Elsevier Editorial System(tm) for Robotics and Computer Integrated Manufacturing
Manuscript Draft

Manuscript Number: RCIM-D-14-00095R1

Title: On-line Knowledge Acquisition and Enhancement in Robotic Assembly Tasks

Article Type: SI:Knowledge Driven Robotics

Keywords: Robotic Assembly, Artificial Neural Networks, Machine Vision, Kitting process

Corresponding Author: Dr. Ismael Lopez-Juarez,

Corresponding Author's Institution:

First Author: Jose L Navarro-Gonzalez, MSc

Order of Authors: Jose L Navarro-Gonzalez, MSc; Ismael Lopez-Juarez; Reyes Rios-Cabrera, PhD; Keny Ordaz-Hernandez, PhD

Abstract: Industrial robots are reliable machines for manufacturing tasks such as welding, panting, assembly, palletizing or kitting operations. They are traditionally programmed by an operator using a teach pendant in a point-to-point scheme with limited sensing capabilities such as industrial vision systems and force/torque sensing. The use of these sensing capabilities is associated to the particular robot controller, operative systems and programming language. Today, robots can react to environment changes specific to their task domain but are still unable to learn skills to effectively use their current knowledge. The need for such a skill in unstructured environments where knowledge can be acquired and enhanced is desirable so that robots can effectively interact in multimodal real-world scenarios.

In this article we present a Multimodal Assembly Controller (MAC) approach to embed and effectively enhance knowledge into industrial robots working in multimodal manufacturing scenarios such as assembly during kitting operations with varying shapes and tolerances. During learning, the robot uses its vision and force capabilities resembling a human operator carrying out the same operation. The approach consists of using a MAC based on the Fuzzy ARTMAP Artificial Neural Network in conjunction with a knowledge base. The robot starts the operation having limited initial knowledge about what task it has to accomplish. During the operation, the robot learns the skill for recognising assembly parts and how to assemble them. The skill acquisition is evaluated by counting the steps to complete the assembly, length of the followed assembly path and compliant behaviour. The performance improves with time so that the robot becomes an expert demonstrated by the assembly of a kit with different part geometries. The kit is unknown by the robot at the beginning of the operation; therefore, the kit type, location and orientation are unknown as well as the parts to be assembled since they are randomly fed by a conveyor belt.

July 21, 2014

Robotics and Computer Integrated Manufacturing Journal
Guest Editor to the Special Issue on Knowledge Driven Robotics and Manufacturing,
Elsevier.

RE Answer to reviewer's comments

Dear Prof. Craig Schlenoff,

I am submitting our reviewed version of the paper entitled **"On-line Knowledge Acquisition and Enhancement in Robotic Assembly Tasks"** for consideration and possible publication in your prestigious journal.

Please find attached the corrections/amendments made to the reviewed version of the paper according to the reviewers' comments. Please note that reviewers' comments are in Times Roman font and ours are in Calibri font.

Should you have any doubt or require further information please do not hesitate to contact me.

Faithfully yours,

Dr. Ismael Lopez-Juarez
Principal researcher
Research Centre for Advanced Studies (CINVESTAV), Mexico

Highlights

- A novel Multimodal Assembly Controller (MAC) was designed having minimal assembly information (-Z assembly direction).
- It was demonstrated that knowledge can be refined when other different matting pair are assembled without the need to acquire another Primitive Knowledge Base (PKB).
- The robot learns a new assembly and improves its skills from experience observed by a reduced number of patterns, lower compliant forces and shorter assembly trajectories.
- The MAC demonstrated that can be used in non-structured environments for the kitting process with uncertainties (in position and geometry) for both, the kits and the pegs.

Answers to No. 1 Reviewer's Questions/Comments

All questions/comments regarding our previously submitted paper were reviewed.
We are profoundly grateful for your valuable comments that helped to significantly improve the paper.

Your questions/comments are in Times New Roman font whereas our answer/amendments are given in Calibri font.

The page and line numbers mentioned in the following sections correspond to the new reviewed version of the paper.

The problem domain addressed is important and the authors' research is focusing on key areas that can make a big impact on the usefulness of robots in manufacturing. The project included consideration for learning multiple different tasks and adaptation to different conditions. The experimental setup included complexities such as vision for part recognition and localization and force sensing for constraining motion and for use in learning task strategies.

The paper overall flow is well-structured at a high-level. There are, however, gaps in the descriptions/results provided, which are listed below. There are also some suggestions on English language usage at the end (these are minor).

Some questions or comments regarding the work presented in the manuscript:

- Please define your task/domain terms. "Kitting" "peg-in-hole", and "assembly" are used interchangeably in the paper, but without defining kitting in general, or in contrast to assembly specifically. This terminology vagueness confuses some readers.

In this regard, a new paragraph was added in page 4, line 9-18 (from our revised paper version):

"The scope of the research presented here needs to define some aspects as to build up a sounded understanding. The task domain of our investigation refers to the assembly of parts having two well defined tasks, the peg in hole task and the kitting task. In this investigation the kitting task is composed by more than one assembly operation. In general sense, the work domain is in the assembly of parts, which is the action of putting together manufactured parts to make a completed product or subassembly. In particular, in this paper we will deal with a special type of assembly composed by a peg and its counterpart named hole and referred to as the Peg-In-Hole (PIH) task. The kitting task can then be understood as the practice of assembling components (pegs in our case) in predetermined quantities that are placed together in specific container (kit). The term of assembly parts, assembly components and mating pairs are used interchangeably within this paper."

In terms of the knowledge base:

- P. 10: you describe the data sent from the Master computer to the KRC2 controller. <DIST> is a distance value, but it sounds like a single scalar value. Would there not be a need to command x, y, and z components for the robot to move to?

The z, y and z components are indicated in the <DIST> byte according to the direction indicated in <CODE>. We amended this paragraph in page 11, lines 16-21 as follows.

“<CODE> is a byte containing the corresponding Command Code (16 possible motion direction commands: +x, -x, +y, -y, +z, -z, +x +y, +x -y, -x -y, -x +y, +Rx, -Rx, +Ry, -Ry, +Rz, -Rz) and 9 control commands that includes do nothing, go to home, world coordinates, tool coordinates, joint coordinates, base coordinates, end communication, open gripper and close gripper.
<DIST> is a byte containing a distance value in the direction indicated by <CODE> and given in tenths of mm.”

- ❖ The knowledge representation is not fully discussed within this paper.

Two paragraphs were rephrased/added to section 2.1 in page 8, lines 9-24 and page 10, lines 1-3 as follows.

“Two different approaches have emerged to represent knowledge. It can be either symbolic as employed in Artificial Intelligence algorithms or numeric (also referred to as subsymbolic) as processed by Artificial Neural Networks. The knowledge embodiment is different however in essence the apparent dichotomy between both approaches is more perceived than real as stated by Honavar [26]. Both approaches to modelling cognition and engineering in intelligent systems can be applied for integrating neural and symbolic processes.

In this paper, it is proposed for the main operation (PIH) to use a numeric knowledge representation using ANN, which can be refined by using on-line training examples (contact conditions during PIH operations). The knowledge is inserted into the network and subsequently refined by ANN training resulting in a Knowledge Based Artificial Neural Networks (KBANN) as pointed by Towell and Shavlik [27]. The core idea here is to create a Multimodal Assembly Controller (MAC) formed by a learning and recognition module in conjunction with a Knowledge Base. The proposal also includes a knowledge refinement module and world effectors so that information from the real-world is considered. The purpose is to train the KBANN under real-world situations providing the robot with the capability of recognising cues or primitive descriptors during early stages of learning, so that initial conditions can be started. During knowledge refinement, and by giving more examples, this knowledge is expected to be enhanced/refined. The Multimodal Assembly Controller is illustrated in figure 1.”

The paragraph in section 3.4, page 20, lines 3-10 was amended as follows:

“As mentioned in section 2.1, knowledge can be inserted in numeric form using KBANN. In this regard, the Adaptive Resonance Theory (ART) is a good candidate since it is a well-established associative brain and competitive model introduced as a theory of the human cognition [10]. The advantage compared to other cognitive models is its fast learning mode for recoding input-output mapping information since it takes usually one epoch to learn. In addition, the network’s incremental learning capability allowed the network to learn newly encountered patterns incrementally without catastrophic forgetting. Both attributes were important in our application since we are interested in controlling the operation in real-time.”

- Need to clarify/expound on the initial knowledge that the robot has (PKB and other) and how it is designed to acquire additional knowledge. For instance, (p 13) you describe that it has process knowledge that includes assembly direction and end-condition, plus rough locations for male and female components, forces and arm motions. How are all these tied together? Is there a structure that holds the different fields that the robot uses to guide its discovery and actions? For instance, what triggers the robot to use its vision system to look for, identify, and localize the components? How does it know that it needs to have 2 components in order to perform the operation? On p. 14, under "task understanding" you describe that the robot "recognizes the assembly it so that it learns the number of parts," etc. Where is this structure for this knowledge held and how is it used to guide the robot's actions?

We really appreciate your valuable point.
We missed to explain the main part of the Multimodal Assembly Controller (MAC) structure that allows tying together the knowledge refinement module in the MAC. This structure is the task planner that integrates forces, arm motions and also triggers when to use the vision templates and the force sensing capability.

Figure 1 and figure 2 were modified. Figure 6 was added to clarify this point.

New paragraphs were added in section 2.1 in page 10, lines 1-17 as follows:

"The Integrator (Task Planner) acts differently according to the stage of the assembly process and will be described in detail in the next section. But, basically depends on the type of motion being carried out (in contact during constraint motion or free motion without any contact). It is important to observe that the Integrator has a bi-directional data flow from/to the world effectors. In our application during free motion the Task Planner requires information about the end-effector position as well as it needs to send information to the motor drives (world effectors) to reposition the arm. During constraint motion, the Task Planner receives information about the end-effector position as well as force sensing information via the F/T sensor located at the robot's wrist.

The Task Planner is in charge of selecting the appropriate information and patterns depending on the involved tasks. During free motion, information about the current position of the arm and the object's template is needed for instance, when approaching the peg to the hole in readiness for insertion. During constraint motion, the Task Planner receives robot's end-effector position as well as the force/torque pattern and selects if the pattern will be learned or recalled. In the following sections, we will describe how the vision system works using template matching for task understanding as well the pattern selection mechanism for deciding if a certain contact force should be learned or forgotten."

A new section 3.3.1 about the task planner was written in page 16, lines 7-24 and page 17, lines 1-4.

- What does the "end condition" knowledge include? (back on p. 13)

This was clarified in page 15, line 8:

"End-condition (for the kitting and PIH operation)."

The kitting end condition is specified in section 3.3.1 (Task understanding) page 17, lines 19-20:

"The order is irrelevant, since the end-condition of the kitting process is when all pegs are inserted in the kit."

The PIH end-condition is specified in section 3.3.1 (Task understanding) page 19, lines 19-22:

"The insertion is solely guided by force and the end-condition is when the peg's body is inserted. It is important to note that in order to avoid a collision with the bottom of the hole, the peg was not allowed to contact. Therefore, the end-condition was considered when the peg's body length (9mm) was fully inserted, which occurred at about 90 downward step motions."

In terms of the learning process:

- The implications of the interplay between learning new patterns and reducing the number of patterns are unclear. It seems straightforward if describing a multi-stage experiment where the types of components are modified in the stated sequence. What happens when you run the final configuration of the system with the original cross-section parts again? Will it have to re-learn patterns that it forgot?

In our experience, yes the system would have to re-learn these new conditions. However this situation was not tested and it is recognised as envisaged work in section 5.

- P. 23: "It was observed that none of the patterns were learnt; however, the MAC still had to forget patterns." Any insights into why this is the case? (both no new patterns learned in particular; having to forget some makes a little bit of sense.) Was this building on the results after the square peg? Is there a combined knowledge base that encompasses the different cases?

This paragraph was modified explaining the reason (page 26, lines 13-15):

"It was observed that none of the patterns were learnt; however, the MAC still had to forget patterns since the new conditions with different offset triggered a further knowledge refinement."

❖ P. 24: "information from the environment was minimal since the robot only knew the assembly direction…" Isn't force/torque feedback also information from the environment? Maybe you should add "a priori" to this statement?

This paragraph was modified (page 28, lines 6-9):

"The a priori information from the environment was minimal since the robot only knew the assembly direction (-Z direction) in conjunction with the associated force/torque information. This PKB was solely used for starting the knowledge acquisition provided by the user to start the operations."

❖ P. 24: "is possible that it can become enhanced while assembling the same part geometry." Not sure that this was discussed in the preceding sections?

It is explained in page 26, lines 10-13:

"The test continued with the assembly of the same squared peg. In this case the knowledge evolved from the same geometry and we referred to as knowledge enhancement (insertions 13-16). The following test was to observe the robustness of the MAC by changing the part geometry again assembling the Radiused square peg (insertions 17-20)."

Other questions/comments

❖ P. 5: "The skill acquisition is evaluated through its performance based on the elapsed time to complete the assembly," Time to complete task is not reported in a meaningful way in the paper.

This was corrected since the elapsed time was not measured directly, but indirectly by the number of elapsed steps during insertions as stated in page 6, line 2:

"The assembly skill is evaluated based on the number of elapsed steps to complete the assembly…"

❖ P. 7: F/T is used without explaining. Some readers may not know that it is shorthand for "force-torque"

It was defined in the reviewed version of the paper at the beginning (page 7 line 20).

❖ P. 8: Figure 1 - this figure requires more explanation.

This figure was redrawn along with better explanation in page 8, lines 16-24 and page 9, lines 1-15:

"In this paper, it is proposed for the main operation (PIH) to use a numeric knowledge representation using ANN, which can be refined by using on-line training examples (contact conditions during PIH operations). The knowledge is inserted into the network and subsequently refined by ANN training resulting in a Knowledge Based Artificial Neural Networks (KBANN) as pointed by Towell and Shavlik [27]. The core idea here is to create a Multimodal Assembly Controller (MAC) formed by a learning and recognition module in conjunction with a Knowledge Base. The proposal also includes a knowledge refinement module and world effectors so that information from the real-world is considered. The purpose is to train the KBANN under real-world situations providing the robot with the capability of recognising cues or primitive descriptors during early stages of learning, so that initial conditions can be started. During knowledge refinement, and by giving more examples, this knowledge is expected to be enhanced/refined. The Multimodal Assembly Controller is illustrated in figure 1.

Figure 1. Learning and knowledge refinement

The learning and recognition module –KBANN- is the heart of the MAC, which includes three additional modules: the PKB, the World Effector and the Knowledge Refinement module. The PKB stores initial information about the environment provided by a human expert. This information is only used during the first stage of training. In this stage the switch SW1 will be open and the switch SW2 closed since the initial training is made only using the PKB. After passing this initial state, the KBANN will predict the next action based on the current input from the sensor (SW1 closed and SW2 open). Knowledge can be refined according to a quality criterion in the Knowledge Refinement module. Should the corresponding force/torque pattern meets the criterion, then the pattern is allowed to be included in the EKB, then SW2 is closed for on-line retraining. Otherwise, the KBANN is in a recall stage until new and useful patterns are encountered."

- P. 8: lines 14-15: "(SW1 closed and SW2)" SW2's state needs to be added (I'm assuming it's "open")

This was corrected. Page 9, line 12.

- P. 12: "It is important to mention that each object descriptor is normalized" Please explain why this is important?

This paragraph was modified explaining the reason of the normalisation (page 13, lines 20-24):

"It is important to mention that each object descriptor is normalised [0, 1] in order to make it invariant to size. In this way, when the BOF is calculated, from centroid to contour (starting at 3pm in anti-clockwise direction), the size is unimportant since the BOF will always be the same. Every object descriptor is also normalised to use only one x, y coordinate for their border per degree. This resulted into a 360 descriptor vector. The normalisation is also important because some objects may not have same number of coordinates on each side (asymmetric objects). This would result into an inaccurate calculation of the rotation and bigger errors in the initial assembly position."

- P. 15: after the robot grasps the peg, then it initiates the "use of its force sensing capability." However, this does not seem to be utilized during the coarse motion towards the target position. Or is it used in case of an unexpected collision. What exactly does the force-sensing measure when not performing the final insertion?

This sentence was amended (page 18, lines 12-15):

“When the peg is selected, it is grasped by the robot in readiness for assembly and the robot’s force sensing capability begins. Strictly speaking, the sensing capability is enabled; however this is not used because the Task planner has not started the PIH assembly yet.”

❖ P. 16: "when the peg is fully inserted”: How does the system know that the peg is fully inserted? What is the feedback? Visual, force? Other?

This sentence was amended (page 19, lines 19-22):

“The insertion is solely guided by force and the end-condition is when the peg’s body is inserted. It is important to note that in order to avoid a collision with the bottom of the hole, the peg was not allowed to contact. Therefore, the end-condition was considered when the peg’s body length (9mm) was fully inserted, which occurred at about 90 downward step motions.”

❖ P. 20: "90 motion steps" - how was the step size arrived at? Is it a fixed distance and why?

This sentence was amended (page 23, line 22 and page 24, lines 1-2):

“The end-condition of the assembly was set to about 9mm depth that corresponded approximately to the length of the peg’s body inside the hole. This experimentally resulted in 90 motion steps in the Z-assembly direction without any offset.”

❖ P. 21, line 16: "same task domain” How is task domain defined?

The task domain was referred to the task carried out with the same peg geometry as defined in the newly rewritten sentence (page 25, lines 4-6):

“The first insertion was completed with the circular peg using the acquired PKB (ACQ-PKB) from the same peg and hole under test, i.e. from the same task domain related to the particular cross-sectional area.”

❖ P. 23: "If the robot has to demonstrate the acquisition of the assembly skill then length of the followed insertion trajectory has to get reduced" "followed insertion trajectory" is not clearly defined. Is this the pre-planned trajectory that is being followed? Why is this a given requirement? A faster motion through the same trajectory could indicate an improvement. Or fewer "bumps" or collisions with the edges may be a better indicator.

The paragraph in reference was improved with details (page 26, lines 21-24 and page 27, lines 1-2):

“If the robot has to demonstrate the acquisition of the assembly skill then length of the followed (learned) insertion trajectory has to get reduced. Consequently, the number of alignment motions. The type of the followed path towards the insertion centre can be viewed as a robot’s performance indicator. An optimum trajectory is a peg straight motion from the offset point to the insertion centre so that there is neither angular nor translational misalignment. Similarly, the contact forces are expected to diminish.”

- P. 23: Figure 8. This figure is unclear; needs to have axes labeled. Are we looking at x, y motion? The trajectories are hard to distinguish. You could provide a sum of the length of the different trajectories in a table/graph form that may get your point across more clearly.

Figure 8 was replaced with 2 figures with explanation (figure 10 and 11) as follows:

“In figure 10 a graph illustrating the length of the trajectories during the circular PIH operation with two different offset is shown. As it can be observed, the length of the insertion trajectory is shorter when using the EKB compared when using the PKB and also the magnitude of the uncompensated remaining positional error.

Figure 10. Trajectory length and uncompensated error

Another form to observe the acquisition of the skill from the same information is to compute the Euclidian distance from final position between the hole’s centre and the peg’s centre as shown in figure 11.

Figure 11. Error decrement vs insertions”

Suggested language edits/comments

- P 2, lines 23 -24: "Considering that …” is not a full sentence. Maybe you mean "Consider that…"

Corrected sentence:

Part assembly accounts for over 50% of total production time according to various surveys [1].

- P 3, lines 6-10: Awkward sentence. Might rephrase/break up as follows: This is perhaps one of the reasons that assembly is not fully automated by industrial robots. The wide variety and the rapid changes in product characteristics require a variety of fixtures and rigs to carry out the complex manufacturing processes. This could be mitigated through enabling self-adaptive intelligent robots within this dynamic manufacturing environment, which is the main motivation in our research.

Sentenced rephrased as suggested:

“This is perhaps one of the reasons that assembly is not fully automated by industrial robots. The wide variety and the rapid changes in product characteristics require a variety of fixtures and rigs to carry out the complex manufacturing processes. This could be mitigated through enabling self-adaptive intelligent robots within this dynamic manufacturing environment, which is the main motivation in our research.”

- P3, line 16: "robot growing population" -> "robot population growth"

Corrected

- P3, lines 20-21, consider breaking awkward sentence in two: "workers. Consequently, it seems natural for robots to acquire ...(etc)"

Corrected as:

“New programming methods are needed; robots have to coexist with skilled workers. Consequently, it seems natural for robots to acquire human-like behaviours to solve human manufacturing tasks.”

- P5, line 15: "the showed" -> "shown"

This whole paragraph was rephrased following the 2nd reviewer’s comments:

“The assembly skill is evaluated based on the number of elapsed steps to complete the assembly, length of the followed assembly path and the compliant behaviour. The compliant behaviour is observed by the reduction in contact force during the insertions. If the respective force/torque pattern in the current assembly trajectory favours the assembly, then this pattern is considered for enhancing the initial knowledge base, hence forming an Enhanced Knowledge Base (EKB).”

- P5, line 16: "its expertise by the assembling" -> "its expertise by assembling"

Same as above

- P11 line 21: "objects to be assembly" -> "objects to be assembled"

Corrected in page 13, line 4.

- Various locations: "matting" -> "mating"

Corrected

- P19, line 5: "good to be included" change to either "good enough to be included" or "good and is included"

Changed to "good enough to be included" in page 22, line 12.

- P23, line 10: "….showing several circular assemblies" is awkward. Maybe use "..illustrating the trajectories followed during several circular assemblies"

Corrected as follows in page 27, line 3:

“… a graph illustrating the length of the trajectories during the circular PIH operation with two different offsets is shown.”

Answers to No. 2 Reviewer's Questions/Comments

We have attended all questions/comments regarding our previously submitted paper.
The reviewer's questions/comments are in Times New Roman font whereas
our answers/amendments are given in Calibri font.

The page and line numbers mentioned in the following sections correspond to the new reviewed version of the paper.

This paper presents an interesting idea about acquiring knowledge to improve the performance of a robot in assembly tasks. However, from the presentation point of view I found the paper lacking of important justifications and some parts badly written, making the paper hard to read.

- For example, in line 4 (page 4) it was mentioned about the development of a unified ontology in the area of manipulative tasks for industrial robots but there are no explanations about it.

A new paragraph with references was included in page 3, lines 20-24 and page 4, lines 1-7:

"We expect to contribute to the understanding of how robots can effectively use their knowledge to accomplish these tasks. The aim is to contribute towards the development of a unified ontology in the area of manipulative tasks for industrial robots considering the complexity of knowledge data present in unstructured scenarios such as the one presented in this paper. Ontology of this type should comprise a set of terms and definitions specified in a machine-readable language and shared by a given community as pointed out by E. Prestes et al. [6] who propose methodologies to encompass this knowledge to enable and facilitate data integration and information exchange between industrial robots systems. Within the Industrial Robots subgroup it is recognised as a good candidate for ontology application the kitting operation [7]. In the current research it is expected to contribute at the lower level of this application in terms of the location of the assembly parts, recognition of the mating pairs and insertion as well as the overall skill acquisition."

- Subsection 3.4 presents ART as a good candidate knowledge representation but the explanation is very poor.

The paragraph was rewritten in page 20, line 3-10:

"As mentioned in section 2.1, knowledge can be inserted in numeric form using KBANN. In this regard the Adaptive Resonance Theory (ART) is a good candidate since it is a well-established associative brain and competitive model introduced as a theory of the human cognition [10]. The advantage compared to other cognitive models is its fast learning mode for recoding input-output mapping information since it takes usually one epoch to learn. In addition, another advantage is the network's incremental learning capability that allows the network to learn newly encountered patterns incrementally without catastrophic forgetting. Both attributes are important in our application since we are interested in controlling the operation in real-time.

- Lately, in section 4 authors mentioned that Fuzzy ARTMAP was set for the tests but without explained why.

The paragraph in page 24, lines 4-11 was amended as follows:

"The FuzzyARTMAP network parameters during experiments were set for fast learning (b = 1). The values for the vigilance were selected with a high value for the network to be as selective as possible to cluster all different patterns. This is achieved by having a high vigilance level for ρmap and ρb; hence, this was the main criterion to select the vigilance and was not related to the task conditions. The value of ρa is small since this is increased internally according to the disparity between the input patterns and the previous recognition categories in the match-tracking mechanism (a detailed description of the Fuzzy ARTMAP architecture can be found in [10]. In our experiments the values for the vigilance were as follows: ρa = 0.2 (base vigilance), ρmap = 0.9 and ρb = 0.9."

- A whole paragraph of the abstract repeats in the introduction. Authors could avoid this repetition.

The paragraph in the abstract (page 2, lines 3-8) was rephrased as follows:

"The approach consists of using a MAC based on the Fuzzy ARTMAP Artificial Neural Network in conjunction with a knowledge base. The robot starts the operation having limited initial knowledge about what task it has to accomplish. During the operation, the robot learns the skill for recognising assembly parts and how to assemble them. The skill acquisition is evaluated by counting the steps to complete the assembly, length of the followed assembly path and compliant behaviour."

The paragraph in the introduction (page 5, lines 18-24 and page 6, lines1-7) was rephrased as follows:

"The proposed learning approach considers a Multimodal Assembly Controller (MAC) and an Artificial Neural Network (ANN) based on the Adaptive Resonance Theory (ART) [10] in conjunction with a knowledge base which is created by the task. At the beginning of the operation, the robot moves with limited initial knowledge, henceforth named the Primitive Knowledge Base (PKB), which is related to the mapping of contact force information and the robot's motion commands. Related to vision, it only considers clues on what to assemble. During the operation, the robot learns the skill that includes the acquisition of knowledge to recognise the part geometry using a template matching approach based on similarity to grasp the required assembly part. The assembly skill is evaluated based on the number of elapsed steps to complete the assembly, length of the followed assembly path and the compliant behaviour. The compliant behaviour is observed by the reduction in contact force during the insertions. If the respective force/torque pattern in the current assembly trajectory favours the assembly, then this pattern is considered for enhancing the initial knowledge base, hence forming an Enhanced Knowledge Base (EKB)."

On-line Knowledge Acquisition and Enhancement in Robotic Assembly Tasks

J. L. Navarro-Gonzalez, I. Lopez-Juarez, R. Rios-Cabrera, K. Ordaz-Hernández

Research Centre for Advanced Studies (CINVESTAV)
Robotics and Advanced Manufacturing Research Group
Industria Metalurgica 1062, Ramos Arizpe. Coahuila. CP 25900. Mexico.
ismael.lopez@cinvestav.edu.mx

ABSTRACT

Industrial robots are reliable machines for manufacturing tasks such as welding, panting, assembly, palletizing or kitting operations. They are traditionally programmed by an operator using a teach pendant in a point-to-point scheme with limited sensing capabilities such as industrial vision systems and force/torque sensing. The use of these sensing capabilities is associated to the particular robot controller, operative systems and programming language. Today, robots can react to environment changes specific to their task domain but are still unable to learn skills to effectively use their current knowledge. The need for such a skill in unstructured environments where knowledge can be acquired and enhanced is desirable so that robots can effectively interact in multimodal real-world scenarios.

In this article we present a Multimodal Assembly Controller (MAC) approach to embed and effectively enhance knowledge into industrial robots working in multimodal manufacturing scenarios such as assembly during kitting operations with varying shapes and tolerances. During learning, the robot uses its vision and force capabilities resembling a human operator carrying out the same operation. The approach consists of using a MAC based on the Fuzzy ARTMAP Artificial Neural Network in conjunction with a knowledge base. The robot starts the operation having limited initial knowledge about what task it has to accomplish. During the operation, the robot learns the skill for recognising assembly parts and how to assemble them. The skill acquisition is evaluated by counting the steps to complete the assembly, length of the followed assembly path and compliant behaviour. The performance improves with time so that the robot becomes an expert demonstrated by the assembly of a kit with different part geometries. The kit is unknown by the robot at the beginning of the operation; therefore, the kit type, location and orientation are unknown as well as the parts to be assembled since they are randomly fed by a conveyor belt.

Keywords: Robotic Assembly, Artificial Neural Networks, Machine Vision, Kitting process.

1. INTRODUCTION

Part assembly accounts for over 50% of total production time according to various surveys [1]. It is therefore easy to understand why this assembly process contributes to 20% of the unit production cost as stated in [2]. Furthermore, despite the technological progress in automation, it is recognised that approximately 50% of all labour work in the mechanical and electrical industries is involved in work closely related to assembly, handling and fitting processes [3]. Nevertheless, this high percentage has a direct impact on total production costs. A reason that assembly plays a crucial role in manufacturing may be due to the diverse variety of products in the market, whose production is driven by consumer’s choice. Nowadays, a wide variety of product is needed so a rapid change in product design and therefore assembly planning is required. This is perhaps one of the reasons that assembly is not fully automated by industrial robots. The wide variety and the rapid changes in product characteristics require a variety of fixtures and rigs to carry out the complex manufacturing processes. This could be mitigated through enabling self-adaptive intelligent robots within this dynamic manufacturing environment, which is the main motivation in our research.

There are multiple manufacturing areas where industrial robots are currently used like in the automotive sector, which is undoubtedly where most industrial robots are used today. In terms of use; tasks such as handling, welding and assembly are the most demanding operations for industrial robots around the world according to the International Federation of Robotics IFR [4]. Statistics show that the trend in robot population growth is expected to continue and has a direct impact also on increasing the number of human skilled jobs related to robotics [5]. This promising scenario in conjunction with the acceptance of skilled workers' interaction with robots will undoubtedly bring new paradigms to solve related to human-robot collaboration. New programming methods are needed; robots have to coexist with skilled workers. Consequently, it seems natural for robots to acquire human-like behaviours to solve human manufacturing tasks. Robots have to adapt to new conditions in unstructured environments and to be self-adaptive making an effective use of their intelligence founded in a solid knowledge representation to accomplish a manufacturing task. Specifically in our research, we are interested in studying the *acquisition, representation and enhancement/refinement* of this knowledge during complex manipulative robot tasks such as kitting and assembly operations. We expect to contribute to the understanding of how robots can effectively use their knowledge to accomplish these tasks. The aim is to contribute towards the development of a unified ontology in the area of manipulative tasks for industrial robots considering the complexity of knowledge data present in unstructured scenarios such as the one presented in this paper. Ontology of this type should comprise a set of terms and definitions specified in a machine-readable language and shared by a given community as pointed out by E. Prestes et al. [6] who propose methodologies to encompass this knowledge to enable and facilitate data integration and information exchange between industrial robots systems. Within the Industrial Robots subgroup the kitting operation is recognised as a good candidate for ontology application [7]. In the current research it is expected to contribute at the lower level of the ontology application in terms of the location of the assembly parts, recognition of the mating pairs and insertion as well as the overall skill acquisition.

The scope of the research presented here needs to define some aspects as to build up a sounded understanding. The task domain of our investigation refers to the assembly of parts having two well defined tasks, the peg in hole task and the kitting task. In this investigation the kitting task is composed by more than one assembly operation. In general sense, the work domain is in the assembly of parts, which is the action of putting together manufactured parts to make a completed product or subassembly. In particular, in this paper we will deal with a special type of assembly composed by a peg and its counterpart named hole and referred to as the Peg-In-Hole (PIH) task. The kitting task can then be understood as the practice of assembling components (pegs in our case) in predetermined quantities that are placed together in specific container (kit). The term of assembly parts, assembly components and mating pairs are used interchangeably within this paper.

The investigation is related to intelligent robots, specifically autonomous industrial robots. An autonomous robot can solve complex tasks by itself and has intentions of its own to produce changes in the world as pointed out in [6]. Accordingly, a robotic agent should emerge from the interaction between the Real-World and the robot manipulator itself. This agent should demonstrate its ‘‘intelligence’’ by using new knowledge, refine and apply it autonomously during skill learning showing the required skill during real world tasks as demonstrated by Lopez-Juarez et al. [8]. The research is founded on previous approach by Lopez-Juarez in terms of creating intelligent robotic agents for assembly using force sensing in conjunction with an image processing method called the Boundary Object Function (BOF) to describe invariantly an object using object's features as initially presented by Peña-Cabrera et al. [9]. Although the work was centred on the PIH operation, the method can easily extend the robot's capability in more complex processes like for instance, the kitting process under high uncertainty.

In real-world scenarios, robots deal with high uncertainty in the environment due to modelling, sensing and control errors. Uncertainties come from a wide variety of sources such as robot positioning errors, gear backlash, arm deflection, ageing of mechanisms and disturbances. Controlling all the above aspects would certainly be a very difficult task; therefore a simpler approach based on reactive learning control is preferred. By using force control the overall effect of the contact force between the environment (assembly parts) and the manipulator are considered as a whole. In order to cope with these situations, we propose the robot to learn the manipulative tasks using multimodal information, i.e. vision templates and force sensing.

The proposed learning approach considers a Multimodal Assembly Controller (MAC) and an Artificial Neural Network (ANN) based on the Adaptive Resonance Theory (ART) [10] in conjunction with a knowledge base which is created by the task. At the beginning of the operation, the robot moves with limited initial knowledge, henceforth named the Primitive Knowledge Base (PKB), which is related to the mapping of contact force information and the robot's motion direction that produced such force condition. Related to vision, it only considers clues on what to assemble. During the operation, the robot learns the skill that includes the acquisition of knowledge to recognise the mating part geometry using a template matching approach based on similarity to grasp the required assembly part and insert in the kit. The assembly skill is evaluated based on the number of elapsed steps to complete the assembly, length of the followed assembly path and the compliant behaviour. The compliant behaviour is observed by the reduction in contact force during the insertions. If the respective force/torque pattern in the current assembly trajectory favours the assembly, then this pattern is considered for enhancing the initial knowledge base, hence forming an Enhanced Knowledge Base (EKB).

The article is organised as follows. After this introduction, section 2 presents the related work and original contribution. In section 3, the methodology that includes the description of the test bed, image processing, task description and knowledge acquisition is presented while the results are explained in section 4. Finally, section 5 provides the conclusions and further work.

2. RELATED WORK

Uncertainties due to manufacturing tolerances, positioning, sensing and control errors make it difficult to perform the assembly. Compliant motion can be applied using passive devices such as the Remote Centre Compliance (RCC) introduced by Whitney [11]. Other alternative is to use active compliance, which actually modifies the position of the manipulated component as a response to constraint contact forces. A detailed analysis of active compliance can be found in works by Mason [12], De Schutter [13] and more recently, using contact based assembly for real industrial assemblies and kitting processes [14]. Alternative solutions using active compliance are based also on connectionist approaches that rely on the information given during the network training stage which implicitly considers all insertion parameters. The use of connectionist models in robot control to solve a canonical assembly task like the PIH operation under uncertainty has been dealt in a number of publications along the years. The reinforcement algorithm implemented by V. Gullapalli demonstrated the ability to learn circular and square peg insertions. However, the network was unable to generalise over different geometries [15]. Cervera used SOM networks and a Zebra robot (same as used by Gullapalli) developing similar insertions improving the autonomy of the system by obviating the knowledge of the part location and used only relative motions [16]. Other interesting approaches have also been used for skill acquisition within the framework of robot programming by demonstration (PbD) that considers the characteristics of human generated data. Work carried out by Kaiser and Dillman [17] shows that skills for assembly can be acquired through human demonstration. The training data is first pre-processed, inconsistent data pairs are removed and a smoothing algorithm is applied.

The PIH operation in unknown conditions is not a solved problem and has recently attracted a lot of attention. Fuzzy logic has been employed in conjunction with Support Vector Machine algorithms to create a knowledge base for assembling cylindrical pegs [18]. Other novel approaches include the use of a set of mirrors in order to determine the misalignment between mating pairs useful even during self-occlusion by the robot [19]. Vibration at the peg's tip has been employed following an elliptical trajectory using a piezoelectric vibrator with frequencies in excess of 7 kHz [20]. High stiffness provided by a parallel actuator at the robot's wrist in conjunction with Force/Torque (F/T) sensing has also been used so that stable contact with stiff environments can be obtained [21]. Learning by demonstration has also been tested using a dextrous hand and tele-operation [22]. If the geometry of the mating pairs is known, the use of CAD data is possible to simulate the generation of F/T maps and 3D poses known as particles, so that the particle's filter theory can be applied [23].

The use of the referred compliant vision using eye-in-hand camera in conjunction with eye-at-hand camera has also been employed for fast alignment of the PIH without force feedback [24]. Some more complex sensing systems have also been tested adding tactile sensing to dextrous hands, so that multimodality from vision, tactile and force can be exploited [25].

2.1 Original Work

Two different approaches have emerged to represent knowledge. It can be either symbolic as employed in Artificial Intelligence algorithms or numeric (also referred to as subsymbolic) as processed by Artificial Neural Networks. The knowledge embodiment is different however in essence the apparent dichotomy between both approaches is more perceived than real as stated by Honavar [26]. Both approaches to modelling cognition and engineering in intelligent systems can be applied for integrating neural and symbolic processes.

In this paper, it is proposed for the main operation (PIH) to use a numeric knowledge representation using ANN, which can be refined by using on-line training examples (contact conditions during PIH operations). The knowledge is inserted into the network and subsequently refined by ANN training resulting in a Knowledge Based Artificial Neural Networks (KBANN) as pointed by Towell and Shavlik [27]. The core idea here is to create a Multimodal Assembly Controller (MAC) formed by a learning and recognition module in conjunction with a Knowledge Base. The proposal also includes a knowledge refinement module and world effectors so that information from the real-world is considered. The purpose is to train the KBANN under real-world situations providing the robot with the capability of recognising cues or primitive descriptors during early stages of learning, so that initial conditions can be started. During knowledge refinement, and by giving more examples, this knowledge is expected to be enhanced/refined. The Multimodal Assembly Controller is illustrated in figure 1.

Figure 1. Multimodal Assembly Controller (MAC)

The learning and recognition module –KBANN- is the heart of the MAC, which includes three additional modules: the PKB, the World Effector and the Knowledge Refinement module. The PKB stores initial information about the environment provided by a human expert. This information is only used during the first stage of training. In this stage the switch SW_1 will be open and the switch SW_2 closed since the initial training is made only using the PKB. After passing this initial state, the KBANN will predict the next action based on the current input from the sensor (SW_1 closed and SW_2 open). Knowledge can be refined according to a quality criterion in the Knowledge Refinement module. Should the corresponding force/torque pattern meets the criterion, then the pattern is allowed to be included in the EKB, then SW_2 is closed for on-line retraining. Otherwise, the KBANN is in a recall stage until new and useful patterns are encountered.

In the particular case of the kitting operation, the world is affected either by visual or force information, depending if the peg is in free motion or constrained motion (while in contact). Hence, a particular feedback is decided by the Integrator (Task Planner) as indicated in figure 2 that shows the proposed MAC's Knowledge Refinement Module.

Figure 2. MAC's Knowledge Refinement Module

The Integrator (Task Planner) acts differently according to the stage of the assembly process and will be described in detail in the next section. But, basically depends on the type of motion being carried out (in contact during constraint motion or free motion without any contact). It is important to observe that the Integrator has a bi-directional data flow from/to the world effectors. In our application during free motion the Task Planner requires information about the end-effector position as well as it needs to send information to the motor drives (world effectors) to reposition the arm. During constraint motion, the Task Planner receives information about the end-effector position as well as force sensing information via the F/T sensor located at the robot's wrist.

The Task Planner is in charge of selecting the appropriate information and patterns depending on the involved tasks. During free motion, information about the current position of the arm and the object's template is needed for instance, when approaching the peg to the hole in readiness for insertion. During constraint motion, the Task Planner receives robot's end-effector position as well as the force/torque pattern and selects if the pattern will be *learned* or *recalled*. In the following sections, we will describe how the vision system works using template matching for task understanding as well the pattern selection mechanism for deciding if a certain contact force should be learned or *forgotten*.

3. METHODOLOGY

3.1. Test bed

The robotic test bed is formed basically by a six DOF KUKA KR16 industrial robot, KRC2 robot controller, KUKA control panel (KCP), master computer, JR3 F/T sensor attached to the robot wrist, and an eye-in-hand Basler 641fc camera as it is shown in figure 3.

Figure 3. Robot test bed

The KRC2 controller houses the components that control and power the robot arm. The master computer hosts the DSP-based JR3 F/T sensor card that communicates and power the sensor. The master computer also communicates with the robot controller via RS232C standard. Data sent from the master computer to the KRC2 controller is transmitted using the following protocol:

<CODE> NUL <DIST> NUL <VEL>

where:

NUL is a byte containing the null ASCII character.

<CODE> is a byte containing the corresponding Command Code (16 possible motion direction commands: +x, -x, +y, -y, +z, -z, +x +y, +x -y, -x -y, -x +y, +Rx, -Rx, +Ry, -Ry, +Rz, -Rz) and 9 control commands that includes do nothing, go to home, world coordinates, tool coordinates, joint coordinates, base coordinates, end communication, open gripper and close gripper.

<DIST> is a byte containing a distance value in the direction indicated by <CODE> and given in tenths of mm.

<VEL> is a byte containing a velocity value given in mm/s.

The information packet is sent to the KRC2 controller using Xon/Xoff flow control. A resident monitor program in KPL (KUKA's language) continuously detects any requested arm motion from the master computer. Depending on the value given in the CODE byte, the robot arm will move in world coordinates during gross motion (e.g. from the pickup zone to the assembly zone) or in tool coordinates at the lower level, moving the arm incrementally while in fine motion during assembly.

The F/T sampling rate by the DSP board was set to 8 kHz (0.125 ms) and the values were read from the sensor every 100 ms by the main program. However arm positions during constraint motion were updated after a delay of 600 ms in order to read the sensor, testing/training of the KBANN on-line and repositioning the arm. The Task Planner and the vision system to recognise the part to be picked up reside in the master computer.

Contact Forces are measured at the JR3 F/T sensor which is attached to the robot's flange edge using an adapter plate. The origin for the F/T coordinate frame is located in the centre of the sensor unit; however, in our experiments the origin was translated and rotated so that it was located at the peg's tip. Programming for the MAC was developed using the C++ builder compiler. The designed interface is shown in figure 4 where basically the F/T readings and the camera workspace are displayed.

Figure 4. User interface

The F/T sensor provides a six value vector corresponding to force/torque data F_{JR3} = (fx, fy, fz, mx, my, mz). This vector is normalized within the range [0,1].

3.2. Image processing and template matching

The robot must be able to identify the target objects to be assembled even if they are not stored in a previously acquired knowledge base. It was implemented a real-time image processing template-based approach, as showed in figure 5. A temporal memory setting was used in which all acquired female components were saved. It was decided to use a template based approach, because it does not need off-line costly training, and it is able to run in real-time. In our setting, we are able to run at 16.6 fps. The first step the robot must accomplish is to identify the female components where it is going to assemble the different male components (Assembly zone). These objects are found in unknown precise location and orientation on the assembly table.

In the Picking up zone on the belt conveyor, the pegs are located standing up and randomly oriented. The next step is to process the peg's image template in order to match with the just found female shapes. The Task planner is in charge of identifying, if the parts were already used, or if they are new patterns. In order to do that, we balanced the image to make it robust to illumination changes. Then we made a contour analysis of the objects. For each object, the centroid and the BOF are calculated as explained in [9].

It is important to mention that each object descriptor is normalised [0, 1] in order to make it invariant to size. In this way, when the BOF is calculated, from centroid to contour (starting at 3pm in anti-clockwise direction), the size is unimportant since the BOF will always be the same. Every object descriptor is also normalised to use only one x, y coordinate for their border per degree. This resulted into a 360 descriptor vector. The normalisation is also important because some objects may not have same number of coordinates on each side (asymmetric objects). This would result into an inaccurate calculation of the rotation and bigger errors in the initial assembly position.

Figure 5. Template matching using an eye-in-hand camera

For the matching process, we propose to minimize the Euclidian distance of the normalized BOF as in equation 1.

$$H\left(BOF_{hyp}, BOF_{list:a=1,\dots,n}\right) = \underset{rot,a}{\arg\min} \sum_{i=1}^{360} \left|BOF_{a,i} - BOF_{hyp,i+rot}\right| \qquad (1)$$

Where: BOF_{list} represents the objects in temporal memory in the Assembly zone on the assembly table. BOF_{hyp} is the current hypothesis taken from the Picking up zone on the conveyor belt and *rot* is the rotation angle in degrees of the normalized boundary object function that can be from 0^o to 360^o. While finding the best match α, we also calculate the rotation needed to match each template. Once the most similar object is found, a threshold value *T* is set to eliminate noisy objects or objects that do not appear in the assembly area. If the current hypothesis is not approved then, it is left for a second round when a new assembly object comes.

$$H\left(BOF_{hyp}, BOF_{list:a=1,\dots,n}\right) > T \qquad (2)$$

3.3. The Kitting process and PIH operation

The unstructured environment has been built so that the robot has limited knowledge of the kitting and PIH operation. The only *process knowledge* that is available to the robot is the following:

1. Assembly direction (-Z).
2. End-condition (for the kitting and PIH operation).
3. Roughly, the location of the male component.
4. Roughly, the insertion area where the female component is located.
5. Primitive Knowledge Base (PKB) composed by force contact conditions and arm motion commands provided by the user.

In other words, the most demanding tasks in this unstructured environment are twofold:

1. The robot ignores the type of operation it has to complete with the kit, since no information about the type of geometry is given. The robot has to discover the geometry and localise the mating pair by searching onto the belt conveyor.

2. The robot starts the operation with minimum information (PKB) building up its knowledge according to the assembly geometry.

The kitting process is carried out by the Task Planner in three stages.

Stage 1. Task understanding

Stage 2. Picking up and motion in free space

Stage 3. PIH assembly

3.3.1 Task Planner

The Task Planner forms part of the refinement module. It integrates the processing of the force sensing signal and the image templates. It is also in charge of managing the knowledge acquisition process and the conditions to start the PIH operations and the kit formation as it is shown in the flow chart provided in figure 6.

Figure 6. Task Planner

The operation of the system starts by handling different initial settings such as moving the robot arm to the global home position, FuzzyARTMAP initial network parameters, force/torque sensor calibration, camera calibration, PKB file initialisation, etc. After this stage is completed, the mapping between contact force patterns and the robot motion directions is carried by the user initialising the PKB. This procedure is accomplished by the user only once and consists of moving the grasped peg against the inner edge of the hole in the four cardinal directions so that the current force/torque-motion direction is recorded. Once the PKB is formed, the learning is guided by the Task Planner that indicates the operation sequence. At the beginning, the task is understood by matching the female component image template against the available peg's template. With this information the operation is structured in such a way that the Task Planner determines the number and type of pegs needed to complete the kit assembly. The order in which the pegs are grasped and inserted in the kit is random since the Task Planner does not know in advance neither the type/number of the needed pegs nor the order in which the pegs will be found on the conveyor belt. The kitting process operation is carried out in three stages by the Task Planner as it is described in detail below.

Task understanding

This first stage is accomplished solely by the vision system. At this time the robot uses only its vision capability since force sensing is not needed. During this initial stage, the robot is aimed to move from the global home position to a predefined assembly home position above the assembly table as showed in figure 7a. At this point the robot recognises the assembly kit so that it learns the number of parts needed in the kit as well as their geometry in the BOF_{list} explained above. The robot learns how many mating pairs are needed to form the kit. In our experiments the kits can be formed by up to four randomly oriented female components. Figure 7b shows some examples of the possible combinations. As it can be observed, the combination considers different location for the female component as well as a random orientation of the kit. In this stage, the robot computes the number of assemblies it has to carry out (1 to 4 PIH operations) and establishes the requirements for the task (i.e. the number and type of male components). The order is irrelevant, since the end-condition of the kitting process is when all pegs are inserted in the kit.

Figure 7. (a) Test bed. (b) Assembly kits.

Picking up and motion in free space

The parts to be assembled are picked up from the conveyor belt in random order since the robot will aim for the closer male component which mates to the female counterpart in the kit. Here, the closer part is the part that has the minimum Euclidean distance measured form the centroid of the male component to the origin in the world frame coordinates. For grasping purposes, the Z coordinate of each male part is the same and every component is randomly located in the x-y plane. In the case that a particular male component geometry is not needed by the kit, the robot obviates it and continues looking for the following part on the belt conveyor until it finds the corresponding mating pair. The end-condition for this stage is accomplished when the robot either finds the last component for the kit or all parts in the conveyor belt were analysed. When the peg is selected, it is grasped by the robot in readiness for assembly and the robot's force sensing capability begins. Strictly speaking, the sensing capability is enabled; however this is not used because the Task planner has not started the PIH assembly yet. The robot's end effector then moves in free space above to the recalled female component's location in the kit (Assembly home position). The end-condition in this stage is when the male component is located above the female counterpart and both BOF's perfectly match; that is, when both patterns are recognised as belonging to the same cluster in the Task Planner. The picking up stage is shown in figure 8.

Figure 8. The pickup stage.

PIH Assembly

This stage properly starts when both mating pairs are roughly aligned vertical, then the end-effector moves in the assembly -Z direction. At this time the robot is solely guided by the visual feedback until the peg contacts the female component. While moving downwards in the assembly direction, the peg is rotated around the Z axis so that the BOF's peg and the BOF's hole match as it is illustrated in figure 9.

Figure 9. Matching the BOF's

The peg's alignment with the hole via the BOF pattern from each object is depicted in the lower part of figure 9. In practice, the end-effector was moved downwards so that the peg's BOF and the hole's BOF were very similar.

Occlusions are unimportant since the motion is directed in the -Z direction until the peg's tip impacts with its counterpart. At this time the Task Planner starts using the force feedback information. The insertion is solely guided by force and the end-condition is when the peg's body is inserted. It is important to note that in order to avoid a collision with the bottom of the hole, the peg was not allowed to contact. Therefore, the end-condition was considered when the peg's body length (9mm) was fully inserted, which occurred at about 90 downward step motions.

3.4 Knowledge Representation

As mentioned in section 2.1, knowledge can be inserted in numeric form using KBANN. In this regard, the Adaptive Resonance Theory (ART) is a good candidate since it is a well-established associative brain and competitive model introduced as a theory of the human cognition [10]. The advantage compared to other cognitive models is its fast learning mode for recoding input-output mapping information since it takes usually one epoch to learn. In addition, the network's incremental learning capability allowed the network to learn newly encountered patterns incrementally without catastrophic forgetting. Both attributes were important in our application since we are interested in controlling the operation in real-time.

The theory has evolved in a series of real-time architectures for unsupervised learning, the ART-1 algorithm for binary input patterns [28]. Supervised learning is also possible through ARTMAP [29] that uses two ART-1 modules that can be trained to learn the correspondence between input patterns and desired output classes. Different model variations have been developed to date based on the original ART-1 algorithm, ART-2, ART-2a, ART-3, Gaussian ART, EMAP, ViewNET, Fusion ARTMAP, LaminART just to name a few.

3.5 Acquisition of the Primitive Knowledge Base (ACK-PKB)

The autonomously acquired PKB henceforth referred to as the ACK-PKB to distinguish it from the original PKB taught by the user, consists of showing the robot how to react to individual components of the F/T vector during operations while the arm is in constraint motion. In our experiments the ACK-PKB was automatically obtained by the Task Planner during the compliant motion. The PKB initiates with only five patterns (assembly direction and four cardinal directions) and depending on the contact states this knowledge base was incremented creating an Enhanced Knowledge Base (EKB).

3.6 Knowledge Enhancement/Refinement

There are potential overtraining problems associated with learning patterns on-line during fine motion, which are solved by the MAC via its Task Planner as indicated in figure 6. The robot should continue to move in the insertion direction if and only if a minimum force value has been reached. In this situation, on-line learning is started to allow the acquisition and learning of the pattern-action pair that produced such contact state and favoured the assembly. In the event of continuous learning after having reached this minimum force value, the performance of the MAC may decay. This situation is similar to what is known as overtraining, overfitting or overlearning in ANNs. At this point, the learning should be stopped because if the robot learns other patterns under the above-mentioned circumstances, eventually the minimum force value will be different leading to wrong motions. The same applies to the condition when the end-effector meets a force higher than the force limit. There should not be any further learning in this situation since learning a higher force value can damage the sensor. The above situations can be resumed in three fundamental questions:

1. What is a good motion?
2. Which motions should or should not be learned?
3. Which motions should be forgotten?

Having an assembly system that is guided by compliant motion, the criterion to decide whether the motion was good enough to be learnt is based on the following heuristic expression:

$$\mathrm{F}_{JR3_{after}} < 0.1\ (\mathrm{F}_{JR3_{\mathrm{before}}}) \qquad (3)$$

where $F_{JR3fter}$ and $F_{JR3before}$ are merit figures calculated before and after a corrective motion is applied and are computed using the following equation:

$$\mathrm{F}_{JR3} = \sqrt{fx^2 + fy^2 + fz^2 + S(mx^2 + my^2 + mz^2)} \qquad (4)$$

If expression (3) is met then that pattern-action will be considered good enough to be included in the Enhanced Knowledge Base (EKB). It is important to mention that in order to compare the robot's behaviour using different knowledge bases, the units have to be the same (e.g. N and N·dm). To avoid the comparison of inconsistent data, a scale factor S has been included in Eq. (4) to allow the use of different units or size components. In our experiments, the scale factor was selected to be equal to 1. Should different units or size components be used, then the scale factor S would have to be modified.

For the sake of clarity, there is a distinction about the type of acquired knowledge that can be simply *enhanced* or alternatively *refined*. We refer to *knowledge refinement* when it results from expanding the EKB during the insertion of a peg with different cross-sectional area and *knowledge enhancement* when this knowledge is expanded using the same cross-sectional area.

3.7 Learning, Recall, Error recovery and Forgetting.

There are three areas for knowledge generation as indicated in the following expressions:

$$0 < F_{current} \leq 0.1(F_{before}) \rightarrow \ learning \qquad (5)$$

$$0.1(F_{before}) < F_{current} \leq F_{limit} \rightarrow \ Recall \qquad (6)$$

$$F_{current} > F_{limit} \rightarrow \ Error\ recovery \qquad (7)$$

Learning only takes place if current values during compliant motion ($F_{current}$) are equal or lower than $0.1F_{before}$. Values higher than $0.1F_{before}$ but equal or lower than the limit values are used only for network recall – testing mode – and learning is not allowed. The third area, error recovery, is a situation where $F_{current} > F_{limit}$. In this case the corresponding pattern that produced this situation within the EKB is deleted.

4. RESULTS

Several tests were carried out to assess the performance of the assembly controller using pegs and female components made from thermoplastic (ABS). The diameter of the circular peg was 25 mm and the side of the square peg was also 25 mm. The dimensions of the radiused-square were a side 25 mm with one corner rounded to a radius of 12.5 mm. During operations, clearances between pegs and mating pairs were 0.3 mm. The end-condition of the assembly was set to about 9mm depth that corresponded approximately to the length of the peg's body inside the hole. This experimentally resulted in 90 motion steps in the Z- assembly direction without any offset.

The FuzzyARTMAP network parameters during experiments were set for fast learning ($\beta = 1$). The values for the vigilance were selected with a high value for the network to be as selective as possible to cluster all different patterns. This is achieved by having a high vigilance level for ρ_{map} and ρ_b; hence, this was the main criterion to select the vigilance and was not related to the task conditions. The value of ρ_a is small since this is increased internally according to the disparity between the input patterns and the previous recognition categories in the match-tracking mechanism (a detailed description of the Fuzzy ARTMAP architecture can be found in [10]. In our experiments the values for the vigilance were as follows: $\rho_a = 0.2$ (base vigilance), $\rho_{map} = 0.9$ and $\rho_b = 0.9$.

4.1 Knowledge Acquisition

Several assemblies were carried out with different offsets at the start of the operation. Using the template matching algorithm, the robot was able to pick up the peg and get to the assembly position in readiness for assembly. The offset position was varied with an error from 0mm to 2mm from the centroid of the female component. Although, the robot was always able to complete the insertions, it was noticed that the assessment was difficult due to the random varying starting points. In order to have the same baseline for comparison purposes, it was decided to initialise the operation with defined offsets by positioning the robot's end-effector manually. The given offset and assembly results are shown in Table 1.

Table 1. EKB acquisition using the same part geometry

The first insertion was completed with the circular peg using the acquired PKB (ACQ-PKB) from the same peg and hole under test, i.e. from the same task domain related to the particular cross-sectional area. As it can be observed in the first assembly, the given offset was -0.5 in X-axis and 0 in Y-axis. After the completion of the assembly, the MAC learned 1 new pattern making a total of 3 alignment motions. This resulted in 3 movements in total apart from the 90 expected movements in the Z- direction. Insertions 2 to 4 have a similar analysis. A second test was carried out to verify the usefulness of the EKB just created. It resulted in learning 3 more patterns during insertions 5 to 8.

4.2 Knowledge Generalisation of the EKB

This generalisation test consisted in using the same EKB with different part geometry in order to observe the capability of the MAC to further refine and enhance its knowledge. As explained earlier, *knowledge refinement* results from expanding the EKB during the insertion of a peg with different cross-sectional area and *knowledge enhancement* occurs when using the same cross-sectional area. The results are shown in Table 2.

Table 2. EKB acquisition using different part geometry

At this point the circular female component was interchanged by a square component (insertions 9–12). The assembly was possible and the number of learned patterns diminished. The MAC did not learn any additional pattern indicating that the network had already acquired the necessary knowledge and used this information effectively (insertions 11-12). Some patterns were forgotten, 1 pattern in insertion number 10 and 4 patterns in insertion number 11. The *forget state* was started should the current pattern produced a high F/T pattern beyond the F_{limit} in eq. (7).

The modification of the EKB during this set of insertions (9-12) corresponds to a *knowledge refinement* process since it evolves from different part geometry (i.e. new knowledge for the insertion of the squared peg using the EKB from the circular peg).

The test continued with the assembly of the same squared peg. In this case the knowledge evolved from the same geometry and we referred to as *knowledge enhancement* (insertions 13-16). The following test was to observe the robustness of the MAC by changing the part geometry again assembling the Radiused square peg (insertions 17-20). It was observed that none of the patterns were learnt; however, the MAC still had to forget patterns since the new conditions with different offset triggered a further knowledge refinement.

4.3 Skill Acquisition

The robot skill acquisition for assembly can be analysed observing the followed trajectory during assembly and the level of compliant motion. If the robot has to demonstrate the acquisition of the assembly skill then length of the followed (learned) insertion trajectory has to get reduced. Consequently, the number of alignment motions. The type of the followed path towards the insertion centre can be viewed as a robot's performance indicator. An optimum trajectory is a peg straight motion from the offset point to the insertion centre so that there is neither angular nor translational misalignment. Similarly, the contact forces are expected to diminish. In figure 10 a graph illustrating the length of the trajectories during the circular PIH operation with two different offsets is shown. As it can be observed, the length of the insertion trajectory is shorter when using the EKB compared when using the PKB and also the magnitude of the uncompensated remaining positional error.

Figure 10. Length of the assembly trajectory and uncompensated error

Another form to observe the acquisition of the skill from the same information is to compute the Euclidian distance from final position between the hole's centre and the peg's centre as shown in figure 11.

Figure 11. Error decrement vs insertions

A quality measure that helps to assess the robot's dexterity is the force and moment traces during assembly and while in constraint motion. This quality measure can be obtained from the continuous monitoring of the force and torque. The F/T data from the first and second assembly of the squared peg are shown in figure 12. It can be observed that the values from the F/T data are lower in the second insertion.

Figure 12. Compliant motion during squared peg assembly

5. CONCLUSIONS AND FUTURE WORK

Results from our experiments demonstrate that industrial manipulators can learn manipulative skills on-line using contact force information. The a priori information from the environment was minimal since the robot only knew the assembly direction (-Z direction) in conjunction with the associated force/torque information. This PKB was solely used for starting the knowledge acquisition provided by the user to start the operations.

It was demonstrated during the experiments that knowledge can be refined when other mating pair with different cross-sectional area was assembled using the previous knowledge without the need to acquire another PKB. Once this knowledge was refined, it is possible that it can become enhanced while assembling the same part geometry.

The robot learned a new assembly and improved its skills from experience. This was observed in three aspects. 1) The reduced number of learned patterns, 2) the compliant forces also diminished during assembly when using the EKB, and 3) the followed trajectories were shorter, hence also the overall assembly time considering that the robot's speed was constant during fine motion.

It was demonstrated that the Multimodal Assembly Controller (MAC) can be used in non-structured environments for the kitting process with positional uncertainties for assembling different type of pegs in different type of kits.

It was readily recognised the need for an appropriate mechanism for knowledge management. In the current state, when the MAC decides to delete (forget) a pattern, this is permanently erased. It is likely that a similar pattern to the forgotten pattern is required and it has to be relearned again. Further work has been envisaged in this direction so that a knowledge management mechanism is included in the MAC to improve its effectiveness.

ACKNOWLEDGEMENTS

Authors want to acknowledge CONACyT for Mr. Navarro-Gonzalez’s scholarship to pursue his doctoral studies.

6. REFERENCES

[1] Nof SY, Wilhelm WE, Warnecke HJ. Industrial Assembly. 1st ed. 1997, XII, 500p.

[2] Martin Vega LA, Brown HK, Shaw WH, Sanders TJ. Industrial perspectives on research needs and opportunities in manufacturing assembly. Journal of Manufacturing Systems 1995, 14 (l) 666 45–58.

[3] Swift K, Booker J. Process Selection from Design to Manufacture. Burlington: Elsevier, 2003.

[4] World Robotics 2013. Industrial Robots. International Federation of Robotics. 2013.

[5] Gorle P, Clive A. Positive Impact of Industrial Robots on Employment. International Federation of Robotics. METRA MARTECH Limited, 2013.

[6] Prestes E, Carbonera JL, Rama Fiorini S, M Jorge VA, Abel M, Madhavan R, Schlenoff C. Towards a core ontology for robotics and automation. Robotics and Autonomous Systems. Volume 61 Issue 11, November, 2013. Pp 1193-1204.

[7] Schlenoff C, Prestes E, Madhavan R, Goncalves P, Li H, Balakirsky S, Miguelanez E. An IEEE standard ontology for robotics and automation. In Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference. Pp. 1337-1342. IEEE.

[8] Lopez Juarez I, Corona Castuera J, Peña Cabrera M, Ordaz Hernandez K. On the design of intelligent robotic agents for assembly. Information Sciences, Volume 171, Issue 4, 13 May 2005, Pp 377-402.

[9] Peña Cabrera M, Lopez Juarez I, Rios Cabrera R, Corona Castuera J. Machine vision approach for robotic assembly. Assembly Automation, 25(3), 2005. Pp. 204-216.

[10] Carpenter GA, Grossberg S, Markuzon N, Reynolds JH, Rosen DB. Fuzzy ARTMAP: A neural network architecture for incremental learning of analog multidimensional maps, IEEE Transactions on Neural Networks. 1992 3 (5), Pp. 698–713.

[11] Whitney DE, Nevis J. What is the remote center compliance (RCC) and what can it do? Proc. of the 9th International Symposium on Industrial Robots 1979, Washington, DC, USA pp. 135–152.

[12] Mason MT, Compliant motion. In: Robot Motion (M. Brady et al., eds.) MIT Press, Cambridge, 1983.

[13] J. De Schutter, Brussel HV. Compliant robot motion I, a formalism for specifying compliant motion tasks. Int. J. Robot. Res. 1988, 7(4), Pp. 3–17.

[14] Sabine S, Rüdiger D. Experience-based optimization of universal manipulation strategies for industrial assembly tasks. Robotics and Autonomous Systems. 2011, 59(11), Pp. 882-898

[15] Gullapalli V, Franklin JA, Benbrahim H. Acquiring Robot Skills via Reinforcement Learning. IEEE Control Syst. 1994, Pp.13–24.

[16] Cervera E, Del Pobil AP. Sensor-based learning for practical planning of fine motions in robotics. Information Sciences, 2002, 145(1-2), Pp. 147-168

[17] Kaiser M, Dillman MR. Building elementary robot skills from human demonstration. Proceedings of the IEEE International Conference on Robotics and Automation. Minneapolis, Minnesota (April 1996), Pp. 2700–2705.

[18] Zivana J, Petar P, Vladimir M, Miroslav P. Fuzzy inference mechanism for recognition of contact states in intelligent robotic assembly. Journal of Intelligent Manufacturing, 2012.

[19] Kim JY. Novel visual sensing systems for overcoming occlusion in robotic assembly. Assembly Automation 2005, 25(1), Pp. 21-29.

[20] Sadauskas E, Bakšys B. Alignment of the parts using high frequency vibrations. Mechanika,2013, 19(2).

[21] Osypiuk R, Kröger T. Parallel Stiffness Actuators with Six Degrees of Freedom for Efficient Force/Torque Control Applications, 2011. In Robotic Systems for Handling and Assembly. Springer Berlin Heidelberg, Pp. 275-291.

[22] Savarimuthu TR, Liljekrans D, Ellekilde LP, Ude A, Nemec B, Kruger N. Analysis of human Peg-in-Hole Executions in a Robotic Embodiment using uncertain Grasps. In Robot Motion and Control (RoMoCo) July, 2013 9th Workshop IEEE. Pp. 233-239.

[23] Thomas U, Molkenstruck S, Iser R, Wahl F M. Multi sensor fusion in robot assembly using particle filters. In Robotics and Automation, 2007 IEEE International Conference. Pp. 3837-3843.

[24] Huang S, Murakami K, Yamakawa Y, Senoo T, Ishikawa M. Fast Peg-and-Hole Alignment Using Visual Compliance. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Tokyo, Japan (November 2013). Pp. 286–292.

[25] Prats M, del Pobil AP, Sanz PJ. Robot Physical Interaction through the combination of Vision, Tactile and Force Feedback. Applications to Assistive Robotics. Springer Tracts in Advanced Robotics. Volume 84, 2013.

[26] Honavar V. Symbolic Artificial Intelligence and Numeric Artificial Neural Networks: Towards A Resolution of the Dichotomy. Computer Science Technical Reports, 1994. Paper 78.

[27] Towell GG, Shavlik JW. Knowledge-based artificial neural networks. Artif. Intell, 1994. 70(1–2). Pp. 119–166.

[28] Carpenter GA, Grossberg S. A massively parallel architecture for a self-organizing neural pattern recognition machine. Comput. Vis. Graph. Image Process 1987. Pp. 54–115.

[29] Carpenter GA, Grossberg S, Reynolds JH. ARTMAP: supervised real-time learning and classification of nonstationary data by self-organizing neural network. Neural Networks, 1991, Pp. 565–588.

Table1

Operation	Insertion	Offset (X,Y) mm	Learned Patterns	Alignment Motions	Knowledge Type
First Circular Peg Insertion (with ACQ-PKB)	1	(-0.5, 0.0)	1	3	Knowledge Acquisition
	2	(1.2, 0.0)	3	14	
	3	(0.5,-0.5)	4	8	
	4	(0.5, 0.5)	3	14	
Second Circular Peg Insertion (with EKB)	5	(-0.5, 0.0)	1	3	Knowledge Enhancement
	6	(1.2, 0.0)	2	12	
	7	(0.5, -0.5)	0	8	
	8	(0.5, 0.5)	0	13	

Table2

Operation	Insertion	Offset (X,Y) mm	Learned Patterns	Forgotten Patterns	Used Patterns From Circular EKB	Used Patterns From Squared EKB	Alignment Motions	Knowledge Type
First Square Peg Insertion (EKB from circle)	9	(-0.5,0)	3	0	4	3	22	Knowledge Refinement
	10	(1.2,0)	1	1	6	1	15	
	11	(0.5,-0.5)	0	4	3	0	38	
	12	(0.5,0.5)	0	0	5	0	10	
Second Square Peg Insertion (EKB from circle-square)	13	(-0.5,0)	1	0	1	1	7	Knowledge Enhancement
	14	(1.2,0)	0	1	3	0	11	
	15	(0.5,-0.5)	1	0	3	1	9	
	16	(0.5,0.5)	0	0	5	0	10	
First Radiused-Square Peg Insertion (EKB from circle-square)	17	(-0.5,0)	0	0	3	0	14	Knowledge Refinement
	18	(1.2,0)	0	0	3	1	13	
	19	(0.5,-0.5)	0	0	4	0	6	
	20	(0.5,0.5)	0	2	5	3	27	

Figure1
Click here to download high resolution image

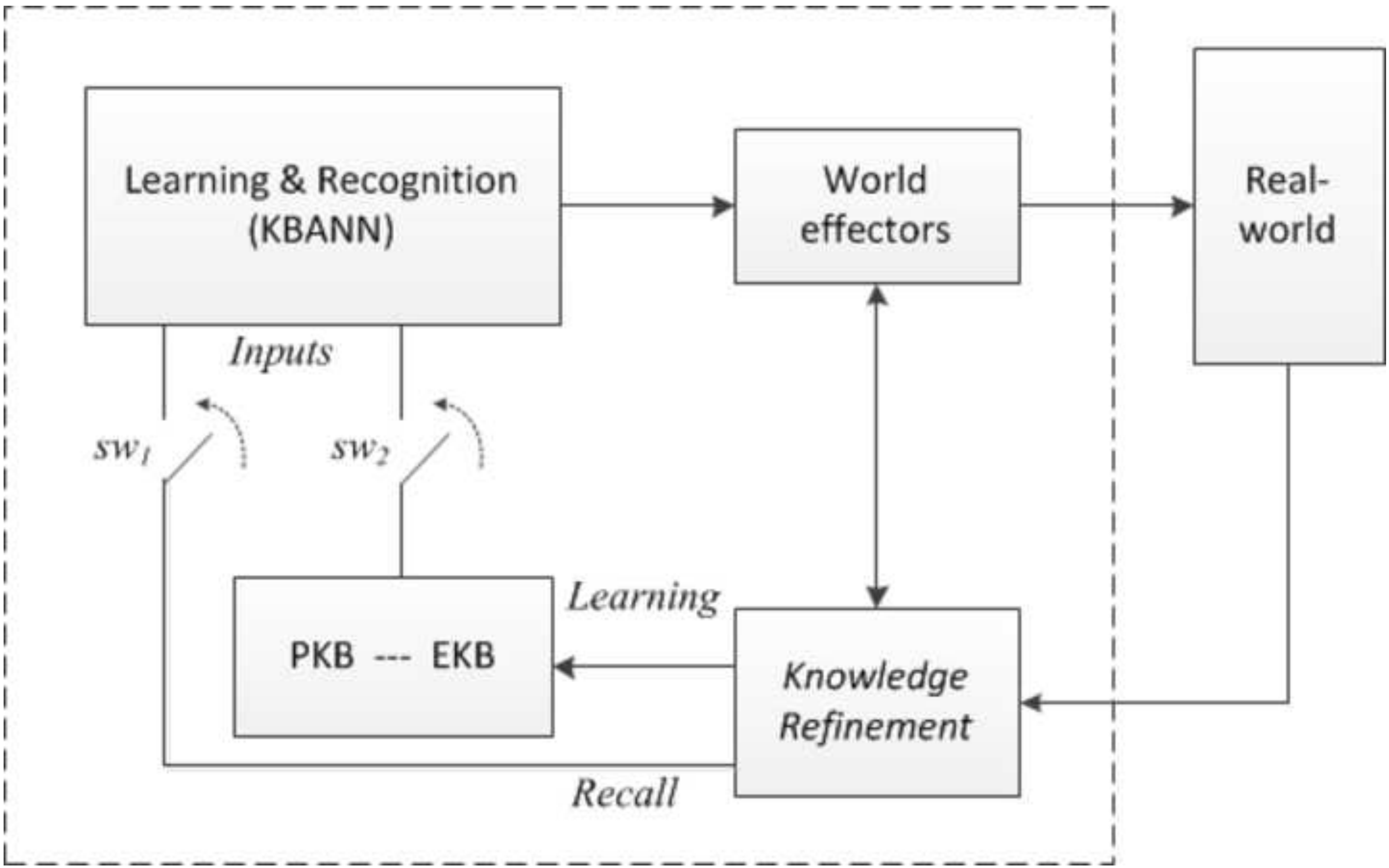

Figure2
Click here to download high resolution image

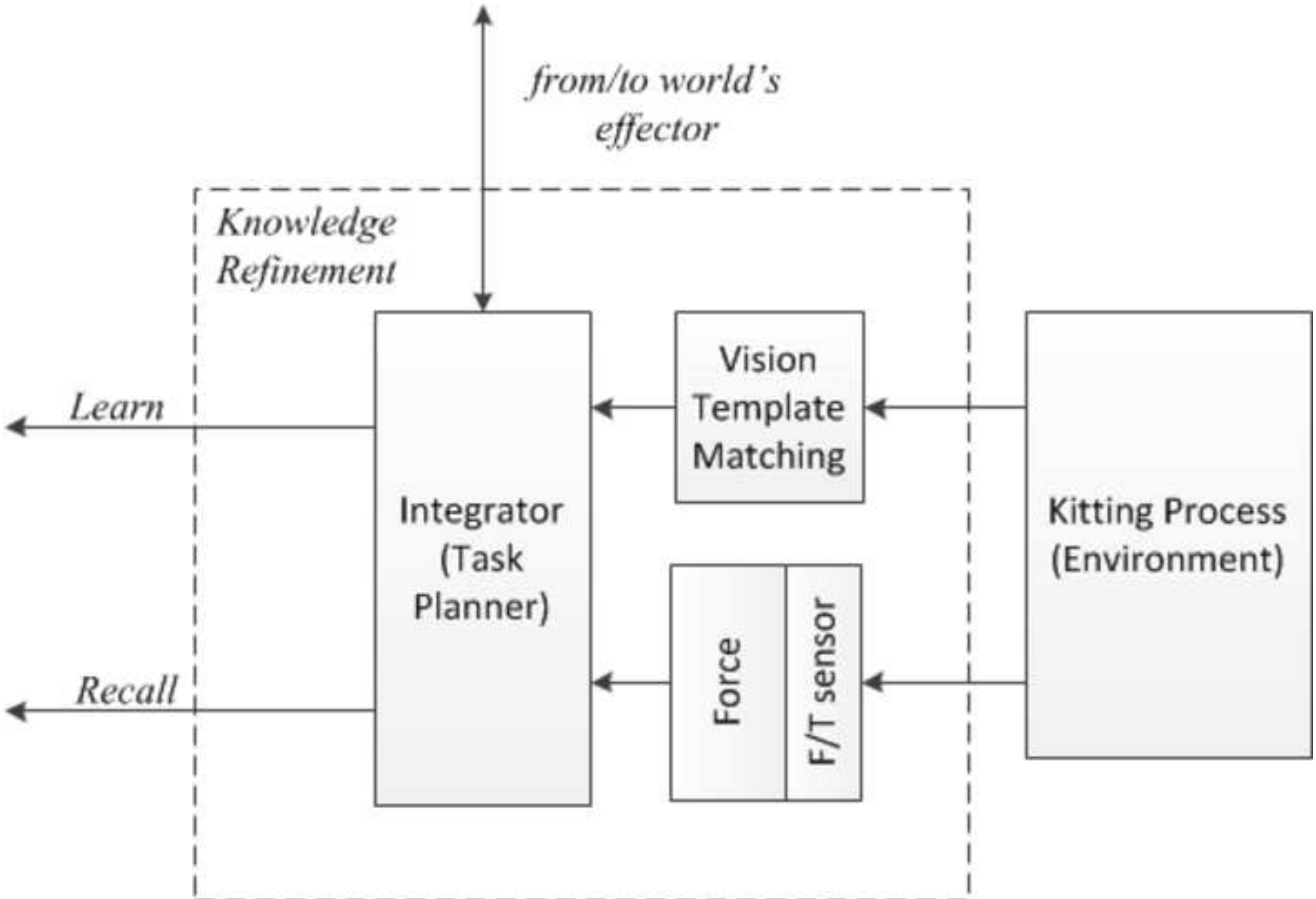

Figure3
Click here to download high resolution image

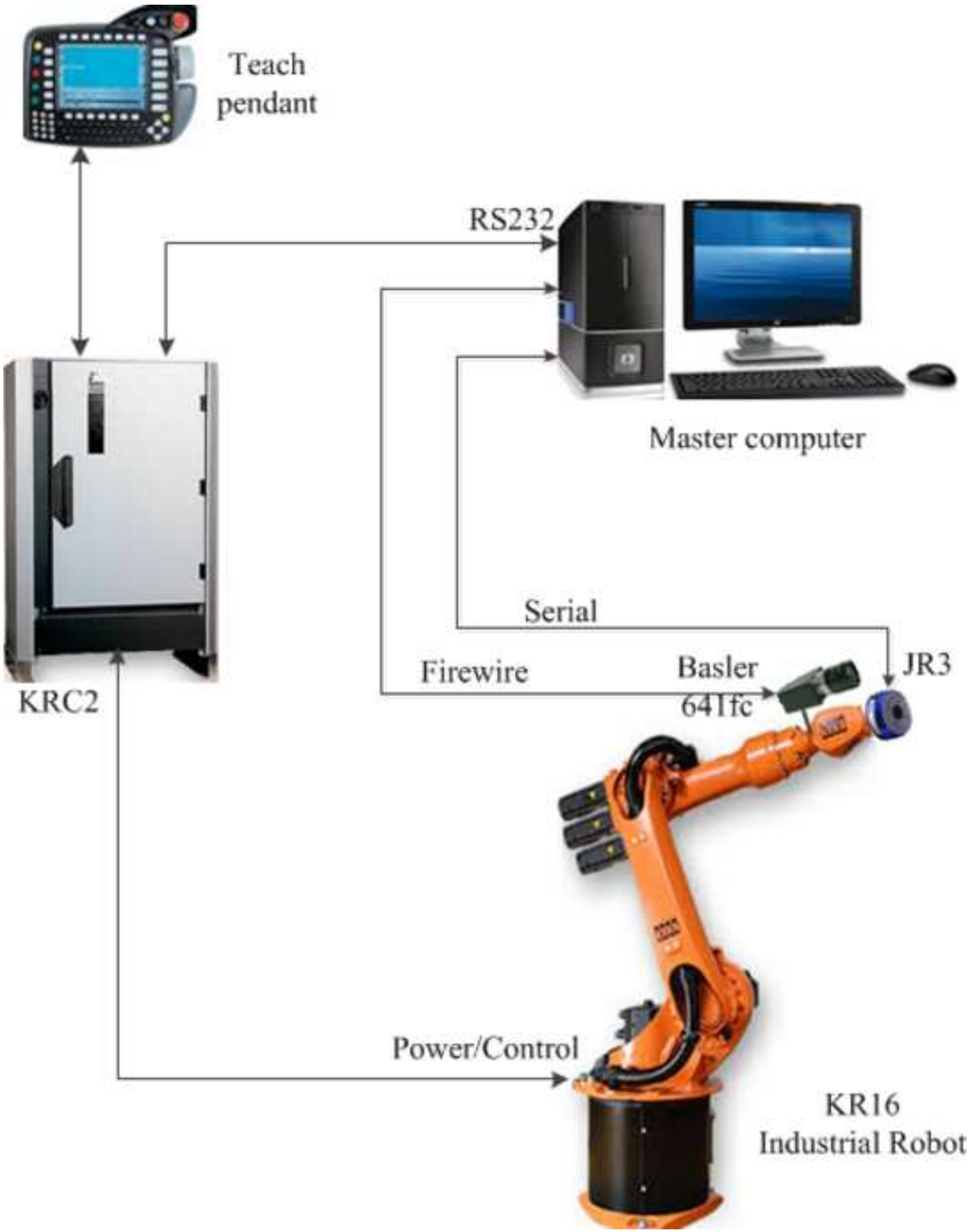

Figure4

Click here to download high resolution image

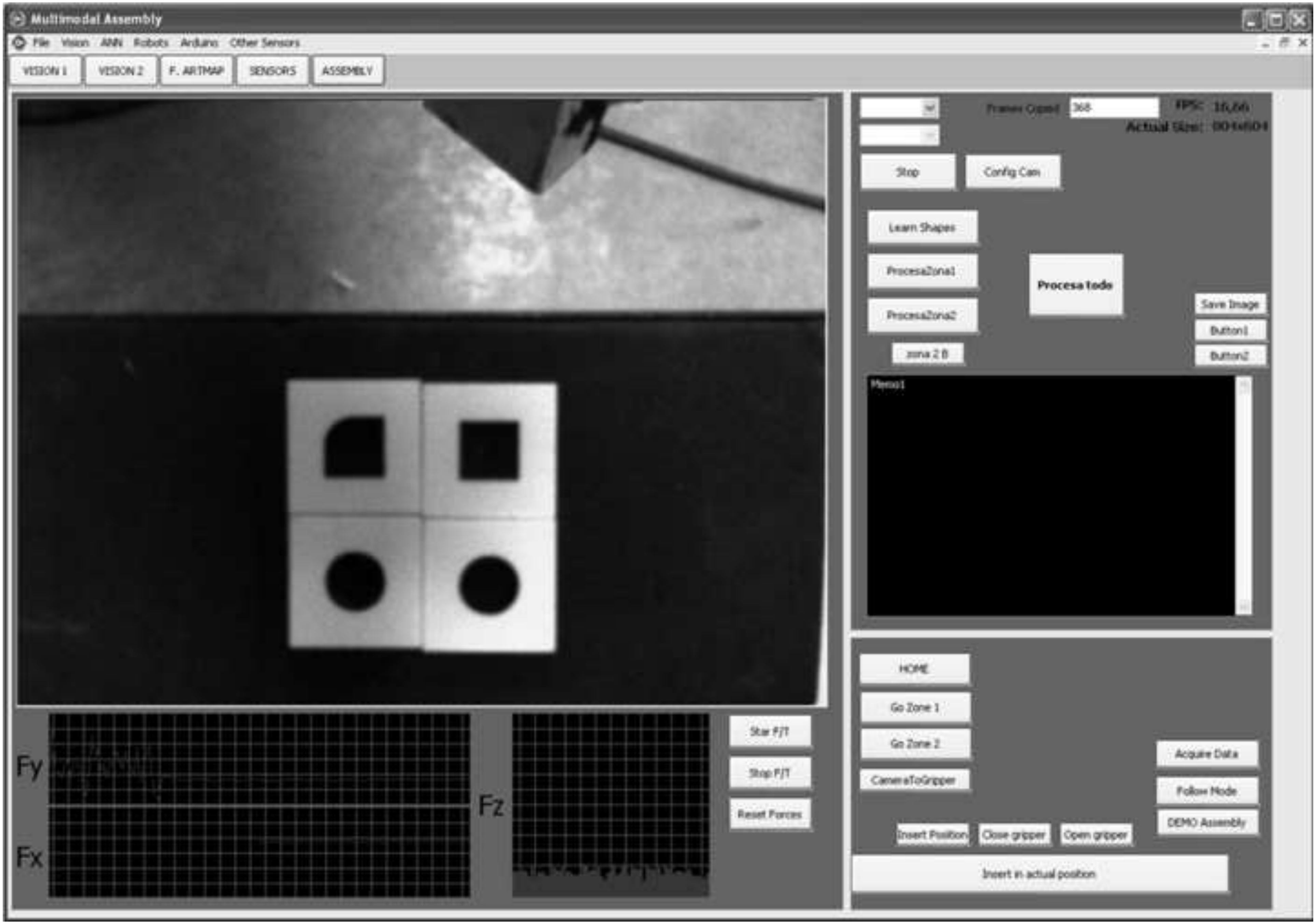

Figure5

Click here to download high resolution image

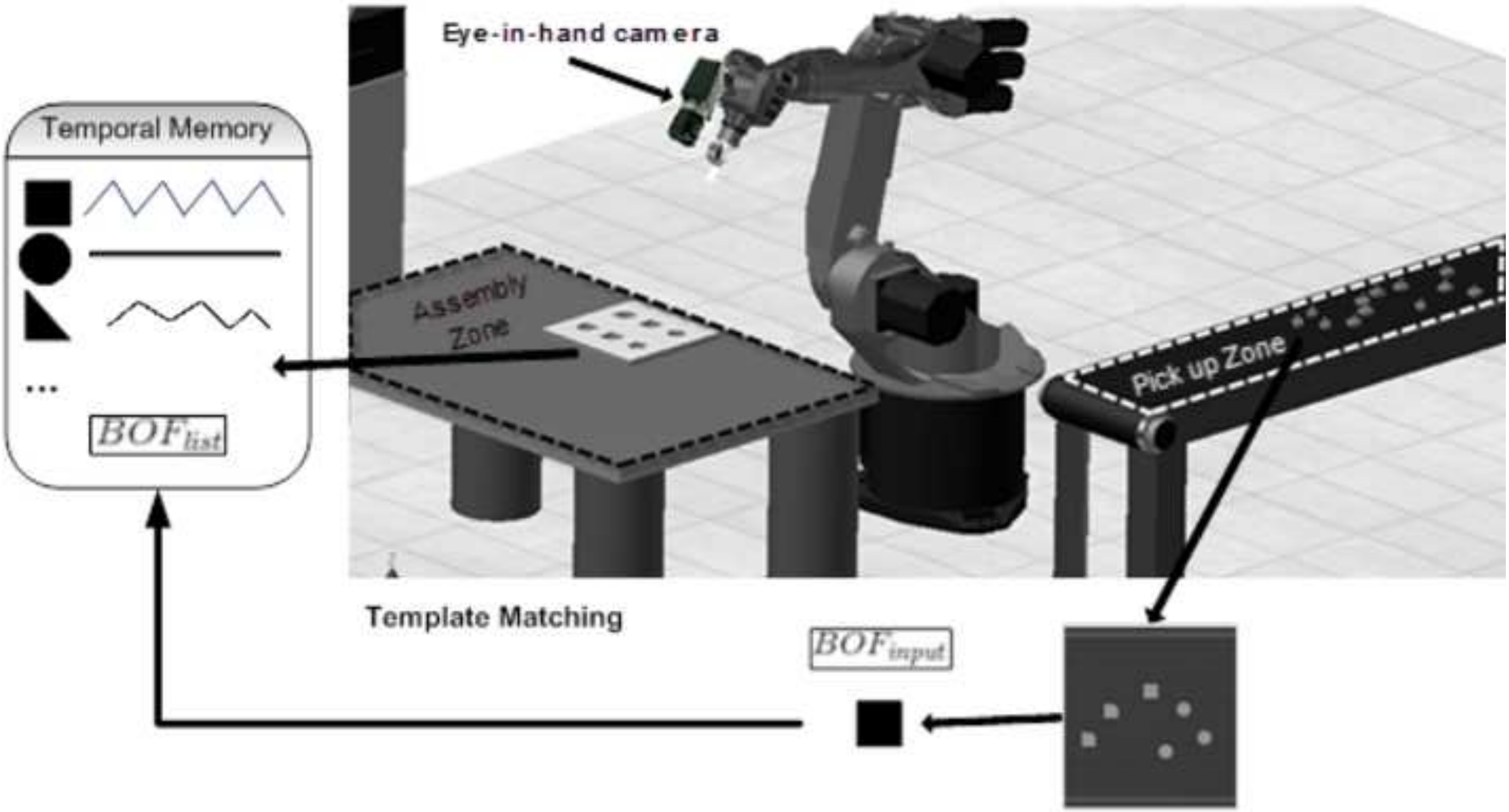

Figure6
Click here to download high resolution image

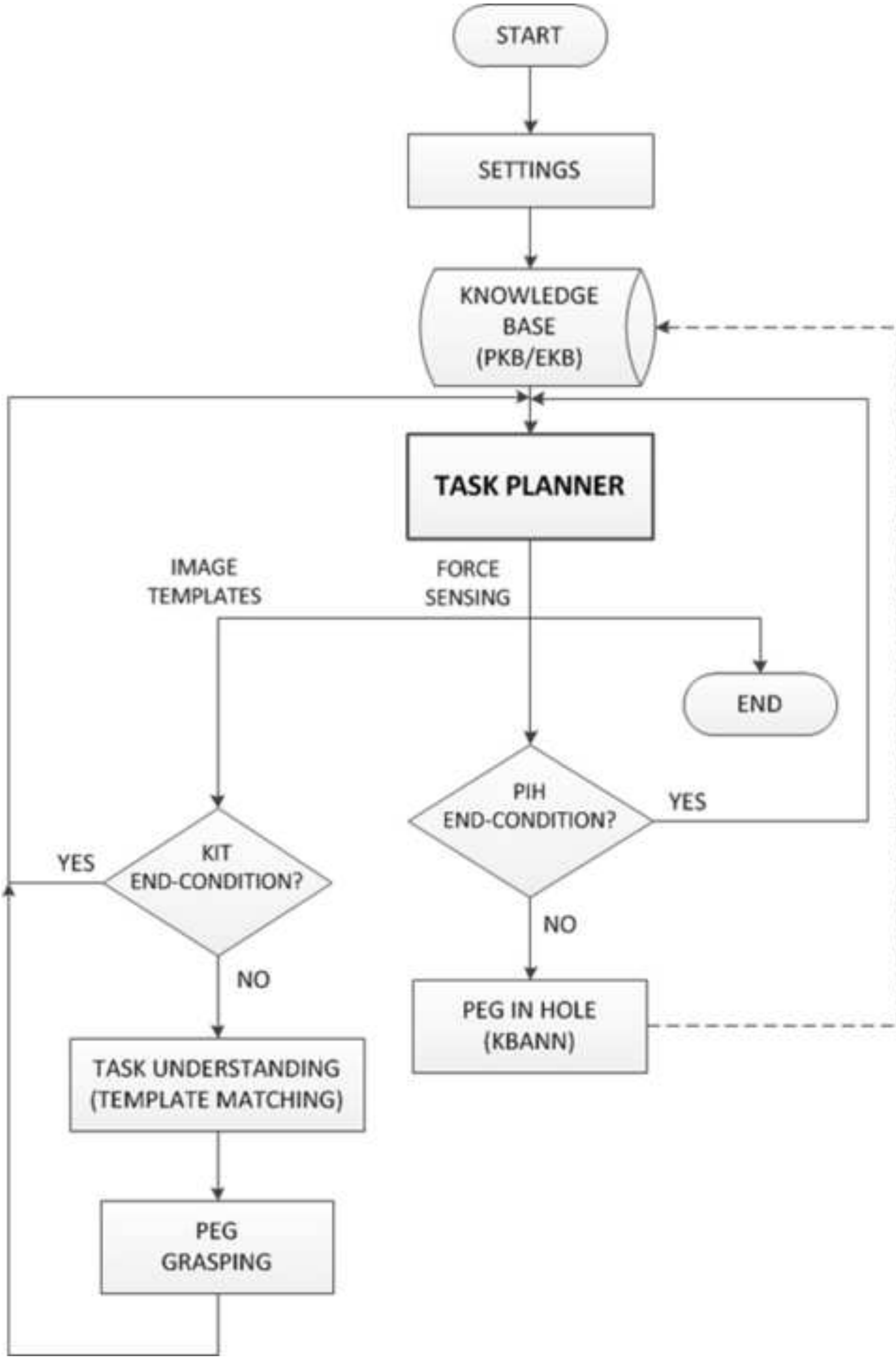

Figure7
Click here to download high resolution image

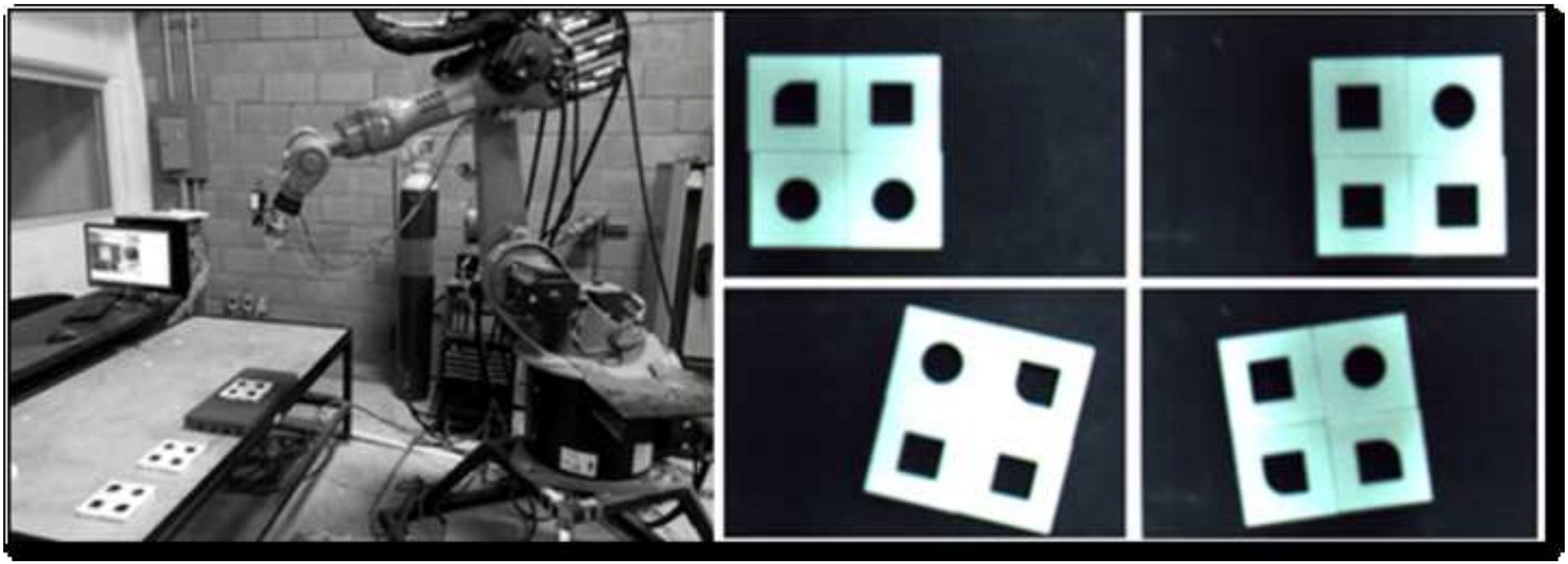

Figure8
Click here to download high resolution image

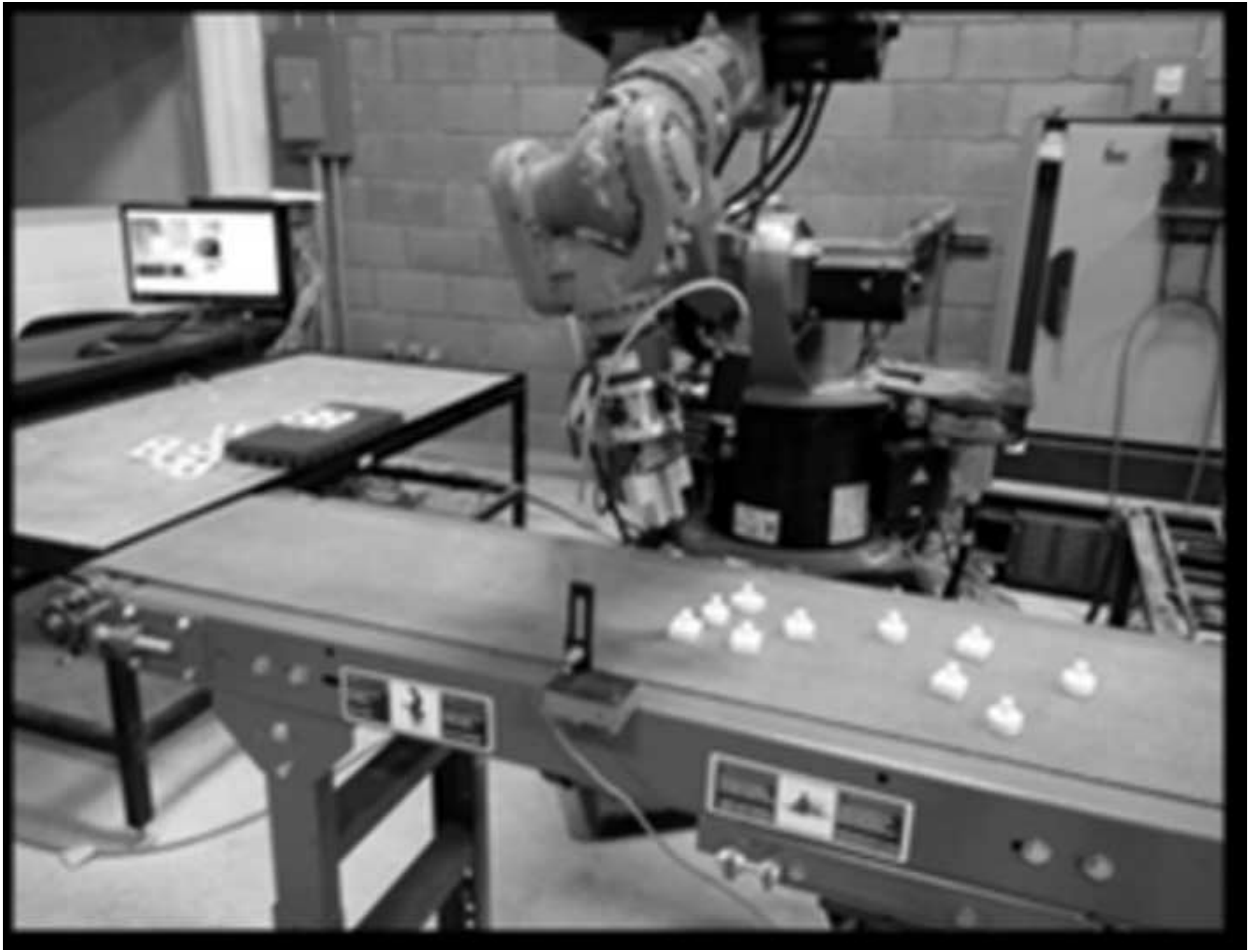

Figure9
Click here to download high resolution image

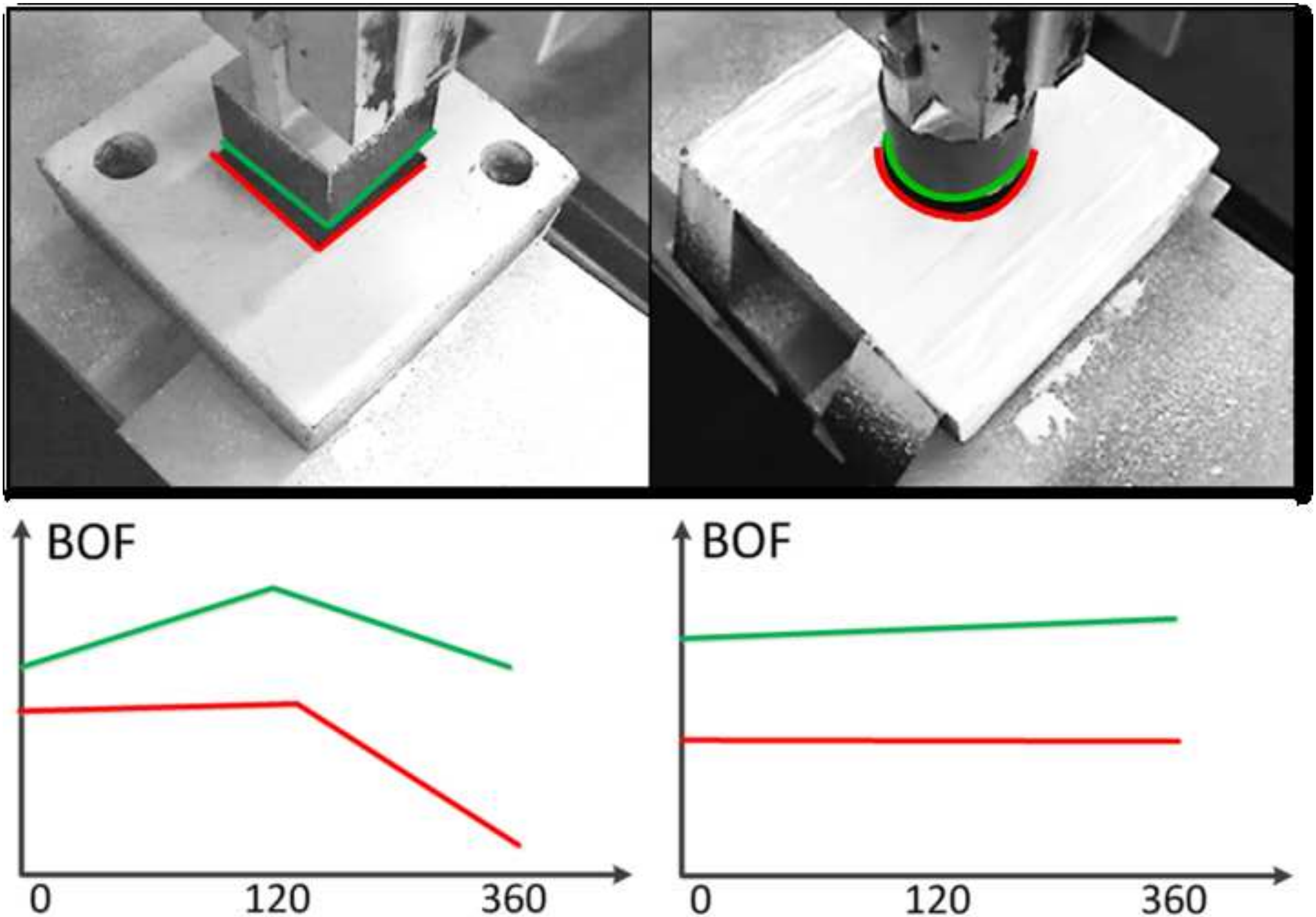

Figure10
Click here to download high resolution image

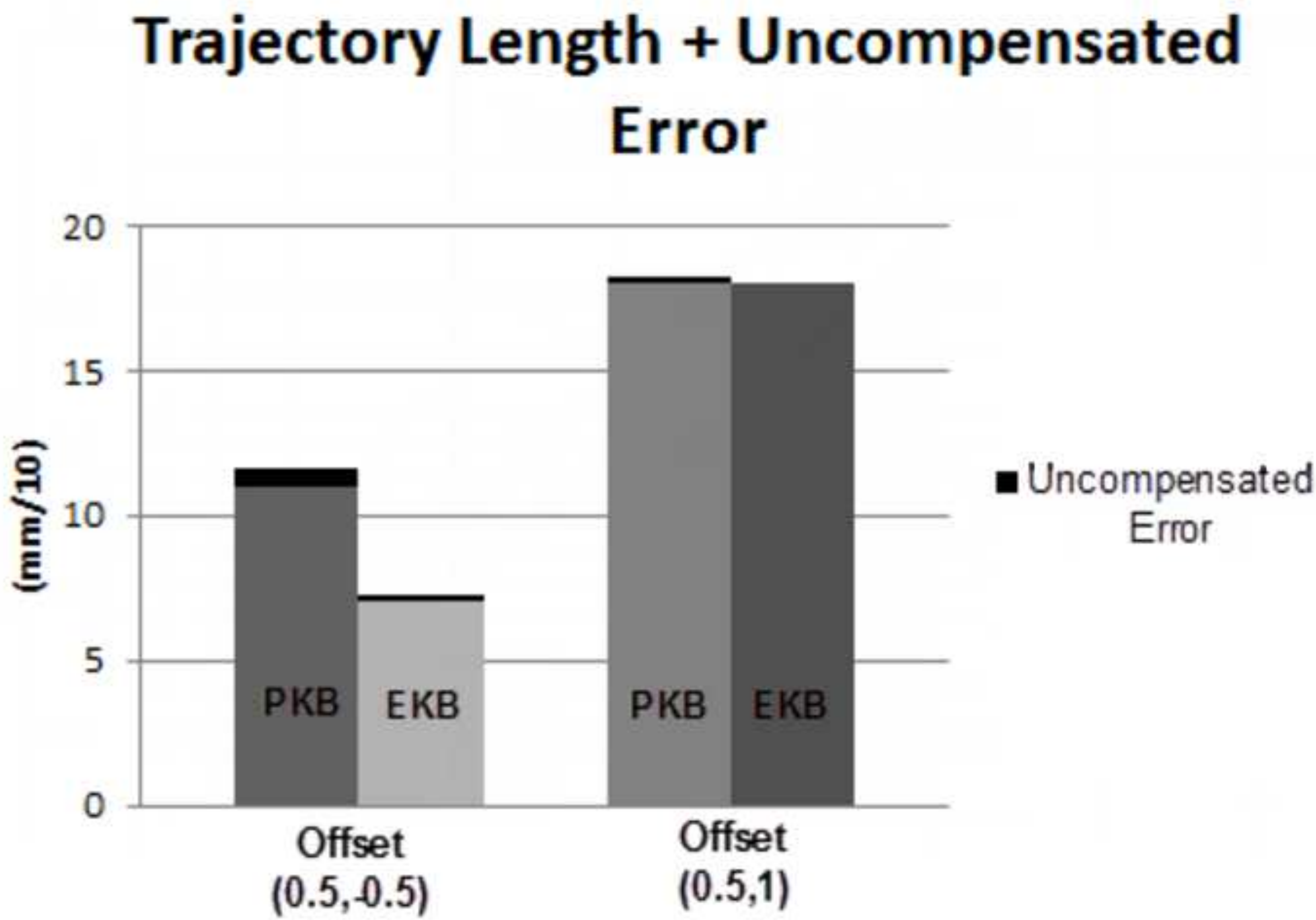

Figure11

Click here to download high resolution image

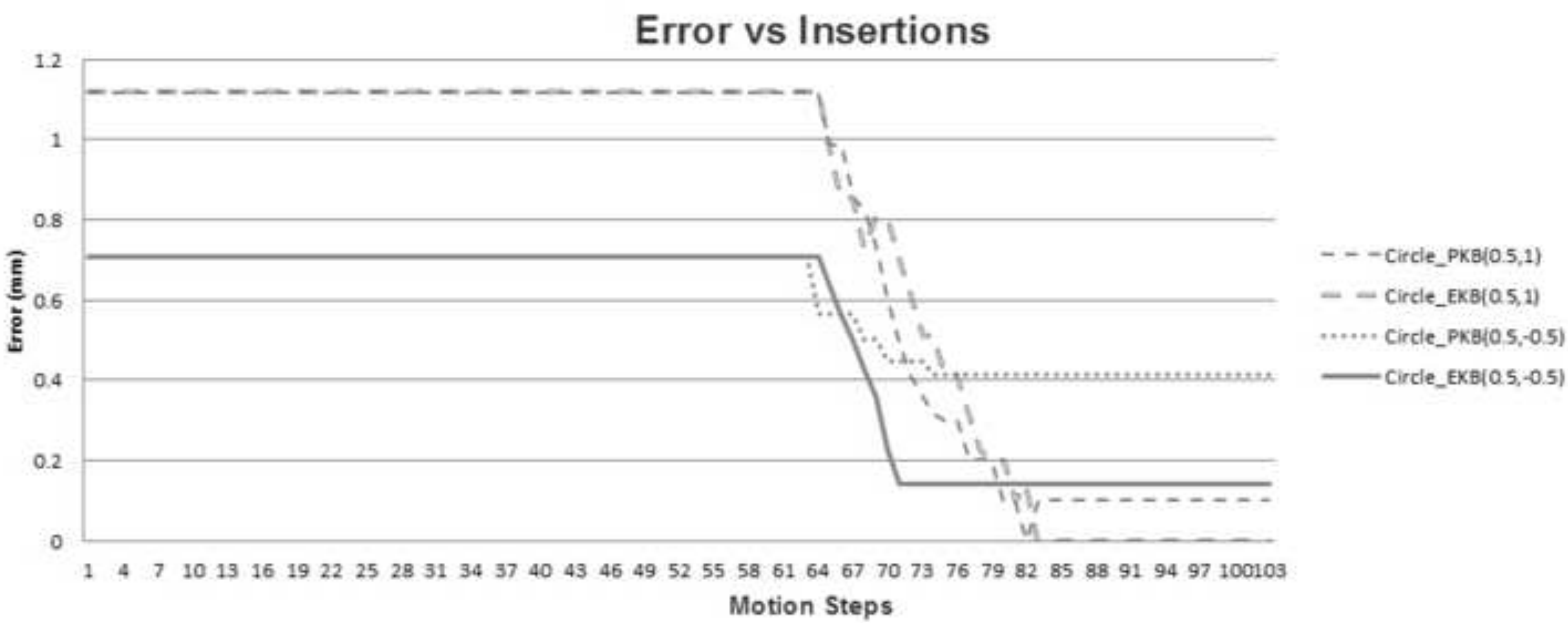

Figure12

Click here to download high resolution image