IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received November 5, 2020, accepted November 16, 2020, date of publication November 19, 2020,
date of current version December 9, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3039413

Image-Based Monte-Carlo Localization With Information Allocation Logic to Mitigate Shadow Effect

SUNG HYUK CHOI[1] AND CHAN GOOK PARK[2,3], (Member, IEEE)

[1]Navigation and Electronic System Laboratory, Department of Mechanical and Aerospace Engineering, Seoul National University, Seoul 08826, South Korea
[2]Navigation and Electronic System Laboratory, Department of Aerospace Engineering, Seoul National University, Seoul 08826, South Korea
[3]Automation and Systems Research Institute, Seoul National University, Seoul 08826, South Korea

Corresponding author: Chan Gook Park (chanpark@snu.ac.kr)

ABSTRACT The development of vision-based navigation algorithms using a camera is becoming more important. The vision-based navigation can be categorized into two types. The first is to use sequential camera images as relative navigation. The second is to estimate the absolute navigation solution using a camera image and database. In absolute navigation, the difference between the database and the camera image is a major obstacle to image registration. One of the factors that make a lot of difference is the shadow effect. This shadow increases the inconsistency between the two images and eventually degrades the localization accuracy. This means that shadows have a significant impact when measuring the similarity of the two templates. To mitigate this effect, we inherited and developed the Monte Carlo Localization (MCL) algorithm based on a new similarity cost function, which is a key contribution to this article. We have established the importance of information with information reallocation logic that considers shadow areas. The proposed algorithm allocates the importance of the information considering a portion of the shaded area in the camera image. First of all, we analyzed the effects of shadows on the camera. To compare the performance of the algorithm, we used not only the shadow restoration algorithm but also various template-based matching algorithms. The proposed algorithm is validated through various simulations and real flight experiments as well.

INDEX TERMS Aerospace engineering, geographic information systems, Monte Carlo methods, particle filters, adaptive systems.

I. INTRODUCTION

Autonomous robots are being studied extensively, especially in drones (UAVs), because they can be accessed directly in a dangerous or inappropriate environment. These drones are known to be useful in many application areas, such as exploration, inspection, guidance generation, and surveillance, as they are significantly free to maneuver [1], [2]. Localization or navigation is an essential research area for an airborne vehicle to perform complex tasks. A common way to do UAV localization is to use the Global Navigation Satellite System (GNSS). In particular, the number of satellites and the quality of the signal play an important role in calculating the UAV's position. However, expected position accuracy can be degraded by multipath problem, fewer satellites, spoofing, and jamming [3]. Furthermore, according to White and Maybeck [4], jamming can cause signal interference, which can make location estimation unreliable as well as spoofing. As a result, aerial drones that rely solely on GNSS signals can be hacked by malicious behavior with catastrophic results. Therefore, the development of vision-based navigation algorithms using a camera is becoming more important. Localization using the vision sensor can be classified into relative navigation and absolute navigation [5]. First, relative navigation calculates the relative velocity to estimate the position of the aerial vehicle using the difference between the sequential images obtained through the vision sensor, generally called a visual-odometry (VO) [6], [7]. However, a VO has the disadvantage of accumulating errors caused by various factors. On the other hand, scene-matching (SM) based on image registration [8], [9] can be an absolute navigation technique that calculates the position of an aerial vehicle by directly matching the database, including the geographic information. It can be said that a kind of database-referenced navigation (DBRN) [10] system. Except for the disadvantage of requiring a large, it can be a good way to compensate

The associate editor coordinating the review of this manuscript and approving it for publication was Rosario Pecora.

for the diverging navigation solution when the use of the global navigation satellite system (GNSS) is limited [11]. The matching result of the absolute localization based on SM can be directly converted into the position information in the database; hence errors are not accumulated [12].

However, there is a problem that the camera image and the database are inconsistent. Many causes make the inconsistency between the two templates. In general, it is known that a large difference is caused by moving objects, the creation and disappearance of buildings, illumination changes [13]. First of all, the moving object and building have a certain structure, some researchers use this structural information to mitigate mismatching. Yang *et al.* focused on regions of interest to increase robustness [14]. Liu *et al.* also proposed an optimization method using spatial-order constraint [15]. A Viewpoint-dependent matching method with a high success rate was proposed by Shan *et al.* [16]. These methods of using structural constraints require prior knowledge of the object. The next, illumination change is hard to predict. One of the ways to solve this problem is the shadow detection and restoration process. Actually many researchers proposed algorithms in this type of researches. Nan *et al.* have been proposed an automatic shadow detection method to mitigate the shadow effect [17]. And the shadow detection and removal method using a new color space are introduced by Liang *et al.* [18]. Ding. *et al.* inherited the computer vision algorithm and developed a new shadow removal algorithm [19]. The performance of these types of using shadow restoration techniques can be improved by using a specific threshold, but it is difficult to apply it to various environments. Recently, machine learning algorithms have been developed to mitigate the shadow effect. Ma provides the convolutional neural network feature and sharpness measure [20]. Krajník and Mu proposed an algorithm that increases the accuracy of the navigation through the feature information learning process according to the environmental change [21], [22]. While this approach has the advantage of being able to use it in a variety of environments, it is hard to explain the process and result.

Besides, there is a new research flow. This is a vision-based localization system that applies the Monte Carlo localization (MCL) measurement model by converting UAV poses into MCL problems [23], [24]. The MCL is a particle filter which is one of the recursive Bayes filters. In particular, to solve the vision-based UAV localization problem, such as satellite images and downward aerial images should be required as the database with geo-localization information. This type of database covers most of the globe, which is advantageous for use in any environment. For example, users can utilize such as Google Maps with geographic information on various systems. The MCL sensor model plays an important role because it directly affects the efficiency and robustness of the localization process. There are several algorithms to evaluate the similarity of image templates. The Sum of squared differences (SSD), Normalized cross-correlation (NCC), Mutual information (MI), and abBRIEF [24] may be used for calculating the likelihood of MCL. These similarity functions are being used to compare the performance of the algorithm in many types of researches [25]–[27]. In particular, the MI algorithm can be said to be the latest algorithm used for unsupervised learning [28]. An improperly designed likelihood function can over-confidently place the vehicle in its position, which leads to filter divergence. However, there are not many studies in this MCL frame-work that have addressed the inconsistency problem.

Therefore, we analyze the effects of shadows on each algorithm and proposes an alternative algorithm. Our proposed measurement model uses an activation function to reduce the effects of shadows. It is the main contribution that the importance of the information is allocated through the portion of the shadow area in the image. We have validated the proposed algorithm comparing with various types of similarity functions and shadow restoration algorithm. Besides, simulation results show that the proposed algorithm works much more effectively than the conventional algorithms and similarity functions. We also analyzed the convergence rate and accuracy of the filter, which varies with the number of particles and other user parameters. And real flight experimental data is used for validating algorithm as well.

This article is organized as follows. Some research related to the proposed algorithm with the Monte-Carlo localization method and basic knowledge is described in section II. In section III, we propose the algorithm with information allocation with the portion of the shadow in the image. In section IV, we describe the simulation and experiment results, and the analysis of the results with discussion is described in section V. The conclusion of this research article is organized in section VI.

II. MONTE-CARLO LOCALIZATION WITH INTENSITY-BASED WEIGHTING

A. MONTE-CARLO LOCALIZATION

To estimate the position of the vehicle, we choose the probabilistic approach, known as the recursive Bayes filter. The likelihood of MCL can be calculated by the prediction and correction process. Which is updated according to (1).

$$p(x_t|z_t,\ u_t) = \alpha \cdot p(z_t|x_t) \cdot p(x_t|u_t,\ x_{t-1}) \cdot p(x_{t-1}) \quad (1)$$

where z is the measurement, the vehicle's position is x in environment m.

And α is a normalization constant that leads to the summation of the density equals one. $p(x_t|u_t,\ x_{t-1})$ represents the probability of the pose given the movement by control input u_t and prior position x_{t-1}. And $p(z_t|x_t)$ is the probability of making observation z_t given the vehicle's current location x_t. The appropriate computation of this quantity is the subject of intensity-based MCL. An improperly designed likelihood function can over-confident the vehicle to its position, thus leading to filter divergence. The whole process of the MCL is summarized in table 1. Therefore, setting the likelihood based on the camera image is an important part of vision-based MCL. This will be covered in section II.C.

TABLE 1. Simple Progress of Monte Carlo Localization (MCL).

Step 1. Initialization
$t = 0$
$for\ i = 1,\ldots, N$
$x_i^{(i)} \sim p(x_0)$:particle generation
Step 2. Importance sampling
$for\ \mathrm{i} = 1,\ldots.N$
Propagation
$x_i^{(i)} \sim p(x_i \mid x_{i-1})$
$x_{0:t}^{(i)} = (x_{0:t-1}^{(i)}, x_{0:t}^{(i)})$
Calculation of importance weight
$w_t^{(i)} = p(z_t \mid x_t^{(i)})$
Step 3. Normalize and resampling
$t = t + 1$
Step 4. Repeat step. 2~3

FIGURE 1. Camera image (-) and reconstructed image (–) with particle (*).

B. CAMERA AND RECONSTRUCTED IMAGE

First, it is about camera image, measurement, from the motion of the vehicle. The captured sequential photographs can be used to calculate the position of the vehicle in the database.

In the model-based particle filter, particles can be propagated to estimate the position of the vehicle. It is possible to estimate more accurately by judging the similarity with reconstructed camera images cropped from the database based on the particles being sprayed. It is called a reconstructed camera image. The particles are scattered over the database with the expected vehicle's position according to a given system model (vehicle's movement), with uncertainty. Sprayed particles represent the locations where the vehicle is expected to be located, so we can create an estimate of each particle. In other words, it means the process of extracting the virtual camera image from the database. We define this as a reconstructed camera image. Camera image (-) and reconstructed camera images (–) are illustrated in fig. 1.

C. INTENSITY-BASED WEIGHTING

In the intensity-based approach, UAV is assumed as the cruising status with constant altitude, and the attitude of the camera is fixed except the yawing, ψ. Because the UAV has a down-looking camera that captures image I with a gimbal. The camera image is I and the intensity-based descriptor is $d_i = (\tau_{i1}, \cdots, \tau_{ic})^T$. Which is developed to reduce the computational load using binary value. c is the number of image color channels and K is the k pairs of target sample position, $p_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2})$. $\tau_{i,j}$ is derived by intensity comparison between $\delta^j(\mathbf{x}_{i,1})$ and $\delta^j(\mathbf{x}_{i,2})$. Target samples are used for calculating the intensity-based descriptor as (2). Fig. 2 shows

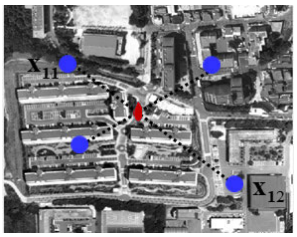


FIGURE 2. Aerial image and sample target pair to calculate intensity-based descriptor.

the conceptual sample target pair in the aerial image.

$$\tau_{i1} = \begin{cases} 1, & if\ \delta^j(\mathbf{x}_{i,1}) > \delta^j(\mathbf{x}_{i,2}) \\ 0, & Otherwise \end{cases} \tag{2}$$

The next step is calculating the expected descriptor, $\hat{d}$, from the reconstructed camera image, $\hat{I}$,. The particles represent the expected positon of the vehicle in the image database, m. When the particle is in the real vehicle's location, the virtual image and the real camera image are almost identical. In the same way, the virtual target sample position can be a $\hat{p}_i = (\hat{\mathbf{x}}_{i1}, \hat{\mathbf{x}}_{i2})$, and the each position $\hat{\mathbf{x}}_{i,b}$is estimated by (3)

$$\hat{\mathbf{x}}_{i,b} = \begin{bmatrix} sR(\psi) & T \\ 0 & 1 \end{bmatrix} \cdot \mathbf{x}_{i,b} \quad \text{with } b = 1,\ 2,\ .. \tag{3}$$

where ψ is the vehicle's yaw angle, s is a scale factor that relates the altitude and the field of view. b represents the number of pairs. Ultimately we can calculate the intensity-based descriptor, $\hat{d}_i = (\hat{\tau}_{i1}, \cdots, \hat{\tau}_{ic})^T$, from the virtual image, $\hat{I}$, as well. So that we are ready to weight the similarity between two templates. The similarity between the camera and the virtual image can be calculated as (4) using the descriptor, d and $\hat{d}$.

$$\xi(d, \hat{d}) = \eta \sum_{i=1}^{k} \sum_{j=1}^{c} \tau_{i,j} \oplus \hat{\tau}_{i,j} \tag{4}$$

where $\eta = (Kc) - 1$, and c is the number of color channels and $\oplus$ is the XNOR operator. The similarity value, the weighting of information is calculated by (4). The similar the two images are, the closer the value is to 1, and the smaller the similarity, the closer to 0. In the ideal case, the captured camera image and virtual image are identical. So the similarity value should be 1. It will be used for calculating the likelihood of the camera image. Therefore, the intensity-based measurement likelihood and its distribution are defined as (5).

$$p(z|x,\ m) = \xi(d, \hat{d}) \sim N(\mu,\ \sigma) \tag{5}$$

where μ is the mean and σ is the standard deviation. And the summarized process of the intensity-based weighting is described in table 2. In this weighting, the calculation budget depends on c and K.

III. SHADOW EFFECT MITIGATION

A. SHADOW EFFECT IN INTENSITY-BASED MCL PROBLEM

The intensity-based MCL algorithm introduced in section II executes image registration depending on the intensity of

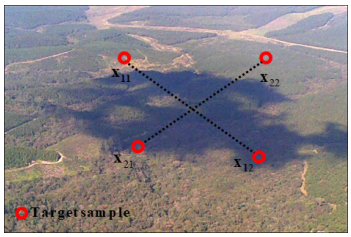


FIGURE 3. Aerial image and sample target pair to calculate intensity-based descriptor with shadow.

TABLE 2. Process of the Intensity Based Weighting.

Step 1. Choose the pixel pair randomly on the camera image then, pixel candidate is automatically selected on the reconstructed boundary

$$\hat{\mathbf{x}}_{i,b} = \begin{bmatrix} sR(\psi) & T \\ 0 & 1 \end{bmatrix} \cdot \mathbf{x}_{i,b}, \quad with\, b = 1,2,\ldots,$$

s : *Scale factor*
R : *Rotation matrix* $\in SO(2)$
ψ : *Rotation angle*
T : *Translation*

Step 2. Intensity comparison to randomly selected pixel pair

$$\tau_{i,j} = \begin{cases} 1, \; if\; \delta^j(x_{i,1}) > \delta^j(x_{i,2}) \\ 0, \; Otherwise \end{cases},$$

c : *color channel*, $\delta^j(x_{i,1}), \delta^j(x_{i,2})$: *Intensity of pixel*

Step 3. Descripting the information obtained from the step. 2

$d_i = (\tau_{i,1}, \cdots, \tau_{i,c})^T$, i : $i-th$ *pair on patch*

Step 4. Similarity check

Performs XNOR operations to randomly selected pair

$$\xi(d, \hat{d}) = \eta \sum_{i=1}^{k} \sum_{j=1}^{c} \tau_{i,j} \bar{\oplus} \hat{\tau}_{i,j}, \quad \eta = \frac{1}{Kc},$$

c : *color channel* $\bar{\oplus}$: *XNOR operator*,

K : *The number of set of random pixel position*

Step 5. Calculate likelihood

$$p(z \mid x, m) = \xi(d, \hat{d}) \sim N(\mu, \sigma)$$

each pixel. However, if the camera image contains shadows as shown in Figure 3, there may be significant intensity differences from the database. As a result, the observed and the expected intensity-based descriptors may not be matched. It may cause a decrease in image matching accuracy and location reliability.

If the intensity difference is used to measure the similarity between the two templates when the target sample is placed in the shadow, the contaminated information is utilized. In the end, even though the similarity between the two templates is high, the similarity function has a low value due to the shadow area. For example, in Fig. 3, $\mathbf{x_{11}}$'s intensity is higher than $\mathbf{x}_{12}$, however, in practice the value of $\mathbf{x}_{12}$ can be greater than $\mathbf{x_{11}}$. In such a case, MCL cannot operate normally. As a result, it can be assumed that the use of unknown information increases localization errors.

We applied a technique to reallocate the value of information located in the shadows. To reduce the effect on shadows, we introduce the image processing technique first. Which include shadow detection and pixel restoration in section III.B. Then, the information allocation method is provided in section III.C. It uses a scheme called the importance of information about target sample points.

FIGURE 4. Aerial image and detected shadow area.

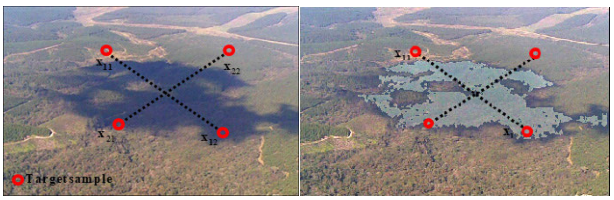


FIGURE 5. Aerial image and sample target pair (Left: with shadow / Right: with restored shadow).

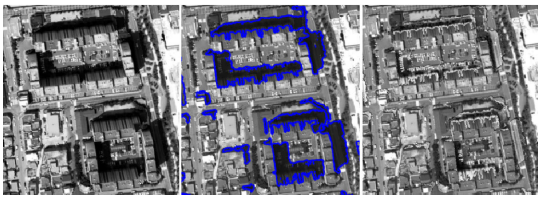


FIGURE 6. Example of shadow detection, removal, and restoration for aerial image.

B. SHADOW DETECTION AND RESTORATION

The computer vision research field has a variety of algorithms for shadow detection and pixel restoration algorithms. We applied [29] and [30] algorithms for detecting shadow easily. And the histogram equalization is used as shadow area restoration [31]. Histogram equalization performs the function of changing the distribution of pixel values in an image to appear evenly over the entire area.

After the series of shadow detection processes, the image containing the shadows changes as shown in Fig. 4. On the left of Fig. 5 represent an image containing shadow, and on the right side of the figure, it represents the restored shadow area using image processing algorithms. If the user utilizes an image that contains the reconstructed pixels, it can be expected to have a similar color to the original image. This helps improve the consistency of the images to be compared.

Fig. 6 shows the actual aerial image with shadow, the image where the shadow was detected, and finally the image where the shadow area was restored. We can expect that after the shadow detection and restoration process, it will generate an accurate matching result in MCL. If the template is well restored, the value of the similarity function is close to 1, otherwise, it will have a value less than 1.

C. ACTIVATION FUNCTION CONSIDERING SHADOW AREA

In this section, we propose an activation function for assigning the importance of particles used in MCL. First, there can be three cases where the sample points in the image containing the shadow are as shown in Fig. 7.

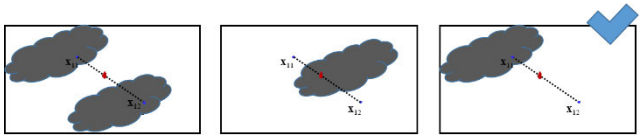


FIGURE 7. Possible location of the target sample pair on the image with a shadow.

From the left side of the figure, this indicates a sample point and its pair are both in the shadow, two sample pairs not in the shadow, and finally only one of the target sample points is in the shadow. We only control the importance of the particle in the third case. Because other cases could be assumed that the intensity of the pixels has changed with a similar level.

To determine if the sample point is in the shadow, we should first detect the shadow area. Then generate the target sample point and count the number of samples existing in the shadow area. If one of the target sample point pairs exists on the shadow, the target sample point pair is deleted. It helps to make a complete similarity check without the use of contaminated information. Next, the similarity function, which is the core of vision-based MCL, is calculated, and the calculated value is multiplied by the activation function to give the importance of the calculated value. The similarity function is described in (6), and lc means the number of points located on the shadow of the number of all target samples K. The ratio of the target sample point located in shadow is reflected and used for similarity tests through the normalization process. In other words, it is a strategy to discard contaminated information and reduce its importance.

$$\xi_m(d,\hat{d}) = \eta_m \sum_{i=1}^{k-lc}\sum_{j=1}^{c} \tau_{i,j} \bar{\oplus} \hat{\tau}_{i,j}, \quad \eta_m = \frac{1}{(K-lc)c} \tag{6}$$

$\bar{\oplus}$: *XNOR operator*, c : *color channel*, lc : *loss count*

k : *The number of set of random pixel position (user parameter)*

The sigmoid function was chosen to readjust the importance of the information. The sigmoid function is bounded, differentiable, and real function that is defined for all real input values and has a non-negative derivative at each point as (7). It is widely used for the machine learning process in backpropagation. The activation function is a modified sigmoid function to multiply the factor with a value between 0 and 1 to adjust the importance of information, which can be expressed as (8). Here a is the slope of the activation function. This serves to make the rate constant, which decreases with the proportion of samples placed in the shadow regardless of the number of samples.

$$\sigma_m(x) = \frac{1}{1+e^{x}} \tag{7}$$

$$\sigma_m(K, lc) = 1 - \frac{1}{1+e^{-a(lc-K/2)}}, \tag{8}$$

Fig. 8 shows the proposed activation functions for a number of different target samples. The horizontal axis represents lc. Which is the number of points located on the shadow

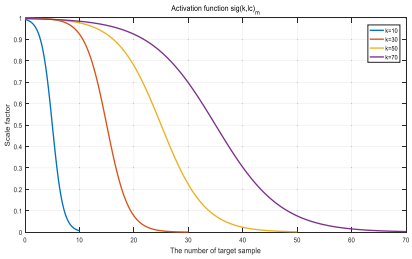


FIGURE 8. Proposed activation function based on the ratio of the target sample located in shadow (k = 10, 30, 50, 70).

TABLE 3. Process of the proposed weighting algorithm considering shadow.

1. Initialization
   - $t = 0$
   - *for* $K = 1,\ldots,N$
   - $x_t^{(k)} \sim p(x_0)$:particle generation
   - $t = 1$
2. Propagation
   - *for* i = 1,....N
   - $x_t^{(i)} \sim p(x_t \mid x_{t-1})$
   - $x_{0t}^{(i)} = (x_{0t-1}^{(i)}, x_{0t}^{(i)})$
   - *Generate reconstruction image*
3. Shadow detection in camera image
   - Count sample target located on shadow
4. Calculation of Importance weight
   - *for* $i = 1,\ldots,N$
   - *Calculate Activation function*: $\sigma(k,lc)$
   - *Similarity check*: $\xi_m(d,\hat{d}) = \eta \sum_{i=1}^{k-lc}\sum_{j=1}^{c} \tau_{i,j} \bar{\oplus} \hat{\tau}_{i,j}$,
   - $w_t^{(k)} = p(z_t \mid x_t^{(k)}) \approx \xi_m(d,\hat{d}) \cdot \sigma(K,lc)$,
5. Normalize importance weights
   - Resampling
   - $t = t+1$
   - Repeat 2~5

of the target samples pair. And the vertical axis represents the importance of information. It will have the scaling value from zero to one. As we can find that the activation function changes depending on the number of used target samples. In other words, it represents the percentage of the total number of points located in the shadow among the target samples used. Since the larger the ratio, the more information is contaminated and the amount of information is reduced, so the importance of the amount of information is also reduced. In the last, the activation function is multiplied by (6). The importance of information can be re-allocated. Finally, the terminal likelihood used for MCL is defined in (9).

$$p(z|x,m) \triangleq \xi_m(d,\hat{d}) \cdot \sigma_m(K, lc), \tag{9}$$

Table 3 shows the whole process of the proposed algorithm applied to the information allocation logic. This is a significantly effective way to increase the reliability of the calculated result. Compared to table 2, the total computational volume is slightly increased, but we can effectively

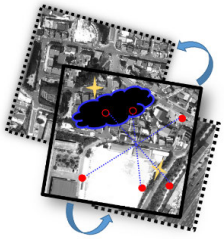


FIGURE 9. The visual concept of the proposed weighting algorithm considering shadow.

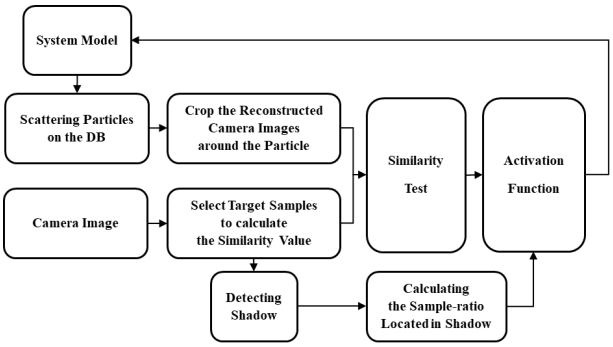


FIGURE 10. Block diagram of the proposed overall system.

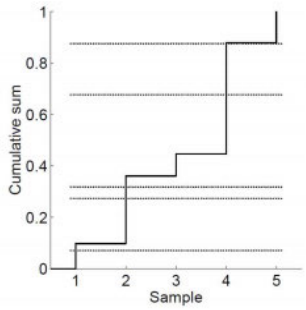


FIGURE 11. The process of multinomial resampling.

control the importance of contaminated information. And the summarized concept and block diagram of the proposed algorithm are described in Fig. 9 and Fig. 10.

There are several types of resampling process. We used a multinomial resampling approach as a standard method. The principle is to pick a realization to generate independent samples from the uniform. To sort samples, they are arranged in ascending order and comparing them with the cumulative sum of normalized weight. These process are illustrated in Fig. 11.

IV. SIMULATION AND EXPERIMENT

A. SIMULATION CONDITION AND RESULT

Simulation conditions are as follows. The projective image with geo-location information is provided Ministry of Land, Infrastructure and Transport, South Korea. There are two kinds of databases with or without shadows. It is assumed that the database does not have a shadow, and the camera image has a shadow. It has the 25 [cm/pix] resolution and the covered area is about 1.2 × 1.4 [km]. And the shadow

TABLE 4. Detailed Simulation Conditions.

Altimeter variance (m)	100 ~ 300
Trajectory length (m)	2,500
Map size (pixel)	4,675 × 5,642
Map size (m)	1,200 × 1,400
Resolution (m/pix)	0.25
Volume (m3)	3.36 × 108

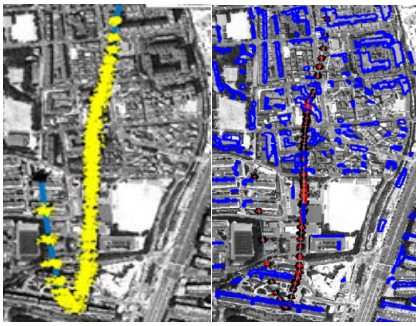


FIGURE 12. Vehicle's trajectory and the database with and without shadow.

does not exist in the database. The total simulation time is 60 [sec]. The system model is assumed as a constant velocity model. We set the constant altitude at 200 m and the initial position uncertainty is assumed as 15m standard deviation with Gaussian distribution. Also, the camera image is cropped from the database with shadow and its size is 600 × 400 [pixel].

For the performance comparison of MCL, three algorithms are used, conventional intensity-based algorithm, shadow removal and restoration, and the proposed weighting considering shadow. The intensity-based algorithm's descriptor $\xi_{intensity}(d, \hat{d})$ with user parameters are following, C = 1, N = 700, K = 10 ~ 70. And all simulation results are from 100 trials.

The simulation results are shown in Fig. 12. First, the blue contour shows the detected shadow area from the camera image and it is mosaicked to the database for the reader's understanding. Besides, the solid line shows the actual trajectory, the black points show the used particles for MCL, and the red points represent the resampled particles after the calculating importance sampling.

We illustrated Fig. 13~16 to easily compare the performance of each algorithm. Each figure shows the result of the operation by increasing K which is the number of the targets sample points. The arrows indicate the abnormal status. (The shadow removal algorithm has a lower performance than the existing algorithm)

Some particular sections can be founded with the down arrow in Fig. 13~15. In other words, when K is above 70, there is a clear tendency, but when it is below 70, anomalies can be seen. We described this event in discussion section V. The K = 70 represents the number of used target sample pairs. So we used only 140 pixels of 600 × 400 [pixel], but we were able to get more accurate results than the shadow restoration algorithm. The main navigation results of various simulations are summarized in table 5. In the case of shadow restoration

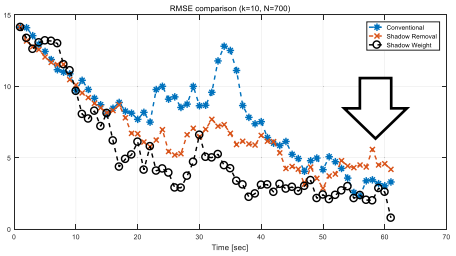


FIGURE 13. Performance comparison with various algorithms (K = 10).

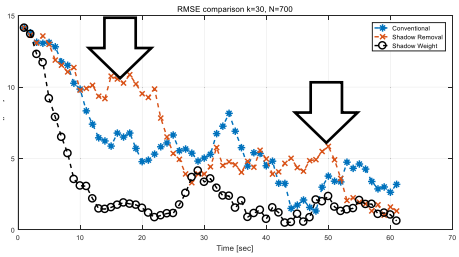


FIGURE 14. Performance comparison with various algorithms (K = 30).

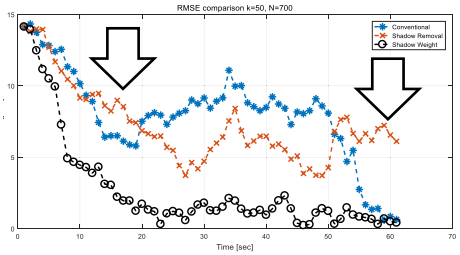


FIGURE 15. Performance comparison with various algorithms (K = 50).

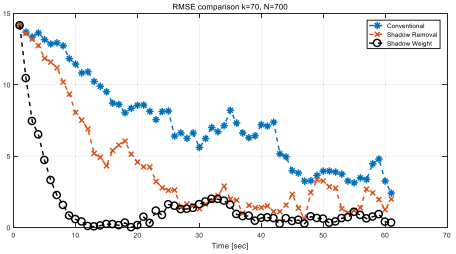


FIGURE 16. Performance comparison with various algorithms (K = 70).

applied algorithm is more accurate than the conventional algorithm.

In addition, the performance of the proposed algorithm was compared with various template-based similarity measurement algorithms to verify the algorithm. The SSD, NCC, and MI are still used as standards to compare the performance of the algorithms. In the case of the proposed algorithm, excellent navigation performance can be guaranteed as shown in Fig. 17, 18, and table 6, including the initial convergence rate. It can be interpreted that the increase in the initial convergence speed is due to discarding uncertain information caused by shadows and measuring the similarity.

TABLE 5. Summarized Navigation Result.

RMSE [m]	K = 10	K = 30	K = 50	K = 70
Conventional algorithm	7.05	8.32	8.51	8.26
Shadow restoration	8.37	7.01	7.13	7.06
Considering shadow	**6.35**	**3.91**	**2.84**	**2.33**

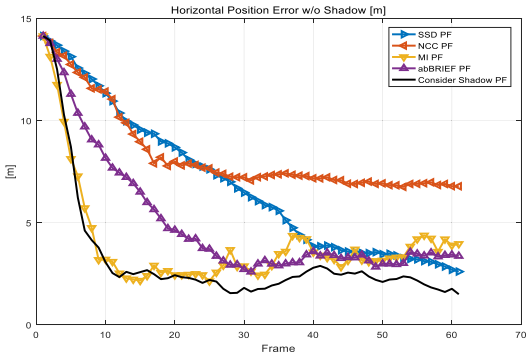


FIGURE 17. Performance comparison with various weighting algorithms (Simulation data).

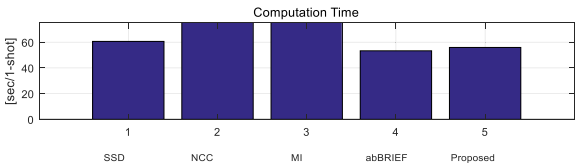


FIGURE 18. Summarized performance with various weighting algorithms.

TABLE 6. Summarized Navigation Result and Computation Time (Simulation).

	SSD	NCC	MI	abBRIEF	Proposed
RMSE [m]	12.5	9.6	5.9	8.9	**5.0**
Computation Time [sec]	60.6	16,401	6,004	**53.3**	56.0

NCC and MI require a lot of computation time that is difficult to use in practice. On the other hand, it can be confirmed that SSD, abBRIEF, and the proposed algorithm are practically available. The proposed algorithm inherits the advantages of the abBRIEF algorithm and can compensate for the performance degradation caused by shadows.

B. EXPERIMENTAL CONDITION AND RESULT

To verify the effectiveness of the proposed algorithm, we used the data obtained through flight experiments. The flight test was conducted in Hanam, Gyeonggi-do, South Korea, and the aircraft used was DJI-Mavic Pro, which is equipped with a 3-axis gimbal, it can always be directed downward. A distance of about 2km was cruising at an altitude of 100m and embedded GPS solution is used as the reference trajectory. The database was processed with Google satellite map (2016y), and continuous 40 frames extracted from the video were used as input images to prevent an increase in the

TABLE 7. Summarized Embedded Camera Specification.

Sensor	1/2.3" (CMOS)
Lens	FOV 78.8°
Video Recording	4K: 3840×2160
Sampling rate	4 Hz

FIGURE 19. DJI-Mavic pro used for experiment with gimbaled down-looking camera.

FIGURE 20. Flight experiment conducted on Hanam, Gyeonggi-do, South Korea.

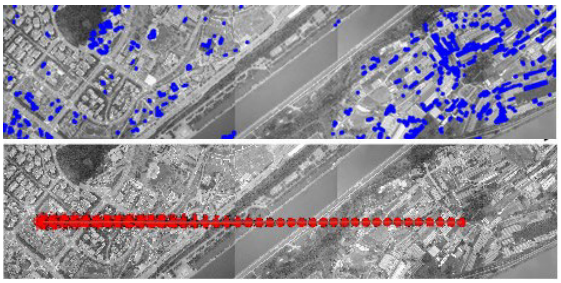


FIGURE 21. Estimated shadow and true trajectory in database.

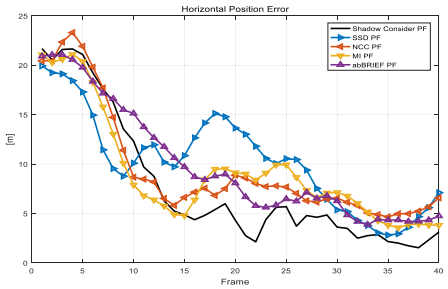


FIGURE 22. Performance comparison with various weighting algorithms (Experiment data).

computational volume. The specifications of the drone and the attached camera are summarized in Table. 7 and Fig. 19.

The detailed trajectory and the area assumed to be a shadow is shown in Fig. 20, 21. The place where the flight test was performed is indicated by a red arrow, the particles used according to the detailed trajectory are indicated by a red dot, and the shadow is indicated by a blue outline.

Simulation results using flight experiment data are shown in Fig. 22 and Table 7. Although the simulation results do not show dramatic results, the tendency of the result is the same as the simulation.

TABLE 8. Summarized Navigation Result (Experiment).

	SSD	NCC	MI	abBRIEF	Proposed
RMSE [m]	11.3	11.0	10.6	11.4	**10.2**

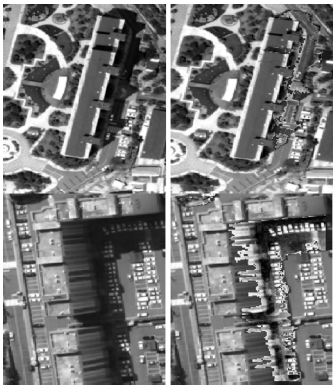


FIGURE 23. Well-restored case (20 ∼ 50 [sec]).

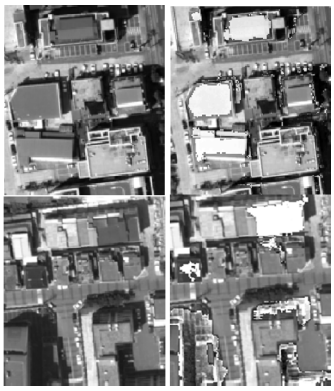


FIGURE 24. Wrong place or over-restored case (10 ∼ 20, 50 ∼ 60 [sec]).

The SSD and NCC algorithms have lower performance and the MI result is good, but difficult to use considering the amount of computation. As a result, we can see that the proposed algorithm is the best in computational time and performance. This can be expected that if the proposed algorithm is applied, it will be suitable for use as an alternative navigation algorithm even when GNSS is not available.

V. DISCUSSION

A. SHADOW RESTORATION ERROR

We will revisit the adverse effect of shadow removal and reconstruction algorithm. Fig. 23 shows an example of a camera image when the performance of the MCL is improved through the shadow detection and restoration process. If the shadow area is well restored as shown in the figure, the similarity between the two images can be expressed as well as the real environment. As a result, more accurate navigation solutions can be estimated compared to conventional algorithms. On the other hand, Fig. 24 shows examples showing why the performance was degraded compared to the conventional algorithm even though the shadow detection

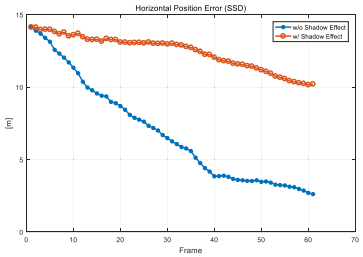


FIGURE 25. Horizontal position error with and without shadow effect (SSD).

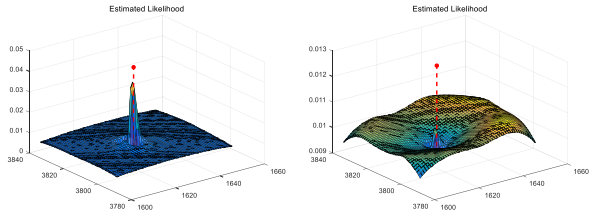


FIGURE 26. Likelihood and true position with(R) and without(L) shadow effect (SSD).

and restoration process was performed. The excessive image restoration causes a large difference from the database which is called the adverse effect.

The conventional algorithm has a fast computational speed using binary descriptor. However, an intensity-based matching algorithm has poor consistency of the similarity value when using shadowed target samples. On the other hand, the shadow removal and restoration algorithm is possible to reduce the difference between the database and camera images. And a slight performance improvement compared to the conventional algorithm. However, the limitations are clear due to the restoration of the wrong place or excessive restoration.

B. SHADOW EFFECT AND LIKELIHOOD WITH VARIOUS ALGORITHMS

The navigation performance degradation and likelihood caused by the shadow are expressed with a true position. The likelihood of SSD is well expressed only in the ideal case. However, it is shown that the likelihood is not well expressed in the shadows. This could be found through Fig. 25 and 26. The navigation accuracy also differs greatly, and it can be seen that it is not good for expressing the similarity of the templates.

NCC is characterized by being the least affected by shadows. This can be seen by the fact that the two likelihoods are very similar in Fig. 28. However, it is difficult to improve accuracy because of the large variance that contains the information. Navigation accuracy is also shown in Fig. 27. It has too enormous error as well, so it is not suitable for use as a cost of MCL.

In the case of MI, as it uses a relatively recent algorithm, simulation result shows fast convergence and excellent navigation accuracy. Also, it belongs to the group with less performance degradation due to the shadow effect as shown

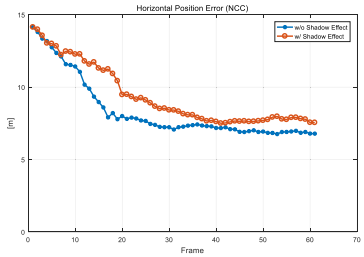


FIGURE 27. Horizontal position error with and without shadow effect (NCC).

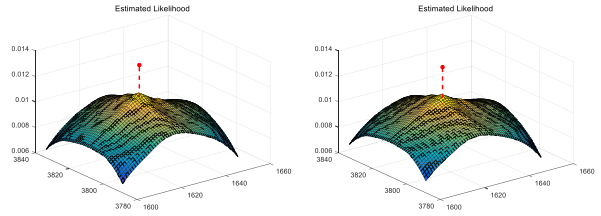


FIGURE 28. Likelihood and true position with(R) and without(L) shadow effect (NCC).

FIGURE 29. Horizontal position error with and without shadow effect (MI).

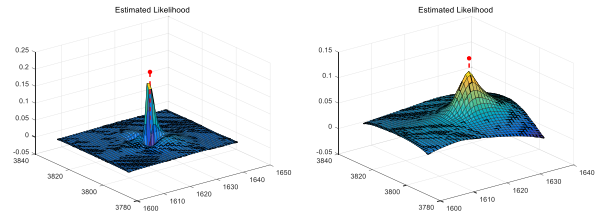


FIGURE 30. Likelihood and true position with(R) and without(L) shadow effect (MI).

in Fig. 28. This can be said to be a robust algorithm for the shadow effect. We can confirm that likelihood does not change significantly, such as Fig. 30.

In the case of calculating the similarity of abBRIEF, in the ideal case without shadows, it doesn't cause a significant problem, but we can see that the performance degradation caused by the shadows occurs significantly. Fig. 31 shows the reduction in navigation accuracy, Fig. 32 denotes the change in likelihood caused by the shadow.

In the case of the proposed algorithm, it is designed to alleviate the effect of shadows while retaining the high speed. It is shown that a certain performance is guaranteed regardless of the presence or absence of shadows. This can be deduced

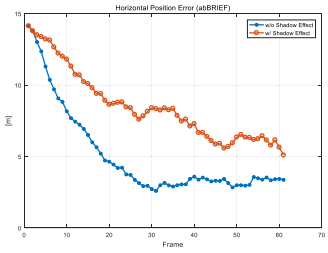


FIGURE 31. Horizontal position error with and without shadow effect (abBRIEF).

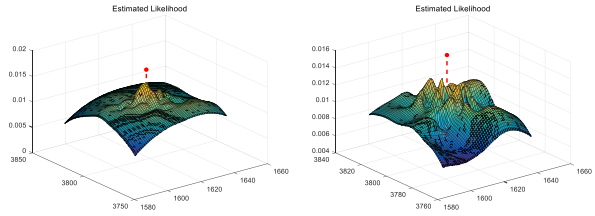


FIGURE 32. Likelihood and true position with(R) and without(L) shadow effect (abBRIEF).

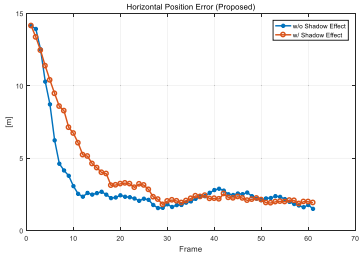


FIGURE 33. Horizontal position error with and without shadow effect (Proposed).

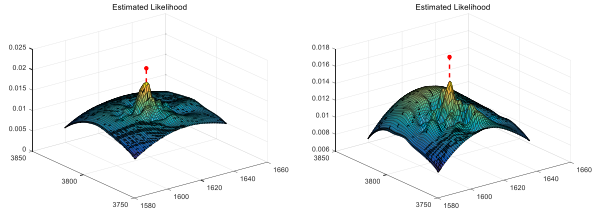


FIGURE 34. Likelihood and true position with(R) and without(L) shadow effect (Proposed).

through Fig. 33. Also, the likelihood is distorted, but the change to the peak value is not large as shown in Fig. 34. This can be interpreted that the proposed algorithm is suitable for use in shaded environments.

The summarized results are shown in table 8 and 9. SSD is fast and express likelihood well in ideal situations, but with real shadows, the performance degradation is very significant. In the case of NCC, the change of likelihood is small, but the speed is slow and the navigation accuracy is poor as well.

MI expresses likelihood well and has excellent performance because it is robust to shadows, but it is difficult to use in practice due to the disadvantage that it takes a lot of calculation time. In particular, abBRIEF algorithm using image

TABLE 9. Summarized Navigation Result and Computation Time with Intensity- based Weighting Algorithms.

	SSD	NCC	MI	abBRIEF	Proposed
Sim. w/o Shadow RMSE [m]	7.7	8.6	4.7	5.9	**4.2**
Sim. w/ Shadow RMSE [m]	12.5	9.6	5.9	8.9	**5.0**
Exp. RMSE [m]	11.3	11.0	10.6	11.4	**10.2**
Computational time [sec]	60.6	16,401	6,004	**53.3**	56.0

TABLE 10. Performance Grade of Navigation Result and Computation Time with Intensity-based Weighting Algorithms.

	SSD	NCC	MI	abBRIEF	Proposed
Accuracy	-	-	**Good**	Bad	**Good**
Robustness	Bad	**Good**	-	-	**Good**
Speed	-	Bad	-	**Good**	**Good**

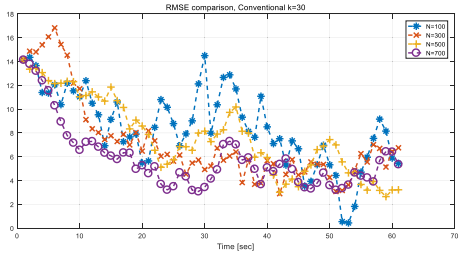


FIGURE 35. Horizontal RMSE with various N = [100, 300, 500, 700] (Conventional).

descriptor, it is possible to express excellent likelihood in ideal case. However, the disadvantage is that the performance is greatly degraded by the shadow. On the other hand, the proposed algorithm properly expresses likelihood in any situation, so that excellent navigation performance can be always guaranteed, and the computational amount is also very good.

C. CHARACTERISTICS OF THE PROPOSED ALGORITHM

Filter performance is depending on N, which is the number of particles used in the MCL, and K, which is the number of sample targets for the similarity test. Therefore, this article completed various performance comparisons using various K and N. Fig. 35∼37 show the characteristic of the algorithm, along with various N.

Fig. 35 shows the horizontal position root mean squared error (RMSE) of the conventional algorithm with Monte-Carlo simulation 100 trials for various N with shadow. As shown in the figure, we can find a slight improvement of the filter's performance as the increasing N, which is the user parameter of MCL. However, it can be confirmed that the position error does not largely converge due to the shadow, which indicates whether the shadow in the camera image is vulnerable to estimating the position of the robot.

The result of applying the method of detecting and restoring shadows is depicted in Fig. 36. It can be confirmed that the convergence of the position error is slightly improved compared with the conventional algorithm. However, we can see

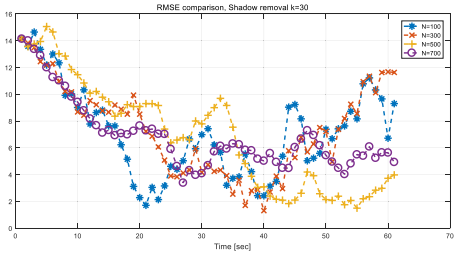


FIGURE 36. Horizontal RMSE with various N = [100, 300, 500, 700] (Shadow removal and restoration).

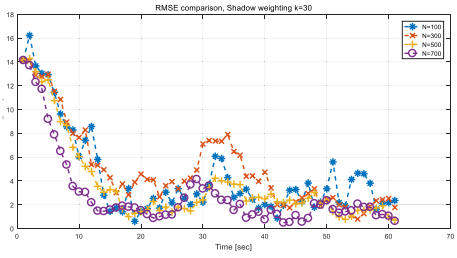


FIGURE 37. Horizontal RMSE with various N = [100, 300, 500, 700] (Weighting considering shadow).

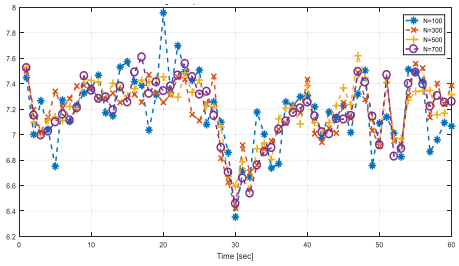


FIGURE 38. Normalize the number of target samples located in the shadow with increasing N.

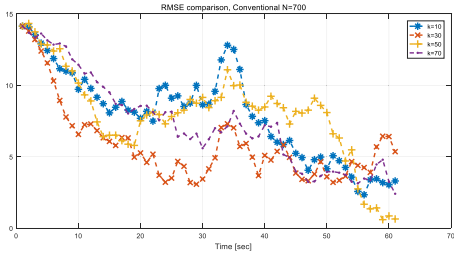


FIGURE 39. Horizontal RMSE with various K = [10, 30, 50, 70] (Conventional).

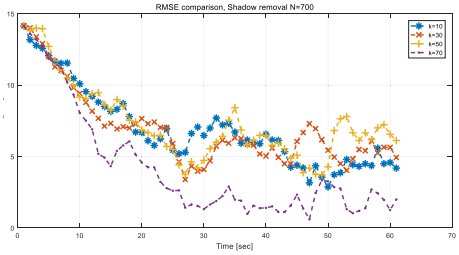


FIGURE 40. Horizontal RMSE with various K = [10, 30, 50, 70] (Shadow removal).

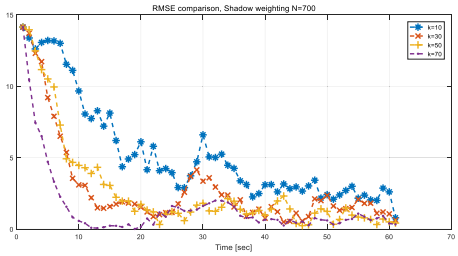


FIGURE 41. Horizontal RMSE with various K = [10, 30, 50, 70] (Weighting considering shadow).

the limitation that the filter's performance does not increase significantly as N increases.

Fig. 37 shows how the proposed algorithm enables effective position estimation. We can find the fastest convergence rate of the filter, and we can confirm that it is possible to operate and estimate the correct solution.

The normalized number of target samples located in shadow is shown in Fig. 38. In other words, the number of used sample targets was divided by the total number of particles and the number of sample targets located in shadow. This normalized number makes the performance comparison intuitively. We can find that there is not a significant performance improvement according to N. That is, we can conclude that N is already saturated. The user does not have to significantly increase the number of N in MCL to mitigate the shadow effect. With this, the proposed algorithm can be expected to be very effective in terms of computational complexity. Next, the filter performance obtained by fixing N and increasing K to 700 is shown in Fig. 39∼41.

Fig. 39 shows the horizontal position accuracy of the conventional algorithm as RMSE. It can be seen that increasing K has a similar trend as increasing N. However, by increasing K, it can be confirmed once again that the running accuracy is not always improved.

The position accuracy obtained through MCL after detecting and restoring the area is illustrated in Fig. 40. Unlike the conventional algorithm, we can see that the filter tends to converge well, unlike when increasing N. Especially, when K = 70, it can be confirmed that the convergence of the filter is greatly improved. Here we can infer that the efficient K is 70. This is not an absolute value. The user should consider the ratio of the total number of pixels in the image.

Finally, the localization result is shown in Fig. 41 when the proposed information allocation logic is applied. Unlike all other cases, we can find that the proposed algorithm not only has a consistent trend but also can make a very accurate navigation solution as K increases. The increase in errors occurring in about 30 seconds can also be greatly reduced.

Fig. 42 show the normalized the number of target sample located in the shadow with increasing K. As a result, our team found that the number of normalized target samples located in the shadows increased as K increased. It can be expected

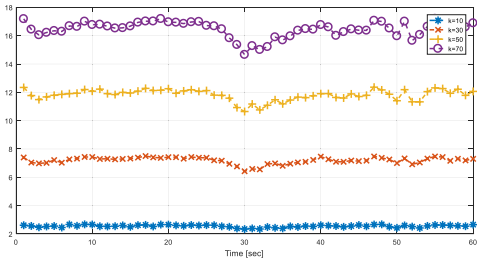


FIGURE 42. Normalized the number of target sample located in the shadow with increasing K.

that the number of target samples is more sensitive than the number of particles.

VI. CONCLUSION

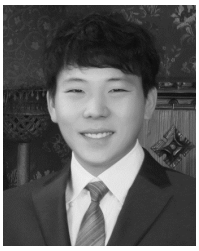
In the case of the shadow restoration algorithm, it has been confirmed that the image is excessively restored, which means that accuracy cannot always be guaranteed. And the conventional similarity test algorithms have at least one of three problems. It has too high a computational burden, too low accuracy, or largely affected by the shadow effect.

We have proposed an algorithm that can mitigate the effects of shadows. The algorithm is designed taking into account the real environment with shadows. The proposed system inherits the intensity-based MCL algorithm and is characterized by distributing the amount of information to effectively filter out contaminated information. The proposed algorithm is designed to be able to set the importance according to the proportion of the target sample located in the shadow. In addition, the activation function is used to effectively express the importance of information.

Simulations have been performed for comparison with various similarity functions and shadow restoration algorithms, and it can be concluded that the accuracy and computation time are suitable for practical use. In addition, through simulation and flight test results, we have confirmed that significant performance improvements have been achieved.

REFERENCES

[1] N. Gageik, P. Benz, and S. Montenegro, "Obstacle detection and collision avoidance for a UAV with complementary low-cost sensors," *IEEE Access*, vol. 3, pp. 599–609, 2015.
[2] Y. Wu, Y. Sui, and G. Wang, "Vision-based real-time aerial object localization and tracking for UAV sensing system," *IEEE Access*, vol. 5, pp. 23969–23978, 2017.
[3] A. Angrisano, C. Gioia, S. Gaglione, and G. del Core, "GNSS reliability testing in signal-degraded scenario," *Int. J. Navigat. Observ.*, vol. 2013, pp. 1–12, Jan. 2013.
[4] N. A. White, P. S. Maybeck, and S. L. DeVilbiss, "Detection of interference/jamming and spoofing in a DGPS-aided inertial system," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, no. 4, pp. 1208–1217, Oct. 1988.
[5] B. J. Kuipers and T. S. Levitt, "Navigation and mapping in large scale space," *AI Mag.*, vol. 9, no. 2, p. 25, 1988.
[6] S. D. Fleischer, "Bounded-error vision-based navigation of autonomous underwater vehicles," Ph.D. dissertation, Dept. Aeronaut. Astronaut., Stanford Univ., Stanford, CA, USA, 2000.
[7] A. Koch, H. Wittig, and F. Thielecke, "A vision-based navigation algorithm for a VTOL-UAV," in *Proc. AIAA Guid., Navigat., Control Conf. Exhib.*, 2006, p. 6546.
[8] A. Cesetti, E. Frontoni, A. Mancini, A. Ascani, P. Zingaretti, and S. Longhi, "A visual global positioning system for unmanned aerial vehicles used in photogrammetric applications," *J. Intell. Robotic Syst.*, vol. 61, nos. 1–4, pp. 157–168, Jan. 2011.
[9] G. Park, B. Lee, D. G. Kim, Y. J. Lee, and S. Sung, "Design and performance validation of integrated navigation system based on geometric range measurements and GIS map for Urban aerial navigation," *Int. J. Control, Automat. Syst.*, vol. 18, pp. 2509–2521, Aug. 2020.
[10] J. P. Golden, "Terrain contour matching (TERCOM): A cruise missile guidance aid," *Proc. SPIE*, vol. 0238, Dec. 1980, Art. no. 959127.
[11] W. Youn, H. Ko, H. Choi, I. Choi, J.-H. Baek, and H. Myung, "Collision-free autonomous navigation of a small UAV using low-cost sensors in GPS-denied environments," *Int. J. Control, Autom. Syst.*, vol. 18, pp. 1–16, 2020.
[12] Y. Bresler and S. J. Merhav, "On-line vehicle motion estimation from visual terrain information part II: Ground velocity and position estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, no. 5, pp. 588–604, Sep. 1986.
[13] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image Vis. Comput.*, vol. 21, no. 11, pp. 977–1000, Oct. 2003.
[14] L. Yang, G. Yang, Y. Yin, and R. Xiao, "Sliding window-based region of interest extraction for finger vein images," *Sensors*, vol. 13, no. 3, pp. 3799–3815, Mar. 2013.
[15] Z. Liu, J. An, and Y. Jing, "A simple and robust feature point matching algorithm based on restricted spatial order constraints for aerial image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 514–527, Feb. 2012.
[16] Q. Shan, C. Wu, B. Curless, Y. Furukawa, C. Hernandez, and S. M. Seitz, "Accurate geo-registration by Ground-to-Aerial image matching," in *Proc. 2nd Int. Conf. 3D Vis.*, Dec. 2014, pp. 525–532.
[17] N. Mo, R. Zhu, L. Yan, and Z. Zhao, "Deshadowing of urban airborne imagery based on object-oriented automatic shadow detection and regional matching compensation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 2, pp. 585–605, Feb. 2018.
[18] Y. Liang, B. Pan, X. Guo, and Y. Liang, "Image shadow detection and removal method based on LAB color space," *Comput. Modernization*, vol. 10, p. 88, Oct. 2019.
[19] J. Ding and G. Zhao, "Gray sequence image shadow removal algorithm based on computer vision," in *Proc. 12th Int. Conf. Measuring Technol. Mechatronics Autom. (ICMTMA)*, Feb. 2020, pp. 366–370.
[20] J. Ma, H. Zhou, J. Zhao, Y. Gao, J. Jiang, and J. Tian, "Robust feature matching for remote sensing image registration via locally linear transforming," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6469–6481, Dec. 2015.
[21] C. Lee and D. Kim, "Visual homing navigation with Haar-like features in the snapshot," *IEEE Access*, vol. 6, pp. 33666–33681, 2018.
[22] J. Ma and J. Zhao, "Robust topological navigation via convolutional neural network feature and sharpness measure," *IEEE Access*, vol. 5, pp. 20707–20715, 2017.
[23] T. Naseer, B. Suger, M. Ruhnke, and W. Burgard, "Vision-based Markov localization for long-term autonomy," *Robot. Auto. Syst.*, vol. 89, pp. 147–157, Mar. 2017.
[24] M. Mantelli, D. Pittol, R. Neuland, A. Ribacki, R. Maffei, V. Jorge, E. Prestes, and M. Kolberg, "A novel measurement model based on abBRIEF for global localization of a UAV over satellite images," *Robot. Auto. Syst.*, vol. 112, pp. 304–319, Feb. 2019.
[25] Y. Ye, L. Shen, M. Hao, J. Wang, and Z. Xu, "Robust Optical-to-SAR image matching based on shape properties," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 4, pp. 564–568, Apr. 2017, doi: 10.1109/LGRS.2017.2660067.
[26] A. Sedaghat and N. Mohammadi, "Evaluation of similarity measures for template matching," *J. Geomatics Sci. and Technol.*, vol. 9, no. 1, pp. 189–206, 2019.
[27] X. Li, Y. Hu, T. Shen, S. Zhang, J. Cao, and Q. Hao, "A comparative study of several template matching algorithms oriented to visual navigation," *Proc. SPIE*, vol. 11550, Oct. 2020, Art. no. 115500E.
[28] B. D. de Vos, B. H. van der Velden, J. Sander, K. G. Gilhuijs, M. Staring, and I. Išgum, "Mutual information for unsupervised deep learning image registration," *Proc. SPIE*, vol. 11313, Mar. 2020, Art. no. 113130R,
[29] C. Unsalan and K. L. Boyer, "Linearized vegetation indices based on a formal statistical framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 7, pp. 1575–1585, Jul. 2004.
[30] *MATLAB and Image Processing Toolbox Release 2016b*, MathWorks, Natick, MA, USA.



SUNG HYUK CHOI received the B.S. degree in mechanical and aerospace engineering from Sejong University, in 2014. He is currently pursuing the Ph.D. degree with the Department of Mechanical and Aerospace Engineering, Seoul National University, South Korea. His current research interests include high-precision inertial navigation, attitude determination for CubeSat, fault detection logic, and vision-based navigation for an aerial vehicle.

CHAN GOOK PARK (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in control and instrumentation engineering from Seoul National University, Seoul, South Korea, in 1985, 1987, and 1993, respectively. He has worked as a Postdoctoral Fellow with Prof. Jason L. Speyer on peak seeking control for formation flight at the University of California at Los Angeles (UCLA), Los Angeles, in 1998. From 1994 to 2003, he was an Associate Professor with Kwangwoon University, Seoul. In 2003, he joined the Faculty of the School of Mechanical and Aerospace Engineering, Seoul National University, where he is currently a Professor. From 2009 to 2010, he was a Visiting Scholar with the Department of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, USA. His current research interests include advanced filtering techniques, high-precision INS, GPS/INS integration, MEMS-based pedestrian dead reckoning, and visual-inertial navigation. He has served as the Chair for the IEEE AES Korea Chapter until 2009.