A Systematic Review on Facial Expression Based Emotion Recognition System for Smart Homes

Nor Azlina Ab. Aziz[1](✉), Tawsif Khan[1], Kamarulzaman Ab. Aziz[2], Muhammad Anas Hasnul[1], and Sharifah Noor Masidayu Sayed Ismail[3]

[1] Faculty of Engineering and Technology, Multimedia University, Melaka, Malaysia
azlina.aziz@mmu.edu.my

[2] Faculty of Business, Multimedia University, Melaka, Malaysia

[3] Faculty of Information, Science and Technology, Multimedia University, Melaka, Malaysia

Abstract. Emotion recognition system (ERS) is a technology that enables machines to recognize human's emotion. The machines use various modalities as input such as facial expression, voice intonation, gait, brain's electroencephalogram (EEG) and heart's electrocardiogram (ECG) to interpret human's current state of emotion. The advancement of learning algorithms supported the growth of this area. The ERS can be adopted in many application areas including smart homes. The incorporation of ERS for smart homes has the potential to help improving occupants' comfort and also quality of life. This paper presents a systematic review that focused on research of facial expression based ERS for smart homes application. The review is limited to works reported over the last decade, from year 2011 till 2021.

Keywords: Facial expression · emotion recognition system · smart homes

1 Introduction

Emotion is defined by the American Psychological Association (APA) as "complex reaction pattern, involving experiential, behavioural, and physiological elements, by which an individual attempt to deal with a personally significant matter or event." Human's decision and action are affected by emotion. Understanding human emotion not only attract the attention of psychologists but it also had attracted the attention of engineers and computer scientist. Engineers and computer scientists work to build emotion recognition system (ERS) for a better human machine interaction. A machine with emotional intelligence can respond to human needs more intelligently and helps to reduce negative emotions so that productivity can be improved.

ERS can be applied in multiple fields including advanced driver assistance system (ADAS) like the personalized driving warning system proposed in [1], the ECG based ERS for driver's emotion and alertness monitoring in [2], and facial expression based ERS to identify the driver's emotion and control the voice of driver assistance system for alertness enhancement in [3]. Emotion recognition enabled system had also been proposed to be applied in education. In [4] the output of an ERS is used to recommend

M. Y. bin Alias (Ed.): MECON 2022, AER 214, pp. 28–37, 2023.
https://doi.org/10.2991/978-94-6463-082-4_5

suitable games including educational games for children with Autism Spectrum Disorder. The ERS uses audio and visual signal for the emotion identification. Smart homes have also benefitted from ERS research. In [5, 6] facial expression based ERS is used to control lighting and electronic devices for the comfort of occupants. An ERS is applied for home care to allow independent living for the elderly while ensuring their safety is proposed in [7]. The system used facial expression images to recognized any signs of distress before alerting the health care provider or next of kin.

Although the applications discussed above are mostly lab scale and prototype, this technology is not far away. During the end of 2019, the Financial Times reported that in the district on Xinjiang, China, security system equipped with emotion recognition had been installed at public places to predict any signs of terrorism such as aggressiveness and nervousness [8]. In Sept 2020, Forbes reported that NtechLab from Russia is going to introduce a similar security system that scans human faces to detect their emotions and any sign of aggressive behaviour [9]. Smyle, a United Kingdom based company launched their "Return of Emotion" solution in Oct 2020 [10]. The "Return of Emotion" is an ERS based on facial expression images, pulse sensors and neurological signal. The solution aims to measure the genuine emotional impact of event attendees.

From the works discussed above, it could be seen that an ERS can be build using various modalities. Among the frequently used modality is through images such as the facial images [11, 12], thermal images [13] and movement images [14, 15]. Facial expression, posture and movements are frequently affected by our feeling. In fact, facial expression is how we "read" the emotion of another human being. Therefore, this had motivated researchers to use these images in development of ERS. Our voice intonation can also be used as a modality of ERS [16]. A person voice pitch might be increased when feeling excited or lowered when feeling down. Other than these physical modalities, physiological signals can also be used for ERS. Among physiological signals commonly used are electroencephalogram (EEG) [17] and electrocardiogram (ECG) [17–20]. Some works combine multiple modalities for their ERS [21, 22].

This study focuses on the application of facial expression based ERS for smart homes. Facial expression based ERS has the advantage of non-intrusive approach [7]. The camera can be positioned in the smart home environment to monitor the occupants without the need for the occupants to be fitted to special devices and sensors which are costly and could cause inconveniences. This modality is also a cost-effective option for monitoring a group of people as a single camera can be used to monitor multiple individuals. The study is restricted to works reported from 2011 till 2021.

The research method is discussed in Sect. 2. In Sect. 3 the results are presented and discussed and finally this paper is concluded in Sect. 4.

2 Research Methodology

A systematic literature review was conducted. Relevant articles are collected from the IEEE Xplore, Science Direct and Springer databases. The search was limited to publications since 2011 to 2021. Only open access or subscribed publications are considered. Conference proceedings and review papers were excluded and only English language articles are selected. The search returned 132 articles, 63 are from Springer database, 51 from Science Direct and 11 from IEEE Xplore.

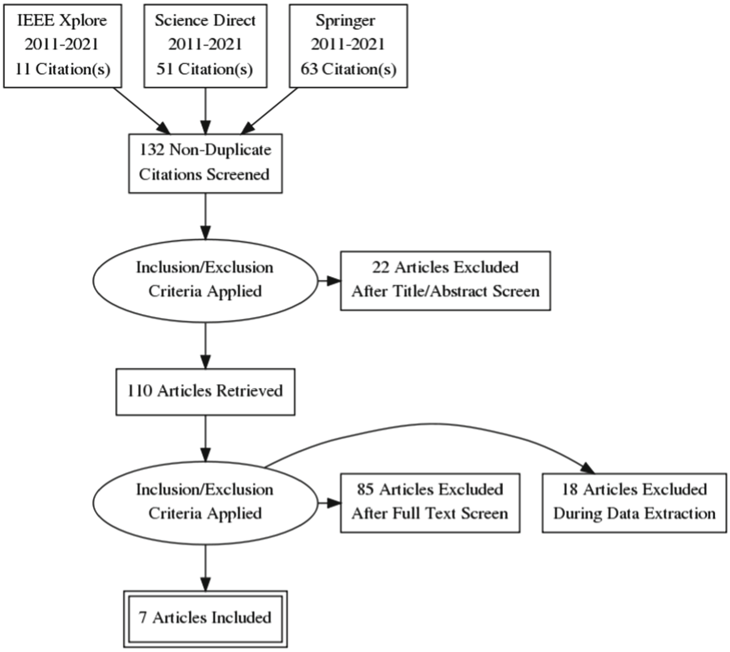

Fig. 1. Flowchart of the Systematic Review

These articles were carefully screened and read. In the first screening the articles with title and abstract that did not match the inclusion criteria or match the exclusion criteria are removed leaving 110 articles. The inclusion and exclusion criteria are as follows;

- The studies must include experimental works, studies that reported reviews of the technology is not included.
- The studies must focus on development of ERS focusing on its application or potential application for smart homes.
- The studies must use facial expression as the modality or part of modalities of the developed ERS.

Additionally, conference papers are excluded, only peer reviewed journal papers are included in this study. The language of the articles must be English and we only focuses on works reported from 2011 onwards. At the end of the screening process only 7 articles were short listed. Figure 1 shows the flowchart of the systematic review conducted.

3 Results

The 7 shortlisted articles are reviewed in detailed, focusing on the specifications of the data used, classifier and the accuracy achieved as well as the applications for smart homes. The findings are summarised and tabulated in Table 1.

Table 1. Selected Articles of Facial Expression Based Emotion Recognition Systems for Smart Homes

Author, Year	Data				Classification		Smart Home	
	Dataset	Subjects	Classes	Additional Information	Classifier	Performance	Area	Details
Yu, 2012	Own, 180 images	4 person	happy, angry, sad, surprised, scared, disgusted and neutral	640 x 480 pixels with 24-bit color intensity encoding.	back-propagation neural network (BPNN)	Happy: 22/30 Angry: 23/30 Sad: 12/30 Surprise: 14/30 Scared: 9/30 Disgusted: 10/30	Furniture – magic mirror table	The mirror is equipped with camera, when negative emotion is d etected positive message will be given by the mirror via text and audio additionally suitable background music will be played to alleviate the mood. Heart rate is used to sense the alleviated mood.
Han, 2015	FERET, 6x10images	10 subjects	happy, angry, sad, surprised, scared, disgusted		LGBP+KNN	Happy: 91.62% Angry: 82.26% Sad: 85.12% Surprise: 89.74% Scared: 79.50% Disgusted: 84.23%	Eldercare robot	Test were conducted involving 15 elders. The robot react to occupant's emotion and controls the smart home environment to improve the emotion.
Rincon, 2016	KDEF Emotional Lab at Karolinska Institute, 4900 images		happy, angry, sad, surprised, scared, disgusted and neutral	90% were used for training, 10% for testing 562×762 pixel	ANN	93.5%	Surrounding	Emotional state of a group of people is used for music selection
Mano, 2020	Training: Radboud Faces (RaFD), 67x7 images Extended Cohn–Kanade (CK+), 169x7 images	Training: 67 American and Moroccan (adults and children) 123 Europeans or Americans, African–Americans and other	happy, angry, sad, surprised, scared, disgusted and neutral	Training: 3115 images, 445 images for each emotion group	Ensemble: kNN, SVM, fuzzy Logic, decision and Bayesian networks with GA to determine weight of classifier	Training: Happy: 426/445 Angry: 312/445 Sad: 316/445 Surprise: 423/445 Scared: 340/445 Disgusted: 366/445 Neutral: 326/445	Surrounding	The emotion recognition system is used to select suitable music to lift up the mood.

(*continued*)

Table 1. (*continued*)

		ethnic groups (adults)						
	IMPA-FACE3D, 38x7 images	38 (adults)						
	FACES, 171x7 images	171 (adults)						
	Testing: Own, 30x7images	30 (adults)		Testing: 210 images, 30 images for each emotion group		Testing: Happy: 30/30 Angry: 19/30 Sad: 22/30 Surprise: 29/30 Scared: 21/30 Disgusted: 21/30 Neutral: 25/30		
Yaddaden, 2018	JAFFE, 207 images	JApanese Female Facial Expression participants with different ages and ethnicities	happy, angry, sad, surprised, scared, disgusted and neutral	256x256	CNN with 10 fold cross validation	JAFFE(95.43%) Happy: 99.32% Angry: 91.83% Sad: 94.83% Surprise: 94.83% Scared: 95.17% Disgusted: 93.04% Neutral: 99%	Assisted living	Error detection of actions, to help and allow independent living among elderly
	RaFD, 469 images			681x1024		RaFD(97.66%) Happy: 99.78% Angry: 97.91% Sad: 95.45% Surprise: 99.63% Scared: 95.82% Disgusted: 99.18% Neutral: 95.82%		
	KDEF, 980 images			562x762		KDEF(90.61%) Happy: 98.70% Angry: 85.04% Sad: 89.86% Surprise: 92.86% Scared: 83.63% Disgusted: 89.61% Neutral: 94.54%		
	MMI, 12484			768x576		MMI(79.40%)		

(*continued*)

Table 1. (*continued*)

	images					Happy: 94.27% Angry: 85.87% Sad: 89.11% Surprise: 92.43% Scared: 29.90% Disgusted: 82.03% Neutral: 82.21%		
	CK+, 3095 images			640x490		CK+(95.38%) Happy: 99.88% Angry: 93.78% Sad: 92.53% Surprise: 99.18% Scared: 96.72% Disgusted: 96.26% Neutral: 89.30%		
Yang, 2020	Own	3 person	happy, angry, sad, surprised, scared, disgusted and neutral		Deep Neural network	Happy: 33.33% Angry: 0% Sad: 33.33% Surprise: 0% Scared: 2.13% Disgusted: 0% Neutral: 31.17%	Entertainment system	Movie recommendation is given based on user's emotion
Mano, 2016	Training: CK+ Testing: Own, 20x100image s	123 actors 20 subjects	happy, angry, sad, surprised, scared and neutral	Only subset of the images is used. Similar or almost identical images are removed.	Ensemble classifiers: kNN, Decision Tree, Fuzzy Logic, Bayesian Networks and SVM	0.7-0.85	In home healthcare	The system aim to allow safe home recovery and independent living especially among elderly. The emotion recognised can be used to alert caregiver of doctor if an event is detected

From the review it can be seen that many researchers developed the system using combination of public dataset and their own dataset. Only two out of seven fully used their own data to train and test the system [23, 24]. Mano et al. in two separate works [7, 25], used public datasets for training of their systems and then tested the built model using their own collected dataset. Meanwhile [26, 27] and [28] fully utilized public datasets for both training and testing. All seven works used discrete multiclass data. Where the facial expressions are classed to seven discrete emotions of happy, angry, sad, surprised, scared, disgusted and neutral. The wide availability of public datasets from various cultures and origins facilitates research in this field.

Variety of artificial intelligence algorithms from machine learning and deep learning categories were applied for the ERS development. The performance is presented either in term of accuracy or the true positive rate. Most of the works reported class based true positive rate. Accuracy and true positive rate are measured differently. The accuracy is calculated using the ratio of all trues (true positive, TP and true negative, TN) over total number of data including the false positive, FP and false negative, FN, as shown in Eq. 1.

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

Meanwhile, the true positive rate which is also known as recall or sensitivity is only measured using data that are truly predicted as true vs the actual number of data that are true, Eq. 2.

$$TP\ rate = \frac{TP}{TP + FN} \tag{2}$$

Both of these evaluation metrics are commonly used to represent performance of learning algorithms. For class imbalance data, accuracy is not sufficient to give an unbiased overview of the performance. This can be supported by the true positive rate that gives overview on rate of correctly classified relevant data. The reported accuracy and true positive rate of the reviewed works ranges broadly from very excellent, 100% to very poor, 0%. This shows that there's room of improvement in this area of research.

The applications for smart homes can be broadly categorized to entertainment advisory system and assisted living. In entertainment advisory system the ERS is used for music and movies recommendation to suit the current state of emotion [24, 25, 27]. A magic mirror that display positive messages and choose suitable background music for mood alleviation is proposed in [23]. In [7, 26] and [28], the facial expression ERS are used for assisted living, enabling independent living among elderly and recovering patients. The success of these works shows the integration of ERS with smart homes has many potential and motivates more works to be conducted in this area.

4 Conclusion

This work systematically reviewed works based on facial expression for ERS that were applied for smart homes. The review was limited to works reported from 2011 to 2021 and extracted from three databases, IEEE Xplore, Science direct and Springer. The results show that research in this area is supported by variety of public datasets from, European, American, Asian, African and Middle Eastern. Nonetheless, researchers also worked on their own collected data. Based on the performance of the ERS reported, particularly the accuracy, it can be seen that there's plenty of room for improvement to create a robust ERS. Meanwhile, smart homes can definitely be enhanced with the application of ERS to improve the comfort and quality of life of the occupants.

Acknowledgments. This project is funded by TM Research & Development Grant (RDTC/190988) which is awarded to the Multimedia University.

Authors' Contributions. The work was designed and conducted by N.A.A. Aziz, K.A. Aziz, T. Khan, M.A. Hasnul, and S.N.M.S. Ismail. Each author contributed to data analysis, interpreting the results, reviewing, and revising the research and manuscript.

References

1. X. Wang, Y. Guo, J. Ban, Q. Xu, C. Bai, and S. Liu, "Driver emotion recognition of multiple-ECG feature fusion based on BP network and D–S evidence," *IET Intell. Transp. Syst.*, vol. 14, no. 8, pp. 815–824, Aug. 2020.
2. K. N. Minhad, S. H. Ali, M. Bin, and I. Reaz, "Happy-anger emotions classifications from electrocardiogram signal for automobile driving safety and awareness," *J. Transp. Heal.*, vol. 7, no. November, pp. 75–89, 2017.
3. S. M. Sarala, D. H. Sharath Yadav, and A. Ansari, "Emotionally adaptive driver voice alert system for advanced driver assistance system (ADAS) applications," *Proc. Int. Conf. Smart Syst. Inven. Technol. ICSSIT 2018*, pp. 509–512, Dec. 2018.
4. N. Heni and H. Hamam, "Design of emotional educational system mobile games for autistic children," *2nd Int. Conf. Adv. Technol. Signal Image Process. ATSIP 2016*, pp. 631–637, Jul. 2016.
5. L. T. Boon, M. H. Husin, Z. F. Zaaba, and M. A. Osman, "Implementation of an Automated SmartHome Control for Detecting Human Emotions via Facial Detection," in *Proceedings of the 5th International Conference on Computing & Informatics*, no. 065, 2015, pp. 39–45.
6. L. T. Boon, M. H. Husin, Z. F. Zaaba, and M. A. Osman, "Eye 2H: A proposed automated smart home control system for detecting human emotions through facial detection," in *2014 the 5th International Conference on Information and Communication Technology for the Muslim World, ICT4M 2014*, 2014.
7. L. Y. Mano *et al.*, "Exploiting IoT technologies for enhancing Health Smart Homes through patient identification and emotion recognition," *Comput. Commun.*, vol. 89–90, pp. 178–190, Sep. 2016.
8. S.-L. Wong and Q. Liu, "Emotion recognition is China's new surveillance craze," *Financial Times*, Nov-2019.
9. T. Brewster, "This Russian Facial Recognition Startup Plans To Take Its 'Aggression Detection' Tech Global With $15 Million Backing From Sovereign Wealth Funds," *Forbes*, Sep-2020.
10. "Smyle Launches Event Attendee Emotion Analysis," *Daily Research News Online*, Oct-2020.
11. A. Cruz, B. Bhanu, and N. Thakoor, "Facial emotion recognition in continuous video," *Proc. - Int. Conf. Pattern Recognit.*, no. 21st, pp. 1880–1883, 2012.
12. E. Pranav, S. Kamal, C. Satheesh Chandran, and M. H. Supriya, "Facial Emotion Recognition Using Deep Convolutional Neural Network," *2020 6th Int. Conf. Adv. Comput. Commun. Syst. ICACCS 2020*, pp. 317–320, Mar. 2020.

13. C. Goulart, C. Valadão, D. Delisle-Rodriguez, E. Caldeira, and T. Bastos, "Emotion analysis in children through facial emissivity of infrared thermal imaging," *PLoS One*, vol. 14, no. 3, pp. 1–17, 2019.
14. N. A. Makhdoomi, T. S. Gunawan, and M. H. Habaebi, "Human gait recognition and classification using similarity index for various conditions," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 53, no. 1, 2013.
15. N. Jamil, N. H. M. Khir, M. Ismail, and F. H. A. Razak, "Gait-Based Emotion Detection of Children with Autism Spectrum Disorders: A Preliminary Investigation," *Procedia Comput. Sci.*, vol. 76, pp. 342–348, 2015.
16. H. R. Ramya and M. R. Bhatt, "Personalised Emotion Recognition Utilising Speech Signal and Linguistic Cues," *2019 11th Int. Conf. Commun. Syst. Networks, COMSNETS 2019*, pp. 856–860, May 2019.
17. S. Katsigiannis and N. Ramzan, "DREAMER: A Database for Emotion Recognition Through EEG and ECG Signals from Wireless Low-cost Off-the-Shelf Devices," *IEEE J. Biomed. Heal. Informatics*, vol. 22, no. 1, pp. 98–107, 2018.
18. S. Tivatansakul and M. Ohkura, "Emotion Recognition using ECG Signals with Local Pattern Description Methods," *Int. J. Affect. Eng.*, vol. 15, no. 2, pp. 51–61, 2015.
19. F. Agrafioti, D. Hatzinakos, and A. K. Anderson, "ECG pattern analysis for emotion detection," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 102–115, 2012.
20. D. Kumar, H. K. Gupta, and R. Shrivastava, "A Review of ECG signals for Human Emotion Detection," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 5, no. 7, pp. 72–76, 2017.
21. M. L. Barron-Estrada, R. Zatarain-Cabada, and C. G. Aispuro-Gallegos, "Multimodal recognition of emotions with application to mobile learning," *Proc. - IEEE 18th Int. Conf. Adv. Learn. Technol. ICALT 2018*, pp. 416–418, Aug. 2018.
22. Y. Dai, X. Wang, P. Zhang, and W. Zhang, "Wearable biosensor network enabled multimodal daily-life emotion recognition employing reputation-driven imbalanced fuzzy classification," *Measurement*, vol. 109, pp. 408–424, 2017.
23. Y. C. Yu, S. D. You, and D. R. Tsai, "Magic mirror table for social-emotion alleviation in the smart home," *IEEE Trans. Consum. Electron.*, vol. 58, no. 1, pp. 126–131, Feb. 2012.
24. J. Yang, R. Wang, X. Guan, M. M. Hassan, A. Almogren, and A. Alsanad, "AI-enabled emotion-aware robot: The fusion of smart clothing, edge clouds and robotics," *Futur. Gener. Comput. Syst.*, vol. 102, pp. 701–709, Jan. 2020.
25. L. Y. Mano *et al.*, "An intelligent and generic approach for detecting human emotions: a case study with facial expressions," *Soft Comput.*, vol. 24, no. 11, pp. 8467–8479, Jun. 2020.
26. Y. Yaddaden, M. Adda, A. Bouzouane, S. Gaboury, and B. Bouchard, "User action and facial expression recognition for error detection system in an ambient assisted environment," *Expert Syst. Appl.*, vol. 112, pp. 173–189, Dec. 2018.
27. J. A. Rincon, V. Julian, and C. Carrascosa, "Developing an emotional-based application for human-agent societies," *Soft Comput.*, vol. 20, no. 11, pp. 4217–4228, 2016.
28. H. Jing, X. Lun, L. Dan, H. Zhijie, and W. Zhiliang, "Cognitive emotion model for eldercare robot in smart home," *China Commun.*, vol. 12, no. 4, pp. 32–41, Apr. 2015.