Annual Reviews in Control 44 (2017) 9–18

Contents lists available at ScienceDirect

Annual Reviews in Control

journal homepage: www.elsevier.com/locate/arcontrol

Review article

Realization of task intelligence for service robots in an unstructured environment☆

Deok-Hwa Kim, Gyeong-Moon Park, Yong-Ho Yoo, Si-Jung Ryu, In-Bae Jeong, Jong-Hwan Kim*

School of Electrical Engineering, KAIST, Daejeon, Republic of Korea

A R T I C L E I N F O

Article history:
Received 14 August 2017
Revised 25 September 2017
Accepted 25 September 2017
Available online 5 October 2017

Keywords:
Task intelligence
Episodic memory
Deep ART network
Motion planning
Task planner

A B S T R A C T

In order to perform various tasks using a robot in a real environment, it is necessary to learn the tasks based on recognition, to be able to derive a task sequence suitable for the situation, and to be able to generate a behavior adaptively. To deal with this issue, this paper proposes a system for realizing task intelligence having a memory module motivated by human episodic memory, and a task planning module to resolve the current situation. In addition, this paper proposes a technique that can modify demonstrated trajectories according to current robot states and recognized target positions in order to perform the determined task sequence, as well as a technique that can generate the modified trajectory without collisions with surrounding obstacles. The effectiveness and applicability of the task intelligence are demonstrated through experiments with Mybot, a humanoid robot developed in the Robot Intelligence Technology Laboratory at KAIST.

Contents

☆ This paper is an extended version of the summary of your plenary talk at IFAC2017.

* Corresponding author.
E-mail addresses: dhkim@rit.kaist.ac.kr (D.-H. Kim), gmpark@rit.kaist.ac.kr (G.-M. Park), yhyoo@rit.kaist.ac.kr (Y.-H. Yoo), sjryu@rit.kaist.ac.kr (S.-J. Ryu), ibjeong@rit.kaist.ac.kr (I.-B. Jeong), johkim@rit.kaist.ac.kr (J.-H. Kim).

https://doi.org/10.1016/j.arcontrol.2017.09.013
1367-5788/

1. Introduction

With ongoing advances of robot technology, various types of robots are being used in the home environment and industrial field. The purposes of the robots are mainly to perform work on behalf of a person or to provide convenient service to a person. However, most of the robots that have been developed to date are focused on providing a single service. Therefore, it is important to study robot task intelligence, which not only allows the provision of various services, but also allows robots to learn task sequences by themselves (Beetz et al., 2012; Jeong et al., 2017). In order to realize task intelligence for robots, it utilizes various artificial intelligence technologies, such as vision-based recognition, robot control, memory-based task sequence learning, adaptive task sequence planning, motion planning, and so on.

Task intelligence has been researched mainly to conduct fixed task sequences (Aboaf, Drucker, & Atkeson, 1989; Beetz et al., 2011; Bollini, Tellex, Thompson, Roy, & Rus, 2013), or to perform planning of sequences to achieve goal states in a symbolical manner (Beetz et al., 2012). Beetz et al. devised a web-based sequence retrieval system to make a pancake (Beetz et al., 2011), and Bollini et al. developed a cooking robot that can interpret recipes (Bollini et al., 2013). Furthermore, in order to learn various tasks, Jeong et al. designed a task intelligence architecture using a neural model-based memory module to memorize behavior sequences. However, these approaches having a fixed sequence for a certain task are not robust to a changeable environment due to a lack of consideration of the working environment. To solve this problem, Beetz et al. presented a cognitive mechanism that incorporates learning, reasoning, and planning for realization of task intelligence (Beetz et al., 2012). In addition, Misra et al. developed a robot having the capability to understand user's instructions in the form of natural language and to generate task sequences considering the working environment (Misra, Sung, Lee, & Saxena, 2016). For the adaptive task sequence generation, it is necessary to define actions, objects, and state changes for each action in a symbolic form. However, there is a disadvantage that defining all actions and all possible states is not scalable. In this light, Kim et al. designed a memory-based task intelligence architecture by fusing the task planner to generate the adaptive task sequences (Kim, Baek, Cho, & Kim, 2016). In their research, the task sequences are learned by a neural network-based memory module, and appropriate action sequences are added according to the situation.

In order to realize task intelligence, a capability to generate adaptive behaviors for a surrounding environment is necessary. In the field of robotics, the research for a precise manipulation for a high degrees-of-freedom (DoF) arm have been conducted (Park, Lee, Cho, Hong, & Kim, 2012), and the learning method of human-demonstrated trajectories for a specific manipulation task has mainly been studied (Jain, Sharma, Joachims, & Saxena, 2015; Sung, Jin, & Saxena, 2015). In this regard, Jain et al. presented a learning method of preferences for manipulation tasks (Jain et al., 2015). Through their research, preference-based manipulation by giving a users feedback could be possible without demonstrations for entire manipulations to obtain optimal trajectories. Sung et al. developed a robot that can make a coffee by using transferred manipulation trajectories for unknown object parts (Sung et al., 2015). In their research, the manipulation trajectories for a part of an object were learned from a crowd-sourcing demonstration of a virtual reality, and the trajectories for an unknown object part could be retrieved by searching similar parts in pre-learned objects. However, these approaches are not suitable for an environment having complex obstacles, because collision avoidance is not considered in their work. To consider the collision avoidance, MoveIt package, which can generate collision-free trajectories based on the open motion planning library (OMPL) (Moll, Şucan, & Kavraki, 2015; Şucan, Moll, & Kavraki, 2012) by using an OctoMap (Hornung, Wurm, Bennewitz, Stachniss, & Burgard, 2013), has been used in many robot applications (Chitta, Sucan, & Cousins, 2012). A sampling-based motion planning approach of the OMPL is often used to generate the collision-free trajectory to reach a target pose or to follow target trajectory. For using this package, the problems to solve the trajectories for each manipulation should be pre-programed. Due to this pre-programming, the sampling-based motion planning approach is not scalable.

In this paper, we propose a novel task intelligence system for service robots having a capability to generate adaptive behaviors in a surrounding environment. First, in the proposed task intelligence, the surrounding environment is perceived by using a vision sensor. In reality, robust recognition is important to extract stable information for objects to be manipulated. In this regard, we propose a robust perception method based on an RGB-D sensor. Furthermore, this paper uses the Deep ART network module to store and retrieve task sequences (Park, Kim, Yoo, & Kim, 2017). The Deep ART network, of which function is similar to that of the episodic memory of humans, can memorize task sequences from human demonstrations. The memorized task sequences are retrieved from the Deep ART network by an input cue to perform an associated task. A task planner, such as fast-forward (FF) planner, is used to follow the task sequences retrieved from the Deep ART network module, because there is a mismatch between the learned environment and the current environment. In order to perform each determined task sequence, this paper proposes an adaptive behavior generation method depending on a geometrical environment. The proposed behavior generation method enables scalable behavior learning and collision-free manipulation.

This paper is organized as follows: Section 2 describes the proposed task intelligence system in detail. To verify the effectiveness of the proposed system, experimental environments and scenarios are presented, and the experimental results are discussed in Section 3. Finally, concluding remarks follow in Section 4.

2. Task intelligence

This paper proposes a task intelligence system for performing the task sequences demonstrated by a user in an unstructured environment. The task intelligence system is based on the intelligent operating architecture (iOA) that is used to realize a general intelligence of robots (Kim, Choi, Park, & Zaheer, 2013). The iOA imitates the basic functions of the human brain in robot aspects. It consists of perception, memory, reasoning, internal state, and execution parts. This paper focuses on the architecture for performing tasks as services to users. The proposed task intelligence system is shown in Fig. 1. First, environment information (objects, contexts) around the robot is obtained from the vision sensor attached to the robot. The context information represents state information of the object (fillable, baked, etc.) and state information between the objects (graspable, reachable, etc.). The recognized information is used for the episodic memory, task planner, and behavior generation modules. When an instruction is given to the robot, the episodic memory module retrieves task sequences based on the recognized information. However, the learning environment and the actual environment may be different; for example, there are no objects that are needed, or the target objects are covered by