

## 6C-1

線分分割による  
線図形データの検索効率改善

松岡正晴 深海 悟  
NTTデータ通信株式会社

## 1. まえがき

近年、地図等の管理システムの需要は増加し、より高性能な管理システムへの要求が高まっている。計算機上で地図等を線分(ベクトル)として表すとき、線分を大量に扱う必要がある。その処理内容としては、空間的位置やその属性に基づく検索などが挙げられる。そこで、この大量な線データを効率良く管理することは、多様な検索を容易でかつ短時間で処理するのに役立ち、重要となる。本稿は、点データ管理の一手法であるBD木[1]を拡張した線データ管理[2]に関し、線分分割による検索効率の改善案とその評価を示す。ただし、ここで対象としたのは、非バケット型で、線に関するデータはすべて主記憶上にある場合を想定している。

## 2. BD木

## 2.1 点データ管理[1]

BD木は完全二分割法による点データ管理手法(図1(b))において、データの存在しない領域(図1(a)の斜線部分)に対する枝を取り除き、その位置関係を領域式で示すことにより、無駄なノードを削除したものである(図1(c))。これにより、検索効率とメモリ効率が向上する。

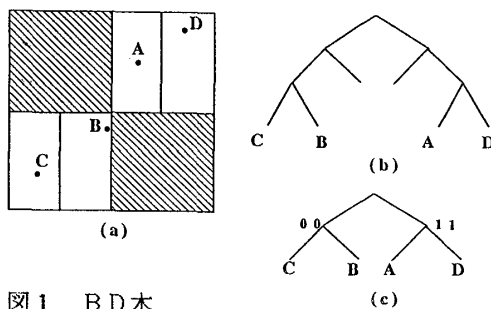


図1 BD木

## 2.2 線データ管理[2]

BD木を線図形データ管理に使用するとき、線分の位置をその中心点に集中して、この点によりBD木を構成する。BD木の各葉ノードには、線分の情報として、この線分を囲む最小の長方形(以下、外接長方形)の情報(左下の点座標と右上の点座標)を置く。また、葉以外のノードには左右の子ノードの外接長方形を

囲む最小の長方形の情報を置く。

例えば、範囲検索(指定された矩形の中を通るすべての線分検出)を行う場合、この外接長方形の情報を用いて、以下の手順を根ノードに対して行う。

- S1. 検索の指定範囲が外接長方形と重なったら、S2へ。そうでなければ、終了。
- S2. 葉ノードなら、指定範囲と線分との交差判定をする。そうでなければ、左右子ノードに対してS1を再帰的に繰り返す。

## 3. BD木の問題点

上述のBD木による線図形データ管理手法の問題点は、BD木の構成に際し、線分の中心点しか考慮していないことにある。点データ管理の時は検索キーと木構造生成キーの対応が取れていたが、線データ管理のときはその対応が取れていない。つまり、線分は長さを持ち、検索のときにこの長さを表す外接長方形が検索キーとなる。しかし、木構造生成にこの長さの情報を反映していない。

例えば、図2(a)で線分Bの外接長方形が他のノードAの示す外接長方形を覆う場合を考える。このとき、矩形Rの範囲検索を行うと、線分Bの指すノードへの無駄な走査をすることになる。とくに、線分AとBのデータ構造上の位置が離れていると、検索効率が下がる。そこで、この無駄な検索走査を減少させることを考える。

## 4. 線分分割による効率改善

線分の外接長方形が他のノードの外接長方形を覆うとき、線分と、ノードの持つ外接長方形の各辺の延長線との交差点の中で、線分の中心点に最も近い点で線分の分割を行う(図2(b))。この操作により、以下の2点の効率改善が見込まれる。

## (1) 検索ノード数の減少

上述の線分分割により、長い線分は分割され短くなる。そしてこの分割した線分を投入するため、各ノードの外接長方形が部分木間を跨ってオーバーラップする状態が減り、部分木上の無駄な走査が減少する。

## (2) 交差判定回数の減少

範囲検索では、線分と指定範囲の交差判定が必要となるが、これは外接長方形どうしの重なり判定より、

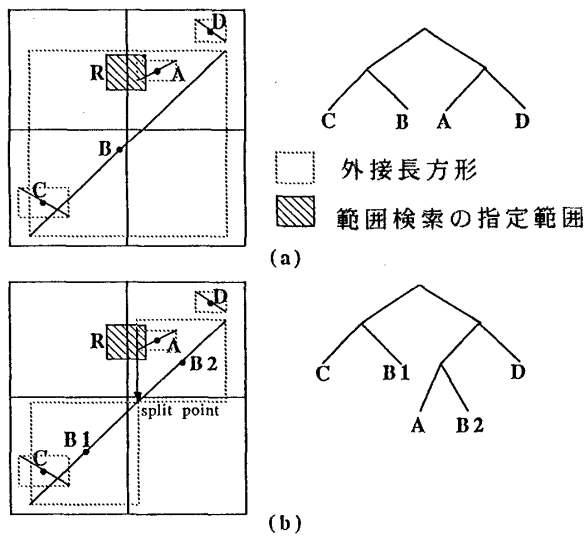


図2 線分分割操作

計算量が多い。この点を考えると、交差判定回数が減ることも効率改善になる。線分分割により、葉ノード上での外接長方形の総面積は小さくなり、葉ノードへの走査回数が減り、効率が上がる。

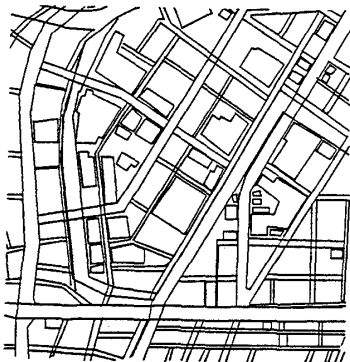


図3 実験用地図データ (データ3)

5. 評価実験

本実験は、3枚の地図データ(一例を図3に示す)を用いて行った。分割操作による各地図データのノードの増加率は、2%、7%、23%であった。各々のデータについて、指定範囲の各辺の長さを、空間全体の1%から20%まで変化させ、指定範囲を任意の位置に設定して、範囲検索を行った。この検索を300回繰り返し、1回の検索に要したノード数と交差判定回数の平均値を測定した。図4は上記(1)の分割前の検索に要したノード数に対する分割後の検索ノードの減少率を示している。また、図5は(2)の分割前の検索に要した交差判定回数に対する分割後の交差判定回数の減少率を示している。分割によりノード数は増えるが葉ノードの外接長方形の総面積が減ったことと、外接長方形の空間的に近いものどうしが木構造上でも近づくことにより、

検索効率が向上している。

ただし、分割点が範囲検索の指定範囲内にあると、線分と指定範囲の交差回数が増える。指定矩形が大きい場合はこの確率が高くなり効率が悪くなる。この場合、ノード増加の影響により、(1)の効率も上がらなくなってくる。

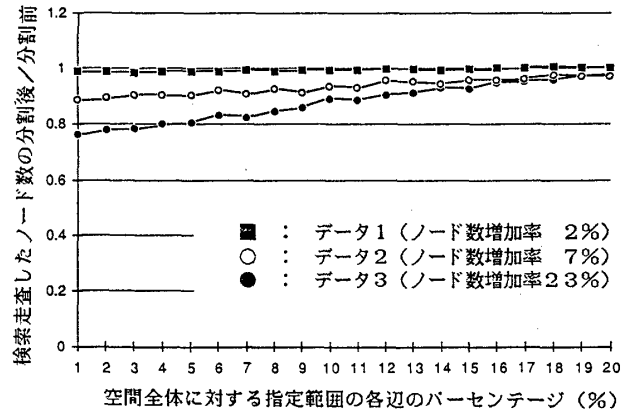


図4 検索走査ノード数の減少率

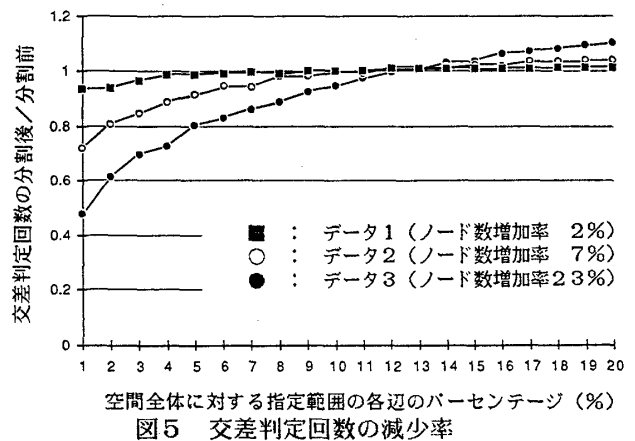


図5 交差判定回数の減少率

6. むすび

BD木を管理構造とした線図形データ管理の効率改善を試みた。分割が多く行われるデータで範囲検索の指定範囲が全空間に対して小さいとき、効率は向上している。ノード数の増加のため指定範囲が大きくなると範囲検索の効率が下がるが、実際の用途を考えると、図4、5の数%の指定範囲で使うことが多いと考えられるため、本方式は効果があると判断される。

参考文献

[1] 大沢、坂内: "良好な動特性をもつ次元点データ管理構造の一提案"、信学論(D)、J66-D、10、pp.1193-1200(昭58-10)  
 [2] 大沢、坂内: "空間的な位置関係に依存した検索に適した線情報の管理方式"、信学論(D)、J69-D、5、pp.724-732(昭61-5)