

3D Matching of resource vision tracking trajectories

Eirini KONSTANTINO¹, Ioannis BRILAKIS²

¹PhD student, Construction Information Technology Group, Laing O'Rourke Centre for Construction Engineering and Technology, Department of Engineering, University of Cambridge, email: ek415@cam.ac.uk

²Laing O'Rourke Lecturer of Construction Engineering, Department of Engineering, University of Cambridge, UK, email: ib340@cam.ac.uk

Abstract

Issues related to management and workforce play a key role in the productivity gap of construction and manufacturing. Both issues are directly related to the way productivity is measured. Current measurement methods tend to be ineffective because they are labour intensive, costly and prone to human errors whereas they are mainly reactive processes initiated after the detection of a negatively influencing factor. So far, research efforts in automating the measuring process have not achieved full automation because they require prior knowledge of the type of tasks performed in specific working zones. This is associated with the lack of depth information. For this purpose, this paper proposes a computationally efficient computer vision method for matching construction workers across different frames based on epipolar geometry, template and motion matching methods. The main result of this process is to provide a method for the acquisition of the 4D features (x, y, z, t) that compose the detailed profile of a construction activity in terms of both time and space.

1. INTRODUCTION

Construction industry, as highlighted by Teicholz (2004) based on data taken from the US Bureau of labour statistics, was seriously lagging in productivity from 1964 to 2003. The highest gap was reached in 2004 where the construction sector was left behind by 100%. In the construction industry productivity is a pace given by the labour output per work hour for each completed task (Thomas et al. 1990). The need to improve the performance of construction projects, has motivated many researchers to study the factors that are responsible for low labour productivity (Dai et al. (2009); Ng et al. (2004); Picard (2004); Wambeke et al. (2011)). A comparison of previous studies (Cundechea (2012); El-gohary and Aziz (2014); Jarkas and Bitar (2012); Kuykendall (2007); Lim and Alum (1995); Makulsawatudom et al. (2004)) concludes that more than 77% of the negatively influencing factors are directly related to activities that take place on site (e.g. congestion, lack of materials, disruption, absenteeism, fatigue). The remaining 23% refers to external offsite factors, including problems out of the range of a jobsite (e.g. age, motivation), and internal offsite factors which have an indirect side-effect relationship to the onsite activities (e.g. weather, law regulations). The interesting observation regarding the highest category is that all these factors are directly connected to the measuring process.

However, currently applied methods for measuring labour productivity rely on observation techniques and manual collection of construction operation

details and they are mainly based on statistical analysis taken from representative work samples (Dozzi and AboutRizk (1993); Navon and Sacks (2006); Shehata and El-Gohary (2011)). As Navon and Sacks (2006) argue the measuring techniques have not changed significantly over the years. The authors also point out that construction managers do not have a clear and solid opinion regarding crucial subjects, such as the amount of time spent or specific activities, the productivity of each task and the reasons behind identified problems. Moreover, collecting data with high frequency and extent becomes a cumbersome process when the collection is manual. Given that in a construction site, multiple activities take place simultaneously and spread along a large area, i.e. excavation works, concrete pouring, etc., the task of recording everything in detail becomes time consuming and labour intensive. Therefore, problems most of the time are first detected (e.g. delays, congestion, lack of materials, absenteeism) and then reported by the surveillance engineers in order to proceed to corrective actions.

As described above, it is clear that there is a relationship between the factors that affect productivity and the level of implementation of automation in construction. To overcome the aforementioned limitations, this paper proposes an automatic method for matching construction resources (e.g. workers, machinery). The motivation that lies behind, is that productivity can be evaluated based on spatiotemporal (x, y, z, t) trajectory analysis. In order to extract these 4D trajectories, workers need to be initially detected, tracked and matched by applying computer vision methods. In this paper, we present the results of this process. The method was tested with data collected from a stereo camera system. The performance featured 86% precision, 79% recall and 78% accuracy. In the following section, we present the current state of research in automating the way productivity is measured while we also discuss studies which are related to our research objective. Then, we analyse the proposed solution using data from real case scenarios with workers intersecting their paths and performing their tasks close to each other and we present the experiments we performed to validate the proposed solution. In the last section, we present the conclusions of this study and a brief description of future work.

2. BACKGROUND

Current research efforts focus on the automation of measuring labour productivity in order to overcome the inefficiencies of current practices. Such efforts are based either on processing the spatiotemporal 3D information taken from GPS, UWB and RFID or on computer vision methods.

The former, are limited by the requirement of advanced knowledge regarding the type of tasks performed at specific work zones (T. Cheng et al. (2011); Tao Cheng et al. (2013); Hildreth et al. (2005); Navon and Goldschmidt (2003); Navon and Sacks (2006); Sedehi (2010)). Moreover, significant manual effort and cost is also required for maintaining and installing multiple tags (RFID JOURNAL 2014).

However, the latter have showed significant advantages, such as lower cost and rich data collection for post process analysis, ability to resolve issues of low-pace performance and train labourers (Bügler et al. 2014; Gong and Caldas (2010); Weerasinghe and Ruwanpura (2010)) as well as the potential to easily detect possible causes of malfunctions. Golparvar-Fard et al. (2013) and Zou and Kim (2007) focused on earthmoving's equipment characteristic posture features,

identifying efficiently the working state (e.g. filling, dumping, hauling, digging, moving or non-working) and transforming it into productivity values. Similarly, Bai et al. (2008) utilize posture data through skeletonization for workers.

Overall, current research in automating productivity measurement is significantly constrained by two parameters. The first is the need of processing only a single type of task (e.g. excavation, hoisting, brick wall). The second is related to the lack of depth information (3D data). Therefore, current studies are obliged to provide prior knowledge about the work zone of each task.

Aiming to overcome the aforementioned limitations, M. Park et al. (2012) propose a triangulation method for extracting 3D tracking trajectories of construction resources based on epipolar geometry. Nevertheless, their methodology is constrained by the need of manually matching the entities with their corresponding 2D coordinates. Except from this stereovision approach that requires at least two cameras, depth information can also be recovered with only one camera through the utilization of world known positions (Lamza, et al. (2013); Tiwari (2010)). In detail, the former study incorporated the camera's height and angle whereas the latter used a predefined set of lines with known distances from the camera, in order to compute the absolute position of the tracked objects with respect to the camera.

Concluding, all of the above approaches, for the evaluation of productivity are related to two main parameters. The first is associated to the unstable pattern of the construction tasks, in time and space i.e. construction activities are described by a multidimensional feature vector, as tasks get completed while the location changes. The second, refers to the diversity of activities e.g. hoisting, concrete pouring and casting. This paper focuses on solving the construction resources correspondence problem in order to achieve the extraction of workers' characteristic activity profile, from 4D trajectory paths (x, y, z, t).

3. PROPOSED SOLUTION

Addressing this paper's objective, we are proposing a method that automatically matches multiple construction workers from two different frames. The overall concept is analytically illustrated below (see Figure 1). The initial input data are videos taken from more than one cameras. In every N number of synchronized captured frames, a worker's detector (Park et al. (2012)) is applied, in order to automatically initialize a 2D kernel based tracker (Ross et al. (2008)). Moreover, the tracker's performance is improved by correcting its position through a worker's detector (ManWoo Park (2012)). Then the centroid of the tracker's bounding box, provides to the system, the workers' 2D position. Having identified the areas of interest in each frame, we then need to match the centroids of the same workers for being able to extract their world position (3D data).

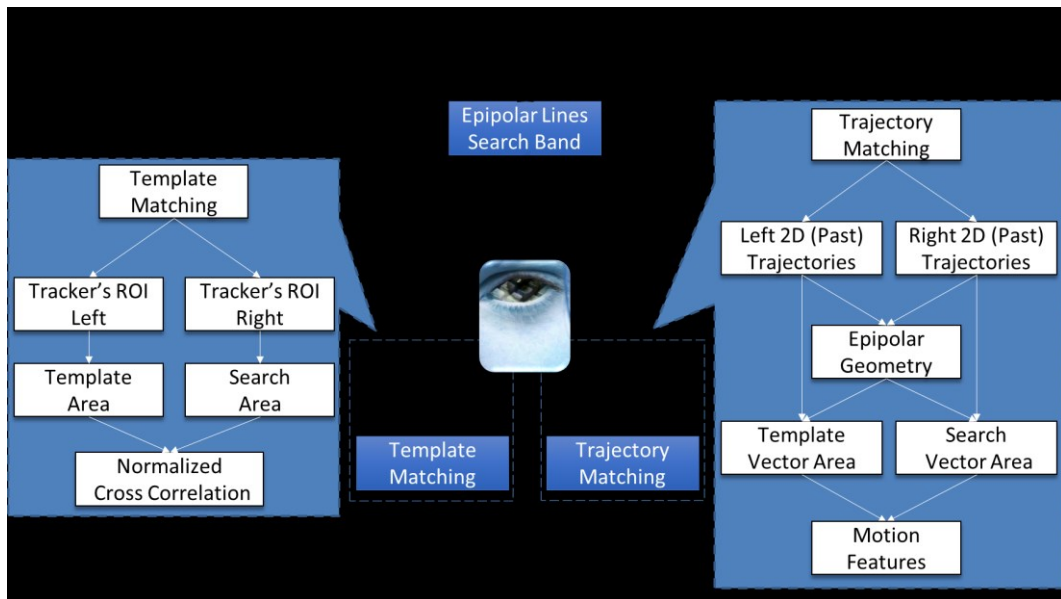


Figure 1. Overall proposed method for 3D matching using a stereo camera.

Epipolar geometry parameters are computed through stereo calibration. A checkerboard with known square size (60mm) was used for that purpose (see Figure 2). For better accuracy in 3D estimation, the calibration board was placed at approximately the same depth with the recorded tasks (Fathi and Brilakis 2014). Using epipolar geometry, the search area for finding one point's correspondence in the other view, is constrained along the epipolar line. However, the use of tracker's bounding box centroid creates some ambiguity, since worker's posture varies across different views. Therefore, the 1D search area was expanded equally (8%) after testing different thresholds, from both sides of each epipolar line (see Figure 3).

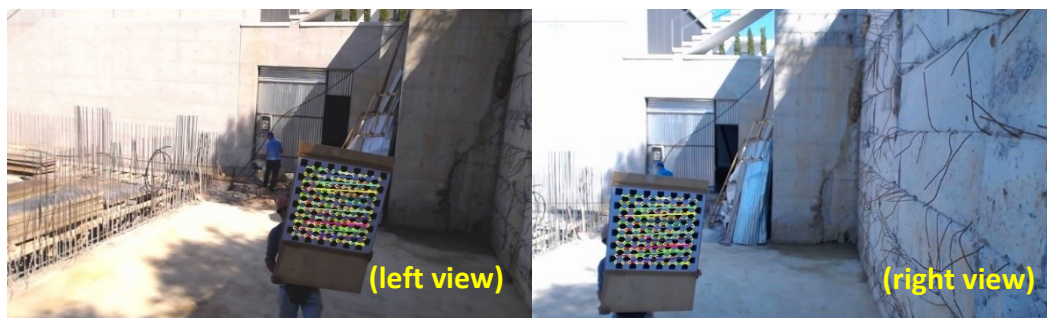


Figure 2. Stereo Calibration



Figure 3. Epipolar line's search band (left: centroid, right: epipolar line)

The matching across different views can easily become a multidimensional problem if more than one centroids lie within one epipolar line's search band. To overcome this complication, two similarity features are utilized based on worker's appearance and motion profile (see Figure 1). With regards to the first, normalized cross correlation was selected for template matching, since it is invariant to brightness and contrast variations (Perveen et al. (2010)). For computation efficiency, the tracker's areas of interest from each view are used as source $I(x)$ and template image $T(x)$. The similarity value is calculated based on the following equation:

$$R(x, y) = \frac{\sum_{x', y'} (T(x', y') * I(x+x', y+y'))}{\sqrt{\sum_{x', y'} T(x', y')^2 * \sum_{x', y'} I(x+x', y+y')^2}}$$

The second feature takes into account worker's past second motion 2D data. As illustrated in Figure 4 (see below), all the subvectors connecting the centroids of previously subsequent frames, are projected on the same plane. Then the average vector from each view is used as a comparison metric. In more detail, the algorithm searches to find a match for each average vector formed in one view (e.g. left), with one of the average vectors of another view (e.g. right frame) which are enclosed within the ABCD polygon, formed from the epipolar lines (see Figure 4). Since vectors are compared in order to have a positive match the angles θ & θ' have to be the equal as well.

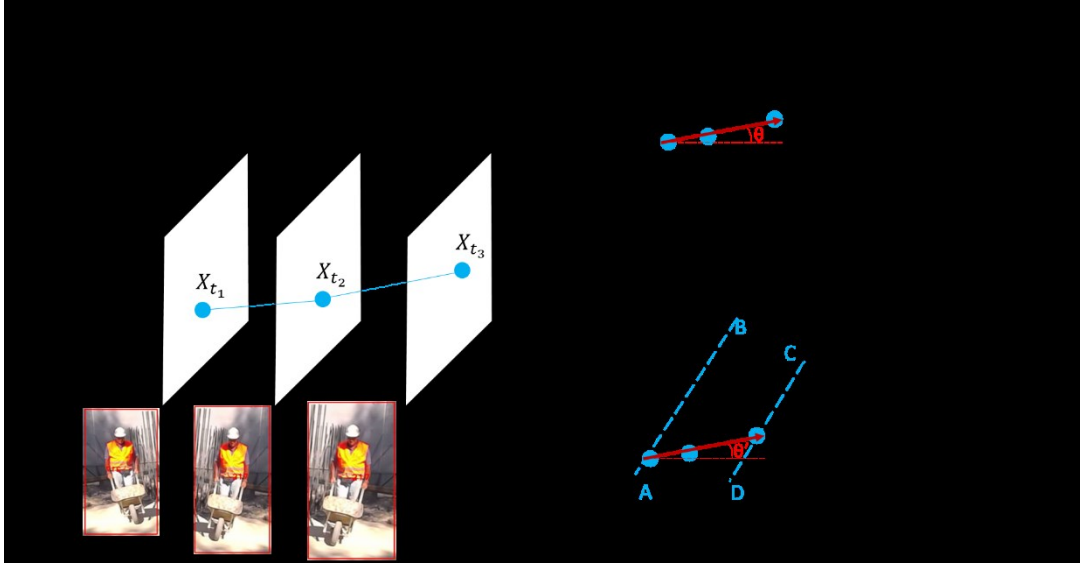


Figure 4. Workers motion matching based on epipolar geometry

The hypothesis tested in this paper is that the proposed method: (1) can be used for calculating workers 4D trajectories in a high rate (2) is computation efficient since only portions of the image are used for solving the correspondence problem. As mentioned above, the advantage of implementing computer vision methods for the calculation of construction recourses world position, is the ability of capturing multiple frames per second. In this way, the fourth dimension of time is captured in great detail.

4. IMPLEMENTATION AND RESULTS

The performance of the method presented in this paper was evaluated with a C# implementation in Microsoft Visual Studio.Net framework. It was developed in Visual Studio 2013 (IDE), running in a Windows 8.1 operating system. In order to call functions from OpenCV publicly available library, the Emgu CV platform was also installed. The data collection was performed with two web cameras Logitech 930e. The frame size provided is 1920x1080 pixels. The cameras were positioned at approximately 2.5 height and their in-between distance was 1m.

To measure the performance of the proposed method three metrics were used: precision, recall and accuracy. Precision is equal to the number of correctly matched workers (TP, True Positive) over the total number of correctly and incorrectly matched workers (TP + FP, True Positive + False Positive). Recall measures the methods' matching completion level and equals the number of workers correctly matched (TP) over the total number of correctly matched and not matched at all (TP + FN, True Positive + False Negative). Accuracy is ultimately extracted, representing the average correctness of the matching method. Accuracy is equal to the number of workers correctly identified (TP) and the number of workers that should not be matched (TN, True Negative), over the total number of the matched workers.

Table 1. Performance results of the 3D proposed matching method

Performance metrics	#	Performance metrics	%
Total TP	439	Accuracy	78%
Total FP	71	Precision	86%
Total TN	212	Recall	79%
Total FN	115		

The 3D matching method was tested by processing 2 videos which 1131 frames each. The stereo data included three workers performing real construction tasks (e.g. cleaning, material transportation), random motion and occlusions due to intersections or limited field of view. The primary results, as shown in Table 1, were promising as they were characterized by 86% precision, 79% recall and 78% accuracy. Some of the results are illustrated below (see Figure 5-6)



Figure 5. One worker ID matching across stereo frames.



Figure 6. Two worker IDs matching across stereo frames.

5. CONCLUSIONS AND FUTURE WORK

In this paper, a computationally efficient method is proposed for solving the correspondence problem aiming to extract detailed (high frequency) 4D trajectories of construction workers. Epipolar geometry is primarily used by exploiting the centroid coordinates provided by a kernel based tracker. The method is enhanced with intensity and motion similarity values. Future work will focus on processing data with better frame rate since the demonstrated results in this paper were based only on 5 frames per second, in an effort to simulate the capturing rate of the construction site's surveillance system.

REFERENCES

- Bai, Yong, Jun Huan, and Abhinav Peddi. 2008. *Development of Human Poses for the Determination of On-Site Construction Productivity in Real-Time*. Lawrence, Kansas.
- Bügler, Maximilian et al. 2014. "A Comprehensive Methodology for Vision-Based Progress and Activity Estimation of Excavation Processes for Productivity Assessment." In *EG-ICE*, Cardiff, United Kingdom, 1–10.
- Cheng, T., M. Venugopal, J. Teizer, and P.a. Vela. 2011. "Performance Evaluation of Ultra Wideband Technology for Construction Resource Location Tracking in Harsh Environments." *Automation in Construction* 20(8): 1173–84.
- Cheng, Tao, Jochen Teizer, Giovanni C. Migliaccio, and Umberto C. Gatti. 2013. "Automated Task-Level Activity Analysis through Fusion of Real Time Location Sensors and Worker's Thoracic Posture Data." *Automation in Construction* 29: 24–39.
- Cundecha, Mahesh Madam. 2012. "STUDY OF FACTORS AFFECTING LABOR PRODUCTIVITY AT A BUILDING CONSTRUCTION PROJECT IN THE USA : WEB SURVEY." North Dakota University.
- Dai, Jiukun, Paul M. Goodrum, William F. Maloney, and Cidambi Srinivasan. 2009. "Latent Structures of the Factors Affecting Construction Labor Productivity." *Journal of Construction Engineering and Management* 135(5): 397–406.
- Dozzi, S.P., and S.M. AbourRizk. 1993. *Productivity in Construction*. Ottawa, Ontario, Canada. www.nrc-cnrc.gv.ca/irc.

- El-gohary, Khaled Mahmoud, and Remon Fayek Aziz. 2014. "Factors Influencing Construction Labor Productivity in Egypt." *Journal of Management in Engineering* (February): 1–9.
- Fathi, Habib, and Ioannis Brilakis. 2014. "Multistep Explicit Stereo Camera Calibration Approach to Improve Euclidean Accuracy of Large-Scale 3D Reconstruction." *Journal of Computing in Civil Engineering*: 04014120.
- Golparvar-Fard, Mani, Arsalan Heydarian, and Juan Carlos Niebles. 2013. "Vision-Based Action Recognition of Earthmoving Equipment Using Spatio-Temporal Features and Support Vector Machine Classifiers." *Advanced Engineering Informatics* 27(4): 652–63.
- Gong, Jie, and Carlos H Caldas. 2010. "Computer Vision-Based Video Interpretation Model for Automated Productivity Analysis of Construction Operations." *Journal of Computing in Civil Engineering, ASCE* 24(6): 252–63.
- Hildreth, John, Michael Vorster, and Julio Martinez. 2005. "Reduction of Short-Interval GPS Data for Construction Operations Analysis." *Journal of Construction Engineering and Management* 131(August): 920–27.
- Jarkas, Abdulaziz M., and Camille G. Bitar. 2012. "Factors Affecting Construction Labor Productivity in Kuwait." *Journal of Construction Engineering and Management* (July): 811–20.
- Kuykendall, Casey J O. 2007. "Key Factors Affeting Labor Productivity in the Construction." University of Florida.
- Lamza, Aleksander, Zygmunt Wróbel, and Andrzej Dziech. 2013. "Depth Estimation in Image Sequences in Single-Camera Video Surveillance Systems." In *Multimedia Communications, Services and Security / 6th International Conference, MCSS, Krakow, Poland, June 6-7*: Springer Berlin Heidelberg, 121–29.
- Lim, Ec, and Jahidul Alum. 1995. "Construction Productivity: Issues Encountered by Contractors in Singapore." *International Journal of Project Management* 13(1): 51–58.
- Makulsawatudom, Arun, Margaret Emsley, and Kuldej Sinthawanarong. 2004. "Critical Factors Influencing Construction Productivity in Thailand." *The Journal of KMITNB* 14(3): 1–6.
- Navon, Ronie, and Eytan Goldschmidt. 2003. "Can Labor Inputs Be Measured and Controlled Automatically?" 127(August): 437–45.
- Navon, Ronie, and Rafael Sacks. 2006. "Assessing Research Issues in Automated Project Performance Control (APPC)." *Automation in Construction* 16(4): 474–84.
- Ng, S.Thomas, R. Martin Skitmore, Ka Chi Lam, and Anthony W.C. Poon. 2004. "Demotivating Factors Influencing the Productivity of Civil Engineering Projects." *International Journal of Project Management* 22(2): 139–46.
- Park, ManWoo. 2012. "AUTOMATED 3D VISION-BASED TRACKING OF CONSTRUCTION ENTITIES." <https://smartech.gatech.edu/handle/1853/45782>.
- Park, Man-woo, Christian Koch, and Ioannis Brilakis. 2012. "Three-Dimensional Tracking of Construction Resources Using an On-Site Camera System." *Journal of Computing in Civil Engineering, ASCE* (August): 541–49.
- Park, Man-Woo, Evangelos Palinginis, and Ioannis Brilakis. 2012. "Detection of Construction Workers in Video Frames for Automatic Initialization of Vision Trackers." In *Construction Research Congress 2012, ASCE*, , 940–49.

- Perveen, Nazil, Darshan Kumar, and Bhardwaj Ishan. 2010. "An Overview on Template Matching Methodologies and Its Applications." *Electronic Instrumentation Customer* 17(10): 3–6.
- Picard, Hans E. 2004. "Driving Fown Construction Project Labor Cost." *Journal of Construction Management Association of America*: 1–10.
- RFID JOURNAL. 2014. "How Much Does an RFID Tag Cost Today?" <http://www.rfidjournal.com/faq/show?85>.
- Ross, David a., Jongwoo Lim, Ruei-Sung Lin, and Ming-Hsuan Yang. 2008. "Incremental Learning for Robust Visual Tracking." *International Journal of Computer Vision* 77(1-3): 125–41.
- Sedehi, Arya John. 2010. "LEVERAGING RADIO FREQUENCY IDENTIFICATION TECHNOLOGY FOR PRODUCTIVITY ANALYSIS IN HIGH-RISE CONSTRUCTION." Georgia Institute of Technology.
- Shehata, Mostafa E., and Khaled M. El-Gohary. 2011. "Towards Improving Construction Labor Productivity and Projects' Performance." *Alexandria Engineering Journal* 50(4): 321–30.
- Teicholz, Paul. 2004. "Labor Productivity Declines in the Construction Industry : Causes and Remedies." *AECbytes Viewpoint #4 4*. http://www.aecbytes.com/viewpoint/2004/issue_4.html.
- Thomas, By H Randolph et al. 1990. "Modeling Construction Labor Productivity." *J. Constr. Eng. Manage* 116(July): 705–26.
- Tiwari, Kaushik K. 2010. "A POLYNOMIAL BASED DEPTH ESTIMATION FROM A SINGLE IMAGE." Cornell University.
- Wambeke, Brad W., Simon M. Hsiang, and Min Liu. 2011. "Causes of Variation in Construction Project Task Starting Times and Duration." *Journal of Construction Engineering and Management* 137(September): 663–77.
- Weerasinghe, I. P. Tharindu, and Janaka Y. Ruwanpura. 2010. "Automated Multiple Objects Tracking System (AMOTS)." *ASCE - Construction Research Congress*: 11–20. <http://ascelibrary.org/doi/abs/10.1061/41109%28373%292>.
- Zou, Junhao, and Hyoungkwan Kim. 2007. "Using Hue , Saturation , and Value Color Space for Hydraulic Excavator Idle Time Analysis." *JOURNAL OF COMPUTING IN CIVIL ENGINEERING, ASCE* 21(August): 238–46.