Video mosaic: employing an efficient ORB feature extraction technique with hamming distance matching for enhanced performance

Shridhar H[1], Sunil S. Harakannanavar[2], Vidyashree Kanabur[3], Jayalaxmi H[4]
[1]Department of Electronics and Communication Engineering, Government Engineering College, Haveri, Visvesvaraya Technological University, Belagavi, India
[2]Department of Electronics and Communication Engineering, Nitte Meenakshi Institute of Technology, Bangalore, India
[3]Department of Electronics and Communication Engineering, National Institute of Technology, Mangalore, India
[4]Department of Electronics and Communication Engineering, Acharya Institute of Technology, Bangalore, India

Article Info

Article history:

Received Mar 31, 2024
Revised Sep 13, 2024
Accepted Oct 30, 2024

Keywords:

Homography
Matrix
Mosaicing
Panorama images
Threshold

ABSTRACT

Video mosaicing is a computer vision and image processing technique used to create a panoramic or wide-angle view from a sequence of video frames. The goal is to seamlessly combine multiple video frames to form a larger and more comprehensive view of a scene. In recent years, the field of image processing has witnessed a growing interest in video mosaic research owing to its application in surveillance and defense applications. This paper introduces an automatic algorithm for video mosaic creation, addressing the alignment and blending of non-overlapping frames within each input video. The proposed algorithm navigates through several key steps to achieve a seamless and continuous mosaic, particularly tackling issues related to camera motion and content variations across frames. The effect of the good number of matches to be chosen while performing frame stitching is evaluated. The proposed algorithm effectively produces a video mosaic with aligned and blended non-overlapping frames, resulting in a visually continuous mosaic. The output video serves as a testament to the algorithm's prowess in addressing challenges related to video frame alignment and blending.

Corresponding Author:

Shridhar H
Department of Electronics and Communication Engineering, Government Engineering College
Haveri, Visvesvaraya Technological University
Belagavi, 590018,
Email: shridhar47@gmail.com

1. INTRODUCTION

Video mosaicing is a computer vision technique used to create a seamless panoramic or mosaic image from a sequence of overlapping video frames. This process is like creating a panorama from multiple still images, but in this case, it's applied to a continuous video stream. Video mosaicing has various applications, including surveillance, robotics, virtual reality, and more. It allows for the creation of a panoramic or wide-angle representation of a scene, which can provide a better understanding of the environment or enhance the visual experience.

Video mosaicing is crucial for numerous applications where creating a broad, comprehensive view from limited perspectives is necessary. It is extensively used in fields such as surveillance, remote sensing, medical imaging, and virtual reality. In surveillance, video mosaicing allows for the monitoring of large areas using fewer cameras, providing security personnel with a comprehensive view of the environment.

In medical imaging, it can be used to combine images from endoscopic or microscopic cameras to provide detailed, large-scale views of tissues or organs, aiding in more accurate diagnoses. For virtual reality, video mosaicing enhances user experience by offering more immersive environments through panoramic views. The impact of video mosaicing lies in its ability to enhance the visualization of spatial environments, improving the effectiveness of analysis and decision-making processes across various industries. Additionally, video mosaicing contributes to resource optimization, reducing the need for multiple camera setups and thereby lowering costs and complexity in imaging systems. This technique is vital for situations requiring enhanced situational awareness, enabling comprehensive analysis from a limited dataset. Video mosaicing involves several key aspects as described in Figure 1.

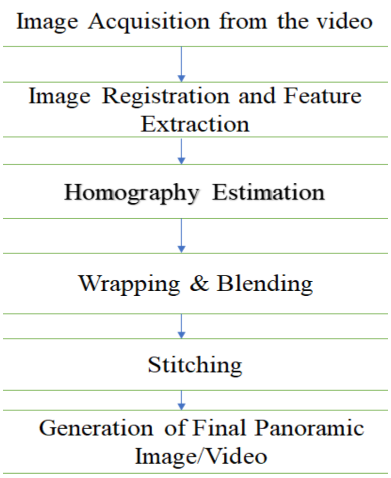

Figure 1. Key concepts and steps involved in video mosaicing

- Image acquisition: video mosaicing typically starts with capturing a sequence of overlapping images or video frames using a camera or other imaging devices.
- Image registration and feature extraction: the first step in video mosaicing is aligning or registering consecutive video frames. This process involves finding correspondence between features or points in the frames and transforming them to match a common reference frame. Common techniques for image registration include feature matching, optical flow, and rigid or affine transformations. These features help to find corresponding points in different frames.
- Homography estimation: a mathematical transformation, known as a homography or a perspective transformation is calculated to map points from one frame to another. This transformation considers the camera’s position and orientation changes between frames. In cases where the camera undergoes projective transformations (e.g., rotations and translations), a homography matrix is used to warp and align frames correctly. This transformation accounts for perspective distortions and ensures that the frames fit together properly.
- Warping and blending: when we stitch two or more images, the edges at the point of stitch are prominent, and so the stitched image looks artificial and not realistic. Therefore, we need image warping and blending operations. This step ensures that the images or frames fit together seamlessly, minimizing visible seams or artifacts. (We can just include the sentence in green instead of red. Even the sentence in black gives the meaning as in green).
- Stitching: the aligned frames are finally stitched together to create a single, continuous panoramic image or video. The goal is to create a single, seamless mosaic image from the overlapping frames. There are different stitching algorithms available, ranging from simple linear blending to more complex methods like graph-cut optimization or bundle adjustment.

The rest of the paper is organized as follows. After formally introducing the topic of video mosaicing, we provide an overview of the various challenges associated with this technique in section 2. We then highlight the motivation behind our work in section 3. Section 4 discusses related work by other researchers and identifies gaps in the existing literature. Sections 5 and 6 present our problem statement and the implementation of our proposed solution, detailing the methodology and techniques employed. We provide a step-by-step analysis of our results in section 7. Finally, we conclude the paper with a summary of our proposed work and a discussion of the scope for future research directions in section 8.

2. CHALLENGES OF VIDEO MOSAICING

Video mosaicing, while a valuable technique, comes with several challenges and limitations that need to be addressed for successful implementation. Here are some of the primary challenges associated with video mosaicking.

- Camera calibration: accurate camera calibration is essential for video mosaicing. Variations in camera parameters, such as focal length, distortion, and sensor characteristics, can lead to inaccuracies in the stitching process. Proper calibration techniques are required to compensate for these variations.
- Motion compensation: both camera and scene motion pose significant challenges. Camera movements, such as panning, tilting, or zooming, need to be accounted for to estimate the correct transformations between frames. Additionally, dynamic scene elements, like moving objects or people, can cause misalignment and artifacts in the mosaic.
- Parallax effects: parallax occurs when objects at different depths in the scene appear to move relative to each other when the camera moves. Handling parallax is challenging, as it can lead to misalignment and ghosting in the mosaic. Advanced techniques may be required to mitigate these effects.
- Feature matching and tracking: feature-based video mosaicing relies on detecting and matching key points in frames. Challenges arise when the scene lacks distinctive features or when there are changes in lighting conditions. Robust feature matching and tracking algorithms are necessary to handle these situations.
- Computational intensity: video mosaicing can be computationally intensive, especially when dealing with high-resolution video streams or many frames. Real-time video mosaicing applications require efficient algorithms and hardware acceleration to process frames quickly.
- Lens distortions: camera lenses introduce distortions that can affect the accuracy of the mosaicing process. Correcting lens distortions is crucial for aligning frames correctly.
- Seamless blending: achieving seamless blending between adjacent frames is a non-trivial task. Mismatches in brightness, color, or exposure can result in visible seams in the mosaic. Sophisticated blending techniques are needed to produce high-quality results.
- Resource constraints: in resource-constrained environments, such as mobile devices or drones, the processing power and memory available for video mosaicing may be limited. Efficient algorithms are required to meet these constraints.
- Robustness and reliability: video mosaicing systems must be robust and reliable in various real-world scenarios. They should handle different lighting conditions, weather conditions, and scene complexities while providing accurate and consistent results.
- User interaction: in some cases, user interaction may be required to correct errors or guide the mosaicing process, especially in challenging scenarios.

Video mosaicing is a powerful technique for creating panoramic or wide-angle views from a sequence of images or video frames. It involves feature extraction, transformation estimation, and image stitching to produce a seamless representation of a scene, with applications ranging from surveillance to entertainment and beyond. Despite these challenges, advances in computer vision and image processing have led to the development of more robust and efficient video mosaicing techniques.

3. MOTIVATION OF VIDEO MOSAICING

The motivation behind video mosaicing stems from the need to capture and represent a broader and more immersive view of a scene or environment than what a single camera frame can provide. This technique addresses various practical and conceptual needs across different fields and applications:

- Enhanced field of view: one of the primary motivations is to expand the field of view beyond the limitations of a single camera frame. Video mosaicing enables the creation of panoramic or wide-angle images or videos, allowing viewers to see more of the scene without the need for specialized wide-angle lenses or equipment.
- Improved visualization: video mosaics provide a more comprehensive and coherent representation of a scene. This can enhance the visualization and understanding of complex environments, making it easier to analyse, navigate, or appreciate the surroundings.
- Immersive experiences: in applications like virtual reality and augmented reality, video mosaicing contributes to creating immersive experiences. By stitching together multiple frames or videos, it allows users to explore and interact with virtual environments in a natural and engaging way.
- Surveillance and security: video mosaicing is valuable in surveillance and security systems. It enables continuous monitoring of large areas using a single camera or a network of cameras. This can be crucial for detecting and tracking intruders or unusual activities.

- Navigation and robotics: video mosaicing is essential for navigation and autonomous robotics. It helps robots and autonomous vehicles understand and navigate their surroundings more effectively, whether indoors (e.g., in warehouses) or outdoors (e.g., for self-driving cars).
- Cultural heritage preservation: video mosaicing is used to capture high-resolution images of historical sites, artifacts, and artwork. It aids in preserving cultural heritage by creating detailed visual records for documentation and restoration purposes.
- Scientific and environmental monitoring: researchers use video mosaicing to study and document natural environments, ecosystems, and geological features. It allows for the creation of panoramic views for scientific analysis and environmental monitoring.
- Entertainment and media: in the entertainment industry, video mosaics are employed for creating breathtaking cinematic shots, enhancing storytelling, and offering viewers a more immersive visual experience.
- Architectural and real estate: video mosaicing helps in showcasing architecture and real estate properties. It enables the creation of immersive virtual tours, allowing potential buyers or clients to explore properties remotely.
- Education and training: video mosaics can be used in educational settings to provide students with interactive and informative visual content, allowing them to explore historical sites, scientific phenomena, and more.
- Emergency response: during disaster management and emergency response operations, video mosaicing can provide an overview of the affected areas, aiding in decision-making and resource allocation.

4. LITERATURE SURVEY

Video mosaicing is a technique that involves stitching together multiple video frames to create a panoramic or mosaic view, which is useful in applications such as surveillance, robotics, and virtual reality. Many researchers have contributed to advancing this field through innovative approaches. For instance, Wang *et al.* [1] developed a method using an improved Harris algorithm for feature point extraction and a two-level Gaussian pyramid for smoothing, achieving high-resolution panoramic images up to 8K. Sumantri and Park [2] proposed a network for synthesizing high-quality 360-degree panoramas with a focus on reducing high-frequency artifacts, thus enhancing visual realism. Bai [3] provided an overview of image mosaic technology, emphasizing algorithms for image preprocessing, registration, and fusion.

Parisotto *et al.* [4] described a primal-dual optimization algorithm, highlighting the significance of optimization techniques in video mosaicing. Changan and Chilveri [5] tailored the Harris corner detection algorithm for stereo image feature matching, while Du *et al.* [6] enhanced medical image visualization by combining smooth, texture, and edge information. X. Lan *et al.* [7] combined GMS characteristics with the RANSAC algorithm to improve image registration accuracy and efficiency. Yang and Mao [8] utilized an improved SIFT algorithm for feature extraction in intelligent surveillance systems.

Han *et al.* [9] introduced a corner detection algorithm combining Harris and SUSAN methods to refine results. Li *et al.* [10] developed a mosaic and hybrid fusion algorithm based on pyramid decomposition, improving color fidelity and ghosting elimination. Kang *et al.* [11] focused on comprehensive panoramic image stitching, addressing challenges related to color fusion and texture features. Xiu *et al.* [12] leveraged NSCT and ANMF algorithms to enhance image fusion efficiency. Ren and Ren [13] used the SURF algorithm for feature extraction, offering better stitching quality with adaptive fusion. R. Ren and Q. Lee *et al.* [14] introduced a real-time panoramic video mosaic system using GPU acceleration, emphasizing practical applications and overcoming feature point scarcity. Nie *et al.* [15] developed Rich360, a system that addresses parallax and enhances panoramic video experiences. Wei *et al.* [16] focused on video stitching and stabilization for handheld cameras, dealing with challenges like shakiness and parallax. Li *et al.* [17] presented a Bayesian fusion technique for high-resolution image recovery from degraded observations. Shridhar *et al.* [18] addressed distortion in fish-eye lens panoramas through multi-band image blending, improving visual detail in virtual reality. Wang *et al.* [19] implemented a pipeline using histogram equalization, feature extraction, RANSAC motion estimation, and image warping to create seamless panoramas. Lin *et al.* [20] automated coastline image stitching with a wavelet fusion approach, mitigating shadow issues and enhancing visual quality. Finally, Shridhar *et al.* [21] proposed a novel stitching method prioritizing natural mosaics and addressing challenges like camera motion and illumination changes, showcasing robustness and automation.

Most existing video mosaicing techniques rely on complex image processing methods. In our proposed work, we address this complexity by providing a simple solution for creating seamless video mosaics. We include histogram matching to align video frames, which helps create a smooth panorama.

Additionally, we select frames with a higher percentage of matches to construct the mosaic. To the best of our knowledge, this novel approach has not been previously explored in the literature.

5. PROPOSED METHOD

The proposed model is shown in Figure 2. The objective of video mosaicing is to create a single, continuous, and panoramic representation of a scene or environment by seamlessly stitching together multiple video frames or images. In addition to this:

- A seamless integration: a video mosaicing model is developed to create a mosaic that appears as a single, coherent image or video, with no visible seams or artifacts at the boundaries where frames are joined. Achieving seamless integration is crucial to provide a natural and immersive viewing experience.
- Evaluation of performance: to evaluate the performance of the proposed video mosaicing model, we focus on its ability to produce high-quality mosaics that maintain visual fidelity and spatial continuity.

The steps involved matching the current frame with the previous frame used in our code to match features between the current and previous frames for video mosaicing:

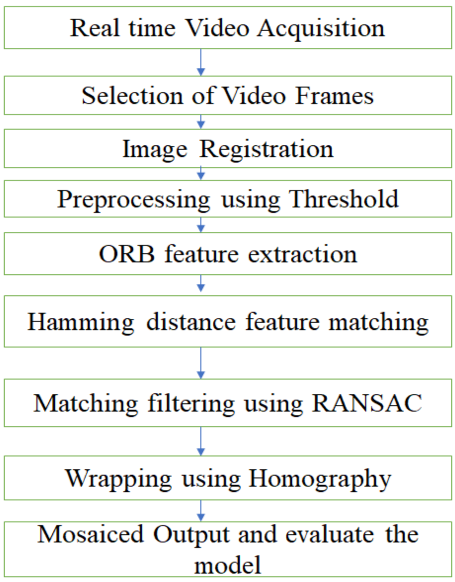

Figure 2. Proposed model

a. ORB feature detection
   - In our code, we use the ORB feature detection algorithm to detect key points (features) in both the current and previous frames.
   - ORB is efficient and well-suited for real-time applications, making it a popular choice for feature detection.

b. ORB feature description
   - For each key point detected by ORB, descriptors are computed. Descriptors encode information about the local image patch around each key point.

c. Feature matching
   - We use a brute-force matcher (cv2.BFMatcher) to match the descriptors of key points between the current and previous frames.
   - The cv2.BFMatcher uses the Hamming distance as the distance metric and performs cross-checking to find mutual matches.

d. Matching filtering with RANSAC
   - After obtaining the initial matches, we apply a filter to remove any incorrect correspondence.
   - Our code checks the number of matches (len(matches)) and uses this as a filtering criterion.
   - If there are more than threshold matches, we use the RANSAC algorithm to estimate a transformation matrix (homography) that aligns the frames.
   - RANSAC helps in filtering out outliers and obtaining a robust transformation model.

e. Warping with homography
  - Once we have estimated the homography matrix using RANSAC, we warp the current frame to align it with the previous frame.
  - The cv2.warpPerspective function is used to apply this transformation and align the frames.

f. Updating previous frame:
  - The current frame, after warping, becomes the new "previous frame" for the next iteration.
  - This step ensures that you maintain a continuous alignment of frames as you process the video.

By following these steps, we can match features between frames, filter the matches, estimate a transformation, and create a mosaic video. The ORB algorithm plays a crucial role in feature detection and description, making it possible to find correspondence between frames and align them effectively for mosaicing. The threshold on the number of matches controls when RANSAC-based transformation estimation is applied, contributing to robustness in the presence of varying levels of overlap between frames.

6. PROPOSED VIDEO MOSAICING ALGORITHM

Step 1: input: provide the path to the input video file (video_path) and set an overlap percentage threshold (overlap_threshold).

Step 2: initialization:
- Initialize the video capture object (cap) to open the input video file.
- Check if the video file opened successfully. If not, return an error.
- Read the first frame from the video to determine the frame dimensions.

Step 3: output video setup
- Define the path and codec for the output video.
- Initialize the video writer object with the specified codec and frame dimensions for creating the output video.

Step 4: frame processing loop
- Initialize variables.
  a. Mosaic: create an empty mosaic and set it to the first frame.
  b. prev_frame: set the previous frame as the first frame.
- Loop through the video frames
  a. Read the next frame from the video.
  b. Calculate the overlap percentage (overlap_percent) between the previous frame and the current frame.
  c. Check if overlap_percent is less than the specified overlap_threshold.

If overlap_percent is below the threshold
- Use ORB feature detection and matching to align and blend the frames.
- If there are enough good matches (e.g., more than 400), compute a homography matrix (M) to align the frames.
- Warp the current frame to align with the previous frame using the homography matrix. Update the mosaic by blending the aligned frame with the previous mosaic.
- Set the previous frame to the non-overlapping frame. Write the mosaiced frame to the output video.
- Display the mosaiced frame and check for a keyboard interruption (e.g., 'q' key) to stop the process.

Step 5: cleanup and finalization
- After processing all frames, release the video capture and writer objects.
- Close any open windows.

Step 6: output: the algorithm produces a video mosaic where non-overlapping frames are aligned and blended to create a continuous mosaic, and the output video is saved.

This algorithm allows you to control which frames are included in the mosaic based on their overlap with the previous frame, making it suitable for cases where you want to mosaic only non-overlapping frames in a video sequence. You can adjust the overlap_threshold and other parameters to customize the behavior of the algorithm.

7. ROLE OF THE NUMBER OF MATCHES ON VIDEO MOSAIC

The number of matches in a feature-based image stitching or mosaicing process can indeed affect the quality of the resulting mosaic.

- Robustness to distortions: a larger number of matches means that more key points in the images have been successfully matched [22]. This is particularly beneficial when the input images have significant

geometric and radiometric variations. Robust matches help ensure that the stitching process can handle distortions such as rotations, scaling, and perspective changes.
- Better estimation of transformations: more matches provide more data points for estimating the transformations (homography or affine transformations) that align the images. When there are many matches, it's more likely that a consistent transformation can be estimated, which helps reduce misalignment and distortions in the mosaic [23].
- Outlier rejection: feature matching often involves the step of rejecting outlier matches. Having more matches allows for a more effective outlier rejection process. Outliers can be caused by factors like moving objects or image noise. A robust matching and rejection process can remove these outliers and improve the overall quality [24].
- Completeness of information: a larger number of matches means that you're using more information from the input images to create the mosaic. This results in a more comprehensive representation of the scene.

8. RESULTS AND DISCUSSION

The proposed model is tested on three datasets in Python software having 32-bit and has been executed in system with configuration i4 processor, 8 GB RAM, 2 GB cache memory and 2.8 GHz processor. Figure 3 shows the images when the threshold for the number of matches is much less than the actual matches between the previous and the current frame. As can be seen from the above figure, the images are misaligned. Figure 4 shows the effect on the mosaic frame when the threshold for the number of matches is much closer to the actual matches between the previous and the current frame. It is observed that the problem of misalignment is overcome by increasing the threshold. Therefore, as the threshold for the number of matches between the frames increases, it is seen that the video mosaic is much clearer, smoother, and aligned well. As the threshold increases, it will increase the computational time but again it's in milliseconds, so it doesn't really make a big impact on the mosaic output.

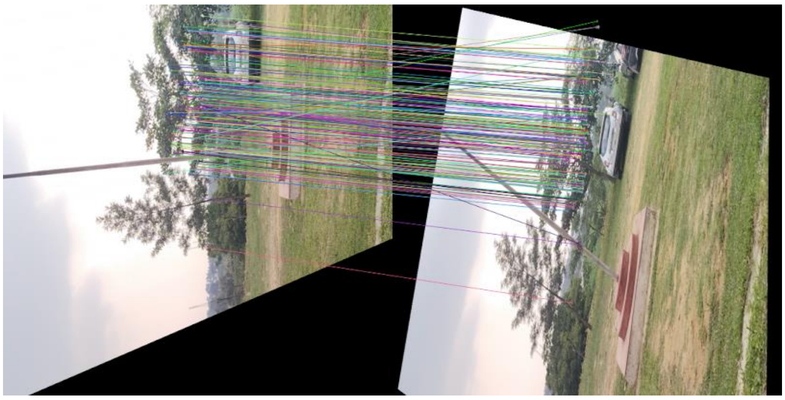

Figure 3. Real time scene showing more matching points

Figure 4. Real time scene showing less matching points

In the proposed work, we choose a video covering largely varying frames. We experimented with various values of threshold of good number of matches between the two adjacent frames. Figure 5 shows an image comparing the same two frames with the threshold on number of good matches being in Figures 5(a) 10, Figure 5(b) 100, and Figure 5(c) 400 respectively.

(a)

(b)

(c)

Figure 5. Threshold on number of good matches (a) 10, (b) 100, and (c) 400

We noticed that when threshold is 10, the current frames are highly misaligned to previous frame, and they cannot be stitched together to get a panoramic view. Also, too many frames are to be included to get an entire video mosaic. When the threshold is 400, we see that the scenes can be captured easily and some of the frames can be easily skipped to get the panoramic view of the entire scene. As we further increase the threshold, the effect is the same.

The evaluation metric [25] measures such as root mean squared error (RMSE) to quantify the accuracy of motion estimation. The RMSE in regenerated image is theoretically calculated using (1) and tabulated in Table 1.

$$MSE = \left(\frac{1}{MN}\right) \sum_{i=0}^{M-1} [x(i,j) - x'(i,j)]^2 \sum_{j=0}^{N-1} [x(i,j) - x'(i,j)]^2 \quad (1)$$

Where MSE is the RMSE obtained, x (i, j) is the original image, x’ (i, j) is the regenerated image, and M*N is the total number of rows and columns of the image. Table. 1 shows the theoretical RMSE calculated on each of the regenerated images.

Table 1. Theoretical RMSE calculated on each of the regenerated images

Figure details	No. of good matches	RMSE
5(a)	10	3.84
5(b)	100	6.12
5(c)	400	7.2

We also tried to see the impact of good matches on RMSE between the two consecutive frames. This gives us ideas of how different these two consecutive frames. The larger the value of RMSE, the more different the frames and less is the probability of including overlapping frames in the output video mosaic. But a too high of RMSE will misalign the video frames and increases the chances of missing relevant frame details. Therefore, we experimented with different numbers of good matches between the frames and their corresponding RMSE as tabulated in Table 1. It is seen that 400 is the optimum number of good matches between frames as it provides a good trade-off between computation efficiency and quality of output video mosaic. Here, it is important to note that an extremely large number of matches aren't always better. To many matches might introduce more potential for errors, especially when dealing with scenes that have repeated patterns or a lot of clutter. Additionally, computational resources can become a limiting factor as the number of matches' increases, which can affect the stitching process's efficiency. So, while a higher number of matches are desirable, there is usually a practical range based on the nature of the scene and the computational resources available. Balancing the number of matches and the quality of matches is key to achieving the best mosaic results for your specific application and that is why we choose 400 as the optimal number of good matches to obtain a video panoramic view.

9. CONCLUSIONS AND FUTURE SCOPE

The proposed algorithm addresses challenges in creating seamless and continuous video mosaics, focusing on camera motion and content variations across frames. The core algorithm resides in a frame processing loop, initializing variables like a mosaic container and the previous frame. Iterating through video frames, it calculates overlap percentages. When below the threshold, the algorithm uses ORB feature detection and matching for frame alignment. The successful matches yield a homography matrix (M). The current frame warps to align with the previous one using the matrix, updating the mosaic through blending. This iterative process continues until the resulting mosaic frame to the output video, displayed, and monitored for interruptions. The algorithm effectively produces a video mosaic with aligned, blended non-overlapping frames, highlighting its prowess in addressing video frame alignment and blending challenges. The output video serves as evidence of the algorithm's capabilities. Our findings demonstrate the algorithm's potential to impact the research field and offer valuable insights for practical applications. By providing a robust and efficient solution, this work paves the way for future advancements in video mosaicing technology, contributing to the broader goals of enhancing visual experiences and expanding the capabilities of computer vision systems. While, the proposed algorithm shows promise, there are opportunities for further research and development. Additionally, expanding the algorithm to handle higher resolutions and more significant parallax effects could enhance its applicability to more challenging scenarios. Exploring hardware acceleration options, such as GPU processing, could also optimize the algorithm for real-time applications in resource-constrained environments.

REFERENCES

[1] K. Wang, T. Yang, X. Zhang, and M. Tang, "An improved ultra-high definition panoramic video mosaic method for airport scene," in *2022 IEEE 8th International Conference on Computer and Communications, ICCC 2022*, Dec. 2022, pp. 2076–2080, doi: 10.1109/ICCC56324.2022.10065899.

[2] J. S. Sumantri and K. Park, "360 panorama synthesis from a sparse set of images on a low-power device," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 1179–1193, 2020, doi: 10.1109/TCi.2020.3011854.

[3] X. Bai, "Overview of image mosaic technology by computer vision and digital image processing," in *Proceedings of 2021 IEEE International Conference on Data Science and Computer Application, ICDSCA 2021*, Oct. 2021, pp. 569–572, doi: 10.1109/ICDSCA53499.2021.9650329.

[4] S. Parisotto, L. Calatroni, A. Bugeau, N. Papadakis, and C. B. Schonlieb, "Variational osmosis for non-linear image fusion," *IEEE Transactions on Image Processing*, vol. 29, pp. 5507–5516, 2020, doi: 10.1109/TIP.2020.2983537.

[5] K. S. Changan and P. G. Chilveri, "Stereo image feature matching using Harris corner detection algorithm," in *International Conference on Automatic Control and Dynamic Optimization Techniques, ICACDOT 2016*, Sep. 2017, pp. 691–694, doi: 10.1109/ICACDOT.2016.7877675.

[6] J. Du, W. Li, and H. Tan, "Three-layer image representation by an enhanced illumination-based image fusion method," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 4, pp. 1169–1179, Apr. 2020, doi: 10.1109/JBHI.2019.2930978.

[7] X. Lan, B. Guo, Z. Huang, and S. Zhang, "An improved UAV aerial image mosaic algorithm based on GMS-RANSAC," in *2020 IEEE 5th International Conference on Signal and Image Processing, ICSIP 2020*, Oct. 2020, pp. 148–152, doi: 10.1109/ICSIP49896.2020.9339283.

[8] P. Yang and Z. Mao, "The research on image mosaics for intelligent video surveillance," in *ICINA 2010 - 2010 International Conference on Information, Networking and Automation, Proceedings*, Oct. 2010, vol. 2, pp. V2-59-V2-62, doi: 10.1109/ICINA.2010.5636784.

[9] S. Han, W. Yu, H. Yang, and S. Wan, "An improved corner detection algorithm based on harris," in *Proceedings 2018 Chinese Automation Congress, CAC 2018*, Nov. 2018, pp. 1575–1580, doi: 10.1109/CAC.2018.8623814.

[10] H. Li, J. Wang, and C. Han, "Image mosaic and hybrid fusion algorithm based on pyramid decomposition," in *Proceedings - 2020 International Conference on Virtual Reality and Visualization, ICVRV 2020*, Nov. 2020, pp. 205–208, doi: 10.1109/ICVRV51359.2020.00049.

[11] H. Kang, L. Xuefei, and Z. Wenhui, "An adaptive fusion panoramic image mosaic algorithm based on circular LBP feature and HSV color system," in *Proceedings of 2020 IEEE International Conference on Information Technology, Big Data and Artificial Intelligence, ICIBA 2020*, Nov. 2020, pp. 94–100, doi: 10.1109/ICIBA50161.2020.9277348.

[12] C. Xiu, J. Fang, and J. Zhang, "Image stitching method based on adaptive weighted fusion," *Proceedings of the 33rd Chinese Control and Decision Conference, CCDC 2021*, pp. 3099–3103, 2021, doi: 10.1109/CCDC52312.2021.9602529.

[13] H. Ren and Q. Ren, "Real-time panoramic video mosaic system based on mapping table and GPU acceleration," in *Proceedings of the 33rd Chinese Control and Decision Conference, CCDC 2021*, May 2021, pp. 812–816, doi: 10.1109/CCDC52312.2021.9601676.

[14] J. Lee, B. Kim, K. Kim, Y. Kim, and J. Noh, "Rich360: optimized spherical representation from structured panoramic camera arrays," *ACM Transactions on Graphics*, vol. 35, no. 4, pp. 1–11, Jul. 2016, doi: 10.1145/2897824.2925983.

[15] Y. Nie, T. Su, Z. Zhang, H. Sun, and G. Li, "Dynamic video stitching via shakiness removing," *IEEE Transactions on Image Processing*, vol. 27, no. 1, pp. 164–178, Jan. 2018, doi: 10.1109/TIP.2017.2736603.

[16] Q. Wei, N. Dobigeon, and J. Y. Tourneret, "Bayesian fusion of multi-band images," *IEEE Journal on Selected Topics in Signal Processing*, vol. 9, no. 6, pp. 1117–1127, Sep. 2015, doi: 10.1109/JSTSP.2015.2407855.

[17] X. Li, W. Zhu, and Q. Zhu, "Panoramic video stitching based on multi-band image blending," in *Tenth International Conference on Graphics and Image Processing (ICGIP 2018)*, May 2019, p. 82, doi: 10.1117/12.2524247.

[18] H. Shridhar, S. S. Harakannanavar, V. Kanabur, and H. Jayalaxmi, "Performance analysis of feature based multiple descriptors with histogram equalization for image mosaicing," *Indian Journal Of Science And Technology*, vol. 16, no. 31, pp. 2380–2387, Aug. 2023, doi: 10.17485/ijst/v16i31.433.

[19] B. Wang, H. Li, and W. Hu, "Research on key techniques of multi-resolution coastline image fusion based on optimal seam-line," *Earth Science Informatics*, vol. 13, no. 2, pp. 333–344, Jun. 2020, doi: 10.1007/s12145-019-00421-z.

[20] C. C. Lin, S. U. Pankanti, K. N. Ramamurthy, and A. Y. Aravkin, "Adaptive as-natural-as-possible image stitching," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2015, vol. 07-12-June-2015, pp. 1155–1163, doi: 10.1109/CVPR.2015.7298719.

[21] H. Shridhar, B. Neelgar, and R. Premananda, "An improved image fusion technique based on wavelet and edge gradients to augment spatial prominence of image features," in *3rd International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques, ICEECCOT 2018*, Dec. 2018, pp. 416–421, doi: 10.1109/ICEECCOT43722.2018.9001470.

[22] A. L. R. Siriani, I. B. de C. Miranda, S. A. Mehdizadeh, and D. F. Pereira, "Chicken tracking and individual bird activity monitoring using the BoT-SORT algorithm," *AgriEngineering*, vol. 5, no. 4, pp. 1677–1693, Sep. 2023, doi: 10.3390/agriengineering5040104.

[23] J. M. Górriz *et al.*, "Computational approaches to explainable artificial intelligence: advances in theory, applications and trends," *Information Fusion*, vol. 100, p. 101945, Dec. 2023, doi: 10.1016/j.inffus.2023.101945.

[24] C. W. Lin, S. Hong, M. Lin, X. Huang, and J. Liu, "Bird posture recognition based on target keypoints estimation in dual-task convolutional neural networks," *Ecological Indicators*, vol. 135, p. 108506, Feb. 2022, doi: 10.1016/j.ecolind.2021.108506.

[25] C. Renteria, J. Suárez, A. Licudine, and S. A. Boppart, "Depixelation and enhancement of fiber bundle images by bundle rotation," *Applied Optics*, vol. 59, no. 2, p. 536, Jan. 2020, doi: 10.1364/ao.59.000536.

BIOGRAPHIES OF AUTHORS

Mr. Shridhar H received the B. E degree in Electronics and Communication Engineering (S.T.J.I.T, Ranebennur) and M. Tech degree in the stream of VLSI Design and Embedded Systems (V.T.U. Regional Center) under Visvesvaraya Technological University, Jnana Sangama, Belagavi, Karnataka, India. He is currently working as an assistant professor in Electronics and Communication Engineering Dept, Government Engineering College Haveri-581110, Karnataka India. He has 16 years of teaching experience. He has published around 3 articles in Peer received International Journal. His research interests include computer vision, image pattern recognition, and biometrics. He can be contacted at email: shridhar47@gmail.com.

Dr. Sunil S. Harakannanavar received the B.E. degree in Electronics and Communication Engineering (S.T.J.I.T Ranebennur) and M.Tech degree in the stream of Microelectronics and Control Systems (NMAMIT, Nitte) under Visvesvaraya Technological University, Jnana Sangama, Belagavi, Karnataka, India. He is currently working as an Associate Professor in Electronics and Communication Engineering Dept Nitte Meenakshi Institute of Technology Yelahanka, Bangalore-560064, Karnataka India. He has 14 years of Teaching Experience. He has published around 25 articles in Peer received International Journal. His research interests include computer vision, image pattern recognition, and biometrics. He is a life member of Indian Society for Technical Education (LMISTE). He can be contacted at email: sunilsh143@gmail.com.

Ms. Vidyashree Kanabur holds a BE from KLECET, Belagavi and M. Tech from SIT, Tumakuru and Pursuing Ph.D. in the stream of Department of Electronics and Communication Engineering, National Institute of Technology, Surathkal, Mangalore, Karnataka, India. She has published around 8 articles in in Peer received International Journal. His research interests include computer vision, speech processing, image pattern recognition, and biometrics. She is a member of IEEE. She can be contacted at email: vidyashreerk1992@gmail.com.

Dr. Jayalaxmi H received the Ph.D. degree in VLSI and Image Processing from JNTU, Hyderabad, India, in 2020. She is currently working in the Department of Electronics and Communication Engineering, Acharya Institute of Technology, Bangalore, Karnataka, India. Her current research interests include image/video processing, VLSI and published papers in IEEE, Elsevier Science Direct, Scopus and ACEEE. She can be contacted at email: jayalaxmi@acharya.ac.in.