 doi:10.20944/preprints202205.0298.v1

Article

Automatic Semantic Segmentation of Salient Patterns in Temporal Images for Digital Image Correlation-based Structural Health Monitoring of Large Structures

Christopher Chun Ki Chan[1], David Kumar[2], Chih-Hung, Chiang[2], and Alexander Ferworn[3]

[1] Chaoyang University of Technology, Department of Information Management; christopherckchan@cyut.edu.tw
[2] Chaoyang University of Technology, Department of Aeronautical Engineering; davidcyut@cyut.edu.tw, chiangc@cyut.edu.tw
[3] Toronto Metropolitan University, Department of Computer Science; aferworn@ryerson.ca
* Correspondence: christopherckchan@cyut.edu.tw; davidcyut@cyut.edu.tw

Abstract:

Large structures such as wind turbines are subject to environmental factors and varying operational loads which may result in structural damage, making components of these large structures prone to performance and mechanical degradation. The use of high-definition optical vision sensors in digital image correlation (DIC) allow for the application of a non-destructive image registration technique in which it measures finite three-dimensional deformations on surfaces through correlations of a unique pattern or set of unique localized patterns. However, the physical placement of an artificial marker such as a unique speckled pattern on the surface of the structure is time-consuming and often impractical for large structures. Therefore, we propose a novel automated methodology that searches and segments salient and unique regions of an image as well as for all subsequent images to assist in performing efficient displacement measurements for vibrational study and structural health monitoring purposes. Our algorithm is validated on a controlled set of images, as well as on a small structure and large real-world wind turbine, which suggests the algorithm's efficacy without the use of artificial markers for large structural health monitoring.

Keywords: digital image correlation; semantic filter; structural health monitoring; unique salient patterns; wind turbine

1. Introduction

Digital image correlation is a nondestructive testing method that involves high-definition optical sensors for precise tracking and measurement of displacements of unique patterns in digital images [1]. By measuring and tracking blocks of pixel displacements from correlated salient and unique patterns, DIC can provide full-field 2D and 3D deformation vector fields and strain maps, which aids in surface-level measurement for examining temporal stress/strain in materials testing, crack tip and crack propagation studies, detecting damage development in composites, structural deflections, high-temperature strain mapping, and dynamic vibrational analysis.

Localized patterns such as circular random dots (or speckled patterns) for digital image correlation (DIC) processing are artificial landmarks that are physically placed in a scene to enable point correlation, or correspondences between images [2], [3]. A localized pattern should be unique in an image and also unique in subsequent images for DIC processing, and the pattern in the image should not be blurry, or relatively 'coarse' as defined in [4], [5], otherwise it would lead to correlation errors or processing failure. As of this writing, the DIC-based structural health monitoring methods (SHM) often include

 doi:10.20944/preprints202205.0298.v1

the physical application of a salient and unique pattern on the actual structure to aid in image correlation operations, little has been done in the automatic finding of unique patterns within a real-world image and a set of subsequent images [6], [7].

Our previous work [4] demonstrated a working classifier that can distinguish the saliency of a projected pattern in a real-world image (sharp), and determine the uniqueness of the projected pattern in a localized region of the image and as well as for all regions in subsequent images. However, two shortcomings of that algorithm are that the search space for the uniqueness algorithm is limited to a region of interest rather than searching the entire image for the sake of computational efficiency, in addition, saliency determination is computationally inefficient, and it does not semantically filter patterns that are part of a target structure, which means background patterns may also be classified as unique and salient.

We augmented our previous algorithm to include the following: (1) Determination of an appropriate pattern/marker size according to the size of the image, (2) Pattern detection with the use of an improved-Speeded-up Robust Feature detector, (3) Pattern region segmentation (4) Salient and Uniqueness Classifier that scans pattern structural similarities in an entire image as well as an image set and lastly, (5) Semantic Filter using YOLOv5 of large structures to obtain only the optimal real-world salient and unique patterns that are part of the structure within the image. Having an automated method to semantically filter a salient and unique pattern within an image and the set of subsequent images is significant in reducing manual labor, errors in DIC processing, and speeding up the overall computational time of the whole DIC process from start to end.

The rest of the paper is designed as follows: (2) Methods, where the algorithm is described in full detail, (3) Results and Discussion, where the algorithm is validated on a control image set, two real-world image sets, and actual DIC processing, and (4) Conclusion.

2. Methods

In our paper, we utilize the terms marker to describe salient patterns, or rather, fiducial markers that are salient and unique in an image and are not artificial, which means that in a real-world image, there would be no artificial markers placed in the scene for image processing.

Figure 1-4 is an overview flowchart of the entire algorithm; the input is an image containing a target structure for structural health monitoring, which may be very large (100m tall, ~3000 pixels x ~4000 pixels) or small (10m tall, ~300 pixels x 400 pixels), and output would be 10 or less potential salient and unique patterns of a fixed size for use in DIC processing.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 May 2022 doi:10.20944/preprints202205.0298.v1

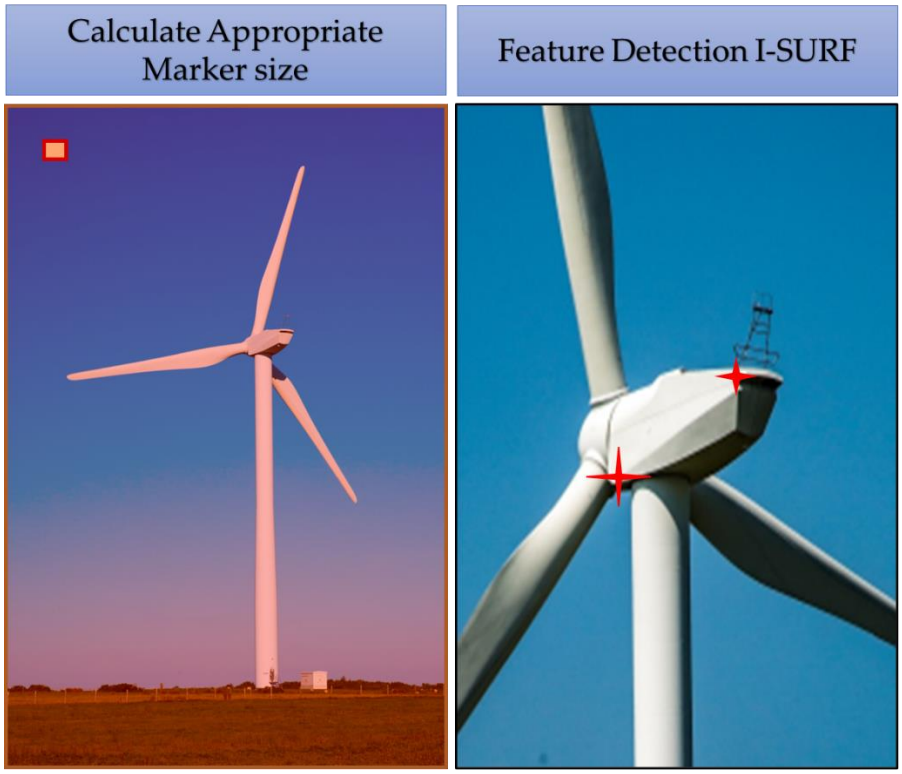

Figure 1. The first two steps of our algorithm which includes the calculation of an appropriate marker size with respect to the entire image, and feature detection with i-SURF, giving us initial features of interest to focus our algorithm on.

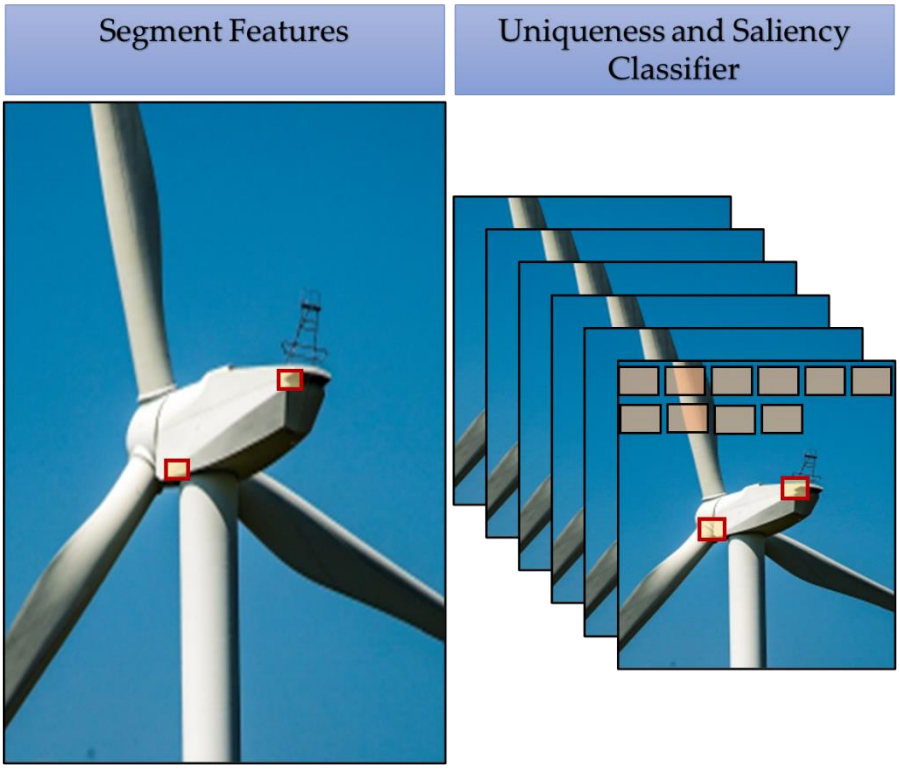

Figure 2. The following 3rd and 4th steps of our algorithm which are segmentation and our ULPA2 classifier, which determines whether the segmented region is salient and unique across all images in the set.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 May 2022 doi:10.20944/preprints202205.0298.v1

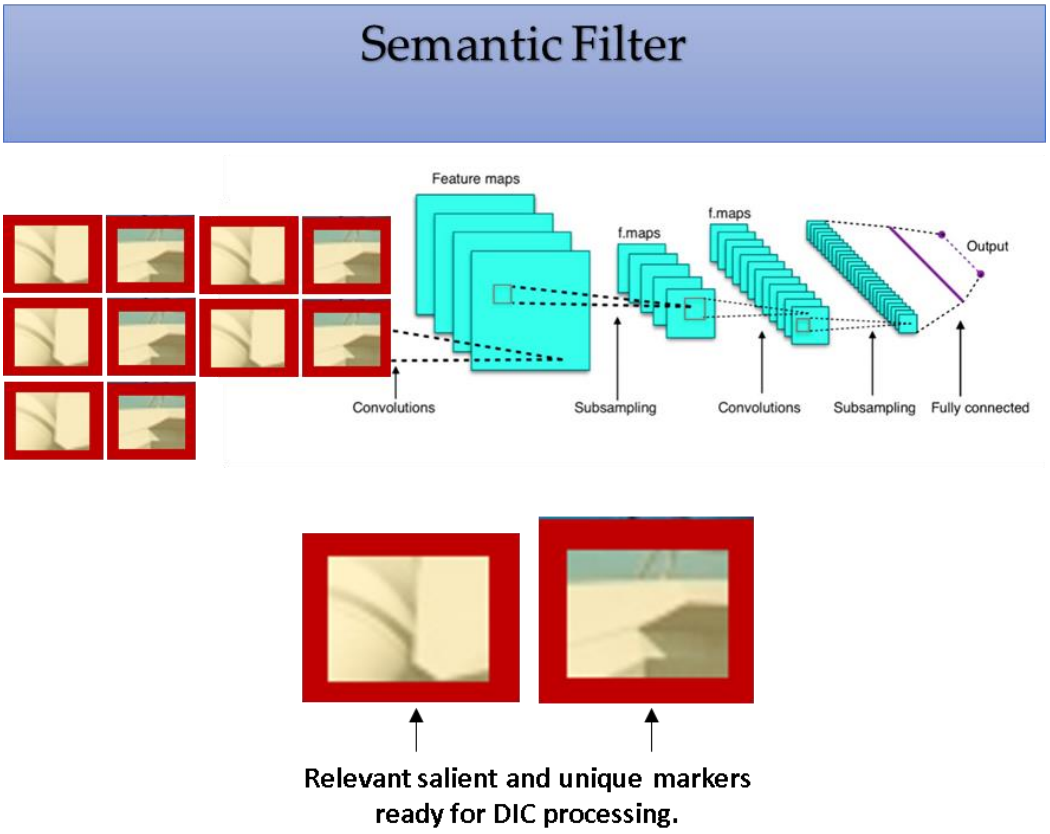

Figure 3. The 5th step of our algorithm which is an intelligent semantic filter.

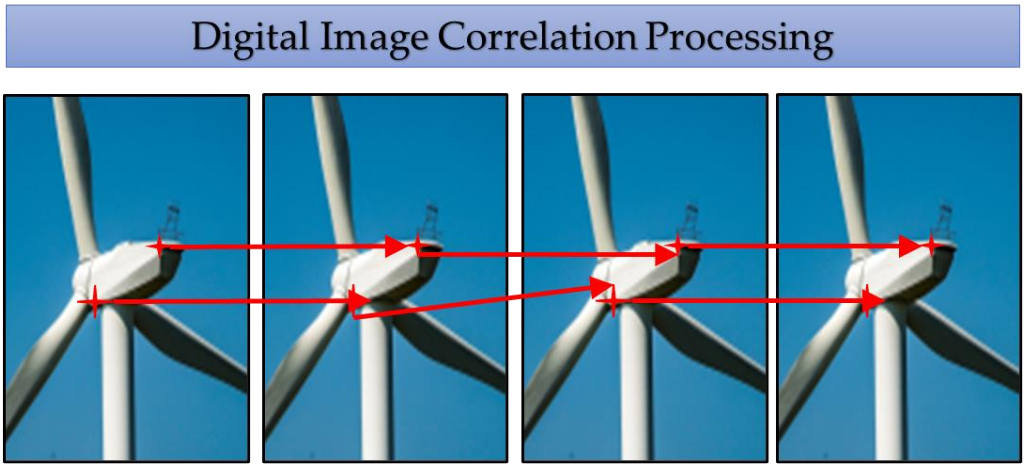

Figure 4. The 5th step of our algorithm which is an intelligent semantic filter.

2.1. Marker size Determination

The size of a unique identifier is calculated with an estimated area of 1.14% of the entire image size. An image size of 471 x 3369 pixels would have a suitable marker size of a total of 18,089 pixels^2. The algorithm takes the ceiling of the square root of 18,089 pixels^2 to get 135 pixels for the marker's length and width. The estimated area size of the marker at 1.14% is most suitable for images containing the structure to be tested at about half or more of the structure shown in a digital image.

 doi:10.20944/preprints202205.0298.v1

2.2. Initial Feature Detection

An improved Speeded-up Robust Features (i-SURF) algorithm [8] is first used to determine the first 10 most significant feature points, determined by the strength of the most salient regions in the image. These most representative key points are obtained with SURF's usual algorithm consisting of the following (1) Construct Hessian matrix; (2) Scale-space generation; (3) Use non-maximal suppression to initially determine feature points; (4) Select the general orientation of feature points; (5) Construct feature point descriptors.

The improved version adds a Frequency-tuned (FT) Salient Region Detector [9] that utilizes a Gaussian differential filter method by only selecting features within a certain frequency range, which speeds up computational time significantly. The Frequency-tuned method [9] uses image-guided filters [10] as local estimates based on the mean and variance of the pixels in the image neighborhood.

Since the set of the images are typically taken in clear daylight with minimal external environmental effects, the intensity frequency ranges of our images and target structure are consistent, which works well with i-SURF's FT feature in which visual saliency is key. The i-SURF algorithm is also used mainly for its computational speed via its box type convolution filter for similarity invariant representation of strong feature points in rotation invariant, blur and warp transform, but without the need for different scale images such as the relatively less efficient SIFT method. The box convolutional filter fits with our marker size equal length and width.

2.3. Segment Features with Marker size

The location of the top ten feature points is set as the center for the designated marker size, assuming that as strong feature points allow for accurate point to point correlations for 2D-Digital Image Processing. These 10 markers are not yet determined to be unique markers of the structure but are considered as fiducial markers of the image. Out of bounds markers are omitted from the set.

For an input image of this small pole structure, the output would result in the following segmentation of features with the appropriate marker size as shown in Figure 5-6.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 May 2022 doi:10.20944/preprints202205.0298.v1

Figure 5. Initial features of interest as detected by i-SURF, without segmentation.

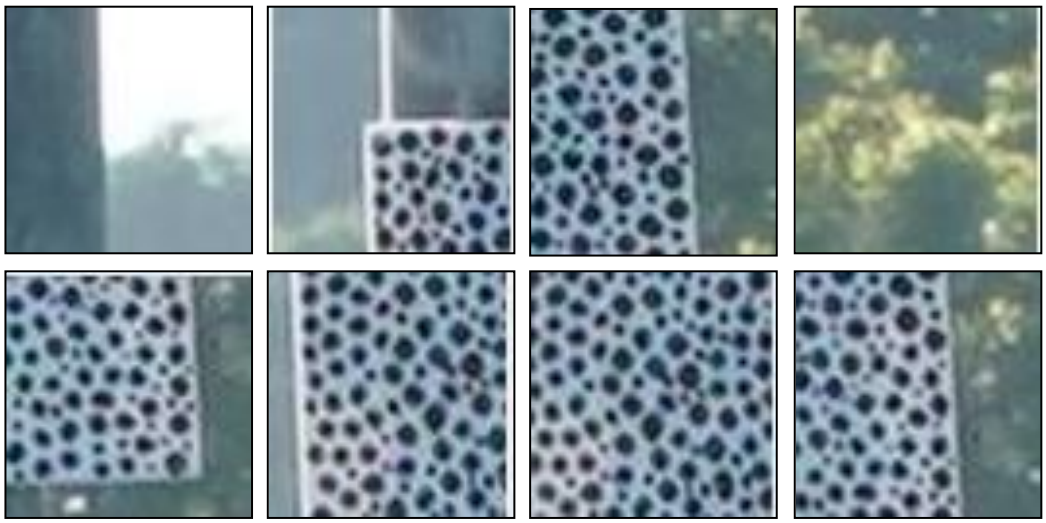

Figure 6. Segmented regions of interest with an appropriate marker size and targeted features from i-SURF.

2.4. Determine Uniqueness and Saliency of each Marker within Image Set

We pass each of the remaining fiducial markers of this image to our proposed pre-processing methodology to automatically classify the saliency and uniqueness of a localized pattern for DIC processing of a large structure for structural health monitoring. This methodology was previously published in our previous paper which combines a localized multi-scale CNN classifier with a Unique Localized Pattern Algorithm (ULPA) that ensures that a marker is salient within a real world image and unique within an image and a set of images. The previous ULPA would have to check each and every region of an image to the set of images to determine a saliant and unique region, which would easily cascade into a computationally expensive task, instead, the updated ULPA2 would only have to check 10 or less segmented markers to the entire image.

The multi-scale CNN utilizes multiscale blur detection networks which are fused together for a unified final output [11]. Three scales are trained at 33% reduction each scale. Blur probability maps are then put into dilation convolutional filters [12].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 May 2022 doi:10.20944/preprints202205.0298.v1

ULPA2 is a uniqueness verification algorithm that ensures the structural similarity (SSIM) index of the marker is above a similarity threshold of 0.9 in every part of an image, and also tests for all parts of all subsequent images. Structural similarity index [13] is similarity measure index value indicates the comparison of two images' signal as input; resulting in signal x and y luminance values. The marker is compared to every identical sized region moving in a sliding window fashion for all images in a set, if there is more than one match in an image, then the current marker is omitted from the results. Having a fixed marker size of 1.14% of the entire image would still be computationally efficient with the algorithm operating on a CPU only, meaning that the search process for each marker on each image of the set would be approximately 99 sliding window iterations depending on the length of the sliding window shift.

Salient and unique markers are then filtered and passed on to the last part of the algorithm for semantic filtering. ULPA can be understood further in the following pseudocode:

Unique Localized Pattern Algorithm v2

Input: A list of 10 target center data points of interest from the iSURF algorithm

obtain markerSizeRow and markerSizeCol (~1.14% of entire image)
for each center data point i **do**
  set top left point = ceiling (x_{1i} – markerSizeRow/2, y_{1i} - markerSizeCol)
  set bottom right point = ceiling (x_{2i} + markerSizeRow/2, y_{2i} + markerSizeCol)
  check if marker is within the image boundaries
  set search window region to be the entire image
  set structural similarity index (SSIM) threshold to 0.9
  for each image in the set, **set** sliding window to be top left corner of image **do**
    for the entire search window (x_3, x_4) and (y_3, y_4) move sliding window **do**
      check if sliding window SSIM value is greater than SSIM value **then**
        increase match counter by 1
      end if (SSIMthreshold checker)
    end for (shift search window by 1 pixel)
  end for (move to next image in the set)
end for (move to next data point i)
if match counter is equal to 0, no match found, pass uniqueness test
  check if semantic filter recognizes marker, if so, then pass semantic test
elseif match counter equal to 1, one match found, pass uniqueness test
  check if semantic filter recognizes marker, if so, then pass semantic test
else multiple matches found, fails uniqueness test

2.5. Semantic Filter of Markers

The last part of the algorithm is an intelligent semantic filter to determine which of the salient and unique markers are part of the target structure of interest, which involves artificial intelligence and convolutional neural networks (CNN). Since our main interest is mainly large structures of interest, we can develop fast YOLOv5 close-up detector [14] that is built to detect local segments of large structures such as windmills and bridges, and disregards background environmental scenes such as forests, trees, sky and grass.

The semantic filter is a YOLOv5 close-up [14] implementation using transfer learning along with our own dataset of large structures, containing bridges, wind turbines, large poles, lighting poles, and buildings. A robust CNN model is trained from over 1500 images per class, and over 10,000 instances per class. We incorporate 10% of the data to

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 May 2022 doi:10.20944/preprints202205.0298.v1

be background images in order to reduce false-positives errors. With a relatively small dataset the training process and transfer learning parameters are adjusted accordingly. The dataset is annotated using YOLO labelling format, with each file containing one bounding-box annotation.

3. Results and Discussion

We validated our algorithm in four parts, the first is a controlled set of simulated images that contain 10 or more patterns, the second is a set of real-world images with artificial markers, and last, are two sets of real-world images without artificial markers. We run a few of these patterns in our 3D-DIC processing algorithm to validate our method.

3.1. Validation Results of Controlled Simulated Images

In this controlled simulated image, we simulated a relatively large solid color vertical bar, covering half of an empty image. Then we incorporated 4 artificially-created unique markers using an in-house marker generator algorithm with the use of random permutations to indicate white or black areas in the marker. Figure 7 shows one of the controlled simulated images in our set.

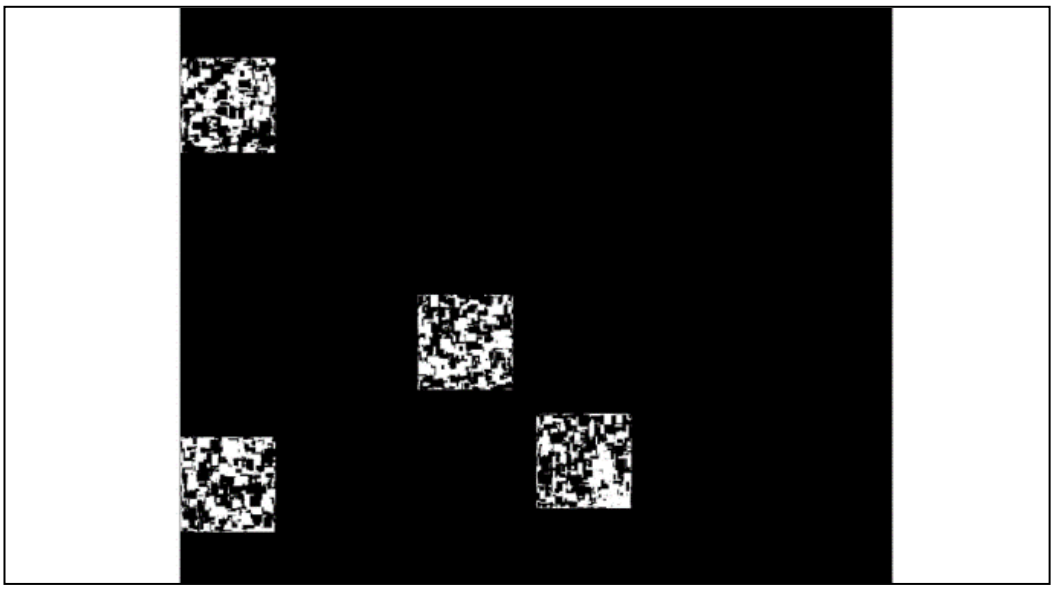

Figure 7. This image is one of the controlled simulated images containing 4 artificial markers, the uncropped image (3000 pixels x 2000 pixels) would contain more white space and the black center would encompass about half the image.

We processed this image with i-SURF to automatically locate these top 10 feature points, the features are not yet segmented, Figure 8 demonstrates the feature detection process on our controlled image.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 May 2022 doi:10.20944/preprints202205.0298.v1

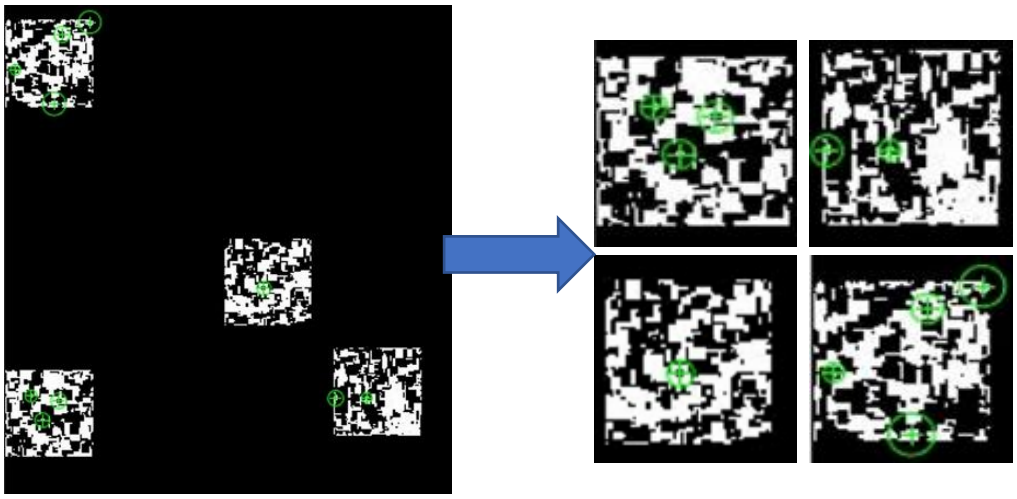

Figure 8. i-SURF quickly finds feature points on our image without semantic segmentation.

Figure 9 demonstrates the automatic segmentation process of the algorithm, resulting in 10 potential markers to be classified using our ULPA2 method.

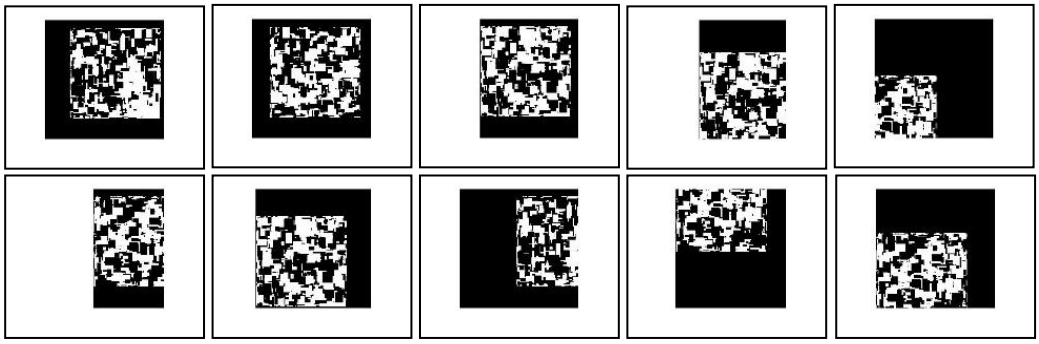

Figure 9. These are 10 segmented markers after i-SURF efficiently searches for top feature points on our image.

The remaining successful marker candidates are shown in Figure 10.

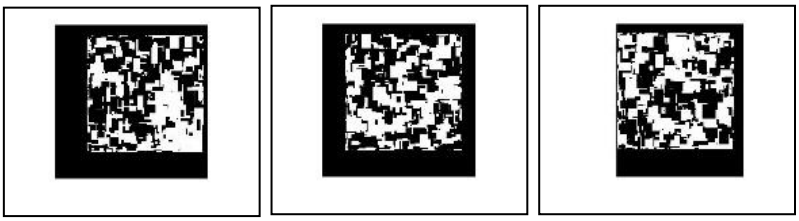

Figure 10. Three salient and unique markers from the output of our algorithm.

3.1.1. Controlled Simulated Images Discussion

Our algorithm was evaluated on 1000 of these controlled simulated images with randomly placed markers in the scene, and it successfully detected relevant markers, which passed our initial visual inspection. However, the semantic segmentation filter for this image set was retrained with a set of 10000 high-definition markers, and it should be noted that this test set does not thoroughly test the robustness of the semantic segmentation filter, which would work better with real-world images as shown in the next validation results.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 May 2022 doi:10.20944/preprints202205.0298.v1

3.2. Validation Results of Experimental Controlled Real-World Images

To further our validation study, we evaluated our algorithm on a set of real-world images with artificial markers. There were a total of 433 controlled real-world images of size 400 pixels x 1200 pixels with at 6 man-made artificial markers placed on a metal pole-like structure.

Of these images, 3 of the artificial markers are considered false-positive in terms of saliency and uniqueness, which will cause errors or program failure in the image correlation process. We incorporated positive and false-positive artificial markers in order to evaluate the algorithm's capability in detecting man-made markers that are not salient and/or unique. The part of our algorithm that should be able detect these false-positives is i-SURF, and the results are shown in Figure 11-13.

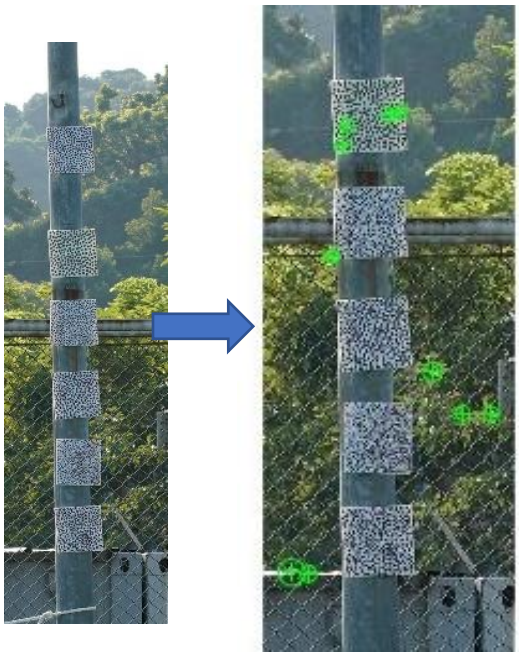

Figure 11. The leftmost image is the input target image, the second image shows detected feature points from i-SURF.

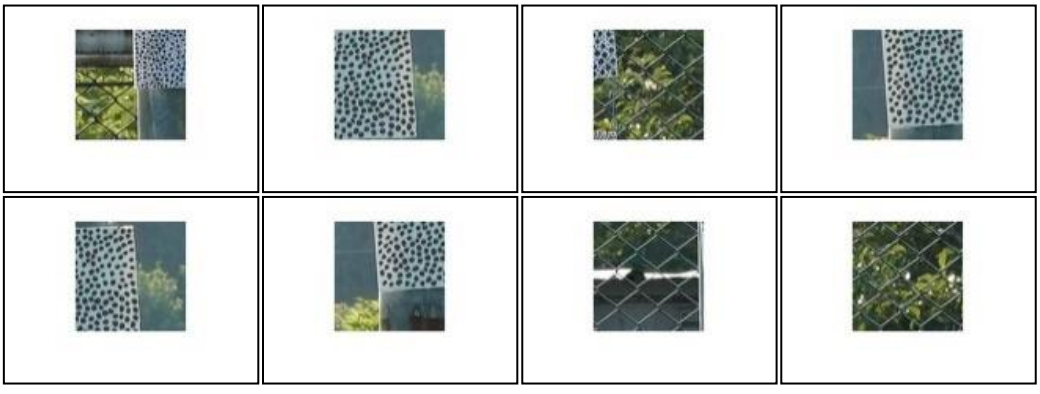

Figure 12. These are 8 segmented markers that are considered in the bounds of the image.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 May 2022 doi:10.20944/preprints202205.0298.v1

Figure 13. These are the resulting 3 segmented regions after the semantic filter is applied, which may be considered as salient and unique markers.

3.2.1. Experimental Controlled Real-world Images Discussion

The algorithm was evaluated on 433 of these experimental controlled real-world images with fixed speckled patterns placed as artificial markers on the scene. The first input image of the set was used to obtain the initial segmented markers, which was then passed through ULPA2 to compare each of the segmented markers with the rest of the regions in 432 images in the set for saliency and uniqueness. In Figure 7, there are only 8 segmented markers because the other 2 were out of image boundaries when segmentation occurred.

With regards to this image set, these speckled patterns were hand-drawn, and not all hand-drawn patterns are salient and unique, which ultimately lead to countless DIC processing errors and/or failure as a result of image to image point miscorrelation. Compounded to this issue, it was difficult to pin-point the reason for failure or which hand-drawn pattern was inadequate for DIC processing. As a result, we developed an algorithm to digitally produce random black and white patterns in a marker to a degree at which it will be unique enough in the entire image [4]. However, this method still required involved the physical placement of an artificial marker on the scene, which is laborious and inefficient. As shown, our algorithm is able to successfully output relevant markers, and in this case, with the placement of artificial markers, but also our algorithm is able to omit irrelevant markers on the structure (false-positives), as well as non-relevant markers (eg. background fence or bushes) for DIC processing.

3.3. Validation Results of Experimental Real-world Images

For our last validation study, the task was to use no artificial markers with a larger-sized image and structure and let the algorithm decide which segmented region is optimal for DIC processing within our parameters and assumptions. Both sets are of a real-world wind turbine, in which the first set contains 382 high-definition digital images (4000 pixels x 6000 pixels), and the other set contains 432 high-definition digital images (6000 pixels x 3376 pixels). In both cases, it would be impractical from its size to place an artificial marker on the structure, or even create a large enough marker. Evaluation results are shown in Figures 14-19.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 May 2022 doi:10.20944/preprints202205.0298.v1

Figure 14. Using real-world images of a wind turbine looking up at the structure, the algorithm is able to locate 10 initial distinct features using i-SURF.

Figure 15. These are the segmented regions of the top 6 features that are not out-of-bounds of the image.

Figure 16. The remaining two segmented markers after semantic filtering is applied.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 May 2022 doi:10.20944/preprints202205.0298.v1

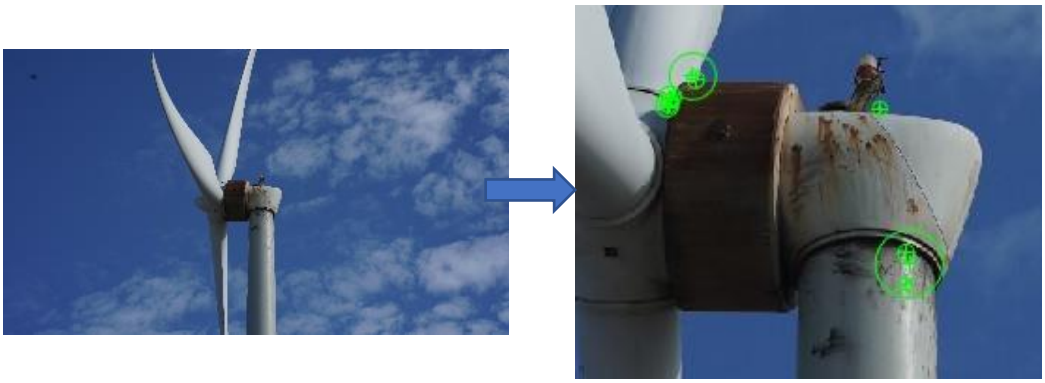

Figure 17. Another set of images of a wind turbine shown across a distance, the algorithm is able to obtain 10 initial feature points using i-SURF.

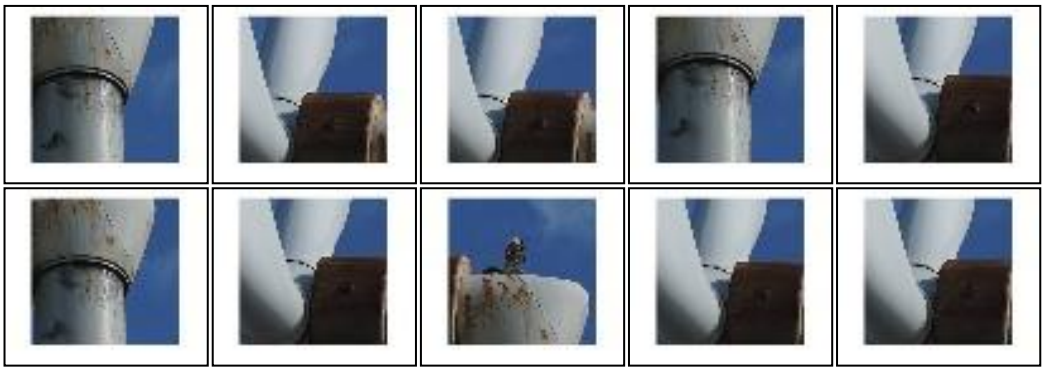

Figure 18. The output after segmentation occurs, which resulted in ten segmented regions.

Figure 19. These are the remaining 3 segmented markers after semantic filtering is applied, which are considered salient and unique within the whole image set.

3.2.1. Experimental Controlled Real-world Images Discussion

In Figure 16, the reason the semantic filter may have omitted the other markers was those images may have contained too much background information, and did not register successfully as a part of a wind turbine.

In Figure 19, we observe an extremely favorable result demonstrating visually relevant, salient, and unique markers. Without ULPA2 and our semantic filter, we would be left with too many false-positive markers that could perhaps cause correlation errors, by going through these filters, we remove the possibility of any issues arising during DIC processing.

3.4. Validation Results DIC Processing

Finally, for the last real-world wind turbine dataset, we evaluated the output of our algorithm, which were the 3 markers as shown in Figure 19, by performing DIC processing to obtain displacements and natural frequencies of the wind turbine. In this study, we employ an in-house 3D-DIC methodology that consists of a MATLAB-based stereo-correlation program and a virtual-physical stereovision calibration technique es-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 May 2022 doi:10.20944/preprints202205.0298.v1

pecially developed for 3D measurements on large structures [15]. The identified markers were supplied to the image correlation program to obtain their pixel coordinates in the images.
This process involves the implementation of a stereovision calibration process in order to obtain intrinsic and extrinsic parameters and use them to transform the pixel coordinates of the markers to their corresponding 3D locations in the real-world coordinates.

Natural frequencies of the wind turbine are then obtained by performing Fast Fourier Transform (FTT) on the measured time-domain displacement results. Figures 21, 23 and 25 show the vibrational results of the wind turbine tower obtained using the identified markers, and Figures 20, 22, and 24 show the displacement results using our markers. We observe that the algorithm is able to produce the displacements and frequencies successfully from DIC processing using our automatically segmented markers.

The complete image correlation process ran smoothly without any errors or failures that further validates our algorithm. It can also be seen that the amplitude of displacements is similar for all three markers as they are acquired from the top of the tower of the wind turbine, which supports and suggests the efficacy of the markers to produce accurate and consistent results without failures.

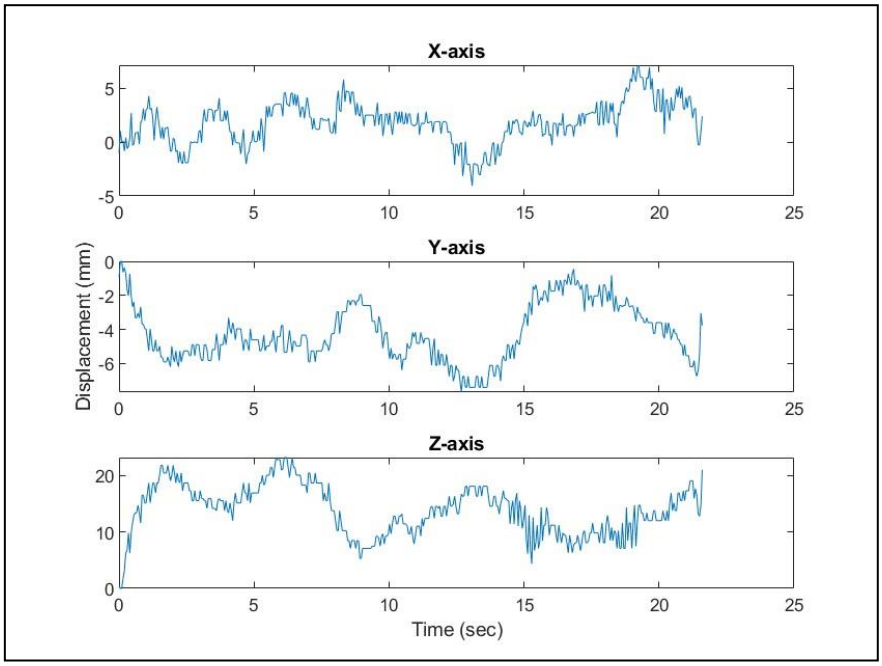

Figure 20. Displacements of the wind turbine tower using Marker-1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 May 2022 doi:10.20944/preprints202205.0298.v1

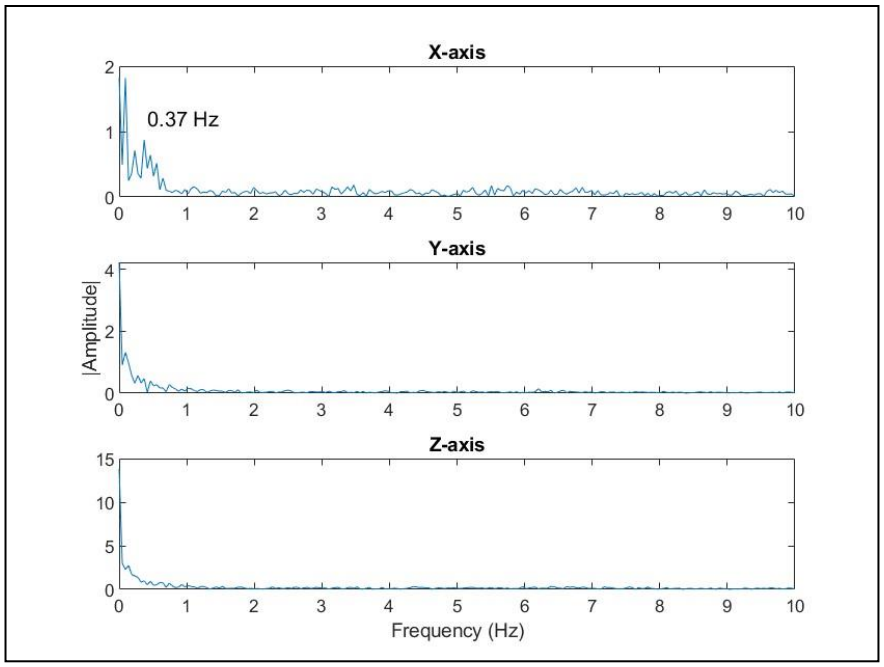

Figure 21. Vibrational results of wind turbine tower using Marker-1.

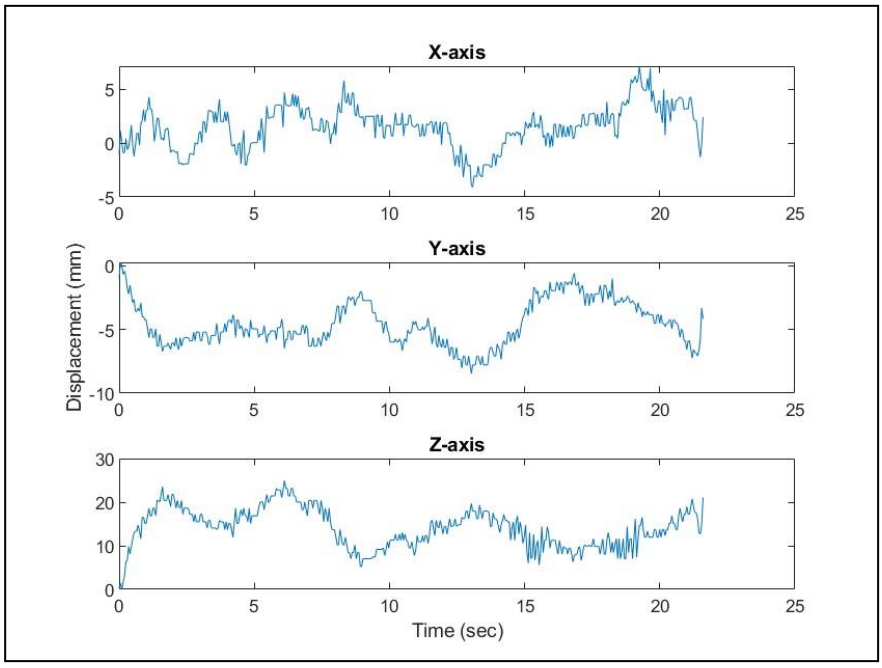

Figure 22. Displacements of the wind turbine tower using Marker-2.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 May 2022 doi:10.20944/preprints202205.0298.v1

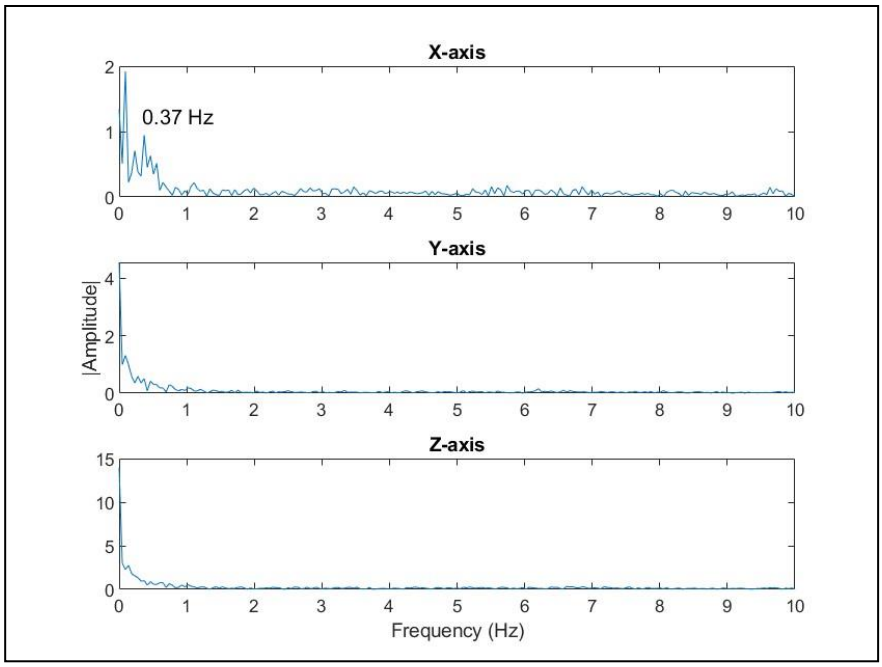

Figure 23. Natural frequency of the wind turbine tower using Marker-2.

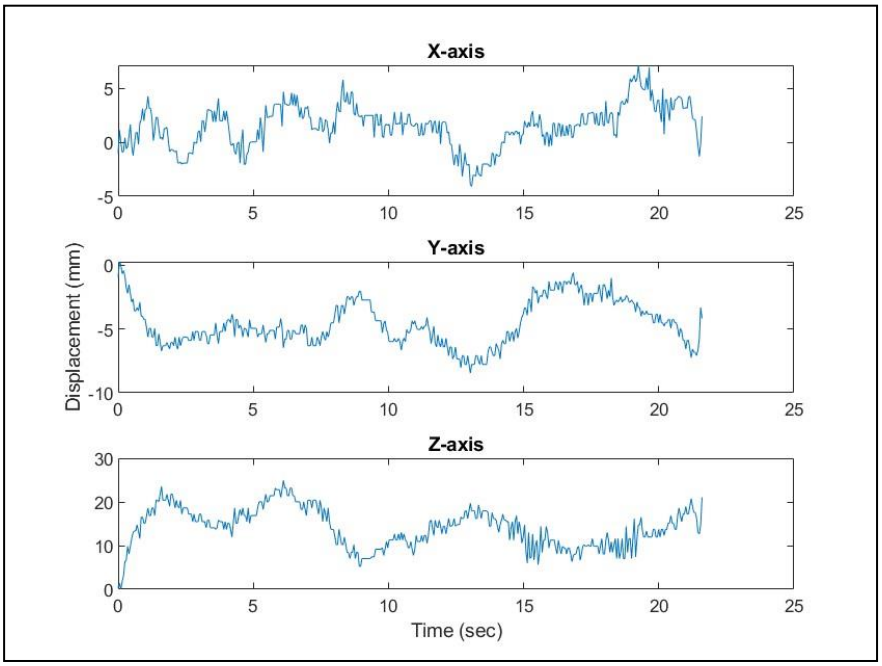

Figure 24. Displacements of the wind turbine tower using Marker-3.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 May 2022 doi:10.20944/preprints202205.0298.v1

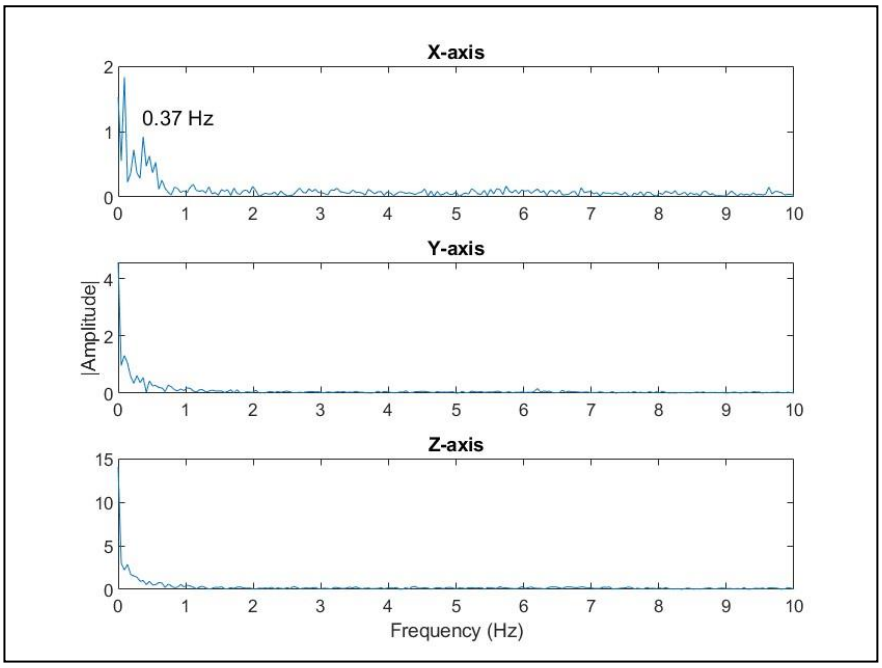

Figure 25. Vibrational results of wind turbine tower using Marker-3.

3.2.1. DIC Processing Discussion

In detail, Figures 20, 22, and 24 are the x,y and z axis displacements as obtained from DIC processing from the set of real-world wind turbine images. DIC is a process that inputs images taken at different intervals but with as little movement as possible from a stationary camera. Minute differences may affect results from DIC so the set of images taken are crucial in terms of accuracy and stillness. Figures 21, 23 and 25 show the vibrational results of the wind turbine, these plots are obtained from calculations of the displacements.

In these results, we observe that our algorithm's output markers are capable of use for error-free DIC processing, and further study in the future may include other large structures of interest for SHM purposes. The process was completed without any noticeable errors and were comparable to results from manual marker segmentation through visual inspection.

5. Conclusions

Our novel algorithm is capable of automatically outputting relevant, salient, unique segmented markers given a dataset of images meant for DIC processing. The outputted segmented markers are considered fiducial unique patterns that can be directly used when preforming 2D or 3D-DIC on large structures. Validation results are favorable and suggest the efficacy of the use of this algorithm as a pre-processing step before SHM DIC processing. This algorithm is extremely significant and useful for future large structural health monitoring purposes as it will reduce manual labor, reduce computational errors, and increase the efficiency of the entire DIC processing pipeline.

Supplementary Materials: The following algorithm and Matlab code implementation can be downloaded at: https://bit.ly/37iaOq3.

Author Contributions: All authors have contributed equally to this study.

Funding: This research received no external funding.

 doi:10.20944/preprints202205.0298.v1

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to acknowledge the support from Chaoyang University of Technology in Taiwan, and Toronto Metropolitan University in Canada.

References

[1] C. Niezrecki, J. Baqersad, and A. Sabato, "Digital Image Correlation Techniques for Non-Destructive Evaluation and Structural Health Monitoring," *Handbook of Advanced Non-Destructive Evaluation*, p. 46, 2018.

[2] D. Reagan, A. Sabato, and C. Niezrecki, "Feasibility of using digital image correlation for unmanned aerial vehicle structural health monitoring of bridges," *Structural Health Monitoring*, vol. 17, no. 5, pp. 1056–1072, 2018.

[3] J. Blaber, B. Adair, and A. Antoniou, "Ncorr: open-source 2D digital image correlation matlab software," *Experimental Mechanics*, vol. 55, no. 6, pp. 1105–1122, 2015.

[4] C. C. Chan, D. Kumar, and C.-H. Chiang, "Coarse and fine localized CNN classifier for intelligent DIC preprocessing in large structure health monitoring sample," in *Nondestructive Characterization and Monitoring of Advanced Materials, Aerospace, Civil Infrastructure, and Transportation XV*, 2021, vol. 11592, p. 115920L.

[5] C. C. K. Chan, C.-C. Chen, S. Delaney, and A. Ferworn, "A Markerless High Resolution Structural Health Monitoring Framework for Smart Cities," in *2021 IEEE Technology & Engineering Management Conference-Europe (TEMSCON-EUR)*, 2021, pp. 1–5.

[6] X.-W. Ye, C.-Z. Dong, and T. Liu, "A review of machine vision-based structural health monitoring: methodologies and applications," *Journal of Sensors*, vol. 2016, 2016.

[7] S. Sony, S. Laventure, and A. Sadhu, "A literature review of next-generation smart sensing technology in structural health monitoring," *Structural Control and Health Monitoring*, vol. 26, no. 3, p. e2321, 2019.

[8] Z. Zhu, G. Zhang, and H. Li, "SURF feature extraction algorithm based on visual saliency improvement," *International Journal of Engineering and Applied Sciences*, vol. 5, no. 3, p. 257267.

[9] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *2009 IEEE conference on computer vision and pattern recognition*, 2009, pp. 1597–1604.

[10] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 6, pp. 1397–1409, 2012.

[11] R. Huang, W. Feng, M. Fan, L. Wan, and J. Sun, "Multiscale blur detection by learning discriminative deep features," *Neurocomputing*, vol. 285, pp. 154–166, 2018.

[12] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[13] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[14] A. Kuznetsova, T. Maleva, and V. Soloviev, "Detecting apples in orchards using YOLOv3 and YOLOv5 in general and close-up images," in *International Symposium on Neural Networks*, 2020, pp. 233–243.

[15] D. Kumar and C.-H. Chiang, "In-house digital image correlation technique for large structures," in *AIAA Scitech 2021 Forum*, 2021, p. 1719.