546908

Robust Matching of Wavelet Features for Sub-Pixel Registration of Landsat Data

Jacqueline Le Moigne
Applied Information &Science Branch
NASA/GSFC-Code 935
Greenbelt, MD 20771
Tel:301-286-8723
Fax:301-286-1776
lemoigne@backserv.gsfc.nasa.gov

Nathan S. Netanyahu
Dept Math. & Comp.Sc. Bar-Ilan Univ., Ramat-Gan, 52900, Israel, and Center For Automation Research, Univ. of Maryland, College Park, MD 20742
nathan@cfar.umd.edu

David M. Mount
Computer Science Dept, University of Maryland, College Park, MD 20742

Jeffrey G. Masek
Biospheric Sciences Branch
NASA/GSFC-Code 923
Greenbelt, MD 20771
jmasek@ltpmail.gsfc.nasa.gov

Samuel Goward
Geography Department,
University of Maryland,
College Park, MD 20742

Abstract – **For many Earth and Space Science applications, automatic geo-registration at sub-pixel accuracy has become a necessity. In this work, we are focusing on building an operational system, which will provide a sub-pixel accuracy registration of Landsat-5 and Landsat-7 data. The input to our registration method consists of scenes that have been geometrically and radiometrically corrected. Such pre-processed scenes are then geo-registered relative to a database of Landsat chips. The method assumes a transformation composed of a rotation and a translation, and utilizes rotation- and translation-invariant wavelets to extract image features that are matched using statistically robust feature matching and a generalized Hausdorff distance metric. The registration process is described and results on four Landsat input scenes of the Washington D.C. area are presented.**

I. INTRODUCTION

Over the next 30 years, NASA will be faced with many new challenges. In Earth Science, it will be the need to predict regional climate change on seasonal and inter-annual time scales, or to understand the interactions between human activity and the changes in the major Earth ecosystems. In Space Science, distant planet exploration and formation flying will be part of many missions. To address such challenges, integrating and creating seamless mosaics of data from multiple times, multiple sensors and multiple viewpoints will be a key component. Very accurate registration of these multi-sensor data is the first requirement of such an integration. But a number of distortions prevent two images acquired either by the same sensor at different times or by two sensors at the same or different times from being "perfectly registered" to each other or to a fixed coordinate system. It is very difficult to determine the exact location within an image using only ancillary data and geo-location is usually computed by combining *navigation* and *registration*. Navigation corresponds to a "systematic correction" based on image acquisition models taking into account satellite orbit and attitude, sensor characteristics, platform/sensor relationship, Earth surface and terrain models and brings the registration accuracy within a few pixels. Image registration, on the other hand, corresponds to a "precision correction" based on landmarks and image features, and refines the geo-location to a sub-pixel accuracy. Registration is either applied after the navigation process, or both processes are integrated in a closed feedback loop. In this paper we will only consider the issue of feature-based, precision-correction automatic image registration.

Our goal is to build an operational system which will provide a sub-pixel accuracy registration of Landsat-5 and Landsat-7 data. Our method assumes a transformation composed of a rotation and a translation, and utilizes rotation- and translation-invariant wavelets to extract image features [2,3], that are matched using statistically robust feature matching and a generalized Hausdorff distance metric [4]. The registration is carried out on carefully chosen "sub-scenes" of a reference and of all incoming scenes. Preliminary results were previously reported at IGARSS'00 [1], and showed results of these individual sub-scene registrations. In this paper, we will first describe how these sub-scenes (or "chips" and "windows") are chosen and extracted. Then, we will summarize the principle of our algorithm, we will present new results, and then show how results of all individual registrations are combined to provide the final registration of an input scene relative to a reference scene.

II. WAVELET-BASED REGISTRATION OF LANDSAT DATA

With the final goal of integrating automatic registration within an automated mass processing/analysis system for Landsat data (REALM), we assume that the input to our registration method consists of scenes that are geometrically and radiometrically corrected. In the future, input scenes will also be pre-processed for detection of clouds and cloud shadows. Such pre-processed scenes are geo-registered by utilizing carefully chosen Landsat "reference chips," or "landmark chips." For the final system, a chip database will be built; we define "reference chips" as 256x256 images representing well-contrasted visual landmarks, such as bridges, city grids, islands, high-curvature points in coastlines. Several reference chips corresponding to different seasons and/or different reflectance conditions will represent each landmark area. For a typical Landsat input scene, we assume that between 5 and 10 well-distributed chips will be available from the database and should hopefully correspond to cloud-free areas of the scene. The choice of the relevant chips will be performed knowing the UTM (Universal Transverse Mercator) projection coordinates of one or the four corners of the scenes, and the UTM coordinates of all database chips. Currently, no chip database is available, therefore for the work presented in this paper, we created chips from a 1999 Landsat-7 scene, and four earlier Landsat-5 scenes are being registered relative to these chips.

For each relevant reference chip that is selected as overlapping the given input scene, a corresponding window is extracted from the scene, using the UTM coordinates of the chip and of the scene corners. The UTM coordinates of the 4 chip corners are projected onto the scene with a simple linear interpolation taking into account the sizes (in meters) of the pixels so that the projections are converted into pixel locations. These pixel locations are the basis to extract a window in the scene that corresponds to the chip.

Then, each chip-window pair is registered using our robust wavelet feature matching [4]. We assume that the transformation between incoming Landsat scenes and reference chips is limited to the composition of a rotation and a translation. Previous registration experiments [3] have shown that orthogonal wavelet filters could be utilized for image registration but were not rotation- and translation-invariant, therefore such a wavelet-based registration was not stable for large transformations and for large amount of noise. We then conducted similar experiments using the steerable filters from an overcomplete frame representation proposed by Simoncelli et al [2], and the registration based on those new filters proved to be much more stable and more accurate than the previous one. For the registration of Landsat data, we decompose both chip and window using four decomposition levels and one steerable band-pass filter. At each level, the results of the band-pass filtering are thresholded to keep only those 10% top pixels whose magnitudes after filtering are the highest. Those top pixels at each decomposition level are used as features in the feature matching process.

The feature matching strategy follows the multi-resolution given by the wavelet decomposition. Starting from the lowest level of decomposition and iteratively refining the matching at each level, strong features are matched using statistically robust feature matching and a generalized Hausdorff distance metric. This matching is based on the principle of point mapping with feedback. Specifically, given a set of control points in the chip and a corresponding set of points in the scene window, and assuming a pre-specified transformation (here composition of translation and rotation), our method represents a computationally efficient algorithm to match these point patterns. An outline of our proposed algorithmic methodology consists of the following:

(0) Monte Carlo sampling of control points.
(1) Application of robust similarity measures (e.g., k-th smallest squared distance to nearest neighbor).
(2) Searching the transformation space through hierarchical spatial subdivisions.
(3) Pruning the search space by "range" similarity estimates.
(4) Employment of fast data structures for nearest neighbor and range queries in image space.

As a summary, our registration algorithm can then be described in five steps:

1. For every new input scene, choose the relevant reference chips that have a sufficient overlap with the scene.
2. For each relevant reference chip, extract a corresponding window in the incoming scene.
3. For each chip-window pair, compute the (rotation, translation) transformation by:
   3.1 Perform a wavelet decomposition of the chip and of the window. Extract the top 10% pixels whose wavelet magnitudes are the highest.
   3.2 Perform a robust matching of the selected pixels (or wavelet features) using a nearest neighbor strategy and a generalized Hausdorff distance.
4. The previous registration is performed and a local transformation is computed for each chip-window pair. From these local transformations, the corrected locations of the four corners of each chip are computed, and the list of all corners of all chips, including old and corrected locations, becomes the input of a Least Mean Square (LMS) computation that computes a global transformation over the entire input scene. If n is the number of chips that is used for the given scene, the global transformation is computed using 4n pairs of points.
5. Using the global transformation computed in step (4), new UTM coordinates for each of the four corners of the scene are being computed. This step provides a new scene indexing, and unless required by the user, no resampling is being performed.

III. Registration Results

The method described in Section II was tested using seven 256x256 chips extracted from a 1999 Landsat-7 scene of the Washington/Baltimore area. Because the navigation system of Landsat-7 is more accurate than Landsat-5's, we chose to utilize a 1999 Landsat-7 scene to extract the reference chips. Figure 1 shows these 7 chips. Four Landsat-5 scenes from 1984, 1987, 1996 and 1997 over the same area were automatically registered to the chips. All scenes were projected using a WGS-84 model. Using the UTM coordinates of the 4 corners of each chip and the UTM coordinates of the 4 corners of the incoming Landsat-5 scene, a corresponding window was extracted for each chip. Figure 2 shows some of the corresponding windows. Then, each chip-window pair was automatically registered by extracting Simoncelli wavelet features up to level 4, and by matching them through robust feature matching. The results of all chip-window registrations are shown in Table 1, including the distance associated with each registration. Then, Table 2 shows the global registrations computed for all four scenes and Table 3 shows the corresponding registrations that were computed "manually" by averaging the locations of two Ground Control Points chosen visually. From these results, we can see that, quantitatively, the average translation error of the automatic registration compared to the manual registration is about 1 pixel in the x-direction and 0.63 pixel in the y-direction. The results also show that the translation errors are the smallest (0.23,0.12) for scene 97275, for which all distances associated with the local registrations are null, that means that those registrations have a very high confidence. This indicates that, if the confidence obtained in each local registration is taken into account when computing the global transformation, we could improve the final results.

IV. CONCLUSION

A system for the automatic registration of Landsat data has been designed. It was tested on four input Landsat-5 scenes registered to 7 chips extracted from a 1999 Landsat-7 scene. First results are encouraging, but further testing need to be performed, including the inclusion of rlocal registration confidence in the global transformation, as well as calculating the reference "ground truth" registration with an standard method such as ENVI. Future work will also include testing the method on a larger number of data, as well as building a well-distributed database of landmark chips.

REFERENCES

[1] J. Le Moigne et al, "Geo-registration of Landsat data by robust matching of wavelet features," *IGARSS'2000.*

[2] E. Simoncelli, and W. Freeman, "The steerable pyramid: a flexible architecture for multi-scale derivative computation," *IEEE ICIP*, October 1995.

[3] J. Le Moigne, and I. Zavorin, "Use of wavelets for image registration," *SPIE/Aerosense'2000,* Orlando, FL, Apr.00.

[4] D. Mount, N. Netanyahu, and J. Le Moigne, "Efficient algorithms for robust feature matching," *Pattern Recognition*, Vol.32, No.1, 1999.

Table 1 - *Local Chip-Window Pairs Transformations Rotations in degrees, Translations in pixels*

Chip #	84240	87136	96193	97275
#1 - Rot:	0	0	1	0
TX	8	18	12	21
TY	-39	-25	-92	-28
Distance	0.00	1.00	1.80	0.00
#2 - Rot:	0	-0.5	1	0
TX	8	11	-66	22
TY	-41	-41	4	-30
Distance	0.00	0.62	1.93	0.00
#3 - Rot:	0	-0.5	-0.3	0
TX	8	11	10.84	21
TY	-40	-41	-96	-29
Distance	1.00	0.89	0.52	0.00
#4 - Rot:	0	0.8	0	0
TX	8	12.34	11	21
TY	-41	--38.12	-94	-28
Distance	1.00	0.96	0.00	0.00
#5 - Rot:	0.4	0.7	0	0
TX	7.8	10.4	-38	20
TY	-40.86	--40.34	52	-31
Distance	0.94	0.93	2.24	0.00
#6 - Rot:	0	-0.8	0.3	0
TX	6	10.61	11.5	22
TY	-41	-41.94	-99	-33
Distance	1.00	0.78	0.70	0.00
#7 - Rot:	0.5	0	0	0
TX	9	12	12	22
TY	-40	-38	-94	-29
Distance	0.56	0.00	0.00	0.00

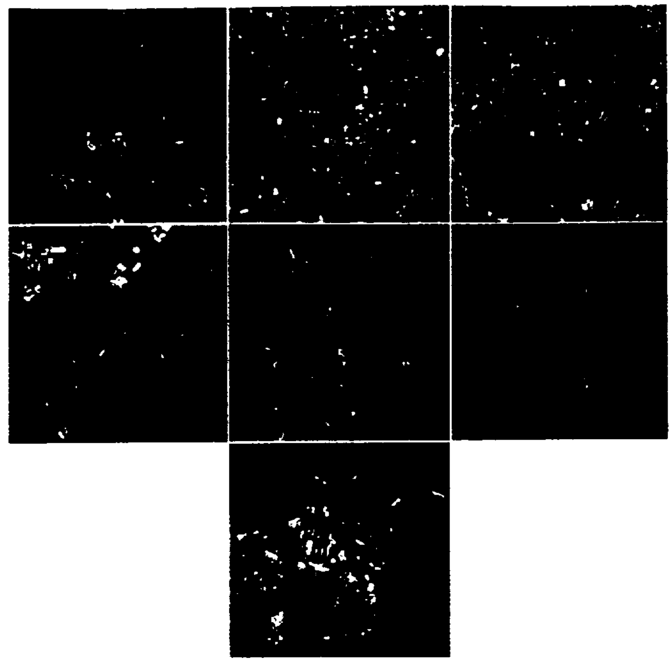

Figure 1

The 7 Chips Extracted from a 1999 Landsat-7 Band 4

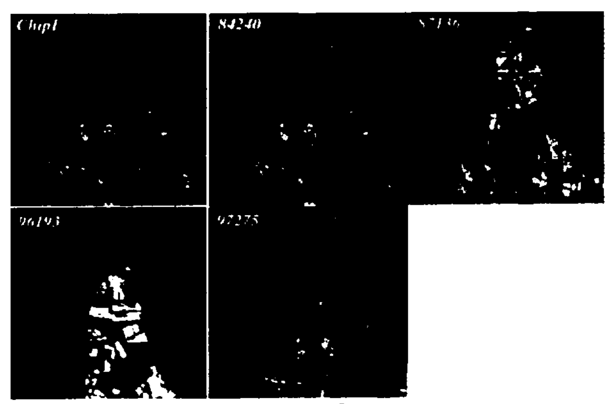

Figure 2

Chip1 and Its Corresponding Windows in the 4 Input Scenes

Table 2

Global Transformation For All Five Scenes

Transf.	84240	87136	96193	97275
Rotation	0.013	0.003	-0.042	-0.143
Transl-x	7.18	11.43	12.61	21.20
Transl-y	-41.12	-40.49	-95.38	-28.85

Table 3

Manual Transformation For All Five Scenes

Transf.	84240	87136	96193	97275
Rotation	0.00	0.00	0.00	0.00
Transl-x	7.18	10.55	9.48	20.97
Transl-y	-40.06	-39.16	-95.16	-28.97