



外邦図のGISデータ化のための自動処理

Automatic processing for GIS data conversion of Gaihōzu

中園悦子 竹内渉
Etsuko Nakazono Wataru Takeuchi



明治時代から昭和初期にかけて作成・修正された外邦図のGISデータ化についての自動処理に取り組み、日本全体の地図を処理した。その方法と結果、問題点について述べる。

We worked on the automatic processing of GIS data conversion of Gaihōzu created and revised from the Meiji Era to the early Showa Era, and processed maps of the entire Japan. Here we describe the method, results, and problems.

はじめに

外邦図は旧日本陸軍によって作成された地図であり、その範囲は当時の日本の領土に対応している。地図に記されている土地被覆情報や等高線などは、約80年のものであり、過去から現在にかけての国土の変化を捉えるためには重要な情報である。Stanford大学によりデジタル化された外邦図のうち、日本の範囲の5万分の1の地図をGISデータ化するにあたって、pythonを使った自動化処理を行うことを目標とした。

データと処理

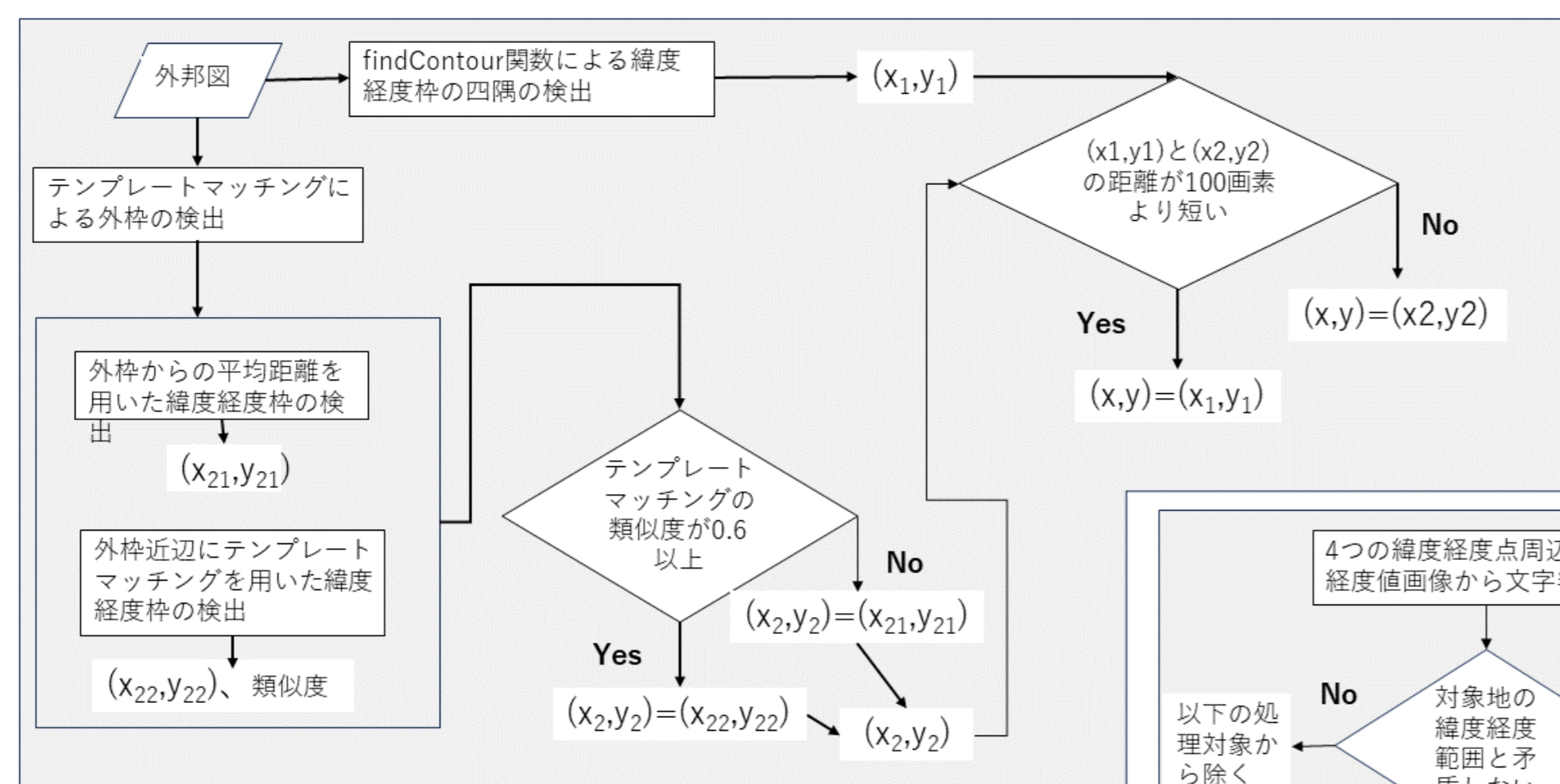
2016年、Stanford大学によって、デジタル化された外邦図が公開された。そのうち現在の日本の範囲の5万分の1の地図をデータとして使用した。この地図から、緯度経度点を検出し、緯度経度の文字を判読する過程の自動化を行った。最終的に、一枚の地図上の緯度経度点全てが検出される確率は86%、その周囲の文字が判読成功する確率は46%になった。

Introduction

Gaihōzu is a map produced by the former Japanese Army, and its extent corresponds to the territory of Japan at that time. The land cover information and contour lines on the map are approximately 80 years old and are important for capturing changes in the country's land from the past to the present. The goal of this project was to perform automated processing using python to convert a 1:50,000 scale map of Japan's extent into GIS data, which was digitized by Stanford University.

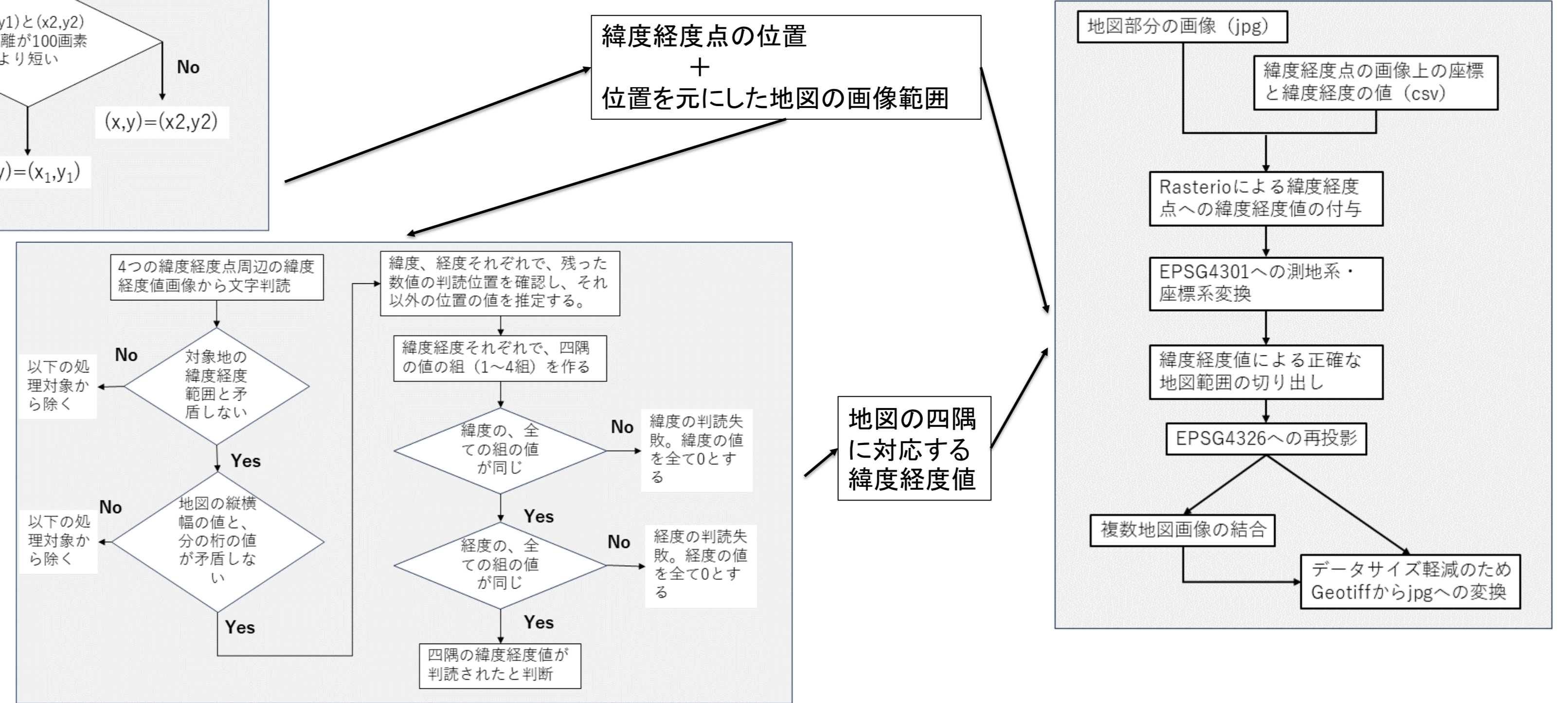
Data and Processing

In 2016, a digitized map of Outer Japan was released by Stanford University. 1:50,000 maps of the current extent of Japan was used as data. The process of detecting latitude and longitude points from this map and deciphering the latitude and longitude characters was automated. In the end, the probability of detecting all the latitude and longitude points on a single map was 86%, and the probability of successfully deciphering the surrounding characters was 46%.



地図の四隅検出フローチャート(左上)と緯度経度値の判読フローチャート(中央)、またこの二つの処理から作成した地図画像とcsvファイルによるGISデータ化処理のフローチャート(右)

Flowcharts for detecting the four corners of a map (top left) and deciphering latitude and longitude values (center), as well as a flowchart of the GIS data conversion process using map images and csv files created from these two processes (right).



The latitude and longitude values of the maps were fixed to a latitude width of 10 minutes and a longitude width of 15 minutes. In addition, at the download site, the maps were numbered regularly within a 4 x 4 sequence. Taking advantage of this, a semi-automatic program was added to change the names when downloading, determine the location within the 4 x 4 range based on the names, and estimate the latitude and longitude values at the four corners. Georeferencing was also performed in python from the two sets of data obtained. The geodetic system and coordinate system were converted from geotiff to jpg for ease of use on WebMap. The outer boundary map was then superimposed on the current high-resolution satellite image to check the overlap with characteristic landforms, such as coastlines and rivers.

地図の緯度経度値は緯度幅10分、経度幅15分と決まっている。またダウンロードサイトでは地図は4×4の並び内で規則的な番号を振られていた。これを利用してダウンロードの際に名前を変更し、その名前から4×4の範囲内の位置を判断し、四隅の緯度経度値を推定する半自動のプログラムを追加した。

また、得られた二つのデータから、pythonによるジオリファレンス処理を行った。その際、WebMap上で使用しやすいように、測地系・座標系を変換し、geotiffからjpgへと変換を行った。その上で外邦図を現在の高解像度衛星画像と重ねて、海岸線や河川など、特徴的な地形との重なり具合を確認した。

地図の枚数	四隅の判読ができていない地図 (文字判読画像を基準にした)	四隅全ての値が判読されているとプログラムが判断した地図	ファイル名を用いた半自動プログラムの結果を追加したもの
九州	190	172	139
四国	59	55	53
関西	136	113	113
関東・東海	333	271	269
東北	145	131	128
北海道	249	221	138
全国	1112	962	840

表: 対象となった地図の枚数と処理が成功した地図の枚数

結果と問題点

日本の範囲の地図1112枚のうち、自動プログラムで444枚、補助のプログラムを使うことで840枚、全体の76%の地図の処理が可能になった。

ただし、以下のような地図については、自動処理ができなかった

- 1: 地図が汚れており、書き込みや破れ、折れ目などの影響が大きい場合。
- 2: 日本の範囲の外邦図は、その多くが明治後期に作成され、昭和初期に修正されている。その間1807年に日本の経度の標準点が修正されたため、地図の経度に10.4"のずれが生じたので、大半の地図の経度に修正値として追加されている。そこで経度の判読結果に一律に10.4"を付与した。しかし九州以南の島々や、1807年以降に作られた北海道東部の地図などには修正値が無い。
- 3: ジオリファレンス処理をpythonで行う際に、gdalとrasterioのモジュールが必要だった。しかしバージョンの影響が大きいので、バージョンを考慮する必要のない環境でプログラムを使用する必要がある。
4. 一部の地図は、長方形から一部がはみ出して描かれている。このプログラムでは検出された四隅の座標に基づいて地図を切り出しているため、この形状からはみ出した場合は手動で処理する必要がある。

Results and Problems

Of the 1112 maps of the Japanese range, 444 could be processed with the automatic program and 840 with the auxiliary program, or 76% of the total. However, the following maps could not be processed automatically

- 1: Maps that were soiled and heavily affected by writing, tears, folds, etc.
- 2: Most of the Gaihō maps of the extent of Japan were created in the late Meiji period and were revised in the early Showa period. During that time, the standard point of longitude of Japan was revised in 1807, which caused a 10.4" discrepancy in the longitude of the maps, which was added as a correction value to the longitude of most of the maps. Therefore, 10.4" was uniformly assigned to the longitude decipherment results. However, islands south of Kyushu and maps of eastern Hokkaido made after 1807 do not have the corrected value.
- 3: The gdal and rasterio modules were required for georeferencing in python. However, because of the large influence of the version, the program had to be used in an environment where the version need not be taken into account.
4. some maps were drawn with parts of the map protruding from the rectangle. The program cuts out the map based on the coordinates of the detected four corners, so if the map extends beyond this shape, it must be handled manually.