Proposal and Verification of AI-Based Automatic Geometric Correction Technology for Satellite Images Using Open Access Basemaps

Han-Gyeol Kim, Jaewon Hur

Data Science Department, TelePIX, 07330 Seoul, Republic of Korea - (khanai, jwhur)@telepix.net

Keywords: Geometric Correction, Rational Polynomial Coefficient, AI Matching, BlueBON, Sentinel-2, SRTM.

Abstract

Geometric correction of satellite images is an essential pre-processing step for accurate geospatial analysis, but non-experts often face practical limitations because detailed sensor models and Ground Control Point data are not readily accessible. Traditional methods rely on physical sensor models or the Rational Function Model (RFM) using vendor-provided Rational Polynomial Coefficients (RPC). However, this information is often unavailable or lacks sufficient accuracy. This paper proposes a two-stage framework that utilizes AI matching technologies and open access data to automatically correct satellite images lacking georeferencing information. In Stage 1, a coarse Affine correction is executed using SuperPoint and LightGlue with an open basemap (Sentinel-2). In Stage 2, precise corresponding points are extracted through patch-based hierarchical LoFTR matching, and 3D GCPs are generated utilizing the SRTM. Subsequently, sensor-independent RPC are robustly estimated through the rpcfit library, and the final geometrically corrected image is generated through resampling. This framework was verified by applying it to 4.8m resolution BlueBON satellite images that lack georeferencing information. In seven experimental regions with diverse geographical characteristics, an average Root Mean Square Error (RMSE) of 8.050m (1.68 pixels based on BlueBON resolution) referenced to the Sentinel-2 basemap, and an average of 9.02m (1.88 pixels) referenced to Google Maps, was achieved. This result demonstrates that it is possible to precisely correct 4.8m medium-resolution images using a 10m open basemap, providing a practical, accessible, and automated geometric correction solution for general users.

1. Introduction

1.1 Background

Geometric correction of satellite images is the process of calculating the ground coordinates corresponding to each pixel of the satellite image and correcting the image, utilizing the sensor model and Ground Control Point (GCP) information. This is the most fundamental pre-processing technology for subsequent analyzes using satellite images, such as map production, change detection, and object recognition (Paul and Pati, 2021). However, geometric correction is often challenging for non-expert users because the required sensor models and GCP databases are not publicly available in most cases. Furthermore, even when RPC are provided for commercial satellite images, their accuracy may be insufficient, or they may not be provided at all for reasons such as security, requiring the user to perform geometric correction themselves.

1.2 Traditional Approaches and Limitations

Traditional geometric correction techniques use methods that either construct the geometric relationship from the satellite sensor to the ground coordinate system as a physical equation, or construct an approximate polynomial, such as the RFM, and perform correction by finding correction coefficients based on GCPs derived from landmarks or survey points. However, physical sensor models are treated as confidential for most commercial satellites. Therefore, the sensor-independent RFM method, which maps 3D ground coordinates X, Y, Z to 2D image coordinates Column, Row, is widely adopted. Correction using GCP is essential to improve the accuracy of RFM, and to automate this, GCP Chips (image patches around reference points) are often built and used. However, these GCP Chips are mostly not disclosed externally for reasons such as national security. Moreover, GCP Chip-based systems have fundamental operational limitations. If clouds exist in the image, preventing the use of a specific GCP Chip, or if the reference point chips are not updated and the terrain characteristics of the area change due to new city development or agricultural land changes, the number or quality of valid GCP Chips available for Rational Function Model (RFM) correction becomes insufficient, which can lead to geometric correction failure.

1.3 SOTA in Operational Correction and AI Introduction

To overcome these limitations, efforts have been underway to build large-scale, consistent reference datasets. Typically, the European Space Agency (ESA) builds and utilizes a database of precise GCPs worldwide, known as the Global Reference Image (GRI), for the automatic geometric correction of Sentinel-2 images. The GRI provides a consistent absolute positioning accuracy of approximately 9.5 m CE95 (Gaudel et al., 2017). The importance of the GRI is well demonstrated by the case where the USGS performed a large-scale bundle adjustment to align Landsat's ground reference points with the Sentinel-2 GRI, in order to ensure the geometric consistency of Landsat Collection-2 data and facilitate fusion analysis with Sentinel-2. Recently, with the development of AI technology in the computer vision field, AI-based matching techniques are also being actively applied to satellite image processing. In particular, matching techniques fine-tuned for the characteristics of satellite images (e.g., large viewpoint differences, illumination changes), such as SatLoFTR and SatMatchFormer, are being reported, overcoming the limitations of traditional feature-based matching methods like SIFT or SURF (Deshmukh and Kak, 2025, Luo et al., 2024).

1.4 Contributions

To address the lack of accessible RPCs and proprietary GCP chips, this study proposes an automated two-stage AI-based framework. By leveraging open-access Sentinel-2 and SRTM data, we provide a reproducible solution that eliminates reliance on restricted sensor models or static reference databases.

- **Accessibility**: We propose a geometric correction technology that anyone can access and reproduce by utilizing open data such as Sentinel-2 and SRTM, and open-source libraries like rpcfit and rpcm. Furthermore, by using the latest Sentinel-2 images as a basemap, geometric correction is performed between images with minimized changes in terrain features
- **Performance and Automation**: Through a two-stage AI matching strategy (SuperPoint+LightGlue, followed by LoFTR), a fully automated pipeline is implemented that demonstrates high precision and speed, even for raw satellite images for which RPC are not provided or initial geometric correction has not been performed
- **Novelty**: We experimentally demonstrate that despite correcting a 4.8m resolution image with a 10m resolution low-resolution basemap, a high geometric accuracy of within 2 pixels (1.68 pixels) can be achieved by leveraging the contextual matching characteristics of context-based matching (LoFTR). This shows the potential to solve the fundamental problems of matching with traditional GCP Chips, such as the presence of clouds and an insufficient number of points, through dense matching with the basemap

Furthermore, considering that most CubeSats launched recently share similar constraints regarding metadata and sensor model availability, the proposed AI-based automatic correction technology provides a highly versatile solution that can be extended to various small satellite constellations.

2. Related work

The automatic geometric correction framework proposed in this study is based on the following technologies.

2.1 Geometric Sensor Models: RFM/RPC

RFM is an approximate model expressed as a ratio of third-order polynomials that maps 3D ground coordinates (X, Y, Z) to 2D image coordinates (Column, Row). RPC refers to the coefficients of this polynomial and is a sensor-independent model that can be estimated if only 3D-2D corresponding points (GCPs) are available, even without knowing the physical characteristics of the sensor. This study aims to automatically estimate these RPC for images lacking georeferencing information.

2.2 Operational Geometric Correction Systems

Current operational geometric correction systems can be broadly divided into methods aiming for a global standard and high-precision methods driven by national needs.

- **(Global) Sentinel-2 and Landsat**: ESA built and utilized the GRI database to consistently maintain the global absolute geometric accuracy of Sentinel-2 (Gaudel et al., 2017). USGS went a step further, dramatically improving the interoperability between two different sensors by aligning the geometric reference of Landsat Collection-2 data with the Sentinel-2 GRI. This shows the importance of data fusion analysis using global standard data (Rengarajan et al., 2020).
- **(National)KOMPSAT and CAS500**: South Korea’s CAS500-1 or KOMPSAT-3/3A satellites utilize their own GCP Chips, built based on domestic survey points and high-resolution aerial photographs within the Korean peninsula, for precise geometric correction (Shin et al., 2018). While this approach ensures very high accuracy in specific regions, it entails operational limitations as mentioned in Section 1.2, such as the non-disclosure of GCP Chips, update cycles, and cloud influence

This study follows the global standard principles of the Sentinel-2/GRI framework, but applies them as a dynamic, AI-based matching methodology rather than relying on fixed GCP-chip databases such as KOMPSAT/CAS500.

2.3 AI-based Matching Techniques

Traditional feature-based matching methods like SIFT, SURF, etc., suffer from a sharp decline in matching performance when there are large differences in resolution, illumination, viewing angles, or sensors between images.

- **Feature analysis based matching (SuperPoint+LightGlue)**: SuperPoint is a keypoint detector trained in a self-supervised learning using synthetic data, and it extracts robust feature points even in real images. LightGlue is an attention-mechanism-based matcher that efficiently matches the feature points extracted by SuperPoint or other detectors. In a comparative study on high-resolution satellite stereo images, the combination of SuperPoint and LightGlue was reported to show the best balance between robustness, accuracy, and efficiency, and notably, to provide a uniform distribution of matching points across the entire image (Luo et al., 2024). This uniformity is ideal for Stage 1 of our study, which needs to estimate the large geometric transformation of the entire image.
- **Context analysis based matching (LoFTR)**: LoFTR is a detector-free context-based matching model (Sun et al., 2021), which excels in identifying correspondences across images with significant resolution disparities. Unlike traditional feature-based matchers that often fail to detect repeatable keypoints between 4.8m and 10m imagery, LoFTR’s dense context analysis enables it to infer semantic relationships from texture patterns even in areas with repetitive or sparse features. Among recent transformer-based models, LoFTR was prioritized for its proven robustness and consistent stability across diverse geographic terrains. This reliability ensures a dense and uniform distribution of GCPs, providing the essential foundation for stable downstream RPC estimation.

The strategic integration of SuperPoint and LoFTR in this framework is specifically designed to address the unique challenges of satellite imagery lacking RPCs. While Stage 1 utilizes SuperPoint to handle large geometric displacements, Stage 2 leverages the detector-free, context-based matching of LoFTR to bridge the resolution gap between 4.8m and 10m imagery. This hierarchical approach is not a mere combination of existing tools but a tailored pipeline optimized for cross-resolution geometric correction.

2.4 Open Source RPC Implementation

This framework is based on open-source libraries to ensure reproducibility.

- **rpcfit**: The open-source library rpcfit is used to estimate RPC from GCPs . rpcfit is a Python library for robustly fitting the coefficients of an RPC model from a set of 3D-2D corresponding points (GCPs). It is proven effective for SAR and Pushbroom sensor modeling (Akiki et al., 2021).
- **rpcm**: The rpcm library is utilized for the image projection/back-projection and final resampling operations using the estimated RPC (De Franchis et al., 2021). The combination of these two libraries enhances modularity by separating model estimation and application.

3. Proposed Method

The workflow of the Two-Stage automatic geometric correction framework proposed in this study is as follows Figure 1. The input data are the BlueBON image, which has no georeferencing information and only provides the center coordinates (lat, lon) of the acquisition, and the basemap data, Sentinel-2 and SRTM 30m Digital Elevation Model (DEM).

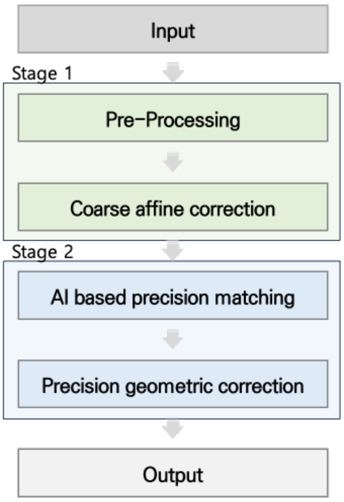

Figure 1. Workflow of proposed method.

- **Input**: Original BlueBON image (including center coordinates), Sentinel-2 L1C (Basemap), SRTM.
- **Stage 1 (Coarse Affine Correction)**: Pre-processing, such as matching coordinates and resolution between the input images, is performed, followed by rapid image matching at low resolution and coarse affine correction. As a result, an initial corrected image is generated.
- **Stage 2 (Precision Geometric Correction)**: Pre-processing is performed again between the 1st corrected image and the reference data, and precise matching points are found through hierarchical matching. Subsequently, GCP conversion of the matching points and RPC estimation are performed, and the final geometrically corrected image is generated.

3.1 Stage 1: Coarse Affine Correction

Stage 1 corrects large initial geometric errors to improve the success rate of precision matching in Stage 2. This coarse alignment is essential, as even robust deep learning models like LoFTR are vulnerable to excessive displacements.

- **Pre-processing**: Using the center coordinates and image size information of the BlueBON image, the corresponding area of the Sentinel-2 basemap and SRTM is cropped. All data are unified to the coordinate system of the basemap (Sentinel-2), and the resolution is upsampled based on the BlueBON image. Figure 2 shows pre-processing.

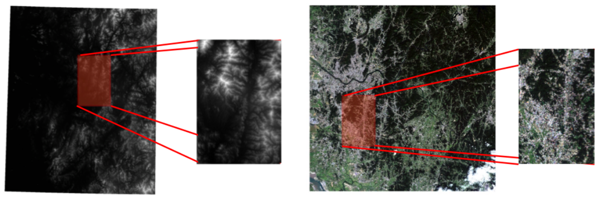

Figure 2. Pre-processed image(DEM and Basemap).

- **Initial Matching**: SuperPoint+LightGlue is applied to extract robust and uniform initial matching points between the original BlueBON image and the Sentinel-2 basemap. At this time, matching is performed on the images downscaled to 1/4 size for fast processing.
- **Coarse Correction**: The RANSAC (RANdom SAmple Consensus) algorithm is applied with the Affine transformation model to remove outliers from the extracted matching points and calculate the optimal Affine transformation matrix. Result of coarse correction is shown in Figure 3.

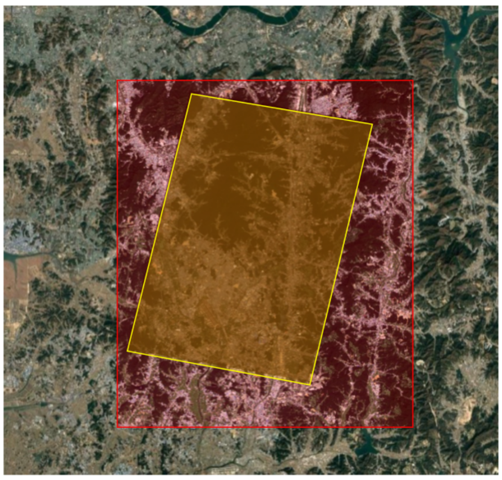

Figure 3. Result of coarse correction
(BlueBON overlaid on Sentinel, on Google Maps).

3.2 Stage 2: Precision RPC Estimation and Correction

The goal of Stage 2 is to extract a large number of precise matching points based on the initial corrected image, convert them into 3D GCPs, and estimate the final RPC model.

- **Precision Matching:** Precision matching is applied between the initial corrected BlueBON image and the Sentinel-2 basemap. Here, hierarchical patch matching is performed: the images are first downsampled to 1/4 resolution for initial matching using SuperPoint+LightGlue. Then, patches are extracted from these matching points to perform LoFTR precision matching at the original resolution. Afterward, matching points judged to be outliers based on confidence are preliminarily removed to overcome the texture differences and noise between the two images, generating dense and accurate corresponding points. Finally, to ensure a uniform GCPs distribution, the image is divided into a 5x5 grid, and filtering is employed so that matching points are evenly distributed across all grids, as illustrated in Figure 4. This hierarchical approach extracts patches around initial matching points obtained from downsampled images. It significantly reduces unnecessary computations and improves accuracy by focusing only on high-probability regions.

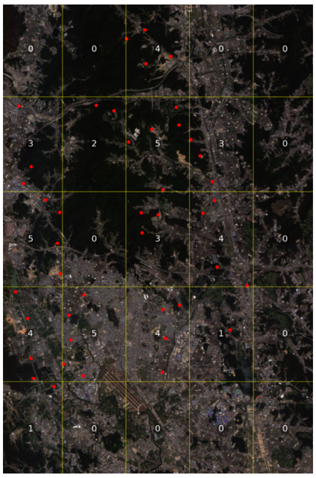

Figure 4. GCP Distribution.

- **3D GCP Generation:** Hundreds of precise matching point pairs generated through LoFTR are acquired. The geographic coordinates X, Y corresponding to the Column, Row coordinates of the corresponding points in the Sentinel-2 image are extracted, and the elevation (Z) value is extracted from the SRTM at the same location. Through this, a 3D GCP set in the format of X, Y, Z, Col, Row is generated.

- **GCP Refinement:** The generated 3D GCP set may contain outliers due to errors in SRTM itself or LoFTR matching errors. These outliers can seriously undermine the stability of RPC estimation, so RFM-RANSAC using the RFM model is performed to remove primary outliers that cause large errors in model fitting.

- **RPC Estimation:** The refined final GCP set is input into the rpcfit library to robustly estimate the RPC coefficients for the BlueBON image.

- **Final Correction:** The final precision geometrically corrected image is generated using the estimated RPC and the rpcm library.

4. Experiments and Analysis

4.1 Datasets and Evaluation Methodology

For the target imagery, 4.8m BlueBON CubeSat data was employed. This dataset represents a typical commercial CubeSat scenario where high-precision sensor models or pre-estimated RPCs are absent due to hardware and mission constraints. By validating our framework on metadata limited solely to acquisition center coordinates, we demonstrate the scalability of the proposed method to various CubeSat platforms facing similar georeferencing challenges. Seven test sites (Jamsil, Andong, Sandonaci, Baghdad, Tubarjal, Chesapeake and Taklamakan) were selected to cover diverse terrains, including high-density urban, mountainous, and desert environments, as well as areas with cloud cover. Sentinel-2 L1C (10m GSD) and SRTM 30m DEM served as the reference basemap and elevation source, respectively. The seven BlueBON test scenes and their corresponding Sentinel-2 basemaps used in the experiments are displayed in Figure 5.

To evaluate geometric precision, 20 checkpoints (10 for Andong due to landmark scarcity) were manually extracted from both Sentinel-2 and Google Maps. Coordinate differences were measured as Root Mean Square Error (RMSE) in meters and pixels. While Google Maps may contain stitching or projection errors, it served as a practical benchmark for visual verification. This dual-validation scheme ensures that the corrected imagery meets both scientific standards for accuracy and practical usability for general geospatial applications.

4.2 Experimental Results

Before geometric correction, the precision matching results with the basemap for the BlueBON images are presented (Figure 6). In this study, based on its superior robustness to scale differences (4.8m vs. 10m) and high matching precision—which are quantitatively verified in Section 4.3—LoFTR was determined to be the most suitable algorithm. Therefore,

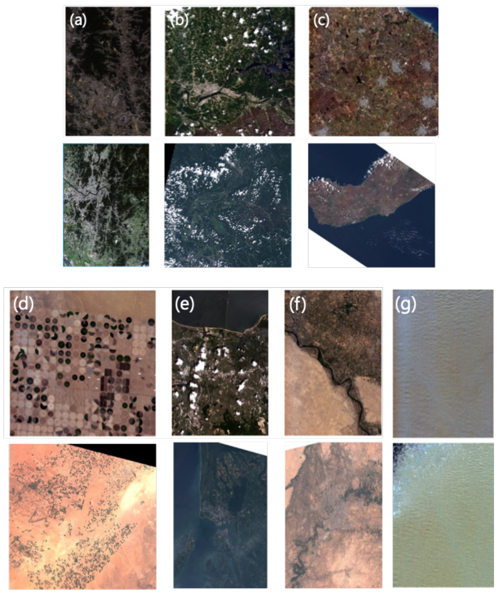

Figure 5. BlueBON(Upper) and Sentinel-2(Lower) images ((a)Jamsil, (b)Andong, (c)Sandonaci, (d)Tubarjal, (e)Chesapeake, (f)Baghdad), (g)Taklamakan.

the geometric correction results using LoFTR are presented first.

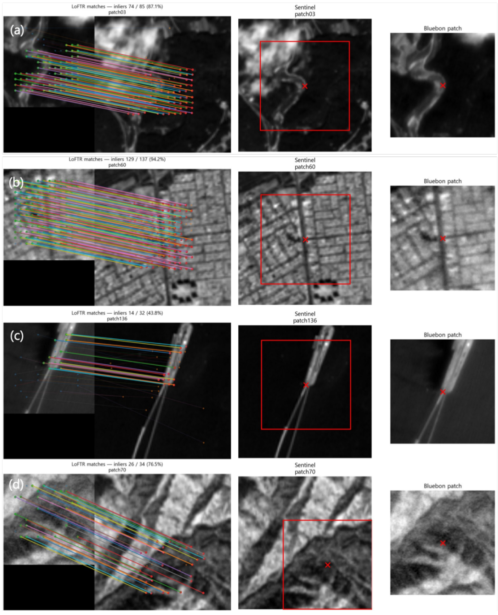

Figure 6. Precision matching results.
((a, b)Matching success, (c, d)Matching failure.

Precision matching after applying LoFTR was considered successful if the inlier ratio used for transformation estimation exceeded a threshold. Figure 6(a, b) show examples of accurate registration with a high inlier ratio. In contrast, Figure 6(c, d) fell below the threshold due to patch characteristics, such as water bodies or vegetation, although the coarse alignment was still consistent. However, due to the nature of precision geometric correction—which aims to use only matching points with minimal error and variables—these points were subsequently filtered out. The geometric correction results generated using the GCPs created from the matching results above can be seen in the following Figure 7. The Figure 7 shows the geometric correction results overlaid on Google Maps, confirming that they are placed in the correct locations.

To evaluate geometric precision, 20 check points were selected from each image and their accuracy was visually measured. The number and location of the checkpoints used for the measurement are shown in the Figure 8 below. In the case of Andong, suitable objects for checkpoints were not identified in the upper part of the image, so checkpoints could not be extracted from that area. However, in the remaining data, checkpoints were extracted with an even distribution across the entire image.

The quantitative evaluation results for geometric correction positional accuracy are presented, and the results are shown in Table 1. When compared with the Sentinel-2 basemap, an average RMSE of 8.05m (1.68 pixels) was observed, and when compared with Google Maps, an average RMSE of 9.02m (1.88 pixels) was observed. The following Table 2 presents the errors for each image.

4.3 Precision Matching Evaluation

In the precision matching stage, to ensure the precision and robustness of corresponding points between BlueBON and Sentinel-2 patches, traditional feature- and area-based matching methods (SIFT and ZNCC) were compared against state-of-the-art (SOTA) AI-based matching models: LightGlue, LoFTR, and SatLoFTR. SatLoFTR has the same architecture as LoFTR, but uses weights trained to reflect the characteristics of satellite imagery, such as illumination, viewpoint, and atmospheric scattering (Deshmukh and Kak, 2025).

	RMSE	
Reference	meters	pixels
Sentinel-2 L1C	8.05	1.68
Google Maps	9.02	1.88

Table 1. Accuracy of proposed geometric correction (RMSE).

	RMSE (BlueBON-S2)		RMSE (BlueBON-GM)	
Location	(meters)	(pixels)	(meters)	(pixels)
Jamsil	5.18	1.08	3.21	0.67
Andong	10.28	2.14	9.54	1.99
Sandonaci	10.12	2.10	8.55	1.78
Tubarjal	8.48	1.77	15.35	3.20
Chesapeake	11.36	2.37	11.57	2.41
Baghdad	6.15	1.28	10.28	2.14
Taklamakan	4.78	1.00	4.66	0.97

Table 2. RMSE results. (S2: Sentinel-2, GM: Google Maps)

4.3.1 Evaluation Metrics The key to precision matching lies not only in securing accurate correspondences but also in how effectively outliers are eliminated. The corresponding points extracted through precision matching are directly utilized in the subsequent 3D GCP generation and RPC estimation; therefore, the precision and stability of the matching have a direct impact on the final geometric correction performance. To evaluate this quantitatively, the following five metrics were defined and used.

- **Matching Success Ratio (MSR):** The ratio of total patch pairs for which valid matching was successfully performed, indicating the algorithm's fundamental matching performance.
- **Global Mismatch Ratio (GMR):** The ratio of total patch pairs classified as mismatches, evaluating the algorithm's relative mismatch frequency.
- **Filtered Patch Ratio (FPR):** The ratio of patches filtered out based on the algorithm's internal confidence criteria.
- **Error Reduction Ratio (ERR):** The ratio of mismatches eliminated through filtering, which evaluates the algorithm's mismatch suppression capability. FPR and ERR have a trade-off relationship.
- **Matching Precision (MP):** The ratio of matching points determined to be true matches (inliers), serving as a key metric that directly evaluates the accuracy of the final registration result.

4.3.2 Results and Analysis The results were analyzed using Jamsil as a representative case(Table 3).

- The Jamsil area is a high-density urban region with numerous complex artificial structures (e.g., buildings,

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-1-2026-159-2026

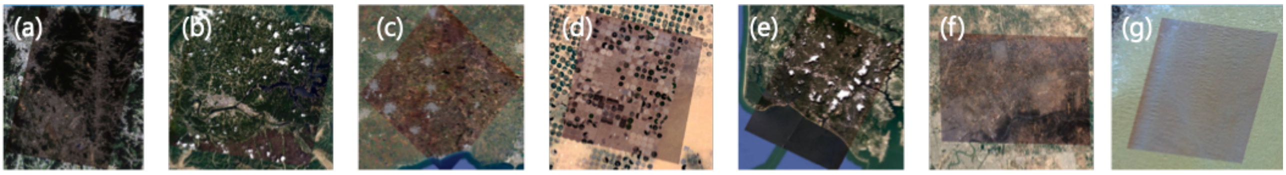

Figure 7. Results of geometric correction.

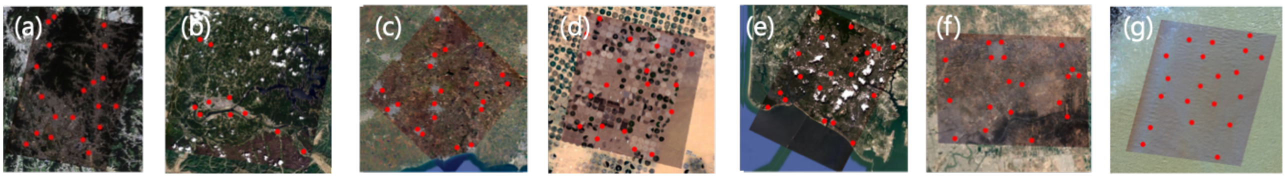

Figure 8. Locations of checkpoints.

bridges). In this environment, feature-based matching methods proved to be vulnerable due to relief displacement. In contrast, SatLoFTR's weights acted robustly against viewpoint changes in urban areas, recording a very high MP.

- LightGlue demonstrated a generally balanced performance with an MP of 92.86%, but its precision was somewhat lower compared to LoFTR. Furthermore, in iterative matching of small patch units, the limitations of its parallel feature-matching structure became apparent due to high computational costs.

Although SIFT achieved the lowest GMR (7.26%), this metric did not translate into reliable precision for geometric correction. Its reliance on descriptor similarity resulted in high-confidence yet incorrect matches (false positives) arising from repetitive building patterns. These matches proved difficult to suppress even with strict thresholds; consequently, its ERR and MP—key indicators for precision correction—were notably low. LightGlue demonstrated more reliable performance by filtering unstable matches via its hard-matching mechanism; however, its final precision was surpassed by the LoFTR-based methods.

LoFTR achieved greater stability by employing inlier-ratio–based filtering, which effectively suppressed mismatches originating from repetitive structures or moderate viewpoint differences. This approach yielded consistently high ERR and MP. SatLoFTR attained the highest overall precision, leveraging satellite-domain pretrained weights optimized for capturing the structural regularities of urban scenes. However, assessing its stability across diverse terrain types necessitates a broader regional evaluation, which is presented in Table 4.

Building on this observation, the regional results in Table 4 reaffirm the advantages of context-based matching. LoFTR achieved the highest MP and ERR across all five regions, maintaining stable precision even in challenging, low-texture areas such as Chesapeake and Tubarjal. SatLoFTR also performed well, particularly in dense urban scenes. However, its satellite-domain pretrained weights, derived primarily from VHR (Very High-Resolution) imagery like WorldView-3, render it highly specialized for fine-grained structural patterns. While this specialization enhances performance in complex urban environments, it reduces its generalizability when matching against midium-resolution basemaps or in homogeneous terrain. This

Algorithm	MSR	GMR	FPR	ERR	MP
SIFT	99.44% (178/179)	7.26% (13/179)	19.66% (35/178)	69.23% (9/13)	88.57% (31/35)
ZNCC	~	19.55% (35/179)	76.54% (137/179)	54.29% (19/35)	88.32% (121/137)
LightGlue	91.62% (164/179)	17.32% (31/179)	42.68% (70/164)	83.87% (26/31)	92.86% (65/70)
SatLoFTR	100% (179/179)	33.52% (60/179)	58.66% (105/179)	**95.00% (57/60)**	**97.14% (102/105)**
LoFTR	99.44% (178/179)	16.76% (30/179)	61.80% (110/178)	83.33% (25/30)	95.45% (105/110)

Table 3. Quantitative evaluation of precision matching performance for the Jamsil region.

led to slight performance degradation in several regions, precluding it from consistently outperforming LoFTR.

In contrast, traditional feature-based approaches (SIFT, ZNCC) exhibited significant fluctuations in MP and ERR based on terrain characteristics, confirming their sensitivity to relief displacement and repetitive patterns. LightGlue demonstrated more balanced performance, but its precision was still insufficient for stable RPC estimation. Taken together, these results indicate that LoFTR provides the most stable and terrain-independent matching performance among all evaluated methods.

Location	Eval.	SIFT	ZNCC	LightGlue	SatLoFTR	LoFTR
Andong	MP	70.00%	84.31%	97.78%	97.40%	**98.75%**
	ERR	65.12%	64.44%	95.65%	96.83%	**97.44%**
Baghdad	MP	80.00%	92.65%	86.96%	98.32%	**98.41%**
	ERR	52.63%	52.38%	52.63%	91.30%	**98.18%**
Chesapeake	MP	89.74%	87.42%	95.90%	97.44%	**98.11%**
	ERR	63.64%	48.65%	77.27%	93.75%	**94.12%**
Sandonaci	MP	88.39%	75.00%	93.33%	96.55%	**96.97%**
	ERR	70.97%	85.71%	89.74%	96.61%	**98.21%**
Tubarjal	MP	79.00%	93.89%	96.92%	**97.46%**	96.91%
	ERR	61.11%	50.00%	94.12%	**93.02%**	90.91%
Taklamakan	MP	93.85%	96.92%	98.46%	96.43%	**100.00%**
	ERR	0.00%	0.00%	0.00%	93.33%	**100.00%**

Table 4. Quantitative evaluation of precision matching.

These results confirm that the type of LoFTR algorithms maintain consistent precision across diverse terrains and image characteristics, demonstrating excellent registration stability and mismatch suppression performance(Table 4). We experimentally demonstrated that the context-based matching characteristic of the LoFTR secures reliable corresponding points in diverse terrain environments. This, in turn, contributed to the improved accuracy of the subsequent 3D GCPs generation and RPC model correction. Although SatLoFTR and LoFTR were confirmed to show similar performance numerically, this paper applied LoFTR because it not only showed superior performance in four of the seven regions, but also consistently achieved a matching precision exceeding 95%, indicating its strong general applicability across satellite images with diverse characteristics. Lower matching quality of baseline methods led to biased RPC estimation, justifying their exclusion from final assessment. Interestingly, the high precision in the Taklamakan region is attributed to the topographic stability and temporal consistency of desert textures across the disparate datasets, which minimizes geometric distortions and facilitates stable matching.

4.4 Additional Experiments

In this study, additional experiments were conducted on two datasets to evaluate the extended applicability of the proposed technology. The purpose of this experiment was to evaluate the possibility of geometric correction through matching between high-resolution satellite images. There is a slight mismatch with the main paper's premise, as the basemap's resolution is higher. However, it was conducted to confirm the feasibility of matching and geometric correction between high-resolution satellite images and to verify the scalability of the research.
The target image used in the experiment was CAS500-1 (0.5m resolution, pan-sharpened), which was chosen as the target because geographic information for high-resolution satellite images is often not provided in South Korea. A WorldView-3 image (30cm resolution) was used as the basemap, and the SRTM was used, as in the previous experiments.

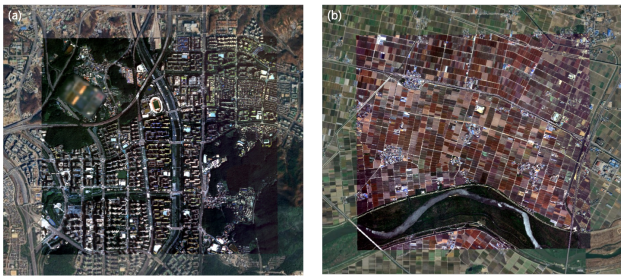

Figure 9. Results of additional test((a)Jamsil, (b)Iksan).

Reference Basemap	RMSE (meters)	RMSE (pixels)
Jamsil(City)		
Google Maps	4.93	9.86
WorldView-3	3.94	7.88
Iksan(Farmland)		
Google Maps	3.13	6.26
WorldView-3	1.93	3.86

Table 5. RMSE comparison for CAS500-1 (GSD 0.5m).

As confirmed in Figure 9 above and the accuracy assessment, automatic geometric correction through matching and RPC estimation based on the basemap is also feasible for high-resolution satellite images. In particular, for the Iksan (farmland) case, it was confirmed that an accuracy in the around 3 pixels range was achieved (Table 5), even without any specific optimization. However, to reach the desired level of accuracy relative to the resolution, it appears necessary to apply advancements in matching technology.

4.5 Discussion and Analysis

The experimental results achieved an average RMSE of 1.79 pixels. The achieved accuracy falls within the 1–3 pixel orthoimage guideline recommended by ASPRS (American Society for Photogrammetry and Remote Sensing) (ASPRS, 2014), leading to the conclusion that the overall performance is excellent. The following analysis is possible based on the experimental results.

- **Validity of Low-Resolution Basemap:** The most noteworthy result is that even though a 4.8m resolution image was corrected with a 10m resolution basemap, an error of 8.05m (1.68 pixels) was achieved, which is more precise than the basemap's pixel resolution (10m). This is judged to be possible because, as discussed in Section 2.3, context-based matching technologies like LoFTR do not simply match the geometric positions of edges or corners, but rather learn and match the contextual relationship between image patches. That is, LoFTR can infer the similarity between the texture pattern of a building cluster represented by 10m pixels and the texture pattern of a building cluster represented by 4.8m pixels to find precise corresponding points. However, this result showed success at approximately a 2x resolution difference. Further verification is needed to determine if the same performance can be achieved with larger resolution disparities.

- **Impact of Terrain and Image Characteristics:** The Jamsil area showed the highest accuracy. This is because the diverse and rich texture patterns, such as building clusters, rivers, and mountains, had a positive effect on LoFTR's contextual analysis, leading to the stable generation of high-quality GCPs. Furthermore, it was confirmed that geometric correction was performed stably even in mountainous terrain or in images with cloud cover, like Andong and Chesapeake, which raises expectations that the problems of geometric correction technology using GCP Chips, as raised in the introduction, can be solved. On the other hand, the Baghdad, Chesapeake, and Tubarjal regions showed relatively lower accuracy. This is because areas with homogeneous textures, such as deserts, seas, and large agricultural lands, were widely distributed within those images, making it difficult to generate matching points (GCPs). This causes the GCPs to be unevenly distributed (unbalanced distribution) in only some areas of the image (e.g., cities), which reduces the geometric stability during RPC estimation and causes large errors in areas without matching points (deserts, seas).

- **Error Budget Analysis**: The error factors in this experiment should be analyzed as a complex interaction of the following factors. First is the propagation of DEM vertical error. The SRTM used for 3D GCP generation has a vertical error of about 5.94m (Elkhrachy, 2018). This vertical error propagates as horizontal error through the RPC model and is a major cause of error, especially in the Jamsil area with many high-rise buildings or in mountainous terrain. However, considering the resolution of the satellite image used, the error is judged to be minimal. Next, matching error must be considered. It can be assumed that there is a matching error from the LoFTR algorithm itself, but checking the matching results indicates that the matching error itself is minimal.

In terms of operational efficiency, the entire pipeline—from data ingestion to final corrected image generation—takes approximately 5 minutes per scene on a standard laptop (MacBook M4 Pro). This efficiency, combined with the automated nature of the two-stage framework, highlights the practical utility of the proposed method for rapid geospatial response and large-scale processing.

5. Conclusions and Future Work

5.1 Conclusions

This paper proposed a fully automated geometric correction framework using open access basemaps (Sentinel-2) and AI matching (SuperPoint+LightGlue, LoFTR) techniques for satellite images where RPC are not provided. Applying the proposed methodology to several actual BlueBON (4.8m) datasets, an average RMSE of 8.05m (1.68 pixels) was achieved referenced to the 10m resolution Sentinel-2 basemap. This demonstrates that precise geometric correction of high-resolution images is possible even with a low-resolution open basemap, and suggests that context-based matching techniques like LoFTR played a key role in overcoming the resolution difference. Furthermore, applying it to various satellite data confirmed the extended applicability of this research, which is judged to be a foundation for establishing a system for fusion analysis between different satellites. Additionally, this study is based on open data and open-source libraries (rpcfit, rpcm), enhancing accessibility, and has the practical advantage of providing users with immediately usable RPC. Notably, the framework demonstrated exceptional stability in desert regions like Taklamakan. This is likely due to the topographic stability and temporal consistency of desert textures, which remain relatively invariant across disparate datasets compared to urban or agricultural areas. Such characteristics allowed the AI-based matching to establish more reliable correspondences despite the lack of prominent features.

5.2 Future Work

To address RPC estimation failures and accuracy degradation caused by GCP quality issues, future research will focus on the following:

- **Ensuring Robustness:** We plan to optimize the robust parameters provided by the rpcfit library and develop a robust RPC estimation model that is resilient to DEM errors or unbalanced GCP distributions.
- **Improving 3D GCP Quality:** We will replace SRTM with more accurate elevation data, such as the Copernicus DEM , and develop AI-based Z-value estimation technology to enhance the vertical accuracy of 3D GCPs (Ghannadi et al., 2023).
- **Solving the GCP Distribution Problem:** To overcome matching failures in extreme desert or ocean regions, we will develop a hybrid strategy combining LoFTR with frequency-based phase correlation to ensure uniform GCP coverage even in textureless environments (Kim et al., 2018).

References

Akiki, R., Marí, R., de Franchis, C., Morel, J., Facciolo, G., 2021. Robust rational polynomial camera modelling for sar and pushbroom imaging. *2021 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Brussels, Belgium, 7908–7911.

ASPRS, 2014. ASPRS Positional Accuracy Standard for Digital Geospatial Data. *Photogrammetric Engineering Remote Sensing*, 81(4), 281–330.

De Franchis, C., Facciolo, G., Meinhardt-Llopis, E., 2021. rpcm: Rational polynomial camera model. GitHub Repository. https://github.com/centreborelli/rpcm.

Deshmukh, R., Kak, A., 2025. Satdepth: A novel dataset for satellite image matching. arXiv preprint arXiv:2503.12706v1.

Elkhrachy, I., 2018. Vertical accuracy assessment for SRTM and ASTER Digital Elevation Models: A case study of Najran city, Saudi Arabia. *Ain Shams Engineering Journal*, 9(4), 2955–2964.

Gaudel, A., Langullie, F., Delvit, J., Michel, J., Cournet, M., Poulain, V., Youssefi, D., 2017. Sentinel-2 global reference image validation and application to multi-temporal performances and high latitude digital surface model. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-1/W1, 447–454.

Ghannadi, M. A., Alebooye, S., Izadi, M., Ghanadi, A., 2023. VERTICAL ACCURACY ASSESSMENT OF COPERNICUS DEM (CASE STUDY: TEHRAN AND JAM CITIES). *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, X-4/W1-2022, 209–214.

Kim, H., Son, J., Kim, T., 2018. Geometric Correction for the Geostationary Ocean Color Imager from a Combination of Shoreline Matching and Frequency Matching. *Sensors*, 18(11), 3599.

Luo, Q., Zhang, J., Xie, Y., Huang, X., Han, T., 2024. Comparative analysis of advanced feature matching algorithms in challenging high spatial resolution optical satellite stereo scenarios. *IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium*, 2645–2649.

Paul, S., Pati, U. C., 2021. A comprehensive review on remote sensing image registration. *International Journal of Remote Sensing*, 42(14), 5396–5432.

Rengarajan, R., Storey, J. C., Choate, M. J., 2020. Harmonizing the Landsat Ground Reference with the Sentinel-2 Global Reference Image Using Space-Based Bundle Adjustment. *Remote Sensing*, 12(19), 3132.

Shin, J.-I., Yoon, W.-S., Park, H.-J., Oh, K.-Y., Kim, T.-J., 2018. A Method to Improve Matching Success Rate between KOMPSAT-3A Imagery and Aerial Ortho-Images. *Korean Journal of Remote Sensing*, 34(6_1), 893–903.

Sun, J., Shen, Z., Wang, Y., Bao, H., Zhou, X., 2021. Loftr: Detector-free local feature matching with transformers. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 8922–8931.