

## (2) 特別研究課題「A I を活用した地物自動抽出に関する研究」終了時評価

○委員長 それでは、議事の2番目でございますけれども、特別研究課題「A I を活用した地物自動抽出に関する研究」について、国土地理院から御説明をお願いいたします。

○発表者 私のほうから、平成30年度から5か年にわたって実施してまいりました「A I を活用した地物自動抽出に関する研究」の成果につきまして御説明いたします。

研究予算は5か年で約5570万円でありました。

私の説明の内容としましては、最初に概要、それから、この研究は中間評価をいただいているのですが、その中間評価時点からの研究環境の変化、それから結論に関してまず簡単に御説明をした上で、研究の成果について御説明するという流れで行いたいと思います。

まず、研究の背景と目的ですが、地図の作成方法の一つに空中写真測量という手法があります。この手法では、まず空中写真の撮影を行いまして、その写真を地図と位置合わせをする標定という工程を経て、測量技術者が地物を判読しながら図化という行為を行います。そして、編集ということを加えて測量成果が出来上がるというのが、空中写真測量で地図を作成する場合の一般的な工程の流れとなっています。

2000年代以降、この測量に使う機材やセンサーというのはデジタル化されていますが、この工程の流れ自体に大きな変化はなく、依然として労働集約型の産業であるという点がこの測量方法の課題となっています。

一方、A I が様々な分野で当たり前に使われるようになってきました。本研究が開始された平成30年度というのは、ちょうどA I の第3次ブームというものの真ただ中にありまして、地図作成にもA I の技術を導入することで図化の工程にブレークスルーをもたらすことを目指して本研究に着手したものであります。

本研究が取り組む図化というのは、公共測量作業規程の準則の中でこのように定義されていますが、空中写真と、それから同時調整等で得られた成果、これは本研究ではカメラパラメータと言っておりまして、写真が空中のどの位置からどの向きを撮影しているかということを表す情報のことなのですが、この2つを入力して地物などの座標値をデータとして記録する作業を言うと定義されています。

この図化結果がそのまま地図データになるというわけではなくて、その一歩手前のデー

タを取得する作業という位置づけなのですが、これまでは測量技術者が立体視をすることができる図化ソフトという、ちょっと特殊なソフトウェアを使って建物や道路などの位置を一つ一つ取得していました。本研究は、これを自動的にできるようにしようという非常にシンプルな研究であります。シンプルなのですが、国土地理院では半世紀近く前から自動図化の問題に取り組んできていますが、成功はしていませんでした。

そのような図化作業に対してA Iでブレークスルーをもたらそうという本研究では、通常の特別研究よりも高い数値目標、具体的には40項目以上の地物に対して、位置精度は基盤地図情報として耐え得る位置精度で、人が判読した場合に匹敵する抽出成功率を達成すること。また、同時に、その目標を達成した地物が抽出対象となる地物数全体の80%程度の抽出という目標を掲げて取り組んでまいりました。

令和3年3月1日に中間評価をいただきまして、その時点ではこのような数値目標の達成状況となっており、残り2年間はこの数値目標①の項目数の増加を図るという計画を立てておりました。

しかし、その中間評価をいただいた直後に、研究官2名体制で実施する計画であったところ、人事異動によりまして研究官は私のみで研究を実行しなければならなくなり、40項目という数値目標の達成は到底困難となってしまいました。そのため、この40項目という数値目標にこだわるのではなくて、本研究の本来の目的である図化工程のブレークスルーを実現しようという方針に切り替え、残り2年間の研究を実施してまいりました。

ここで数値目標の達成を断念したのは40という数値のみでありまして、その他の位置精度、それから、人が判読した場合に匹敵する抽出成功率については当初目標どおり達成を目指すということにしまして、その結果、項目によって性能にばらつきはあるのですが、人が介在しない自動図化というものを実現することができましたので、以降御説明をいたします。

こちらのスライドは本研究の流れを示しています。図化の部分をA Iで実施する場合の工程を細分化したものがピンク色の矢印の部分です。この種の研究では、あらかじめ地図と重なるように加工されたオルソ画像をA Iに入力するというのが一般的な研究アプローチなのですが、オルソ画像を用いた図化では抽出漏れが多くなる傾向がありますし、空中写真測量にそもそもオルソ画像というのは使用できませんので、本研究では空中写真の生画像を直接A Iに入力するという手法を取っています。これは本研究の特徴の一つとなっています。その画像から目的の地物を抽出するA Iを開発するという部分が1つ目の研究

要素。そして、A Iによって抽出されたものをどんどん重ねてマージしていくのですが、それによって抽出漏れの低減を図るという手法が本研究の2つ目の特徴的な研究要素となっています。

説明時間が限られておりますので、本研究におけるクリティカルなこの2つの研究要素に絞って御説明をいたします。

まず、1つ目の研究要素であります地物抽出用A Iの開発について御説明をいたします。地物抽出用A Iというのは畳み込みニューラルネットワーク——以降、CNNと略します——と深層学習を用いまして、特定の土地利用の部分を色分けしようというものです。本研究ではシングルカテゴリーで、つまり1つのA Iでは、ターゲットとする地物1種類の土地利用の部分のみを色づけするという手法を取っています。これはターゲットとして道路を抽出するA Iの例ですけれども、18層のCNNを学習させたA Iを使いますと、現状、この右上の図のように、道路が写っている画素が赤色に塗り潰された推論画像というものが得られます。推論画像のところに $F=0.801$ という数字が書いてありますが、これは抽出性能を定量的に評価する数値でありまして、満点は1.0ですので、100点満点中80点程度の抽出性能であるということを示しています。

地物抽出用A Iというのは、適切な学習データを使って学習させることで性能を高めることができます。学習データには、対象の地物をラベリングしたTRUE画像と、それから誤抽出を避けるために対象の地物と間違えやすいものをそのままラベリング画像としたFALSE画像の2種類があります。このグラフは、道路に関して学習データが一定数たまるごとに学習をさせて、その時点の抽出性能の推移を示したものです。赤い線のところがF値でありまして、この値の変化を見ていただきたいのですが、学習データが700枚から900枚近辺でF値の傾きが変わっています。このような変曲点というのはほかの多くの項目でも見られているのですが、どうやらこの変曲点のあたりまでというのは学習データの質と関係なく、抽出性能は高まってくるようなのです。

なので、道路の場合で言いますと、Fが0.70から0.75程度というのは、いかなる手法の研究であっても容易に達成可能なレベルであると見ています。しかし、それ以降の0.75から0.80というレベルを実現するには、それなりに工夫をしないと性能は高まっていかないというようでありました。

そのようにして学習データを作成しながら、様々な項目のA Iを構築してまいりました。これは、それらの例であります。普通鉄道とか駐車場なども、目標としては $F=0.8$ という

のを定めていて目標値には到達はしていませんが、良い抽出性能を発揮するモデルが構築できています。F 値が0.500以上となっている項目の推論結果の例とスコアの推移というのは、数十ページあるのですが、参考資料のほうに掲載しております。

自然植生にも取り組んでおります。竹林とか針葉樹林あたりは、F 値としては0.6クラスなのですが、人間により判読された地図データとしての植生記号よりも正しく抽出されているのではないと思われるケースが多く見られます。特に竹林については、地図記号としては表示されていないことが多くて、空中写真から推論をしてみると、実際はかなりの面積の竹林が抽出されてきます。

こちらが本研究で取り組んだ地物抽出用 A I の項目と抽出性能 F 値の一覧であります。この2年間の研究期間の大部分はこの学習データ作成に費やしたという状況でありました。

本研究の特徴として、多くの研究者が対象とするような主要地物以外のマイナーな地物を広く対象に含んでいるという点が挙げられます。

こちらには途中断念分も含めて記載をしています。断念した地物の理由は、リソース上の問題によるものと、そもそも本研究の手法が適合しなかったというものの2種類がありました。砂礫地については学習データを積み増したのですが、中間評価時から性能を伸ばすことはできませんでした。本来の数値目標であった F 0.8以上というものをクリアしたのは、この項目の数え方では中間評価時点の3項目が9項目に、準則の標準図式ベースで5項目が12項目となっています。

次に、本研究の独自性の強い2つ目の研究要素であります A I 単独の抽出結果の抽出漏れをさらに低減させる手法について御説明いたします。

これは中間評価時のスライドから持ってきたものなのですが、空中写真というのは、60%ぐらいの範囲を重複させることでステレオモデルが構築できるように撮影をされています。その重複範囲を利用して、同じ場所が写っている何枚かの写真から A I で抽出した結果を重ね合わせますと、1枚の写真では抽出漏れとなってしまった部分でも、ほかの写真の抽出結果で補完をすることで、全体として抽出漏れが軽減できるのではないかとというアイデアでしたが、その実現のための実装を行ってまいりました。通常、1か所が2~3枚の写真には写っていますし、それから、撮影方法によってはさらに多くの写真に写っていることもありますので、それを利用して抽出漏れを減らそうという本研究独特の手法です。

スライド14に示した空中写真を例に具体的に説明いたします。このスライドはアニメーションを含んでいますので、できれば画面のほうで御覧いただければと思います。

1枚の空中写真に写っているのは、この画面中央の赤枠に写っている範囲であります。しかし、実際は、これの周囲にこれと重なり合うように、このように8枚の写真が撮影されています。この図の灰色が濃い部分ほど多くの空中写真に撮影されていまして、丸数字が重複して撮影されている写真の枚数を表しています。赤い範囲である目的の写真の部分ではほとんどの部分で5枚以上、多いところでは6枚の写真に撮影されているということが分かります。

単画像のAIの抽出結果をこのようにどんどん重ねていけば、この枚数分だけ抽出漏れを低減させられるチャンスが生まれるということになります。

この赤い範囲のAIの抽出結果を実際に重ねてみるとどうなるのかというのを見てみたいと思います。

まず、写真1枚だけから道路を図化した結果だと、このような結果になります。中央の紫色の線の部分は、建設中の北陸新幹線の高架橋の位置を示しています。特に赤丸の部分、図化されてない、白く抜けたままの道路が目立つ部分がありますが、これからこの場所が写っている隣の写真5枚分のAIの図化結果を次々に重ねていきます。このように、1枚の写真からだけでは漏れてしまう部分でも、この手法によって幾分改善できるということが分かります。

こちらは全ての研究成果を適用して、空中写真から道路を自動図化させた結果です。黒い枠線の範囲が先ほどの空中写真1枚分の範囲に当たりますが、隣接する写真を含めて9枚の空中写真の図化結果を全てマージしています。また、この程度の範囲の図化であれば、1時間足らずで結果を得ることができます。もちろん人が介在せずに自動的に得られた図化結果であります。

今回、後処理部分の説明は割愛いたしましたが、実際は自動図化の流れの中でノイズを除去したり、あるいは面積を閾値にして誤抽出を削除したりというようなことを地物ごとに後処理として工夫しています。例えば太陽光発電設備ですと、AIでは個々のソーラーパネルが抽出されてくるのですが、地図記号としては、個々のパネルではなくてメガソーラーとしての範囲を図化結果として欲しいわけなので、抽出された画素を膨張させるという後処理を行うことでパネル間の空間を埋めて、メガソーラーとしての一体となる範囲に変換するといった処理を加えています。表示しているものはメガソーラーとしての範囲を

示したものになります。それらの後処理も含めることによって、自動的に各地物の図化結果を得ることができるようになり、最終的に自動図化機構として完成させることができました。

本研究では、研究成果の社会実装を意識して研究してまいりましたので、この自動化機構というのは、空中写真画像とカメラパラメータを持ってきてダブルクリックを一発するだけ。すなわち判読等の技術を全く持たない人でも使えるようにしていますし、これ以外に災害用のリアルタイム図化システムというのがあるのですが、これでは撮影開始とともに自動的に自動図化が始まる仕組みというのも実装して、そもそも人間が必要ないというものも出来上がっています。

自動図化機構というのは災害対応との親和性が高いのも特徴です。空中写真数枚程度であれば、あっという間に図化してしまいます。これは令和3年に発生した熱海の土砂災害の際に被災建物を特定できるかどうかというのを試行してみた事例です。実際に災害を撮影した写真を入手してから40分で消失建物を推定することができました。この技術は被害判読に人員をほとんど必要としないため、例えば南海トラフ地震のような大規模な災害の際でも少ない人員で大きな効果を発揮できると思われれます。

次に、社会実装に向けた動きですけれども、本研究で構築したAIの学習データは国土地理院の地図記号には含まれないものもありまして、また、深層学習用の教師データとして貴重なものでもあると思いますので、抽出性能が一定以上となった項目の学習データは「GSIデータセット」という名前で国土地理院のホームページから公表を進めています。

本研究で使用した学習データは、基本的に公表した項目のものは全て公開しておりますので、本資料に記載している程度の自動図化であれば、このデータを学習させることで実現できると思います。公開したデータは出典明示していただければ商用利用も可能であります。

本来はそれぞれを論文として発表したかったところなのですが、リソースを極限まで減らされてしまった中で、まず自動図化を実現することというのを優先したため、現状では論文までは手が回りませんでした。

また、本研究の成果は、現状から言うと大分突き抜けた成果を含んでいます。(ハイパーパラメータは公開していないので) 学習データだけ公開しても、本研究で実現した抽出性能を再現できない可能性もありますので、推論に使ったプログラムやツール、それから、実際に目標性能を達成した学習済みモデルそのもの、あと自動図化機構についても技術資

料として公開する方向で準備中です。これらを実際に使っていただければ、F=0.8クラスの自動図化というのは、お手元でと言うと環境にもよる（多数の空中写真を用いた自動図化を行うにはそれなりの性能を持ったGPUやCPUを有するPCを使用する必要があり、通常のPCで自動図化を行うのは処理時間を要して現実的ではない）ので難しいかもしれませんが、基本的に国土地理院以外でも自動図化が行えるようになるはずです。

本研究の最終的な目標達成状況を自己評価したものがこちらとなっています。数値目標につきましては、リソースの問題から40項目という数字は達成できませんでしたけれども、位置精度としては、全ての地物で2500レベルの自動図化を実施できるようになりました。また、本研究が対象とする地物数全体に対する抽出シェアというのは79.15%でありまして、この数字には太陽光発電設備の数が含まれていないのですが、それを含まないと目標は達成できたのではないかと考えています。

最後にまとめることを申し上げますと、本研究の目的であります図化工程のブレークスルーに関しては、ブレークスルーそのものというのは、これから社会実装が進んだ結果、そういう評価が得られるかどうかということです。現時点でブレークスルーを果たしたとは申しませんが、測量技術者でないとできないと思われていたことをAIによって自動的に行える、つまり自動図化ができるようにしたということで、その達成への道筋をつけられたのではないかと考えています。本研究で我々としては長年夢物語であった自動図化というのを実現することができました。これによって、これからの地図の作り方と人との関係というのを根本的に変えていく嚆矢になったのではないかと考えています。

測量成果としての自動図化まで持ち込むにはもう一步研究が必要ではありますが、AIの研究はこれで一区切りをつけたいと思います。今後は事業部に研究成果を引き継ぎ、技術開発という位置づけで教師データの作成が続けられる予定です。

私からの説明は以上です。

○委員長 御説明ありがとうございました。

次に、資料2-1の終了時評価表について、事務局から御説明をお願いいたします。

○研究企画官 では、事務局から資料2-1、終了時評価表の説明をいたします。

終了時評価表でございますけれども、1番から7番までは今の研究概要の説明と内容が重複いたしますので、そこは省略させていただきます。

8番、必要性、効率性、有効性等の観点からの評価というところにつきまして、内容をかいつまんで御紹介したいと思います。

まず、必要性の観点からの評価でございますけれども、空中写真測量において、図化の工程というのは、従来は判断スキルを有する測量技術者が実施する必要があったものでございまして、自動化というものは非常に困難なものと思われてきたところでございます。AI技術を用いた試みも幾つかありましたけれども、あまり自動図化までたどり着いたものはなかったかという話をさせていただいたかと思えます。

その中でこの研究につきましては、AI技術のみではなく、いわゆる従来技術も使うことによって、ある一定程度の目的地物の自動化を実装することができたという成果を得られております。その成果に関しましては、今後の技術的な発展の端緒をつかんだという意味で、この研究の必要性を示すことができたというところを評価をしているところでございます。

効率性の観点からの評価でございます。本研究につきましては、効率性を上げるために幾つかの試みをしてございます。幾つか列挙させていただきますけれども、簡単に御紹介いたしますと、まず1点目は、効率的に行うために夜間にコンピューターの資源を利用するし、自動的な処理というものについては、夜間を利用して効率的に進めたというところでございます。

また、学習モデルの初期段階において現れる性能曲線の変曲点の出現位置及び性能曲線の傾きから対象項目の絞り込みを実施して、効果が出やすい項目にリソースを集中するというを行いました。

また、自動図化がある程度稼働した時点ですけれども、自動図化結果から誤抽出や抽出漏れの箇所を発見できるようになったということを利用して、やはり対象画像の選定について、さらなる効率化を図ったということを行ってございます。

最終的に目標達成の見込みがある項目に対象を絞り込みをするということで、リソースを集中して成果を効率的に得られるように工夫したということもございます。

また、資金面でございますけれども、深層学習用のPCを、既製品を購入するのではなくて組立てをするということで、研究費を効率的に執行するというような工夫もしてきたところでございまして、効率性の面からも効率的に執行ができた、研究を進めることができたというところを評価をしているところであります。

有効性の観点からの評価でございます。有効性につきましても、やはり何といたっても自動図化が実現できたということは有効性が非常に大きいことを示すことができたと考えて

いるところであります。今後の事業を展開するに当たってのまず第一歩を示したと思っております。

成果の有効性という意味では、いろいろなところに波及効果もあるだろうと考えています。ただ、測量成果ほどの精度が上げられないような図化の部分もありますので、そういったところに関しましては、研究概要の説明の中でもありましたけれども、災害対応用の自動図化というような分野で実用化のレベルには達しておりますので、そういったところでの活用というものが広がれば有効性はさらに大きくなるものと考えているところであります。

9番に関しましては、概要説明の中で紹介いたしましたので省略をさせていただきます。

最後に、地理分科会におきまして総合評価としてコメントいただいておりますので、私のほうから紹介をさせていただきます。

地理分科会では、十分に目標を達成できたという評価をいただいております。コメントといたしまして読み上げさせていただきます。数値目標等は途中で変更したが、変更された目標に対して十分に達成できたと判断した。

分科会のコメントとしては以下の3点である。1、今回の研究は非常にレベルの高い研究内容と認める。2、今後、この内容を企業や大学などと共同で研究を進めることを期待する。3、本研究の成果は高く評価できるが、プロジェクトの途中で人事異動があり、研究体制が縮小されたことは残念である。

以上でございます。

○委員長 御説明ありがとうございました。

ただいまの事務局からの説明について、地理分科会の主査から何かコメントございますでしょうか。よろしく願いいたします。

○委員 ただいま御紹介いただきましたとおり、途中で体制の変更があったにもかかわらず、大変高い研究成果を上げられたと私どもは判断いたしました。

○委員 特にはございません。むしろ、こんな途中の体制の変化にもかかわらず、最後のスライドとスライド6番目、私にとっては非常に印象的だったのですが、御苦労されて、ここまで御研究をされたということが非常に高く評価できるところではないかと思えます。

以上です。

○委員 私のほうからも、大変興味深い研究であったと思います。特に付け加えることはありません。

○委員長 ありがとうございます。

それでは、委員の先生方から御質問や御意見をいただきたいと思っておりますけれども、お願いいたします。

○委員 大変興味深い研究であったと思います。聞きたいことはたくさんあるのですが、1人で時間を使うわけにいかないのです、教師データについて2つぐらいお伺いしたいんです。当初掲げられていた40種類の地物の中でうまく判別できないものもあったということですが、この教師データを時間をかけて作りさえすれば、今は判別できないけれども、教師データさえよくなれば、どんな地物でも理論的には判別できる見込みということでしょうか。

○発表者 御質問ありがとうございます。実は両面ありまして、まず、特徴量がきちんと判別できるものに関しては、適切に学習データを追加していくことができれば、適切にというのは枚数という意味ではなくて、スライド9を写していただけますか。

これで変曲点という話をいたしましたけれども、700枚から900枚付近に変曲点があるわけですが、実は何も考えずに枚数だけ投下しても性能というのは上がらなくて、この変曲点のあたりに結局戻ってきてしまうみたいなのです。なので、性能を上げたいのであれば、やはり投入する学習データの質というのは考えないといけない。我々は結局推論を繰り返して、うまく出てないところ、あるいは出過ぎたところを反証するものを投入したり、あるいは、実はFALSEとTRUEの比率というのが結構効いてくるということと、TRUEを入れるについても、純粋なTRUEのところのみではなくて、できるだけ誤抽出が出ているところを含むTRUEを入れる。つまりTRUEだけでも、FALSEと同じ働きをする部分も加えるというようなところを工夫しながらでないとも性能としては上がってこない。

ただ、逆に言えば、そういうことを注意しながら枚数を増やしていけば、特徴量が明確なものについては、全ての地物について、Fの傾きはまだ上方に向いていますので、限界がどのぐらいまでいくのかというのは、そこまでは到達できませんでしたので、そこについては申し上げることはできないのですが、少なくともうまく質のよい学習データを追加

すれば性能は向上できると思います。

ただ一方で、特徴量が明確でない地物というのがあります。評価表のほうには実例を幾つか書いたのですけれども、例えば畑に関しては、キャベツ畑もあれば、落花生畑もあれば、いろんな畑が混ざっています。特徴量が非常に混ってきているものというのは、枚数とか、単純にデータの質ということではどうも向上しないようで、定義自体を分けるというようなことを適切に考えていかないと、多分、よい性能のものには、何枚学習データを作成したとしてもたどり着かないのではないかなというのがこの5年間で到達した結論になります。

○委員 ありがとうございます。作成した教師データが、恐らく日本国内の写真に応用しようと思っているのだらうと思うのですけれども、それは、例えば大きく地質状況の違う海外にも応用できるのでしょうか。

というのは、今回のトルコ地震の件を想定しているのですけれども、今学習させたものをトルコに応用することによって、例えば被害抽出などが素早くできるだろうかという疑問があるわけです。

○発表者 私も当然考えました。ただ、今回構築した学習データがターゲットとしたものは、地上部画素寸法20cmクラスで撮影された、かつ測量用のデジタルカメラの画像というのをターゲットにしています。実は地上画素寸法のサイズというのは性能に大きく効いてくるものですから、海外のそういった事例のケースで、こういった航空測量に用いるようなデジタルカメラの画像というのはなかなか入手できないので、通常であれば地上画素寸法1mクラス。もし「だいち」3号がうまく打ち上がっていたのであれば80cmクラスの画像が得られるわけなのですけれども、その画像に対して今回構築した学習データを使ってモデルをつくっても、恐らくよい性能は出ないと思います。

特に空中写真用のデジタルカメラの色の合成の仕方というのと、それから、衛星はパンシャープンで色を乗せていくのですけれども、そのときの色の乗り方というのが違うのですね。実はデジタルカメラとアナログカメラをスキャンしたもので違うのです。これは私も一目で大体見分けつくぐらいの違いがあるので、そういった部分が結構よく効いてくるというか、影響が大きいものですから、もし海外というものをターゲットにするのであれば、やはりそれなりのターゲットとする衛星、あとターゲットする地上画素寸法のものを作成しないといけないのかなと。

○委員 ありがとうございます。

○委員長 それでは、ほかの先生方いかがでしょうか。

○委員 私、専門外なのですけれども、楽しく勉強させてもらいました。ありがとうございます。

シンプルな質問が1つあるのですけれども、今後、航空写真の撮り方に対して、この研究がどういうインパクトがあるのか。今までの航空写真の撮り方は、恐らく人を見ることを前提に撮っていたと思うのですけれども、今後、変化を期待したいかということと、どういう点なのかということをお教えください。

○発表者 御質問ありがとうございます。ちょっと事業面にも絡むのであんまり勝手なこととは言えないのですけれども、この研究の成果をうまく使うための撮影方法という観点から言えば、オブリークカメラというのがあります。通常のカメラは真下のみを撮るカメラになっているのですけれども、オブリークカメラというのは、それに加えて4方向を傾いて撮影するカメラというのを飛行機につけて一緒に撮影しようというカメラなのです。そういったもので、真上からではない、横の方向からの撮影した画像というのを入手できるようにすると、この研究の特徴の一つが、1枚だけの写真からの抽出結果だけでは不十分なので、別のいろんな方向から撮影された写真から抽出した結果をどんどんマージしていくことで性能をもっとよくしようというアプローチで行っていますので、その場合に、真上からではない方向というのが幾らかでも多く撮影できる写真のほうがやはり有利です。

衛星に広げるときも、衛星というのは基本的に真下は撮らなくて斜め方向を撮るのですが、「だいち3号」に対して、ちょっと残念ではあるのですが、本来であれば、「だいち1号」が持っていた3方向視の撮影方法で撮影してくれると、この方法は非常によくはまったかもしれないです。そういう部分があります。

○委員 ありがとうございます。

○委員長 ほかの先生方、いかがでしょうか。

○委員 ありがとうございます。今後に向けて大きな成果なんじゃないかなと思って伺いました。

ちょっと個人的な興味で質問なのですが、先ほど御説明の中で竹林を例に挙げられていましたけれども、多分人が読んでも落としてしまうようなところもかなりのF値で拾っているなという印象です。水田などもそうだと思うのですが、植生などで季節によるばらつきみたいなものはあると思うのですが、その辺、教師データのバランスみたいなものは考えられていたのでしょうか。

○発表者 ありがとうございます。季節による影響は非常に大きいものがあります。当然水田ですと、夏場は稲が植わっているので緑色に青々としたものが見られますし、冬であれば刈取りが終わって切り株が並んでいるような状況がありますので、基本的に全ての季節のデータから学習データを作成しています。

どうやって作成したかという、恣意的に選ぶと偏ってしまうので、国土地理院がデジタルカメラで撮影した20cm級の画像というのが大体30万枚あるのですけれども、それを全てデータベース化しまして、ランダムに選んでくるというスクリプトをつくって、数多くの抽出したものを推論にかけて、そこから対象部分を選択していくということを全ての項目について実施しています。

なので、自然植生も、例えば広葉樹林はちょっとうまくいかなかったんですけども、夏場の葉っぱがついているとき、紅葉のとき、あと冬の落葉したときというものも全てデータとしては投入するようにしていますし、雪の影響というのも考えて、積雪の写真も何枚か入れているものがあります。そういったあらゆる条件をできるだけ考慮する形で作成しています。

○委員長 そのほか、よろしいでしょうか。

それでは、2番目の件については、これで終わりにさせていただきます、これから委員の先生方で講評について議論したいと思います。ちょっと遅れているということもございますので、このまま休憩を取らずに、それほど長い時間をかけずにやりたいと思いますので、隣の部屋にお移りいただければと思います。よろしく願いいたします。