ANALYSIS OF PERFORMANCE BETWEEN KINECT V1 AND KINECT V2 FOR VARIOUS FACIAL PART MOVEMENTS USING SUPERVISED MACHINE LEARNING TECHNIQUES

HENG SHENG GUANG

MASTER OF SCIENCE

UNIVERSITI MALAYSIA PAHANG

SUPERVISOR'S DECLARATION

I hereby declare that I have checked this thesis and in my opinion, this thesis is adequate in terms of scope and quality for the award of the degree of Master of Science.

(Supervisor's Signature)

Full Name : Dr. Rosdiyana Binti Samad

Position : Senior Lecturer

Date : 26/10/2022

STUDENT'S DECLARATION

I hereby declare that the work in this thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Universiti Malaysia Pahang or any other institutions.

(Student's Signature)

Full Name : Heng Sheng Guang

ID Number : MEG17001

Date : 12 March 2022

ANALYSIS OF PERFORMANCE BETWEEN KINECT V1 AND KINECT V2 FOR VARIOUS FACIAL PART MOVEMENTS USING SUPERVISED MACHINE LEARNING TECHNIQUES

HENG SHENG GUANG

Thesis submitted in fulfillment of the requirements
for the award of the degree of
Master of Science

Faculty of Electrical and Electronics Engineering Technology

UNIVERSITI MALAYSIA PAHANG

DECEMBER 2022

ACKNOWLEDGEMENTS

I would like to acknowledge and extend my heartfelt gratitude to my supervisor, Dr. Rosdiyana Binti Samad for her constant encouragement in completing this research. This research would not come to an end without her expert guidance. The precious thought provided from my supervisor is the key to across the difficult faced in the middle part of the study. Her valuable discussions and academic suggestions enriched my knowledge and understanding.

Besides, I am so thankful for the funding of Fundamental Research Grant Scheme FRGS/1/2016/TK04/UMP/02/1 that provides me financial support in completing my research.

Furthermore, I would specially thank the members of my family who assisted me in this research. Many thanks to their support during my hard time working on this study. Besides, I am grateful to my friends who helped me in the sample data collection during the analysis phase. This study would not be succeeded without their enthusiastic assistance.

It is difficult to list out the names of all people involved in the completion of this research. Much appreciated to everyone who provided me academic concepts and emotional support.

ABSTRAK

Senaman untuk wajah adalah siri pergerakan bahagian wajah seperti membesarkan dan mengecilkan zon muka bahagian atas, tengah dan bawah supaya nampak lebih muda dan meremajakan otot muka. Sebelum ini, keberkesanan senaman wajah untuk tujuan pemulihan dan peremajaan masih menjadi kontroversi kerana kekurangan kajian kuantitatif. Walau bagaimanapun, kajian terbaru telah membuktikan keberkesanan senaman wajah melalui alat bantu dan bimbingan daripada ahli fisioterapi. Proses pemulihan dan peremajaan melalui senaman wajah adalah membosankan dan memakan masa. Kekurangan motivasi dan kesabaran adalah punca utama kegagalan rawatan. Oleh itu, pengalakkan dan rawatan senaman wajah yang menarik adalah mustahak untuk meningkatkan tahap keberkesanan. Tujuan kajian ini adalah untuk menganalisis prestasi sensor pengerakan Kinect versi 1 dan 2 untuk pembangunan aplikasi senaman wajah berdasarkan analisis prestasi pengesanan wajah. Kamera 2D biasa kekurangan dalam maklumat kedalaman (depth). Oleh itu, ia akan mengakibatkan pengesanan titik wajah yang tidak tepat. Dalam kajian ini, kamera 3D seperti sensor pengerakan Kinect versi 1 (v1) dan versi 2 (v2) digunakan untuk menggantikan kamera 2D biasa supaya dapat mengesan titik wajah dengan lebih tepat. Kedua-dua sensor Kinect digunakan untuk menggunakan kaedah Model Penampilan Aktif (AAM) untuk mengesan dan mengekstrak ciri muka. Kemudian, pelbagai kaedah klasifikasi klasik seperti rangkaian saraf, pokok keputusan kompleks, mesin vector sokongan padu (SVM), SVM Gaussian halus, jiran k-terdekat (kNN), dan analisis diskriminasi kuadratik digunakan untuk menganalisis ketepatan pengesanan kedua-dua sensor pengerakan Kinect. Setelah itu, versi sensor Kinect yang mempunyai prestasi lebih baik dipilih untuk aplikasi senaman wajah. Set data pengujian dengan ketepatan tertinggi dianalisis untuk membina sistem berasaskan peraturan untuk aplikasi senaman wajah. Akhirnya, Kinect v2 telah mengungguli hampir setiap tugas membandingkan dengan Kinect v1 kecuali senaman angkat kening. Selanjutnya, senaman dan ujian ketepatan data wajah yang diperoleh dengan menggunakan Kinect v2 adalah jauh lebih tinggi daripada Kinect v1. Klasifikasi kNN adalah kaedah yang paling sesuai antara kaedah klasifikasi yang digunakan kerana ia menandakan ketepatan ujian dan senaman yang memuaskan untuk Kinect v2, iaitu 97.8% dan 94.3%. Oleh itu, set data yang diramalkan oleh kaedah klasifikasi kNN dengan betul digunakan untuk analisis plot kotak untuk mendapatkan parameter ambang, iaitu min, nilai kuartil bawah dan nilai kuartil atas pergerakan bahagian muka. Kemudian, parameter ambang digunakan untuk membina sistem berasaskan peraturan untuk aplikasi senaman wajah. Aplikasi senaman wajah dilengkapi dengan sistem penggredan House-Brackmann yang memberikan penilaian skor dan gred setelah menyelesaikan setiap set senaman wajah yang dipilih. Kesimpulannya, aplikasi senaman muka berasaskan Kinect yang dibangunkan berfungsi dengan jayanya dan mampu memberi maklum balas pemarkahan kepada pengguna. Aplikasi ini akan dapat memberi motivasi dan menggalakkan pengguna untuk melaksanakan senaman wajah di rumah.

ABSTRACT

Facial exercises are a series of facial movements such as exaggeration and deflation in upper, middle, and lower face zone for promoting youth and rejuvenating facial muscles. Previously, the effectiveness of facial exercises for rehabilitation and rejuvenation purposes is still controversial due to lack of quantitative study. However, recent studies have proven the efficacy of the facial exercises through assistive device and guidance from physiotherapist. The process of rehabilitation and rejuvenation through facial exercises is uninteresting and time consuming. Frequently, lack of motivation and patient are the main reasons for treatment failure. Hence, interesting course of facial exercise treatment and encouragement are essential to increase the success rate and effectiveness. The aim of this study is to analyse the performance of Kinect motion sensor version 1 and 2 for development of facial exercise application based on the analysis of face tracking performance. Common 2D cameras are lack of depth information, hence will result in inaccurate facial point detection. In this study, 3D cameras such as Kinect motion sensors version 1 (v1) and version 2 (v2) are used instead of ordinary 2D camera for more accurate facial points detection. Both Kinect sensors are used to apply the Active Appearance Models (AAM) method to detect and extract the facial features. Then, various classic classification methods such as neural network, complex decision tree, cubic support vector machine (SVM), fine Gaussian SVM, fine k-nearest neighbours (kNN), and quadratic discriminant analysis are applied to analyse the detection accuracy of both Kinect motion sensors. After that, the version of Kinect sensor which has better performance is adopted for facial exercises application. The dataset with highest testing accuracy is analysed for constructing rule-based system for the facial exercises application. Eventually, Kinect v2 has outperformed in almost every task by comparing to Kinect v1 except the raising eyebrows exercise. Furthermore, the training and testing accuracies of face data acquired by using Kinect v2 are significantly higher than Kinect v1. The kNN classification is the most suitable method among the applied classification methods as it marks the satisfying training and testing accuracy for both Kinect v2, which are 97.8% and 94.3% respectively. Hence, the correctly predicted dataset using kNN classification method is used for box plot analysis to obtain the threshold parameters, which are mean, lower quartile values, and upper quartile values of the facial part movements. Then, the threshold parameters are used to construct a rule-based system for the real-time facial exercises application. The facial exercises application features with House-Brackmann grading system which gives evaluation of score and grade after completing each set of selected facial exercises. In conclusion, the developed Kinect-based facial exercise application performs successfully and able to give scoring feedback to the user. The application will be able to motivate and encourage users for home facial exercise purpose.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

REFERENCES

Abate, A. F., Nappi, M., Riccio, D., & Sabatino, G. (2007). 2D and 3D face recognition: A survey. *Pattern Recognition Letters, 28*(14), 1885-1906. doi:10.1016/j.patrec.2006.12.018

Assegie, T. A., & Nair, P. S. (2019). Handwritten digits recognition with decision tree classification: a machine learning approach. *International Journal of Electrical Computer Engineering, 9*(5), 4446-4451.

Ayoobkhan, M. U. A., Chikkannan, E., & Ramakrishnan, K. (2018). Feed-forward neural network-based predictive image coding for medical image compression. *Arabian Journal for Science, 43*(8), 4239-4247.

Beale, M. H., Hagan, M. T., & Demuth, H. B. (2010). Neural network toolbox user's guide. *The MathWorks*, 77-81.

Bernard, J. A. (1988). Use of a rule-based system for process control. *IEEE Control Systems Magazine, 8*(5), 3-13.

Berwick, R. (2003). An Idiot's guide to Support vector machines (SVMs). *Retrieved on October, 21*, 2011.

Betke, M., Gips, J., & Fleming, P. (2002). The camera mouse: visual tracking of body features to provide computer access for people with severe disabilities. *IEEE Transactions on neural systems, 10*(1), 1-10.

Brethes, L., Menezes, P., Lerasle, F., & Hayet, J. (2004). *Face tracking and hand gesture recognition for human-robot interaction.* Paper presented at the IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004.

Brown, C. (1992). Assistive technology computers and persons with disabilities. *Communications of the ACM, 35*(5), 36-45.

Bryant, D. P., & Bryant, B. R. (2003). *Assistive technology for people with disabilities*: Allyn and Bacon Boston.

Carvelho, T., Allison, R. S., Irving, E. L., & Herriot, C. (2008). *Computer gaming for vision therapy*: IEEE.

Červená, L., Kříž, P., Kohout, J., Vejvar, M., Verešpejová, L., Štícha, K., . . . Mareš, J. (2021). Advanced Statistical Analysis of 3D Kinect Data: A Comparison of the Classification Methods. *Applied Sciences, 11*(10), 4572.

Charbuty, B., & Abdulazeez, A. (2021). Classification based on decision tree algorithm for machine learning. *2*(01), 20-28.

Chee, G. H., & Nedzelski, J. M. (2000). Facial nerve grading systems. *Facial plastic surgery, 16*(04), 315-324.

Cintra, M. E., Monard, M. C., & de Arruda Camargo, H. (2010). *An evaluation of rule-based classification models induced by a fuzzy method and two classic learning algorithms.* Paper presented at the 2010 Eleventh Brazilian Symposium on Neural Networks.

Cootes, T. F., Edwards, G. J., & Taylor, C. J. (2001). Active appearance models. *IEEE Transactions on pattern analysis machine intelligence, 23*(6), 681-685.

Cootes, T. F., & Taylor, C. J. (2004). Statistical models of appearance for computer vision. In: Technical report, University of Manchester.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning, 20*(3), 273-297.

Croxson, G., May, M., & Mester, S. J. (1990). Grading facial nerve function: House-Brackmann versus Burres-Fisch methods. *The American journal of otology, 11*(4), 240-246.

Demuth, H., Beale, M., & Hagan, M. (1994). *Neural network toolbox*: Mathworks.

Edwards, G. J., Cootes, T. F., & Taylor, C. J. (1998). *Face recognition using active appearance models.* Paper presented at the European conference on computer vision.

Esty, W. W., & Banfield, J. D. (2003). The box-percentile plot. *Journal of Statistical Software, 8*(17), 1-14.

Gareth, J., Daniela, W., Trevor, H., & Robert, T. (2013). *An introduction to statistical learning: with applications in R*: Spinger.

Gips, A., DiMattia, P. A., & Gips, J. (2004). *The effect of assistive technology on educational costs: Two case studies.* Paper presented at the International Conference on Computers for Handicapped Persons.

Grosan, C., & Abraham, A. (2011). *Intelligent systems*: Springer.

Gu, H., Su, G., & Du, C. (2003). Feature points extraction from faces. *Image vision computing NZ, 26*, 154-158.

Gunawan, A. A. (2015). Face expression detection on Kinect using active appearance model and fuzzy logic. *Procedia Computer Science, 59*, 268-274.

Guo, G., Wang, H., Bell, D., Bi, Y., & Greer, K. (2003). *KNN model-based approach in classification.* Paper presented at the OTM Confederated International Conferences" On the Move to Meaningful Internet Systems".

Hansen, L. K., & Salamon, P. (1990). Neural network ensembles. *IEEE Transactions on pattern analysis machine intelligence, 12*(10), 993-1001.

Heng, S. G., Samad, R., Mustafa, M., Abdullah, N. R. H., & Pebrianti, D. (2019). *Analysis of performance between Kinect v1 and Kinect v2 for various facial part movements.* Paper presented at the 2019 ieee 9th international conference on system engineering and technology (icset).

Heng, S. G., Samad, R., Mustafa, M., Zain, Z. M., Abdullah, N. R. H., & Pebrianti, D. (2021). *Classification of Facial Part Movement Acquired from Kinect V1 and Kinect V2.* Paper presented at the Proceedings of the 11th National Technical Seminar on Unmanned System Technology 2019.

Hoda, M., Hoda, Y., Hage, A., Alelaiwi, A., & El Saddik, A. (2015). Cloud-based rehabilitation and recovery prediction system for stroke patients. *Cluster Computing, 18*(2), 803-815.

House, J. W. (1983). Facial nerve grading systems. *The laryngoscope, 93*(8), 1056-1069.

Huang, C. L., & Chen, C. W. (1992). Human facial feature extraction for face interpretation and recognition. *Pattern Recognition Letters, 25*(12), 1435-1444.

Huang, S., Cai, N., Pacheco, P. P., Narrandes, S., Wang, Y., & Xu, W. (2018). Applications of support vector machine (SVM) learning in cancer genomics. *Cancer genomics proteomics, 15*(1), 41-51.

Islam, M. R., Kamal, A. R. M., Sultana, N., Islam, R., & Moni, M. A. (2018). *Detecting depression using k-nearest neighbors (knn) classification technique.* Paper presented at the 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2).

Kanade, T., Cohn, J. F., & Tian, Y. (2000). *Comprehensive database for facial expression analysis.* Paper presented at the Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580).

Kayri, M. (2016). Predictive abilities of bayesian regularization and Levenberg–Marquardt algorithms in artificial neural networks: a comparative empirical study on social data. *Mathematical Computational Applications, 21*(2), 20.

Kayri, M., & Kayri, İ. (2015). The comparison of Gini and Twoing algorithms in terms of predictive ability and misclassification cost in data mining: an empirical

study. *International Journal of Computer Trends and Technology (IJCTT), 27*, 10.

Khella, R. (2017). Rule based system for chest pain in infants and children. *International Journal of Engineering Information Systems, 1*(4), 138-148.

Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering, 160*(1), 3-24.

Krishna, S., Little, G., Black, J., & Panchanathan, S. (2005). *A wearable face recognition system for individuals with visual impairments.* Paper presented at the Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility.

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (Vol. 26): Springer.

Kwon, S. K. (2019). Face recognition using depth and infrared pictures. *Nonlinear Theory and Its Applications, IEICE, 10*(1), 2-15.

Lacava, P. G., Golan, O., Baron-Cohen, S., & Smith Myles, B. (2007). Using assistive technology to teach emotion recognition to students with Asperger syndrome: A pilot study. *Remedial Special Education, 28*(3), 174-181.

Li, S. Z., Zhang, H., Yan, S., & Cheng, Q. (2002). *Multi-View Face Alignment Using Direct Appearance Models.* Paper presented at the FGR.

Li, T., Zhou, J., Tuya, N., Du, C., Chen, Z., & Liu, S. (2017). *Recognize facial expression using active appearance model and neural network.* Paper presented at the 2017 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC).

Li, T., Zhu, S., & Ogihara, M. (2006). Using discriminant analysis for multi-class classification: an experimental investigation. *Knowledge information systems, 10*(4), 453-472.

Liu, H., Gegov, A., & Cocea, M. (2016). Rule-based systems: a granular computing perspective. *Granular Computing, 1*(4), 259-274.

Lucey, S., Matthews, I., Hu, C., Ambadar, Z., De la Torre, F., & Cohn, J. (2006). *AAM derived face representations for robust facial action recognition.* Paper presented at the 7th International Conference on Automatic Face and Gesture Recognition (FGR06).

McGill, R., Tukey, J. W., & Larsen, W. A. (1978). Variations of box plots. *The American Statistician, 32*(1), 12-16.

Miskelly, F. G. (2001). Assistive technology in elderly care. *Age ageing, 30*(6), 455-458.

Mothes, O., & Denzler, J. (2017). *Anatomical Landmark Tracking by One-shot Learned Priors for Augmented Active Appearance Models.* Paper presented at the VISIGRAPP (6: VISAPP).

Mousavi, S. M. H. (2021). A Spatial And Frequency Based Method For Micro Facial Expressions Recognition Using Color And Depth Images. *Journal of Software Engineering & Intelligent Systems*, 6(1), 17.

Neely, J. G., Cherian, N. G., Dickerson, C. B., & Nedzelski, J. M. (2010). Sunnybrook facial grading system: reliability and criteria for grading. *The laryngoscope, 120*(5), 1038-1045.

Nixon, M., & Aguado, A. (2019). *Feature extraction and image processing for computer vision*: Academic press.

Ojeda Noda, G. (2016). *Kinect driven facial animation.* Universitat Politècnica de Catalunya,

Pavlakos, G., Choutas, V., Ghorbani, N., Bolkart, T., Osman, A. A., Tzionas, D., & Black, M. J. (2019). *Expressive body capture: 3d hands, face, and body from a single image.* Paper presented at the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition.

Potter, K., Hagen, H., Kerren, A., & Dannenmann, P. (2006). Methods for presenting statistical information: The box plot. *Visualization of large unstructured data sets, 4*, 97-106.

Prabasheela, B., Sakithya, V., Nandhini, V., & Logeshwari, M. (2017). Understanding Bell's palsy–a review. *Pharm. biol. eval, 4*(3), 130-134.

Pterneas, V. (2015). How to use Kinect HD FACE. *Vangos Pterneas*. https://pterneas.com/2015/06/06/kinect-hd-face

Quinlan, J. R. (1986). Induction of decision trees. *Machine learning, 1*(1), 81-106.

Ronzhin, A., & Karpov, A. (2005). *Assistive multimodal system based on speech recognition and head tracking.* Paper presented at the 2005 13th European Signal Processing Conference.

Ross, B. G., Fradet, G., & Nedzelski, J. M. (1996). Development of a sensitive clinical facial grading system. *Otolaryngology—Head Neck Surgery, 114*(3), 380-386.

Shih, Y. S. (1999). Families of splitting criteria for classification trees. *Statistics Computing, 9*(4), 309-315.

Siqueira, L. F., Júnior, R. F. A., de Araújo, A. A., Morais, C. L., & Lima, K. M. (2017). LDA vs. QDA for FT-MIR prostate cancer tissue classification. *Chemometrics Intelligent Laboratory Systems, 162*, 123-129.

Tran, H. T., Na, I. S., Kim, Y. C., & Kim, S. H. (2017). Human Face Tracking and Modeling using Active Appearance Model with Motion Estimation. *Smart Media Journal, 6*(3), 49-56.

Umirzakova, S., & Whangbo, T. K. (2018). *Study On Detect Stroke Symptoms Using Face Features.* Paper presented at the 2018 International Conference on Information and Communication Technology Convergence (ICTC).

Wang, Q., & Yu, Y. (2015). AAM Based Facial Feature Tracking with Kinect. *Cybernetics Information Technologies, 15*(3), 127-139.

Wasenmüller, O., & Stricker, D. (2016). *Comparison of kinect v1 and v2 depth images in terms of accuracy and precision.* Paper presented at the Asian Conference on Computer Vision.

Webb, J., & Ashley, J. (2012). *Beginning Kinect Programming with the Microsoft Kinect SDK*: Apress.

Wu, H.-H., & Bainbridge-Smith, A. (2011). Advantages of using a Kinect Camera in various applications. *University of Canterbury*.

Xu, C. H., Wang, Y. H., Tan, T. N., & Quan, L. (2004). Depth vs. intensity: Which is more important for face recognition? *Proceedings of the 17th International Conference on Pattern Recognition, Vol 1*, 342-345. doi:Doi 10.1109/Icpr.2004.1334122

Xu, L., Wang, J., Cui, J., Hu, M., Zhang, K., & Teng, W. (2018). Dynamic expression recognition based on dynamic time warping and active appearance model. *40*(2), 338-345.

Yan, C., & Su, G. (1998). Facial feature location and extraction from front-view images. *Journal of Image Graphics of China, 3*(5), 375-380.

Yi, C., Tian, Y., & Arditi, A. (2013). Portable camera-based assistive text and product label reading from hand-held objects for blind persons. *IEEE/ASME Transactions On Mechatronics, 19*(3), 808-817.

Younis, O. (2018). A predictive model for solar photovoltaic power using the Levenberg-Marquardt and Bayesian regularization algorithms and real-time weather data. *International Journal of Advanced Computer Science Applications, 9*(1).

Zhai, M.-y., Feng, G., & Gao, Z. (2008). Facial grading system: physical and psychological impairments to be considered. *Journal of Otology, 3*(2), 61-67.

Zhang, X., Kulkarni, H., & Morris, M. R. (2017). *Smartphone-based gaze gesture communication for people with motor disabilities.* Paper presented at the Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.

Zhao, W., Feng, H., Lun, R., Espy, D. D., & Reinthal, M. A. (2014). *A Kinect-based rehabilitation exercise monitoring and guidance system.* Paper presented at the 2014 IEEE 5th International Conference on Software Engineering and Service Science.

Zhu, X., Lei, Z., Liu, X., Shi, H., & Li, S. Z. (2016). *Face alignment across large poses: A 3d solution.* Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.