frac{\alpha+\beta}{2}\right) = \sqrt{\frac{1 - \left(\frac{S}{L_1} \cdot \frac{h_3}{L_3} - \frac{S}{L_3} \cdot \frac{h_1}{L_1}\right)}{1 + \left(\frac{S}{L_1} \cdot \frac{h_3}{L_3} - \frac{S}{L_3} \cdot \frac{h_1}{L_1}\right)}}$$

$$\tan\left(\frac{\alpha+\beta}{2}\right) = \sqrt{\frac{L_1 L_3 - S \cdot h_3 + S \cdot h_1}{L_1 L_3 + S \cdot h_3 - S \cdot h_1}}$$

$$= \sqrt{\frac{\frac{L_1}{S} \cdot \frac{L_3}{S} - \frac{h_3}{S} + \frac{h_1}{S}}{\frac{L_1}{S} \cdot \frac{L_3}{S} + \frac{h_3}{S} - \frac{h_1}{S}}}$$

여기서 $L_1^2 = S^2 + h_1^2$, $L_3^2 = S^2 + h_3^2$ 이므로,

$$\tan\left(\frac{\alpha+\beta}{2}\right) = \sqrt{\frac{\sqrt{1 + \frac{h_1^2}{S^2}} \cdot \sqrt{1 + \frac{h_3^2}{S^2}} - \frac{h_3}{S} + \frac{h_1}{S}}{\sqrt{1 + \frac{h_1^2}{S^2}} \cdot \sqrt{1 + \frac{h_3^2}{S^2}} + \frac{h_3}{S} - \frac{h_1}{S}}}$$

여기에 $a_1 = \frac{h_1}{S}$, $a_3 = \frac{h_3}{S}$ 를 대입하면,

$$\tan\left(\frac{\alpha+\beta}{2}\right) = \sqrt{\frac{\sqrt{a_1^2+1} \cdot \sqrt{a_3^2+1} - 1 + a_1 \cdot a_3}{\sqrt{a_1^2+1} \cdot \sqrt{a_3^2+1} + 1 - a_1 \cdot a_3}}$$

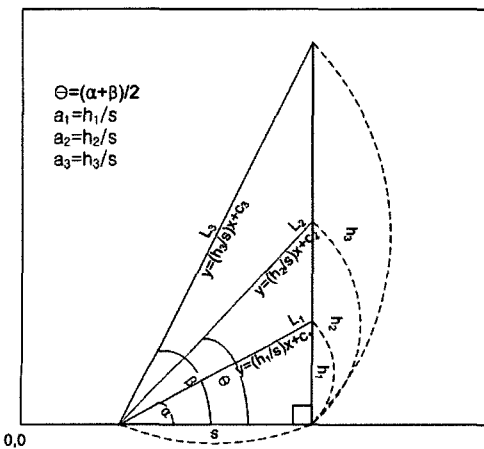
$a_1 \neq a_3$ 일 때

$$a_2 = \sqrt{\frac{\sqrt{a_1^2+1} \cdot \sqrt{a_3^2+1} - 1 + a_1 \cdot a_3}{\sqrt{a_1^2+1} \cdot \sqrt{a_3^2+1} + 1 - a_1 \cdot a_3}}$$

y 축의 절편 c_2 는

$$c_2 = \frac{a_1(c_3 - c_1) + c_1(a_1 - a_3) - a_2(c_3 - c_1)}{(a_1 - a_3)}$$

($a_1 \neq a_3$)이다.



0.0

그림 12 삼각형의 각의 이등분선

그러므로 선분 L_1 과 L_3 로 형성된 각의 이등분선 L_2 는 선분 L_1 과 L_3 의 기울기 a_1 , a_3 과 y축의 절편 c_1 , c_3 으로 다음과 같이 표현할 수 있다.

$$y = \sqrt{\frac{\sqrt{a_1^2+1} \cdot \sqrt{a_3^2+1} - 1 + a_1 \cdot a_3}{\sqrt{a_1^2+1} \cdot \sqrt{a_3^2+1} + 1 - a_1 \cdot a_3}} \cdot x + \frac{a_1(c_3 - c_1) + c_1(a_1 - a_3) - a_2(c_3 - c_1)}{(a_1 - a_3)}$$

3.3.3 연립부등식을 이용한 영역 밖의 내접원 제거

좌표 평면상에 최적화된 스카이라인 영역(OSR)을 형성하는 선분들은 직선의 방정식 $b \cdot y = a \cdot x + c$ 로 나타낼 수 있으며 선분에 의하여 만들어지는 영역은 부등식으로 나타낼 수 있다. 직선의 윗부분 영역의 부등식은 $b \cdot y > a \cdot x + c$ 이고 아랫부분 영역은 $b \cdot y < a \cdot x + c$ 이다.

3.3.2에서 계산한 삼각형의 내접원들 중에서 내심이 다각형의 밖에 있는 내접원은 제거한다. 그림 13에서 최적화된 스카이라인 영역(OSR_i)을 결정할 때 계산된 선분 a, b, c, d의 연립부등식을 이용하여 삼각형 P1, P2의 내접원 c1, c2의 내심은 연립 부등식 검사에서 제거된다. c1의 내심(x1,y1)은 영역의 변 b 와의 부등식 검사에서 제거되고, 내접원 c2의 내심(x2,y2)은 영역의 변 a 와의 부등식 검사에서 제거된다.

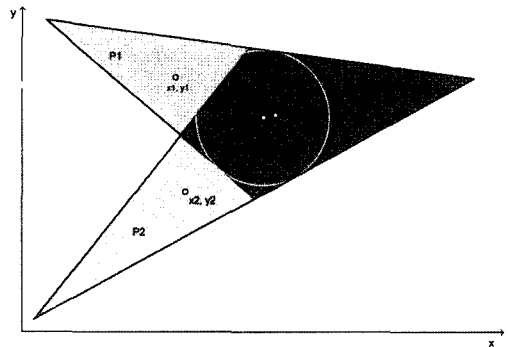


그림 13 삼각형의 내접원

3.3.4 영역 내의 최대내부원 추출

3.3.3에서 부등식을 만족하지 않는 내접원을 제거한 뒤 내접원의 반경이 내심에서 나머지 변까지의 거리보다 큰 내접원을 제거하고 남은 내접원들 중에서 반경이 가장 큰 내접원을 다각형의 최대내부원으로 선택한다.

그림 14에서 내접원 c3, c4는 3.3.3에서 영역 밖의 내접원을 제거한 뒤 남은 내접원들이다. c3은 선분 b, c, d로 형성된 삼각형 P3의 내접원이고 R3은 내접원의 반경이다. c3의 내심(x3,y3)에서 영역 OSR_i을 형성하는

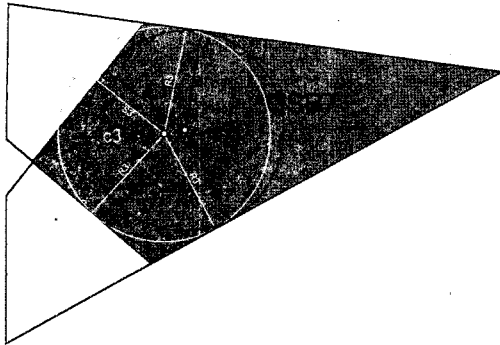


그림 14 조건을 만족하는 삼각형의 내접원

다각형의 나머지 변 a 까지의 거리(S_3)는 반경 R_3 보다 작기에 내접원 $c3$ 은 삼각형 $P3$ 에는 내접하지만 영역 OSR_3 을 형성하는 다각형의 내부원이 될 수 없으므로 제거한다. 마찬가지로 $c4$ 는 선분 a, c, d 로 구성된 삼각형 $P4$ 의 내접원이고 반경은 R_4 이다. 내접원 $c4$ 의 내심 (x_4, y_4) 에서 다각형의 나머지 변 b 까지의 거리 S_4 는 반경 R_3 보다 크므로 내접원 $c4$ 는 삼각형 $P4$ 에 내접하면서 영역 OSR_4 을 형성하는 다각형의 내부원이 될 수 있다. 내접원의 내심에서 다각형의 나머지 변까지의 거리는 선분의 두 꼭지점과 내심좌표를 이용하여 구할 수 있다.

1단계부터 4단계까지 거친 뒤 남은 삼각형의 내접원들 중에서 반경이 가장 큰 내접원이 볼록 다각형의 최대내부원으로 결정된다. 그림 14에서는 조건을 만족하는 내접원이 $c4$ 뿐이므로 최적화된 스카이라인 영역(OSR_4)을 형성하는 다각형의 최대내부원은 $c4$ 이며 중심좌표는 (x_4, y_4) 이고 반경은 $R4$ 이다.

4. 실험 및 성능 평가

이 논문에서 제안한, $IVcircle$ 을 최적화된 스카이라인 영역의 유효 영역으로 결정한 이동객체의 연속적인 스카이라인 질의 처리 기법의 성능을 평가하기 위해 대상 객체와 이동 객체의 모의 데이터를 생성하여 실험하였

다. 기존의 $Vcircle$ 을 유효 영역으로 결정한 기법과 비교하기 위하여 영역 내에서 발생된 재질의 횟수를 비교 측정하였으며 그 결과를 다각형별로 통계하였다. 대상 객체의 분포에 따른 성능 변화도 측정하기 위하여 대상 객체의 Uniform 분포뿐만 아니라 Gaussian 분포, Skew 분포에 대해서도 영역 내 재질의 횟수를 비교 측정하였다. 또한 대상 객체수의 변화에 따른 최적화된 스카이라인 영역을 형성하는 다각형별 분포에 대해 측정하였다.

4.1 실험 환경

$IVcircle$ 을 최적화된 스카이라인 영역의 정적 유효 영역으로 결정한 이동객체의 연속적인 스카이라인 질의 처리 기법은 Pentium-IV 1.73Ghz 프로세서와 1GB의 메인 메모리, 80GB의 하드 디스크를 가진 Windows XP 운영체제의 NoteBook PC에서 구현하였다. 이동 객체의 위치 좌표의 범위를 5,000m×5,000m로 가정하였고 2차원 영역에 균등(Uniform) 분포된 100, 200, 300개의 대상 객체와 5,000개의 균등 분포된 이동 객체를 생성하였다. 이동 객체의 이동 방향은 랜덤 함수를 이용하였다. 이동 객체의 분포에 따른 성능 비교 실험에서는 100개의 대상 객체를 Gaussian 분포, Skew분포된 영역에서 영역 내 재질의 횟수를 비교 측정하였다.

이 논문에서 제안한 $IVcircle$ 을 최적화된 스카이라인 영역의 정적 유효 영역으로 결정한 이동 객체의 연속적인 스카이라인 질의 처리 기법의 성능을 평가하기 위하여 기존의 $Vcircle$ 을 유효 영역으로 결정한 기법과 비교 실험하였다. 기존에 연구된 이동 객체의 연속적인 스카이라인 질의 처리 기법은 질의 위치에서 영역 변까지 가장 가까운 거리(MINDIST)를 반경으로 하는 동적 크기의 원($Vcircle$)을 유효 영역으로 결정하였다. 이동 객체의 위치에 따라 원($Vcircle$)의 크기가 결정됨으로 최적화된 스카이라인 영역을 효율적으로 활용하지 못하여 영역 내에서의 빈번한 재질이 발생하였다. 이 논문에서 제안하는 $IVcircle$ 을 유효 영역으로 이용한 영역 내 질의 발생 감소율을 측정하기 위하여 5,000개의 랜덤 좌표를 가진 이동 객체를 임의의 방향으로 영역을 이탈할 때까지의 질의 횟수를 비교 측정하였다. 표 1은 실험에

표 1 실험 파라미터

실험 파라미터	설명	값
UserNumber	이동 객체의 수	5,000
ObjectNumber	대상 객체의 수	100, 200, 300
UpperBound_x	공간 좌표의 x좌표의 최대 값	5,000
UpperBound_y	공간 좌표의 y좌표의 최대 값	5,000
LowerBound_x	공간 좌표의 x좌표의 최소 값	0
LowerBound_y	공간 좌표의 y좌표의 최소 값	0
ObjectDistribution	대상 객체의 분포 형태	Uniform, Gaussian, Skew

사용된 파라미터이다.

4.2 다각형별 영역 내 질의 횟수의 비교 측정

이 실험에서는 최적화된 스카이라인 영역(OSR)의 최대내부원(IVcircle)을 이용한 정적 유효 영역의 효율성을 평가하기 위하여 5,000m×5,000m의 영역에 균등 분포된 100개의 객체를 대상으로 실험하였다.

표 2는 랜덤 좌표를 가진 5,000개의 이동 객체가 처음 질의 위치에서 영역을 이탈할 때까지의 평균 질의 발생 횟수를 Vcircle을 유효 영역으로 고려했을 때와 IVcircle을 유효 영역으로 결정했을 때의 결과를 다각형별로 비교하여 보여준다. 실험 결과, Vcircle을 유효 영역으로 고려했을 때보다 평균 52.55% 감소하였다.

그림 15는 표 2의 수치를 그래프로 표시한 결과이다. IVcircle을 유효 영역으로 결정했을 때, 이동 객체가 위치한 최적화된 스카이라인 영역이 4각형일 때 영역 이탈까지 평균 3.43번의 질의가 발생하였고 8각형일 때 평균 1.00번의 질의가 발생하였다.

그림 16은 다각형별 영역 내 평균 질의 감소율을 보여준다. 실험 결과, 연속적인 스카이라인 질의에서 Vcircle을 유효 영역으로 고려한 경우보다 IVcircle을 유효 영역으로 이용한 경우가 질의 발생 빈도는 최소 36.76%에서 최대 66.52%까지 감소하였다.

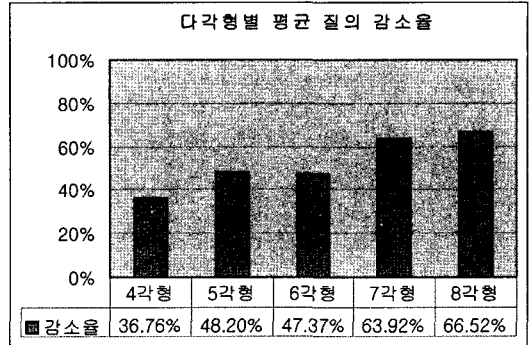


그림 16 다각형별 영역 내 평균 질의 감소율(대상 객체 Uniform 분포)

예상대로, 이동 객체가 위치한 최적화된 스카이라인 영역이 4각형일 때보다 8각형일 때 더 높은 감소율을 보여 다각형의 각의 수가 많을수록 더욱 효율적인 것을 알 수 있었다. 이것은 영역을 형성하는 다각형의 각의 개수가 많을수록 원에 근접하므로 IVcircle의 면적이 전체 다각형에서 차지하는 비율이 높아 더 좋은 성능을 보여 주는 것으로 분석된다.

실험결과, IVcircle을 이용한 최적화된 스카이라인 영역의 정적 유효 영역 결정 기법으로 이동 객체에 대한 연속적인 스카이라인 질의를 처리함으로써 이동 객체의 영역 내에서 발생하는 재질의 횟수를 감소시켜 네트워크 부하와 서버의 재계산 비용을 감소시킬 수 있음을 알 수 있었다.

4.3 대상 객체의 분포에 따른 질의 횟수 비교 측정

이 실험에서는 실제 환경과 최대한 비슷한 상황에서 평가하기 위하여 대상 객체를 Gaussian분포와 Skew분포로 나누어 성능을 비교해 보았다. 그림 17(a),(b)는 각각 서로 다른 분포를 가진 대상객체에 대한 실험 결과이다. 예상했던 대로 Vcircle을 유효 영역으로 고려한 경우는 다각형의 면적의 크기에 따라 질의 횟수가 증감한다. 하지만 IVcircle을 유효 영역으로 고려한 경우는 다각형의 면적보다 다각형의 각의 개수, 즉 최대내부원이 다각형에서 차지하는 비율에 따라 질의 횟수가 증감한다.

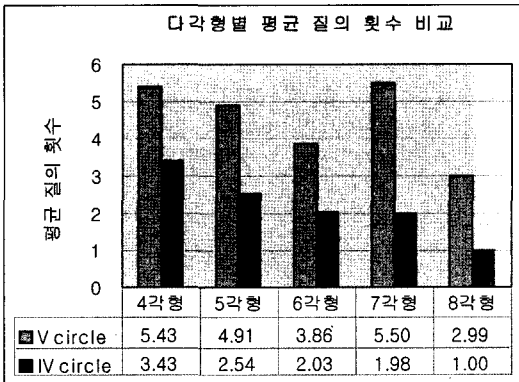
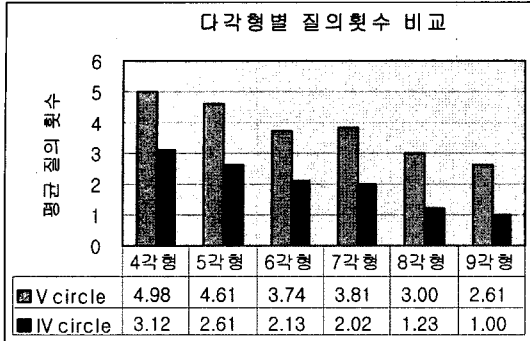


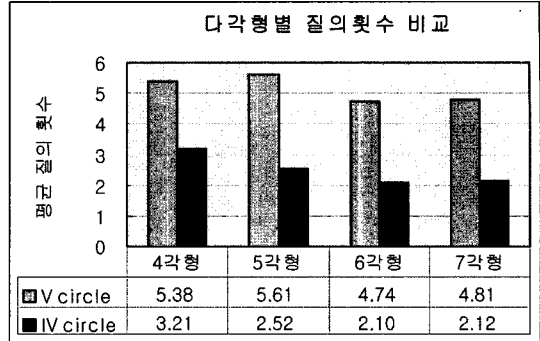
그림 15 다각형별 영역 내 평균 질의 발생 빈도(대상 객체 Uniform분포)

표 2 다각형별 평균 질의 횟수의 비교 측정(대상 객체 Uniform 분포)

다각형	다각형의 출현 빈도	평균 질의 횟수		평균 감소율
		Vcircle	IVcircle	
4각형	1,159	5.43	3.43	36.76%
5각형	1,994	4.91	2.54	48.20%
6각형	1,208	3.86	2.03	47.37%
7각형	562	5.50	1.98	63.92%
8각형	77	2.99	1.00	66.52%
합계	5,000	4.80	2.52	52.55%



(a) 대상객체 Gaussian분포



(b) 대상객체 Skew분포

그림 17 대상 객체의 분포에 따른 영역 내 질의 횟수 비교

4.4 대상 객체 수에 따른 수치 측정

이 실험에서는 제안한 기법의 성능을 평가하기 위하여 균등 분포된 대상객체를 100, 200, 300 개로 증가하면서 나타나는 다각형의 분포를 측정해 보았다.

그림 18은 대상 객체가 증가함에 따라 다각형별 출현 빈도를 표시한다. 그림에서처럼 최적화된 스카이라인 영역을 형성하는 블록 다각형들 중에서 4각형, 5각형, 6각형이 대상객체가 100개일 때 86%, 200개일 때 84%, 300개일 때 80%를 이루고 있음을 알 수 있었다.

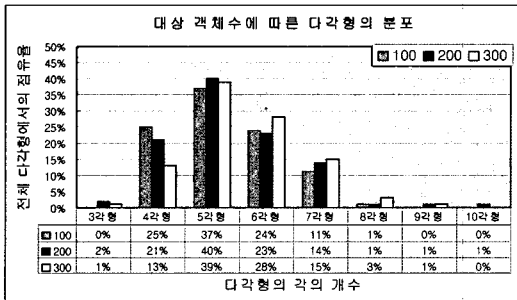


그림 18 대상 객체 수에 따른 다각형별 점유율

5. 결론

이동객체에 대한 스카이라인 질의는 다양한 위치 기반 서비스를 제공할 수 있다. 기존의 이동객체의 스카이라인 질의 처리에 대한 연구는 스카이라인 질의 결과 값이 변하지 않는 영역을 미리 계산하고 이동객체가 질의를 요청할 때 질의 위치에서 영역 변까지의 거리(Vcircle)를 반경으로 한 동적 영역을 유효 영역으로 결정하여 이동객체에 전송한다. 이동객체는 '이동할 때마다 유효 영역 이탈 여부를 판단하여 질의 여부를 결정하였다.

이 논문에서는 모든 대상 객체에 대하여 스카이라인 결과 값이 불변의 최적화된 스카이라인 영역(OSR)을

미리 계산하고, 이동 객체의 위치와 상관없는 영역의 최대내부원을 정적 유효 영역으로 결정하는 기법을 제안하였다. 즉, 최적화된 스카이라인 영역(OSR)을 형성하는 블록 다각형의 특성을 이용하여 유효 영역의 면적을 최대화하여 영역 내에서의 불필요한 질의 횟수를 줄일 수 있었다. 실험을 통해 비교 분석한 결과 영역내의 평균 질의 발생 빈도를 기존의 Vcircle을 이용한 동적 유효 영역 결정 기법보다 평균 52.55% 감소하였다.

참고 문헌

- [1] 나경석, 김진호, 박영배, "연속적인 스카이라인 질의를 위한 효율적인 영역 결정기법", 한국정보과학회 데이터베이스연구회 KDBC 2006 학술대회 우수논문상, 2006. 5.
- [2] Zheng B., Lee, D. "Semantic Caching in Location - Dependent Query Processing," SSTD, p.97-116, 2001.
- [3] Song, Z., Roussopoulos, N. "K-Nearest Neighbor Search for Moving Query Point," SSTD, 2001.
- [4] Brzsonyi, S, Kossmann, D., Stocker, K. "The Skyline Operator," In ICDE, pp.421-430, 2001.
- [5] Tan, K., Eng, P. Ooi, B. "Efficient Progressive Skyline Computation," In VLDB, pp.301-310, 2001.
- [6] D. Kossmann, F. Ramsak, S. Rost, "Shooting Stars in the Sky: an Onlie Algorithm for Skyline Queries," In VLDB, pp.275-286, 2002.
- [7] Benetis, R., Jensen, C., Karcauskas, G., Saltenis, S. "Nearest Neighbor and Reverse Nearest Neighbor Queries for Moving Objects," IDEAS, 2002.
- [8] Tao, Y., Papadias, D., Shen, Q. "Continuous Nearest Neighbor Search," VLDB, 2002.
- [9] J. Zhang, M. Zhu, D. Papadias, Y. Tao, and D. L. Lee. "Location-based spatial queries," In SIGMOD, pp.443-454, 2003.
- [10] Papadias, D., Tao, Y., Fu, G., Seeger, B. "An Optimal and Progressive Algorithm for Skyline Queries," In SIGMOD, pp.443-454, 2003.



이 중 혁

1998년 중국 대련경공업학원 화학공정학과(학사). 2006년 명지대학교 컴퓨터공학과(석사). 2006년~현재 명지대학교 컴퓨터공학과 데이터베이스 연구실(박사과정) 관심분야는 Mobile DB, 위치기반 서비스



박 영 배

1993년 서울대학교 컴퓨터공학과(박사) 1990년~1992년 명지대학교 전자계산소장. 1997년~2001년 명지대학교 산업대학원장. 1981년~현재 명지대학교 컴퓨터공학과 교수. 관심분야는 Mobile DB, Spatial DB, 한국어 정보처리, Large

Fingerprint DB