# OBJECT DETECTION UNDER MOVING CLOUD SHADOWS IN WAMI

P. U. Hatipoglu[1]*, R. T. Albayrak[1], A. A. Alatan[2]

[1] ESEN System Integration, 06800 METU Technopolis Ankara, Turkey - (poyraz.hatipoglu, tufan.albayrak)@esensi.com.tr
[2] Dept. of Electrical & Electronics Engineering, METU, 06800, Ankara, Turkey - alatan@metu.edu.tr

**KEY WORDS:** Moving Cast Shadow Detection, Cloud Shadow, Background Subtraction, Moving Object Detection, Gaussian Mixture Models, Video Surveillance, Wide Area Motion Imagery

**ABSTRACT:**

For a reliable and robust moving object detection, the subtraction of a precisely modeled background is crucial in wide-area motion imagery (WAMI). Even the most successful background subtraction algorithms that are designed to model highly-dynamic environments cannot cope with rapidly changing scenery, such as moving cloud shadows, which has different characteristics from dynamic textures. This paper presents a novel method to detect moving objects and to eliminate false alarms under moving cloud shadow regions in gray-level video sequences. The proposed method uses the relation between reflectance values of the shadowed and well-illuminated sequences of the regions in the video frame. A modified adaptive region growing approach, which extends from seed points, is designed to obtain the moving parts of the cloud shadows without presuming the geometric structure of the clouds. In order to determine the moving border of the cloud shadows, where false alarms typically occur, the cloud shadow motion should be detected. As the last stage of the proposed method, real moving objects in the scene are tried to be discriminated from false alarms by exploiting the relation of intensity ratios between the object candidate and its surroundings. The accuracy and computational efficiency of the proposed approach make it a reliable and feasible approach to be used in real-time surveillance solutions.

## 1. INTRODUCTION

Moving object detection and tracking are constantly developing active research areas in remote sensing and computer vision. WAMI is one of the most common wide-area surveillance data sources and has drawn attention in the last couple of decades. With the development in both imaging technology and unmanned aerial vehicle platforms, the attention to the fully automatic and real-time WAMI tracking systems have increased. WAMI solutions can be integrated into various platforms such as unmanned aerial vehicles (Lin, Medioni, 2007), aerostats (Nagendran et al., 2010), etc. and for numerous civil and military applications.

Targets whose resolution is considerably low are tried to be detected and be tracked in very large-scale WAMI videos. Even there are a few multi-spectral solutions, most of the WAMI solutions use monochromatic (AFRL, 2009), (Force, n.d.), (Perera et al., 2006) imaging format depending on the application. Without the color information, there is a chance in that the intensity values of both target object and background can be quite similar or even the same. Hence detection and tracking of targets in a monochromatic solution can be challenging even if the data is captured in favorable weather and illumination conditions. Moreover, due to the negative effects of atmosphere related distortions, the object boundaries can be seen unclear or even completely blended to the background. Since a WAMI solution can monitor around tens of km2 region with hundreds of mega-pixel frame resolution, reducing the false alarm rate is quite critical. The reliability and usability of the product directly depend on both the detection ability of the targets and the accuracy of the detection. Hence any major false alarm sources are needed to be attacked to achieve a more robust and reliable solution.

Reliable background subtraction is the key operation to obtain moving foreground objects in the scene with high precision. To obtain an initial estimate and extract information of non-stationary objects numerous background subtraction methods with different working mechanisms have proposed (Piccardi, 2004), (Bouwmans et al., 2017), (Zivkovic, 2004). The precise modeling and constant updating of the model of background is the initial step of robust tracking performance (Sommer et al., 2016). Not only the discriminating power of moving objects from the background but also reducing the false alarm rate is the other requested ability of successful background subtraction technique. Since WAMI solutions try to monitor a large-scale area persistently, the preferred background subtraction technique needs to be work even for a highly dynamic environment. As waving tree branches (Elgammal et al., 2000), optical turbulence deformations (Oreifej et al., 2012), stabilization related defects, and illumination changes (Pilet et al., 2008), fast-moving cast shadows is also one of the major challenges that background subtraction method needs to deal with by using its adaptivity property. At the same time, the preferred background detection algorithm needs to be computationally efficient to work operationally in real-time solutions.

Even spatially strengthened versions of the Gaussian mixture model (Sommer et al., 2016), (Reilly et al., 2010) cannot cope with fast-changing stationary signals, such as moving cloud shadows. To prevent the generation of false objects (false alarms) caused by the motion of cloud shadows, we need to identify the moving section of cloud shadows. During the last decade, various cloud shadow detection methods (Li et al., 2017), (Zhu, Woodcock, 2012), (Simpson, Stitt, 1998) have been designed and implemented to improve the performance of the different applications, such as feature extraction, segmentation, classification (Li et al., 2017). However, nearly all the current cloud detection algorithms have used either multispectral information (Luo et al., 2008), (Simpson, Stitt, 1998) or the geometrical properties of the cloud and orientation of imaging system (Braaten et al., 2015), (Huang et al., 2010).

---

* Corresponding author

In our study, the moving part of the cast shadow of the clouds has tried to be detected by using low fps-rate monochromatic large scale video sequences without using any prior location information of camera and cloud regions. The main contribution of this study is that moving parts of the cloud regions have been identified using adaptive double thresholding methods and the false alarms generated by the cast cloud shadow have been eliminated. Furthermore, the proposed algorithm is quite fast, efficient and generalized to work in different weather and seasonal conditions with minimal assumptions.

The paper of our study is organized as follows: The first section is designed to reveal the motivation and aim of this study. Section 2 introduces the literature of the cast shadow concept and the shadow detection algorithms using gray-level video sequences. In Section 3 the proposed method with all subsections is presented. In the first subsection the system overview, in the second one the assumptions of the proposed algorithm, in the third subsection the methodology to find the moving parts of the cloud shadow, and in the last subsection, the elimination method of false alarms generated by the cloud cast shadow is clarified. The datasets, experiments and performance results are described with the evaluation criteria in Section 4. Finally, in the last section, the study is concluded with discussions.

## 2. RELATED WORKS

Shadows occur when a light source is occluded by an object partially or completely. As shadow can be divided into two categories which are self-shadow and cast shadow (Stander et al., 1999). Cloud shadow in the outdoor or aerial data can be described as a cast shadow, which is a terminology used for the shadows which are generated by an object and projected to another object at the scene.

### 2.1 Cast Shadow

If the direct light is blocked by object completely, that section of the cast shadow is classified as umbra whilst if a light source is blocked partially, the darkening region of the shadow is called the penumbra (Stander et al., 1999).

The opacity of the occluding object and both the location and the geometry between the light source and occluding object determine the penumbra region of the cast shadow. The luminance transition in the penumbra region of the cast shadow can be assumed as linear for an opaque and solid occluding object. (Stander et al., 1999). However, due to having non-uniform density and random 3D geometry, even if we can assume that the luminance of the cloud shadows rises from the inside to the outside, the structure of the penumbra regions of the cloud shadow cannot be represented mathematically.

The intensity (brightness) value of a point $(x, y)$ which is the 2D projection of the object surface at point $q$ and time instant $t$, can be expressed as:

$$C_t(x,y) = k_{c,t}(x,y) \cdot E_t(x,y), \qquad (1)$$

where 
$C_t$ : intensity value of a pixel at time instant $t$,
$x, y$ : image coordinate of the object surface point $q$,
$k_{c,t}$ : camera gain at time instant $t$,
$E_t$ : luminance at time instant $t$.

By using the reflection model, the luminance at time instant $t$ can be modeled as:

$$E_t(x,y) = \rho_t(x,y) \cdot S_t(x,y), \qquad (2)$$

where 
$\rho_t$ : the reflectance of the object at point $q$ and time $t$,
$S_t$ : irradiance at time $t$.

Depending on the illumination conditions, in the (Stander et al., 1999) the reflectance $S_t$ has been represented as:

$$S(x,y) = \begin{cases} c_P \cdot cos(\theta) + c_A & \text{, no shadow} \\ c_P \cdot cos(\theta) \cdot k(x,y) + c_A & \text{, penumbra} \\ c_A & \text{, umbra.} \end{cases} \qquad (3)$$

According to Lambert's cosine law (Basri, Jacobs, 2003), the angle $\theta$ between the direction of the incident light and the surface normal defines the contribution amount of the direct light source to the irradiance of the surface at point $q$. Depending on the light transition of the penumbra region of cast shadow, the $k(x,y)$ value varies in the range of $[0, 1]$.

### 2.2 Shadow Detection Methods

According to a taxonomy (Al-Najdawi et al., 2012), the shadow detection algorithms can be clustered concerning their dependencies on objects and the environment. Moreover, the number of spectral bands used and the implementation domains have also used to categories the shadow detection algorithms.

Since the proposed algorithm in this study has been designed without concerning the special geometry of the occluding object that forms the cast shadow, the object independent and monochromatic pixel-based methods from the literature have been investigated (Stander et al., 1999), (Toth et al., 2004), (Vargas et al., 2010), (Chien et al., 2002), (Xu et al., 2004), (Jung, 2009), (Jung, 2009), (Jacques et al., 2005), and (Lin et al., 2010).

The reflection model presented in (Stander et al., 1999) has formed the basis of many methods (Toth et al., 2004), (Lu et al., 2006), (Vargas et al., 2010). The idea is the usage of the ratio between the pixel intensity values in the current frame (collected $u$ seconds later than the reference) and that in the reference frame as shown in (4).

$$\xi_{t+u,t}(x,y) = \frac{C_{t+u}(x,y)}{C_t(x,y)}. \qquad (4)$$

According to (1) and (2), (4) can be expanded as:

$$\xi_{t+u,t}(x,y) = \frac{k_{c,t+u}(x,y) \cdot \rho_{t+u}(x,y) \cdot S_{t+u}(x,y)}{k_{c,t}(x,y) \cdot \rho_t(x,y) \cdot S_t(x,y)}. \qquad (5)$$

Since the reflectance value of the background region does not change in time ($\rho_{t+u}(x,y) = \rho_t(x,y)$) and one can control the gain value, $k_c$, of the camera we can simplify (5) as:

$$\xi_{t+u,t}(x,y) = \frac{S_{t+u}(x,y)}{S_t(x,y)}. \qquad (6)$$

Since the umbra and penumbra regions of the shadow have different illumination characteristics, when shadow-free background region in the background frame is covered by moving cloud regions in the current frame the ratio of irradiances ($\xi'_{t+u,t}$) should be calculated as follows (Al-Najdawi et al., 2012):

$$\xi'_{t+u,t}(x,y) = \begin{cases} \frac{c_A}{c_P \cdot cos(\theta)+c_A} & \text{, no shadow to umbra} \\ \frac{c_P \cdot cos(\theta) \cdot k(x,y)+c_A}{c_P \cdot cos(\theta)+c_A} & \text{, no shadow to penumbra.} \end{cases} \qquad (7)$$

In (Stander et al., 1999) the ratio (6) is used to detect the shadow regions. However, the study assumes that the intensity values of the background in a defined neighborhood remain constant. This assumption cannot hold for the very complex background environment visualized in wide area surveillance. In fact, even in indoor environments, it is quite hard to rely on this assumption. Furthermore, this study also assumes that the object occludes the direct light source is opaque and hence, the intensity change in the penumbra field of the object shadow is approximately linear. However, due to the unique random structural density of each cloud bank, the penumbra region of the shadows might show a unique transition property.

In one of the studies, the author (Toth et al., 2004) calculates the $\xi_{t+u,t}$ values for foreground objects by taking the average of the ratio (6) over sliding window pixels. Then Gaussian white noise is added to the $\xi_{t+u,t}$ values to test the stability of the designed method. By using the shadow-free background and calculated $\xi_{t+u,t}$ value the foreground image is tried to be estimated. The major contribution of the study (Toth et al., 2004) is that a significance test is derived to extract the shadow regions. However, the algorithm ignores the penumbra region of the shadow by stating that the penumbra region is very small and sometimes not recognizable. According to (Stander et al., 1999), the statement cannot be valid unless the distance between the occluding object and background is negligible compared to the distance between the light source and occluding object. Moreover, the occluding object must be opaque to confirm the statement of (Toth et al., 2004). Since cloud regions do not comply with these two assumptions, the developed approach cannot be used as a moving cloud shadow detector.

In a different approach, the author (Jacques et al., 2005) uses the normalized cross-correlation (NCC) statistic between the background pixels and the foreground pixels in a close neighborhood. The NCC metric can produce reliable scores for the umbra regions, since the NCC score is not affected by the multiplication of each pixel with a positive constant value. However, the intensity change ratio is not the same for every pixel in a neighborhood of a penumbra region due to the variable $k(x,y)$ value as shown in (7). Hence, the performance of the shadow detection algorithm is quite poor for the penumbra regions as stated in (Al-Najdawi et al., 2012).

According to (Al-Najdawi et al., 2012), the algorithms developed by (Xu et al., 2004), and (Chien et al., 2002) are applicable for just specific indoor environments. The algorithm of (Jung, 2009) is too complicated to work in real-time applications and highly parameter-dependent.

As mentioned earlier, since it is not known whether the randomly selected reference frame has cloud shadow regions or not, the irradiance ratio calculation for newly-illuminated areas should also be one of the main concerns of the proposed study. The irradiance ratio calculation for the first-shadowed-then-well-illuminated areas (FSTI) should be calculated as demonstrated in (8).

$$\xi''_{t+u,t}(x,y) = \begin{cases} \frac{c_P \cdot cos(\theta)+c_A}{c_A} & \text{, umbra to no shadow} \\ \frac{c_P \cdot cos(\theta)+c_A}{c_P \cdot cos(\theta) \cdot k(x,y)+c_A} & \text{, penumbra to no shadow} \end{cases} \qquad (8)$$

## 3. PROPOSED METHOD

### 3.1 System Overview

A moving object detection under moving cloud shadow algorithm is proposed as two main sub-procedures being linked to each other. The functionality of the first main block is detecting moving cloud shadow regions and its moving border regions. Next, the moving border masks is used to filter out the foreground objects located under the moving border regions. The filtered foreground includes both real target objects and the false alarms generated by abrupt intensity change caused by fast-moving cast cloud shadows. The decision for the elimination of possible false alarms is performed in the latter main block. The overall system is shown in Figure 1 and the details of system sub-blocks can be viewed in Figure 6 and Figure 7.
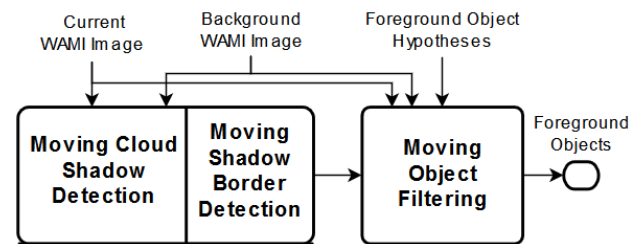


Figure 1. Proposed Method Overall Chart

### 3.2 Assumptions

In the proposed method, two major assumptions are made to detect moving parts of the cloud shadows. As in (Sexton, Zhang, 1993) and most of the reflectance ratio related studies, the intensity of the direct light source, $c_P$, is assumed to be high compared to the ambient light source, $c_A$. If it is not the case, the shadow regions cannot be differentiated properly, but since the background subtraction algorithm can suppress the small intensity changes, the object detection, and tracking algorithm will not be affected. The second assumption is that the camera is static and the WAMI video frames can be registered to the reference (background) frame with negligible error. Without this latter assumption, pixel-wise or region-wise temporal information cannot be exploited.

In a video sequence, the changes due to a shadow can be analyzed by computing the ratio of intensity values in the current frame with the intensity values in the reference frame. If the reference frame could be selected among cloud shadow-free frames, the proposed algorithm would work as a cloud shadow detection algorithm instead of a detection algorithm

for the moving part of the cloud shadow. Since the reference frame can have cloud shadows, one can analyze the first-well-illuminated-then-shadowed (FITS) and FSTI regions only to find the moving parts of the cloud shadows. In other words, there is no chance to detect the stationary parts of the cloud shadows without having the prior information of the shadow map of the reference frame.

### 3.3 Moving Cloud Detection Algorithm

Since in addition to the stable background regions, moving objects also can be covered by the cloud shadows either on the reference or current image, the reflectance ratio calculated using (7) or (8) might yield discontinuities even for the umbra regions of cloud shadows. To get rid of the discontinuities, a smoothing operation should be performed. Since the WAMI solutions are designed to work as a real-time application, the complexity reduction is always one of the key criteria at each stage of the algorithm. Hence in the proposed method, a downscaling operation has been applied for both the reference and the current images to reduce computational complexity and to get rid of the discontinuities caused by moving objects under the shadow region. In this presented approach, 10 times downscaling operation was applied for each dimension of the video sequences.

In this study, we have exploited a few general properties of the cloud shadows. One of the major advantages of dealing with the cloud shadows is that the cloud shadows usually cover reasonable large areas and cannot be vanished by downscaling operation. Experimentally it is observed that tiny cloud banks create a slight intensity change in their shadow regions, sometimes they cannot create even any change. Hence in the proposed algorithm after obtaining the quotient (reflectance ratio) image by dividing the current downscaled image to the reference downscaled image large spatially connected regions are searched which satisfy the desired reflectance ratio. To form masks for FSTI and the FITS regions similar adaptive-thresholding operations are applied for both such regions independently as explained in subsection 3.3.1.

#### 3.3.1 Adaptive Thresholding on Quotient Image:

In this step, it is desired to build a time-efficient and generalized method for the detection of the moving cloud shadow. Although there are many different ideas to detect moving cloud shadow regions using the reflection ratios, either due to the complexities of them or their assumptions made them inapplicable to our problem. Hence a modified adaptive thresholding approach has been designed. The double-thresholding approach explained in (Lyons, 2004) and the region growing method introduced in (Matas et al., 2004) are linked to each other to form this approach.

As mentioned earlier, since large spatially connected moving shadow regions are detected to be found, in the first stage of the proposed thresholding method it is focused to find out core regions within the large moving cloud shadow areas. To acquire those core areas, an initial thresholding operation with predefined values is applied for both of the FSTI and the FITS regions. Since it was assumed that the intensity of the direct light source, $c_P$, is high to be compared to the ambient light source, $c_A$, the pixel intensity ratios demonstrated in both (7) and (8) give a clue to determine the initial thresholds.

For the FITS regions, (7) can be arranged as:

$$\frac{c_A}{c_P \cdot cos(\theta) + c_A} \leq \frac{c_P \cdot cos(\theta) \cdot k(x,y) + c_A}{c_P \cdot cos(\theta) + c_A} \leq 1 \quad (9)$$

For the FSTI regions, (8) can be arranged as:

$$1 \leq \frac{c_P \cdot cos(\theta) + c_A}{c_P \cdot cos(\theta) \cdot k(x,y) + c_A} \leq \frac{c_P \cdot cos(\theta) + c_A}{c_A} \quad (10)$$

In (Toth et al., 2004), it is stated that the intensity value ratio for the umbra part of FITS regions varies between 0.77 and 0.97. After extensive studies, empirically it is found that unless $cos(\theta)$ term of both (9) and (10) takes very small values (e.g in dusk), we can specify the predefined thresholds. Hence it is decided that 0.85 is a slack enough to be safe starting threshold to find core areas of the FITS regions. As (8) is the inverse of the (7), for the FSTI regions the initial threshold was defined as $1/0.85$ (1.176).

The regions with a lower reflectance value than the initial threshold constitute the core areas of the FITS region. Small regions that were not spatially connected are discarded from the mask of the core regions using basic morphological operations. After obtaining the mask of the core regions, they are tried to be grown using the following process after utilizing multiple thresholds as follows:

---
**Algorithm 1:** Moving Cloud Shadow Region Growing

---
**Input:** Quotient Image, Initial Threshold, Maps of Core Region
**Output:** The Final Threshold , Grown Cloud Regions
1  Start with the predefined initial thresholds
2  Measure the ratio of,
   (I) the number of connected component regions with cores to the number of connected component regions.
   (II) newly added pixels to connected components with the already available pixels.
3  If either *(I)* or *(II)* of the *Step 2* exceed the predefined ratio levels, Stop.
   Otherwise, increase the threshold with a small step size and repeat procedure starting from the Step 2.
4  Get the enlarged regions including core regions as the moving section of the cloud shadow regions.

---

The same procedure explained in Algorithm 1 was repeated for FSTI regions by decreasing the second threshold value. The illustrative results of the frames can be examined in Figure 5.(e) for one of the datasets.

#### 3.3.2 Moving Shadow Border Detection:

The absolute difference between the FITS masks of the current ($t$) and earlier ($t - v$) frame gives the moving border mask of FITS regions as presented below:

$$BM(x,y) = |SM_t(x,y) - SM_{t-v}(x,y)|, \qquad \forall(x,y) \quad (11)$$

where
$w$ : width of the video frame,
$h$ : height of the video frame,
$x, y$ : image pixel coordinate,
$BM$ : moving border mask,
$SM$ : moving cloud shadow mask.

The moving border mask for the FSTI regions can also be detected with the same operation. In other words, the final changed parts of the cloud shadow regions are marked as the border regions. The moving border regions of the moving cloud shadows are marked with red color in Figure 2.

The $v$ value in (11) needs to be determined by concerning the behavior of the background subtraction algorithm used to get foreground object candidates. In the presented approach, a spatially strengthened version of (Zivkovic, 2004) is applied and especially the learning rate and sigma distance that are used in (Zivkovic, 2004) drastically affect the region of the border in order to filter out object candidates.

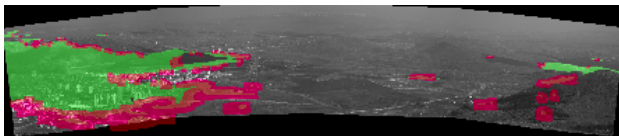Based on the value of $v$, it is possible to obtain quite narrow or wide boundary regions.



Figure 2. Moving Border of the Cloud Shadow Illustration

### 3.4 Moving Object Filtering under Cloud Regions

To analyze the foreground object candidates under the border regions of the moving cloud shadows, the candidates were filtered using the sum of border masks. The real moving objects are tried to be selected among all candidates by analyzing the relation between the object candidate and cloud border regions (background) surrounding the candidate on the quotient image. It should be noted that in this part of the study, all the operations have performed on the quotient image on the original scale.

The reflectance ratio distribution of the background border regions is highly consistent in a close neighborhood even for the penumbra region. If the candidate object is a false alarm belonging to the stable background the distribution of the candidate and the surrounding region shows very similar characteristics. The key idea behind this procedure is that one should exclude the other candidate objects in the background to get reliable statistics belonging to the surrounding region of the object candidates. Otherwise, the real moving objects located in the surrounding region of the candidate can misguide the object-surround analysis. If the candidate object is a real moving object, it is assumed that the distributions of the candidate and surrounding region shows distinct characteristics on the quotient image. Illustrations for both a false alarm and a real object samples can be seen in Figure 3 and 4 respectively.

For the sake of simplicity and computational efficiency, for discriminating the real objects from the false alarms, the mode values of the distributions are used. The key idea is that the distribution of the object does not resemble the FITS or FSTI version of the background. Hence the mode value of the object and its surrounding regions is expected to have a different mode intensity value on quotient image. However, if it is not the case, the object candidate cannot be differentiated from the background and it will be eliminated. In order to eliminate the non-real object candidates, the thresholding operation to the absolute difference of the mode values is shown as follows:

$$|Mo(cen) - Mo(sur)| \overset{?}{\geq} thr\_mode, \qquad (12)$$

where    $cen$ : image patch showing the object,
           $sur$ : image patch showing surrounding of the object,
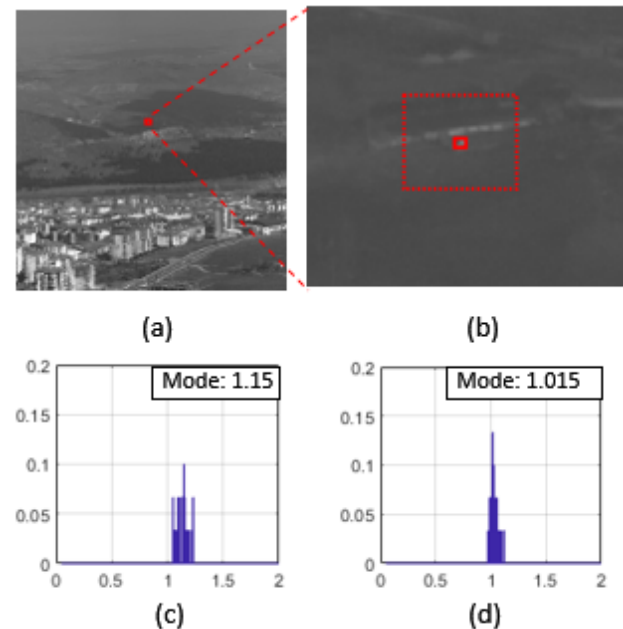           $thr\_mode$ : real moving object threshold value.



Figure 3. Deleted Detection and Distribution of Reflectance Ratios: (a) Zoom-Out Image Patch, (b) Candidate Object & Surroundings, (c) Distribution of the Candidate, (d) Distribution of the Surroundings
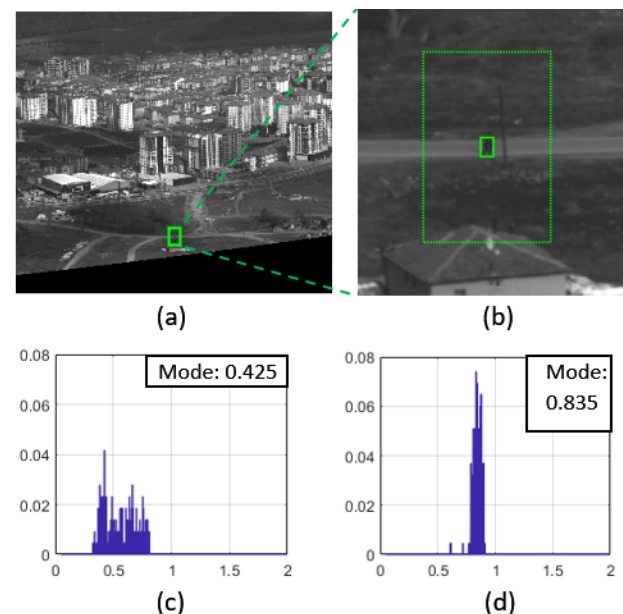


Figure 4. Surviving Detection and Distribution of Reflectance Ratios: (a) Zoom-Out Image Patch, (b) Candidate Object & Surroundings, (c) Distribution of the Candidate, (d) Distribution of the Surroundings

## 4. EXPERIMENTAL RESULTS

### 4.1 Evaluated Datasets

All of the three datasets are captured by using an Aerostat WAMI solution with an 8-bit gray-level imaging format at different times and different locations. The video frames that the proposed algorithm has been tested with were selected from a time span between 5 minutes to 20 minutes later than the reference (background) frame was captured for each dataset. Each of the three datasets has a single reference image and three testing images taken from the video sequences.

### 4.2 Evaluation Metrics & Performance Results

To evaluate the performance of the moving cloud shadow detection stage of the proposed algorithm in a quantitative way the data is annotated by drawing ground truths for each image in the datasets as shown in Figure 5.(c). We used different color annotations for the FSTI and the FITS regions. The former ones have labeled with yellow and the last ones have labeled with red color. The drawing of the ground truths is quite a challenging operation; therefore, in the truths there could be some minor deflections due to human error and the lack of a rigid structure of the cloud regions. The total detection performance of both regions defines our scene based performance.

The recall, precision and F1 score (Goutte, Gaussier, 2005) are well-known and common evaluation metrics used to validate the pixel-wise performance of the application of detection, segmentation, and classification, etc. Therefore, these metrics were used to evaluate the performance of the first stage of the proposed method. We have repeated the same detection procedure for the 3 testing scenes of each dataset then the mean performance results of each dataset have reported in the first three rows of Table 1. In the last two rows of Table 1, the mean elapsed times during the detection procedure and resolutions of the video frames are shown respectively.

| Datasets | Set-1 | Set-2 | Set-3 |
|---|---|---|---|
| Mean Recall | 0.8477 | 0.8082 | 0.8361 |
| Mean Precision | 0.8583 | 0.8449 | 0.8686 |
| Mean F1 score | 0.8524 | 0.8216 | 0.8517 |
| Detection Time (in msec) | 22.5 | 24.45 | 5.48 |
| Resolution (in megapixels) | 130.14 | 130.14 | 28.83 |

Table 1. Performance Results of Proposed Moving Cloud Detection Algorithm

The proposed method is implemented using OpenCV including a few common libraries of C++ and executed in an Intel Core i7-7700 3.60GHz 16GB RAM, GeForce GTX 1050 Ti PC system.

The total execution time for filtering out moving objects in the elimination operation of the false alarms generated by cloud shadows is varying depending on the number of objects subjecting to the mode value comparison operation. The average time spent for the boundary identification stage is 15.8983 msec per video frame and mode value comparison stage for a candidate object lasts approximately 0.0613 msec. Although the test applied to eliminate false alarms takes 0.0613 msec the total time can rise to tens of milliseconds depending on the number of object candidates in borders of the shadow regions.

## 5. DISCUSSIONS & CONCLUSIONS

In the proposed approach, it is aimed to detect dynamic target objects under moving cloud shadow regions. Since the mo-

tion characteristics of the cloud shadows do not similar to the shadow regions of the static man-made or terrain objects, background subtraction algorithms usually cannot model it successfully. Especially, the border regions of the cloud shadow cause the production of highly error-prone and misleading outputs due to the deficiencies of the background subtraction algorithms. Hence, detection of the moving part of the cloud regions in a precise manner is an essential stage to obtain reliable object detection and tracking solutions in WAMI systems.

Even it is assumed that the camera position is static, due to atmospheric distortions and the dynamic behavior of the observed scenes, an initial preprocessing step is required before performing any detection algorithm for the moving cloud shadows regions. Since the real-time WAMI applications have a heavy computational burden, a downscaling operation is performed not only to eliminate the disadvantages of aforementioned dynamic distortion effects, but also to reduce the time elapsed during cloud shadow detection.

Utilization of the proposed adaptive region growing approach for cloud shadow detection and determining the moving border of the cloud regions are two main contributions of the proposed technique. Moreover, in this study, it is desired to successfully detect the moving targets that are located under the border regions of the cloud shadows, while trying to eliminate the false alarms generated by the cloud itself.

As in many gray-level object independent cloud shadow detection methods, the quotient (i.e. ratio) images that are obtained by dividing the current video frame to the background image are the major input source for the proposed approach. The distribution of the reflectance ratio values that form the quotient image is used in the last stage of our methodology to discriminate the real objects from the false alarms due to the boundaries of the moving cloud. If the mode value of the reflectance ratio distribution of the real object turns out to be similar to that of the background, the object unfortunately is missed. Even in such a scenario, the detection of such moving targets is delayed until the objects move out from the border regions of the cloud.

## REFERENCES

AFRL, U., 2009. Wright-patterson air force base (wpafb) dataset.

Al-Najdawi, N., Bez, H. E., Singhai, J., Edirisinghe, E. A., 2012. A survey of cast shadow detection algorithms. *Pattern Recognition Letters*, 33(6), 752–764.

Basri, R., Jacobs, D. W., 2003. Lambertian reflectance and linear subspaces. *IEEE transactions on pattern analysis and machine intelligence*, 25(2), 218–233.

Bouwmans, T., Sobral, A., Javed, S., Jung, S. K., Zahzah, E.-H., 2017. Decomposition into low-rank plus additive matrices for background/foreground separation: A review for a comparative evaluation with a large-scale dataset. *Computer Science Review*, 23, 1–71.

Braaten, J. D., Cohen, W. B., Yang, Z., 2015. Automated cloud and cloud shadow identification in Landsat MSS imagery for temperate ecosystems. *Remote Sensing of Environment*, 169, 128–138.

Chien, S.-Y., Ma, S.-Y., Chen, L.-G., 2002. Efficient moving object segmentation algorithm using background registration technique. *IEEE Transactions on Circuits and Systems for Video Technology*, 12(7), 577–586.

Elgammal, A., Harwood, D., Davis, L., 2000. Non-parametric model for background subtraction. *European conference on computer vision*, Springer, 751–767.

Force, U. A., n.d. Columbus large image format dataset 2007.

Goutte, C., Gaussier, E., 2005. A probabilistic interpretation of precision, recall and f-score, with implication for evaluation. *European Conference on Information Retrieval*, Springer, 345–359.

Huang, C., Thomas, N., Goward, S. N., Masek, J. G., Zhu, Z., Townshend, J. R., Vogelmann, J. E., 2010. Automated masking of cloud and cloud shadow for forest change analysis using Landsat images. *International Journal of Remote Sensing*, 31(20), 5449–5464.

Jacques, J. C. S., Jung, C. R., Musse, S. R., 2005. Background subtraction and shadow detection in grayscale video sequences. *XVIII Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI'05)*, IEEE, 189–196.

Jung, C. R., 2009. Efficient background subtraction and shadow removal for monochromatic video sequences. *IEEE Transactions on Multimedia*, 11(3), 571–577.

Li, Z., Shen, H., Li, H., Xia, G., Gamba, P., Zhang, L., 2017. Multi-feature combined cloud and cloud shadow detection in GaoFen-1 wide field of view imagery. *Remote sensing of environment*, 191, 342–358.

Lin, C.-T., Yang, C.-T., Shou, Y.-W., Shen, T.-K., 2010. An efficient and robust moving shadow removal algorithm and its applications in ITS. *EURASIP Journal on Advances in Signal Processing*, 2010, 1–19.

Lin, Y., Medioni, G., 2007. Map-enhanced uav image sequence registration and synchronization of multiple image sequences. *2007 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 1–7.

Lu, Y., Xin, H., Kong, J., Li, B., Wang, Y., 2006. Shadow removal based on shadow direction and shadow attributes. *2006 International Conference on Computational Inteligence for Modelling Control and Automation and International Conference on Intelligent Agents Web Technologies and International Commerce (CIMCA'06)*, IEEE, 37–37.

Luo, Y., Trishchenko, A. P., Khlopenkov, K. V., 2008. Developing clear-sky, cloud and cloud shadow mask for producing clear-sky composites at 250-meter spatial resolution for the seven MODIS land bands over Canada and North America. *Remote Sensing of Environment*, 112(12), 4167–4185.

Lyons, R. G., 2004. *Understanding digital signal processing, 3/E*. Pearson Education India.

Matas, J., Chum, O., Urban, M., Pajdla, T., 2004. Robust wide-baseline stereo from maximally stable extremal regions. *Image and vision computing*, 22(10), 761–767.

Nagendran, A., Harper, D., Shah, M., 2010. New system performs persistent wide-area aerial surveillance. *SPIE Newsroom*, 5, 20–28.

Oreifej, O., Li, X., Shah, M., 2012. Simultaneous video stabilization and moving object detection in turbulence. *IEEE transactions on pattern analysis and machine intelligence*, 35(2), 450–462.

Perera, A. A., Srinivas, C., Hoogs, A., Brooksby, G., Hu, W., 2006. Multi-object tracking through simultaneous long occlusions and split-merge conditions. *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, 1, IEEE, 666–673.

Piccardi, M., 2004. Background subtraction techniques: a review. *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583)*, 4, IEEE, 3099–3104.

Pilet, J., Strecha, C., Fua, P., 2008. Making background subtraction robust to sudden illumination changes. *European conference on computer vision*, Springer, 567–580.

Reilly, V., Idrees, H., Shah, M., 2010. Detection and tracking of large number of targets in wide area surveillance. *European conference on computer vision*, Springer, 186–199.

Sexton, G., Zhang, X., 1993. Suppression of shadows for improved object discrimination. *IEE Colloquium on Image Processing for Transport Applications*, IET, 9–1.

Simpson, J. J., Stitt, J. R., 1998. A procedure for the detection and removal of cloud shadow from AVHRR data over land. *IEEE Transactions on Geoscience and Remote Sensing*, 36(3), 880–897.

Sommer, L. W., Teutsch, M., Schuchert, T., Beyerer, J., 2016. A survey on moving object detection for wide area motion imagery. *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, IEEE, 1–9.

Stander, J., Mech, R., Ostermann, J., 1999. Detection of moving cast shadows for object segmentation. *IEEE Transactions on multimedia*, 1(1), 65–76.

Toth, D., Stuke, I., Wagner, A., Aach, T., 2004. Detection of moving shadows using mean shift clustering and a significance test. *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, 4, IEEE, 260–263.

Vargas, M., Toral, S., Milla, J., Barrero, F., 2010. A shadow removal algorithm for vehicle detection based on reflectance ratio and edge density. *13th International IEEE Conference on Intelligent Transportation Systems*, IEEE, 1123–1128.

Xu, D., Liu, J., Liu, Z., Tang, X., 2004. Indoor shadow detection for video segmentation. *2004 IEEE International Conference on Multimedia and Expo (ICME)(IEEE Cat. No. 04TH8763)*, 1, IEEE, 41–44.

Zhu, Z., Woodcock, C. E., 2012. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote sensing of environment*, 118, 83–94.

Zivkovic, Z., 2004. Improved adaptive gaussian mixture model for background subtraction. *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, 2, IEEE, 28–31.
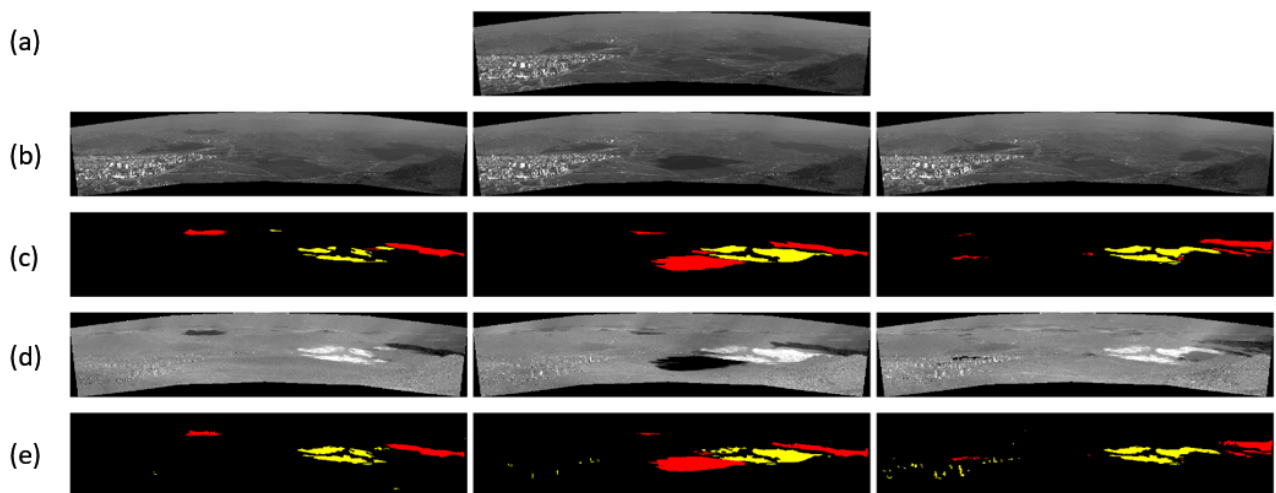
**APPENDIX**

Figure 5. Moving Cloud Shadow Detection Stages for a Single Dataset: (a) Background Image, (b) Images Including the Moving Cloud Shadows, (c) Ground Truth Masks, (d) Quotient (Reflectance Ratio) Images, (e) Moving Cloud Region Detection Results
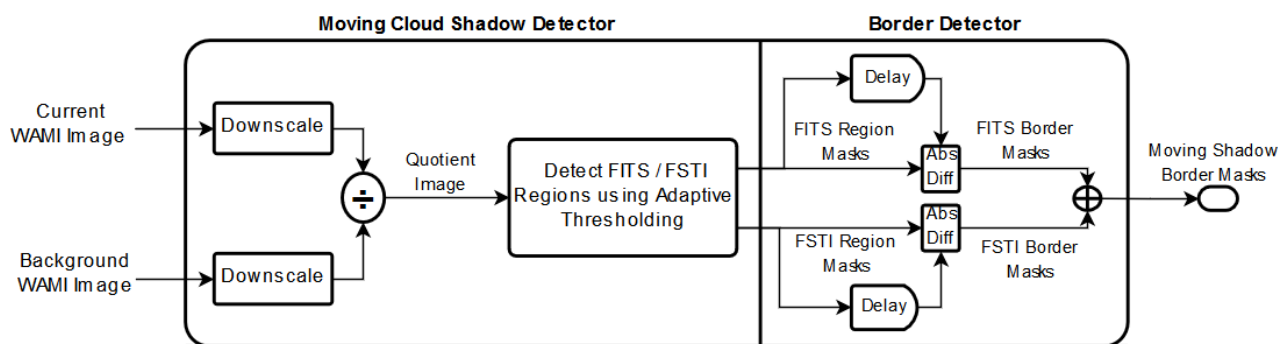


Figure 6. Moving Cloud Shadow Detector Flow Chart (FITS: First-well-illuminated-then-shadowed, FSTI: First-shadowed-then-well-illuminated, The overall system chart can be seen in Figure 1)
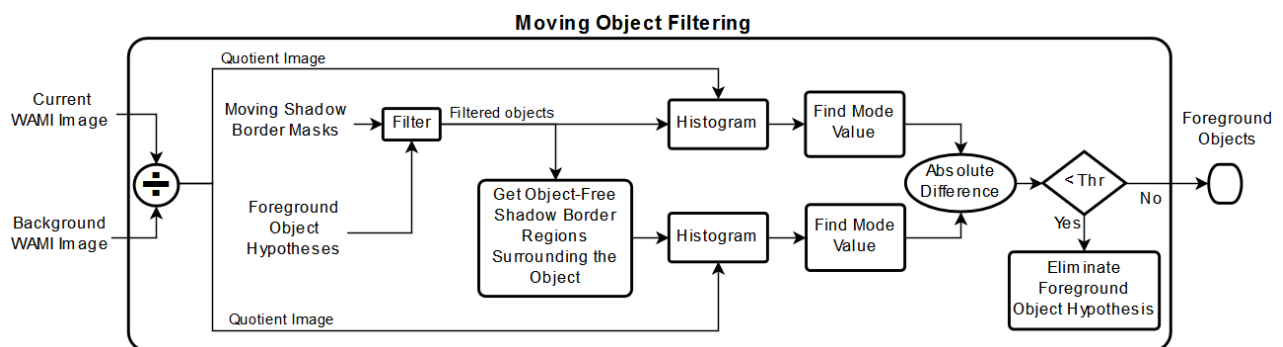


Figure 7. Moving Object Filtering Flow Chart (The overall system chart can be seen in Figure 1)