

空中写真を用いた地物抽出性能評価用データセットの構築

Construction of Dataset for Feature Extraction Performance Evaluation using Aerial Photographs

大野 裕幸*¹
Hiroyuki Ohno

遠藤 涼*¹
Ryo Endo

中埜 貴元*¹
Takayuki Nakano

篠田 昌子*¹
Masako Shinoda

*¹ 国土地理院
Geospatial Information Authority of Japan

Several datasets are known as datasets for feature extraction. However, they are targeted to outside Japan and are not necessarily dataset with high regional diversity. Therefore, it is not suitable as an evaluation dataset for feature extraction in mapping performed as survey in Japan. So, we carried out this research with the aim of enabling the evaluation of feature extraction performance for high regional diversity and aerial photograph actually used in mapping in Japan. As a result, a dataset for evaluation was constructed using aerial photographs of 1328 regions taken in Japan since 1967. Furthermore, we compared the evaluation value of the prediction result using our dataset with that using the existing dataset using pix2pix and U-Net, and concluded that our dataset can perform sufficiently reliable evaluation.

1. はじめに

地図を作成する主要な手段の一つが写真測量である。写真測量では、空撮画像を元に、画像に写っている道路や建物等の地物の位置や形状等を読み取って地図データを作成する。そのような空撮画像から地物の位置と範囲を特定する工程を「判読」、判読した地物を地図上の座標を持ったデータにする工程を「図化」というが、これらの判読、図化の作業はこれまで測量技術者による手作業に頼らざるを得なかった。

一方で、CNN を始めとする画像認識技術が年々高度化しており、高分解能衛星の画像を用いた道路と建物抽出用のデータセット SpaceNet や、空撮画像を元に作成された建物抽出のベンチマーク用データセット Inria Aerial Image Labeling Dataset [Maggiori 17] (以下、「Inria データセット」という。)などが公表され、衛星画像を用いた地物抽出 (Feature Extraction) のコンテスト DeepGlobe 2018 [Demir 18] も実施されるなど、空撮画像から Feature Extraction を行う研究が盛んになりつつある。Feature Extraction は、測量の地図作成における判読及び図化の工程に類似していることから、これを測量の工程に導入して判読及び図化の一部でも自動化することができれば、地図作成の生産性向上に繋がるだけでなく、災害時の被害判読等の迅速化にも応用が可能と考えられる。そこで、筆者らは地上画素寸法 10cm~40cm の範囲で撮影された測量用の空撮画像を対象として、我が国における測量への導入を目標とした高性能な Feature Extraction の実現に取り組んでいる。

Feature Extraction の性能は、評価用のデータセットを用いて F 値や IoU などの数値によって定量的に評価することが一般的である。測量という実業務への導入を念頭に置いた場合、測量で用いられているデータと同じものを評価対象として、様々な季節、地域、気象条件の下で安定して均質な性能を発揮できるか否かを適切に評価しなければならない。そのためには、測量用に撮影された空撮画像から作成され、十分な多様性を有する評価用のデータセットが必要である。しかし、既存の評価用データセットは、日本国外の都市を対象としたものであるうえ、衛星画像から作成されたものが多く、地域多様性の観点からも十分とは言えない。そのため、まず、我が国を対象として十分な地域多様性を備え、測量用の空撮画像を対象とした Feature Extraction の性能評価用データセットを構築したので報告する。

2. 測量に用いられる空撮画像

測量に用いる空撮画像の撮影に関する仕様は、鳥瞰写真などとは異なり、要件が厳密に作業規程の準則[国土交通省 08]によって規定されている。そのため、測量に用いる空撮画像を一般的な空撮画像と区別して「空中写真」と記述することとし、まず初めに、空中写真と、既存の評価用データセットでよく用いられているオルソ画像について説明する。

2.1 空中写真の撮影方法

空中写真は、航空機に搭載された専用のカメラを用いて直下に向けて撮影される。かつては 23cm 幅のフィルムを使用したアナログカメラが用いられていたが、現在では全面的にデジタルカメラによる撮影に移行している。カメラパラメータは、焦点距離などの内部パラメータが事前に精密に検定されているうえ、航空機に搭載された GNSS (いわゆる GPS 等の衛星測位システム) と慣性計測装置 (IMU) によって、外部パラメータである写真撮影時のカメラの三次元位置とレンズ光軸の角度も極めて高精度に計測される。さらに、撮影後に画像とカメラパラメータを用いてバンドル法等で調整計算を行うことで測量の精度を保持するように設計されている。空中写真撮影用のデジタルカメラは主なもののだけでも 10 種類以上が使用されている。

2.2 空中写真の特性とオルソ画像

空中写真はカメラのレンズの中心に光束が集まる中心投影で撮影されるため、画像が写真の中心から周囲に向かって倒れ込むように写るという特性があり、土地の起伏に伴って歪みも生じるため、そのままでは地図とは重ならない。これに対し、中心投影による倒れ込みと、地形の起伏による歪みを取り除く画像処理 (正射変換) を行い、画像のどこであっても真上から見たような傾きの無い正しい大きさや位置に表示された画像に変換したものをオルソ画像という (図 1)。オルソ画像は、地図とぴったりと重ねることができる。そのため、オルソ画像を元画像とすれば、既往のポリゴン型の地図データを参照用のデータとして用いることが可能である。ただし、撮影年や撮影縮尺が異なるなど、参照用の地図データの作成に用いられた画像とは異なる写真から作成されたオルソ画像を用いた場合は、地図データと元画像に写っている内容の不一致が生じる場合がある。

一方で、空中写真では、中心投影であることを利用して左目用の画像と右目用の画像を用意し、簡易立体視鏡や立体視に

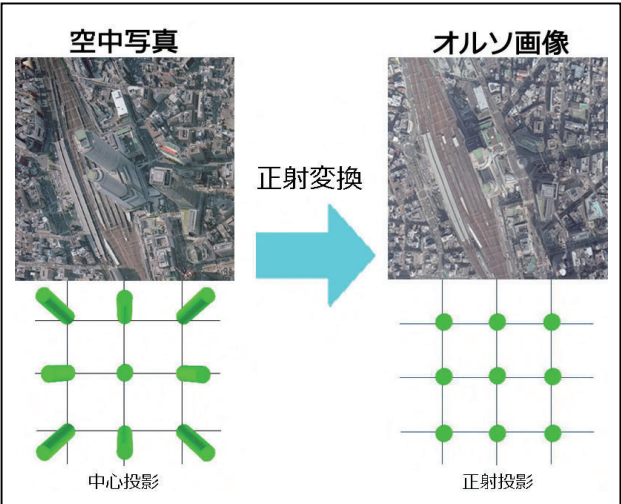


図 1. 空中写真とオルソ画像の違い

対応した眼鏡と図化ソフトを用いることで立体視をすることができる。元画像に対する適切な参照データが入手できない場合、立体視によって地物を判読し、使用する画像と完全に合致する高精度な参照データを作成することができる。筆者らは、立体視によって作成された参照データが評価用データセットとして最適であると考えている。

3. 多様性を持った評価用データセットの構築

空中写真から作成された既存の Feature Extraction 用データセットとしては Inria データセットが面積の観点から最大と思われる。Inria データセットは、地上画素寸法 30cm のオルソ画像を使用し、アメリカ 6 都市とオーストリア 4 都市の合計 10 都市 810 km² のデータが公開されており、5 都市 405 km² 分はトレーニング用データとして参照データとセットで入手することが可能である。

しかし、日本における測量に用いることを目標とした Feature Extraction のトレーニング／評価データという観点で見ると、既存のデータセットには次のような問題点がある。

- すべて日本国外の都市のデータであり、日本とは地勢や地物の写り方が異なる。
- 数～10 都市程度のデータであり、地域多様性が少ない。
- 使用されている地上画素寸法が、日本の測量で一般的に用いられるサイズと異なる。日本では、20cm が用いられることが多く、10cm も用いられ、災害時は 40cm で撮影される場合がある。
- 多くが「建物」又は「道路」カテゴリのデータのみである。

そのため、国土地理院が保有し、測量用カメラで撮影された空中写真を元に、日本における測量に適した Feature Extraction 性能評価用のデータセット(以下、「GSI データセット」という。)を構築することとした。

3.1 GSI データセットの仕様

GSI データセットは、地域多様性の確保の観点から、原則として国土地理院がこれまでに測量用カメラで撮影した全ての地区から均等に範囲を抽出して構築した。過去に撮影された空中写真に対応するポリゴン型の地図データは作成されていないことから、オルソ画像ではなく、空中写真の画像データをそのまま使用し、比較的歪みが少ない各空中写真の中央を中心とする範囲を切り出して元画像とした。アナログカメラで撮影された空中写真は、1200dpi 又は 1270dpi でスキャンされた画像を使用し、[大野 17]の手法により写真中心 が画像中心となるよう補正した

画像から、同様に写真の中央を中心とする範囲を切り出して元画像とした。

元画像の主な仕様は次のとおりである。

- 地区数 549 地区(デジタルカメラ)
779 地区(アナログカメラ)
- 1地区あたりの切り出し枚数 1 枚
- 1枚あたりの切り出し範囲 572×572pixel
- 地上画素寸法 10cm～40cm
(各空中写真の撮影縮尺に基づく)
- カテゴリ分類 道路、建物、水域等 36 種類
(表 1 に示す)

表 1. GSI データセットのカテゴリ分類

番号	RGB	カテゴリ	番号	RGB	カテゴリ
2	#7F7F7F	舗装道路	22	#FF7F00	果樹園
3	#BF7F3F	未舗装道路	23	#7F0000	荒地
4	#5F5F5F	駐車場	24	#3F3F3F	空地
5	#FF007F	分離帯	27	#BFBF00	太陽光発電設備
6	#3F0000	軌道	28	#BF00BF	輸送管
7	#0000FF	水域	36	#BF7FBF	雪覆い等
8	#FF0000	堅牢建物	38	#7FBFBF	プラントホーム
9	#FF00FF	普通建物	39	#FF7FBF	タンク
10	#FF7F7F	無壁建物	44	#FFBFBF	防波堤
11	#7F7F00	被覆	54	#BFFFBF	芝地
12	#BFBFBF	人工斜面	58	#FFBF00	砂礫地
13	#007FFF	水制	61	#BFBFFF	たたき
14	#7FBF00	歩道橋	62	#7F3F3F	岩がけ
15	#007F00	自然植生	65	#00007F	湿地
16	#FFFF00	畑	66	#BF00FF	材料置場
17	#00FFFF	水田	67	#7FBF7F	園庭
18	#007F7F	収穫後の水田	68	#7F7FFF	墓地
21	#7FFF7F	茶畑	255	#FFFFFF	未分類

各地区の写真は、写真枚数を n とした場合のファイル名順で n/2(小数部分切り捨て) 番目の画像を機械的に抽出した。そのため、全域が海域や森林という画像も一定数含まれる。

カテゴリは、まず道路(舗装道路、未舗装道路を合わせて 1 カテゴリとしたもの)について全データのラベル付けを完了しており、建物、水域、その他の順で順次追加中である。

各画像のラベル付けは、1328 枚のすべてを測量士で図化経験がある者が目視で判読及び確認することによって実施した。ラベル付けされた画像には、性能確認用の先行整備データとして位置付け、立体視を行わずに判読・分類したカテゴリ別画像と、精度よく性能を確認する正式データとして位置付け、すべて立体視により判読・分類した全カテゴリ画像の 2 種類がある。なお、全カテゴリ画像からあるカテゴリだけを抽出して作成したカテゴリ別画像は正式データとして位置づけられる。図 2 に元画像とカテゴリのラベル付けが完了した画像の例を示す。ラベル付けは、各カテゴリに表 1 に基づく Index 番号と塗分け色を定義を行い、全カテゴリ画像は、RGB と Index Color の両方で、カテゴリ別画像は RGB 画像で保存した。道路のカテゴリ別画像は、舗装道路と未舗装道路のカテゴリの分類色をいずれも RGB=#FF0000(赤色)としている。また、図 3 にデジタルカメラによる 549 地区分の元画像抽出位置の分布を示す。赤い点が抽出された画像の位置を示しており、北海道、秋田県、長野県、

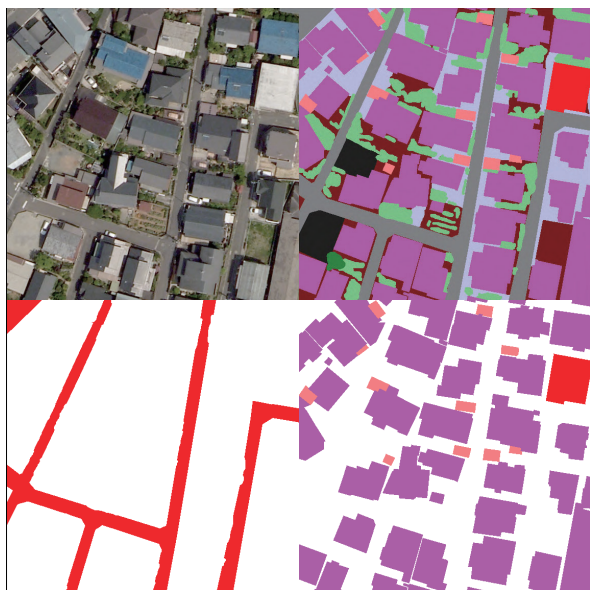


図 2. 元画像とラベル付け画像の例
(左上:元画像, 右上:全カテゴリラベル付け画像, 左下:
道路のカテゴリ別画像, 右下:建物のカテゴリ別画像)

岐阜県の山岳部を除き、ほぼ万遍なく抽出されていることが分かる。日本最東端の南鳥島は、デジタルカメラでの撮影が行われていない。三陸沿岸と熊本県にやや集中が見られる理由は、災害状況把握のための空中写真撮影が実施されたためである。

3.2 評価値の算出方法

カテゴリ別の性能評価は、F 値により行う。複数カテゴリの性能評価は IoU により行うことを原則とする。

4. Feature Extraction と性能評価の試行

4.1 使用したネットワーク

表現の異なる画像の相互変換に高い性能を示している pix2pix [Isola 17]と、pix2pix の Generator 部分に使用されている U-Net 構造の CNN の 2 種類を用いて Feature Extraction を実施した。いずれの入力画像、推論結果とも 256×256 pixel の画像である。推論では、入力画像をいったん縦横とも 256 の倍数になるよう線形内挿法でリサンプリングし、 256×256 pixel に分割したうえでネットワークに入力し、推論結果は分割画像をマージして最近隣法で元の画像サイズにリサンプリングしている。損失関数は、pix2pix は cGAN $\times 0.5 + L1$ 、U-Net は L1 とした。

4.2 トレーニングデータ

トレーニングデータは、GSI データセット構築に用いた空中写真をあらかじめ除外し、1967 年以降にカラーフィルム及びデジ

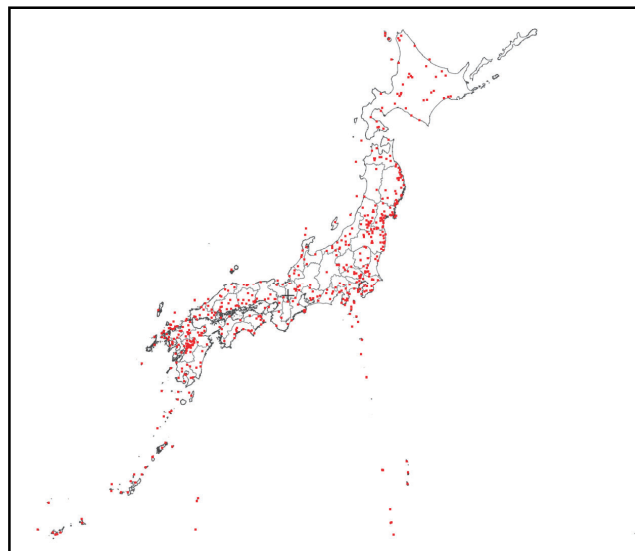


図 3. デジタルカメラ 549 地区分の抽出位置の分布

タルカメラで撮影された空中写真約 79 万枚の中から 286×286 pixel のサイズで、道路、建物 2 種類のそれぞれについて先行整備データと同じ判読方法で作成した。トレーニング時は、ランダムに 256×256 pixel のデータにクリッピングされる。作成したデータ数は、道路 3100 枚、建物 1700 枚である。さらに 90 度、180 度、270 度に回転させた画像を追加し、合計で道路 12400 枚、建物 6800 枚のトレーニングデータとした。

なお、トレーニングデータに Inria データセットの画像は一切使用していない。

4.3 性能評価の試行と考察

性能評価の試行は、GSI データセットによる性能評価値と、既存の評価用データセットによる性能評価値の傾向を比較することによって実施した。既存の評価用データセットには、Inria データセットでトレーニング用データセットとして提供されている 5 都市分の建物データを用いた。

Feature Extraction の推論結果を図 4 に、GSI データセットと Inria データセットで算出した建物の評価値を表 2 に、GSI データセットによる道路の評価値を表 3 に示す。なお、建物に関してはデータの整備途中であるため、デジタルカメラ 314 地区、アナログカメラ 633 地区による評価値である。また、GSI データセットの建物はコンクリート等の堅牢建物、普通建物、ビニールハウスやカーポートのような壁の無い無壁建物、建物以外の 4 カテゴリ分類としているが、建物と建物以外の 2 カテゴリの Inria データセットと仕様を揃えるため、堅牢建物、普通建物、無壁建物の 3 カテゴリの推論結果を建物 1 カテゴリとみなして評価した。

まず、pix2pix と U-Net による性能比較であるが、[Isola 17]はセマンティックセグメンテーションでは損失関数を cGAN よりも L1 とした方が良好な結果が得られるとしているように、本試行で

表 2. 建物抽出の評価結果

		GSI データセット			Inria データセット					
		Digital	Analog	All	Austin	Chicago	Kitsap	Tyrol-W	Vienna	All
Pix2pix	Precision	82.1	69.7	74.9	80.2	78.8	71.6	68.7	79.2	77.3
	Recall	67.4	44.7	53.0	45.8	47.7	56.6	71.2	43.3	48.0
	F 値	74.0	54.5	62.0	58.3	59.4	63.2	69.9	56.0	59.3
U-Net	Precision	86.6	74.9	79.5	83.1	79.4	75.0	73.9	79.6	80.8
	Recall	73.8	49.0	57.2	48.8	49.0	59.5	73.4	51.2	48.3
	F 値	79.7	59.2	66.5	61.5	60.6	66.3	73.7	62.3	60.5

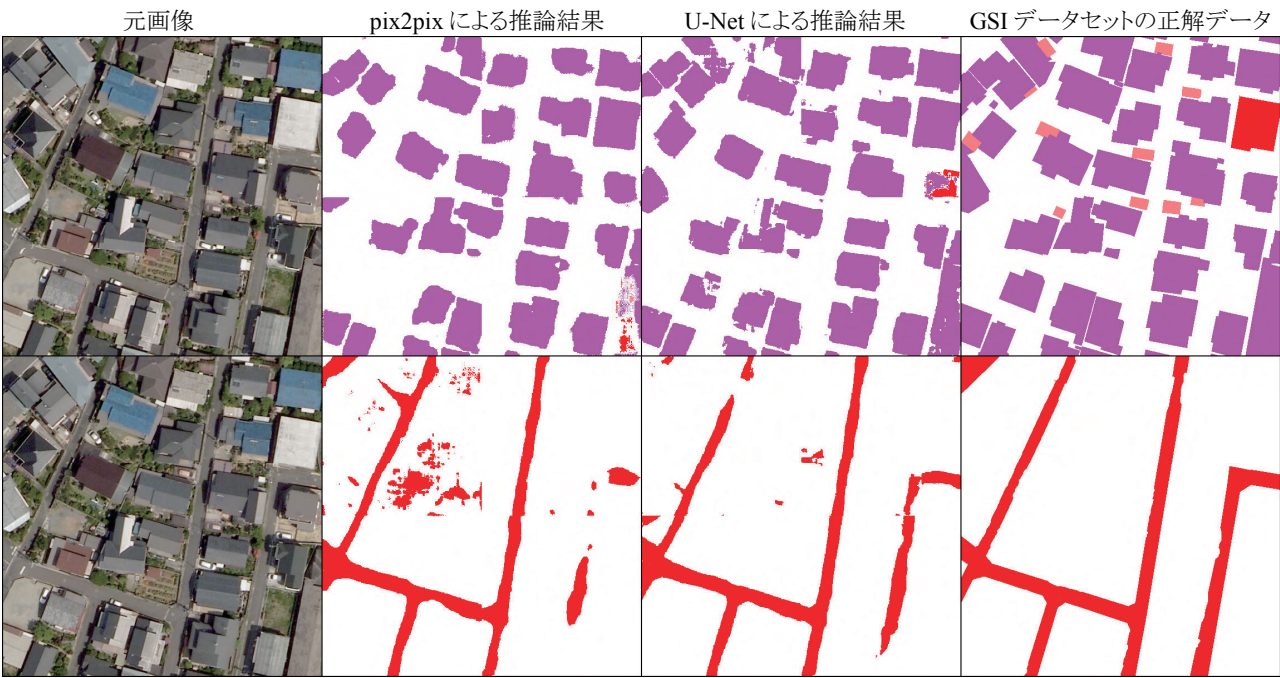


図 4. Feature Extraction の推論結果(上:建物, 下:道路).指定 RGB の各値から±5 の範囲を有効な pixel とみなしている.

表 3. 道路抽出の評価結果

		GSI データセット		
		Digital	Analog	All
Pix2pix	Precision	76.0	71.0	73.7
	Recall	67.2	47.9	57.0
	F 値	71.3	57.2	64.3
U-Net	Precision	82.6	75.9	79.4
	Recall	70.7	53.5	61.6
	F 値	76.2	62.7	69.4

もすべてのケースで U-Net の評価値が高い結果が得られた。

次に、建物の推論結果における、GSI データセット(デジタルカメラ)と Inria データセットでの評価値の比較では、GSI データセットの方の Precision がやや高く、Recall は大きく異なるという結果が得られた。Precision がやや高いのは、参照データに既往の地図データを用いず、すべての画像で測量技術者が判読してラベル付けを行ったことにより、参照データの位置ズレや不整合の量が相対的に少ないことが要因の一つと考えられる。一方、Recall の値が大きく異なるのは、トレーニングデータに日本国内の画像しか使用していないため、日本の都市の建物形状と大きく異なる形態、例えば、Vienna(ウィーン)はヨーロッパによくみられる古い石造りの街並みであるし、Austin は樹木に囲まれた住宅が点在する街並みであるなど、を呈する都市では、適切に推論できず、抽出されなかった建物が多いためであると推測される。この推測は、日本の建物や土地利用と似通っている Tyrol-W の Recall の値が GSI データセットの Recall の値と大きく変わらない点からも整合的である。GSI データセットと Inria データセットの間での評価値の傾向は一貫しており、面積は少ないものの多様性を重視した GSI データセットによる評価値が一定の信頼性を有することを証明していると考えられる。

なお、今後の空中写真撮影でアナログカメラが使用されることは無いことから、デジタルカメラによる評価値の重要性が高いことを付記する。

5. おわりに

測量における地図作成という実業務に Feature Extraction の導入を目指した研究に必要となる性能評価用のデータセットを構築し、その評価結果を既存のデータセットの評価結果と比較することで評価結果に信頼が置けるかどうかを考察した。このデータセットは、今後の研究における性能評価の指標として用いる予定である。

また、道路と建物が Feature Extraction の対象地物として用いられることが多いが、地図作成のためにはその他の多くの種類の地物の Extraction にも対応しなければならない。国家測量機関として基盤地図情報を整備・提供している国土地理院が地図作成能力評価のためのマルチカテゴリの性能評価用データセットを構築、公開することの意義は少なからずあると思われる、可及的速やかに GSI データセットの公開につなげたいと考えている。また、データセットの対象面積も順次拡大していく予定である。

参考文献

[Demir 18] Ilke Demir, Krzysztof Koperski, David Lindenbaum, Guan Pang, Jing Huang, Saikat Basu, Forest Hughes, Devis Tuia, Ramesh Raskar: DeepGlobe 2018: A Challenge to Parse the Earth through Satellite Images, arXiv:1805.06561v1, 2018

[Isola 17] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros: Image-to-Image Translation with Conditional Adversarial Networks, CVPR2017, 2017.

[国土交通省 08] 国土交通省:作業規程の準則, 国土交通省告示第 413 号, 2008.

[Maggiori 17] Emmanuel Maggiori, Yuliya Tarabalka, Guillaume Charpiat, Pierre Alliez: Can Semantic Labeling Methods Generalize to Any City? The Inria Aerial Image Labeling Benchmark, IEEE International Symposium on Geoscience and Remote Sensing (IGARSS),2017.

[大野 17] 大野裕幸: 空中三角測量の全自動化によるオルソ画像作成の効率化に関する研究, 国土地理院時報 129, 国土地理院, 2017.