DETECTION BASED ORTHOMOSAIC REGISTRATION LITERATURE REVIEW

TREEREG Project

Damian Wilson
Department of Computer Science
University of Cape Town
wlsdam001@myuct.ac.za

ABSTRACT

Aerobotics, a company in the Precision Agriculture sector, provides agricultural trend analysis to farmers through the analysis of large orthomosaic maps of orchards over time. In order for successful trend analysis to occur, orthomosaic maps of the same orchard area, taken at different times, need to be aligned correctly. The process of aligning different images of the same scene is called image registration. The field of image registration has grown a lot since the late 20th century, with many techniques being developed and modified for a wide range of practical applications. In this review, existing registration methods are analyzed and evaluated with the purpose of finding the best possible overall image registration strategy to align the multitemporal orthomosaics produced by Aerobotics. Aerobotics have already developed a robust feature extractor, a tree instance segmentation model, to aid this process. Taking this into account, the best overall registration strategy was found to be a feature-based registration method, using the correspondence of invariant feature descriptions to correctly align the input images. A concern regarding the robustness and accuracy of this proposed method arose. This was due to trees being selected as the detected feature to be used as the basis for matching in the registration. Their inherent temporally unstable nature has the potential to violate the invariant requirement of the feature description. Therefore, an additional feature tracking step, making use of a geospatial database and probabilistic model, was proposed to ensure the accuracy of the overall registration.

KEYWORDS

Image Registration, Image Mosaicking, Orthomosaic Registration

1 INTRODUCTION

Precision Agriculture is a current trend in the agricultural sector to optimize farming management practices. This is done through the analysis of real-time observational data of agricultural fields obtained with the aid of various Information Technology (IT) systems. Modern advancements in the field of the Internet of Things (IoT), specifically the widespread use of sophisticated Unmanned Aerial Vehicles (UAVs) or drones, are a driving force behind the viability of current Precision Agricultural practices. This is due to their ability to take high spatial and temporal resolution images of farmland on demand [1, 4]. Which are subsequently processed, aligned and combined to produce orthomosaic maps of a large stretch of farmland, a key input in agricultural analysis applications. The process of aligning overlapping individual images is called *image registration*. Despite recent experiments proving the high level of positional precision in orthomosaic maps derived from images obtained by UAV photogrammetry [2], the level of proliferation of UAVs into the Precision Agriculture sector has not met expectations. This is partly due to the difficulty of choosing and implementing data acquisition and image processing techniques, as this is a relatively new field with no standardized workflow or best practice [1]. As such, each company in this field must develop and implement their own registration methodology for the creation and analysis of orthomosaic maps that meets their requirements. Aerobotics, a stakeholder in this project, is one such company.

Aerobotics is a South African company in the Precision Agriculture sector that uses multitemporal UAV based imagery to provide farmers with information of trends on their farms over time. This is done through the creation and subsequent analysis of orthomosaic maps of orchards over time. Due to the inherent error in the GPS tagging of the UAV captured images, orthomosaic maps may be misaligned between dates, hindering the analysis process. To aid the registration process of the orthomosaics, they have developed a tree instance segmentation model that creates a bounding polygon around individual tree canopies on the map. This salient feature extractor is intended to aid the feature matching stage of the overall image registration process [3].

The aim of this paper is to review existing image registration and mosaic construction techniques to improve the process of orthomosaic registration for Aerobotics, by reducing misalignment. Image Registration is an established field of research encompassing many various implementations which date back to the late 20th century with many practical applications, of which mosaic construction is but one [3]. As such, experiments regarding similar use cases in the agricultural field will be analysed with the hope of extracting relevant methods that can be used for this project. Consideration will be taken for how these extracted techniques can best be integrated with Aerobotics' existing feature extractor. Three topics, relevant to the aim of the project will be discussed. These are: Image registration, where different classical and modern registration methods are discussed. Next will be image mosaic construction, where different methodologies for the creation of image mosaics (using various registration methods) will be analysed, and relevant techniques to this case will be extracted. And finally, the topic of object tracking over time will be discussed, which will concern the analysis and tracking of individual trees over different orthomosaic maps in this case. This last section is relevant to the project as trees are not static objects, they grow, shrink, wither, bloom or move occasionally. Therefore, special consideration has to be taken to ensure that the same tree can be tracked over time for the sake of aligning multitemporal orthomosaics. The results and interaction of these topics will be discussed and from this, conclusions regarding the overall findings of this review will be drawn.

2 IMAGE REGISTRATION

Image registration is the process of aligning and overlaying two separate images of the same or adjacent scene(s). These images could be captured from different viewing angles, sensors or times [3, 5, 7] These two images are the reference and sensed image. The end result of image registration is a final, combined image in which the sensed image has been transformed to geometrically align with the reference image and now overlays it [3]. Image registration is a crucial component in the overall aim of this project for two reasons. Firstly, image registration is used to construct the orthomosaic maps of the agricultural land used by Aerobotics. These orthomosaics are produced through the process of stitching together multiple overlapping images taken of a particular scene of interest during one or more UAV flights over the area. The images captured on the flight are each a part of the required whole map. Due to the nature of the UAV flight, each of these images capture the land at a different viewpoint, time and rotation. Therefore, these images have to be individually transformed through registration to produce a consistent and viable orthomosaic. The images are registered among each other and overlayed, resulting in a single orthomosaic map covering the entire area of interest [1]. The second reason image registration is crucial is that the resulting orthomosaic maps of different dates then have to be registered among each other. This is so that they are correctly aligned to enable subsequent trend analysis of the relevant changes observed on the farm over time, a key part of Precision Agriculture [1].
Image registration is a well-researched field, with many methods published dating back to the late 20th century [6]. The first comprehensive survey of image registration methods was produced by Brown in 1992 [6]. The next robust survey on image registration techniques was produced in 2003 by Zitova & Flusser [3]. This survey refined the methods outlined in the original 1992 survey and included new techniques published since then. The final survey on image registration techniques being reviewed is by Saxena & Singh, produced in 2014 [5]. This paper covers more recent developments in the field of image registration methodology since the 2003 survey. Due to the diversity of possible images to be registered, and all the possible levels of distortion in each image, it is impossible to define a universal method for image registration [3]. However, in the majority of the registration methods described in these papers, the process of image registration can be divided in to 4 sequential steps: *feature detection*, *feature matching*, *transform model estimation* and *image transformation and resampling* [3, 5, 6, 7]. Zitova & Flusser classify all registration methods in to two classes based on how the features are detected in the first step [3]. Registration methods are either area-based or feature-based.
Relevant literature on each of these steps, extracted from the robust survey papers listed above, will be reviewed below. Various approaches to each of these steps, relevant to the aim of this project, will be analyzed without going into significant depth on their individual algorithmic implementation.

2.1 Feature Detection

Feature detection is the first step of the image registration process. In this step salient and distinctive objects in the image are detected. These features could be closed-boundary regions (i.e., a tree canopy) lines, line intersections, corners etc. [3]. Aerobotics has already developed a feature detector to aid this project, called a tree instance segmentation model. This model detects trees in the image and identifies them by producing a bounding polygon over their canopy. In that case, the trees are the features. These detected features are represented by geometric point representatives on the image, usually being the end points on a line feature or the center of mass on a region feature (such is the case for a bounding polygon over a tree canopy) [3, 5]. Note that these detected features are referred to as *control points* (CPs) in the reviewed literature.

Feature detection methods differ according to the classification of the overall registration method (area-based or feature-based).

2.1.1 Area-based detection

Area-based registration methods put more emphasis on the feature matching step of the process than the detection phase [5]. These types of methods use an entire area of the image (in the form of a block of pixels) as the "feature" to be matched in the next step. As such, no features are detected in this step and this stage is skipped [3]. Refer to the area-based matching subsection (2.2.1) for more details on how these registration methods are implemented. Note that area-based registration methods are generally only used when the images being registered lack distinctive shapes or structures to enable feature-based registration [3]. In this project, the input images contain distinctive shapes in the forms of tree canopies and a feature detector has already been developed for the project. Therefore, a feature-based (or possibly a hybrid feature and area based) approach should be used for this project.

2.1.2 Feature-based detection

Feature-based registration methods require the extraction and description of salient features (CPs) in this detection step, as described above. Common classes of features to be detected here are regions, lines and points [7]. Zitova & Flusser state that features should be distinct, spread out and efficiently detectable, despite any potential distortion or degradation in the reference and sensed image [3]. An adequate level of comparability between the feature sets of the images, required in the feature-based matching stage in the next step, is dependent on the invariance and accuracy of the feature detector and a sufficient level of overlap or intersection of the detected feature sets of the images [3]. Taking this requirement in to account, a logical choice for the choice of detected feature in this project is a tree canopy. Presumably, the only distinct and comparable element of the individual UAV-captured images of orchards would be the tree canopy (note that at this point of the project I have not seen the input data images, hence the presumption). This sentiment is clearly shared by Aerobotics as they have already created a feature detector to detect and segment tree canopies. A tree canopy falls under the region class of features due to its size, relatively high contrast and closed boundary nature in the drone captured aerial imagery of orchards [8]. Methods for detecting these types of features will be briefly reviewed, as this stage of the process has presumably already been implemented by Aerobotics adequately with their tree instance segmentation model. Region features are detected via segmentation methods, with the result usually being a bounding polygon or convex hull shape around the detected feature. With the latter result requiring a clearer input image and more processing time [9]. The accuracy of the segmentation is of vital importance to the subsequent feature matching step. The more precise and refined the segmentation is, the more accurate the matching will be. A classical approach to ensuring accuracy and improving quality in region-based segmentation was introduced by Goshtasby et al. [8]. The segmentation of the image was done iteratively, estimations of the control points (CPs) correspondence to one another were used to refine the segmentation parameters in each iteration. They claimed this process could achieve sub-pixel accuracy in the derivation of control points of the regions (which would be the centers of mass for the tree region), which would improve the precision of the contours generated around the region. This level of accuracy is especially important in this project as the shape of the trees in the input image are generally very similar, therefore utmost precision in the segmentation process is essential for accurate feature matching. Other reviewed approaches to region-based detection include the use of virtual circles as region features, proposed by Alhichri & Kamel [10]. The advantage of this approach being that it improves registration efficiency for images with translation and scale differences. However, a main differentiating factor of the features detected is the radius of the enclosing circle, which proves a challenge when comparing multiple tree canopies with very similar radii during feature matching. A more recent approach to feature detection, called shape feature detection was outlined in a survey by Minqiang et al. [11]. These methods have objects identified by their unique shape outline (i.e., a tree crown) and has the advantage of being relatively invariant to noise or deformation and efficient in registering images differentiated by translation, rotation and scaling. If the existing, provided segmentation model proves insufficient for feature matching, these approaches can be looked at for improving the precision of detected features.

2.2 Feature Matching

This step in the registration process involves matching the corresponding features in the intersection of detected feature sets in the reference and sensed image [3]. This means the process is concerned with determining if there are common features in the two images and matching the ones that correspond. The results of this step are used to estimate the parameters of the mapping function in the next step. These matching methods are divided into two major categories, according to the type of registration occurring. Relevant literature and techniques will be reviewed for each, without going in depth on their respective algorithmic implementations.

2.2.1 Area-based matching

All area-based feature matching techniques omit the previous detection step of registration as they do not require the extraction of salient features as an input [5]. Instead, windows of predetermined size or entire images are used for correlation estimation based on their intensity values in close neighbourhoods [3, 12]. A classic example of area-based methods is the cross-correlation (CC) method [3]. This method calculates a similarity metric between pairs of windows from the referenced and sensed image. The pair of windows for which the maximum level of similarity is found are set as matching [5]. This method is mainly used for pattern matching and works best for when the difference between the image pair is only a matter of translation, but can work for when scaling and rotation transformations are present too. However, the generalized CC methods to match the more geometrically deformed images come at a significantly higher computational cost [3]. As such, this method is not especially suited to the context of this project, as relatively significant geometrical deformations are implicit with the aerial drone images Aerobotics uses. An improvement to the traditional CC method was proposed by Evangelidis & Psarakis in 2008 called the Enhanced Correlation Coefficient algorithm [13]. This newer method has the desirable modification of being invariant with respect to a wide range of photometric distortions. Theoretically making it better suited to register the required UAV-captured images for this project and a legitimate contender for our overall registration strategy. The ECC method's practical applications and validity for this context have been shown in the framework for registering UAV-based imagery for crop-tracking proposed by Lopez et al. [4]. This framework proposed a multi-layer model for registering hyperspectral aerial images of crop fields. They proved that the ECC method is well suited to register heterogeneous images (specifically images of different spectral bands). The ECC method was shown to handle a wide range of typical distortions of UAV-based imagery after preprocessing the data to remove any lens based distortion (i.e.. fish-eye distortion). Assuming Aerobotics uses hyperspectral images to construct their orthomosaics, this is an extremely viable strategy for registration. In the likely case that plain RGB images will be the entire scope for this project, the ECC method can still be used effectively, provided any lens-based distortion is removed first [4].

2.2.2 Feature-based matching

Feature based matching involves analyzing the two sets of salient features detected in the reference and sensed image from the previous step [3]. The features are represented by their control points (CPs), which would be the center of mass of the region covered by individual tree canopies in this case. The aim of this step in the process is to use the find the correspondence between the elements of the two detected feature sets, thereby finding matching pairs. This is done through analyzing the spatial relations between the CPs or the various descriptors of the features themselves [3, 5]. Therefore feature-based matching methods being reviewed can be divided into methods using spatial relations and methods using invariant feature descriptors.

Spatial relation-based feature matching methods are typically used when the detected features are ambiguous or if their nearby neighborhoods are distorted [3]. In this case, the information about the distance between the CPs are used to match features. In the case that the segmentation of tree instances done in the detection step does not produce sufficiently unique feature descriptors (this would be the bounding polygon produced from Aerobotics' own feature detector), spatial relation or even area-based matching methods would have to be used. Various spatial-based methods have been introduced over the years. Goshtasby proposed a method for matching CPs from two different planes using a graph matching algorithm, in which the transformation parameters are estimated through the iterative transformation of a subset of points from the sensed to the reference image, deriving which parameters result in the best match [14]. This method does not guarantee the utmost accuracy in the subsequent transformation step as only a subset of points are used to estimate the transformation parameters, if the whole set were to be used, it would be the equivalent of an expensive exhaustive search algorithm. Öfversted et al. introduced a more robust matching method using spatial relations [15]. This newer method uses both spatial and image intensity information to match features. This approach can be thought of as a combination of area-based and spatial relation-based matching methods and the result is a faster, more accurate and noise resistant approach to registration (relative to stand alone are or spatial based methods). This approach has great potential in the context of this project. Area-based registration methods, like ECC mentioned above, have proven to produce accurate orthomosaics from UAV-based imagery, especially for heterogeneous images [4]. If the images used in this project turn out fairly homogenous (all RGB with very similar image intensity values) then the addition of a spatial information between the detected features should make a hybrid area-based registration approach viable and efficient.

Invariant descriptor-based feature matching methods match features based on the correspondence of their description [3, 5]. In this case, using the tree instance segmentation model developed by Aerobotics, each detected feature's description would be the bounding polygon surrounding it, generated by the model. The features (trees in this instance) would be matched based on the similarity of their bounding polygon. Zitova & Flusser state that the feature description should satisfy 4 conditions in order for descriptor-based feature matching to be successful [3]. These being invariance (the description of a common feature in the reference and sensed image should be the same), uniqueness (different features should have different descriptions), stability (deformations in the source image should not drastically change the generated description) and independence (vector-based descriptions should be functionally independent). If Aerobotics' provided feature detector meets these requirements, descriptor-based matching methods are a viable approach to this registration project. A classic, simple approach to feature descriptors is using the image intensity function of the feature are as it's description [16]. More modern and robust feature descriptors include the Scale Invariant Feature Transform (SIFT) method [5, 17] and newer faster Speeded Up Robust Feature (SURF) method [5, 18]. Note that since a robust feature descriptor has already been developed for this project, these alternative descriptor methods will not be reviewed in depth. If the provided descriptor does not meet the accuracy requirements, these alternative approaches will be investigated.

2.3 Transform Model Estimation

After correspondence between matching features in the reference and sensed image has been established in the previous step, a mapping function be constructed. This function should transform the sensed image so that it is overlayed on the reference one, while ensuring that once overlayed, the matching features in the images are as close as possible [3]. An appropriate mapping function must be chosen, with it's parameters estimated to ensure this requirement is best met. The choice of mapping function depends on the assumed level of geometric distortion in the images, the method of image capture and the required accuracy of the registration [3]. Images being captured from UAV's are prone to scale, translation as well as rotational transformational differences from one another [4]. And the level of required accuracy of the registration must be sufficiently high enough to enable the effective trend analysis performed by Aerobotics. Therefore, mapping functions that account for the transformational differences mentioned, while ensuring the best possible level of accuracy will be considered. Mapping function models can be divided in to 2 categories, *local mapping models* and *global mapping models* [3, 5]. The key difference being that global mapping models use all corresponding sets of control points to estimate a single set of mapping function of parameters that are used to transform the entire sensed image. Whereas local mapping models are based on treating the input image as a collection of individual patches, with each patch requiring it's own set of mapping function parameters to be accurately transformed [3, 5]. Local mapping models are only generally used when significant, localized distortions in the image are observed, requiring separate mapping parameters in each of the differently distorted areas. The UAV-captured images for this project presumably do not have these localized distortions present. The results of experimentation with UAV-based imagery by Lopez et al. showed that the distortion experienced in these types of images mostly relate to the sensor taking them (i.e., The fish-eye effect), which are not localized to specific image regions [4]. They were able to filter out the global, sensor-induced distortion and then proceed with a global mapping function to produce accurate results. Therefore, global mapping functions should be used in this project too. Bivariate polynomials of low degrees are frequently used as global transformation models. Three types of these models, from least general and computationally intensive to most, are *similarity transform*, *affine transform* and *perspective projective transform* [3, 19]. Similarity transform works well for simple scaling, translation, and rotation transformations only. Affine transform handles multi-view registration better, given that the sensor is a suitable distance away from the field, proportional to the size of the field. Perspective projection models are used when the distance requirements of affine transformation are not met [3]. Due to this, depending on the altitude of the UAV taking the pictures, either affine transformation or perspective projection models should be used for this project.

2.4 Image Transformation

The final step in the overall registration is the transformation and resampling of the sensed image to overlay the reference image, using the mapping functions defined in the previous step [3, 5, 19]. This step involves transforming each pixel in the sensed image using either a single, global mapping function or a local mapping function (the parameters of which depend on region of the image the pixel falls under) [5]. An intuitive, forward-based approach to this image transformation is to directly transform each pixel in the sensed image, using the mapping function, to lie on the reference image coordinate system [3]. However, this approach can lead to overlapping transformed pixels or holes in the transformed image. This is due to rounding or discretization in the coordinate mapping process [3]. This is undesirable in this project as Aerobotics requires accurate and complete othomosaics for analysis purposes. Therefore, an alternative interpolative approach to transformation and resampling should be used. This approach uses the target pixel coordinates of the mapping function (on the reference coordinate system), as well as the inverse of the mapping function to transform the sensed image pixels [3]. This enables interpolation of the sensed pixels on the reference image grid, ensuring no overlaps or holes can occur. Various interpolation methods have been used for this step of image registration. Common ones in include the *nearest neighbour*, *bilinear* and *bicubic* functions [20]. Nearest neighbour methods are relatively out of date with newer interpolative approaches. Bilinear functions are shown to provide a good tradeoff between computational complexity and output accuracy, with higher order functions like cubic methods, providing better accuracy at a greater computational cost [3]. Depending on the computational and time complexity requirements of this project either bilinear or bicubic functions (along with their modern refinements and alternatives) should be considered for this stage of the registration.

3 IMAGE MOSAICKING METHODOLOGIES

UAV-based orthomosaic images are created through the alignment of several overlapping images taken of a particular scene of interest through the registration of these individual images, this is called image mosaicking [21]. The result of this process is an undistorted map with a field of view encompassing the input images that preserves the original resolution of these input images [22]. As the aim of this project is mainly about the registration of pre-constructed orthomosaics to enable subsequent trend analysis, literature over image mosaicking methodologies will be lightly reviewed without going into significant depth over their implementation. Instead, a brief overview over different types of mosaicking methodologies will be provided for the sake of providing useful context to the orthomosaic theme of this project. The overall mosaicking process can be divided in to 4 steps of image processing: *registration*, *reprojection*, *stitching* and *blending* [21]. The reprojection and stitching steps mentioned here relate to the transform model estimation and transformation steps of the registration process mentioned above. The blending step is applied after the individual input images have been stitched together to minimize any discontinuities in the global appearance of the mosaic (caused by misalignment errors) [21]. Ghosh & Kaabouch propose that mosaicking algorithms can be classified according to their registration and blending method [21]. As such, the orthomosaic method used by Aerobotics are presumably in the feature-based registration class of mosaicking algorithms, assuming their provided feature descriptor was used in the feature matching process (2.2). Existing mosaicking algorithms rely on inter-image holography computation, to produce an accurate registration, based on the assumption the area of interest is a planar scene [23]. However, this assumption only holds when the images are taken from a sufficiently high altitude. Li & Isler proposed a new novel mosaicking technique that removes this limitation and showed that it produced efficient and accurate results, specifically in an agricultural context [23]. This method could be employed in the mosaicking process used by Aerobotics to ensure accurate results are produced from images captured at low altitudes. More recent trends in the mosaicking field include the use of enterprise software to automate the mosaicking process, from drone flight planning to automated orthomosaic generation [2, 24, 25]. The image processing platform “Pix4D Mapper Pro” seems to be a popular choice for this, which is based on the structure-from-motion algorithm. Hung et al. showed that orthomosaics derived this way exhibit a high level of positional precision, making it a viable option for producing orthomosaics, if the project requires [2]. Another recent trend in the mosaicking field is the emergence of hyperspectral or multispectral orthomosaics, combining images taken from different sensors [4, 24, 25, 26]. This is especially used in the field of Precision Agriculture as the wide array of spectral and thermal information provided by these images provide valuable insights in to crop health and growth. Lopez et al. proposed a multi-layered framework for the registration and mosaicking of multispectral images using the ECC feature matching algorithm (see section 2.2.1). The incorporation of multispectral mosaicking is left as proposed future work, currently out the scope of this project.

4 FEATURE TRACKING

The recommended orthomosaic registration approach for this project is a feature-based registration method using an invariant descriptor for feature detection (which is Aerobotics' own tree instance segmentation model). An issue arising from this approach stems from the rule that the choice of detected feature should be stable in time and stay in fixed positions throughout all images [3]. This property is required for the feature descriptions generated by the feature detector to be invariant in all the multitemporal orthomosaic input images being registered. This is necessary for an accuracy in the feature matching step of registration [3]. However, trees are not fixed in size and position by nature, they can grow, wither and shift over time. Therefore, in order to ensure the accuracy of the registration, additional feature tracking should be implemented to ensure that the feature descriptions of the same tree remain invariant over time. A basic requirement for this tracking of trees is a storage model for the tree description. Lopez et al. proposed a geospatial database model for this purpose [4]. This model records details about the feature as well as it's generated identifying description (which was a bounding polygon in their experiment, just like the descriptions generated by our descriptor). They showed that this model is effective for tracking growth and change in crops overtime. Therefore, a similar model to theirs can be used effectively in our project for feature tracking. Using this database model, each unique feature detected in the input mosaics being registered would have a unique ID with a field for it's description (which would be the contours of it's bounding polygon in this case). Additional fields that should be added to improve tracking are GPS coordinates of the feature's CP (control point) and it's spatial relation other CPs. Hung et al. showed that GPS coordinates of individual CPs in a mosaic can be accurately derived using the GPS metadata of the input images [2]. And spatial relation between CPs can be used to determine matching CPs (section 2.2.2) [3, 5]. The population of this database is intended to occur during the feature matching stage of registration. As images are detected, the mentioned fields will be populated. Some sort of statistical and probabilistic analysis should then determine if the detected features match any other features in the database based on the similarity of their data fields. A probabilistic model that could be used for this is a *Bayesian Network*, as it has been shown to accurately track and match objects in videos [27-29]. This video object tracking is a relatively similar concept to the multitemporal feature matching in images required for this project.

5 CONCLUSIONS

Various registration techniques, relevant to the aim of this project were reviewed and their suitability to the requirements of this project were discussed. The results of these discussions formulated the following conclusions: The best method for the registration of the orthomosaics produced by Aerobotics is a feature-based registration method. Specifically, one that uses the provided tree instance segmentation model to detect salient tree objects in the images and produce a bounding polygon description of the tree in the feature detection phase. Following this, it should then perform feature matching based on the similarity of the invariant description of the detected tree instances (this being the bounding polygon) in the feature matching step. Next, an affine transformation model should be used to transform the sensed image on to the reference in the transform model estimation step. And finally, a bilinear or bicubic interpolation model should be used to resample the sensed on to the reference image during the final transformation stage of the registration process.

An observed gap in the literature reviewed is that all the feature-based registration methods that were analyzed required that the feature being detected be stable through time. Meaning it's generated description in the detection phase must be invariant in all the images being registered, even if they're taken at different dates. This is not the case in this project due to trees being used as our features. Trees could very well change over time as they have the potential to grow, wither or even move. More research could be done in the field of registering images using features with significant potential to change over time. An additional step of feature tracking is proposed as a potential solution for this challenge. This involves storing each uniquely detected feature, along with it's characteristics like it's bounding polygon description, coordinates and inter-feature spatial relations, in a geospatial database. Statistical and probabilistic models (like a Bayesian Network) will then use this data to aid the feature matching process, ensuring a robust and accurate feature-based registration process despite the use of temporally unstable trees as features.

In conclusion, there is a lack of feature-based registration techniques using temporally unstable features as control points. This presents an opportunity for this project to contribute to the image registration field by filling this knowledge gap with the work it will be conducting.

REFERENCES

[1] Tsouros DC, Bibi S, Sarigiannidis PG. A Review on UAV-Based Applications for Precision Agriculture. Information. 2019; 10(11):349. DOI: 10.3390/info10110349

[2] Hung I-K, Unger D, Kulhavy D, Zhang Y. Positional Precision Analysis of Orthomosaics Derived from Drone Captured Aerial Imagery. *Drones*. 2019; 3(2):46. DOI: 10.3390/drones3020046

[3] Zitova, B., & Flusser, J. (2003). Image registration methods: a survey. Image and vision computing, 21(11), 977-1000. DOI: 10.1016/S0262-8856(03)00137-9

[4] López, A., Jurado, J. M., Ogayar, C. J., & Feito, F. R. (2021). A framework for registering UAV-based imagery for crop-tracking in Precision Agriculture. *International Journal of Applied Earth Observation and Geoinformation*, *97*, 102274. DOI: 10.1016/j.jag.2020.102274.

[5] Saxena, S., & Singh, R. K. (2014). A survey of recent and classical image registration methods. *International journal of signal processing, image processing and pattern recognition*, *7*(4), 167-176.

[6] Lisa Gottesfeld Brown. 1992. A survey of image registration techniques. ACM Comput. Surv. 24, 4 (Dec. 1992), 325–376. DOI: 10.1145/146370.146374

[7] K. Sharma and A. Goyal, "Classification based survey of image registration methods," 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), 2013, pp. 1-7, DOI: 10.1109/ICCCNT.2013.6726741.

[8] A. Goshtasby, G. C. Stockman and C. V. Page, "A Region-Based Approach to Digital Image Registration with Subpixel Accuracy," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-24, no. 3, pp. 390-399, May 1986, DOI: 10.1109/TGRS.1986.289597.

[9] Pal, N. R., & Pal, S. K. (1993). A review on image segmentation techniques. *Pattern recognition*, *26*(9), 1277-1294. DOI: 10.1016/0031-3203(93)90135-J.

[10] Alhichri, H. S., & Kamel, M. (2003). Virtual circles: a new set of features for fast image registration. *Pattern Recognition Letters*, *24*(9-10), 1181-1190. DOI: 10.1016/S0167-8655(02)00300-8.

[11] Mingqiang, Y., Kidiyo, K., & Joseph, R. (2008). A survey of shape feature extraction techniques. *Pattern recognition*, *15*(7), 43-90.

[12] R. J. Althof, M. G. J. Wind and J. T. Dobbins, "A rapid and automatic image registration algorithm with subpixel accuracy," in *IEEE Transactions on Medical Imaging*, vol. 16, no. 3, pp. 308-316, June 1997, DOI: 10.1109/42.585765.

[13] G. D. Evangelidis and E. Z. Psarakis, "Parametric Image Alignment Using Enhanced Correlation Coefficient Maximization," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 10, pp. 1858-1865, Oct. 2008, DOI: 10.1109/TPAMI.2008.113.

[14] A. Goshtasby and G. C. Stockman, "Point pattern matching using convex hull edges," in *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-15, no. 5, pp. 631-637, Sept.-Oct. 1985, DOI: 10.1109/TSMC.1985.6313439.

[15] J. Öfverstedt, J. Lindblad and N. Sladoje, "Fast and Robust Symmetric Image Registration Based on Distances Combining Intensity and Spatial Information," in *IEEE Transactions on Image Processing*, vol. 28, no. 7, pp. 3584-3597, July 2019, DOI: 10.1109/TIP.2019.2899947.

[16] S. Abdelsayed, D. Ionescu, D. Goodenough, Matching and registration method for remote sensing images, Proceedings of the International Geoscience and Remote Sensing Symposium IGARSS'95, Florence, Italy, 1995, pp. 1029–1031.

[17] Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision* **60,** 91–110 (2004). DOI: 10.1023/B:VISI.0000029664.99615.94

[18] Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer vision and image understanding*, *110*(3), 346-359. DOI: 10.1016/j.cviu.2007.09.014

[19] Deshmukh, M., & Bhosle, U. (2011). A survey of image registration. *International Journal of Image Processing (IJIP)*, *5*(3), 245. DOI=10.1.1.740.3626

[20] N. A. Dodgson, "Quadratic interpolation for image resampling," in *IEEE Transactions on Image Processing*, vol. 6, no. 9, pp. 1322-1326, Sept. 1997, DOI: 10.1109/83.623195.

[21] Ghosh, D., & Kaabouch, N. (2016). A survey on image mosaicing techniques. *Journal of Visual Communication and Image Representation*, *34*, 1-11. DOI: 10.1016/j.jvcir.2015.10.014.

[22] Alessandro Bevilacqua and Pietro Azzari. 2007. A Fast and reliable image mosaicing technique with application to wide area motion detection. *In Proceedings of the 4th international conference on Image Analysis and Recognition (ICIAR'07)*. Springer-Verlag, Berlin, Heidelberg, 501–512.

[23] Z. Li and V. Isler, "Large Scale Image Mosaic Construction for Agricultural Applications," in *IEEE Robotics and Automation Letters*, vol. 1, no. 1, pp. 295-302, Jan. 2016, DOI: 10.1109/LRA.2016.2519946.

[24] Jurado, J. M.; Ortega, L.; Cubillas, J. J.; Feito, F. R. 2020. "Multispectral Mapping on 3D Models and Multi-Temporal Monitoring for Individual Characterization of Olive Trees" *Remote Sens.* 12, no. 7: 1106. DOI: 10.3390/rs12071106

[25] Zheng, Hengbiao; Cheng, Tao; Li, Dong; Zhou, Xiang; Yao, Xia; Tian, Yongchao; Cao, Weixing; Zhu, Yan. 2018. "Evaluation of RGB, Color-Infrared and Multispectral Images Acquired from Unmanned Aerial Systems for the Estimation of Nitrogen Accumulation in Rice" *Remote Sens.* 10, no. 6: 824. DOI: 10.3390/rs10060824

[26] Tucci, Grazia; Parisi, Erica I.; Castelli, Giulio; Errico, Alessandro; Corongiu, Manuela; Sona, Giovanna; Viviani, Enea; Bresci, Elena; Preti, Federico. 2019. "Multi-Sensor UAV Application for Thermal Analysis on a Dry-Stone Terraced Vineyard in Rural Tuscany Landscape" *ISPRS Int. J. Geo-Inf.* 8, no. 2: 87. DOI: 10.3390/ijgi802008

[27] A. Dore, M. Soto and C. S. Regazzoni, "Bayesian Tracking for Video Analytics," in *IEEE Signal Processing Magazine*, vol. 27, no. 5, pp. 46-55, Sept. 2010, DOI: 10.1109/MSP.2010.937395.

[28] Meng Howe Tan and S. Ranganath, "Multi-camera people tracking using Bayesian networks," *Fourth International Conference on Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint*, 2003, pp. 1335-1340 vol.3, DOI: 10.1109/ICICS.2003.1292681.

[29] M. Kafai and B. Bhanu, "Dynamic Bayesian Networks for Vehicle Classification in Video," in *IEEE Transactions on Industrial Informatics*, vol. 8, no. 1, pp. 100-109, Feb. 2012, DOI: 10.1109/TII.2011.2173203.