

Human Skill Capture: A Hidden Markov Model of Force and Torque Data in Peg-in-a-Hole Assembly Process

Y.C. Zhao*, A. Al-Yacoub, Y.M. Goh, L. Justham, N. Lohse, M.R. Jackson

*EPSRC Centre for Innovative Manufacturing in Intelligent Automation, Loughborough University, UK
y.zhao@lboro.ac.uk; a.al-yacoub@lboro.ac.uk; y.goh@lboro.ac.uk; l.justham@lboro.ac.uk; n.lohse@lboro.ac.uk; m.r.jackson@lboro.ac.uk

***Abstract**— A new model has been constructed to generalise the force and torque information during a manual peg-in-a-hole (PiH) assembly process. The paper uses Hidden Markov Model analysis to interpret the state topology (transition probability) and observations (force/torque signal) in the manipulation task. The task can be recognised as several discrete states that reflect the intrinsic nature of the process. Since the whole manipulation process happens so fast, even the operator themselves cannot articulate the exact states. Those are tacit skills which are difficult to extract using human factors methodologies. In order to programme a robot to complete tasks at skill level, numerical representation of the sub-goals are necessary. Therefore, those recognised ‘hidden’ states become valuable when a detail explanation of the task is needed and when a robot controller needs to change its behaviour in different states. The Gaussian Mixture model (GMM) is used as the initial guess of observations distribution. Then a Hidden Markov Model is used to encode the state (sub-goal) topology and observation density associated with those sub-goals. The Viterbi algorithm is then applied for the model-based analysis of the force and torque signal and the classification into sub-goals. The Baum-Welch algorithm is used for training and to estimate the most likely model parameters. In addition to generic states recognition, the proposed method also enhances our understanding of the skill based performances in manual tasks.

I. INTRODUCTION

Industrial robots are delivering more and more manipulation services in manufacturing. It is common to decompose the task for a robot into sub-goals where each goal is supported by an executable action module [1]. However, when the task is complex, it is difficult to programme a robot to fulfill all the requirements because even a relatively simple sub-task such as a Peg-in-hole (PiH) insertion contains uncertainties, e.g. clearance, initial grasping position and insertion path. Generally, manipulation tasks contain gross motion states and fine motion states. For the design of a robotic solution, a gross motion such as approaching a workpiece can be controlled using vision feedback [2] whilst a fine motion needs to rely on force and torque (F/T) feedback. Some intermediate passive solutions such as using a compliant

gripper have been proposed [3], but they cannot sufficiently address the uncertainties in very complex situations. F/T feedback, on the other hand, introduces noise and uncertainties from the sensor signal. Although humans can adapt to uncertainties easily, most of the time, the operators cannot easily articulate the skills they used to perform a task. We call these performances skill based performances. The reason humans are good at performing the PiH insertion task is because we have strong haptic feedback and uncertainties are compensated with both active and passive compliance [4]. Even though the automation solution may not fully imitate human motion since some of them are not necessary, it would be useful if the skill based performance from a human could be firstly interpreted by model and secondly transferred to a robot in a numerical way.

This paper is focusing on achieving the first goal. A general methodology is introduced and used to model the uncertainties in the sensor data captured from observing human operators perform a task.

Recently, robot learning from demonstration is gaining interest [5] as a framework to transfer skills from human to robot. It introduces probability encoding approaches such as the Gaussian Mixture Model (GMM) and the Hidden Markov Model (HMM) for modelling observations and state transition uncertainties. The advantages of using these probabilistic approaches are firstly that the results (transition probabilities and observation distributions) are easy to store and secondly that a robot can easily understand the underlying states by querying the model parameters. Gaussian Mixture Regression (GMR) together with GMM and HMM are extensively used for trajectory generalisation and force-based skill analysis [6, 7]. In general, the above probabilistic approaches have two aims: i) a generalised trajectory that a robot can implement; ii) a sequence of recognised states that a robot can understand.

In this paper, the state recognition capability of a probabilistic approach is investigated for the analysis of manual PiH assembly tasks. State recognition is important because it can help to interpret the skill based performance which cannot be directly articulated even by the operator themselves into meaningful states (sub-goals). Also it provides useful state transition information to the robot controller so that the

* The authors acknowledge support from the EPSRC Centre for Innovative Manufacturing in Intelligent Automation, in undertaking this research work under grant reference number EP/IO33467/1.

corresponding control strategies can change accordingly [8].

In general, the state recognition for a PiH process can be roughly classified into the following major states: approaching, insertion, and extraction. Since it is not possible to see inside the hole, vision cannot be considered as the primary source of input (also it is not accurate enough for fine motion). F/T sensors are used and installed near the contact area instead.

In the literature, state recognition is mainly achieved through either an analytical approach or a learning based approach. Analytical approaches are based on thorough analysis of the geometrical and dynamical constraints and they are limited by its computation speed since they require a large number of parameters to be measured and calculated simultaneously [9]. Learning based approaches are an alternative.

Various learning based approaches have been proposed in the literature. For instance, a HMM based recognition has been used for force and torque information analysis during tele-manipulations [10]; however, the tight clearance situation has not been addressed in this paper because the states they were interested in are more generic. Another research on developing HMM for states recognition explored the use of a virtual environment [11]. The accuracy of this approach is limited by the accuracy of the virtual environment used for state exploration. Jasim and Plapper [12] successfully implemented Expectation Maximization (EM) and Gaussian Mixture Model (GMM) to classify the F/T information into contact states. But the possibility to use feature transformation to reduce the dimensionality of the F/T signal was not addressed. It is not always required to reduce data dimensions but it will provide more compact training data sets and reduce the required training time. A Support Vector Machine (SVM) was used for contact state (CS) recognition during a PiH assembly [13]. The feature used for learning was pre-defined and inspired by the quasi-static force insertion model by Whitney and Dowe [14]. It is always beneficial if a feature can be used to reduce the number of variables required for training; however, how these features can be generalised for more complex situations was not discussed in the paper.

HMM has also been used for state recognition in other applications. For instance, in [15], they proposed a method to recognise the states in valve opening. They used a symbolic representation of the F/T signal where the number of data is hugely reduced to a pre-defined number of segments. This method converges faster than using the numerical representation of the training input. However, there is a risk that important information is lost when the signal is segmented into symbols. Therefore, justifications of how many segments are essential for a specific application are needed. HMM is also used in speech recognition where each word is encoded into a sequence of states [16]. Therefore, depending on the information of interest HMM can be used as a generic state recognition approach by considering both discrete symbolic and continuous numerical inputs.

To summarise, even though analytical and learning based approaches have been applied on the robot in the PiH application, the human skills are less understood. In another word, there is a lack of focus on generating numerical representation of the hidden information in manipulation process. Also more explorations are needed on interpretation of the 'hidden' information in a skill based PiH manipulation where a tight clearance is considered. Therefore, this paper proposed a methodology to extract the human skills from PiH process and focused on interpretation of the underlying states. How robot will use those extracted skills is not the main focus of this paper, but it will be explained in the end. In this paper, the insertion state is investigated because the most valuable information for the F/T-based control is contained during this period of the signal. Therefore, the information from approaching and extraction phase has been excluded from this study. The focus is on recognising the sub-states which occur during the insertion phase. The tight clearance situation has been addressed specifically as this requires the intricate skill from an operator. In order to verify the findings, different subject has been involved in the experiments.

II. METHODOLOGY

A. Overview

The proposed methodology is shown in Fig.1. The F/T signal feeds into a pre-processing stage which contains filtering, normalization and Principle Components Analysis (PCA). Then, a state extraction pipeline is used for encoding and decoding the process. In this stage, a K-means algorithm is used for initialization of GMM. A GMM is then trained for the initial guess of observation densities for each state. A Bayesian Information Criterion (BIC) is used for model selection where the candidate number of components is determined. HMM then takes the output from GMM and further encodes the states into transition probabilities. Finally the Viterbi algorithm [17] is used for state recognition (decoding) with a given set of model parameters. After recognition, the state is interpreted and some of the parameters such as the transition matrix (a matrix that describes the probabilities a state transit to other states) can be refined after interpretation.

HMM has been chosen as a general encoding structure in this paper to avoid too many assumptions on spatial- temporal nature of the data set. Here, a fully connected continuous HMM, with full covariance matrix, describing the distributions of the output variables is considered. However, this leads to a model which requires the estimation of a large number of parameters. On the other hand, for demonstration purpose, it is desirable that the operator should not have to demonstrate the task more than a few times (5~10). This leads to more parameters needing to be estimated compared to the amount of training data.

In order to overcome this issue, the Baum-Welch (BW) algorithm [18], an HMM extension of Expectation-Maximization (EM) optimization algorithm, is used for parameter estimation purpose. However, the algorithm

cannot guarantee a global maximum and may become trapped in a local maximum of the likelihood function. Thus, the initializations have great influence on the model performance. Consequently, k-mean algorithm has to run multiple times to guarantee a good convergence.

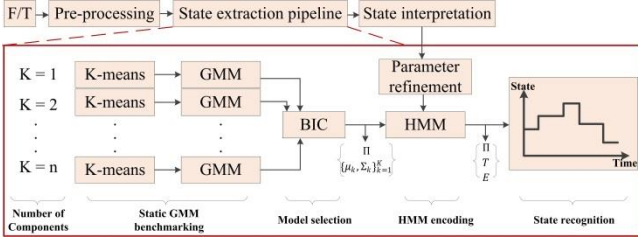


Figure 1. An overview of methodology.

B. Principle Component Analysis (PCA)

PCA [19] is a data dimensionality reduction method. Considering an M dimensional dataset $O = \{O_1, O_2, \dots, O_N\}$, where N is the number of samples, PCA will project all the observations based on their Eigen vectors and Eigen values to full or lower dimensions. The equation $Y = A \times X$ shows the same idea, where Y is the observations and X are the latent states. A is a mixing matrix with dimension $M \times N$ ($N \leq M$) that defines the rules of projection. In PCA, only the co-variance between the variables (6 channels of F/T data) are considered and re-ordered from the most important components to the least important component. As a consequence, the data dimension is reduced and the new dataset is re-ordered based on their importance. Here, PCA is a pre-process and re-representation of the data in terms of its importance.

C. Gaussian Mixture Model

A dataset of N data points of dimensionality D , $X = \{\vec{x}(t_1), \vec{x}(t_2), \dots, \vec{x}(t_N)\}$ with $\vec{x}(t_n) \in R^D$ is modelled by a multivariate Gaussian mixture of K -components.

$$p(\vec{x}(n)) = \sum_{k=1}^K \pi_k N(\vec{x}(t_n); \vec{\mu}_k, \Sigma_k) \quad (1)$$

Where $\pi_k \in \Pi$ is the prior probability on the Gaussian component k , and $N(\vec{x}(t_n); \vec{\mu}_k, \Sigma_k)$ is the D -dimensional Gaussian density of component k . $\vec{\mu}_k, \Sigma_k$ are the mean and covariance matrix of the multivariate Gaussian k . $\{\pi_k, \vec{\mu}_k, \Sigma_k\}$ are estimated using the EM algorithm.

D. Model selection

The optimal number of components K in a model is not known beforehand. A method is needed that evaluate the trade-off between optimising the model's likelihood (a metric of how well the fitting is) and minimizing the number of parameters to estimate. Even though this number can be learned heuristically, a formalized approach is preferred.

In order to select the optimal number of components K , a Bayesian Information Criterion (BIC) [19] is used after GMM in benchmarking stage:

$$S_{BIC} = -L + \frac{n}{2} \log(N) \quad (2)$$

Where L is the log-likelihood of the model, n is the number of free parameters required for a mixture of K components with a full covariance matrix, i.e. $n = (K - 1) + K \left(D + \frac{1}{2} D(D + 1) \right)$. N is the number of D -dimension data points. The first term of the equation measures how well the model fits the data, while the second term has two parts: the number of parameters to estimate transition matrix and the observation densities.

E. Hidden Markov Model

A Hidden Markov Model [20] not only uses a mixture of multivariate Gaussians to describe the distributions of the observation, the temporal variations were also encapsulated in the transition probabilities.

Let $\{\Pi T E\}$ be, the initial guess of state distribution, the transition probabilities between the states or components and the multivariate data distribution, respectively. $E = \{\vec{\mu}_k, \Sigma_k\}_{k=1}^K$ are returned from the GMM and directly used for initialising the HMM. The prior state distribution Π can also be suggested by the GMM as well. Therefore, the HMM only needs to estimate state transition probabilities T and refine parameter Π and E . The Baum-Welch algorithm is used to estimate those parameters [17].

III. EXPERIMENT SETUP

The F/T data acquisition device is 6-axis ATI force torque sensor and it is sampled at 200 Hz. The experiment setup is shown in Fig.2 to record the data in tight clearance manipulation. The human operator is standing in front of the test rig performing a PiH assembly. The F/T sensor is fixed installed on the plate. The nominal diameter of the hole is 16.2 mm. The peg diameter used in the experiments is 15.90mm. The operator starts with holding the peg and performs the task in the following phases: approaching, insertion, releasing and waiting. The whole process takes less than 25 seconds and was repeated 11 times for one experiment data set. Therefore, there are in total 11 demonstrations from the operator. The process was repeated on another operator for comparison purpose

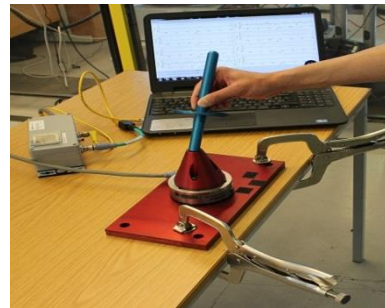


Figure 2. Experiment setup. An operator is holding a peg and demonstrating insertion process. The Force and Torque sensor is fixed statically on the fixture.

IV. RESULT ANALYSIS

In this section, the results from implementing the proposed framework are discussed by firstly fixing the subject (one subject) and peg clearance. The number of states is determined based on model selection. PCA was

used to reduce the dimension of the signals by retaining at least 98% of the information. There is no constraint on the state topology. Two people were involved in the experiment. The first operator was more familiar with the process than the second operator. The state recognition results from these two subjects were shown for comparison purpose.

A. Model selection result

The model selection results are shown in Fig 3. Recall, the functionality of BIC is to calculate the logarithmic likelihood over the observations while penalising the complexity of the parameters for the model estimation. Assuming there are K components and M dimension observable variables in total. Therefore, the total number of GMM parameters is $(K - 1) + K(M + M(M + 1)/2)$. From the results in Fig.3, it indicated that starting from 8 states the BIC scores did not improve that much. On the other hand, too many states might over complicate the problem; therefore, the number of components (states) is fixed at 8.

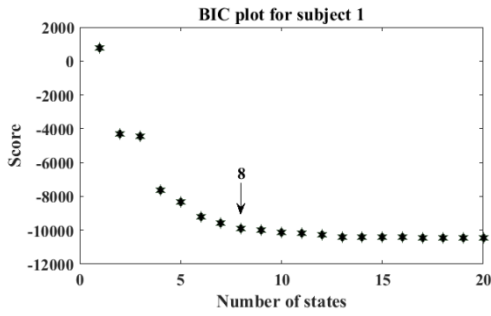


Figure 3. BIC selection for the number of components

B. Result discussion on a fixed number of states

After fixing the number of states, the state recognition results are discussed and interpreted by a human operator. The results from three demonstrations are plotted in Fig.4. The x axis is the discretised time line of data sampling points. Y axis has multiple meanings: i) it assigns numbers to states, ii) it represents scaled observations in the range $[-10, 10]$. The different curves from top to bottom indicate the filtered F_x , F_y and F_z signals. T_x , T_y and T_z signals are not plotted here for the purpose of visualization. The thick segments indicate the classified states and each of them was assigned a state labels from S1 to S8.

These three plots give a typical representation of the 11 demonstrations. Since the main purpose of this paper was to extract the tacit skills from manual operation, a human operator was asked to evaluate the state recognition results. To begin with, it is hard to tell exactly what happened in each state especially as 8 states have been chosen as the model. This is because by nature, tacit skills are hard to articulate. However, looking closely at the data and classification results, states can be largely grouped into three phases which are highlighted in different colour in Fig.4 and concluded in Table I. For instance, during the first demonstration, the first phase contains state 1, 2, 3, 4; the second phase contains state 5 and 6; the third phase contains state 7 and 8. The same

result happened in every demonstration which meant that the operator was doing the task in a highly repetitive manner but with variations in each state.

TABLE I. STATES RECOGNITION RESULTS IN 11 DEMONSTRATIONS

Demonstrations	Phase 1	Phase 2	Phase 3
1,2,3,4,5,6,7,8,9,10,11	1,2,3,4	5,6	7,8

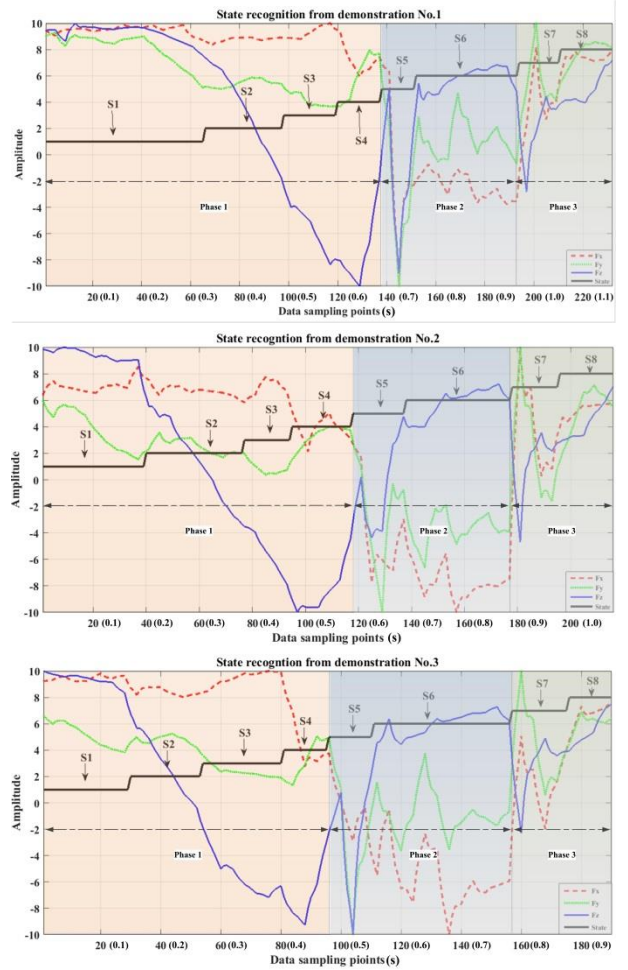


Figure 4. State recognition results from F/T sensor. The data samples are coming along x axis in time sequence and normalised for visualisation purpose.

C. Phase interpretation and summary of result

The three phases were identified from an operator. They can be interpreted and defined in the following: In phase one, the operator was attempting to insert the peg from the hole chamfer. The dominant force was the F_z value since the operator was trying to push firmly. Then in phase two, the peg was already inside the hole, but the lateral angles still needed to be adjusted (wobbling movement) so that the peg was well aligned inside the hole. This phase is due to the small wobbling movement. A final insertion attempt was made after this to make sure that the peg was fully inserted into the hole. The absolute peak value of this phase is smaller than the absolute peak value in phase 1 which contains the first insertion attempt. The overall process was very fast (within 2s);

therefore it was really difficult for the operator to articulate what happened within 2s.

To summarise all the 11 demonstrations, the results indicated that each phase could be clearly defined by a certain fixed state. This result could be used as initial indication of the skill based performance. Further indications cannot be articulated by the operator. However, the states are already encoded in the probabilistic. Therefore, states such as 1,2,3,4 inside phase 1 are ‘hidden’ in the skill but it can be ‘articulated’ by the HMM model. Also, as the states in each demonstration were recognised, the decoded states could be used to align the sequences so that the temporal distortions of the sequences were removed in this way. This is useful because the unwanted samples can be trimmed and all the sequences are re-represented in the same length.

D. Result discussion on a different subject

The results above were obtained from operator 1 who was more experienced with the process. State recognition results from operator 2 with less previous experience are introduced in this section. 3 out of 11 demonstration results are shown in Fig.5. We started from data normalisation and PCA reduction where 98 % percent of information was reserved. The number of components returned by BIC was 8. The results indicate that the operator takes slightly longer to insert the peg and he adopts a different PiH manipulation pattern. But the three phases remained as shown in Fig.5. However, by looking at the recognised states in Fig.5, the individual states do not look the same. This indicates that operator 2 did PiH in different manner. Also, state 1 was recognised in more than one phase in the all three plots. But the same state accounted for initial insertion stage and release stage respectively. Intrinsically, HMM tries to recognise the state which depends on the Multi-variant Gaussian distributions from the observations. The observations in those stages were close to each other so that they confused the algorithm. That is the reason why human interpretation is needed to classify the states returned by the HMM model. For instance, the first state 1 in the top Fig.5 was clearly different from the state 1 at the end. This is because they represented the initial chamfer crossing state and the force releasing state respectively. The same phenomenon can be observed in the bottom plot where not only state 1, but also states 4, 5 and 6 are repeated. This indicates that the operator 2 was not doing the manipulation in a way as consistent as the more experienced operator 1.

Fig.6 is the Gaussian plot where each Gaussian was drawn as an ellipsoid against the first 2 principle components. The arrow indicated the direction of the state transitions. The left one is from subject 1 and the right one is from subject 2. By comparing the two plots, the Gaussians are more distinctive from each other, thus the expert operator from experiment 1 have more consistent way of doing the PiH manipulation than operator 2.

From the above results, the skills from novice and expert operators are represented by probabilistic models,

however, how to transfer the model to the robot is unclear. This can be done by doing statistical inference from the existing HMMs. Since each state is responsible of encoding a specific period of PiH process in terms of its variations, the robot can query the probabilistic model and the uncertainties that a robot might encounter in this specific state can be inferred. The problem remained is designing specific control strategy of dealing with uncertainties in each state.

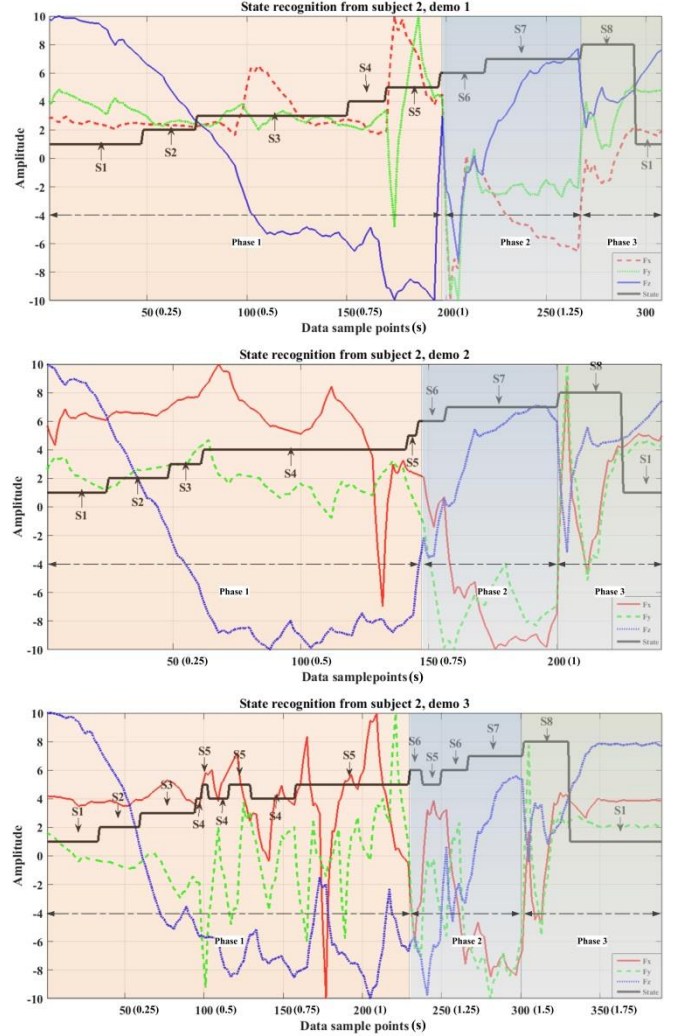


Figure 5. State recognition results from 3 demonstrations by operator 2.

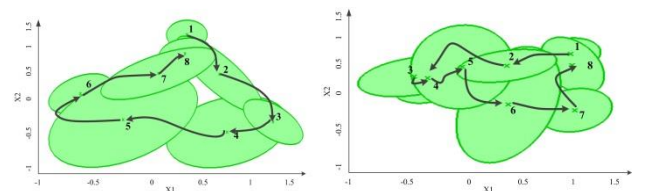


Figure 6. Gaussian plots for subject 1 (left) and subject 2 (right).

V. CONCLUSION

In this paper, a Hidden Markov Model has been used to analyse the hidden states in the PiH assembly process. The methodology started with data pre-processing where a more compact representation of the data was returned by PCA. Then the GMM/HMM encoding process was used for state recognition. Finally, human interpretations

were introduced to classify the recognised states. The discussions focused on interpretations of the ‘hidden’ states embedded in the skill based performance (the PiH insertion task). It was shown that it is possible to further interpret the insertion task into 3 major phases which are defined by their corresponding sub-states. This gives a detail explanation of the intricate skills a robot should interpret and implement. The results from different operators indicated that operators might carry out the task following different patterns but still go through common phases. This indicates that the methodology can be implemented for encoding processes from different subject. The models learned from different operators can be saved in a knowledge base and they can be queried when required.

From the results, the definition of the phases from observing operator 2 was less clear due to the repeated states. This might appear because the observations were encoded by Gaussians and two close features will be recognised as the same state. This is mainly due to operator 2 being less experienced than operator 1. This is creating a higher chance some states might be repeated during the demonstration. However, if the operator is an expert and the repeated states were appeared in the most of the demonstrations, this became a necessary state transition of this specific operator. It is recommended that the demonstrator should demonstrate the process with enough variations in each state, and the state transitions should be similar. The skilled demonstrator needs to be confident and familiar with the process. The trained model parameters from the skilled demonstrator should return fewer repeated states and can be used to evaluate the performance from a novice demonstrator. The proposed work used one peg clearance and limited by two subjects in the experiments. Therefore, the future work needs to explore the generalisation capability when different peg clearance is used and more subjects are involved. Also, the recognised states in the fine motion level need further explanation by the researchers from human factors. They will evaluate whether the ‘hidden’ states are useful for them to interpret the manual process. This paper can be regarded as a methodology for detail skill extraction on skill based manipulation, the knowledge from different person can be encoded and saved as candidate model that help to interpret the manual process. In order to use the model in the robot so that it can take various inputs, different control strategies need to be designed. Also, the performance of these should be evaluated against with the existing control strategies such as impedance control.

REFERENCES

- [1] M.Tenorth, D.Jain, and M.Beetz, “Knowledge Processing for Cognitive Robots,” *Kunstl Intell*, vol.24 (3), pp.233-240, 2010.
- [2] R.Kouskouridas, A.Amanatiadis and A. Gasteratos, “Guiding a robotic gripper by visual feedback for object manipulation tasks,” in Proc. 13th Int. Conf. Mechatronics, pp. 433-438, April 2011.
- [3] S.Yun, “Compliant manipulation for peg-in-hole: Is passive compliance a key to learn to the contact motion?” in Proc. 2008 Int. Conf. Robotics and Automation. (ICRA), pp.1647-1652, 2008.
- [4] K.T.Park, Y.J.Shin, C.H.Park, Y.S.Mo, D.C.Jeong and Y.S.Shin, “Robot application for assembly process of engine part,” in Int.Conf.Control, Automation and Systems (ICCAS), pp.2107-2110, Oct 2008.
- [5] S.Caliron, F.Guenter and A.Billard, “On learning the statical representation of a task and generalizing it to various contexts,” in Int.Conf.Robotics and Automation (ICRA), pp. 2978-2983, Mey 2006.
- [6] P.Kormushev, S.Caliron and D.G.Caldwell, “Imitation learning of positional and force skills demonstrated via kinesthetic teaching and haptic input,” in *Advanced Robotics*, vol.25(5), pp.581-603, 2011.
- [7] S.Caliron, P.Evrard, E.Gribovskaya, A.Billard and A.Kheddar, “Learning collaborative manipulation tasks by demonstration using a haptic interface,” in Int.Conf.Advanced Robotics (ICAR), pp.1-6, June 2009.
- [8] T. Lefebvre, J. Xiao, H. Bruyninckx and G. De Gersem, “Active compliant motion: a survey,” *Advanced Robotics*, vol.19(5), pp. 479-499, 2005.
- [9] Z.Jakovljevic, P.Petrvic and J.Hodolic, “Contact states recognition in robotic part mating based on support vector machines,” in *Int.J.Adv.Manuf.Tech.*,vol.59(1-4), pp.377-395, 2012.
- [10] B.Hannaford and P.Lee, “Hidden markov model analysis of force/torque information in telemanipulation,” in *Experimental Robotics I*(V.Hayward and O.Khatib,eds), vol.139 of *lecture Notes in Control and Information Sciences*, Springer Berlin Heidelberg, pp.145-149, 1990.
- [11] S.Dong and F.Naghdy, ”Application of hidden markov model to acquisition of manipulation skills from haptic rendered virtual environment,” in *Robot.Cim.Int.Manuf*, vol.23 (3), pp.351-360, 2007.
- [12] I.Jasim and P.Plapper, “Contact-state monitoring of force-guided robotic assembly tasks using expectation maximization-based Gaussian mixtures models,” in *Int.J.Adv.Manuf*, vol.73 (5-8), pp.623-633, 2014.
- [13] Z. Jakovljevic, P. Petrovic, and J. Hodolic, “Contact states recognition in robotic part mating based on support vector machines,” in *Int. J. Adv. Manuf. Tech*, vol. 59(1-4), pp. 377-395, 2012.
- [14] D. Whitney and D. Dowe, *Mechanical Assembly: Their design manufacture and the role in product development*. Oxford University Press, 2004.
- [15] N.Jamali, P.Kormushev and D.Caldwell, “Robot-object contact perception using symbolic temporal pattern learning,” in Int.Conf.Robotics and Automation (ICRA), pp.6542-6548, 2004.
- [16] K.S. Lee, “EMG-based speech recognition using hidden markov models with global control variables,” *IEEE trans. Biomed. Eng.*, vol. 55(3), pp. 930 – 940, March 2008.
- [17] L.R. Rabiner, “A tutorial on hidden Markov models and selected applications in speech recognition,” in Proc. IEEE., vol. 77(2), pp. 257-286, 1989.
- [18] L.R. Welch, “Hidden Markov models and the Baum-Welch algorithm,” *IEEE Information Theory Society Newsletter.*, vol. 53(4), pp.10-13, 2003.
- [19] W.J. Krzanowski, *Principles of multivariate analysis*. Oxford University Press, 1988.
- [20] D. Posada and R.B. Thomas, “Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests,” *Syst. Biol.*, vol. 53(5), pp. 793-808, 2004.