Developing a Video Synchronization Framework for Wind Turbine Inspection Videos

Armagan Elibol^{a,*} and Nak Young Chong^a

^aSchool of Information Science, Japan Advanced Institute of Science and Technology
Nomi City, Ishikawa, 923-1292, Japan

*Corresponding Author: aelibol@jaist.ac.jp

Abstract

Robotic platforms have been continuously and rapidly advancing, especially over the last two decades. Apart from their usage in automating repetitive duties in a controlled environment, they have also been employed for exploration, mapping, inspection, monitoring purposes, particularly in areas that cannot be easily accessible to humans (e.g., underwater, space and many others). One of the leading research directions in such robotic fields is Unmanned Aerial Vehicles (UAVs). Visual inspection is one of the most important and automated application areas in which such robotic platforms have been often used in open areas for inspection, monitoring, mapping, and similar applications with a certain distance to the area of interest. Wind turbines are one of the widely used renewable energy sources and their sustainable operations highly depend on periodical inspection and maintenance in which UAVs have been actively deployed to collect optical data. The optical data gathered by UAVs need to be processed for visual inspection of wind turbine blades. In this paper, we discuss a methodology for identifying images in which blades are near the same position obtained at a different time on the way to make a visual inspection and detecting temporal changes process in a fully automatic manner. We present preliminary experimental results obtained on synthetically generated data.

Keywords: UAV, Visual Inspection, Wind Turbine.

1. Introduction

With the recent developments of camera-carrying robotic platforms, gathering optical data from areas beyond human reach or areas not easy access to humans have

become available. Visual inspection is one of the most important and automated application areas in which such robotic platforms have been often employed. While the existing approaches are able to cope with mapping/inspecting static areas or surfaces (e.g., infrastructure⁽¹⁾, tunnel⁽²⁾, pipeline^(3,4), hull^(5,6), dam^(7,8), bridge^(9,10) and many others), there is an increasing demand for inspecting areas/parts moving constantly. This has been recently included in the context of dynamic scene mapping^(11,12).

Wind turbines have been one of the leading renewable energy sources converting the kinetic energy of wind to electrical energy. According to the statistics announced by World Wind Energy Association (WWEA), Wind Power Capacity Worldwide has reached 597 GW covering approximately 6% of the global electricity demand¹. Their inspection and maintenance are essential factors for their service in an efficient and long-lasting manner^(13,14). Their inspection is labor-intensive, and they offer dangerous environments for humans to carry out inspection taking into account their size. Their inspection may take a long time, and this leads to also long time without operation causing energy loss. Unmanned Aerial Vehicles (UAVs) have been shown as a useful tool for making the inspection process easier, faster and safer for the visual inspection of surfaces⁽¹⁴⁾. Most of the existing recent studies focus on developing methods for detecting cracks, debris and/or imperfections on blade images using machine learning/deep learning methods^(13,15,16). Although these methods are

¹Retrieved from

<https://wwindea.org/blog/2019/02/25/wind-power-capacity-worldwide-reaches-600-gw-539-gw-added-in-2018/>
on September 2019

efficient, they require a relatively big training dataset, computational resource and a certain amount of expertise even using pre-trained models for testing. Also, their performance relies on training data quality. Inspection systems usually are composed of two main steps, namely, mapping and automatic comparison, to detect changes over time. The mapping part requires efficient methods to synchronize the videos acquired by a UAV from different viewpoints considering the dynamic scene. Existing methods have an assumption of a synchronized multi-camera setting (e.g., surveillance)⁽¹²⁾.

In this paper, it is aimed to discuss an end-to-end combination of image processing methods to synchronize video files for visual inspection of wind turbines automatically while they are operating using only image information obtained by low-cost UAV. In our context, the term synchronization is referred to as identifying frames among video file(s) in which wind turbine blades are in the same angular composition with respect to the camera. Our proposal takes a video file in which the wind turbine is imagined nearly fronto-parallel to the camera as an input and extracts angular information of blades to be used later in order to track the changes on blades obtained through videos captured in a different time. Unlike most of the aforementioned existing methods, our framework does not make use of any method requiring pre-training and/or learning, which needs a certain amount of well-defined training data and computational power. Therefore, the proposed framework has a low computational cost and does not use any other sensor information apart from the optical data allowing it to be used with low-cost UAVs, e.g., only flying cameras.

2. Video Synchronization Framework

Our framework assumes video files in which the camera is almost fronto-parallel to the rotating wind turbine, and the camera undergoes small motions due to the stabilization and environmental effects. The proposed framework makes use of image-only information to find similar positions of blades of a wind turbine in the images from video files obtained at different times. The overview of the proposed framework is summarized in Algorithm in 1. In order to eliminate the background, We adopted a well-known image stabilization through image alignment/mosaicking^(17,18) method. We extract scale-invariant features from images and compute the transformation between images using RANSAC in order to

compute background motion (referred to as *approxH* in Alg. 1). Since the objects closer to the camera move more than the ones are further, we apply a sampling-based searching mechanism in order to register the center part of the wind turbine correctly. This is a crucial procedure as image subtraction will be applied afterward. We generate a small number of (e.g., 5) different motion matrices by slightly changing the translation part of the *approxH* and test them to select *optimalH* with minimum residual. Later, we apply morphological operations to extract regions, mainly blades of the wind turbine. We obtain some important properties of the extracted regions, such as the angle between the x-axis and the major axis of the enclosing ellipse, angle of the bounding box, and angles of min. and max Feret diameters. The proposed framework is motivated by the fact that in the case of two images where blades are in different position extracted properties through the framework given in Algorithm 1 are expected to be similar. An example of the situation mentioned is given in Fig. 2. Image 1 is compared against images 9 and 12 extracted from the same video. After the image subtraction step, morphological properties would remain the same in both cases for Image 1. Once the region properties are found as a result of comparing all-images-against-all, histogram analysis is carried out in order to find representative region properties for each image. During this procedure, images in which blades of the wind turbine has the same configuration can also be identified within the same video file. The same process is repeated for each video file obtained at a different time. Representative region properties for each image in video files are compared. Thus, images are identified among different video files. Since the approach relies and operates on mainly angle information, it is invariant to background and image resolution.

3. Experimental Results

To test the proposed framework, we created synthetic images using the algorithm outlined in Algorithm 2. We generated three different video files (Merged Video is available at <https://streamable.com/sc04y>). Example frames from each generated video are given in Fig.1. Each video has a different blade rotation speed, background, and the total number of frames, respectively, 201, 144, and 101. The image resolution for the first video is 742x875 while for the second and the third video files is 1484x1750. Datasets are created taking into account that objects that are closer to the camera move more. Therefore, two different

random translations were used; one for the wind turbine and one for the background.

We first test our framework to find similar frames in which blade angles are near identical within the same video file. Finding these pairs provides valuable information, e.g., rotation speed. We compute the histogram and select peak 3 angles in order to choose corresponding blade angles to associate to the frame using the angle information extracted in Algorithm 1. We used these angles for searching the same frames both within the same video file and among the video files. Table 1 presents the number of correct and wrong identified image pairs and the ratio of correct frames within the same video file, while Table 2 is for among the video files for different bin-width during histogram analysis. From the tables, it can be seen that increasing bin-width within the same video file reduced overall performance as the rotation speed of blades is constant within the video file. The total number of identified image pairs among the different video files is relatively low. This is mainly due to the fact that the same angular status for blades is less likely to exist as the rotation speed is different for each video. Also, while increasing bin-width for identifying among different video files improved the results for the first and the second video files, the ratio for the third video reduced. From our experiments, we observed that primary error sources were singularities in angle representation, background elimination and also the cases one of the blades overlaps with the pole either fully or partially. Since the third video file has a relatively more complex background comparing two other tested video files, this background caused imperfections during background elimination. Results presented in the tables were obtained using solely aforementioned extracted information without using any filtering, thresholding, non-maximal suppression among frames, any shape prior (e.g., the angle between blades is the same) and frame rate information. Some of the resulting sample frames are given in Fig. 3.

Fig. 1 Sample images from the tested video files

```

Input: Video file,  $V$ 
Extract images  $(I_i, i = 1, 2, \dots, n)$  from  $V$ 
foreach Image  $I_i, i = 1, 2, \dots, n$  do
    Extract features from  $I_i$ 
    Save features to  $featI(i)$ 
end
foreach Image  $I_r, r = 1, 2, \dots, n - 1$  do
    Set  $featIr$  to  $featI(r)$ 
    foreach Image  $I_s, (s = 1, \dots, n) \setminus (r)$  do
        Set  $featIs$  to  $featI(s)$ 
        Apply Feature matching and RANSAC using
         $featIr$  and  $featIs$  to compute approximate
        motion  $approxH$ 
        Search for  $optimalH$  using  $approxH$ 
        Warp  $I_r$  using  $optimalH$  to obtain  $warpIr$ 
        Subtract Images  $warpIr$  and  $I_s$  to get  $subsI$ 
        Apply morphological operations to  $subsI$  to
        retrieve region properties
        Save region properties to  $propIr(r, s)$ 
    end
end

```

Algorithm 1: The overview of the proposed framework

```

Input: Background image file, wind turbine image file,
Rotation degree per second, Video file frame rate,
Length of the video, Initial angle  $iniA$ 
Output: Video file
Compute the total number of frames  $totalFrame$ 
Compute rotation angle increment  $incAngle$  for each
successive image
Set the initial position of wind turbine  $iniPos$ 
for  $counterIm = 1$  to  $totalFrame$  do
    Create empty image matrix  $currIm$ 
    Compute angle of rotation  $rotAngle$  for blades
     $rotAngle = iniAngle + (counterIm - 1) \times incAngle$ 
    Apply  $rotAngle$  to the wind turbine image
    Apply random translation  $rndTrans1$  to  $iniPos$ 
    Apply random translation  $rndTrans2$  to the
    background image
    Place the background and wind turbine image in
     $currIm$  accordingly
    Save  $currIm$ 
end
Create file video using generated images

```

Algorithm 2: Main steps of Experimental Data Generation

4. Conclusions and Future Work

Wind turbines play an essential role in generating electricity from wind, and their importance is more and more as a source of renewable and clean energy generation. They can be regarded as a mechanical system composed of different parts. Their periodic inspection is indispensable for their continuous and sustainable service. UAVs carrying different sensors have been deployed for creating a visual record of wind turbines in order to facilitate the inspection process.

Fig. 2 Sample images from the same video file and their overlaid versions. The first row denotes images 1 and 9 along with their overlaid version without registration step while the second row shows images 1 and 12 along with their registered overlaid version.

Table 1. Total number of identified image pairs within the same video file

Video File	Bin Width degree	Correct	Wrong	Total	Ratio
1-1	2.5	165	3	168	98.21%
2-2		114	5	119	95.80%
3-3		66	23	89	74.16%
1-1	5	88	6	94	93.62%
2-2		90	2	92	97.83%
3-3		58	30	88	65.91%
1-1	10	67	4	71	94.37%
2-2		103	3	106	97.17%
3-3		71	29	100	71.00%
1-1	20	67	22	89	75.28%
2-2		99	18	117	84.62%
3-3		82	41	123	66.67%

Generally, existing approaches provide tools for data gathering, and the data is further studied by human experts. In this paper, we present an approach to synchronize videos of a wind turbine obtained both while operating and in different time intervals using the image only information. Our approach can be seen as the first step of fully automatic

inspection. The proposed approach makes use of the shape/morphological properties of the blades to find a similar configuration in the other video. The approach takes video files in which the wind turbine is nearly fronto-parallel to the camera as input and does not use any other information. We present our preliminary findings through experimental results on synthetically generated realistic data (available at <https://streamable.com/sc04y>), which takes into account a certain level of vibration motion that might be caused due to environmental conditions (e.g., wind) and/or stabilization of UAV. Our approach is built upon well-established computer vision and image processing algorithms, which have relatively low computational costs and can be run on off-the-shelf computers without requiring extensive experience.

The presented framework could be enhanced by incorporating different information, which can be derived from the video file, e.g., rotation speed, and related others. As future work, we will evaluate our framework with real data. We will also develop methods for detecting temporal changes on blades as images, in which blades have similar orientations are identified, extracted from videos acquired at different times.

Table 2. Total number of identified image pairs among different video file

Video File	Bin Width degree	Correct	Wrong	Total	Ratio
1-2	2.5	9	4	13	69.23%
1-3		1	0	1	100.00%
2-3		2	0	2	100.00%
1-2	5	15	8	23	65.22%
1-3		20	12	32	62.50%
2-3		19	10	29	65.52%
1-2	10	21	10	31	67.74%
1-3		22	15	37	59.46%
2-3		26	8	34	76.47%
1-2	20	16	4	20	80.00%
1-3		11	4	15	73.33%
2-3		17	9	26	65.38%

Acknowledgment

This research was supported under the JAIST Research Grant program funded by JAIST Support Foundation.

Fig. 3 (Top) An example of correctly identified images from the first video file. (Middle) An example of correctly identified image pair from the second and third video. (Bottom) Misidentified sample images from the first and the third video file.

References

- (1) S. Agnisarman, S. Lopes, K. C. Madathil, K. Piratla, and A. Gramopadhye: "A survey of automation-enabled human-in-the-loop systems for infrastructure visual inspection," *Automation in Construction*, vol. 97, pp. 52–76, 2019.
- (2) R. Montero, J. Victores, S. Martinez, A. Jardon, and C. Balaguer: "Past, present and future of robotic tunnel inspection," *Automation in Construction*, vol. 59, pp. 99–112, 2015.
- (3) M. Tavakoli, L. Marques, and A. de Almeida: "Development of an industrial pipeline inspection robot," *Industrial Robot*, vol. 37, no. 3, pp. 309–322, 2010.
- (4) P. Hansen, H. Alismail, P. Rander, and B. Browning: "Visual mapping for natural gas pipe inspection," *The International Journal of Robotics Research*, vol. 34, no. 4-5, pp. 532–558, 2015.
- (5) A. Kim and R. M. Eustice: "Real-time visual SLAM for autonomous underwater hull inspection using visual saliency," *IEEE Transactions on Robotics*, vol. 29, no. 3, pp. 719–733, June 2013.
- (6) A. Milella, R. Maglietta, M. Caccia, and G. Bruzzone: "Robotic inspection of ship hull surfaces using a magnetic crawler and a monocular camera," *Sensor Review*, vol. 37, no. 4, pp. 425–435, 2017.
- (7) P. Ridao, M. Carreras, D. Ribas, and R. Garcia, "Visual inspection of hydroelectric dams using an autonomous underwater vehicle," *Journal of Field Robotics*, vol. 27, no. 6, pp. 759–778, 2010.
- (8) H. Sugimoto, Y. Moriya, and T. Ogasawara: "Underwater survey system of dam embankment by remotely operated vehicle," in *2017 IEEE Underwater Technology (UT)*, pp. 1–6, Feb 2017.
- (9) R. S. Adhikari, O. Moselhi, and A. Bagchi: "Image-based change detection for bridge inspection," in *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, vol. 30. IAARC Publications, 2013.
- (10) N. Pham, H. La, Q. P. Ha, S. Dang, A. Vo, and Q. Dinh, "Visual and 3d mapping for steel bridge inspection using a climbing robot," in *Proceedings of the ISARC 2016 - 33rd International Symposium on Automation and Robotics in Construction and Mining*, Alabama, USA, pp. 141–149, July 2016.
- (11) M. R. U. Saputra, A. Markham, and N. Trigoni: "Visual SLAM and structure from motion in dynamic environments: A survey," *ACM Comput. Surv.*, vol. 51, no. 2, pp. 37:1–37:36, Feb. 2018.
- (12) A. Mustafa, H. Kim, J. Guillemaut, and A. Hilton: "General dynamic scene reconstruction from multiple view video," in *2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 900–908, Dec 2015.
- (13) L. Wang and Z. Zhang: "Automatic detection of wind turbine blade surface cracks based on UAV-taken

- images,” *IEEE Transactions on Industrial Electronics*, vol. 64, no. 9, pp. 7293–7303, Sep. 2017.
- (14) G. Morgenthal and N. Hallermann: “Quality assessment of unmanned aerial vehicle (UAV) based visual inspection of structures,” *Advances in Structural Engineering*, vol. 17, no. 3, pp. 289–302, 2014.
- (15) Y. Wang, R. Yoshihashi, R. Kawakami, S. You, T. Harano, M. Ito, K. Komagome, M. Iida, and T. Naemura: “Unsupervised anomaly detection with compact deep features for wind turbine blade images taken by a drone,” *IPSJ Transactions on Computer Vision and Applications*, vol. 11, no. 1, p. 3, 2019.
- (16) A. Shihavuddin, X. Chen, V. Fedorov, A. Nymark Christensen, N. Andre Brogaard Riis, K. Branner, A. BJORHOLM DAHL, and R. Reinhold Paulsen: “Wind turbine surface damage detection by deep learning aided drone inspection analysis,” *Energies*, vol. 12, no. 4, p. 676, Feb 2019.
- (17) R. Hu, R. Shi, I. Shen, and W. Chen: “Video stabilization using scale-invariant features,” in *2007 11th International Conference Information Visualization (IV '07)*, pp. 871–877, July 2007.
- (18) T. Buyukyazi, S. Bayraktar, and I. Lazoglu: “Real-time image stabilization and mosaicking by using ground station CPU in UAV surveillance,” in *RAST 2013 - Proceedings of 6th International Conference on Recent Advances in Space Technologies*, pp. 121–126, 2013.