2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 25-29, 2020, Las Vegas, NV, USA (Virtual)

Prediction of Backhoe Loading Motion via the Beta-Process Hidden Markov Model

Kento Yamada[1], Kazunori Ohno[1,2], Ryunosuke Hamada[1], Thomas Westfechtel[1], Ranulfo Plutarco Bezerra Neto[1], Naoto Miyamoto[1], Taro Suzuki[3], Takahiro Suzuki[1], Keiji Nagatani[4], Yukinori Shibata[5], Kimitaka Asano[6], Tomohiro Komatsu[7] and Satoshi Tadokoro[1]

***Abstract*— Backhoe loads sediment onto the bed of dump trucks during earthmoving work. The prediction of backhoe loading time is essential for ensuring safe cooperation between the backhoe and dump trucks. However, it is difficult to predict the instant at which the backhoe is ready to load sediment, because of the similarity in motions observed during gathering sediment. Moreover, since operators have different skill levels, the prediction requires a unique model for each operator. In this study, we attempt to predict the instant at which the backhoe is ready to load sediment into the dump truck. For this purpose, the beta-process hidden Markov model (BP-HMM) is employed to build a backhoe motion model for a specific operator. Time series data of backhoe loading motions for crushed rocks and wood chips, which were measured using 6-axis inertial measurement unit (IMU) sensors equipped at the cab, boom, and arm of the backhoe, were used for modeling with the BP-HMM. Several primitive motions of the backhoe, which occur at the completion of preparation before the loading process begins, were discovered as a result of the motion modeling based on the BP-HMM. We developed the prediction of the instant using three primitive motions. At best, the proposed method could predict the instant with a probability of 67% and 100%, at 6.0 s and 0.7 s before the loading motions began, respectively. This phased prediction can be used to reduce the idle time and risk for dump trucks during earthmoving work with the backhoe.**

I. Introduction

We researched automated dump trucks that are used for earthmoving work in co-ordination with a backhoe operated by an individual [1] (Fig. 1). Recently, the automation of construction machinery has been rising [2], [3], [4]. In some developed countries, the number of workers in the construction industry is decreasing, owing to declining birth rates and aging populations. Automated construction machinery could compensate for the decrease in the number of workers and reduce the time required for construction.

A prediction of the instant at which a backhoe is ready to load sediment into a dump truck is crucial for cooperation between an automated dump truck and the backhoe operated by a human. The dump truck waits until the completion of a series of motions, performed by the backhoe, at a certain distance (Fig. 2 right). The backhoe scoops sediment and moves the bucket above the loading spot to signify the completion of the loading preparation. The dump truck then moves into the loading spot after acknowledging the signal from the backhoe. The prediction of the instant at which the backhoe is ready to load sediment into the dump truck makes earthmoving work more efficient, especially at large construction sites. At this instant, the backhoe operator has difficulty using voice or other commands to convey the completion of preparation for loading to the dump truck, because these additional tasks complicate the operation of the backhoe.

Fig. 1. Earthmoving work involving a backhoe and dump truck

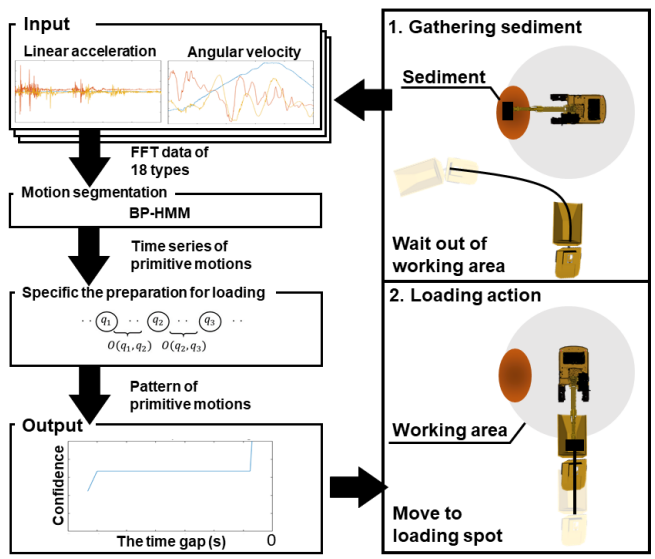

Fig. 2. Prediction as a trigger to move a dump truck for loading spot

[1]K. Yamada, K. Ohno, R. Hamada, T.Westfechtel, R. P. Bezerra Neto, N. Miyamoto, Takahiro Suzuki, S. Tadokoro are with Tohoku University, Japan
[2]K. Ohno is with the RIKEN Center for Advanced Intelligence Project
[3]T. Suzuki is with Chiba Institute University, Japan
[4]K. Nagatani is with Tokyo University, Japan
[5]Y. Shibata is with Sato komuten Co., Japan
[6]K. Asano is with Sanyo-Technics Co., Japan
[7]T. Komatsu is with KOWATECH co., Japan

However, it is difficult to predict this instant for a backhoe operated by a human. The manipulator tip position and its direction cannot be used for the prediction because the loading spot and its direction change according to the earthmoving progress. Moreover, the backhoe gathers sediment before loading action with the dump truck, and this process includes motions like scooping and rotation, which are similar to those carried out during preparation for loading. Therefore, these similar actions complicate the classification of these processes. This classification is crucial to predict the instant at which the dump truck should move into the loading spot.

This paper proposes the prediction of the instant at which a dump truck should move into the loading spot using a combination of primitive motions of the backhoe during earthmoving work in co-ordination with the dump truck. We employ the beta process hidden Markov model (BP-HMM) [5] to develop a model that isolates the primitive motions from the time-series data of the backhoe motions by a specific operator. The time-series data is measured during experiments using a backhoe and a dump truck with wood chips and crushed rocks. We extracted a combination of primitive motions that identifies when the backhoe is ready to load sediment into the dump truck, by analyzing the frequency of occurrence of primitive motions which is calculated for manually labeled data for "preparation for loading", "gathering sediment", "loading sediment".

The main contributions of this paper are as follows:

1) **Identification of primitive motions (extracted by the BP-HMM) that are common to preparation for loading by a specific operator**
   The primitive motions are extracted from backhoe motion data, operated by a specific operator, by capturing the earthmoving of wood chips and crushed rocks over 12 iterations. We visualize the frequency of the appearance of primitive motions for each process, namely, gathering sediment, preparation for loading, and loading sediment. Some primitive motions could be observed during each preparation for loading process.
2) **Extraction of a combination of primitive motions that correctly indicates preparation for loading with a probability of with high confidence over 95%**
   A combination of primitive motions is observed during preparation for loading and is rarely observed while gathering sediment. Therefore, this combination could classify these two processes.
3) **Prediction of the completion of preparation for loading with confidence several seconds before**
   We predict the instant at which a backhoe is ready to load sediment into a dump truck using a combination of primitive motions. This combination includes three kind of primitive motions. The confidence in the prediction increases by tracking the combination. At best, the instant is predicted with a probability of 67% and 100% at approximately 6 seconds before and 0.7 seconds after the instant at which the backhoe is ready to load sediment to the dump truck, respectively.

II. Related works

A few researchers have adopted learning methods to predict the backhoe behavior. Jinwoo Kim et al. combined Convolutional Neural Networks (CNN) and Double-layer Long Short Term Memory (DLSTM) to classify the behavior of a backhoe using five distinct actions from a camera [6]. Chieh-Feng Cheng et al. classified the backhoe motion into normal motions and strange motions using a Support Vector Machine (SVM) from auditory information [7]. Zongyao Jin et al. proposed Bayesian nonparametric clustering with ordering [8] for segmenting the backhoe operations of expert operators into primitive motions using string potentiometers and radio control inputs. The primitive motions are used for guiding beginner loading tasks. The method [8] is similar to our approach in the related studies described above. However, the method [8] is not suitable for modeling our earthmoving task because it involves a strong assumption about the order of operations, moreover the transition probability to a non-adjacent state is fixed by a small constant. Our target data are iterative data that contain the backhoe motions in different order depending on each process. The transition does not occur in the same order and the transition probability is not constant. Consequently, it is necessary to construct other data-driven approaches that can treat such primitive transitions. Therefore, we developed a BP-HMM based approach for the time prediction of the completion of preparation for loading.

Unsupervised parametric estimation based HMM is often used for the segmentation or prediction of time-series data. For example, it is applied to areas such as the prediction of stock prices [9], and speech recognition [10]. For the prediction of the completion of preparation for loading, we extract primitive motions from a specific operator using the BP-HMM and raw sensor data from IMU.

Trajectory data are analyzed for segmenting motions. A. Asahara et al. predicted the next step of pedestrians using their trajectory [11]. Y. Tanaka et al. predicted the next action of a worker depending on the order of the work and the area in which the worker was located [12]. Rudolf Lioutikov et al. segmented robot motions by analyzing trajectories and extracting primitive libraries without a transition of primitives [13]. Trajectory data analysis is also an interesting approach for analyzing motion. However, these methods cannot be applied to our task because our data is not trajectory but raw IMU sensor data. In our target scenario, backhoe tip position estimation using GNSS and IMU sometimes fails because of multipass from a forest.

III. Prediction for Backhoe Loading Motion

In this paper, we predict the instant at which a backhoe completes preparation for loading sediment into a dump truck. The earthmoving work by the backhoe comprises of three processes, namely, gathering sediment, preparation for loading, and loading sediment. The earthmoving between the backhoe and the dump truck requires the prediction of the completion of preparation for loading. The operators of the dump truck observe the backhoe motion and predict the

instant at which preparation for loading is completed. This prediction is necessary for the operators to determine the instant at which the dump truck needs to be positioned at the loading spot near the backhoe. In addition, the prediction before several seconds could reduce the idle time of the dump truck, and the backhoe expended while waiting.

Furthermore, this prediction should be advanced and employed as a trigger to move the dump truck into the loading spot (Fig. 2). The prediction of the completion of preparation for loading is essential to the cooperation between the backhoe and the dump truck. An incorrect prediction could lead to the dump truck positioning itself for loading while the backhoe is being operated and could lead to an accident. A delayed prediction could result in the dump truck arriving at the loading spot after the backhoe has finished preparation for loading, which would render the earthmoving work inefficient. The automated dump truck requires several seconds for path planning to the loading spot [14]. 10 times path planning took 2.8 seconds in the average. Therefore, we should predict the completion of preparation for loading several seconds before with probability over 50% which prevents the randomness of the prediction.

In this paper, we aim to predict several seconds before the completion of preparation with the confidence over 50% and detect with the higher confidence close to 100% after 1 second.

IV. Prediction of Backhoe Loading Motion based on the BP-HMM

A. Outline of Prediction besed on the BP-HMM

For measuring the backhoe motion, we attached IMU sensors, which can measure three-axis linear acceleration ± 16 G and angular velocity ± 250 deg/s at 200 Hz, to three parts of the backhoe showed in Fig. 3 (the arm, boom, and cab). Because, during earthmoving work, the arm, boom, and turning actuators are mainly employed. We also considered attaching an IMU sensor to the bucket to improve loading motion analysis. However, due to potential collision risk, the construction workers did not permit us to attach the IMU to the bucket.

Features of the backhoe motion $\boldsymbol{X}$, $18 \times L$ are extracted using the 18 axes inertial data $\boldsymbol{D}_{raw}$, $18 \times T$, L and T correspond to the length of the time windows and the time length of $\boldsymbol{D}_{raw}$, respectively. $\boldsymbol{D}_{raw}$ are measured using the three IMU sensors, attached to the cab, boom, and arm. Therefore, $\boldsymbol{D}_{raw}$ includes the three-axis linear acceleration $\boldsymbol{a}_x$, $\boldsymbol{a}_y$, $\boldsymbol{a}_z$, and angular velocity $\boldsymbol{\omega}_x$, $\boldsymbol{\omega}_y$, $\boldsymbol{\omega}_z$. To extract $\boldsymbol{X}$, the intensity of the frequency component $\boldsymbol{H}(f)$, $18 \times L$ of $\boldsymbol{D}_{raw}$ are compared to the stationary situation and the active situation of the backhoe.

The use of frequency components allows for the noise caused by engine or body vibrations that is unrelated to the backhoe motion to be ignored, as well as the effect of gravity on the accelerations. The discrete fast Fourier transition (DFFT) is used to calculate $\boldsymbol{H}(f)$ (Eq. 1). The length of the window L and the gap between each window are set to 128 (0.64 seconds) and 64 (0.32 seconds), respectively. The

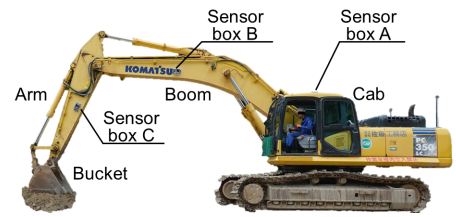

Fig. 3. Position of each sensor box and parts of the backhoe

Hann window changes $\boldsymbol{D}_{raw}$ to $\boldsymbol{D}$ for the detection of the peak in $\boldsymbol{H}(f)$ (Eq. 2).

$$\boldsymbol{H}(f) = \sum_{j=1}^{L} \boldsymbol{D}(j) w_L^{(j-1)(f-1)}, w_L = \exp(-\frac{2\pi j}{L}) \tag{1}$$

$$\boldsymbol{D}(j) = \frac{1}{2}(1 - \cos(2\pi \frac{j}{L}))\boldsymbol{D}_{raw}(j), 0 \leq j \leq L \tag{2}$$

The error in the angular velocity $\boldsymbol{\omega}$ for the IMU sensor increases owing to its drift. The offset $\boldsymbol{\omega}_{static}$ is measured using the average of each axis' angular velocity in the stationary situation and is subtracted from the raw data.

The direction of the joint rotation of the manipulator is essential to classify the backhoe motion because each direction is different for each motion, like scooping or loading. To represent this differences, the sign of the average Yaw-axis angular velocity in the window $\boldsymbol{\omega}_{Yaw}^{mean}$ is used. The features $\boldsymbol{X}_{Yaw}^{boom}$, $\boldsymbol{X}_{Yaw}^{arm}$ are defined with the frequency component of Yaw-axis angular velocity $\boldsymbol{H}(f)$ (Eq. 3).

$$\boldsymbol{X}_{Yaw} = \boldsymbol{H}(f) * sgn(\boldsymbol{\omega}_{Yaw}^{mean}) \tag{3}$$

The model of the backhoe motion during earthmoving is built using the BP-HMM from the features $\boldsymbol{X}$. The BP-HMM estimates the sequence of the primitive motions $\boldsymbol{Z}^{(i)}$, $1 \times L$, the probability of the transition between these primitive motions $\boldsymbol{\pi}^{(i)}$, $K \times K$, for each time series i, and the Gaussian distribution $\boldsymbol{\theta}_k$ of each primitive motion k, K corresponds to the number of the primitive motions. The BP-HMM is suitable to segment the primitive motions from the the backhoe motion during earthmoving work. The backhoe repeats three processes, namely, gathering sediment, preparation for loading, and loading sediment during earthmoving. Each branch and order between the motions is static and not random. Therefore, a model with the distribution of the transition of the primitive motions like the HMM is effective in expressing the earthmoving work of the backhoe. We employ the BP-HMM, which estimates the unknown number of primitive motions using the beta-process, which is data-driven using $\boldsymbol{X}$.

We then extract a pattern of a combination of primitive motions $\boldsymbol{S}_{key}$, which predicts the instant at which the backhoe finishes preparation for loading, from the sequence of primitive motions $\boldsymbol{Z}^{(i)}$. Some motions in each process have the same primitive motions. For example, gathering sediment and preparation for loading include the same motion in which the backhoe brings sediment closer together. Therefore, these two processes are considered to include the same primitive

motions. We focus on $\boldsymbol{S}_{key}$ to predict the completion of preparation for loading, which classifies this process with gathering sediment. We extract $\boldsymbol{S}_{key}$ with the probability of the appearance of the primitive motions for each process.

The motion model of the backhoe using the BP-HMM and the pattern $\boldsymbol{S}_{key}$ is used to predict the completion of preparation for loading using the evaluation data. The sequence of the primitive motion transitions $\boldsymbol{Z}^{(i)}$ from the evaluation data is estimated using the motion model of the backhoe. To predict the completion of preparation for loading, we observe $\boldsymbol{Z}^{(i)}$ and detect the instant at which $\boldsymbol{S}_{key}$ appears.

B. BP-HMM based segmentation of time series data

The BP-HMM segments the primitive motion for each time window l based on the Gaussian distribution of each primitive motion $\boldsymbol{\theta}_k$ from the features $\boldsymbol{X}^{(i)}$ of the backhoe motion. Fig. 4 shows the graphical model of the BP-HMM. We assume that the primitive motions of the backhoe follow the beta-Bernoulli distribution $\boldsymbol{b}_i$ and the number of the primitive motions K is infinite for the data-driven estimation of the unknown number using $\boldsymbol{X}^{(i)}$. The list of primitive motions $\boldsymbol{f}_i$ which appear in the time series i is estimated using the beta-process sampling from $\boldsymbol{b}_i$ based on hyperparameters γ, c. The distribution of the transitions $\boldsymbol{\pi}_i$, between primitive motions in time series i is calculated from the normalized $\boldsymbol{\eta}_i$ with existing primitive motions in $\boldsymbol{f}_i$. $\boldsymbol{\theta}_k$ is sampled using the Dirichlet process based on a hyperparameter λ. The sequence $\boldsymbol{Z}^{(i)}$ is composed of the time series of the primitive motion $z_l^{(i)}$, which is estimated from the posterior primitive motion $z_{l-1}^{(i)}$, the features $\boldsymbol{x}_l^{(i)}$, $\boldsymbol{\pi}^{(i)}$, $\boldsymbol{\theta}_k$. The number of the initial primitive motions K_{init} must be decided beforehand. The earthmoving work of the backhoe is divided into three processes P_h, and each process includes the motions mentioned below.

- P_1) **Gathering sediment** : Rotation with leveling, Scooping, Rotation with hauling, Rotation for scooping, Pushing, Loading, Sliding
- P_2) **Preparation for loading** : Scooping, Rotation with hauling
- P_3) **Loading sediment** : Scooping, Rotation with hauling, Loading, Rotation for scooping, Consolidating

Each motion is considered to include certain primitive motions like raising and slowly changing the heading direction. Therefore, we estimate that the earthmoving work of the backhoe consists of more than nine primitive motions that cover the stationary situation. In this paper, the number of initial primitive motions K_{init} is set to 17, other parameters are empirically determined, and the BP-HMM segments the primitive motions of the backhoe.

C. Prediction of timing with confidence

The pattern $\boldsymbol{S}_{key}$, which predicts preparation for loading (process P_2), is extracted from the sequence $\boldsymbol{Z}^{(i)}$. The probability of the appearance of the primitive motion in the process P_h, $p_{appear}(P_h, k)$ is calculated for the classification between P_2 and P_1. It is assumed that the number of

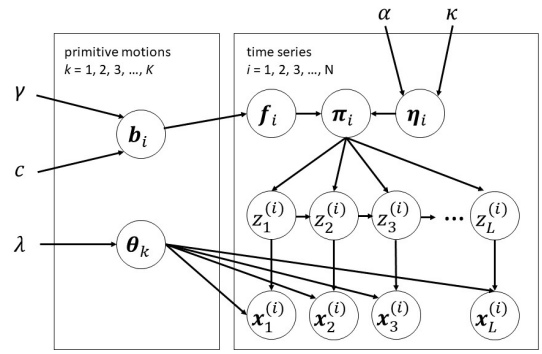

Fig. 4. Graphical model of the BP-HMM

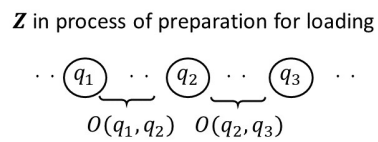

Fig. 5. Transition between primitive

processes, where the primitive motion k appears in the process P_h, is $n(P_h, k)$, and the number of the process P_h is $n(P_h)$, $p_{appear}(P_h, k)$ follows with Eq. 4.

$$p_{appear}(P_h, k) = \frac{n(P_h, k)}{n(P_h)} \tag{4}$$

To extract the primitive motions that characterize P_2, we focus on below three actions A_m of the backhoe in P_2.

- A_1) Scoop sediment
- A_2) Rotate to loading spot with hauling sediment
- A_3) Weaken velocity of rotation to stop at the loading spot

To constitute the pattern $\boldsymbol{S}_{key}$, we extract primitive motions q_m that appears in each action A_m and satisfies $p_{appear}(P_2, k) = 1$.

The feature of primitive motions that appear between q_m and q_{m+1}, $O(q_m, q_{m+1})$ is extracted by observation of the sequence $\boldsymbol{Z}$ (Fig. 5). The pattern $\boldsymbol{S}_{key}$, which indicates P_2, is given as Eq. 5 with a natural number u and the primitive motion $r_{mu} \in O(q_m, q_{m+1})$.

$$\boldsymbol{S}_{key} = [q_1, r_{11}, r_{12}, ..., r_{1v}, q_2, r_{21}, r_{22}, ..., r_{2w}, q_3] \tag{5}$$

V. Evaluation of Prediction

A. Backhoe loading motion data

The motion of the three parts of the backhoe was measured using the IMU sensors to build the model of the earthmoving motion of the backhoe, operated by a specific operator. We installed attachable sensor boxes to measure the inertial data of the 18 axes $\boldsymbol{D}_{raw}$ [15]. Fig. 6 shows the linear acceleration of the X-axis and the angular velocity of the Yaw-axis measured from each part of the backhoe. The details of the data from $\boldsymbol{D}_{raw}$ which includes four sets of process P_2 is shown in Table. I.

The intensity of the frequency component of $\boldsymbol{D}_{raw}$, $\boldsymbol{H}(f)$ was calculated using DFFT. Fig. 7 shows $\boldsymbol{H}(f)$ of the linear acceleration of the X-axis and the angular velocity of the Yaw-axis at the cab in the stationary and active situation. In

the stationary situation, the intensity of the high frequency, for example, 60 Hz of the X-axis linear acceleration, change more substantially than the intensity of the low frequency and resembles the noise from engine. The low-frequency of $\boldsymbol{H}(f)$ like 0-10 Hz is effective in capturing the motion of the backhoe avoiding the high frequency noise. These frequency components also help avoid the influence of the oscillation of the sign of $\boldsymbol{X}_{Yaw}^{boom}$ and $\boldsymbol{X}_{Yaw}^{arm}$ which caused by sensor noise and Eq. 3, because the intensities of these frequency components are close to zero and not large in the stationary situation. However, the difference in the posture of the backhoe changes H(f) of the linear acceleration close to 0 Hz. This range of the frequency divides the stationary situation of the backhoe into multiple primitive motions, which makes the segmentation of the motion complicated.

Therefore, we selected the range around 3.14 Hz as the feature of the linear acceleration and the range near 0 Hz as the feature of the angular velocity to extract the feature of the backhoe motion $\boldsymbol{X}$.

To evaluate our prediction methodology, the three datasets are used as the training data to build the model of the backhoe motion, and the other dataset is used as the evaluation data. The four datasets A, B, C, D was built using combinations of the four backhoe motion data (Table. I). For example, the dataset A used the data **b**, **c**, **d** as the training data, and the data **a** as the evaluation data.

TABLE I
DATA OF BACKHOE MOTION

Data	a	b	c	d
Sediment type	Wood chips	Wood chips	Crushed rocks	Crushed rocks
Time length (s)	2098.2	1302.7	1984.3	1745.0

B. Evaluation of extraction of primitive motions shared in preparation for loading

We compare the BP-HMM and HMM to evaluate the extraction of three primitive motions $\boldsymbol{Q}$, which is constituted of $[q_1, q_2, q_3]$, using the beta-process. The primitive motions, which satisfy the probability of the appearance $p_{appear}(P_2, k) = 1$ is extracted from the sequence of the primitive motions $\boldsymbol{Z}^{(i)}$ from the features $\boldsymbol{X}$ of the backhoe motion, operated by a specific operator. The number of the primitive motions K for the HMM is set 17 (HMM$_{init}$) which is the same as the number K_{init} and the number of the estimated primitive motions using the BP-HMM (HMM$_{esti}$).

C. Evaluation of prediction of timing and its confidence

To evaluate the prediction of the completion of process P_2, we extract the pattern $\boldsymbol{S}_{key}$ with the primitive motions $\boldsymbol{Q}$, which satisfies $p_{appear}(P_2, k) = 1$ with the methodology given in chapter IV-C. The sequence of the primitive motions for the evaluation data $\boldsymbol{Z}^{(i)}$ is estimated with the likelihood L_k between $\boldsymbol{x}_l$ and each primitive motion k, and the probability of the transition $\pi_{j,k}$ from primitive motion j to k. The likelihood L_k is calculated (Eq. 7) from the Gaussian

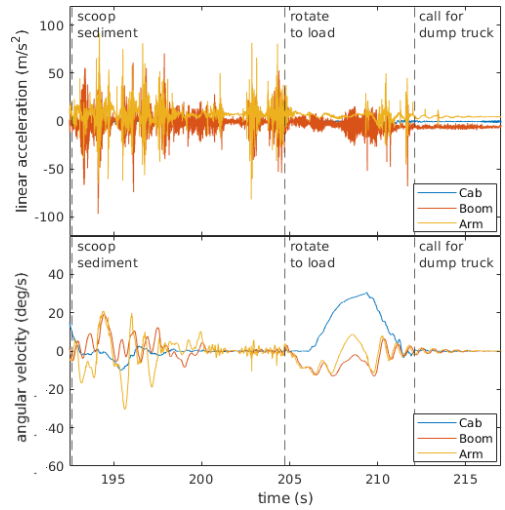

Fig. 6. Inertial data $\boldsymbol{D_{raw}}$ of the X-axis linear acceleration and the Yaw-axis angular velocity of the three parts of the backhoe

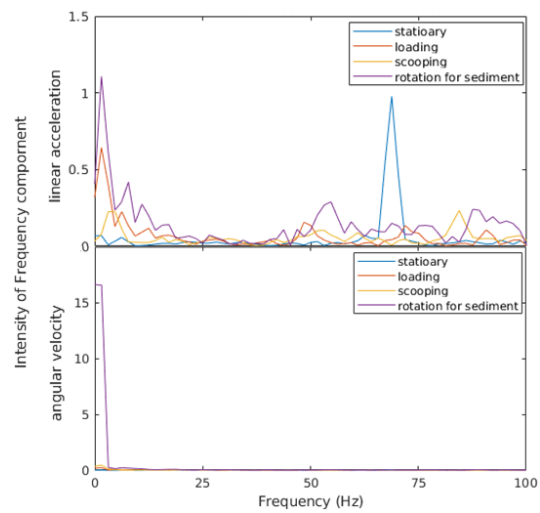

Fig. 7. Intensity of frequency component $\boldsymbol{H}(f)$ of the X-axis linear acceleration and the Yaw-axis angular velocity of the cab

distribution of each primitive motion $\boldsymbol{\theta}_k = \boldsymbol{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$, the Mahalanobis distance d_k (Eq. 6) between the feature $\boldsymbol{x}_l$ and $\boldsymbol{\theta}_k$. The primitive motion z_l is estimated using $p(z_l = k|\boldsymbol{x}_l)$, which is calculated using L_k and $\pi_{j,k}$ (Eq. 8, 9). Then, the angular velocity drift of the IMU is updated using the average of several seconds during the primitive motions shows of stationary state of the backhoe.

$$d_k = (\boldsymbol{x}_l - \boldsymbol{\mu}_k)^T \Sigma_k^{-1} (\boldsymbol{x}_l - \boldsymbol{\mu}_k) \tag{6}$$

$$\log(L_k) = -\log\left(\frac{\exp(-\frac{d_k}{2})}{\sqrt{|2\pi\boldsymbol{\Sigma}_k|}}\right) \tag{7}$$

$$p(z_l = k|\boldsymbol{x}_l) = \frac{L_k - \max(L_k)}{\sum_{i=1}^{K}(L_i - \max(L_k))}\boldsymbol{\pi}_{j,k} \tag{8}$$

$$z_l = \arg\max_k (p(z_l = k|\boldsymbol{x}_l)) \tag{9}$$

With the number $n_p(P_h)$ that the pattern $\boldsymbol{S}_{key}$ indicates process P_h, and the number of process $N_p(P_h)$ in training datasets, the confidence p_{conf} which the pattern $\boldsymbol{S}_{key}$ clas-

sifies P_2 from P_1 follows with Eq. 10.

$$p_{conf} = \frac{n_p(P_2)}{n_p(P_1) + N_p(P_2)} \tag{10}$$

We evaluate the confidence p_{conf} for each condition of the pattern $\boldsymbol{S}_{key}$ and the time gap from the actual point in time, which is calculated using the prediction of preparation for loading with estimated $\boldsymbol{Z}^{(i)}$ and the pattern $\boldsymbol{S}_{key}$ from the evaluation data (Eq. 10).

To evaluate the proposed method's performance, we compare the proposed BP-HMM based method with a threshold based method, and two LSTM-based methods, one trained to recognize the process of "preparing for loading" and the other trained for regression prediction of the time till loading movement. For the recognition of preparation, we prepared two types of labeling data, one contains the time series data with four different labels ("gathering sediment", "preparation for loading", "waiting", "loading sediment"), and the other contains the time series data with three different labels ("gathering sediment", "loading", "waiting"). After comparing both types of labeling, we found the latter set conveyed better performance. Therefore, we employed the data annotated with three different labels for the comparison. For the regression prediction of the time gap via LSTM, the time gap before loading $t^{pre} = T_i^{start} - t$ is prepared as input data for training. (Let T_i^{start} be the time before the maximum duration of the time used to prepare for loading. Here, 22.7 seconds was considered for the evaluation). Otherwise, t^{pre} was used as input data. In the LSTMs, the number of the hidden unit was set to 500, 300 for each of the recognition preparation and the regression prediction respectively, using the grid-search between 10-1300, and the number of iterations was set to 10,000 times. In addition, a simple threshold based approach was also evaluated and compared with the proposed method. We focused on the operation of lifting the boom while turning, which are characteristic motions that occur, before loading motion begins. On the basis of the data, thresholds that could detect these motions, were selected. After the angular velocity of the turning motion exceeds 18.0 deg/s, when the boom has the angular velocity which exceeds 8.80 deg/s or the angle of elevation is equal to greater than 64.3 degree, the instant is judged as "before loading". The four backhoe loading motion datasets were used as the input data. Average precision, recall, and the confidence were calculated using the four datasets for the evaluations.

VI. Results

The sequence of the primitive motions $\boldsymbol{Z}^{(i)}$ was estimated from each dataset with the model using the BP-HMM, two types of HMM, that are employed a different number of primitive motions. Fig. 8 shows the sequence $\boldsymbol{Z}^{(i)}$ from the dataset D using each model and the primitive motions q_m extracted from the model with the BP-HMM. Table II shows the number of the model that could extract q_m from each dataset. The BP-HMM could extract every q_m, but the two types of HMM did not extract. Especially, the two types of

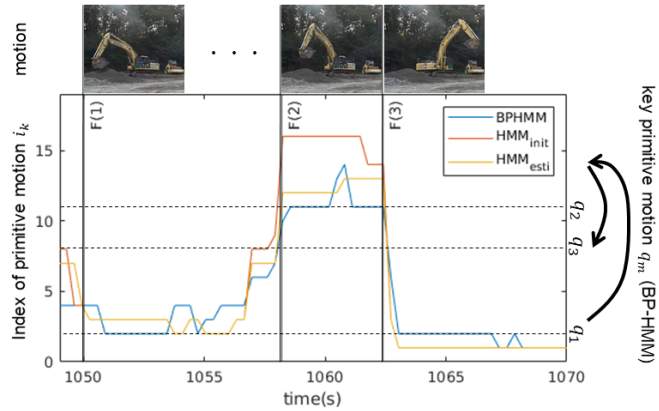

Fig. 8. The sequence of the primitive motion $\boldsymbol{Z}^{(i)}$ of the dataset D with each model and key primitive motions q_m extracted from BP-HMM model

TABLE II

THE NUMBER OF MODELS WITH EXTRACTED EACH q_m

	q_1	q_2	q_3
BP-HMM	4	4	4
HMM$_{init}$	3	4	0
HMM$_{esti}$	2	4	0

HMM did not extract q_3 which appeared in action A_3, when the backhoe reduces the velocity of rotation to stop at the loading spot to which a dump truck could easily come.

With the index of primitive motion i_k and the primitive motion r_{ku}, each feature of primitive motions which appear in between q_m and q_{m+1}, $O(q_m, q_{m+1})$ as follows:

$O(q_1, q_2)$: $i_{r_{1u}} > 1$

$O(q_2, q_3)$: $i_{r_{2u}} > i_{q_2} - 1$ detected, $i_{r_{2u}}$ kept down to i_{q_3}

The pattern $\boldsymbol{S}_{key}$ was constituted of the primitive motions $\boldsymbol{Q}$ and $O(q_1, q_2)$, $O(q_2, q_3)$. Fig. 9 (up) shows the confidence and the time gap from the completion of process P_2 detected from the model with the pattern $\boldsymbol{S}_{key}$. The pattern $\boldsymbol{S}_{key}$ from these datasets completely could classify P_2 from P_1. The pattern $\boldsymbol{S}_{key}$ from the dataset B detected wrong timing and missed P_2 onetime each other.

With the pattern $\boldsymbol{S}_{key}$ and the sequence of the primitive motions $\boldsymbol{Z}^{(i)}$ which was estimated using the model, we predicted the completion of P_2 from the evaluation with each dataset. Fig. 9 (down) shows the confidence and the time gap from the completion of P_2. From three datasets, P_2 was completely predicted. From only the dataset C, the pattern $\boldsymbol{S}_{key}$ detected P_1 only once, incorrectly.

Table. III shows the results of the proposed method, the LSTM-based recognition of preparation for loading, the LSTM-based regression prediction of the time till loading movement, and the threshold-based method. The proposed method exhibited better performance in terms of both precision and recall than that of two LSTM-based methods and the threshold-based method. The highest p_{conf}, 0.95 was also obtained from the proposed method using the BP-HMM model. From this result, it was confirmed that the proposed method performed the best of all the methods considered.

VII. Discussion

In section V-A, all of the primitive motions $\boldsymbol{Q}$, which were shared during every P_2 by a specific operator, were

TABLE III

THE COMPARISON WITH TWO TYPE LSTM AND THRESHOLDING

method	precision	recall	p_{conf}
Proposed method	0.95	1.0	0.95
LSTM (pattern recognition)	0.46	0.94	0.43
LSTM (regression prediction)	0.47	0.44	0.29
Thresholding	0.46	1.0	0.46

only extracted with the BP-HMM. The HMM$_{esti}$, which the number of the primitive motions was the same to the number estimated with the BP-HMM, missed the primitive motion q_3 which appeared when the backhoe weakened the velocity of rotation to stop at the loading spot, Therefore, the BP-HMM, that estimates the primitive motions with the beta-process, was suitable for the extraction of the primitive motions $\boldsymbol{Q}$ which characterized the motions in P_2. The point, which the beta-process was good at the extraction of unusual primitives [5], was considered to realize to extract the primitive motion q_3.

In section V-B, the pattern $\boldsymbol{S}_{key}$ from the three models completely classified P_2 from P_1. The three primitive motions $\boldsymbol{Q}$ in the pattern $\boldsymbol{S}_{key}$ realized the phased prediction, which was advanced and had a high confidence for determining the instant of the completion of P_2. The primitive motion q_3 and the feature $O(q_2, q_3)$ realized the confidence of the prediction 100%. The motion primitive q_3 was rarely detected in P_1. Therefore, the primitive motion q_3, which weakens the velocity of rotation to stop loading spot, is considered to especially characterize P_2 and to be useful for classification P_2 with P_1.

In section V-C, with the pattern $\boldsymbol{S}_{key}$ based on the BP-HMM, the completion of P_2 could be predicted with high confidence. From the three models, the pattern $\boldsymbol{S}_{key}$ predicted the completion of P_2 with the confidence of 1.0. The incorrect prediction probability of 40% when q_2 detected mainly consisted of motions during which the backhoe rotated to avoid a collision with the dump truck. In addition, q_2 was sometimes detected at which the backhoe rotated the upper manipulator from a low position owing to gathering sediment. However, q_3 appeared in P_2 with high probability and in P_1 with low probability. Therefore, the phased prediction using the three primitive motions, which includes q_3, helped upper the confidence of the prediction.

The proposed methods could predict the instant with high confidence because using three primitive motions. Otherwise, the confidence over 50 % was detected at the time gap over 3.0 in every dataset. With the velocity of the rotation of the heading direction at a value twice that of the specific operator, the time gap of early prediction with low confidence is considered to be within 1.5 seconds. However, 1.5 seconds is effective for dump trucks to prepare for moving early and helps reduce of the idle time of backhoe operators. Finally, the dump truck could move into the loading position with a high confidence over 95% on average.

Therefore, we considered that this prediction could be used as a trigger to move a dump truck to the loading spot with the model that includes enough motion data. Finally, the dump truck moves into the loading spot with high confidence early.

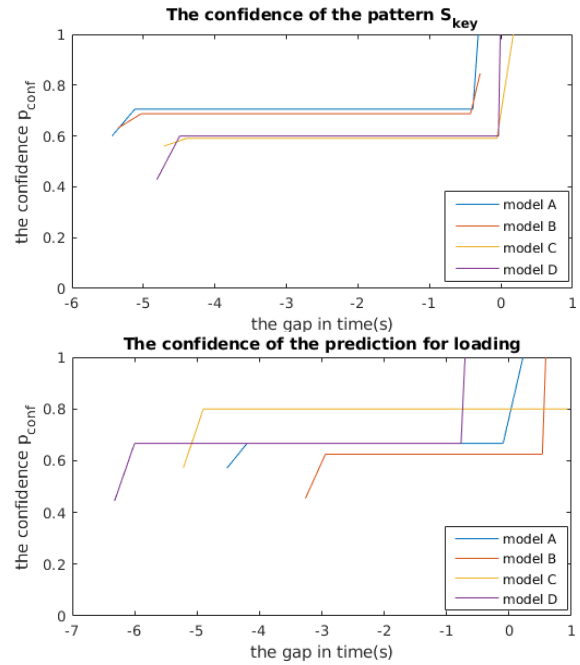

Fig. 9. The confidence p_{conf} and the gap in time from the completion (up : from training data, down : from evaluation data)

This prediction is useful for improving earthmoving work. The first application is the improving the safety in earthmoving work. The time prediction can be used for notifying dump trucks or construction workers not to get close to the backhoe. The second application is the automatic monitoring of the progress. The prediction with high confidence over 95% helps to estimate the amount of sediment which was loaded by the backhoe. The third application is improving the running efficiency of dump trucks through the operation management. Dump trucks sometimes need to wait on the way to avoid congestion around the loading spot. Our method could automate the management by informing of the instant when the dump truck moves to the loading spot.

The other relative methods are difficult to use the prediction of the instant at which the dump truck should move for the loading spot, because these predictions with the low precision induce the dangerous movement by the dump truck with a wrong moment. The results of the pattern recognition using the LSTM and the threshold-based method show the simple classification using labels of some processes or with a specific motion is difficult to use the prediction, because these methods could not extract the unique feature in the process P_2 which is different from in P_1. The regression prediction of the time gap via the LSTM found also difficult. The relation between the time gap until the backhoe finishes P_2 and the backhoe motion is unstable dependently on the speed of P_2. The condition of work environment like the relative position between sediment and the dump track changes this speed. Therefore, This unstable relation should influence on the regression prediction using the backhoe motion negatively. However, the proposed approach based on the BP-HMM could predict overcoming these difficulty. Using multiple shared features was effective to extract a

unique feature which is useful for the prediction. In addition, the proposed method is not influenced by the speed of the work, because not using the time gap but the transition between the primitive motions.

The primitive motions $\boldsymbol{Q}$ could be extracted from the motion data that was generated during earthmoving with the wood chips and the crushed rocks and were combined. During scooping motions, the intensity of the frequency component of the Y-axis acceleration for the arm was very different when during to sediment loading (Fig. 10), which resulted from the hardness of sediment. However, this did not strongly influence the prediction because the pattern $\boldsymbol{S}_{key}$ constituted of the primitive motions q_2, q_3 which appeared after scooping. Therefore, our prediction was considered to be robust for the given type of sediment.

In the future, we would like to measure the motion of the backhoe using several different operators and attempt to build a model that can be used for different operators. Our next challenge is to confirm that the proposed approach can be used to model backhoe motions by different operators and analyze the key motions. In this paper, three key motions, which can be represented using simple features, were extracted for one operator. If we can extract similar key motions, we can build a general model that can be applied to different operators. In addition, our proposed BP-HMM based method will identify key motions from iterative data that contain the primitive motions. in different order. This might be used for prediction of timing in the other applications if the iterative data exhibits similar characteristics.

VIII. CONCLUSIONS

In this paper, we developed a motion model for a backhoe operated by a specific operator using time series data with the BP-HMM and predicted the completion of preparation for a loading sequence using a sequence of primitive motions.

At best, our method could predict the completion of preparation for loading with a probability of 67% and 100% at approximately 6 and 0.71 seconds before.

The proposed technique can predict the instant at which the dump truck is required to move into the loading spot and help reduce the idle time that the backhoe expends waiting until the dump truck arrives at the loading spot. Thus, the proposed method can lead to reduced construction times.

ACKNOWLEDGMENT

This research was partially supported by NEDO's commissioned research project (No. 18065741)

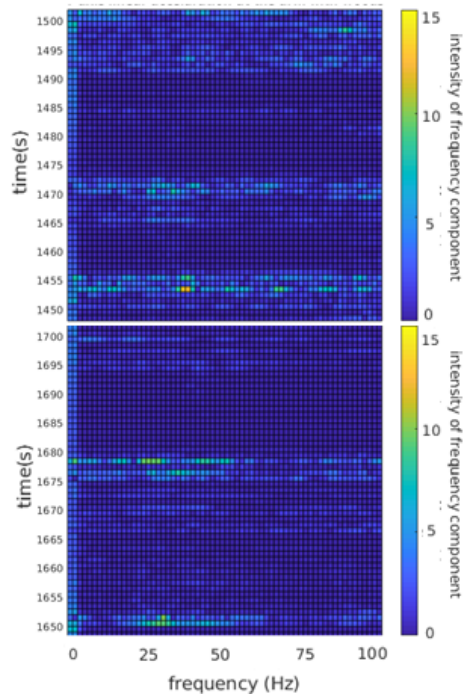

Fig. 10. Intensity of the frequency component of the Y-axis acceleration of the arm in scooping with wood chips (up) and crushed rocks (down)

REFERENCES

[1] T. Komatsu, Y. Konno, S. Kiribayashi, "Autonomous Driving of Six-Wheeled Dump Truck with Retrofitted Robot," 2019 12TH CONFERENCE ON FIELD AND SERVICE ROBOTICS (FSR), Tokyo, 2019.

[2] D. Jud, G. Hottiger, P. Leemann, and M. Hutter, "Planning and Control for Autonomous Excavation," in IEEE Robotics and Automation Letters, vol. 2, no. 4, pp. 2151-2158, Oct. 2017.

[3] J. Jongluxmanee and M. Yamakita, "Improved Soil Shape on the Dump Truck for Soil Loading Operation of Excavator," 2019 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), Hong Kong, China, 2019, pp. 845-850.

[4] T. Takanobu, K. Shinohara, and H. Yoshinada. "Research on effective teleoperation of construction machinery fusing manual and automatic operation." ROBOMECH Journal 4.1 (2017).

[5] E. Fox, M. Hughes, E. Sudderth, and M. Jordan. "Effective Split-Merge Monte Carlo Methods for Nonparametric Models of Sequential Data," 8(3):1281–1313, 2014.

[6] J. Kim, S. Chi, "Action recognition of earthmoving excavators based on sequential pattern analysis of visual features and operation cycles,"Automation in Construction, Volume 104, 2019, Pages 255-264.

[7] Chieh-Feng Cheng, A. Rashidi, Mark A. Davenport, David V. Anderson, "Activity analysis of construction equipment using audio signals and support vector machines," Automation in Construction, Volume 81, 2017, Pages 240-253, ISSN 0926-5805

[8] Z. Jin, P. R. Pagilla, H. Maske and G. Chowdhary, "Task Learning, Intent Prediction, and Adaptive Blended Shared Control With Application to Excavators," in IEEE Transactions on Control Systems Technology, doi: 10.1109/TCST.2019.2959536.

[9] A. Gupta and B. Dhingra, "Stock market prediction using Hidden Markov Models," 2012 Students Conference on Engineering and Systems, Allahabad, Uttar Pradesh, 2012, pp. 1-4.

[10] S. Renals, N. Morgan, H. Bourlard, M. Cohen, and H. Franco, "Connectionist probability estimators in HMM speech recognition," in IEEE Transactions on Speech and Audio Processing, vol. 2, no. 1, pp. 161-174, Jan. 1994.

[11] A. Asahara, K. Maruyama, A. Sato, and K. Seto. 2011. "Pedestrian-movement prediction based on mixed Markov-chain model." In Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (GIS '11). Association for Computing Machinery, New York, NY, USA, 25–33.

[12] Y. Tanaka, J. Kinugawa, Y. Sugahara and K. Kosuge, "Motion planning with worker's trajectory prediction for assembly task partner robot," 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, Vilamoura, 2012, pp. 1525-1532.

[13] R. Lioutikov, G. Neumann, G. Maeda, and J. Peters. 2017. Learning movement primitive libraries through probabilistic segmentation. Int. J. Rob. Res. 36, 8 (7 2017), 879–894.

[14] N. Mizuno, K. Ohno, R. Hamada, "Enhanced path smoothing based on conjugate gradient descent for firefighting robots in petrochemical complexes," 2019 Advanced Robotics, Vol33, 2019,

[15] K. Yamada, K. Ohno, N. Miyamoto, "Attachable Sensor Boxes to Visualize Backhoe Motion," 2020 IEEE/SICE International Symposium on System Integration (SII), Honolulu, 2020, pp. 1-6.