iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** V **Month of publication:** May 2018

DOI: http://doi.org/10.22214/ijraset.2018.5228

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue V, May 2018- Available at www.ijraset.com

Touch Free Human Computer Interaction Using Economical Gesture Recognition Solution

Shashank K S[1], Manoj Kumar Chakladar[2], Shreyansh Patni[3], Swapna B[4], Mrs Smitha Sasi[5]
[1, 2, 3, 4, 5]*Department of Telecommunication Dayananda Sagar College of Engineering Bangalore, India.*

Abstract: ***Interesting research has been done in the past on human-computer interaction (HCI). Present day players in the industry like Intel, Google, Apple, Microsoft, Facebook, Snapchat, etcetera are carrying intense research and have invested huge figures in this domain among others. Some of the applications include manipulating medical equipment by single skilled doctor located remotely, manipulating industrial equipment in space, learning in a tactile manner, interactive games, virtual personal assistants, etc. Gesture recognition has evolved as a new input paradigm as it enables fluent communication in a natural manner, more natural than code or even natural language for tactile tasks like moving objects. This becomes more relevant and intuitive in any environmental setting involving picking an object from one point and dropping off at another, or where any kind of locomotive interaction involved. Presently, the major application areas include Automotive, Consumer Electronics, Transit and Gaming sectors. The project employs Minimal equipment with minimal cost, open source technologies, can be made enterprise-ready to be deployed as a SaaS(Software as a Service). Minimal costs due to economical, options in choosing sufficiently powerful hardware (Raspberry Pi, Banana Pi, Nano Pi, Onion Omega, C.H.I.P), use of open source technologies (Python, OpenCV - an open source library used for enterprise computer vision applications launched as early as the end of 20th century as a research initiative by Intel and is publicly available since 2000.) rather than proprietary resources.***
Keywords: Computer vision, Image processing, OpenCV, Python, Raspberry Pi, Node.*js*

I. INTRODUCTION

Several avenues of research have been explored in the field of human-computer interaction (HCI) by industry pioneers like Intel, Google, Apple, Microsoft, Facebook, Snapchat, etc. These industry giants among others are carrying intense research and have invested huge figures in this domain among others. Some of the applications include manipulating medical equipment by single skilled doctor located remotely, manipulating industrial equipment in space, learning in a tactile manner, interactive games, virtual personal assistants, etc. The uses of human-computer interaction are vast. In this project we build a particularly cost-effective yet efficient solution for gesture recognition. Traditional focus on gesture recognition has been geared towards building custom hardware and software, without harnessing the power of existing hardware and software. In this project we follow an alternative approach – we aim to harness the immense but highly underrated power and capabilities of open-source technologies to build a prototype for an eventually enterprise-ready application that accepts human gestures as input via a low-cost webcam and is capable of performing pre-programmed actions based on such a user-provided input. The prototype will involve a sudoku like gamified mathematical puzzle. This game is a web application freely available on the internet, and is often used to sharpen the intellect by puzzling the user with a interesting mathematical puzzle, which involves physically moving tiles with numbers engraved upon them which are displayed on the screen. Like-numbered tiles merge if moved in a particular way as can be seen if the game is played. When like-numbered tiles merge, they become a single tile and the number engraved upon this single tile is now double the number originally present on each of the individual tiles before. In this project, we build a web application from scratch, host it on cloud services procured from an appropriate cloud service provider, build the mechanics of the “2048” application from scratch, and on top of this layer, we deploy our gesture-recognition software in order to perform the physical action of moving tiles described in the previous paragraph, through a user’s gestures. Although this prototype illustrates the gesture-recognition solution we developed by means of a simple yet elegant mathematical puzzle involving user-input physical motion, the gesture recognition solution itself can be deployed in a variety of use cases specially due to its economical nature as a result of using low-cost and easily accessible hardware, use of open source technologies, and use of cloud technologies. The use of opensource technologies makes further development of this minimal prototype far more promising than if enterprise closed-source technologies were used. The hope is that

©IJRASET: All Rights are Reserved

the immense support that Python and OpenCV have enjoyed in the open-source community can be tapped to build a robust community of active open-source contributors around the current gesture recognition solution as well. As we know, contributors to patches and new features of highly popular projects like Python and OpenCV tend to be coders who are at the top of their respective domains. An effective and economical solution to gesture-recognition using Python and OpenCV has immense potential to tap into this existing community of elite contributors in the open source community already existing around Python and OpenCV. Slowly but surely, traditionally closed source companies are realizing the importance of transitioning at least some of their closed source technologies to open-source and are also moving towards building many of their closed-source offerings using open-source technologies, for the simple reason that the offering's usability increases by leaps and bounds due to the robust community already existing around any popular open-source technology. As mentioned in abstract, gesture recognition has the potential to evolve as a new input paradigm and outdo even traditional input methodologies like the keyboard, mouse and even natural language commands communicated to the computer by means of a digital microphone. This is so because it enables effortless and fluent communication in a natural manner when it comes to physical movement-related tasks. In this domain, gesture recognition as a mode of input is more natural than lines of code written to perform the same task, or even natural language directions as mentioned. This becomes more relevant and intuitive in any environmental setting involving picking an object from one point and dropping off at another, or where any kind of locomotive interaction involved. The potential for expanding the present use case is immense, though active research would be necessary for any high-accuracy application like remote surgery or industrial production. The current prototype however illustrates that minimal hardware and cloud technologies is enough to create an immensely robust gesture recognition product. Once expanded to cover other use cases, the core gesture-recognition functionality can be offered as a SaaS (Software as a Service) and even as a PaaS (Platform as a Service) for building gesture recognition-based applications purely on the cloud, without spending one's own resources to procure even the minimal hardware that we use in this prototype. This would enable any user to quickly, efficiently and economically use gesture-recognition in their own custom use-cases.

II. LITERATURE SURVEY

Extensive literature exists on the topic of computer vision. Human visual systems, from the engineering view-point, traditionally seek to automate tasks which are monotonous, dreary or easily automatable. This is well illustrated in the following famous texts – the landmark textbook by Dana H. Ballard and Christopher M. Brown (1982) titled "Computer Vision", published by Prentice Hall, a paper by Huang, T. (1996-11-19), Vandoni, Carlo, E, bearing the title Computer Vision: Evolution and Promise [2], published in the conference proceedings of the 19th CERN School of Computing. In the 1960s, the field of computer vision was in its infancy. "The Summer Vision Project" conducted by MIT aimed to use the summer workers of MIT to help build a computer attached to a camera, and enable the camera to describe what it visualized. The description of this project can be found in Papert, Seymour (1966-07-01), "The Summer Vision Project"[3], MIT AI Memos (1959– 2004) and Margaret Ann Boden (2006), "Mind as a Machine: A History of Cognitive Science" [4], Clarendon Press. p.781. There exist many reviews on the state of research in the broad field of Computer Vision. The following review papers shed some light on the state of research as well as on the state of education in the field of Computer Vision: Review of Computer Vision Education [1](https://www.cs.unm.edu/~williams/cs591/TransEduc.pdf) by George Bebis, Dwight Egbert and Mubarak Shah which reviews the state of Computer Vision education [1] in the present times. A more recent research's and review's illustrates that the current state of the field of Gesture Recognition. Image processing and computer vision finds immensely beneficial use in industrial automation through robotics, as illustrated in Steger, Carsten; Markus Ulrich & Christian Wiedemann (2008) [5]. Machine Vision Algorithms and Applications, Weinheim: Wiley-VCH.; Beyerer, Jürgen; Puente León, Fernando & Frese, Christian (2016) in Machine Vision – Automated Visual Inspection: Theory, Practice and Applications [6], Berlin: Springer; Graves, Mark & Bruce G. Batchelor (2003) Machine Vision for the Inspection of Natural Products, Springer. OpenCV is an open-source computer vision library which due to being a popular and widely used open-source library, enjoys top-notch support from a highly active and skilled open-source developer community. This library is aimed mainly at realtime computer vision as opposed to a proliferation of other non-real-time applications. This is illustrated in Pulli, Kari; Baksheev, Anatoly; Kornyakov, Kirill; Eruhimov, Victor (1 April 2012), Realtime Computer Vision with OpenCV. A survey of available literature related to OpenCV is incomplete without a reference to the famous book Bradski, Gary; Kaehler, Adrian (2008), Learning OpenCV: Computer vision with the OpenCV library, O'Reilly Media, Inc. p.6. as well as a vast variety of others, which we do not mention for want of space and for purposes of brevity.

©IJRASET: All Rights are Reserved

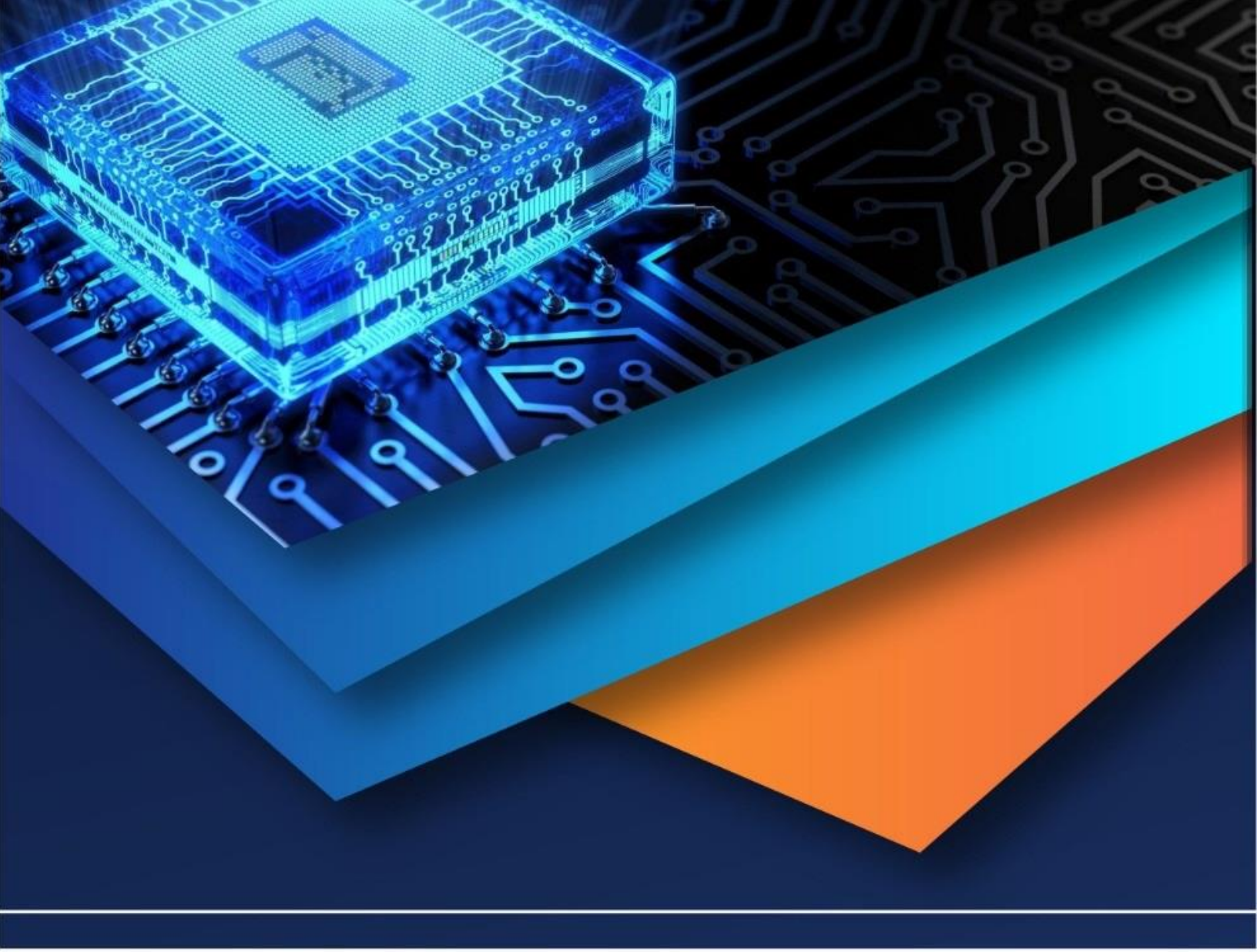
III. ACKNOWLEDGEMENT

We thank Head of our Department for giving us this opportunity, we also thank our guide for helping us and guiding us through this endeavour.

REFERENCES

[1] George Bebis., Dwight Egbert., Mubarak Shah: Review of Computer Vision Education, https://www.cs.unm.edu/~williams/cs591/TransEduc.pdf
[2] Huang, T., Vandoni, Carlo, E: Computer Vision: Evolution and Promise. In: 19th CERN School of Computing. Geneva: CERN. pp. 21–25 (1996)
[3] Papert, Seymour: The Summer Vision Project (1966) http://hdl.handle.net/1721.1/6125
[4] Margaret Ann Boden: Mind as a Machine: A History of Cognitive Science, p. 781. Clarendon Press (2006)
[5] Steger, C., Markus U., Christian W: Machine Vision Algorithms and Applications Weinheim, Wiley-VCH, p. 1 (2008)
[6] Beyerer, Jürgen., Puente León, Fernando., Frese, Christian: Machine Vision – Automated Visual Inspection: Theory, Practice and Applications, Berlin: Springer(2016)
[7] Pallotta, V., Bruegger, P., Hirsbrunner, B: Kinetic User Interfaces: Physical Embodied Interaction with Mobile Pervasive Computing Systems. Advances in Ubiquitous Computing: Future Paradigms and Directions, IGI Publishing (2008)
[8] Jaimes, A., Sebe, N: Multimodal human–computer interaction: A survey, Computer Vision and Image Understanding, Special Issue on Vision for Human-Computer Interaction. Volume 108, Issues 1-2, pp. 116-134 (October–November 2007)
[9] AK, Sofia: 9 Minority Report Inspired Touchless Technology. HKDC. (2017)
[10] Quek, F: Toward a vision-based hand gesture interface, In: Proceedings of the Virtual Reality System Technology Conference, pp. 17-29, Singapore (1994)
[11] Kolsch, M., Turk, M: Fast 2D Hand Tracking with Flocks of Features and Multi-Cue Integration., CVPRW '04. In: Proceedings Computer Vision and Pattern Recognition Workshop (2004)
[12] Fujimura, K: Hand gesture recognition using depth data. In: Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 529- 534 (2004)
[13] Official website of the open source computer vision library, https://opencv.org/
[14] OpenCV repository, www.github.com/opencv/opencv/

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)