 Received 03-09-21; Accepted 26-09-21; Published 30-09-21

Generation of Head Mirror Behavior and Facial Expression for Humanoid Robots

YizhouChen,
XiaofengLiu, XFLIU@HHU.EDU.CN
JieLi, TingtingZhang
College of IoT Engineering,
Hohai University, Changzhou, 213100, China.
AngeloCangelosi
Cognitive Robotics Lab, Department of Computer Science,
University of Manchester, M13 9RL, U.K.

Corresponding Author: Xiaofeng Liu.

Abstract

Understanding of human expressions and action intentions is crucial to improve the intelligence of the social robots. The social robots can also perform mirror behaviors to learn from and understand others. Here we present an intelligent framework in which a humanoid social robot can simulate human facial expressions and emulate human head motions. To fulfill these requirements, an intelligent interactive robot control framework which can simulate human facial expressions and head motions is proposed. We developed a physical animatronic robotic head platform with soft skin, whose expressions can be adjusted through customized mechanical dynamic structure. A vision-based facial expression and head motion recognition method is proposed to establish the understanding mechanism of the expression and action intention of the intelligent robot using deep learning framework. Finally, the recognition results are employed to determine the optimal motor displacements for transferring the head states of a human to the robot. The experimental results demonstrate that our method can maintain high accuracy when facing different human subjects. The comprehensive evaluation shows that our method can perform accurate and real-time facial expression generation and head mirror behavior, which has a promising application for home assistant robots.

Keywords: Facial Expression, Head Mirror Behavior, Humanoid Robot, Deep Learning.

1. INTRODUCTION

Humanoid robots with the functions of facial expression generation and head mirror behavior, are widely used to assist the humans in their daily lives, such as education, nursing and elderly companion [1-3]. The facial expressions are now considered to be an important way of delivering

Citation: YizhouChen, et al. Generation of Head Mirror Behavior and Facial Expression for Humanoid Robots. Advances in Artificial Intelligence and Machine Learning. 2021;1(2):10.

human emotional information and harmonizing interpersonal relationships [4-6], which is essential for the assistant robot to understand human emotions and behaviors. The head mirror behavior, which is used to express more complex emotional states and intentions together with the facial expressions, is mixed with rich emotional intentions [7], and reflects a sense of self and covaries with the quality of the human-robot interaction. Therefore, designing a humanoid robot that can imitate human facial expressions and head motions will contribute to the development of social robots, and the robots can infer the mental states of people to provide more natural and trustworthy social services and human-robot interaction.

Current humanoid robotic head system cannot accurately and adaptively imitate human facial expressions and head motions. The key limitation is lacking mature and reliable robot expression generation control algorithms. Although the previous studies [8-10], which mapped human expressions to 2D or 3D virtual avatars, provided a reference for the humanoid robot to imitate the facial expressions and head motions [11-13], it still remains a challenging for the robots to realize the accurate tracking of facial expressions and head motions of human beings.

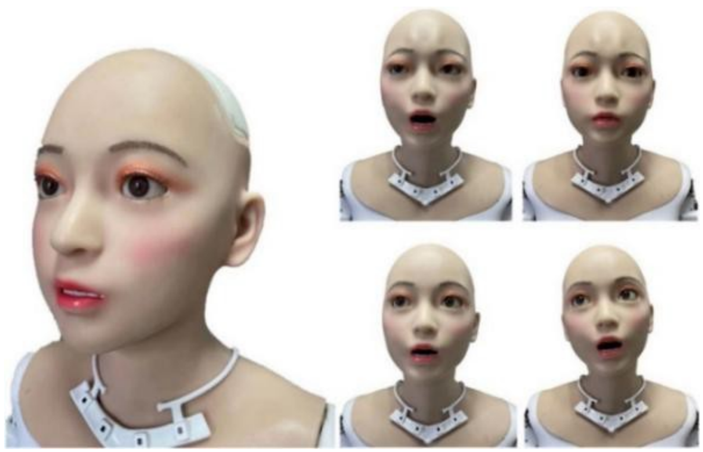


Figure 1: Human-like robotic head platform.

In this paper, as shown in FIGURE 1, a human-like robotic head platform for facial expression and head motion imitation was developed. The motor control protocol of the robotic head platform was further optimized to control the movement of the neck, mouth, cheeks, eyes, eyelids, and eyebrows by synthesizing a set of motion commands through ID number, position and speed of motor. Through training deep learning models for facial expression and head motion recognition, and establishing the control mechanism between the motor system and recognition model, our proposed method can improve the accuracy of the robotic head platform in imitating human facial expressions and head motions. The contributions of this paper can be concluded as follows.

- A mechanical control structure for a robotic head platform with soft skin was developed.
- A vision-based deep learning framework for expression generation method of intelligent robot was proposed.
- A vision-based intelligent robot head motion tracking method was proposed to realize the head mirror behavior.

2. RELATED WORKS

Understanding human facial expressions and action intentions is very important in social human-robot interaction [14, 15]. In previous studies, many researches have reported the progress of facial expression recognition and generation in intelligent robots [16]. For instance, Kobayashi et al. [17], developed the real time machine recognition of facial expressions by using a layered neural network. Wairagkar et al. [18], mapped the space of potential emotions of the robot to specific facial feature parameters. Chen et al. [19], developed a vision-based self-supervised learning framework for facial mimicry. However, none of them has considered the importance of head motion information for human-robot interaction. The focus of attention established by head motion can be used as a marker of visual attention [20], and the head orientation can provide information about the environment and the direction of human interests [21]. Therefore, head motion information is an important factor in increasing the variety of human-robot interaction.

Traditional head motion recognition methods [22], mainly rely on a fixed set of pre-programmed head motions, or use the positions of features such as the eyes, mouth, and nose tip to determine their relative configurations of poses [23]. However, these methods suffer from localization errors and low accuracy, and are poorly adapted in different human subjects. Therefore, in this paper, we use the 3D Morphable Face Model (3DMM) to reconstruct the human face by fitting the average face model in 2D photographs to obtain an accurate 3D face model, which can improve the accuracy of head motion prediction.

In facial expression recognition, most traditional approaches use mathematical methods [24], based on expression feature extraction. Some use Bayesian network or distance metric for feature classification [25, 26]. These methods are not very effective in feature extraction and also contain tedious manual feature extraction steps. Meanwhile, due to the continuous optimization of Graphic Processing Unit (GPU) and network architectures, deep learning methods which can substantially improve recognition results have gradually become mature and perfect its wider application [27, 28]. In this paper, a deep CNN-based facial expression recognition method was used to obtain facial expression recognition results and corresponding probabilities, thereupon then deriving the rotation position and speed of each motor.

This paper focuses on the imitation learning of the robotic head platform. With the help of two learning frameworks, facial expression recognition and head motion recognition, we trained our model on the public datasets (FERG-DB, CK+, and JAFFE) and adapted to different human face configurations. Eventually, the simulation of the facial expressions and head motions for the intelligent robot were realized.

3. METHODOLOGY

A robotic head platform which can imitate the head state of different human subjects was developed. In this paper, a webcam was used to capture the human face images, which were fed into the trained network to obtain the encoding information that could represent the facial expressions and head motions.The encoding was converted into motor signals to control the robotic head platform, and eventually achieved the target of simulating the facial expressions and head motions for the robotic head platform.

3.1 Design and Control of Robotic Head Platform

The robotic head platform integrates microprocessor, servo motors control system and mechanical structure. The microprocessor STM32F103C8T6 can control the twelve micro servo motors to drive the robotic head platform. The soft skin is connected to twelve micro servo motors via nylon cords. The overview of the hardware design is shown in FIGURE 2.

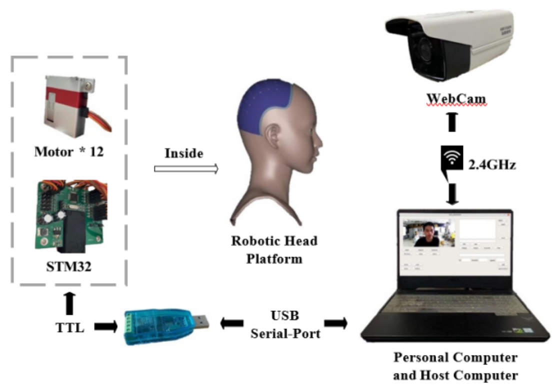


Figure 2: Overview of hardware design.

a) Design of robotic head platform: The robotic head platform can be further divided into head frame, face module, and neck module. The head frame is 3D printed according to human face, the interioris used to store the individual modules, and the exterior is closely fitted with soft skin to make the surface of the robotic head platform look more natural. As shown in FIGURE 3, a pair

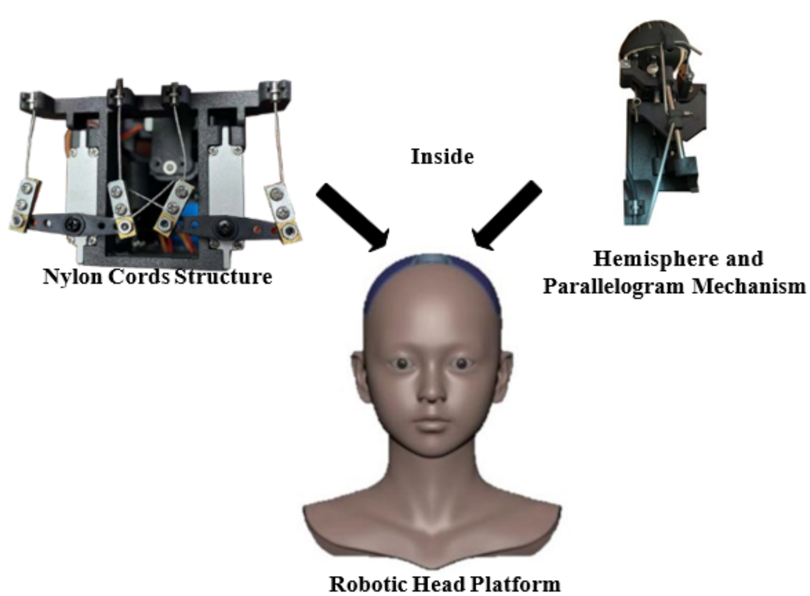


Figure 3: Mechanical structure in the robotic head platform.

of hemispheres and two pairs of parallelogram mechanisms enable the 4-DoF eyeball module to move freely horizontally and vertically within a 20-degree range. The eyebrow area, eyelid area and cheek area can be deformed to different degrees by pulling different nyloncords. The neck module achieves 6-DoF of rotation through the synergistic control of 4 motors, corresponding to three Euler angles of head motion.

b) Control of robotic head platform: The robotic head platform uses the STM32F103C8T6 micro-processor as the control core. The micro-processor and the motors communicate with each other in polling mode. Each motor has a uniqueID number in the bus-topology control network. The host computer generates a set of specific codes by recognizing facial expressions and head motions, and sends them to the microprocessor through the serial interface. The microprocessor drives the corresponding motors through the received codes to realize the simulation of facial expressions and head motions. The specifications of the robotic head platform are shown in TABLE 1.

Table 1: Specifications Of The Robotic Head Platform

Unit	Parameters
Voltage/Currentlevel	DC7.5V/10A
Baudrate	115200
Maximum movement of forward flexi on and backward extension	±43°
Maximum movement of left rotation and right rotation	±46.5°
Maximum movement of left and right head swing	±25°
Maximal opening of the jaw joint	25°
Maximum eye rotation left and right	±20°
Maximum eye rotation up and down	±20°
Maximum pull back of the corners of the mouth	5mm
Maximum dramatic eyebrow upward	5mm
Maximum width of the eyebrows to the center of the	5mm

As shown in TABLE 1, the robotic head platform has 14 degrees of freedom, including the control of the mouth, cheeks, eyelids, eyeballs eyebrows and the 3D rotation angles of the neck. The rotation angle of the motor in the command determines the amplitude of the robotic facial expression and the mobility of the neck. The rotation speed of the motor in the command determines the control speed of 14 degrees of freedom. Therefore, the motor ID number, position and speed information are used to control the robotic head platform. The control commands are automatically generated based on the three key pieces of information by the program to simplify the complexity of generation process, thereupon then making the control elements can better adapt to our training model framework.

3.2 Facial Expression and Head Posture Recognition Methods

A vision-based facial expression generation and head motion tracking method is proposed to obtain the head states of different human subjects for the robotic head platform to imitate the facial expressions and head motions of different human beings. In our method, the RGB webcam is used to capture the images in real time. When the webcam detects a human face, the recognition algorithm will segment the face area and feed the processed face image into the facial expression and head motion recognition networks to obtain the rotation angle and speed of the motor, which can be used to generate control commands. Eventually, the robotic head platform imitates the head state of the human subject according to the control commands.

a) Facial expression recognition: The schematic diagram of the facial expression recognition framework is shown in FIGURE 4. It can be divided into three processes: preprocessing, deep feature learning and facial expression classification,and control command generation.

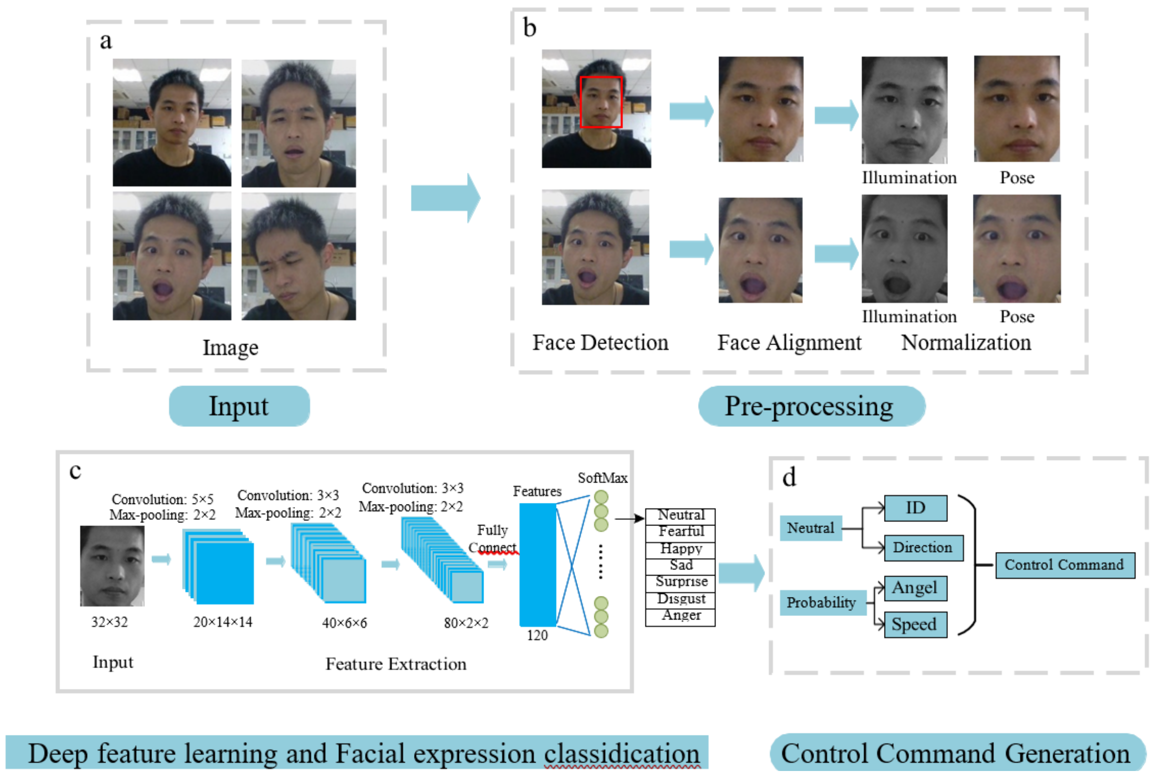


Figure 4: The schematic diagram of facial expression recognition system. (a) Input layer. Include7 expression emotions. (b) Pre-processing layer. Perform face alignment and normalization. (c) Deep feature learning and facial expression classification. (d) Control command generation.

As shown in FIGURE 4(b), in the preprocessing, we use the Haar-like feature classifier to implement face detection and the face key point detection algorithm from Dlib library for face alignment [29]. Finally, we normalize the face images, including geometric normalization and grayscale normalization. The purpose of geometric normalization is to transform the expression of sub-images to a uniform size, which is beneficial to the extraction of expression features. While grayscale normalization is used to increase the contrast of the images and perform lighting compensation to make the details of the images clearer to attenuate the effects of light and lighting intensity.

The facial expression feature extraction and expression classification were fused into an end-to-end network. The structure of the deep convolutional neural network is shown in FIGURE 4(c), which contains three convolutional layers, three pooling layers, one fully connected layer, and one SoftMax layer. The input image is a 32 × 32 grey face block. The number of features in each layer is decreasing and finally it is fixed at 120 in fully connected layer, which can express rich facial information. A dropout strategy before the fully connected layer is added to increase the model robustness. The SoftMax layer following the fully connected layer directly divides seven expression categories for recognition and the probability is normalized. Furthermore, the functions were added to calculate the loss and accuracy of the model code. After 50,000 epochs of training, the accuracy of our model can research up to 95%.

The predicted emotional categories were corresponded to a fixed set of pre-programmed facial expressions to obtain the ID number and rotation direction for driving the motors. The rotation angle and speed of the motors were obtained according to the normalized probability. Finally, the ID number, position and speed information of each motor are stored in the container in turn to generate control commands.

b) Head mirror behavior: The three Euler angles of the head can be calculated through the rotation matrix from the camera coordinate system to the world coordinate system by a series of mathematical equations, and the calculation of the rotation matrix requires multiple sets of corresponding points between the camera and world coordinate systems. Therefore, the facial features are used as the corresponding points for computing the rotation matrix. In the camera coordinate system, the face key point detection algorithm from Dlib library is used to detect the coordinates of facial feature points of 2D images and get the coordinates of face feature points in the camera coordinate system according to the camera internal reference and distortion coefficients.

Considering the different facial feature distributions of different human subjects, there was a large error if the average 3D face model was used to calculate the head motions in the world coordinate system. Therefore, the 3DMM method [30] was used to re-model the human face. Each human subject had a corresponding 3D face model as a way to improve the accuracy and precision of head motion recognition and make the robot head closer to the real head state when imitating.

The 3DMM method is a linear representation of a face, where a new face model contains both shape and expression parameters. The linear representation of the face model is described as Eq. (1):

$$\sum\nolimits_{fModel} = \bar{s} + \sum\nolimits_{i=1}^{m-1} \alpha_i s_i + \sum\nolimits_{i=1}^{n-1} \beta_i e_i \tag{1}$$

where s_{fModel} is the new face model, $\bar{s}$ is the average face model, s_i is shape and e_i is expression. Then the face reconstruction problem can be converted into a problem of finding the α and β.

In this paper, cascade regression is used to predict α and β. Cascade regression will progressively refine the specified initial predictions by a series of regressors. Each regressor relies on the output of the previous regressor to perform the next step. We choose an average face model and regress the parameters to the ground-truth by multiple regressions. The mathematical expression is described as Eq. (2):

$$p^{K+1} = p^k + Reg^k(Fea(I, P^k)) \tag{2}$$

where p is the regression target and Fea is the image feature. The p contains the following elements:

$$P = [f, R, t, \alpha, \beta] \tag{3}$$

where f is the scaling factor, R is the rotation matrix and t is the translation matrix. The rotation matrix is represented by the Euler angles, which are the attitude angles of yaw, pitch and roll.

Each pose Euler angle has a corresponding motor and ID number on the robotic head platform. The positive and negative of the three Euler angles determine the direction of the head rotation and the values of the three Euler angles determine the angle of the head rotation. Therefore, the control commands of the neck module in the robotic head platform can be obtained based on the three Euler angles.

4. EXPERIMENTS

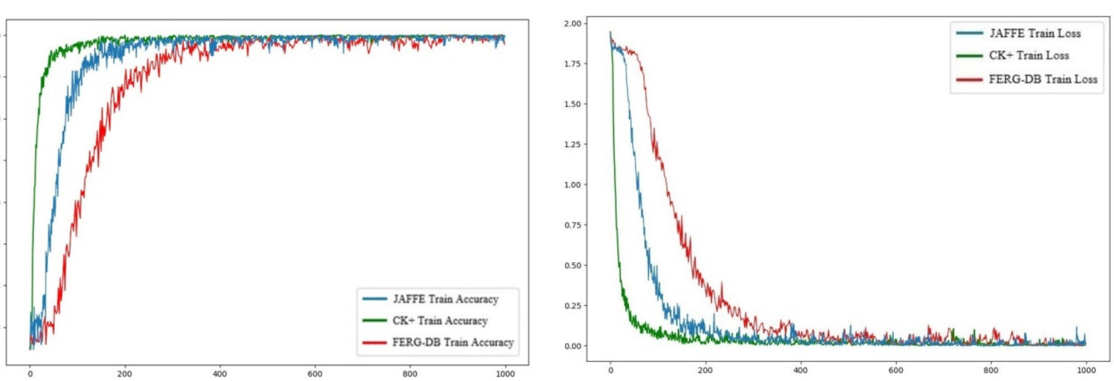
In order to evaluate the effectiveness of our method, the accuracy of the algorithms for facial expression recognition and the head motion tracking is evaluated in this section. Meanwhile, the effectiveness of the robotic head platform for facial expression generation and head mirror behavior of human subjects is evaluated.

4.1 Evaluation Metrics

For the facial expression recognition algorithm, three datasets FERG-DB, CK+ and JAFFE were chosen for validation. In order to know whether our model can accurately recognize the facial expressions of different human subjects, we visualized the accuracy, sensitivity and specificity of each expression category and show the results of comparison experiments. For the head motion recognition algorithm, the effectiveness of face reconstruction, the accuracy of head motion recognition and the maximum angle of head rotation were tested to demonstrate the effectiveness of our method.

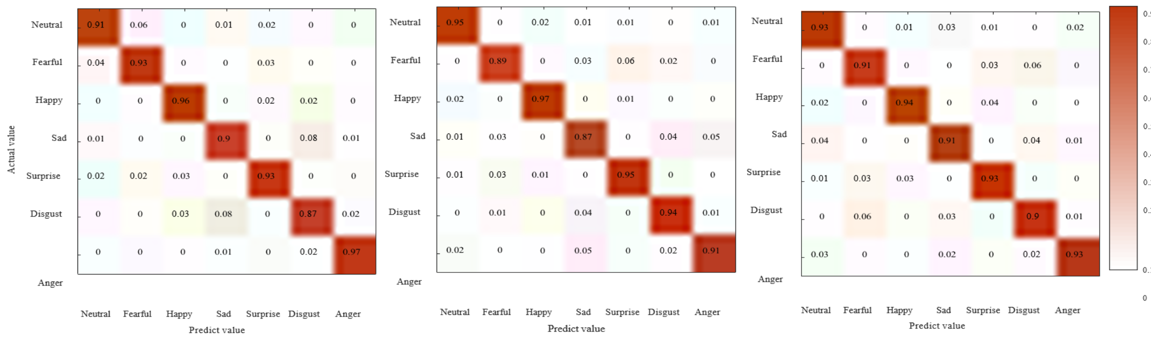
4.2 Facial Expression Recognition Performance

In order to evaluate the facial expression recognition algorithm, three datasets FERG-DB, CK+ and JAFFE were used to train and test our network respectively. In the experiment, we defined a total of seven expressions, which are neutral, fearful, happy, sad, surprise, disgust and anger. FIGURE 5 shows the training accuracy and training loss of the recognition algorithm. FIGURE 6(a) shows



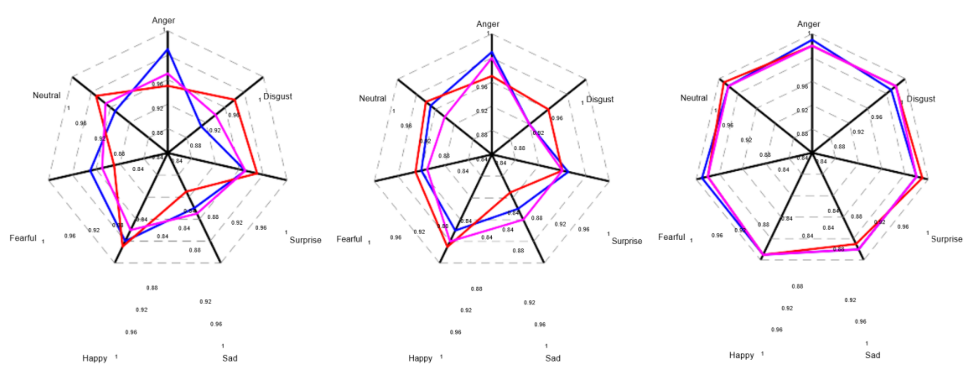
(a) Training Accuracy **(b) Training Loss**

Figure 5: Training accuracy and loss of our facial expression recognition.



(a) Confusion matrix of FERG-DB **(b) Confusion matrix of CK+** **(c) Confusion matrix of JAFFE**

Figure 6: Confusion matrix of our facial expression recognition.



(a) Accuracy **(b) Sensitivity** **(c) Specificity**

Figure 7: The accuracy, sensitivity and specificity performance of seven expressions in three datasets. The red line represents the dataset FERG-DB, the bule line represents the dataset CK+ and the purple line represents the dataset JAFFE.

the confusion matrix of the recognition algorithm used in dataset FERG-DB. FIGURE 6(b) shows the confusion matrix of the recognition algorithm used in dataset CK+. FIGURE 6(c) shows the confusion matrix of the recognition algorithm used in dataset JAFFE. The accuracy, sensitivity and specificity of seven expressions in three datasets are shown in FIGURE 7. The red line represents the dataset FERG-DB, the bule line represents the dataset CK+ and the purple line represents the dataset JAFFE. Our method performs consistently on different datasets. The training accuracy and sensitivity of different expressions are above 92% and the training specificity can reach more than 97%, which proves the effectiveness and stability of our method.

In order to further illustrate the effectiveness of our method, we compare the recognition method with the state-of-the-art. TABLE 2 shows the comparison of the classification accuracy of facial expression. It can be found that our method performs consistently and accurately in the three

Table 2: Classification accuracy on the FERG-DB, CK+ and JAAFE datasets.

Dataset	Algorithm	Accuracy
FERG-DB	Our Method	94.42%
	Deep Expr [31]	89.02%
	Fisher face[32]	89.2%
CK+	Our Method	93.57%
	3DCNN[33]	88.99%
	LBP-TOP[34]	85.90%
JAAFE	Our Method	93.24%
	Salient Facial Patch[35]	91.8%
	Deep Features+HOG[36]	90.58%

datasets. In other word, for different human subjects, our method can effectively understand their facial expression states and ensure the accuracy of the facial module in the robotic head platform to imitate facial expressions.

4.3 Head Motion Recognition Performance

For the head motion recognition algorithm, the effectiveness of 3DMM face reconstruction were firstly evaluated. FIGURE 8(b) shows the results of 3DMM face reconstruction, and FIGURE 8(c) shows the visualized face model on MATLAB. These results show that the reconstructed face model can well describe the contours of the face and the positions of the five senses and is applicable to various human subjects, which can help us to obtain more accurate Euler angles for the head motion.

(a) Image **(b) 3DMM Face Reconstruction** **(c) Face Model Visualization**

Figure 8: Result for 3DMM face model reconstruction

The accuracy and precision of the head motion recognition algorithm were evaluated. FIGURE 8 shows the recognition results of our head motion recognition algorithm for 6-DoF of the head. OpenCV were used to describe a quadrilateral in the image to represent the head orientation. The recognition results show that our head motion recognition algorithm can recognize all poses and orientations within the rotatable range of the human head with a high accuracy. Then the precision of the head posture recognition algorithm was measured. According to the collected yaw, pitch and roll angles and the results of the actual camera tests, the precision of the measurement can reach 0.1 degree. As shown in FIGURE 9, the left and right head rotation is about 90 degrees, the up and down head rotation is about 60 degrees and the left and right swing is about 45 degrees. Even with these large head rotation angles, our algorithm can still perform accurate recognition, which indicates that our method can effectively improve the accuracy of head motion recognition.

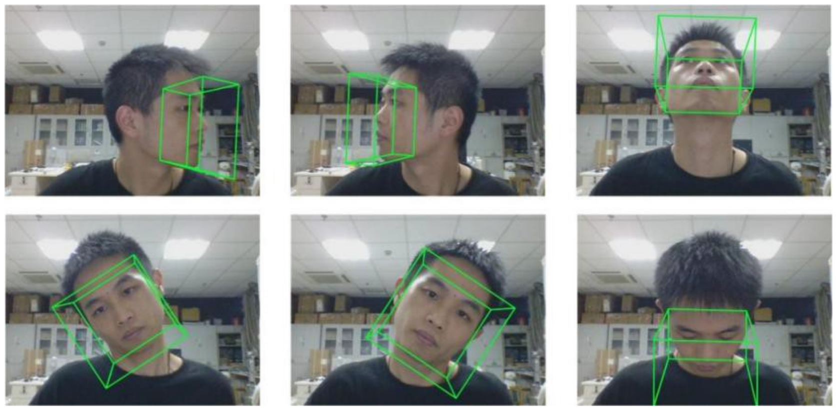


Figure 9: Results of head motion recognition algorithm for 6DoF of the head.

4.4 Facial Expression Generation and Head Mirror Behavior

The optimal motor displacements, which are obtained from the outputs of facial expression recognition and head motion recognition, are executed on our robotic head platform. To provide a qualitative evaluation on the similarity with a human subject, ten sets of representative frames were intercepted randomly during the test to evaluate the facial expression generation and head mirror behavior on the robotic head platform.

First, we invited three participants to evaluate the performance of the robotic head platform. The participants were asked to perform the same five sets of facial expressions and head motions. Five sets of comparison frames for each participant, which can reflect the mapping from the real human face to the robotic head platform, are visualized in FIGURE 10. The results demonstrate that the robotic head platform can imitate different human head state according to the control commands sent by the recognition algorithm with real-time and similarity. Although FIGURE 10 reflects the performance of the imitation similarity, it measures neither the generation similarity of facial expression nor the mapping similarity of the head mirror behavior.

Since these indicators are important for evaluating the reliability of the robotic head platform and interaction experience of humans, we further evaluated the performance of the facial expression generation and head mirror behavior from a human subject to the robotic head platform. As shown in FIGURE 11, ten sets of mapping comparison frames were fed into our facial expression recognition algorithm. The vertical coordinate represents the probability of emotion recognition, and different colors represent different emotion categories. The results demonstrate that the robotic head platform can imitate the human facial expressions with same recognition result and similar recognition probability, which reflect the imitation similarity from the robotic facial expressions to the human facial expressions.

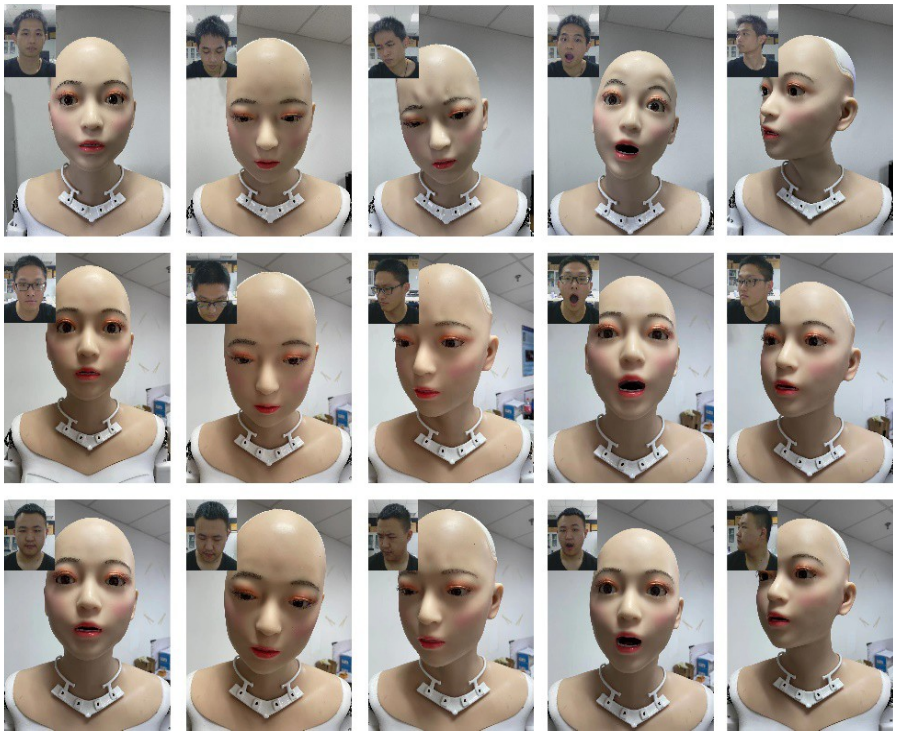


Figure 10: Facial expression generation and head mirror behavior from different human subjects to the robotic head platform.

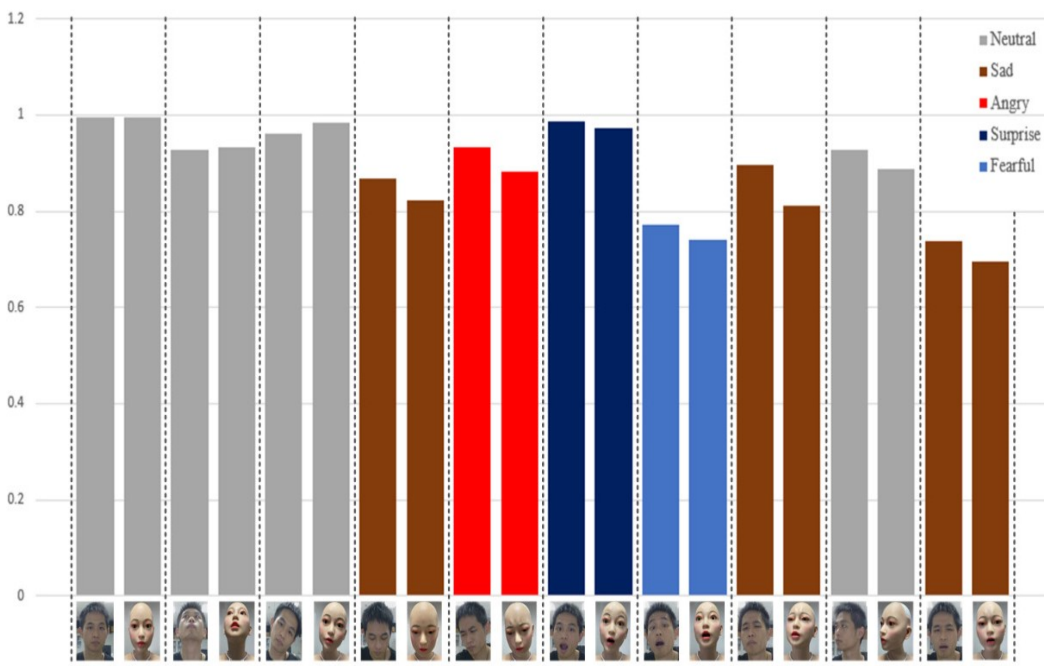


Figure 11: Similarity of probabilities on facial expression generation from humans to humanoids.

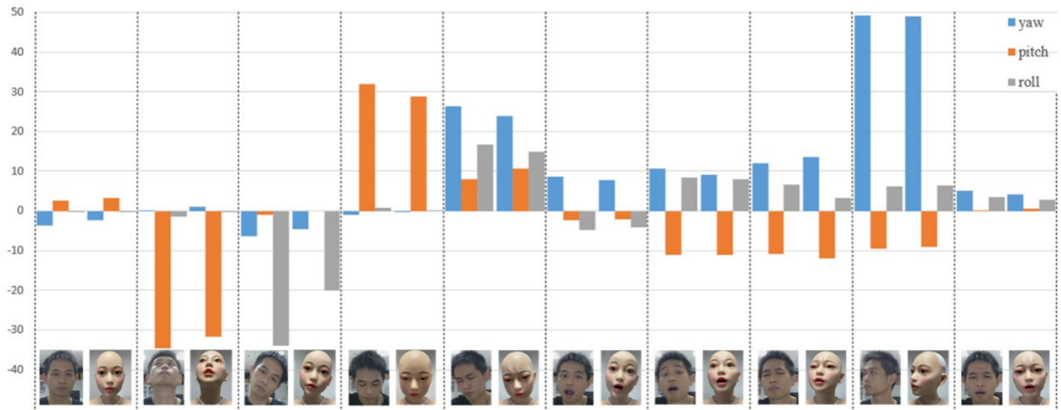


Figure 12: Similarity of yaw, pitch and roll on head mirror behavior from humans to humanoids.

In the end, the performance of the head mirror behavior is evaluated through head motion recognition algorithm. As shown in FIGURE 12, ten sets of mapping comparison frames were fed into our head motion recognition algorithm. The vertical coordinate represents the value of three Euler angles, which contains yaw, pitch and roll. According to the similarity of the three Euler angles in each set, it can reflect that the robotic head platform can accurately obtain human head motions and complete the imitation.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we present a facial expression generation and head mirror behavior for humanoid robots. A deep learning framework is proposed to recognize the human facial expressions and head motions. A control mechanism between the motor system and recognition model is established to drive the robotic head platform. The evaluation results reflect the accuracy and reliability of the facial expression and head motion recognition algorithms. In addition, the imitation similarity of facial expressions and head motions on the robotic head platform is discussed, which have a certain reference value for the research and development of emotional interaction of assistant robots.

Some limitations also exist in our study. First, current algorithm suffers from the low accuracy for multi-target face recognition. Second, our hardware system is still relatively rough and has many possibilities for improvement, especially in terms of data transmission, device packaging, and system power consumption. In addition, although the evaluation results reflect the good performance of the imitation on the robotic head platform, the naturalness and identity of facial expression generation and head mirror behavior are closely related to the subjective perceptions during natural human-robot interaction. Therefore, current research lacks the evaluation of actual interaction scenarios.

In future research, electroencephalography, speech et al., which can reflect emotional changes, will be considered for information fusion, realizing multimodal emotion expression and recognition to make the expression generation of the robot more natural and efficient. Simultaneously, WiFi or Bluetooth wireless communication channels are also being considered to replace the serial port to increase the controllability of the system. In addition, actual human-robot interaction experiment about affective communication will be considered to enhance the depth of research. Of course, commercial applications are also being pursued based on this research, including the development of home assistant robots and human-robot interaction.

6. CONFLICT OF INTEREST

No potential conflict of interest was reported by the authors.

7. ETHICAL GUIDELINE

There were no any examinations on a human or living creatures in this paper.

8. ACKNOWLEDGEMENT

This work was supported in part by National key R&D program of China 2018AAA0100800, the Key Research and Development Program of Jiangsu under grants BK20192004B and BE2018004, Guangdong Forestry Science and Technology Innovation Project under grant 2020KJCX005, International Cooperation and Exchanges of Changzhou under grant CZ20200035, Changzhou Technical key Project under Grant No. CE20205041, the Fundamental Research Funds for the Central Universities under Grant No. B200202209.

References

[1] Cristina S, Camilleri KP.Model-based Head Pose-free Gaze Estimation for Assistive Communication. Computer Vision and Image Understanding. 2016;149:157-170.

[2] Hara F, Kobayashi H.Use of Face Robot for Human-computer Communication.In IEEE International Conference on Systems, 1995.

[3] Zhong J, Ling C, Cangelosi A, Lotfi A, Liu X.On the Gap Between Domestic Robotic Applications and Computational Intelligence. Electronics.2021;10:793.

[4] Mori H, Miyawaki K, Sano M, Nishiguchi S.A Study of Expression Change Detection Method for Smooth Communication. IEICE Technical Report. 2008;108:159-162.

[5] Breazeal C, Buchsbaum D, Gray J, Gatenby D, Blumberg B.Learning From and About Others: Towards Using Imitation to Bootstrap the Social Understanding of Others by Robots. Artificial Life. 2005;11: 31–62.

[6] Shiomi M, Kanda T, Miralles N, Miyashita T, Ishiguro H.Face-to-face Interactive Humanoid Robot," in Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on, 2004.

[7] Murphy-Chutorian E, Trivedi, Manubhai M.Head Pose Estimation in Computer Vision: a Survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009.

[8] Oh JH, Hanson D, Kim WS, Han Y, Park IW.Design of Android Type Humanoid Robot Albert HUBO, in IEEE/RSJ International Conference on Intelligent Robots and Systems, 2007.

[9] Hashimoto T, Hiramatsu S, Kobayashi H.Dynamic Display of Facial Expressions on the Face Robot Made by Using a Life Mask, in Humanoid Robots, 2008. Humanoids 2008. 8th IEEE-RAS International Conference on, 2009.

[10] Lin CY, Huang CC, Cheng LC.An Expressional Simplified Mechanism in Anthropomorphic Face Robot Design.Robotica, 2016.

[11] Borghi G, Fabbri M, Vezzani R, Calderara S, Cucchiara R.Face-from-depth for Head Pose Estimation on Depth Images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017;42:596-609.

[12] Balasubramanian VN, Ye J,Panchanathan S.Biased Manifold Embedding: A Framework for Person-independent Head Pose Estimation. IEEE Conference on Computer Vision and Pattern Recognition, 2007: 76-82.

[13] Valenti R, Sebe N, Gevers T.Combining Head Pose and Eye Location Information for Gaze Estimation, IEEE Transactions on Image Processing. 2012;21:802–815.

[14] https://arxiv.org/pdf/2004.06558.pdf

[15] Newman R, Matsumoto Y, Rougeaux S, Zelinsky A.Real-time Stereo Tracking for Head Pose and Gaze Estimation. IEEE International Conference on Automatic Face and Gesture Recognition, 2000:1:122.

[16] Hashimoto T, Hitramatsu S, Tsuji T, Kobayashi H.Development of the Face Robot Saya for Rich Facial Expressions, in SICE-ICASE, 2006. International Joint Conference, 2007.

[17] Kobayashi H, Hara F.Facial Interaction Between Animated 3d Face Robot and Human Beings, in Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation., 1997 IEEE International Conference on, 1997.

[18] https://arxiv.org/ftp/arxiv/papers/2012/2012.04511.pdf

[19] https://arxiv.org/pdf/2105.12724.pdf

[20] Jean-David B, Ugo P, Amelie L, Gerrard B, Frederic E, et al.I Reach Faster When I See You Look: Gaze Effects in Human-human and Human-robot Face-to-face Cooperation. Frontiers in Neurorobotics. 2012;6:3.

[21] Kompatsiari K, Ciardo F, De Tommaso D, Wykowska A.Measuring Engagement Elicited by Eye Contact in Human-robot Interaction.In 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2019: 6979-6985.

[22] HaggertySE, Michael KW.The Interaction of Pre-programmed Eye Movements With the Vestibulo-ocular Reflex. Frontiers in Systems Neuroence. 2018;12:4.

[23] Qian C, Wu H, Shioyama S, Shimada T.Head Pose Estimation Using Both Color and Feature Information.In International Conference on Pattern Recognition, 2000.

[24] Kim TY, Cho NH, Jeong GB. Bengtsson E, Choi HK.3D Texture Analysis in Renal Cell Carcinoma Tissue Image Grading. Computational and Mathematical Methods in Medicine, 2014;2014: 536217.

[25] https://patentimages.storage.googleapis.com/26/57/1b/d0ff7eca724124/US20050060295A1.pdf

[26] Fagnani F, Fosson SM, Ravazzi C.A Distributed Classification/estimation Algorithm for Sensor Networks. Siam Journal on Control and Optimization.2014;52: 189–218.

[27] Zheng Y, Zhu C, Luu K, Bhagavatula C, Le T, et al.Towards a Deep Learning Framework for Unconstrained Face Detection.In IEEE, 2016: 1-8.

[28] https://arxiv.org/pdf/1803.05846.pdf

[29] Sharma S, Shanmugasundaram K, Ramasamy SK. FAREC-CNN based efficient face recognition technique using Dlib, in 2016 International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), 2016.

[30] Amberg B, Knothe R, Vetter T. “Expression invariant face recognition with a 3dmm,” 2008.

[31] Aneja D, Colburn A, Faigin G, Shapiro L, Mones B. Modeling stylized character expressions via deep learning, in Asian Conference on Computer Vision, 2016.

[32] Abidin Z,Harjoko A. A neural network based facial expression recognition using fisher face. International Journal of Computer Applications.2012;59.

[33] Liu M, Li S, Shan S, Wang R, Chen X.Deeply Learning Deformable Facial Action Parts Model for Dynamic Expression Analysis.In Asian Conference on Computer Vision, 2014.

[34] ZhaoG, Pietikainen M.Dynamic Texture Recognition Using Local Binary Patterns With an Application to Facial Expressions, IEEE Transactions on Pattern Analysis and Machine Intelligence. 2007;29:915-928.

[35] Happy SL, Routray A.Automatic Facial Expression Recognition Using Features of Salient Facial Patches. IEEE Transactions on Affective Computing.2015; 6:1-12.

[36] Wang H, Wei S, Fang B.Facial Expression Recognition Using Iterative Fusion of Mo-hog and Deep Features, The Journal of Supercomputing, 2020;76:5.