

多様な観測データや事前知識からの地物の時空間位置の再現

関本 義秀^{†*} 柴崎 亮介[†]

近年, ITS(Intelligent Transportation System)・マーケティング・防災等の分野で経路情報, 滞留情報など時空間データに対するニーズが増加している。すなわち, 時々刻々とダイナミックに変化する人間や車両あるいは建物などの地物に関する情報を集めることが必要とされている。一方, 様々なセンサー技術あるいはネットワーク技術の発展により, GPS, PHS, トラフィックカウンタ, 監視カメラなどの従来の技術に加え, シュードライト, 電子タグ, 情報家電の稼動状況など多様な時空間データの取得可能性が広がってきている。しかし, そのように地物に関して取得されたデータはあくまで断片的であり誤差も含むため, 簡単には全体像が分からない。したがって情報提供の機能に重点を置いた従来の GIS の枠組みではこれら多様な誤差を含むデータを処理するのは限界がある。そこで次世代的な GIS の一つの方向性として, 多様な時空間データを整理し, それらから誤差を包含しつつ妥当な地物の時空間的状态を推論する方法論が必要となるはずである。理論的には地物の観測データに対して尤度最大化基準で推定し, しかも地物に関する知識を事前確率という形で入れ込む尤度最大化基準に基づくベイズ推定法が妥当であるものの, 多様なデータから知りたい地物の状態に対して動的に最適化問題を構成し, 適当なパフォーマンスで解くことが自体が大きな問題である。そこで本研究では, 地物の時空間的状态を再構成を行う観点から多様な観測データを構造化して最適化問題を構成し, 構造化されたデータのタイプから効率的に解く方法を示す。

Reconstruction of spatio-temporal position of feature with various observation data and knowledge about feature

YOSHIHIDE SEKIMOTO^{†*} and RYOSUKE SHIBASAKI[†]

Recently, in the field of ITS(Intelligent Transportation System), marketing and disaster prevention, data with finer spatio-temporal resolution with respect to time, space and attribute will be needed for spatio-temporal database compared with the resolution of existing maps or statistical data. In other words, we need to collect information of features such as human, car or building changing dynamically. On the other hand, many kinds of spatio-temporal data are being repeatedly acquired or generated, such as pseudolite and electric tag, etc. in addition to existing GPS, PHS, traffic counter and monitoring camera, mainly because of the rapid progress of data acquisition technology and network technology in these days.

However, those data are still so fragmentary and may have some errors that we cannot always directly monitor dynamic changes of the spatial objects or features with sufficient spatio-temporal resolution. Therefore we need to develop a method for arranging various kinds of spatio-temporal data and inferring the most probable spatio-temporal state of feature.

While inference of the state from observation data, based on maximum likelihood criteria, is feasible theoretically, it is necessary to form optimization problem dynamically and solve this problem efficiently according to data type in the next stage. In this research, we show the method of structuring various spatio-temporal data, forming optimization problem and solving the problem efficiently, in terms of reconstruct spatio-temporal state of feature.

1. はじめに

近年, ITS(Intelligent Transportation System)・マーケティング・防災等の分野で経路情報, 滞留情報など時空間データに対するニーズが増加している(表 1)。すなわち, 時々刻々とダイナミックに変化する人間や車両あるいは建物などの地物に関する情報を集めることが必要とされている。したがって, 既存の地

[†] 東京大学空間情報科学研究センター
Center for Spatial Information Science, University of Tokyo

^{*} 現在, 国土交通省国土技術政策総合研究所
Presently with National Institute for Land and Infrastructure Management

図や統計データよりも時間、空間、属性の点から精度あるいは分解能の高い地物の時空間データベースを用意することが必要といえる。実際に GIS(Geographic Information System) はカーナビゲーションや施設管理のように地物の情報を提供することを主眼として進歩しており、今後より高度なサービス展開を期待されている。

一方、様々なセンサー技術あるいはネットワーク技術の発展により、GPS、PHS、トラフィックカウンタ、監視カメラなどの従来の技術に加え、シュードライト、電子タグ、情報家電の稼働状況など多様な時空間データの取得可能性が広がってきている(表2)。

しかし、そのように地物に関して取得されたデータはあくまで断片的であり誤差も含むため、数箇所の断面交通量データから全体的な交通流が分からないように、簡単には全体像が分からない。したがって情報提供の機能に重点を置いた従来の GIS の枠組みではこれら多様かつ誤差を含むデータを処理するのは限界がある。そこで次世代的な GIS の一つの方向性として、多様な時空間データを整理し、それらから誤差を包容しつつ妥当な地物の時空間的状态を推論する方法論が必要となるはずである。

理論的には地物の観測データに対して尤度最大化基準で推定し、しかも地物に関する知識を事前確率という形で入れ込む尤度最大化基準に基づくベイズ推定法が妥当であるもの^{7),13)}、多様なデータから知りたい地物の状態に対して動的に最適化問題を構成し、適当なパフォーマンスで解くこと自体が大きな問題である。そこで本研究では、地物の再構成を行う観点から多様な観測データを構造化して最適化問題を構成し、構造化されたデータのタイプから効率的に解く方法を示す。

このように多様で誤差を含む断片的なデータから地物の時空間位置を再現できることにより、人の流動や車の混雑具合などが再現しやすくなり、マーケティングにおける販売戦略、施設立地計画や ITS における交通施策あるいは防災における避難誘導、ブース配置などにもより高度なレベルで反映させることが可能になる。

表1 各分野における時空間データのニーズ
Table 1 Needs for spatio-temporal data in various fields

分野	データの種類
ITS	各車の位置情報、渋滞情報
マーケティング	各人の滞在情報
防災	分布情報、避難経路情報

表2 様々な時空間データの取得技術

Table 2 Many kinds of acquisition technology for spatio-temporal data

	非集計データ	集計データ
既存のデータ	GPS、PHS、ジャイロ+加速度計によるトラッキング、チケット予約データ、クレジットカード情報、携帯・PHS のログ情報など	自動改札データ、トラフィックカウンタデータ、監視カメラ、POS データ、パーソントリップ、国勢調査、イベントデータなど
将来利用可能性のあるデータ	シュードライトや電子タグによるトラッキングデータ、自動車の稼働状況データ(プローブカー)など	列車搭載重量データ、成層圏プラットフォームからの画像データ(人、車の流れ)、情報家電の稼働状況データなど

2. 本研究での手法

この章では、様々な時空間データから地物を再構成するアルゴリズムについて述べる。まず 2.1 では我々が提案してきた時空間データ概念モデルである FEO モデルをその適用例とともに簡単に説明し、2.2 ではそのモデルに基づく再構成のフレームワークについて詳細な説明をする。

2.1 FEO モデルを用いた時空間データの構造化

近年、GIS、CG などの分野では、従来の空間的表現に時間的要素を導入した地物の時空間表現に関する研究が進展してきており^{1),4),6),11),12),14),15)}、ISO や OGC 等における地理情報の標準化活動もまとまりつつある。しかし基本的にはどれも「正しいはず」のデータをどう表現するかに焦点があてられており、そもそもどのように時々刻々とダイナミックに変化する真の時空間的状态を求めるかについては言及されていない。

そのため、筆者らは地物だけでなく地物を観測するもの、すなわち地物のある部分を写し取るものと、地物に影響を与えるもの、すなわち地物のある部分に変化を与えるもの、といった外的要素を加えた FEO (Feature/Event/Observation) モデルを提案してきた^{5),9)}(図1)。ここでは観測モデルの内容とその適用例として表3でGPSによる測位と自動改札による人数把握について記述した。

開始・終了についてはGPSでは「測位開始時刻・終了時刻」とし、自動改札では「駅の開始時刻・閉鎖時刻」とした。変化等の自身のふるまいについてはこの場合は特になく、観測モデル固有な観測メソッドでは対象時空間範囲、対象地物のタイプ、関連タイプ、属性タイプ、また観測手順、観測データ、観測誤差など

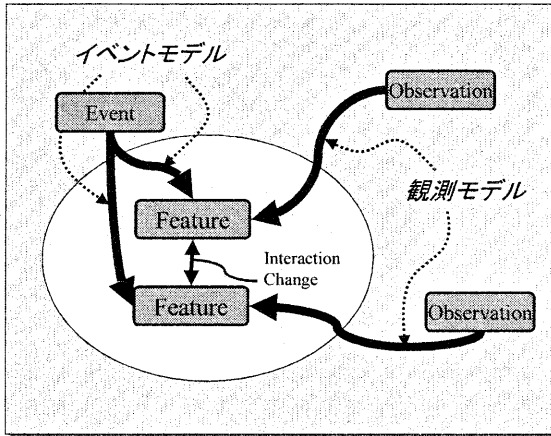


図1 FEO(Feature/Event/Observation) Model
Fig.1 FEO(Feature/Event/Observation) Model

を記述し、この例では、「電波受信可能空間で任意の時間」あるいは「改札間を1時間の間」などを時空間範囲とする。また、GPSは特定のある人を対象としているのに対し、自動改札の場合は不特定多数の人間を対象地物とする。それらの対象地物に対し、位置座標や通過フラグが属性タイプとなり、観測手順では誤差モデルの適用や集計を行うかどうかなどを記述している。そして各観測データや誤差などの特徴を書いている。これらのモデルにより、観測状況やイベント状況などの再現も可能となる。

2.2 再構成の枠組み

2.2.1 最適化の枠組み

ここでは前章に基づき分類し、構造化された時空間データから再構成を行う枠組みを示す。基本的には地物モデルの持つデータに基づき表現された時空間位置を、観測モデルが持つ観測データから推定を行う。さらにベイズの定理を用いることにより、地物に関する知識を生起確率という形で組み込むことができる。すなわち再構成すべき時空間位置は以下のような数式で表現できる。

$$\hat{\theta} = \arg \max_{\theta} p(\theta | d) \quad (1)$$

ただし、

$$p(\theta | d) = \frac{p(\theta)l(d | \theta)}{\int p(\theta)l(d | \theta)} \propto p(\theta)l(d | \theta) \quad (2)$$

表3 観測モデルとその適用例

Table 3 Observation model and its application examples

観測モデル	GPSによる位置特定	自動改札による人数特定
開始・終了	位置特定の開始・終了	駅の開始・閉鎖
ふるまい		
自身の変化	なし	なし
観測		
時空間範囲	電波受信可能空間でいつでも	改札間を例えば1時間
対象地物		
タイプ	特定のある人間	人間
関連タイプ	なし	なし
属性タイプ	位置座標	通過フラグ
観測手順	位置誤差を伴い観測する	人数誤差を伴い集計観測する
観測データ	例) $x=200, y=238$	例) 120人(1時間)
観測誤差	例) 標準偏差 10m	例) 2,3人程度
関連	なし	なし
属性	X社製	Y電鉄所有

表4 各関数の定義場所

Table 4 Definition place of each element

式	所属オブジェクトとその形式
尤度関数 $l(d \theta)$	観測オブジェクト中で定義
観測データ d	観測オブジェクト中で保持
生起確率 $p(\theta)$	地物オブジェクト中で定義等

$$\left\{ \begin{array}{l} \theta \equiv (x, y, z, t)^t \quad \text{Sequence of} \\ \quad \text{spatio-temporal position} \\ d \quad \text{Observation data} \\ p(\theta) \quad \text{Occur Probability of} \\ \quad \text{spatio-temporal position} \\ l(d | \theta) \quad \text{Likelihood of observation} \\ \quad \text{data to true value} \end{array} \right.$$

また、主要な関数は2.1で見えてきたように各モデル中で構造化したデータを用いて構築できるようになっている(表4)。例えば尤度関数は観測モデル中で定義され、観測データは観測オブジェクト中で保持される。知識に基づく地物の時空間位置に関する生起確率は一般的な知識であれば地物オブジェクト中で保持する。

2.2.2 観測モデルによる疑似観測と尤度計算

尤度関数に入力し尤度を自動計算するにも、観測データに対応するデータを候補の中から自動的に見つける必要がある。そこで本研究では、解の候補に対して観測オブジェクトが疑似観測を行い「観測される

はず」のデータを取得する。それから実際の観測データと比尤度計算を行う。すなわち、図2に示したように解候補として表現されたある地物の時空間位置に対して、観測オブジェクトが実際の観測データを取得した際の時空間範囲を用いてそこに引っかかるものが擬似観測によって得られた位置データということになり実観測データと比較を行う。その時空間範囲も2.1で記述していたものである。

2.2.3 知識による解候補の生起確率計算

各知識では、地物がとり得る時空間的狀態を記述したいいくつかの選択肢をそれぞれの生起確率とともに持たせる。それによって、解候補が属すべき選択肢が特定された後に自動的に生起確率を求めることができる。しかし、その属すべき選択肢を特定するアルゴリズムは、今の所、知識の特徴に依存しており、今後、より汎用的にすべき検討事項である。

2.2.4 遺伝的アルゴリズムを用いた最適解の探索

2.2.2.2.3で算出した尤度と生起確率を合わせた評価値に基づき、最適解を探索する。この最適化問題に我々は遺伝的アルゴリズム(GA)を使うこととし、その理由として以下が挙げられる。

- (1) 広大な解空間中でもグローバルな最適解を見つけやすい
- (2) 時空間位置を個体の配列表現で表しやすい(図3)
- (3) シンプルな配列表現により多様な観測や知識との比較が行いやすい

具体的なGAの手順は基本的なものである。初期生成、評価、淘汰、交叉、突然変異(以降、評価から突然変異まで世代数分繰り返す)とする。もちろん、GAを

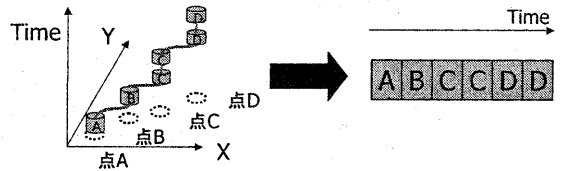


図3 GAの個体における時空間表現
Fig. 3 Spatio-temporal representation in GA individual

用いて最短経路等を求めた研究は^{2),3),8),10)}等多数あるものの、経路情報から経路を決定するだけでなく、地物を断片的に観測したデータにも整合性が取れるように推定するという意味では既存のものと根本的に異なる。

2.2.5 再構成した時空間位置の評価

前節までの流れで求めた最適解を最後に評価しておく必要がある。シミュレーション実験のように真値がわかっている場合には再構成した時空間位置と真のものを比較できる。もし軌跡などのように非集計的に時空間位置を評価する必要があるときには、同じ時刻での位置ずれを累積したものを評価値とできる(たとえば位置誤差の平均値は5.3mといったように)。

しかしふつうは、全員の非集計データを十分に得られないため、非集計で評価してしまうと、高い精度を得るのはかなり難しい。そこで、もともと非集計的な時空間位置にはそれほど興味がなく時間ごとの通過人数、滞在人数などが単純に知りたい場合は、いくつかのチェックエリアや区間を設け、そこを通過、滞在人数を集計し比較しても満足できる。

3. 数値実験

この章では、人の流れや分布について2種類の再現をシミュレーション実験で行う。一つは一人の人間の時空間位置を再現するもので、もう一つはより大規模な人数の人間を再現するものである。シミュレーション実験の場所として展示場を想定し、人は点在するブースを歩き回っているとす。

3.1 一人の時空間位置の再現

まず一人の時空間位置の再現を行い、以下のように3つの条件で比較実験をした。

- (1) 十分に非集計の観測データが得られる場合(GPSデータが1分ごとに得られる)
- (2) 十分には非集計の観測データが得られない場合(GPSデータが3分ごとにしか得られない)
- (3) 十分には非集計データが得られないので他の集計データ(6箇所のトラフィックカウンターによる1分ごとの通過人数データ)といくつかの

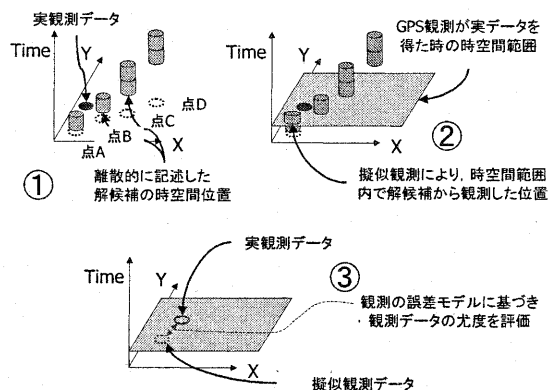


図2 観測モデルに基づく擬似観測と尤度評価
Fig. 2 Pseudo observation and likelihood evaluation by Observation model

表 5 実験における共通条件

Table 5 Common conditions in experiments

実験エリア	60m 四方
再構成の時間間隔	10 分間の行動に関して 1 分間隔
GPS トラッキング	標準偏差 2m の正規分布を持つ誤差関数
再構成後の評価法	1 分ごとの平均位置誤差
GA における個体数	200
GA における世代数	200
再構成実験の試行数	10 回

表 6 異なる条件での再構成実験の比較

Table 6 Comparison with three experiments

実験 ID	計算時間	平均精度	最良精度	精度分散
No.1	41 sec	3.8m	3.8m	0m
No.2	20 sec	14.6m	9.9m	0.8m
No.3	1358 sec	13.4m	6.2m	1.6m

人間の動きに関する知識

条件 3 の知識については次のようなものである。

- (1) 展示場における歩行速度は最大毎分 25m である。
- (2) 起終点の位置は既知である。
- (3) 人間は 30 度以内の鋭角ではターンしない。
- (4) 壁は通過しない。

他の共通な条件については、表 5 に記述し、結果は表 6 にまとめ、各条件での時空間位置の再構成解と真値の比較は図 4、図 5、図 6 に図示した。もちろん実験 1 では最適解は GPS の位置座標を結んだものとなり、真の位置にかなり近い。しかし実験 2 のように十分な観測が得られない時はローカルな最適解に陥り悪い結果になる。しかし、トラフィックカウンターによる集計データやいくつかの知識を付け加えた場合は計算時間はかかるものの精度はかなり向上する。

3.2 多人数の時空間位置の再現

この節では、同様の実験を 10 人や 100 人の多人数で行った。その結果を表 7 にまとめたが、非集計的な精度や計算時間はかなり悪化してしまうものの集計精度に関してはそれほど悪化しないことで今後、規模をより拡大することへの道筋はつかめるといえよう。実際に図 7、図 8 の非集計的な時空間位置では真値と再構成解での差がかなり見られるが、図 9 の集計的な表現の場合はそれほどではない。

4. おわりに

本研究では多様な時空間データからダイナミックに変化する地物の時空間的状態を再構成する汎用的なアルゴリズムを提案し、シミュレーション実験により、

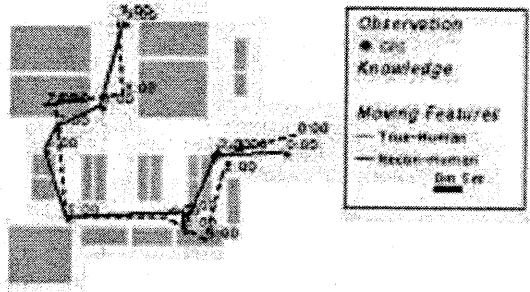


図 4 時間的な密な非集計データから再構成した時空間位置と真値の比較
Fig. 4 The comparison between true trajectory and reconstructed trajectory with enough GPS data (one observation per 1 minute)

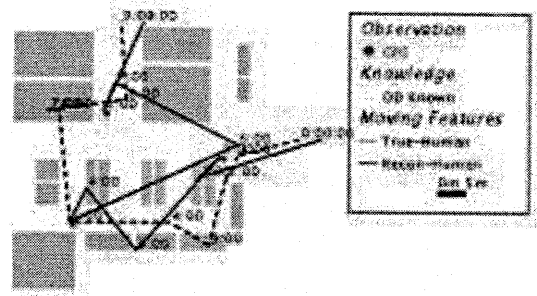


図 5 時間的に疎な非集計データから再構成した時空間位置と真値の比較
Fig. 5 The comparison between true trajectory and reconstructed trajectory with insufficient GPS data (one observation per 3 minute)

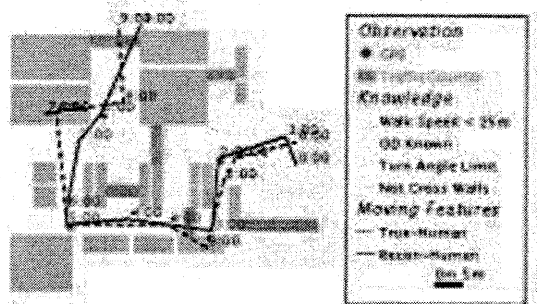


図 6 他の観測データや地物に関する知識を加えて再構成した時空間位置と真値の比較
Fig. 6 The comparison between true trajectory and reconstructed trajectory with other observations (6 traffic counter observations per one minute) and knowledge about human

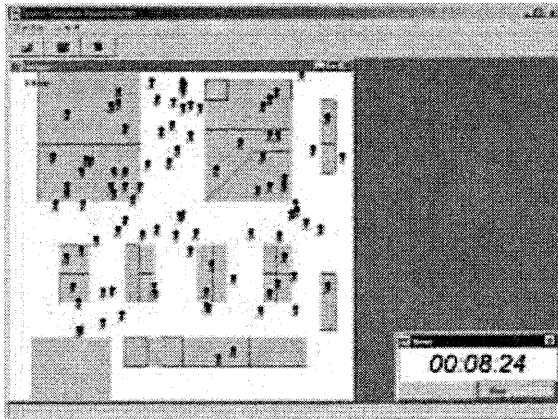


図 7 100 人の真の時空間位置
Fig. 7 True spatio-temporal distribution of one hundred people

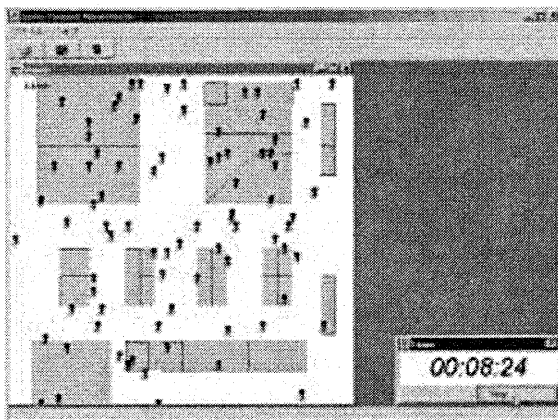


図 8 100 人の再構成された時空間位置
Fig. 8 Reconstructed spatio-temporal distribution of one hundred people

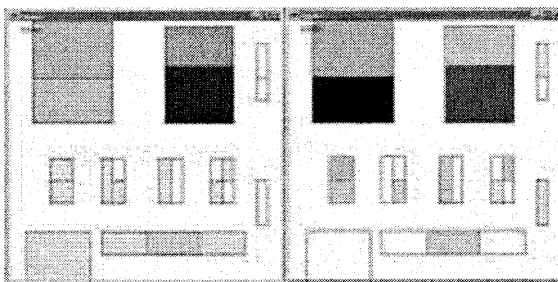


図 9 ブースごとに集計した人数の比較
Fig. 9 The comparison of aggregated number of human distribution in each booth between true one and reconstructed one

表 7 多人数の場合との比較

Table 7 Comparison with a great number of people

人数	収束までの計算時間	非集計精度	集計精度
100 人	150 時間	10.1m	5.6 人
10 人	2.1 時間	5.5m	0.6 人
1 人	21 分	5.2m	0.1 人

いくつかの時空間観測データや人間に関する知識から人間の時空間の状態を再現できることを確認した。しかし今後規模を拡大する上で、より再構成の精度を高め計算時間を短縮することが必要となる。

精度を向上しようとした場合、ローカルな最適解に陥らないためにはある程度以上の観測、あるいは地物に関する知識が必要であり、それらの程度と問題の規模との相関関係を把握する必要がある。それにより逆に正しい地物の時空間状態を再現するための必要最低限の観測計画を立てることも可能になる。

解を得るまでにかかる時間は、1 世代あたりにかかる計算時間と最適解に収束するまでの世代数を掛け合わせたものである。1 世代あたりの計算時間は評価部分でかなり時間をとり、一般に観測数が多いほど評価に時間がかかるものの、規模との相関関係はまだ定量的には把握できていない。また収束するまでの世代数は最適化問題の構造に依存しており、その特性は現時点では経験的に把握しているのみで、今後定量的に把握する必要がある。

さらに再構成すべき地物の性質や要求精度、観測の性質に応じて効率的な評価手法などを用意することが、今後、より規模を拡大していく上で必須であろう。

参 考 文 献

- 1) Cameron, S.: Collision Detection by Four-Dimensional Intersection Testing, *IEEE Trans. Robotics and Automation*, Vol.6, No.3, pp.291-302 (1990).
- 2) Goldberg, D. and Lige, R.: Traveling Salesman Problem, *Proceedings of International Genetic Algorithm 1* (1985).
- 3) Grefenstette, J., Gooal, R., Rosamaita, B. and VanGucht: Genetic Algorithm for Traveling Salesman Problem, *Proceedings of International Genetic Algorithm* (1985).
- 4) Langran, G.: A review of temporal database research and its use in GIS applications, *International Journal of Geographical Information System*, Vol. 3 (1989).
- 5) Sekimoto, Y. and Shibasaki, R.: Conceptual Data Modeling for Dynamic Revision of

Spatial-Temporal Database, *Proceedings of 9th International Symposium on Spatial Data Handling(SDH2000)*, pp. 42-56 (2000).

- 6) 飯村威, 高澤信司, 久保紀重, 平井政二, 大伴真吾, 荒井徹哉: 空間情報と時系列情報の統合化に関する研究—土地・建物情報管理のためのプロトタイプ—, 地理情報システム学会講演論文集, pp. 113-117 (1998).
- 7) 麻生英樹, 赤穂昭太郎, 本村陽一: 統計的推論とAIの推論, 人工知能学会誌, Vol. 12, pp. 196-203 (1997).
- 8) 山村雅幸, 小野貴久, 小林重信: 形質遺伝を重視した遺伝的アルゴリズムに基づく巡回セールスマン問題の解法, 人工知能学会誌, Vol. 7, No. 6 (1992).
- 9) 関本義秀, 柴崎亮介: 時空間データベースのダイナミックな更新を目指した概念データモデルの提案, GIS—理論と応用, Vol. 8, No. 1, pp. 63-73 (2000).
- 10) 狩野均, 中村信昭, 中村友洋: 知識の集団を用いたGAによる不特定な立ち寄り地を含む経路探索, 人工知能学会誌, Vol. 17 (2002).
- 11) 太田守重: GISのための時空間スキーマ, GIS—理論と応用, Vol. 7, No. 1 (1999).
- 12) 黒木進, 牧之内顕文: 位相空間データモデル Universe での空間、時間、時空間データ表現, 電子情報通信学会データ工学研究専門委員会、情報処理学会データベースシステム研究会 合同研究会 データベースワークショップ (1998).
- 13) 大津展之: 認識と理解のための柔らかな論理, 電子情報通信学会誌, Vol. 71, pp. 1231-1240 (1988).
- 14) 畑山満則, 松野文俊: 災害時での利用を考慮した時空間地理情報システムにおけるデータ構造に関する考察, 情報処理学会論文誌: データベース, Vol. 41 (2000).
- 15) 大沢裕, 金景月: 離散的な時系列管理方式の一提案, 地理情報システム学会講演論文集, pp. 107-112 (1998).