

A FRAMEWORK FOR AUTOMATED PROGRESS MONITORING BASED ON HOG FEATURE RECOGNITION AND HIGH RESOLUTION REMOTE SENSING IMAGE

Xu Ruhang^{1,2}

¹ School of Economics and Management, North China Electric Power University,
Beijing, PCR

² Beijing Key Laboratory of New Energy and Low-Carbon Development (North China
Electric Power University), Changping Beijing, 102206;

ABSTRACT

Construction target monitoring became one of the key concerns of the construction field. Computer vision based monitoring methods like 2D image, video and 3D laser scan methods have difficulties in meeting real-time and data objectivity requirements. This paper points out the applicability of high resolution remote sensing image (HRRSI) in this task. This paper proposes a way for enhancing Histogram of Gradients (HOG) feature for HRRSI. HOG is applied as the key feature for machine learning, which is carried out by the boosting method. To overcome the complexity of HOGs in HRRSIs, this paper builds a feature framework according to the top, side and height characteristics of engineering ground objects. Based on the feature framework, this paper presents a HRRSI based framework for construction progress monitoring. Through an application case, the performance of the framework is presented. Judging from the result, this method is robust under different seasonal and illuminative conditions.

KEYWORDS

Network Object recognition, HOG feature, High resolution remote sensing image, Progress monitoring, Construction project

1. INTRODUCTION

Construction target monitoring became one of the key concerns of the construction field. Generally, monitoring targets include construction targets which may have critical influences inside and outside the construction projects, like progress, quality, safety and environment. Some researchers acquire real-world construction site data through several kinds of sensors, like RFIDs (radio frequency identification devices), CSSs (chirp spread spectrum) and GPSs (global positioning system) [1-4]. The computer vision technology also shows great potential in construction target monitoring applications. Firstly, most of the computer vision data acquiring ways are not expensive. Secondly, vision data can provide information not only about building objects, but also about moving workers, moving equipment, occupied space and so on [5-7].

The on-site construction progress monitoring can benefit project control by detecting and correcting progress deviations. There are 3 major technical approaches for computer-vision-based construction target monitoring.

1.1.2D IMAGE BASED

Xiaonan et al. emphasized that the computer-vision-based construction progress automation system can greatly help not only progress but also cost control [8]. Photographs are combined with BIM to generate as-built 3D model of interior construction works. Seungjun et al. proposed a way to build this kind of model by applying the photograph-based object recognition algorithm [9]. Quality monitoring, the quality of the tile construction work is assessed by analyzing computer-vision geometric characteristics [10]. Kevin et al. emphasized the importance of the on-site operation-level progress monitoring. It is claimed that the operation-level real-time progress monitoring can reveal engineering details and provide sufficient deviation information for project control. They combined BIM with site photo logs to create a detailed real-time on-site progress model [11]. Firstly, the BIM plan should be built. To monitor the interior progress of the construction works, the inside BIM model should be fully built. Secondly, objects in the real world should be detected and attached with corresponding plan objects. Normally, this progress is done by applying computer vision methods based on photographs or videos. Thirdly, the progress is analyzed by comparing the BIM-plan with photograph-real-world information [12]. Real-time 3D modeling of built objects is another way to achieve real-time progress monitoring. Pablo et al. proposed an as-built 3D reconstruction method by processing 2D images acquired from an unmanned aerial system [13].

1.2. VIDEO BASED

To perform progress tracking, some researchers paid attention to moving object tracking other than building part tracking like worker and equipment movement tracking. Gong et al. proposed a worker movement tracking method through the vision analysis based on video and photograph [14]. Man-Woo et al. carried out a comparative study on the vision tracking methods for construction site object tracking. They categorized 2D vision tracking methods into contour-based, kernel-based and point-based methods. They concluded that for current technical situation, the kernel-based 2D tracker may be the most suitable one for construction site object tracking [15]. Man-Woo combined the background subtraction method, histogram of oriented gradients and the HSV color histogram together and proposed a construction site worker tracking method [16]. Video-based analysis was also applied in quality control. Mahmoud et al. applied a video processing method to detecting defects in the city sewer system.

1.3. 3D LASER SCAN BASED

Some researchers applied laser scanners to acquiring the as-built raw data. Frédéric presented a 3D computer-aided-design object recognition method in laser scanning results and emphasized the advantage of this method in construction dimensional compliance control [17]. Chao et al. proposed a way to automatically extract building geometries from unorganized laser-scanning point clouds, and attach them to as-planned BIM objects, so as to detect progress deviations [18]. As for 3D as-built object reconstruction, there are two main approaches: the image-based and the laser-scan-based. Mani et al. pointed out that the former may not be as accurate as the latter in defect detection or alignment tasks. They also concluded that the former has an advantage in operation cost [19]. Some researchers considered the possibility of combining 2D and 3D sensing technologies to build a multifunctional progress monitoring system, which is still in its initial stage yet [20-21].

It can be summarized from previous studies in the field that each kind of technology has its advantages and disadvantages. 2D image-based construction as-built modeling is less expensive. Compared with 3D laser scanners, 2D image sensors with satisfying technical parameters are far less expensive. At the same time, the usage of 2D image sensors, such as a digital camera, is

simpler than a 3D laser scanner, which requires operation training. As mentioned in literatures, 2D real-world images can provide more information than a simple 3D point cloud. The information of elements like shadows, surface textures, movements and specific patterns can be provided in a visible light exposure 2D image, which can be significant for modeling. The deployment of a 2D image as-built sensing system may be simpler than a 3D-laser-scanning one because 2D-based image monitoring system application is more mature in many fields like security monitoring.

At the same time, the disadvantages of a 2D image-based approach are obvious. The information, especially distance information provided in a 2D photograph is less accurate. Due to the light sensing mechanism of a 2D image sensor, the results can be interfered by many factors like illumination, view point and temperature. The advantages of a laser-scanning based monitoring method are obvious. Firstly, 3D laser scanning can acquire precise coordinates of each point, which is significant in quality deviation monitoring. Secondly, 3D-scanning results can be directly transformed into a 3D model. With the precise coordinates, this 3D modeling way can be faster and more accurate than a 2D-way. Besides, an as-built 3D model can provide visualized construction process information to project managers and customers. However, disadvantages came along. 3D-scanning equipment is much more expensive than a 2D camera under current technical conditions. What's more, 3D scanning maybe more difficult to deploy and provide mostly position information at the same time.

In general, there are obstacles for both the 2D and the 3D ways in practical construction monitoring application. A deployment solution should be provided for the whole construction period. Actually, monitoring equipment is hard to be fixed in one location during the construction process. Because the interior environment changes a lot while the engineering proceeds. If the data cannot be automatically acquired by previously deployed equipment, a manned acquiring way is the only option, which can still be error-prone and costly.

Practically, there is virtually no proven deployment solution for either a 2D or 3D monitoring way for a practical entire construction process. It can be concluded that there are obstacles for data acquisition when the size of the demand data granularity is too small. In fact, in practical applications, some monitoring tasks only require a bigger size of data granularity.

The project monitoring in the Chinese government-leading investment is a good example. In last decades, the government's investment in fixed assets played an important role in economy simulation in China. Great amount of GDP growth has been stimulated by the government fixed assets investment since 2000. To handle the international economy crisis, the Chinese government decided to carry out a 4,000 billion RMB investment in 2008, mostly in public infrastructures. Although Public Private Partnership became population in recent years, there is no doubt that government-leading fixed assets investment plays a significant role in the economy development. There is a unique character of government-leading fixed asset investment. That is the public resource utilization. This kind of utilization comes in many forms, like public finance, land use and franchise. So, the utilization process must be monitored so as to guarantee that public resources are put into proper use.

In fact, problems came along with the government investment in China in last decades' practices. Some local projects start construction before legal procedures, which means unauthorized public resource occupation. Even in some authorized projects, misuse of the public resources happens. For example, land-use ways can be changed. In the practice, some developers change some of the development targets and made them more profitable for themselves. For example, they may add a commercial building in the development plan on the land which they were authorized to construct for public benefit. In some cases, the developer changed the construction target, to make it better-built meanwhile more expensive for their own benefit. Some targets were changed during the

construction procedure, so that the government had to increase the budget. For the government, there can be great amount of projects in construction, and these projects vary in stages, types and locations, which makes the monitoring very difficult [22].

The structure of this paper is as follows. Section 1 is the background introduction. Section 2 discusses the applicability of HRRSI in progress monitoring of construction projects. Section 3 gives the feature framework for the engineering ground object recognition task in a HRRSI based on HOG feature. Section 4 gives the progress monitoring framework of a construction project. Section 5 presents an example of the methods mentioned in this paper. Section 6 gives the conclusion.

2. THE APPLICABILITY OF HIGH RESOLUTION REMOTE SENSING IMAGE IN PROGRESS MONITORING OF CONSTRUCTION PROJECTS

At present, some problems still exist in nowadays technical framework for effective monitoring of construction projects, as shown below:

- Real-time data acquisition. Traditional monitoring methods have a common problem, that is, how to effectively obtain related data in real time. In most of these methods, data need to be subjectively measured, which creates two derivative problems. Firstly, if the data are provided by the under-monitoring entity, the authenticity of the data is difficult to be guaranteed. Secondly, if the supervisor uses the method of manual data collection, it will face a lot of data acquisition costs and it is too difficult to achieve the real-time performance.
- Data objectivity. There is a major problem with current monitoring that the use of information asymmetry by the supervised may lead to the damage to project owner's interests and further cause a series of conspiracy problems. Besides, many of the indicators in traditional monitoring approaches are subjective. Thus, there is a common problem in the monitoring approach outlined in this section. As long as the subjective data acquisition is involved, the original problem will perhaps become a more complex game problem. Therefore, using subjective data for monitoring has inherent defects.

The combination of the Computer Vision (CV) technology and High Resolution Remote Sensing Image (HRRSI) is suitable for the above problem scenarios. The phrase "high resolution" here refers specifically to "high spatial resolution" and "high spectral resolution". For example, the spatial resolutions of American IKONOS, QuickBird and WorldView-02B satellite panchromatic image have respectively reached 1m, 0.61m, 0.5m. The adaptability is manifested in the following three aspects:

(1) Satellite data has a variety of application scenarios. The monitoring application is only a small part of its application value. Using high-resolution satellite remote sensing data for construction project monitoring has the advantage of low cost from the cost-sharing perspective.

(2) Satellite remote sensing data has a natural advantage in wide geographical distributed project monitoring. Supervisors do not need to install and maintain any sensing equipment at the project site, nor do they need to send personnel to the site to collect data. They only need to process the data sent by the satellite, which will greatly improve the monitoring efficiency.

(3) Satellite remote sensing data meets the monitoring requirements of the data size. Supervisors do not need to deal with too specific project data.

The technical principle of the monitoring framework based on the satellite remote sensing image and the CV processing method will be discussed in Section 3.

3. THE MONITORING OBJECT RECOGNITION METHOD BASED ON THE ENGINEERING APPARENT VISION PROCESS FEATURE

3.1. ENHANCED HOG FEATURE ALGORITHM FOR GROUND OBJECT

HOG refers to Histogram of Oriented Gradient. Navneet Dalal and Bill Triggs firstly mentioned this method for pedestrian detection in [25]. This method presents the feature of a certain object in an image through the histogram of different gradient orientations [26].

The basic steps of the method are described as follows. First, an image is grayscale and normalization processed and the gradients of the image is calculated. Then, the image is divided into cells by certain window, for example, 16×16 or 32×32 . The histograms of gradients in cells can be calculated. The feature vector of a cell is the weighted average of gradients in different directions. Next, certain amount of cells is combined into blocks. The block connect feature vector of cells in the block is normalized. Feature vector of blocks in the image is the HOG feature of the image.

The algorithm of feature vector F for a certain image $I(x, y)$ ($x \in [1, m], y \in [1, n]$) is described in Algorithm 1.

Algorithm 1: HOG feature for ground object

Input: $I(x, y)$

Output: F

$I_y \leftarrow I(x, y) \otimes [-1, 0, 1]$

$I_x \leftarrow I(x, y) \otimes [-1, 0, 1]^T$

$w \leftarrow \text{window width}$

$o \leftarrow 360/9$

for $i \leftarrow 1:m/w$ **do**

for $j \leftarrow 1:n/w$ **do**

$xw \leftarrow x \text{ direction gradient in the window}$

$pw \leftarrow \text{phase in the window}$

for $p \leftarrow 1:w$ **do**

for $q \leftarrow 1:w$ **do**

$ang \leftarrow \text{residual value of } \frac{\arctan(pw(p, q)) \times 180}{\pi \times 360}$

if $xw(p, q) < 0$

if $ang < 90$

$ang \leftarrow ang + 180$

if $ang > 270$

$ang \leftarrow ang - 180$

$hist(N, Ang)_{i, j} \leftarrow \text{gradient direction histogram; } N \text{ is number; } Ang = \{1, 2, \dots, 9\}$

//block width is 2

$F_{k, s} \leftarrow \{hist(N, Ang)_{i, j}, hist(N, Ang)_{i, j+1}, hist(N, Ang)_{i+1, j}, hist(N, Ang)_{i+1, j+1}\}$

This paper hopes to recognize engineering ground object by top and side outline features. The most outstanding characteristics of these features are the closed geometric figures in a HRRSI. However, the inside and outside textures of a target can also have closed geometric figures. To suppress the disturbance brought by this phenomenon, this paper proposes Algorithm 2. Inspired

by literature [27], every pixel is smoothed by its peer group so as to make engineering ground object more outstanding considering 3 tunnel information.

Algorithm 2: Enhancement of outline feature by peer group filter in 3 tunnels

Input: $I(x, y)$

Output: $I(x, y)$

$w \leftarrow$ window width

for $i \leftarrow 1:m/w$ **do**

for $j \leftarrow 1:n/w$ **do**

$R_{i,j}^{(r)} \leftarrow$ red tunnel in the window

$R_{i,j}^{(g)} \leftarrow$ green tunnel in the window

$R_{i,j}^{(b)} \leftarrow$ blue tunnel in the window

$d \leftarrow (R_{i,j}^{(r)} - R_{w/2,w/2}^{(r)})^2 + (R_{i,j}^{(g)} - R_{w/2,w/2}^{(g)})^2 + (R_{i,j}^{(b)} - R_{w/2,w/2}^{(b)})^2$

$d_s \leftarrow$ sort d by ascending sequence

$k \leftarrow$ length of d_s

for $i \leftarrow 1:k$ **do**

$temp \leftarrow 0$

for $i \leftarrow 1:i-1$ **do**

$temp \leftarrow d_s(j,1) + temp$

$a_{1i} \leftarrow temp / i$

for $i \leftarrow 1:k$ **do**

$temp \leftarrow 0$

for $i \leftarrow 1:i-1$ **do**

$temp \leftarrow d_s(j,1) + temp$

$a_{2i} \leftarrow temp / (k+1-i)$

for $i \leftarrow 1:k$ **do**

$temp \leftarrow 0$

for $i \leftarrow 1:i-1$ **do**

$temp \leftarrow (d_s(j,1) - a_{1i})^2 + temp$

$s_{1i} \leftarrow temp$

for $i \leftarrow 1:k$ **do**

$temp \leftarrow 0$

for $i \leftarrow 1:i-1$ **do**

$temp \leftarrow (d_s(j,1) - a_{2i})^2 + temp$

$s_{2i} \leftarrow temp$

$J \leftarrow (a_1 - a_2)^2 / (s_1 + s_2)$

$mJ \leftarrow$ max of J

$PeerGroup(x, y) \leftarrow$ whichever $I(x, y) < mJ$ in the window

$I(x, y) \leftarrow PeerGroup(x, y)$ convolution with Gaussian filter

3.2 TRAINING OF THE FEATURE-BASED CLASSIFIER

This section intends to adopt HOG feature to recognize monitoring objects.

After determining what features are used to recognize objects, a classifier shall be set up to identify the target image. The classifier training method used in this paper is based on 25 (PAC), called boosting. The specific method of setting up such kind of classifier is consistent with [23, 24]. The classifier is a multi-stage one. Each stage can be regarded as a sub-classifier and each sub-classifier is composed of multiple "weak classifiers" based on trained weights. According to the PAC theory, a "weak classifier" can be trained into a "strong classifier". The staged classifier can speed up processing. Firstly, the classifier chooses different window sizes to slide traverse through the object. Then, the classifier sequentially uses each stage to perform the recognition judgment on the images within the window. Once this stage has rejected the object with regarding it as a non-recognition object in the current window, the classifier will move to the next window. This process is shown in Fig.1.

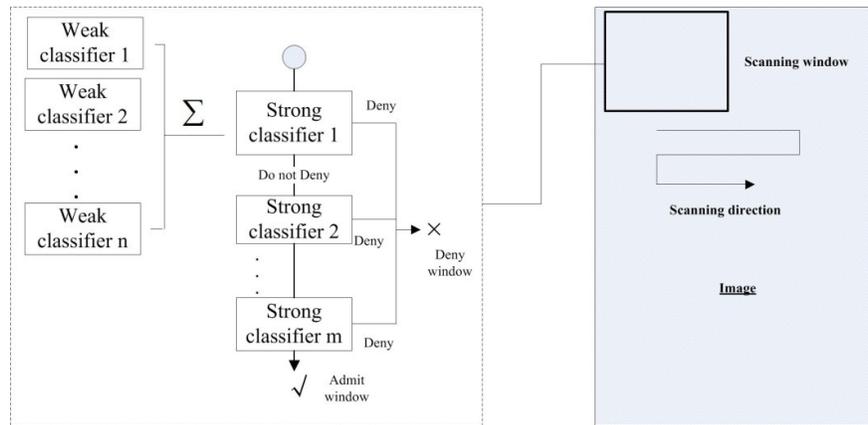


Figure 1. The structure of the classifier

Each of the classifiers has a certain false positive rate, which should be controlled primarily. Firstly, the false positive rate of the classifier in different stages is set as $f_i, i = 1, 2, \dots, N, f_i \in [0, 1]$. Then through N stages, this rate can be controlled at $\prod_i f_i$. But it should be noted that when setting the true positive rate of each sub-classifier in different stages as $k_i, i = 1, 2, \dots, N, k_i \in [0, 1]$, the total true positive rate will be lower than the true positive rate of each sub-classifier. The training of classifiers requires a large number of samples and there are two types of such samples. One is the positive sample and the other is the negative sample. The number of samples is mainly determined by the number of stages of the selected classifier and the total true positive rate. Both positive samples and negative samples should be collected under a variety of different conditions. Thousands of positive samples and negative samples are needed for creating a classifier with a higher true positive rate. It need to be pointed out that this recognition technology has a basic uniformity requirement of the object's appearance characteristics. For the case where there are variations in observation angles, the recognition cannot be achieved by a single classifier. Therefore, it is necessary to establish a logical framework for ground engineering object features so as to establish a classifier system.

3.3 THE GROUND ENGINEERING OBJECT FEATURE FRAMEWORK

3.3.1 TOP FEATURES

To reduce the priori determination of the concept of ground objects, the feature description method is adopted in this section. Computer learning of the features of a ground object makes it

possible for the computer to automatically determine the presence of the ground object in a remote sensing image. This paper employs HOG feature as the description feature.

From the human eye perspective, differences may occur in images collected at different time for the same ground engineering object because of the following four reasons: ① The ground object presents different colour features due to different illumination conditions; ② The ground object presents different shadow features due to different illumination conditions; ③ The ground object with a certain height may have different side features due to different satellite imaging angles; ④ Imaging quality may be different, that is, the images may have different degrees of noise due to the impact of uncertainty factors.

The differences under the above conditions can be seen from the building image taken at different times, shown in Fig.2. It can be seen that there are significant differences in building angles of each image. Fig.2a differs from the other figures without apparent shadows. The building image in Fig.2e is noticeably darker and there are apparent differences in hue because of the difference in imaging conditions. A 32×32 size cell HOG feature is calculated for these images. Seen from the feature map, the HOG feature is not sensitive to illumination conditions and shadow differences. When the imaging noise is small, the HOG feature has robustness to some extent.

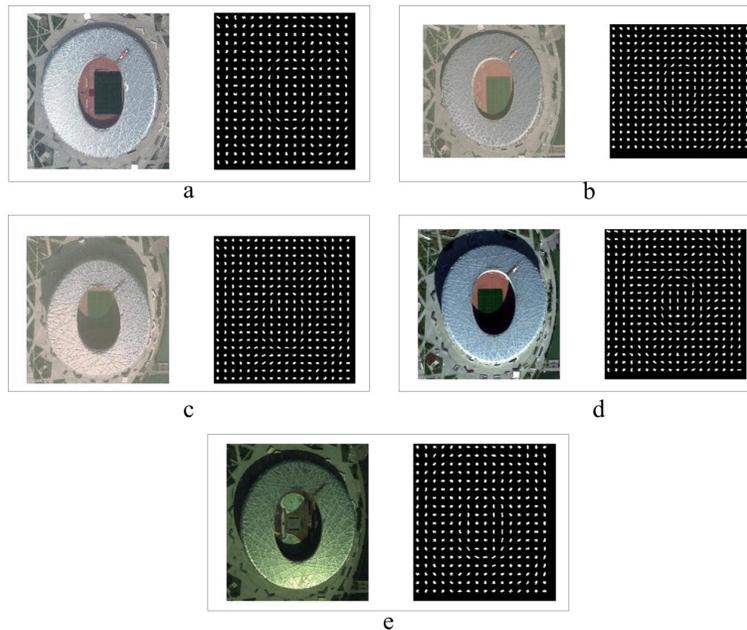


Figure 2. The HOG feature of a building under different conditions of illumination, colour temperature and noise.

Then, another group of building images are compared as shown in Fig.3. It can be seen that when the height of the building is high, its side features show a great difference. There are mainly three cases (excluding buildings with peculiar forms): ① The east, south, west, north side, and the top of the building are individually obviously imaged; ② Only two sides, that is the southeast, southwest, northeast, northwest, and the top of the building are obviously imaged; ③ Only the top of the building is obviously imaged. The HOG features of these images are calculated separately. Seen from the result in Fig.3, the effects of different imaging angles are relatively large, mainly in the presence of different side features. Hence, the feature collection should be carried out from different imaging angles for a certain ground object when a ground engineering object feature library is established.

For most buildings, the most obvious facet in a satellite image is their top outline. The HOG feature used in this paper is sensitive to orientation, so the influence of building orientation on feature recognition under the same external condition is studied firstly.

First of all, the standard observation window used in this paper is north points up and south points down. This is essentially different from the classic recognition of pedestrians. As pedestrians are standing on the ground, the up direction of the image is vertically upward and the down direction of the image is vertically downward in images taken for pedestrians. It is necessary to regulate the orientation of the observation window because if not, the HOG feature of the same object in the same image may be different.

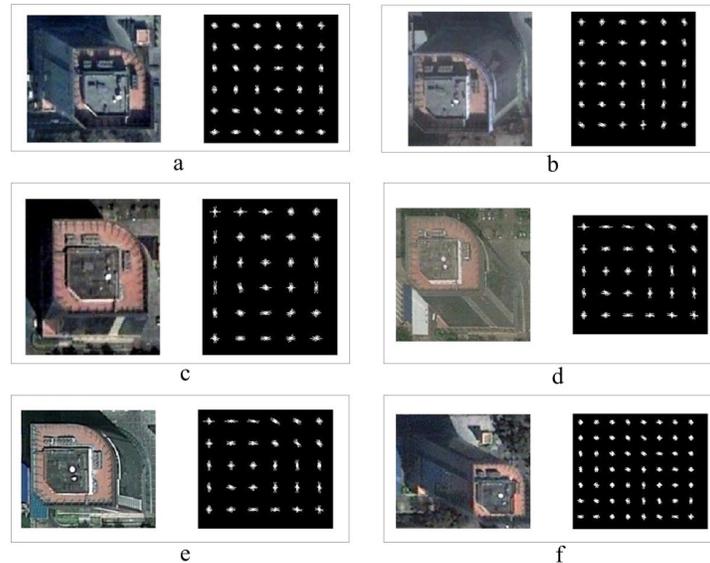


Figure 3. The HOG feature of a building under different side features.

At present, most of ground buildings can be divided into two kinds: ① the north-south orientation; ② the east-west orientation. In feature training, some ground objects can be attributed to the north-south orientation, and the others can be attributed to the east-west orientation, as shown in Fig.4. In order to improve the feature recognition rate, it is necessary to set up sample libraries against the two orientations to carry out feature training. In reality, for example in China, most ground buildings are in the north-south orientation. Thus, the north-south orientation can be regarded as the main recognition orientation for the feature library and classifier when the object recognition is performed.

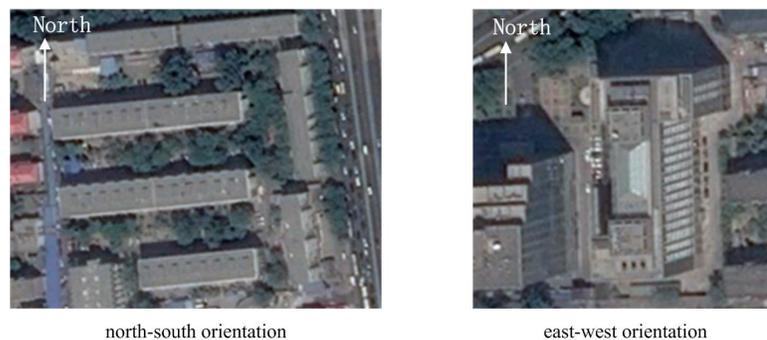


Figure 4. The orientation feature of buildings.

Secondly, the top outline of building ground objects can be divided into several types with differences in the HOG feature. The types are: ① the rectangular feature; ② the round feature; ③ the triangle feature; ④ the combination feature; ⑤ the special type. The first three are basic shapes.

(1) The rectangular feature On the basis of the intuitive appearance, the buildings with the rectangular feature can be divided into three types. ① Standard rectangle. The top outline of this type of building attribute to the standard rectangle according to human eye recognition. ② Micro-deformation rectangular. The top outline of some buildings changes slightly on the basis of the rectangle. ③ Square. The top outline of many towers is intuitively similar to the square. Since the HOG feature is concerned with the direction distribution of the gradient in the image, the building attributed to the micro-deformation rectangular type and the square type can still be trained in accordance with the HOG feature library of the standard rectangle type, and identified by the corresponding classifier. Some images of the buildings with the rectangle outline feature and their HOG features are shown in Fig.5.

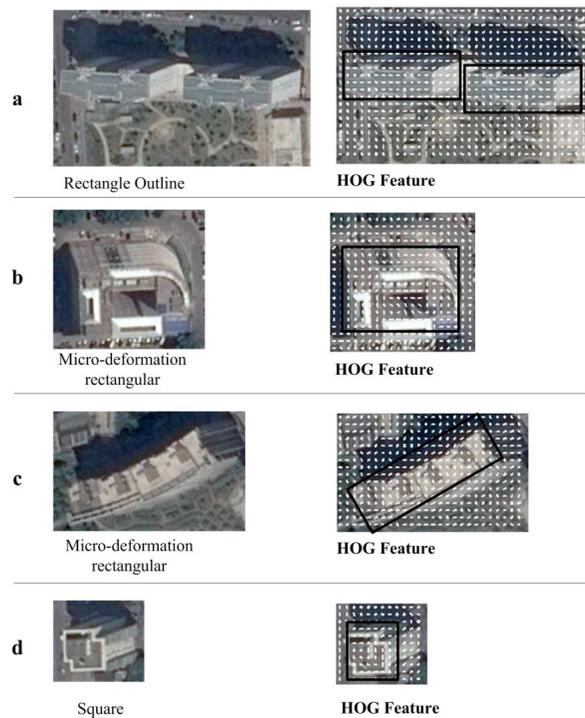


Figure 5. Rectangle outline features and their HOG feature.

(2) The roundness feature The basic intuitive feature of this type of building is that there are no obvious straight edges, as well as obvious corners. The HOG features of circles and rectangles are very different, making it difficult to establish a unified classifier. Therefore, this type of building requires an independent classifier. There is a shape transformation from a standard circle to an oval, and the main change is the length ratio variation of the long axis and short axis. The top outlines of such buildings have a quite uniform feature. Some images of the buildings with the roundness outline feature and their HOG features are shown in Fig.6.

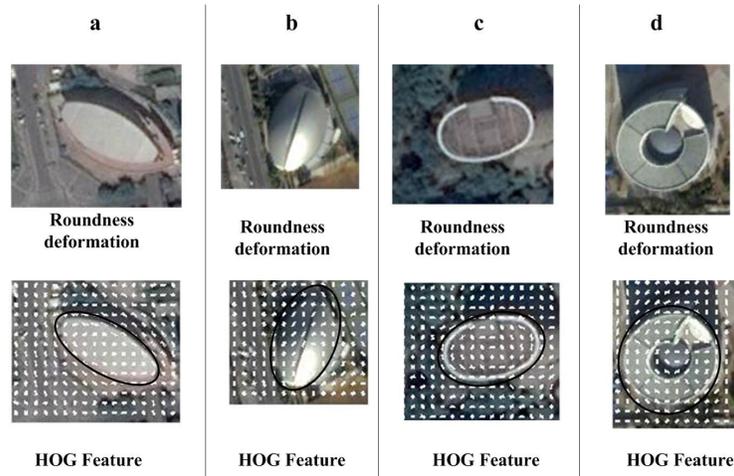


Figure 6. Roundness outline features and their HOG feature.

(3) The triangle feature The basic intuitive feature of this type of building is that there are three distinct edges that form a triangle. It's necessary to establish a classifier for this type of building as the features of basic triangles differ from rectangles and circles. Similarly, the top outlines of this type of building also contain a series of transformations. For example, some of the three angles may appear to be round. But in general, the top outlines of such buildings have a quite uniform feature. Some images of the buildings with the triangle outline feature and their HOG features are shown in Fig.7.

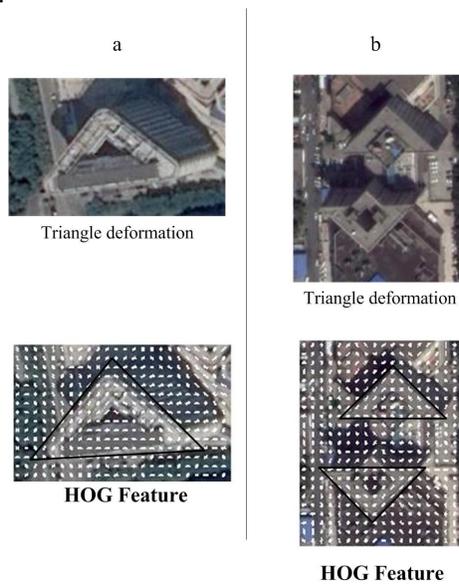


Figure 7. Triangle outline features and their HOG feature.

(4) The combination feature and peculiar types

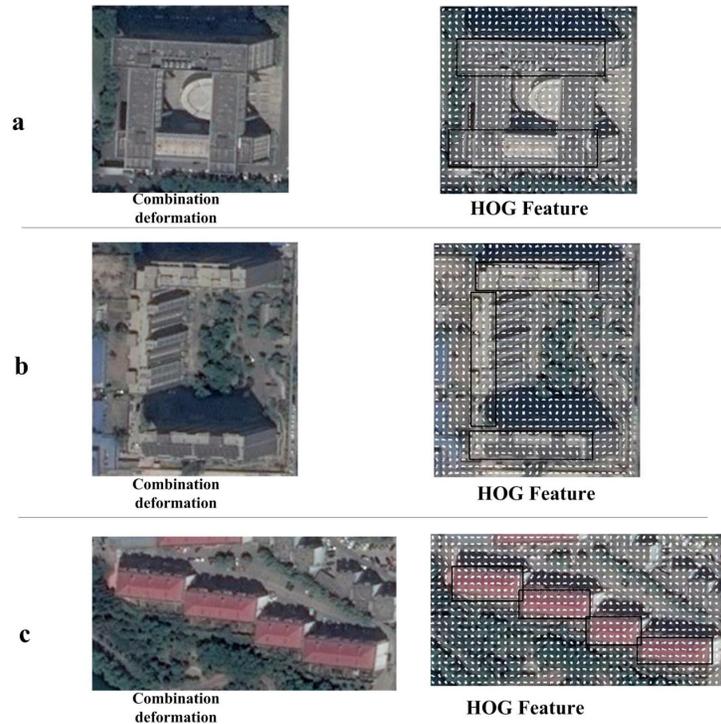


Figure 8. Combination outline features and their HOG feature.

Another type of the top outline is the combination of one or multiples of rectangles, rounds, and triangles, which can be clearly and directly seen, as shown in Fig.8. On the whole, the basic outline of the combination type building changes a lot because a variety of combinations can be made according to the design scheme and the land use requirements. From this perspective, it is difficult to establish a unified comprehensive characterization of combination-type buildings. However, the top outline of this type consists of some basic outlines and their features are not affected and remain unchanged. Seen from the HOG feature images in Fig.8, the outlines contain some different orientation rectangular features. Hence, the building with the combination outline feature can be identified divisionally by the classifier based on basic top outline types.

Moreover, there is a type of outline shape that can hardly be classified as either basic top outline types or the combination type, as shown in Fig.9. Thus, this type of building is classified as the peculiar type in this paper. Such complex buildings usually have a special function like landmark and their overall construction volume is usually large. However, the objects in the construction process still have some basic outlines. For example, they always have rectangular foundation pits. Therefore, the method mentioned in this paper can still be adopted for the recognition of such objects.



Figure 9. Some peculiar building outlines.

3.2.2 THE SIDE FEATURE

Due to satellite orbits and satellite imaging angles, the sides of a ground object can be noticeable in the satellite image sometimes. According to this nature, ground building objects can be divided into two categories: ① the ground building object with obvious sides; ② the ground building object without an obvious side.

In the process of collecting feature samples through artificial visual observation, ground building objects with apparent sides shall be distinguished from ground building objects without an apparent side so as to establish different classifier sample libraries. As the current collection of samples can only be conducted with artificial visual observation, there is no quantitative standard for the distinguishing process. Here this paper proposes an empirical standard. When a side area of a building is larger than 1/5 of its top area, it is considered to have a non-negligible side. According to empirical observations, in northern China, the buildings with apparent sides in satellite imaging are usually high-rises. In the Chinese standard, the definition of a "high-rise building" is: the residential building with a height greater than 27 meters and the non-monolayer factory building, warehouse and other civil buildings with a height greater than 24 meters. Some ground building object images with apparent sides and their HOG feature are shown in Fig.10.

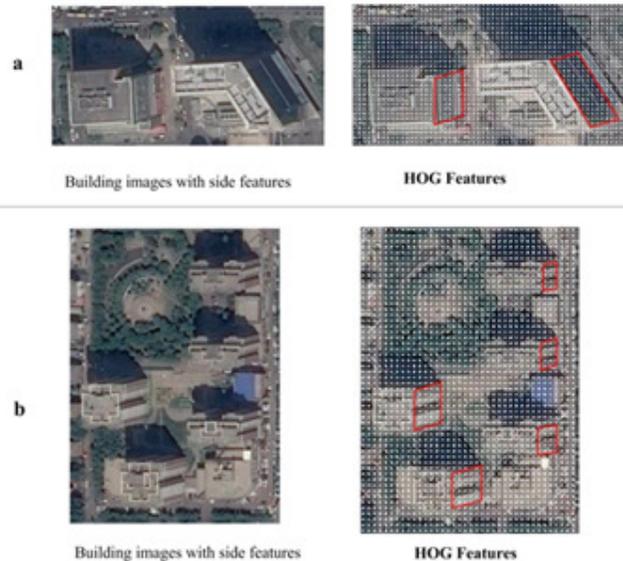


Figure 10. Building images with side features and their HOG feature.

Buildings with no apparent side are usually low in height. Many buildings have this feature, such as low-rise housing areas, Beijing's traditional residential areas (such as Hutong district), temporary living and working areas in the construction site, large farmer product market trading areas and so on. Some ground building object images with no apparent side are shown in Fig.11. It should be noted that when the satellite optical axis is parallel to the vertical direction, high-rise buildings may not have apparent sides. However, it still has some other features, such as a clear shadow. In this case, the building may be recognized on the basis of its top outline. Thus, the classifier should be selected in accordance with reality in a specific object recognition task.

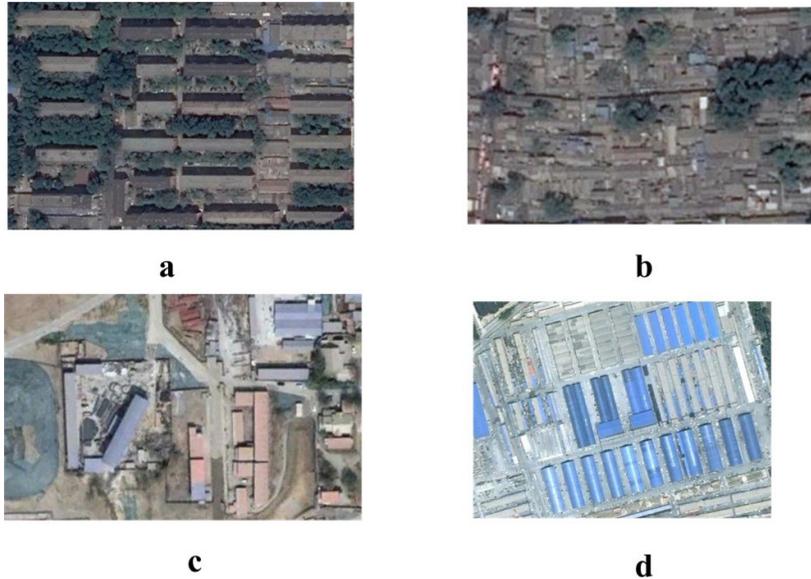


Figure 11. Some samples of ground engineering objects without apparent side surfaces.

3.3 THE FRAMEWORK OF A CLASSIFIER

A ground building object will undergo two kinds of transformations in a construction process. One kind is top texture transformation, because top textures are related to specific work contents and construction organizations. Only when a building's main structure is completed and the exterior decoration is finished will the transformation stop. The other kind is the side transformation, which cannot be observed if the ground object always has no apparent side from the start to the completion of the construction. If a ground object is a high-rise building, it will experience the process from none-side existence to side existence.

In summary, it is difficult to establish a unified feature classifier for ground engineering objects with a variety of internal and external differences. Because the recognition method based on the HOG feature can only establish a classifier for objects that have similar features. Therefore, from the broadest definition, classifiers should be established separately for the various types of features so as to achieve random remote sensing image recognition. According to this, the total number of classifiers that should be established is $2 \times 9 \times 3 \times 2 - 2 \times 3 \times 8 - 2 \times 3 = 54$, as shown in Fig.12. Among them, 'the middle-rise or low-rise building with no apparent side' has no side feature, and the "no-side" feature is not necessary to be established for 'the high-rise building with obvious sides'.

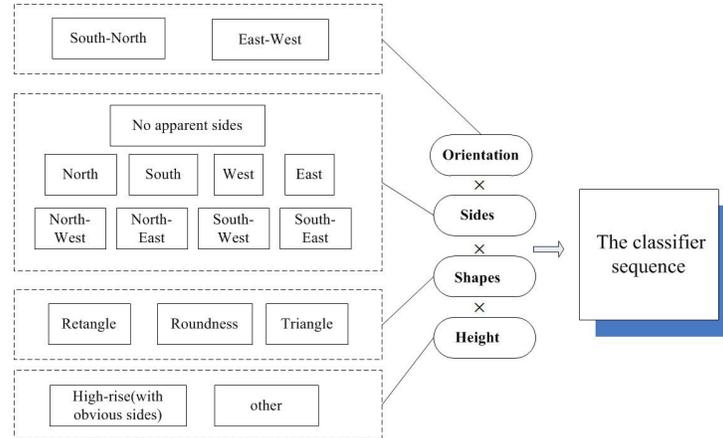


Figure 12. The structure of the classifier sequence.

If the false positive rate of each classifier is 2% and the number of objects identified by each classifier is $n_i, i = 1, 2, \dots, 54, \sum_i n_i = M$, then the average number of false-positively identified

objects is $2\%M$. When the number M is large, the number of false-positively identified objects will be large. If some external information is added in the recognition process, the required number of classifiers can be greatly reduced, thereby reducing the value of M . For instance, if the attitude angle of the satellite with respect to the ground during the imaging can be determined before analysing the remote sensing image, possible side features can be determined. In this case, only a total of 12 classifiers are needed. If the orientation of the target in the recognition window can also be determined, only six classifiers are required then. The apparent progress development will be studied in Section 4. Combined with the apparent progress and some external information, the accuracy of regulation and recognition will be greatly enhanced. For example, the most important thing in the apparent progress is to determine the existence of the object, the planned coordinates and the planned orientation. Hence, the existence of the target object can be easily determined using the classifiers if the satellite imaging posture and the nature of the object to be identified is determined before recognition.

The classifier has a true negative rate and for this situation, it is necessary to manually compare with the apparent progress to find out that whether the recognition is missing or the corresponding project object is not completed.

4. A FRAMEWORK FOR CONSTRUCTION PROJECT PROGRESS MONITORING BASED ON THE HIGH RESOLUTION REMOTE SENSING SATELLITE

4.1 THE MAPPING OF THE PROJECT PROGRESS TO THE APPARENT PROGRESS

For the project progress, what needs to be defined firstly is the project period. As this paper focuses primarily on the apparent nature changes of the construction project, the project progress is only discussed from the start to the end of the construction. The basic objective of monitoring is to compare the actual progress of the project with the planned progress. For the purpose of achieving the automated progress monitoring, the progress of a project needs to be mapped to its apparent progress. As the monitoring based on the apparent nature requires medium level resolution information of the project, it mainly focuses on the following nature of the project: ① Whether a particular object appears in the entire observation scene; ② Whether the location of the object is consistent with the plan; ③ Whether the time when the object appears and disappears is consistent with the plan.

In this paper, the remote sensing image with 0.6m resolution is adopted. In order to facilitate the comparison, an apparent planning system for the project should be established. Firstly, a ground coordinate system is established. According to the general layout plan of different construction phases and the resolution of the specific observation window, the apparent coordinate system can be set up with a point O of the site as the reference point and two fixed directions respectively as the X-axis and Y-axis, as shown in Fig.13.



Figure 13. The coordinate system of a construction project.

Based on this system, the coordinates of the engineering objects in the observation window can be obtained. In order to reduce the omission, engineering objects should be identified from the project work breaking-down system (WBS). According to the project progress plan, the time point when each engineering object appears, transforms and disappears can be acquired. By sorting out this information, a project apparent plan can be obtained, as shown in Fig.14. The apparent plan should indicate that: in what coordinates, what object should appear at what time point; what object should be present at what time point; what object should disappear at what time point. In Fig.14, blue windows indicate the positions of the planned objects.

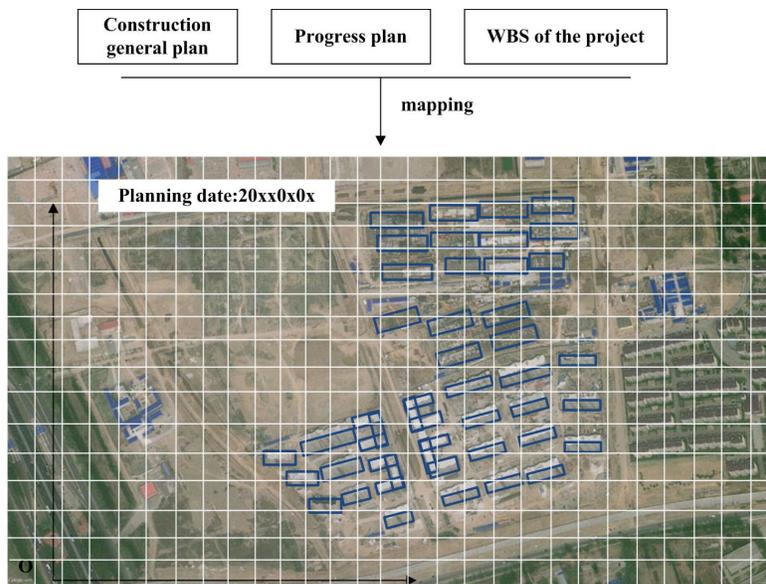


Figure 14. The apparent progress plan of a construction project at a certain time point.

By comparing the actually acquired images that are also identified using the classifier with the apparent progress, the automated construction project progress comparison can be achieved, as shown in Fig. 15.

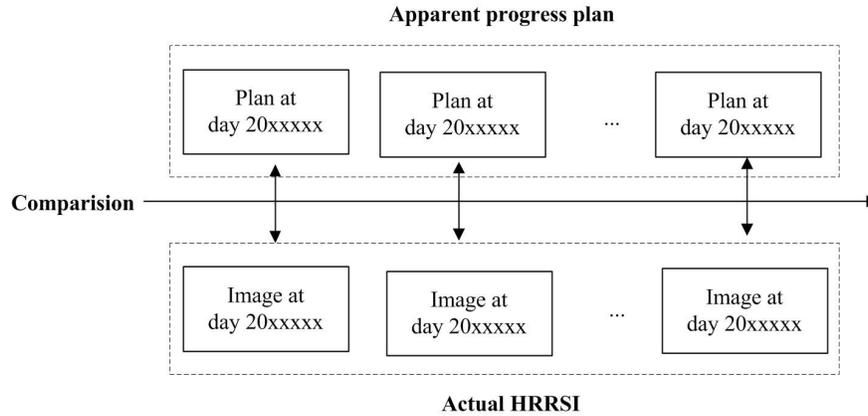


Figure 15. The comparison between apparent progress plan and actual HRRSI.

4.2. OBJECT ALIGNMENT AND OBJECT SEARCH OF THE REMOTE SENSING IMAGE AND THE APPARENT PLAN

In order to achieve the comparison between the actual apparent progress and the apparent progress plan of a construction project, it is necessary to carry out the 'object alignment' process of the collected HRRSIs and the apparent plans.

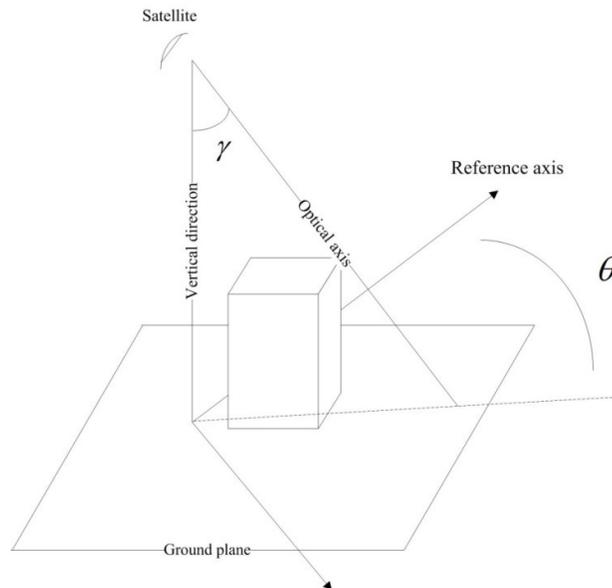


Figure 16. The basic satellite imaging model of engineering ground project.

Assume that the imaging scenes of the target engineering ground object comply with the following assumptions: ① The scene of the ground object is relatively flat so that the terrain factor does not have a big impact on the satellite imaging accuracy; ② The engineering ground objects are perpendicular to the earth surface. The basic imaging model of engineering ground objects is shown in Fig.16, where γ is the angle between the optical axis of the satellite and the

vertical direction. Because of this angle, there will exist not only the top image but also the image of its partial bottom and side in the engineering ground object imaging when the value of γ is significant, as shown in Fig.16. θ is the azimuth angle of the satellite imaging. It is the clockwise angle between the north direction and the imaging direction line of the satellite.

The satellite images of the same feature at different time will be slightly different due to the difference in satellite gestures. Assume that the basic imaging plane remains unchanged and set the coordinates of a point in the three-dimensional space as (x, y, z) when it has a certain height. When γ of the satellite imaging is significant, the relationship between the imaging coordinates (x', y') and the original coordinates is shown in Fig.17.

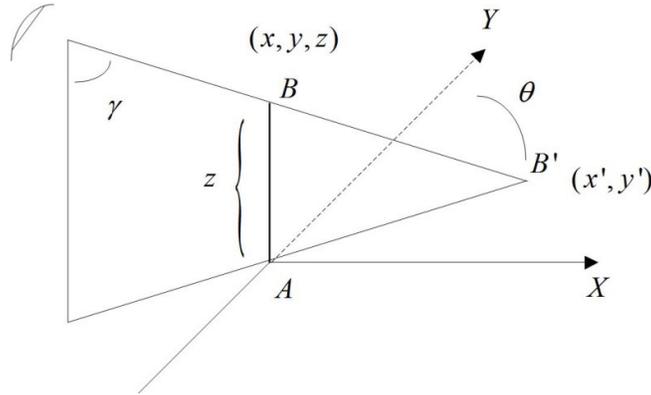


Figure 17. Coordinates relation in satellite imaging.

It can be seen from Fig.17 that the relationship between imaging coordinates and original coordinates is:

$$x' = x + z \cdot \tan \gamma \cdot \sin \theta \quad (1)$$

$$y' = y + z \cdot \tan \gamma \cdot \cos \theta \quad (2)$$

Assume that the observed ground is flat. When the satellite is performing remote sensing imaging, there will be two different imaging points of the same target location. Their imaging distance will be different supposing that the satellite azimuth keeps unchanged while optical axis inclinations are different. As shown in Fig.18, the imaging distance of point A and point B on the ground is AB when the satellite imaging plane is parallel to the ground. The imaging distance of point A and point B is $A'B'$ when the satellite imaging plane is not parallel to the ground. There is an inclination angle κ . As can be seen from Fig.18, $A'B' = AB \cos \kappa$.

Hence, when there is no significant height of the ground object, its imaging patterns will also show a certain difference because of the existence of the optical axis inclination. The endpoint B of line AB , is used as the reference point. Let $B = B'$. Thus, although line AB is a line on the ground without a height, its endpoint A cannot completely coincide. If the satellite images twice, the angles between the imaging plane and the ground plane are respectively κ and κ' , and the imaging lines are respectively AB and $A'B'$. If $B' = B''$, the distance difference between A' and A'' is:

$$\Delta d = |A'B' - A''B''| = |AB(\cos \kappa - \cos \kappa')| \quad (3)$$

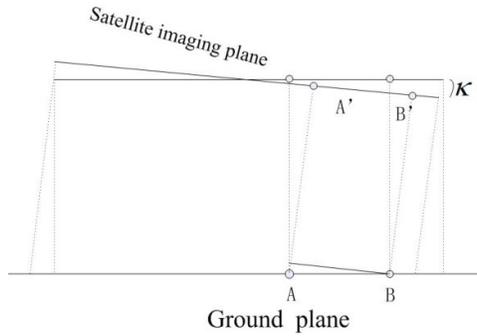


Figure 18. The influence of illumination inclination on imaging.

In the three-dimensional case, when the intersection of the satellite imaging plane and the ground plane is known, the projection position of the new point should be found in line with the intersection' normal directions in each plane, as shown in Fig.19.

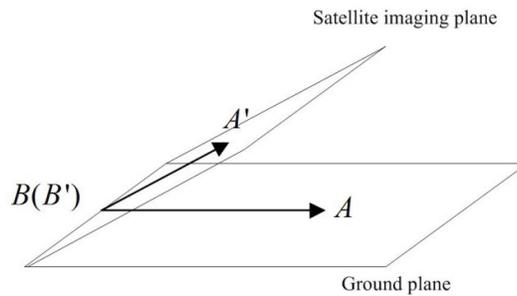


Figure 19. The alignment error direction in 3D condition.

To sum up, in the satellite imaging of the ground three-dimensional object, a difference between the object's top imaging coordinates and the object's bottom imaging coordinates is caused because of the object's own height, the optical axis inclination and the azimuth of the satellite. Besides, a basic stretch or compression along the normal direction with the ratio $(\cos \kappa - \cos \kappa')$ appears because of the inclination difference of the imaging plane at each imaging time causing. For instance, assuming the satellite respectively takes image one time on the monitored area in two monitoring time points and the imaging inclinations are respectively 25° and 30° , then the basic error of the two imaging is $(\cos 25^\circ - \cos 30^\circ) = 0.86\%$.

The so-called object alignment is to align the images collected by remote sensing satellites with the corresponding objects of the project apparent progress. For simplicity, a fixed search radius r can be determined since the basic error is often relatively small where $r > \max(\Delta d)$ and Δd is the possible error distance under the current satellite parameter and window size condition. For collected remote sensing images and the apparent progress, the coordinate system needs to be aligned firstly, and then the searching for the actual apparent object's existence within the search radius r should be conducted with a base point A of a certain apparent progress object as the standard. The point A chosen should be the ground point of the apparent object. Fig.20 shows the search direction alignment.

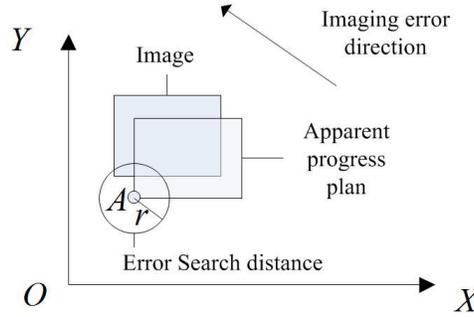


Figure 20. The search direction of appearance progress plan and real image alignment.

In summary, the alignment and recognition of a ground object in the apparent progress coordinates is shown in Fig.21. The yellow window represents the automatic recognition result of certain ground objects, and the blue window identifies the apparent plan location of the objects at the corresponding time point.

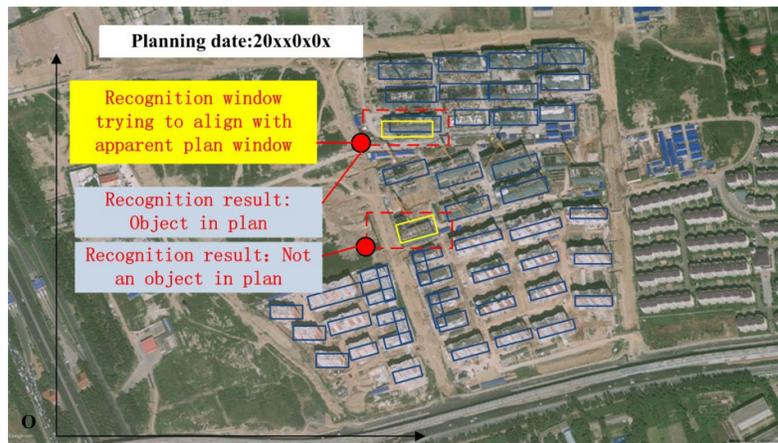


Figure 21. Monitoring compare in the appearance progress plan of a construction project at a certain time point.

5. CASE STUDY

Taking a residential construction project in Beijing as an example, this paper established corresponding classifiers for ground building objects. Now, the side feature is taken as the example to describe the sensitivity of classifiers to different feature types. For the project construction site, it can be seen from naked eyes that the buildings in the image are mainly high-level rectangular outline buildings in the north-south direction with no obvious side. Classifier A for the objects has feature sequence: 'high-level', 'north-south orientation', 'no obvious side' and 'rectangular outline'. Classifier B has feature sequence: 'high-level', 'north-south orientation', 'a southwest side' and 'rectangular outline'. The comparative result is shown in Fig.22. As can be seen, Classifier A correctly identified a large number of ground building objects while the correct recognition rate of Classifier B was significantly lower. The results are compared as shown in Table.1.

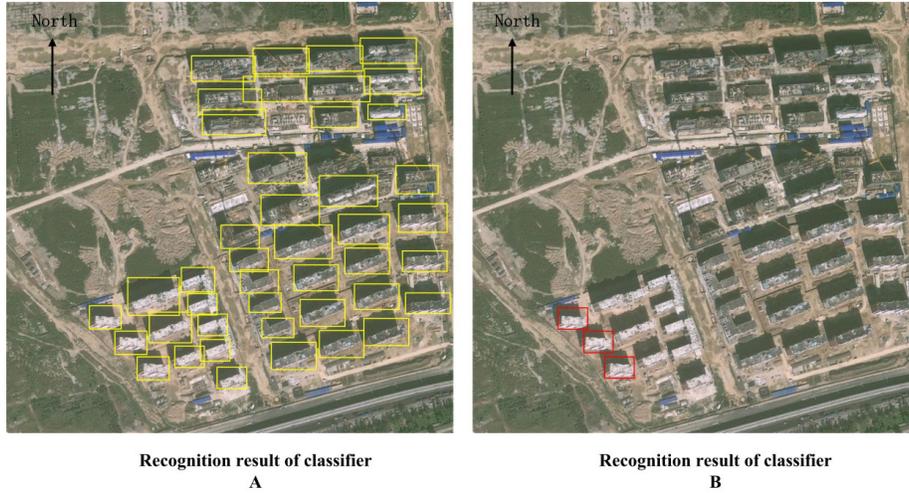


Figure 22. The recognition result of different classifier.

Table 1. The recognition result of different classifier.

	Classifier A	Classifier B
Total number of objects	43	43
The number of objects correctly recognized	42	3
False positive rate	0%	0%
False negative rate	2.33%	93.02%

Seen from the comparison results, only when the appropriate classifier is applied can the objects be identified.

The robustness of this recognition method to the conditions such as the season and illumination is observed. Firstly, HRRSIs of different seasons for the same construction site are obtained. It can be seen that the two images are collected in summer and winter respectively. The summer image is characterized by the presence of large green-vegetation-covered portions while the winter image is characterized by the presence of large snow-covered portions. Then, the classifier: 'high-level', 'north-south orientation', 'a northwest side' and 'rectangular outline', and the classifier: 'high-level', 'north-south orientation', 'a southwest side' and 'rectangular outline' are respectively used to identify these two images. The comparison results are shown in Fig.23 and Table.2.

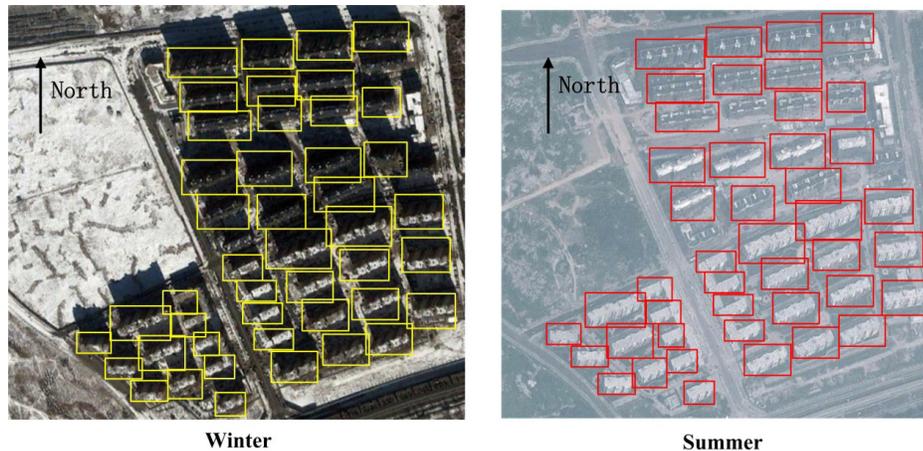


Figure 23. The influence of seasons on the recognition result

Table 2. The influence of seasons on the recognition result.

	Winter	Summer
The total number of objects	45	45
The number of objects correctly recognized	43	42
False positive rate	0%	0%
False negative rate	4.44%	6.67%

It can clearly be seen that the method has strong robustness to seasons. The main reason for the error is the similarity between the target building and the surrounding ground objects, which can be called "camouflages".

Next, the robustness of the recognition method to illumination conditions is observed. First of all, HRRSIs for the same construction site under different illumination conditions are obtained. These two images have significant shadows and colour differences due to differences in illumination conditions. Then, the classifier: 'high-level', 'north-south orientation', 'a northwest side' and 'rectangular outline', and the classifier: 'high-level', in 'north-south orientation', with 'a northeast side' and 'rectangular outline' are applied to identify the images. The comparison results are shown in Fig. 24 and Table.3.

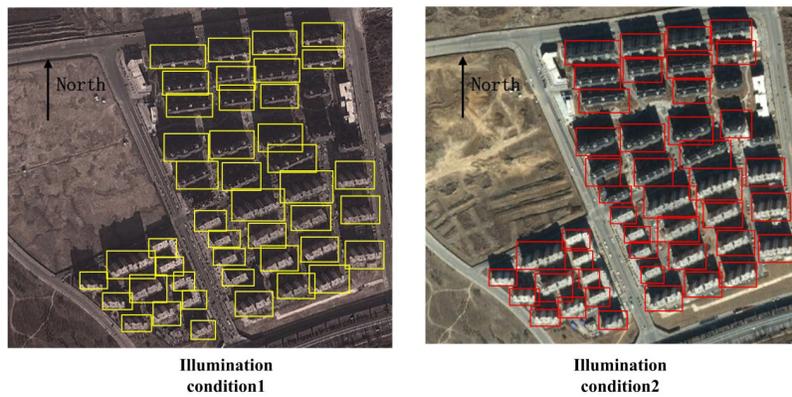


Figure 24. The influence of illumination on the recognition result.

Table 3. The influence of illumination on the recognition result.

	Illumination 1	Illumination 2
The total number of objects	45	45
The number of objects correctly recognized	42	43
False positive rate	0%	0%
False negative rate	6.67%	4.44%

6. CONCLUSION

This paper presents the advantages and disadvantages of computer vision based construction monitoring by 2D image, video and 3D lase scan. For a monitoring task, traditional methods bear the problems of real-time requirements and date objectivity requirements. This paper points out the applicability of HRRSI in this task. HOG feature is applied as the key feature for machine learning, which is carried out by the boosting method. To overcome the complexity of HOGs in HRRSIs, this paper builds a feature framework according to the top, side and height characteristics of the engineering ground object. To realize progress monitoring, this paper firstly

defines the basic tasks of automated progress monitoring: ① Whether a particular object appears in the entire observation scene; ② Whether the location of the object is consistent with the plan; ③ Whether the time point when the object appears and disappears is consistent with the plan. Secondly, this paper gives a framework for real-time HRRSI and apparent progress plan comparison to realize monitoring. Finally, an object alignment and search method are derived. Through an application case, the performance of this framework is presented. Judging from the result, we can see that this method is robust under different seasonal and illuminative conditions. Recently, deep learning method gives good performance in object recognition. For further studies, the combination of the framework given in this work and deep learning is encouraged.

ACKNOWLEDGEMENTS

This paper is supported by the Fundamental Research Funds for the Central Universities (2016MS70) and the Research Funds of Beijing Social Science (16GLC069).

REFERENCE

- [1] Alsemmeiri R A, Bakhsh S T, Alsemmeiri H, (2016). Free Space Optics Vs Radio Frequency Wireless Communication, *International Journal of Information Technology and Computer Science* 8(9), pp:1-8.
- [2] Faridaddin Vahdatikhakia, Amin Hammadb, Hassaan Siddiquic, (2015). Optimization-based excavator pose estimation using real-time location systems. *Automation in Construction* 56, pp: 76–92.
- [3] YManoj P K, Manoj K S, (2013). GPS Tracking System Coupled with Image Processing in Traffic Signals to Enhance Life Security. *International Journal of Computer Science & Information Technology* 5(4), pp:131-142.
- [4] Yuxi Liu, Denghua Zhong, Bo Cui, Guiliang Zhong, Yongxin Wei, (2015). Study on real-time construction quality monitoring of storehousesurfaces for RCC dams. *Automation in Construction* 49, pp: 100–112.
- [5] Fei Daia, Man-Woo Parkb, Matthew Sandidge, Ioannis Brilakis, (2015). A vision-based method for on-road truck height measurement in proactive prevention of collision with overpasses and tunnels. *Automation in Construction* 50, pp: 29–39.
- [6] Stefania C. Radopouloua, Ioannis Brilakis, (2015). Patch detection for pavement assessment. *Automation in Construction* 53, pp: 95–104.
- [7] Jane Matthews, Peter E.D. Love, Sam Heinemann, Robert Chandler, Chris Rumsey, Oluwole Ol, (2015). Real time progress management: Re-engineering processes for cloud-based BIM in construction. *Automation in Construction* 58, pp: 38–47.
- [8] Xiaonan Zhanga, Nick Bakisa, Timothy C. Lukins, Yahaya M. Ibrahim, Song Wu, Mike Kagioglou, Ghassan Aouad, Ammar P. Kaka, Emanuele Trucco, (2009). Automating progress measurement of construction projects. *Automation in Construction* 18 pp: 294–301.
- [9] Seungjun Roha, Zeeshan Aziza, Feniosky Pena-Mora, (2011). An object-based 3D walk-through model for interior construction progress monitoring. *Automation in Construction* 20 pp: 66–75.
- [10] Kuo-Liang Lina, Jih-Long Fang, (2013). Applications of computer vision on tile alignment inspection. *Automation in Construction* 35 pp: 562–567.
- [11] Kevin K. Hana, Mani Golparvar-Fardb, (2015). Appearance-based material classification for monitoring of operation-level construction progress using 4D BIM and site photologs. *Automation in Construction*, 53 pp: 44-57.
- [12] Kuo-Liang Lina, Jih-Long Fang, (2013). Applications of computer vision on tile alignment inspection. *Automation in Construction* 35 pp: 562–567.
- [13] Pablo Rodriguez-Gonzalez Diego Gonzalez-Aguilera, Gemma Lopez-Jimenez, Inmaculada Picon-Cabrera, (2014). Image-based modeling of built environment from an unmanned aerial system. *Automation in Construction* 48 pp: 44–52.
- [14] Jie Gong, Carlos H. Caldas, (2011). An object recognition, tracking, and contextual reasoning-based video interpretation method for rapid productivity analysis of construction operations. *Automation in Construction* 20 pp:1211–1226.

- [15] Man-Woo Parka, Atefe Makhmalbaf, Ioannis Brilakis, (2011). Comparative study of vision tracking methods for tracking of construction site resources. *Automation in Construction* 20 pp: 905–915.
- [16] Man-Woo Parka, Ioannis Brilakis, (2012). Construction worker detection in video frames for initializing vision trackers. *Automation in Construction* 28 pp: 15–25.
- [17] Frédéric Bosché, (2010). Automated recognition of 3D CAD model objects in laser scans and calculation of as-built dimensions for dimensional compliance control in construction. *Advanced Engineering Informatics* 24 pp: 107–118.
- [18] Chao Wang, Yong K. Cho, Changwan Kim, (2015). Automatic BIM component extraction from point clouds of existing buildings for sustainability applications. *Automation in Construction* 56 pp: 1–13.
- [19] Mani Golparvar-Farda, Jeffrey Bohner, Jochen Teizer, Silvio Savarese, Feniosky Pena-Mora, (2011). Evaluation of image-based modeling and laser scanning accuracy for emerging automated performance monitoring techniques. *Automation in Construction* 20 pp: 1143–1155.
- [20] Samir El-Omari, Osama Moselhi, (2008). Integrating 3D laser scanning and photogrammetry for progress measurement of construction work. *Automation in Construction* 18 pp: 1–9.
- [21] Ioannis Brilakis, Manolis Lourakis, Rafael Sacks, Silvio Savarese, Symeon Christodoulou, Jochen Teizer, Atefe Makhmalbaf, (2010). Toward automated generation of parametric BIMs based on hybrid video and laser scanning data. *Advanced Engineering Informatics* 24 pp: 456–465.
- [22] Chang Chen-Yu, Chen Shi, (2016). Transitional Public-Private Partnership Model in China: Contracting with Little Recourse to Contracts. *Journal of construction engineering and management* 142(10) pp: 1-34.
- [23] Sascha E.A. Muenzinga, Bram van Ginneken, Max A. Viergever, Josien P.W. Pluim, (2014). DIRBoost—An algorithm for boosting deformable image registration: Application to lung CT intra-subject registration. *Medical Image Analysis* 18(3) pp: 449-459.
- [24] Jingjing Cao, Sam Kwong, Ran Wang, (2012). A noise-detection based AdaBoost algorithm for mislabeled data. *Pattern Recognition* 45(12) pp: 4451-4465.
- [25] Dalal N, Triggs B, Schmid C (2006). Human Detection Using Oriented Histograms of Flow and Appearance. *European Conference on Computer Vision*. Springer-Verlag pp: 428-441.
- [26] Prates R F, Cámara-Chávez G, Schwartz W R, et al, (2014). Brazilian License Plate Detection Using Histogram of Oriented Gradients and Sliding Windows. *International Journal of Computer Science & Information Technology* 5(6):39-52.
- [27] Smolka B (2011). Adaptive Edge Enhancing Technique of Impulsive Noise Removal in Color Digital Images. *International Conference on Computational Color Imaging*. Springer, pp:60-74.

AUTHORS

Xu Ruhang received his PhD in management science in School of Economics and Management in North China Electric Power University (NCEPU), China 2015. From 2015 until now he is a researcher and lecturer in North China Electric Power University. His research interest includes computing in construction and image processing.

