Liu et al., Gerontechnology (2023) 22:1
https://doi.org/10.4017/gt.2023.22.1.827.07

Facial emotion recognition tool for teleoperated robot reminiscence group therapy for people with dementia

Liu Xiangyu MA[a,*], Yamazaki Ryuji PhD[b], Kase Hiroko PhD[c]

[a]*Waseda University, Japan;* [b]*Institute for Open and Transdisciplinary Research Initiatives, Osaka University, Japan;* [c]*School of Social Science, Western Sydney University, Australia;*
**Corresponding author: dukeedison@fuji.waseda.jp*

Abstract

Background: Historically, questionnaires and laboratory tests have been used to verify reminiscence group therapy (RGT) efforts which are widely used methods for the non-pharmacological treatment of dementia. Today, facial expression recognition tools may be applied in a noncontact, more objective, and efficient manner.
Objective: We propose to enhance such applications with more accurate databases and higher performance-trained models to aid PWD in the future.
Method: Our study used the OpenCV-based AI facial emotion recognition tool to explore the feasibility of emotional improvement for dementia RGT. Data from two intervention experiments of teleoperated robot RGT conducted on people with dementia (PWD) were analyzed. Experiments were conducted on six PWD in Tokorozawa, Saitama, and eight PWD in Nakameguro, Tokyo. The Emo-Rec application showed a difference in trends between the crossover method and the controlled trial.
Results: The trend for Experiment Tokorozawa was relatively flat compared with that of Experiment Nakameguro. The overall group means were mostly below 10, indicating that RGT was relaxing and enjoyable, and that participants' moods improved over time.
Conclusion: Based on our data analysis of long-term interventions, we believe that a facial emotion recognition tool is more objective for quantifying participants' emotions, without the need to use other instruments. More importantly, it does not suffer from cognitive barriers for self-reports, questionnaires, or interviews.

Keywords: facial emotional recognition, reminiscence group therapy, teleoperated robot, dementia

Introduction

Reminiscence group therapy (RGT) is one of the most effective psychosocial interventions for people with dementia (PWD). It is easy to organize, labor-saving, and has been used in many elderly care facilities with additional props such as memory books (Chao et al., 2006). RGT can also reduce the behavioral and psychological symptoms of dementia (BPSD) in participants, facilitating group communication, improving their mood, and slowing down PWD's social regression. Most studies verifying the effect of therapy on participants have been conducted using questionnaires, such as the Mini-Mental State Examination (MMSE), the ADAS-cog, and the MoCA, and for nursing staff, questionnaires like the NPI DBD, among others (Cotelli et al., 2012). Studies have verified the improvement in patients with dementia using emotional questionnaires (e.g., the Face Scale) and linguistic analysis to validate the effects of therapy (Kase et al., 2019).

Elderly adults with dementia also experience rare medical conditions, such as proteinopathy and dysfunctional brain proteins. This may cause physical dysfunction, communication barriers, or comprehension barriers. Consequently, challenges to the questionnaire survey process may emerge, thereby affecting the final results, thus, causing bias (Irazoki et al., 2017; Rosa-Neto, 2021). The emotional recognition of RGT has received support as a new verification tool. Researchers can use the expression recognition results of RGT to grasp participants' emotions and conditions and evaluate the effects of various interventional experiments (Liu, 2021).

A comfortable and social atmosphere should be maintained for PWD during RGT, and a contactless emotion analysis tool could help researchers better assess the effects of RGT in more dimensions. We propose mechanisms that could improve the efficiency of evaluating RGT effectiveness to facilitate further application of this therapy in robotic cognitive care.

Figure 1. Telenoid and experiment T

Method

Our study aimed to validate the usability of a facial emotion recognition tool for RGT using teleoperated robotic experiments. Two experiments were conducted. Expression data obtained via teleoperation robot experiments with humans and robots were compared and analyzed to discuss the differences in emotion recognition between human and robotic interactions. Two robots, Telenoid and Pepper, were used.

The candidates were required to have basic communication abilities and an overall tendency to have milder MMSE scores. Experiment T was conducted with six PWD (age = 87.5±5.76), with an average MMSE score of 19 (range, 27–13), at a low-cost nursing home in Tokorozawa, Saitama as shown in *Figure 1*. From July 2018 to August 2018, all six PWD participated in RGT sessions for 20 minutes, as follows: teleoperation robot session (3 times) and human intervention (3 times). The other group's sessions were conducted in reverse order for the cross-over interventional experiment.

Experiment N was conducted with eight PWD (age=85.75±3.7(G1), 89.00±2.37(G2); MMSE=15.25±4.85(G1), 17.25±12.21(G2).) in a nursing home in Nakameguro, Tokyo, from December 4, 2020, to January 12, 2021, as shown in *Figure 2*. Experiment N was conducted using a controlled trial of human or teleoperated robot groups, and PWD received a single interventional approach. We compared the data from the two interventional experiments separately for humans and teleoperated robots, and the Emo-Rec app was used to measure the emotion ratings of the participants.

Emotional change is difficult to accurately capture over time, and the speed and direction of change are difficult to predict (Albers & Bringmann, 2020). Nevertheless, by sampling multiple times per experiment, it was possible to provide a relatively accurate description of the overall trend in participants' emotions during each RGT session. During the sampling phase, we marked the timeframe in each RGT video by referring to the random sequence given by SPSS 12 Com-

Figure 2. Pepper and experiment N

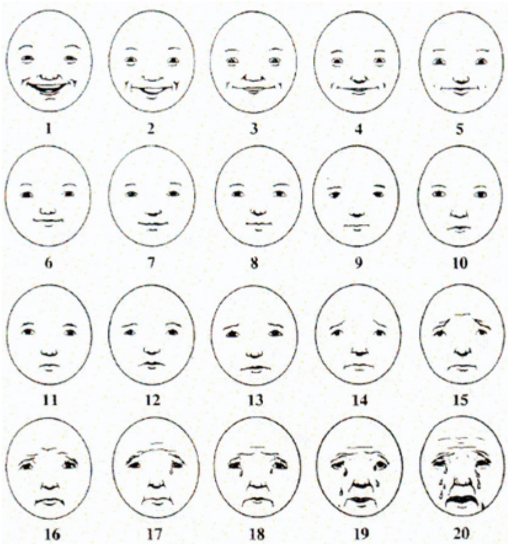


Figure 3. The Face Scale

patible and classified the frames of each video using the Emo-Rec app to derive the proportion of each emotion expressed by the participants. According to the length of the video, we need to generate a series of random numbers to determine the specific number of seconds in the video for sampling. We have adopted the random number generation method in SPSS 12. For example, if the duration of the video data is 10 minutes, 9 random numbers from 1 to 600 will be generated. And use the number of seconds corresponding to the random number to extract the time frames to complete the sampling.

OpenCV-based open-source tools were used to quantify and analyze PWD's mood swings in social environments. The results were divided into seven categories of expression: neutral, anger, contempt, disgust, fear, happiness, and sadness. We converted the seven emotions detected by the Emo-Rec app into a binary ratio of "happy–sad" dimensions that were retained and then divided into a scale, scored from 1 to 20 according to the Face Scale (Lorish & Maisiak, 1986). Finally, the Emo-Rec point was calculated using the following formula:

Emo-Rec point=(anger*10) + (disgust*10) + (scare*10) + (happy*1) + (sad*20) + (surprise*10) +(neutral*10)

The output result of Emo-Rec App is expressed in percentage matching degree. The ratio given by the Emo-Rec App can ensure that the sum of the ratios of each emotion is always 100%. Through this formula and the ratio given by the app, we can get a range of 1-20. For example, happy100% is 1, and sad 100% is 20. We used this formula under the consideration of converting the output result into the more commonly used Face Scale. This Scale is to obtain the true feelings of PWD through a questionnaire survey as shown in *Figure 3*.

Face Scale is also a scale commonly used to describe emotions and feelings in Japanese dementia patients. We adopted this scale to facilitate

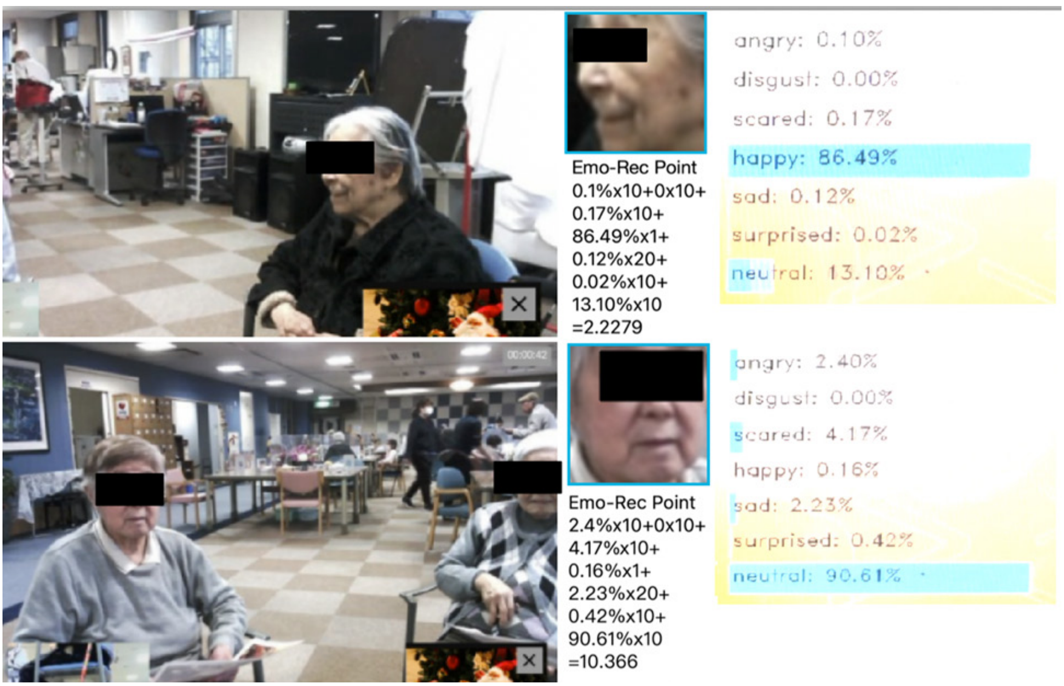


Figure 4. Actual use of Emo-Rec Point

Table 1. The Emo-Rec Point of interventions (Under 10)

Experiment		Under 10	% within	X^2	p
T(n=258)	Human	112	69.1%	0.057^a	0.811
	Robot	110	67.9%		
N(n=324)	Human	56	44.4%	0.596^b	0.440
	Robot	65	49.2%		
Human in Total		168	58.3%	0.085^c	0.770
Robot in Total		175	59.5%		

*0 cells (0.0%) have expected count less than 5. The minimum expected count is 51.00[a]. 59.09[b]. 118.27[c].

communication and coordination with facility staff. We can imagine the 10 of the Emo-Rec Point as 0, and the resulting range is between -10 and 10. In order to compare the final Emo-Rec Point with the widely used Face Scale, to communication with facility nursing staff and to continue discussing the result, we set it to 1-20. The reason for setting this weight is also to try to filter other emotions other than happy-sad, to better match the Face Scale.

In the examples shown in the *Figure 4*, we calculated the Emo-Rec Point through the formula based on the emotion ratios given by the Emo-Rec App on the right side, and then we can find the approximate expression through the Face Scale. For the upper example, the proportion of happy is relatively high, and the Emo-Rec Point is 2.2279. While, the one below is relatively calm, and the result is 10.366.

The classification accuracy rate of the FER2013 dataset (another relatively complete classification database) using the Emo-Rec application was 86%. The validity of the scores obtained by the formula were verified by conducting a questionnaire survey with experienced nurses and researchers to manually score the participants on the frames. The correlation of Spearman's test in the two groups was 0.863, and the intraclass correlation coefficient (ICC) was 0.835 (Liu et al., 2022).

All analyses were conducted using the IBM SPSS Statistics R26 program, 64-bit version. Descriptive data were used to test the differences in overall mood between the two interventions. Finally, the mean scores of the participants from each RGT group were pooled to identify mood improvement trends. This study complied with the requirements of the Waseda Ethical Review Committee (ethical review numbers 2017-173 and 2019-328).

Results

We pooled the overall Emo-Rec Points under 10 for the different intervention methods in the two experiments. The *Table 1* indicated that participants' facial expressions were likely to exhibit more relaxed emotions when faced with a teleoperated robot. The total under 10 proportion for face-to-face participants were less than those of the participants in the robot group. Additionally, the people who tend to be emotionally pleasant of Experiment T, were slightly more than Experiment N.

We aggregated the average Emo-Rec Points of all participants in each RGT in the two experiments in *Table 2*, summarized them in the trend graph, and expected to observe the effects of two different experimental settings on the participants' emotions. We visualized the trend for Experiment T was relatively flat compared with that of Experiment N through the *Figure 5*. The overall group means were mostly below 10, indicating that RGT was relaxing and enjoyable, and that participants' moods improved over time.

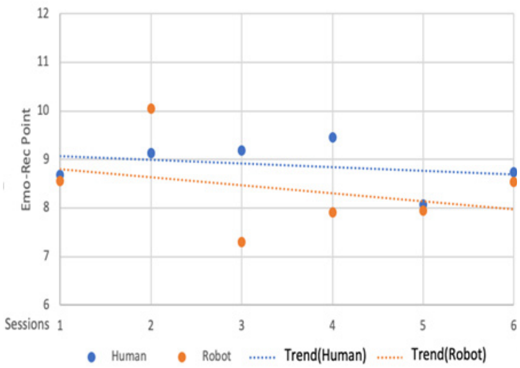
Discussion

According to the statistical results in Table 1 for points under 10, the number of people who felt pleasure was significantly higher in the crossover setting of experiment T. However, the robot group performed better in experiment N of the controlled trial. The Robot group was also in a dominant position in the overall data. This suggests that interaction with a robot brings about a more pronounced positive emotion than interaction with a human operator. This is in line with Heerink's theory of "Perceived Enjoyment" , that the pleasure of use arises when a decision has been made to use a robot (Heerink et al., 2010). Heerink organized a robot experiment to verify the relationship between user enjoyment (Perceived Enjoyment) and the willingness of the elderly to use robots. In the model of Heerink's theories, Perceived Enjoyment was the key factor affecting the decision to use robots. This explains the relatively favorable affective tendency of the robot group found in the experiment. The hedonic effects of new technologies and products on people are also mood-enhancing effects of robots.

The participants in Experiment T experienced a crossover set-

Table 2. The overall average Emo-Rec Point of participants in each RGT

		G1	G2	G3	G4	G5	G6
Experiment T	Human	8.687	9.129	9.191	9.446	8.078	8.739
	Robot	8.561	10.052	7.298	7.913	7.941	8.539
Experiment N	Human	-	11.091	9.759	10.711	8.335	10.576
	Robot	9.683	10.663	-	7.959	9.026	6.486



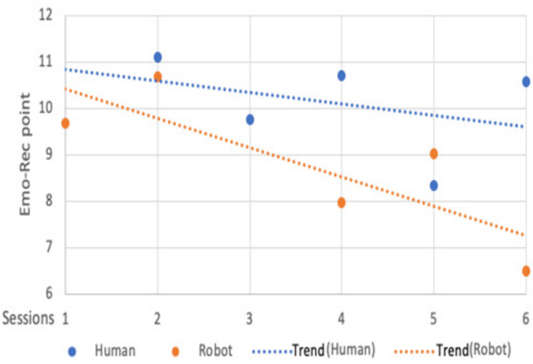


Figure 5. Trends of experiments T & N

ting of human and teleoperated robot experiments in which the effect of alternating stimuli was more pronounced and the mood tended to be significantly more relaxed. Experiment N involved a single stimulus in the controlled trial, with both groups receiving a single intervention with either a human or robot. Hence, the participants in Experiment N were at a disadvantage in terms of their overall Emo-Rec point scores.

We visualized the trend for Experiment T was relatively flat compared with that of Experiment N through the *Figure 5*. Differences between the two interventions led to different outcomes. Overall, there was an improvement in participants' mood during long-term RGT. Mood improvement was more pronounced in the robot group than in the experimental group. This shows that the process analysis of different experimental setups can be achieved using this tool.

Yamazaki et al. reported the therapeutic outcomes of telenoid robots in facilitating PWD conversations to improve dementia (Yamazaki et al., 2019). According to Shibata et al., the improvement of the face scale is also considered important for alleviating BPSD in the study of teleoperated robots (Shibata & Wada, 2011). The tool makes the process of emotional improvement more intuitive and convenient. Helps care giver adapt cognitive improvement strategies. However, the facial recognition app did not perform well on the attention metric. In the future, it is expected that eye movement recognition equipment or other sensors can effectively reflect this type of data to provide researchers with more dimensional references.

Conclusions

Based on our data analysis of long-term interventions, we believe that a facial emotion recognition tool is more objective for quantifying participants' emotions, without the need to use other instruments. It is simpler and easier to use than contact-based electroencephalogram (EEG) and blood tests, and can be adjusted to achieve more accurate results by adjusting the feature library of the training datasets. More importantly, it does not suffer from cognitive barriers for self-reports, questionnaires, or interviews.

It has been reported that PWD who have undergone tracheostomy or have other vocal problems can also benefit from this tool. They were able to use expressions and gestures to engage in RGT and achieve some degree of emotional improvement (Liu et al., 2022). We propose further enhancement of such applications with more accurate databases and higher-performance-trained models to aid PWD in the future.

Abbreviations

PWD, People With Dementia; RGT, Reminiscence Group Therapy; BPSD, Behavioral and Psychological Symptoms of Dementia; MMSE, Mini-Mental State Examination; NPI, Neuropsychiatric Inventory; DBD, Disruptive Behavior Disorder; MoCA, Montreal Cognitive Assessment; ADAS-Cog, The Alzheimer's disease assessment scale-cognitive subscale.

Conflicts of interest statement

All authors disclose no relevant relationships.

Acknowledgements

The authors would like to express their gratitude to Mr. Masato Yamazaki for his contribution in the development of Pepper's teleoperation system.

Funding/Support statement

This work was partially supported by JSPS KAKENHI [grant numbers 19K11395, 19H01577].

References

Chao, S. Y., Liu, H. Y., Wu, C. Y., Jin, S. F., Chu, T. L., Huang, T. S., & Clark, M. J. (2006). The effects of group reminiscence therapy on depression, self esteem, and life satisfaction of elderly nursing home residents, Journal Of Nursing Research, 14, 36-45. https://doi.org/10.1097/01.jnr.0000387560.03823.c7.

Cotelli, M., Manenti, R., & Zanetti, O. (2012). Reminiscence therapy in dementia: A review. Maturitas, 72(3), 203-205. https://doi.org/10.1016/j.maturitas.2012.04.008.

Kase, H., Yamazaki, R., Zhu, W., & Nishio, S. (2019). Tele-operated android robot reminiscence group therapy and human coordinated RGT for older adults with dementia: a comparative study. International Journal Of Gerontology, Special issue s51(5), https://doi.org/10.6890/IJGE.201910/SP.0008.

Irazoki, E., Garcia-Casal, J. A., Sanchez-Meca, J., & Franco-Martin, M. (2017). Efficacy of group reminiscence therapy for people with dementia: systematic literature review and meta-analysis. Revue Neurologique, 65, 447–456. https://doi.org/10.33588/rn.6510.2017381.

Rosa-Neto, P. (2021). Journey through the diagnosis of dementia, https://www.alzint.org/resource/world-alzheimer-report-2021/.

Liu, X. Y. (2021). Validation research on the application of depthwise separable convolutional AI facial expression: recognition in non-pharmacological treatment of BPSD. Journal Of Clinical Nursing Research, 5, 31-37. https://doi.org/10.26689/jcnr.v5i4.2325.

Albers, C. J., & Bringmann, L. F. (2020). Gradual and abrupt changes in emotion dynamics with the time-varying change point autoregressive model. European Journal of Psychological Assessment, 36, 492-499. https://doi.org/10.1027/1015-5759/a000589.

Lorish, C. D., & Maisiak, R. (1986). The face scale: a brief, nonverbal method for assessing patient mood. Arthritis Rheum, 29, 906–909. https://doi.org/10.1002/art.1780290714.

Heerink, M., Krose, B., Evers, V, & Wielinga, B. (2010). Assessing acceptance of assistive social agent technology by older adults: the Almere model. International Journal of Social Robotics, 2, 361–375.

Yamazaki, R., Kase, H., Nishio, S., & Ishiguro, H. (2019). A conversational robotic approach to dementia symptoms: measuring its effect on older adults. the 7th International Conference HAI'19, (Sep)110–117. https://doi.org/10.1145/3349537.3351888

Shibata, T., & Wada, K. (2011). Introduction of field test on robot therapy by seal robot, PARO. Journal of the Robotics Society of Japan, 29(3), 246-249. https://doi.org/10.7210/jrsj.29.246

Liu, X. Y., Yamazaki, R., & Kase, H. (2022). How to measure the effect of reminiscence group therapy on mute people with dementia?: A trial using a facial emotion recognition method. Gerontechnology, 21(s), 671.https://doi.org/10.4017/gt.2022.21.s.671.pp4