Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 207 (2022) 3820–3829

Procedia Computer Science

www.elsevier.com/locate/procedia

26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)

Context change and triggers for human intention recognition

Tong Tong[a,*], Rossitza Setchi[a], Yulia Hicks[a]

[a]*Research Centre in AI, Robotics and Human-Machine Systems (IROHMS), Cardiff University, Cardiff CF24 3AA, UK*

Abstract

In human-robot interaction, understanding human intention is important to smooth interaction between humans and robots. Proactive human-robot interactions are the trend. They rely on recognising human intentions to complete tasks. The reasoning is accomplished based on the current human state, environment and context, and human intention recognition and prediction. Many factors may affect human intention, including clues which are difficult to recognise directly from the action but may be perceived from the change in the environment or context. The changes that affect human intention are the triggers and serve as strong evidence for identifying human intention. Therefore, detecting such changes and identifying such triggers are the promising approach to assist in human intention recognition. This paper discusses the current state of art in human intention recognition in human-computer interaction and illustrates the importance of context change and triggers for human intention recognition in a variety of examples.

Keywords: human intention recognition; context change; triggers; multimodal perception.

1. Introduction

A robot plays a key role in human-robot interaction and is likely to become a part of the daily life of everyone who needs service, cooperation, and collaboration in the future. Significant research effort is dedicated to designing a robotic system applicable to robots cooperating with humans. The purpose is to accomplish the task defined by humans or help humans autonomously [1]. For this purpose, robots are designed to understand humans and complete

* Corresponding author.
E-mail address: tongt6@cardiff.ac.uk

10.1016/j.procs.2022.09.444

tasks together. For the human-robot interaction to be effortless, it is necessary that the robot understands the human and can predict the intention of humans [2]. Fully understanding the ubiquitous intentions of humans in human-robot interaction is the challenge that needs to be realised today.

In psychology, intention refers to the idea or plan that an individual wants to implement [3]. Humans can express their intention through various clues, such as actions, gestures, body movements, and even a glance. People always exhibit some characteristics that can be used as a reasonable basis for recognising intentions. Thus, these characteristics are valuable clues for robots to confirm a human's intention. Human intention recognition is essential for human-robot interaction [4], for example, when people need help or complete tasks. To explore human-like artificial intelligence, it is vital to develop a human-like robot to understand human intention. Therefore, the type of intention is an important aspect to clarify. A person's intention can be expressed in many ways. Collecting, conveying and processing these features and information are essential for human intention recognition. It is necessary to recognise the human intention correctly and efficiently to achieve fluent and efficient human-robot interaction in different scenarios.

Human intention is closely related to context awareness, which includes many aspects. There are many definitions on what context awareness is but the synergy between intention and context-awareness has not been explored so far. Wei Liu et al. [24] divide the context into the user and physical contexts. These two parts include the clues that can infer human intention. In fact, the environment, the humans and all data that can be collected constitute the context in human-robot interaction. The environment significantly impacts human intention, and people will also have different goals for the same action according to different scenarios. The relationship between the physical and user contexts can logically conclude human intention. Context awareness provides many task-specific clues in human-robot interaction and supports recognising human intentions [5]. The purpose of context awareness is to make robots capable of spontaneously recognising human intention. Changes in the current environment (parts or objects) can assist in recognising and predicting a person’s action or thought. Context awareness of objects, human actions, and environmental features can effectively identify human intentions or assist in completing tasks [5].

Although there are many methods for human intention recognition or human action recognition, many are limited to ordered or single tasks. Thus, they cannot be applied in more complex environments. Human intention is different with context constantly, even in the same environment and for the same action. Context should be used for preventing misrecognition, which is one of the most important requirements for human-robot interaction in everyday situations. Context awareness can also assist a robot in interpreting high-level relationships expressed in many variables and disambiguating and reducing uncertainty in human-robot interaction tasks [2]. Therefore, human intention recognition based on context awareness is necessary for complex environments and scenarios.

Many factors influence human intention in real life, especially in sudden situations. The variation of these factors should be detected by a robot. The context change can represent these factors that are essential for human intention recognition. The context will change all the time, but not all changes are the key to recognising human intention. Therefore, defining the significant context change that impacts human intention is crucial. The significant context change that can indicate the human intention can be called the trigger. For example, a robot may open a window because the temperature makes the human feel hot, not only because of the temperature rise. The human feeling hot is the trigger to open the window. Thus, the context change detection and the definition of the trigger are vital for human intention recognition.

This paper aims to explore the context change and triggers to recognise the human intention in the human-robot interaction. It is hypothesised that context changes and triggers are significant for a robot to recognise human intention and can provide direct evidence to reason the human intention. Intention recognition is the crucial element for the human-robot interaction in real life, which is necessary for the robot to recognise the intention of the human's next action and thought. The factors influencing human intention can be obtained from the context changes. Multimodal data collection, context change detection, trigger definition, and semantic reasoning will form a complete system for human intention recognition.

2. State of the art

2.1 *Human intention recognition*

There are many studies on human intention recognition and prediction, such as motion-based, speech-based, gesture-based, gaze-based, and based on other similar features. Motion is a direct feature to recognise the next human movement, which is used in many fields. In human-robot cooperation, parts are assembled by recognising human movements [5][7]. The robot recognises and analyses the motion of the human hand to predict the object that the human will take next and cooperate to complete the task [6]. The trajectory of human action is analysed to predict the following movement so that the robot and human can complete the delivery task [9].

Human gestures are one of the features for expressing an intention. Gestures also can visualise the human intention, like waving, pointing, etc. Since the gestures have the characteristics of flexibility and fast delivery, they can be used as essential feature variables for intention recognition. Gesture features can be used in practical applications, whether obtained via glove sensors or cameras [14], and gestures are also a direct and effective way to express human intentions. For example, a Kinect camera is used to detect human poses and gestures, recognise 3D gestures and calculate the angle between shoulders and hands to determine the direction and plan the robot's navigation [14]. Although gestures contribute significantly to intention recognition, using gestures as a single variable for intention recognition is inaccurate. The literature [16] summarises gesture-based intention recognition and comprehensively analyses the advantages and disadvantages. General disadvantages, such as low resolution and low robustness, are also identified for non-wearable devices. There is a greater dependence on the stability of the environment, light, and other factors for wearable devices.

For most people, the gaze is critical in transmitting information to the observer. From common sense, when people want an object, their eyes will stay on the object for more time. Controlling robots using eye movement to complete complex grasping tasks is an interesting direction of research in intention recognition [20]. Gaze provides information that other body parts cannot. Human-robot interaction and collaboration can be smoother and more efficient [21]. Gaze is more like the intention externalisation, which is the first captured feature.

There are many ways to recognise human intention with good accuracy. At the same time, understanding human intention is not practical using a single kind of feature as if these features cannot be detected, it would make the system unreliable.

2.2 *Multimodal perception*

In recent research on human-robot interaction and human-computer collaboration, the extraction of multiple features and the fusion of multiple data streams in a model have become a trend. Such models also have better stability and robustness than human intention recognition model based on single feature type. Assembly task is the hot topic for the human intention recognition in human-robot collaboration research, combining the human actions, objects, and knowledge base to assemble a chair [10]. Human intentions are multifaceted, not just in handover tasks and assembly tasks. Intention recognition is also important in, for example, collision avoidance, where researchers were using MYO armband, Leap Motion, Kinect camera and RGB camera sensors to collect a large amount of data, analyse the human point cloud, and then predict the following human action [11].

The robot can help a doctor to complete surgeon tasks in the shared workspace [8]. In the kitchens, the robot extracts and trains the features of speech, gestures, and eyes and then votes through the independent opinion pool to determine the desired items [2]. Wearable sensors are efficient equipment to get information from humans, which provide strong proof for human intention recognition. Natural language processing tools and wearable devices have also been used to recognise human intention and complete the delivery task [19]. An assembly line, including obstacle avoidance and other intention recognition, is realised on the basis of analysis of gestures, human body recognition and object detection [13]. Xia Z et al. [14] have also emphasised that the system combining gesture and vision can satisfactorily complete the work of intention recognition in human-robot interaction scenarios.

A complex human-robot environment would have ambiguities about the object choosing. Combining gesture and facial feature recognition allows the robot to detect accurately which object a human want [17]. Human intention can be detected not only based on their apparent characteristics, but also considering their external environment. For example, factors such as age, gender, temperature, humidity, etc., have been used to judge if a human would like to drink water [18]. This also proves that by combining all the factors, it is possible to identify the human preference for something, which has an essential role in promoting the identification of human intentions.

As shown above, multimodal perception is an effective way to collect many features useful for recognising human intention. And due to the many features, more clues can be the triggers to confirm human intention, which helps to exclude the inadequate features and increase the robustness of the system.

3. Concept of the human intention recognition system

Human intentions may be expressed in diverse ways in different scenarios. In a handover task, human gaze is used to identify the tool they would like to obtain [9]. In autonomous driving, the pedestrian's heading and posture are useful for detecting human intention [27]. In the service scenario, human needs may be indicative of the required services [28]. Therefore, the hypothesis presented in this article is that robots should be able to detect and analyse all features related to a human and the environment to recognise human intention successfully independently of scenario.

Figure 1 shows the relationship between the human, robot, and environment when working on the tasks collaboratively. Before recognising the human intention, the robot should detect the human and analyse the environment to collect the features. Meanwhile, the environment will constantly influence human intention because human intention depends on the different circumstances. As the monitor, the robot needs to get feedback from the human. It can give more clues to analysing the human intention correctly. Finally, the robot recognises the human intention and completes the task.

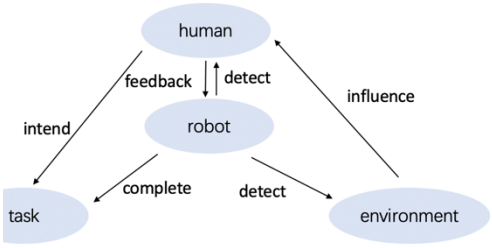

Fig.1. Relationship in the human, robot, and environment in the intention recognition system

In many cases, the difficulty of human intention recognition is that there is no standard to evaluate [2]. Therefore, collecting as much information as possible is essential to recognise intention correctly. The robot fuses the information and extracts the change in real-time. The robot needs to observe the human change, and the environment changes constantly. According to the level of change and reasoning, the robot defines the triggers. The context change, the triggers and the reasoning are the pivotal parts of human intention recognition. It is necessary to fully collect more features and reasoning according to the context change and the triggers.

To recognise the human intention, data collection is the primary step. Human intention recognition is relevant for both humans and robots, the latter of which should pay attention to the former's surroundings. Therefore, multimodal perception is an effective approach to collecting data. The single features are not suitable for more complex conditions. Some features are not captured in different scenarios. In these situations, it needs other features to support the human intention recognition. More features can bring more clues. The advantage of the information obtained through multimodal perception is that it can reduce the reasons for the low robustness due to a single variable and reduce the uncertainty under the true intention [23].

The specific features are acquired from the various sources, and they have diverse types, dimensions, and contents, such as gestures, gaze etc. Combining them is essential to human intention recognition. Data fusion is the processing of integrating multiple data sources to produce more consistent, accurate, and useful information than that provided by any individual data source. Data fusion is a combination of multiple sources to obtain improved information [26] for human intention recognition. Data fusion increases the robot's understanding of context.

With the complexity of human-robot interaction scenarios, context understanding contributes to high-level semantic reasoning [25]. Humans and the environment together make up the context. A context model is created by combining all environmental and human features. It is used to make robots understand the current scenario and

combine reasoning to recognise the human intention. Context models also give robots a higher-level understanding of human intention reasoning. The context model is continuously optimised and updated with the incoming data, which improves the accuracy of the model and assists in generalisation and context adaptation [29]. The context model synchronises all features to ensure the robot can extract the significant context change and identify the triggers.

Human intention and context are strongly related. Context change is important to predict human intention. Detecting these changes is the precondition of reasoning. Any changes in human-robot interaction can be context change, such as human actions, human gaze, environment sounds, objects, etc. The aim of detecting context change is to find the effective evidence to recognise the human intention and narrow the reasoning range. The context changes constantly, but not all changes affect human intentions, and relying on context changes to determine human intentions will have the problem of overfitting [30]. Therefore, the robot also needs to determine the triggers, which can serve as direct evidence for recognising human intention.

The triggers can be determined from the context changes and are the significant changes that indicate the human intention. In human-robot interaction, triggers can be implemented as the nodes of semantic reasoning networks that lead to correct recognition of human intention. The trigger is the evidence that leads to recognition of human intention. For example, in the handover task, the human pointing or gaze is the trigger to decide that the human wants the tool [5] [14]. Not all context changes serve as triggers and lead to correct reasoning. Recognising the trigger can lead to an action from the robot. The benefit of defining and recognising triggers is providing the robot with the facility to reason and react quickly. The robot should detect the context changes and determine if any of the changes constitute triggers, *i.e.,* the evidence leading to recognition of human intention in combination with other contextual information.

Semantic reasoning supports the robot recognising the human intention according to the context and the data. Semantic reasoning relies on the concept and relationship. In addition, the context changes and triggers provide adequate information for the reasoning. When a robot detects context changes and defines the necessary triggers, it should read the same depending on its relationship with the human in question. In human-robot interaction, the robot can interact autonomously by understanding the overall scene, including the real-time state of people and the relationship between environmental objects [22]. The figure following shows the complete concept of human intention recognition.

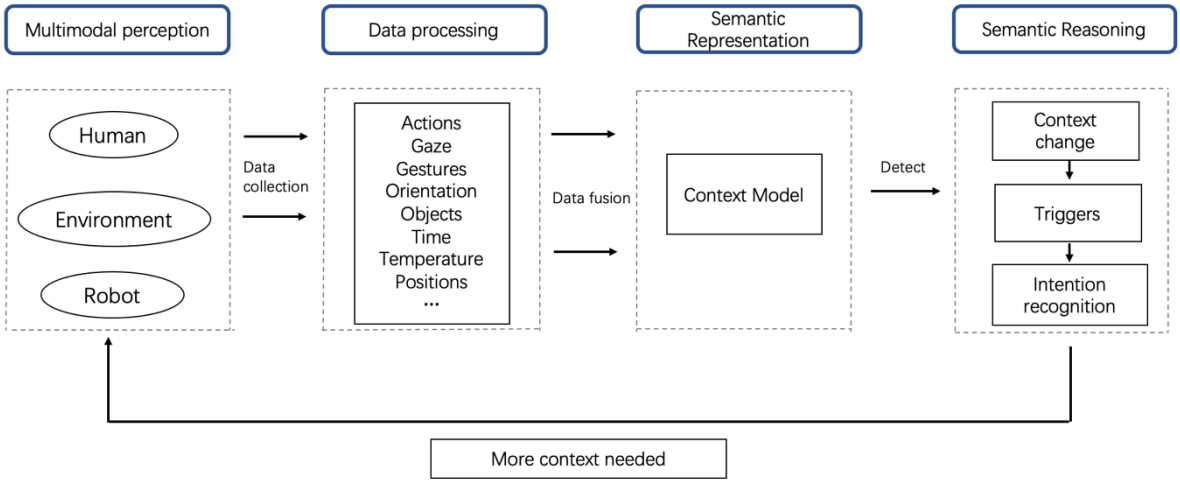

Fig. 2. The concept of the human intention recognition system

It can see that multimodal perception can acquire the features from the environment and the human. There are many features should be collected. In real life, not all features, such as occlusions, can be observed. At this time, human intentions can judge based on other features, which have better robustness. Data processing can extract the specific features, which are the clues of human intention. The robot should fuse these features, and make them

normalisation, which is used to create the context model. The context model is the important part for the intention recognition, which includes all features from the environment and the human. The robot will understand the contextual awareness better based on the context model. The robot will detect the context change from the context model. Depending on the importance of the context change, the robot should define the triggers, which is the direct evidence to reason the human intention. Finally, the robot recognises the human intention through reasoning the relationship between the triggers and the human.

The context changes and the triggers are the crucial elements in this system. First, human intention is influenced by the environment. Second, the context change can provide effective clues to identify the human intention. Then, the trigger can support the robot in reasoning directly. Finally, the robot can recognise the human intention based on the context change and triggers.

4. Human intention recognition

In order to illustrate the model of human intention recognition, there is a scenario for human intention recognition, which is based on humans, objects and the environment. According to the features of a human and the environment, the robot recognises human intention in daily life. This includes multimodal perception (collecting data from multi-sensors), context change, triggers, and reasoning. The robot can collect different data using many sensors. These can be used to analyse the environment, which can help the robot understand human intentions. Humans can express their intention in many ways, but some of them cannot be acquired by the robot in some special situations. An example of such a situation is the robot may not be able to recognise gestures when the hands are not free. So, the robot will need other features to analyse human intention.

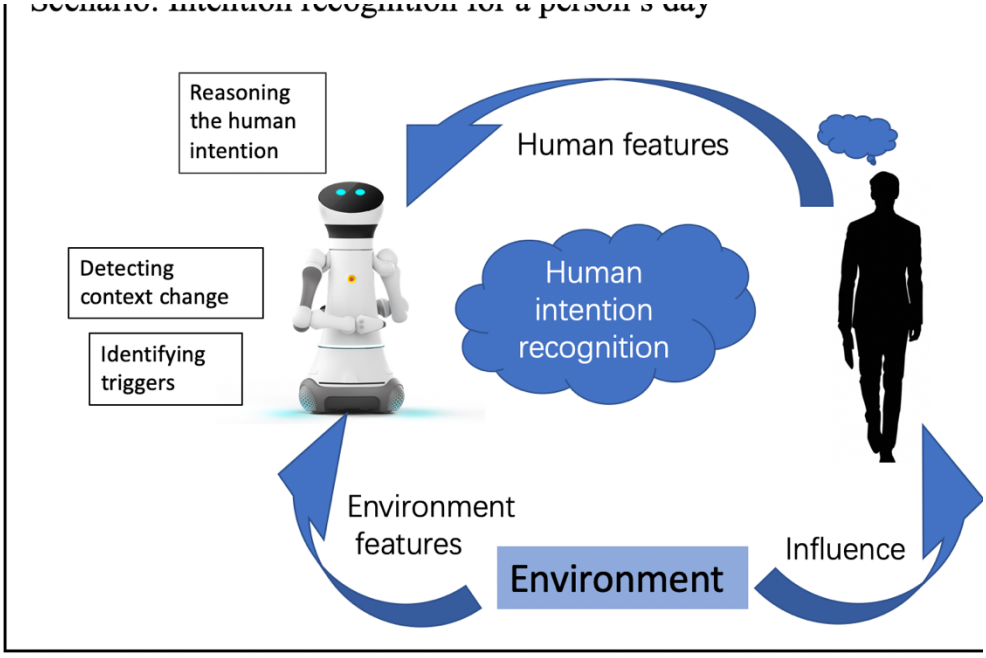

Fig. 3. A scenario of the human intention recognition

This scenario requires many different features from humans and the environment. These features can allow the robot to analyse the current human states and the current environment, which are important for recognising human intention. Then, the robot should detect the context changes and detect the triggers. Both of them are the key to

recognising human intention. The context will influence the human constantly, and the context is the evidence when judging human intention.

Examples

Environment: the temperature has dropped (context change).
Robot: the robot continues observing the person and the environment.
Person: the person folds their arms as if they are cold and looks at the temperature control (trigger).
Robot: the robot controls the indoor temperature (intention recognition).

Person: the person turns on the computer (context change).
Robot: the robot continues observing the person and the environment.
Person: the person looks at the pen and the paper out of reach (trigger).
Robot: bring the pen and the paper to the person (intention recognition).

Person: the person holds rubbish in their hand (context change).
Robot: the robot continues observing the person and the environment.
Person: the person looks at the bin (trigger).
Robot: the robot takes the rubbish from person's hand and throws the rubbish into the bin (intention recognition).

Environment: somebody knocks on the door (context change).
Robot: the robot continues observing the person and the environment.
Person: the person looks at the door (trigger).
Robot: the robot continues observing the person and the environment.
Person: the person appears to be busy (trigger).
Robot: the robot goes to open the door (intention recognition).

Environment: the room has become dark (context change).
Robot: the robot continues observing the person and the environment.
Person: the person is reaching out for a book (trigger).
Robot: the robot turn on the light (intention recognition).

Person: the person goes to bed (context change).
Environment: time has become late (context change).
Robot: the robot continues observing the person and the environment.
Person: the person closes the eyes (trigger).
Robot: the robot turns off the light (intention recognition).

These examples involve a person's daily life. To recognise the human intention in a complex environment, the precondition is the robot knows the person's habits. It should be recorded by detecting human behaviours. It is noticed that the nonverbal features are used in these scenarios, which are convenient and effective to capture. The robot will explore the features to recognise the human intention. For example, when there is a knock on the door, the context changes. Meanwhile, the robot will continue to observe the other context features because it cannot recognise the human intention based on the knock on the door. The person looks at the door, and the robot will set this action as the trigger. The robot detects the human is busy now. It is also a trigger because the robot will confirm whether the robot should open the door. According to the human state, context change and the triggers, the robot reasons that it should open the door. It is suitable for normal situation and imitates human-human interaction. In the above examples, all triggers happen to be human actions, but they do not have to be. In other scenarios, other sources of information/contaxt change could serve as the triggers.

The context changes and triggers are the evidence used to recognise the human intention. It includes the human features change and the environment change. A person's state change is the signal to analyse the human intention.

When a person intends to put rubbish into a specific bin, they will hold the rubbish. Holding a rubbish is the context change. It can be detected by the sensors. The robot captures this action and tries to understand the human intention. Identifying the context change is the vital step (the person holds the rubbish). Then, the robot keeps observing the person and the environment. The robot detects the person who is looking at the bin. The gaze can be acquired from the sensors and transferred to the robot. According to the semantic network containing a suitable rule, the person looking at the bin can lead the robot to identify the trigger. Finally, the robot recognises the human intention according to semantic reasoning.

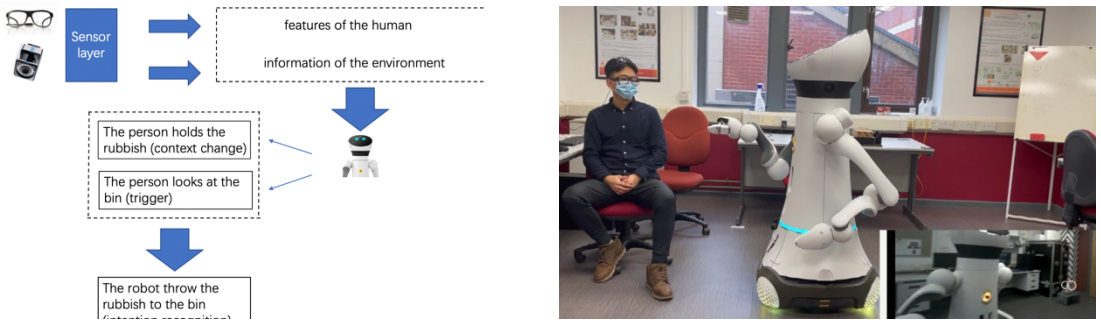

Fig. 3. An example of human intention recognition

The robot should judge the level of the context change when it is detected. The robot will continue observing the human and the environment. If the context change is significant enough, the robot will set it as the trigger. In other words, trigger is a context change which leads to a change in recognised human intention and may require an action from a robot. For example, when the temperature drops, the robot will provide warm clothes and adjust the indoor temperature only if it makes people feel cold. Not all context changes will affect human intention. The robot should divide them and identify the triggers. Then, the robot will understand the human intention.

These examples can be extended to many situations. Context changes and triggers should be considered constantly by observing many variables. Focusing on context changes and triggers can assist in recognising human intention.

5. Conclusion

With the rapid development of human-computer interaction, the requirements of many industries in this field have also increased. Intention recognition is central to human-robot interaction in various areas such as medical care, service, and industry. The ability of robots to help people is the basis of their design. This paper explores a method for pro-active human intention recognition based on context changes and triggers. In real life, human intention recognition is based not only on voice commands to robots but also on various non-verbal characteristics of humans and the environment. The robot can detect and reason about the context changes and the triggers, which assists in smooth human-robot interaction. This paper illustrates the proposed model of human intention recognition using several everyday life examples involving human-robot interaction, suggesting that this approach is feasible.

Acknowledgements

The authors would like to acknowledge the support of the Centre for Artificial Intelligence, Robotics and Human-Machine Systems (IROHMS), operation C82092, part-funded by the European Regional Development Fund (ERDF) through the Welsh Government

References

[1] Setchi R, Dehkordi M. B, and Khan J S. (2020) “Explainable robotics in human-robot interactions.” Procedia Computer Science 176: 3057-3066.

[2] Trick Susanne, Koert Dorothea, Peters Jan, and Constantin A. Rothkopf. (2019) “Multimodal uncertainty reduction for intention recognition in human-robot interaction.” IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) pp. 7009-7016.

[3] Kim S, Yu Z, and Lee M. (2017) “Understanding human intention by connecting perception and action learning in artificial agents.” Neural Networks 92:29-38.

[4] Guo L, Lu Z, and Yao L. (2021) “Human-machine interaction sensing technology based on hand gesture recognition: A review.” IEEE Transactions on Human-Machine Systems 51:300-309.

[5] Wang P, Liu H, Wang L, and Gao R. X. (2018) “Deep learning-based human motion recognition for predictive context-aware human-robot collaboration.” CIRP annals 67(1): 17-20.

[6] Chang M. L, Gutierrez R. A, Khante P, Short E. S, and Thomaz A. L. (2018) “Effects of integrated intent recognition and communication on human-robot collaboration.” IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 3381-3386.

[7] Tortora S, Michieletto S, Stival F, and Menegatti E. (2019) “Fast human motion prediction for human-robot collaboration with wearable interface.” IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM) 457-462.

[8] Zhou Tian, and Wachs Juan Pablo. (2018). “Early prediction for physical human robot collaboration in the operating room.” Autonomous Robots 42(5): 977-995.

[9] Zhang J, Liu H, Chang Q, Wang L, and Gao, R. X. (2020) “Recurrent neural network for motion trajectory prediction in human-robot collaborative assembly.” CIRP annals 69(1):9-12.

[10] Darvish K, Bruno B, Simetti E, Mastrogiovanni F, and Casalino G. (2018) “Interleaved online task planning, simulation, task allocation and motion control for flexible human-robot cooperation.” IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN) 58-65.

[11] Liu Hongyi, and Wang Lihui. (2021) “Collision-free human-robot collaboration based on context awareness. Robotics and Computer-Integrated Manufacturing.” 67:101997.

[12] Z Ji, Renxi Qiu, Alex Noyvirt, Anthony Soroka, Michael Packianather, Rossi Setchi, Dayou Li, and Shou Xu. (2012) “Towards automated task planning for service robots using semantic knowledge representation.” International Conference on Industrial Informatics 1194-1201.

[13] de Gea Fernández J., Mronga D, Günther M, Knobloch T, Wirkus M, Schröer M, Trampler M, Stiene S, Kirchner E, Bargsten V, Bänziger T, Twiwes J, Krüger T, and Kirchner F. (2017) “Multimodal sensor-based whole-body control for human–robot collaboration in industrial settings.” Robotics and Autonomous Systems 94:102-119.

[14] Xia Z, Lei Q, Yang Y, Zhang H, He Y, Wang W, and Huang M. (2019) “Vision-based hand gesture recognition for human-robot collaboration: a survey.” International Conference on Control, Automation and Robotics (ICCAR) 198-205.

[15] Lai Y, Wang C, Li Y, Ge S.S, and Huang D, (2016). “3D pointing gesture recognition for human-robot interaction.” Chinese Control and Decision Conference (CCDC) 4959-4964.

[16] Liu Hongyi, and Wang Lihui. (2018) “Gesture recognition for human-robot collaboration: A review.” International Journal of Industrial Ergonomics 68:355-367.

[17] Canal G, Escalera S, and Angulo C. (2016) “A real-time human-robot interaction system based on gestures for assistive scenarios.” Computer Vision and Image Understanding 149: 65-77.

[18] Hao M, Cao W. H, Wu M, Liu Z. T, and Li S. H. (2018) “An initiative service method based on fuzzy analytical hierarchy process and context intention inference for drinking service robot.” IEEE Transactions on Cognitive and Developmental Systems 11(2):221-233.

[19] Wang W, Li R, Chen Y, Sun Y, and Jia Y. (2021) “Predicting Human Intentions in Human-Robot Hand-Over Tasks Through Multimodal Learning.” IEEE Transactions on Automation Science and Engineering 1-15.

[20] Shafti A, Orlov P, and Faisal A. A. (2019) “Gaze-based, context-aware robotic system for assisted reaching and grasping.” International Conference on Robotics and Automation (ICRA) 863-869.

[21] Saran A, Majumdar S, Short E. S, Thomaz A, and Niekum S. (2018) “Human gaze following for human-robot interaction.” IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 8615-8621.

[22] Li S, Wang R, Zheng P, and Wang L. (2021) “Towards proactive human–robot collaboration: A foreseeable cognitive manufacturing paradigm.” Journal of Manufacturing Systems 60: 547-552.

[23] Ernst M.O. and Banks M.S. (2002) “Humans integrate visual and haptic information in a statistically optimal fashion.” Nature 415(6870):429-433.

[24] Liu W, Li X, and Huang D. (2011) “A survey on context awareness.” International Conference on Computer Science and Service System (CSSS) 144-147.

[25] Xu B, Li J, Wong Y, Zhao Q, and Kankanhalli M S. (2019) “Interact as you intend: Intention-driven human-object interaction detection.” IEEE Transactions on Multimedia 22(6): 1423-1432.

[26] Castanedo F. (2013) “A review of data fusion techniques.” The scientific world journal.

[27] Mínguez R Q, Alonso I P, Fernández-Llorca D, and Sotelo M A. (2018) “Pedestrian path, pose, and intention prediction through gaussian process dynamical models and pedestrian activity recognition.” IEEE Transactions on Intelligent Transportation Systems 20(5): 1803-1814.

[28] Wang Z, Tian G, and Shao X (2020) “Home service robot task planning using semantic knowledge and probabilistic inference.” Knowledge-Based Systems 204: 106174.

[29] Oliver G, Lanillos P, and Cheng G. (2021) “An empirical study of active inference on a humanoid robot.” IEEE Transactions on Cognitive and Developmental Systems.

[30] Da Costa L, Lanillos P, Sajid N, Friston K, and Khan S. (2022) “How active inference could help revolutionise robotics.” Entropy 24(3): 361.