VIDEO ANNOTATION FRAMEWORK FOR ACCELEROMETER PLACEMENT IN WORKER ACTIVITY RECOGNITION STUDIES

Liju Joshua * and Koshy Varghese

Department of Civil Engineering, Indian Institute of Technology Madras, Chennai, India
* Corresponding author (liju_joshua@smail.iitm.ac.in)

ABSTRACT: Automated recognition of worker activities has the potential in aiding quick assessment of labour productivity on construction sites. A novel method called accelerometer based activity recognition has been investigated and preliminary results show that it has good potential for deployment in construction environment. The major decisive factor influencing the performance of the activity recognition system is the location of the accelerometer on the human body. The objective of this study is to determine a-priori, the appropriate accelerometer location using videos of construction activities. A framework was developed to track the movement of body segments through observation using Anvil, a generic annotation tool. Locations of accelerometer are selected after evaluating the information gain of each body segment towards activity classification with due consideration to subject comfort and integration possibilities. A study of masonry activity using the framework identified the placement locations as right lower arm, left lower arm and waist. An experimental setup was arranged for determining performance of the accelerometer based activity recognition system for a mason working in uninstructed environment with accelerometers attached at selected locations. Activity recognition performance of the locations was evaluated with ten runs of 10-fold cross validation using multilayer perceptron algorithm. The results showed that classifier performances for the three locations have the same order of ranking as predicted by the framework. The activity recognition performance for the selected locations gave accuracies above 80% and it can be concluded that the proposed framework can be used for placing accelerometers at appropriate locations for activity recognition.

Keywords: Activity Recognition, Video Annotation, Information Gain, Masonry, Classifier Performance, Body Segments, Accelerometer Placement

1. INTRODUCTION

Accurately recognizing worker activities is an important task in work sampling studies for measuring labour productivity. Manual methods of recognizing and recording activity are both laborious and tedious. Currently, vision based methods have been investigated for automated detection of worker status in productivity studies [1]. But the accuracy of recognition is severely affected by moving backgrounds and varying light conditions of the construction environment. Moreover, image processing techniques are computationally intensive and manually dependant and, this limits the wide application of video in construction [2].

A novel method based on accelerometers has been investigated in construction and preliminary results have shown that it has good potential to be used for activity recognition in construction [3]. In this method, machine learning algorithms called as classifiers are trained using the features generated from the accelerometer data to classify and recognize construction activities.

The location at which an accelerometer is placed is an important consideration in accelerometer based activity recognition studies. Experimental studies in activity recognition show that the classifier performance is extremely location sensitive [4, 5]. Hence selecting the appropriate location of the accelerometer becomes critical in activity recognition studies. Bouten et al. [6] considers the place of attachment of accelerometers as an important issue and rates subject comfort as the first criteria in deciding the location. In studies where it was limited to use

single sensor, waist has been the preferred choice as it caused minimal discomfort. Knowledge about the ideal location will enhance the efficiency of the activity recognition system and this necessitates the need of the present study.

As per Godfrey et al. [7], the accelerometers are normally attached to the part of the body whose movement is being studied. Hence the study of body motion during activities will help in determining the body segments whose movements are conspicuous. Motion capture is a quantitative method that uses reflective marker's point of reference for documenting movement. Detailed human movement can be captured using this method, but it is a costly and lengthy process [8]. Notation systems like Labanotation also provide rich description of the kinematic structure of body movement [9]. But the notations are too complex and not suited for personal vocabularies of motion [10].

Video based annotation tools are simple and low cost alternative for studying human movement patterns. Video recordings provide maximal data on the subject and the situation and it can be replayed and reinterpreted. The content analysis of the video provides the ability to generate quantitative data to support the inferences from the qualitative observation [11]. Observation based methods OWAS [12] & PATH [13] have been used in construction, primarily for studying body postures.

This paper investigates a video annotation based evaluator for selecting locations through observation. A framework was developed for observing, tracking and evaluating body segment movements. Both video annotation and activity recognition field studies were carried out to test the framework for masonry activity. The paper is organized as follows. The framework for evaluating sensor locations is described in section 2. The video annotation and activity recognition studies carried out in masonry and the results are presented in section 3. The discussion of the results is given in section 4, followed by conclusion of the work.

2. FRAMEWORK FOR EVALUATING BODY SEGMENT MOVEMENTS

Figure 1 shows the framework developed for evaluating the body segment movements in order to identify the location for sensor placement. The framework up to step 6 is implemented in Anvil, a generic annotation tool [14]. Anvil's overall design is object oriented and is written in Java language. The annotation scheme has to be written in an XML specification file, according to a formal description of the tracks, elements, attributes and their possible values. The information gain evaluation and ranking is performed using Weka, which is a collection of machine learning algorithms and data preprocessing tools [15]. The step wise explanation of the framework is given below.

1. The construction activity of interest is captured through video.

2. The video file is imported in Anvil to carry out frame by frame observation. The work categories are identified and the frame labeling is done accordingly.

3. The frames are divided into segments each with a length of 4 seconds. This duration is appropriate for observing and annotating body movements.

4. Then the annotation of body segment movements is performed for each frame segment. The anatomy of able-bodied individuals can be adequately modeled with 23 segments and 14 joints [16]. Hence the events occurring in these segments and joints would provide an adequate basis for coding body movements [17]. The body segments are arranged in a hierarchical manner consisting of parent and child segments. In this phase, the tracking and annotation of active/inactive status of the child segments is performed independently by two observers.

5. The reliability of annotation is tested using Cohen's kappa statistic [18]. An inter-observer agreement value greater than 0.75 is acceptable or other wise the observers consult each other to resolve differences in the agreement matrix.

6. The annotation output is imported to spreadsheet and a truth table is generated with true or false values and this table represents the body segment movement pattern for the various activities.

7. The truth table is imported into Weka for carrying out information gain evaluation. Here every frame segment act as a training instance and the status of body segments represent the attributes with binary values of true or false.

```
         1. Video Capture
               │
               ▼
Identify work   2. Frame by frame   Frame labeling
categories ◄──   observation    ──►
               │
               ▼
         3. Frame
         segmentation
               │
               ▼
         4. Body segment
         annotation
               │
               ▼
         5. Test annotation
         agreement
               │
               ▼
         6. Body segment
         pattern generation
               │
               ▼
         7. Information gain
         evaluator
               │
               ▼
         8. Rank segments
               │
               ▼
Subject comfort ──► 9. Assign single ◄── Integration
                    locations          possibilities
```
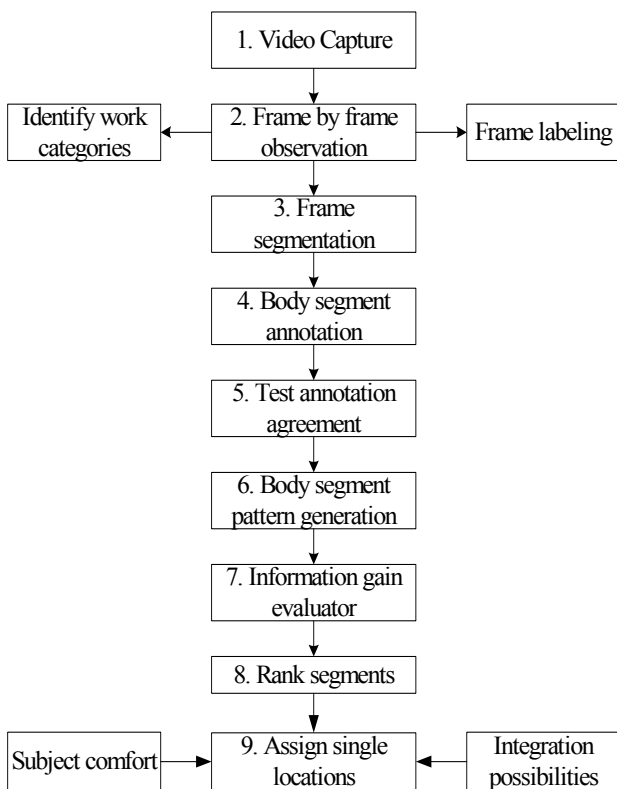
Fig.1 Framework to evaluate locations

Information gain is a statistical property that measures how well a given attribute separates the instances according to the target classification of work categories [19]. If the attribute in training instances can take 'c' different values, then the entropy of S relative to this c-wise classification is defined as given below, where P_i is the proportion of S belonging to class i. Information gain measures the expected reduction in entropy caused by classifying the training instances according to an attribute A. Thus information gain, Gain(S, A) of an attribute A, relative to a collection of training examples S, is defined as given below, where *Values (A)* is the set of all possible values for attribute A, and S_V is the subset of S for which attribute A has value v [19].

$$Entropy(S) \equiv \sum_{i=1}^{c} - P_i \log_2 P_i$$

$$\text{Gain(S, A)} \equiv Entropy(S) - \sum_{v \in Values(A)} \frac{|Sv|}{|S|} Entropy(Sv)$$

8. After evaluating the information gain the body segments are ranked.

9. The final assignment is done taking into consideration both the ranking of the body segments and, subject comfort and integration possibilities.

3. EXPERIMENTAL STUDY AND RESULTS

The experimental study was conducted in two stages. In the first stage a video annotation study was carried out for masonry work using the framework to identify the appropriate locations for the accelerometer. The results of this study were evaluated in the second stage consisting of both video annotation and activity recognition studies.

3.1. Video annotation study in masonry

A video annotation study was carried on recorded video of masonry work. The work categories identified are fetch and spread mortar, fetch and lay brick, adjust and level brick, and fetch and fill joint. The result of the video annotation study for the top ranked body segments is given in Table 1.

Table 1 Result of video annotation study (stage-1)

Rank	Body segment	Information gain (bits)	Remarks
1	Hand(R)	0.566	Not selected
2	Lower arm(L)	0.350	Selected
3	Lower arm(R)	0.280	Selected
4	Upper arm(R)	0.196	Not selected
5	Waist	0.182	Selected
6	Hand(L)	0.114	Not selected

The locations of right lower arm, left lower arm and waist were selected as appropriate locations for accelerometer placement to classify masonry activities. The locations of right hand, left hand and upper arm were not selected as they interfere with the activity of the worker.

3.2 Field experimental study

The experimental set up for the second stage study is shown in Fig. 2. A mason was given the task of finishing three layers of brick laying with the accelerometers placed on the right lower arm, left lower arm and waist as indicated in the figure

Fig. 2 Experimental set up

Accelerometer data logger (Fig. 3) with sample rate of 40 Hz and range of ± 6g from Gulf coast Data Concepts, LLC were used for the study. The accelerometers were firmly attached to arm bands at both lower arms and to waist belt at low back.

Fig. 3 Accelerometer
(Courtesy: Gulf Coast Data Concepts, LLC)

The details of the video annotation and activity recognition studies are given in the subsequent subsections.

3.2.1 Video annotation study

Video annotation study was again carried out to evaluate the different accelerometer placement locations on the human body. The work categories identified were identical to the stage-1 annotation study and the result of the video annotation study is given in Table 2.

Table 2 Result of video annotation study (stage-2)

Rank	Body segment	Information gain (bits)	Remarks
1	Hand(L)	0.301	Not selected
2	Upper arm(R)	0.280	
3	Hand(R)	0.225	
4	Waist	0.174	Selected
5	Lower arm(L)	0.149	
6	Lower arm(R)	0.057	

It can be observed that the list of top ranked segments is identical for both the annotation studies even if there is a difference in the rank order. The three locations of waist, left lower arm and right lower arm were verified as appropriate locations for masonry activity.

3.2.2 Activity recognition study

Activity recognition performance at the selected locations was evaluated through a process briefly described below, whose details are discussed in the previous work [3]. The accelerometer data features were trained with multilayer

perceptron algorithm, a neural network classifier. Ten runs of 10-fold cross validation are performed and the activity performance is obtained. The percent correct classifications, which is the activity recognition performance for the three different locations and the possible combinations is given in Table 3

Table 3 Activity recognition performance

Sr. No.	Sensor location	Percent correct
1	Lower arm(R)	80.57
2	Lower arm(L)	81.38
3	Waist	82.96
4	Lower arm(R&L)	87.65
5	Lower arm(L) & Waist	83.02
6	Lower arm(R) & Waist	82.94
7	Lower arm(R&L) and Waist	89.72

4. DISCUSSION

It can be seen from this study that 'locating accelerometer' is a critical decision in activity recognition studies. The results show that the selected locations for accelerometer placement were identical for the annotation studies carried out on two different masonry videos. The categories of masonry work identified were similar for both the annotations studies and this call for similar body actions resulting in comparable movement patterns. But differences can be observed in the value of information gain and order of ranking across the two annotation studies. This may be attributed to differences in the personal traits and work situations, but further studies are required to verify these differences.

It can be seen that ranking of selected locations using the framework and the activity recognition performance for these locations are in agreement. This shows that information gain, the statistical property could identify the body segments whose movements have the capability to discriminate among the classes of masonry activity.

The selected locations are appropriate for keeping the sensor as the classifier performance for all of them gave an accuracy of above 80%. The top ranked body segments could have given even better results but was not selected as placing accelerometer at hands and upper arm will interfere with the activities of the worker. In Table 3 it is seen that the activity recognition performance improved when data from more sensors were used for classifier training. This shows that construction activity recognition cannot rely on single accelerometer due to the complicated tasks involved. Ensuring reliable observations is a big challenge in observation studies. Employing trained observers and multiple replays of the frames may be required to code accurately the actively moving body segment. Further studies are planned to validate the framework across other trades in building construction. It has also been planned to identify accelerometer locations for different trades by conducting video annotation studies across large cross section of workmen in specific trades.

5. CONCLUSIONS

A video annotation framework for selecting location of wearable accelerometer in worker activity recognition studies was developed. Video annotation study in masonry and convenience of placement showed that right lower, left lower arm and waist are the appropriate placement locations of accelerometer. The results of the video annotation study were verified by conducting activity recognition studies at the selected locations. Preliminary results show that Anvil based framework can be used for selecting locations in activity recognition studies. It is proposed to carry out video annotation studies across different trades to determine a-priori, the appropriate accelerometer location for activity recognition studies.

REFERENCES

[1] Peddi, A., Huan, L., Bai, Y., and Kim, S.,

"Development of human pose analyzing algorithms for the determination of construction productivity", *Proceedings of Construction Research Congress 2009*, ASCE, Reston, Virginia, pp.11-20, 2009.

[2] Gong, J., and Caldas, C.H., "Computer vision-based video interpretation model for automated productivity analysis", *Journal of Computing in Civil Engineering*, Vol. 24 (3), pp.252-263, 2010.

[3] Joshua, L., and Varghese, K., "Accelerometer based activity recognition in construction", *Journal of Computing in Civil Engineering,* doi: 10.1061/ (ASCE) CP.1943-5487.0000097 (Article in press)

[4] Bao, L., and Intille, S.S., "Activity recognition from user-annotated acceleration data", *Pervasive 2004,* LNCS 3001*, pp.1-17, 2004.

[5] Lester, J., Choudhury, T., and Borriello, G., "A practical approach to recognizing physical activities", *Pervasive 2006*, LNCS 3968, pp.1-16, 2006.

[6] Bouten, C.V.C., Koekkoek, K.T.M., Verduin, M., Kodde, R., and Janssen, J.D., "A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity", *IEEE Transactions in Biomedical Engineering*, Vol. 44(3), pp.136-147, 1997.

[7] Godfrey, A., Conway, R., Meagher, D., and Olaighin, G., "Direct measurement of human movement by accelerometry", *Journal of Medical Engineering and Physics*, 30 (2008), pp. 1364-1386, 2008.

[8] Ebenreuter, N., "Dance movement: A focus on the technology", *IEEE Computer Graphics and Applications*, Vol. 25(6), pp.80-83, 2008.

[9] Foroud, A., and Whishaw, I.Q., "Changes in the kinematic structure and non-kinematic features of movements during skilled reaching after stroke: A Laban movement analysis in two case studies", *Journal of Neuroscience Methods*, 158(2006), pp.137-149, 2006.

[10] Kahol, K., Tripathi, P., and Panchanathan, S., "Documenting motion sequences with a personalized annotation system", *IEEE Multimedia*, Vol.13 (1), pp.37-45, 2006.

[11] Leicht, R. M., Hunter, S.T., Saluja, C., and Messner, J.I., "Implementing observational research methods to study team performance in construction management", *Journal of Construction Engineering and Management*, Vol.136 (1), pp.76-86, 2010.

[12] Karhu, O., Harkonen, R., Sorvali, P., and Vepsalainen, P., "Observing working postures in industry: Examples of OWAS application", *Applied Ergonomics*, Vol.12 (1), pp.13-17, 1981.

[13] Buchholz, B., Pacquet, V., Punnet, L., Lee, D., and Moir, S., "PATH- A work sampling-based approach to ergonomic job analysis for construction and other non-repetitive work", *Applied Ergonomics*, Vol. 27(3), pp.177-187, 1996.

[14] Kipp, M., "Anvil-A generic annotation tool for multimodal dialogue", *Proceedings of the 7th European Conference on Speech Communication and Technology (Eurospeech 2001)*, pp. 1367-1370, 2001.

[15] Witten, I.H., and Frank, E., *Data mining: Practical machine learning tools and techniques,* Morgan Kaufmann Publishers, San Francisco, CA, 2005.

[16] Kroemer, K.H.E., Kroemer, H.B., and Kroemer-Elbert, K.E., *Ergonomics: how to design for ease and efficiency*, Prentice Hall, 1997.

[17] Kahol, K., Tripathi, P., and Panchanathan, S., "Computational analysis of mannerism gestures", *Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04), IEEE,* Vol.3, pp.946-949, 2004.

[18] Frings-Dresen, M.H.W, and Kuijer, P.P.F.M., "The TRAC-system: An observation method for analyzing work demands at the workplace", *Safety Science*, 21(1995), pp.163-165, 1995.

[19] Mitchell.T.M., *Machine Learning*, McGraw-Hill, Singapore, 1997.