Toward the use of humanoid robots as assemblies of content technologies

— Realization of a biped humanoid robot allowing content creators to produce various expressions —

Shin'ichiro Nakaoka *, **Kanako Miura, Mitsuharu Morisawa, Fumio Kanehiro, Kenji Kaneko, Shuuji Kajita and Kazuhito Yokoi**

[Translation from *Synthesiology*, Vol.4, No.2, p.80-91 (2011)]

A significant feature of humanoid robots is their potential to make various expressions as humans do, and this feature will allow the use of humanoid robots as assemblies of content technologies. Technical issues required for the practical use of humanoid robots are discussed in terms of robot hardware, motion expression generation, vocal expression generation and integrated GUI (Graphical User Interface), and the development of technologies to solve the issues and their integration have been carried out. As a result, we have produced HRP-4C, a life-size biped humanoid robot with realistic human-like appearance, and Choreonoid, an integrated software interface that allows us to choreograph motions with robots as done with CG characters. Experiments on creating contents with these technologies verified the potential of humanoid robots as assemblies of content technologies.

Keywords : Biped humanoid robots, content technology, entertainment, cybernetic human HRP-4C, motion creation, key pose, Choreonoid, VOCALOID

1 Humanoid robot as content technology

Among several types of robots, the humanoid robot enchants people, because of the sense of wonder created by the fact that an artifact made in the image of humans can actually move like a human, the expectation that it may take over the various tasks like a human, and because it has been referenced in various works of fiction. Driven by this fascination, various humanoids have been developed. In 1996, Honda developed the P2, a life-size humanoid capable of bipedal walking[1], and the development of humanoids have become active ever since. These robots made frequent appearances in public events and the media, and as a result, the expectation for the practical utilization of humanoids has increased.

One of the applications of robots for which people have high expectations is to have the robot assist various tasks in everyday activities, and there have been many researches on humanoids to fulfill this demand[2][3]. If the objective is to freely move around within the human environment and to use the tools and devices that humans use, it rationally follows that the robot should have a similar form as humans. However, there is a large gap between the ability of the tasks that people expect from robots and the current technological level of robots, and at this moment the practical utilization in this direction is no where in sight.

However, there is a great potential for the practical utilization of humanoids, where the robots are made to perform certain acts to be viewed and heard by an audience. There are many thoughts on how such performances could be used, but seen from the technological perspective, many performances can be captured within the framework of "content technology". Here, "content" means a set of information and experiences that may be valuable to the audience, viewer, or consumer. If the humanoid can be used as an expression of contents, this can be called the "content technology" that supports the expression and execution of the contents.

Humanoids possess potential attractiveness as an assembly of the content technologies. In fact, "the ability to enchant people", as mentioned above, is directly linked to the value of the contents. The robot is a machine that is controlled by a computer, and it is possible to do things that flesh-and-blood humans cannot do in the production, expression, and execution of the contents, such as combining with various information technologies and implementing special physical functions. Since most of the contents that people demand are geared to humans, the humanoids are more befitting in terms of general interest contents compared to the robots with non-human forms.

The above characteristic overlaps with the character animation in computer graphics (CG). In fact, the CG characters are utilized in various forms, and are becoming essential parts of the content technology. The robots are distinct from CG characters in that they are entities in the

Intelligent Systems Research Institute, AIST Tsukuba Central 2, 1-1-1 Umezono, Tsukuba 305-8563, Japan * E-mail: s.nakaoka@aist.go.jp

Original manuscript received November 8, 2010, Revisions received February 21, 2011, Accepted March 7, 2011

real world. This allows a sense of reality and involvement as well as physical interaction that is not possible with CG characters.

We believe that the use of humanoids in the content technology is a valuable application that allows optimization of its characteristics, and therefore we must engage in such a realization. As accomplished by the digital content technologies such as CG, computer music, and game devices, contents with new values can be created, and the industries of related technologies can be vitalized. If the humanoids can be widely utilized in such applications, continued investments can be expected, and that may lead to the development of other applications including the aforementioned support of daily human activities.

2 Objectives and issues of the research

The contents featuring humanoids did exist before. The technical demonstrations of the state-of-art humanoid were one such example, as the audiences were surprised and entertained by some specific actions of the humanoid. However, these were basically created by the robot developers and focused on the technological prowess of the robot. Also, it did not have the value or the breadth that would create sufficient income.

In practice, the producer of the contents should be the "creators" who are the specialists of content production. The "establishment of the content technology of the humanoids" is to be able to incorporate the humanoids into the contents that are created by the creators of various fields, to express things that do not merely fall in the realm of robot technology. Unless this is achieved, the humanoids are unlikely to be used widely, as seen from the situation of the current content technology. However, the conventional humanoids and the peripheral technologies were not realistically usable for the creators due to the difficulty of use and the limitations of expressions.

The objectives of this research are to improve the situation by developing, integrating and verifying the elemental technologies that would be the foundation and to set a path for the commercialization of the humanoids as a content technology.

We set the following technological issues as those we must solve.

Robot hardware
Develop a life-size, bipedal humanoid with human-like appearance throughout the body, as hardware with expressive ability unseen before.

Motion expression support technology
Create a technology as a method for generating diverse motion expressions, to enable choreography of the whole body movement of the bipedal humanoid through operation similar to the key frame animation for the CG characters.

Voice expression support technology
Incorporate the voice synthesis technology to enable the easy creation of the diverse expressions of speaking and singing, along with the movement of the mouth.

Integrated interface
Develop a software interface that enables integration of the above technologies and the current information and content technologies.

To solve these issues, it is necessary to engage comprehensively in the development, selection and integration of the various elemental technologies as shown in Fig. 1.

After overcoming these technological issues, another important issue of this research is to verify whether the expressions produced by a creator who is not a robot researcher are acknowledged as new contents unseen before. By combining the technological foundations that we created and the ideas and skills of the creator, it is possible to pioneer diverse contents utilizing the humanoid. This is our expected scenario.

In this paper, we describe the individual technological issues, the background for setting them and the technologies that were actually developed and integrated to solve the issues. Then the work on the verification of the technologies as well

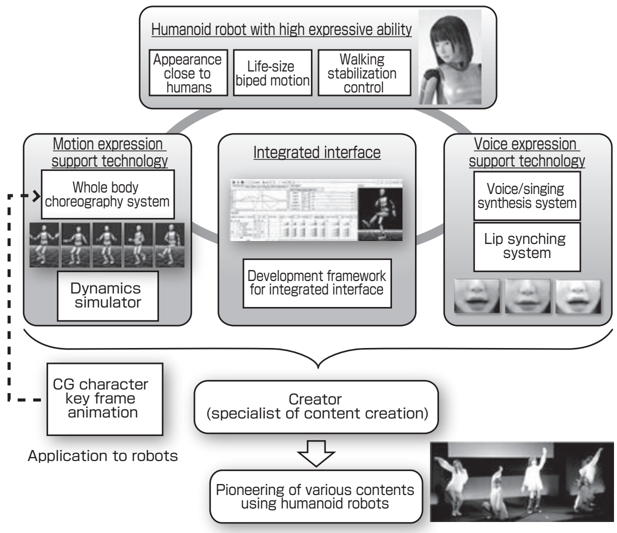

Fig. 1 Outline of research

as the future prospects will be explained.

3 Robot hardware

3.1 Issues of the form and appearance of the robot

While many conventional humanoids claimed to have "human form", most of them could be identified as robots from their form and appearance. While these were effective in the contents that intentionally featured the "robot-like form and appearance", such contents were for minor audiences. If a robot has the form and appearance like humans and possessed expressive ability like humans, it can be used in many categories of contents.

Considering the above points, we set the following two conditions as the issues for robot hardware. The robot:
(1) must be life-sized and capable of stable bipedal motion, and
(2) must have an appearance similar to humans throughout its entire body.

Advanced technology is necessary to fulfill the condition (1), and many humanoids do not fulfill this condition. Some of the examples are humanoids supported by stands and wheels, those connected by cables to external devices, and those that are small in size. While these robots may not possess expressive abilities similar to humans in terms of the whole body action and scale, many robots that fulfill the condition (1) have recently been developed by robotics institutions[1][4]-[9]. However, the appearances of these robots are very "robot-like", and do not fulfill the condition (2). We present examples of the conventional bipedal humanoid that our research group had developed in Fig. 2. Albert HUBO[12], a bipedal humanoid with a realistic head, was developed, but this had the conventional robot-like appearance except for the head. On the other hand, looking only at condition (2), robots that are so close to humans in appearance, so much so that one may not be able to make the distinction, have been developed[13][14]. However, only the upper body moved in such robots and they do not fulfill the condition (1). Therefore, this issue could be cleared if the two technologies were integrated in one robot.

The reason why the robot that fulfills the condition (1) has a robot-like appearance is not only because it was designed so intentionally, but also it is due to the mechanical limitations of size and form of the torso, appendages, and joints. In the conventional robots that fulfill the condition (2), the controller and the power source that drove the multiple joints at high speed to ensure human-like appearance were placed externally to the robot body. To fulfill the two conditions simultaneously by incorporating the mechanism for autonomous bipedal motion in a slender human body to give them a human-like appearance is a difficult technological issue.

The mechanical dolls, as exemplified by the "Audio-Animatronics" of Disneyland, may be called humanoids used in the context of content technology. However, these do not fulfill the condition (1) as they are limited to specific motions and are set in specific places, and do not have the general-use capacity that allow them to be used outside of their original settings. Therefore, they are different from the "content technology" that we have in mind.

3.2 Development of cybernetic human HRP-4C

We set out to solve these issues, and succeeded in the development of cybernetic human[Note 1)] HRP-4C[15][16] as shown in Fig.3. As seen in Fig. 3, HRP-4C is a life-size (height 158 cm) humanoid with an appearance similar to a human being throughout the body. Moreover, HRP-4C has all the mechanisms necessary for motions contained within its body, and is an autonomous robot with bipedal motion. The size and form of the body is close to the average young Japanese woman. It has a slender body compared to a conventional bipedal humanoid. Its total weight is 47 kg[Note 2)].

There is an 8 degrees-of-freedom (DOF) in the head, 3 in the neck, 6 in each arm, 2 in each hand, 3 in the waist, and 7 in each leg. The total joint freedom is 44 ways. The 8 DOF in the head allows changes in the facial expression, movement of the line of sight, and movement of the mouth when speaking. The 3 DOF in the waist allows its smooth movement. In terms of movement, these allow expressive ability close to humans compared to the conventional robots.

As mentioned earlier, the creation of such hardware was a technologically difficult issue. For the development of HRP-4C, a small distributed motor driver combined with a distributed control system, and newly developed ankle joint driving mechanisms were introduced to the design technology nurtured in the development of HRP-2[6] and HRP-3[11]. The actuator and battery were downsized to enable the design with reduced weight capacity. As a result of the integrated effort, we solved the issues and

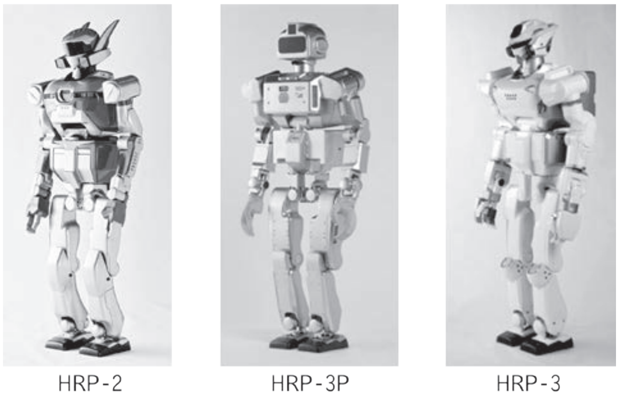

Fig. 2 Conventional biped humanoid robots
From left, HRP-2[6], HRP-3P[10], and HRP-3[11].

achieved the size and weight reductions of the mechanical and electric systems.

3.3 Improvement of the walking stabilization system

To achieve a form similar to humans for the legs in HRP-4C, as seen in Fig. 4, the sole of the feet is smaller compared to the conventional bipedal robots, and the ankle-center is placed close to the heel. In bipedal robots, when the position of the zero-moment point (ZMP) of the sole approaches the edge of the sole, it topples along the edge[17]. Therefore, to prevent falling, it is necessary to keep the ZMP between the sole and floor accurately within the sole, but this becomes difficult to control as the surface area of the sole becomes smaller. We succeeded in obtaining sufficient stability by introducing the new walking stabilization system[18] based on the linear inverted pendulum tracking control that we have been studying as basic research to increase the ability of the bipeds to cover uneven ground.

4 Motion expression support technology

4.1 Issues in choreography of the movement

In addition to the form and appearance of the robot, the movement of the body is, of course, an important element of expression for the content. The function that will be basic for the content technology is that the robot engages in a series of movements designated by the content creator.

The problem is how to choreograph such movements with the robot. In the life-size bipedal robot, the conventional method was to individually develop a program for some specific movements, or to set the commands for the basic preset movements. However, these were non-intuitive tasks that required specialized skills, and the resulting movements tended to be monotonous. Instead of such methods, an easy to understand, efficient method was needed to enable choreography of the various movements as desired by the creator.

We shall take another look at the animation technology for CG characters. The humanoids and CG characters are similar in the point that the desired motion is choreographed to the human-like physical model. The CG character animation is a practical technology that has been used in many moving image contents over the years. Therefore, the technology to choreograph the humanoids in similar ways to the CG characters may be used realistically in the content technology of the humanoids.

The basic methodology in CG animation is called the “key frame animation”. Here, “frame” means the images in sequential order that switches at tens of frames per second to generate the animation. In this method, the key frames are set when the character’s key poses are selected. The poses that fill the areas between the key poses are automatically created for the other frames. This results in the character’s motion. This allows choreographing detailed motions to the character directly, while skipping the between work. The setting of the poses that determine the motion is intuitive and easy to understand.

While this method had been used in the software system intended for use in robots, it was insufficient for “bipedal” robots.

In fact, most systems do not take into account the physical interaction between the robot and the floor in the real world, in processing the generation of the movement from the key poses. In such cases, the movement may be physically impossible for the robot to stay in balance on the feet, and the robot may fall when it is made to carry out the movement. This is the point that differs greatly between the robot and the CG character. In this system, if the robot is small and light and the relative sole size is large, the range where it can stay in balance without falling may be wide, and falls can be prevented depending on the adjustments of the key poses. Because of the size condition, the system was limited to the field of hobby robots, and it was not realistic for our purpose.

The only system that considered the physical behavior

Fig. 3 Cybernetic human HRP-4C
Name is “Miim”.

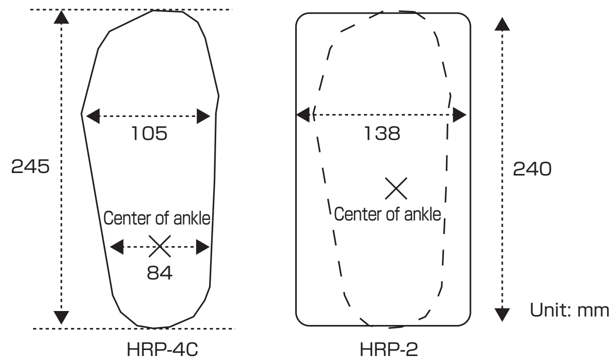

Fig. 4 Size of left soles and the central position of the ankles of HRP-4C and HRP-2
For comparison, the sole of HRP-4C is shown as dotted lines over the sole of HRP-2 on the right.

between the robot and the floor was the SDR Motion Creating System[19] that enabled various movements in a small, 58-cm-tall biped humanoid QRIO[20]. However, in this system, movements could be made from the key poses only for the upper body, and the movement of the lower body could be set only by the specific command and the parameters provided by the system. In this case, the movement of the lower body was limited to the ones allowed by the preset commands, and the task of creating the movement of the whole body was complex. In another light, the fact that a stable movement could be possible only by such a method highlighted the difficulty of the problem.

Until now, the system that allowed the creation of whole body motion of a life-size bipedal humanoid in a way similar to the key frame animation did not exist. We thought such a system would be the fundamental technology for the motion expression of humanoids, and set our topic as the integration of the key frame animation technology and the bipedal robot technology.

4.2 Development of the whole body motion choreography system

As a technology to solve the aforementioned issues, we succeeded in developing the whole body motion choreography system for bipedal humanoids[21]. The interface for the system and the example of the motion created by this system are shown in Fig. 5.

The interface of the system handles the whole body, without dividing the body into the upper and lower parts. As shown in the middle row of Fig. 5, the user can set the key poses on the CG model of the robot. The result will be a stable movement where the robot keeps balance on its feet. The robot can execute the motion without falling as long as there is no self-collision or exceeding of the limits of the joint angle speed (the user will be prompted to correct them when such events occur).

Fig. 5 Editing screen for the whole body choreography system and examples of the created key poses
In this example, eight patterns of key poses are needed to create an action of about 7 seconds where the robot makes the pose, takes a step, and does a kicking motion.

To realize this, we devised an interface design never seen before. The greatest characteristic is that the system determines the horizontal position of the hips at each key pose so the robot can stay in balance. This determination is done instantly as the user enters or corrects the key pose. The result is presented on the spot to the user in the form of the correction of the hip position in the key pose. The supplementation between the key poses is also done to generate a balanced movement. In another word, the system allows only well-balanced choreography. While this maneuver is done, the user does the same operation as in the ordinary key frame animation, and the robot can be choreographed in a manner just like a CG character. It is also possible for the user to explicitly indicate the weight placement of the robot to the floor as the ZMP between the sole and the floor in the key pose, to obtain the desired horizontal movement of the hips within the balanced range.

Such design was not self-evident, and the fact that we reached this design was important in overcoming the issue. Also, implementation of this design was difficult. While providing a simple interface to the user, the system must integrate various complex computations such as the detection of sole landing status, the transition of target ZMP, the transition of supplementary space, the addition of supplementary key poses, and the calculation of center of gravity path that matches the target ZMP. Moreover, these have to be done at high speed.

In overcoming such implementation issues, we were able to utilize the technology of OpenHRP3[22], the dynamics simulator for robots that we have been developing. The various calculation processes for robotics implemented in the simulator were developed to handle practical simulation in terms of execution speed and precision, and it was useful for the implementation of this system. The dynamics simulation method that we developed for OpenHRP3[23] could accurately verify the behavior of the bipedal robot on the floor. By incorporating this into our system, we could directly verify the adequacy of the implementation to the system, and were able to increase the development efficiency.

This system was confirmed to be effective for HRP-4C. By combining the whole body motion created with the walking stabilization system[18] mentioned in subchapter 3.3, stable execution was confirmed to be possible for HRP-4C, as in the example of Fig. 5. To be able to construct such complex movements in a life-size humanoid with small soles, as in HRP-4C, was a major accomplishment. Also, as shown in Fig. 6, it was possible to create the changes in facial expressions[24]. In this system, the operation of the actuators in the head can be reflected in real time on the robot, and therefore, the changes of the fine facial expressions, which was difficult to render perfectly by CG, could be made directly. This system can bring out the expressive ability of the movement of HRP-4C.

4.3 Method of using the motion capture

In the CG character animation, a method with which actions of an actual person are incorporated by motion capture is also widely used. A method for applying captured human whole body motion to the whole body motion of the bipedal humanoid was developed[25]. Using this method, the whole-body performance of a Japanese folk dance called Aizu Bandaisan was done by HRP-2[6].

Comparing the motion capture method and the method developed in this research the former is, of course, more applicable in recreating the human actions. However, it must be noted that due to the limitations of the method and the limitations of the robot's movements, the human action cannot be completely reproduced in the robot. On the other hand, to express the action unique to a robot or to create high quality action within the limit of the robot's motion capacity, the latter method that choreographs the robot directly is more applicable. Moreover, the former requires a skilled performer that can move in a certain way, as well as specialized equipment and a studio, while the latter does not require such equipment and can be done easily on a PC.

Considering the above characteristics, for the objective of the creation of new contents using the humanoid and the diffusion of its use, the method of this research will be used as the base. It is highly significant that this was realized for the first time. On the other hand, the motion capture method can be also useful, and one of the future topics will be to integrate the two methods.

5 Voice expression support technology

The voice of the robot is an important expression element for the content. Research has been done on the speech mechanism simulating the human vocal cord[26]. However, this is a large mechanism including the lungs, and cannot be currently installed in a humanoid like HRP-4C. Therefore, production of some voice source through the speaker is adequate as the source of speech. In this case, to make it look as if the robot is speaking, the robot's mouth must move according to the voice source (lip synching). To obtain the vocal source, there are methods of using the human speech or using the voice synthesis technology. The difference of the characteristics of the two is similar to the difference between the two methods for creating the actions as mentioned in subchapter 4.3. In that sense, use of the voice synthesis will be more appropriate for our purpose.

(a)Smile (b)Surprise (c)Anger

Fig. 6 Examples of expression creation

From the above considerations, the issue for voice expression would be to enable various speech and singing expressions with the voice synthesis technology in linkage with the mouth movement.

To solve this issue, we developed a system using the VOCALOID[27], the song synthesis technology of Yamaha Corporation with whom we worked jointly, as the voice expression of HRP-4C[28]. The VOCALOID was developed to synthesize the singing voice, and is a technology that generates a singing voice very close to humans. Moreover, the VOCALOID-flex technology, where the technology is applied to produce natural speech with rich intonations, is available, and diverse voice expressions are possible. The system allows the robot to lip synch the voice data of the VOCALOID, and this enables easy creation of natural speech and singing performances by the robot.

6 Integrated interface

For the implementation of the whole body motion choreography system mentioned in subchapter 4.2, it was necessary to implement the various functions including the management of various data, the display and operation of the 3D models, the sequential display of the key poses, and the dynamic simulations in a collaborative format, in addition to the essential key pose processing. To create the integrated expressions of the robot and to have the robot perform them, it is necessary to link the choreography system and voice expression support technology to the robot hardware, and to provide an interface that allows easy use by the user. Moreover, the usefulness as a content technology will increase further if the information and media technologies for the robot expression including the existing technologies and those that will be developed in the future can be collaboratively used. The motion capture technology mentioned in subchapter 4.3 is an example of such a useful technology.

To realize the above, we developed the "Choreonoid framework", a software framework for the development of the integrated interface. The interface for the technology developed and selected in this study was implemented on the framework, and resulted in the integrated software Choreonoid[Note 3].

The Choreonoid framework is based on the C++ language, is compatible with programs written in C or C++, and is capable of realizing the algorithms and interface that require high-speed processing. Also, it is designed based on the architecture called Model View Controller and the signal mechanism, and while keeping the maintainability and expandability by increasing the independency of the objects, it is possible to execute complex associations between the objects[29]. New functions can be added as plug-ins and the new plug-ins can easily collaborate with the existing functions and other plug-in functions. Due to such characteristics, the Choreonoid surpasses the framework of the content technology, and is expected to be used widely as an environment to develop higher layer robot software.

7 Experiment in content production

In this research, to promote the technology, it was necessary to actually produce the contents using the Choreonoid and HRP-4C, and to verify and improve the technology. Considering the objective of this study, it was important to have the content created by professional creators rather than by us. By doing so, the practicality of the system could be verified, and new contents could be pioneered.

As a joint activity with Dr. Masaru Ishikawa, Project Researcher of the Information and Robot Technology Research Initiative, The University of Tokyo, we asked Mr. SAM, a famous dancer and dance creator, to create an entertainment content using HRP-4C. This was supported by the DC-EXPO (DIGITAL CONTENT EXPO)[Note 4] 2009 and DC-EXPO 2010, and the resulting performance was presented at the exposition. We shall now describe the efforts at the DC-EXPO 2010 that was done after the development of Choreonoid.

At the DC-EXPO 2010, to fully verify the technologies described in this paper, we tried to develop a content where the HRP-4C would sing and dance, as shown in Fig. 7. In this content, HRP-4C gave a three-minute demonstration of a dance, choreographed by Mr. SAM, and a song “Deatta Koro No Youni (Like When We First Met)” by Every Little Thing, a Japanese music group.

Fig. 7 Demonstration of singing and dancing by HRP-4C

The dance motion was created entirely by Choreonoid based on Mr. SAM’s choreography. We told Mr. SAM the robot’s motion capacity, and then had him choreograph freely within the possible motion range. The dance motion included various poses and actions using the entire body as shown in Fig. 8, and the richness of expression was unseen in any previous performance by a robot. We changed the wig of HRP-4C and dressed HRP-4C in a costume to match the dance, and the stage lighting was also carefully planned. Mr. SAM’s choreography included the total combination performance with four human back dancers.

For singing, we asked the cooperation of YAMAHA Corporation, and obtained the voice data and lip synch data through the VOCALOID. The VOCALOID sound source was CV-4Cβ of the Crypton Future Media, Inc. Using the VocaListener technology of Nakano and Goto[30], tuning was based on the song by Ms. Kaori Mochida, the singer of the original song (voice track of the vocal part by Ms. Kaori Mochida supplied by avex trax). This ensured expressive singing close to that of a human being. The lip synch data was produced automatically and was created efficiently from the voice data by using the VOCALOID linkage function explained in chapter 5.

This dance demonstration generated news mainly on the Internet since it was presented on October 16, 2010. As an example for its newsworthiness, the movies uploaded to YouTube by some audiences recorded over 2 million view counts and over 1,500 comments in 10 days after the presentation. One of the movies had the 6th highest view counts for YouTube Japan in the year 2010. Immediately after

Fig. 8 Example of various movements and poses actually done in the dance

the great response on the Internet, we received several offers for appearances of HRP-4C from various places including overseas. This showed that the expression created by the creator using the technology in this study was recognized as a totally new content, and demonstrated the effectiveness of our scenario.

The entry of the key poses on the Choreonoid was done by one of the authors, and the time required for entry was about 80 hours. A rather large amount of time was required for the three-minute performance, but this was mainly because the person did not have any experience creating CG character animation. Needless to say, verification by professional creators, including the key pose entry work, is necessary and we plan to have professional CG creators work directly on Choreonoid to create the contents.

8 Future prospects

Since the press release in March 2009, HRP-4C has received requests and suggestions for its use from various places. It gave the opening speech in the "SHINMAI Creator's Project" of the 8th Japan Fashion Week in Tokyo held in March 2009, was a model for the wedding dress in Yumi Katsura Paris Grand Collection in Osaka held in July 2009, and dressed up as a VOCALOID character such as "Hatsune Miku" and gave singing demonstrations at the Yamaha Corporation booth in CEATEC JAPAN 2009 in September 2009[31][29].

What is significant here is that the requests from the general public for HRP-4C do not stop at asking the robot developer to operate the robot, but are active requests where the requesters propose what they wish the robot to perform. This was unseen in the previous robots such as HRP-2, and in this sense, the strategy for the form and appearance of the robot mentioned in subchapter 3.1 was successful.

With the motion expression and voice expression support technologies that were developed in this study, the contents utilizing the expressive ability of the humanoid, as in the dance at the DC-EXPO 2010, became possible, and we are now able to respond to the various proposals for the use of HRP-4C. To verify this further, we believe we must continue the work, as mentioned in chapter 7, with several creators.

While engaging in such promotional activities, there are still many things that must be done technologically. In the area of bipedal motion, movements that include sliding and jumping as well as more human-like walking with the knees straight cannot be done in the current Choreonoid, but these are necessary to increase the range and quality of the contents. While there are examples where such motions are accomplished individually[32]-[34], the creation of movements that freely combine these elements still remains as a difficult issue.

The improvement of autonomy is an issue necessary to present a more natural motion. For example, the movement of the eyes must be automated to look more natural. In striking a pose, rather than stopping completely, it will look more natural if slight swaying was generated automatically.

Further improvements are necessary in the robot technology as a whole. The robot must move adaptively to the surrounding environment in some contents, and research results on environmental recognition and motion planning based on it or on human interactions are needed. Also, in contents that require use of props, manipulation ability is necessary.

The various technologies for the robot can be developed one step at a time as the contents are created. Therefore, the effort of pioneering the contents using the humanoid is effective in developing the robotics technology and applying them to industry. If the ability of the robot increases as a result, the path to the practical use of humanoids outside the content technology, including daily activity support, may be opened.

With the above experience behind us, we plan to continue the R&D for the humanoid as a content technology.

Acknowledgements

The development of HRP-4C was conducted as part of the "Development of User Centered Robot Open Architecture (UCROA)", AIST Innovative Research Initiative, an AIST-industry-academia collaborative project. The development of Choreonoid was based on the "motion pattern design tool" developed with the support of NEDO "Development Project for Intelligent Technologies of Next-Generation Robots".

For the execution of the UCROA Project, we received great support from Hirohisa Hirukawa, director of the Intelligent Systems Research Institute, AIST and Junji Ito, former vice-president of AIST. We received useful comments from Kiyoshi Fujiwara, Kensuke Harada, Takeshi Sakaguchi, Hitoshi Arisumi, Nobuyuki Kita, Isao Hara and Tokuo Tsuji of AIST, as well as Neo Ee Sian, formerly of AIST. The content for the DC-EXPO 2010 was produced through the cooperation of: Masaru Ishikawa, researcher, Information and Robot Technology Research Initiative, The University of Tokyo; SAM, choreographer; Osamu Oshima, Hideki Kenmochi, Makoto Tachibana of Yamaha Corporation; and many other people.

Notes

Note 1) This is a term that denotes a humanoid robot that has an appearance and form similar to humans, is capable of walking and moving in a manner extremely close to humans, and is capable of interaction with people using voice recognition and other functions.

Note 2) This is the latest specification as of writing of this paper. This stands true for the other descriptions in this paper.

Note 3) The name "Choreonoid" is a combination of "choreograph" and "humanoid", and expresses the choreography function which is the core of this software.

Note 4) The DIGITAL CONTENT EXPO (DC-EXPO) is an international event organized by the Ministry of Economy, Trade and Industry and the Digital Content Association of Japan. It has been held in October since 2008, at the National Museum of Emerging Science and Innovation and the Tokyo International Exchange Center.

Note 5) More accurately, we used the "NetVocaListener" service developed by the Yamaha Corporation that obtained the license for this technology from AIST.

References

[1] K. Hirai, M. Hirose, Y. Haikawa and T. Takenaka: The development of Honda humanoid robot, *Proceedings of the 1998 IEEE International Conference on Robotics and Automation, Leuven, Belgium*, 1321-1326 (1998).

[2] E. S. Neo, T. Sakaguchi, K. Yokoi, Y. Kawai and K. Maruyama: A behavior level operation system for humanoid robots, *Proceedings of the 6th IEEERAS International Conference on Humanoid Robots, Genoa, Italy*, 327-332 (2006).

[3] K. Okada, M. Kojima, S. Tokutsu, Y. Mori, T. Maki and M. Inaba: Integrating recognition and action through task-relevant knowledge for daily assistive humanoids, *Advanced Robotics*, 23 (4), 459-480 (2009).

[4] Y. Sakagami, R. Watanabe, C. Aoyama, S. Matsunaga, N. Higaki and K. Fujimura: The intelligent ASIMO: System overview and integration, *Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems, EPFL, Lausanne, Switzerland*, 2478-2483 (2002).

[5] S. Kagami, K. Nishiwaki, J. Kuffner, Y. Kuniyoshi, M. Inaba and H. Inoue: Online 3D vision, motion planning and bipedal locomotion control coupling system of humanoid robot H7, *Proceedings of the 2002 IEEE/ RSJ International Conference on Intelligent Robots and Systems, EPFL, Lausanne, Switzerland*, 2557-2562 (2002).

[6] K. Kaneko, F. Kanehiro, S. Kajita, H. Hirukawa, T. Kawasaki, M. Hirata, K. Akachi and T. Isozumi: Humanoid Robot HRP-2, *Proceedings of the 2004 IEEE International Conference on Robotics and Automation, New Orleans, LA, USA*, 1083-1090 (2004).

[7] I. W. Park, J. Y. Kim, J. Lee and J. H. Oh: Mechanical design of humanoid robot platform KHR-3 (KAIST humanoid robot -3: HUBO), *Proceedings of the 5th IEEE-RAS International Conference on Humanoid Robots, Tsukuba, Japan*, 321-326 (2005).

[8] Y. Ogura, H. Aikawa, K. Shimomura, H. Kondo, A. Morishima, H. ok Lim and A. Takanishi: Development of a humanoid robot WABIAN-2, *Proceedings of the 2006 IEEE International Conference on Robotics and Automation, Orlando, Florida, USA*, 76-81 (2006).

[9] S. Takagi: Toyota partner robots, *Journal of the Robotics Society of Japan*, 24 (2), 208-210 (2006) (in Japanese).

[10] K. Akachi, K. Kaneko, N. Kanehira, S. Ota, G. Miyamori, M. Hirata, S. Kajita and F. Kanehiro: Development of humanoid robot HRP-3P, *Proceedings of the 5th IEEE-RAS International Conference on Humanoid Robots, Tsukuba, Japan*, 50-55 (2005).

[11] K. Kaneko, K. Harada, F. Kanehiro, G. Miyamori and K. Akachi: Humanoid Robot HRP-3, *Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, Nice, France*, 2471-2478(2008).

[12] J. H. Oh, D. Hanson, W. S. Kim, I. Y. Han, J. Y. Kim and I. W. Park: Design of android type humanoid robot Albert HUBO, *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China*, 1428-1433 (2006).

[13] H. Ishiguro: Android science: Conscious and subconscious recognition, *Connection Science*, 18 (4), 319-332 (2006).

[14] D. Sakamoto, T. Kanda, T. Ono, H. Ishiguro and N. Hagita: Android as a Telecommunication Medium with a Human-like Presence, *Arlington, Virginia, USA*, 193-200 (2007).

[15] K. Kaneko, F. Kanehiro, M. Morisawa, K. Miura, S. Nakaoka and S. Kajita: Cybernetic Human HRP-4C, *Proceedings of the 10th IEEE-RAS International Conference on Humanoid Robots, Paris, France*, 7-14(2009).

[16] K. Kaneko, F. Kanehiro, M. Morisawa, T. Tsuji, K. Miura, S. Nakaoka, S. Kajita and K. Yokoi: Hardware Improvement of Cybernetic Human HRP-4C for Entertainment Use, *Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Francisco, California, USA*, (2011), accepted.

[17] M. Vukobratovic, B.Borovac, D.Surla and D.Stokic: Biped Locomotion: Dynamics, Stability, Control and Application, *Scientific Fundamentals of Robotics,* 7, Springer-Verlag (1990).

[18] S. Kajita, M. Morisawa, K. Miura, S. Nakaoka, K. Harada, K. Kaneko, F. Kanehiro and K. Yokoi: Biped walking stabilization based on linear inverted pendulum tracking, *Proceedings of the 2010 IEEE/ RSJ International Conference on Intelligent Robots and Systems, Taipei, Taiwan* (2010).

[19] Y. Kuroki, B. Blank, T. Mikami, P. Mayeux, A. Miyamoto, R. Playter, K. Nagasaka, M. Raibert, M. Nagano and J. Yamaguchi: Motion creating system for a small biped entertainment robot, *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems, Las Vegas, Nevada*, 1394-1399 (2003).

[20] Y. Kuroki, M. Fujita, T. Ishida, K. Nagasaka and J. Yamaguchi: A small biped entertainment robot exploring attractive applications, *Proceedings of the 2003 IEEE International Conference on Robotics and Automation, Taipei, Taiwan*, 471-476 (2003).

[21] S. Nakaoka, S. Kajita and K. Yokoi: Intuitive and flexible user interface for creating whole body motions of biped humanoid robots, *Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, Taipei, Taiwan*, 1675-1682 (2010).

[22] S. Nakaoka, N. Yamanobe, H. Hirukawa, K. Yamane and Y. Kawasumi: OpenHRP3: A robot simulator implemented as distributed objects, *Journal of Robotics Society of Japan*, 26, 5, 399-406 (2008) (in Japanese).

[23] S. Nakaoka, S. Hattori, F. Kanehiro, S. Kajita and H. Hirukawa: Constraint based dynamics simulator for humanoid robots with shock absorbing mechanisms, *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Diego, CA, USA*, 3641-3647 (2007).

[24] S. Nakaoka, F. Kanehiro, K. Miura, M. Morisawa, K. Fujiwara, K. Kaneko, S. Kajita and H. Hirukawa: Creating facial motions of cybernetic human HRP-4C, *Proceedings of the 10th IEEE-RAS International Conference on Humanoid Robots, Paris, France* (2009).
[25] S. Nakaoka, A. Nakazawa, F. Kanehiro, K. Kaneko, M. Morisawa, H. Hirukawa and K. Ikeuchi: Learning from Observation Paradigm: Leg Task Models for Enabling a Biped Humanoid Robot to Imitate Human Dances, *International Journal of Robotics Research*, 26(8) , 829-844, SAGE Publications (2007).
[26] K. Fukui, Y. Ishikawa, K. Ohno, N. Sakakibara, M. Honda and A. Takanishi: Three dimensional tongue with liquid sealing mechanism for improving resonance on an anthropomorphic talking robot, *Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, Kobe, Japan*, 5456-5462 (2009).
[27] H. Kenmochi and H. Ohshita: Singing synthesis system "VOCALOID" - Current situation and to-do lists, *Information Processing Society of Japan Technical Reports*, 2008-MUS-74-9, 12, 51-58 (2008) (in Japanese).
[28] M. Tachibana, S. Nakaoka and H. Kenmochi: A Singing Robot Realized by a Collaboration of VOCALOID and Cybernetic Human HRP-4C, *InterSinging 2010, Tokyo, Japan*, WA01 (2010).
[29] Shin'ichiro Nakaoka: Choreonoid as a Software Framework for Implementing Graphical Robotics Applications, *Proceedings of the 29th Annual Conference of the Robotics Society of Japan, Tokyo, Japan* (2011), to appear.
[30] Tomoyasu Nakano and Masataka Goto: VocaListener: A Singing-to-Singing Synthesis System Based on Iterative Parameter Estimation, *Proceedings of the 6th Sound and Music Computing Conference (SMC 2009)*, 343-348 (2009).
[31] K. Miura, S. Nakaoka, S. Kajita, K. Kaneko, F. Kanehiro, M. Morisawa and K. Yokoi: Trials of cybernetic human HRP-4C toward humanoid business, *Proceedings of the 6th IEEE Workshop on Advanced Robotics and its Social Impacts, Seoul, Korea*, WA01 (2010).
[32] K. Miura, S. Nakaoka, M. Morisawa, F. Kanehiro, K. Harada and S. Kajita: Analysis on a Friction Based ``Twirl'' for Biped Robots, *Proceedings of the 2010 IEEE International Conference on Robotics and Automation, Anchorage, Alaska, USA*, 4249-4255 (2010).
[33] Ryosuke Tajima and Keisuke Suga: Motion having a Flight Phase: Experiments Involving a One-legged Robot, *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China*, 1726-1731 (2006).
[34] Y. Ogura, K. Shimomura, H. Kondo, A. Morishima, T. Okubo, S. Momoki, H. ok Lim and A. Takanishi: Human-like walking with knee stretched, heel contact and toe-off motion by a humanoid robot, *Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China*, 3976- 3981 (2006).

Authors

Shin'ichiro NAKAOKA

Received the B.S. degree in information science from Tokyo Institute of Technology, Japan, in 2001, and the Ph.D. degree in computer science from the University of Tokyo, Japan, in 2006. Since 2006, he has been a researcher with the National Institute of Advanced Industrial Science and Technology (AIST), Japan. His research interests include motion presentation by humanoid robots and robotics software platforms. He was the recipient of the Best Paper Award of the Robotics Society of Japan in 2008. In this research, he was mainly in charge of the development of motion expression support technology, voice expression support technology, and integrated interface, as well as the experiments in content production.

Kanako MIURA

Received the Ph.D. degree in information science from the Tohoku University, Japan, in 2004, and also the Ph.D. degree in Electronique, Electrotechnique, Automatique from l'Université Louis Pasteur, France, in 2004. From 2004 to 2005, she was a Postdoctoral Fellow with the Tohoku University, Japan. From 2005 to 2007, she was a researcher with NTT Docomo, Japan. Since 2007, she has been a researcher with the National Institute of Advanced Industrial Science and Technology (AIST), Japan. Her research interests include human movement analysis and motion planning for humanoid robots. In this research, she was mainly in charge of the development of HRP-4C.

Mitsuharu MORISAWA

Received the Ph.D. degree in engineering from Keio University, Japan, in 2004. Researcher of the Intelligent Systems Research Institute, AIST from April 2004. Visiting researcher at the LAAS-CNRS, France for one year from April 2009. Engages in research of parallel mechanism, motion control, and humanoid robots. In this research, was mainly in charge of the development of HRP-4C.

Fumio KANEHIRO

Received the Ph.D. degree in engineering from The University of Tokyo, Japan, in 1999. Special researcher of the Japan Society for the Promotion of Science in 1998. Joined the Electrotechnical Laboratory, Agency of Industrial Science and Technology in 2000. Due to reorganization, became the researcher of the Intelligent Systems Research Institute, AIST from 2001. Senior researcher in October 2007. Visiting researcher at the LAAS-CNRS, France from April 2007. Interested in the system composition and whole body motion control of the humanoid robot. In this research, was mainly in charge of the development of HRP-4C.

Kenji KANEKO

Received the B.E., M.E., and Ph.D. in electrical engineering from Keio University in 1988, 1990, and 1997, respectively. In 1990, he joined the Mechanical Engineering Laboratory, Agency of Industrial Science and Technology, Ministry of Industrial Science and Technology (AIST-MITI). From September 1995, he was a visiting researcher at the Carnegie Mellon University, U.S.A. for half a year. From September 1999, he was a visiting researcher at the Paris Robotics Laboratory (LRP), CNRS, France for one year. He is currently the senior researcher of the Intelligent Systems Research Institute, AIST. His research interests include motion control and humanoid robot. He received the IEEE-RAS International Conference on Humanoid Robots Best Paper Award (Oral Category) in 2009. In this research, he was mainly in charge of the development of HRP-4C.

Shuuji KAJITA

Received the master's degree (1985) and the doctor of engineering degree (1996) from Tokyo Institute of Technology, Japan. In 1985, he joined the Mechanical Engineering Laboratory, Agency of Industrial Science and Technology, Ministry of International Trade and Industry (AIST-MITI). Meanwhile he was a visiting researcher at California Institute of Technology, 1996-1997. Currently he is a senior researcher at the National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan, which was reorganized from AIST-MITI in April 2001.His research interests include robotics and control theory. He is a member of Society of Instrument and Control Engineers, Robotics Society of Japan and IEEE (Robotics and Automation Society). Received the best paper award of the Society of Instrument and Control Engineers in 1996 and the best paper award of the Robotics Society of Japan in 2005. In this research, was mainly in charge of the development of HRP-4C and the walking stabilization system.

Kazuhito YOKOI

Received his M. E. and Ph. D. degrees in Mechanical Engineering Science from Tokyo Institute of Technology in 1986 and 1994, respectively. Joined the Mechanical Engineering Laboratory, Agency of Industrial Science and Technology in 1986. Senior researcher of the Intelligent Systems Research Institute, AIST from 2001. Research group leader in 2004, and deputy-director of the Intelligent Systems Research Institute, AIST in 2009. Visiting researcher at Stanford University in 1995~1996; co-director of the CNRS-AIST Joint Robotics Laboratory in 2003~2008; and adjunctive professor of the Graduate School of Systems and Information Engineering, University of Tsukuba (Cooperative Graduate School) from 2005. Engages in the research for enhancing the autonomy of humanoid robots. Received the Young Engineers award of the Japan Society of Mechanical Engineers in 1993; Best Paper Award of the Society of Instrument and Control Engineers in 1993; Best Video Award of the IEEE ICRA for 2003 and 2004; Best Paper Award of the Robotics Society of Japan in 2005; Robotics and Mechatronics Academic Merit Award, Japan Society of Mechanical Engineers in 2006; Best Paper Award of the Japan Society of Mechanical Engineers in 2008; and many others. In this research, was mainly in charge of the project management and the contents production experiment.

Discussions with Reviewers

1 Overall composition in terms of synthesiology

Comment (Masaaki Mochimaru, Digital Human Research Center, AIST)

[Composition of the paper in terms of synthesiology]

In terms of synthesiology, I see this paper as an integration of the three elemental technologies of life-sized, realistic humanoid (platform as hardware), motion expression support technology (whole body motion generation + dynamic simulator), and singing expression support technology (voice synthesis technology + lip synch technology). Verification was then done for whether the performance created by the integrated technologies would be recognized as an unprecedented new media content. As a synthesiology paper, I think you should clarify this position and organize the chapters along this line.

Comment (Naoto Kobayashi, Center for Research Strategy, Waseda University)

[Diagram that allows understanding of the whole picture in terms of synthesiology]

The elemental technologies that include the completion of cybernetic human HRP-4C, the creation of Choreonoid, the whole body movement software, the realization of facial expression creation function, and the development and use of dynamic simulator are extremely important. As I see it, the final objective was achieved by integrating these technologies. I think the reader's understanding will be improved if you provide a diagram that describes the elemental technologies in blocks, showing the relationship between the elemental technology blocks and how they were combined and synthesized to approach the final goal. I think this is also extremely important from the perspective of "synthesiology" as an academic study.

[Description of technological difficulties in integration]

I think you should explain the major difficulties that had to be overcome in integrating the technologies.

[Feedback from integration to elemental technology]

In conducting the integration, I think you had to have feedback for the correction and improvement of the individual elemental technologies. Please give us an actual example of a feedback that provides a clear case study that will be useful to the readers.

Answer (Shin'ichiro Nakaoka)

[Composition of the paper in terms of synthesiology]

The description was altered based on the comment from the reviewer. For the composition of the chapters, the four technological domains that were the issues in achieving the objective of this research were briefly summarized in chapter 2. The details of the issues and the efforts to solve them were described in the chapters for the four technological domains.

[Diagram that allows understanding of the whole picture in terms of synthesiology] Figure 1 was added to show the outline of this research.

[Description of technological difficulties in integration]

In the large framework of "robot hardware", "motion expression support technology", "voice expression support technology", and " integrated interface" as elemental technologies, the difficulties were to select the necessary elements, and to figure out how to integrate them as a whole. On the other hand, in the individual framework of robot hardware and motion expression, it was difficult to integrate the "human-like appearance" and "bipedal motion". It was also difficult to integrate the "CG key frame animation" and "maintaining stable motion of the bipedal robot". The descriptions of these difficulties were included in the corresponding chapters.

[Feedback from integration to elemental technology]

There were actual case studies of feedbacks. Particularly, from such feedbacks, it became necessary to improve the walking stabilization system, and I added the details in subchapter 3.3.

2 Contents industry for humanoids

Question (Naoto Kobayashi)

You described that you are aiming to commercialize the humanoid as a content medium. At this point, other than the "singing and dancing", "fashion show host", and "model" that were given as content examples in this paper, what other demands do you expect in the future, and what is the projection of the market scale of this industry (humanoid content industry or humanoid amusement industry)? What other technologies, opportunities, or additions do you think would help such an industry to grow further?

Answer (Shin'ichiro Nakaoka)

Other than the examples in the paper, the specific contents using HRP-4C include theatrical performance, message video for a coming-of-age ceremony, presentation at trade shows, and others. The HRP-4C greeted the heads of the nations in the exhibit of Japan's state-of-art technology during the summit meeting of the Asia Pacific Economic Cooperation (APEC) forum held in 2010. Including the examples in the paper, the range of application of the HRP-4C in "communication", "presentation" ,"expression", and "performance" is wide, and there is no end to the contents when meeting specific demands. This part is the work of the creators, and we would like to consult with the creators of various fields.

In terms of industry, we are considering the use of robots in the framework of the current contents industry or the amusement industry. This expectation is described as the "idol robot" in the Technological Strategy Map 2010 of the METI, and this extends to the industries of music, motion pictures, dramas, amusement parks, and tourism. The market for live entertainment is 1 trillion yen (2007). Within this market in Japan, the initial goal for the share of robots (robot hardware, software, operation service, etc.) will be to attain a multibillion-yen scale within Japan. After that, with the technological advancement mentioned in "8 Future prospects," this industry can evolve further by expanding the application of the robots. On the point of opportunities and nurturing the industry, here again, collaboration with the creators will be important.

3 Uniqueness of the Japanese culture

Question (Naoto Kobayashi)

Do you think the affinity people have for humanoids may be rooted in the unique Japanese culture? In other cultures, the responses people have for this type of robots may be different. What are the positive and negative responses in international academic societies?

Answer (Shin'ichiro Nakaoka)

The responses we received directly from the researchers at the international academic societies were mostly positive, regardless of nationality. I think this is mainly because the researchers evaluated the technology itself, and the robot researchers love robots to begin with. Therefore, the responses of different countries cannot be measured from the responses at the academic societies. On the other hand, we received comments on the Internet through the video upload sites, and they were not entirely positive. Several English comments written most probably by people overseas said, "This robot will eventually start attacking people" or "It will be used for war". There were hardly any such comments written in Japanese, and I did feel the difference in culture here. We do have many, many comments, and if we do some statistical analysis, we might get an accurate picture of the differences in perception.

4 Technological evolution of the humanoid and the limitation on the attractiveness of the contents

Question (Naoto Kobayashi)

Currently, people are attracted to the fact that a humanoid is moving like a person, and this is provided as a content. In the future, as the technology advances, the humanoids will be able to do things humans cannot do. It was great news when the computer won the chess game against a human for the first time, but people gradually lost interest. I feel that the contents featuring humanoids may reach a point where people start losing interest. What do you think is such a point and when do you think it would arrive? What kind of measures should you take to avoid this situation?

Answer (Shin'ichiro Nakaoka)

The factors that attract people to the contents featuring humanoid robots include the "interest for the robot itself" and the "interest in the contents". For the former, people will be continually interested in how the robots will become closer and closer to humans, and then eventually surpass humans. For the latter, the characteristics and the functions of the robot can be used in the expression and execution of the interesting and attractive contents. As the robots continue to evolve and become capable of doing what humans cannot do, the range of expression and execution of the contents will expand, and the content itself will become more attractive.

5 Strategy of humanoid R&D

Question (Naoto Kobayashi)

One of the strategies and policies of the humanoid R&D is to start with the use of the humanoid featured in the contents media, as in this study. The basic performances such as the motion ability, sturdiness, and safety of the humanoids are enhanced as they are used, and the improvement in the basic performance can be applied to other uses including the support of daily human activities. I think this strategy is appropriate. On the other hand, I think there is a powerful argument that a robot does not have to be humanoid in form when supporting human functions. Considering the psychological aspects, why do you think the humanoids are necessary?

Answer (Shin'ichiro Nakaoka)

If we increase the tasks that can be taken over by robots to support human activities, I think the physical form of the robot will necessarily become human-like, due to the reason given in paragraph 2 of "1 Humanoid robot as content technology". I also think the human-like form is effective for supporting people from the psychological aspects. It is a matter of how much one expects from robots. To respond to the high demands of human activity support, I think the "humanoid form" becomes necessary in addition to the various developments of the elemental technologies.