

## Research Article

# Self-Generated Dataset for Category and Pose Estimation of Deformable Object

Yew Cheong Hou, Khairul Salleh Mohamed Sahari\*

*Department of Mechanical Engineering, Universiti Tenaga Nasional, Selangor, Malaysia***ARTICLE INFO***Article History*

Received 24 July 2018

Accepted 16 November 2018

*Keywords*Deformable object  
robotic manipulation  
computer vision  
particle-based model**ABSTRACT**

This work considers the problem of garment handling by a general household robot that focuses on the task of classification and pose estimation of a hanging garment in unfolding procedure. Classification and pose estimation of deformable objects such as garment are considered a challenging problem in autonomous robotic manipulation because these objects are in different sizes and can be deformed into different poses when manipulating them. Hence, we propose a self-generated synthetic dataset for classifying the category and estimating the pose of garment using a single manipulator. We present an approach to this problem by first constructing a garment mesh model into a piece of garment that crudely spread-out on the flat platform using particle-based modeling and then the parameters such as landmarks and robotic grasping points can be estimated from the garment mesh model. Later, the spread-out garment is picked up by a single robotic manipulator and the 2D garment mesh model is simulated in 3D virtual environment. A dataset of hanging garment can be generated by capturing the depth images of real garment at the robotic platform and also the images of garment mesh model from offline simulation respectively. The synthetic dataset collected from simulation shown the approach performed well and applicable on a different of similar garment. Thus, the category and pose recognition of the garment can be further developed.

© 2019 The Authors. Published by Atlantis Press SARL.

This is an open access article distributed under the CC BY-NC 4.0 license (<http://creativecommons.org/licenses/by-nc/4.0/>).**1. INTRODUCTION**

Handling deformable object by robot is a challenging problem in robotic research. To teach a general household robot to handle these deformable objects is difficult compared with rigid object due to the variety of sizes, textures and poses for these objects. [Figure 1](#) shows a pipeline for the autonomous garment handling by a robot. To manipulate it from initial state to final desired state successfully, the category and pose of the garment should be recognized by the robotic system. The computer vision technique integrated with the robust machine learning algorithm through garment perception should be acquired by a robot to manipulate these objects dexterously by recognized their deformed configuration. In this research, we are focusing the task of classify the category and recognizing the pose of hanging garment grasped by single robotic manipulator. The main idea of this work is to generate a dataset of hanging garment that combine the real and synthetic images from the real garment and deformable model respectively. The deformation poses of the garment can be reduced by hanging the garment under gravity. A dataset of hanging real garment by robotic manipulator can be obtained by capturing the depth images. For deformable object modeling, the particle-based model is constructed to the garment in an image and then used into the simulation to acquire the synthetic dataset. This work is extended to include the 2D garment mesh model that able to represent the real garment that not

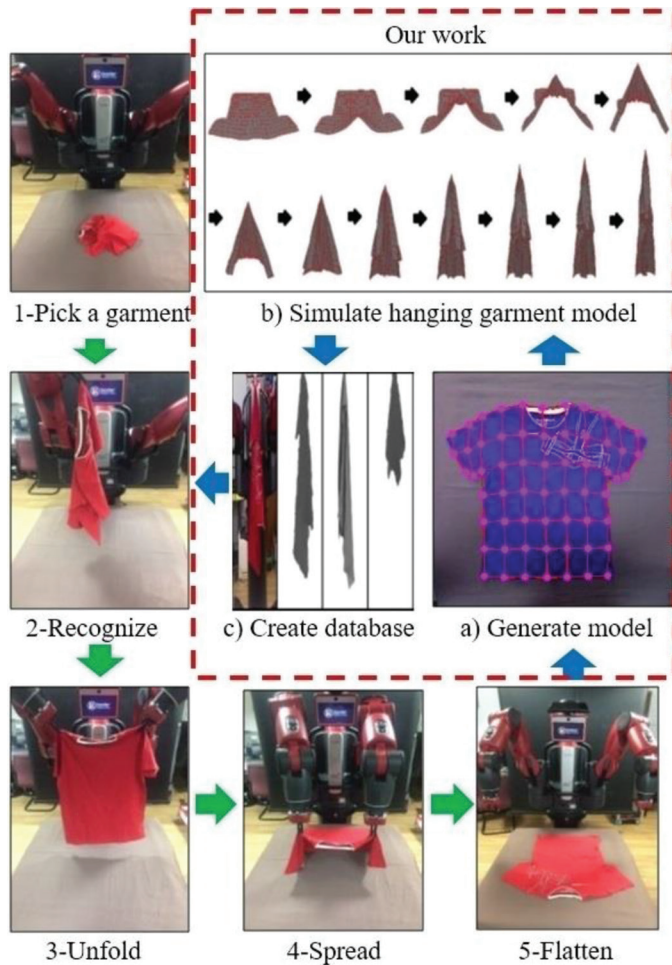
found in the database. The dataset will take advantages of this algorithm to generate large number of data for learning purposes without using professional graphic software. The main contribution in these works are:

- The simple 2D garment mesh model is extracted from the real garment that crudely spread-out on the flat platform. The multiple poses of hanging garment model under gravity are simulated, and different viewpoints of the model are captured as an image.
- Different types and sizes of garment are simulated and the dataset is generated from the synthetic data of hanging garment mesh model integrated with the real data.

**2. RELATED WORK**

Modeling and simulation are commonly used to compute the deformation of the deformable objects and becomes essential in future autonomous robotic manipulation since the result produced from simulation is seeming real and applicable. In robotic manipulation, the research on autonomous laundry service for general household robot has gained attention and increase significantly. The end-to-end task of garment folding had been decomposed into several sub-tasks, the existing approaches had been proposed from different aspects to solve the particular problem to achieve its accuracy and robustness. Generally, the procedure of autonomous garment handling can be categorized into several research areas. The majority

\*Corresponding author. Email: [khairuls@uniten.edu.my](mailto:khairuls@uniten.edu.my)



**Figure 1** | The pipeline for garment handling in autonomous robotic manipulation.

contributions in the literature of garment handling mainly focus on the grasping skills, recognition, and motion planning strategy of unfolding and folding.

There has been previous work on the recognition and manipulation of deformable objects. A series of works had been proposed Kita et al. [1] in pose estimation for a hanging garment grasped by a single end-effector using single view image and stereo image. The hanging garment models are obtained from the physics simulation by a set of predefined grasping points to generate a deformable model database. The approach is observed the garment from different views and checking the consistency during pose recognition. Maitin-Shepard et al. [2] propose an approach for folding a pile of crumpled towels using PR2 robot. Later, Miller et al. [3] proposed a parametrized polygonal model to describe the garment perception in the folding process. The contour of a piece of garment had been fitted by the predefined polygonal model iteratively in order to identify its category and parameters from the polygonal model.

Doumanoglou et al. [4] proposed a dual-arm industrial robot to unfold different garments such as sweater and trouser. They focused on classification and pose recognition of hanging garment and then proceeded to unfold it into desired configuration. Stria et al. [5] proposed an approach using polygonal model incorporating relative angle and length of contour segments to fit the model

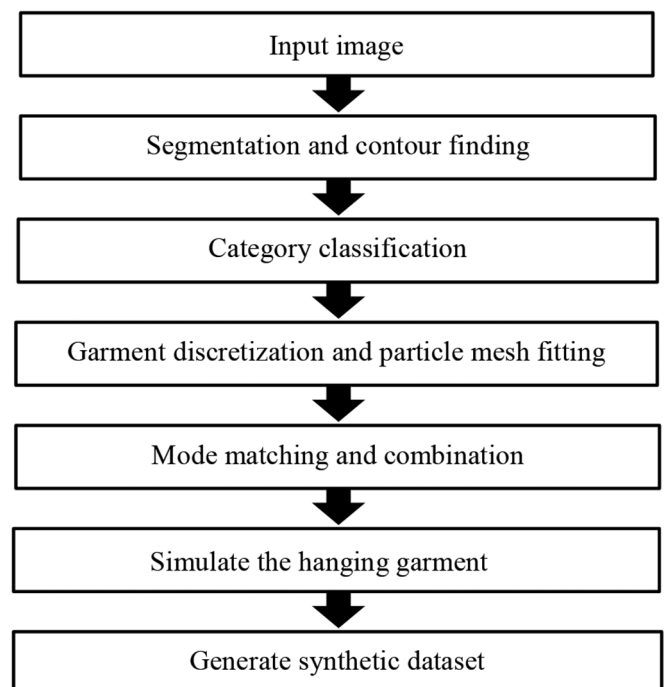
into the real garment. Mariolis et al. proposed an approach of recognize the pose for garment using convolutional neural network. Their model is trained by the combination of real and synthetic dataset in 2D garment model [6]. Li et al. presented a series of laundry works in recognition and manipulation of garment. Initially, the model-driven approach based on precomputed deformable model of hung garments has been proposed. A set of deformable models had been trained using the method of support vector machine to recognize the category and pose of garment [7]. They improve their recognition approach using KinectFusion technique to reconstruct a deformable 3D model [8].

### 3. PARTICLE-BASED POLYGONAL MODEL

We suppose a piece of garment that is crudely spread-out and uniformly placed on a flat platform. We perform a computer vision analysis to a single color image that captured from the top view from the garment. By applying the dynamic programming algorithm, the real garment is matched by the particle-based polygonal model and its parameters are extracted. The flowchart for the model generation is described as shown in Figure 2.

#### 3.1. Segmentation and Contour Finding

From the input image, the color of garment had been assumed sufficiently different compared with background color. Hence, the object in an image that represents garment can be acquired using GrabCut algorithm [9]. Contour finding is used to mainly extract the contour and centroid of the garment object in the binary image as shown in Figure 3 [10]. Additional parameters such as contour area, straight bounding rectangle and rotated rectangle with minimum



**Figure 2** | Flowchart for the particle-based polygonal model generation for synthetic dataset.

area can be obtained. The category of garment can be estimated based on these related information.

### 3.2. Category Classification

For simplicity, the garment is supposed uniformly placed on the platform by robot. The category of garment such as towel, shirt and trouser can be estimated using multiple lines detection and comparison to the contour as shown in Figure 4 [10]. The length of width and height at different layers can be measured, the category can then be estimated based on the relationship between each layer.

### 3.3. Garment Discretization and Particle Mesh Fitting

The garment can be discretized into separated parts such as main part and sleeves as shown in Figure 5. From the information obtained in the category classification, the width and height

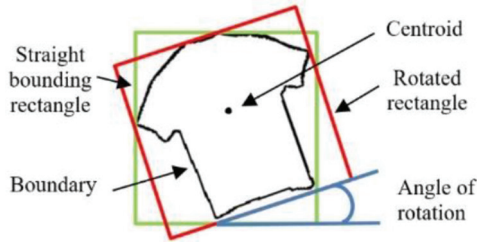


Figure 3 | Contour details of the garment in an image.

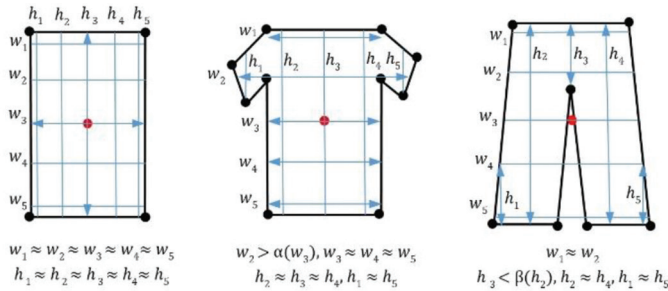


Figure 4 | Polygonal model for each category of garment. Measure the width and height at different layers by subtracting the outer points of contour either horizontally and vertically.

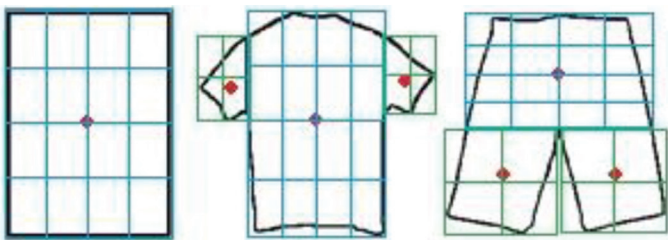


Figure 5 | Discretize the towel, shirt, trouser into main body and sleeves respectively. Particle mesh fitting to the rectangle for each category of garment.

of rectangle can be measured and the main part of the garment can be extracted and later the other parts such as sleeved can also be extracted using the contour finding discussed in Section 3.1. For each type of garment, they can be separated into several parts depending on their categories. Each separated part has parameters such as rectangle and centroid. Thus, rectangle can be transformed to a mesh lattice for each contour in the garment.

### 3.4. Model Matching and Combination

The particles at the boundary in particle mesh lattice are matched to the contour for each discretized part as shown in Figure 6 [11]. The particles in the internal structure of mesh lattice will realign to more organized and keep the distance between each other to prevent particles combine together. Lastly, the mesh lattice of main part and sleeve are combined to generate a completed particle-based polygonal model as shown in Figure 7.

### 3.5. Simulate the Hanging Garment

The particle-based modeling is applied into the dynamic programming and used to simulate the movement and deformation of the garment mesh model. The approach in the previous work is adopted and the mass-spring model is utilized in this work [12]. The hanging process of garment mesh model generated from a real garment is simulated as shown in Figure 1, a piece of crudely spread-out garment mesh model is placed flat on the

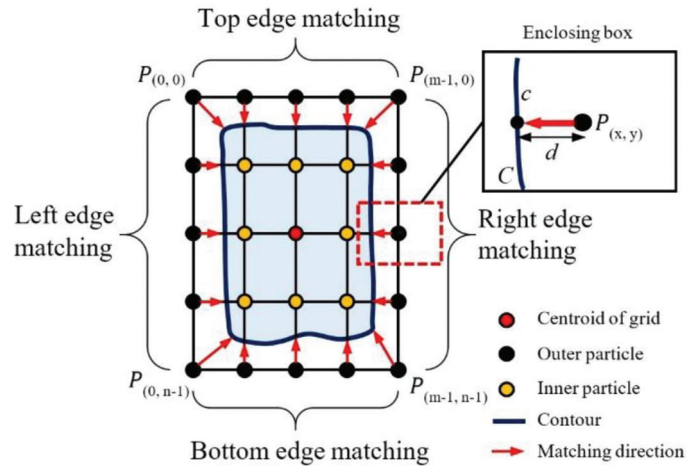


Figure 6 | The contour are matched by particle mesh lattice.

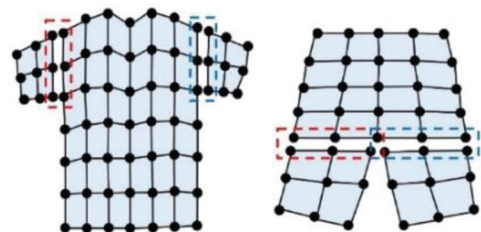


Figure 7 | Model combination between discretized parts of garment to form a completed particle-based polygonal model.

simulated environment. For each step, a specified particle from the garment mesh model will pick up by robot as a grasping point then moves it to the hanging position following the simulated trajectory.

### 3.6. Generate Synthetic Dataset

One of the hardest problems need to be solved in deep learning is to collect the large scale data into the desired format. Collecting the data for different categories of real garment are very time consuming particularly encompassing the different types, sizes and material properties with its deformation. Hence, a training set that gather relevant information of garment which can be used in visual recognition purposes is required. The features such as colors and textures of the garments are less significant in this work. The dynamic programming modeling is applied to the garment mesh model, the deformation poses of the model hanging by robotic manipulator is simulated. By applying different physical properties of the garment mesh model, the model look visually similar with depth images can be produced from the virtual simulation. An amount of images that simulate the deformation of real garment can be acquired. A series of images are captured for each pose of garment, the model is rotated vertically with  $360^\circ$  and the virtual camera captures the image of the rotated garment by increasing one rotating degree until full rotation is achieved.

## 4. PRELIMINARY RESULTS

### 4.1. Robotic Manipulation Setup

Robotic manipulations have been conducted using a Baxter robot with two Kinect depth sensor. The depth images of the hanging

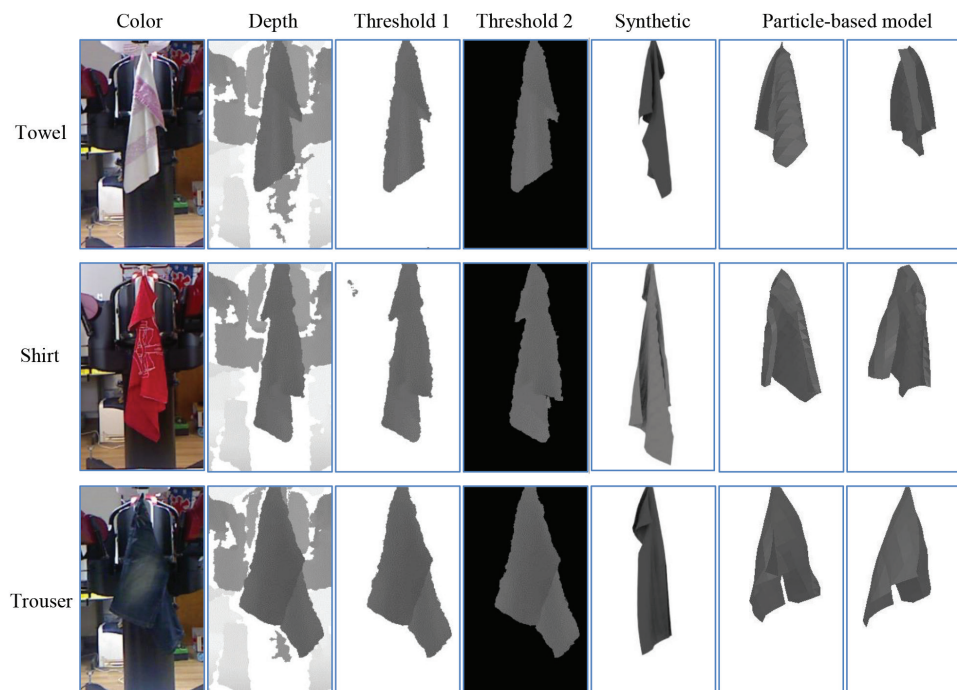
garments have been acquired by a Kinect depth sensor placed in front of the Baxter robot at a fixed height. The second Kinect sensor is placed at the head of Baxter robot which used to capture the images for a piece of crudely spread out garment on the flat platform. As shown in Figure 8, the garment mesh model extracted from the real garment, the selected particles are predefined and simulated in virtual simulation and the deformation poses of the garment model can be acquired.

### 4.2. Real and Synthetic Dataset

To collect the real garment dataset, a piece of the garment placed on the flat platform is hung by a Baxter robot and a Kinect depth sensor located in front of the Baxter robot. The hanging garment is then rotated  $360^\circ$  vertically and the color and depth images are captured. As shown in Figure 9, the depth images are thresholded to obtain the garment from the background. The  $200 \times 340$  pixels depth images that only consists of a piece of hanging garment are



**Figure 8** Sample of particle-based polygonal model is extracted for each category of garment. Selected particles defining the poses of the hanging garments. Orange markers define the poses on real garments from the particle-based model.



**Figure 9** Example images of color, depth, threshold 1 (white), threshold 2 (black), synthetic and particle-based model of hanging garment grasped by specified point.

extracted from  $640 \times 480$  original images. For category classification, three categories which are towels, shirts and trousers are selected in this current work. Total nine different real garments have been used as the samples, three real garments for each category. Each garment have been hung two or three times and about 360 depth images have been acquired for each grasping point. For pose recognition, 21 poses for shirts, 15 pose for trousers and nine poses for towels are selected by defined by the markers. Total around 48,600 depth images including towels, shirts and trousers are collected, the procedure of image acquisition is considered slow and time consuming.

For synthetic dataset, a large dataset of synthetic depth images are constructed using computer graphic software and dynamic programming algorithm respectively. Using Blender software, the total nine garment models which three models of towels, three models of trousers and three models shirts are constructed with different in size and shape. For simplify task in designing and model simulation, the garment models are constructed in 2D by assuming the front and back sides of shirt and trouser are not separated during grasped by robotic arm. In addition, the 2D garment model also can be generated using the proposed dynamic programming algorithm. The particle-based polygonal model can be generated from a crudely spread out garment placed on the flat platform. The towel models consists of 81 particles, the trouser models consists of 121 particles and the shirt model consists of 209 particles are constructed. Due to the symmetry axes of the garment and reduce the dense of the particles in the model, only selected particles in the garment model that matching with the markers in the real garment are simulated. The hanging garment models can be simulated by each of their particles. Different views of the hanging garment model can be acquired by rotating the model  $360^\circ$  in front of virtual camera that placed at fixed height as similar as possible with the setup of Kinect sensor for image acquisition in robotic platform. The approximately 97,200 virtual depth images have been generated. For synthetic dataset, the condition in the virtual environment is set as close as possible to the depth images in real dataset. Hence, the images acquired from the virtual simulation are comparable with real depth images, then allow to use as training dataset in the learning model.

## 5. CONCLUSION

In this work, an approach to generate the dataset of hanging garment using both real and synthetic images is proposed. The garment mesh model is extracted from the crudely spread-out garment using particle-based polygonal model, then the different poses of hanging garment can be simulated in the virtual environment. Experiment results demonstrate the proposed approach is applicable in dataset collection that need a large volume of data to train the learning model. Using this proposed approach, different types and sizes of garment mesh model can be extracted and simulated rather than design them in the graphics software. For future improvement, this proposed approach will be including more complex and different category of garment such as long-sleeved shirts and socks, and evaluate its performance once the datasets are extended. Thereafter, the further works such as recognition and unfolding a piece of garment can be considered in the garment handling process. In addition, the proposed

approach may also benefit to the pipeline of autonomous garment folding in robotic system which real garment can be represented by the particle-based polygonal model and the robotic grasping point can be defined.

## ACKNOWLEDGMENT

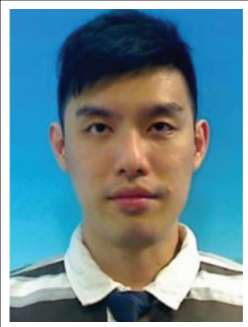
This work is supported by the Ministry of Higher Education, Malaysia through research grant: 2017/FRGS/V3500.

## REFERENCES

- [1] Y. Kita, T. Ueshiba, E.S. Neo, N. Kita, Clothes state recognition using 3D observed data, 2009 IEEE International Conference on Robotics and Automation, IEEE, Kobe, Japan, 2009, pp. 1220–1225.
- [2] J. Maitin-Shepard, M. Cusumano-Towner, J. Lei, P. Abbeel, Cloth grasp point detection based on multiple-view geometric cues with application to robotic towel folding, 2010 IEEE International Conference on Robotics and Automation, IEEE, Anchorage, AK, USA, 2010, pp. 2308–2315.
- [3] S. Miller, J. van Den Berg, M. Fritz, T. Darrell, K. Goldberg, P. Abbeel, A geometric approach to robotic laundry folding, *Int. J. Robot. Res.* 31 (2012), 249–267.
- [4] A. Doumanoglou, A. Kargakos, T.K. Kim, S. Malassiotis, Autonomous active recognition and unfolding of clothes using random decision forests and probabilistic planning, 2014 IEEE International Conference on Robotics and Automation, IEEE, Hong Kong, China, 2014, pp. 987–993.
- [5] J. Stria, D. Průša, V. Hlaváč, L. Wagner, V. Petřík, P. Krsek, et al. Garment perception and its folding using dual-arm robot, 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, Chicago, IL, USA, 2014, pp. 61–67.
- [6] I. Mariolis, G. Peleka, A. Kargakos, S. Malassiotis, Pose and category recognition of highly deformable objects using deep learning, 2015 International Conference on Advanced Robotics (ICAR), IEEE, Istanbul, Turkey, 2015, pp. 655–662.
- [7] Y. Li, C.F. Chen, P.K. Allen, Recognition of deformable object category and pose, 2014 IEEE International Conference on Robotics and Automation (ICRA), IEEE, Hong Kong, China, 2014, pp. 5558–5564.
- [8] Y. Li, Y. Wang, M. Case, S.F. Chang, P.K. Allen, Real-time pose estimation of deformable objects using a volumetric approach, 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, Chicago, IL, USA, 2014, pp. 1046–1052.
- [9] C. Rother, V. Kolmogorov, A. Blake, “GrabCut”: interactive foreground extraction using iterated graph cuts, *ACM Trans. Graphics* 23 (2004), 309–314.
- [10] C.H. Yew, K.S.M. Sahari, Real-time modeling and parameter approximation of dexterous garment folding by robot, *Artif. Life Robot.* 2018, 1–8.
- [11] Y.C. Hou, K.S.M. Sahari, “Real-time modeling and parameter approximation of dexterous garment folding by robot,” *Artif. Life Robot.* 2018.
- [12] Y.C. Hou, K.S.M. Sahari, Y.W. Leong, D.N.T. How, H. Seki, Particle-based perception of garment folding for robotic manipulation purposes, *Int. J. Adv. Robot. Syst.* 14 (2017), 1–14.

## Authors Introduction

**Yew Cheong Hou**



He obtained his bachelor and master degrees from Universiti Tenaga Nasional in 2009 and 2012 respectively. He is currently pursuing his PhD at Universiti Tenaga Nasional.

**Khairul Salleh Mohamed Sahari**



He is currently a Professor at the Department of Mechanical Engineering, Universiti Tenaga Nasional. He received his PhD from Kanazawa University, Japan in 2006. His research interest includes robotics, mechatronics, Artificial Intelligence and Machine Learning.