ASIMO and Humanoid Robot Research at Honda

Satoshi Shigemi

Contents

1 Introduction 56
2 Creation of Mobile Entities that Embody New Value: Knowing and Learning from Humans 57
3 Mobility Capabilities 58
3.1 Honda's Unique Bipedal Walking Technology 58
3.2 From Walking to Running 66
4 Task-Performing Capabilities 68
5 Communication Capabilities 73
5.1 Voice Recognition 74
5.2 Image Recognition 75
5.3 Physical Expression 76
6 From Automatic to Autonomous 78
6.1 Sensing 79
6.2 Situation Estimation 79
6.3 Behavior Generation 82
6.4 Field Experiments 84
7 Technology for People 86
References 89

Abstract

One of the major characteristics of robot development at Honda is "knowing and learning from humans." In 1986, Honda started a research on robot whose bipedal walking was modeled after humans.

In this chapter, capabilities of Honda humanoid robots such as mobility, task-performing, and communication are introduced, and technologies which realized the above capabilities are explained.

S. Shigemi (✉)
Honda R&D Co., Ltd., Wako-shi, Saitama, Japan
e-mail: satoshi_shigemi@n.f.rd.honda.co.jp

A. Goswami, P. Vadakkepat (eds.), *Humanoid Robotics: A Reference*,
https://doi.org/10.1007/978-94-007-6046-2_9

With walk stabilization control technology which includes ground reaction force control, model ZMP control, and foot landing position control, biped robots could stably walk on uneven or slanted floors. Gait generation technology, which limits slipping and spinning, made it possible to assure dynamic stability during running.

In terms of task performance, by fusion of physical capabilities with recognition of the external environment using sensors of various kinds, the robot completed several tasks such as handing over a tray, pushing a cart, and pouring a drink.

Voice and image recognition technologies and an abundance of physical expressions enabled robots to interact with people in a natural way.

In order to behave properly in a real-world environment that is constantly changing, autonomous behavior generation technology has been developed. A system architecture called the intelligence loop was devised for this technology. The robot demonstrated this autonomy in two field experiments in the science museum where the robot made autonomous explanation to the visitors.

As applications of the robotics technology created in the course of humanoid robot research, High-Access Survey Robot which was sent to Fukushima Daiichi Nuclear Power Station and new mobility devices are briefly described.

1 Introduction

Taking on the challenge of creating a new mobile entity to follow from motorcycles, automobiles, and power products, Honda has been engaged in the development of humanoid robots since 1986 [1]. Figure 1 shows a brief history of Honda humanoid robot development. At first, the bipedal robot E0 could walk in a straight and static way. In 1993, a torso and two arms were added to complete the first truly humanoid robot P1. Subsequently, measures to reduce weight and size were started, and development of walking technology was continued for the robot to be useful in human living spaces.

This chapter introduces the aims of humanoid robot research at Honda beginning with the Advanced Step in Innovative Mobility (ASIMO) (Fig. 2), together with the core technology involved. Section 2 describes the motivation and the purpose of robot development at Honda. In Sect. 3, Honda's unique bipedal walking and

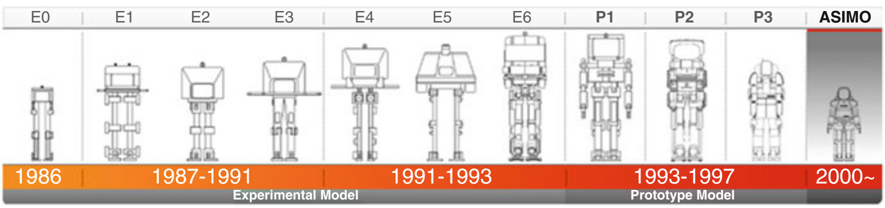


Fig. 1 History of Honda humanoid robot development

Fig. 2 Latest ASIMO

running technologies are introduced. Next, task-performing and communication capabilities are described in Sects. 4 and 5, respectively. In Sect. 6, the most important characteristics of the latest ASIMO are presented. Finally, the outlook section introduces recent robotic platforms [2] beyond ASIMO.

2 Creation of Mobile Entities that Embody New Value: Knowing and Learning from Humans

Motivated by the company's founding inspiration of technology for people, Honda is taking on the challenges of creating new products and evolving new technologies. It is also the case in humanoid robot research that Honda is aiming to create a mobile entity that is useful to people and society and that embodies new value. Development is advancing in two directions. One is toward an "assistant robot" that supports people in their living spaces, as the ASIMO does, and the other is toward robots that can substitute humans in areas that are dangerous or inaccessible. Although their purposes and applications may differ, the two robot types have one characteristic in common: that they learn from humans in order to be useful to people.

Humans possess high-level physical capabilities and intelligence that enable them to move freely in locations of any kind, to do a variety of different work, and to provide support to people. If a robot could be developed with humanlike ability to walk on two legs; to run using the whole body, including the arms and torso; to climb; and to travel, then the robot would be able to move freely. It could move not only in homes, offices, and other human living spaces but also in factories, plants, constricted spaces, and places with many ladders. Furthermore, if a robot could be endowed with hands, arms, and intelligence like humans, then it could travel to various places and carry out different kinds of work there.

What Honda is aiming for is a scenario in which robots that possess humanlike physical capabilities and intelligence in that way act as partners with people, coexist and cooperate with people, and contribute to people's life and to society.

3 Mobility Capabilities

3.1 Honda's Unique Bipedal Walking Technology

In 1986, research started on a major characteristic of the Honda humanoid robot: bipedal walking modeled after humans [1, 3, 4]. For Honda, everything about robot development was uncharted territory, including bipedal walking. The work started, therefore, with observation and experiments on every manner of walking, investigating the principle of bipedal walking.

The E0, which was the first robot developed by Honda, achieved static walking in which the legs move forward in alternation. The E2 of 1991 succeeded in dynamic walking, in which the robot walks on a flat surface while changing its center of gravity out of foot soles (Fig. 3). In 1993, Honda finally developed the proprietary walk stabilization control technology. When humans walk, they maintain balance between their own center of gravity and the force they receive from the floor as they walk. Walk stabilization control technology realizes this balance in a robot. This is a core technology in posture control that enables robots to walk on a variety of floor surfaces, including uneven and slanted floors and floors with uneven levels where dynamic walking had not formerly been possible.

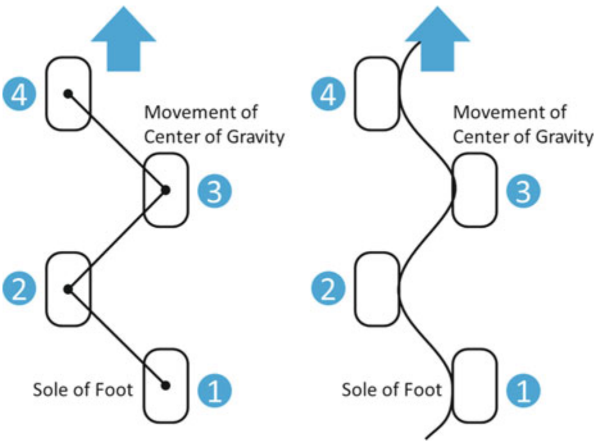


Fig. 3 Static walking (*left*) and dynamic walking (*right*)

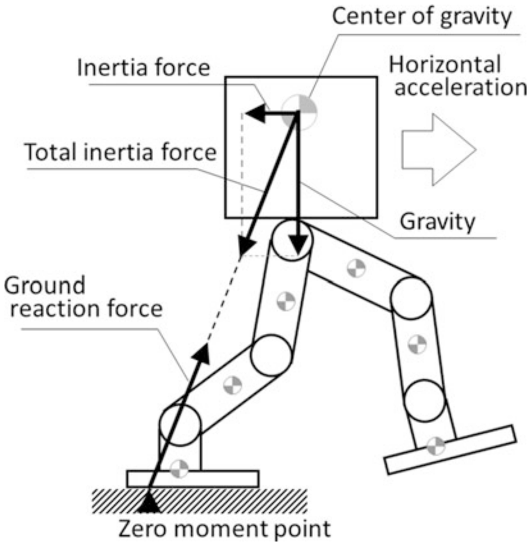


Fig. 4 Zero moment point

This is accomplished by actively changing the robot's stride length or center of gravity.

The walk stabilization control technology generates a target walking pattern based upon a concept in dynamics called the zero moment point (ZMP) [5]. The actual robot is then caused to follow that pattern. The resultant of the gravity and inertia force is the total inertia force, and the point at which the line of action of this force meets the floor surface is the ZMP (Fig. 4). If the ZMP is located within the contact patch of the supporting leg during the single-support phase, and if it is within the support polygon formed by the contact patches of both legs during the double-support phase, then the robot can be considered to be walking stably in terms of dynamics (Fig. 5). The target ZMP is therefore defined so as to satisfy the above conditions. The target walking pattern is then generated to realize that trajectory. The target walking pattern is described by the trajectory of the front of the foot that is necessary to determine the angles of the robot's joints, the position of its upper body, and the trajectory of its posture. The target walking pattern is generated by adjusting the horizontal acceleration of the upper body so that the moment around the target ZMP will be zero.

When the environment and the design coincide perfectly, the robot is assured of being able to move as instructed. Under these ideal conditions, successively outputting the target walking pattern will keep the robot walking. In actual environments, however, floors may be uneven or slanted or have uneven levels as noted above. Simply outputting the target walking pattern will result in the robot falling over. Here the three posture control technologies of ground reaction force control, model ZMP control, and foot landing position control come into play (Fig. 6). These are used to modify the walking pattern with respect to changes in the environment and cause the robot to walk stably.

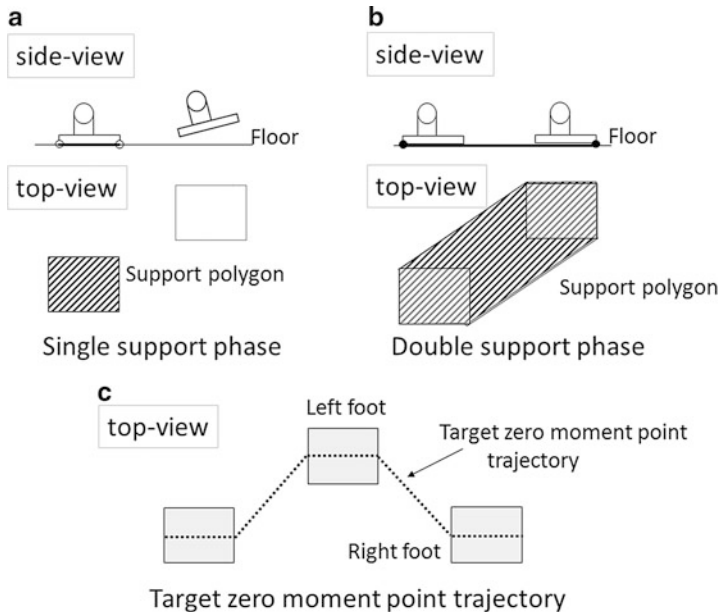


Fig. 5 Target ZMP trajectory

When humans feel themselves about to fall while they are walking or standing straight, they will (a) press down hard on the part of the sole of one foot, and if that is not enough to keep them from falling, they will (b) change the way they are moving their feet, legs, and torso or (c) take a step forward to try and recover a balanced posture.

Ground reaction force control is control of (a) above. It involves absorbing the unevenness of the floor while pressing hard on the sole of the foot when feeling on the verge of falling. When a robot is walking in an ideal fashion, the target total inertia force is on the line of action of the actual ground reaction force. However, if the robot steps on an uneven floor surface, for example, the lines of action diverge, balance is lost, and a tipping moment is exerted. In ground reaction force control, the six-axis force sensors below the ankles detect the central point of the actual ground reaction force. While doing this, the positions of the front of the feet and the posture are changed by rotating the feet so that the target ZMP is at the center of the movement. This causes the heel to step down hard and thus controls the central point of the actual ground reaction force so that it is in an appropriate location (Fig. 6a).

Model ZMP control is control of (b) above. When the robot cannot press hard enough on the sole of the foot, this control maintains a balanced posture by accelerating the upper body in a direction that seems likely to cause the robot to fall over. For example, when the robot appears likely to fall over forward, the control accelerates the upper body trajectory of the target walking pattern somewhat

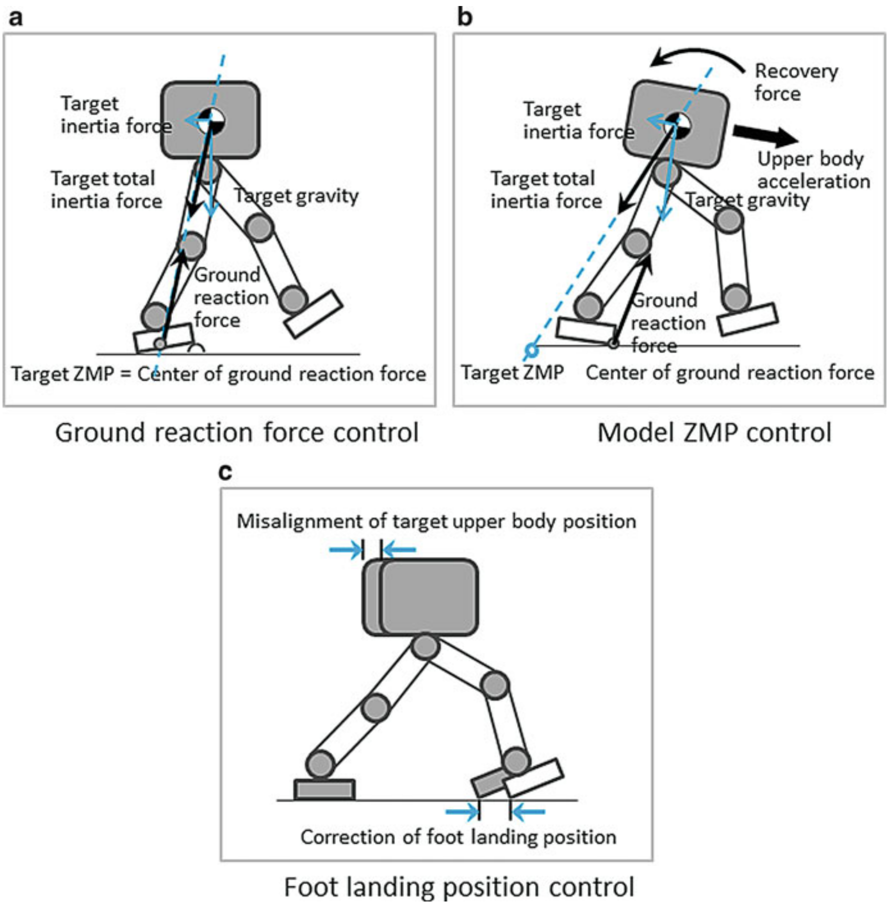


Fig. 6 Three posture controls to achieve stable walking

more forcefully forward than the ideal trajectory. This shifts the target ZMP to the rear of the central point of the actual ground reaction force so that a force tilting the robot backward is applied and the balanced posture is restored. In other words, this form of control deliberately unbalances the target walking pattern so that the robot's balanced posture is recovered (Fig. 6b). In contrast to original target ZMP control, model ZMP control can shift the target ZMP outside the support polygon.

Foot landing position control operates by adjusting the stride length to correct the displacement of the upper body caused by model ZMP control. The operation of model ZMP control causes the position of the upper body to be displaced in the direction of acceleration beyond its position in the target walking pattern. Therefore, if the other foot steps forward in the stride length defined by the original target walking pattern, a displacement between the upper body and the position of the feet occurs. That is when foot landing position control applies an appropriate correction to the stride length so that the upper body and the feet are brought back into an ideal positional relationship (Fig. 6c).

These three posture control technologies were integrated in the walk stabilization control technology developed in 1993. This enabled the E6 to achieve walking that includes turning, climbing steps, and stepping over things (Fig. 7).

Continuing from the development of walk stabilization control technology in 1993, development efforts were directed toward combining the head portion and the upper body in a humanoid robot with two arms and two legs. In 1996, Honda announced the world's first autonomous bipedal walking humanoid robot, the P2 (Fig. 8). With the walk stabilization control technology described above as a base, the P2 achieved the ability to operate without falling over by moving a foot in the direction of the fall even if an external force was applied (Fig. 9).

Subsequently, measures to reduce weight and size of the robot were started, and development of more advanced walking technology was continued. These are indispensable for the robot to be useful in human living spaces. In 2000, these efforts produced the Advanced Step in Innovative Mobility or ASIMO (Fig. 10) [6].

The first-generation ASIMO had a body that was 120 cm tall and weighed 43 kg. This size was decided upon after investigation of ways of enabling the robot to perform work of various kinds in human living spaces and also making the robot easier for people to relate to. Specifically, the minimum size necessary for performing tasks such as reaching door knobs, light switches, and electric sockets in the same way as humans do was verified. This also included the ability to move freely in narrow passageways, on stairs, and in other such places. A variety of different work postures were also taken into consideration, such as work at tables, work with arms extended, work while crouching down, and so on. The location of

Fig. 7 Experimental robot: E6

Fig. 8 Prototype robot: P2

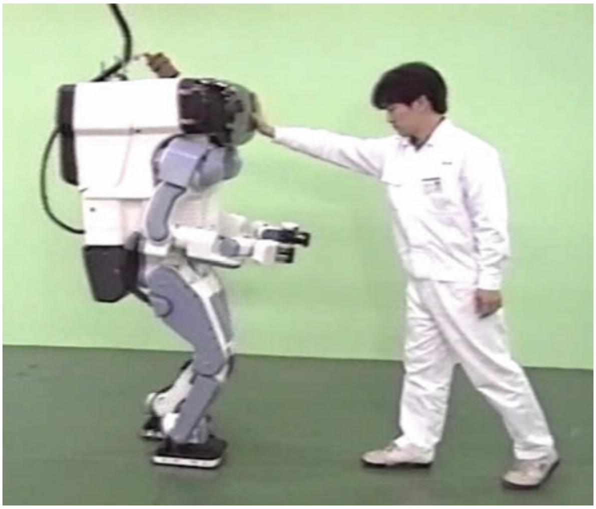


Fig. 9 Control against external disturbance

Fig. 10 First-generation ASIMO

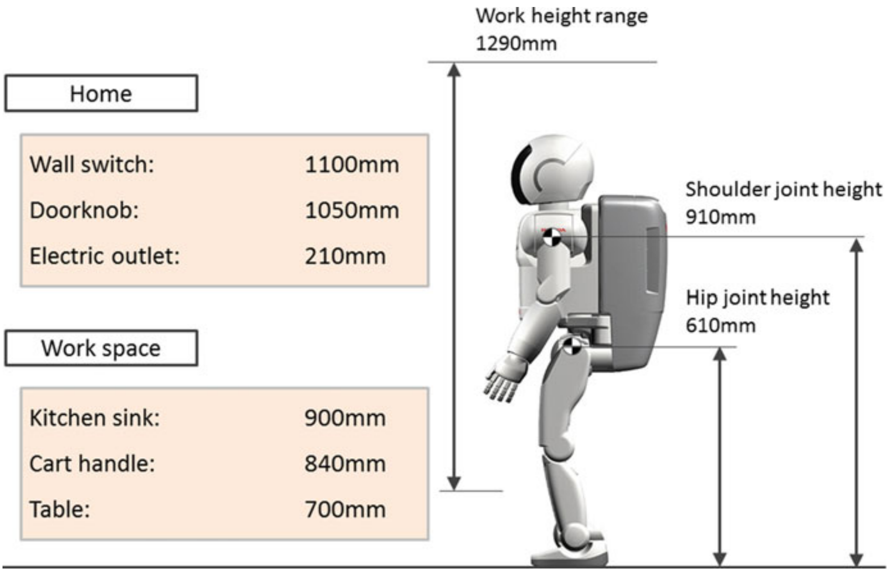


Fig. 11 Investigation of robot dimension

shoulders, length of arms, leg dimensions, hip joint location, body width and depth, and other such dimensions were decided in this way (Fig. 11).

In order to enable the robot to move freely through narrow passageways and on stairs as humans do, a new technology for intelligent, real-time, flexible walking called i-WALK was developed, as well. This builds on previous walking control technologies (Fig. 6) with the addition of predictive movement control. It is the core technology for autonomous bipedal walking.

When humans shift from walking in a straight line to turning a sharp corner, they first shift their center of gravity to the inside of the corner. This allows them to continue walking without stopping. When going from a turn to walking in a straight line, they similarly shift their center of gravity first and continue walking. This kind of real-time flexible walking was achieved with i-WALK.

Up until the P3, the robots had several tens of standard types of walking pattern designed off-line and then loaded into their memory in the form of time-series data. The robot system would then generate a flexible walking pattern as a synthesis of those patterns in combination. This principle, however, meant that when robots shifted from moving straight forward to turning, they would have to stop once. The walking cycle was also limited to several types. These were among the issues with that system.

Figure 12 shows the change in the center of gravity that occurs when shifting from straight walking to turning. When shifting from a straight line to a turn with the i-WALK system installed on the ASIMO, a predictive calculation is performed before making the turn. This seeks an optimal change in the center of gravity while generating a walking pattern in real time. The foot landing position, the turning angle, and even the walk cycle can freely be changed. This approach realizes a more natural, smoother walk that can continuously and flexibly change how the robot is walking at any time without pausing.

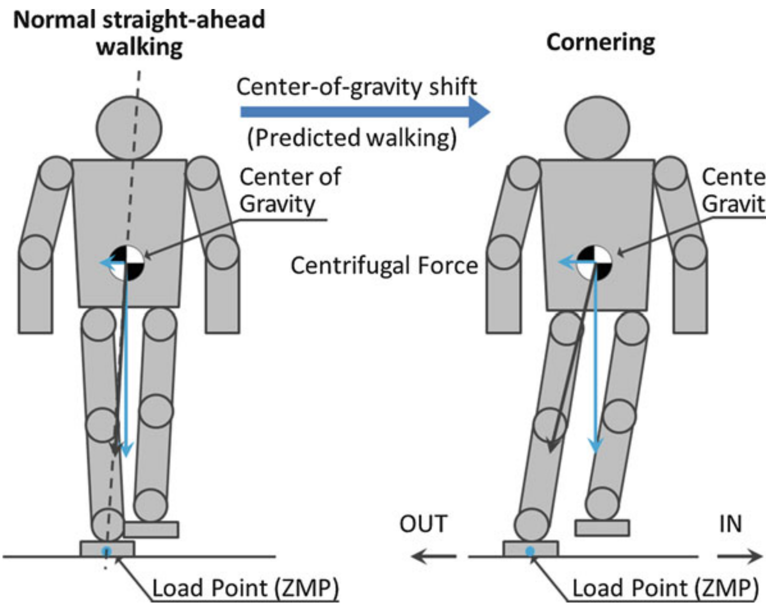


Fig. 12 Center-of-gravity shift

3.2 From Walking to Running

The ASIMO first displayed running in 2004 [7, 8]. Since that time, it has achieved a higher running speed every year, and the 2011 model ASIMO can run at a speed of 9 km/h (Fig. 13). It is a great challenge to make a human-sized robot like the ASIMO run, to go forward by kicking and jumping against the floor like a human being. To make that possible, the robot has to be able to rapidly produce the motions of kicking the feet, swinging the legs forward, and bringing the feet down to the ground and, in addition, absorbing the instantaneous shock caused by landing. It is additionally important to reduce the slipping of the feet and spinning of the upper body that occur in conjunction with rapid movement. And it is essential to kick off firmly from the floor surface and to land on it securely.

A new technology was applied to the hardware and to the gait generation technique [9–12] in order to resolve these issues. The hardware included high-speed processing circuits, highly responsive and high-power motor drive units, and lightweight, highly rigid leg structures. These modifications improve the overall responsiveness of the system to situational changes.

Under rapid movements, because of the increase of the inertia force and moment, keeping the target ZMP inside the support polygon is not enough to prevent instability. Thus, the new gait generation technology employs an allowable range of the horizontal inertia force and moment. This allowable range varies continuously in response to the ground load, which changes while switching to the flight phase. The method used, in order that the horizontal component of the ground reaction force generated by the target walking pattern will not exceed this allowable range, involves adjusting the horizontal acceleration of the upper body as well as the bending and twisting of the upper body.

Specifically, when generating a target walking pattern, the target ZMP trajectory during the landing phase is determined by a target gait generation method based on

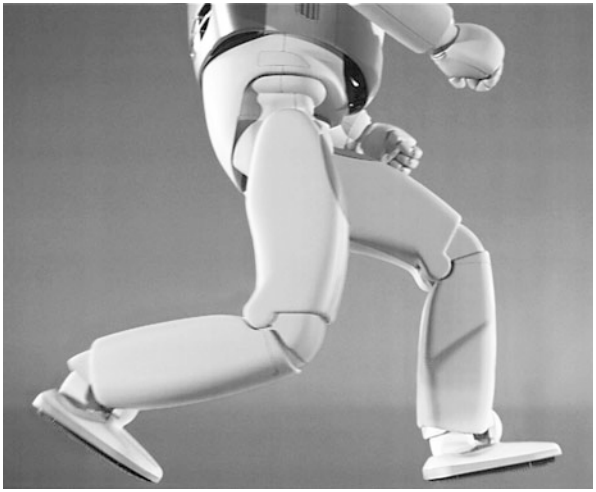


Fig. 13 Leg movement during 9 km/h running

the former ZMP. During the flight phase when the ground reaction force becomes zero, the target ZMP trajectory is set so that movement will proceed virtually continuously to the next foot landing position. After that, the ground load pattern is designed using the height of the target center of gravity during the flight phase, during the landing phase, and during the instant of landing. Next, the allowable range for the horizontal component of the ground reaction force is determined from this ground load.

In an approach similar to that for walking, when the horizontal acceleration of the upper body in the target walking pattern is adjusted so that the moment around the target ZMP will be zero, the horizontal component of the ground reaction force can be dependently obtained. If the horizontal component of the ground reaction force is within the allowable range, both dynamic stability and avoidance of slipping can be achieved. If it exceeds the allowable range, then the horizontal component of the acceleration of the center of gravity for the robot as a whole is kept from changing while causing the upper body to rotate (bend). This keeps the horizontal component of the ground reaction force within the allowable range while maintaining the moment around the target ZMP at zero, thereby avoiding slipping (Fig. 14).

There is also the matter of spin that occurs when the legs are swung forward rapidly when running and when walking at high speed. In an approach similar to that for avoiding slipping, the allowable range for the yaw moment of the ground reaction force with the target gait is determined from the ground load and the coefficient of floor friction. If the yaw moment of the ground reaction force with the target walking pattern is within the allowable range, then both dynamic stability and avoidance of spin can be achieved. If the allowable range is exceeded, then spin can be avoided by causing the upper body to rotate (twist) around the vertical axis (Fig. 15).

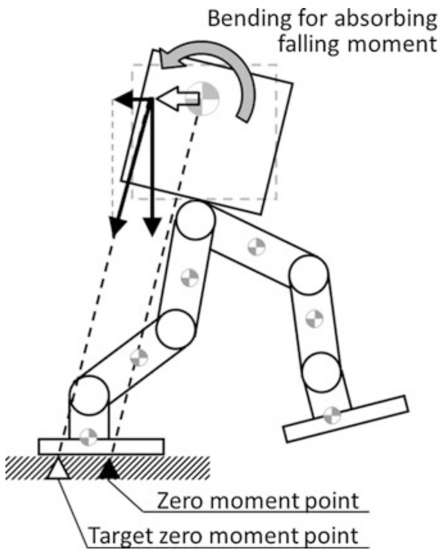


Fig. 14 Balancing motion via upper body bending

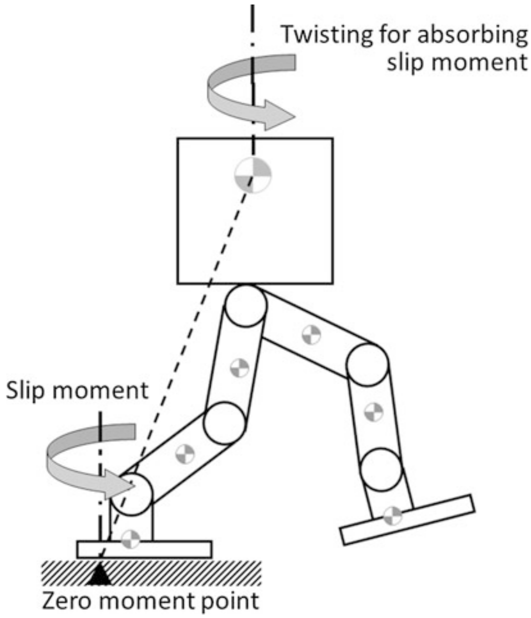


Fig. 15 Vertical moment control

The development of this new gait generation technology made it possible to assure dynamic stability during running while limiting slipping and spinning at the instants just before the foot lifts off the floor and just after it lands on the floor. There are also cases of movement such as slow jogging when the ground load does not drop completely to zero, or walking on a floor with a low coefficient of friction, or other of various such dynamic walking patterns. These can now be handled in a uniform manner just by changing the allowable range of the floor friction force.

The 2011 model ASIMO has the ability to run at a speed of 9 km/h with its increased leg strength, expanded range of leg movement, and development of new control technology that enables it to change its foot landing position while running. Beyond that, the robot can also run backwards, hop on one leg, hop on both legs, and perform other such movements continuously and smoothly one after the other. The ASIMO has also become better able to respond flexibly to changing external circumstances, such as by traversing uneven road surfaces while maintaining a stable posture (Fig. 16).

4 Task-Performing Capabilities

Together with mobility, the robot's task-performing capabilities have also been evolved. Task-performing capabilities are capabilities made possible by fusion of physical capabilities with recognition of the external environment using sensors of various kinds.

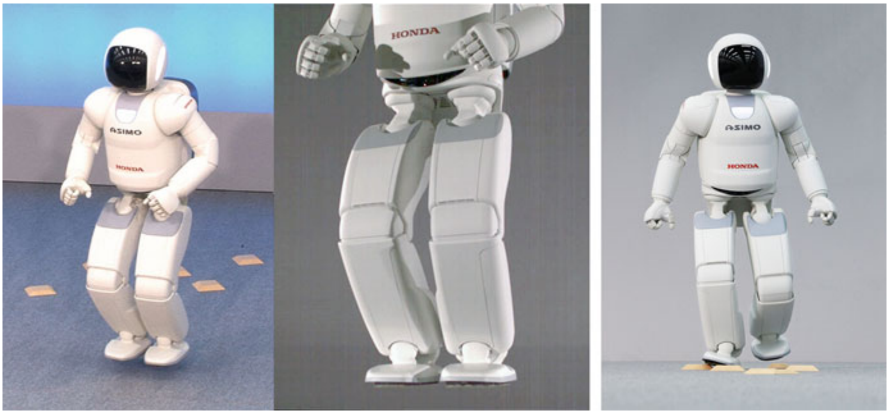


Fig. 16 Advancements of physical capabilities

Fig. 17 Tray handling

For example, in handing over a tray, the eye camera installed in the robot's head and force sensors in the wrists are used. The camera and the sensors detect the movement of the tray so that the robot can match its actions to the other party and hand over the tray without fail (Fig. 17).

Another task envisioned for performance in an office is pushing a cart. This entails doing things that are very advanced for a robot, such as absorbing the upper body rocking that occurs during walking and keeping the hand that is holding onto the cart in a fixed position and steering the cart while pushing and pulling it. This further involves predicting the position of the cart from the force sensors in the wrists and changing the force in both arms as well as the movement of the legs and feet (Fig. 18). Furthermore, if there is interference with the movement of the cart, the robot is able to respond flexibly, by slowing down, changing direction, or some other such action, while continuing to move.

These tasks of handing over a tray while matching actions with the circumstances of the other party, pushing a wagon while responding to changing conditions in

Fig. 18 Cart handling

the actual environment, and other such actions are things that humans do without thinking about them. To propose engineering hypotheses for such tasks and put them into practice using a robot is an activity that vividly reveals how highly efficient and intelligent human beings are.

In connection with task-performing capabilities, a cooperative work function and independent self-charging function were also developed in anticipation of robot operation in real-world environments.

The cooperative work function is a function that enables tasks to be performed cooperatively by multiple ASIMO robots. The robots are linked together by a network with a server computer that tracks the task status of each ASIMO and distributes tasks among the various robots in an optimally efficient manner (Fig. 19).

The independent self-charging function is a system that makes use of a charging station (Fig. 20). When the remaining battery charge falls below a certain value, the ASIMO automatically detects and goes to the nearest unoccupied charging station, where it recharges. This has not only made it unnecessary to change batteries, but it has also made it possible to operate continuously, which is important for performing day-to-day tasks.

The 2011 model ASIMO was also evolved in terms of task-oriented functions that use the hands and fingers. Figure 21 shows the ASIMO performing the task of opening a drinking flask that had been placed on a table and pouring the beverage in the flask into a paper cup. This is another task that is simple for human beings, but that requires hands and fingers that move flexibly in order for the ASIMO to perform it.

Fig. 19 Work in collaboration

A hand that integrates various types of sensors and control systems was developed for that purpose. Tactile sensors for use in grasping objects were placed under the skin on the palms of the ASIMO's hands, and six-axis force sensors used in gripping objects were installed on all five fingertips of each hand. A new master-slave hydraulic system capable of controlling the finger joints independently was also developed, providing 13 degrees of freedom in each hand. These advances achieved a highly functional, compact hand with multiple degrees of freedom.

When performing the task shown in Fig. 21, the ASIMO first uses its head camera to detect the presence of a cylindrically shaped drinking flask on the table. It extends its hand toward that object, and it uses the signals from the tactile sensors in the palm of its hand integrated with the signals from the force sensors at its fingertips to pick up the drinking flask. In order for the robot to pick up the drinking flask in a stable manner, the force of each fingertip needs to be distributed appropriately so that the drinking flask does not slip. The ASIMO grasps the object so that the force in the tips of the thumb and the other four fingers are in equilibrium in the horizontal direction. When it lifts the object, it adjusts the force in each of the fingers so that the force exerted in the vertical direction is matched with gravity. The system distributes the internal forces so that the grasp remains stable even if there is a change in the number of grasping fingers (Fig. 22).

When the robot opens the lid, it turns the lid with equal force in all five fingers. When the robot places the lid on the table, it uses the change in signals from the force sensor in its wrist to judge when the lid has been placed on the table.

Fig. 20 Independent self-charging

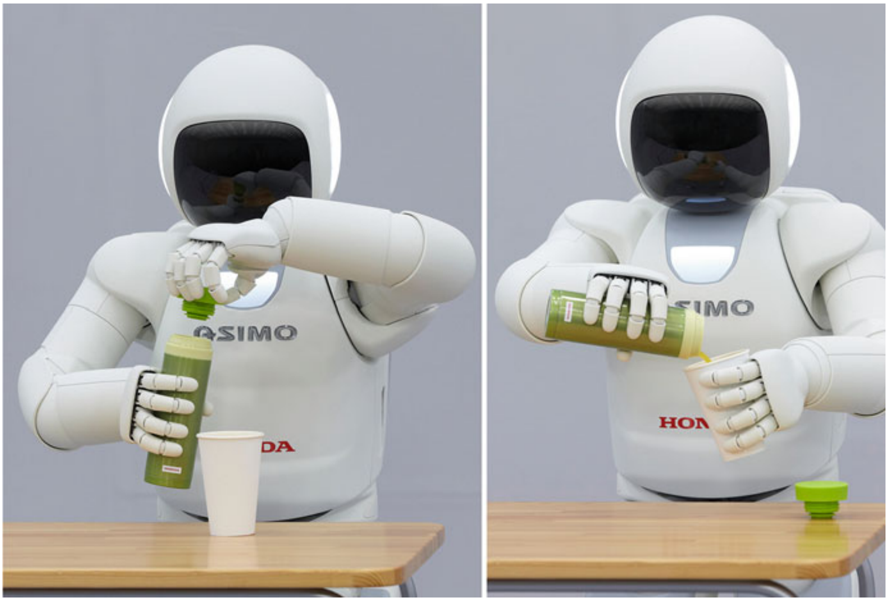


Fig. 21 Drink pouring

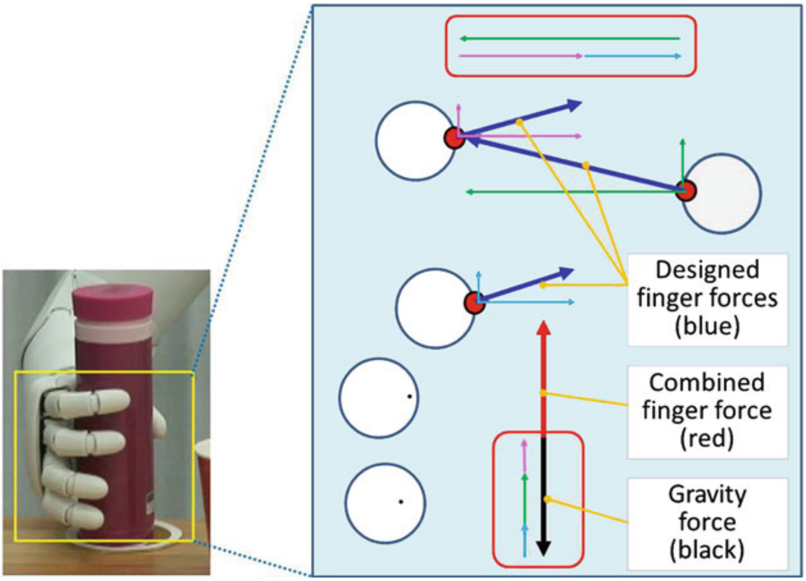


Fig. 22 Multi-finger grasping

In order for the robot to handle a paper cup or other such soft object, it needs to distribute force appropriate to the fingertips so as not to let the object slip or to crush the object. The ASIMO is able to make fine adjustments of the strength in its fingertips not only for hard objects such as a drinking flask but also for soft objects. Another point is that when a beverage is poured into the paper cup, the drinking flask becomes lighter and the paper cup grows heavier. The ASIMO responds to these various changes by adjusting the grip force in the fingertips of its left and right hands in order to keep a stable grip on both the drinking flask and the paper cup.

5 Communication Capabilities

In order for robots to engage in activities in human living spaces while coexisting and cooperating with people, the robots need high-level mobility and task-performing capabilities. They also need advanced communication capabilities to enable them to interact with people and to take into account people's feelings and intentions when they act.

When humans exchange messages, not only do they use language, but they also use facial expressions, eye contact, body and hand gestures, physical postures, and so on. Humans use these kinds of nonverbal communication regardless of whether they are acting consciously or unconsciously.

Cooperation in language and physical expression has been engaged as a theme for the ASIMO since the beginning of development.

The greatest advantage for a humanoid robot in terms of communication is the ability to use this nonverbal communication effectively. As a humanoid robot has an appearance similar to humans, it can nod in agreement with what the other party says, tilt its head if it didn't hear what was said, use its finger to point toward a direction, make eye contact, turn its gaze away, shake hands, hold hands, and so on. The robot can communicate through various actions like these and by means of the expressions of feelings that humans employ unconsciously. Communication of true variety and abundance becomes possible when physical expression and verbal expression are combined or deliberately differentiated in use. For example, expressions of thanks can be voiced while looking the other person in the eye, and appreciation can be expressed while dancing for joy. When acting in light of a person's feelings or intentions, the robot can apprehend not just the voiced information uttered by the person but also the person's nonverbal expression through posture, gestures, and so on, as image information. In this way, the robot can act in a more appropriate manner that is harmonious to the gestures and activities of the other party.

Recognition technologies and body functions related to the ASIMO's communication capability are introduced in the following.

5.1 Voice Recognition

Voice recognition is configured, with proxemics as a reference, so that the distance between the robot and people it converses with is 1–2 m (Fig. 23). The objective is to recognize only words spoken by people who are at this distance from the robot.

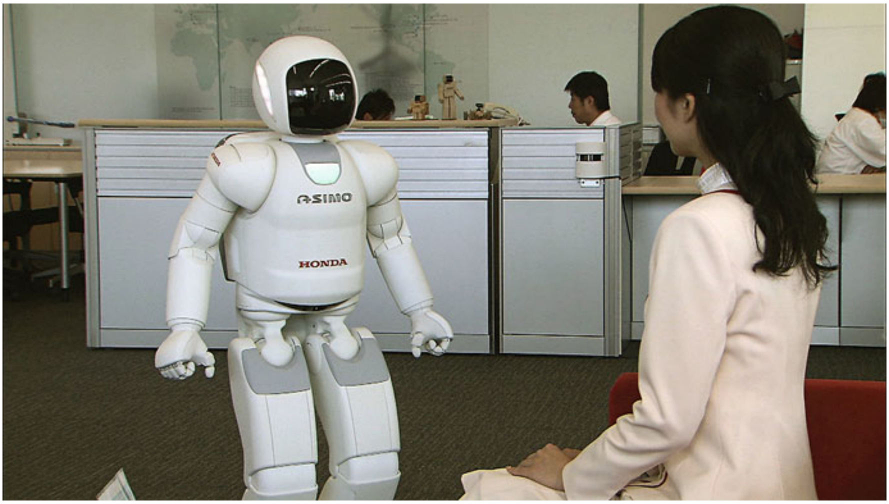


Fig. 23 Voice Recognition

Since distances may differ, and people interacting with the robot will not all speak at the same volume level, the audio intensity of input to the robot will vary considerably. In order to deal with these variations in circumstances, a technology was developed that uses a single model to learn sounds over a wide range of signal-to-noise ratios. This approach has achieved a favorable recognition rate.

Speech features, sound volume, and fundamental frequency are also used to detect “umm” sounds and other such sections of word lengthening. The purpose is to eliminate misrecognition caused by the filled pauses that occur frequently in speech.

Open-source robot audition software [13] developed by Honda Research Institute Japan Co., Ltd. and Kyoto University was applied for the voice recognition system. The robot audition program takes input from multiple microphones (a microphone array) to perform sound source localization, sound source separation, and even recognition of separated voices. It employs middleware that configures and connects a variety of different functional modules in a GUI programming environment. By changing the functional modules, the system can support robots with different shapes and microphone layouts, and it can build robot auditory systems that are matched to applications.

Robots that are situated in human living spaces are required to hear and differentiate the voices of different people even in real-world environments where there are noises and sounds of people talking. The 2011 model ASIMO has an eight-channel microphone array installed around the periphery of the head. It can pick up mixed voice signals even from multiple people speaking simultaneously and estimate the number of speakers and the direction of each speaker based on the transfer function of each sound source direction. It then separates and extracts the voice of each speaker, performs independent component analysis of the extracted voices, and deletes common components found among the separated voices in order to achieve high-precision separation.

For voice recognition of words spoken by people, the system selects words that are closest to the voice data in the system, which is entered in order of vocabulary that is most likely to be spoken. Thus the system has vocabulary set to words with a higher frequency of use in the location where the ASIMO is to be used and with its purpose there. This gives the system a higher recognition rate.

5.2 Image Recognition

The ASIMO has a stereo pair of cameras mounted in the head. One of them is made into a multiple resolution camera using a prism, therefore achieving a balance between recognition of nearby objects across a wide field of view and recognition of people at a distance. Image recognition with greater accuracy is also realized by automatically adjusting shutter speeds in response to lighting environments that change with weather, time, and other such factors.

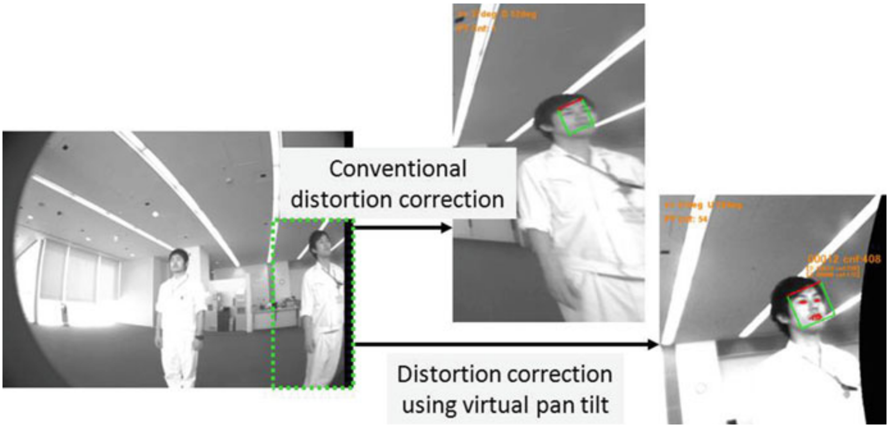


Fig. 24 Correction of image distortion

When a robot interacts with multiple people, it will not necessarily be the case that the face of the person it needs to respond to is always directly in front of the robot. Therefore a facial image recognition technology was developed that would enable face detection, face direction estimation, and face identification of a person regardless of that person's location in the camera's field of view. Rather than the previous method, which took the center of the image as the focal point for application of distortion correction to the entire image field, this system uses a newly developed method of image distortion correction known as virtual pan and tilt, which virtually directs the camera to the face location and uses the image that has that face at its center. This has realized facial recognition across wide field of view (Fig. 24).

5.3 Physical Expression

Figure 25 shows the configuration of joints in the 2011 model ASIMO. It has a total of 57 degrees of freedom. The joint configuration is a large factor not only in physical capabilities but also in the physical expression that affects nonverbal communication. In the ASIMO, whole-body cooperative control is able to flexibly coordinate the movement of every joint, giving it the ability to communicate with abundant expressiveness.

The highly functional, compact hand with multiple fingers has 13 degrees of freedom, providing high-level working functionality as described above. This has not only enabled a variety of different expressions using the fingers in the area of physical expression, but it has also enabled the robot to express itself in sign language, which requires complex movements of the fingers (Fig. 26).

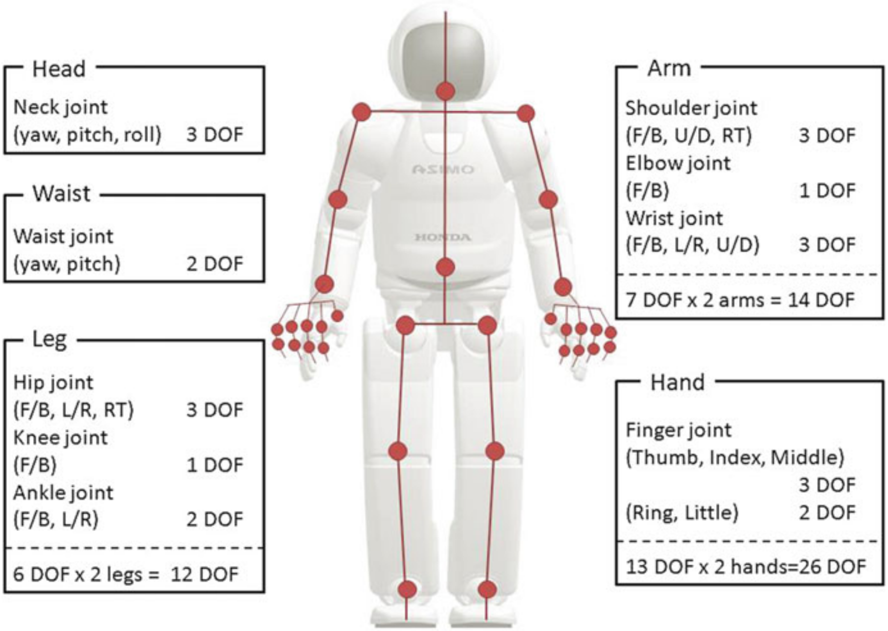


Fig. 25 Joint configuration and DOF (degree of freedom)

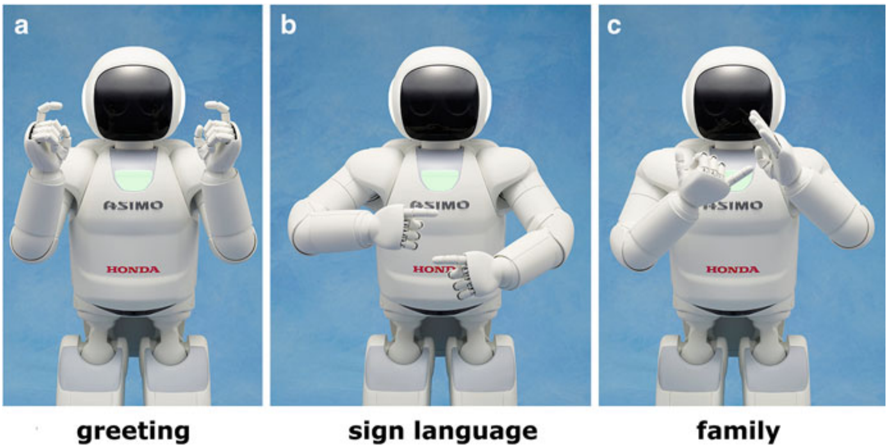


Fig. 26 Japanese sign languages

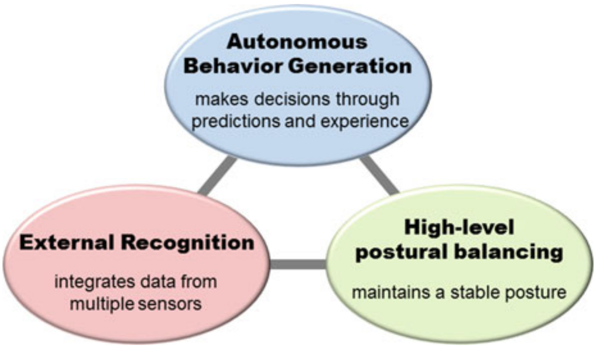
The robot was also given a neck structure with the three degrees of freedom needed for nodding and tilting the head. For head movement in the up and down directions, the use of a six-link construction has expanded the angle of movement to 26 degrees up and 32 degrees down, enabling more variously nuanced expression.

6 From Automatic to Autonomous

In a real-world environment that is constantly changing, there are many cases when people do not take the action they were envisioned in advance. The previous ASIMO acted automatically according to scenarios that were fixed in advance, but such an approach does not enable appropriate behavior. In addition to high-level postural balancing and external recognition, development of the 2011 model ASIMO further pursued the third element of autonomous behavior generation in order to overcome such issues [14]. This element is needed by a robot for it to be an autonomous machine (Fig. 27).

A system architecture called the intelligence loop was devised for such an autonomous behavior generation technology. The intelligence loop is made up of the four functions of sensing, situation estimation, behavior generation (planning), and movement expression, which are implemented on rapid cycles in order to act in response to changing situations. It also includes the function of learning, which provides for better outcomes through repeated performance of actions (Fig. 28).

Fig. 27 Three elements for autonomous machine



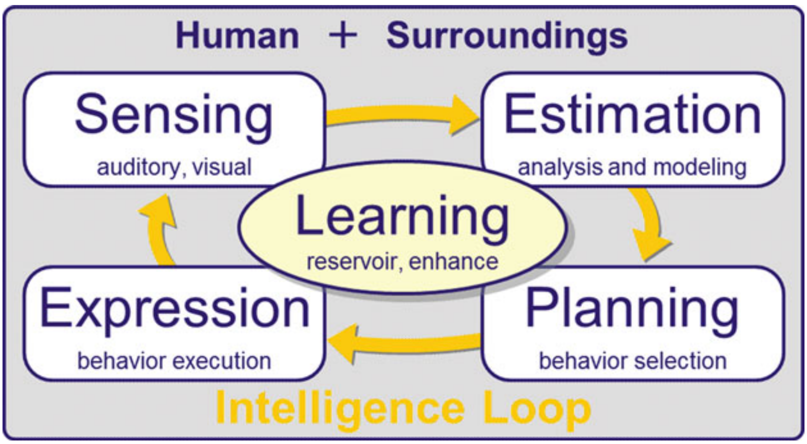


Fig. 28 Intelligence loop

The 2011 model ASIMO estimates its situation based on input from multiple sensors. It selects the optimal combination of actions from among many alternatives and executes those actions.

Figure 29 is a diagram of the system architecture as adapted to office reception guide operations. The reception guide operations here include meeting and greeting visitors, guiding them to a sofa or conference room, offering them something to drink, and providing information by means of a display unit (Fig. 30).

6.1 Sensing

In order for the ASIMO to behave appropriately in a constantly changing real-world environment, external recognition capability that accurately grasps surrounding circumstances is important.

The system shown in Fig. 29 employs laser range finders (LRFs) installed in the reception area for spatial sensing, the ASIMO eye camera for image recognition, and the eight-channel microphone array also installed on the ASIMO for voice recognition. From these three, the system perceives circumstances in the area of ASIMO activity.

For spatial sensing, the LRFs with one-dimensional scanning capability are directed at the wall surfaces in the area where the ASIMO will be active. The LRFs with two-dimensional scanning capability are located on the ceiling. The two types of LRFs can detect visitors and their movements around corners and at distances not visible to the eye camera, providing a grasp of circumstances in the space as a whole.

For image recognition, nearby visitors are detected and related circumstances are observed using the eye camera.

This application for reception guide operations also includes voice recognition, and it simultaneously recognizes multiple people's faces and voices. A demonstration is conducted in which three visitors speak simultaneously and ask for something to drink. The ASIMO can differentiate the respective orders and serve the beverages accordingly (Fig. 31).

6.2 Situation Estimation

In situation estimation, the ASIMO uses voice, image, and spatial sensing information to estimate the attributes of people that are needed to deal with those people. A person's attributes refers to the person's interest in the ASIMO, the object of that person's attention, and the extent of that person's perplexity, affirmation, or negation. These are estimated from changes in where and for how long that person stands still, the person's distance from the ASIMO, the orientation of the person's face, and so on. These attributes of people are estimated from voice, image, and spatial sensing information using a Bayesian network to rank their likelihood.

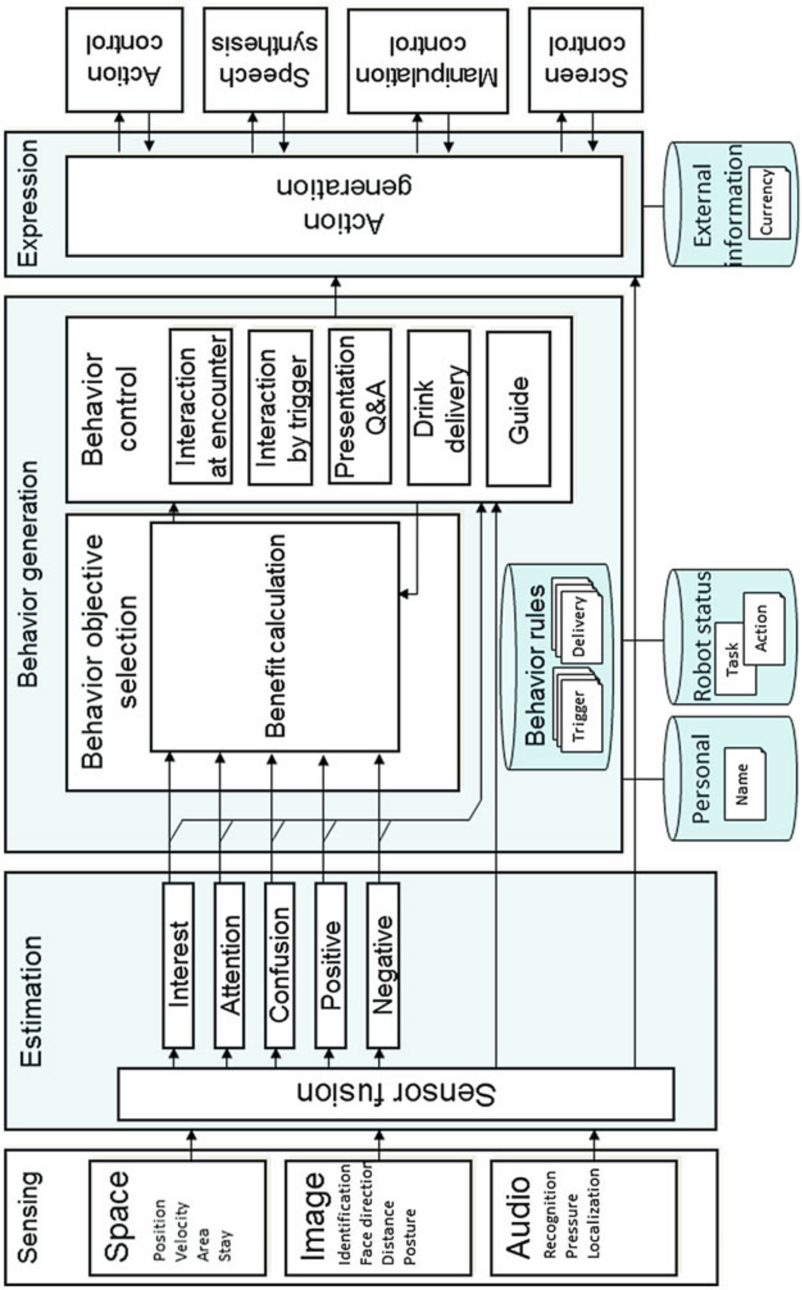


Fig. 29 Interaction system architecture

Fig. 30 Reception guide operations

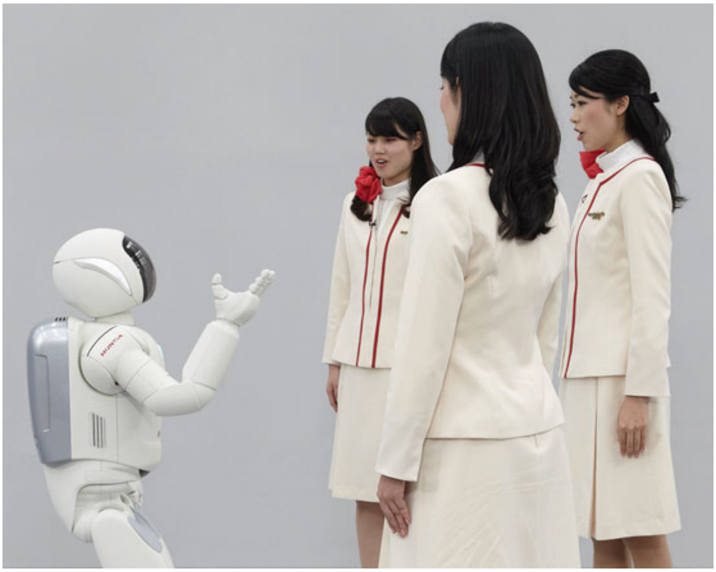


Fig. 31 Simultaneous speech recognition

6.3 Behavior Generation

Figure 32 shows the portion of Fig. 29 that relates to behavior generation in the application for reception guide operations in an office. Behavior generation is made up of behavior objective selection and behavior control.

With previous robots, the behavior objective was given in advance, but an autonomous robot needs to be able to decide the behavior objective on its own. The behavior objective selection unit encompasses multiple people and multiple behavior objectives in combination. From among these, the unit selects the optimal behavior objectives with respect to the various people. In this example, the ASIMO is dealing with five candidates designated A to E. The behavior objectives for the ASIMO are of five types, namely, interaction at encounter, interaction by trigger, acting as a guide, giving a presentation, and delivering a drink. Here the choice is interaction at encounter with person C.

The question of which person to be associated with which behavior objective is decided by the behavior value of the target person and the behavior objective with respect to that person (the person-behavior pair). The behavior value is calculated by the person attributes that are output from situation estimation, by the appropriateness calculated from behavior objectives, and by the behavior effectiveness calculated on the basis of the robot's behavior history. Out of the person-behavior pairs, the pair with the highest behavior value is decided upon first, and if there are more that can be carried out, then the person-behavior pairs are decided upon in order of their behavior value.

If, for example, the target person that the ASIMO is addressing shows little response, and another person appears on the scene, then the behavior value of the other person who appeared will be higher than that of the target person who is presently showing little response. As a result, the robot will discontinue its present behavior and address the newly appeared person. This design enables the robot to change to a different behavior according to the response of the person it is interacting with, change the target person for behavior that was adapted to the person's movement or circumstances, or change the nature of the interaction.

Once the behavior objective has been decided, then the specific behavior to be carried out is decided next in the behavior control unit. Behavior will be carried out according to behavior rules that are formulated in advance for each behavior objective. Since the system allows for parallel processing and interrupt processing of multiple behaviors, the robot is also able to continue a behavior it is currently carrying out or to suspend it temporarily and carry out a separate behavior.

For example, if the visitor says "thank you" while the robot is placing a tray on a table during drink delivery, the resulting person attribute estimated from that voice information will be a trigger for suspending the present behavior objective. The robot will therefore be switched over to a new behavior objective with regard to the guest who said "thank you," and it will respond "you're welcome."

In this way, too, if a guest asks a question during a presentation, the robot can suspend the presentation and respond to the question. The system also possesses

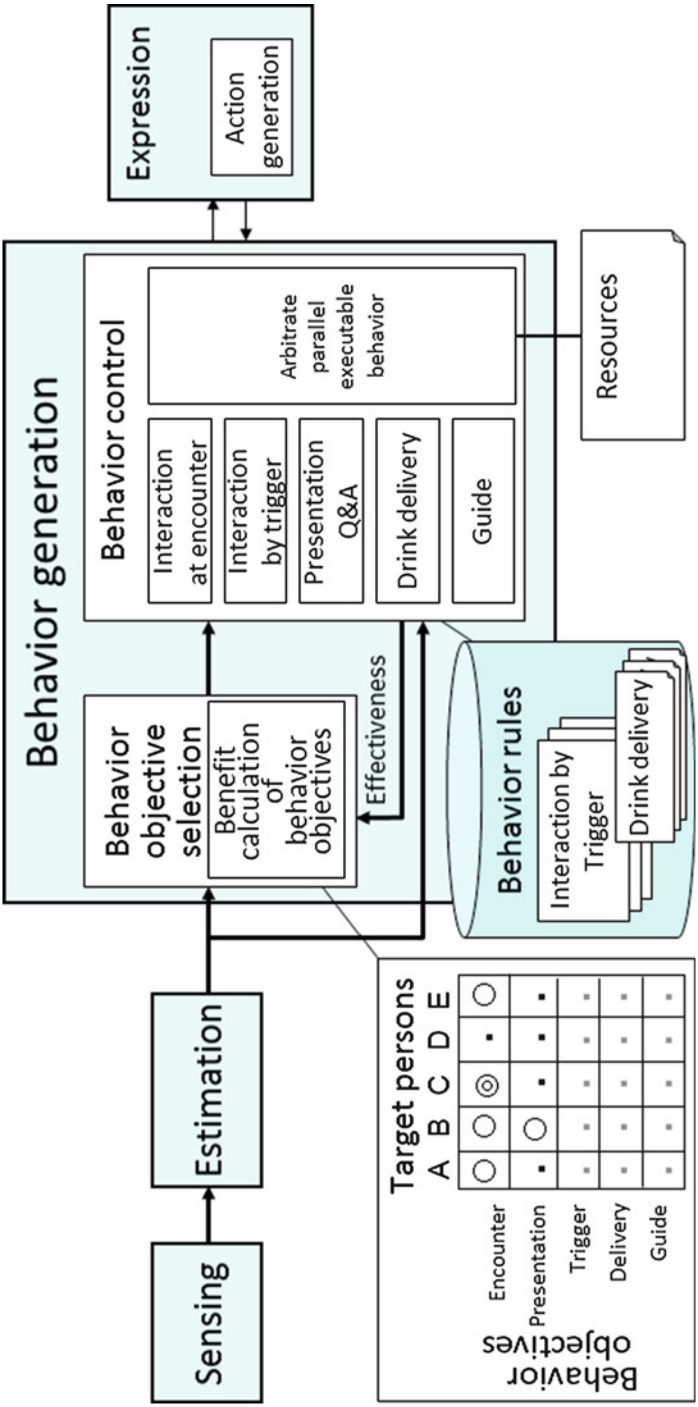


Fig. 32 Behavior generation

the functionality to manage how the robot will return to the presentation according to the circumstances, such as by continuing from the point where it suspended the presentation, returning to a slightly earlier point and continuing, or starting over from the beginning.

The behavior history can also be used to set a constraint condition so that the robot will not repeat the same phrase. When dealing with guests who have come back for another visit, the robot can autonomously change how it addresses them or proceed to give a new presentation explaining a different matter.

6.4 Field Experiments

During 2013, two field experiments were carried out at Miraikan - the National Museum of Emerging Science and Innovation, which has one million or more visitors per year.

In the experiment conducted for a 1-month period from July, the ASIMO engaged interactively with large numbers of visitors gathered in an exposition area while explaining its own functions to them.

The experiment involved the ASIMO posing questions to the visitors. From their response by a show of hands, the system would then estimate the visitors' inclination and provide an explanation accordingly. In doing so, the ASIMO was sensing the behavior of the visitors through the network, so that it instantly recognized the reaction from several tens of visitors. It then proceeded to carry out the explanation autonomously (Fig. 33).

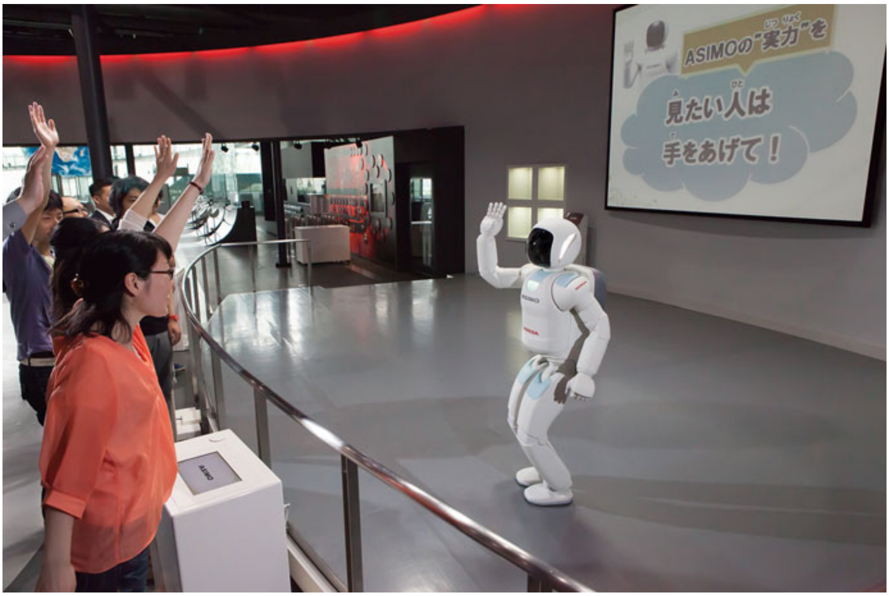


Fig. 33 Explanation for large number of visitors

Fig. 34 Explanation for single visitor

The experiment made effective use of the robot's memory, which is a characteristic strength, to store explanations with a variety of different content. The system then estimated the visitors' needs from a number of sensors, and the value of an "autonomous explaining robot" that can communicate content in an easily understandable way was demonstrated in the real-world context of people and the environment undergoing constant changes.

The second field experiment was held over a 20-day period in October as part of a collaborative research with Advanced Telecommunications Research Institute International (ATR). The experiment involved having the ASIMO match its movements to those of alone visitor who was looking at exhibits arranged in the exhibition hall. The ASIMO would approach that visitor and autonomously provide explanations of the exhibits (Fig. 34).

The ASIMO would detect the direction in which visitors were walking or the direction in which the bodies or faces of visitors viewing exhibits were facing, by means of spatial sensors located in the exhibit hall. From that information, it would ascertain what exhibits the visitors were directing their interest toward, and the ASIMO would approach visitors by its own judgment to provide them autonomously with explanations of the exhibits that were the objects of the visitors' interest. When doing this, the ASIMO would direct its gaze toward the exhibit and point with its finger to give its explanation. It would also choose an optimal position in which to stand that would make the exhibit easier for the visitors to see and its own explanation easier to hear. The ASIMO would also remember the content of the explanations it gave to individual visitors. When those people came again, it would not only greet them differently but also give them explanations with content different from before.

This experiment demonstrated the advanced autonomous explanatory function whereby the intentions of visitors were estimated from their behavior, the robot would walk up to visitors and give them explanations suited to what they were interested in, in a manner that made it easier for the visitors to see and hear, and for repeat visitors, it would change its explanation on each occasion.

7 Technology for People

Field experiments of the ASIMO to date have been held at public places such as the National Museum of Emerging Science and Innovation and the reception lobby of the Honda head office and other such locations. From the start of development, the consistent objective throughout has been a robot that would be able to function in ordinary households, helping the family in various ways as a member of the family. The robot would be that family's robot.

The 2011 model ASIMO is presently engaged in permanent ongoing demonstrations at three locations in Japan. This activity includes field experiments, and the various kinds of data obtained through interactive exchange with visitors are provided as feedback for development of the next phase ASIMO. It is being channeled toward realization of the dream of one robot per family.

In the spring of 2011, a new objective which is to create a humanoid robot for the purpose of disaster prevention, disaster reduction, and other such response to disaster was added. What occasioned this addition was the accident at the Fukushima Daiichi Nuclear Power Station that occurred with the Great East Japan Earthquake in March 2011. As news reports on disaster circumstances came in day after day, Honda was inundated with comments from customers. What they were saying was, "Can't the ASIMO be sent into the nuclear power station?"

The ASIMO had a different purpose and design requirements from the start, so it could not be dispatched to the nuclear power station. The notion was already present, however, that the technology cultivated in the course of ASIMO development could provide a basis for creating a robot that would be able to perform tasks inside a nuclear power station. There was a need not only for a robot like ASIMO, that would be considered invaluable in the future, but for a robot that would be considered invaluable right now in the present. Honda felt this need very keenly.

In April 2011, 1 month after the earthquake, Honda called together members of the ASIMO development team to develop a robot intended for the Fukushima Daiichi Nuclear Power Station. Initially, the primary function of the robot would be opening and closing valves on pipes within the station, and the prototype was completed in 3 months. However, discussions with TEPCO's Fukushima Daiichi Stabilization Center revealed a much more important objective: inspection and

Fig. 35 High-Access Survey Robot

monitoring of ceilings and other inaccessible locations. Although the purpose of the robot was altered and many specification changes were made along the way, High-Access Survey Robot was completed in June 2013 (Fig. 35). This robot had the capability to perform detailed survey tasks under remote control even in the complexity of structure inside a nuclear reactor building. The robot was twice sent into the Fukushima Daiichi Nuclear Power Station, and it performed survey tasks successfully.

The High-Access Survey Robot intended for use in nuclear power stations was not the only aim. Work has also been accelerated on the development of a humanoid robot to act in place of a human in providing an initial response when disasters occur in thermal power plants, factories, and other such dangerous places, as well as to make inspection rounds and perform other such functions under ordinary conditions.

Honda has also applied the robotics technology created in the course of humanoid robot research to develop new mobility devices. These include Walking Assist Device (Fig. 36), Bodyweight Support Assist (Fig. 37), the UNI-CUB personal mobility device, and the U3-X (Fig. 38).

Motivated by the company's founding inspiration of technology for people, Honda is committed to continue meeting people's demands and expectations with a focus both on the present and on the future.

Fig. 36 Walking Assist Device

Fig. 37 Bodyweight Support Assist

Fig. 38 UNI-CUB (*left*), U3-X (*right*)

References

1. M. Hirose, K. Ogawa, Honda humanoid robots development. Phil. Trans. R. Soc. A **365**, 11–19 (2007). https://doi.org/10.1098/rsta.2006.1917
2. Honda Worldwide | Honda Robotics, http://world.honda.com/HondaRobotics/. Accessed 13 Mar 2015
3. K. Hirai, Current and future perspective of Honda humanoid robot, in *Proceedings of IROS*, 1997, Grenoble, France pp. 500–508. https://doi.org/10.1109/IROS.1997.655059
4. K. Hirai, M. Hirose, Y. Haikawa, T. Takenaka, The development of Honda humanoid robot, in *Proceedings of ICRA*, 1998, Leuven, Belgium pp. 1321–1326. https://doi.org/10.1109/ROBOT.1998.677288
5. M. Vukobratovic, J. Stepanenko, On the stability of anthropomorphic systems. Math. Biosci. **15**, 1–37 (1972). https://doi.org/10.1016/0025-5564(72)90061-2
6. M. Hirose, T. Takenaka, Development of the humanoid robot ASIMO. Honda R&D Tech. Rev. **13**(1), 1–6 (2001)
7. S. Shigemi, Y. Kawaguchi, T. Yoshiike, K. Kawabe, N. Ogawa, Development of new ASIMO. Honda R&D Tech. Rev. **18**(1), 38–44 (2006)

8. T. Takenaka, T. Matsumoto, T. Yoshiike, S. Shirokura, Running gait generation for biped robot. Honda R&D Tech. Rev. **20**(2), 101–107 (2008)
9. T. Takenaka, T. Matsumoto, T. Yoshiike, Real time motion generation and control for biped robot -1st report: walking gait pattern generation, in *Proceedings of IROS*, 2009, St. Louis, USA pp. 1084–1091. https://doi.org/10.1109/IROS.2009.5354662
10. T. Takenaka, T. Matsumoto, T. Yoshiike, S. Shirokura, Real time motion generation and control for biped robot -2nd report: running gait pattern generation, in *Proceedings of IROS*, 2009, St. Louis, USA pp. 1092–1099. https://doi.org/10.1109/IROS.2009.5354654
11. T. Takenaka, T. Matsumoto, T. Yoshiike, Real time motion generation and control for biped robot -3rd report: dynamics error compensation, in *Proceedings of IROS*, 2009, St. Louis, USA pp. 1594–1600. https://doi.org/10.1109/IROS.2009.5354542
12. T. Takenaka, T. Matsumoto, T. Yoshiike, T. Hasegawa, S. Shirokura, H. Kaneko, A. Orita, Real time motion generation and control for biped robot -4th report: integrated balance control, in *Proceedings of IROS*, 2009, St. Louis, USA pp. 1601–1608. https://doi.org/10.1109/IROS.2009.5354522
13. K. Nakadai, H.G. Okuno, H. Nakajima, Y. Hasegawa, H. Tsujino, Design and implementation of robot audition system "HARK" – Open source software for listening to three simultaneous speakers. Adv. Robot. **24**, 739–761 (2010). https://doi.org/10.1163/016918610X493561
14. S. Shigemi, K. Kawabe, T. Nakamura, Development of new ASIMO-realization of autonomous machine. Honda R&D Techn. Rev. **24**(1), 37–45 (2012)