

曖昧な位置に基づく空間問合せ処理手法の実現

飯島 裕一[†] 石川 佳治[‡]名古屋大学 工学部 電気電子・情報工学科 情報工学コース[†] 名古屋大学 情報連携基盤センター[‡]

1 はじめに

近年、移動ロボットやセンサ環境などの分野において、曖昧な位置情報に基づくデータベース問合せ処理技術の必要性が高まってきている。現実環境中を動き回る移動ロボットにとって、自身の位置の推定は円滑なサービス提供を行う上で欠かせないものであるが、センサの分解能やモータの制御ノイズなどのために、正確な位置の推定は容易ではなく、誤差を伴った推定となる。また、多くのセンサが各々の周辺の環境情報を収集するセンサ環境では、各センサの位置を把握する手段として GPS を利用する方法が考えられるが、電波状況によっては期待通りの測位精度を得られないという問題がある。加えて、GPS による位置取得は多くの電力を消費するため、各センサが電池で駆動しているような場合には極力避けたいという要求もある。以上のように、移動ロボットの位置やセンサ環境における各センサの位置は曖昧であるため、そのことを踏まえた問合せ処理手法が必要である。

このような背景から、本研究グループはオブジェクトの位置が曖昧な位置情報で表現されているような場合における、空間データベース問合せの処理手法について提案している [1]。この提案手法では、曖昧な位置情報で表現されているオブジェクトが、その周辺に位置するオブジェクトを検索するために空間問合せを発行するという状況について考えており、特に、問合せオブジェクトの位置が正規分布で表現され、問合せ対象のオブジェクトは確定的な位置で表される点データである場合を想定している。対象とする問合せは、ユークリッド距離に基づく通常の空間問合せを拡張したものである。本稿では、提案手法の問合せ処理アルゴリズムを実装して評価実験を行った結果を報告する。

2 確率的範囲問合せ

提案手法では、問合せオブジェクトの位置の曖昧性に対応するために、空間データベースにおける従来の範囲問合せの概念を拡張した確率的範囲問合せ (probabilistic range query, PRQ) の概念を以下のように定義している。

d 次元空間において、問合せオブジェクト q の位置が d 次元ベクトルの座標値 \mathbf{x} を持つ確率が、 d 次元の正規分布により

$$p_q(\mathbf{x}) = \mathcal{N}(\mathbf{x}, \Sigma) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{x} - \mathbf{q})^t \Sigma^{-1} (\mathbf{x} - \mathbf{q}) \right] \quad (1)$$

Implementation of Processing Spatial Query for Objects with Imprecise Location

[†] Yuichi Iijima, Department of Information Engineering, School of Engineering, Nagoya University

[‡] Yoshiharu Ishikawa, Information Technology Center, Nagoya University

で表されるとする。ただし、 Σ は $d \times d$ の共分散行列であり、 $|\Sigma|$ は Σ の行列式を表す。 q は確率分布の平均の位置に対応している。このような q が与えられたとき、 q とのユークリッド距離が δ 以下である確率が θ 以上であるようなオブジェクトの集合を $PRQ(q, \delta, \theta)$ で表す。

3 確率的範囲問合せ処理のアルゴリズム

提案手法による確率的範囲問合せ処理のアルゴリズムについて、その概要を紹介する。詳細については [1] を参照されたい。

3.1 問合せ処理の基本戦略

基本的なアイデアとしては、問合せオブジェクトと問合せ対象オブジェクトの役割を逆転させて考える。つまり、「問合せオブジェクト q から距離 δ 以内に問合せ対象オブジェクトが位置する確率」を考えるのではなく、「各問合せ対象オブジェクトに対し、そのオブジェクトから距離 δ 以内に問合せオブジェクトが位置する確率」を評価し、それが θ 以上となるオブジェクトを結果集合に加えるというアプローチをとる。しかし、そのような確率の計算は数値積分を伴うため、すべての問合せ対象オブジェクトに対して確率を計算するとコストが高くなってしまう。そのため、提案手法では問合せ処理について以下の 3 ステップからなる基本戦略を採用している。

1. q との距離が δ 以下になる確率が θ 以上となる可能性のあるオブジェクトが存在しうる検索領域を求める。
2. 検索領域内のすべてのオブジェクトを索引を用いて検索する。
3. 検索された各オブジェクトに対し、 q との距離が δ 以下である確率が θ 以上であるかどうかを明らかにできない場合には、数値積分により確率を求め、問合せ結果集合に含めるか否かを決定する。

さらに、この基本戦略を実現する具体的な問合せ戦略として 2 つの戦略を導入している。

3.2 問合せ戦略 1

問合せ戦略 1 では、 θ 領域という、その領域の外に問合せオブジェクト q が存在する確率が $1 - 2\theta$ (ただし、 $0 < \theta < 1/2$ とする) であるような領域を考え、それを基に検索領域を求めるという戦略をとる。

θ 領域は、図 1 の灰色部分で示されるような d 次元楕円体の形状を持つが、楕円体の形状では R 木などの索引構造を利用して適切に検索を行うことができないため、 θ 領域に外接する矩形領域を考えることにする。その矩形領域から距離 δ 以内の領域の外に位置するオブジェクトについては、解となる可能性がないという性質があるので、図 1 の実線で示されるような検索領

域を導出することができ、数値積分を必要とするオブジェクトの個数を減らすことができる。

3.3 問合せ戦略 2

問合せ戦略 2 では、 $p_q(x)$ の上限の関数と下限の関数を考え、それらを基にして、検索領域と、検索領域内部の領域であって、その領域内に位置するオブジェクトは無条件で解であると判断できるような領域の 2 つの領域を求めるといった戦略をとる。

任意の x について

$$p_q^-(x) \leq p_q(x) \leq p_q^+(x) \quad (2)$$

が成り立つような $p_q^-(x)$ と $p_q^+(x)$ を考える。すると、 $p_q(x)$ の上限を与える関数である $p_q^+(x)$ によって、 q から α^+ 以上離れたオブジェクトは解となる可能性がないと判断できるような α^+ を、また、 $p_q(x)$ の下限を与える関数である $p_q^-(x)$ によって、 q から α^- 以下の距離にあるオブジェクトは無条件で解であると判断できるような α^- を求めることができる。図 2 に示すように、まず、 q からの距離が α^+ 以内のオブジェクトを検索し、さらに、検索されたオブジェクトのうちで q からの距離が α^- 以内のオブジェクトを問合せ結果集合に含めることにより、数値積分を必要とするオブジェクトの個数を減らすことができる。

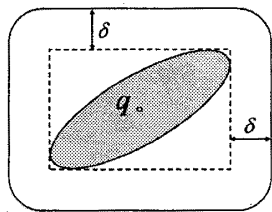


図 1: 問合せ戦略 1

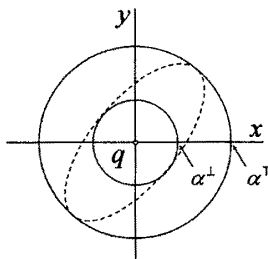


図 2: 問合せ戦略 2

4 評価実験

提案手法の問合せ処理アルゴリズムを 2 つの問合せ戦略のそれぞれについて実装して評価実験を行い、比較を行った。

4.1 実験方法

$[0, 1000] \times [0, 1000]$ に正規化された 2 次元の地図データ (50747 個の点データから成る) が格納されている R 木を使用して、問合せオブジェクト q から距離 50 以内にある確率が 1% 以上であるオブジェクトの集合を求める処理、すなわち、 $PRQ(q, \delta = 50, \theta = 0.01)$ を求める処理を、問合せ戦略 1 を用いて実装した場合と問合せ戦略 2 を用いて実装した場合のそれぞれについて行い、処理時間と数値積分を必要としたオブジェクト数を測定した。また、問合せ戦略 2 の場合については、積分を行うことなく解と判定できたオブジェクト数も測定した。以上の測定を q の位置をランダムに変えて 5 回行い、平均の測定値を求めた。なお、(1) 式の共分散行列 Σ のパラメータは以下のように設定した。

$$\Sigma = 100 \begin{bmatrix} 7 & 2\sqrt{3} \\ 2\sqrt{3} & 3 \end{bmatrix}$$

これにより、 $p_q(x)$ の等確率面の形状は 30° の傾きを持った、長軸と短軸の比が 3 : 1 の楕円となる。

4.2 実験結果と考察

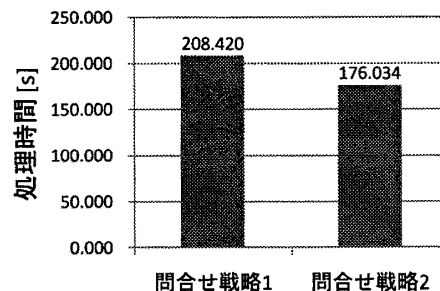


図 3: 処理時間の比較

処理時間の比較を図 3 に示す。問合せ戦略 2 を用いて実装した場合の方が、問合せ戦略 1 を用いて実装した場合よりも、15% ほど処理時間が短くなった。処理時間の内訳としては、数値積分計算の処理が全体の大部分を占めており、両戦略ともに処理時間の 98% 以上は数値積分計算のために費やされていた。数値積分を必要としたオブジェクト数が、問合せ戦略 1 の場合では 4414 個、問合せ戦略 2 の場合では 3731 個であったことから、今回の実験については、問合せ戦略 2 の場合の方が、数値積分を必要とするオブジェクトの個数を問合せ戦略 1 の場合よりも少なくすることができたため、より短時間で問合せ処理を行うことができたものと考えられる。問合せ戦略 2 の場合においては、積分を行うことなく解と判定できたオブジェクトが 533 個存在したので、そのことが数値積分を必要とするオブジェクトの個数を減らすことに大きく貢献したようである。最終的に解と判定されたオブジェクト数は、問合せ対象オブジェクトの総数 50747 個のうち 2898 個であった。問合せ戦略 1 と問合せ戦略 2 のどちらが有利かについては、問合せ対象オブジェクトの分布の仕方や、確率的範囲問合せのパラメータ q, δ, θ の値などによって変化するため、一概に結論付けることはできない。状況に応じて両者をうまく組み合わせることで、さらに効率的な問合せ処理を実現できるであろう。

5 まとめ

本稿では、問合せオブジェクトの位置が曖昧な位置情報で表現されているような場合における空間データベース問合せの処理手法について、本研究グループの提案している手法を、2 つの問合せ戦略のそれぞれについて実装し、評価実験を行った。現在、本研究グループでは提案手法の最近傍問合せへの拡張を進めており、その評価実験などが今後の課題である。

謝辞

本研究の一部は、文部科学省科学研究費 (19024037) の助成による。

参考文献

- 石川佳治, 飯島裕一. 曖昧な位置に基づく空間問合せ処理の効率化. 電子情報通信学会第 19 回データ工学ワークショップ (DEWS2008), 2008.