University of Alberta

HSV Color-Space Digital Image Processing for the Analysis of Construction Equipment Utilization and for the Maintenance of Digital Cities Image Inventory

By

Junhao Zou ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

In

Construction Engineering and Management

Department of Civil and Environmental Engineering

Edmonton, Alberta

Fall, 2006

Your file Votre référence
ISBN: 978-0-494-22420-5
Our file Notre référence
ISBN: 978-0-494-22420-5

University of Alberta

Library Release Form

Name of Author: Junhao Zou

Title of Thesis: HSV Color-Space Digital Image Processing for the Analysis of Construction Equipment Utilization and for the Maintenance of Digital Cities Image Inventory

Degree: Master of Science

Year this Degree Granted: 2006

Junhao Zou

Date: 2006/05/30

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **HSV Color-Space Digital Image Processing for the Analysis of Construction Equipment Utilization and for the Maintenance of Digital Cities Image Inventory** submitted by **Junhao Zou** in partial fulfillment of the requirements for the degree of **Master of Science in Construction Engineering and Management.**

David Zhu

Mohamed Al-Hussein

Sherif Ghali

Hyoungkwan Kim

Date: 2006/05/30

Abstract

The understanding of equipment utilization conditions is crucial for construction equipment deployment in construction projects. However, considering a project's budget, it is not feasible to visually observe the operation of construction equipment. An image processing-based methodology is presented in this thesis to automatically quantify the excavator idle time and truck load cycle time. The image color space, HSV (hue, saturation and value), is used as the base for image segmentation. Two applications are used to test this proposed methodology. Meanwhile, the proposed methodology is applicable for other applications. A study on the feasibility of building a city image database system is presented. The plan is to take images from a vehicle driving on streets, to detect objects automatically or semi-automatically by image processing, then to extract 3-D information of objects from these images. Finally, a 3-D model of the city will be built to record the city's infrastructure changes.

Acknowledgment

My supervisors, Dr Kim and Dr Al-Hussein, are greatly appreciated for their guidance during this research. I would also like to give special thanks to the construction engineering management group, at the University of Alberta, for their efforts in collecting image data applied in my first two research cases. Finally, I want to express my appreciation to Fangyi Zhou, Tong Zhou, Hongqin Fan and other individuals who help me during my research.

Table of Contents

List of Tables

List of Figures

Fig. 4. RGB ranges of a red excavator, a yellow excavator, and their background. 13

Chapter 1 Introduction

1.1 Motivation

Generally, labor, material and equipment are the three major resources required for construction projects. The success of construction projects depends greatly on the accurate construction equipment planning and on efficient site utilization. Minimizing construction equipment's idle time, for any deployed equipment, has been of major interest in the industry. Determining truck load cycle time is essential for optimizing the bunching of excavators and trucks in large scale excavations. Although empirical formulations are available to estimate the load cycle time, the result may not correlate with real-time construction. The collection of real-time data, regarding truck load time, is the prerequisite for accurate, instant and real-time optimum arrangement of trucks and excavators. Using manual observation to record the above information is possible technically. But it is not feasible in the aspect of economic efficiency. Cameras installed on construction sites monitor the construction activities produce abundant overlook images of the construction site. This provides an opportunity to apply an image processing-based approach to automatically measure construction equipment idle time and truck load cycle time.

An image database and 3-D model of a city is considered to be of assistance to a municipality's operation and control managers. To build such a model, a survey should be performed to acquire the location and dimensional information of buildings, trees, streets and traffic signs. Those survey results should be inputted into a database management system, which will allow for a convenient query and an accurate 3-D model to be built. Thus, a municipality has to allocate a great deal of time and funds in order to build such a system. If the municipality wants to update the data and identifies the changes in a time interval, all the work should be repeated again. As a result, maintaining and operating such a system may be unaffordable or, at the very least, bring a huge burden to the tax payer. The development of image processing techniques makes it possible to build an efficient and cost saving image database and 3-D model to perform the same function. A methodology is proposed to extract 3-D information from images that are shot from a vehicle at various streets. If the whole system can be implemented successfully, it will bring tremendous convenience to city management.

1.2 Objectives

- ✧ Develop a methodology to identify objects in color images.
- ✧ Test the feasibility of applying image processing to construction equipment management.
- ✧ Study the feasibility of a city image database system.

1.3 Methodology

In this report, the applications of image processing, which are excavator idle time analysis, truck load cycle time analysis and city image database systems, follow a general procedure to reach application purposes, as shown in Fig. 1. Collecting image data is the first step of the image processing application. The image data can be images taken from a fixed location with a constant time interval, or images taken from mobile locations. The next step is extracting mobile or fixed targets based on objects' features, such as color (hue, saturation and value), shape, size and location information. The subsequent step is determining the existence status of targets and measuring the location and/or geometric dimensions of targets. Only location information is extracted in the first two applications and both location information and dimension information are discovered in the last application. The final step is reaching meaningful explanation, such as idle time of an excavator, load cycle time of different classes of trucks and status of buildings and traffic signs. This is accomplished by analyzing the location information and geometric dimension of targets and reasonable inference. The details of the methodologies will be discussed in chapter 3.

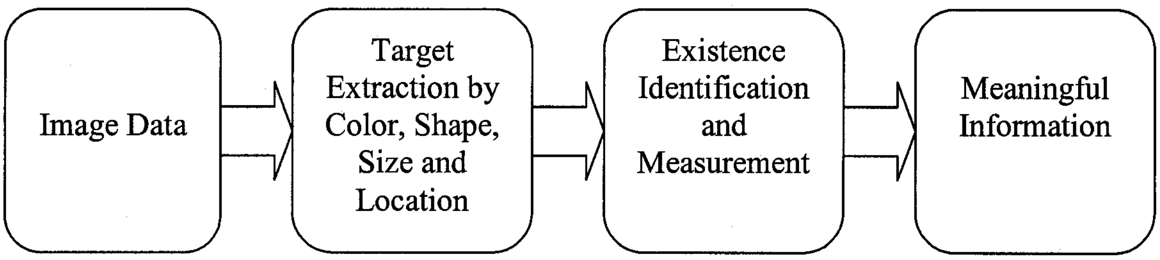


Fig. 1. My research's general procedures.

1.4 Report Organization

This report includes all materials used in this research. Since three application cases are introduced in this thesis, each application is summarized in a chapter. Excavator idle time analysis is described in Chapter 2, truck load cycle time analysis is presented in Chapter 3 and the feasibility study of city image database is summarized in Chapter 4. The literature review and current practice, the proposed methodology and the implementation, and the application results analysis of each application are discussed respectively in those chapters. The thesis is concluded in Chapter 5 and future works are proposed in the same chapter.

Chapter 2 Excavator Idle Time Analysis

2.1 Literature Review and Current Practice

Image processing technology is developing rapidly; image segmentation/feature extraction for measurement, as one of image processing areas is applied in a variety of application fields, such as astronomy, semiconductor production, traffic control, oil exploration, agriculture, medicine diagnosis and weather forecasting.

Today, researchers are attempting to introduce image segmentation/feature extraction for measurement within the civil engineering industry in order to fulfill some tasks with a higher efficiency. At present, the main focus is on the area of material characterization and road surface monitoring (Haas et al. 2001; Kim et al. 2002; Kim et al. 2004; Chandan et al. 2004; Offrell et al. 2005). In this research, image segmentation is applied to construction equipment management area, which is a new area for this technique. The innovative image segmentation and the latest target tracking are reviewed in the following paragraphs.

2.1.1 Review of the State of the Art of Image Segmentation

"Image segmentation is a fundamental step in most of the applications of image analysis" (Iannizzotto and Vita, 2000). A variety of researchers contributed to image segmentation techniques, such as combining thresholding segmentation, intensity based segmentation, spatial relationship based segmentation and region growing algorithm. Other researchers focused on image segmentation strategies to create techniques that can increase the stability of image processing applications. For example, utilizing image segmentation results by different image segmentation algorithms, a strategy which is named as data fusion, is used in robotic and automated sensing systems (Haas, 1990), in 3D image segmentation of aggregates from laser profiling (Kim, 2003) and in semi-automatic ongoing construction progress assessment system (Wu, 2005).

Despite a variety of researches that are carried out in image segmentation, there are few research efforts in automatic construction site monitoring since target extraction is hard to achieve by virtue of the complex background in the field. Due to the types of cameras being used and the research's tendency to avoid complex imaging computation in favor of rapid modeling, color information from construction fields is not incorporated into developing methodologies due to the types of cameras being used and the research's tendency to avoid complex imaging computation in favor of rapid modeling. For example, Wu and Kim (2004) used a grayscale image-based data fusion approach to detect concrete columns on a general building construction site. In order to reduce the effects of snow and poor light conditions, an image lighting compensation methodology was introduced. Data fusion was applied based on the following two powerful algorithms: canny edge detection and watershed transformation. An AutoCAD 3D perspective view of objects of interest was created to work as an image filtering mask. This experiment testifies the effectiveness of an image segmentation methodology in such an application.

However, color can provide rich information about a construction field to facilitate a correct understanding of the scene. Considering the enormous benefits of using color to accurately represent outdoor environments, therefore, this valuable technique cannot be ignored. Neto and Arditi (2002) used color information in the RGB (red, green and blue) color space to identify a structural component (pier structure) on a bridge construction site. An RGB-based edge detect algorithm is created for image segmentation. "A library of RGB ranges for different types of construction materials such as concrete, steel and aluminum" (Neto and Arditi, 2002) which is called "Construction Materials Research (CMR)" (Neto and Arditi, 2002) is built to detect various materials on the construction site. When an image is read, CMR can detect the edge of the wanted construction elements through existing RGB information about the specific material. If the CMR cannot detect the object, a user is allowed to crop a small sample from the original image in order to store the RGB ranges for this kind of material until the objective is detected successfully. As a result, this new RGB ranges for this specific material becomes a part of the library for future use. The problem regarding this methodology is the RGB ranges of the same material change tremendously in different light conditions.

When using color information as a measure of image segmentation, various color spaces play an important role in image segmentation. RGB, L*a*b*, L*u*v*, YES, YIQ and HSV color coordinates are all tried to perform the base of image segmentation in a variety of research areas. A recent research (Takahashi et al. 2005) compares the effects of different color coordinate systems in "uniformity of the color distribution, color concentration, color separation ability and local color distribution in regions of images" (Takahashi et al. 2005) shows that "there does not exist one single color coordinate system that is appropriate for all image processing tasks and that the most suitable color coordinate system will be dependent on the particular processing task at hand" (Takahashi et al. 2005).

HSV, as a color space close to human perception, catches more attention in recent image segmentation researches. One research project (Kramberger 2005) which was designed to detect skin color objects applied the HSV color coordinates as the base of image segmentation. The research shows skin color only occupies a small range of hue. With the help of the HSV color space, a real time video skin feature extraction was achieved.

2.1.2 Review of the State of the Art of Target Tracking

With the help of images shot by a fixed camera, it is easy to identify the movement of an object if there is only one extracted object in an image data series. However, if there are numerous objects extracted by target segmentation algorithm, a target tracking strategy has to be designed. Target tracking is a common topic of image processing application in traffic control areas. As a result of that there are several vehicles that are commonly extracted in an image. If the size and shape of a target do not change tremendously, normalized cross-correlation (NCC) can work well. Behrad et al. (2001), in their research, attempted to find corresponding feature points on detected horizontal and vertical lines by virtue of a vehicle's "geometrically well-defined structures" and to calculate the disparity vectors.

Lou et al. (2005) designed a system to track a specific vehicle by matching the detected edge of an object with the 3-D wire-frame model of a specific vehicle. Then, an estimate of the vehicle's position in the next image is made by calculating the speed and orientation of that specific vehicle in current image. Thanks to the high frame rate of the testing image data (15 frame/s), the program achieved a satisfying experimental result.

Although there are a few image processing trials within the civil engineering industry, there is still few image processing research regarding mobile targets within construction management applications. One important feature common to Wu (2005) and Neto and Arditi's (2002) studies is that the objects of interest in the image data are stationary. For example, in the case of column construction, once the location of a particular column is identified, and if the monitoring camera's angle and location are not changed, the same image region can be assumed to contain the same column's information. However, this assumption is invalid in identifying mobile objects in image data. Mobile objects, such as construction equipment, change their location in image data over time. Moreover, since three-dimensional equipment is being reflected in the two-dimensional image space, the shapes that are projected in the two-dimensional image can also vary over time when the equipment changes its direction of movement or rotates. These characteristics of mobile object monitoring, especially in an outdoor environment, present a large challenge in the effort to automatically analyze equipment idle time by using image processing.

2.2 Proposed Methodology and Implementation

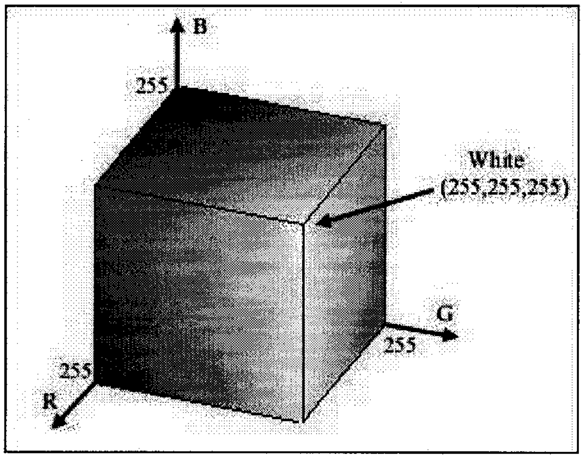
An image processing-based methodology for analyzing the idle time of construction equipment is presented in this section. Image data obtained from the early phase of the construction of the Natural Resources Engineering Facility (NREF) at the University of Alberta in Edmonton, Canada, are used to test the presented methodology. During the foundation excavation activity, images of a hydraulic excavator operating on the construction site were taken every 10 seconds. By detecting the motion of the equipment on the consecutive images, an estimation of equipment idle time was achieved. The main technical issues to be considered in accurately estimating equipment idle time is how to distinguish the hydraulic excavator from its complicated background region, including the land with natural soils, the land covered with snow, and other equipment; how to trace the hydraulic excavator through all of the images taken at 10-second time intervals and how to use a certain indicator to determine the status of equipment movement, i.e., whether the equipment is idle or not. The details of the proposed methodology are described in the following parts.

2.2.1 Comparison of RGB and HSV Color Spaces

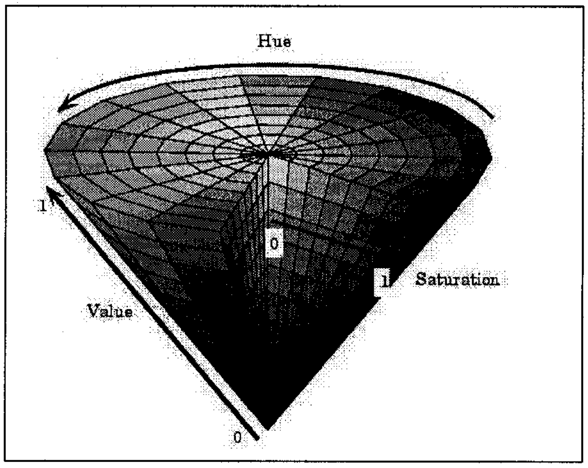
There have been several attempts to use color image information represented in RGB space for construction project monitoring (Neto and Arditi 2002; ElGhandour 2003). In RGB images, the color of a pixel is determined by the combination of different red, green and blue values. An RGB image with $m \times n$ pixels has an $m \times n \times 3$ data array since each pixel should have the three different color components. If an eight-bit of memory space is allocated for each color component of a pixel, a total of 16,777,216 colors (256^3) – so called "true color" – are represented. Depending on the application need, a different amount of memory space is allocated for each pixel, which results in a more detailed color information. RGB color space is the most widely used color model since it can be easily implemented with electronic display devices using its additive nature. Fig. 2(a) shows how different combinations of RGB components can create true color.

In the HSV color space, hue refers to the color that the viewer recognizes, such as red, yellow and green, which can be regarded as representing the main wavelength of the perceived color. Saturation refers to the degree the color is mixed with a white component. It ranges from unsaturated (shades of gray) to fully saturated (no white component), giving rise to, for example, "light purple" and "dark purple" (Sonka et al.

1999; The MathWorks 2003). Value indicates the brightness of a color. Fig. 2(b) illustrates the cone shape of HSV color space, with hue, saturation and value varying from 0 to 1.



(a) RGB

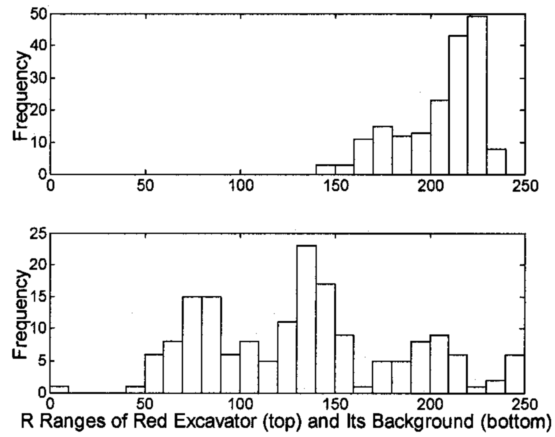


(b) HSV

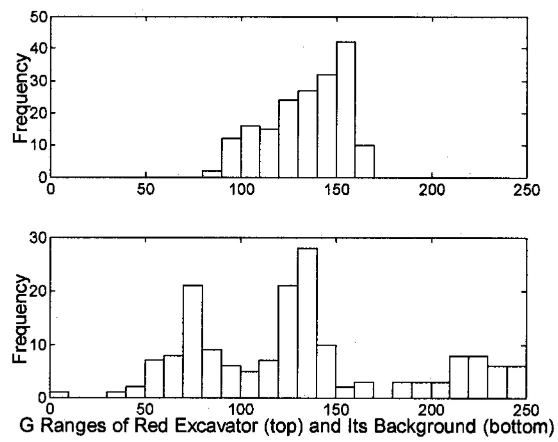
Fig. 2. Comparison of RGB and HSV color spaces (The MathWorks 2003).

RGB color space is easy to understand and effective in displaying sophisticated colors by combining the three primary components: red, green, and blue. However, the additive nature of the RGB color model is not suitable for major image processing tasks such as object identification and extraction. To successfully identify and extract an object of interest in image processing, it would be ideal for the object of interest to have a range of colors significantly different from those of the background region. That way, a simple threshold value can be chosen to separate the object of interest from other regions, resulting in successful object extraction. Fig. 3 shows histograms of RGB values for an object of interest (a red excavator) and its surrounding background region in an image of the NREF building construction site. The RGB color distributions of the red excavator region overlap with its background region, making it difficult to choose a discriminating threshold value in any domain of the three colors. Fig. 4 shows two three-dimensional RGB color representations of another NREF building construction image. This image contains both a red excavator and a yellow excavator operating at the same site. Each point in Fig. 4 represents the color of a pixel of the image which belongs to either the red excavator, yellow excavator or background regions. The straight lines in Fig. 4 represent planes parallel to our line of sight that separate the color space into two regions. The plane in Fig. 4(a) divides the color space into excavators and background, whereas the plane in Fig. 4(b) separates both the red and yellow excavators. Both of the two planes were obtained by manually rotating the RGB color space to find the optimum division results. Fig. 4 indicates that although it is technically possible to find some optimum division spaces to classify different image objects, it would require a complex solution, such as non-linear artificial intelligent function, and require a significant amount of training data. Fig. 5 illustrates the significant changes in RGB values and the difficulty in finding ratios of RGB value to represent the color of the excavator when lighting condition changes from brightness (sunny) to darkness (cloudy). This is an indication that

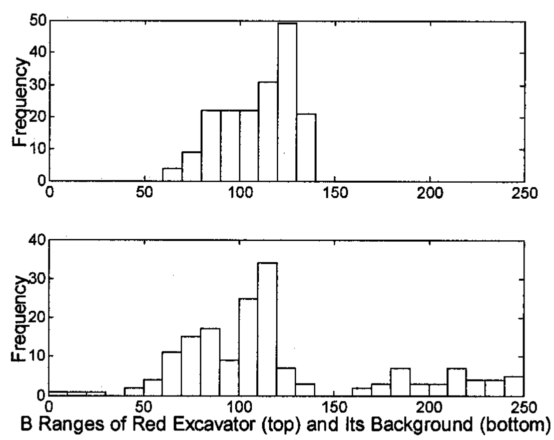
RGB color space is a feeble option for the purpose of object identification and extraction in an open construction field.



(a) Red

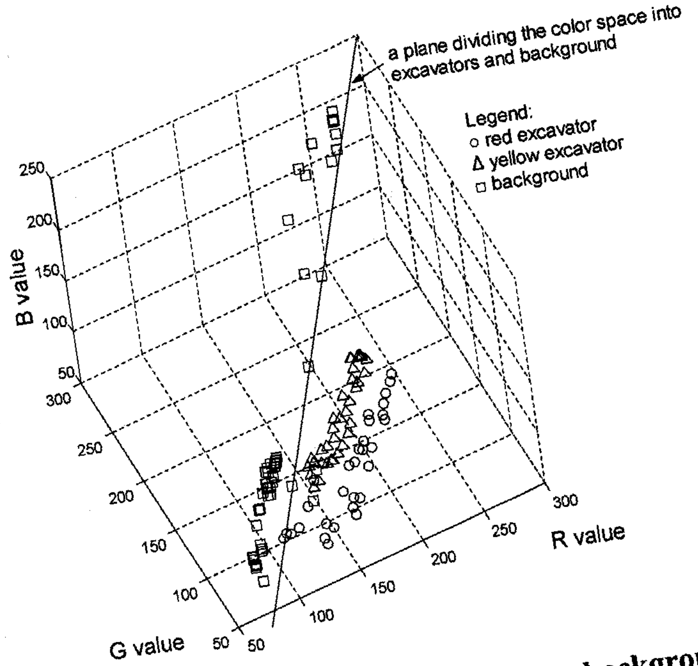


(b) Green

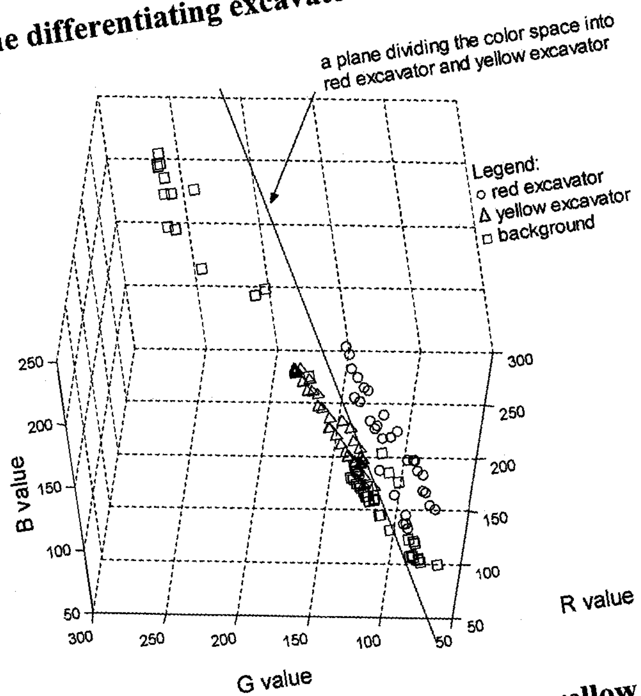


(c) Blue

Fig. 3. RGB ranges of a red excavator and its background.



(a) With a plane differentiating excavators from background



(b) With a plane differentiating red excavator from yellow excavator

Fig. 4. RGB ranges of a red excavator, a yellow excavator, and their background.

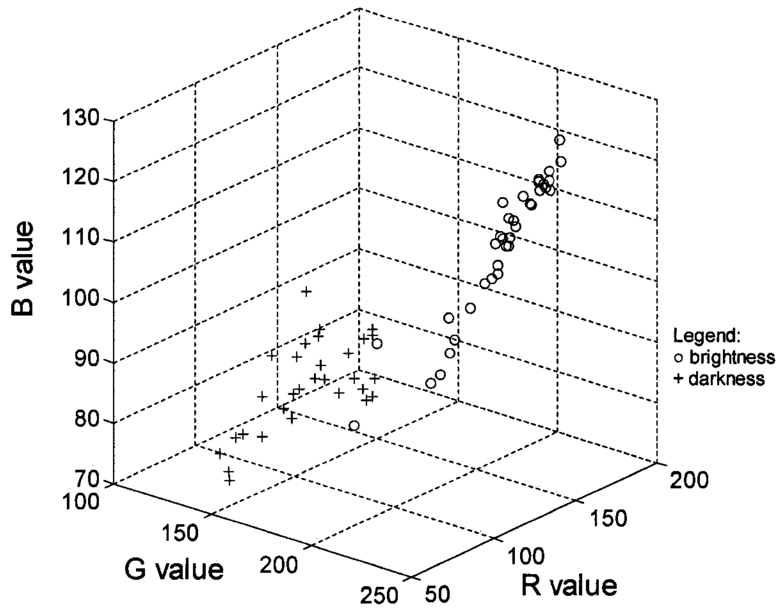
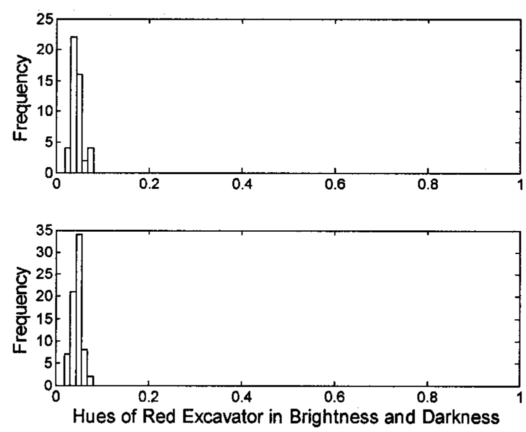
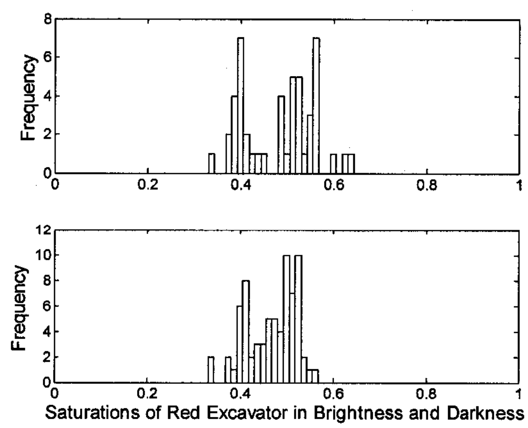


Fig. 5. RGB of the red excavator in brightness and darkness.

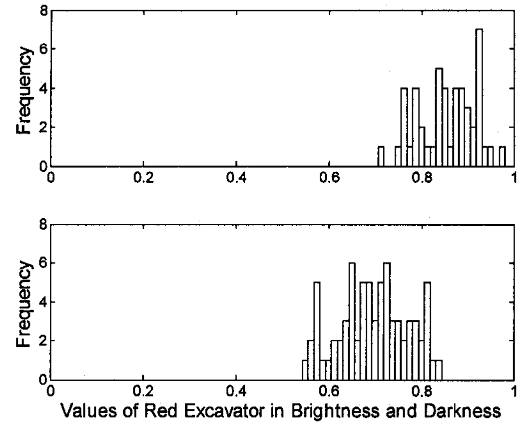
In HSV color space, the hue component plays a major role in identifying color information since it represents the dominant wavelength of the perceived color. The characteristics of hue are relatively less sensitive to different lighting conditions, making it more suitable than RGB components for identifying and extracting objects of interest in images captured in open fields. In this study, the hues of an object of interest (red excavator) in 30 images taken in cloudy weather are compared with the hues of the same object in 30 images taken in sunny weather. A two-sample Kolmogorov-Smirnov test was conducted to see if the two distributions (hues of the red excavator in dark lighting condition vs. bright lighting condition) are significantly different. The null hypothesis that the two datasets of hues are not significantly different from each other was accepted at a significance level of 5%. The means of hues, on a 0 to 1 scale, were 0.0458 and 0.0458, in bright and dark conditions respectively, while the means of values were 0.8551 and 0.6966, which shows more discrepancy caused by different lighting conditions. Thus, hue is considered to be a more robust indicator to identify the same color. Figs. 6(a), (b), and (c) are the quantitative distributions of hue, saturation, and value, respectively, of the red excavator in bright and dark conditions (using the same pixel samples as in the Kolmogorov-Smirnov test), which illustrate that hue is a more stable indicator of the red excavator. Hue is a useful feature for separating objects with different colors. Fig. 7 is the quantitative distributions of hue for a red excavator and a yellow excavator in an image. This illustrates that there is little overlap of hues between the two different colors of excavators, indicating that a simple threshold method can differentiate the two pieces of construction equipment.



(a) Hue's quantitative distributions



(b) Saturation's quantitative distributions



(c) Value's quantitative distributions

Fig. 6. HSV's quantitative distributions of a red excavator in brightness and darkness.

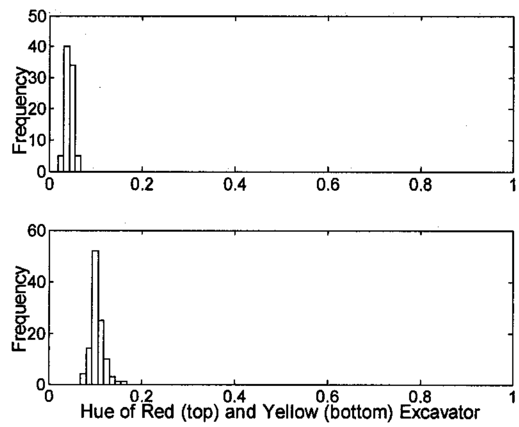


Fig. 7. Hue's quantitative distributions of a red excavator and a yellow excavator.

In many images of the NREF building foundation excavation activity, the main components of the background region are dark colored soil and white snow, as shown in Fig. 8(a). Since HSV color space allocates almost identical saturation values (close to zero) for soil (black) and snow (white), identifying and extracting the construction equipment object from the background region in image processing can be achieved easily with saturation (Fig. 8(c)). Fig. 9 illustrates another advantage of using saturation for equipment object extraction in images of early phase construction activities that contain soil and snow. Fig. 9(a) shows an original image of an excavator in the typical background of black soil and white snow. Using a simple threshold method, Fig. 9(b) and 9(c) show the excavators extracted by hue and saturation. The equipment object detected by the hue has a zigzag outline, while the equipment edge detected by saturation is much smoother. The comparison of Fig. 9(b) with 9(c) demonstrates that saturation can be a more stable indicator than hue in differentiating construction equipment from the background of soil and snow. It is important to consider that saturation may not always be the best indicator in all cases.

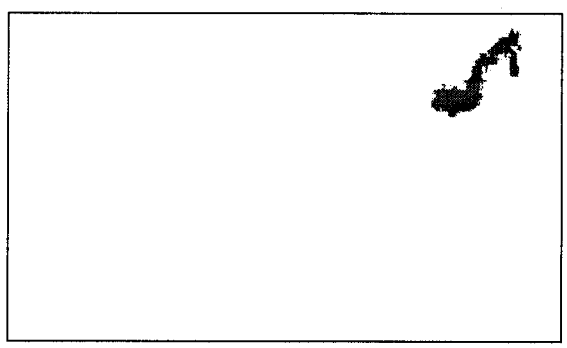
(a) Original image

(b) Grayscale image

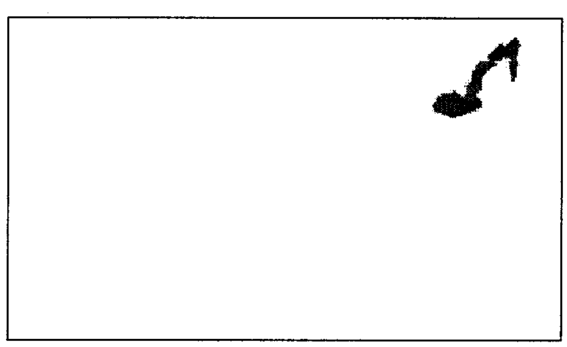
(c) Saturation image

Fig. 8. A sample image of an excavator in the background of soil and snow.

(a) Original image



(b) Extracted excavator by hue



(c) Extracted excavator by saturation

Fig. 9. A sample image of an excavator.

To sum up, hue is less sensitive to different lighting conditions and has the ability to separate objects of different colors. Saturation is also helpful in distinguishing construction objects from the background, since construction objects such as equipment and some materials tend to have higher saturation numbers than the background. Value is helpful to distinguish black from white objects whose hue and saturation have a value close to zero. Based on the above analyses, it is suggested that HSV color space, rather than RGB color space, is best suited for object segmentation in images of open fields. In Takahashi's (2005) research, HSV was also found to be an effective color space.

2.2.2 Detail Methodology for Excavator Idle Time Analysis

Before the details of the proposed methodology for equipment idle time analysis are described, it is important to review some characteristics and assumptions of this research.

First, the resolution of the original image data is 640×480 and the image quality is not desirable due to the small number of pixels that belong to the excavator. In this situation some noise reduction techniques, which are intended for improving image quality, can actually tarnish the important features of the excavator. Thus, image filtering operations are performed as sparingly to keep more accurate information about the hydraulic excavator. The performance of fewer image processing operations is also desirable for improving image processing efficiency.

Second, since the time interval between every two successive images is 10 seconds, it is assumed that the color range of the hydraulic excavator does not change significantly during the time interval. The color information can then be used to identify and trace the equipment movement in consecutive images.

Third, it is assumed that hydraulic excavators move sufficiently slowly so that an excavator on an image is the same as the excavator found in the previous image, as long as they have reasonably similar color and size features. This assumption eliminates the possibility that another excavator with the same color replaces the excavator of interest during a given 10-second time interval.

Finally, the most important assumption made for idle time estimation is that the hydraulic excavator is treated as working for 10 seconds, so long as the next consecutive image shows a significantly different position or orientation. This may cause some errors in the estimation results compared to the actual working time, which could be less than 10 seconds. However, the errors can easily be reduced by shortening the time interval between two consecutive images, that is, by increasing the frequency of image capturing.

1) **System Main Process**

In this system, the input parameters include image data with shooting time, and target features information. Image data in this application are images captured by a fixed digital camera on a building near the NREF construction site. Target features information means the information about the HSV ranges, size and location of the target red excavator. Tools and techniques in this system include image processing tools such as object segmentation tools, based on HSV, size information and object tracking tools. Criteria refer to the thresholds that help to identify targets. They include thresholds of hue, saturation, value, size and location. The output of the system is the excavator idle time information. The relationship of this proposed system is shown in Fig. 10.

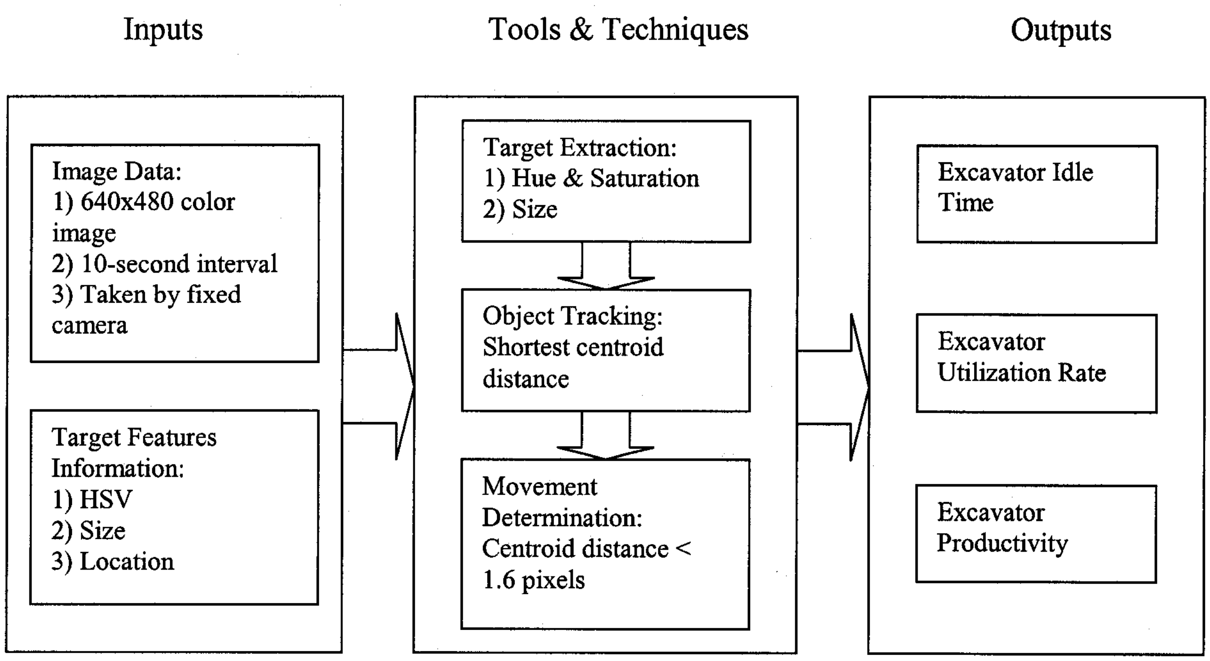


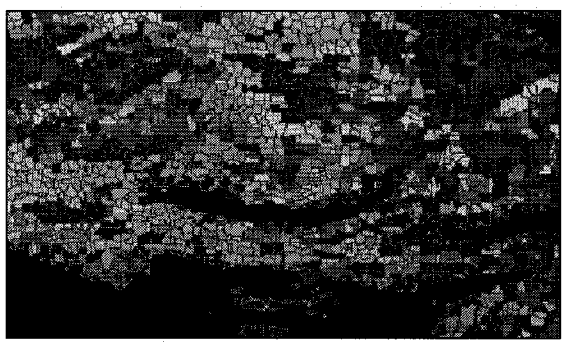
Fig. 10. System main process of excavator idle time analysis.

2) **Target Extraction**

Target (object) extraction is one of the most challenging and important tasks, as is the case in many other image processing researches. Successful separation of the object of interest from its background region can significantly increase the accuracy of the image processing analyses. Traditionally, color information from construction fields is not incorporated into developing methodologies due to the types of cameras being used and the research's tendency to avoid complex imaging computation in favor of rapid modeling. However, color can provide rich information about a construction field to facilitate a correct understanding of the scene. The enormous benefits of incorporating color to accurately represent outdoor environments cannot be ignored. Figs. 8(a) and 8(b) are sample color images of a hydraulic excavator and its grayscale version. As can be seen from Fig. 8(b), it is extremely challenging to extract the equipment in the grayscale image since the grayscale of the hydraulic excavator is very close to that of its background. Thus, adjusting the grayscale of images to highlight the hydraulic excavator is not considered to be an appropriate approach.

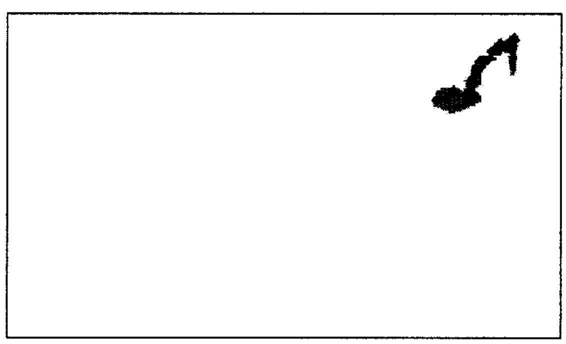
Due to the disadvantages of RGB color space that were previously discussed, HSV color space is used to extract the excavator of interest in the image data of foundation excavation activity. Saturation information is used to separate the excavator from the background, while hue information is used to separate excavators with different colors. As shown in Fig. 8(c), saturation images facilitate the target segmentation process by suppressing most of the noise and making the hydraulic excavator the most noticeable object. The application of other sophisticated algorithms such as Canny Edge Detector (Canny 1986) and Watershed Transformation (Beucher 1991) to the saturation images could not achieve a satisfactory segmentation result unless some pre- and/or post-image processing were conducted. These preliminary tests indicated that setting a threshold in the HSV color space is the most efficient way to extract the hydraulic excavator. Figs. 11(a), 11(b) and 11(c) show the segmentation results of the aforementioned three methods: Canny Edge Detector, Watershed Transformation and saturation-based thresholding, respectively. After removing the noises with fewer pixels and filling the holes in the object of interest, the target extraction is achieved. Fig. 11(d) shows the final result.

(a) Canny edge detector



(b) Watershed transformation

(c) Thresholding



(d) Thresholding with noise removal

Fig. 11. Comparison of image processing results on the saturation image.

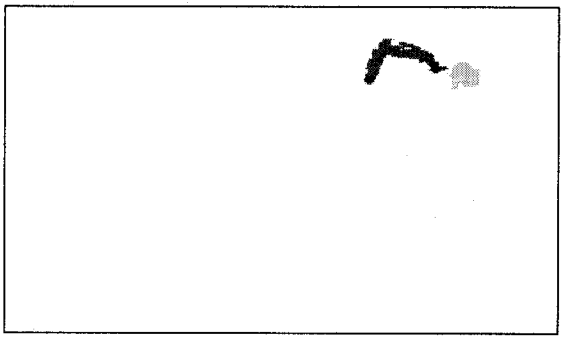
3) **Object Tracking**

After target extraction, tracing the hydraulic excavator becomes the main problem. Generally, such features as the unique color range, shape and size of the object of interest are identified and used for this type of object tracking. However, color range information could not be solely used as an object identifying feature in this application since other equipment with a similar color range can appear on the job site. The frequently changing shape of the hydraulic excavator with different orientations on the image data also indicates that shape feature extraction techniques, such as normal cross correlation function, would not work for the purpose of object tracking. The area (number of pixels) of the hydraulic excavator also varies significantly when the equipment changes its direction of movement or appears in different distances from the camera.

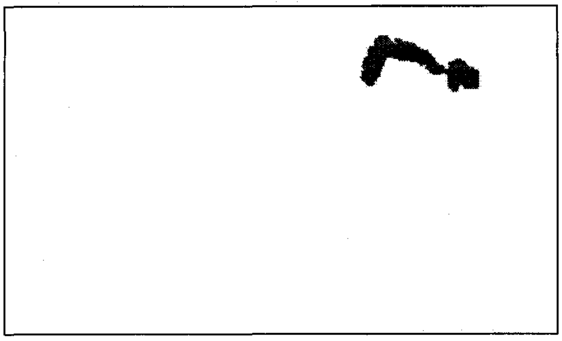
To address the aforementioned concerns, the concept of distance between objects is used in this study. Once the location of the hydraulic excavator in the current image is identified, it is assumed not to change significantly in the next image (10 seconds later). Therefore, the object in the current image that has the closest centroid distance to the hydraulic excavator in the preceding image is identified as the same hydraulic excavator. In other words, the centroid of each object in the current image is calculated and the object with the centroid closest to that of the hydraulic excavator in the previous image is identified as the same piece of equipment. Although this approach generally showed good results, when two hydraulic excavators with similar color were positioned very close to one another, their identities were sometimes confused. To avoid these extreme situations, image data based on shorter time intervals (such as 5 seconds), or on additional features of the equipment should be used. In the study reported here, this extreme situation does not happen. For safety reasons, other excavators are not allowed to operate in close range of the boom of the excavator of interest.

Other problems can arise in object tracking, as illustrated in Fig. 11. When the hydraulic excavator swings its boom in a certain direction (Fig. 12(a)), the cab may divide the hydraulic excavator into two parts, as shown in Fig. 12(b). To remove this effect, which influences the final equipment idle time estimation, image dilation is performed when the segmented regions are in close proximity to one another. In this study, a 3x3 structuring element was used for image dilation. Fig.12(c) shows the correct segmentation result after the dilation operation.

(a) Original image



(b) Extraction error



(c) Image dilation result

Fig. 12. Extraction of a hydraulic excavator in a certain direction.

4) **Movement Indicator**

Determining if the equipment is actually operating is another important step in measuring equipment idle time. For this purpose many region properties, such as the coordinates of top-left, top-right, bottom-left and bottom-right, and the length and orientation of major and minor axes could theoretically be used. However, in reality, these options are unreliable movement indicators since any small effects caused by different lighting and weather conditions can lead to unexpected analysis results. Therefore, the centroid coordinates of the segmented region were tested as the simplest indicator of movement.

Selection of the movement indicator threshold value is another important factor that affects the accuracy of idle time estimation. Since perfectly stationary equipment can have slightly different centroid coordinates in consecutive images due to minor trembling of the camera, a certain threshold value in the form of a pixel distance must be chosen to determine whether or not the equipment moves. Thus, when the distance between the locations of an object's centroid differs by more than the threshold value between two consecutive images, then the object is considered to have moved. Too big a threshold value will result in exaggerated equipment idle time, while too small of a value will generate an underestimated idle time. After trials of 300 images in three different time sections, the criterion for equipment movement was set to be a distance of 1.6-pixels.

5) **System Architecture**

An image processing system was developed using MATLAB to automatically estimate the idle time of equipment. A total of 60 sample images were studied to calibrate the parameters required for the system. The processing steps are as follows (Fig.13):

1. The object of interest (equipment) on the first original image is manually selected on a computer screen. The system uses the image pixel coordinates, which are indicated by the mouse cursor location, to identify the equipment to be analyzed. This step also requires the user to specify the number of images that will be used in the analysis. As previously mentioned, one image in this study is considered to represent the equipment status for a period of 10 seconds.
2. The next image (the image captured 10 seconds later) is selected to assess the operation status.
3. Part of the image, with 450×260 pixels, which covers the whole construction site, is kept for further processing. The rest of the image, which includes several nearby buildings and the roads outside the construction field, is removed for faster and more efficient image processing.
4. Other equipment with colors different from that of the equipment of interest is removed using hue feature-based thresholds.
5. The saturation image is extracted from the original color image using HSV color space.

6. The saturation image is converted to a binary image using a predetermined threshold value.
7. Small noises left in the binary image are removed using the size feature of each noise.
8. An examination to determine if there are two objects in very close proximity to each other is conducted. If yes, image dilation is processed. This step is to prevent one object from being identified as more than one.
9. Holes in identified objects are filled using image morphological functions.
10. The centroid coordinates of the identified objects are calculated.
11. The centroid coordinates of the objects in the current image are compared with the centroid coordinates of the equipment of interest in the previous image. The object in the current image that is closest to the equipment in the previous image is now regarded as the equipment of interest.
12. The working status of the equipment is determined using a predetermined threshold value of centroid distance.
13. Step 2 to step 12 are repeated until the last image in the analysis period is processed.
14. The total idle time of the equipment is calculated based on the number of images that show idle status.

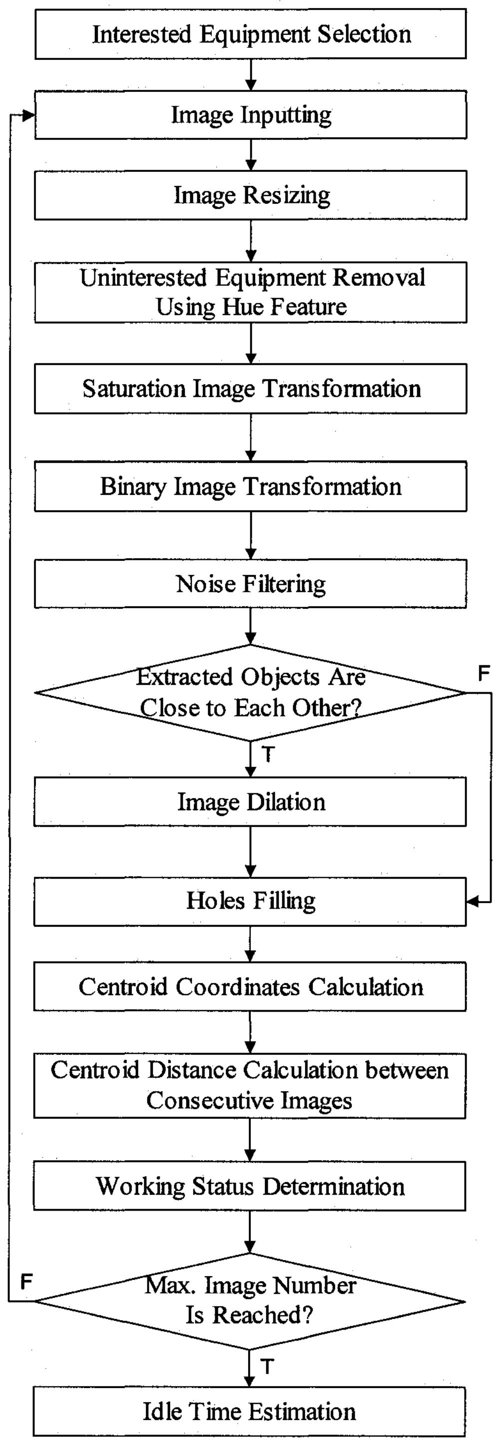


Fig. 13. System architecture of excavator idle time analysis.

2.3 Application Results Analysis

In the first application, Excavator Idle Time Analysis, a three-hour long image data obtained from the early phase of the construction of the NREF at the University of Alberta in Edmonton, Canada is used to test the algorithms for this application. The methodology of the application determines that the idle time of the excavator will be slightly underestimated. When the excavator stops during the interval between two successive images, the program will count the whole interval as working time for the difference of those two images. Fortunately, that factor would not bring the program much error since the interval is only 10 seconds.

The proposed methodology described in Section 3.2, provided a high degree of accuracy in excavator idle time analysis for a three-hour long image series (a total of 1080 images). The results show that the hydraulic excavator works 7210 seconds (2 hours 10 seconds) during the 10800-second (three-hour) investigation period, with a working rate of 66.8% and an idle time of 3590 seconds (59 minutes 50 seconds). The real working rate by visual observation of the original images is 67.0%. This gives a discrepancy of only 0.2%. However, a further analysis of the processing results of each image shows that there are 36 images (3.3% error rate) that are incorrectly processed. In those errors, some idle periods are treated as working periods, while some actively working periods are identified as idle. The explanation for the overall error rate being only 0.2% is that those two kinds of errors neutralize each other. In fact, some errors are unavoidable because it is difficult to determine the suitable threshold value for the movement indicator. This threshold value depends on various factors, such as the size of equipment, camera resolutions, lighting conditions and image processing methodology.

Furthermore, the distance between the hydraulic excavator and the camera also affects the threshold value. For example, a threshold value suitable for detecting small movement of the hydraulic excavator 50 meters away from the camera may not be suitable for an excavator 100 meters away. Therefore, a careful calibration process should be conducted before determining the threshold value in various cases. For this study, either 3.3% or 0.2% is considered to be accurate enough for project management practice.

Chapter 3 Truck Load Cycle Time Analysis

3.1 Literature Review and Current Practice

Cost related to the excavators, trucks and their operators is a main component in large scale earth work projects and mining operations. However, it is still common to see excavators or trucks stopping on the field for various reasons. Improving the equipment utilization rate is an important development in reducing equipment costs. Due to the complexity of equipment deployment, few researches have focused on this topic. Current practice of construction equipment planning is based on empirical formulations. For example, truck load time is estimated by multiplying the number of bucket swings by bucket cycle time and the balance of excavator and truck fleet is reached based on the estimation of load cycle time, haul time, return time and dump time.

The disadvantage of the above method is that it could not account for abnormal or unforeseen conditions, which always arise on real constructions sites. A new developed technique is to optimize the equipment deployment by building a simulation model, which can introduce some uncertainties in the real practice. Smith et al. (1995) applied discrete-event simulation to earth-moving systems. In that research six different factors, which were number of trucks, passes per load, mean of load pass time, variability of load pass time, mean of travel time and variability of travel time, were introduced to this research, and then levels of each factor were set. Utilizing these simulation tools, the most sensitive factors were identified for the purpose of efficient construction planning. Currently, with the development of computer technologies and simulation programs, it is easy to simulate construction progress by combining specific statistical distributions of all possible factors. Another recent research project that was carried out by Feng and Wu (2006) explores how to make a more efficient delivery plan for a ready mixed concrete plant. In this research, the CYCLONE simulation program is introduced to optimize the design of a truck delivery schedule for orders from a variety of sites to meet the up-to-date requirements of "JIT (Just in Time), QR (Quick Response) and BTO (Build to Order) in the construction industry. With the achievement of that reported research, the batch plant manager can obtain an optimized dispatching schedule to reach a more efficient truck deployment.

Simulation software, such as Simphony and CYCLONE, allows simulation to become more popular and easy to implement. The trend of construction plan simulation allows for considerations, regarding uncertain factors, by introducing probability distribution. The determination of the forms and parameters of probability distributions of the factors in the simulation model is mainly based on manually recorded statistical data or experience. As a result, if there is a methodology that can collect statistical data automatically, the application of simulation will be further facilitated.

A research carried out by El-Ghandour et al. applied image processing to automated acquisition of the statistical data of truck load cycle time. They proposed a target extraction methodology based on thresholds of three rates of "H1=R/G, H2=G/B, H3=R/B" (El-Ghandour et al.) and tested it by image data of a blue truck. But from Fig. 5, the line of best fit of RGB scatter of the red excavator in different light conditions does not pass through the original point of (0, 0, 0); thus those rates between R, G and B spread in big ranges. That means more noises might be included in target extraction result. So this methodology might not be suitable for extracting objects with other colors.

3.2 Proposed Methodology and Implementation

A 2-hour long NREF building construction image data series is used to experiment with the methodology for truck load cycle time analysis. The original images are also 640×480 images with an interval of 10 seconds between two consecutive images. The research purpose of this application is to collect the real load cycle time data of trucks with different dump box sizes. For this study, truck load cycle time is defined as the time period between a truck's coming into the location for loading and its leaving for unloading. Thus, it includes the time period that the excavator, or the truck, stops for any reason. This statistical data can include some abnormal situation which can happen on the site, but it is ignored in the calculation.

3.2.1 System Main Process

In this system, the input parameters include image data with shooting time, and target features information. Image data are images captured by the same digital camera as in the first application. Target features information refers to the information of the HSV ranges, size, shape and location of the trucks. Tools and techniques in this system are image processing tools, which include object segmentation tools based on HSV, shape and size information and object tracking tools. Criteria refer to the thresholds which help to identify targets. They include thresholds of hue, saturation, value, size, shape and location. The output of the system is the truck load cycle time information. The relationship of this proposed system is shown in Fig. 14.

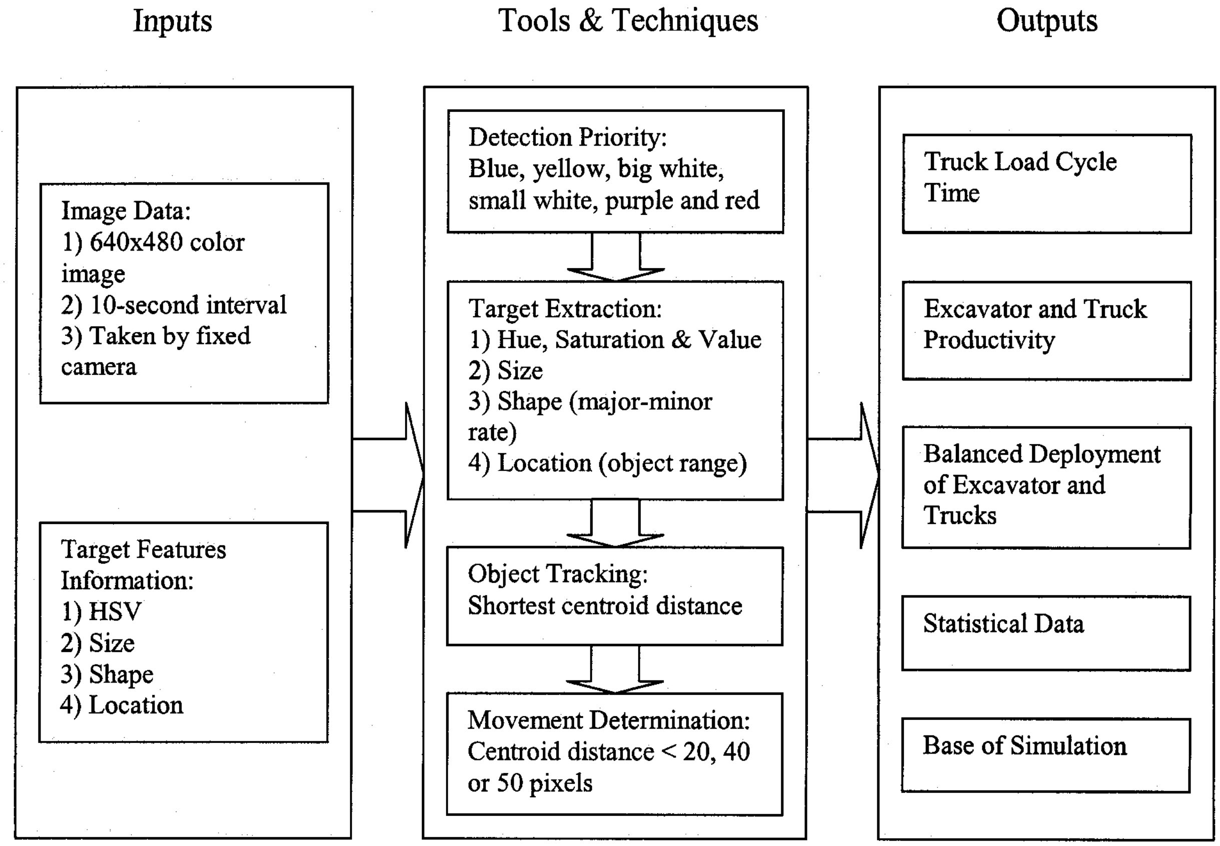
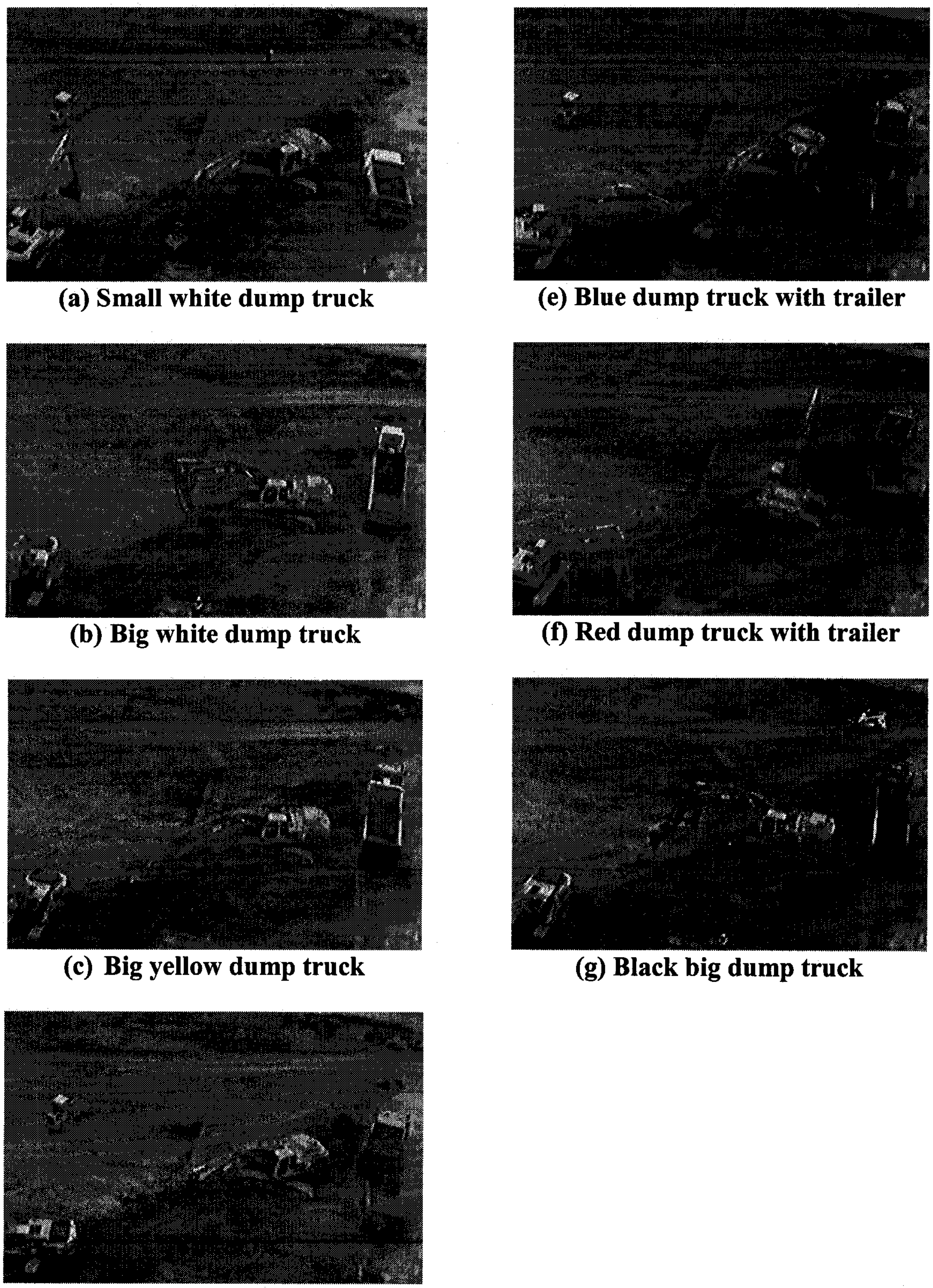


Fig. 14. System main process of truck load cycle time analysis.

3.2.2 Target Extraction

In this image data, based on the colors of cabs of the trucks and the size of their dump boxes, trucks that appeared on the site can be categorized into the following seven different types: small white dump truck, big white dump truck, big yellow dump truck, purple dump truck with trailer, blue dump truck with trailer, red dump truck with trailer and big black dump truck. Those trucks are shown respectively as in Fig. 15(a), (b), (c), (d), (e), (f) and (g).



(a) Small white dump truck

(b) Big white dump truck

(c) Big yellow dump truck

(d) Purple dump truck with trailer

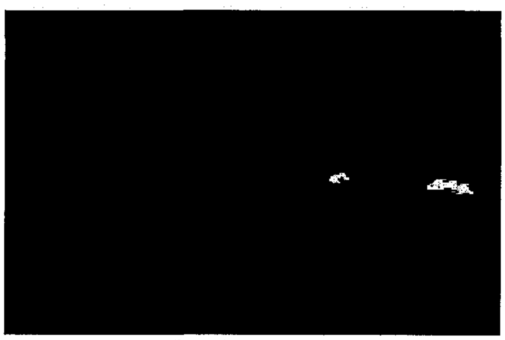
(e) Blue dump truck with trailer

(f) Red dump truck with trailer

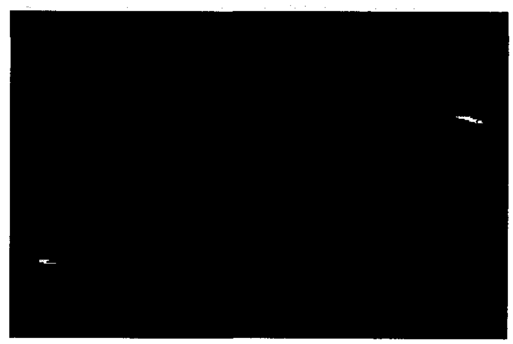
(g) Black big dump truck

Fig. 15. Original image of seven different types of trucks (on the right side).

For the purpose of this research, those types of trucks are further classified into three classes based on the size of their dump boxes including: small trucks, big trucks and small trucks with trailers. From Fig. 15, those seven types of trucks have cabs with six different colors, which are white, yellow, black, blue, red and purple. Thus, color information is used in order to perform the indicator of object extraction. HSV color space is tested again for the purpose of object segmentation. In this application, saturation is no longer the only parameter since saturation and value are not good for distinguishing objects with different colors. Hue does not work when separating white and black objects. Combining hue, saturation and value thresholds and removing noises with fewer or more pixels, objects are extracted except for the big black dump truck, as illustrated in Fig. 16(a), (b), (c), (d), (e) and (f). It is hard to detect the big black dump truck's cab from the noises of black earth in that image.

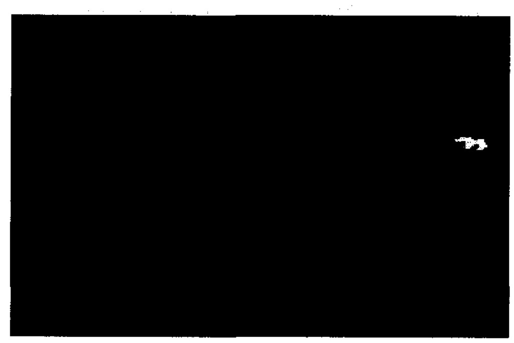


(a) Small white dump truck

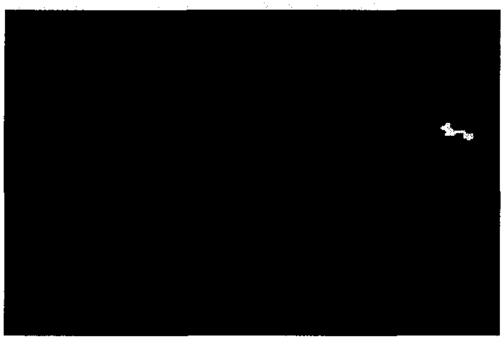


(c) Big yellow dump truck

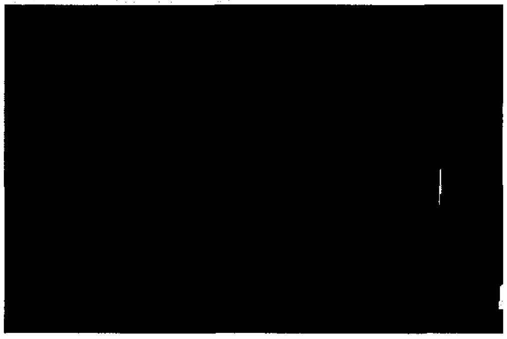
(b) Big white dump truck



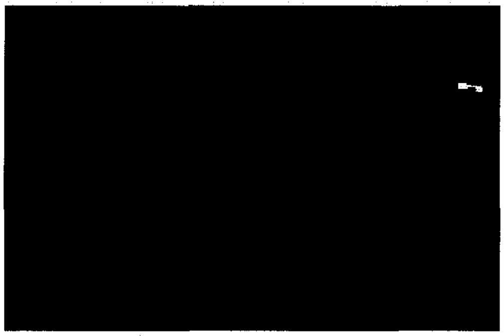
(d) Purple dump truck with trailer



(e) Blue dump truck with trailer



(g) Big black dump truck



(f) Red dump truck with trailer

Fig. 16. Target extraction results of seven different types of trucks.

A shared feature of big dump trucks, which is their white dump box wall, was noticed. To extract dump boxes of big dump trucks means to identify big dump trucks. Then the problem becomes how to distinguish the extracted white box wall from white truck-cab. The extracted results show a thin and long shape of the dump box wall and relatively wider shape of the white cab. A new parameter is introduced into this methodology to separate wide objects and thin objects, known as the major-minor rate, which is the rate of lengths (in pixels) of major axis and minor axis of the objects. Here, "major axis" means the major axis of the ellipse that has the same second-moments as the extracted objects; "minor axis" means the minor axis of the ellipse that has the same second-moments as the extracted objects (Fig. 17). The larger the major-minor rate, the longer and thinner the extracted object. Based on this parameter, the object with a wide shape can be filtered out. The image processing result of the big black dump truck is shown in Fig. 16(g). Big black dump trucks and big white dump trucks share one target extraction function. As a result, there are six target extraction functions to detect the seven types of trucks.

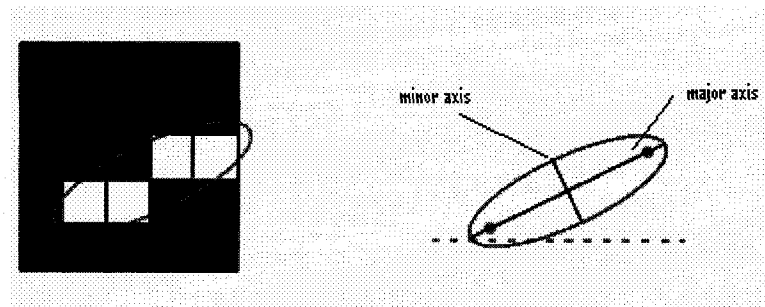


Fig. 17. Major axis and minor axis (The MathWorks 2003).
("All rights reserved by The MathWorks, Inc.")

In this application, the excavator does not change its location and trucks always come into site from the right bottom corner of the image and go out from the right top corner in this two-hour long image data series. This phenomenon accords with excavator's operation customs and, at the same time, it reduces the image processing time and reduces the disturbance caused by the possible noises in other locations of the original images. In this case, only a small portion of the whole image, a 120×160 image, is cropped and processed and the centroids of objects of trucks are expected to appear in a smaller range between two black lines as shown in Fig. 18(b). A sample of processed image range and object range is shown in Fig. 18(a) and (b) respectively. Any objects with centroid points, which are not located in the range between two black lines, are ignored and classified as noises by this program.

(a) Image range **(b) Object range**

Fig. 18. A sample of processed image range and object range.

All threshold values applied in this research application for target extraction are shown in Tab. 1. 36 images from other image series are used to calibrate HSV thresholds and 40 images which are a part of 720 testing images are used to set the thresholds of size, location and shape parameters.

Tab. 1. Feature thresholds for object segmentation.

thresholds	blue truck	purple truck	yellow truck	red truck	white truck	big truck	soil
hue	>0.45 & <0.65	>0.6 & <0.9	>0.1 & <0.17	>0.95 or <0.03	<0.02 or >0.17	<0.02 or >0.17	>0.06 & <0.14
saturation	>0.2	>0.05	>0.2	>0.16 & <0.6			>0.1 & <0.3
value				>0.55	>0.9	>0.8	>0.5 & <0.9
size (pixel number)	>30	>30	>30	>20 & <300	>30	>10 & <280	
location (x)	>70 & <110	>70 & <110	>70 & <110	>70 & <110	>70 & <110	>60 & <100	
major-minor rate			<6	<6	<6	>6.9	

3.2.3 Object Tracking

Object tracking is also necessary in this application. In some occasions, a truck finishes loading and drives away, but it still appears in the processed image portion. In other occasions, two small white trucks appear in an image. In the first example, if the moving of that truck is not detected, the loading time of that truck will be exaggerated. In the second example, the program will be confused by two objects without an object tracking procedure. The same method, as in the equipment idle time analysis, is transferred to this case, that is, two objects with the same color information in two consecutive images with the smallest centroid distance are identified as the same truck. Based on the practice in this application, this method shows its applicability again and does not cause any confusion.

Another problem that happens in object tracking occurs when the hydraulic excavator swings its boom in a certain direction and the bucket covers the cab, as shown in Fig. 19(a), (b) and (c). Comparing Fig. 19(a), (b), (c) with Fig. 20(a), another phenomenon noticed is that no tire rut can be detected when a truck is over it, while a tire rut can be detected by an image processing program, as in Fig. 20(b) if there is no truck. To reduce the chance of missing objects in those situations and to prevent an exaggeration of the truck loading cycle time, it is assumed that if the loading cycle time does not reach the minimum load time and the tire rut is not detected, the truck is still there. Although, the object is not detected in the image processing, the minimum load time refers to the time period in which loading could not be finished for each class of truck. The specific minimum load time of small trucks, big trucks and trucks with trailers are determined as 40 seconds, 90 seconds and 100 seconds, respectively.

3.2.4 Movement Indicator

As in the first application, determining whether or not a truck actually moves is another important step in measuring equipment idle time. Centroid distance is applied for this purpose. But different movement indicator threshold values are applied in this research because dump trucks with trailers move a short distance when loading two different dump boxes, as shown in Fig. 21(a) and (b), while small dump trucks and big dump trucks do not move when loading. After trials of 30 images which are included in the 40 images for calibrating size, location and shape parameters, for a more accurate analysis result, a distance of 50-pixel is determined to be the movement indicators of trucks with a trailer, 40-pixel is determined to be the movement indicators of big white and black dump trucks, and 20-pixel is for small dump trucks and big yellow dump trucks. If the centroid distance between objects in two sequential images is greater than its movement indicator, or no same objects as in immediate preceding image are detected, the load cycle of the truck in the preceding image is finished.

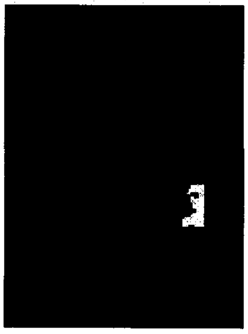
(a) The bucket covers small white dump truck's cab

(b) The bucket covers blue dump truck with trailer's cab

(c) The bucket covers big yellow dump truck's cab

Fig. 19. Samples of images of the bucket covering trucks' cabs.

(a) Tire rut



(b) extracted Tire rut

Fig. 20. A sample of tire rut and the extraction result of tire rut.

(a) When its dump box is loaded

(b) When its trailer is loaded

Fig. 21. A sample of dump trucks with trailers moving during loading.

3.2.5 Detection priority

Since those images were taken from a long distance, the targets generally have a small number of pixels. Thus, it is not possible to remove all noises left in the extracted results by setting a greater area threshold. In this algorithm, six target extraction functions are designed to detect different types of trucks. Once a function detects qualified objects in an image, the program will think the type of truck corresponding to that function is loading in that image. Occasionally, one image gets several positive results from those six functions, which will cause some confusion to this program. To solve this problem, the function corresponding to the type of truck in the immediate preceding image is run first. If the same type of truck is detected again, and the change of the location of that truck does not exceed its movement indicator, the program will identify the same truck as the one in the previous image. If no such object or some movement of that object is detected, the program will rerun all six functions separately until one positive result is achieved. Thus, the sequence of those six functions becomes important for the final result. In this application, the functions to detect objects with colors, which are easy to be identified, run first while the functions that tend to include noises run later. The function that detects big white dump trucks has to run before the function that detects small white dump trucks. Otherwise, big white dump trucks will be treated as small white dump trucks because of their white cabs. The final sequence of those six functions is determined, from the first to the last, as follows: the function for the blue trucks, the function for the yellow trucks, the function for the big white and the black trucks, the function for the small white trucks, the function for the purple trucks, and the function for the red trucks.

3.2.6 System Architecture

Another image processing system was developed using MATLAB to automatically estimate the truck loading cycle time for various types of trucks. A total of 76 sample images were studied to calibrate the parameters required for the system. The processing steps are as follows (Fig.22):

1. The truck of interest on the first 120×160 image, which is cropped from the original image, is manually selected on a computer; the type and the centroid of the object are then identified.
2. The next image (the image captured 10 seconds later) is read for analysis.
3. Corresponding Target extraction function of the preceding image is run to identify whether the same type of objects is in this image, or if there is one of the six types of trucks detected in the preceding image. If yes, obtain the centroid information; otherwise, if there is no object detected in the preceding image, all six functions are run until one function can identify objects and then record the type and centroid information of that object.
4. The cenroid distances are calculated and the object with the shortest centroid distance is identified as the truck.
5. If the type of truck is the same as the type in the preceding image, and the centroid distance less than the criteria, add 10 seconds to the loading time of that truck; otherwise, another truck begins to be loaded.
6. If no objects are identified in step 3, and if the load cycle of the truck in the preceding image is less than the minimum criterion of its class and no soil can be detected, add 10 seconds to the loading time of that truck; otherwise, the load cycle

of that truck is complete.

7. Step 2 to step 6 are repeated until the last image in the analysis period is processed.
8. Remove the load cycle time data, which is less than the minimum criterion of the corresponding class of truck. Statistics information is calculated for small trucks, big trucks and small trucks with a trailer.

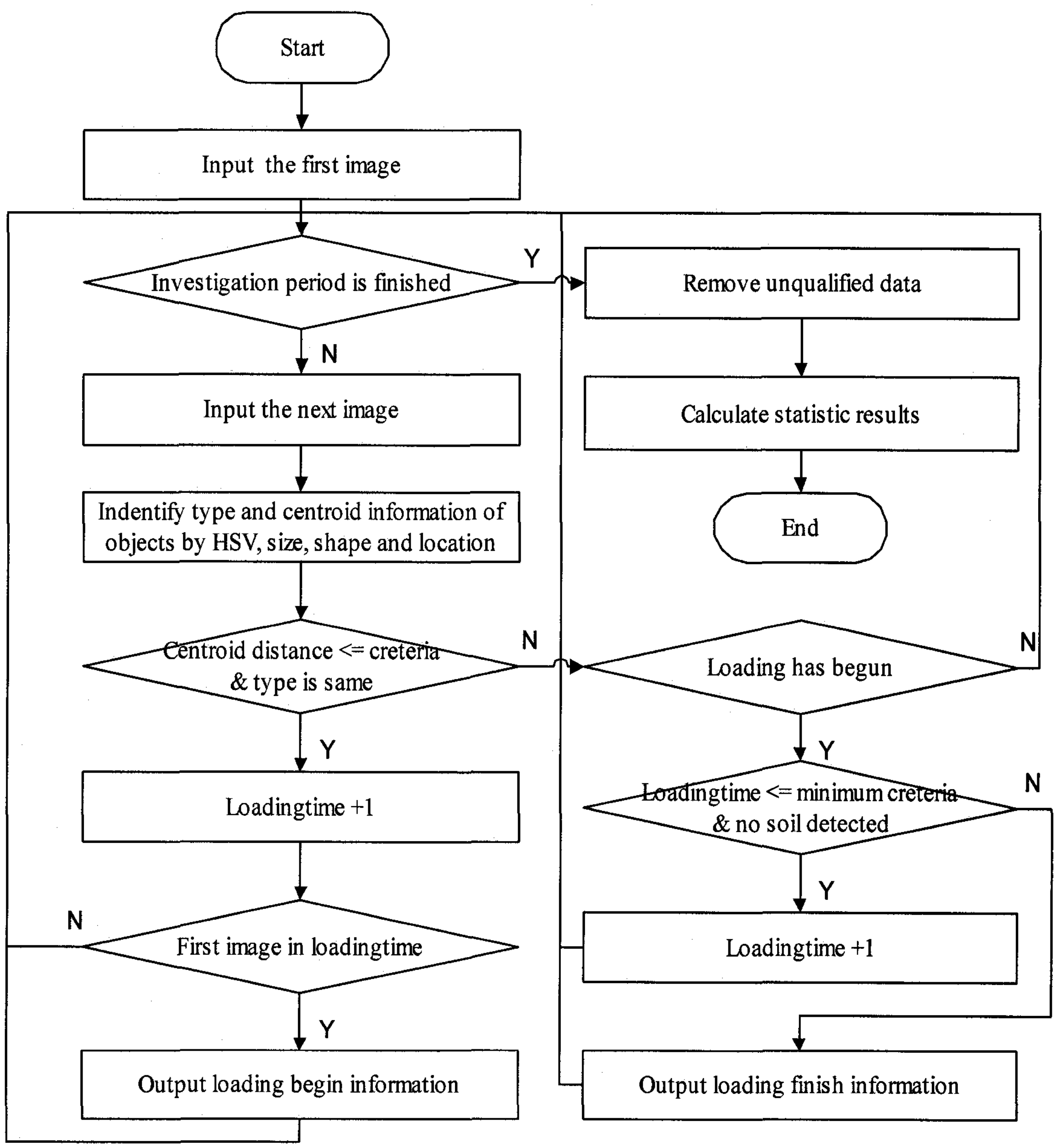


Fig. 22. System architecture of truck load cycle time analysis.

3.3 Application Results Analysis

In this application, a two-hour long image series is used to test the proposed algorithm. This image data contains 720 images and, during a two-hour working cycle, 37 trucks finished loading. This is based on visual observation and includes 18 small white dump trucks, 2 big yellow dump trucks, 8 big white/black dump trucks, 5 purple dump trucks with a trailer, 2 blue dump trucks with a trailer and 2 red dump trucks with a trailer. There are 36 whole load cycles during that two-hour duration. It is best not to record the whole load cycle of the first small white dump truck, which ended its loading at 07:00:22.

For the same reason as in the first application, this methodology underestimates the truck load cycle time due to the intervals of 10 seconds between two successive images. Compared with the average load cycle of a small dump truck, which is about 50 seconds, the error caused by that factor is greater than in the first application. However, the results are still acceptable and the effects can be diminished by adding two random numbers from a uniform distribution between 0 and 10 seconds.

The image processing result shows that there are 34 trucks detected to finish loading and 29 qualified load cycle times are included in the final statistical results by removing those load cycles, which do not meet the minimum load cycles of their classes. The detection rate is nearly 92% and, compared with 36 whole load cycles in the image data, the effective statistical data rate is about 81%. The details of detected load cycle times are shown in Tab. 2, 3, 4, 5, 6 and 7.

Tab. 2. Visual observation and image processing results of small white trucks.

small white truck	visual observation result			image processing result			difference (seconds)
	start time	finish time	load cycle time	start time	finish time	load cycle time	
1	07:00:02	07:00:22	00:00:20	07:00:02	07:00:22	00:00:20	
2	07:03:32	07:04:42	00:01:10	07:03:32	07:04:42	00:01:10	0
3	07:08:02	07:09:12	00:01:10	07:08:22	07:09:12	00:00:50	-20
4	07:09:32	07:10:42	00:01:10	07:09:32	07:10:42	00:01:10	0
5	07:11:02	07:12:02	00:01:00	07:11:02	07:12:02	00:01:00	0
6	07:29:32	07:30:22	00:00:50	07:29:32	07:30:22	00:00:50	0
7	07:33:32	07:34:12	00:00:40	07:33:42	07:34:12	unqualified	
8	07:36:32	07:37:22	00:00:50	missed			
9	07:42:12	07:43:02	00:00:50	07:42:12	07:43:02	00:00:50	0
10	07:47:32	07:48:22	00:00:50	07:47:22	07:48:22	00:01:00	10
11	08:00:52	08:01:42	00:00:50	08:01:02	08:01:42	00:00:40	-10
12	08:08:22	08:09:12	00:00:50	08:08:32	08:09:12	00:00:40	-10
13	08:09:32	08:10:22	00:00:50	08:09:32	08:10:22	00:00:50	0
14	08:20:02	08:20:52	00:00:50	08:20:02	08:21:02	00:01:00	10
15	08:47:22	08:48:22	00:01:00	08:47:32	08:48:32	00:01:00	0
16	08:48:42	08:49:32	00:00:50	missed			
17	08:54:42	08:55:52	00:01:10	missed			
18	08:56:02	08:56:52	00:00:50	08:56:22	08:56:42	unqualified	

Tab. 3. Visual observation and image processing results of big yellow trucks.

big yellow truck	visual observation result			image processing result			difference (seconds)
	start time	finish time	load cycle time	start time	finish time	load cycle time	
1	07:12:42	07:14:42	00:02:00	07:12:42	07:14:42	00:02:00	0
2	08:10:42	08:12:32	00:01:50	08:10:42	08:12:32	00:01:50	0

Tab. 4. Visual observation and image processing results of big white and black trucks.

big white and black truck	visual observation result			image processing result			difference (seconds)
	start time	finish time	load cycle time	start time	finish time	load cycle time	
1	07:15:02	07:17:12	00:02:10	07:15:02	07:17:12	00:02:10	0
2	07:17:32	07:19:42	00:02:10	07:17:32	07:19:42	00:02:10	0
3	07:22:42	07:24:42	00:02:00	07:22:42	07:24:42	00:02:00	0
4	07:58:52	08:00:42	00:01:50	07:58:52	08:00:52	00:02:00	10
5	08:12:52	08:14:42	00:01:50	08:12:52	08:14:42	00:01:50	0
6	08:15:12	08:17:12	00:02:00	08:15:12	08:17:12	00:02:00	0
7	08:17:32	08:19:42	00:02:10	08:17:32	08:19:42	00:02:10	0
8	08:50:02	08:51:52	00:01:50	08:50:02	08:51:52	00:01:50	0

Tab. 5. Visual observation and image processing results of purple truck with trailers.

purple truck with trailer	visual observation result			image processing result			difference (seconds)
	start time	finish time	load cycle time	start time	finish time	load cycle time	
1	07:00:42	07:03:12	00:02:30	07:00:42	07:03:12	00:02:30	0
2	07:20:12	07:22:12	00:02:00	07:21:02	07:22:12	unqualified	
3	08:02:02	08:04:02	00:02:00	08:02:02	08:03:52	00:01:50	-10
4	08:21:22	08:23:52	00:02:30	08:21:32	08:22:22	unqualified	
5	08:52:12	08:54:22	00:02:10	08:52:12	08:54:22	00:02:10	0

Tab. 6. Visual observation and image processing results of blue trucks with trailers.

blue truck with trailer	visual observation result			image processing result			difference (seconds)
	start time	finish time	load cycle time	start time	finish time	load cycle time	
1	07:05:02	07:07:52	00:02:50	07:05:02	07:07:52	00:02:50	0
2	08:05:32	08:08:02	00:02:30	08:05:22	08:08:02	00:02:40	10

Tab. 7. Visual observation and image processing results of red trucks with trailers.

red truck with trailer	visual observation result			image processing result			difference (seconds)
	start time	finish time	load cycle time	start time	finish time	load cycle time	
1	07:26:12	07:28:52	00:02:40	07:26:22	07:28:52	00:02:30	-10
2	08:24:22	08:46:52	00:22:30	08:24:02	08:46:52	00:22:50	20

The average absolute value of the differences between 29 qualified image processing results and their corresponding visual observation results is only 3.8 seconds.

The average load cycle times of small dump trucks, big dump trucks and trucks with a trailer are shown as in Tab. 8. The differences in small dump trucks and big dump trucks are ignorable. The reason for the big difference in dump trucks with a trailer is that in the image processing results, seven truck load cycles, including one of 22 minutes and 30 seconds, are identified as qualified data instead of nine truck load cycles by visual observation. This was caused by two load cycles of purple trucks not meeting the minimum load cycle time for trucks with trailers (100 seconds). The average cycle time by image processing is much greater than the actual average. With more image data being analyzed by this methodology, it is believed that the difference will be diminished. Based on the detection rate of approximately 92%, and the average errors about 3.8 seconds, the application of image processing in truck load cycle time analysis is a successful attempt. After the reliability and flexibility of this system is improved through future works, it will become an efficient real-time automatic construction equipment management tool.

Tab. 8. Average load cycle time of visual observation and image processing results.

	visual observation result (s)	image processing result (s)
small dump truck	53	55
big dump truck	119	121
dump truck with trailer	278	320

Chapter 4 Feasibility Study of City Image Database System

4.1 Literature Review and Current Practice

The proposed city image database/inventory is an integrated application of database, image processing and 3-D CAD. Traditional databases have been developed over the last 40 years, but image database, as a new member of databases, is gaining popularity among the research community due to the recent increasing requirements of multimedia applications. Currently, one of the challenges faced is how to transfer functions of traditional database to the database for nonstructural data, such as image, audio and video. Image database is an interdisciplinary endeavor in computer vision, image processing, image retrieval, and database. Today, with the requirements of multimedia information, image database becomes an active research topic. The latest development of image database will be discussed in the following paragraphs.

It is easy to perform content retrieval by searching for specific words in data items in traditional databases since all data are in text form. Images in image databases can be stored with some text indexes, such as time and location, as the data of classical databases. If every image in the image database has its content indexes, it is easy to execute a typical query on the content of image data. Unfortunately, images are not a kind of structural information and it is infeasible to manually attach indexes of contents to images, especially when there is a large quantity of image data or too many contents are included in each image. To automatically capture contents in images is difficult except in some simple scientific application, such as, to extract the forests in satellite images or erythrocytes in microscope images of blood sample.

"Techniques that seek to index unstructured visual information are grouped under the collective name of content-based retrieval" ("Image Databases", 2002). In content-based retrieval, the system searches images to find images containing objects with some visual features identified by the query. For example, "pixel level", "feature level" and "semantic level" (Image Databases, 2002) are different levels of abstraction in image content. "Pixel level" retrieval, which is rarely used, can help a user find point information. "Feature level" retrieval is widely used to search objects with required features such as color, texture and shape. "Semantic level" retrieval is the most ideal query for users. For example, a user can search mobiles, buildings and traffic signs by name. However, this retrieval is difficult to achieve because some limited conditions, objects in images can be identified as specific types of objects by its feature level contents. Its success, to some extent, depends on the accuracy of the image processing result.

For feature level retrieval, "multidimensional indexing" is necessary for "each feature space (e.g., a color histogram space) can be viewed as a multidimensional space, in which a feature vector representing an object corresponds to a point" ("Image Databases", 2002). Since traditional database management systems (DBMS) do not support multidimensional indexing or do not support it well, some researches are now carried on to include multidimensional indexing search engine into DBMS.

4.2 Proposed Methodology and Implementation

To build a city image database system poses a challenging task. A reported research for the similar purpose is using a satellite image to discover the change of a city in 2-D space. In this study, the proposed city image database system is built by images shot from the streets to record the change of a city in 3-D space, such as a new building, loss of traffic signs, growing of trees and even the city's pattern change. The images used to test the proposed methodology are captured by a digital camera (Canon Powershot S20) with a focal length of 6.5-13mm. The system design and some detail methodologies will be discussed in the following paragraphs.

4.2.1 System Main Process and System Main Components

1) System Main Process

In the proposed system, the input parameters include image data with times, locations and orientations of shots, and target's features information. Image data in this proposed system refers to images taken by a digital camera that is fixed onto a vehicle which drives along streets. Every image has a record of shooting time, and the location and orientation of that shot. A target's feature information includes HSV ranges, shape, size and location information. For those targets with fixed colors and shapes as traffic signs, their HSV and shape thresholds can be preset in the target information system. For those targets with unfixed colors and shapes as buildings, their HSV information can be cropped from any image that contains those targets and it will be stored in the target information database for identifying the same target in future image data. Tools and techniques in this system are image database management techniques, image processing tools, target 3-D information extraction tools and 3-D model building tools. Image database management techniques, which are not the focus of this study, may include image storage management techniques, image retrieval techniques and file system architecture. Image processing tools includes object segmentation tools based on HSV, shape and size information. Target 3-D information extraction tools are a set of algorithms that calculate targets' location and elevation information by combining information of two or more images. With the help of information attained from previously mentioned tools, 3-D model building tools are tools that assist in the building of a 3-D model. Criteria include thresholds which help to identify targets, such as thresholds of hue, saturation, value, shape parameters, size and location. The outputs are updated image database, identified changes in the city and 3-D model of targets and city. The relationship of this proposed system is shown in Fig. 23.

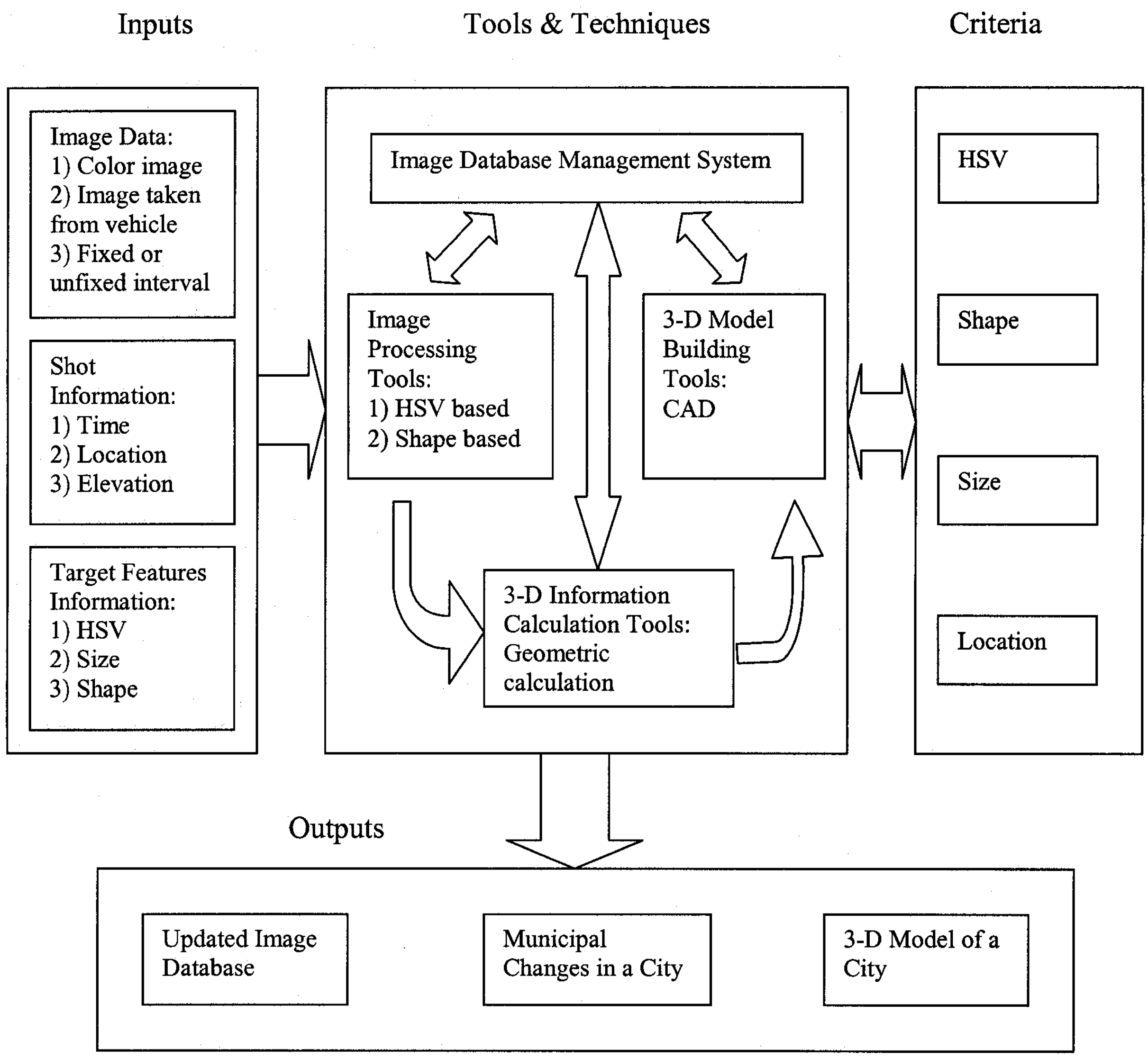


Fig. 23. System main process of city image database.

2) **System Main Components**

The proposed system contains several modules for performing a variety of functions, which are a user multimedia interface module, an image data collecting module, a database module, an image processing module, a 3-D information extraction module and a 3-D model building module. A user multimedia interface module creates a friendly input/output interface for users to obtain required information. An image data collecting module is an equipment system that takes qualified images of a city from the streets and records the shot information. A database module performs the functions related to image storage function, image retrieval function, image management function and targets' features management function. An image processing module has two sub-modules, which are an image processing module for objects with fixed colors and shapes, such as traffic signs, and the module for objects with unfix colors and shapes. The function of this module is to detect images with the features of required targets and to extract those targets for further processing. A 3-D information extraction module performs a function to calculate the 3-D information by the object segmentation result. The methodology will be discussed in details later. A 3-D model building module is responsible for the actual building of a 3-D model through 3-D information acquired from the above module. Those modules are shown as in Fig. 24.

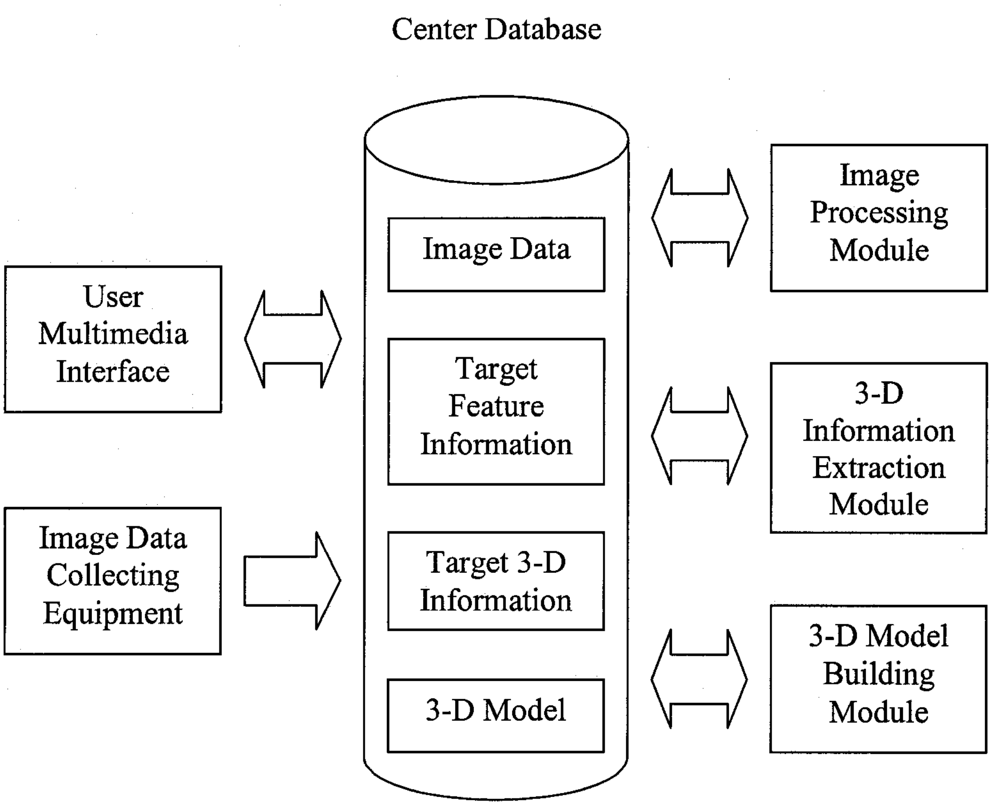


Fig. 24. System main components of city image database.

4.2.2 Data Collecting Equipment Design

In this proposed research, image data collecting equipment plays an important role in the success of the whole system. Before describing the details of image data collecting equipment, the structure of digital cameras should be reviewed. The lens of a digital camera, which is generally composed by several pieces of glasses, is simplified as a convex lens. The CCD (Charge Coupled Device) sensor is responsible for detecting the light and output digital signal. Generally, the CCD sensor of a digital camera is an array of millions of square semiconductor cells. How a camera functions is shown as Fig. 25.

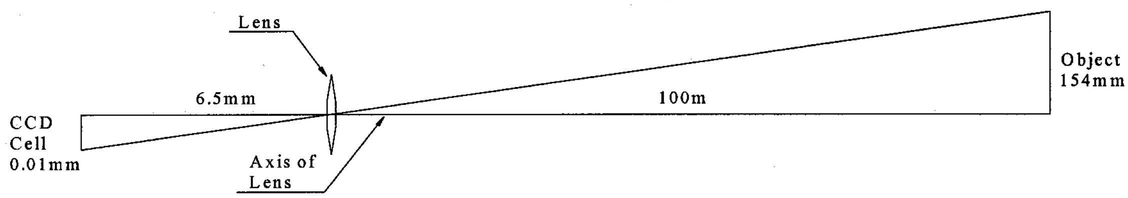


Fig. 25. Sketch of how digital camera works.

To collect image data effectively and efficiently for this research, the recommended camera is an industrial digital camera with a fixed focal length. It is assumed that the camera has a lens with a fixed focal length of 6.5mm and the size of pixel is 10μm. Based on Fig. 25, it is easy to estimate that 1 pixel covers 154mm in a distance of 100 meters. Generally, image processing would not miss an object with a dimension of 1.5 pixels, which means the proposed camera can detect an object 100 meters away with a dimension bigger than 231mm. The size of the image is recommended to be 2 Mpixel, which can provide a good resolution and a good image processing speed based on processing capability of the current PC.

Assuming the speed of a vehicle is 30km/hour and the frame rate is one frame per second, the camera takes one picture for every 8.3 meters of vehicle travel. With this frame rate, assuming the camera's horizontal angle of view is 60°, targets with a distance to the vehicle's route which is greater than 7.2 meters can be shot at least twice for a driving if the orientation of the shot is perpendicular to the vehicle's route. Generally, a 2 Mpixel image has a size of 0.8 MB. The data rate is about 0.8MB/second, which can be transferred by a FireWire port without an image grabber. A 200 GB hard drive can save nearly 70-hour long image data.

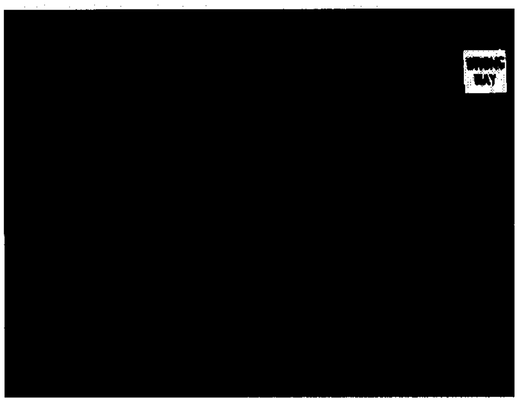
For calculating the 3-D information of targets, the locations of the shots are necessary. With the help of a GPS device on the vehicle, which are common devices found on many newer model vehicles, the coordinates and the elevations of the shots could be recorded and saved with corresponding images. Since the digital camera is fixed on the vehicle and the vehicle do not always keep driving in a straight line, the orientations of shots may vary; thus the orientations of shots have to be recorded. This orientation can be measured by installing a gyroscope in the vehicle. A factor that affects the elevation calculation of the targets is that the pitching of the vehicle could not keep the axis of the camera horizontal. Of course, it is also feasible to record the pitching angle and to consider it in algorithm if the pitching angle is small. But when the vehicle drives in an adverse road condition, the camera could miss the top of a building for pitching. As a result, the best choice is to fix the camera on a platform that can always keep the axis of lens horizontal.

4.2.3 Image Processing Tools

1) **Image Processing for targets with fixed colors and shapes**

Due to the limited and fixed colors of those objects, preliminary object segmentation is achieved by presetting HSV thresholds. A sample of a traffic sign is tested to investigate the feasibility of the proposed methodology as in Fig. 26(a) and (b).

(a) Original image



(b) Preliminary object segmentation by HSV with small noises removed

Fig. 26. A sample of traffic sign.

After preliminary object segmentation, shape features are introduced to separate traffic signs with the same color. For example, a parameter, which is the distance between top-left point and left-top point (Fig. 27) of the extracted objects, can be used for distinguishing rectangle signs from octagon signs. To avoid the effect of size on the threshold setting, that parameter is modified by dividing the square root of the object's area. This methodology is proved in Fig. 28(a), (b) and (c), which are the original images. Preliminary object segmentation results by color information and final extracted results of "stop" signs by a function designed for "wrong way" signs. With the help of color, shape and size features, it is feasible to distinguish different geometric shapes of signs and even the words on the signs by the virtue of the limited types of traffic signs. Semantic level retrieval is then made possible as a result of the link between a type of traffic sign and a combination of features such as color, shape and size. In that case, all images with stop signs can be searched, if a user executes a query function by typing "stop sign". Of course, to design the detail specific algorithms for every traffic sign is a content of further researches.

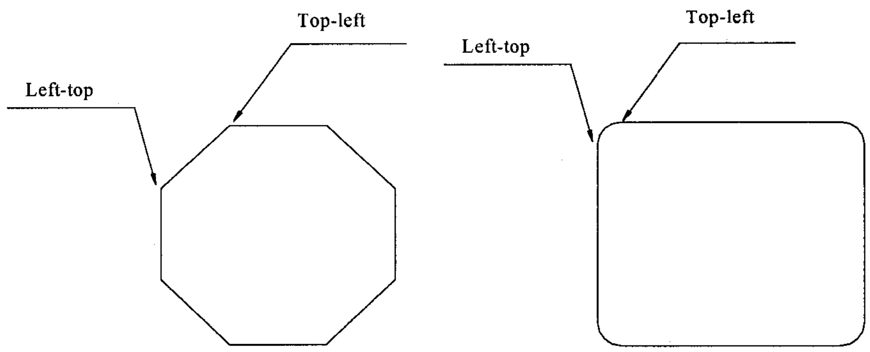
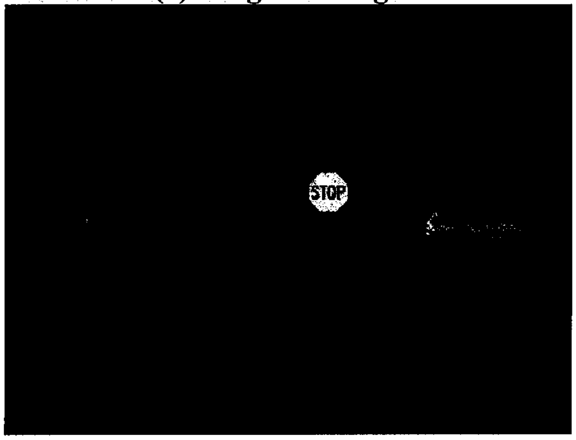


Fig. 27. A sample of distinguishing octagon from rectangle with filleted corners.

(a) Original image

(c) Final object segmentation result



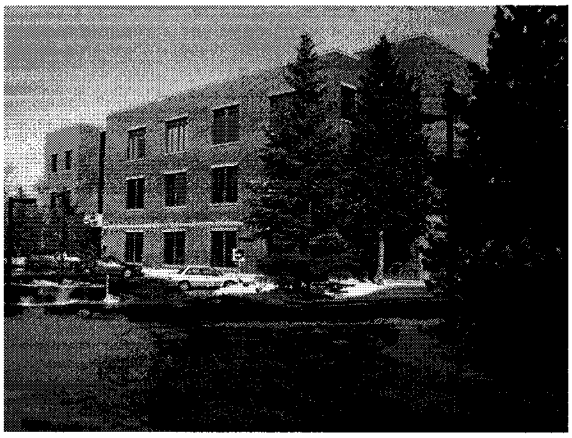
(b) Preliminary object segmentation result

Fig. 28. A sample of eliminating octagon signs based on shape feature.

2) **Image processing for targets without fixed colors and shapes**

Targets, such as buildings in a city, have a variety of colors and shapes. As a result, every object should be cropped manually on an original image. Then, HSV information of that object is saved in the system for future automatic target segmentation. For accommodating the variations of HSV information caused by different weather and light conditions when images are taken, in this proposed algorithm, the hue range is extended by $\pm 10\%$ based on the hue range of the cropped pixels, saturation range is extended by $\pm 20\%$ and value range is extended by $\pm 30\%$ since hue is much more stable than saturation and value in different conditions. An experiment shows the adaptability of this target segmentation algorithm. Fig. 29(a) and (b) are the sample images of the same building taken in different time and orientation. Fig. 29(c) is the extracted result of Fig. 29(a) and Fig. 29(d) is the extracted result of Fig. 29(b) based on the HSV information cropped in Fig. 29(a).

(a) A sample image of a building



(b) A sample image of the same building taken in different time

(c) The extracted result of (a)

(d) The extracted result of (b)

Fig. 29. A sample of the original images and the extracted results of a building.

Another problem faced is how to detect the edges of a building in the condition that some edges of facade are sunk in the extracted target. According to observation, it is easy to detect the silhouette of a building which has a cubic shape as in Fig. 29 By the reason that the elevation of the camera on a vehicle at the street is generally lower than the top of a building, the top-left point or top-right point of an extracted object are the top corners of a building's façade. Those top points are generally easy to detect since they are less likely to be covered by trees, in comparison with the lower points at the edge of a building. For a cubic shape building, once the top corners are located, the edge of the building is determined.

4.2.4 3-D Information Extraction Tools

In the proposed study, a 3-D information retrieval algorithm is introduced into the system and plays an important role in content retrieval, meaningful result output and 3-D model. Image processing results can be helpful in detecting targets. If it is not possible to identify which target is detected, the system could not give a user a definite explanation. For example, if a user wants the information of a building, there may be several buildings with similar color that are extracted after image processing. As a result, the system could not tell which particular building is the one the user wants. Furthermore, if the location information can be mined based on image processing results, it is easy and economical to build a 3-D model of a city by the proposed system.

An image survey algorithm is presented to detect the 3-D information of a target as long as the target is captured by digital camera in two different locations. Based on how the camera functions, a simplified sketch of the proposed algorithm is shown in Fig. 29. In the equipment design, how to obtain coordinates and elevations of shots is discussed. Assuming that the coordinates of shot 1 and shot 2, orientations of shot 1 and shot 2, lens of camera and the dimension of CCD cell are known, the coordinates of point 3 (X_3,Y_3) are easy to achieve by geometric calculation.

The procedures are as follows:

1. The distance from point A_1 to point B_1 equals the horizontal CCD cell distance from P_3 in image to the center line on image multiplied by the real cell size.
2. Angle b_1 is calculated by the distance from A_1 to B_1 and B_1 to S_1 (lens) as Eq. (1). In the same way, angle b_2 is calculated as Eq. (2).
3. Angles c_1 and c_2 are equal to a_1 minus b_1 and a_2 minus b_2 respectively as Eq. (3) and (4).
4. Angle c_3 is equal to 180° minus c_1 and c_2 as Eq. (5).
5. The horizontal distance from shot 1 to shot 2 in XY plane is calculated by their coordinates as Eq. (6).
6. With the angles of c_1 and c_2, the horizontal distance from P_3 to S_1 is calculated as Eq. (7).

7. Coordinates of P_3 are calculated as Eq. (8) and (9).

The calculation is shown as follows ($\overline{AB}$ means the distance from point A to point B):

$$b_1 = \tan^{-1}\left(\overline{A_1B_1} / \overline{B_1S_1}\right) \quad (1)$$
$$b_2 = \tan^{-1}\left(\overline{A_2B_2} / \overline{B_2S_2}\right) \quad (2)$$
$$c_1 = a_1 - b_1 \quad (3)$$
$$c_2 = a_2 - b_2 \quad (4)$$
$$c_3 = 180° - c_1 - c_2 \quad (5)$$
$$\overline{S_1S_2} = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2} \quad (6)$$
$$\overline{S_1P_3} = \frac{\overline{S_1S_2} \times \sin c_2}{\sin c_3} \quad (7)$$
$$X_3 = X_1 + \overline{S_1P_3} \times \cos c_1 \quad (8)$$
$$Y_3 = Y_1 + \overline{S_1P_3} \times \sin c_1 \quad (9)$$

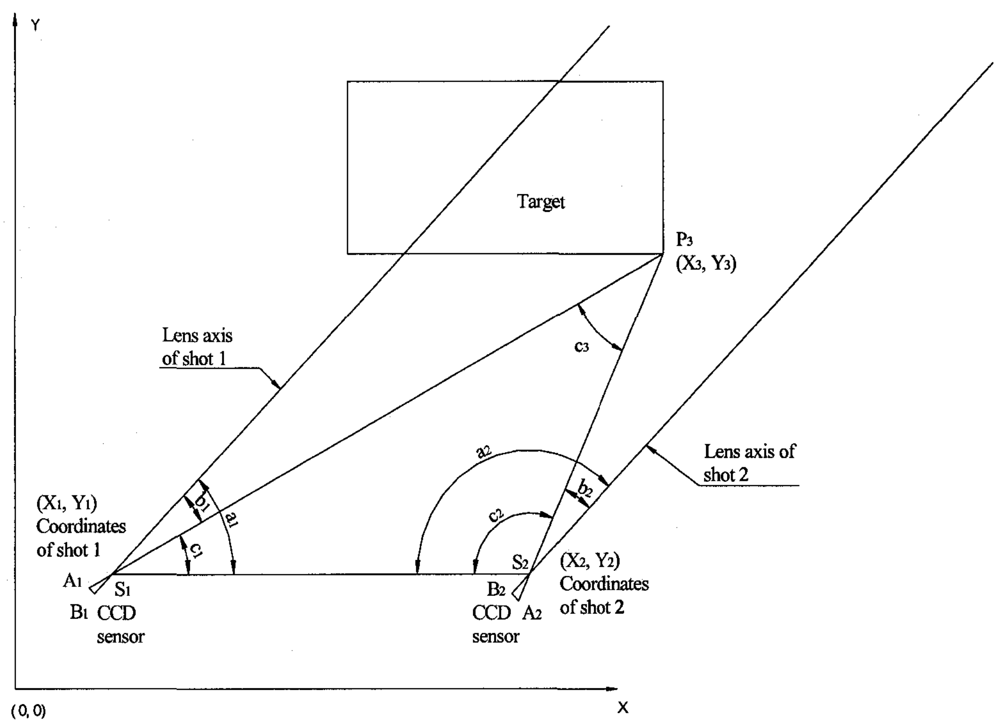


Fig. 30. Sketch of proposed algorithm for calculating target's coordinates in XY plane.

In the above algorithm, there is an assumption that the distance from the CCD sensor to the center of the lens equals the focal length. Generally, a digital camera has a function called automatic focus, which adjusts the location of CCD sensors and produces a clear image. If the object is close to the camera, the CCD sensor is located in the place furthest away from the lens than the focal length. But, if the object is far away from the camera (about 10 meters), the CCD sensor is located very close to the focus in order to capture an image of objects with sharp edges. The sketch of the calculation is in Fig. 31. It is assumed that the height of the object is 10 meters, the distance from the object to lens is 10 meters and the focal length of the lens is 6.5mm. The ideal distance from CCD sensor to lens is 6.504mm, which is almost the same as the focal length. The details of the calculations are as follows ($\overline{AB}$ means the distance from point A to point B):

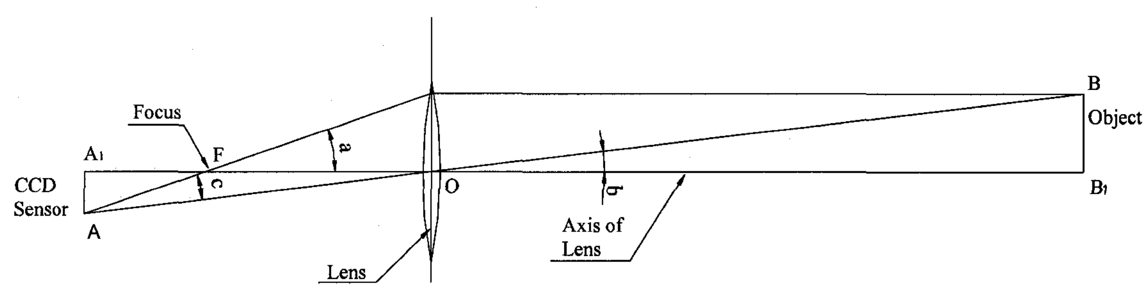


Fig. 31. Sketch of calculation of location of CCD sensor.

$$a = \tan^{-1}\left(\frac{\overline{BB_1}}{\overline{OF}}\right) = \tan^{-1}\left(\frac{10\times10^3}{6.5}\right) = 89.963°$$

$$b = \tan^{-1}\left(\frac{\overline{BB_1}}{\overline{OB_1}}\right)\tan^{-1}\left(\frac{10\times10^3}{10\times10^3}\right) = 45°$$

$$c = 89.963° - 45° = 44.963°$$

$$\overline{OA} = \frac{6.5\times\sin(180° - a)}{\sin c} = 9.198mm$$

$$\overline{OA_1} = \overline{OA}\times\cos b = 9.198\times\cos 45° = 6.504mm$$

It is unlikely that a camera can adjust the CCD sensor to reach such a precision of 0.004mm. For the purpose of this system, it is suggested to fix the CCD sensor at the location of focus to avoid the automatic focus function, which affects the image quality when it adjusts the location of the CCD sensor. In Eq. (1), the $\overline{B_1S_1}$ is equal to the focal length of that digital camera.

Based on Eq. (1) to (9), the coordinates of P_3 are calculated. The height of a target (as in Fig. 31) can be estimated as follows if the axes (the line passes through A_1 and A_2) of shots are kept horizontal ($\overline{AB}$ means the distance from point A to point B):

$$\overline{A_2B_2} = \frac{\overline{A_1B_1}}{\overline{A_1O}}\times\overline{A_2O} \tag{10}$$

In the Eq. (10), $\overline{A_2O}$ is the horizontal distance from object to lens, which can be obtained from calculation of coordinates in XY plane. The elevation of point B_2 is equal to the elevation of point O plus $\overline{A_2B_2}$.

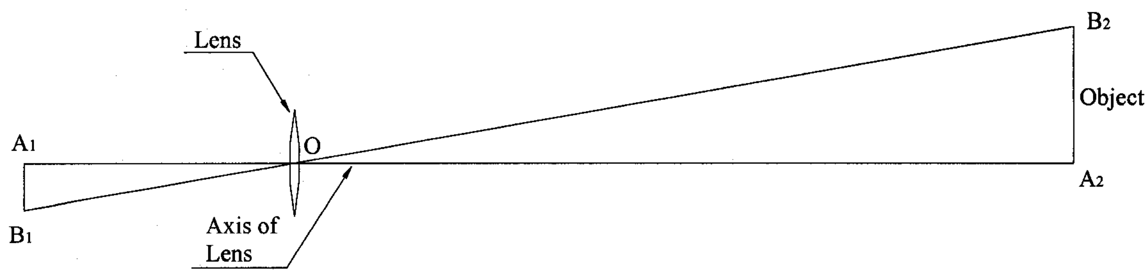


Fig. 32. Sketch of proposed algorithm for discovering target's elevation.

An experiment is designed to test the feasibility of the proposed algorithm. A piece of paper with a marked green line and a piece of paper with a marked purple line are posted on the wall. The distance between the left edges of the two marked lines is 374 mm. The distance between two shots is 500 mm. The axes of the two shots are perpendicular to the wall and the distances between lenses and wall are both 1000 mm. Fig. 33 illustrates the layout of the experiment with the coordinates of objects and shots. All images are taken with a focal length of 6.5mm. The actual distance from lens to CCD sensor is about 6.54mm. In the calculation, 6.5mm is used as an approximation of 6.54mm. The images taken in shot 1 and shot 2 are shown in Fig. 34 (a) and (b). The extracted targets are shown in Fig. 34 (c) and (d).

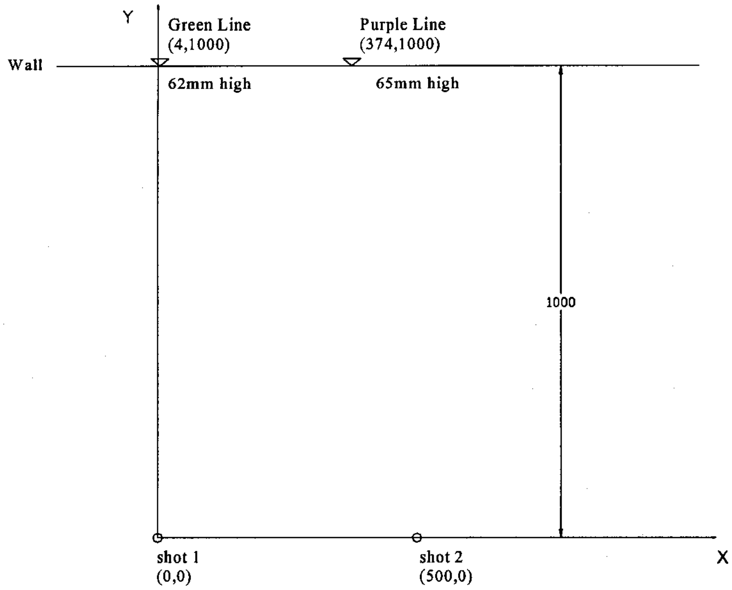
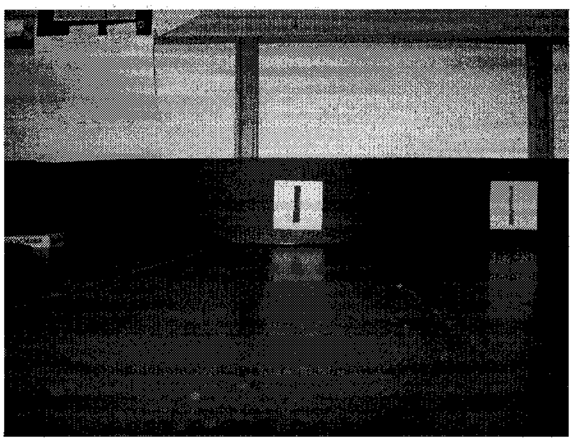
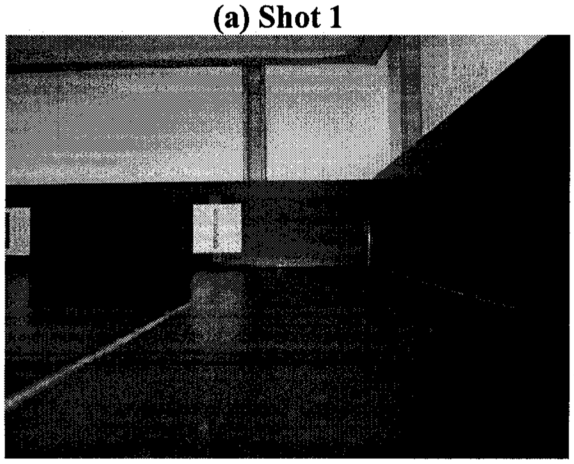


Fig. 33. Layout sketch of experiment.

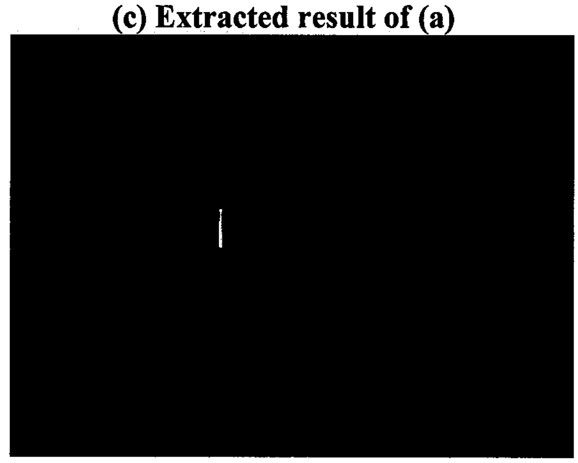


(a) Shot 1

(c) Extracted result of (a)



(b) Shot 2



(d) Extracted result of (b)

Fig. 34. Experimental original images and extracted results.

The pixel size on the CCD array is unknown, but the distance between the green marked line and the lens of camera is about 1000mm. The measured length of the green line is 62mm and the pixel distance in the image processing result is 41.05. The real pixel size on the CCD array is estimated as follows:

By Eq. (10), the pixel size of CCD array $= \dfrac{6.5 \times 62}{1000 \times 41.05} = 9.82 \times 10^{-3}\, mm$

In Fig. 34 (c), the pixel coordinates of the top-left point of the purple line is (572.5, 239.5), while the pixel coordinates of the same point is (237.5, 241.5) in Fig. 34 (d). In most cases, the width of the CCD cell is the same as the height of the CCD cell, although there are only a few CCD sensors with different widths and heights. It is assumed that the widths of CCD cells of this camera are equal to the lengths of CCD cells. Then, the coordinates of left edge of purple line are calculated as follows:

$$b_1 = \tan^{-1}\left(\overline{A_1B_1} / \overline{B_1S_1}\right) = \tan^{-1}\left(\frac{(572.5 - 320) \times 9.82 \times 10^{-3}}{6.5}\right) = 20.9°$$

$$b_2 = \tan^{-1}\left(\overline{A_2B_2} / \overline{B_2S_2}\right) = \tan^{-1}\left(\frac{(320 - 237.5) \times 9.82 \times 10^{-3}}{6.5}\right) = 7.1°$$

$$c_1 = a_1 - b_1 = 90° - 20.9° = 69.1°$$

$$c_2 = a_2 - b_2 = 90° - 7.1° = 82.9°$$

$$c_3 = 180° - c_1 - c_2 = 28°$$

$$\overline{S_1S_2} = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2} = 500mm$$

$$\overline{S_1P_3} = \frac{\overline{S_1S_2} \times \sin c_2}{\sin c_3} = \frac{500 \times \sin 82.9°}{\sin 28°} = 1056.9mm$$

$$X_3 = X_1 + \overline{S_1P_3} \times \cos c_1 = 0 + 1056.9 \times \cos 69.1° = 377.0mm$$

$$Y_3 = Y_1 + \overline{S_1P_3} \times \sin c_1 = 0 + 1056.9 \times \sin 69.1° = 987.4mm$$

The calculated coordinates are (377.0, 987.4). Compared with the measurement result of that point, which is (374, 1000), the error in distance is 13mm.

4.3 Application Results Analysis

In this study, the proposed city image database system is studied for feasibility, and to build such a system is not the target of this research. In fact, this system cannot be constructed from assembling image collecting equipment to building image database and 3-D model in a short period of time. The equipment discussed above is only a preliminary design of the equipment that is needed for performing the proposed function. It is possible to neglect some problems in this design and to meet some difficulties in the implementation. However, it can be a reference for when a real image collecting equipment is built.

The potential of image processing based on HSV color space information is tested again as a tool to automatically extract objects of interest, such as buildings, trees and traffic signs. For distinguishing different traffic signs, after preliminary segmentation based on HSV, shape features are introduced and successfully act as the indicator of different traffic signs. Although not all the traffic signs are tested by this methodology, to find shape feature of each traffic sign and design a distinct algorithm is feasible since every traffic sign is a combination of a vivid color and an identifiable geometric shape. For a building with complex edges, this proposed automatic image processing methodology may have difficulties in detecting all the edges of that building. Using a pastel to mark the edges of that building first, and then extracting edges by image processing, can reach a more accurate result.

A 3-D information extraction algorithm is a key component of this system. Without it, it is difficult to identify a specific building. As the experiment shows, the accuracy of this algorithm might reach 1.5 meters in a distance of 100 meters. With that accuracy, it is impossible to treat one building for another by mistake, even if two buildings have the same color and shape features. With the help of that information, it is possible to link the name of a building with objects in the image database according to location. Then, semantic retrieval and query can make meaningful application possible. The proposed algorithms for location and elevation calculation will cause errors when they are applied to a tall building in a short distance for perspective issues. But these kind of errors can be compensated in algorithms.

Finally, with both the location and elevation information, it is easy to build a 3-D model. This system can be an efficient tool to build a large scale 3-D model for any municipality.

Chapter 5 Conclusion and Recommendation Remarks

5.1 General Conclusion

This thesis presents an image segmentation methodology that relies on a thresholding method of HSV (hue, saturation and value) color space, shape features and size features. This methodology, combined with other tools, such as object tracking methodology, can automatically discover helpful information. Two applications in construction equipment management testify the effectiveness of this promising methodology when it is applied in automatic construction management.

Similar image processing strategy is studied in order to detect objects besides other equipment on the construction site, such as buildings and traffic signs. Satisfying image processing results are achieved in tested targets and the potential of using a similar image processing strategy on three different applications illustrates its adaptability.

A city image database system is also tested to be a feasible research topic with practical functions. The proposed 3-D information extraction methodology can be applied to future 3-D image processing applications. Image database, with semantic retrieval functions and a 3-D model of a city, is the product of the proposed system. Based on the above two products, the changes of buildings' heights and traffic signs' locations in the city can be automatically recorded and tracked. This will improve the efficiency of city infrastructure management. Future works are also listed for the implementation of this research.

5.2 Research Contribution

First, the comparison between HSV and RGB color spaces shows the advantage of HSV color space when it is applied in object segmentation of images taken in open fields. The second contribution includes introducing an automatic equipment utilization analysis methodology based on image processing techniques, which can be a part of the efforts to achieve automatic construction. The third contribution includes presenting a method to automatically distinguish different traffic signs by their color and shape features. It may be helpful for the researches of automatic driving system, based on visible light detectors. The fourth contribution includes the methodology of extracting 3-D spatial information from 2-D images, which can build the foundation for further mining of information in image data. Finally, if the proposed city image database can be built after continuous researches, the system has potential to be an assistant to more scientific city management.

5.3 Recommendation for Future Work

5.3.1 Future Work of Excavator Idle Time Analysis

The proposed algorithm works very well in this application. High effectiveness and efficiency are achieved by the program. This program is suitable for idle time analysis of other equipment, with a vivid color, by resetting the object HSV information.

5.3.2 Future Work of Truck Load Cycle Time Analysis

The proposed methodology does not reach a result as ideal as in the first application. One reason is the sizes of targets in an image are too small to remove all noises, which bring a lot of inconvenience when image processing is applied. Moreover, the orientation of the camera causes the bucket of an excavator to occasionally cover parts of a target, which are small even without covering. It is believed to achieve a better result with image data taken from a shorter distance in a better orientation, since the tested images are not prepared for this application.

Due to the quality of images, location thresholds are introduced into this methodology for filtering noises. Those features have limits when extending the program to other image data to solve problems in the tested image data series. Generally, an efficient excavator does not move while loading. But the excavator may still change location during a long duration. When location limits, which are based on the prior knowledge, are introduced the program's expendability will be affected. A possible solution might be to detect objects in the small range around the excavator by tracking the excavator based on the same methodology in the first application.

5.3.3 Future Work of City Image Database System

As previously mentioned, this study was only conducted to determine the feasibility of the system. Many examinations are left for future researches. The first step is to build and test the image collecting equipment. The next step is to perfect image segmentation methodology by testing it on more objects. The third step is to design the specific algorithms for all traffic signs. As for 3-D information extraction algorithm, it should be tested and perfected through experiments of real buildings and the accuracy of real application should be estimated. The algorithm for locating buildings, with circle shape in layout, should be considered since no fixed corner points can be detected in that kind of building. A 3-D model building tool can be designed based on AutoCAD, Micro

Station or other programs. The most arduous work might be collecting image data and combining all above tools in a database system with good adaptability to accommodate other applications.

References

- Behrad A., Shahrokni A., and Motamedi S. A., "A robust vision-based moving target detection and tracking system". In the Proceeding of Image and Vision Computing Conference, University of Otago, Dunedin, New Zealand, November 2001.
- Beucher, S. (1991). "The watershed transformation applied to image segmentation." *Proc., 10th Pfefferkorn Conference on Signal and Image Processing in Microscopy and Microanalysis,* Cambridge, UK.
- Canny, J. (1986). "A computational approach to edge detection." *Transactions on pattern analysis and machine intelligence*, IEEE, 8(6), pp. 679-698.
- Chandan, C., Sivakumar, K., Masad, E., and Fletcher, T. (2004). "Application of Imaging Techniques to Geometry Analysis of Aggregate Particles." *J. of Computing in Civil Engineering*, ASCE, 18(1), pp. 75–82.
- ElGhandour, W. (2003). "Automated Real-Time Data Acquisition System", *10th Annual Canadian Construction Research Forum*, Edmonton, Alberta, Canada, Aug.
- Feng, C. and Wu, H. (2006) "Integrating fmGA and CYCLONE to optimize the schedule of dispatching RMC trucks", *Automation in Construction,* 15 (2006) pp. 186 – 199
- Haas, C. T., Saidi, K., Cho, Y. K., Fagerlund, W., and Kim, H. (2001). "Implementation of an Automated Road Maintenance Machine (ARMM)." 80th Annual Meeting, TRB, Washington, D.C.
- Iannizzotto, G. and Vita, L. (2000), "Fast and Accurate Edge-Based Segmentation with No Contour Smoothing in 2-D Real Images", IEEE TRANSACTIONS ON IMAGE PROCESSING, 9(7), pp. 1232-1237.
- Castelli, V. and Bergman, L. (2002) "Image Databases: Search and Retrieval of Digital Imagery" ISBNs: 0-471-32116-8 (Hardback); 0-471-22463-4 (Electronic)
- Kim, H., Haas, C. T., and Rauch, A. F. (2004). "Automated Quality

Assessment of Stone Aggregates Based on Laser Imaging and a Neural Network." *J. of Computing in Civil Engineering*, ASCE, 18(1), pp. 58–64.

- Kim, H., Haas, C. T., Rauch, A. F., and Browne, C. (2002). "Dimensional Ratios for Stone Aggregates from 3D Laser Scans." *J. of Computing in Civil Engineering*, ASCE, 16(3), pp. 175–183.
- Kramberger, I. (2005) "Real-time skin feature identification in a time-sequential video stream", *Optical Engineering* 44(4), 047201 (April 2005)
- Lou, J., Tan, T., Hu, W., Yang, H. and Maybank, S. J. (2005) "3-D Model-Based Vehicle Tracking", *IEEE TRANSACTIONS ON IMAGE PROCESSING*, VOL. 14, NO. 10, OCTOBER 2005, pp 1561-1569
- Neto, J. A., and Arditi, D. (2002). "Using Colors to Detect Structural Components in Digital Pictures", *Computer-Aided Civil and Infrastructure Engineering*, Vol. 17, pp. 61-67.
- Offrell, P., Sjogren, L., and Magnusson, R (2005). "Repeatability in Crack Data Collection on Flexible Pavements: Comparison between Surveys Using Video Cameras, Laser Cameras, and a Simplified Manual Survey," ASCE, *Journal of Transportation Engineering*, Vol. 131, No. 7, pp. 552-562.
- Smith, S. D., Osborne, J. R. and Forde, M. C. "Analysis of Earth-Moving Systems Using Discrete-Event Simulation", *Journal of Construction Engineering and Management,* December 1995, pp 388-396
- Sonka, M., Hlavac, V., and Boyle, R. (1999). "Image Processing, Analysis, and Machine Vision", *PWS Publishing*, Pacific Grove, CA, USA.
- Takahashi, K., Matsuura, M., Sugiyama, T. and Abe, K. (2005) "A Comparison of Color Coordinate Systems Using Color Classification and Image Segmentation Results of Humans", *Systems and Computers in Japan*, Vol. 36, No. 10, 2005, Translated from *Denshi Joho Tsushin Gakkai Ronbunshi*, Vol. J84-D-II, No. 7, July 2001, pp. 1378–1388
- The MathWorks (2003), *Image Processing Toolbox for Use with MATLAB: User's Guide Version 4*, The MathWorks, Inc., Natick, MA, USA.
- Wael E., Amir E. and Simaan A. "Automated Real-Time Data Acquisition System".

- Wu, Y., and Kim, H. (2004). “Digital Imaging in Assessment of Construction Project Progress”, Proc., 21st International Symposium on Automation and Robotics in Construction, IAARC, Jeju, Korea.
- Zou, J. and Kim, H. (2005). “Image Processing for Construction Equipment Idle Time Analysis”, Proc., 22nd International Symposium on Automation and Robotics in Construction, IAARC, Ferrara, Italy.