Robust Feature Extraction from Omnidirectional Outdoor Images for Computer Vision Applications

ASHWANI KUMAR AGGARWAL[1], AJAY PAL SINGH CHAUHAN[2]

[1]Electrical and Instrumentation Engineering Department
Sant Longowal Institute of Engineering and Technology
SLIET, Longowal, 148106, Punjab
INDIA

[2]Department of Electronics and Communication Engineering
Sant Longowal Institute of Engineering and Technology
SLIET, Longowal, 148106, Punjab
INDIA

Abstract: - Robust feature extraction from digital images is a challenging task for many computer vision applications. Several methods are available for extraction of features based on texture, shape, color, and geometry in the digital images captured using conventional camera. In the outdoor environment, the feature extraction becomes increasingly challenging because of dynamic scene, presence of outliers, occlusions, and changing illumination conditions. Omnidirectional cameras are becoming popular to capture images in outdoor environment as these gather scene information from a wide angle. This paper addresses the challenges of dynamic scenes, occlusions, outliers, and changing lighting conditions. The methodology used in this paper integrates SIFT, deep learning-based feature maps, accurate feature detection and descriptor matching under outdoor conditions. In this paper, robust feature extraction methods which consider the pixel formation in the omnidirectional images, are used to extract features from omnidirectional outdoor images for many computer vision applications. Such feature extraction methods can be useful in applications like intelligent transportation systems, mobile robots, and location-based services. The findings have significant implications for intelligent transport systems. This paper presents an approach towards enhancing robustness of image feature detection methods in dynamic environments.

Key-Words: - Omnidirectional image, feature descriptor, image matching, outliers, illumination variation, feature selection, noise removal.

Received: May 14, 2024. Revised: December 25, 2024. Accepted: January 28, 2025. Published: May 16, 2025.

1 Introduction

Omnidirectional images are widely used in many computer vision applications such as intelligent transportation systems, mobile robotics, and location-based services. In the outdoor environment, due to changing illumination conditions, presence of outliers, occlusions, and dynamic scene, the feature extraction from such images becomes a challenging task. A robust method for feature extraction helps in achieving the desired results in terms of correct image matches in the outdoor environment.

An omnidirectional image is obtained using a spherical camera of a pack of cameras where each individual camera covers a fixed field of view. When images from individual cameras are stitched together, an omnidirectional image with a wide field of view is obtained. The feature points and feature descriptors are obtained from omnidirectional images using many methods available in computer vision literature such as Scale-Invariant Feature Transform (SIFT), Speeded-Up Robust Features (SURF), Oriented FAST and Rotated BRIEF (ORB), and Binary Robust Independent Elementary Features (BRIEF), etc. The transform-based methods such as Fourier Transform (FT), Discrete Wavelet Transform (DWT), and Hough Transform etc. are also used. These methods for feature extraction are then used for various computer vision tasks such as object detection, scene segmentation, and navigation, etc. The challenges associated with omnidirectional images are distortion, varying illumination, and dynamic backgrounds, etc.

The contribution of the paper is analyzing robust feature extraction techniques for omnidirectional images.

2 Related Work

Several studies have investigated the use of feature extraction methods from omnidirectional images. A hierarchical framework for mobile robot localization using omnidirectional images and artificial intelligence techniques with various classifiers and global-appearance descriptors is proposed in [1]. The approach may limit adaptability to highly dynamic environments. A comprehensive review and taxonomy of deep learning methods for omnidirectional vision is provided in [2]. The works on omnidirectional images using deep learning methods is discussed with adequate details. The spherical camera model, planar representations, and deep learning challenges in omnidirectional visual computing are discussed in [3]. The paper focuses primarily on deep learning-based approaches, and not the traditional methods that may be relevant for certain applications. A multi-camera omnidirectional odometry system with feature prioritization and multi-view pose refinement is proposed in [4]. The approach increases system complexity, limiting real-time applications on resource-constrained platforms. Siamese Neural Networks with CNN-based descriptors for indoor robot localization in omnidirectional images are used in [5] achieving superior performance in varying lighting conditions. The changing lighting conditions affect the performance of the image features drastically and makes the computer vision tasks challenging. A review of deep learning methods for omnidirectional vision, including imaging principles, convolution techniques, and learning strategies is given in [6]. The convolution techniques used in various omnidirectional images are discussed for computer vision applications in the paper. A deep learning framework combining distortion-aware learning and bidirectional LSTM for structure recovery from omnidirectional images in VR and MR applications is proposed in [7]. The VR and MR applications are used in various domains of computer vision such as entertainment, medical training, and intelligent transportation systems. An intrinsic image decomposition method for omnidirectional images to improve the separation of reflectance and shading is discussed in [8]. The method is designed for low dynamic range images, which may limit its effectiveness when applied to high dynamic range omnidirectional images. A transformer-based multi-task learning framework that jointly infers depth, normals, semantic segmentation, reflectance, and shading from a single indoor panoramic image is proposed in [9]. The image is taken using an omnidirectional camera or by stitching several overlapping images captured using a conventional camera. A hierarchical extract-and-match transformer that improves efficiency, robustness, and precision in local feature matching is proposed in [10]. The robustness of feature extraction method is very important in computer vision applications in case of dynamic scenes and changing lighting conditions. The model's effectiveness in highly dynamic environments may be limited due to inherent challenges in feature matching. A deep learning framework for omnidirectional depth estimation using a multi-fisheye system, using multi-scale feature extraction, fusion cost volume, and cascaded cost regularization is used in [11]. The method's performance may be affected under extreme lighting conditions, where fisheye distortions and feature mismatches become more challenging. A feature extraction method for place recognition and point registration in autonomous driving is discussed in [12]. The place recognition in autonomous driving becomes challenging due to the presence of occlusions in the environment in which the vehicle is moving. A real-time ground target position estimation system combining fisheye camera and LiDAR data with Kalman fusion is used in [13]. The registration of camera images and LiDAR data becomes challenging due to change in viewpoints. The data collected from two different modalities are registered to estimate the target position. The system's accuracy may degrade in highly cluttered environments where occlusions and sensor noise affect target detection and fusion reliability. Bayesian optimization and data augmentation is used to achieve robust mobile robot localization from raw omnidirectional images without panoramic conversion in [14]. The approach may not work effectively where moving objects affect localization accuracy. A real-time, lightweight deep learning network for omnidirectional depth estimation using multi-scale feature extraction for diverse weather conditions is discussed in [15]. The method relies on synthetic data, which may not generalize to real-world scenarios with complex lighting and texture variations. A vision-based depth estimation system that converts omnidirectional depth estimation into a simpler binocular problem is discussed in [16]. A two-stage omnidirectional depth estimation framework using spherical feature learning is proposed in [17]. The conventional methods for feature extraction suffer from many drawbacks when applied on omnidirectional images. The depth estimation is done using feature extraction methods that work efficiently on omnidirectional images. An omnidirectional self-driving robot for indoor surveillance is discussed in [18]. The method features holonomic motion, ROS integration, and AI-driven monitoring for elderly fall detection and intrusion detection. A comparative analysis of keypoint matching algorithms for omnidirectional images and discussion on tangent plane projections to mitigate spherical distortions for improved pose estimation is presented in [19]. A framework using an LSTM-based classifier is used to predict and control robot movements in real-time [20]. Robot movement is affected by various factors including the path, the type of sensors used, and the adopted control strategy. A data-driven deep learning approach for jointly estimating depth and room structure from a single spherical panorama is discussed in [21]. This method may limit its applicability to irregular interior layouts. A benchmark dataset with 360° images and ground truth poses for visual localization is introduced in [22]. The proposed virtual camera approach may not eliminate the domain gap between different camera types, affecting localization accuracy in real-world applications.

3 Methodology

3.1 Omnidirectional Image Acquisition

A conventional camera has a limited field of view whereas an omnidirectional camera has a wider field of view. The omnidirectional image is obtained by stitching several images captured using conventional cameras which have sufficient overlap between them. There are omnidirectional cameras also available which directly generate panoramic images using multiple cameras mounted on a single assembly in a particular geometry.

3.2 Preprocessing Steps

The images captured using omnidirectional cameras are sometimes distorted due to several reasons. These images are given to image preprocessing tasks which rectify the distortions in the images. The distortions in the omnidirectional images may appear as pincushion distortion or barrel distortion among many others. The techniques sued to undistort such distortions from the omnidirectional images are based on geometry, warping, and deep learning-based methods.

3.3 Feature Extraction Techniques

Several methods are available in computer vision literature to extract features from digital images. Among various methods, the traditional methods for image feature extraction are SIFT, SURF, and ORB, etc. SIFT feature point and feature descriptor around the feature point assists in image matching, and localization-based services. The SIFT feature descriptor is a 128-dimentional vector. SURF feature extraction is preferred sometimes over SURF because it is faster than SIFT. SURF feature descriptor is a 64-dimensional vector.

$$D_{SIFT} = [h_1, h_2, \cdots, h_{128}] \in R^{128} \tag{1}$$

The SIFT descriptor is mathematically represented by using equation (1). ORB is a 32-byte descriptor. Several deep learning-based feature extraction methods are also used to extract features from omnidirectional images. CNNs, transformers, and self-supervised feature extraction methods are some of the methods using deep learning techniques. Hybrid approaches are useful in some cases where traditional feature extraction methods are combined with deep learning-based feature extraction methods.

$$D_{HYBRID} = \lambda\, D_{SIFT} + (1 - \lambda) D_{DEEP} \tag{2}$$

The hybrid feature descriptor which combines SIFT feature descriptor and a deep learning based feature descriptor can be represented mathematically using equation (2).

3.4 Feature Selection and Optimization

Sometimes all the features extracted by feature extraction methods are not useful for computer vision tasks. Techniques such as PCA, LDA, and attention mechanisms are used on the extracted features to refine these extracted features. PCA helps in reducing the dimension of the features and hence reducing the computation cost. LDA is used to reduce the dimensionality of the features while maximizing the separability between the classes. Attention network-based feature reduction methods reduce the dimensionality of the features by paying more attention to useful portions in the image instead of focusing on the entire image.

4 Experiments and Results

4.1 Dataset Description: Omnidirectional outdoor image dataset is used for evaluation. The dataset contains several omnidirectional images. Each image contains the outdoor environment scene. The images included in the dataset contain several types of image distortions. The images are histogram equalized before applying feature extraction methods.

4.2 Experimental Setup: The performance of the method is evaluated using performance evaluation metrics such as accuracy. The images are pre-proceeded before these are given to feature extraction methods. The extracted features are then given to feature reduction methods. Principal Component Analysis (PCA) is used to keep the most informative features eliminating the redundancy in the data.

4.3 Results and Analysis: The quantitative analysis of extracted features is done. Figure 1. shows the original image and the undistorted image.

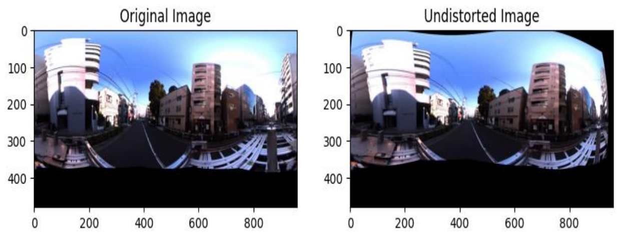

Fig. 1. Original image and undistorted image

The features are extracted from the undistorted images. Figure 2. Shows the extracted features from the omnidirectional image. The performance comparison is done in different environmental conditions such as lighting variations, and occlusions.

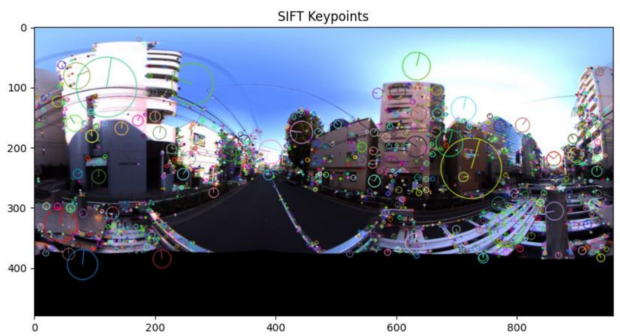

Fig. 2. Features extracted from omnidirectional image

The feature distribution analysis is done. The histogram of feature values is shown in Figure 3.

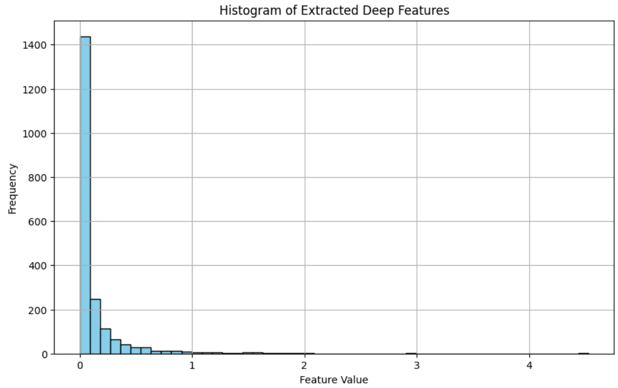

Fig. 3. Histogram of extracted deep features

The box plot of feature values is also obtained and is shown in Figure 4.

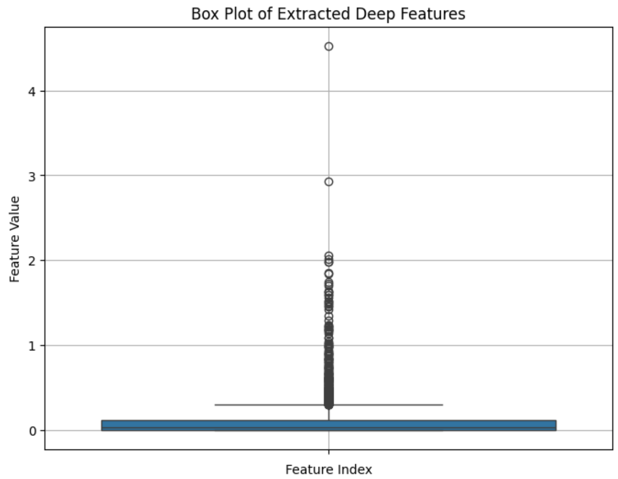

Fig. 4. Box plot of extracted features

5 Discussion

Among many feature extraction techniques, each method has its advantages and disadvantages over others. Some feature extraction methods have suitable for images where there are large occlusions while others perform better where the illumination variations are large. The various feature descriptors used around each feature point also have their own advantages and disadvantages. Some feature descriptors are high dimensional which make them computationally expensive to use while others are not suitable for large occlusions and illumination variations. In this paper, the image data collection is done using omnidirectional cameras under outdoor environments. The goal of the work is to evaluate the performance of image feature extraction methods applied on omnidirectional images under changing lighting conditions and dynamic environments. The image data is collected from different outdoor locations to capture diverse features including trees, vehicles, pedestrians, and building facades. The camera for capturing omnidirectional images is moved smoothly during different times of day in order to capture images under changing lighting conditions. The feature extraction is carried out on the captured omnidirectional images after histogram equalization. The robustness of feature extraction methods was done using matching accuracy. The feature extraction methods for omnidirectional images captured in outdoor environment are used in this work for computer vision applications such as image matching, localization, and intelligent transportation system.

6 Conclusion

The feature extraction methods used on omnidirectional images captured in outdoor environments are used. The distorted images are rectified using undistortion methods. The histogram of extracted deep features and box plot of extracted feature values are also plotted in this paper. This paper has discussed that conventional feature extraction methods are not robust enough for several computer vision applications. By considering the pixel geometry in the omnidirectional images, the proposed method significantly improves the robustness of feature extraction, and the reliability of the methods used for various computer vision applications. Omnidirectional cameras may be used instead of conventional cameras for intelligent transportation systems in urban environments.
.

Acknowledgement:
The author is indebted to his colleagues for proofreading this manuscript.

References:

[1] S. Cebollada, L. Payá, A. Peidró, W. Mayol, and O. Reinoso, ‘Environment modeling and localization from datasets of omnidirectional scenes using machine learning techniques’, Neural Computing and Applications, vol. 35, no. 22, pp. 16487–16508, 2023.

[2] H. Ai, Z. Cao, and L. Wang, ‘A Survey of Representation Learning, Optimization Strategies, and Applications for Omnidirectional Vision’, arXiv preprint arXiv:2502. 10444, 2025.

[3] T. L. T. da Silveira and C. R. Jung, ‘Omnidirectional visual computing: Foundations, challenges, and applications’, Computers & Graphics, vol. 113, pp. 89–101, 2023.

[4] Z. Javed and G.-W. Kim, ‘OmniVO: toward robust omni directional visual odometry with multicamera collaboration for challenging conditions’, IEEE Access, vol. 10, pp. 99861–99874, 2022.

[5] J. J. Cabrera, V. Román, A. Gil, O. Reinoso, and L. Payá, ‘An experimental evaluation of Siamese Neural Networks for robot localization using omnidirectional imaging in indoor environments’, Artificial Intelligence Review, vol. 57, no. 8, p. 198, 2024.

[6] H. Ai, Z. Cao, J. Zhu, H. Bai, Y. Chen, and L. Wang, ‘Deep learning for omnidirectional vision: A survey and new perspectives’, arXiv preprint arXiv:2205. 10468, 2022.

[7] M. Meng, Y. Zhou, D. Zuo, Z. Li, and Z. Zhou, ‘Structure recovery from single omnidirectional image with distortion-aware learning’, Journal of King Saud University-Computer and Information Sciences, vol. 36, no. 7, p. 102151, 2024.

[8] R.-K. Xu, L. Zhang, and F.-L. Zhang, ‘Intrinsic omnidirectional image decomposition with illumination pre-extraction’, IEEE Transactions on Visualization and Computer Graphics, 2024.

[9] U. Shah et al., ‘MultiPanoWise: holistic deep architecture for multi-task dense prediction from a single panoramic image’, in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 1311–1321.

[10] Q. Wang, J. Zhang, K. Yang, K. Peng, and R. Stiefelhagen, ‘Matchformer: Interleaving attention in transformers for feature matching’, in Proceedings of the Asian Conference on Computer Vision, 2022, pp. 2746–2762.

[11] X. Su, S. Liu, and R. Li, ‘Omnidirectional depth estimation with hierarchical deep network for multi-fisheye navigation systems’, IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 12, pp. 13756–13767, 2023.

[12] J.-U. Im, S.-W. Ki, and J.-H. Won, ‘Omni point: 3d lidar-based feature extraction method for place recognition and point registration’, IEEE

Transactions on Intelligent Vehicles, vol. 9, no. 8, pp. 5255–5271, 2024.

[13] Y. Ding, J. Che, Z. Zhou, and J. Bian, ‘Omni-OTPE: Omnidirectional Optimal Real-Time Ground Target Position Estimation System for Moving Lightweight Unmanned Aerial Vehicle’, Sensors, vol. 24, no. 5, p. 1709, 2024.

[14] J. J. Cabrera, S. Cebollada, M. Flores, Ó. Reinoso, and L. Payá, ‘Training, optimization and validation of a CNN for room retrieval and description of omnidirectional images’, SN Computer Science, vol. 3, no. 4, p. 271, 2022.

[15] S. Liu, R. Li, and F. Xue, ‘Real-time omnidirectional depth perception based on multi-view wide-angle vision system’, in 2023 35th Chinese Control and Decision Conference (CCDC), 2023, pp. 882–887.

[16] S. Xie, D. Wang, and Y.-H. Liu, ‘Omnividar: Omnidirectional depth estimation from multi-fisheye images’, in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 21529–21538.

[17] M. Li, X. Jin, X. Hu, J. Dai, S. Du, and Y. Li, ‘MODE: Multi-view omnidirectional depth estimation with 360$\circ$ cameras’, in European Conference on Computer Vision, 2022, pp. 197–213.

[18] D. Dhanush, K. R. Raghul, S. R. Kr, P. G. Mukund, and M. Ganesan, ‘An omni-directional self-driving robot for indoor surveillance’, in 2024 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IConSCEPT), 2024, pp. 1–6.

[19] J. Murrugarra-Llerena, T. L. T. Da Silveira, and C. R. Jung, ‘Pose estimation for two-view panoramas based on keypoint matching: A comparative study and critical analysis’, in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 5202–5211.

[20] H. B. Mahajan et al., ‘Automatic robot Manoeuvres detection using computer vision and deep learning techniques: a perspective of internet of robotics things (IoRT)’, Multimedia Tools and Applications, vol. 82, no. 15, pp. 23251–23276, 2023.

[21] G. Pintore, F. Bettio, M. Agus, and E. Gobbetti, ‘Deep scene synthesis of Atlanta-world interiors from a single omnidirectional image’, IEEE Transactions on Visualization and Computer Graphics, vol. 29, no. 11, pp. 4708–4718, 2023.

[22] H. Huang, C. Liu, Y. Zhu, H. Cheng, T. Braud, and S.-K. Yeung, ‘360loc: A dataset and benchmark for omnidirectional visual localization with cross-device queries’, in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 22314–22324.