Robotica (2016) volume 34, pp. 652–670. © Cambridge University Press 2014
doi:10.1017/S0263574714001787

An expressional simplified mechanism in anthropomorphic face robot design

Chyi-Yeu Lin†*, Chun-Chia Huang‡ and Li-Chieh Cheng§

†*Department of Mechanical Engineering, National Taiwan University of Science and Technology, Taipei, 10607 Taiwan*
‡*Department of Mechanical Engineering, National Taiwan University of Science and Technology, Taipei, 10607 Taiwan Email: D9603110@mail.ntust.edu.tw*
§*Department of Mechanical Engineering, National Taiwan University of Science and Technology, Taipei, 10607 Taiwan Email: D9503101@mail.ntust.edu.tw*

(Accepted Jun 09, 2014. First published online: July 9, 2014)

SUMMARY
The goal of this research is to develop a low-cost face robot which has a lower degree-of-freedom facial expression mechanism. Many designs of facial robots have been announced and published in the past. Face robots can be classified into two major types based on their respective degrees of freedom. The first type has various facial expressions with higher degrees of freedom, and the second has finite facial expressions with fewer degrees of freedom. Due to the high cost of the higher-degree-of-freedom face robot, most commercial face robot products are designed in the lower-degrees-of-freedom form with finite facial expressions. Therefore, a face robot with a simplified facial expression mechanism is proposed in this research. The main purpose of this research is to develop a device with a lower degree-of-freedom mechanism that is able to generate many facial expressions while keeping one basic mouth shape variation. Our research provides a new face robot example and development direction to reduce costs and conserve energy.

KEYWORDS: Face robot; Facial expression; Facial mechanism; Expressional simplified mechanism.

1. Introduction

Because of the needs of humans and the rapid advancements in science, more and more service robots have been created and used in daily life. Therefore, many human-robot-interaction (HRI) related researches have been processed. The majority of these researches are studies of androids. There are two mainstreams of android robot research with expressional faces. One is with hard-shell type robotic head. Typical robots with hard-shell appearances are Kismet,[1] WE-4RII,[2] Nexi,[3] iCub,[4] Simon,[5] and Flobi.[6] This kind of robot simplified its facial expressions by only showing key facial features such as eyes, eyebrows, and mouth. The surface of its face cannot be deformed like human skin, but the mechanism designs behind the hard-shells have inspired some researchers who devote themselves to develop face robots with soft polymeric facial skin.

The other mainstream of android robot research with facial expressions is to cover the robot face with soft, elastic, and realistic polymeric facial skin. The first face robot covered with elastic skin and with human-like touch was developed by Hara and Kobayashi[7] in Japan in the 1990s. The authors used a kind of flexible micro-actuator (FMA) driven by air pressure to control the artificial skin made of a silicone membrane. The softness and elasticity of the skin made it easy to generate 6 typical facial expressions including surprise, fear, disgust, anger, happiness and sadness by pulling 19 control points (CPs) located underneath the skin. The authors also developed another kind of face robot that adopted shaped memory alloys as actuators[8,9] and showed 6 typical facial expressions in their papers.

* Corresponding author. E-mail: jerrylin@mail.ntust.edu.tw

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press

In order to improve on the previous generations of their face robots, the authors used the Mckibben pneumatic actuators as artificial muscles to generate facial expressions.[10] In this study, they used wires to connect the actuators and 18 CPs that were set underneath the artificial skin. The face robots were also able to generate at least 6 typical facial expressions.

Hanson *et al*.[11,12] applied electro-active polymers (EAP) as artificial muscles to change the shape of the artificial facial skin. The length of the skin can vary by changing the level of the input voltage. Finite facial expressions are shown in his paper. Furthermore, Ishiguro *et al*.[13–15] built some android robots from 2004 to 2006 such as Repliee R1, Repliee Q1 and Repliee Q2. The head of Repliee R1 used DC motors to generate motions. There are 5 degrees of freedom for its eyes and 1 degree of freedom for its mouth. The other robots used air actuators; each head of these two face robots had 13 DOFs (degrees of freedom) and both robots emphasized the eye motions and the speaking motions. In addition, the author also co-operated with the KOKORO Company[16] to build a series of female androids. These robots were designed to imitate the movements of the human body and to perform them naturally. They can also reveal unconscious human movements just like a person. The robots used motors as the actuators in their system so that they can show smiley faces and speaking abilities.

In 2004, Wu *et al*.[17] proposed a combined design of a facial expression driving unit. This design is applied to a portrait robot structure. This kind of portrait robot can demonstrate 8 kinds of facial expressions. In their robot, 14 CPs became 6 driving units after some CP combinations and reductions. Their method effectively reduced the number of actuators.

In 2005, the Korea Advanced Institute of Science and Technology (KAIST) developed Albert HUBO.[18] This humanoid is equipped with 31 servo motors inside its head to generate various facial expressions. There are 28 servos for his facial movements and 3 servos for his neck movements. The head of Albert HUBO was made by HANSON ROBOTICS.[19] This company has developed many different kinds of face robots and these robots inherited similar facial expression mechanism designs. The Korea Institute of Industrial Technology (KITECH) published a couple of papers about the EveR humanoid series.[20] In these papers, the robot with 22 DOFs on its face is described. The EveR series can generate 6 facial expressions by servo motors under its face. Moreover, in 2005, Beans and Braun[21,22] developed the ROMAN robot which could show at least 6 facial expressions. They used 17 motors (6 for the eyes) to produce the facial expressions.

In 2006, Hashimoto *et al*.[23] published a receptionist robot called SAYA. The head of SAYA was built using Mckibben pneumatic actuators as its artificial muscles. The author arranged 18 CPs on the underside of the facial skin. The combinations of these CPs could generate 6 typical facial expressions. In the same year, Minoru *et al*.[24] tried to mimic the human muscular structure to generate facial expressions. In his study, he designed some muscle groups to connect the artificial skin and the skeleton, and controlled the shrinkage of the muscle groups to generate facial expressions. The authors declared that their robot could generate 6 typical facial expressions by using a motor device to pull 11 muscle groups in the face, but only 2 facial expressions were shown in the paper.

In 2008, The Japanese National Institute of Advanced Industrial Science and Technology (AIST) announced the HRP4 C humanoid.[25,26] Her head was co-developed by the KOKORO Company. This kind of robotic head drives its facial linkage mechanism using motors. Eight motors are used to generate several facial expressions and to enable her to speak. Four facial expressions and some mouth motions were introduced in the paper published by this organization. In 2009, Allison *et al*.[27] proposed a design of facial muscle actuation system based on nodal dependency to generate facial expressions. Their robot could generate 6 typical facial expressions by using 4 motors to pull 4 muscle groups on the face. Their method effectively reduced the number of actuators.

Our laboratory has been devoted to studying face robot since 2006 and has developed some face robots for entertainment performances.[28–30] Servo motors are used to produce the motions of faces in each model. Taking the last model "Eva" as an example, we set a total of 23 servo motors inside: 4 motors for the neck, 4 motors for the eyes and 15 motors for facial expressions. Our robot head can generate 6 typical facial expressions, can grimace and speak.

After summarizing the literature reviewed above, we find that the majority of researchers used motors as actuators to drive their face robots because motors have a high response speed, and that they are powered by electricity, etc. No additional auxiliary machinery such as an air compressor is required in the mechatronic system. Therefore, our face robots use motors as the main actuators. The above approaches are summarized in Table I in which the amount of actuators installed in different robot head can be seen.

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press

Table I. A CP/actuator quantitative comparison between previous approaches.

Year	Institution / Reference No.	Model Picture	Control Point (CP) Distribution	CP/ Actuator Quantity	Year	Institution / Reference No.	Model Picture	Control Point (CP) Distribution	CP/ Actuator Quantity
1996	Tokyo University of Science and Technology / [7]			18 / 24	2006	Korea Advanced Institute of Science and Technology / [18]			25 / 28
2000	Tokyo University of Science and Technology / [8]			18 / 21	2006	Shinshu University / [24]			22 / 22
2002	University of Texas at Dallas / [12]			25 / 28	2006	Taiwan University of Science and Technology / [28]			19 / 23
2003	Tokyo University of Science and Technology / [10]			18 / 21	2008	Taiwan University of Science and Technology / [29]			16 / 24
2004	Harbin Institute of Technology / [17]			14 / 14	2008	Korea Institute of Industrial Technology / [20]			14 / 26
2006	University of Kaiserslautern / [22]			8 / 11	2008	State University of New York / [27]			8 / 14
2006	Tokyo University of Science and Technology / [23]			18 / 21	2009	Japan National Institute of Advanced Industrial Science and Technology / [26]			5 / 8

According to human anatomy and physiology, the facial expression muscle system is composed of 44 basic sub-units called action units (AUs) [31,32] and each AU is independent. In some face robot studies, 14 is the minimum number of AU combinations for 6 typical facial expressions. It means that if more AUs are used, then more facial expressions can be shown. Face robots can be classified into two major types based on their respective DOFs. The first type has various facial expressions and higher DOFs, and the second has finite facial expressions and fewer DOFs. The type with fewer DOFs costs less than those with the higher one so that most commercial face robot products are designed with lower DOFs.

Actuators always make up the majority of hardware cost in robotic products. If the same facial expression capability can be achieved by using the minimum amount of actuators, it will have a beneficial contribution to the mass production of face robots. Therefore, we combined the advantages of these two types of face robots to develop a low cost face robot with fewer DOFs, while also having a higher variation of facial expressions. Our research makes it possible to develop and popularize anthropomorphic face robots. The major difference between this study and other similar researches such as the face robots developed by Wu *et al.* and Allison *et al.* is the expandability of the kinds of facial expressions without increasing the actuators. Section 3.2 will clearly explain this novel difference in detail.

2. Facial Expression Analysis

2.1. Movement classification on a human face

Figure 1 shows the biological structure of a human face. Referring to anatomy,[33] facial expressions are made through the contraction of many muscles. However, not every facial muscular expression is independent. Some can be easily classified in the same group in terms of their expressional functions

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press

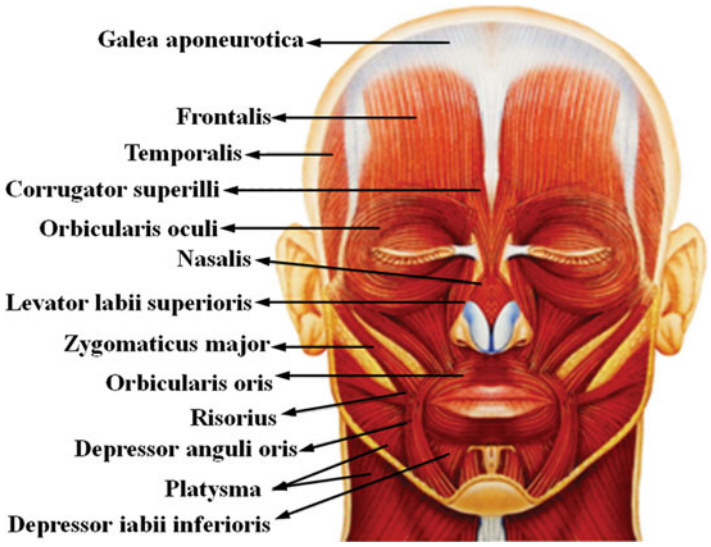


Fig. 1. Dissection chart of facial expressional muscles.[33]

and interlocking relations, but the others cannot. If we ignore all the insignificant relationships between muscles, then we can divide facial movements into three major groups. The first group includes the *frontalis, orbicularis oculi*, *corrugator supercilli*, *levator labii superioris* and some other motion-related muscles. The movement region produced by this group covers the surrounding areas of the eyebrows, eyes and nose. The second group includes muscles distributed around the eyeballs. This group of muscles is responsible for the rotations of the eyeballs and can be treated as an independent group. The third group includes *orbicularis oris, risorius* and the muscles distributed around the mouth. This group is in charge of the motions near the lips.

To assess people's feeling and the influence of facial expressions on the acceptability for polymeric skinned robot head, we conducted an online questionnaire survey on internet. Provided with some brief explanation of the face robot, one of our questions is "Do you think the face robot should be given the ability to show their facial expressions?" The questionnaire collection time is 1 week, and the number of effective samples is 85. The age range of people who replied these samples is from 5 to 64 years old, including 31 female samples and 53 male samples. The result of this survey is shown in Fig. 2. From this result, we know more than half of people (58%) agree that face robot can show some facial expressions. Furthermore, Cheng[34] had conducted a related evaluation about what reactions most people will have when they face a skin-deformable face robot. The result he got is shown in Fig. 3. It illustrates that skin-deformation-type face robot brings uncanny feeling to some people, but it is limited to minority of people. These results support that developing skin-deformation-type face robot will not be repelled by most users at the current stage, not to mention that the realistic level of the face robot will keep improving along the line.

Using the concept of AU, we can also classify the AUs on a human face into three major groups with respect to their corresponding muscle groups, as shown in Fig. 4. These three groups do not interfere with each other even if they are acting at the same time. In other words, the eye, eyebrow and mouth movements are all independent. For example, a human can blink his/her eyes and close his/her mouth at the same time.

In order to reduce the gap between humans in face-to-face communications, most people communicate with joyful faces. Based on this reason, most designers let their android robots have a smiley face as their default facial expression. Other facial expressions will be presented in accordance with the different scenarios that follow. Actually, facial expressions which are a response to different scenarios by a human often last for only a few seconds. The time for every expressional response is short. We assume that this situation is true in robot applications as well. According to psychologist Paul Ekman's definition,[32] the AU for eye blinking is independent from other AUs. By observing human faces, we discovered that the eyebrows and forehead muscle groups are motionless and that the eyelids muscle groups are in motion when people smile. Obvious movements occur intensively

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press

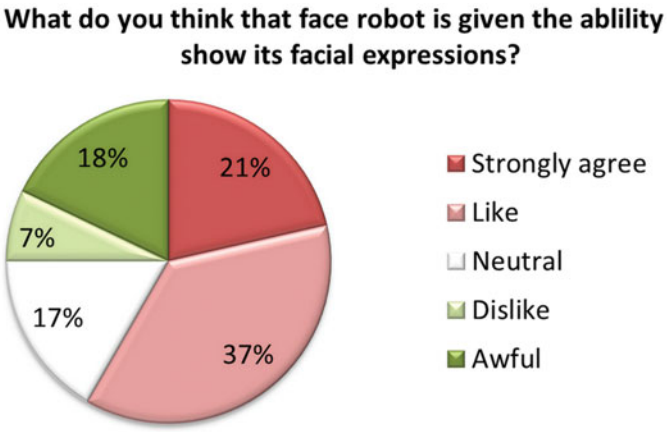


Fig. 2. Opinions on face robot with facial expressions.



Fig. 3. Opinions on robot head with soft polymeric facial skin.

around the mouth in a smiley face. The other typical facial expressions not only relate with the muscle group around the mouth, but also involve the eyebrows and the forehead muscle groups. Moreover, if a frown or a raise of one eyebrow is just momentary, eye blinking never happens synchronously. In other words, the eye blinking motion is only produced for the smiley face, and the eyebrow and forehead motions are produced for other typical facial expressions. Therefore, it is necessary for us to design a mechanism to take charge of eyebrow motions and eyelid motions, and to perform them alternately at a time to achieve the goal of mechanical simplification.

2.2. Movement simplification on the robotic face

In this research, referring to the related literatures[1–27] and our face robot development experiences,[28–30] we chose fourteen positions on the facial skin of the face robot as CPs. The combination of different groups of CPs can generate six typical facial expressions. The allocation of these CPs is also shown in Fig. 4. The directions of the arrows represent the movable directions of the CPs. The number next to each arrowhead indicates the serial number for each CP. In the last section, we specifically divided the human face into three major areas. We can sort fourteen CPs into three major groups by those areas, too.

For facial expression mechanism simplification, we must analyze whether the facial movements can be combined, reduced or omitted. First, we have to know how these facial movements work. Considering the movements of the eyes, each eye has two DOFs for eyeball rotation. To most healthy people, their double eye motions are always symmetrical. Unless they rotate each of their eyes in different directions deliberately, the directions of eye motions are almost the same. Therefore, we can reduce the eye DOFs from 4 to 2. In this research, we set both eye staring directions to be the same at every single moment.

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press

Table II. Control areas of a face robot.

No.	Facial Expressions	Control Unit (CU)	
		Group1	Group3
1.	Surprised	A(1, 2), B(3, 4)	
2.	Happy		K(11), L(12)
3.	Angry	G(5), H(6)	
4.	Sad	C(2, 3)	M(13), N(14)
5.	Disgusted	D(9, 10)	
6.	Natural		
7.	Mouth Shape		E(15, 16), F(17, 18)
8.	Wink	I(7), J(8)	

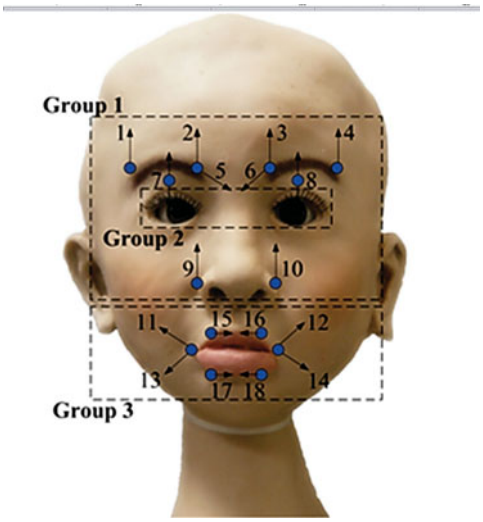


Fig. 4. Group classification of facial expressional muscles.

The facial expressions of the face robot are composed of different combinations of control units (CUs) in Group 1 and Group 3. In this research, we defined a CU based on the CPs and their corresponding pulling actuators. The concept of the CUs is similar to AU, but physically different because the elements in the AU system physically mean muscles, but the elements in our CU system physically mean actuators. A single or a group of CPs can be pulled by one actuator only so that they can be treated as an individual CU. In Table II, CU A or CU B includes two CPs at both ends of each eyebrow. Activating one of them can raise the eyebrow where its CU is pulled. CU C includes two CPs at the eyebrow ends which are near the bridge of the nose. While CU C is being activated, two eyebrows will form an inverted "V" shape. CU D includes two CPs located on both sides of the nose. Activating this CU can lift the nose and produce some winkles around the saddle of the nose. CU E represents the upper lip with direction 15 and direction 16. CU F represents the lower lip with direction 17 and direction 18. CU G and CU H include the CPs at the eyebrow ends which are near the bridge of the nose. Pulling these two CPs to direction 5 and direction 6 can generate a frown. In this research, there are four CPs which can be moved in two different directions. CU A and CU G, or CU B and CU H share the same CP at the inside end of the eyebrow. CU K and CU M, or CU L and CU N share the same CP at the corner of the mouth. When these CPs are pulled in on one direction, the wire connected to the other must be released. The stroke of every wire should be reserved a section of distance for the other control-point-shared-direction pulling. CU I or CU J can control the CP at the upper eyelid to move up and down for showing winking. For example, a sad face is composed of CU C in Group 1 and CU M and CU N in Group 3. CU C means direction 2 and direction 3 are pulled by the same actuator. CU M or CU N means it is pulled by a particular actuator. Furthermore, if some CPs have the same direction of motion or an approximate displacement in the same facial expression, we can combine them into the same CU such as A, B, D, E, and F.

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press

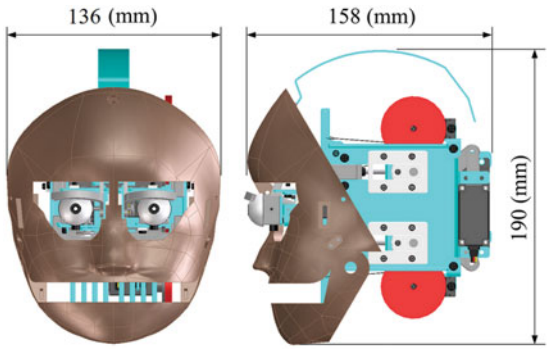


Fig. 5. External dimensions of the simplified mechanism.

Since CUs can be decided by their corresponding actuators, we can also consider how to design a mechanism for each independent group to supply a similar function to replace the function performed by the output performance of several actuators. We can use the following equations to describe our simplification idea but excluding eyeball motions:

$$\begin{aligned} E_i &= M^{\mathrm{i}}_{\mathrm{top}} + M^{\mathrm{i}}_{\mathrm{down}} \\ M^{\mathrm{i}}_{\mathrm{top}} &= R^{\mathrm{i}}_{\mathrm{top}} + T^{\mathrm{i}}_{\mathrm{top}}, \quad i = 1 \sim 8 \\ M^{\mathrm{i}}_{\mathrm{down}} &= R^{\mathrm{i}}_{\mathrm{down}} + T^{\mathrm{i}}_{\mathrm{down}} \end{aligned} \tag{1}$$

Symbol E represents the facial expressions of a face robot, and index i stands for the type of facial expression. The eight types of facial expressions constitute the specifications of our systems capability. Variable $M^{\mathrm{i}}_{\mathrm{top}}$ and $M^{\mathrm{i}}_{\mathrm{down}}$ represent the expressional motions of the upper face and the lower face in the ith facial expression. The expressional motion of each facial region can be designed as a combination of a rotation motion (Symbol R) and a translation motion (Symbol T). The rotation motion is for the rotary facial expression selection. The translation motion is for pulling the CUs. Therefore, 2 DOFs are required for each half face. The total DOFs of facial expressions are 4. In other words, only 4 servo motors are required in our design. The concept is just like in a music box structure, every row represents a set of CU displacement settings. There is a row of 8 holes on the roller surface at every 45 degrees. The total number of rows is 8. Translation motion supplies the power to pull the wires of selected CUs. The detailed mechanism is explained in more detail in the next section.

3. Mechanism Design of the Face Robot

Following the concept described in the last section, we designed a kind of mechanism for face robot simplification. Its structural appearance and main external dimensions are shown in Fig. 5. The shape of the head was copied from that of a ten-year-old child. We used a 3-D scanner to scan that child's head and used the facial profile data for the computer-aided-design system to reconstruct a 3-D virtual geometric head model. The reason for choosing a child's head as a mimic object to build the face robot is that we wanted the whole simplified mechanism to be as small as possible so that this mechanism can be applied to all sizes of the human face. Fig. 6 shows the main body of the simplified expressional mechanism. In this section, we divided the topic into two parts to be explained separately. The first part is the eye module, and the second part is the facial expression module, which is the main point in this study.

3.1. Eye module

Continuing the idea of eye design in the previous section, we must reduce the eye DOFs from 4 to 2. To satisfy this idea of simplification, a simplified eyes module solution has been figured out, as shown in Fig. 7. This eye module has two motors installed: $\mathrm{Motor}_{\mathrm{ER}}$ and $\mathrm{Motor}_{\mathrm{ES}}$. $\mathrm{Motor}_{\mathrm{ES}}$ is for driving two eyeballs to move up or down, and Motor_{ER} is for rotating the eyeballs to move left or right. This design enables both eyeballs to move at the same time. Each eyeball is equipped with a

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press

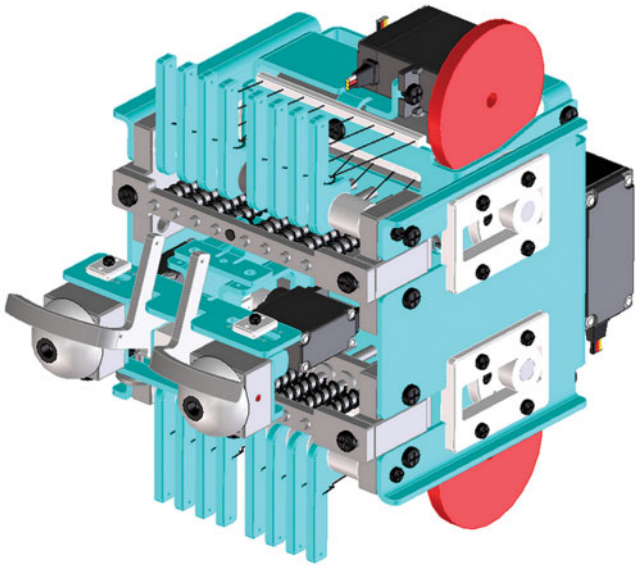


Fig. 6. Main body of the simplified expression mechanism.

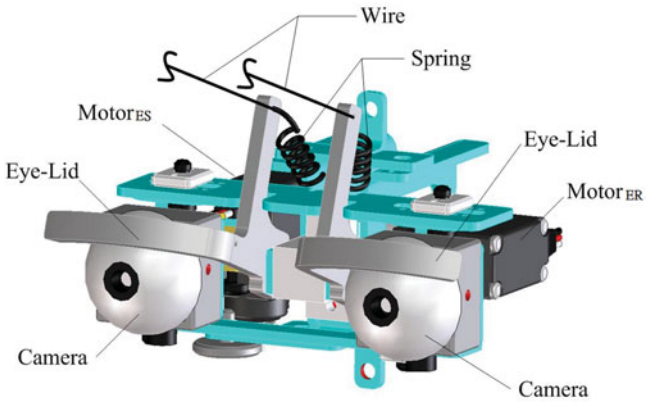


Fig. 7. Eye module.

pinhole camera inside for environmental image recognition. The eyeball mechanism is also a kind of DOF reduction design compared with human eyes' DOFs, but the eyeballs are not installed and included in our simplified expressional mechanism. The reason why the eyeballs are put in Figs. 5–7 is just for the structural completeness and for showing the aspects of the mechanisms. One more degree of freedom is added for each eyelid. To sweep above the eyeball surface by rotating, each part of the eyelid is connected with a steel wire to the front and with a spring to the back at the same end hole of its stick-like protrusive portion. Each of those steel wires goes to the front through one small hole of the skull structure and along a short distance on the skull surface to the top. Then, it goes continuously to the back through another small hole on the skull surface and is finally connected with the simplified expressional mechanism and shares the actuators installed in the mechanism. There is a pin through another hole on the tail of each part of the eyelid and is fixed on the eye module frame to be a rotary shaft for eyelid rotation and sweeping. If the steel wire is pulled tightly, that part of eyelid rotates forward (counterclockwise). It causes this eye to close. In contrast, if the pulling force is removed and the spring is released, that part of eyelid rotates backward (clockwise). It causes this eye to open.

N. A. Dodgson[35] pointed out that the mean inter-pupillary distances of the vast majority of adults and children in the world is around the range of 40–75 mm. By referring to this range, we made the inter-pupillary distance of our face robot to be adjustable. The main purpose behind this kind

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press

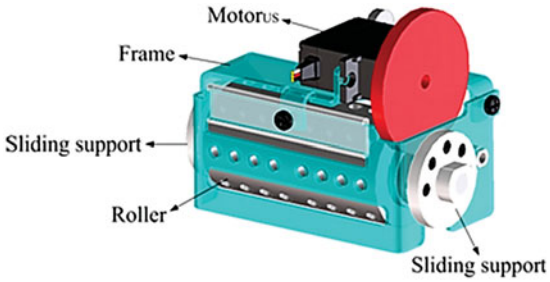


Fig. 8. Selection module of facial expressions.

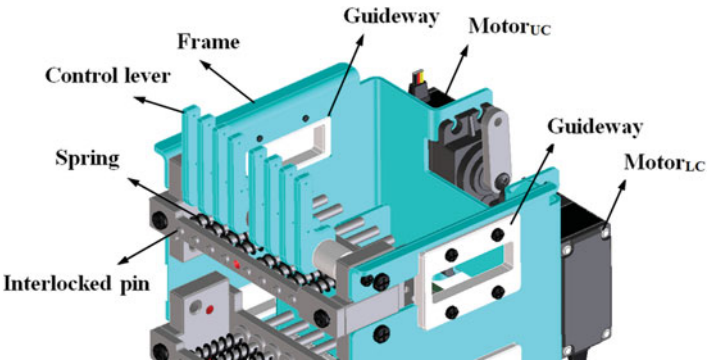


Fig. 9. Clutch module.

of design was to allow the application of this eye module to any human face even with a different inter-pupillary distance.

3.2. Modularization of facial expressions

Our simplified face robot has a total of two facial expression modules. These two modules take control of the upper facial movements and the lower facial movements, respectively. They are copies of the same design to reduce the cost. Each of them is composed of two sub-assemblies which are the facial expression selection module and the clutch module. Both of these sub-modules are described as follows.

Fig. 8 shows the selection module of facial expressions. This module is for selecting the facial expression which the face robot is going to manifest. The selection module is assembled with an external frame, a servo motor, a roller and some auxiliary parts. Its sliding support parts on either side are supported by two guide ways of the clutch module. According to Cheng *et al.*'s experimental result on human-face-deformation measuring,[36] the maximum displacement tangent to the facial surface is not over 15 mm in some typical facial expressions compared with a neutral face. It means the maximum stroke of the wire will not exceed 15 mm in most cases. Therefore, the maximum depth of the hole on the selection roller can be set to be slightly greater than 15 mm because of considering the assembly errors and the elongation of an elastic steel wire. If a powerful motor is selected for Motor UC or Motor LC in the clutch module, as shown in Fig. 9, the effect at every CP basically equals to that made by the single-motor-per-DOF mechanism. The diameter of the selection roller with that kind of the depth of the hole fits the space in the life-scale skull.. This design makes the whole selection module move along a straight line back and forth. The principle of the facial expression selection is shown in Fig. 10. Similar to a music box structure, a roller with rows of holes can be rotated to the prearranged positions by controlling the rotation angle of the motor. We designed 8 rows of holes on the roller with equal angular space to each other. Each row of the upper selection module represents the CU combination in Group 1 and each row of the lower selection module represents the CU combination in Group 3. As shown in Table II, we allocated 5 rows for upper facial expressions and 1 row for eyelid motions. Two rows remain for the expansion of upper facial expressions. We allocated 2 rows for lower facial expressions and 1 row for lip motion. Five rows remain for the expansion of lip motions. There are 8 holes in each row. The maximum number

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press

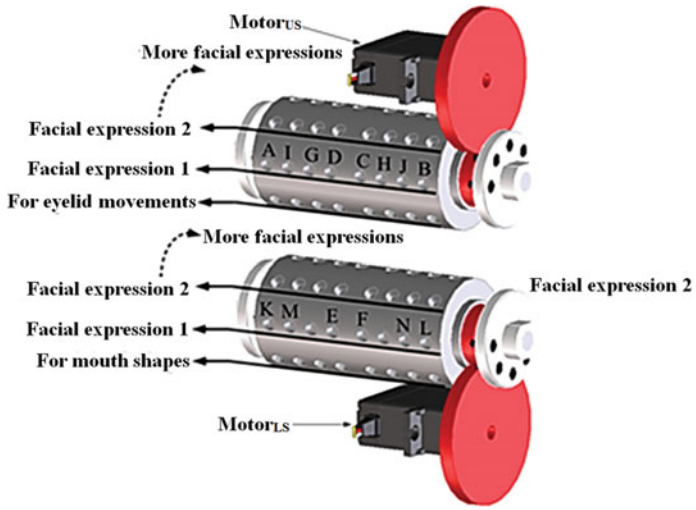


Fig. 10. Principle of facial expression selection.

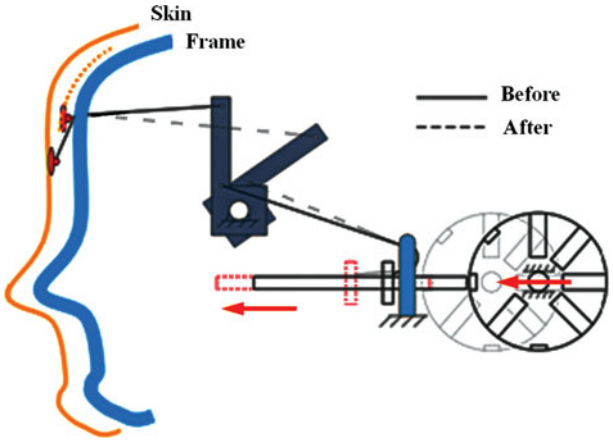


Fig. 11. Generation mechanism of facial expressions.

of holes used is 8. The maximum number of holes is in the upper selection module because 8 CPs must be pulled at the same time on the upper face for some facial expressions. Therefore, we set 8 holes in a row with each hole having a different depth. Every depth setting of a hole relates to the stroke depth setting of its corresponding control lever in the clutch module. The clutch module is for moving the selection module back and forth to push or release the interlocked pins. Every interlocked pin is placed on a spring to apply a release force. The clutch module is composed of a frame, a motor, a couple of guide ways, 8 interlocking pins, 8 control levers, 8 springs and some auxiliary parts. One motor is fixed on the side of a frame and is used to apply force to push and pull the selection module back and forth along the direction of the guide way. The other side of the frame is assembled with all interlocking pins and control levers.

The stroke depth of each control lever determines the pulling distance of the corresponding CP which is connected to this control lever by a steel wire. When a row of holes with different depths in the upper/lower module is inserted pushing the interlocking pins, their corresponding control levers are forced to rotate and pull the steel wires. CPs belonging to the same set is pulled at the same time as shown in Fig. 11. In the meantime, a selected upper/lower facial expression is generated. Every row of holes on the roller represents a pulling combination of CPs. The combination of the respective effects from the two rows in the upper and lower modules represents a complete specific facial expression. In other words, the depth of each hole is related to the displacement of each facial CP on the reverse side of the artificial facial skin. Therefore, we can just adjust or modify the depth of each hole to change the displacement combination of CPs in every single facial expression. If more

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press

facial expressions need to be designed in this robot, only a small modification is required. Changing the number of the holing rows on the facial expression selection roller is the faster and cheaper way to achieve this. The addition of new CP combinations will not affect other existing ones. Increasing the number of actuators is not necessary. In Wu, *et al.*'s study, 6 actuators govern 6 different groups of CPs. If their robot needs to be added with more facial expressions where their CPs will require pulling actions from other different directions for some special reasons, some additional actuators will be needed. A similar situation which requires additional actuations happens in Allison *et al.*'s cases.

We can clearly summarize all the control processes of facial expression generation in sequence as follows:

- First, drive the motors in the upper and lower selection modules to rotate the rollers to select the upper and lower facial expression rows;
- Second, drive the motors in the upper and lower clutch modules to push the upper and lower selection modules forward until all interlocking pins are inserted into the selected rows of roller holes and are pushed forward. This will cause the control levers to rotate in different angles, pulling the CPs on the reverse of the robotic facial skin to move;
- Third, drive the motors in the upper and lower clutch modules to release the upper and lower selection modules backward to return to the default facial expression.

3.3. Artificial facial skin

To make a smaller simplified mechanism, the face of our robot is copied from a child's face. The photos in different stages of the procedure of making the artificial facial skin in the shape of real facial skin are shown in Fig. 12 and the detailed sequence is stated as follows:

- First, use alginate to make the gypsum mold to reproduce the facial shape;
- Second, use this gypsum mold to generate the silicon mold and clay model to manufacture the external and internal mold for skin material injection;
- Third, inject the urethane rubber into the mold;
- Fourth, open the mold and take out the skin after the material solidifies;
- Fifth, the back side of the artificial skin is sewn on some metallic rings that are connected with steel wires. These metallic rings are arranged based on the results of the facial expression analysis. Some Velcro® tapes are also set on the back of the artificial facial skin to fix the outer surface of the skull and the inner surface of the immovable skin area together.

3.4. System configuration

The main objective of this research is to simplify the complexity of the mechanical design, to save on costs and to show finite, but life-like expression performances. Instead of pursuing perfection, the design goal was set to be cheap and effective. Therefore, the design strategy tends to make things simpler, but still acceptable. Fig. 13 shows the system configuration of our simplified face robot. According to the design concept, we developed the robot with a sense of vision only and neglected the sense of hearing, smell, taste and touch. So, our face robot is simply equipped with two pinhole cameras in its eyeballs. We wanted this face robot to have some vocal communication ability so there was also a speaker installed inside. For facial expression generation, there are 7 servo motors (4 for facial expressions movements, 1 for jaw motions, 2 for eye motions). All these components are governed by a computer through an RS232 interface. A high performance processor in the computer is necessary because the image processing and object recognition functions require it. Moreover, a camera controller is required because the volume of the pinhole camera casing is so small that some control circuits must be connected externally. Also, the 8-PWM-channel motor controller is installed for servo motor driving. Thus, the mechatronic system of our face robot is quite simple and easy to maintain.

4. Performance Evaluation

Theoretically, if one wants a face robot to generate a realistic facial expression, it is better to have the same number of motors as the number of facial expression muscles. We tried to come up with a design with only 4 motors excluding the eyeball motors and jaw motors to produce several typical

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press

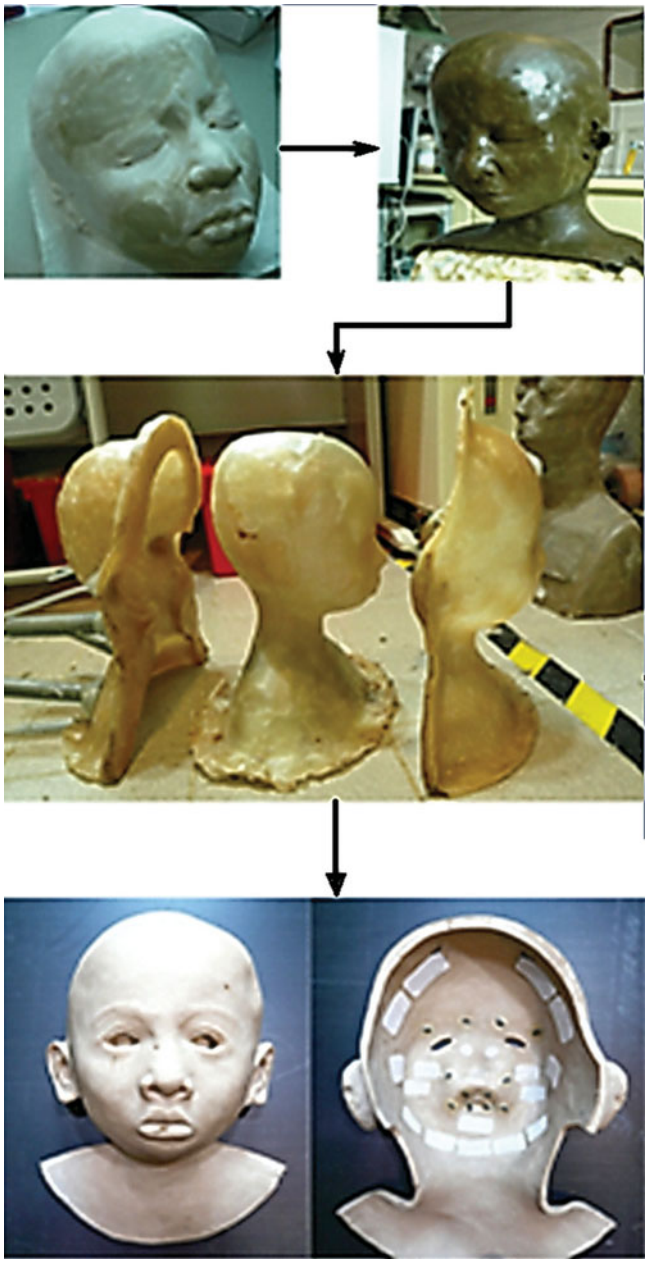


Fig. 12. Procedure for making artificial facial skin.

facial expressions whose genuineness is close to the manifestations on a human face. However, it is really difficult to achieve the effects produced by a large number of motors. In terms of our objectives, we must define a standard of evaluation to check whether or not our design is good. In this section, we evaluated our face robot performances based on three important aspects: correct recognition rate of facial expressions, deformation capability and power consumption.

4.1. Correct recognition rate

We randomly interviewed 100 college students on our campus as the subjects to recognize 5 typical facial expressions excluding the neutral face, shown in Fig. 14. None of the students have been trained in the recognition of facial expressions before. There are two parts in our recognition experiment. In the first part, we prepared five still color pictures and showed them to each subject, and then we asked each person what expression was made by the face robot in every picture. In the second part, we recorded a 1 min long video containing these five facial expressions in sequence and displayed it to our

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press

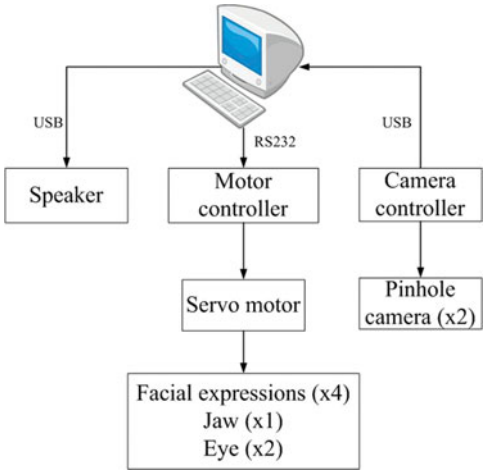


Fig. 13. System Configuration.



Fig. 14. Six typical facial expressions.

subjects. Then, we asked the same questions of our subjects and took down their answers. We wanted to undertake a performance evaluation by watching static and dynamic images to obtain the results more objectively. Table III shows the statistical results of the static facial expression recognition rate. We found that the sad and surprised face can be recognized correctly by these 100 participants and the angry and disgusted face also have a good recognition rate, but the happy face has just a 54% correct recognition rate. Table IV shows the statistical results of the dynamic facial expression recognition rate. We found that the sad and surprised faces still have absolutely high correct recognition rates. The correct recognition rate of the happy face went up to 67%, and the correct recognition rate of the angry and disgusted faces also went up to 80% and 75%, respectively. These results mean that a dynamic facial expression generation process makes it easy for people to tell the differences between two facial states. It also affirms our face robot's ability to generate acceptable, non-confusing, and life-like facial expressions by a simpler solution in reality.

4.2. Facial deformation ability

The purpose of evaluating the robot's facial deformation ability is to make sure that our simplified face robot (called Simple type, S-type) has the ability to make similar facial deformation as the face robot without any mechanical simplifications (called Complex type, C-type). Because we have had some experiences in designing face robots before, we defined the C-type face robot as a form of our

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press

Table III. Static recognition rate of facial expressions.

Picture (Static Image)		Recognition results (%)				
		Sa.	Su.	Ha.	An.	Di.
Facial expressions	Sad (Sa.)	**83%**	1%	1%	2%	13%
	Surprised (Su.)		**99%**		1%	
	Happy (Ha.)	22%	4%	**54%**	2%	18%
	Angry (An.)	9%	1%	1%	**73%**	16%
	Disgusted (Di.)	1%	1%		24%	**74%**

Table IV. Dynamic recognition rate of facial expressions.

Video (Dynamic Image)		Recognition results (%)				
	Sa.	Su.	Ha.	An.	Di.	
Facial expressions	Sad (Sa.)	**82%**	1%	6%	3%	8%
	Surprised (Su.)		**99%**			1%
	Happy (Ha.)	23%	1%	**67%**	5%	4%
	Angry (An.)	5%			**80%**	15%
	Disgusted (Di.)	1%	1%		23%	**75%**

previous face robot models.[26–28] If both of their facial deformation abilities were similar, it means that our efforts on mechanical simplification have been successful. We measured and compared the displacements of key control points (KCPs) to understand how the face of each type of robot deforms for every facial expression and what differences lie between the S-type and C-type face robots. The distribution of the KCPs is shown in Fig. 15. First, we used the same child face robot mentioned previously and the robot's facial expression mechanism was assembled in the S-type formation. In front of the robot, there is a transparent plastic plate that was installed vertically with a graph-paper-like black mesh drawn on it, with each grid size being $1\times$ 1 mm. Furthermore, there is also a high resolution camera in front of the plastic plate. Through the mesh on the transparent plastic plate, we took pictures to observe and measure the position of all KCPs. After taking all the pictures required for the S-type, we assembled the facial expression mechanism of the child robot in the C-type form and had it assembled the same way as our previous face robot models. The C-type requires one or two servo motors for each KCP so that there was a total of 13 motors installed in the C-type face robot. The plastic plate was set as close to the robot as possible. Since we just wanted to find out the related position variation, we dismissed the error of the projection effect and took down the X-Y positions of all KCPs by checking these pictures. Table V lists all the displacement differences of all KCPs in the X and Y directions under the five typical facial expressions. We found that there was almost no difference between the sad and disgusted faces, and that there were only slight differences between the surprised and happy faces. The biggest difference was equal to or lesser than one grid size scale ($\leqq$1 mm, after dismissing the errors). It was quite small compared to the whole face. Thus, we can conclude that the facial expression performance of our simplified face robot is similar to that of the face robot without any mechanical simplifications. In other words, our 4-motor design produced similar facial expression effects as the 9-motor design.

4.3. Power consumption

We are also concerned with the power consumption of the two types. Theoretically and intuitively, the power consumption of 5 motors is absolutely less than the power consumption of 13 motors when using the same type of motors. How much power wattage can we save from our simplified face robot? Is the power saving rate of the different facial expressions equal? We are only interested in these questions. Therefore, we designed an experiment to measure the power consumption of these two kinds of robots under different status of facial expressions. With every facial expression, we controlled the robot to do the same facial expression 5 times and we measured the power consumption each time with a power meter, and calculated the average. Table VI shows the average power consumption for

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press

Table V. Displacement difference of each KCP (unit: mm).

	Facial expressions				
Measuring direction of KCP	Sad	Surprised	Happy	Angry	Disgusted
1	–	0	–	0.5	–
2-X	–	0	–	–	–
2-Y	0	0	–	1	–
3-X	–	0	–	–	–
3-Y	0	1	–	–	–
4	–	0	–	−0.5	–
5-X	–	–	–	−0.5	–
5-Y	–	–	–	0	0
6-X	–	–	–	−0.5	–
6-Y	–	–	–	0	0
7-X	0	0	0	–	–
7-Y	0	0	0	–	–
8	–	0.5	0	1	–
9-X	0	0	−0.5	–	–
9-Y	0	0	1	–	–

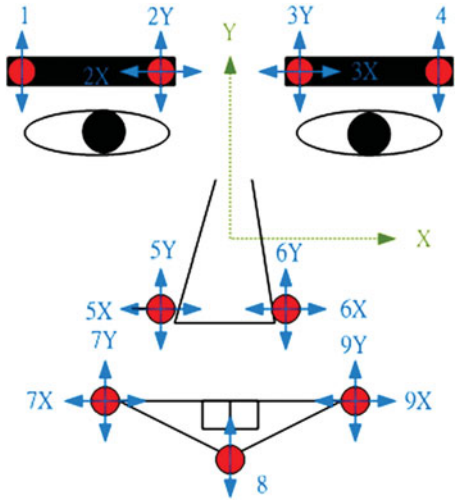


Fig. 15. Key control points for displacement check.

every facial expression made by the S-type face robot. Table VII lists the average power consumption comparison between the S-type and the C-type face robots under the five typical states of facial expressions. From these results, we found that the rate of power saved in every facial expression was not the same. It meant that having different numbers of working CPs between two different facial expressions also affected the results.

5. Discussions

We found that three facial expressions: anger, surprise and sadness had high recognition rates (more than 80%) by the subjects in our facial expression recognition experiment. The other two facial expressions had lower recognition rates. Due to our disposition of the mechanism, the deformation around the eyebrows is more obvious than the deformation around the lips. Therefore, the facial expressions involving the eyebrow motion can get a higher correct recognition rate. The subjects found it difficult to recognize the facial expressions when the happy or disgusted face was being generated. Moreover, in another experiment, we provided static pictures and dynamic videos to the subjects and had them recognize the facial expressions. The results showed that in order to raise the correct recognition rates for the happy facial expressions, dynamic images had to be presented to

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press

Table VI. Average power consumption of the simplified face robot.

Facial Expressions	Average motor power (W)					Total Power (W)
	M_{US}	M_{UC}	M_{LS}	M_{LC}	M_{JAW}	
Surprised (Su.)	0.022	9.735	0.022	0.990	2.200	12.969
Happy (Ha.)	0.022	0.357	0.022	2.970	0.022	3.393
Angry (An.)	0.022	4.565	0.022	0.022	0.022	4.653
Sad (Sa.)	0.022	6.600	0.022	1.870	0.022	8.536
Disgusted (Di.)	0.022	4.510	0.022	0.022	0.022	4.598

M_{US} = Motor in upper selection module, M_{UC} = Motor in upper clutch module.
M_{LS} = Motor in lower selection module, M_{LC} = Motor in lower clutch module.
M_{JAW} = Motor for jaw driving.

Table VII. Average power consumption comparison between the C-type and the S-type.

Facial expressions		Su.	Ha.	An.	Sa.	Di.
Power (W)	C-type face robot	14.50	6.30	12.30	12.30	6.30
	S-type face robot	12.97	3.39	4.65	8.54	4.60

the subjects. The reason why the correct recognition rate of a facial expression is raised by showing dynamic images to the subjects is that the subjects can observe every step in the process of a facial expression variation. For instance, even if the deformation around the lips is not obvious in the happy face, the subjects can still automatically absorb the information of the deformation such as the surface deformation direction and the changes of the surface shape by comparing the slight differences in the videos between the frame of the present moment and the frame following it. Thus, by displaying dynamic images, the correct recognition rate of the happy face goes up to 13%, individually. Also, there is another factor that influences the correct recognition rate of the facial expressions. Dailey *et al.*[37] and Aviezer *et al.*[38] indicated that not everyone can differentiate an angry face from a disgusted face very clearly. Taking the disgusted face as an example, there were 23% of people who thought that the face was an angry face in our experiment. Moreover, Jack *et al.*[39] pointed out that between Orientals and Westerners, the methods used to distinguish the facial expressions are quite different. Orientals emphasize and recognize the variations around the eyes and look at the mouth when they want to interpret someone's facial expression. Our subjects were all Orientals and that is why the facial expressions involving the eyebrow motions were recognized correctly and quickly. It is possible that assembly errors make the displacements of the CPs unable to achieve the design value and consequently make the deformation around the lips become not as obvious. Because the medium of transmission is steel wire, the robotic facial skin is an elastomer, and the steel wire is connected directly to the robotic facial skin, it is difficult to assemble the lengths of steel wires as our design values. This kind of error is an urgent issue of further improvement for us.

To develop a simplified mechanism to simulate a complex system in reality must take some basic simplification assumptions to keep some major and key feature variables and at the same time ignore other minor and unknown variables in this system. In facial expression analysis research field, many studies have based on the simplification assumption which Kanade *et al.*,[40] Cohen *et al.*,[41] and Pantic and Patras[42] have mentioned about. This simplification assumption is "expressions are singular and begin and end from a neutral position." Thus, the transition problem in this simplified mechanism design is supposed to be regarded as a facial expression analysis problem, and we adopt the same simplification assumption. The transition between two different facial expressions always goes through the neutral facial expression state.

In our experiment of measuring the displacement of facial expressions, we know that the facial deformation abilities of the S-type and C-type are extremely close. The maximum displacement difference in the same position CP between the S-type and C-type is only about 1 mm. Other than the angry expression, the CP displacement differences all have a positive value. It shows that the S-type face robot can produce a larger displacement with each CP. The CP displacement differences between

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press

these two types are all very small. This shows that S-type face robots have similar facial deformation abilities as the C-type face robots. It also shows that our design idea is effective and successful.

In our power consumption experiment, the surprised face consumed more power than the other facial expressions produced by the S-type face. This is due to the fact that the jaw must be rotated with a larger angle in the surprised expression. The mechanism responsible for rotating the chin is supposed to withstand a load that equals the weight of the whole chin, including the weight of the mechanical parts and the skin. Thus, we designed the motor for jaw driving with the specification of a higher power output. For generating a sad face, the upper and lower expression modules should be driven at the same time. As a result, the power consumption of the sad face is also greater. The results tell us that the S-type face robots can save more power. The power saving means that the S-type face robot can be equipped with a battery of a smaller electrical capacity or that when under the same electrical battery capacity conditions, the working time of the S-type face robot lasts longer than the C-type robots.

Regarding to the cost, comparing with our previous model "Judy"[34] which is a C-type design, the motor cost of this S-type face robot is roughly 65% of that of Judy robot. Although all motors are upgraded with power specification in the S-type face robot, the quantitative advantage on motors (5:17, excluding neck DOFs) plays a decisive role. Wang[43] had investigated that the motor cost is almost 1/3 of the total material cost of an industrial robot. It is a similar ratio that motors in Judy robot share about 25% cost. It means if the 25% 33% cost can be reduced significantly, the total cost of the whole robot can also be reduced accordingly. In this case, 9% cost has been saved due to the motor quantity reduction. The other parts such like structural parts or mechanical parts can be further cut down by using different manufacturing processes or various designs because of the technical level of the design engineer. For this reason, it is hard to discuss this issue in this paper. But, the important thing we know for sure from this case is the total cost has been successfully cut down by the simplification.

Summing up the results from our three evaluation experiments, we may conclude that the direction of our face robot simplification research is on the right track and that our simplified face robot designs are practical and successful. Overall, our simplified face robot is of an acceptable standard.

6. Conclusions

The objectives of this research are not only to reduce the production costs of face robots, but also to present acceptable facial expression manifestations while generating good enough mouth shape variations. We developed a simplified facial expression mechanism in the face robot to achieve these goals. By reducing the number of actuators substantially and modularizing the mechanical structure, the face robot can save on electrical power and effectively reduce the production cost. The results of our evaluative experiments tell us that the facial expressions generated by our simplified face robot are of a good standard. For most of the facial expressions we assigned, their correct recognition rates were quite high. The measurement also shows us that our simplified facial expression mechanism can perform similar facial deformation as the mechanism without any simplification. Thus, we know that this idea of simplification works. However, there are still some issues waiting to be resolved in this field, such as skin-skull attachment problems, reliability problems, optimum design problems of CP positions, and so on.

In the future, we are aiming to simplify the robot assembly process and we are also going to develop a mass production version of this type of simplified face robot. The only way to make robots popular and universal in the world is to make the structure of robots simple and robust, to enhance the functionality of robots and to keep the price of robots at a minimum. We believe that simplifying the face robot design creates the possibility of achieving that goal.

References

1. C. Breazeal, Sociable Machines: Expressive Social Exchange Between Humans and Robots, *Ph.D. Dissertation* (Cambridge, MA: Dept. EECS, MIT, 2000).
2. M. Zecca, S. Roccella, M. C. Carrozza, H. Miwa, K. Itoh, G. Cappiello, J. J. Cabibihan, M. Matsumoto, H. Takanobu, P. Dario and A. Takanishi, "On the development of the emotion expression humanoid robot WE-4RII with RCH-1," *Proceedings of the 4th IEEE-RAS International Conference on Humanoid Robots,* Santa Monica, CA, (2004) vol. **1** pp. 235–252.

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press

3. M. M. Lab, "MDS Head & Face," Retrived March 2, 2014. [Online] http://robotic.media.mit.edu/projects/robots/mds/ headface/headface.html.
4. R. Beira, M. Lopes, M. Praga, J. Santos-Victor, A. Bernardino, G. Metta, F. Becchi and R. Saltaren, "Design of the robot-cub head," *Proceedings of the IEEE International Conference on Robotics and Automation,* Orlando, FL (2006) pp. 94–100.
5. C. Diana and A. L. Thomaz, "The shape of Simon- creative design of a humanoid robot shell" *Proceedings of the* 29th Annual CHI Conference on Human Factors in Computing Systems, Conference Proceedings and Extended Abstracts, Vancouver, Canada (2011) pp. 283–298.
6. I. Lütkebohle, F. Hegel, S. Schulz, M. Hackel, B. Wrede, S. Wachsmuth, and G. Sagerer, "The bielefeld anthropomorphic robot head- Flobi," *Proceedings of the IEEE International Conference on Robotics and Automation*, Anchorage, AK (2010) pp. 3384–3391.
7. H. Kobayashi, and F. Hara, "Study on face robot for active human interface-mechanisms of face robot and expression of 6 basic facial expressions," *Proceedings of the IEEE International Workshop on Robot and Human Communication*, Tokyo (1993) pp. 276–281.
8. H. Kobayashi, "Study on Face Robot Platform as a KANSEI Medium," *Proceedings of the IEEE International Conference on Industrial Electronics*, Nagoya, Japan (2000) pp. 481–486.
9. F. Hara, H. Akazawa and H. Kobayashi, "Realistic Facial Expressions by SMA Driven Face Robot," *Proceedings of the IEEE International Workshop on Robot and Human Communication*, Bordeaux-Paris, France (2001) pp. 504–511.
10. H. Kobayashi, Y. Ichikawa, M. Senda and T. Shiiba, "Realization of Realistic and Rich Facial Expressions by Face Robot," *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Las Vegas, NV (2003) pp. 1123–1128.
11. D. Hanson, G. Pioggia, Y. B. Cohen and D. De Rossi, "Androids: Application of EAP as Artificial Muscles to Entertainment Industry," *Proceedings of the SPIE – The International Society for Optical Engineering*, Newport Beach, CA (2001) pp. 375–379.
12. D. Hanson, G. Pioggia, S. Dinelli, F. Di Francesco, R. Francesconi and D. De Rossi, "Bio-inspired Facial Expression Interfaces for Emotive Robots," *Proceedings of the AAAI Mobile Robot Competition*, (2002) pp. 72–82.
13. T. Minato, M. Shimada, H. Ishiguro and S. Itakura, "Development of an android robot for studying human-robot interaction," *Proceedings of the 17th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems*, Ottawa, Canada (2004) pp. 424–434.
14. D. Matsui, T. Minato, K. MacDorman and H. Ishiguro, "Generating natural motion in an android by mapping human motion," *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Edmonton, AB (2005) pp. 3301–3308.
15. M. Shimada, T. Minato, S. Itakura and H. Ishiguro, "Evaluation of android using unconscious recognition," *Proceedings of the IEEE- RAS International Conference on Humanoid Robots*, Genoa, Italy (2006) pp. 157–162.
16. "Kokoro Company Ltd," Retrived March 2, 2014. http://www.kokoro-dreams.co.jp.
17. W. G. Wu, Q. M. Men and Y. Wang, "Development of the Humanoid Head Portrait Robot System With Flexible Face and Expression," *Proceedings of the IEEE International Conference on Robotics and Biomimetics*, Shenyang, China (2004) pp. 757–762.
18. J. H. Oh, D. Hanson, W. S. Kim, Y. Han, J. Y. Kim and I. W. Park, "Design of Android type Humanoid Robot Albert HUBO," *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Beijing, China (2006) pp. 1428–1433.
19. Hanson Robotics Company Ltd, Retrieved on March 2, 2014 from http://www.hansonrobotics.com.
20. D. W. Lee, T. G. Lee, B. So, M. Choi, E. C. Shin, K. W. Yang, M. H. Back, H. S. Kim and H. G. Lee, "Development of an Android for Emotional Expression and Human Interaction," *Proceedings of the International Federation of Automatic Control*, Seoul, Korea (2008) pp. 4336–4337.
21. K. Berns and T. Braun, "Design concept of a human like robot head," *Proceedings of the 5th IEEE-RAS International Conference on Humanoid Robots*, Tsukuba, Japan (2005) pp. 32–37.
22. K. Berns and J. Hirth, "Control of facial expressions of the humanoid robot head roman," *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Beijing, China (2006) pp. 3119–3124.
23. T. Hashimoto, S. Hiramatsu and H. Kobayashi, "Development of Face Robot for Emotional Communication between Human and Robot," *Proceedings of the IEEE International Conference on Mechatronics and Automation*, Luoyang, China (2006) pp. 25–30.
24. H. Minoru, Y. Chisaki and S. Tsugutake, "Development and Control of a Face Robot Imitating Human Muscular Structures," *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Beijing, China (2006) pp. 1855–1860.
25. K. Kaneko, F. Kanehiro, M. Morisawa, K. Miura, S. Nakaoka and S. Kajita, "Cybernetic Human HRP-4 C," *Proceedings of the 9th IEEE-RAS International Conference on Humanoid Robots*, Paris, France (2009) pp. 7–14.
26. S. Nakaoka, F. Kanehiro, K. Miura, M. Morisawa, K. Fujiwara, K. Kaneko, S. Kajita and H. Hirukawa, "Creating facial motions of Cybernetic Human HRP-4 C," *Proceedings of the 9th IEEE-RAS International Conference on Humanoid Robots*, Paris, France (2009) pp. 561–567.

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press

27. B. Allison, G. Nejat and E. Kao, "The Design of an Expressive Human-like Socially Assistive Robot," *ASME J. Mechanisms Robotics*, **1**(1), (2009).
28. C. Y. Lin, C. K. Tseng, W. C. Teng, W. C., Lee, C. H. Kuo, H. Y. Gu, K. L. Chung and C. S. Fahn, "The realization of robot theater-humanoid robots and theatric performance," *Proceedings of the International Conference on Advanced Robotics*, Munich, Germany (2009) pp. 1–6.
29. C. Y. Lin, L. C. Cheng and C. K. Tseng, "The development of mimetic engineering for theatric android head," *Proceedings of the International Conference on Service and Interactive Robotics*, Taipei, Taiwan (2009).
30. C. Y. Lin, L. C. Cheng, C. K. Tseng, H. Y. Gu, K. L. Chung, C. S. Fahn, K. J. Lu and C. C. Chang, "A face robot for autonomous simplified musical notation reading and singing," *J. Rob. Auton. Syst.* **59**(11), 943–953 (2011).
31. P. Ekman and W. V. Friesen, *Unmasking the Face* (Prentice-hall, Englewood Cliffs, 1975).
32. P. Ekman and W. V. Friesen, *The facial action coding system* (Consulting Psychologists Press, Palo Alto, 1978).
33. G. W. Jenkins, C. P. Kemnitz and G. J. Tortora, *Anatomy and physiology: from science to life* (John Wiley & Sons Inc. Press, New York, 2007).
34. L. C. Cheng, "*Development, Evaluation and Improvement of Humanoid Robot Head for Theatric Performance*," PhD dissertation, Dept. ME, NTUST, Taipei, Taiwan, ROC (2012).
35. N. A. Dodgson, "Variation and extrema of human interpupillary distance," *Proceedings of SPIE-The International Society for Optical Engineering*, San Jose, CA (2004) pp. 36–46.
36. L. C. Cheng, C. Y. Lin, amd C. C. Huang, "Visualization of Facial Expression Deformation Applied to the Mechanism Improvement of Face Robot," *Int. J. Soc. Rob.* **5**(4), 423–439 (2013).
37. M. N. Dailey, G. W. Cottrell, C. Padgett and R. Adolphs, "EMPATH: A neural network that categorizes facial expressions," *J. Cognitive Neurosci.* **14**(8), 1158–1173 (2002).
38. H. Aviezer, R. Hassin, J. Ryan, G. Grady, J. Susskind, A. Anderson, M. Moscovitch and S. Bentin, "Angry, disgusted or afraid? Studies on the malleability of emotion perception," *Psychological Science* **19**(7), 724–732 (2008).
39. R. Jack, C. Blais, C. Scheepers, P. Schyns and R. Caldara, "Cultural confusions show that facial expressions are not universal," *Curr. Biol.* **19**(18), 1543–1548 (2009).
40. T. Kanade, J. F. Cohn and Y. Tian, "Comprehensive Database for Facial Expression Analysis," *Proceedings of the 4*th *IEEE International Conference on Automatic Face and Gesture Recognition,* Grenoble, France (2000) pp.484–490.
41. I. Cohen, N. Sebe, A. Garg, L. S. Chen, and T. S. Huang, "Facial expression recognition from video sequences temporal and static modeling," *Comput. Vis. Image Understanding* **91**(1–2), 160–187 (2003)
42. M. Pantic and I. Patras, "Dynamics of Facial Expression Recognition of Facial Actions and Their Temporal Segments From Face Profile Image Sequen," *IEEE Trans. Syst. Men Cybernetics B: Cybernetics* **36**(2), 433–449 (2006).
43. S. Y. Wang, "The introduction of patent analysis on global industrial robot-The patents for mechanical transmission techniques,"(in Chinese), Retrived March 2, 2014. [Online] http://www.robotworld.org.tw/index.htm?pid=10& News_ID=4409.

https://doi.org/10.1017/S0263574714001787 Published online by Cambridge University Press