


# A concise review on sensor signal acquisition and transformation applied to human activity recognition and human–robot interaction

International Journal of Distributed  
Sensor Networks  
2019, Vol. 15(6)  
© The Author(s) 2019  
DOI: 10.1177/1550147719853987  
journals.sagepub.com/home/dsn  


Lourdes Martínez-Villaseñor and Hiram Ponce 

## Abstract

Human activity recognition deals with the integration of sensing and reasoning aiming to understand better people's actions. Moreover, it plays an important role in human interaction, human–robot interaction, and brain–computer interaction. When these approaches have to be developed, different efforts from signal processing and artificial intelligence are considered. In that sense, this article aims to present a concise review of signal processing in human activity recognition systems and describe two examples and applications both in human activity recognition and robotics: human–robot interaction and socialization, and imitation learning in robotics. In addition, it presents ideas and trends in the context of human activity recognition for human–robot interaction that are important when processing signals within that systems.

## Keywords

Human activity recognition, human–computer interaction, sensor signals, machine learning

Date received: 29 December 2018; accepted: 30 April 2019

Handling Editor: Paolo Bellavista

## Introduction

Current trends in computer science and the integration of signal processing in embedded devices have been rapidly growing. Interestingly, mobile devices and the miniaturization of sensors allow research and development in the field of context-aware systems for different domain applications. Particularly, the identification of daily activities carried out by people is of big importance. For instance, applications related to this activity recognition are relevant in pervasive and mobile computing, surveillance-based security, context-aware computing, robotics, health, and ambient assistive living,<sup>1</sup> among others.<sup>2–4</sup>

In this regard, human activity recognition (HAR) deals with the integration of sensing and reasoning, aiming to understand people's actions better.<sup>5</sup> The goal of HAR is to classify and identify activities based on the collected data from different devices such as sensors

or cameras, mainly processed by machine learning methods and pattern recognition techniques.<sup>3,6</sup>

HAR plays an important role in human interaction, human–robot interaction (HRI), and brain–computer interaction (BCI). It results significant because it provides information related to the profile of a person, its personality and psychological state, and its behavior with the environment.<sup>7</sup> Directly or indirectly, HAR allows building more robust and complete tasks to understand and implement computer interactions with humans.

---

Facultad de Ingeniería, Universidad Panamericana, Ciudad de México, México

### Corresponding author:

Hiram Ponce, Facultad de Ingeniería, Universidad Panamericana, Augusto Rodin 498, Ciudad de México 03920, México.  
Email: hponce@up.edu.mx



Depending on the level of recognition, activities can be classified as simple and complex tasks. Simple activities consider daily life actions that most of people perform easily, such as “walking” or “running,” and they are relatively easy to recognize.<sup>7</sup> Complex activities are typically composed of simple activities and they are considered high-level or context tasks, for example, “drink a cup of coffee” or “cooking sushi.”<sup>7</sup> Normally, complex activities are very difficult to recognize, even if the activity can be decomposed into simple tasks.

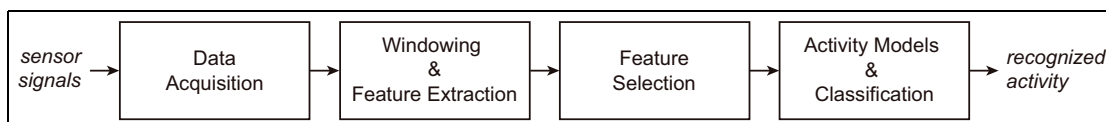
In order to recognize activities, HAR systems can be developed under the following approaches:<sup>2,8,9</sup> sensor-based systems, vision-based systems, acoustic-based systems, and multimodal. Each approach has advantages and weaknesses<sup>9</sup> as well as challenges such as the number of sensors used, the localization of the devices, the organization of the data retrieved, the signal processing methods used, and others. From the point of view of signal processing, in general, HAR systems perform the following steps: data acquisition, windowing, feature extraction, feature selection, building activity models, and classification.

In that sense, this article aims to present a review on the different steps of signal processing in the context of HAR. It provides two examples and applications in signal processing dedicated to HAR and robotics: (1) HRI and socialization, highlighting the importance of recognizing human activities for the context of robots and how those interact in a social way with humans; and (2) imitation learning, as one high-level task in robots to extract knowledge from humans to learn and later to recognize and to perform demonstrated actions for better interaction between humans and robots.

There are many surveys on HAR,<sup>4,9,10</sup> HRI,<sup>11,12</sup> and imitation learning.<sup>13</sup> Hence, to our knowledge, there is no review that focuses on the application of HAR in these fields.

## Signal processing in HAR

The intention of this section is to provide a general overview of signal processing in HAR. In that sense, we adopted the methodology reported in Bulling et al.<sup>14</sup> and complemented in Ponce et al.,<sup>5</sup> namely, activity recognition chain (ARC) approach, that describes the full process in HAR. The HAR methodology is shown in Figure 1. Each step is explained below.



**Figure 1.** Block diagram of the main steps in a human activity recognition system.

## Data acquisition

The source of data is relevant in HAR. Particularly, there are different types of sources: inertial sensors, microphones, infrared sensors, video, Kinect, brainwave-based helmet, force sensors, smart-watches, and mobile phones, among others. In any case, different challenges must be faced to acquire data and extract relevant information from them. These challenges are related to signal processing and inherent to HAR. Challenges specific to HAR are the selection of number and characteristics of the subjects mainly in real-world environments. Regarding sensors, we can mention selection of type of sensors, location of the sensors, calibration, synchronization of several sensors, and noise.<sup>2</sup>

There are several issues related to determining the source of data.<sup>2</sup> There are multiple affordable modalities and types of sensors well-suited for HAR as mentioned above: wearable and body sensors, smart phones and bracelets, cameras, depth sensors, motion capture sensors, ambient sensors, and so on. An important challenge is which type of sensor or combination of sensors to select for a given activity recognition task. In addition, the localization of them (in the body, in the environment, in devices, etc.) is another important task to be performed beforehand. Each activity recognition system can rely on different sensor signals with particular settings and locations.

Data synchronization, data pre-processing, and data inconsistency are inherent challenges during data acquisition stage.<sup>15,16</sup> Since multiple heterogeneous sensors are used for activity recognition, data synchronization and data buffering are always a challenge given the diverse sampling rates, sensing periodicity, and sensor platforms.

Data inconsistency is also a common issue to be addressed due to wireless communication, deprivation (e.g. sensor failure), limited spatial coverage, imprecision from individual sensors, and uncertainty when features are missing.<sup>17</sup> Multi-sensor fusion has been presented as a solution to deal with individual sensor issues.<sup>17</sup>

Several studies<sup>18</sup> have shown that noise in data in particular affects the performance of the overall system, falling in poor performance of activity recognition. This is not only a concern to HAR developers but also, in general, to people working on signal processing.

**Table 1.** Features extracted in time domain, adapted from Ponce et al.<sup>5</sup>

Features	References
Mean	Bulling et al., <sup>14</sup> Phinyomark et al., <sup>19</sup> Avci et al., <sup>20</sup> Dargie, <sup>21</sup> Rasekh et al., <sup>22</sup> Atallah et al., <sup>23</sup> Preece et al. <sup>24</sup>
Standard deviation	Avci et al., <sup>20</sup> Dargie, <sup>21</sup> Preece et al. <sup>24</sup>
Root mean square	Phinyomark et al. <sup>19</sup>
Maximal amplitude	Dargie, <sup>21</sup> Atallah et al. <sup>23</sup>
Minimal amplitude	Dargie, <sup>21</sup> Atallah et al. <sup>23</sup>
Median	Rasekh et al., <sup>22</sup> Preece et al. <sup>24</sup>
Number of zero-crossing	Phinyomark et al., <sup>19</sup> Dargie <sup>21</sup>
Skewness	Atallah et al. <sup>23</sup>
Kurtosis	Bulling et al., <sup>14</sup> Atallah et al. <sup>23</sup>
First quartile	Rasekh et al., <sup>22</sup> Preece et al. <sup>24</sup>
Third quartile	Rasekh et al., <sup>22</sup> Preece et al. <sup>24</sup>
Autocorrelation	Dargie, <sup>21</sup> Atallah et al. <sup>23</sup>

Furthermore, real-world applications of HAR consider noisy data caused by several reasons, such as device miscalibration, dead or blocked devices, localization of devices (body and environmental), noisy environments, and interleaved activities.

### Windowing and feature extraction

Once the data are acquired by the sensors, this information has to be processed. To do that, the sensor signals can be divided in variable or fixed time windows,<sup>2,3</sup> and time windows can also be overlapping or disjoint.<sup>2</sup>

Fixed windowing is the most reported approach in HAR.<sup>2,3,5,14,18</sup> At each window, there is information contained that is then employed to infer a human activity. Actually, this information from window to window should be compared among them to determine the right activity. However, comparing segments of sensor signals contained in different windows is considered complicated. Instead, representative information is measured inside windows, so they can be quantitatively comparable.<sup>2</sup> Particularly, the latter is known as feature extraction.

There are different features extracted for HAR systems reported in the literature. Table 1 summarizes common time-domain features and Table 2 lists common frequency-domain features. As noted, these measurements are mostly statistical quantities that contain the overall information of the signal in a window.

Although features enable the system the capability to model and detect activities in a more efficient way than in raw signals, conventional feature extraction has several drawbacks identified in Wang et al.:<sup>10</sup> (1) feature extraction relies on human expertise and implies longer time to build an activity recognition system; (2)

**Table 2.** Features extracted in frequency domain, adapted from Ponce et al.<sup>5</sup>

Features	References
Mean frequency	Phinyomark et al., <sup>19</sup> Dargie <sup>21</sup>
Median frequency	Phinyomark et al. <sup>19</sup>
Entropy	Avci et al., <sup>20</sup> Rasekh et al. <sup>22</sup>
Energy	Avci et al., <sup>20</sup> Rasekh et al., <sup>22</sup> Bulling et al. <sup>14</sup>
Principal frequency	Rasekh et al., <sup>22</sup> Atallah et al., <sup>23</sup> Preece et al. <sup>24</sup>
Spectral centroid	Dargie, <sup>21</sup> Rasekh et al. <sup>22</sup>

shallow features (see Table 1) can be used to recognize low-level activities, but it is difficult to recognize high-level activities; and (3) conventional approaches need labeled data to train models in contrast to deep generative networks which train models from unsupervised examples. For instance, deep learning can alleviate the effort on designing features.

### Feature selection

Once the feature extraction is done, there might be some issues: (1) redundancy in features and (2) too many features computed. In the first case, there might be different features that are redundant so they can influence positively or negatively in the classification of activities. In practice, there are different methods that are used for selecting a suitable set of features. For example, literature reports the usage of Bayesian information criterion, minimum description length, minimum redundancy and maximum relevance, and correlation-based feature selection, among others.<sup>2</sup>

Not only the redundancy but also the large number of features can affect the recognition accuracy of activities. In that sense, feature reduction, such as principal component analysis,<sup>5</sup> is employed to determine how the set of features can be minimized in the number of elements, or to compose another subset with less features.

At the end, feature selection and reduction are applied to the feature extraction in order to get a new dataset of relevant features.

### Building activity models and activity classification

Next, an activity model has to be designed. This model contains the information and the methods to classify the sensor signals for determining the different activities performed by a human.

In this regard, machine learning has been widely used in HAR systems to build patterns and to analyze, describe, and predict data.<sup>2</sup> Particularly to HAR, supervised learning is applied and implemented by using a set of training data with instances (observations from

sensors) and labels (e.g. activities) that allow the system to learn patterns for classification.<sup>25,26</sup> At the end, these learning methods build activity models.

Literature reports different supervised machine learning methods employed in HAR:<sup>2,3,5,27,28</sup> stochastic gradient boosting, AdaBoost, decision trees,<sup>29</sup> rule-based classifier, single rule classification, support vector machines (SVM), random forest,  $k$ -nearest neighbors (KNN),<sup>30</sup> discriminant analysis, adaptive regression splines, naive Bayes,<sup>31</sup> artificial neural<sup>29</sup> deep learning,<sup>10</sup> and more recently artificial hydrocarbon networks.<sup>3,5</sup>

Finally, these models are employed to recognize the activities performed by the users. Since windows are normally small in contrast to the time that an activity is performed, then a classical methodology to summarize the estimated activity per window is the majority voting approach.<sup>2</sup> It considers to output the most frequent estimated activity in a sequence of windows. This technique greatly improves the accuracy in recognition, as shown in Ponce et al.<sup>5</sup>

## HAR in robotics

The interest in HAR research is growing rapidly in the past decades due to the great versatility in the application of action, activity, and behavior understanding. In different surveys of HAR, Lara and Labrador<sup>2</sup> and Aggarwal and Ryoo<sup>32</sup> discuss the applicability of HAR in medical, security, entertainment, military and tactical, and human interface systems.

In the healthcare and medical domain, HAR enables monitoring patients, elderly and/or children. HAR is an important task for assistive living applications where daily and abnormal activities are identified to aid elderly people. Avci et al.<sup>20</sup> also reviewed several medical applications of activity recognition for healthcare, well-being, and sports systems.

Recognition of human activities is necessary in security systems for intrusion detection, public places surveillance, abnormal activity detection, and video analysis. In Vishwakarma and Agrawal,<sup>33</sup> authors provide a survey for activity recognition in video surveillance. The authors identify diverse applications of HAR for surveillance, namely, behavioral biometrics, content-based video analysis, automatic recognition of abnormalities, interactive applications, learning, and simulation environments.

Tactical and military applications are closely related to the security applications. HAR can be useful for training of military soldiers and public service personnel, soldiers' activity recognition, location, and health condition.

All these previously mentioned applications can be considered from the robotics point of view since robots

must recognize human activities in order to socialize with persons<sup>34–36</sup> and for imitation learning purposes.<sup>37</sup> Robots perform many tasks to improve human life in health care, security, entertainment, military, and tactical domains. In different situations, robots must perceive, represent, recognize, and understand from simple human actions and gesture to complex activities and behaviors. Ke et al.<sup>38</sup> described important actions and behaviors to be recognized in vision systems. As robots are intended to interact in surveillance, assistive health and safety, and entertainment situations, similar human activity and behaviors must be identified and understood by robots. Table 3 summarizes activity recognition of one person, multiple people, and crowd behavior for different application systems.

Krüger et al.<sup>37</sup> analyzed action recognition from three points of view: computer vision, robotics, and artificial intelligence (AI) community. They discussed the interpretation and recognition of actions in robotics for robot movement and imitation learning. Other authors highlight the importance of HAR in robots for a successful HRI and socialization.<sup>35,39,40</sup>

In the rest of the article, we analyze two applications in which HAR is indispensable from the robotics' point of view: (1) HRI and socialization and (2) imitation learning.

## HRI and socialization

Goodrich and Schultz<sup>34</sup> define that the HRI problem is to “understand and shape the interactions between one or more humans and one or more robots.” The authors distinguish applications that require mobility, physical manipulation, or social interaction. Industrial and agricultural areas also benefit by applying HRI.<sup>11,12</sup> Action and activity recognition is relevant in these categories of applications, but they are crucial for robots' social interaction.

Depending on the autonomy of the robot, the interaction with a human can vary from direct control and teleoperation to dynamic autonomy, in which robots interact as partners, peers, or assistants. In the full range of possible scenarios of HRI, action and activity recognition can be useful—from simple command understanding for robot control to peer-to-peer interaction.

Social robots must be able to perceive and interpret the world as humans do.<sup>40</sup> Service robots interact directly with people, so it is important to find natural and easy-to-use interfaces.<sup>36</sup> Fong et al.<sup>40</sup> emphasize that social robots' perceptions need to be human-oriented: “optimized for interacting with humans and on a human level.”<sup>40</sup> Depending on the application nature, robots must be able to track human bodies, faces, and hands. They need to represent, recognize,

**Table 3.** Human activities recognized in different application domains.

HAR approach	Surveillance	Health/safety domains	Entertainment
Single target	<ul style="list-style-type: none"> <li>• People detecting and tracking</li> <li>• Loitering</li> <li>• Staking</li> <li>• Following</li> <li>• Dropping items</li> <li>• Human pose estimation</li> <li>• Gait recognition</li> </ul>	<ul style="list-style-type: none"> <li>• Falling detection</li> <li>• Human pose estimation</li> <li>• Abnormalities</li> <li>• Daily life activity monitoring</li> <li>• Respiration behavior</li> <li>• Hand and mouth movements</li> <li>• Medication intake</li> <li>• Drowning</li> <li>• Object interaction</li> <li>• Rehabilitation activities/exercises</li> <li>• Non-verbal commands</li> <li>• Walking gait monitoring</li> <li>• Group interactions</li> </ul>	<ul style="list-style-type: none"> <li>• Trajectory tracking</li> <li>• Human pose estimation (hand, head, feet)</li> <li>• Sports actions/sequences</li> <li>• Dance movements</li> <li>• Gestures</li> <li>• Tracking people and their actions for games</li> </ul>
Multiple targets	<ul style="list-style-type: none"> <li>• People counting</li> <li>• Group tracking</li> <li>• Violent behavior</li> <li>• Criminal activities</li> </ul>		<ul style="list-style-type: none"> <li>• Group interactions</li> </ul>
Crowd behavior	<ul style="list-style-type: none"> <li>• People count estimation</li> <li>• Crowd understanding</li> <li>• Normal/abnormal behavior</li> </ul>		

HAR: human activity recognition.

and interpret speech, facial expressions, gestures, and human activity. Human-oriented perception implies detecting and understanding, according to the needs, from simple commands with gestures or actions to complex activities and behaviors.

According to Goodrich and Schultz,<sup>34</sup> one important robot design decision is the way information is exchanged between a human and a robot. The primary media of communication between human and robots are based on three senses: seeing, hearing, and touching.<sup>34</sup> Despite the differences of the nature of interaction between human and robot, in all scenarios, the robot needs sensory inputs to capture movements and intelligence to understand the meaning of those movements.<sup>37</sup> Some applications are based on speech recognition, but in this section, we focus on human action and activity recognition performed by a robot. We present an overview of representative works that use different types of perceptions.

### *Vision-based approaches*

Vision-based approach is the most frequently used for activity recognition by robots. A good example of action recognition for HRI is robots used for assisted living. Robots monitor persons in daily life and identify their activities and abnormal events, and improve human life particularly in the case of older adults. Assisted robots usually rely on vision systems. Stavropoulos et al.<sup>41</sup> present a HAR method for assisted robots based on EigenJoints descriptor.<sup>42</sup> The authors use information extracted from depth input video from color and depth (RGB-D) cameras.<sup>43</sup>

Assisted robots have also been used as therapy tools for autism.<sup>44</sup> In cases of autism, the robot must sense and recognize body parts and motion of the child in order to imitate the movement accurately.<sup>44</sup> Similarly, Fasola and Mataric<sup>45</sup> present the design, implementation, and evaluation analysis of an assisted robot that motivates elderly users to engage in physical exercise. They used a vision-based robot to recognize arm gestures and poses in real time.

In order for a robot to interact with humans, it must understand humans' intentions, behaviors, and even emotions. This recognition and interpretation is performed having robot as an observer enabling it to react if it is threatened by human actions or has an adequate social response to human behavior. Xia et al.<sup>46</sup> referred to this problem as robot-centric activity recognition and presented a framework and algorithm to analyze RGB-D videos captured by the robot while interacting with humans in daily living environments.

Sidobre et al.<sup>47</sup> presented a solution to exchange objects between a human and a robot based on motion capture technology. They show that in order to exchange objects between a human and a robot in a natural way, the robot must be capable to adapt to human motion and gasp in real time. Robot motion has to be executed in human-aware way to avoid collision and ensure human safety.

Gesture recognition is an important task in HRI that requires recognition of motions of human body parts. For long time, gesture recognition has been used to allow HRI.<sup>48</sup> For instance, Mitra and Acharya<sup>49</sup> identify the different body parts involved in gesture identification, for example, (1) motions of hands and arms are

helpful for interpretation of sign language, (2) head and face motions are related to gestures about nodding, or (3) body motions as a whole can be useful for tracking or analyzing people moving or interacting. HRI using gesture recognition can be found when developing medical hearing impaired devices and automatically monitoring emotional states and stress levels in patients, or when monitoring drivers' alertness and drowsiness, as well as lie detection.<sup>49</sup>

### Sensor-based approaches

There are two main approaches for gesture recognition: computer vision techniques and cameras,<sup>50</sup> and sensor-based approaches.<sup>51,52</sup> As an example, Rautaray and Agrawal<sup>50</sup> identified hand gesture recognition as a core application for controlling and programming robots, such that these can imitate motions and interact with humans.

A wide variety of sensors are used to perceive face, hands, and body motion.<sup>53</sup> Sensor-based approaches are also commonly used for HAR. Zhu and Sheng<sup>54</sup> proposed a robot-assisted living system for elderly people, patients, and disabled. Their method is based on the fusion of wearable inertial sensors to achieve activity recognition combining neural networks and hidden Markov models. A new robot-automated semantic mapping system was presented in Sheng et al.<sup>55</sup> based on wearable sensors. The aim of the system is to enable a robot to build metric maps and identify furniture. Wearable motion sensors attached to the human body were used to recognize several daily human activities. The authors avoid using vision approaches to overcome their high computational cost, environment, light, and occlusion possible problems.

### Acoustic-based approaches

Although vision approaches are most commonly used by robots for activity recognition, sound produced by humans or human-object interaction provides rich information about ongoing context, events, and behaviors.<sup>56</sup> Acoustic signals can be used as a complement of vision information or as the main source of data for activity recognition. Maxime et al.<sup>56</sup> address the problem of sound and recognition of domestic events with humanoid robot NAO.<sup>57</sup> They performed a comparative analysis of several classification methods in order to identify events occurring in the domestic environment. Stork et al.<sup>35</sup> developed a method to classify human activities from the sounds of them in order to enable a robot to infer those activities. They proposed the non-Markovian ensemble voting method<sup>35</sup> to classify 22 human activities performed in the bathroom and kitchen contexts.

### Tactile-based approaches

Tactile HRIs have also gained interest as a way of keeping a safe operation of the robot around humans, as a way of the human to guide or partner with the robot, and as a necessary element for robots' behavior development.<sup>58,59</sup>

A good review of what is detected and how the perceptions are used can be found in Argall and colleagues.<sup>58,59</sup> Human and robot tactile interactions can be unexpected or unintended interfering with the robot's behavior execution. The robot must be able to recover from disturbance and react to physical contact. In these cases, the robot may passively react to the interference.<sup>60</sup> On the contrary, robots can actively predict the effects of human contact and identify the best behavior to operate safely and efficiently around humans.

When human tactile interaction with the robot is deliberate, the interpretation of human intentions by the robot is very important.<sup>58,59</sup> In social robots, particularly those made for "psychological enrichment,"<sup>61</sup> sensing and interpretation is very important to improve the human behavior.<sup>62</sup> Paro De Seal<sup>63</sup> and Kasper<sup>64</sup> are good examples.

### Multimodal-based approaches

The majority of related works only analyze human actions and events from one point of view. Rodomagoulakis et al.<sup>8</sup> presented a multimodal action recognition system in assistive HRI for elderly persons. Based on inputs from microphone and visual input of high definition and depth cameras, the robot can recognize audiovisual human commands. Accordingly, Han et al.<sup>65</sup> discussed methods and technologies for non-verbal HRI with the NAO robot.<sup>57</sup> They reviewed build-in technologies to detect face and head, and track people; sonar sensors and speech direction detection; and tactile sensors and object recognition. Likewise, Pieropan<sup>66</sup> argued that robots must be able to perceive and understand the environment autonomously out of multiple perception modalities, as humans do. The authors proposed a method for audiovisual recognition of human manipulation actions. In addition, they presented an RGB-D-audio dataset of humans making milk and cereals.

### Imitation learning

In the last decades, robots have been gaining ground in human lives. Robots cannot only be found in the industry, but a new generation of service and social robots appear in everyday life environments.<sup>67,68</sup> Social robots must relate to humans and cooperate in new scenarios, doing different tasks in possible unknown situations.

Hence, robots should be able to adapt and learn new behaviors.

Bandera et al.<sup>68</sup> say that robots must be provided with a natural and intuitive learning system for humans that enable them to expand their behavior repertoire quickly and efficiently. Try and error approaches require reward functions for each task, and even in simple tasks, the possibilities grow exponentially.<sup>13</sup> One of the needs for research on robot imitation learning is the intuitive way of communication between those and humans. Imitation learning is a key technology for applications, where robots are expected to work closely with humans, such as manufacturing, elder care, and service industry. Osa et al.<sup>69</sup> and Schaal<sup>70</sup> state that imitation learning means to speed up learning of humanoid robots. “The goal of imitation learning is to develop robot systems that are able to relate perceived actions of another (human) agent to its own embodiment in order to learn and later to recognize and to perform the demonstrated actions.”<sup>37</sup>

According to Krüger et al.,<sup>37</sup> for a robot to imitate other agent’s (human or not) actions and activities, it must be able to perceive, analyze, and recognize continuous human movements and actions. This entails also identifying objects relevant to a task, and analyzing the changes in the environment caused by human actions.

The perception of movement is most commonly done with vision-based inputs;<sup>68,71,72</sup> hence, these inputs can be complemented or replaced with magnetic tracking systems,<sup>73</sup> motion capture technologies, or proximity and infrared sensor technologies.

Mainly, a robot has to identify what, how, when, and who to imitate. The expert is also acting in a given context and environment sometimes interacting with other people or objects.

### What to imitate

First of all, a robot must decide what to imitate when perceiving a demonstration. An important research problem is how the robot determines what observed movements and actions of the selected model are relevant to the task, and what movements are only circumstantial.<sup>74</sup> The robot must also understand what effects certain actions have on the environment of the actor.<sup>37</sup> It has to recognize objects and their manipulation.

An interesting approach to determine what is observed was presented in Pieropan.<sup>66</sup> He proposed an audiovisual approach to understand an observed action. This approach tries to mitigate visual limitations (occlusion, luminosity, actions performed out of sight) with multimodal perception. He also focused on learning manipulation activities and object discovery, tracking, and object affordance for objects that can be grasped.

### When and who to imitate

According to Breazeal and Scassellati,<sup>74</sup> the robot must not only decide what to imitate but also when and who to imitate. Depending on the social context, availability of a good model, and robot’s internal motivation, the robot decides to engage in imitation. In the data recollection phase, common problems are noise and sensor errors, and incomplete or inaccurate demonstration.

In Burns et al.,<sup>75</sup> the experimenter aims to determine the effect of robotic imitation to increase the engagement of the human participant during a structured social setting. This experiment was not a task-oriented interaction.

### How to imitate

Once the robot has identified what to imitate, it has to determine how to imitate the observed action; it has to map these movements and actions into movement that its body can perform. Calinon et al.<sup>72</sup> addressed the problems of “what to imitate” and “how to imitate.” The first problem is to identify which features are relevant for achieving the task, and the second problem refers to transferring the perceived action into primitive movements performed with the robots body. The authors presented a method to extract the constraints of a task in order to determine an imitation strategy. They used a humanoid platform to show a goal-directed imitation.

Nehaniv and Dautenhahn<sup>76</sup> define this as the correspondence problem. In order to physically imitate certain action, the robot has to map the observed action-generating movements (primitive movements)<sup>58,59</sup> that the robot is capable to perform with its embodiment. Argall et al.<sup>58,59</sup> use the word *demonstration* when there is no difference in the agents’ body, and therefore, it is no embodiment mapping issue between the teacher and the learner. The author refers to *imitation* when the agents are not identical, so embodiment issues do exist between the teacher and the learner. The authors divide approaches in those obtaining imitation data from sensors on the teacher, and approaches that obtain imitation data from external observation (sensors that may or may not be in the learner).

Some authors apply different kinematic-based approaches to deal with the correspondence problem. These approaches perform offline optimization to determine the corresponding configurations<sup>77</sup> or real-time techniques for imitation.<sup>78</sup> Jin et al.<sup>79</sup> proposed a framework based on sparsely sample correspondences extracted from raw data that allow imitation in real time.

### Where to imitate

One step further for the robot is to understand the context in which the actions are been made. Chella et al.<sup>80</sup>

proposed a framework to not just reproduce movements of a human teacher but also understand the environment and perceived actions. Their cognitive architecture for imitation learning is able to learn natural movements and generate action plans.

A different approach for physical HRI is<sup>81</sup> an efficient machine learning algorithm and two human-in-the-loop learning scenarios inspired by human parenting behavior. The test subject was asked to assist a robot in a standing-up and walking assistance task improving the HRI.

Finally, the robot must be able to learn and improve its performance over time. Eventually, the robot must be able to evaluate its actions and establish the similarity of the outcome of its actions in comparison to the observed demonstration.

## Challenges and trends

According to Fong et al.,<sup>40</sup> social interactive robots are important because in some domains, robots must interact with humans as peer to perform a specific task or when it is used to change human attitudes or behavior. In order to achieve this peer-to-peer relation with humans, robots must develop social skills. Fong et al.<sup>40</sup> identified important design issues in social robot systems. In particular, human-oriented perception, natural HRI, and real-time performance are design issues closely related to the domain of HAR.

### Human-oriented perception

A great variety of activities have to be recognized by robots according to different application systems (see Table 3). As social robots are becoming part of our daily lives, they have to be able to recognize and to understand simple gestures related to human activities, object manipulation, complex human behaviors, and social conventions.

Several approaches based on different types of perceptions are used for HRI, and each of them has advantages and limitations. RGB-D systems are of low cost, but they are highly sensitive to lighting changes and environmental factors. These approaches are limited on viewpoint changes. When cameras are integrated to robots, a special challenge is the moving cameras, as reported in Yazdi and Bouwmans.<sup>82</sup> Motion capture systems find good representation of the action, but they are expensive and have high computational costs. The systems require markers and calibration, but they are usually invasive and have to be used in controlled spaces.<sup>67,68</sup>

Wearable sensors are obtrusive and uncomfortable if they have to be worn for prolonged periods of time. They are limited for certain activities, sensitive to sensor location, and can produce noisy data. On the contrary,

sensors do not have all vision drawbacks (e.g. occlusion, fixed location and fixed views, blurring, external conditions such as lightning, high amount of data to process), and they can be used in different environments. We can observe a trend to combine different types of perceptions in multimodal systems<sup>83</sup> to overcome the mentioned limitations at the cost of increasing the computational complexity.

### Natural HRI

In order to interact with humans as peers, robots must communicate and act in a friendly way to humans. Human behavior is complex and rich, and robots must understand and act according to social conventions and norms.<sup>40</sup> The robot's behavior must be believable. Natural embodiment, natural language and dialog, smooth natural motions, and understanding of human's emotions are desirable. Some of these requirements become relevant depending on the application, and they increase the performance complexity. According to Ishiguro and Nishio,<sup>84</sup> HRI researchers have been neglected the appearance of a robot prioritizing behavior over appearance. There is a shift in recent years to build more humanoid robots who interact with natural communication with humans. Sophia humanoid from Hanson Robotics is the most mentioned example of this trend.<sup>85</sup>

Transfer learning for HAR<sup>86</sup> can be helpful to cope with the correspondence problem in HRI.

### Real-time performance

“Socially interactive robots must operate at human interaction rates.”<sup>40</sup> This means that a robot must be able to identify and interpret activities and situations, plan and take decisions, and learn as rapidly as humans do.

Regarding the learning techniques, hidden Markov models are frequently used to train the robot's policy when direct state-action is required, but this is often not enough for different behaviors.<sup>13</sup> Bandera et al.<sup>68</sup> described the following drawbacks for this approach: (1) the complexity of training and inference limits the number of states that can be modeled and (2) the huge amount of training data needed for this approach. Some commonly used methods for classification and regression are artificial neural networks (ANN), KNN, locally weighted regression (LWR), and SVM.

Imitation learning is an important attempt to speed up a robot's learning.<sup>70</sup> Hussein et al.<sup>13</sup> reviewed the learning methods used for imitation learning which are generic to robot motion learning tasks. The authors claim that there is a need for specialized learning algorithms to represent and predict human action in order to be able to emulate these motor functions. In several



applications of human activity, recognition data are collected with vision, sensor, or multimodal perception, and this information is processed to obtain features associated to labels. These features and labels are the learning examples for activity recognition.

In imitation learning, there are states that represent the status of an agent and actions. Imitation learning depends on the expert demonstrations, repeated interactions from which sequential prediction must be achieved.<sup>87</sup> The learning process for imitation learning is capturing actions of the teacher via different sensing methods. The state-action is captured using wearable sensors, motion capture systems, force gloves, movement and position sensors, tactile devices, and cameras, among others. This information is also processed to extract features that describe the state of the performer and task-related information of the surrounding. From these features, the robot must learn a policy to imitate the demonstrated behavior.<sup>13</sup> The policy can be refined for continuous improvement.

According to Hussein et al.,<sup>13</sup> classification, regression, and apprenticeship learning methods are used for learning the policies that determine action units or movement primitives<sup>59</sup> that the robot must do. For policy refinement, active learning, apprenticeship learning, reinforcement learning, transfer learning, structured prediction, and optimization can be used.<sup>13</sup>

There is a need on the timing response of learning techniques for HRI, since the performance should be done in very limited amount of time, typically in real time. In that sense, learning techniques have issues on training time, mostly offering only offline options for robots. But, there is an urge to propose novel methods for online training. In terms of implemented learning models previously trained, there are too many methods with fast and efficient response. In this regard, future on HRI is mostly focused on new paradigms not only in training learning techniques but also in the implementation of signal processing techniques at low level.

The important design issues discussed above lead us to think about the balance between robot's complexity and computational and investments costs. The more complex multimodal perceptions and more natural and complex communication and behavior of the robot demand more computational complexity and more costs. This also means that real-time performance is still an issue for complex robot needs. There will always be a trade-off between the complex multimodal perceptions, and natural robot's behavior and real-time performance, which has to be analyzed for each application.


### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research has been funded by Universidad Panamericana through the grant “Fomento a la Investigación UP 2018,” under project code UP-CI-2018-ING-MX-04.

### ORCID iD

Hiram Ponce  <https://orcid.org/0000-0002-6559-7501>

### References

1. Loreti D, Chesani F, Mello P, et al. Complex reactive event processing for assisted living: the habitat project case study. *Expert Syst Appl* 2019; 126: 200–217.
2. Lara OD and Labrador MA. A survey on human activity recognition using wearable sensors. *IEEE Commun Surv Tutor* 2013; 15(3): 1192–1209.
3. Ponce H, Martínez-Villaseñor L and Miralles-Pechuán L. A novel wearable sensor-based human activity recognition approach using artificial hydrocarbon networks. *Sensors* 2016; 16(7): E1033.
4. Kong Y and Fu Y. Human action recognition and prediction: a survey, 2018, <https://arxiv.org/abs/1806.11230>
5. Ponce H, Miralles-Pechuán L and Martínez-Villaseñor L. A flexible approach for human activity recognition using artificial hydrocarbon networks. *Sensors* 2016; 16(11): 1715.
6. Ponce H, Martínez-Villaseñor L and Miralles-Pechuán L. Comparative analysis of artificial hydrocarbon networks and data-driven approaches for human activity recognition. In: *International conference on ubiquitous computing and ambient intelligence*, Puerto Varas, Chile, 1–4 December 2015, pp.150–161. Berlin: Springer.
7. Vrigkas M, Nikou C and Kakadiaris I. A review of human activity recognition methods. *Front Robot AI* 2015; 2: 28.
8. Rodomagoulakis I, Kardaris N, Pitsikalis V, et al. Multimodal human action recognition in assistive human-robot interaction. In: *International conference on acoustics, speech and signal processing*, Shanghai, China, 20–25 March 2016, pp.2702–2706. New York: IEEE.
9. Ramasamy Ramamurthy S and Roy N. Recent trends in machine learning for human activity recognition: a survey. *Wiley Interdiscip Rev* 2018; 8(4): e1254.
10. Wang J, Chen Y, Hao S, et al. Deep learning for sensor-based activity recognition: a survey. *Patt Recog Lett* 2019; 119: 3–11.
11. Villani V, Pini F, Leali F, et al. Survey on human-robot collaboration in industrial settings: safety, intuitive interfaces and applications. *Mechatronics* 2018; 55: 248–266.
12. Vasconez JP, Kantor GA and Cheein FAA. Human-robot interaction in agriculture: a survey and current challenges. *Biosyst Eng* 2019; 179: 35–48.
13. Hussein A, Gaber MM, Elyan E, et al. Imitation learning: a survey of learning methods. *ACM Comput Surv* 2017; 50(2): 21.

14. Bulling A, Blanke U and Schiele B. A tutorial on human activity recognition using body-worn inertial sensors. *ACM Comput Surv* 2014; 46: 1–33.
15. Chen C, Jafari R and Kehtarnavaz N. A survey of depth and inertial sensor fusion for human action recognition. *Multimed Tools Appl* 2017; 76(3): 4405–4425.
16. Koshmak G, Loutfi A and Linden M. Challenges and issues in multisensor fusion approach for fall detection. *J Sensors* 2016; 2016: 6931789.
17. Gravina R, Alinia P, Ghasemzadeh H, et al. Multi-sensor fusion in body sensor networks: state-of-the-art and research challenges. *Inform Fusion* 2017; 35: 68–80.
18. Nettleton D, Orriols-Puig A and Fornells A. A study of the effect of different types of noise on the precision of supervised learning techniques. *Artif Intell Rev* 2010; 33: 275–306.
19. Phinyomark A, Nuidod A, Phukpattaranont P, et al. Feature extraction and reduction of wavelet transform coefficients for EMG pattern classification. *Elektron Elektrotech* 2012; 122(6): 27–32.
20. Avci A, Bosch S, Marin-Perianu M, et al. Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: a survey. In: *Proceedings of the 23rd international conference on architecture of computing systems*, Hannover, 22–23 February 2010, pp.1–10. New York: IEEE.
21. Dargie W. Analysis of time and frequency domain features of accelerometer measurements. In: *Proceedings of 18th international conference on computer communications and networks (ICCCN)*, San Francisco, CA, 3–6 August 2009, pp.1–6. New York: IEEE.
22. Rasekh A, Chen CA and Lu Y. Human activity recognition using Smartphone, 2014, <https://arxiv.org/abs/1401.8212>
23. Atallah L, Lo B, King R, et al. Sensor placement for activity detection using wearable accelerometers. In: *International conference on body sensor networks*, Singapore, 7–9 June 2010, pp.24–29. New York: IEEE.
24. Preece SJ, Goulermas JY, Kenney LP, et al. A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data. *IEEE Trans Biomed Eng* 2009; 56(3): 871–879.
25. Wang X, Xu Y, Hu H, et al. Feedback-based metric learning for activity recognition. *Expert Syst Appl*. Epub ahead of print 10 September 2018. DOI: 10.1016/j.eswa.2018.09.021.
26. Hossain HMS, Roy N and Khan MAAH. Active learning enabled activity recognition. In: *International conference on pervasive computing and communications (PerCom)*, Sydney, NSW, Australia, 14–19 March 2017, pp.312–330. New York: IEEE.
27. Preece SJ, Goulermas JY, Kenney LP, et al. Activity identification using body-mounted sensors—a review of classification techniques. *Physiol Measure* 2009; 30(4): R1–R33.
28. Roggen D, Calatroni A, Rossi M, et al. Collecting complex activity datasets in highly rich networked sensor environments. In: *Proceedings of the 7th international conference on networked sensing systems (INSS)*, Kassel, 15–18 June 2010, pp.233–240. New York: IEEE.
29. Dohnálek P, Gajdoš P, Moravec P, et al. Application and comparison of modified classifiers for human activity recognition. *Prz Elektrotechniczny* 2013; 89(11): 55–58.
30. Altun K. *Intelligent sensing for robot mapping and simultaneous human localization and activity recognition*. PhD Thesis, Bilkent University, Ankara, 2011.
31. Guneyasu A and Arnrich B. Socially assistive child-robot interaction in physical exercise coaching. In: *Proceedings of the 26th international symposium on robot and human interactive communication (RO-MAN)*, Lisbon, 28 August–1 September 2017, pp.670–675. New York: IEEE.
32. Aggarwal JK and Ryoo MS. Human activity analysis: a review. *ACM Comput Surv* 2011; 43(3): 16.
33. Vishwakarma S and Agrawal A. A survey on activity recognition and behavior understanding in video surveillance. *Visual Comput* 2013; 29(10): 983–1009.
34. Goodrich MA and Schultz AC. Human-robot interaction: a survey. *Found Trend Human Comput Inter* 2007; 1(3): 203–275.
35. Stork JA, Spinello L, Silva J, et al. Audio-based human activity recognition using non-markovian ensemble voting. In: *IEEE RO-MAN: the 21st IEEE international symposium on robot and human interactive communication*, Paris, 9–13 September 2012, pp.509–514.
36. He W, Li Z and Chen CP. A survey of human-centered intelligent robots: issues and challenges. *J Automatica Sinica* 2017; 4(4): 602–609.
37. Krüger V, Kragic D, Geib C, et al. The meaning of action: a review on action recognition and mapping. *Adv Robot* 2007; 21(13): 1473–1501.
38. Ke SR, Thuc HLU, Lee YJ, et al. A review on video-based human activity recognition. *Computers* 2013; 2(2): 88–131.
39. Piyathilaka L and Kodagoda S. Human activity recognition for domestic robots. In: Mejias L, Corke P and Roberts J (eds) *Field and service robotics*. Berlin: Springer, 2015, pp.395–408.
40. Fong T, Nourbakhsh I and Dautenhahn K. A survey of socially interactive robots. *Robot Autonom Syst* 2003; 42(3): 143–166.
41. Stavropoulos G, Giakoumis D, Moustakas K, et al. Automatic action recognition for assistive robots to support MCI patients at home. In: *Proceedings of the 10th international conference on pervasive technologies related to assistive environments*, Rhodes, 21–23 June 2017, pp.366–371. New York: ACM.
42. Yang X and Tian Y. Eigenjoints-based action recognition using naive-bayes-nearest-neighbor. In: *Proceedings of the international conference on computer vision and pattern recognition*, Providence, RI, 16–21 June 2012, pp.14–19. New York: IEEE.
43. Buys K, Cagniart C, Baksheev A, et al. An adaptable system for RGB-D based human body detection and pose estimation. *J Visual Commun Image Represent* 2014; 25(1): 39–52.
44. Scassellati B, Admoni H and Mataric M. Robots for use in autism research. *Ann Rev Biomed Eng* 2012; 14: 275–294.

45. Fasola J and Mataric MJ. Using socially assistive human–robot interaction to motivate physical exercise for older adults. *Proc IEEE* 2012; 100(8): 2512–2526.
46. Xia L, Gori I, Aggarwal JK, et al. Robot-centric activity recognition from first-person RGB-D videos. In: *IEEE winter conference on applications of computer vision*, Waikoloa, HI, 5–9 January 2015, pp.357–364. New York: IEEE.
47. Sidobre D, Broquere X, Mainprice J, et al. *Human–robot interaction*. Berlin: Springer, 2012.
48. Liu H and Wang L. Gesture recognition for human-robot collaboration: a review. *Int J Indus Ergonomic* 2018; 68: 355–367.
49. Mitra S and Acharya T. Gesture recognition: a survey. *IEEE Trans Syst Man Cybernet C* 2007; 37(3): 311–324.
50. Rautaray SS and Agrawal A. Vision based hand gesture recognition for human computer interaction: a survey. *Artif Intell Rev* 2015; 43(1): 1–54.
51. Zhang X, Chen X, Li Y, et al. A framework for hand gesture recognition based on accelerometer and EMG sensors. *IEEE Trans Syst Man Cybernet A* 2011; 41(6): 1064–1076.
52. Zhu C and Sheng W. Wearable sensor-based hand gesture and daily activity recognition for robot-assisted living. *IEEE Trans Syst Man Cybernet A* 2011; 41(3): 569–573.
53. Chaudhary A, Raheja JL, Das K, et al. Intelligent approaches to interact with machines using hand gesture recognition in natural way: a survey, 2013, <https://arxiv.org/abs/1303.2292>
54. Zhu C and Sheng W. Human daily activity recognition in robot-assisted living using multi-sensor fusion. In: *International conference on robotics and automation*, Kobe, Japan, 12–17 May 2009, pp.2154–2159. New York: IEEE.
55. Sheng W, Du J, Cheng Q, et al. Robot semantic mapping through human activity recognition: a wearable sensing and computing approach. *Robot Autonom Syst* 2015; 68: 47–58.
56. Maxime J, Alameda-Pineda X, Girin L, et al. Sound representation and classification benchmark for domestic robots. In: *International conference on robotics and automation*, Hong Kong, China, 31 May–7 June 2014, pp.6285–6292. New York: IEEE.
57. Gouaillier D, Hugel V, Blazevic P, et al. Mechatronic design of NAO humanoid. In: *International conference on robotics and automation*, Kobe, Japan, 12–17 May 2009, pp.769–774. New York: IEEE.
58. Argall BD, Chernova S, Veloso M, et al. A survey of robot learning from demonstration. *Robot Autonom Syst* 2009; 57(5): 469–483.
59. Argall BD and Billard AG. A survey of tactile human–robot interactions. *Robot Autonom Syst* 2010; 58(10): 1159–1176.
60. Wosch T and Feiten W. Reactive motion control for human-robot tactile interaction. In: *International conference on robotics and automation*, Washington, DC, 11–15 May 2002, Vol. 4, pp.3807–3812. New York: IEEE.
61. Shibata T. An overview of human interactive robots for psychological enrichment. *Proc IEEE* 2004; 11: 1749–1758.
62. Silvera-Tawil D, Rye D and Velonaki M. Artificial skin and tactile sensing for socially interactive robots: a review. *Robot Autonom Syst* 2015; 63: 230–243.
63. Wada K and Shibata T. Living with seal robots—its sociopsychological and physiological influences on the elderly at a care house. *IEEE Trans Robot* 2007; 23(5): 972–980.
64. Robins B, Amirabdollahian F, Ji Z, et al. Tactile interaction with a humanoid robot for children with autism: a case study analysis involving user requirements and results of an initial implementation. In: *International symposium on robot and human interactive communication*, Viareggio, 13–15 September 2010, pp.704–711. New York: IEEE.
65. Han J, Campbell N, Jokinen K, et al. Investigating the use of non-verbal cues in human-robot interaction with a NAO robot. In: *Proceedings of the 3rd international conference on cognitive infocommunications*, Kosice, 2–5 December 2012, pp.679–683. New York: IEEE.
66. Pieropan A. *Action recognition for robot learning*. PhD Thesis, KTH Royal Institute of Technology, Stockholm, 2015.
67. Bandera JP. *Vision-based gesture recognition in a robot learning by imitation framework*. PhD Thesis, Universidad de Malaga, Málaga, 2010.
68. Bandera JP, Rodriguez JA, Molina-Tanco L, et al. A survey of vision-based architectures for robot learning by imitation. *Int J Humanoid Robot* 2012; 9(1): 1250006.
69. Osa T, Pajarinen J, Neumann G, et al. An algorithmic perspective on imitation learning. *Found Trend Robot* 2018; 7(1–2): 1–179.
70. Schaal S. Is imitation learning the route to humanoid robots? *Trend Cognitive Sci* 1999; 3(6): 233–242.
71. Mühlhig M, Gienger M, Hellbach S, et al. Task-level imitation learning using variance-based movement optimization. In: *International conference on robotics and automation*, Kobe, Japan, 12–17 May 2009, pp.1177–1184. New York: IEEE.
72. Calinon S, Guenter F and Billard A. Goal-directed imitation in a humanoid robot. In: *International conference on robotics and automation*, Barcelona, 18–22 April 2005, pp.18–22. New York: IEEE.
73. Asfour T, Gyurfas F, Azad P, et al. Imitation learning of dual-arm manipulation tasks in humanoid robots. In: *Proceedings of the 6th IEEE RAS international conference on humanoid robots*, Genova, 4–6 December 2006, pp.40–47. New York: IEEE.
74. Breazeal C and Scassellati B. *Challenges in building robots that imitate people*. Cambridge, MA: The MIT Press, 2002.
75. Burns R, Jeon M and Park C. Robotic motion learning framework to promote social engagement. *Appl Sci* 2018; 8(2): 241.
76. Nehaniv C and Dautenhahn K. *The correspondence problem*. Cambridge, MA: The MIT Press, 2002.
77. Ude A, Atkeson CG and Riley M. Programming full-body movements for humanoid robots by observation. *Robot Autonom Syst* 2004; 47(2): 93–108.
78. Dariush B, Gienger M, Arumbakkam A, et al. Online transfer of human motion to humanoids. *Int J Humanoid Robot* 2009; 6(2): 265–289.

79. Jin S, Dai C, Liu Y, et al. Motion imitation based on sparsely sampled correspondence. *J Comput Inf Sci Eng* 2016; 17: 041009.
80. Chella A, Dindo H and Infantino I. A cognitive framework for imitation learning. *Robot Autonom Syst* 2006; 54(5): 403–408.
81. Ikemoto S, Amor HB, Minato T, et al. Physical human-robot interaction: mutual learning and adaptation. *IEEE Robot Automat Mag* 2012; 19(4): 2435.
82. Yazdi M and Bouwmans T. New trends on moving object detection in video images captured by a moving camera: a survey. *Comput Sci Rev* 2018; 28(5): 157–177.
83. Katsamanis A, Pitsikalis V, Theodorakis S, et al. Multimodal gesture recognition. In: Oviatt S, Schuller B, Cohen PR, et al. (eds) *The handbook of multimodal-multisensor interfaces*. New York: Association for Computing Machinery; Morgan & Claypool, 2017, pp.449–487.
84. Ishiguro H and Nishio S. Building artificial humans to understand humans. *J Artif Organs* 2018; 10: 133–142.
85. Hanson Robotics. Hi, I Am Sophia, <https://www.hanson-robotics.com/sophia/> (2019, accessed 9 April 2019).
86. Cook D, Feuz KD and Krishnan NC. Transfer learning for activity recognition: a survey. *Know Inform Syst* 2013; 36(3): 537–556.
87. Kober J, Bagnell JA and Peters J. Reinforcement learning in robotics: a survey. *The Int J Robot Res* 2013; 32(11): 1238–1274.