

インフラとしてのデータ整備と 解析技術による利活用の展開 ～社会基盤情報学に向けて～

東京大学大学院工学系研究科 社会基盤学専攻准教授 布施 孝志

1. はじめに

ただ今ご紹介いただきました、東京大学の布施と申します。本日はどうぞよろしくお願ひいたします。近年、計測技術の進展に伴いまして、アクティブなデータ、パッシブなデータ問わずに大規模、さらには多種多様、そしてそれが高頻度でデータが取得されています。そのような流れを受けて、ビッグデータという言葉が流行りましたが、それに対応した社会の動きを、国の白書でも見ることができます。平成 24 年版の情報通信白書では、まだまだビッグデータという言葉が出始めたということもありまして、「ビッグデータとは?」という項目で、どのようなものを指してビッグデータと呼ぶのかということが説明されています。ただし、もうそれからだいぶ時間が経ちまして、ビッグデータという言葉がかなり浸透しましたので、今や、あまりビッグデータという言葉自体を騒ぐということは減ってきているのかと思います。

その流れは情報通信白書を見ているとわかりまして、平成 25、26 年版でも一部ビッグデータの利活用に関する項目がありますが、最新の平成 27 年版の情報通信白書を見ますと、もう目次の中でもビッグデータという言葉は 1 箇所だけです。情報通信白書ですから、ICT 活用を推進していこうということは当然なのですが、その中の一つとしてビッグデータが位置付けられるようになってきています。

さまざまなデータとともに、ICT を使って新たな未来の社会を作っていく、このような流れになっているのかと思います。特にこのビッグデータ

だけに限らず、データに関しましては、当然データだけで使えるというのではなく、そこから重要な、あるいはこれまで気づかれていないような情報に変換していくということが重要になってきます。

2. 社会基盤分野における ICT 活用の動き

ICT 活用の推進に関しましては、我々と近い社会基盤の分野でもいろいろな取り組みがなされています。国土交通省のものだけをピックアップしますが、先ほどのお話でもありました i-Construction が昨年提唱されたところです。つい先日、CIM 活用推進委員会も発足しまして、国土交通省全体でこの CIM、あるいは i-Construction を推進していこうということになっています。もう皆さんは、i-Construction はご存知のところですから、細かくは説明しませんが、特にこの測量、設計・施工計画、施工、施工検査というすべてのところで ICT を使って情報を共有しよう、あるいは効率的にデータを取得しようということが本筋かと思っています。

そのなかで、測量では、ドローン等による三次元測量ということが言われています。それに関しまして、私は少しだけマニュアル作りに携わりましたので、本当に基本的な考え方のところだけをご紹介させていただきたいと思います。この測量のところに関連する「UAV を用いた公共測量マニュアル(案)」が国土地理院から出されました。i-Construction に係わる新しい基準が 15 個作られたものの一つになります。このマニュアルでは、大きく二つの内容を扱っており

まして、一つは「UAVを用いた写真測量を公共測量へ導入する」ことです。これは、これまでの2500分の1から、より大縮尺、つまり500分の1、250分の1などの縮尺にも写真測量を適用しようというものです。そこでもUAVを使えますという話です。もう一方のほうが、これまであまりなかったのですが、先ほどのご発表でもありましたが、「公共測量の成果にUAV写真による三次元点群データを追加する」というものです。出来形のほうに展開していこうという狙いがあります。

今回、このマニュアルを3月に出しました。その基本的な考え方だけをご説明させていただきますと、これまでの国土地理院が出しているマニュアルと意味が若干違います。それはどういうことかということ、一つは、従来のマニュアルですと、やはりプロが見てそれを理解して適用するというものだったのですが、今回は専門家以外も対象になります。特に、三次元点群を使うところで、工事測量でも使うということが重要になってきます。そのようななかで、専門家以外の方も対象になるため、特に一般的な説明を細かくしています。図解を入れたり、あるいは何編にも分かれているなかで、重複を許して繰り返しの説明をしています。また、これまでですと、こうこうこういう手順で行うとこの精度が出ますということだったのですが、今回はより柔軟性を持たせるために、一般的な方法論を記述しているということです。マニュアルのなかには精度の具体的な数値が出ていますが、これはあくまで例でしかありません。当然、高精度でとりたいところもあれば、工事の進捗管理など、やや精度を落としてもいいところもあり、精度がいろいろ違いますので、適用対象によって高い自由度が持てるようにしております。その目標精度によって方法が調整可能であるようにしているものです。

もう一つ、このほうがより重要な考え方なの

かもしれません。ここでは「進化するマニュアル」と書いてありますが、UAVを使った測量も、現在発展途上にあるということです。それは、ベースにしている技術がStructure from Motion (SfM)であったり、Multi-View Stereo (MVS)であったり、UAVのハード的な面も含め、進展が進んでいるところです。そのため、UAVによる三次元点群を取得する事例が、実は十分ではありません。このマニュアルでは、現状の技術を勘案した手法や許容範囲について、ここまでがんばればできるのではないかといいところにしてあります。今後いろいろな適用事例を通して、どんどんとマニュアル自体を発展させていこうという考え方に基づいています。そのため、一般的な方法論は示したものの、今後検証データを蓄積してより良いマニュアルに変更していこうという考え方になっています。これに関しましては、このマニュアルはより適切な内容へと改善することを、概説の「はじめに」に明記しています。ですから、このマニュアルが出たからこれだけでやらなければならないということではなく、より効率的な方法は何なのかということ、みなさんのご協力でご提案していただければと思っています。

国土交通省のなかで、道路局でもデータを活用していこうということが進められています。これは道路局内にある「地域道路経済戦略研究会」という研究会の中間提言です。この中間提言も先の3月に公開されています。このなかでは、IT技術の飛躍的な進化や、ETC2.0を含む多様なビッグデータということが前提にあるわけですが、大きく3本の柱で成っております。これに関しても私はとりまとめ担当で関わらせていただきましたので、少しだけご紹介させていただきます。

この3本柱の一つは、ビッグデータ活用のためのプラットフォームです。まさにここが最も関係するところかと思えます。そのほかは、公共

交通や新たなモビリティとの連携や道路空間のマネジメントという3本柱になっています。このなかでビッグデータ活用のためのプラットフォームについて、いくつか挙げておりますが、まず、道路関係データ自体をしっかりと体系化しようというものです。これまでのデータを捨てるのではなく、それも有効利用しながら新しいデータを使っていきます。そのための既存データの整理や先進技術の活用ということになります。特にこのETC2.0と関係するところでは、動的データに対してプライバシーが関係してきますし、大容量のデータになりますので、それをいかに効率にかつ安全に管理をしていくかも重要になります。あるいは、いかにしてデータをハンドリングしていくかということにも触れております。

あとは、マルチモーダル道路空間マネジメントのためのプラットフォームの構築と運用課題の抽出ということになっております。ここで中間提言に挙げたものは、我々のような大学の人間だけが「こうしましょう」と言うだけではなく、道路局プラス地方整備局の方々もご参加いただいて、今、各地方でタグを組んで、実際に研究あるいは実用化への取り組みを進めているところです。そして2020年に向けて実装していくことを目指しています。実際にこれから4年ほどしかございませんが、そこに向けて進んでいくのだという意思表示の中間提言になっております。先ほどの関係するところで言いますと、新たなセンシング技術、ここは皆さまの方がお詳しいところかと思いますが、こういう内容も視野に入れていきます。あるいは、近年の自動運転との関連も議論がされつつあるところです。もう一つは、データがあれば良いというだけではなく、それをきちんと共有する必要があるため、データ共有のための共通基盤の構築も、道路局としてしっかり取り組むべき項目になっています。ここでのデータ共有に関しては、物理的なデータベースあるいはプラットフォームというだ

けではなく、さまざまな主体、関係機関が手を組みながら情報を共有していこうという社会的な組織論のところまでも踏み込んでいる内容となっています。最近の国の動きで、私が関係させていただいたところを簡単にご紹介させていただきますが、これ以降が今日のお話のメインになってきます。

3. データ増加と取り巻く環境の変化

先ほども言いました通り、ビッグデータと言われていますが、それはデータだけの問題ではなく、その背景には計測技術の進展や計算機能力の向上、あるいは解析技術の発展などがあり、全体としての環境を指しています。データを取り巻く環境というものが現在のその流れを作っているのではないかと考えられます。ここでデータは、ほとんど位置と時間を持っていますが、データはそのままでは情報ではないということです。「情報とは何か」ということを書いている方もたくさんいらっしゃいますが、そのなかでよく言われるのが、情報とはデータに加えて目的に対する有用性を持つものだということです。つまり、生データがあればそのまま情報になるわけではありません。それがすぐに使えるわけではないということです。そのためにデータの高度な解析を行って、判断のために必要な知識に変換していく。ここのステップが大切になるわけです。現在ですと、このデータの高度な解析というところでは、特にもうビッグデータの次の流行り言葉になっているかもしれませんが、人工知能とか機械学習というものが非常に脚光を浴びています。

人工知能自体は1956年のダートマス会議で提唱された言葉ですが、それから第二次のブームも過ぎ、それが再び脚光を浴びるようになってきています。現在は第3次のブームと言われていますが、そのブームを牽引しているのがいわゆるディープラーニングです。おそらくディ-

プラーニングが最も早く話題になったのは、グーグルのオフィシャルブログで紹介された、普通に撮影した非常に大量なデータからネコの自動認識に成功したことだと思います。このブログが2012年に出てから、非常に脚光を浴びるようになってきました。そのあと、ILSVRCという画像の一般物体認識のコンテストで話題になりました。これは一般的ないろいろな画像が提示されて、それが自動的に何であるかを認識しようというコンテストです。このなかで、2012年以降はほとんどディープラーニングを使った手法が上位を占めています。さらには、最近では囲碁でプロに勝利したことも話題になりました。あれはディープラーニングだけではなく強化学習というプラスアルファの手法を使っているのですが、ディープラーニングがさらに脚光を浴びるようになってきています。

また、ディープラーニングを適用する環境も整ってきました。私の研究室でも簡単に、簡単にという言い過ぎかもしれませんが、少し苦労すればこのディープラーニングを適用することができます。たとえば、地形図に使えるかということで試してみました。何に使ったのかというと、旧版地形図の地図記号認識です。現在の地図はもうすべてベクターデータ化されていますが、旧版地形図は画像でしかないわけですから、今後のベクター化のための第一歩として、地図記号だけでも自動認識できないかと考えました。ある地形図を見ると、これだけたくさんの地図記号があるわけですから、それを目視で認識することはやはり非現実的です。そこを自動化しようというわけです。ただし、旧版地形図は手書きで地図記号が書かれていますので、この水田の地図記号を見ましても、かすれがあったり左右で長さが異なるようなものが非常にたくさんあります。そのため、単純にこれまでの画像のマッチング手法を使うだけでは、この自動認識は困難です。それでは、そ

こをディープラーニングで学習させて、認識できないかということを試しました。ディープラーニングのいくつかある特性のうち非常に重要な特性は、特徴量を自動的に作ってくれるということです。これがディープラーニングの非常に有効なところ。地図記号を認識するときには、これまでの画像マッチングですと画素値を使います。しかし、画素値のみで地図記号の特徴を表現することは、非常に難しいです。それだけではマッチングできません。またパイロット的な実験をしたところ、何かしら特徴量があったとしても、閾値だけに頼る単純な識別だけでは困難だということがわかりました。そこで、先ほどのディープラーニングの一つの有力な機能である自動的に特徴量を見つけることに着目すると、オートエンコーダ(AE)という手法が提案されています。このベースはニューラルネットワークですから、入力層からいろいろつながって中間層があり出力層があります。これは一番単純な3層のニューラルネットワークです。それに対して、入力層にインプットデータを入れて、さらに出力層にも同じデータを入れます。インプットとアウトプットは同じものを入れていきますから、それをこのニューラルネットワークで学習してあげれば、再現するために必要となる特徴量が自動的に抽出されます。これは人間が特徴量を職人的に作るのではなく、ニューラルネットワークが自動的に抽出してくれるというものです。この特徴量を使って、より柔軟な識別法として、昔から使われているサポートベクターマシン(SVM)を用いて機械学習を行うということになります。これが一つ目の方法です。

ただしもう少し細かく見てみると、たとえば果樹園とか広葉樹林とかありますが、両方とも大域的な特徴と局所的な特徴とが階層的にあると考えられます。こういう階層的な特徴の表現が好ましいということが考えられます。さらに識別性の自由度を高めるということも望まれます。こ

れらに該当する方法が畳み込みニューラルネットワーク (CNN) と呼ばれるものです。最近の画像処理で最も高精度な手法、成果をあげている手法がこの CNN、あるいはこの CNN の発展版です。CNN は階層的ですから、先ほどのニューラルネットワークを何層にもつなぎ合わせたものになります。何層にもつなぎ合わせて層が深くなるということが、ディープラーニング、深層学習と呼ばれる所以です。CNN では、何か入力画像があったときに、特徴量を抽出するためのいくつかのフィルターをかけます。また、位置ズレに頑健になるように、プーリング層と呼ばれる層を通して、さらにそれらをつなげて、最終的なニューラルネットワークの構造を作ります。そして、パラメータを学習データから学習させて、自動的に認識をすることになります。

この例では、特に、土地利用と関係する地図記号のうち、11 種類に関して適用しました。

今回はサンプル数をいくつか試してみましたが、大体 5000 個以上になるとそれほど性能が変わりませんでしたので、5500 個のサンプルデータを使って、そのなかから 500 個を抜いてクロスバリデーションによって精度を評価しました。その結果、従来の画素値を使った手法に対して、オートエンコーダと SVM を使うと、地図記号が正しく認識された率が上がりました。さらに階層化させた CNN では、98.7% の精度が出ました。よくわからないけれども、とにかく正解を導くことができるということがディープラーニングの特徴です。

一度、地図記号を認識すれば、その土地利用がどう変化するかという簡単な応用もできるわけです。認識されたポイントデータに対して、カーネル密度推定法によって色付けをしていきます。これは明治 44 年の石和の旧版地形図ですが、ディープラーニングで地図記号を認識して、色

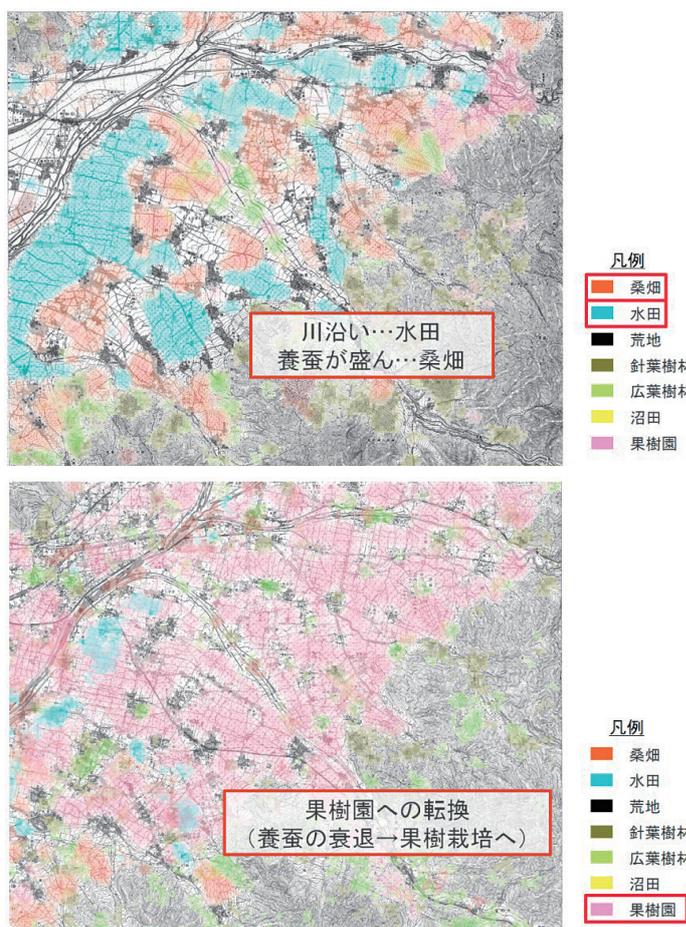


図1 ディープラーニングによる地図記号の自動認識とカラー化 (上: 明治 44 年, 下: 昭和 45 年)

付けをすると、図のようになります。ここは石和の典型的な扇状地ですが、明治44年のときには、この青が水田ですが、川沿いには水田が広がって、さらに自然堤防では桑畑になっています。このオレンジが桑畑になります。今度は昭和45年の地形図に同様に適用して色付けしてみると、これらのところがピンクに変わっています。これは実は果樹園でして、この昭和45年には養蚕が衰退してほとんどの農家が果樹園に転換したことも一目瞭然になるわけです。

ここまでは地図記号を認識するという話でしたが、地形図全体に適用すると、先ほどの98%の認識率が落ちます。70%くらいになります。その理由は、ディープラーニングはあくまで識別することが得意であって、どこに物体があるかということまでは探し当ててくれません。画像処理の研究でも最近主に用いられていますが、この一般物体認識と呼ばれる問題に対し、まず物体らしいものがどこにあるのかという物体の位置検出を最初に行います。その上で、この物体らしい四角で囲まれたところにディープラーニングを適用して、これが車なのか人なのか看板なのかを識別する手法が、最も成功している例になっています。

では少し視点を変えて、衛星画像や航空写真でこの物体位置検出を試みようと考えます。そうしますと、当然、画像のサイズが大きいですから、高速ではないといけないということや、スナップ写真と違って物体の回転がいろいろあり、ボケもあり、それらに頑健であることが求められます。あるいはバンド数やビット数も、衛星画像では最近増えていますので、それへの対応も求められます。レビューをした上で最も向いているであろうというのがBING(ピング)と呼ばれる手法、これは2014年に提案されている手法です。BINGとはどういうものかといいますと、複数サイズのウィンドウを準備してお

きまして、それにより画像を全探索します。そして、その探索した結果、各ウィンドウに対して物体らしさという特徴量を算出して、その物体らしさの高い順に物体の位置を矩形で提案する手法になります。その物体らしさの特徴量を計算するものはNG特徴量と呼ばれ、輝度勾配を使ってそれをスケールに関係なく8×8サイズのフィルターを通してスコアを計算するものです。これを学習させて、今度は新規にテスト画像を入れたときに物体らしいところを矩形、この矩形をプロポーザルと呼んでいます。これを提案していくということになります。

今回はこの方法をWorldView-3の衛星画像とDMCの航空写真の両方に適用しました。今回は物体を車両に限定しておりますが、それぞれ2600台くらいと1200台くらいになります。検証項目はいろいろ考えられますが、今回は特にこの航空写真と衛星画像の違いは何なのか、あるいは事前に学習をさせて新しいデータに対して答えを出すため、学習用画像の分解能の相違の影響などを検討しました。どの方法がいいのかは、検出率という指標で見ます。先ほどの物体認識の例で言いますと、最初に位置を抽出してそのあと識別するわけですから、抽出のところでなるべく取りこぼしが少ないほうがいいということになります。一回抽出さえしてしまえば、ディープラーニングで、物体ではないプロポーザルをはじくことができるわけですから、まずは取りこぼしの少なさを表す検出率を設定しています。実際に存在する物体の数に対する予測が的中した比率ということ。そして横軸にそのプロポーザルの数を、縦軸に検出率を示していきますと、プロポーザルの数を増やせば増やすほど、当然検出率が上がっていくのですが、この立ち上がりが良いほど性能が良いということになります。また、最終的に収束したところの検出率が高いほど性能が良いということになります。まずは航空写真と衛星画像で比較して

みますが、衛星画像は立ち上がりが悪いのですが、航空写真は解像度が高いということもあり、航空写真のほうが検出がしやすいということがわかるかと思えます。ただし衛星画像に関してはプロポーザルの数を増やしていけば90%程度までは上がります。また、4バンドや11ビットを試しましたが、あまり大きな変化はなかったことを確認しました。

今度は学習の画像分解能が変わった場合にどうなるのかを見てみます。矢印の左側が学習データの空間分解能です。適用先は航空写真に対してだけです。そして右側がテストデータの空間分解能です。同じ分解能のもので学習してテストをするという方が性能がいいということは全体的な傾向ですが、そのような状況ばかりとも限らないため、これらをミックスした場合にどうなるのかを検証しました。特に注目されるのは、15cmのテストデータに対して、7cmや30cmなど、異なる分解能画像を混ぜると、少し性能は上がります。頑健性を増やすためにも、分解能の異なる画像で学習させておくということが重要になります。これはまだ途中段階のもので、このあとまた発展させていこうと思っているところです。

ただしこのニューラルネットワークには限界がありまして、基本的に中で何が起きているのかわかりません。ブラックボックスです。勝手に学習をしてくれるだけです。ですから、適切ではない結果が出てきたときには、もう原因を把握することは困難です。どうすればいいのかという指針を示すことすら難しいということです。最近の画像処理の研究では、途中で何が起きているのかを人間が理解しやすくするようにインバージョンという方法もとられているのですが、これはまだ緒についたばかりで今後の発展が期待されます。

4. データの統合・データとモデルの統合

代わりに、データの構造をしっかりとみていこうという動きがあります。このニューラルネットワークの中はブラックボックスだけれども、とにかく当てにいこうということに対して、よりデータ間の関係性を見ようというアプローチがこのデータの構造化です。そして、最近非常に着目されている一つが、いわゆるスパースモデリングと呼ばれるものです。これは、たとえばアマゾンのおすすめ商品の推薦のために使われている技術です。他には、ベイジアンネットワークもよく用いられ、バイオインフォマティクスと呼ばれる遺伝子の構造を理解しようという分野において、その構造を逆推定していこうという方法です。もう一つは先ほど少し申し上げました、情報はデータに対して何か有用性が必要であり、その有用性を作るためにモデルを使うというアプローチです。

モデルはいわゆる3次元モデルだけではなく、社会的な現象を数学的に表現した、抽象化したものになります。このデータとモデルを組み合わせる、モデルは既存の知識ですが、それを組み合わせて情報にしていこうという考え方がデータ同化の基本にあるところです。そして、データ同化の技法を使えば、多様なデータを統計的に合理的に統合することが可能になります。さらにモデル化した知識とデータを統合することも可能です。それを支えているものが、一般状態空間モデルと呼ばれる枠組みです。その背景にある理論がベイズ統計学になります。データもこれだけたくさんあるわけですから、多様なデータを組み合わせることがポイントになるわけです。そしてそれはデータの統合だけではなく、データとモデルの統合も考えましょうということになります。

ではこれから、データの統合、データとモデルの統合へ向けた、このデータ同化について、私たちの研究室で行っている例を紹介してい

たいと思います。まずはデータの統合、データとモデルの統合の基礎理論になります。その前提となる考え方ですが、まずデータ取得や解析環境の進展はこれまで申し上げたとおりです。次に、データがたくさんあればそれで良いのか？ということ。データ量が増えること自体は良いのですが、データの種類もどんどん増えているわけです。これはデータ空間として描いてみましたが、データの増加として、たとえばこの青の点しかなかったところに対して、たくさん赤い点が増えています。これはこれで良いのですが、種類も増えますから、データ空間の次元も増えていきます。このデータ空間の次元の増え方というのは、実はデータの増え方以上に影響があります。そのため、このデータ空間の次元が増えてしまうと、どんなにデータがたくさん集まってもスカスカな、いわゆるスパースなデータになります。こういう状況を指して、ビッグデータの価値密度は低いとおっしゃっているのが、統計数理研究所の樋口知之先生です。樋口先生はデータ同化を世界的に牽引されている研究者です。

このスカスカなデータに対して、何か未知の点のデータを知りたいといったときに、普通考えることは、周りの点から内挿していくということです。ただし、あくまでスパースなデータですから、内挿の精度はどれくらいあるのだろうかというところが疑問になるわけです。もう一つ、データが周辺にないような、あるいは未来を予測するというような外挿の世界になってくると、どうすればいいのかということになります。外挿の世界になってきますと、予測のモデル、あるいはシミュレーションで対応していくということがよく行われています。近年ですと、シミュレーションもだいぶ進展してきているところです。データから何かを予測する帰納的な考え方と、シミュレーションによって演繹的に何かを予測する考え方、この二つを組み合わせると内挿あ

るいは外挿の精度を向上させることがこのデータ同化になるわけです。その帰納的な世界が観測データに対応しますし、演繹的な世界が予測モデルに対応するということになります。

このデータ同化の意義は、大きくは二つあります。一つは、たとえばこの黒い線とか薄いグレーの線がシミュレーションだとしたら、シミュレーションとしては境界条件などがありますから、いろいろな候補があるわけです。それに対して、観測された赤い点を最もよく通るようにシミュレーションのモデルを作る、つまり、実測データを用いて数値シミュレーションモデルの精度を改善するということが一つです。もう一つは逆に、赤い点に対して青い点、まだ観測していない点を予測したいときに、シミュレーションモデルを用いることによって観測の不足を補ったり、観測誤差を修正することがこの意義になります。このデータ同化の方法には、逐次型と非逐次型がありますが、今回は時々刻々とデータが入ってくるわけですから、それに対応した逐次型に限定します。その逐次型のモデルとして、抽象的に表したものが、この一般状態空間モデルと呼ばれるものです。この z が観測データを表しています。添字が時刻になりますが、観測データは時々刻々と入ってきます。一般的には、それらを時系列にモデリングすれば良いのですが、今回はより表現力を高めるために、本来観測できない隠れ変数と呼ばれる x を導入します。これを状態ベクトルと呼んでいますが、観測できないあるいは我々が知りたい量です。隠れ変数を導入することによって、かなり柔軟にモデルを作ることができるようになります。隠れ変数である状態ベクトル x と観測データ z の関係性を示すモデルを観測モデルと呼びまして、状態が逐次変化していく関係性を示すモデルをシステムモデルと呼んでいます。

一度この状態ベクトルと観測ベクトルを定義して、観測モデル、システムモデルをそれぞれモ

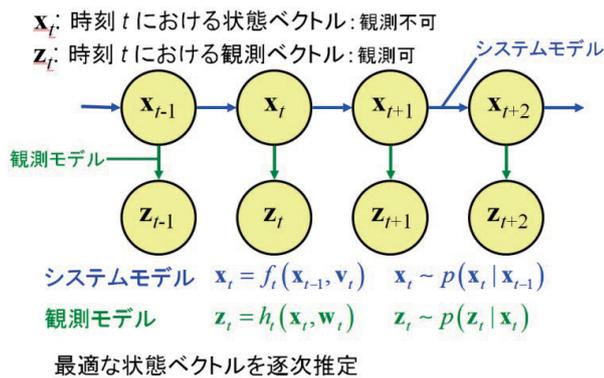


図2 データとモデルの統合のための基本的な枠組み

デル化してあげれば、これは確率的にモデリングができますが、あとはベイズ推定の理論を使って時々刻々と最適な \mathbf{x} を推定するということが可能になります。具体的にもう少し言いますと、時刻 $t-1$ の状態がわかっているとしましたら、システムモデルを使って状態の変化を予測します。今度はデータが入ってきた観測モデルに従って \mathbf{x}_t を修正します。そのときに最終的には事後確率の期待値などを使って、もっともらしい \mathbf{x}_t を求めていくということです。この後、私たちの研究室で行った応用例を紹介させていただきますが、すべてこの枠組でモデリングをしております。ただし適用対象は違います。

ではまず、モバイルセンサの自己位置推定を考えたいと思います。今回、対象としておりますモバイルセンサは、スマホなどですが、特に屋内での利用を想定しています。いろいろなセンサがありますが、何も周りに設置せずに屋内で利用しようということを考えます。いまやスマホにはIMUやカメラ等々がついておりますので、それらを統合します。これはデータの統合の一つになります。また、逐次型として、自己位置推定をリアルタイムに行っていきます。時刻 $t-1$ と時刻 t でスマホを持ちながら動いた場合に、今回知りたい状態ベクトルのメインは位置です。その他の状態を推定するために、速度、加速度、角度、角速度なども使っています。これに対して先ほどのシステムモデルを作れば、時刻 t

ではどこに移動するかをある程度予測できます。さらにこの時刻 t でIMU、あるいは画像からデータが得られた場合には、それを用いてこの位置を修正していきます。そのシステムモデルに関しては、等速直線運動をしているという仮定を今回はおいています。IMUの観測モデルに関しましては、加速度センサと地磁気センサとジャイロセンサを使ってこの特徴量を求めていきますが、ここは線形の方程式になっていますので、いわゆるカルマンフィルタで解くことができます。かたや、画像に関しては、非線形の問題になりますので、カルマンフィルタを適用できません。ですからカルマンフィルタの一般形として、離散的に近似するパーティクルフィルタという手法を画像のほうでは使っています。

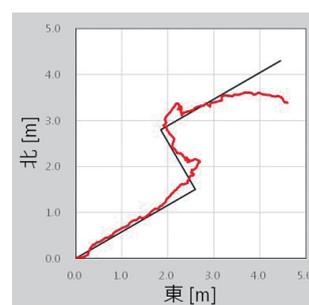
そのときの観測モデルは、どういうものがあるのかというと、写真測量では一般的に相対的な位置・傾きを求めるには共面条件を用います。ただし共面条件は、同一平面上にあるという条件ですから、スケールに対しては不定となり、進行方向で少しずれても、同様な確からしさが得られてしまいます。そのため、確からしさを計算すると、進行方向にばらつきます。一方で、基線長が非常に短い場合にもよく使われる方法として、コンピュータビジョンの分野で提案されているInverse Depth (インバースデプス) モデルというものがあります。これは逆に特徴点の方向にカメラが移動しても、もっともらしさは変わりません。そのため、特徴点の方向にばらつきが出てきます。今回これらを両方とも統合することにより、かなり精度は上がりました。

提案手法を適用した結果がこちらになります。我々の学生部屋で実験したものです。部屋を歩いている様子がこの軌跡ですが、三次元モデルの中に推定結果をプロットしたものがこちらの図になります。これはほぼリアルタイムでできるようになっています。平面で見るとこうなります。実際に正解値と比べてみますと、

軌跡の距離は短いのですが、平均誤差としては40cmくらいが出ています。こちらを見ていただくとわかるとおり、最後のところがずれていますが、これは地磁気の乱れによるものです。ただしこの大きな乱れを加えても40cmということで、精度的にはかなり上がっているのではないかと考えられます。

ではもう少し長い範囲、こちらは大学の廊下になりますが、ここでも同様に推定をした結果がこちらの図になります。途中でずれているところがあるかと思えます。実はこの部分では推定は安定しています。あくまでこれは相対的にしか測っておりませんので、絶対値を参照していません。どこが悪さをしているのかというと、ここです。この辺りから変化が起きています。実際にこの変化が起きたあと、推定が安定しているところを回転してみますと、しっかりと廊下を歩いているという結果になります。これは、蓄積誤差のためにこのような結果になり、この部分がそのまま影響しています。先ほども少し申し上げました通り、これは磁場の方向をはかったものですが、スタート時点から磁場が30度ぐらいずれていまして、それが全く同じように軌跡でもずれているという、その影響になります。当然これに関しましては、何かしら参照点を使わないと対応しようがありません。ここでは、画像を使っていますので、看板などからある特徴点の座標をデータベースとして持っておけば、新たに参照点を設置する必要はなく、参照点として認識して補正をするということも考えられるようになります。これが自己位置推定の話になります。

位置を推定することに関しては、別のデータなのですが、波形記録式レーザでも全く同じ枠組みが適用できます。波形からの地物検出を考えたとき、照射波形があった場合に、単純に閾値だけで閾値以上のところが地物だと認識する方法もあると思います。しかし、当然ノイズ



— 実際の経路
— 提案手法の適用結果



図3 モバイルセンサの自己位置推定
(上:室内、下:廊下)

が入ってきますから、そのノイズと地物の識別には課題が残っています。それに対応するために、何らかの関数でフィッティングをさせる手法も非常によく使われています。この方法ですと、事前に地物がいくつあるかということを設定する必要があります。さらにスキャニングデータにも関わらず、単一の照射のみしか使っていないという、もったいないことも考えられます。そういう意味で、波形データの解析に求められることは、単一の照射であっても、地物からの反射、ノイズの識別が必要であるということ、さらにある地物から近傍の地物の位置に関しては予測が可能であることになります。

このようなスキャニングデータとしての特性、連続性も利用しようとする、先ほどの時系列のフィルタリングになります。さらに、ここでは地物数の推定も同時に行うことになります。そのために、消失や出現に対しても適切に表現することが、ここでの目的になります。ただし枠組みは変わりません。求めたい状態ベクトルは

地物の位置です。さらに今回は、地物の数も変わるために、その地物の数も状態として同時に推定します。観測ベクトルは波形データになります。システムモデルはその隣にスキヤニングしていくことから作成します。観測モデルも観測の照射の強度から推定するということになっています。そうしますと、予測のところでは、特にポイントが残存したり消失したりとか、あるいは新たに出現するという項を明示的に入れて予測のモデルを作ります。フィルタリングの方にも、検出や誤検出があることも使っています。

こちらは、自然教育園でとった波形記録式レーザのデータですが、これに適用した結果が次の図になります。この左側のプロットした点は、横軸がスキャナからの水平距離で縦軸がスキャナからの鉛直距離になります。つまり、一本のレーザがあった場合に、得られた波形から、先ほどの手法により地物と認識されたところを

プロットしたものです。赤い点が推定で、黒い点がマニュアルで位置を特定したものです。これに対して、スキヤニング全部をプロットしたものがこちらの図になります。マニュアルと比較して50cm以内のところを囲った部分で、全体的な傾向としては捉えられていることがわかるかと思います。特に、スキヤニングラインの番号kが4~7番のところでは、どういうことが起こっているのかを説明したいと思います。そのために、まず図の見方を説明したいと思います。これは、横軸が経過時間です。つまり、どのように波形が変わっていくのかというものです。縦軸がその波形のデータの強度であり、さらにこの色付きのものがそれぞれ地物らしいところになります。これはラベル変数が振ってありまして、違う色は違う地物であり、数までも推定しているというものになります。そうすると、k=4のところでは、明確に3つの山があり、3つの地物

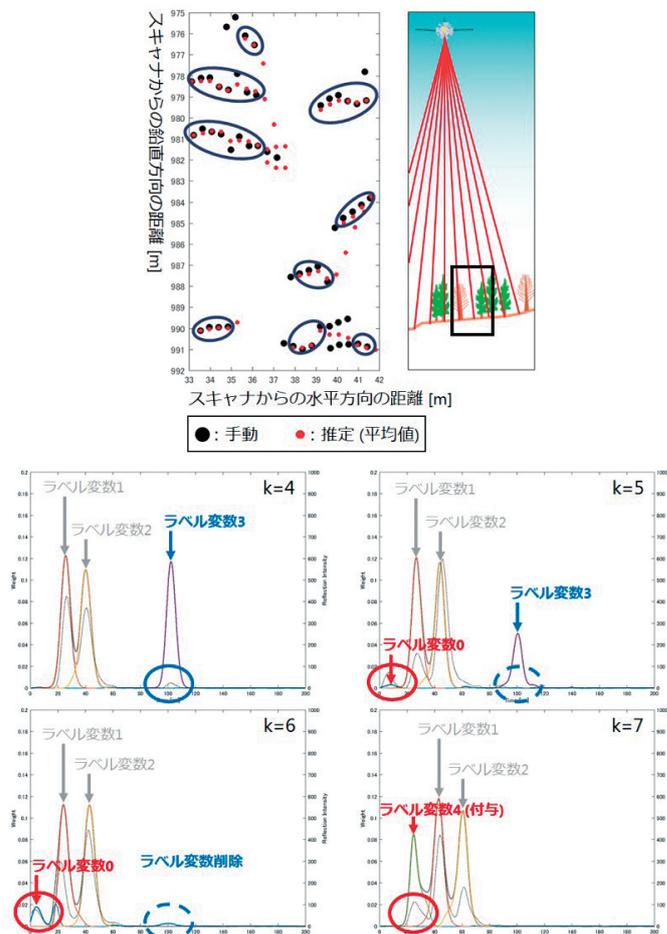


図4 波形記録式レーザの物体数・位置推定
(上: 推定結果のプロット、下: 推定結果の時間推移)

があったということになりますが、この $k = 4$ から中央部の反射強度が小さくなっていきます。 $k = 5$ でほとんど反射強度がなくなり、少し時間差がありますが、この地物は消失したものとしてラベルが消えるということが自動的に行われています。一方で左のほうでは、 $k = 5$ のところから少し反射強度が強くなってきていますが、 $k = 7$ のところで新たに物体が出現したということも自動的に捉えられています。

今度はまた対象を変えまして、画像からの人物追跡の話をしていきます。今回利用したセンサはステレオビデオカメラです。これによって色情報と距離情報というデータを統合しようということになります。一方でシミュレーションの方は、最近ではマイクロな歩行者の挙動モデルなどがいろいろと提案されています。次の一歩をどこに踏み出すかという非常にマイクロなモデルです。これは、交通分野で開発されているものです。この両方を使いまして、群衆の中でオクルージョンがあった場合に対して、色や距離を使って可能な限り認識しようということと、シミュレーションの挙動モデルを使って、これをデータ同化によって統合して、人物追跡を行おうというものになります。今回推定したい状態は、人物の位置と、人物を楕円体で近似した形状になります。観測データの方はステレオカメラから得られる距離画像の色と三次元座標です。システムモデルは先ほどのシミュレーションを使いまして、人物モデルも色の類似度と楕円体の形状の類似度を使います。

この歩行者挙動モデルは、離散選択モデルというモデルですが、次の一歩にどのマスを選ぶかという問題を考えます。そのときにどこが効用が高いか、満足度が高いかの確率をマスにつけていくことになります。どういうときに確率が高いのかというと、進行方向を維持する、速度を変えないなどが普通使われます。このモデルの良いところは、歩行者間の相互作用も考

慮できることです。先に歩いている人がいましたらついて行く、あるいはすれ違うときにはその人たちを回避するということまでもきちんと定式化しているというものです。

ある駅の朝ラッシュ時の改札部分のところをステレオビデオカメラで撮りました。色の情報と距離の情報がそれぞれ得られています。こちらが適用結果になります。この赤い点が実際に自動認識された人の腰の辺り、楕円体の中心の位置になります。この番号は同一人物であることを追っています。かなり複雑な部分も追えている結果が出ています。特に、急な方向転換や遮蔽近接がある場合にも精度が出ているというのがこの結果です。ちなみに今回はシステムモデルとして先ほどのシミュレーションを使いましたが、単にノイズのみ、いわゆる千鳥足だった場合のモデルを入れてみますと、最初のほうから人が追えなくなっていますので、このシミュレーションの効果が出ているということが確認できます。認識率としても80%弱であり、これだけの混雑の中で80%という追跡精度を達成しております。

最後の例になりますが、もう少し広い範囲での人の動きを見たいと思います。都市レベルでの人口分布からの異常検知です。これはまだパイロットスタディレベルです。途中段階ですが、ご紹介したいと思います。ベースになるのは

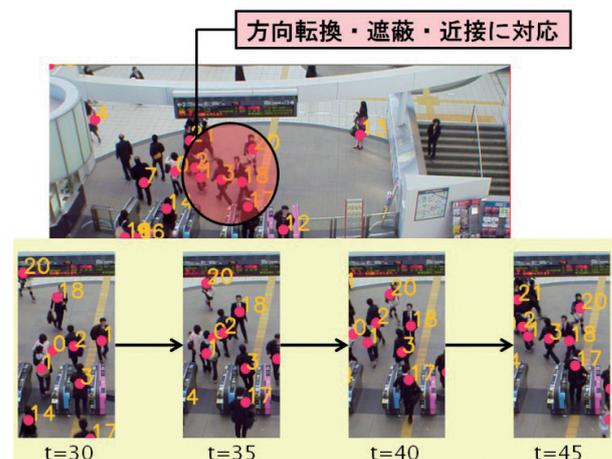


図5 群集における人物追跡

GPS のログデータです。当然、個人情報保護の意識の高まりから、メッシュデータに変換して公開するという流れになっておりまして、このメッシュデータをどう使っていくかという話になります。各地域のメッシュ人口データが、時系列で得られるため、その変化を見ることができます。しかし、それがいつも通り多いのか、いつも通り少ないのか、あるいはいつも通り増加しているのかということにはわかりません。あるいは、普段と異なって人口が多いのか、普段と異なって人口が減少しているのか、などを判別したいという問題を考えます。今知りたいのは、普段通りの状態で増加して多いのか、あるいは普段と違った状態なのかということを離散的にこの状態ベクトル x に当てはめます。つまり、離散的な状態になるということです。

観測値のほうはメッシュの人口データになります。メッシュの人口データは、よくわからない背景にある隠れ変数から結果として得られたものだと考えます。これも全く同じデータ同化の枠組みですが、状態変数が離散的になると、いわゆる隠れマルコフモデルと呼ばれるモデルになります。これは自然言語処理でかなり使われていて、発展してきた手法です。パイロットスタディですから、シミュレーションでどれくらいの精度が出るのかということを試しました。実際のデータを真似て、通常の一つのメッシュの人口の変動データを作成しました。横軸が時間で縦軸がその1メッシュの人口を表しています。実際のデータを参考にデータを作り、ノイズを付加したものを学習データとして学習させます。これが正常だということを学習します。それに対して、検証データとしては、異常値を付加しています。これを自動的に検出できるかという問題になります。離散的に表したどの状態かということ、番号で示しております。学習したモデルに対して検証データを当てはめれば、どの状態になるのかということがわかるため、状態の

番号が違うところが異常として検出可能になるわけです。実際に異常値を付加したところは90%で検出はできています。ただし、異常ではないところでもまだ検出がされているという状況が続いていまして、これを今後改良していきたいと思っています。

5. おわりに

それでは、時間も近づいてまいりましたので、最後にまとめたいと思います。本日はいろいろなデータを組み合わせる、データとモデルあるいは知識を組み合わせるという話でした。今後いろいろなデータの取得可能性があるとはいえ、やはりこの多様なデータ取得手法の整理ということは重要だと思います。それがないと、統合によってどう効率化できるのか、精緻化できるのかということにはわかりません。そのためにもデータ特性の解釈として、各データの有用性や限界、さらには代替可能性を整理する必要があります。たとえば、これまでのパーソントリップ調査がGPSで代替できるのかどうか、バスのトリップだけ見るにはどうすればいいのか、そういう話になってくるかと思います。そこでは解釈するための解析も関係します。その解析手法自体の理解も重要になってきます。その上で、データの統合手法の検討が可能になると思います。そのためにも、インフラとしてデータをどんどん蓄積していくべきでありますし、これをどう使っていくのかということも重要になってくると思います。さらに、データと知識、あるいは今回のシミュレーションなどを統合することができます。分析した結果が、今度は別な情報となって蓄積されていけば、当然その分野の知見として蓄えていくことができるわけです。それをまたさらに解析技術にフィードバックしてあげることが考えられます。

これらを支えることを目標にして、社会基盤情報学という分野があるべきではないかと考

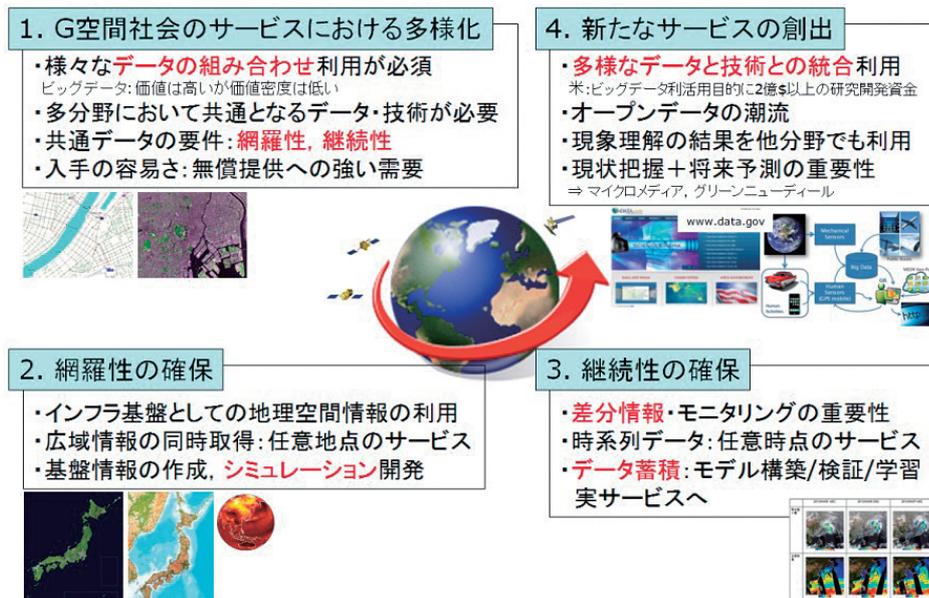


図6 社会基盤情報学へ向けて

えています。特に、計測からデータを蓄積する、またはデータベースを構築する、さらにはそれを利活用する、そして、すべてのフェーズで解析技術というものが関係します。これら全体を担える分野は、やはり空間情報の分野だと思っています。まずはデータの組み合わせから始めて、さらにサービス展開する上で、当然のことながらインフラという意味でも、網羅性や継続性を確保するということが重要になってきます。それぞれシミュレーションやデータの蓄積も重要になってきて、最後に多様なデータと技術、これは知見でも知識でも構いませんが、それらを統合して、スパイラルアップして発展していければと思っています。

今後、空間情報学をベースにした社会基盤の分野での情報学の構築、これに関していろいろと皆様からもご支援をいただければと思います。時間になりますので、私のお話はここまでにさせていただきます。どうぞご静聴ありがとうございました。

講演者
布施 孝志 (ふせ たかし)
 東京大学大学院工学系研究科
 社会基盤学専攻 准教授

本稿は平成 28 年 7 月 1 日、当協会の第 38 回測量調査技術発表会において特別講演をいただいた布施准教授のご講演内容を本誌編集委員会にてまとめたものです。