Teaching Robots to Perform Construction Tasks via Learning from Demonstration

C. J. Liang[a], V. R. Kamat[a,b] and C. C. Menassa[a]

[a]Department of Civil and Environmental Engineering, University of Michigan, USA
[b]Robotics Institute, University of Michigan, USA
E-mail: cjliang@umich.edu, vkamat@umich.edu, menassa@umich.edu

Abstract –

Robots are expected to be widely used on future construction sites to assist human workers in the performance of repetitive physically-demanding tasks. Unlike typical manufacturing assembly lines, where parts are delivered to robots and workers in stationary workstations, construction robots and human workers must accumulate all necessary resources and repeatedly navigate to desired assembly locations on-site to perform useful work. The condition of such resources and the geometry of the environment are constantly changing and generally unstructured. As a result, the motion trajectories of any robot arms cannot be programmed beforehand. The robots must define the trajectory based on the encountered workspace geometry. Learning from Demonstration (LfD) methods have the potential to be used in teaching robots specific skills through human demonstration such that the robots can repeat the same process in different conditions. In this research, we explore the LfD method to teach robots to perform repetitive but geometrically adaptive construction tasks of installing suspended ceiling tiles within pre-assembled ceiling grids. The developed method translates the work context from the set of training videos demonstrated by humans to the target scene, then applies the reinforcement learning method to generate the policy for the robot to perform the subsequent ceiling tile installation. The first phase of the proposed method, i.e., the context translation model, is implemented and evaluated by characterizing whether robot-installed ceiling tiles are successfully moved to the grid area. The experiments demonstrate promising results that show the applicability of the LfD method in teaching robots to perform geometrically-adaptive construction tasks.

Keywords –

Robot Learning from Demonstration; Computer Vision; Autoencoder; Ceiling Tile Installation

1 Introduction

Human-robot and co-robot collaborative teams are broadly envisioned to be deployed on future construction sites [1]. The construction industry accounts for 4.3% of GDP in the U.S. [2] and is predicted to reach over 1.4 trillion U.S. dollars in volume by 2021, and $15.5 trillion worldwide by 2030 [3]. Construction is the economic locomotive of the society, whose competitiveness is affected by speed and quality of the construction process. However, the construction industry is confronting major issues of the aging workforce and the lack of skilled labor [44]. The U.S. labor force growth is forecasted to be lower than 0.5% by 2030 due to demographic transition [4]. The average age of the construction worker in the U.S. is 43 years old, and the younger generation is reluctant to enter the construction industry [5] due to its dangerous and demanding working environment, which causes a significant decrease in the employment-population. As a result, applying co-robots to assist or relieve human workers from hazardous, dangerous, and repetitive construction tasks has emerged as a key prospect to help mitigate issues faced by the construction industry [6–9].

According to Manzo et al. [10], approximately 50% of the construction tasks can be automated and may potentially replace nearly 2.7 million construction workers with robots by 2057. The human workers will work alongside robots or supervise them on the future construction site. By applying robots on construction sites, the human workers will switch their duty to perform the planning and cognitive tasks as supervisors, and train the robots to perform the repetitive physical work.

Similar to the manufacturing assembly line setting, the construction process is performed by several repetitive basic tasks [11,12], as shown in Figure 1. Take suspended ceiling tile installation process as an example: workers have to measure the ceiling tile layout, maneuver and position the tile, place the tile, and inspect the alignment. However, unlike the manufacturing environment, the construction site is an unstructured and dynamic environment [45]. The robots have to understand the environment, and then plan and execute the task accordingly. They need to navigate to the next assembling location as well, which makes it impossible for robots to perform the exact same task repeatedly [13].

Figure 1. The construction process involves repetitive tasks with different objects, such as ceiling tile installation, the tiles are maneuvered and placed repetitively at different locations.

On the other hand, construction projects are unique and unidentical. The loose tolerances and the relatively low quality control of the construction project causes each workpiece to be not strictly identical with each other and requires additional adjustments or improvisation on-site. Even though the workpieces such as tiles are designed to be the same size, the actual pieces may not be the same. Therefore, construction robots must learn how to plan and execute the construction process while overcoming the general uncertainty of the environment and the workpieces they handle.

2 Robot Learning from Demonstration

When training the human worker to perform a task, the human will learn the task by observing the demonstration from experts and practicing the action. The practical knowledge will be absorbed during the practice. Thus, similar to the human learning procedure, the Learning from Demonstration method (LfD) [14,15] can be utilized for co-robots to learn the collaborative task. One approach is imitation learning, where the agent attempts to clone the behavior of the expert [16]. Inverse reinforcement learning (IRL), on the other hand, estimates a reward function based on the expert's demonstration [17,18], then utilizes the learned reward function to obtain the policy. Existing demonstration methods include visual demonstration [19–23], force demonstration [24–26], visual and force demonstration [27,28], and trajectory demonstration [29–35].

First, the visual demonstration methods utilize several or limited video clips of expert's demonstration [36–38]. These methods usually apply to move, push or place simple and light objects, which is not suitable for the construction process, that require handling complex and heavy objects of various geometries. Second, the force demonstration methods are done by force sensors attached to robots and human experts. The robot will learn the contact force on the object and try to replicate the same action. These methods have only been demonstrated on simple tasks such as bolting or unscrewing bottle caps. They do not consider the environment and only try to repeat the action.

Third, the trajectory demonstration methods utilize demonstrating trajectory from expert to train the robot, such as driving simulation or path planning. One of the examples is the kinesthetic demonstration [25], where the robot is dragged by the human to finish the task while memorizing the trajectory. However, these methods require sufficient trajectory data from experts, which is difficult to obtain on construction sites. Finally, the visual and force demonstration methods combine the advantage of both methods to teach the robot to manipulate or grasp objects [28]. For construction tasks, in addition to visual observation, the workers have to feel the haptic or force feedback from the object and react accordingly, yet the existing visual and force demonstration methods only consider the pose of the human instead of the objects.

The objective of this research is to investigate the robot Learning from Demonstration method (LfD) [36] and evaluate the feasibility of applying LfD for teaching construction tasks to co-robots. The trained robots can collaborate with human workers while the human workers focus on the planning and cognitive tasks on the future construction site. The suspended ceiling tiles installation is used as the target construction process to describe the developed methods. The video of the human worker performing the tile installation is recorded in the laboratory and utilized to train the robot. The pose of the object, i.e., ceiling tile, and the target location, i.e., suspended grid, are extracted and tracked in the training video, then encoded as knowledge for the robot to learn. The success rate of the robot performance evaluates the learned skill of the robot.

3 Construction Task Learning

The method for learning the construction task is adapted from a context translation and imitation method by Liu et al. [36]. The practical knowledge of the skill is extracted from the training video as context and translated to the target scene. The robot can further learn the results of the translation through reinforcement learning method [39]. The detailed problem definition of the construction task and the context translation method are discussed in the following sections.

3.1 Problem Definition and Assumption

When teaching a robot to perform a specific task, the knowledge, or context, from the expert's demonstration must be defined in order to let the robot know what information needs to be tracked and absorbed, as well as how to determine the action to take for achieving the task. The context ω can be defined as the pose of the object and expert, the viewpoint of the camera, the condition of the environment, or the target location. In this research, we assume the camera is fixed in two different viewpoints, and the task is to perform in the same environment in order to reduce the complexity.

The demonstration of the task is defined as [36] (1):

$$\begin{bmatrix} D_1 \\ D_2 \\ \vdots \\ D_n \end{bmatrix} = \begin{bmatrix} O_0^1 & O_1^1 & \cdots & O_T^1 \\ O_0^2 & O_1^2 & \cdots & O_T^2 \\ \vdots & \ddots & & \vdots \\ O_0^n & O_1^n & \cdots & O_T^n \end{bmatrix} \tag{1}$$

where O_t is the observation at time t, which is generated from the Partially Observable Markov Decision Process (POMDP) [40]. The probability observation distribution $p(O_t|s_t,\omega_i)$, dynamics $p(s_{t+1}|s_t,a_t,\omega_i)$, and the policy of the expert $p(a_t|s_t,\omega_i)$ are utilized to define the POMDP, where s_t and s_{t+1} are the current and next state (e.g., unknown Markovian state), a_t is the action of the agent (e.g., maneuvering direction), and ω_i is the i-th context (e.g., pose of the ceiling tile, viewpoint). The context is unknown to the robot learner. Since the robot learner might try to track the mismatch context from the demonstration (e.g., follow the pose of the ceiling tile but consider as the pose of the expert), the context ω_i is assumed to be sampled independently and the robot learner has fixed context ω_l sampled from the same distribution [36].

Due to the undetermined feature of the context variables, the robot learner cannot know whether the context from the demonstration is the same as the context in their learning domain. This can be overcome by applying the context translation model [36] to translate the context from the source to the target, then defining the reward function based on the tracked context feature for reinforcement learning method, such as Trust Region Policy Optimization (TRPO) [41] or Deep Deterministic Policy Gradient (DDPG) [42], to learn the policy. The source and the target demonstrations are defined as follows (2)(3):

$$D_s = [O_0^s \quad O_1^s \quad \cdots \quad O_T^s] \tag{2}$$
$$D_t = [O_0^t \quad O_1^t \quad \cdots \quad O_T^t] \tag{3}$$

where D_s is the demonstrations from the unknown context of source video, and D_t is from the unknown context of target video. After training with sufficient examples, the context translation model is capable of translating the new demonstration D_n into the robot learner's context ω_l so that the robot can track and learn the feature.

3.2 Context Translation Model

The objective of the context translation model is to learn the translation function that can translate the source demonstration $D_s = [O_t^s]$, $t = 0,1,\dots T$ to the target context ω_t with the first observation O_0^t in the target demonstration D_t, that is, the first frame of the target demonstration video. The full translation function is defined as (4):

$$M(O_t^s, O_0^t) = (\hat{O}_t^t)_{trans} \tag{4}$$

where $(\hat{O}_t^t)_{trans}$ represents the translated observations in the robot learner's context.

The context translation model is constructed by several encoders, decoders, and autoencoders [43], which includes a source encoder $En_s(O_t^s)$, a target first observation encoder $En_t(O_0^t)$, and a target context decoder $De_t(z_{trans})$, as shown in Figure 2. On the left side is the framework of translating the source observations to the target observation through the translation function $T(z_s, z_t) = z_{trans}$, where z_s, z_t and z_{trans} represent the features of the encoded source, target, and translation. The loss function for training the translation is defined as L_2-norm (5):

$$L_{trans} = \left\| (\hat{O}_t^t)_{trans} - O_t^t \right\|_2^2 \tag{5}$$

Since the unknown context is translated from source to target, the features need to train on the target video in order to ensure the consistency of the feature representation between the encoder En_s and decoder De_t. The right side of Figure 2 is the framework of the autoencoder for training the En_s and De_t with a reconstruction loss, which is defined as (6):

$$L_{rec} = \left\| De_t(En_s(O_t^t)) - O_t^t \right\|_2^2 \tag{6}$$

The next step is to align the feature representation of the autoencoder with features z_{trans}. The loss function for the alignment is defined as (7):

$$L_{align} = \| z_{trans} - En_s(O_t^t) \|_2^2 \tag{7}$$

3.3 Network Architecture

The network of the encoder and decoder is illustrated in Table 1. In the encoder network, four convolutional layers with different filter size and stride followed by two linear layers with the size of 100 and 0.5 dropout are applied to the training video frame. In the decoder network, one linear layer with 0.5 dropout followed by four deconvolutional layers with different filter size and stride are applied to the translated feature. All the convolutional and deconvolutional layers are followed by LeakyReLu activation function with 0.2 leak except the last deconvolutional layer in the decoder. The filter size of the linear layer in the decoder is dependent on the size of the input image. Batch normalization is applied to the network. The input image is cropped to size of 48 by 48 pixels for training and testing.

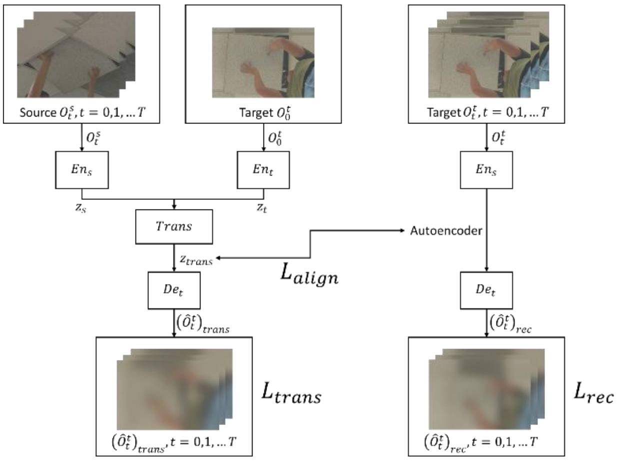

Figure 2. Architecture of the context translation model.

Table 1. Network architecture of the encoder and decoder.

Type	Layer	Filter size	Stride	Other
Encoder	Conv	32	1	LeakyReLu leak = 0.2
	Conv	16	2	
	Conv	16	1	
	Conv	8	2	
	Linear	100	n/a	Dropout = 0.5
	Linear	100	n/a	
Translation function	Linear	100	n/a	Dropout = 0.5
Decoder	Linear	*	n/a	Dropout = 0.5
	Deconv	16	½	LeakyReLu leak = 0.2
	Deconv	16	1	
	Deconv	32	½	
	Deconv	3	1	n/a

*Depends on the size of the input image

3.4 Reward Function for Robot Learning

After the source context is translated to the target observation, a reward function must be defined for the robot to learn the policy through reinforcement learning method. The reward function contains a feature tracking reward function (8):

$$\hat{R}_f(O_t^l) = -\left\| En_s(O_t^l) - \frac{1}{n}\sum_i^n T(z_s, z_t) \right\|_2^2 \tag{8}$$

and an image tracking reward function [36] (9):

$$\hat{R}_i(O_t^l) = -\left\| O_t^l - \frac{1}{n}\sum_i^n M(O_t^s, O_0^t) \right\|_2^2 \tag{9}$$

where $En_s(O_t^l)$ encodes the learner's observation O_t^l to z_l. Finally, the total reward function is defined as (10):

$$\hat{R}(O_t^l) = \hat{R}_f(O_t^l) + \omega \hat{R}_i(O_t^l) \tag{10}$$

where ω represents the weight.

4 Experiments

In order to evaluate the feasibility of applying the robot LfD method for construction tasks, the ceiling tile installation demonstration video was collected in the laboratory and utilized to train the context translation model. The performance of the model was evaluated by the success rate of the installation task.

4.1 Implementation and Training Details

The context translation model was implemented by modifying the original network using TensorFlow. A total of 55 videos were utilized to train the network, and 20 initial observations were used as the testing data, i.e., first frame of the video as the starting point for testing. In the demonstration video, the camera was set up at two fixed viewpoints for reducing the complexity, as shown in Figure 3. The network was trained by ADAM optimizer with learning rate 10^{-4} and the loss function described above.

Figure 3. Example of the ceiling tile installation demonstration video with two different camera viewpoints.

4.2 Results

We use the success rate to evaluate the performance of the context translation model on the ceiling tile installation. The success metric is defined as whether the final distance between the ceiling tile and the target grids is within a predefined threshold. After the source context is translated to the target scene, we identified the final location of the ceiling tile and the grid, then determined whether they were within the predefined threshold in order to calculate the success rate. Figure 4 is an example of the demonstration source video of ceiling tile installation. The human worker demonstrates how to install the ceiling tile.

Figure 5 shows examples of the target observations with two different viewpoints (iso view and bottom view). The target observation is the first frame of the demonstration video, which is utilized for translating the context from the source video. Figure 6 shows the results of the translated scene. The top row is the successful result, and the bottom row is the unsuccessful result. The red rectangle represents the ceiling tile, and the green rectangle represents the target grid. The distance between the ceiling tile and the grid is over the predefined threshold; thus it is determined as unsuccessful result. We compare the success rate of the translation with the different type of viewpoint, as shown in Table 2. The success rate of the bottom view is 25%, and the iso view is 43%. The overall success rate is 35%.

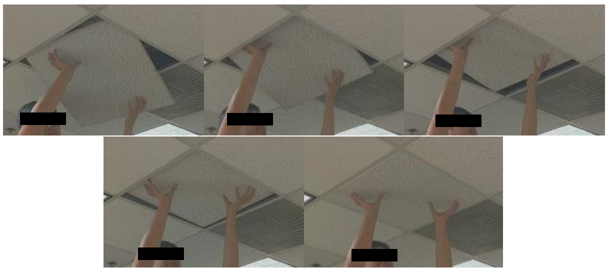

Figure 4. Example of a source video of ceiling tile installation.

Figure 5. Example of target observations, which is the first frame of the video.

Translated Success

Translated Failure

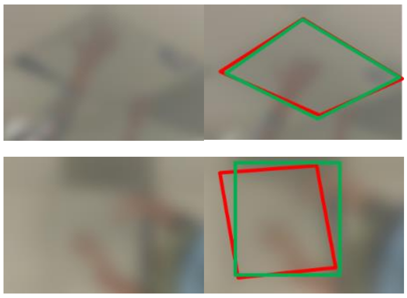

Figure 6. Results of the translated scene. Top - successful result; Bottom - unsuccessful result. Red rectangle represents the ceiling tile, and the green rectangle represents the target grid.

Table 2. Success rate of the translated result.

Viewpoint	Success	Failure	Success rate
Bottom view	2	6	25%
Iso view	5	7	42%
Overall	7	13	35%

In comparison with the Liu et al. [36], where they used the context translation model to teach the robot to ladle almonds into a frying pan and sweep the almonds into a dustpan, the success rate is 66% with 60 training videos of the almonds pouring and 75% with 100 training videos of the almonds sweeping. Thus, the performance can be improved by providing more demonstration videos. In addition, the bottom view has a lower success rate since most of the trajectories of ceiling tile installation are vertical types and the bottom view cannot provide sufficient information. This can be addressed by avoiding the vertical type viewpoint.

5 Conclusion and Future Work

In this research, we proposed and evaluated a Learning from Demonstration (LfD) method to train the construction robot to perform repetitive construction tasks, in which we utilized the ceiling tile installation as the target task. We adopted a visual LfD method, i.e., the context translation model, for our application. The context translation model only uses the visual demonstration as input to train the robot. The model is trained on a set of ceiling tile installation demonstration videos. The results showed that the model could translate the work context from the source video to the target observation with acceptable success rate when using the iso camera view, which can further apply the reinforcement learning method for robots to determine the control policy. In ongoing work, we are collecting more demonstration videos from new viewpoints. We are also implementing the reinforcement learning method for robots to learn and perform the policy.

References

[1] You S., Kim J.-H., Lee S., Kamat V.R. and Robert L.P. Enhancing perceived safety in human–robot collaborative construction using immersive virtual environments, *Automation in Construction*. 96:161–170, 2018.

[2] Statista. U.S. construction industry - statistics & facts. Online: https://www.statista.com/topics/974/construction/, Accessed: 12/11/2018.

[3] GCP and Oxford Economics. Global construction 2030. Online: http://www.globalconstruction2030.com/#products, Accessed: 19/11/2018.

[4] Harris K., Kimson A. and Schwedel A. Labor 2030: the collision of demographics, automation and inequality. Online: https://www.bain.com/insights/labor-2030-the-collision-of-demographics-automation-and-inequality/, Accessed: 2/1/2019.

[5] Belton P. Why robots will build the cities of the future. Online: https://www.bbc.com/news/business-46034469, Accessed: 19/11/2018.

[6] Lundeen K.M., Kamat V.R., Menassa C.C. and McGee W. Scene understanding for adaptive manipulation in robotized construction work. *Automation in Construction*, 82:16–30, 2017.

[7] Liang C.-J., Lundeen K.M., McGee W., Menassa C.C., Lee S. and Kamat V.R. Stacked hourglass networks for markerless pose estimation of articulated construction robots. In *Proceedings of the 35th International Symposium on Automation and Robotics in Construction*, pages 1–7, Berlin, Germany, 2018.

[8] Liang C.-J., Kang S.-C. and Lee M.-H. RAS: a robotic assembly system for steel structure erection and assembly. *International Journal of Intelligent Robotics and Applications*, 1(4):459–476, 2017.

[9] Liang C.-J. and Kang S.-C. Development of a steel beam hauling system for automatic steel beam assembly. In *Proceedings of the 2014 International Conference on Computing in Civil and Building Engineering*, pages 1295–1302, Orlando, FL, USA, 2014.

[10] Manzo J., Manzo F. and Bruno R. The potential economic consequences of a highly automated construction industry. Online: https://midwestepi.files.wordpress.com/2018/01/the-economic-consequences-of-a-highly-automated-construction-industry-final.pdf, Accessed: 30/12/18.

[11] Everett J.G. and Slocum A.H. Automation and robotics opportunities: construction versus manufacturing, *J. of Construction Engineering and Management*, 120(2):443–452, 1994.

[12] Kamat V.R. and Martinez J.C. CEPM4: comparison of simulation-driven construction operations visualization and 4D CAD. In *Proceedings of the 34th Conference on Winter Simulation: Exploring New Frontiers*, pages 1765–1770, San Diego, CA, USA, 2002.

[13] Lundeen K.M., Kamat V.R., Menassa C.C. and McGee W. Autonomous motion planning and task execution in geometrically adaptive robotized construction work. *Automation in Construction*, 100:24–45, 2019.

[14] Argall B.D., Chernova S., Veloso M. and Browning B. A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57(5):469–483, 2009.

[15] Zhu Z. and Hu H. Robot learning from demonstration in robotic assembly: a survey. *Robotics*, 7(2):17, 2018.

[16] Schaal S. Is Imitation Learning the Route to Humanoid Robots? *Trends in Cognitive Sciences*, 3(6):233–242, 1999.

[17] Ng A.Y. and Russell S.J. Algorithms for inverse reinforcement learning. In *Proceedings of the International Conference on Machine Learning*, pages 663–670, San Francisco, CA, USA, 2000.

[18] Levine S., Popovic Z. and Koltun V. Nonlinear inverse reinforcement learning with gaussian processes. In *Advances in Neural Information Processing Systems*, pages 19–27, Granada, Spain, 2011.

[19] Acosta-Calderon C.A. and Hu H. Robot imitation: body schema and body percept. *Applied Bionics and Biomechanics*, 2(3-4):131–148, 2005.

[20] Finn C., Yu T., Zhang T., Abbeel P. and Levine S. One-shot visual imitation learning via meta-learning. In *Proceedings of the 1st Annual Conference on Robot Learning*, pages 357–368, Mountain View, CA, USA, 2017.

[21] Kalashnikov D., Irpan A., Pastor P., Ibarz J., Herzog A., Jang E., Quillen D., Holly E., Kalakrishnan M., Vanhoucke V. and Levine S. Scalable deep reinforcement learning for vision-based robotic manipulation. In *Proceedings of The 2nd Conference on Robot Learning*, pages 651–673, Zurich, Switzerland, 2018.

[22] Kim B., Farahmand A., Pineau J. and Precup D., Learning from limited demonstrations. In *Neural Information Processing Systems*, pages 9, Lake Tahoe, NV, USA, 2013.

[23] Krishnan S., Fox R., Stoica I. and Goldberg K. DDCO: discovery of deep continuous options for robot learning from demonstrations. In *Proceedings of the 1st Annual Conference on Robot Learning*, pages 418–437, Mountain View, CA, USA, 2017.

[24] Abu-Dakka F.J., Rozo L. and Caldwell D.G. Force-based variable impedance learning for robotic manipulation. *Robotics and Autonomous Systems*. 109:156–167, 2018.
[25] Calinon S., Guenter F. and Billard A., On learning the statistical representation of a task and generalizing it to various contexts. In *Proceedings of the 2006 IEEE International Conference on Robotics and Automation*, pages 2978–2983, Orlando, FL, USA, 2006.
[26] Li W. and Fritz M. Teaching robots the use of human tools from demonstration with non-dexterous end-effectors. In *Proceedings of the IEEE-RAS 15th International Conference on Humanoid Robots*, pages 547–553, Seoul, South Korea, 2015.
[27] Calandra R., Owens A., Upadhyaya M., Yuan W., Lin J., Adelson E.H. and Levine S. The feeling of success: does touch sensing help predict grasp outcomes? In *Proceedings of the 1st Annual Conference on Robot Learning*, pages 314–323, Mountain View, CA, USA, 2017.
[28] Edmonds M., Gao F., Xie X., Liu H., Qi S., Zhu Y., Rothrock B. and Zhu S.-C. Feeling the force: Integrating force and pose for fluent discovery through imitation learning to open medicine bottles. In *Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3530–3537, Vancouver, BC, Canada, 2017.
[29] Abu-Dakka F.J., Nemec B., Kramberger A., Buch A.G., Krüger N. and Ude A., Solving peg-in-hole tasks by human demonstration and exception strategies. *Industrial Robot: An International Journal*, 41(6):575–584, 2014.
[30] Peternel L., Petrič T. and Babič J. Robotic assembly solution by human-in-the-loop teaching method based on real-time stiffness modulation, *Autonomous Robots*, 42(1):1–17, 2018.
[31] Piot B., Geist M. and Pietquin O. Boosted bellman residual minimization handling expert demonstrations. In *Proceedings of the ECML PKDD 2014*, pages 549–564, Nancy, France, 2014.
[32] Ravichandar H., Salehi I. and Dani A. Learning partially contracting dynamical systems from demonstrations. In *Proceedings of the 1st Annual Conference on Robot Learning*, pages 369–378, Mountain View, CA, USA, 2017.
[33] Savarimuthu T.R., Buch A.G., Schlette C., Wantia N., Roßmann J., Martínez D., Alenyà G., Torras C., Ude A., Nemec B., Kramberger A., Wörgötter F., Aksoy E.E., Papon J., Haller S., Piater J. and Krüger N. Teaching a robot the semantics of assembly tasks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*. 48(5):670–692, 2018.
[34] Tanwani A.K. and Calinon S., Learning robot manipulation tasks with task-parameterized semitied hidden semi-markov model. *IEEE Robotics and Automation Letters*. 1(1):235–242, 2016.
[35] Wang G.-F., Fang Z. and Li P. Shaping in reinforcement learning by knowledge transferred from human-demonstrations of a simple similar task, *Journal of Intelligent & Fuzzy Systems*. 34(1):711–720, 2018.
[36] Liu, Y., Gupta A., Abbeel P. and Levine S. Imitation from observation: learning to imitate behaviors from raw video via context translation. In *Proceedings of the 2018 IEEE International Conference on Robotics and Automation*, pages 1118–1125, Brisbane, Australia, 2018.
[37] Shukla N., He Y., Chen F. and Zhu S.-C. Learning human utility from video demonstrations for deductive planning in robotics. In *Proceedings of the 1st Annual Conference on Robot Learning*, pages 448–457, Mountain View, CA, USA, 2017.
[38] Xie A., Singh A., Levine S. and Finn C. Few-shot goal inference for visuomotor learning and planning. In *Proceedings of The 2nd Conference on Robot Learning*, pages 40–52, Zurich, Switzerland, 2018.
[39] Sutton R.S. and Barto A.G. *Reinforcement learning: an introduction*, MIT Press, Cambridge, MA, 2018.
[40] Kaelbling L.P., Littman M.L. and Cassandra A.R. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*. 101(1):99–134, 1998.
[41] Schulman J., Levine S., Abbeel P., Jordan M. and Moritz P. Trust region policy optimization. In *Proc. of International Conference on Machine Learning*, pages 1889–1897, Lille, France, 2015.
[42] Lillicrap T.P., Hunt J.J., Pritzel A., Heess N., Erez T., Tassa Y., Silver D. and Wierstra D. Continuous control with deep reinforcement learning. In *Proc. of the 4th International Conference on Learning Representations*, San Juan, Puerto Rico, 2016.
[43] Vincent P., Larochelle H., Bengio Y. and Manzagol P.-A. Extracting and composing robust features with denoising autoencoders. In *Proceedings of the International Conference on Machine Learning*, pages 1096–1103, Helsinki, Finland, 2008.
[44] Behzadan A. H., Iqbal A. and Kamat V. R. A collaborative augmented reality based modeling environment for construction engineering and management education. In *Proceedings of the Winter Simulation Conference (WSC)*, pages 3568-3576, Phoenix, AZ, USA, 2011.
[45] Akula M., Lipman R. R., Franaszek M., Saidi K. S. Cheok G. S. and Kamat V. R. Real-time drill monitoring and control using building information models augmented with 3D imaging data. *Automation in Construction*. 36: 1–15, 2013