www.adeepakpublishing.com

Manandhar, P. et al. (2020): JoSS, Vol. 9, No. 2, pp. 897–910
(Peer-reviewed article available at www.jossonline.com)

www. JoSSonline.com

Towards Rectifying the Effect on Image Registration of Pointing Inaccuracies in Nanosatellites Using Deep Convolutional Neural Networks

Prajowal Manandhar, Adham M. Alkhaja, Prashanth R. Marpu, and Zeyar Aung

Khalifa University
Masdar City Campus
Abu Dhabi, United Arab Emirates

Abstract

With the advent of satellite imaging with the use of nanosatellites, there is also huge potential to create a large volume of remote sensing images acquired over the same region in different periods of time by different satellites. Usually, remote sensing image datasets are geo-referenced. However, slight displacement is often seen with images from different sensors, or even within the same sensor due to pointing inaccuracies because of problems in attitude control systems. Here, we demonstrate a deep learning framework to solve such problems relating to identifying the location of satellite images by maintaining a separate repository of images from known reference satellites. Our idea is to detect the geographical location of remote sensing images using deep convolutional neural networks (CNN). We trained the VGGNet-16 model using Fully-Connected2 (FC2) features based on a reference Worldview-2 dataset to predict the geographical location based on the closest match found; in addition, it can be used to perform image registration. Performance evaluation of our proposed model is performed on different satellite images using a reference WorldView-2 satellite acquired in 2014, along with DubaiSat-2, image tiles from Google Earth, images from Sentinel-2 and Landsat-8 images acquired over an area of Abu Dhabi, United Arab Emirates (UAE).

1. Introduction

Nanosatellites are the current trend for performing diversified space science missions such as remote sensing and technological demonstrations. For remote sensing missions, it is crucial to achieve high pointing accuracies, which can be a challenge for nanosatellites. Due to the nonlinear nature of spacecraft dynamics, coupled with many uncertainties and external/internal disturbance sources, satellites requiring high pointing accuracy need efficient attitude and determination control systems, which are often restricted by severe constraints on space, power, and mass restrictions of nanosatellites (Nudehi et al.,

Corresponding Author: Prajowal Manandhar – prajowal.manandhar@ku.ac.ae

2008). The required pointing accuracy can be achieved by integrating different configurations for sensors and actuators for spacecraft active attitude control. Many options for actuator configurations exist (such as magnetic coils and reaction wheels); the choices are mainly a trade-off between power consumption and control performance (Kristiansen et al., 2005).

Among the many sets of configurations for active attitude control, reaction wheels have proven to be most accurate for three-axis stabilization (Kristiansen et al., 2005). They can control the attitude of the spacecraft independent of its altitude and geographical position (Kristiansen et al., 2005). The control performance of the reaction wheels can be degraded in the presence of internal disturbances, which can contribute vastly to the attitude error, and greatly influence the accuracy and stability of the satellite; hence, they must be accounted for in the control scheme design (Ge and Cheng, 2006). Another configuration used for active attitude control are magnetic coils. The magnetic coils produce a maximum torque from the magnetic dipole in the direction normal to the magnetic field vector (Zhou, 2010). A major drawback of using such configuration is that the maximum torque produced is zero when it is parallel to the local magnetic field vector (Zhou, 2010). In addition, controllability in all three axes of the spacecraft are lost at two points in the orbit; the yaw axis is uncontrollable at the poles of Earth and the roll axis will be uncontrollable at the equator (Kristiansen et al., 2005).

Extensive research has been conducted in spacecraft attitude control, mainly for cases involving an under-actuated system (Ping et al., 2016). Furthermore, adaptive control, model uncertainties, and disturbance attenuation are also trends of research on attitude control schemes for space systems (Ping et al., 2016). In adaptive control scheme development, the control objective is to achieve high stability attitude tracking while adapting to uncertain conditions such as failures in one or more actuators (de Ruiter, 2016). The performance of the adaptive control schemes presented in Yao et al. (2016) has been proven to guarantee stability during actuation failure cases. However, the required pointing accuracy for remote sensing missions still cannot be achieved in a constrained system.

Pointing accuracy determines attitude determination and control system (ADCS) overall performance; it involves pointing knowledge accuracy and pointing control accuracy, which comprise the knowledge of orientation based on sensors and maneuvering to desired orientation by actuators, respectively (Xia et al., 2017b). The control algorithm utilized for the mission also can have a major effect on the net accuracy of the ADCS; some examples of control algorithms include spin-stabilization, gravity-gradient stabilization, and momentum-bias stabilization (Chen et al., 1999). The pointing accuracy is generally determined by the configuration of actuators, sensors, and control method chosen. Two main factors that contribute to limiting the pointing control accuracy of a stabilized spacecraft are the modelling errors (e.g., geomagnetic field model or variation in inertia matrix due to propellant consumption) and external disturbances present in the space environment (Chen et al., 1999). Various examples of pointing accuracies that can be achieved by different ADCS configurations can be seen in Table 1 (Wertz et al., 2011).

Such errors in pointing will lead to inaccurate representation of ground coverage, thereby affecting the usability of the acquired data. In this work, we propose

Table 1. Attitude Control Methods and Pointing Accuracies

No.	Attitude Control Methods	Maneuverability Options	Pointing Accuracy (in ° degrees)
1	Gravity-Gradient (GG) stabilization	Gravity gradient boom	± 5
2	Spin stabilization	Thrusters	± 1 to ± 5
3	Momentum-bias stabilization	Momentum wheels	± 0.1 to ± 1
4	Three-axis stabilization	Control moment gyroscope (CMG) or reaction wheels	± 0.1

to use a post-processing approach for improving the accuracy of the geolocation of the acquired remote sensing data, particularly Earth Observation images.

Here, we demonstrate a deep learning framework to solve the problem related to rectifying geolocation of satellite images by maintaining a separate repository of images from known reference satellites. Our idea is to detect the geographical location of remote sensing images using deep learning with convolutional neural networks (CNN). We trained the VGGNet-16 model, employing Fully-Connected2 (FC2) features based on a reference Worldview-2 dataset to predict the geographical location based on the closest match found; in addition, it can be used to perform image registration. Performance evaluation of our proposed model is performed on different satellite images using a reference WorldView-2 satellite acquired in 2014, while test satellite images are obtained from DubaiSat-2, Sentinel-2, and Landsat-8, along with image tiles from Google Earth acquired over an area of Abu Dhabi, United Arab Emirates (UAE).

2. Related Works

The attitude pointing accuracy of satellites is often limited leading to large errors in the on-ground geolocation information in the provided metadata. At the same time, it is often an issue that the metadata is not available or not correctly matched with the satellite data, especially in the case of large volumes of archived data. This leads to several issues while trying to use such data for monitoring applications and for historical records. Image retrieval has been a fundamental task in remote sensing to efficiently retrieve data of interest from large collection of such data. One such big repository is mentioned in Cheng et al. (2017), which provides optical images in the Red, Green and Blue (RGB) channels for classification purposes. Researchers often face issues with mismatched labeling, even in the benchmark dataset, which hinders performance and results. Hence, the identification of location of such particular region will allow further investigation in other different images of the same region and also, pave a way towards correction through public feedback. The task of image retrieval in remote sensing imagery over the period has shifted from hashing methods to deep feature learning neural networks (Li et al., 2017b). Thus, the task of identification of image location using baseline remote sensing imagery helps to deal with change detection (De et al., 2017) more efficiently and precisely.

Deep Learning has been widely applied to various computer vision tasks (Krizhevsky et al., 2012; LeCun et al., 2015; Ma et al., 2019) with considerable success because of its superiority in terms of feature representation. In remote sensing, deep learning methods have been able to provide remarkable success in the fields of remote sensing such as land cover/use classification (Cheng et al., 2017; Chaib et al., 2017; Xia et al., 2017a; Scott et al., 2017); SAR image classification (Liu et al., 2017; Geng et al., 2017); and hyper-spectral image classification (Chen et al., 2016; Jiao et al., 2017); as well as object recognition in remote sensing images (Cheng et al., 2016; Long et al., 2017). The convolutional neural network (CNN) is able to deliver such desirable success as it is able to learn the filters performing convolutions in the image domain (Zhu et al., 2017). The GoogLeNet, AlexNet, VGG-16, VGG-19, and ResNet are the most widely used CNN framework in different computer vision tasks that address problems based on performing transfer learning or with convolution feature extraction (Zhu et al., 2017). Usually, training the CNN model can be performed in either patch-based or pixel-to-pixel-based mode. The method that uses a patch-based mode commonly starts with small image patches and trains the CNN model, following which, predictions of the class of each pixel are made based on a sliding window approach. In addition, the fully connected layers can be converted to convolution layers, without overlapping at pixel level (Paisitkriangkrai et al., 2015; Sherrah, 2016), to detect large urban objects. Pixel-based methods use an end-to-end CNN, where usually encoder-decoder architectures are used by applying methods like up-sampling, interpolation, etc. (Jiang et al., 2015; Langkvist et al., 2016). The latter approach is essential to trace fine detail of the input images. With deep learning, researchers also started focusing towards GANS for image registration (Ma et al., 2019), where the images are first translated into others by training the GANS, enabling

the two images to have similar intensity or feature information (Merkle et al., 2018).

Since many applications pertaining to climate change, agricultural monitoring, and change detection require the use of multi-temporal satellite datasets; hence, the issue of mis-registration can lead to false results. When we talk about images from different satellites, the co-registration between the images is very important for proper analysis. Image registration is a common problem in satellite image acquisition due to the differences in geometry of satellite look angle and illumination. For example, when performing change detection, the perfect co-registration of the images at very high resolution is difficult to achieve as the buildings viewed are tilted by the look angle of the satellites. Hence, the image co-registration is essential for proper analysis. The most commonly used approach for dealing with image registration in Remote Sensing is by estimating the geometric transformation between the images after identifying stable matching features. Such matching features could be point, edge, contour, or surface. Various feature descriptors such as maximally stable extremal regions (MSER) (Matas et al., 2004), Speeded-Up Robust Features (SURF) (Bay et al., 2008), Scale-Invariant feature Transform (SIFT) (Zhu et al., 2013), along with its other variants ((PSO)-SIFT (Ma et al., 2017), R-SIFT (Paul and Pati, 2016), FSC-SIFT (Wu et al., 2015), and Affine-SIFT (Morel and Yu, 2009)) have been developed for dealing with extraction of tie or key matched points. SIFT is a robust feature descriptor as it is invariant under translation, rotation, and scale change on images. Since, robustness and accuracy are the two main challenging problems in feature-based image registration process, various algorithm such as Random Sample Consensus (RANSAC) (Kim and Im, 2003), adaptive RANSAC (Hossein-Nejad and Nasri, 2017), and Fast Sample Consensus (FSC) (Wu et al., 2015) have been developed to deal with reduction of false match points by treating them as outliers. The RANSAC algorithm has been used alongside with SIFT features in dealing with cases of matching SAR images (Dellinger et al., 2015), while another use of the RANSAC algorithm is with Gaussian-Hermite moments for the purpose of image registration (Li et al., 2017a). One problem with the algorithm is the need to choose the threshold value for inlier match point detection (Ye and Shan, 2014).

Our main motivation is that accurate geographic image identification allows study and analysis of further changes or developments across the particular region using various images acquired over the time by different nanosatellites. Similarly, it also provides a platform for making the best use of archived data that lacks geolocation information. In the past, automated approaches for co-registration of images have been examined (Skakun et al., 2017; Scheffler et al., 2017; Ye et al., 2018), but the process consists of manual selection of two similar images of different satellites. Thus, our integrated model using CNN with FC2 features and the use of SIFT feature descriptors followed by application of the RANSAC algorithm allows us to accurately deal with the problem of image co-registration automatically.

3. Data and Tools Used

We have used images acquired over a study area of Abu Dhabi city using the WorldView-2, as shown in Figure 1, with 0.46 meters ground sampling distance for Panchromatic and 1.84 meters for multispectral, all images composed of eight multi-spectral bands (Coastal, Blue, Green, Yellow, Red, Red Edge, and Near Infra-Red (NIR 1 and NIR 2)) with a radiometric resolution of 11 bits per pixel. We used a Python-based VGGNet-16 CNN model in TensorFlow (Simonyan and Zisserman, 2014) to implement the identification of image location. We only used three (R, G, and B) channel satellite images to perform this experiment.

In this experiment, we used images from the Worldview-2 satellite acquired in 2014 as a reference image repository used for training purposes. The training set was a collection of ‘480 by 640’ pixel tile images obtained from the mosaic of Abu Dhabi City area, making a collection of more than 3000 images. The query images were generated from random positions, to depict real scenarios. We also generated other sets of query images, varying the resolution factor by two (meters per pixels) from that of the trained set, each containing 50 images, respectively. We also

Figure 1. Satellite image acquired over Abu Dhabi, UAE.

picked 20 such query images from DubaiSat-2, Landsat-8, and Sentinel-2 satellite and tiles from Google Earth. In addition, we also used a different scaled version as well as a slightly rotated version of these images to make the total images in each satellite category to be 50 images. The use of different satellite images that have more enhanced or diminished resolution than the WorldView-2 satellite images allows us to showcase the versatility of our approach. The images from DubaiSat-2 were acquired in 2016, image tiles from Google Earth were retrieved in 2018, images from Sentinel-2 were acquired in 2017, Landsat-8 images acquired in 2017, and Pleiades-1A images acquired in 2016 over an area of Abu Dhabi, UAE.

4. Methodology

This paper proposes a method for mapping satellite images after identification of the known closely matched referenced satellite images. The block diagram of our approach is shown in Figure 2. The input or query image consists of three channels with 480 x 640 pixel dimension where we supply RGB image. The VGG16 framework is trained using the known

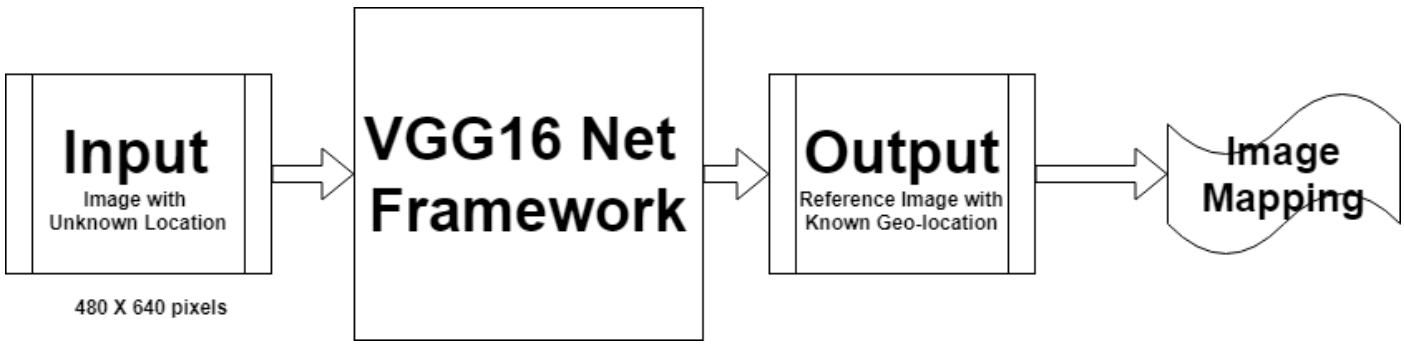

Figure 2. Block Diagram of the Methodology

geo-referenced image tile of same dimension 480 x 640 pixels. And the architecture of VGG16 is provided in Figure 3. Our approach consists of two parts: satellite image location identification and the mapping (or image co-registration).

The main contribution of this paper is to either locate the un-referenced image or rectify it by using baseline georeferenced repository using deep convolutional neural network features to find the closest match. Once the closest match image is obtained, then the SIFT features from the corresponding two images are used for co-registration process. The novelty of our work lies in the idea of using basemap tiles of known geo-referenced satellite images to either locate the geo-location of unknown satellite images or to correct the geo-location and automatically co-register the image.

4.1. Satellite Image Location Identification Using CNN

A typical CNN consists of multiple layers which repeats in turn with convolution and pooling layers, and one or more fully connected (FC) layers as seen in Figure 3. The convolution layer produces the feature maps of the previous layers with filters. The pooling layer receives smaller size rectangular blocks from the convolution layer and further sub-samples it to obtain a single output from each block. The max and average pooling are its two types. In the former, the maximum value observed in the window is sent to the next layer, whereas in the latter, average of the observed value is sent. The fully connected layer has connections to all neurons of previous layers, with each connection having its own individual weight (Kim, 2017). The VGG networks make use of filters of a small receptive field of '3×3', as compared with '5×5' and '7×7' in Alexnet (Zhu et al., 2017), and initially has the same feature map size as the number of filters in each convolutional layer of the same block; then, the size of the features in the deeper layers increases by almost twice after each max-pooling layer. Thus, the CNNs with deeper architectures can allow higher feature representations of visual data, which in response improves the accuracy of the matching task.

We began with a mosaic image of the study area that is divided into images of equal grid size which are fed as training inputs to the CNN model, where we employed an FC2 feature. VGG-16 CNN model consists of 16 layers, out of which 13 are convolutional

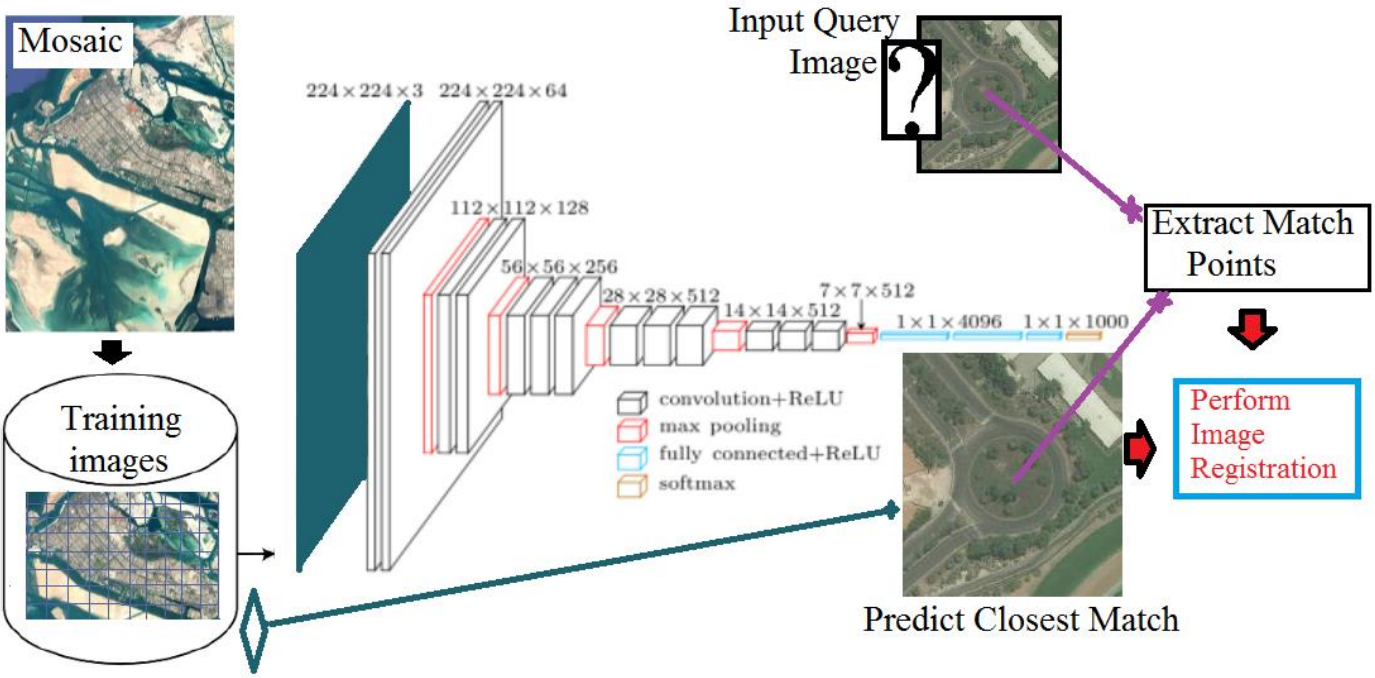

Figure 3. Architecture of the Proposed Approach.

(along with ReLU) layers, five are max-pooling layers and three are fully connected (along with ReLU) layers.

We then proceeded with a pre-processing layer, where we took RGB images with pixel values in the range of 0–255 and subtracted the mean image values calculated over the entire training set. We used a VGG-16 model with pre-defined weights to obtain the closest match between the query and trained images with the minimum Euclidean distance. Thus, this process gives us the closest match image from the referenced image dataset for the query image based on least Euclidean distance found during the matching process.

4.2. Extraction of Robust Key Match Points

We employed a Scale-Invariant Feature Transform (SIFT) on the query image and output image of CNN to detect the key match points. To select the robust key match points, we employed the RANSAC algorithm with a homography model (Naixin et al., 2017), thereby avoiding the misleading points.

RANSAC computes a homography and provides a prediction for pairs that are inliers and outliers. In SIFT, we first generated keypoints and the feature vector for each keypoint. In addition, if *k* is a keypoint in image *x* and *k′* is a keypoint in image *x′*, the feature vectors provide a way to tell whether *k, k′* are good matches. Then, for each *k*, we looked for a good match: a *k′* from *x′* that is significantly closer to k than any other keypoint in *x′*. The set of good matches forms a set of possible inliers.

Algorithm 1 shows steps of the RANSAC algorithm. Initially, we randomly picked four non-colinear feature pairs, computed a homography from these four, and then tested the suitability of this homography by checking whether the number of good matches were consistent with this homography. If *k, k′* is a good match, we assume that in most cases the homography will map k to something near *k′*. Good matches that are consistent with the homography would be treated as inliers, and those that are not consistent are considered as outliers (for this homography) with respect to the threshold chosen. By using a hit and trial method, the threshold value of 65 is found to be a good choice for dealing with the cases of our experiment of different satellite images. This process was repeated 100 times, picking a set of four good matches each time. Thus, in each iteration, we derived a homography and counted the number of outliers associated with it. Finally, we extracted the homography with the smallest number of inliers to represent our tie-points for image co-registration.

Algorithm 1: RANSAC Algorithm

1: For each four non-colinear feature pairs selected randomly:
2: Compute homography H for these features;
3: Compute inliers where Sum of Squared Distances (p_i', H_{pi})<*threshold*;
4: Keep largest set of inliers;
5: Re-compute least-squares H estimate on all of the inliers.

4.3. Evaluation Criterion

To validate the effectiveness of the proposed method, we considered Root Mean Squared Error (RMSE) taking into account all true matched points (n). The RMSE is calculated as follows, where $(x''i, y_i'')$ denotes the transformed coordinates of (x_i', y_i'):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}((x_i'' - x_i'')^2 + (y_i' - y_i'')^2)}$$

5. Experiment Results and Analysis

In this section, we demonstrate the output of the proposed method. The matched output for the query image can be seen on the right side of Figures 4 and 7. We use the term ‘accuracy’ to validate whether the matched regions between query image and output of CNN are identified. Euclidean distance is used to find the difference between the trained and queried images along the matched regions. Table 2 shows that the VGG-16 model with FC2 features is able to perform well, with 100% accuracy in the standard size dataset,

Figure 4. CNN Result of query match of images of same resolution as that of trained images (Worldview-2). Images on the left are from the query set, while images on the right are from the training set.

Table 2. Comparison of Performance with Different Dataset for Predicting the Closest Match with Reference Image

Query Sets	Year	Sensor Resolution GSD	Accuracy (%)	Avg. Eucli. Distance (pixels)
WV-2	2014	1.84 meters	100 %	28.7
WV-2	2016	1.84 meters	96 %	29.5
Pleiades-1A	2016	2 meters	96 %	30.1
DubaiSat-2	2016	5 meters	92 %	39.1
Sentinel-2	2017	10 meters	92 %	37.1
Landsat-8	2017	30 meters	94 %	38.1
Google Earth Tiles	2018	15 meters	94 %	37.6

96% accuracy in the 2016 dataset, and the Pleiades dataset. We were able to achieve 94% accuracy on the test query set of tiles from Google Earth and 92% on DubaiSat-2 images. Figures 5 and 8 show the potential against true matched key (or tie) points between the query and reference training image. We can see that the RANSAC algorithm is able to precisely predict the true tie points, removing the outliers.

When dealing with the same satellite images, both the potential and true tie points are observed to be distinct, as can be seen from first three horizontal labels in Figure 10. However, when images of Google Earth and DubaiSat-2 are tested, the tie points were found to be mostly random, as drastic changes existed between the images with respect to color and developments.

The images from WorldView-2 (2016) and Pleiades-1A (2016), being closest to the reference WorldView-2 (2014), showed a great match with 96% accuracy, as the original multispectral resolution of the images used from both satellites in this experiment is about two meters. It was also observed that the images that have similar resolution had a comparatively larger number of true tie or match points with regard to the WorldView-2 satellite in comparison with other satellites. It was also observed that average Euclidean distance of the matched images was proportionate to the ground sampling distance (GSD) of respective multispectral resolution of these satellites (as shown in Table 2, except for the case of the DubaiSat-2 satellite). The error incurred during the registration is depicted in Table 3. It was clearly observed in Table 3 that the RMSE of the registration increased as the time of acquisition of the images differed more from that of the reference image.

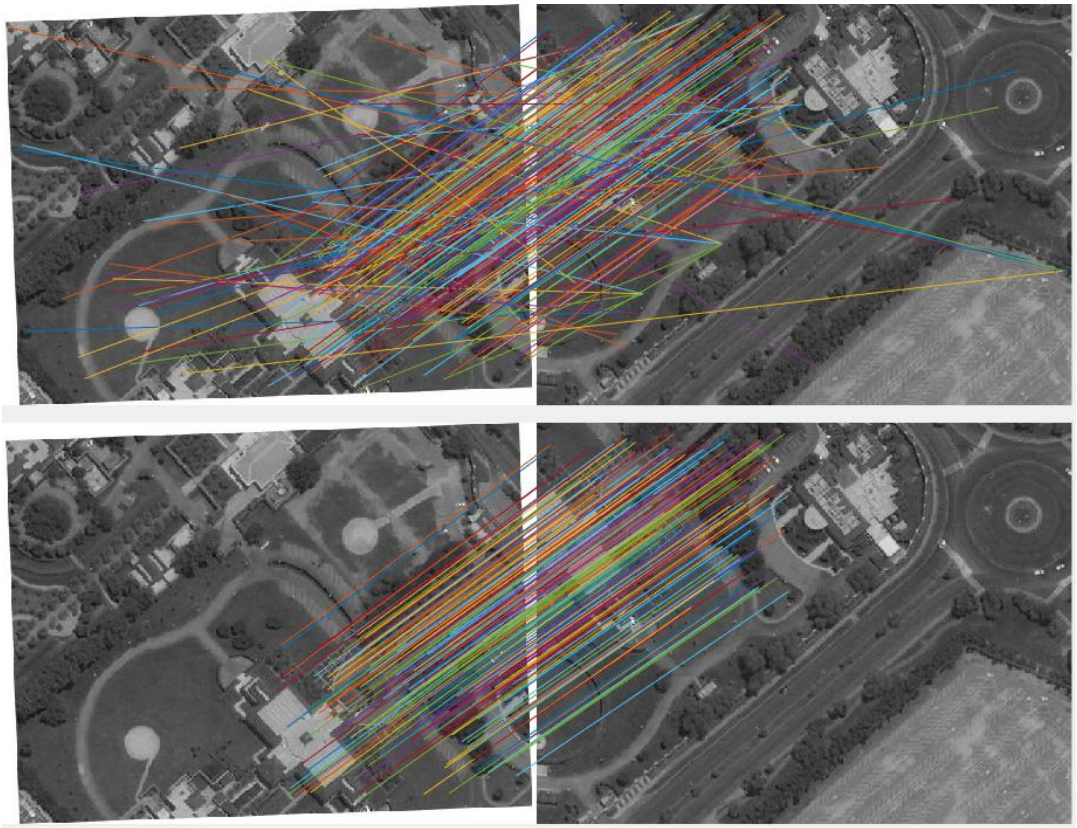

Figure 5. Top: Mapping of the potential matched points between the rotated query image in Figure 4, using SIFT features. Bottom: True matched points after RANSAC application.

Figure 6. Mosaic of rotated images of Figure 4 with the reference image.

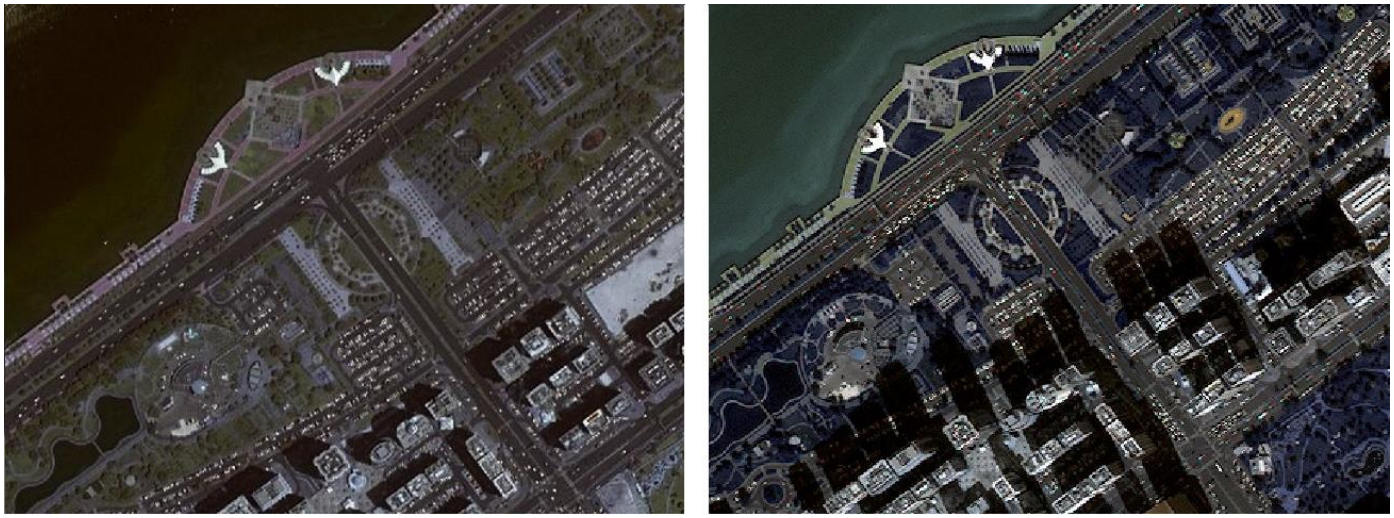

Figure 7. CNN result of query match images from DubaiSat-2 with Worldview-2. Images on the left are from the query set, while images on the right are from the training set.

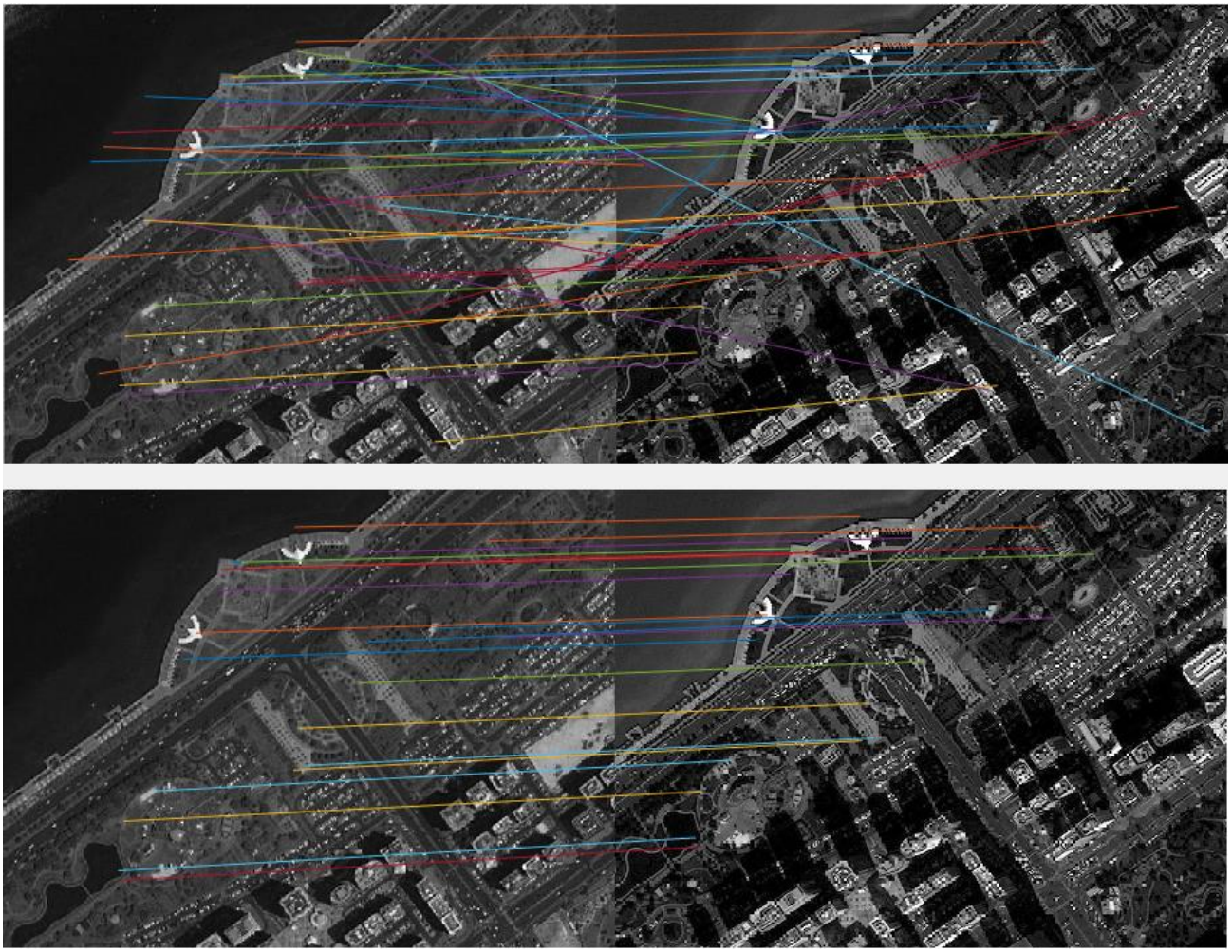

Figure 8. Top: Mapping of the potential matched points between images using SIFT features. Bottom: True matched points after RANSAC application.

Figure 9. Mosaic of images from DubaiSat-2 and Worldview-2, as shown in Figure 7.

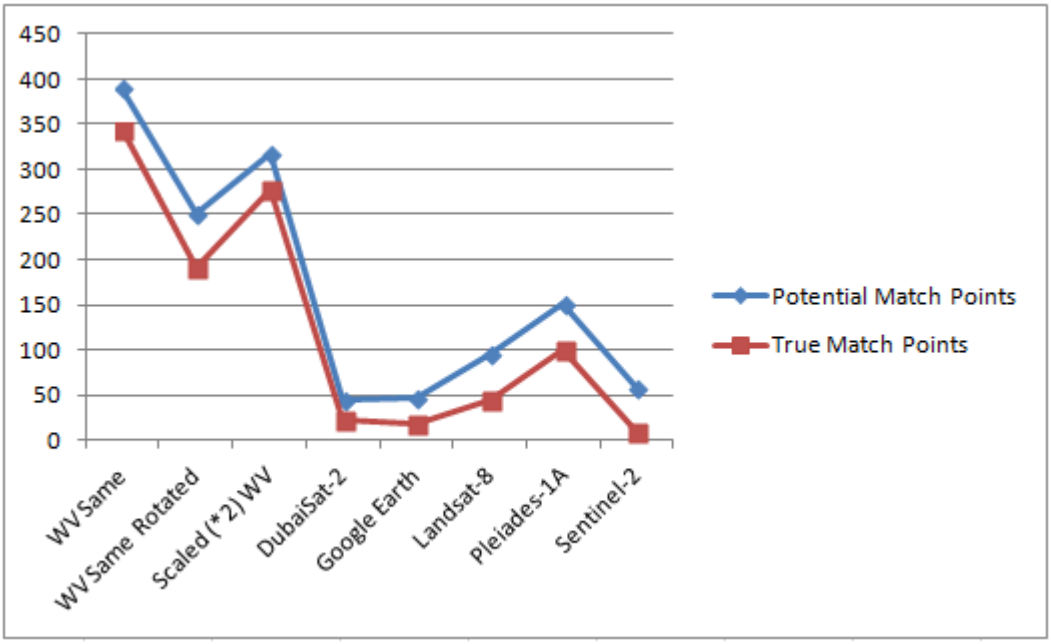

Figure 10. Average potential match points vs. True match points in different query sets.

Table 3. Comparison of RMSE for Image Registration

Query Sets	Year	RMSE (pixels)
WV-2	2014	0.45
WV-2	2016	0.71
Pleiades-1A	2016	0.88
DubaiSat-2	2016	0.91
Sentinel-2	2017	0.95
Landsat-8	2017	1.21
Google Earth Tiles	2018	1.25

6. Conclusion

In this paper, we present a novel way to rectify the problem caused by pointing inaccuracies of attitude control system in nanosatellites. We present an illustration of mapping of different satellite images with a known baseline referenced image based on fully connected features of CNN, and then provide co-registration of the image after using SIFT feature descriptors, followed by the application of the RANSAC algorithm. The output of the CNN model provides the closest match of the query image, whose location is further obtained from the known geo-coded image. We were able to obtain accuracy of above 92% for all the tested datasets. The accuracy of the results, irrespective of the date of acquisition of these satellite images, shows the versatility of our approach for image identification and co-registration. For future work, we would like to further fine tune the VGG-16 model by obtaining the weights to improve our image matching rate and also explore the RANSAC algorithm that is able to calculate the threshold value automatically.

Acknowledgement

We would like to acknowledge the support of various agencies for the data used in this work. WorldView data was acquired through Abu Dhabi Digital Authority (ADDA) and DubaiSat-2 data from Mohammed Bin Rashid Space Centre (MBRSC). We acquired Sentinel-2 and Landsat-8 data from United States Geological Survey explorer while Pleiades and Google Earth tiles were obtained from Google Earth.

References

Bay, H., Ess, A., Tuytelaars, T. et al. (2008): Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, Vol. 110(3), pp. 346–359.

Chaib, S., Liu, H., Gu, Y. et al. (2017): Deep Feature Fusion for VHR Remote Sensing Scene Classification. *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 55(8), pp. 4775–4784.

Chen, X., Steyn, W. H., Hodgart, S. et al. (1999): Optimal Combined Reaction-Wheel Momentum Management for Earth-Pointing Satellites, *J. of Guidance, Control, and Dynamics*, Vol. 22(4), pp. 543–550.

Chen, Y., Jiang, H., Li, C. et al. (2016): Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks. *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 54(10), pp. 6232–6251.

Cheng, G., Han, J., and Lu, X. (2017): Remote Sensing Image Scene Classification: Benchmark and State of the Art, in *Proc. of the IEEE*, Vol. 105(10), pp. 1865–1883.

Cheng, G., Zhou, P., and Han, J. (2016): Learning Rotation-Invariant Convolutional Neural Networks for Object Detection in VHR Optical Remote Sensing Images. *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 54(12), pp. 7405–7415.

De, S., Pirrone, D., Bovolo, F. et al. (2017): A Novel Change Detection Framework Based on Deep Learning for the Analysis of Multi-Temporal Polarimetric SAR Images, in *IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, Fort Worth, TX, July 23–28. Paper INSPEC17414201 doi:10.1109/IGARSS.2017.8128171.

de Ruiter, A. H. (2016): Observer-Based Adaptive Spacecraft Attitude Control with Guaranteed Performance Bounds. *IEEE Trans. on Automatic Control*, Vol. 61(10), pp. 3146–3151.

Dellinger, F., Delon, J., Gousseau, Y. et al. (2015): SAR-SIFT: A SIFT-Like Algorithm for SAR Images. *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 53(1), pp. 453–466.

Ge, S. and Cheng, H. (2006): A Comparative Design of Satellite Attitude Control System with Reaction

Wheel, in *Proc., First NASA/ESA Conf. on Adaptive Hardware and Systems*, AHS 2006, pp. 359–364.

Geng, J., Wang, H., Fan, J. et al. (2017): Deep Supervised and Contractive Neural Network for SAR Image Classification. *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 55(4), pp. 2442–2459.

Hossein-Nejad, Z. and Nasri, M. (2017): An Adaptive Image Registration Method Based on SIFT Features and RANSAC Transform. *Computers & Electrical Eng.*, Vol. 62, pp. 524–537.

Jiang, Q., Cao, L., Cheng, M. et al. (2015): Deep Neural Networks-Based Vehicle Detection in Satellite Images, in *2015 Int. Symp. on Bioelectronics and Bioinformatics (ISBB)*, pp. 184–187.

Jiao, L., Liang, M., Chen, H. et al. (2017): Deep Fully Convolutional Network-Based Spatial Distribution Prediction for Hyperspectral Image Classification. *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 55(10), pp. 5585–5599.

Kim, P. (2017): Convolutional Neural Network, in *MATLAB. Deep Learning*, pp. 121–147, Springer.

Kim, T. and Im, Y. J. (2003): Automatic Satellite Image Registration by Combination of Matching and Random Sample Consensus. *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 41(5), pp. 1111–1117.

Kristiansen, R., Egeland, O., and Nicklasson, P. J. (2005): A Comparative Study of Actuator Configurations for Satellite Attitude Control. *Modeling, Identification and Control*, Vol. 26(4), pp. 201–219.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012): Imagenet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, pp. 1097–1105.

Längkvist, M., Kiselev, A., Alirezaie, M. et al. (2016): Classification and Segmentation of Satellite Orthoimagery Using Convolutional Neural Networks. *Remote Sensing*, Vol. 8(4), p. 329.

LeCun, Y., Bengio, Y., and Hinton, G. (2015): Deep Learning. *Nature*, Vol. 521(7553), pp. 436–444.

Li, Y., Li, F., Yang, K. et al. (2017a): Remote Sensing Image Registration Based on Gaussian-Hermite Moments and the Pseudo-RANSAC Algorithm. *Remote Sensing Letters*, Vol. 8(12), pp. 1162–1171.

Li, Y., Zhang, Y., Huang, X. et al. (2017b): Large-Scale Remote Sensing Image Retrieval by Deep Hashing Neural Networks. *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 56(2), pp. 950–965.

Liu, H., Yang, S., Gou, S. et al. (2017): Polarimetric SAR Feature Extraction with Neighborhood Preservation-Based Deep Learning. *IEEE J. of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 10(4), pp. 1456–1466.

Long, Y., Gong, Y., Xiao, Z. et al. (2017): Accurate Object Localization in Remote Sensing Images Based on Convolutional Neural Networks. *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 55(5), pp. 2486–2498.

Ma, L., Liu, Y., Zhang, X. et al. (2019): Deep Learning in Remote Sensing Applications: A Meta-Analysis and Review, *ISPRS J. of Photogrammetry and Remote Sensing*, Vol. 152, pp. 166–177.

Ma, W., Wen, Z., Wu, Y. et al. (2017): Remote Sensing Image Registration with Modified SIFT and Enhanced Feature Matching, *IEEE Geoscience and Remote Sensing Letters*, Vol. 14(1), pp. 3–7.

Matas, J., Chum, O., Urban, M. et al. (2004): Robust Wide-Baseline Stereo From Maximally Stable Extremal Regions, *Image and vision computing*, Vol. 22(10), pp. 761–767.

Merkle, N., Auer, S., Müller, R. et al. (2018): Exploring the Potential of Conditional Adversarial Networks for Optical and SAR Image Matching. *IEEE J. of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 11(6), pp. 1811–1820.

Morel, J. M. and Yu, G. (2009): ASIFT: A New Framework for Fully Affine Invariant Image Comparison. *SIAM J. on Imaging Sciences*, Vol. 2(2), pp. 438–469.

Naixin, Q., Shengxiu, Z., Lijia, C. et al. (2017): A Robust to Outliers Method for Estimating the Homography, in *Proc., 29th Chinese Control and Decision Conf. (CCDC)*, pp 1586–1590. Chongqing, China, May 28–30. Paper INSPEC17041812 doi:10.1109/CCDC.2017.7978770.

Nudehi, S. S., Farooq, U., Alasty, A. et al. (2008): Satellite Attitude Control Using Three Reaction Wheels, presented at the *IEEE American Control Conf.*, Seattle, WA, June 11–13. Paper INSPEC10060718 doi:10.1109/ACC.2008.4587262.

Paisitkriangkrai, S., Sherrah, J., Janney, P. et al. (2015): Effective Semantic Pixel Labelling with Convolutional Networks and Conditional Random Fields, in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2015 IEEE Conf.*, pp. 36–43.

Paul, S. and Pati, U. C. (2016): Remote Sensing Optical Image Registration Using Modified Uniform Robust SIFT. *IEEE Geoscience and Remote Sensing Letters*, Vol. 13(9), pp. 1300–1304.

Ping, Z., Dong, Y., Tang, H. et al. (2016): Robust Approach for Attitude Tracking and Nonlinear Disturbance Rejection of Rigid Body Spacecraft. *IET Control Theory & Applications*, Vol. 10(17), pp. 2325–2330.

Scheffler, D., Hollstein, A., Diedrich, H. et al. (2017): AROSICS: An Automated and Robust Open-Source Image Co-Registration Software for Multi-Sensor Satellite Data. *Remote Sensing*, Vol. 9(7), p. 676.

Scott, G. J., England, M. R., Starms, W. A. et al. (2017): Training Deep Convolutional Neural Networks for Land—Cover Classification of High-Resolution Imagery. *IEEE Geoscience and Remote Sensing Letters*, Vol. 14(4), pp. 549–553.

Sherrah, J. (2016): Fully Convolutional Networks for Dense Semantic Labelling of High-Resolution Aerial Imagery. Available at: https://arxiv.org/abs/1606.02585 (accessed May 24, 2020).

Simonyan, K. and Zisserman, A. (2014): Very Deep Convolutional Networks for Large-Scale Image Recognition. Available at: https://arxiv.org/abs/1409.1556 (accessed May 24, 2020).

Skakun, S., Roger, J. C., Vermote, E. F. et al. (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. *Int. J. of Digital Earth*, Vol. 10(12), pp. 1253–1269.

Wertz, J. R., Everett, D. F., and Puschell, J. J. (2011): *Space Mission Engineering: The New SMAD*, Microcosm Press.

Wu, Y., Ma, W., Gong, M. et al. (2015): A Novel Point-Matching Algorithm Based on Fast Sample Consensus for Image Registration. *IEEE Geoscience and Remote Sensing Letters*, Vol. 12(1), pp. 43–47.

Xia, G. S., Hu, J., Hu, F. et al. (2017a): AID: A Benchmark Data Set for Performance Evaluation of Aerial Scene Classification. *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 55(7), pp. 3965–3981.

Xia, X., Sun, G., Zhang, K. et al. (2017b): NanoSats/CubeSats ADCS Survey, presented at the *IEEE Control and Decision Conf. (CCDC)*, Chongqing, China, May 28–30. Paper INSPEC17041432 doi:10.1109/CCDC.2017.7979410.

Yao, X., Tao, G., Ma, Y. et al. (2016): Adaptive Actuator Failure Compensation Design for Spacecraft Attitude Control. *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 52(3), pp. 1021–1034.

Ye, F., Su, Y., Xiao, H. et al. (2018): Remote Sensing Image Registration Using Convolutional Neural Network Features, *IEEE Geoscience and Remote Sensing Letters*, Vol. 15(2), pp. 232–236.

Ye, Y. and Shan, J. (2014): A Local Descriptor Based Registration Method for Multispectral Remote Sensing Images with Non-Linear Intensity Differences. *ISPRS J. of Photogrammetry and Remote Sensing*, Vol. 90, pp. 83–95.

Zhou, Z. (2010): Spacecraft Attitude Tracking and Maneuver Using Combined Magnetic Actuators, in *AIAA Guidance, Navigation, and Control Conf.*, p. 7899.

Zhu, G., Wang, Q., Yuan, Y. et al. (2013): SIFT on Manifold: An Intrinsic Description. *Neurocomputing*, Vol. 113, pp. 227–233.

Zhu, X. X., Tuia, D., Mou, L. et al. (2017): Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience and Remote Sensing Magazine*, Vol. 5(4), pp. 8–36.