

Measuring Bridge Construction Efficiency Using the Wireless Real-Time Video Monitoring System

Yong Bai, F.ASCE¹; Jun Huan²; and Seonghoon Kim³

Abstract: To enhance the efficiency of bridge construction, the wireless real-time video monitoring system (WRITE) was developed. Utilizing the advanced technologies of computer vision and artificial neural networks, the developed system first wirelessly acquired a sequence of images of work-face operations. Then human pose analyzing algorithms processed these images in real time to generate human poses associated with construction workers who performed the operations. Next, a portion of the human poses were manually classified into three categories as effective work, contributory work, and ineffective work and were used to train the built-in artificial neural networks (ANN). Finally, the trained neural networks were employed to decide the ongoing laborers' working status by comparing the incoming images to the developed human poses. The developed system was tested for accuracy on a bridge construction project. Results of the test showed that efficiency measurements by the system were reasonably accurate when compared to the measurements produced by the manual method. Thus, the success of this research indicates promise for enabling project managers to quickly identify work-face operation problems and to take actions immediately to address these problems. DOI: 10.1061/(ASCE)ME.1943-5479.0000061. © 2012 American Society of Civil Engineers.

CE Database subject headings: Bridges; Construction; Neural networks; Monitoring; Imaging techniques.

Author keywords: Bridge; Computer vision; Construction; Efficiency; Neural networks; Real time.

Introduction

Construction of highway bridges has been closely monitored by government agencies, the engineering and construction communities, and the general public, particularly after the September 11, 2001, terrorist attacks, Hurricane Katrina, the I-35W bridge collapse, the 2004 tsunami in South Asia, and the 2010 Haiti earthquake. Highway bridges are the critical components of the nation's transportation network because a bridge controls the capacity of the highway system; a bridge is the highest cost per mile element of the highway system; and if a bridge fails, the highway system fails (Barker and Puckett 1997).

Previous research results show that to enhance the efficiency of highway bridge construction, there is a need to develop innovative technologies to provide real-time information to engineers and project managers so that they are able to make quick decisions regarding the efficiency of work-face operations (Bai and Burkett 2006). To address this urgent need, the wireless real-time video monitoring system (WRITE) was developed. Although the prototype WRITE is more advanced than the traditional efficiency measurement techniques employed in the construction industry,

engineers and project managers still need to look at the images to manually interpret the working status of a construction laborer (e.g., effective work, contributory work, or ineffective work) (Kim et al. 2009). Because the process of determining working status is time-consuming and subject to human errors and biases as a result of a lack of the required skills, the ideal situation would be to fully and automatically detect the working status of a laborer on the basis of the human poses shown on the captured images. This paper presents the development of human pose analyzing algorithms using artificial neural networks (ANN) that can be employed by the WRITE to automatically determine the efficiency of work-face operations. The developed algorithm was tested in a highway bridge construction project in Kansas.

The rest of the paper is organized as follows. First, the research objective will be outlined. Second, the results of previous research on construction efficiency measurement methods and human tracking will be briefly reviewed. Third, the development of the WRITE will be described. Fourth, the human pose analyzing algorithms using ANN will be presented. Fifth, the experimental study of the WRITE and the developed algorithms in a bridge construction project will be discussed. Finally, the conclusions and recommendations will be provided.

Objective and Scope

The primary objective of this research project was to use ANN to develop human pose analyzing algorithms that can be employed to automatically determine the efficiency of work-face operations on the basis of the images of laborers captured by the WRITE in the construction site. The developed algorithms were tested in a steel girder bridge reconstruction project to determine its accuracy. The work breakdown structure (WBS) was used to break down the steel girder bridge into four levels, including Level 1 (project), Level 2 (work zone), Level 3 (activity), and Level 4 (operation). WBS is defined as "a deliverable-oriented grouping of project elements,"

¹Associate Professor, Dept. of Civil, Environmental and Architectural Engineering, The Univ. of Kansas, Lawrence, KS 66045 (corresponding author). E-mail: ybai@ku.edu

²Associate Professor, Dept. of Electrical Engineering and Computer Science, The Univ. of Kansas, Lawrence, KS 66045. E-mail: jhuan@ku.edu

³Assistant Professor, Dept. of Construction Management and Civil Engineering, Georgia Southern Univ., Statesboro, GA 30460. E-mail: shkim@georgiasouthern.edu

Note. This manuscript was submitted on February 22, 2010; approved on June 27, 2011; published online on July 11, 2011. Discussion period open until September 1, 2012; separate discussions must be submitted for individual papers. This paper is part of the *Journal of Management in Engineering*, Vol. 28, No. 2, April 1, 2012. ©ASCE, ISSN 0742-597X/2012/2-120-126/\$25.00.

Table 1. Examples of WBS for Steel Girder Bridge Reconstruction

Level 1 (project)	Level 2 (work zone)	Level 3 (activity)	Level 4 (operation)
Steel girder bridge	General	Mobilization	Set up crane
	Abutment	Traffic control	Moving concrete safety barrier
	Pier 1	Demolition	Driving pile
	Pier 2	Excavation	Forming
	Pier 3	Abutment 1	Structural excavation
	Pier n	Abutment 2	Slope protection (filter fabric and rock)
	Span 1	Pier drill shafts	Set bearing devices
	Span 2	Pier columns	Unload beams
	Span 3	Pier cap	Set beams
	Span n	Slope protection	Install diaphragms
		Beam setting	Bolting and tightening splice
		Deck forming	Ground splice
		Reinforcing deck	Prepare deck material
		Bridge barrier rail	Prepare deck forming
		Concrete barrier	Overhangs
		Backfill abutments	Strip
		Approach road	Place backwall (strip drain and backfill)
			Tying rebar
			Pouring and curing
			Strip and check elevation
		Other	

which organizes and defines the hierarchical structure of the entire project (Jung and Woo 2004). It is often used in complex construction projects to identify project information and improve the efficiency of control processes. Examples of the levels of the steel girder bridge WBS are shown in Table 1. With the WBS, it was possible to systematically measure on-site labor working efficiency at the operational level. For this research project, the operation of tying rebar was selected to develop and test the human pose analyzing algorithms because this operation was a labor-intensive task. If additional resources are available in the future, the human pose analyzing algorithms should be tested using the other bridge operations listed in Table 1.

The success of this research project contributes to the advancement of the bridge construction industry in two ways. First, it applies advanced technologies, such as computer vision and ANN, in construction operations. Second, the results of this research project make it possible to automatically determine the efficiency of work-face operations in real time. Thus, engineers and project managers will be able to quickly identify on-site labor efficiency problems and to take actions immediately to address these problems. As a result, the developed WRITE and algorithms enhanced the contractors' capability of managing bridge construction projects.

Literature Review

Literature review for this research project included two major topics: on-site construction efficiency measurement methods and human tracking. Efficiency or productivity data have been widely used as performance indicators to evaluate construction operations throughout the entire phase of construction (Thomas et al. 2003). Construction companies must continuously track operation efficiency to monitor their performance (Goodrum and Haas 2004). Hence, measuring efficiency at a project site has been an important task in the construction industry. Over the years, several on-site efficiency measurement techniques have been developed, including the stopwatch study and the time-lapse video. These techniques can

also be utilized to measure on-site construction productivity. In this research project, the efficiency measurement is viewed as an indirect measure of productivity.

Stopwatch study, also known as time study, has been used as the fundamental approach to measure efficiency since its invention in 1880 by Frederick W. Taylor, the father of scientific management (Meyers 1992). This is the oldest and most simple measurement method of efficiency for recording the duration of various construction activities.

The time-lapse video technique, one of the photographic techniques, has been used since the 1960s to view lengthy construction operations in a short period of time (Fondahl 1960). Pictures are taken using a special camera with 1–5 s intervals and are then converted to a video (Sprinkle 1972; Oglesby et al. 1989). Not only can a time-lapse video reduce the time spent viewing a construction operation, but it can also provide an accurate interpretation of a construction operation (Adrian 2004).

Since 2000, wireless technologies, such as the global positioning system (GPS), have been utilized to track the status of resources and construction operations. Navon and Shpatnitsky (2005) used GPS technology to automatically measure earthmoving performance by identifying the locations of equipment at regular time intervals and converting the information into project productivity.

All available efficiency measurement methods have some common limitations. First, these methods are unable to provide data for engineers and project managers to use to analyze the on-site construction operations in real time. Therefore, actions to address the on-site efficiency problems cannot be taken immediately. Although GPS technology can be used to transfer construction operational data in real time, these data cannot be utilized to analyze the on-site labor efficiency. Second, current methods do not provide a platform to share the real-time efficiency data among owners, engineers, contractors, and material suppliers. As a result, communication and coordination among project participants suffers.

Human tracking is a branch of computer science known as computer vision and has been a pressing research topic for many years.

Results of previous research indicate that existing computer vision technologies have the required capabilities of successfully recognizing objects such as flowers, soccer balls, tools, and materials (Brilakis et al. 2005; Shapiro and Stockman 2001); human parts such as the limbs, torso, and head (Mori and Malik 2002; Ramanan and Forsyth 2003); and human motions such as walking, running, kicking, and jumping (Bohn and Teizer 2009; Ioffe and Forsyth 2001; Lan and Huttenlocher 2005; Lim et al. 2006). Building on these successes, computer scientists have also developed technologies to automatically detect human actions such as dancing and playing tennis and football (Efros et al. 2003; Han et al. 2006; Yilmaz and Shah 2005). However, technologies that can automatically detect a construction activity and lead to the determination of the efficiency of work-face operations do not exist. A construction operation could be very complicated because of the mixed motion of human parts, tools, and materials that takes place at any given time and place.

To address the current shortfalls in construction efficiency measurement, the WRITE, along with human pose analyzing algorithms, was developed using advanced technologies such as computer vision and ANN. The framework of the WRITE will be introduced first, followed by the development of the human pose analyzing algorithms.

Development of the WRITE

The WRITE can provide pictorial data through a wireless network so that anyone in the construction field office or home office can monitor construction activities and analyze the efficiency of work-face operations in real time as long as there is an Internet connection. The WRITE offers several unique advantages. First, there is no disruption to construction operations. Second, the on-site construction efficiency can be determined in real time so that the engineers and project managers will be able to take actions immediately if necessary. Third, all participants involved in a

construction project can share the collected data through the Internet at any time and at any location.

The WRITE is composed of four major components, including the video camera house, the data processor, the ac transformer, and the computer. Along with these major components, additional items are required to operate the WRITE. These items include the wireless modems, a generator, a steel pole, and cables to connect the components. The framework of the WRITE, which shows the connection of the major components, is presented in Fig. 1.

Once the video camera has taken pictures from a construction site, the data processor immediately saves the pictorial data into files. These files are then transmitted in real time using wireless modems. An engineer or a project manager with the IP address at another location can access the data files with a wireless modem or a local area network (LAN) to conduct efficiency analysis. After finishing the data analysis, efficiency data and live pictures can be presented on a website so that other users, including the owners, engineers, contractors, and material suppliers, can share the information as long as they have an Internet connection. Although the proto type WRITE is more advanced than the traditional efficiency measurement methods employed in the construction industry, engineers or project managers still need to look at the video images to manually interpret the working status of a construction laborer. Because the process of determining working status is time-consuming and subject to human errors and biases, the ideal situation would be to the automatic detection of the working status of a laborer on the basis of the human poses shown in the captured images.

Development of Human Pose Analyzing Algorithms

In this research project, the human pose analyzing algorithms were developed to automatically determine the efficiency of a work-face operation (specifically, tying rebar in concrete bridge deck construction) using the images captured by the WRITE. The development of the human pose analyzing algorithms includes three

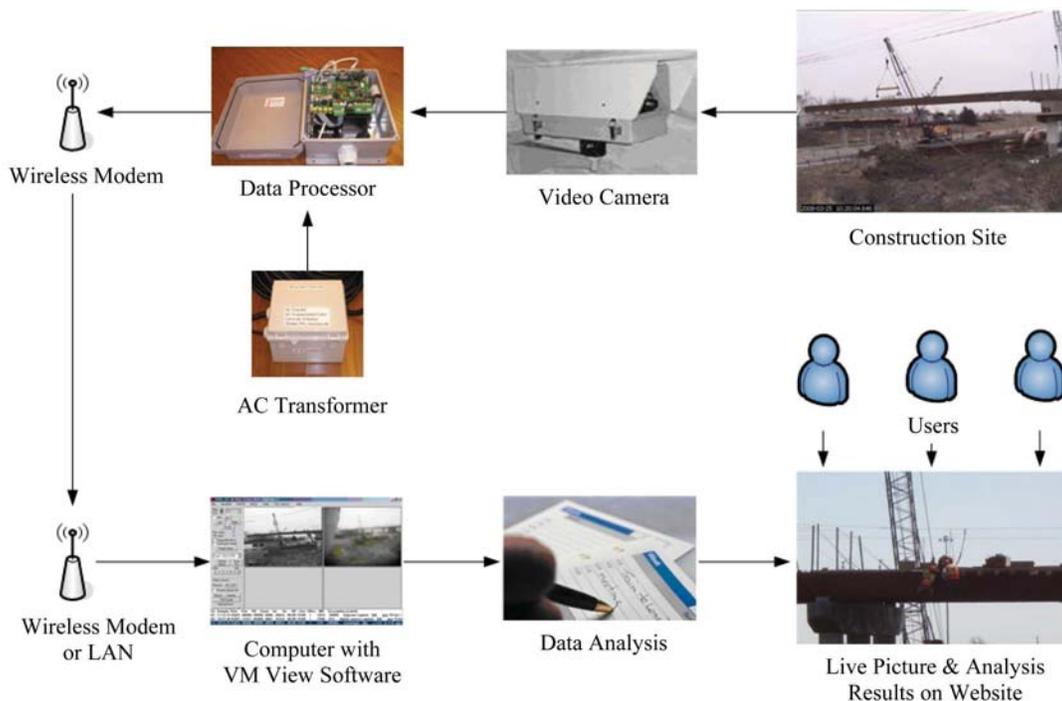


Fig. 1. Framework of the WRITE (images by Seonghoon Kim)

major steps: (1) image processing, (2) human pose creation, and (3) human pose classification.

Image Processing

Fig. 2 shows an example of tying rebar images acquired by using the WRITE. The incoming images were processed to identify and extract the construction workers captured in the photographs. When identifying workers, it was assumed that the workers were in motion over a period of time when performing a construction operation. Hence, a motion segmentation algorithm similar to the one in Lin et al. (2006) was developed to identify all the moving objects in these images.

However, as this approach captures several other moving objects that are irrelevant to the study (such as swaying trees and birds) an algorithm to filter the workers from a set of moving objects was also developed. First, if moving objects are of interest in an image, then all the other regions that are static are collectively classified as background. A moving average model, defined as follows, is used to compute this background.

Definition 1: If $I^N(x, y)$ is the intensity of the pixel at location (x, y) on the N th frame, then the intensity of the pixel on its background (BG) image for BG^N th at the same location (x, y) , computed over K frames is defined as

$$I^{BG}(x, y) = 1/K \left[\sum_{i=N-K/2}^{i=N+K/2} I^i(x, y) \right]$$

Second, before filtering out the irrelevant objects such as swaying trees and flying birds, the environmental noises need to be removed. When the background is computed, it is subtracted from the current frame (N th) to extract the workers from the image. This method focuses on a simple pixel-to-pixel subtraction holding a fixed threshold, which is good for noiseless environments. However, the construction operation (tying rebar) was performed outdoors, and the wind created a jitter in the camera's focus as a result of its mounted elevation. To overcome the inevitable sources of noise, a mask level background subtraction was developed.

Definition 2: If I_{avg}^N and I_{avg}^{BG} denote the average of pixel intensities of the mask M (typically 5) in the N th frame and its background, respectively, then the intensity of a pixel I^{SG} of the segmented image (SG) over the same mask is defined as

$$I^{SG}(x - i, y - i) = 1; \quad \text{if } 1/M |I_{avg}^N - I_{avg}^{BG}| > T = 0; \quad \text{otherwise}$$

where



Fig. 2. An image of construction workers tying rebar for the concrete deck (image by Seonghoon Kim)

$$I_{avg}^N = \sum_{j=-M/2}^{j=M/2} I^N(x - j, y - j)$$

$$I_{avg}^{BG} = \sum_{j=-M/2}^{j=M/2} I^{BG}(x - j, y - j)$$

where $i = [-M/2, M/2]$ and T is a predetermined threshold (typically 25).

Finally, once the moving objects are segmented and the environmental noises are removed, the irrelevant objects such as swaying trees and flying birds were filtered through the use of pattern matching. The assumption in filtration was that every worker in the crew wore identifiable clothing. The pattern of the clothing was determined, and a match was run over all the objects, considering a threshold T_V (typically 40). The pattern of an object was modeled as its statistical color variance described in definition 3 as follows.

Definition 3: If V denotes the variance over a mask M_V (typically 7) in the segmented image, and $V_{Standard}$ denotes the standard variance, then the intensity of a pixel I^F of the filtered image F over the same mask is defined as where

$$I^F(x - i, y - i) = 1; \quad \text{if } (1/M_V) |V - V_{Standard}| \leq T_V = 0; \quad \text{otherwise}$$

where

$$V = \sum_{k=-M_V/2}^{k=M_V/2} \left(I^{SG}(x - k, y - k) - 1/M \sum_{j=-M/2}^{j=M/2} I(x - j, y - j) \right)^2$$

where $i = [-M_V/2, M_V/2]$; and T_V is a predetermined threshold (typically 40).

After filtering the irrelevant objects, an image of a person can be described by using a 256×256 matrix. Each entry of this matrix could be any number from 0–255. The next step was to characterize the workers and to assign them a specific identity. The outline of each worker was identified, and a silhouette was created by clustering pixels in the outline with the region growing technique, in which similar pixels are grouped together to form a single region and are then given a distinct color (Adams and Bischof 1994). Certain attributes for each silhouette (such as centroid location, area, and the minimum and maximum along both the axes) were computed and stored dynamically in a database. Fig. 3 presents silhouettes corresponding to three workers shown in Fig. 2 (where they are identified by an arrow).



Fig. 3. Silhouettes corresponding to workers in Fig. 2

Human Pose Creation

The poses of construction workers were extracted in every image by skeletonizing their corresponding silhouettes using the fast parallel algorithm for image silhouette thinning developed by Zhang and Suen (1984). This algorithm involves two subiterations, one for deleting the northwest corner points and southeast boundary points and another for deleting the southeast corner points and northwest boundary points. Each silhouette is thinned to form a pose (skeleton) that is 1 pixel thick. The end points and the pixel connectivity are preserved so that a continuous pose is obtained and the distortion is minimal. Each pose was represented using a 1×256 matrix. Fig. 4 shows poses obtained from silhouettes in Fig. 3.

The extracted poses are categorized into three groups on the basis of performance, including effective work, contributory work, and ineffective work. Any motion that is essential for progress and that adds to the completion of a construction operation is defined as effective work. Conversely, any motion that does not add to the completion of a construction operation is referred to as ineffective work. Contributory work is defined as any motion that is essential for progress but does not directly add to the completion of a construction operation. Because there are no precise, agreed-upon definitions of effective, contributory, and ineffective human postures in the construction industry, in this research project, effective, contributory, and ineffective poses were classified based on experience, on-site observations, images of workers, and consultations from industry experts. In addition, the same standard was kept throughout the entire project. For the construction operation of tying rebar, the ironworkers tie rebar by hand with pliers and tie wire. This work requires repeated fast hand and arm movements while applying high force. If ironworkers tie rebar at ground level (e.g., for a concrete bridge deck or concrete highway pavement), they also have to work in a stooped position with their bodies bent deeply forward. Through field observations, an ideal working cycle for tying rebar, along with working status, can be described in Table 2. In the real world, there will be ineffective work actions

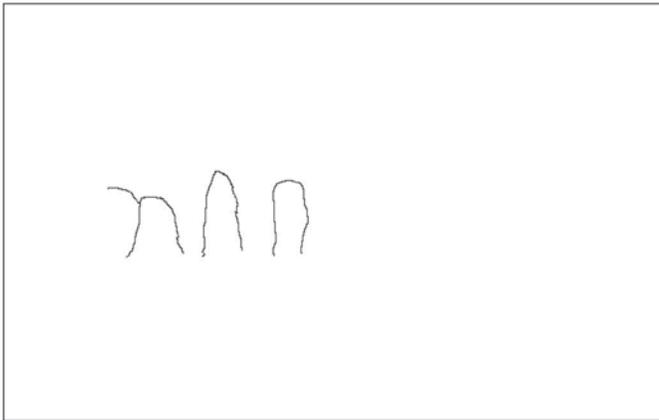


Fig. 4. Corresponding pose images

Table 2. Ideal Working Cycle for Tying Rebar

Number	Description of action	Working status
1	Back bending body to a tie position	Effective work
2	Tying rebar	Effective work
3	Body adjustment (such as bending back the body)	Contributory work
4	Move to the next tie position	Effective work

in between the effective work and contributory work operations. An example of ineffective work is the act of standing at a location doing nothing.

Human Pose Classification

An ANN was employed to classify human poses into effective, contributory, and ineffective categories. Several reasons exist for why ANN technology is suitable for this research project. First, an ANN can be used to model complex relationships between inputs and outputs and to find patterns in data. Second, algorithms can be designed to alter the weights of the connections in the network to produce a desired result. Finally, there is the possibility of learning, which has attracted the most interest in ANN. The learning capability enables researchers to use a small amount of captured data to predict the characteristics of a large amount of new data.

An ANN contains a large number of processing elements called neurons that are grouped in layers and linked together by weighted connections called synapses. A neuron performs two basic functions. It sums up the values at each input multiplied by the weight associated with each interconnection, and then it generates an output by passing the sum through an activation function. The interconnecting weights begin at random and are adjusted when the network undergoes training for a specific application. There are several algorithms available to adjust these weights and minimize the training error. The algorithm developed for this research project is based on the back-propagation learning method (Amin and Shekhar 1994). In this algorithm, the training of the network begins by feeding the inputs through the input layer. The inputs were a sequence of human poses, and each pose was represented in a 1×256 matrix. The network output is then computed and compared with the desired output, which was manually annotated as effective work, contributory work, or ineffective work. The resulting error is fed back to the network using the input layer. This process is repeated iteratively until the resulting training error is acceptable or the specified number of iterations has been completed.

The implementation of the developed algorithm undergoes two working phases: the learning phase and the execution phase. In the learning phase, training data sets (human poses) are first manually annotated such that each pose is assigned to one of the three classes: effective work, contributory work, or ineffective work. The annotated training data sets are the used to train the neural networks. In the execution phase, the trained neural networks will classify the incoming human poses (new poses) into the categories of effective work, contributory work, or ineffective work. During this process, the developed algorithm will update the records of a construction worker with the total number of effective, contributory, and ineffective corresponding poses. The efficiency of the worker $P_A(W)$ is defined and computed as:

$$P_A(W) = \left[\frac{1}{2} \right] [P_{A1}(W) + P_{A2}(W)] = \sum P_{An}(W)/n \quad (1)$$

where

$$P_{A1}(W) = \frac{[I_{\text{effective}}(W) + I_{\text{contributory}}(W)]}{[I_{\text{effective}}(W) + I_{\text{contributory}}(W) + I_{\text{ineffective}}(W)]}$$

$$P_{A2}(W) = \frac{[I_{\text{effective}}(W)]}{[I_{\text{effective}}(W) + I_{\text{contributory}}(W) + I_{\text{ineffective}}(W)]}$$

$I_{\text{effective}}(W)$ = total effective instances of the worker W
 $I_{\text{ineffective}}(W)$ = total ineffective instances of the worker W
 $I_{\text{contributory}}(W)$ = total contributory instances of the worker W .

Experimental Study

A bridge construction project was utilized for the experimental study to determine the accuracy of the developed algorithms. The construction operation, tying rebar, was recorded using the WRITE as a series of images at a rate of 1 frame per s. These images were 720×480 pixels in size, with a resolution of 96 dots per inch (dpi). Two workers performing the same operation, shown in Fig. 2, were chosen and labeled as W1 (left) and W2 (right), respectively. A set of 1,000 image frames for both W1 and W2 was selected for the experimental study. The respective poses of each worker were manually classified to generate the ground truth data shown in Table 3.

The poses obtained from W1 were used to train the neural networks. These poses were manually divided into three groups as effective, contributory, and ineffective. Each pose was represented by using a 1×256 matrix, and the matrix was input into the neural networks for training. Training data sets were created by randomly selecting an equal number of poses from each group. The neural networks were first trained over these selected training sets. Then the trained neural networks were tested for accuracy by using them to analyze and classify all of the poses of W1 and W2 (training data sets were returned to the original groups). Because these neural networks were trained on W1 poses, testing them on W1 poses was straight testing and testing them on W2 poses was cross testing. It should be noted that W2 performed the same operation as W1. This entire procedure of training and testing formed a single experimentation cycle. For this research project, each cycle was repeated 200 times. Through the trial-and-error method, it was determined that when the size of the training data set was 26 from each group, the developed algorithm produced reliable, stable, and maximum performance (Peddi et al. 2009). Continuously increasing the size of the training data set had no significant improvement on the final results and only increased the time necessary to complete the experimentations.

Work-face efficiency measurements for W1 and W2 using the manual method and the developed algorithms using ANN with a training data set size of 26 are presented in Table 4. As indicated, by using the developed algorithms to predict the efficiency of W1 (straight-testing method), the result of the algorithm method was almost the same as the outcome of the manual method (74 versus 73%). When using the algorithms to predict the efficiency of W2 (cross-testing method), the result of the algorithm method was 81% of the measurement of the manual method (60% versus 74%). In

Table 3. Ground Truth Data for Experimental Study

Worker ID	Number of frames	Effective frames	Contributory frames	Ineffective frames
W1	1,000	637	180	183
W2	1,000	717	48	235

Table 4. Efficiency Measurements Using Manual and ANN Methods

Worker ID	Number of frames	Analysis method	Effective frames	Contributory frames	Ineffective frames	Efficiency (%)
W1	1,000	Manual	637	180	183	73
		ANN	583	312	105	74
		% of accuracy	92	58	57	99
W2	1,000	Manual	717	48	235	74
		ANN	531	147	322	60
		% of accuracy	74	33	73	81

Note: % of Accuracy = (ANN)/(Manual), if (ANN) < (Manual); % of Accuracy = (Manual)/(ANN), if (ANN) > (Manual); Manual and ANN efficiency rates were calculated using Eq. (1).

general, the straight-testing method produced more accurate results than the cross-testing method based on the percentage of accuracy measurements. The developed algorithms had the highest accuracy in predicting the effective work and the lowest in predicting contributory work.

Conclusions and Recommendations

All of the available efficiency measurement methods in construction have some common limitations. First, these methods are unable to provide data for engineers and project managers to analyze the efficiency of work-face operations in real time. Therefore, actions to address the on-site efficiency problems and to improve construction operations cannot be taken immediately. Second, current methods do not provide a platform to share the real-time efficiency data among owners, engineers, contractors, and material suppliers. As a result, communication and coordination among construction project participants suffers. Third, existing efficiency measurement methods require a continuous human involvement to analyze the data. Such methods are prone to errors because of human biases and limitations, in addition to the time delay. To address these shortcomings, the WRITE and the human pose analyzing algorithms using ANN were developed to automatically determine the efficiency of work-face operations in the construction project site.

The WRITE, with the human pose analyzing algorithms, was utilized in the following fashion. Images of construction workers were acquired using the WRITE at a construction site. The captured images were first processed to identify and extract the construction workers in the images and to filter out irrelevant objects. Second, a silhouette of each worker was created by clustering pixels with the region growing technique, in which similar pixels were grouped together to form a single region and were given a distinct color. Third, human poses associated with construction workers were extracted in every image by skeletonizing the corresponding silhouettes using the image silhouette thinning algorithm. A portion of the extracted poses were manually categorized on the basis of performance as effective work, contributory work, or ineffective work. Fourth, the classified human poses were utilized to train the neural networks. Finally, the trained neural networks were used to determine the efficiency of construction workers shown in the incoming images on the basis of the pattern match.

The developed algorithms were tested for accuracy using straight-testing and cross-testing methods. Two construction workers (W1 and W2) performing the activity of tying rebar in a bridge construction project were selected for the experimental study. A total of 1,000 of each worker were taken to form the ground truth data. A total of 26 human poses in each performance category taken from W1 were utilized to train the neural networks. After the training, the neural networks were used to predict the efficiency for both W1 and W2, and results were compared to the outcomes of the manual method. By using the developed algorithms to predict

the labor efficiency in the straight-testing, the result of the algorithm method was almost the same as the outcome of the manual method. When using the algorithm to predict labor efficiency in the cross testing method, the result of the algorithm method was 81% of the measurement of the manual method. Examining the individual performance category, the results indicated that the developed algorithm had the highest accuracy in predicting the effective work and the lowest in predicting contributory work.

Real construction operations are very complicated as a result of the many factors involved. Several issues, such as the number of cameras needed to cover the entire construction site and obstructions in complex job sites, have not been addressed in this research project because of the resource limitation. All of these topics need to be investigated in the future. In addition, the developed algorithms can be extended in several ways. First, the developed human poses were two-dimensional (2D) images, and these images might not be able to clearly differentiate the distinctive features of different human beings. Therefore, there is a need to conduct further research to improve the human pose representations. Second, only human poses related to tying rebar were created in this research. Therefore, there is a need to develop human poses related to other construction operations. Third, further research is needed to improve the accuracy of predicting contributory work because the developed algorithms had the poorest performance in identifying this work category in the experimental study. Fourth, the training poses were manually categorized into three groups as effective, contributory, and ineffective, and the procedure was subjective. Improvement should be made in future research to minimize human biases in this procedure by using the motion of consecutive human poses. Finally, future research is needed to implement the developed WRITE and human pose algorithms in a real-time application and to determine whether the WRITE can detect labor efficiency problems and make suggestions for project managers to initiate improvements.

Acknowledgments

This research was funded in part by the Transportation Research Institute at the University of Kansas from Grant No. DT0S59-06-G-00047, provided by the Research and Innovative Technology Administration, the U.S. Department of Transportation. This research was also funded in part by the National Science Foundation under Grant No. 0741374. The financial support from these agencies is greatly appreciated. The field experimental site was provided by the Kansas Turnpike Authority. The authors would also like to thank Mr. Ken Johnson and Mr. Mike Laird from BRB Contractors, Inc. and Mr. Ray Rinne and Mr. Christopher Rech from A & M Cohron Inc. for their support during the field experimental study. Finally, the authors want to thank Mr. Abhinav Peddi, a graduate research assistant at the University of Kansas, for his contribution to the development of human pose analyzing algorithms.

References

- Adams, R., and Bischof, L. (1994). "Seeding region growing." *IEEE Trans. Pattern Anal. Mach. Intell.*, 16(6), 641–647.
- Adrian, J. J. (2004). *Construction productivity: Measurement and improvement*, Stipes, Champaign, IL.
- Amin, M. B., and Shekhar, S. (1994). "Customizing parallel formulations of backpropagation learning algorithm to neural network architectures: A summary of results." *Proc., 6th Int. Conf. on Tools with Artificial Intelligence*, IEEE Computer Society, Los Alamitos, CA, 181–189.
- Bai, Y., and Burkett, W. R. (2006). "Rapid bridge replacement: processes, techniques, and needs for improvements." *J. Constr. Eng. Manage.*, 132(11), 1139–1147.
- Barker, R. M., and Puckett, J. A. (1997). *Design of highway bridges*, Wiley, New York.
- Bohn, J. S., and Teizer, J. (2009). "Benefits and barriers of monitoring construction activities using hi-resolution automated cameras." *Proc., of 2009 Construction Research Congress*, Vol. 1, ASCE, Reston, VA, 21–30.
- Brilakis, I., Soibelman, L., and Shinagawa, Y. (2005). "Material-based construction site image retrieval." *J. Comput. Civ. Eng.*, 19(4), 341–355.
- Efros, A. A., Berg, A. C., Mori, G., and Malik, J. (2003). "Recognizing action at a distance." *IEEE Int. Conf. on Computer Vision*, Narosa Publishing House, New Delhi, India.
- Fondahl, J. W. (1960). "Photographic analysis for construction operations." *J. Constr. Div.*, 86(1), 9–25.
- Goodrum, P., and Haas, C. (2004). "The long-term impact of equipment technology on labor productivity in the U.S. Construction industry at the activity level." *J. Constr. Eng. Manage.*, 130(1), 124–133.
- Han, T. X., Ning, H. Z., and Huang, T. S. (2006). "Efficient nonparametric belief propagation with application to articulated body tracking." *Proc., 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, Los Alamitos, CA.
- Ioffe, S., and Forsyth, D. (2001). "Human tracking with mixtures of trees." *IEEE Int. Conf. on Computer Vision*, IEEE Computer Society, Los Alamitos, CA.
- Jung, Y., and Woo, S. (2004). "Flexible work breakdown structure for integrated cost and schedule control." *J. Constr. Eng. Manage.*, 130(5), 616–625.
- Kim, S., Bai, Y., Huan, L., and Peddi, A. (2009). "Measuring construction productivity using the write system." *Proc., 2009 Construction Research Congress*, Vol. 1, ASCE, Reston, VA, 131–140.
- Lan, X. Y., and Huttenlocher, D. P. (2005). "Beyond trees: Common-factor models for 2D human pose recovery." *Proceeding of the 10th IEEE Int. Conf. on Computer Vision*, IEEE Computer Society, Los Alamitos, CA.
- Lim, H., Morariu, V. I., Camps, O. I., and Szaier, M. (2006). "Dynamic appearance modeling for human tracking." *Proc., 2006 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, IEEE Computer Society, Los Alamitos, CA.
- Lin, H. W., Yang, S. Q., Xia, Z. J., and Kang, C. Y. (2006). "A moving objects detection approach for smart sensor." *Proc., Int. Conf. on Machine Learning and Cybernetics*, IEEE Computer Society, Los Alamitos, CA, 3751–3754.
- Meyers, F. E. (1992). *Motion and time study: work methods and management*, Prentice Hall, Upper Saddle River, NJ.
- Mori, G., and Malik, J. (2002). "Estimating human body configurations using shape context matching." *Proc., 7th European Conf. on Computer Vision*, Springer, London.
- Navon, R., and Shpatnitsky, Y. (2005). "Field experiments in automated monitoring of road construction." *J. Constr. Eng. Manage.*, 131(4), 487–493.
- Oglesby, C. H., Parker, H. W., and Howell, G. A. (1989). *Productivity improvement in construction*, McGraw-Hill, New York.
- Peddi, A., Huan, L., Bai, Y., and Kim, S. (2009). "Development of human pose analyzing algorithms for the determination of construction productivity in real-time." *Proc., 2009 Construction Research Congress*, Vol. 1, ASCE, Reston, VA, 11–20.
- Ramanan, D., and Forsyth, D. A. (2003). "Finding and tracking people from the bottom up." *Proceedings of CVPR*, IEEE Computer Society, Los Alamitos, CA.
- Shapiro, L. G., and Stockman, G. C. (2001). *Computer vision*, Prentice Hall, Upper Saddle River, NJ.
- Sprinkle, H. B. (1972). "Analysis of time-lapse construction films." *J. Constr. Div.*, 98(2), 183–199.
- Thomas, H. R., Minchin, R. E., and Chen, D. (2003). "Role of workforce management in bridge superstructure labor productivity." *J. Manage. Eng.*, 19(1), 9–16.
- Yilmaz, A., and Shah, M. (2005). "Recognizing human actions in videos acquired by uncalibrated moving cameras." *Proc., 10th IEEE Int. Conf. on Computer Vision*, IEEE Computer Society, Los Alamitos, CA.
- Zhang, T. Y., and Suen, C. Y. (1984). "A fast parallel algorithm for thinning digital patterns." *Commun. ACM*, 27(3), 236–239.