Performance Investigation on Panoramic Image Stitching Techniques

Bittu Kumar(✉), N. Varshith, K. Siva Nikhitha, Nareddy Sony, Gannoj Koushik, and R. Loukik Karthikeya

Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, Hyderabad 500075, Telangana, India
`bittu.mlrit@gmail.com`

Abstract. In this paper, the performance of various panoramic image stitching techniques applicable across diverse domains, including virtual reality, robotics, remote sensing, and multimedia content creation, is investigated. The assessment involves methodologies such as, Oriented FAST and Rotated BRIEF (ORB), Speeded Up Robust Features (SURF), Binary Robust Independent Elementary Features (BRISK) and Scale-Invariant Feature Transform (SIFT). The study utilizes images of Koneru Lakshmaiah Education Foundation, Hyderabad's main building and library building as input, along with generated noisy images derived from these captures. These images undergo processing using different panoramic image stitching techniques in quiet and noisy environments. Evaluation of these techniques considers simulation time, memory consumption, and a step-by-step analysis of the output panoramic images in noisy environments. From the results, it is observed that the stitched image using BRISK does not produce a correct panoramic image, indicating that under noisy conditions, the BRISK algorithm may not function effectively.

Keywords: Image Stitching · Oriented FAST and Rotated BRIEF · Scale-Invariant Feature Transform · Binary Robust Independent Elementary Features · Speeded Up Robust Features

1 Introduction

Panoramic image stitching represents a revolutionary approach in computer vision and image processing. This technique seamlessly merges numerous individual images, each captured from overlapping perspectives, to create a unified wide-angle view. The underlying process emulates the innate way human vision operates, as our brain instinctively merges visual information from our eyes, crafting a coherent and expansive perception of our surroundings [1].

A panoramic photograph provides a comprehensive and wide-angle view of a subject, capturing more of its surroundings than a typical snapshot. This effect is achieved by seamlessly stitching together multiple shots from various perspectives and situations

P. Pareek et al. (Eds.): IC4S 2024, LNICST 599, pp. 177–190, 2025.
https://doi.org/10.1007/978-3-031-77081-4_15

[2]. Panoramic images come in various forms, including cylindrical panoramas, which offer a 360-degree horizontal view with a limited vertical perspective, and spherical panoramas, providing a complete 360-degree horizontal and vertical view.

The applications of panoramic image stitching are diverse and widespread. In virtual and augmented reality, panoramic imagery is the foundation for immersive experiences, offering users a seamless and all-encompassing view of virtual environments [3]. In robotics, it enhances situational awareness, enabling autonomous systems to navigate and map their surroundings effectively.

In recent years, panoramic image stitching [4] has gained considerable prominence due to its versatile applications across various industries. It has become a foundational technology in virtual reality, robotics, remote sensing, and multimedia content creation. This introductory overview aims to illuminate the fundamental principles and key steps of panoramic image stitching.

The process of panoramic image stitching involves several crucial stages. It commences with extracting distinctive features or critical points from individual images, acting as reference points for matching corresponding elements across images. Subsequently, alignment techniques are applied to ensure precise registration of the images, compensating for variations in viewpoint, scale, and rotation. Finally, blending methods seamlessly merge overlapping regions, resulting in a visually coherent panorama [5].

A pivotal aspect of panoramic image stitching lies in selecting appropriate feature detection and matching algorithms. Techniques such as SIFT [6], SURF [7], ORB and BRISK [8] are employed to identify key-points in the images. The images are transformed into panoramic views based on the matched key-points. This paper includes the performance evaluation of various techniques such as SIFT, SURF, ORB, and BRISK under both quiet and noisy environments.

The remainder of the paper unfolds: Sect. 2 provides a concise literature survey on algorithms relevant to creating panoramic images. Section 3 delves into the methodology employed. Following that, Sect. 4 covers the image database and results. Subsequently, Sect. 5 encapsulates the conclusion drawn from the results and outlines the future scope of the research work.

2 Literature Survey

The objective of this study is to comprehensively evaluate the efficacy of various image stitching techniques, including Oriented FAST and Rotated BRIEF (ORB), Speeded Up Robust Features (SURF), Binary Robust Independent Elementary Features (BRISK) and Scale-Invariant Feature Transform (SIFT). Each technique involves distinct steps, encompassing feature point extraction, description, and matching. This segment integrates a concise literature survey to provide a thorough analysis, delineating these methods' evolutionary trajectory and advancements over time. By exploring the historical context and technological progress, we aim to gain insights into the strengths and limitations of each approach, facilitating informed decision-making in image stitching applications.

In their work [8], the authors present an automatic method for panoramic image stitching utilizing invariant features. This method addresses challenges encountered in traditional alignment techniques and achieves robust and visually appealing results. The approach involves feature detection, matching, and geometric transformation, with its effectiveness demonstrated through experiments, showcasing its impact on computer vision and panoramic imaging technology.

Image stitching is a foundational task within the realm of computer vision and image processing, involving the alignment and fusion of multiple images that exhibit varying perspectives, lighting conditions, and distortions. In a referenced study [9], comprehensive research on image stitching takes a theoretical and comparative approach towards various pre- and post-processing methodologies. The investigation employs a range of techniques including SIFT, its variants, SURF, as well as edge or corner detection algorithms to ensure effective feature extraction. To enhance efficiency in aligning multiple photographs, a hybrid image registration technique is developed. Moreover, the study meticulously evaluates advanced blending techniques using both qualitative and quantitative performance metrics to produce high-resolution panoramic images with a wide field of view.

In [10], Rublee et al. introduce the Oriented FAST and Rotated BRIEF algorithm as a proficient alternative to SIFT and SURF for feature extraction. ORB combines FAST corner detection with BRIEF descriptor computation, balancing speed and robustness. The algorithm's rotation invariance extends its applicability across various computer vision tasks. Experimental results showcase ORB's competitiveness in speed and accuracy, establishing it as an effective tool for feature-based applications.

In their ground-breaking work [11], the authors introduce an innovative approach for generating stereoscopic panoramas, mitigating picture stitching errors and ensuring consistent depth to enrich the stereoscopic viewing experience. Two panoramas are individually stitched through sophisticated techniques before being merged using a customized algorithm, culminating in a seamless and immersive stereoscopic image. This novel method outperforms traditional approaches regarding picture stitching accuracy and the quality of the 3-D effect.

Additionally, in their study [12], researchers propose a cost-effective real-time method for image stitching to capture high-resolution wide-angle videos. Initially, images from multiple fixed-position cameras are combined using standard techniques, followed by remapping through a GPU-based approach. The resulting high-resolution full-field panoramic image is especially advantageous for deep learning-based autonomous video capture. Experimental findings demonstrate that a standard desktop computer achieved a frame rate of 36 frames per second, with favorable results observed in tests assessing image quality.

In [13], Achanta et al. delve into using deep learning techniques within image and video mosaicing. The study thoroughly explores diverse methods and approaches that harness deep learning to enhance the quality and efficiency of mosaicing processes. The survey accentuates their effectiveness across various mosaicing applications, encompassing a wide array of techniques are available convolutional neural networks and recurrent neural networks. The paper addresses challenges, identifies trends, and outlines potential future directions in this dynamic and evolving field.

3 Methodology

The process of panoramic image stitching involves the following key steps:

Step:1. *Image Acquisition:* Capture a series of overlapping scene images using a camera or other imaging device. It's essential to have sufficient overlap between consecutive images for successful stitching.
Step:2. *Image Alignment:* Detect and match key features or interest points between the images. This step establishes correspondences, which are used to align the images. Standard feature detection algorithms include SIFT, SURF, ORB, etc.
Step:3. *Homography Estimation:* Use the matched points to estimate the transformation (homography) that aligns one image to another.
Step:4. *Warping:* Apply the computed homography to warp or transform the images to align properly. This step compensates for viewpoint changes, scale differences, and possible distortions.
Step:5. *Output:* Combine the blended images to produce the final panoramic image. This image provides a wide-angle view of the scene, seamlessly merging the information from the individual frames.
Step:6. *Post-processing:* Apply additional enhancements or corrections to the stitched panorama. This may include color correction, exposure adjustments, or other image processing techniques.
Step:7. *Evaluation:* Assess the quality of the stitched panorama. This can be done qualitatively by visually inspecting the result and quantitatively by measuring factors like alignment accuracy or blending quality.

Oriented FAST and Rotated BRIEF (ORB), Speeded Up Robust Features (SURF), Binary Robust Independent Elementary Features (BRISK) and Scale-Invariant Feature Transform (SIFT) are prominent feature detection and matching algorithms extensively utilized in computer vision, with panoramic image stitching being a notable application domain. These algorithms play crucial roles in identifying distinctive image features and facilitating their alignment for seamless stitching. In the subsequent subsections, we offer comprehensive insights into these methods, elucidating their underlying principles, key characteristics, and notable advancements. By delving into the intricacies of these algorithms, we aim to provide a deeper understanding of their functionalities and potential applications in various image-processing tasks beyond panoramic stitching.

3.1 Scale-Invariant Feature Transform

SIFT is renowned for its robustness in handling variations in scale, rotation, and illumination, making it a cornerstone in computer vision tasks. SIFT operates by identifying key points or features within images that remain invariant despite these transformations. This is achieved through a multi-scale approach, where SIFT detects keypoints across different scales by convolving the image with Gaussian filters of varying sizes. Once key points are identified, SIFT computes descriptors for each keypoint based on gradient information, capturing its local appearance and texture properties. During the matching phase, SIFT utilizes a nearest-neighbor algorithm to pair keypoints across different images. In the context of panoramic image stitching, SIFT's ability to robustly detect and describe distinctive features is instrumental in accurately aligning and blending overlapping regions, ultimately leading to seamless and visually pleasing panoramic images.

3.2 Speeded up Robust Features

It is an algorithm engineered with computational efficiency in mind, rendering it well-suited for real-time applications. Like SIFT, SURF exhibits robustness against variations in scale and rotation, ensuring reliable performance across diverse scenarios. What sets SURF apart is its utilization of integral images to expedite feature descriptor computation, significantly enhancing computational speed. Moreover, SURF leverages Haar wavelet responses for feature extraction, further contributing to its efficiency. This unique combination of techniques enables SURF to outperform traditional methods in terms of speed while maintaining robustness to transformations. As a result, SURF emerges as a compelling choice for applications requiring rapid processing, such as real-time panoramic image stitching tasks. Its ability to deliver swift and accurate results makes it invaluable in scenarios where computational resources are limited or time-sensitive considerations are paramount.

3.3 Oriented FAST and Rotated BRIEF

ORB algorithm is engineered with a primary focus on efficiency and speed, making it a standout algorithm in the realm of computer vision. By integrating the FAST (Features from Accelerated Segment Test) key-point detector with the BRIEF (Binary Robust Independent Elementary Features) descriptor, ORB achieves a harmonious balance between feature detection and description. The FAST algorithm enables ORB to swiftly identify key-points within images, while the BRIEF descriptor facilitates the calculation of binary feature descriptors with remarkable efficiency. A notable enhancement introduced by ORB is its incorporation of rotation invariance, further expanding its applicability across diverse scenarios. This feature makes ORB particularly well-suited for real-time applications where rapid performance is paramount, such as robotics or mobile devices. Additionally, ORB's capabilities extend to feature matching tasks in panoramic stitching applications, where its efficiency and reliability contribute to seamless image alignment and merging processes.

3.4 Binary Robust Invariant Scalable Key-Points

This method is designed for efficiency and robustness. It combines a scale-space pyramid and an adaptive non-maximum suppression algorithm. BRISK seamlessly integrates a detector to pinpoint key-point locations with a binary descriptor. Operating on a scale-space pyramid ensures resilience to scale variations. BRISK shines in scenarios where efficiency and robustness are paramount, making it a valuable choice for feature detection and matching in panoramic image stitching. Its utility becomes particularly evident when computational resources are constrained [15].

4 Results and Discussion

This section unveils the simulated results obtained through various techniques: SIFT, SURF, ORB, and BRISK. The evaluation encompasses testing these techniques under a noise-laden environment. Input images or unstitched images are systematically stored in a designated folder. This strategic arrangement facilitates seamless access when initiating the stitching process.

Fig. 1. (**a**). Left-side unstitched source-1 image **(b)** Right-side unstitched source-1 image

Fig. 2. **(a).** Left-side unstitched source-2 image **(b)** Right-side unstitched source-2 image

The input images comprise four distinct types, capturing the main building block (source-1) and the library block (source-2) of Koneru Lakshmaiah Education Foundation, Aziz Nagar, Hyderabad, Telangana-500075, India, using a standard mobile camera. Fig. 1 and Fig. 2 showcase the unstitched source-1 images (left-side and right-side) and unstitched source-2 images (left-side and right-side), respectively. The source-1 image exhibits a pixel size of (865 × 1080) with a memory size of 954 KB for the left-side image and (902 × 1080) pixel size with a memory size of 844 KB for the right-side image. On the other hand, the other source image possesses a (1037 × 822) pixel size with a memory size of 1.29 MB for the left-side image and (1170 × 818) pixel size with a memory size of 1.62 MB for the right-side image.

We introduced noise [16–18] to the unstitched images to glean more nuanced insights into the techniques. Fig. 3 and Fig. 4 depict the unstitched source-1 noisy images (left and right) and unstitched source-2 noisy images (left and right), respectively. The source-1 noisy image features a pixel size of (865 × 1080) with a memory size of 2.57 MB for the left-side image and (902 × 1080) pixel size with a memory size of 2.62 MB for the right-side image. Conversely, the other source image showcases a (1037 × 822) pixel size with a memory size of 2.50 MB for the left-side image and (1170 × 818) pixel size with a memory size of 2.85 MB for the right-side image.

Fig. 3. **(a).** Left-side unstitched source-1 noisy image **(b)** Right-side unstitched source-1 noisy image

Fig. 4. **(a).** Left-side unstitched source-2 noisy image **(b)** Right-side unstitched source-2 noisy image

Table 1 presents the output's simulation time and memory size, i.e., stitched images generated by different methods (SIFT, SURF, ORB, and BRISK) for source-1 images. For these results, both clean and noisy unstitched source-1 images were selected as inputs. The simulation time and memory were calculated once the stitched images (output) were obtained from all methods.

Upon examination of the table, it is evident that the simulation time for ORB is 1.34 s for the clean image and 1.23 s for the noisy image, showcasing lower values compared to

Table 1. Simulation time and memory for source-1 image

Image Type (Clean/noisy)	Techniques	Simulation Time (sec)	Memory (MB)
Clean Image	SIFT	2.44	1.37
	SURF	2.32	1.37
	BRISK	3.03	1.36
	ORB	1.34	1.38
Noisy Image (corrupted by pixel noise with 10 dB)	SIFT	3.04	4.03
	SURF	3.52	4.03
	BRISk	2.06	2.34
	ORB	1.23	2.44

other methods. Regarding memory, the output of SIFT and SURF contains less memory for the clean image. However, for the noisy image result, the image produced by BRISK occupies less memory compared to other methods.

Table 2 shows the simulation time and memory size of output, i.e. stitched image by different methods - SIFT, SURF, ORB, and BRISK for source-2 images. This table shows that the simulation time for ORB is 1.68 s for a clean image, and for a noisy image, the simulation time for BRISK is 1.98 s. Its values are lower than those of other methods. In the case of memory size, the output of ORB contains 1.9 MB memory size for a clean image. Its value is marginally equal with other methods, but in the case of noisy images, the image result of BRISK holds less memory size than other methods.

Table 2. Simulation time and memory for source-2 image

Image Type (Clean/noisy)	Techniques	Simulation Time (sec)	Memory (MB)
Clean Image	SIFT	3.09	1.99
	SURF	3.24	1.99
	BRISK	2.73	1.98
	ORB	1.68	1.96
Noisy Image (corrupted by pixel noise with 10 dB)	SIFT	4.35	4.12
	SURF	3.29	4.12
	BRISK	1.98	2.56
	ORB	2.11	4.09

The feature points (key points) extraction process is conducted on the source-2 clean images (unstitched clean images) using various algorithms such as BRISK, SIFT, SURF, and ORB. The individual output images of the key point extraction steps are depicted in Figs. 5, 7, 9, and 11 for BRISK, ORB, SIFT, and SURF algorithms. Examining these figures reveals that the ORB algorithm extracts fewer key points (feature points) than the other algorithms [19].

Figures 6, 8, 10, and 12 showcase the final output images, i.e., stitched images, for BRISK, ORB, SIFT, and SURF algorithms, respectively. These stitched images show that the output image provides a more comprehensive perspective of the scene, capturing every detail and offering the viewer an immersive experience.

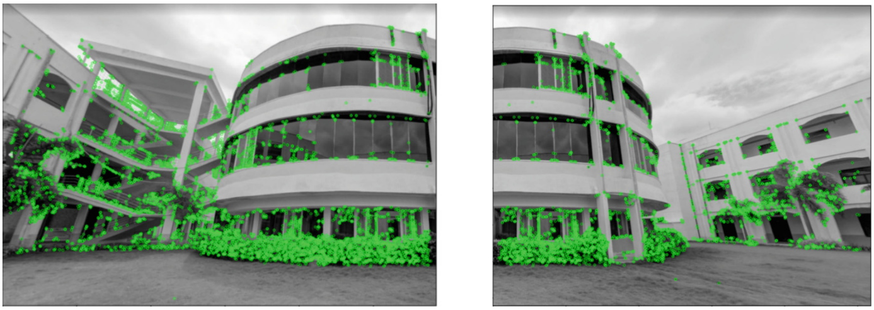

Fig. 5. Key points drawn on (a) Left-side source-2 image (b) Right-side source-2 image using BRISK

Fig. 6. Warp transformed stitched Image by BRISK.

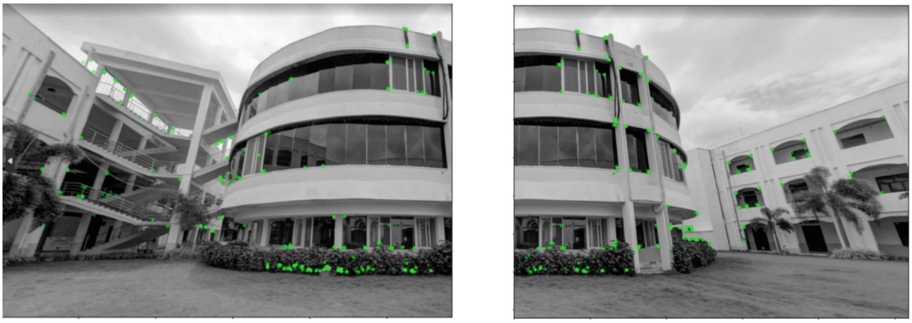

Fig. 7. Key points drawn on (a) Left-side source-2 image (b) Right-side source-2 image using ORB

Fig. 8. Warp transformed stitched Image by ORB.

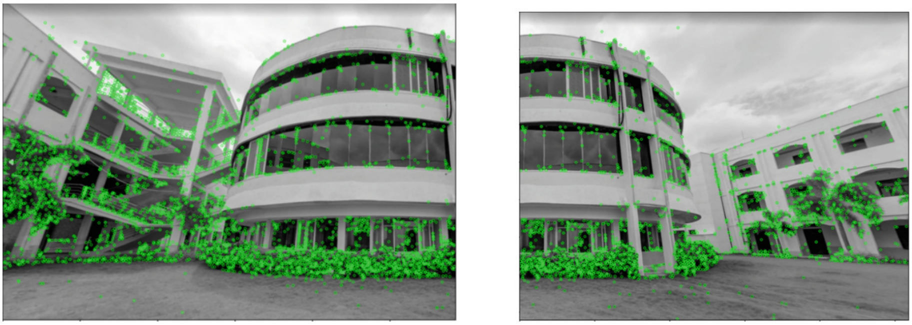

Fig. 9. Key points drawn on (a) Left-side source-2 image (b) Right-side source-2 image using SIFT

Fig. 10. Warp transformed stitched Image by SIFT.

In Figs. 5, 6, 7, 8, 9, 10, 11 and 12, using various algorithms, only clean images are utilized to generate the stitched image. In the subsequent experimentation, noisy images are introduced to assess the performance of algorithms – BRISK, ORB, SIFT, and SURF. For this, pixel noise is added to the stitched image (source-2 image) with a signal-to-noise ratio (SNR) of 10 dB. After introducing the noise, the noisy image is processed by different algorithms. The results, presented in terms of images, are shown in Figs. 13, 14, 15 and 16 for BRISK, ORB, SIFT, and SURF algorithms, respectively. From these figures, it is evident that the stitched image using BRISK does not produce a

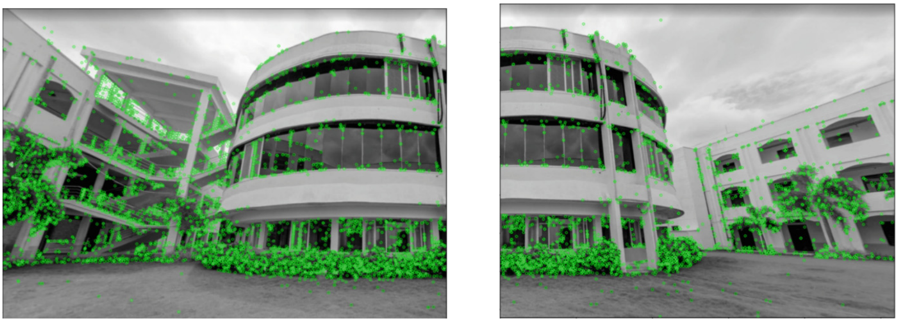

Fig. 11. Key points drawn on **(a)** Left-side source-2 image **(b)** Right-side source-2 image using SURF

Fig. 12. Warp transformed stitched Image by SURF.

correct panoramic image, indicating that under noisy conditions, the BRISK algorithm may not function effectively.

Fig. 13. Warp transformed stitched Image by SIFT.

Fig. 14. Warp transformed stitched Image by BRISK.

Fig. 15. Warp transformed stitched Image by ORB.

Fig. 16. Warp transformed stitched Image by SURF.

5 Conclusion

This paper comprehensively assesses key panoramic image stitching algorithms, offering valuable insights for researchers and practitioners in computer vision and related fields. The performance of prominent algorithms, including SIFT, SURF, ORB, and BRISK, was thoroughly evaluated. Through rigorous experimentation, we have highlighted their strengths and weaknesses in creating seamless wide-angle views in quiet and noisy environments. Memory size and time consumption for different algorithms—SIFT, SURF, ORB, and BRISK—were also calculated. The evaluation revealed that the chosen algorithms performed well for clean images, but the BRISK algorithm encountered challenges in the case of noisy images, particularly those affected by pixel noise. At the same time, the rest of the methods proved effective. Looking ahead, future research could explore the integration of deep learning techniques to enhance the performance and robustness of panoramic image stitching. Additionally, investigating novel approaches for multi-sensor fusion and addressing challenges related to dynamic scenes present promising avenues for advancement in this field.

References

1. Ravi, K.C., et.al.: Implementation of panoramic image stitching using Python. In: 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, pp.1–7. IEEE (2023)
2. Sakharkar, V.S., Gupta, S.R.: Image stitching techniques-an overview. Int. J. Comput. Sci. Appl. **6**(2), 324–330 (2013)
3. Reddy, G.K., et al.: Advanced object enhancement in an image. In: 2022 IEEE Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI), Gwalior, pp. 1–5. IEEE (2022)
4. Jian, W., Cui, Z., et al.: A comparative study of SIFT and its variants. Meas. Sci. Rev. **13**(3), 122–131 (2013)
5. Qu, Z., Wang, T.F., An, S.Q., Liu, L.: Image seamless stitching and straightening based on the image block. IET J. Mag. **12**(8), 1361–1369 (2018)
6. Javier, V.C., Marcelo, B.: Angular-based pre-processing for image de-noising. IEEE Signal Process. Lett. **25**(2), 219–223 (2018)
7. Brown, M., Lowe, D.G.: Automatic panoramic image stitching using invariant features. Int. J. Comput. Vis. **74**(1), 59–73 (2007)
8. Fu, M., et. al.: Image stitching techniques applied to plane or 3D models a review. IEEE Sen. J. **23**(8), 8060–8079 (2023)
9. Arulananth, T.S., et al.: Python based smart trial room. In: 2022 IEEE Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI), Gwalior, pp. 1–6. IEEE (2022)
10. Rublee, E., et al.: ORB: an efficient alternative to SIFT or SURF. In: 2011 International Conference on Computer Vision, Barcelona, Spain, pp. 2564–2571. IEEE (2011)
11. Chang, C.H., Sato, Y., Chuang, Y.-Y.: Shape preserving half-projective warps for image stitching. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA, pp. 3254–3261. IEEE (2014)
12. Zhang, J., Xiu, Y.: Image stitching based on human visual system and SIFT algorithm. Vis. Comput. **40**(1), 427–439 (2024)
13. Garcia-Garcia, A., et al.: A survey on deep learning techniques for image and video semantic segmentation. Appl. Soft Comput **70**, 41–65 (2018)
14. Lin, C.C., et al.: Adaptive as-natural-as-possible image stitching. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, USA, pp. 1155–1163 (2015)
15. Li, J., et al.: Parallax-tolerant image stitching based on robust elastic warping. IEEE Trans. Multimed. **20**(07) (2017)
16. Vivekanand, K.B.: Melody extraction in noisy music of different genre. In: 2017 International Conference on Recent Innovations in Signal Processing and Embedded Systems (RISE), Bhopal, pp. 629–632. IEEE (2017)

17. Prasad, S.V.S., et al.: An analysis of image quality enhancement techniques in an optical based satellite image. In: 2022 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC), Bhubaneswar, pp. 1–5. IEEE (2022)
18. Kumar, B., et. Al.: Comparative studies of single-channel speech enhancement techniques. IETE J. Res., 1–17 (2023)
19. Chia, T.L., Shin, Y.D., Huang, P.S.: Identification of panoramic photographic image composition using fuzzy rules. Sensors **24**(1195) (2024)