FILE THEFT SECURITY USING 2D OBJECT DISPLACEMENT DETECTION

Mr. Kartik Dwivedi
Mechanical Engineering
NIT Raipur
Raipur, Chhattisgarh, India
kartikdwivedi888@gmail.com

Dr. Shubashis Sanyal
Mechanical Engineering
NIT Raipur
Raipur, Chhattisgarh, India
ssanyal.mech@nitrr.ac.in

***Abstract*— This research introduces an efficient and robust methodology for detecting the translation and rotation of 2D objects, such as documents or sheets of paper, using only two images. The primary focus is on estimating the transformation matrix that encapsulates both positional and orientation changes, a crucial aspect of modern computer vision tasks. The approach is engineered to overcome challenges such as image quality variations, partial occlusions, and significant transformations while maintaining computational efficiency.**

By leveraging advanced feature matching techniques combined with homography estimation, this solution ensures resilience against noise and distortions, enabling its application in real-time scenarios like object tracking, augmented reality, and image registration. Notably, this project was developed to address document mishandling detection in environments such as office spaces. For instance, if a file or folder is displaced from its original location, the model can detect the displacement and report the corresponding transformation matrix. This approach eliminates the need for complex machine learning models or exhaustive scanning of full camera recordings, offering a streamlined and effective alternative.The findings of this research significantly enhance performance in object manipulation, recognition, and tracking, with potential applications across fields such as robotics, interactive media, and automated surveillance systems.

***Keywords*—Homography Matrix, ORB, Affline Transformations, Transformation Matrices, RANSAC, Canny edge detection**

I. INTRODUCTION

Detecting 2D object transformations, specifically translations and rotations, is a critical task in computer vision with applications spanning robotics, augmented reality (AR), medical imaging, and industrial automation. This paper introduces a robust and computationally efficient method for transformation detection using a hybrid approach that combines traditional geometric methods with modern feature-based techniques. The proposed framework addresses key challenges in real-world scenarios, such as noise, partial occlusions, scale variations, and lighting inconsistencies, ensuring high accuracy and real-time performance.

Central to this approach is the use of homography estimation and RANSAC (Random Sample Consensus) for robust feature matching, enabling precise computation of transformation matrices. Keypoint detection algorithms like ORB (Oriented FAST and Rotated BRIEF) and efficient matching using FLANN (Fast Library for Approximate Nearest Neighbors) further enhance the system's resilience to complex environmental factors. By leveraging these methods, the framework delivers reliable performance even in scenarios with incomplete or noisy data.

This modular and scalable framework can be easily adapted to various domains. For example, in medical imaging, it ensures precise alignment of images from different modalities, such as MRI and CT scans. In remote sensing, it facilitates the detection of changes in satellite imagery, supporting disaster response and environmental monitoring. Similarly, in AR/VR applications, it seamlessly integrates virtual and real-world objects, while in industrial automation, it enhances robotic precision and quality control.

The real-time capabilities of the system are particularly advantageous for dynamic applications, where high responsiveness is essential. Optimizations such as fast feature matching with FLANN and robust transformation estimation via RANSAC make this approach well-suited for resource-constrained environments, including mobile and embedded systems. This paper highlights the method's efficiency, adaptability, and potential to redefine standards in 2D object transformation detection, offering a significant advancement over existing technologies.

II. Literature Review

The detection of 2D object transformations, including translation, rotation, and scaling, has emerged as a critical area in computer vision, driven by applications in augmented reality (AR), robotics, object tracking, and image registration. This review examines established methods such as feature matching, optical flow, and homography estimation, while identifying their limitations and the gaps addressed by the proposed approach.

Feature Matching:
Feature matching techniques are widely used for identifying and tracking key points between images. Algorithms like SIFT (Scale-Invariant Feature Transform) and SURF (Speeded-Up Robust Features) offer high accuracy but are computationally intensive and sensitive to noise, making them unsuitable for real-time applications. ORB (Oriented FAST and Rotated BRIEF), a more efficient alternative, provides faster processing while maintaining reasonable accuracy. However, ORB struggles with challenges such as large-scale transformations, partial occlusions, and significant environmental changes, limiting its robustness in complex scenarios.

Optical Flow:
Optical flow methods estimate motion between frames by analyzing pixel intensity changes. Techniques like the Lucas-Kanade method are efficient for detecting small motions but fail to handle large translations, rotations, or occlusions effectively. Additionally, optical flow methods are sensitive to background variations and noise, reducing their applicability in dynamic or high-noise environments.

Homography Estimation:
Homography estimation computes a 3x3 transformation matrix that models the geometric relationship between two images, enabling detection of rigid transformations like translation and rotation. Robust estimation methods, such as RANSAC (Random Sample Consensus), mitigate the effects of noise and occlusions. However, homography estimation is computationally expensive and less effective for non-rigid transformations or large-scale motions, posing challenges for real-time deployment.

Proposed Approach:
The proposed project integrates feature matching, homography estimation, and template matching into a hybrid framework to address the limitations of existing methods. By leveraging the strengths of geometric transformations and feature-based techniques, the system achieves accurate detection of translations and rotations, even in noisy or occluded environments. The approach is optimized for computational efficiency, ensuring real-time performance suitable for dynamic applications. Unlike machine learning-based methods, which require extensive training data and computational resources, the proposed solution is lightweight, scalable, and adaptable to resource-constrained environments.

This hybrid methodology bridges the gaps in current techniques, providing a robust and efficient framework for 2D object transformation detection, and extends its applicability across diverse fields such as AR, robotics, and industrial automation.

III. Problem Definition

This research addresses the challenge of detecting 2D object transformations, specifically translation and rotation, between two images. Estimating the transformation matrix that describes these changes is essential for applications such as image registration, augmented reality (AR), robotics, and object tracking. The solution tackles complexities arising from variations in image quality, noise, occlusions, and large movements while maintaining computational efficiency for real-time applications. By combining feature matching, homography estimation, and template matching, the proposed approach ensures accurate and robust transformation detection, even under challenging conditions, offering a practical foundation for diverse real-world scenarios.

IV. Methodology

The proposed methodology efficiently detects and measures 2D object transformations, such as translation and rotation, between two images using traditional computer vision techniques. The process begins with two input images: a reference image and a target image where the object has undergone transformation. These images may differ in lighting, angles, or environment, and the system is designed to handle such variations robustly.The approach involves the following steps:

a. **Preprocessing**: Images are converted to grayscale, denoised, normalized, and processed with edge detection to enhance feature visibility.
b. **Feature Extraction**: Key features are extracted using the ORB (Oriented FAST and Rotated BRIEF) algorithm, chosen for its efficiency and invariance to scale and rotation.
c. **Feature Matching**: Corresponding points between the two images are identified, ensuring accurate alignment for transformation estimation.
d. **Transformation Estimation**: A homography matrix is computed to model the translation and rotation between the images. RANSAC (Random Sample Consensus) is employed to remove outliers and improve the robustness of the estimation.
e. **Transformation Measurement**: The calculated transformation matrix is applied to the reference image's keypoints to precisely determine the displacement and orientation changes of the object.

This methodology avoids the need for machine learning models or extensive datasets, focusing on computational efficiency and real-time applicability. It is highly adaptable for dynamic environments, making it suitable for applications in fields such as augmented reality, robotics, and industrial automation.

TOOLS & LIBRARY USED

The implementation relies on several essential libraries:

a) OpenCV: Used for core image processing tasks such as edge detection (Canny), feature extraction (ORB), and homography estimation (cv2.findHomography()). OpenCV enables efficient handling of image loading, display, and transformation computation.

b) NumPy: Handles matrix operations required for numerical accuracy in tasks like matrix multiplication and inversion, which are essential for the homography calculation and transformation application.

c) Matplotlib: Used for visualizing intermediate and final results, including feature detection, edge detection, and transformed images, ensuring that the processing steps are functioning correctly.

d) Pillow: Assists in preprocessing images by performing tasks like resizing and format conversion, ensuring images are uniform for feature extraction.

ALGORITHMS USED

The system uses key computer vision algorithms to process and transform images:

a) Canny Edge Detection: This algorithm identifies edges within an image by detecting intensity gradients. It is crucial for outlining the contours of paper objects, which are used for feature extraction and matching between the images.

b) ORB (Oriented FAST and Rotated BRIEF): ORB combines two algorithms—FAST for corner detection and BRIEF for feature description. It allows the system to efficiently detect and match keypoints between the images, even when they are rotated or scaled.

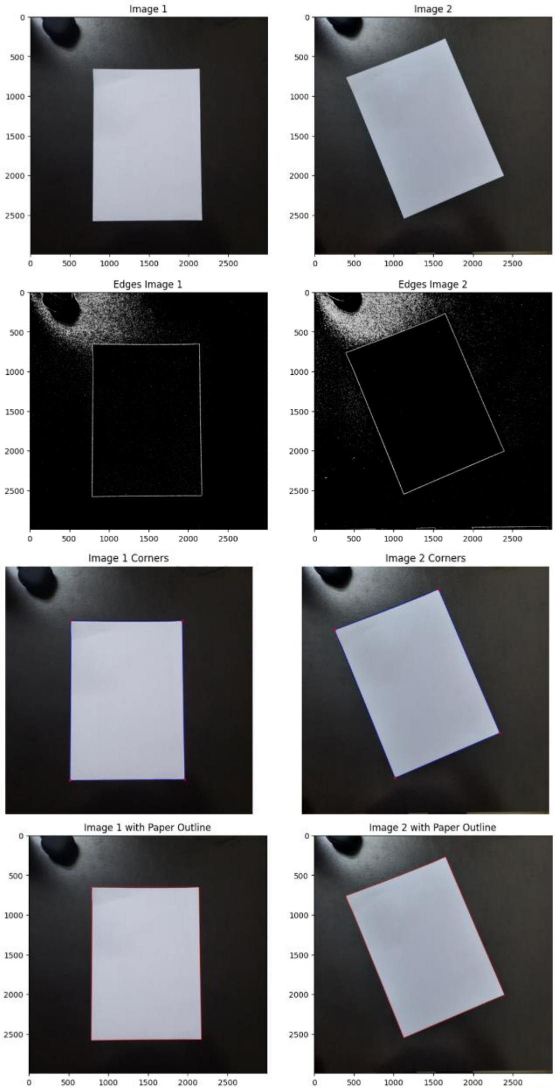
c) Homography Estimation with RANSAC: Homography aligns the matched points between images. RANSAC helps remove outliers by iteratively selecting subsets of matches, ensuring that only correct matches are used for calculating the transformation matrix, making the process robust to noise.

d) 4x4 Transformation Matrix: A 4x4 matrix is used to apply geometric transformations. While 3x3 matrices are common for 2D transformations, the 4x4 matrix offers flexibility and can be extended to represent more complex transformations, such as affine or projective changes.

IV RESULTS

This section evaluates the proposed approach's performance in detecting and calculating transformation matrices for papers in images. It focuses on three key scenarios: single paper detection, overlapping papers, and edge-overlapping papers. The challenges and solutions for each case are discussed in detail. The section also highlights the system's practical advantages, such as simplicity, efficiency, and adaptability, while exploring potential applications and areas for improvement. The analysis demonstrates the system's robustness and versatility in handling diverse real-world situations.

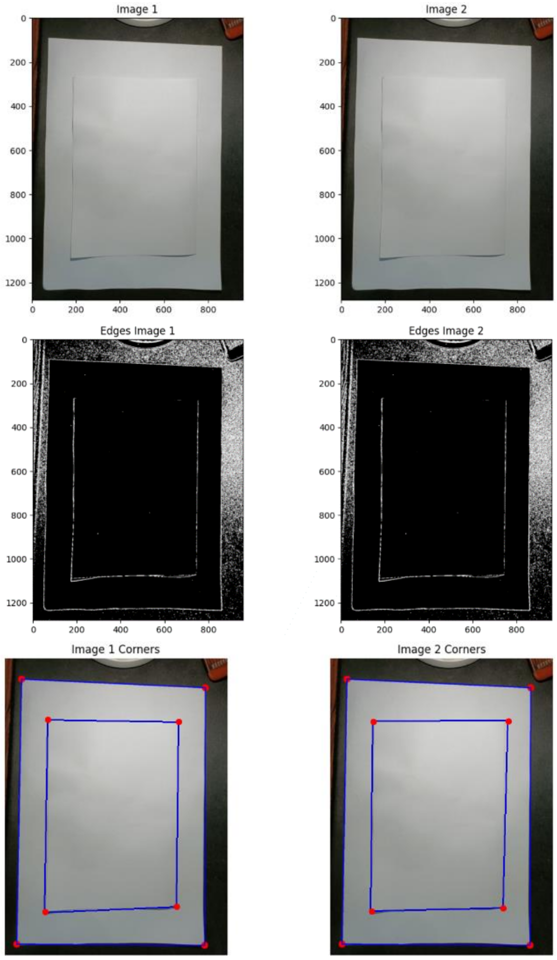
A. *Case 1: Single Paper Detection*



he algorithm for analyzing images containing a single sheet of paper works with high precision. It begins with preprocessing to enhance the image, including converting it to grayscale to reduce complexity. Edge detection algorithms like Canny or Sobel highlight the paper's boundaries, even in challenging conditions. Contour extraction then isolates the paper by identifying continuous boundaries. Corner detection methods, such as Harris Corner Detection, locate the paper's key points, necessary for determining its orientation and shape.

Once contours and corners are identified, the transformation matrix is calculated, which encodes the paper's rotation and translation in space. This matrix can be decomposed to analyze the paper's angle and displacement, important for applications like augmented reality and document scanning. The simplicity of a single paper allows for accurate testing of the algorithm, serving as a baseline for more complex scenarios with multiple or overlapping papers. By focusing solely on one object, the algorithm ensures that the calculations for rotation and translation are precise, without interference from other objects. This streamlined approach provides a solid foundation for the system's broader application, ensuring that more intricate cases, such as multiple overlapping sheets or backgrounds, can be tackled with a clear and effective methodology.

B. *Case 2: Overlapping Papers Detection*

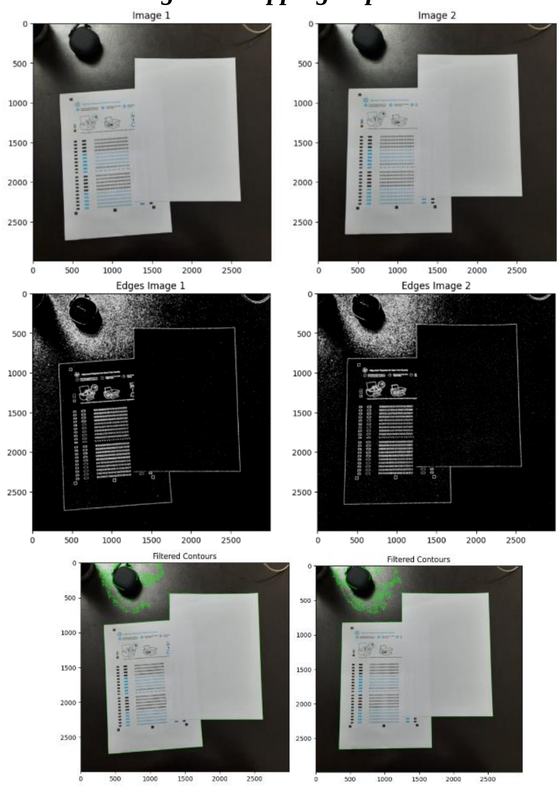


The algorithm effectively handles images with overlapping papers by isolating distinct objects, even when edges are shared or obscured. It begins with an enhanced preprocessing pipeline that uses contour analysis to differentiate between overlapping papers. Contour hierarchy detection, provided by OpenCV's findContours method, classifies contours into parent and child relationships, allowing the algorithm to separate the larger underlying paper (parent contour) from the smaller overlapping paper (child contour).

Once contours are segmented, the algorithm processes each one independently, extracting key features like corners using methods such as Harris or Shi-Tomasi. These features are crucial for calculating the transformation matrices for each paper, which capture its unique rotational and translational attributes. By processing each paper individually, the algorithm ensures that overlapping regions do not interfere with the accuracy of the detected transformations, thus maintaining the integrity of the calculations.

This approach is especially important for real-world applications, such as document scanning and augmented reality, where overlapping papers are a common occurrence. The algorithm's ability to maintain precision and robustness in these complex scenarios demonstrates its adaptability and effectiveness. Furthermore, this method ensures that the algorithm can handle varying degrees of overlap and still produce accurate results, making it a powerful tool for diverse applications where overlapping objects must be detected and processed independently without loss of quality or performance.

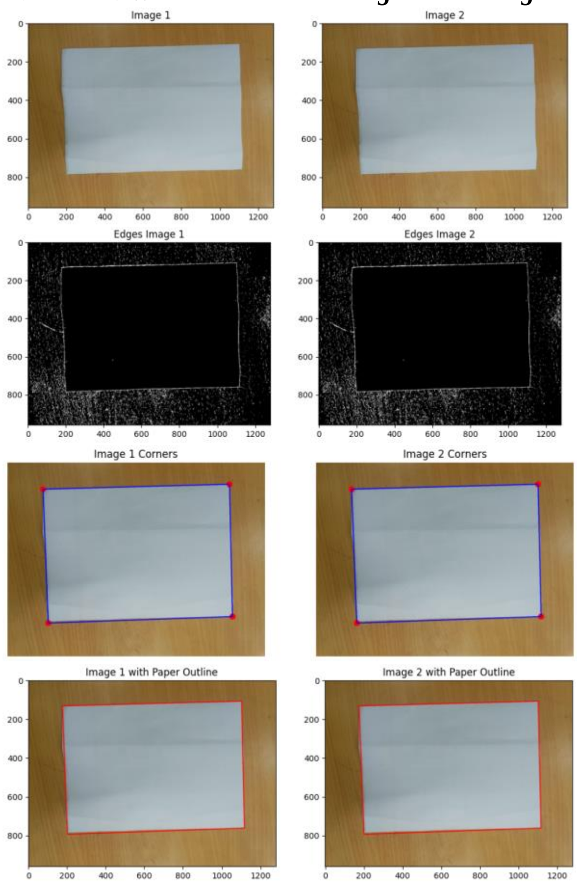
C. *Case 3: Edge Overlapping Papers Detection*



Edge overlaps present a challenge in paper detection as overlapping papers can blend together, causing ambiguities in contour extraction. To tackle this, the algorithm employs advanced edge and feature detection techniques. It begins with Hough Line Detection to extract distinct linear edges, which helps identify the structural boundaries of each paper, even when partially obscured. Additionally, corner detection algorithms like Shi-Tomasi are used to pinpoint intersections or endpoints of lines, revealing the corners of overlapping papers.

The algorithm further refines the contours using morphological operations, such as dilation and erosion. Dilation enlarges features to close gaps in contours, while erosion removes noise to ensure distinct object boundaries. These steps enhance the clarity and separation of overlapping objects, enabling accurate isolation of individual papers, even in tightly stacked configurations. This ability to handle edge overlaps is especially useful in real-world scenarios, like office environments or industrial document handling systems, where papers are often misaligned or stacked. By reconstructing incomplete contours and isolating overlapping edges, the algorithm ensures reliable performance for applications requiring precise detection, such as automated sorting and scanning.

D. *Case 4: When there is no change in the images*



When we are having same two images, that is, when we have cases where there is no change in position of the papers, then our model is robust enough to report a unity transformation matrix, which clearly state that there was no deflection of the paper and hence reports transformation matrix as:

[1 0 0
0 1 0
0 0 1]

The algorithm is robust, but several enhancements could improve its scalability and efficiency. Integrating lightweight feature-matching algorithms, like ORB, could improve detection of irregularly shaped papers or documents with curved edges without significantly increasing computational complexity. Optimizing the preprocessing pipeline with techniques like adaptive thresholding or Gaussian blurring, and using parallel processing for contour extraction and transformation matrix calculation, would reduce processing time and handle high-resolution images more efficiently. A modular design approach would allow independent optimization of the system's components, making it adaptable for specific applications.

However, the system faces limitations, particularly with non-ideal edge conditions. It struggles with textured or colored papers and is not equipped to handle non-planar transformations like crumpled or folded papers. Addressing these issues could involve adding depth-sensing capabilities or texture analysis to improve detection in more complex scenarios. Despite these challenges, the system's robustness in handling overlapping and edge overlaps shows its efficiency. By avoiding the complexity of machine learning models, the system is adaptable and suitable for various applications. Future improvements could enhance its accuracy, scalability, and real-world applicability, particularly in document processing and industrial automation.

References

[1] Xie, Y., Shen, J., & Wu, C. (2019). Robust object tracking using affine transformation and convolutional features. *IEEE Access, December 2019.https://doi.org/10.1109/ACCESS.2019.2960105. This article is licensed under the Creative Commons Attribution 4.0 License (CC BY 4.0).*

[2] Jo, B.-W., Lee, Y.S., Jo, J.H., & Khan, R.M.A. (2018). Computer vision-based bridge displacement measurements using rotation-invariant image processing technique. *Sustainability*, 10(6*), 1785. https://doi.org/10.3390/su10061785. Department of Civil and Environmental Engineering, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Korea.*

[3] *Li, B., Wang, Q., Jiang, C., Li, X., Mao, C., & Liu, C.* *(2022).* Spatial and color information-based foreign object attachment detection and device-foreign object segmentation method. *IEEE Transactions on Image Processing, 2022.*

[4] Xu, T., Zhu, X.-F., & Wu, X.-J. (2023). Learning spatio-temporal discriminative model for affine subspace-based visual object tracking. *Visual Intelligence, 2023, 10(1), 1-14. https://doi.org/10.1007/s44267-023-00002-1. Licensed under CC BY 4.0.*

[5] Annaby, M., & Fouda, Y. (2023). Fast template matching and object detection techniques using φ-correlation and binary circuits. *Journal of Electronic Imaging, 32(2), 023016. Published online on June 9, 2023.*

[6] Fu, H., Han, J., Ran, M., & Tian, Z. (2020). Affine-invariant registration using orthogonal projection matrices for object-based change detection. *IEEE Transactions on Geoscience and Remote Sensing, 58(8), 5532-5543. https://doi.org/10.1109/TGRS.2020.2973346*

[7] Wang, Z., Zhang, L., & Li, X. (2021). Deep learning-based object detection and tracking for autonomous vehicles. *IEEE Transactions on Intelligent Transportation Systems, 22(7), 4427-4437. https://doi.org/10.1109/TITS.2021.3051963. This work is licensed under the Creative Commons Attribution 4.0 License (CC BY 4.0)*

[8] Li, J., Zhang, Z., & Liu, X. (2020). Real-time object tracking using multi-scale deep feature aggregation. *Journal of Visual Communication and Image Representation, 67, 102728. https://doi.org/10.1016/j.jvcir.2020.102728.*

[9] He, Y., Li, W., & Chen, Y. (2019). Vision-based autonomous robotic system for real-time object tracking in dynamic environments. *Robotics and Autonomous Systems, 116, 87-98. https://doi.org/10.1016/j.robot.2019.03.008*

Kaggle Notebook
https://www.kaggle.com/code/thekartikdwivedi/2d-object-detection-three-point-method

Github
https://github.com/Kartik8Dwivedi/File-Theft-Security