IMAGE COREGISTRATION AND MOSAICKING OF DIWATA MICROSATELLITE IMAGERY

Panji P. Brotoisworo (1), Arlo Jayson C. Sabuito (1), Harry C. Merida (1), Mark Jayson B. Felix (1), Gay Jane P. Perez (1)

[1] STAMINA4Space Program, Institute of Environmental Science and Meteorology, University of the Philippines, Diliman, Quezon City, 1101, Metro Manila, Philippines
Email: pbrotoisworo@stamina4space.upd.edu.ph, acsabuito@stamina4space.upd.edu.ph, merida@stamina4space.upd.edu.ph, mjbfelix@stamina4space.upd.edu.ph, gpperez1@upd.edu.ph

KEY WORDS: Image Coregistration, Microsatellites, Phase Cross-Correlation, Seamless Mosaic

ABSTRACT: Image coregistration and mosaicking is an essential part of satellite data processing and transforms raw single-band data into usable products that are ready for research and analysis. The images of the Spaceborne Multispectral Imager (SMI) onboard Diwata-2 presents unique challenges compared to standard push-broom or target and shoot detectors in other satellites because the imagery is taken using only one detector for visible bands and one for infrared bands, both equipped with Liquid Crystal Tunable Filter (LCTF). The two detectors capture multispectral images ranging from 440nm to 1050nm which leads to a staggered image acquisition due to the spatiotemporal offset as the detector needs to sequentially go through the target wavelengths. This study aims to create an automated process to align and geometrically correct Spaceborne Multispectral Imager (SMI) and High Precision Telescope (HPT) images from the Diwata microsatellite. The HPT or High Precision Telescope, is another imaging payload of Diwata-2, which features a fixed filter for its four spectral band detectors, enabling the imaging acquisition of each band almost simultaneously, so the challenge present in SMI images is not necessarily observed in HPT imagery. Nonetheless, HPT and SMI imagery go through the same process to test the coregistration method presented in this study. Images are first assigned a dummy UTM coordinate system with a cell size of 1 and then go through global coregistration. This involves using a using a phase cross-correlation algorithm to calculate vertical and horizontal shifts with sub-pixel accuracy to align the images in the coordinate system. Then a local coregistration is performed which corrects for local geometric distortion via a grid of ground control points to calculate local image shifts using a moving window. After image alignment, a color correction algorithm is used to create seamless single band mosaics which will be used to create a composite multi-band image. The success rate of SMI imagery was 73% while HPT imagery featured a success rate of 57%. Successfully coregistered image composites featured a final RMSE value of 0.635 for SMI imagery and 0.563 for HPT imagery which is less than the target threshold of 1 pixel. This algorithm performed well but struggles if there is a lack of distinct features such as when there is a dense cloud cover or if a large portion of the image is water.

1. INTRODUCTION

1.1 - Background

The Diwata microsatellites are the Philippines' first optical imaging microsatellites capable of capturing images in red, green, blue, and near-infrared bands. Due to microsatellites' limited payload which is only up to 100 kg microsatellites feature less equipment when compared to their full size counterparts. Microsatellites feature cheaper launch costs at the cost of fewer hardware

redundancies and limited thermal control which require more frequent in-orbit calibration (He et al., 2020).

The initial methodology of creating remote sensing products involves manual georeferencing of individual satellite imagery and then generating a composite, a time-consuming task that is difficult and costly to scale up if the number of images increases. Quick turnaround times for remote sensing products are especially important in time-sensitive missions such as disaster-response or national defense wherein the situation can change very rapidly.

Image coregistration is a standard process used to align different images and minimize any distortion effects due to temporal or geometric effects. This type of work is often used in remote sensing, medical imaging, and computer vision research and is often used to align images for change detection or to create a composite image (Scheffler et al., 2017). The use of image coregistration with Diwata imagery for remote sensing purposes will create a coregistered and mosaicked stack of images. Like the Sentinel-2 satellites, the Diwata microsatellites feature a temporal offset when capturing imagery using different bands (European Space Agency, n.d.). The temporal offsets lead to Diwata imagery having a staggered image acquisition wherein some parts of the slave image may not be present in the reference image. Unintended satellite rotation during the image acquisition process can also cause geometric distortions as the satellite captures photos.

Image coregistration is a crucial step when generating satellite imagery products because misaligned images can distort the image's visual appearance and it can negatively affect any research work done using those misaligned images. This research aims to automate the image coregistration process and thus speed up the distribution of satellite products of the Diwata microsatellites to relevant end-users. It is important to note that the output imagery generated by this study is still considered ungeoreferenced but seeks to cut down georeferencing time by allowing the processing team to georeference a single mission mosaic instead of every unique image taken during the mission.

1.2 - Related Work

Previous research on image coregistration methods for Diwata imagery that was done by (Tupas et al., 2016) focused on creating keypoints that were generated using the Scale-invariant Feature Transform (SIFT) and Features from Accelerated Segment Test (FAST) algorithms. The results were improved by combining feature matching with pass stitching. Using Landsat imagery as a reference image, the study evaluated the performance of the SIFT and FAST algorithms in generating keypoints between Diwata and Landsat imagery. These keypoints acted as ground control points (GCPs) that allowed the slave image to be warped to the master image. The output of this study aimed to have imagery that was already georeferenced and was able to be composited based on the output of the SIFT and FAST algorithms. The study achieved a reasonable success rate of 60% but had root mean square errors larger than the half-pixel requirement set by the researchers. This result was not accurate enough for the algorithm to be used for operational purposes.

The Phase Cross-Correlation algorithm by Guizar-Sicairos et al. (2008) is an image coregistration algorithm based on discrete Fourier transforms (DFT) and works by doing image reconstruction through phase retrieval by taking an image $g(x, y)$ of an object $f(x, y)$. An upsampled matrix multiplication DFT is used to find the peak of cross correlation $r_{fg}(x_0, y_0)$. The location of this peak in (x, y) coordinates will be the offsets that are applied to the target image. In equation (1),

N (columns) and M (rows) represent the image dimensions, uppercase letters represent the discrete Fourier transform (DFT) of the image and object variables, and summations are taken using image points (x, y).

$$r_{fg}(x_0, y_0) = \sum_{x,y} f(x,y)g^*(x - x_0, y - y_0)$$
$$= \sum_{u,v} F(u,v)G^*(u,v)exp\left[i2\pi\left(\frac{ux_0}{M} + \frac{uy_0}{N}\right)\right] \quad (1)$$

The phase cross-correlation algorithm has been used in other remote sensing applications in the context of detecting image offsets such as Rex & Hirt (2014) for computing georeferencing accuracy for different DEM's and Skakun et al. (2017) for coregistration of Sentinel-2 and Landsat-8 imagery. This is a subpixel algorithm which allows for increase accuracy since vertical and horizontal offsets can be calculated with higher accuracies and the magnitude of offset movement will not be limited by the pixel size. The strength of this algorithm is that it is self-correcting as it attempts to calculate the offset of each image pair. The pass stitch method used by Tupas et al. (2016) still required a significant amount of manually georeferenced imagery to ensure the coordinates generated by the stitching process were valid.

Geometric correction is done using software called an Automated and Robust Open-Source Image Coregistration Software (AROSICS) by Scheffler et al. (2017). Geometric correction is based on a phase cross-correlation algorithm that is done using a moving window that allows for calculation of local image shifts. The results of the moving window are saved into a tie point grid that will use a polynomial warp to align the target image with the reference image.

AROSICS has been used in different remote sensing contexts such as high resolution multi-temporal optical image coregistration of Sentinel-2 by Yan et al. (2018) and image coregistration featuring different resolutions and satellites using Landsat-8 and Sentinel-2 satellite imagery by Stumpf, Michéa, & Malet (2018).

2. MATERIALS AND METHODS

2.1 - Images

The images used in this study are Level-0 10-bit imagery taken by the Spaceborne Multispectral Imager (SMI) and High Precision Telescope (HPT) sensors captured by the Diwata satellite. SMI imagery feature a unique and more challenging feature which results in large offsets between images. This is because instead of using a standard push-broom or target and shoot detector the SMI sensor captures images using a wide range of wavelengths ranging from 440nm to 1050nm using only two detectors both equipped with a Liquid Crystal Tunable Filter (LCTF) which allows a single detector to capture different wavelengths. The Diwata satellite features one detector for visible bands and one for near-infrared bands. Due to the two sensors simultaneously taking images for visible and infrared bands there are several instances where near-infrared imagery will be coregistered with non near-infrared imagery. There is an image offset due to the LCTF being adjusted as the detector goes through all the target wavelengths. The script will coregister all available bands sorted by the time of capture to maximize the overlap between the two image pairs. Only operational bands will be selected later for mosaicking. Operational bands for SMI include 490nm, 550nm, 670nm, 778nm, 780nm. Near-infrared imagery sometimes alternates from 778nm to 780nm. HPT features a detector for each of the target bands thus features a near

simultaneous image acquisition for all four bands and large overlap between single band images. The specifications of these imagery are seen in Table 1 below.

Table 1: Overview of the imagery used in this study

Sensor	Number of Testing Samples	Image Size (Columns, Rows)	Number of Bands	Ground Sampling Distance (Meters)	Bit Depth
Spaceborne Multispectral Imager	30	659, 494	4	126.9	10-bit
High Precision Telescope	30	659, 494	4	4.7	10-bit

Additional preprocessing is done on the images for artifact correction and improvement of coregistration performance using a combination of median filters, histogram matching, and spectral whitening. The output of the algorithm is then applied to the original unmodified imagery. The images used for the coregistration study consist of only one composite image taken from sample missions for the purpose of analyzing image coregistration performance. Mosaic imagery used in this paper uses the entirety of the mission in order to create a mission mosaic.

2.2 - Blending Algorithm

Seamless mosaicking for each unique band is done after the coregistration and geometric correction of single band imagery. At this stage, the images feature arbitrary geotransform data based on the results from the phase cross-correlation results. Thus a seamless mosaic using a blending algorithm can be generated based on the geographic overlap of the unique images. The blending algorithm uses a distance transform function on the binary mask of two adjacent images to be mosaicked. A weight factor array will be generated from the distance maps using the expressions:

$$w = \frac{dmap_1}{dmap_1 + dmap_2} \tag{2}$$

$$Overlap(x,y) = \big(w(x,y) * image_2(x,y)\big) + \Big(\big(1 - w(x,y)\big) * \big(image_1(x,y)\big)\Big) \tag{3}$$

Where w is the weight factor array, and $dmap_1$ and $dmap_2$ are the separate distance transform maps of the adjacent images excluding their area of intersection. Figure 1a shows the footprint of two adjacent images in white and their overlapping region in gray. Figure 1b shows the resulting weight factor array using the formula (2). The resulting array contains a gradient map as seen in Figure 1c that decreases with increasing proximity to the first image and increasing value as it approaches the second image. The values for the non-overlapping regions of the images will be copied from the original images while the values for the intersection will be computed using the formula (3).

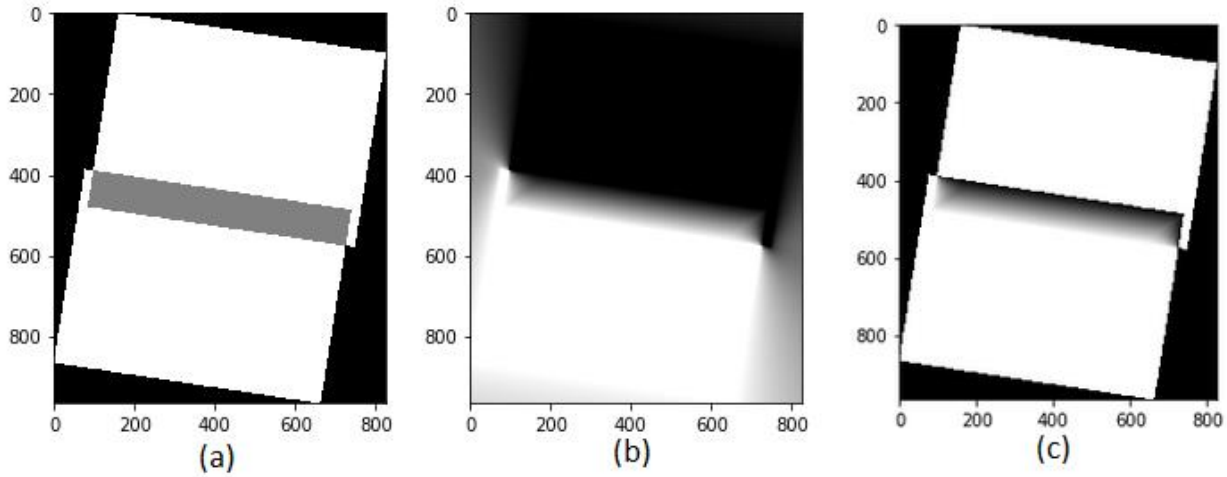

Figure 1: Images showing how the images are blended together in a seamless mosaic

The darker areas in the overlap indicate increasing influence of the first image to the computed values while lighter intensities signify increasing influence of the second image. The resulting image is then used as the new $image_1$ and the next adjacent image is used as the new $image_2$. This process is repeated until all images for the target band are mosaicked as seen in Figure 2. The final matrix is written in GeoTiff format and saved in the directory for mosaicked images and is ready to be made into a composite raster.

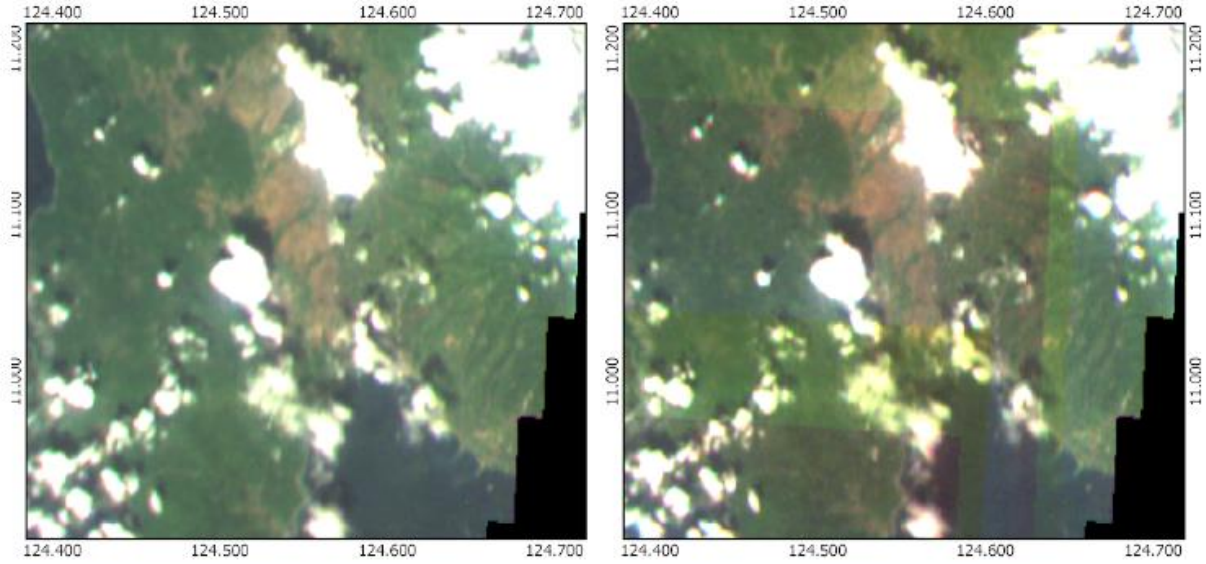

Figure 2: Output of the blending algorithm after being processed into an RGB mosaic (left) and the same mosaic using no blending algorithm (right)

2.3 – General Methodology

The script sorts the images according to the oldest image first and the newest image last. Then each image is assigned an index number i in ascending order where the oldest image starts with an index number of zero. Then using a loop, it goes through the available index numbers sequentially and it processes images in pairs where the offset calculations are applied to the geotransform of the target image $i + 1$ with respect to the reference image i. The first image and oldest image in the mission is assigned an arbitrary UTM coordinate system where the origin is set as 1000, 1000, and the pixel size is set to one. This allows the outputs of the phase cross-

correlation algorithm that use pixel units to be easily applied to the geotransform data and it allows for easier understanding of offsets and geometric shifts in the performance metrics.

This methodology is used for both the image offset calculation and geometric correction algorithms. Performance metrics are tracked for each image pair and include horizontal and vertical offset movement for phase cross-correlation analysis and geometric displacement information for geometric correction analysis. It should be noted that HPT is the only sensor that requires two rounds of geometric correction. After the image coregistration process is done, the script proceeds to filter for the operational bands and then proceeds to create a seamless mosaic using the blending algorithm before finally proceeding to create a composite image using the coregistered bands. The visualization of this algorithm is seen in Figure 3.

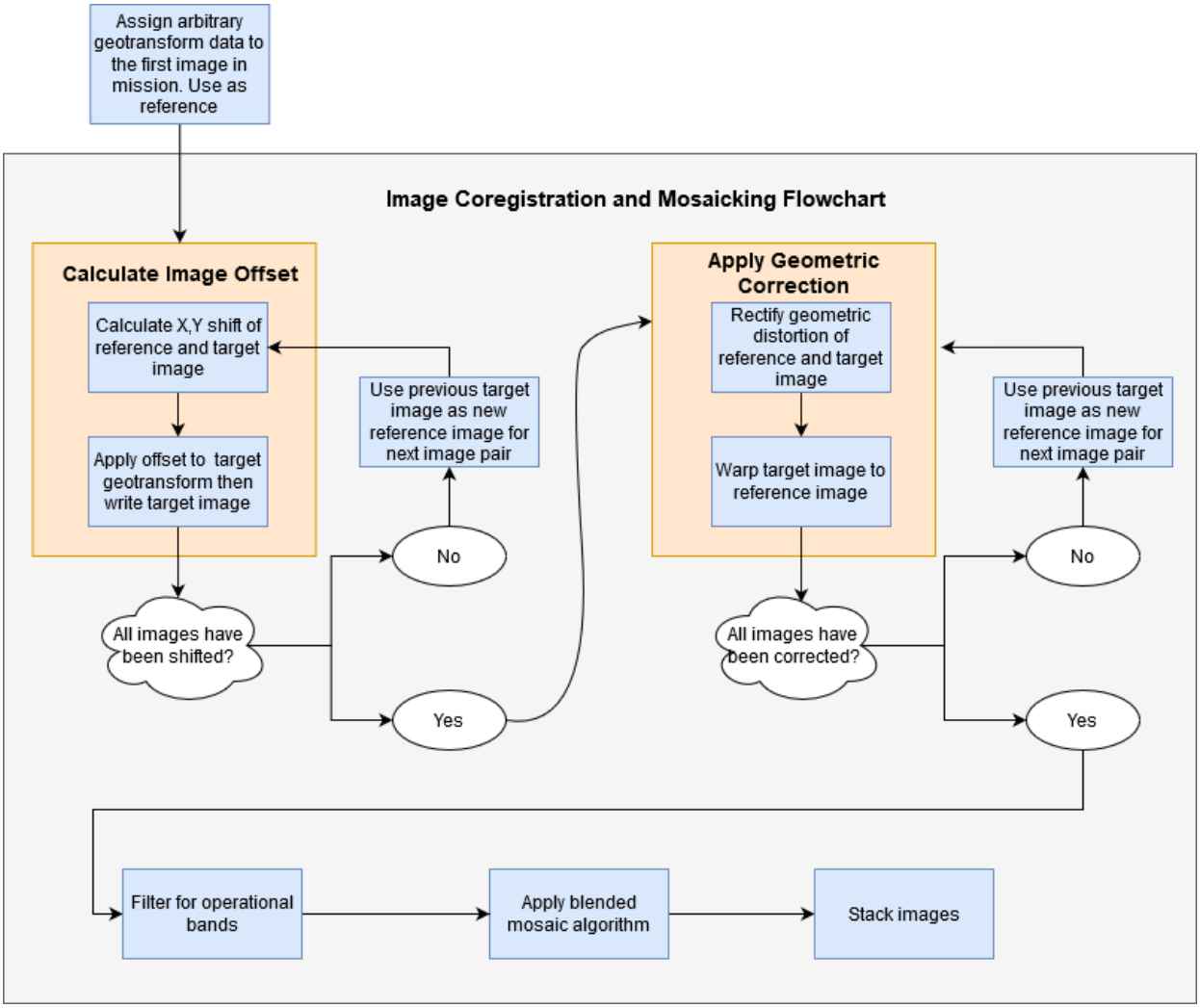

Figure 3: Methodology of image coregistration process

In each mission there will be summary statistics where it takes the average, standard deviation, minimum values, and maximum values of each metric. Data of each image pair is also available for finer analysis.

3. RESULTS AND DISCUSSIONS

3.1 – Image Coregistration

Each image was a four-band image which contains the blue, green, red, and near-infrared band. Successful coregistrations for NIR and RGB composites using those bands were tracked separately for both SMI and HPT imagery as well as the full four-band composite. An image is counted as successful if all bands are aligned and fails if it suffers a global coregistration error (Improper horizontal or vertical shift) or local coregistration error (geometric distortion still present). RMSE is calculated from the final coregistered product by calculating absolute shifts of the tie point grid used in geometric correction to warp the target image to the reference image. Only the RMSE of successful coregistrations were able to be tracked since AROSICS did not attempt geometric corrections if the alignment is not precise.

Table 2: Results of image coregistration

	SMI			HPT		
	RGB Composite	NIR Composite	4-Band Composite	RGB Composite	NIR Composite	4-Band Composite
Success Rate (%)	73	86	73	93	56	57

As seen in Table 2, SMI had a higher accuracy when looking at full four-band composites, however HPT RGB composites featured the highest accuracy rate at 93%. It is hypothesized that HPT NIR composites had a high failure rate due to lack of similar features between some NIR band imagery and non-NIR band imagery. Some areas of the NIR imagery were saturated and had few distinct features, however if you looked at the non-NIR band imagery there are distinct features present in the same area.

Table 3 shows the average metrics of successful coregistration output. Successful image coregistration featured an average RMSE value of 0.635 for SMI and 0.563 for HPT. This value fits well within the acceptable value that the Diwata processing uses when georeferencing imagery which aims for an RMSE value of less than one. Total offset displacement was calculated from horizontal and vertical displacement through the Pythagorean theorem. Due to the HPT imagery requiring two rounds of geometric correction, HPT metrics are calculated by taking the average of the averages from the two geometric corrections. In some cases where geometric correction by AROSICS cannot proceed due to lack of tie points after the filtering process or other errors, the geometric correction results will not be counted in the metrics.

Table 3: Analysis results for SMI imagery showing the displacement information for the correction process

	Successful SMI Coregistration			Successful HPT Coregistration		
	Geometric Correction RMSE	Geometric Correction Tie Points	Offset Displacement	Geometric Correction RMSE	Geometric Correction Tie Points	Offset Displacement
Mean	0.635	272.135	42.733	0.563	402.755	21.418
Standard Deviation	0.183	87.023	12.051	0.7	173.914	18.197

After coregistration, the images are mosaicked and if the coregistration is successful it will result in a full mission seamless mosaic which is ready for further preprocessing as seen in Figure 4.

Figure 4: Diwata-2 SMI RGB mosaic after merging the equivalent of 13 unique image composites

3.2 – Comparison of Automated Image Coregistration vs. Manual Method

Initial tests were conducted to study the drift of the pixel values of the coregistered images compared to existing individually georeferenced files which serves as the reference. Ideally, the values should be identical. However, values will deviate from the reference due to georeferencing errors as well as the shifts involved in the warping process, which both algorithms use to coregister the images. As a result, the analysis involves investigating which algorithm produces results with values closest to the reference.

A sample HPT image with the four operational bands from the test dataset was coregistered and georeferenced using the automated process detailed in this paper and compared to an image that was manually stacked and georeferenced. All images were utilizing reflectance values. Both the coregistered images are georeferenced in QGIS using Polynomial 2 and Bilinear for transformation. The coregistered images are resampled to the reference with bilinear resampling. The percent difference was calculated using the formula (4).

$$Percent\ Difference = \left(\frac{image}{reference} - 1\right) * 100 \quad (4)$$

Percentage values above zero indicate that the value is above the reference while the inverse indicates that the value is below the reference. In an ideal case, the percentage difference to the reference should be zero.

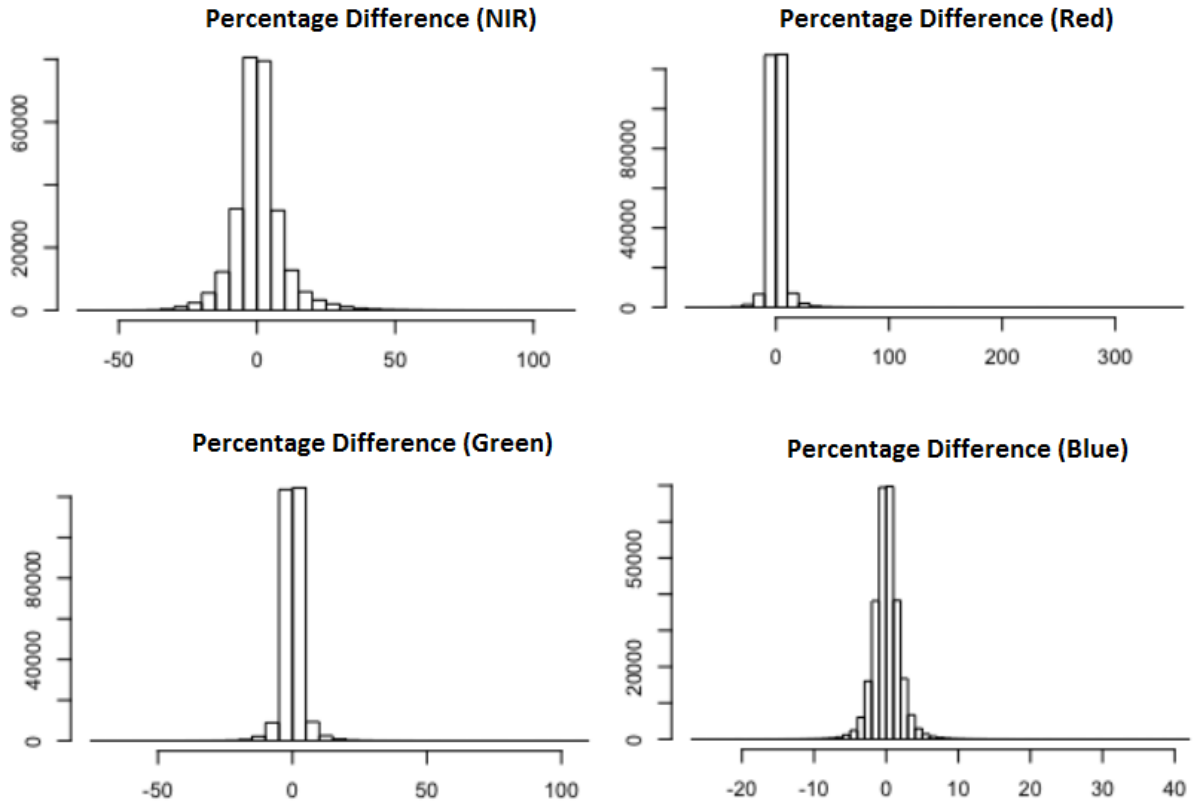

Figure 5: Histogram distribution of algorithm-based image coregistration when compared to reference image with pixel count on the Y-axis and percent difference on the X-axis

Table 4: Comparison of manual and automated processes of image coregistration

Band	Manual Coregistration Percent Difference	Automated Coregistration Percent Difference
Blue	0.015	0.049
Green	-0.031	0.086
Red	-0.064	0.252
NIR	0.173	0.542

Initial results for the test indicate that both the manual and algorithm-based coregistration produce good results that have low mean percent difference values as seen in Table 4 and Figure 5 shows that most of the pixel values feature close to zero percent difference. Close inspection shows that the area with the least variation are the fields and most of the drifting values from the reference are situated along the roads as well as some built-up areas. Some of the errors are easily traceable to the road outlines and built areas. This is expected as the georeferencing of the images are not perfect and this was minimized by achieving a georeferencing RMSE of below one pixel.

4. CONCLUSION

Our research has shown that our presented algorithm-based image coregistration method can be a viable method with a 73% success rate with SMI imagery and a 57% success rate with HPT imagery. The main benefit of this workflow is the time savings due to the laborious process of georeferencing data. If an image is unsuccessful in coregistration the user only needs to manually georeferenced a few images instead of georeferencing the entire mosaic. These algorithm-based image coregistration methods can allow for faster distribution of satellite imagery by minimizing the need to manually georeferenced every unique imagery.

However, it cannot be completely automated at this point due to challenges such as coregistering near-infrared bands with the non near-infrared bands. This can vary from sensor to sensor as it was observed that, for SMI imagery, NIR composites had a higher success rate while HPT imagery had a higher success rate with the RGB composites. This is a topic of ongoing research and research on how to extract matching features between infrared and non-infrared imagery is on-going.

Other future work includes coregistration of Diwata imagery to other optical satellite sensors such as Landsat or Sentinel-2, and more comparisons with the automated coregistration versus manual coregistration.

References

European Space Agency n.d., *Sentinel-2 MSI Technical Guide*, viewed 01 October 2020, https://earth.esa.int/web/sentinel/technical-guides/sentinel-2-msi/msi-instrument

He, L., Ma, W., Guo, P. and Sheng, T., 2020. Developments of attitude determination and control system of microsats: A survey. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, p.0959651819895173.

Guizar-Sicairos, M., Thurman, S.T. and Fienup, J.R., 2008. Efficient subpixel image registration algorithms. Optics letters, 33(2), pp.156-158.

Rexer, M. and Hirt, C., 2014. Comparison of free high resolution digital elevation data sets (ASTER GDEM2, SRTM v2. 1/v4. 1) and validation against accurate heights from the Australian National Gravity Database. Australian Journal of Earth Sciences, 61(2), pp.213-226.

Scheffler, D., Hollstein, A., Diedrich, H., Segl, K. and Hostert, P., 2017. AROSICS: An automated and robust open-source image co-registration software for multi-sensor satellite data. Remote Sensing, 9(7), p.676.

Skakun, S., Roger, J.C., Vermote, E.F., Masek, J.G. and Justice, C.O., 2017. Automatic sub-pixel co-registration of Landsat-8 Operational Land Imager and Sentinel-2A Multi-Spectral Instrument images using phase correlation and machine learning based mapping. International Journal of Digital Earth, 10(12), pp.1253-1269.

Stumpf, A., Michéa, D. and Malet, J.P., 2018. Improved co-registration of sentinel-2 and landsat-8 imagery for earth surface motion measurements. Remote Sensing, 10(2), p.160.

Tupas, M.E.A., Dasallas, J.A., Jiao, B.J.D., Magallon, B.J.P., Sempio, J.N.H., Ramos, M.K.F., Aranas, R.K.D. and Tamondong, A.M., 2017, October. Feature extraction and descriptor calculation methods for automatic georeferencing of Philippines' first microsatellite imagery. In Image and Signal Processing for Remote Sensing XXIII (Vol. 10427, p. 104271M). International Society for Optics and Photonics.

Yan, L., Roy, D.P., Li, Z., Zhang, H.K. and Huang, H., 2018. Sentinel-2A multi-temporal misregistration characterization and an orbit-based sub-pixel registration methodology. Remote Sensing of Environment, 215, pp.495-506.