Arcada Working Papers 3/2019
ISSN 2342-3064
ISBN 978-952-7365-02-1

Recent Developments in Computer Vision Based Real-Time Monitoring in Health and Well-Being

Göran Pulkkis, Jonas Tana, Thomas Hellstén, Jonny Karlsson

www.arcada.fi

Arcada Working Papers 3/2019
ISSN 2342-3064
ISBN 978-952-7365-02-1

Recent Developments in Computer Vision Based Real-Time Monitoring in Health and Well-Being

Göran Pulkkis[i], Jonas Tana[ii], Thomas Hellstén[iii], Jonny Karlsson[iv]

Abstract

Computer vision and its application is promising in several areas of health and well-being. Research within computer vision has made substantial progress during the past several decades, and will likely bring forward rapid advancements, especially when it comes to fast, real-time detection and recognition of different patterns or objects within health and well-being. This kind of real-time monitoring within computer vision has expanded rapidly during the last decades. The aim of this article is to present some of these advancements. Focus lies on current gesture recognition for real-time detection of human emotions and alertness, sign language translation, detection of safety critical incidents such as fall incident detection, functional vision aids for partially and fully blind persons, tele-surgery, computer vision based diagnostic health examination methods (endoscopy, photoplethysmography, and digital mammography), and computer vision based aids within rehabilitation. Recent research shows that computer vision based real-time monitoring is a valuable aid within many fields of health and well-being that enables and aids people. Computer vision is also an emerging research field and direction that requires multi-disciplinary collaboration and deployment of advanced machine learning methods.

Keywords: Computer Vision, Real-Time Monitoring, Surveillance, Tracking, Gesture Recognition, Emotion Detection, Alertness Detection, Sign Language, Bionic Eye, Visual Prosthesis, Tele-surgery, Endoscopy, Photoplethysmography, Digital Mammography

[i] Arcada UAS, Finland, [goran.pulkkis@arcada.fi]
[ii] Arcada UAS, Finland, [jonas.tana@arcada.fi]
[iii] Arcada UAS, Finland, [thomas.hellsten@arcada.fi]
[iv] Arcada UAS, Finland, [jonny.karlsson@arcada.fi]

1 INTRODUCTION

The field of computer vision is the science and technology that focuses on the development and implementation of algorithms, which allow computers to "see" or "understand" and extract data from an image or video that is necessary to solve a task [1] [2]. Computer vision can be seen as a process where input of images, either still or video, produces an output with different characteristics or parameters related to the images in a rapid, accurate and comprehensive way [3] [4]. This kind of image or object detection, interpretation, recognition, understanding or tracking is highly relevant in modern intelligent control, and has been applied in many different parts of modern society, like robotics, surveillance, monitoring and security systems [3] [5]. Computer vision also includes methods for acquiring, processing and under-standing complex constructions to model, replicate and even exceed human vision to perform useful tasks [4] [6]. Technologies that rely on computer vision are already, and will most likely be, used in a myriad of services and devices in the future, including smartphones, cameras, wearable technology and monitoring equipment [6].

Computer vision and its application is promising in several areas of health care and will likely bring rapid advancements, especially when it comes to fast detection and recognition of different objects or patterns [2] [6]. Research with-in computer vision has made substantial progress during the past several decades. According to [7], there are four major healthcare research domains, where computer vision has, or likely will have, a bigger impact:

1. analysis of medical images,
2. computer vision techniques for predictive analytics and therapy,
3. fundamental algorithms for medical images, and
4. machine learning techniques for medical images.

Research on this type of task-oriented image analysis or understanding has the potential to offer assistance to human perception, cognition and decision-making, by developing computer vision aided diagnosis systems [1]. From a more practical perspective, computer vision techniques have been gaining more importance within assisted living systems, and more specifically within patient monitoring [8]. Within this area, computer vision can offer a complementary, non-invasive, monitoring system to the direct monitoring of physiological parameters. This is in relation to the rapid advances in technology with decreasing costs, increasing miniaturization pervasiveness of technologies such as smartphones and tablets with high-resolution cameras as well as high speed computing systems that allow for continuous or real-time monitoring [8]. The area of computer vision based real-time monitoring, especially for the betterment of human life, or within healthcare applications, has expanded rapidly during the last decades, with an increasing demand for new technological solutions [9] [10]. This calls for an exploration of recent advances.

The aim of this paper is to present recent developments where eHealth and welfare applications utilize computer vision based real-time monitoring. Section 2 presents current gesture recognition for real-time detection of human mental states and sign language translation. Section 3 describes a computer vision based safety service in the living environment of a person. In Section 4 is described how computer vision can restore functional vision to partially and fully blind persons. Section 5 describes tele-surgery. Section 6 presents computer vision based diagnostic health examination methods. Section 7 discusses computer vision based aids for rehabilitation. Finally, Section 8 presents conclusions and outlines potential directions for future research.

2 GESTURE RECOGNITION APPLICATIONS

Gesture recognition is interpretation of a person's gestures using mathematical algorithms. Most gestures originate from hand movements and facial expressions. Gesture recognition is a real time application with images recorded with a camera as input. Current health and welfare applications of gesture recognition are emotion recognition, alertness detection, and sign language translation.

2.1 Emotion Detection

Autistics and people with communication disabilities have difficulties in reading other persons' emotions. If an affected person could read emotion, then interaction with other persons and their feelings would be much easier. Therefore, a useful aid would be a real-time computer vision system for detection of the emotion of another person. An example of an application to detect the emotion of a person is SHORE (Sophisticated High-speed Object Recognition Engine) [11], which applies Google Glasses [12] to detect emotions through a built-in camera. Image analysis occurs in real-time using face and emotion recognition algorithms. Facial features such the state (open or closed) of eyes and mouth are detectable in real time. Emotions, such as being angry, happy, sad, or surprised are detectable.

2.2 Alertness Detection

Autistics and people with communication disabilities have difficulties in reading other persons' emotions. Monitoring and ensuring the alertness of a human being is important within many domains, both in professional and daily life. One domain is car safety as various studies have shown that around 20% of all road accidents are due to driver drowsiness [13]. Drowsiness is detectable with high accuracy by measuring and monitoring physiological signals like brain waves, pulse rate, respiration, and heart rate but these methods are not practical, since they require various sensors connected to the body.

Computer vision is applicable for driver drowsiness detection as a low cost and real-time monitoring solution. Furthermore, it is implementable without any built-in system requirements on the car. All it takes is a camera pointed at the driver's face in combination with drowsiness detection software. Computer vision for driver drowsiness detection has been an intensive research topic already for many years [14].

Research in this area is still highly topical. In recent proposals for computer vision based drowsiness detection both face monitoring and pressure management reveal driver drowsiness [15] [16]. The drowsiness detection technique consists of two steps. In the first step, a force sensitive resistor measures pressure to check the physiological performance of the driver. In the second step, facial images captured utilizing the viola-jones object detection algorithm [17] analyse the driver's eye and yawn state. Finally, the drowsiness detection algorithm combines the results from both the facial feature analysis and the pressure measurement.

2.3 Sign Language Translation

Autistics and people with communication disabilities have difficulties in reading other persons' emotions. A sign language, which is used in communication with seriously hearing impaired persons, consists of hand and face gestures. There is no universal sign language, since a sign language is specific for a natural language. For example, even British English and American English have their own sign languages [18] [19]. A basic sign language consists of specific gestures for each character in the alphabet and for decimal numbers. Figure 1 shows the hand gestures representing the 7 last characters of the alphabet in the American Sign Language (ASL). Figure 2 shows the hand gestures representing the decimal digits 0 … 9. A real sign language has also signs for common words and phrases such as 'Merry Christmas', 'thank you', and 'excuse me', where a sign may consist of gestures with both hands in combination with some facial expressions. For example, the ASL dictionary consists of signs for about 2000 common words and phrases [19].

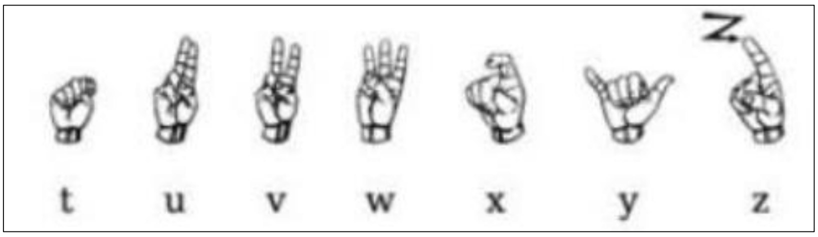

Figure 1. Hand gestures representing 7 characters of the alphabet in ASL

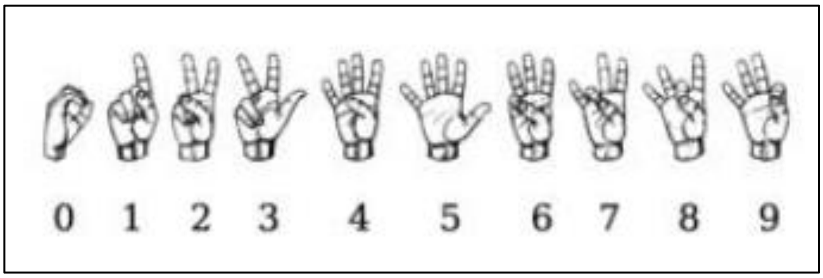

Figure 2. Hand gestures representing decimal digits in ASL

Recent computer vision based real time sign language recognition implementations use before operation supervised machine learning requiring training with a given gesture dataset. The real time sign language recognition implementation presented in [20] classifies camera-captured gestures with a multi-layered random forest (MLRF) [21] after a training session. Implementations in [22] [23] [24] use a neural network for gesture recognition after a training session.

In [25] is described the use of a smartphone for computer vision based real time sign language translation. The sampling rate of gesture images with the smartphone camera is five frames per second. The LABVIEW software Vision Assistant [26] processes captured images to recognize unique sign features. The smartphone loudspeaker provides audio output after sign recognition. The described sign language translator can translate the ALS alphabet and decimal digit signs.

Commercial computer vision based real time sign language translator solutions are also available [27] [28] [29].

3 FALL INCIDENT DETECTION

The Intelligent video surveillance of monitoring elderly and disabled persons in their home environment ensures safer living, since the surveillance system can trigger alerts to family members and health professionals by detection of safe-ty critical incidents such as vision based fall detection and suspiciously long permanence in the same laying down or sitting static state. A real time surveillance system called Ghost, described in [30], detects persons' postures from constructed silhouette based body models. Four main postures - standing, sitting, crawling-bending, and laying down - are considered. Each main posture can have three view-based appearances: front view, left side, and right side. The silhouette-based body model locates six primary body parts - head, two hands, two feet, and torso. On the silhouette boundary, 10 secondary body parts can support localization of primary parts. The Ghost algorithm computes similarities of horizontal and vertical histograms for the detected silhouette and the main postures. The estimated posture is the most similar one. Possible body parts are localised with a recursive convex-hull algorithm in the estimated posture. Repeated evaluations of the Ghost algorithm on monitored images reveals suspiciously long permanence in the same laying down or sitting static state for a person, but it is not possible to distinguish a fall incident from an event when a person lays down for rest or sleep.

More recent real-time surveillance systems [31] [32] detect also fall incidents with the use of falling time registration. The method used in [31] is classification of persons' postures with machine learning based on the k-NN algorithm [33]. Representations of persons are segmented moving foreground objects extracted from the background using the approach proposed in [34]. A horizontal and a vertical projection histogram is extracted from the segmented foreground object using calculation of pixels row wise and column wise. The object is rescaled first to a fixed vertical length of 128 pixels and after this to a fixed horizontal length of 128 pixels. Both histograms constitute a feature vector of 256 elements for input to the k-NN classifier. Offline training sessions require set consisting of 256 pre-computed element tem-plates. Each pre-computed element template is stored in video sessions for each one of five different postures: standing, sitting, bending, lying on a side, and lying forward. Classification of detected postures is possible after training. The k-NN algorithm can then find the highest similarity with the templates stored in the training. The classification result in combination with an evidence accumulation procedure determines the choice of posture type. When the time difference between an estimated lying posture and the last estimated standing posture is less than a threshold value setting, then a real fall incident has apparently occurred. More than 90 % of 743 detected postures have been experimentally correctly classified [31]. The fall incident detection method reported in [32] is simpler than the method reported in [31] and fall indent detection rate was 80% or less in experiments. A time interval of 0.4 to 0.8 seconds between the lying feature and the last standing feature indicated a fall incident.

4 BIONIC EYE

A bionic eye is a visual prosthesis, a device for restoring functional vision to partially or fully blind persons by electrical stimulation of an eye globe layer, the optic nerve, or the brain with signals created by processing images received by a digital camera. A visual prosthesis system consists of an external module and an implantable module as is depicted in Fig 3. The external module consists of a camera, a unit for pre-processing camera images, and a unit for wireless transmission of energy and pre-processed image information to the implantable module. The implantable module has an embedded processing unit connected to a wireless interface for reception of energy and pre-processed image information. Energy and data from the external module to the implantable module is wireless transferred as induction be-tween closely coupled coils or by a capacitive link [35]. The embedded processing unit controls stimulation circuitry, which stimulates a microelectrode array to issue electrical signals to the target interface. [36]

A bionic eye is a visual prosthesis, a device for restoring functional vision to partially or fully blind persons by electrical stimulation of an eye globe layer, the optic nerve, or the brain with signals created by processing images received by a digital camera. A visual prosthesis system consists of an external module and an implantable module as is depicted in figure 3. The external module consists of a camera, a unit for pre-processing camera images, and a unit for wireless transmission of energy and pre-processed image information to the implantable module. The implantable module has an embedded processing unit connected to a wireless interface for reception of energy and pre-processed image information. Energy and data from the external module to the implantable module is wireless transferred as induction between closely coupled coils or by a capacitive link [35]. The embedded processing unit controls stimulation circuitry, which stimulates a microelectrode array to issue electrical signals to the target interface. [36]

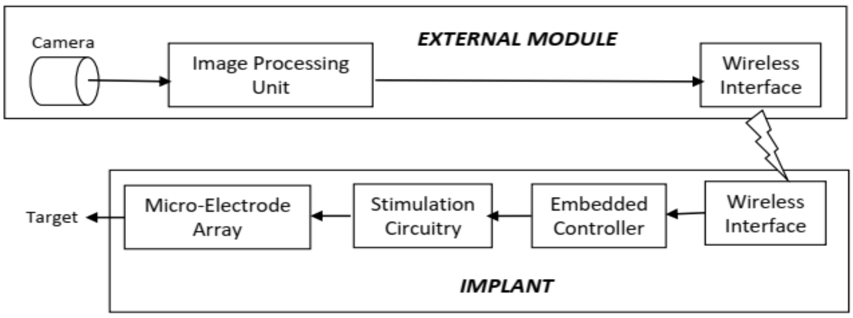

Fig. 3. General architecture of a visual prosthesis system (adapted from [36])

In the retinal approach to a bionic eye, an implant stimulates electrically retinal cells. The stimulation is epi-retinal for an implant attached to the inner surface of the retina as shown in figure 4 and sub-retinal for an implant with a microelectrode array implanted between layers inside the retina. Microelectrode array implants in eye globe layers behind the retina is also a considered approach to bionic vision. These layers are the choroid and the sclera as is shown in figure 4. The microelectrode array implant would then be behind the choroid or within the sclera. An implanted cuff electrode, which encircles the optic nerve and receives stimulation signals from a processing implant as shown in figure 5, implements electrical stimulation of the optic nerve. A microelectrode implant at the visual cortex as shown in figure 4 or within the relay center of the thalamus could implement electrical stimulation of the brain. [36] [37] [38]

There are several on-going bionic vision research projects around the world [36]. Commercial visual prosthesis products are currently also available [39] [40]. The first reported implantation of the retinal approach to a bionic eye in 2012 enabled the patient to see light but gave no aid for orientation in the environment. Mose recently implanted more developed prostheses have enabled recognition of abstract images in patients' environments. [41]

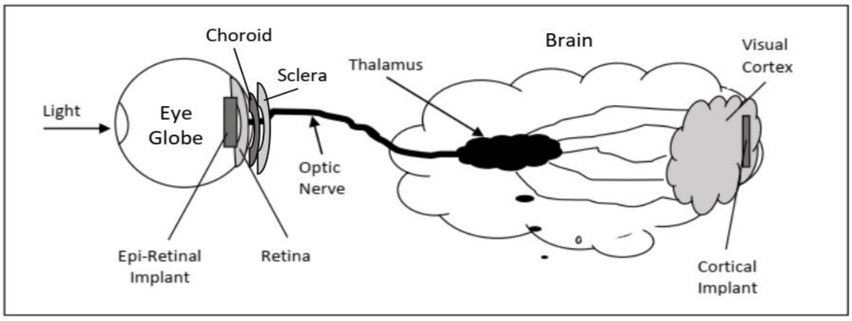

Figure 4. *Retinal and cortical approach to a bionic eye (adapted from [36] and modified).*

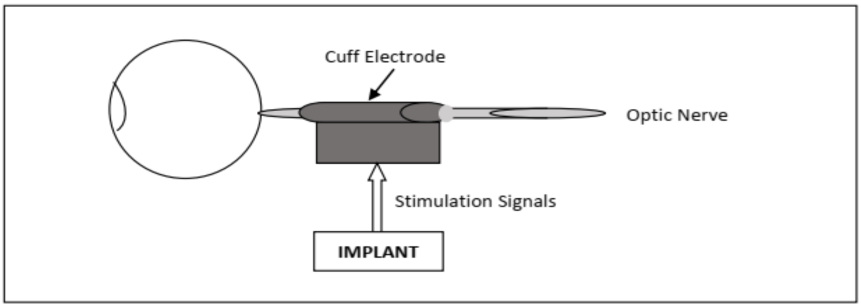

Figure 5. *Electrical stimulation of the optic nerve (adapted from [36]).*

5 TELE-SURGERY

In tele-surgery, surgeons operate on distantly located patients by steering remote robotic devices in an operation environment provided by networking and real-time computer vision [42]. A surgical team in New York, USA conducted the first tele-surgery in the world using the Zeus tele-surgery technology, which consists of the patient's and the surgeon's subsystems [43]. The patient's subsystem consists of a camera, two robotic arms, and an automated endoscope system. The surgeon's subsystem has handles to control the robotic arms and the endoscope system as well as a con-sole showing in real-time the actions of the robotic arms and the input to the endoscope system. Virtual Interactive Presence (VIP) is a novel computer vision based technology for tele-surgery [44]. VIP enables collaboration between remote and local surgeons by a shared 3-Dimensional display showing a merged surgical field view of all surgeons' hand motion. Figure 6 shows a VIP based tele-surgery setup.

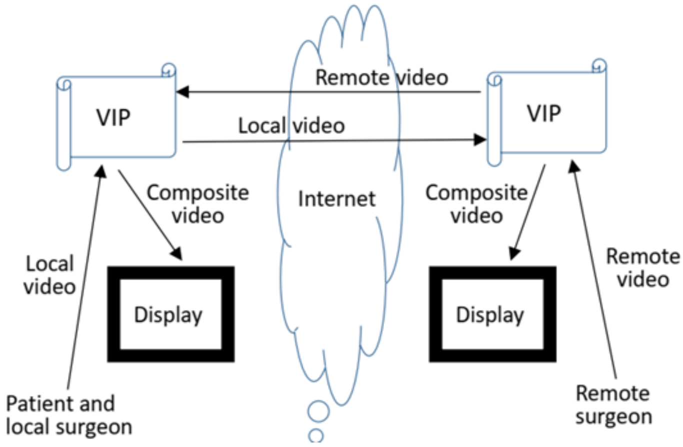

Figure. 6. A Virtual Interactive Presence (VIP) based tele-surgery setup (adapted from [44])

6 HEALTH EXAMINATION METHODS BASED ON COMPUTER VISION

Several computer vision based method for diagnostic health examination are currently available. Endoscopy makes examinations inside a person's body, photoplethysmography examines infrared light reflected or transmitted from a person's skin, and digital mammography examines a person's breasts with x-rays.

6.1 Endoscopy

Endoscopy can carry out diagnostic examinations and small-scale surgery inside a person's body using an instrument, which is a long and thin tube with a frontend consisting of a high-intensity light source, a high-resolution camera, and some surgery tools. As an endoscope moves inside a person's body, a video monitor receives real time images from the camera. There are several endoscopy types:

- Laparoscopy is endoscopy on abdominal organs with an instrument called a laparoscope. Insertion of a laparoscope requires only a small incision in the abdominal wall. Laparoscopy allows a doctor to make real time diagnostic examinations and surgery inside a patient's abdomen.
- Colonoscopy is endoscopic examination of a person's bowel.
- Gastroscopy is endoscopic examination of a person's gullet, stomach or first part of the small intestine.
- Hysteroscopy is endoscopic examination inside of a woman's womb.
- Cystoscopy is endoscopic examination inside of a person's bladder
- A special endoscopy type is wireless capsule endoscopy where a person swallows a camera capsule, which transmits wirelessly images to a video monitor from inside of the person's stomach and digestive system. The capsule size is as a large pill and the capsule leaves the person's body when the person is on the toilet.

[45] [46]

6.2 Photoplethysmography (PPG)

In PPG the skin of a person is illuminated with an infrared light source and the amount of either transmitted or reflected light is measured. The infrared light penetrates the skin whereby bones, skin pigments, and blood in veins and arteries absorb light. Measurements of transmitted or reflected light register absorption level. The measured and processed light input is a photoplethysmogram. An IoT device with an infrared LED actuator and a photodiode sensor could therefore implement a PPG device. A photoplethysmogram consists of signals with an AC component and a DC component obtained by amplifying and filtering the photodetector input. The AC component shows blood volume changes with each heartbeat in veins just under the skin. The slowly varying DC component relates to respiration, activity of the sympathetic nervous system, and thermoregulation. [47] [48]

6.3 Digital Mammography

Standard mammography uses x-rays to create detailed images of the breasts on a film cassette. Digital mammography, also called computerized mammography, captures breast images with an electronic x-ray detector converting the images to digital pictures, which a computer displays on a connected monitor. Advantages of digital mammography in comparison with standard film mammography are faster image acquisition time with a smaller radiation dose and the possibility to process the magnification, brightness, and contrast of captured images for better image examination results. Advantages of standard film mammography in comparison with digital mammography are higher image detail resolution and lower image acquisition costs. [49]

7 AIDS FOR REHABILITATION

A new practice in computer vision supported rehabilitation in motion analysis is the marker-less approach that is totally self-going and non-invasive. This implementation would be a big step forward in rehabilitation to analyse the motion. Marker-based analysis in rehabilitation that is in use today needs long preparation time or laboratory facilities with several cameras and markers. With a well working marker-less approach the motion analysis could take place for ex-ample at the client's home [50]. In normal face-to-face clinical cases the cost, travel that is necessary for the session or the chance for therapy can be the main barriers for a client to get help [51]. Several computer vision based real time monitoring aids for rehabilitation have been proposed and implemented. This section presents some recent examples. Balance is a central task for people and especially for elderly. It is important to be able to identify fallers, because the injuries that becomes when a person fall can be severe. Most of the injuries are mild but five to ten percent of the injuries are severe for people that are older than 65 years [52]. To be able to identify possible fallers Nalci et al. [53] analysed with help of marker-less computer vision standing on one foot with eyes open and closed and compared the results with a golden standard balance board that analysed the sway. The results had a high correlation and shows that computer vision can in the future be a possible equipment used in balance measuring in rehabilitation to identify possible fallers.

Homebased exercises supervised by a therapist are one of the most important treatment in the recovery phase in several diagnosis and especially in osteoarthritis (OA). The goal of home based exercises in OA is to decrease pain in the joint and get better functionality [54]. Computer vision aided system that can capture exercise and give feedback can be a key to effective and successful rehabilitation process. Without feedback, the therapeutic pro-gram are difficult to personalized and motivate the client to do the exercise [55]. There is also a possibility that if the clients do not get feedback they do the exercises wrong. Especially, after surgery this can be harmful and slow down the recovery process [56]. Dorado et al. [57] present a developed easy to use computer vision system called ArthriKin, which offers a possibility to interact directly with a therapist to effective the homebased exercise program.

In [58] is demonstrated real-time measurements of a patient's vestibular exercises with a smartphone camera. During Vojta therapy [59] body movements of patients having motor disabilities have been monitored and analysed in real-time with computer vision based experimental methods [60]. A computer vision based action identification system for upper extremity rehabilitation in patients´ home environments has been proposed [61]. The proposed system captures sequences of colour images with colour and depth of a patient's upper extremity actions for identification of movements. The image sequence with a rate up to 125 images/s is processed and analysed to distinguish between correct and wrong rehabilitation actions in action training. Neuro-rehabilitation of stroke patients has also used computer vision [62]. Computer vision models and tracks objects to assist patients, which try to reach for and grasp these objects with the aid of a robotic device. Rehabilitation of patients with an injured arm or wrist has likewise used computer vision [63]. A web camera records a cuboid object and software calculates the object's position of in real-time when the patient tries to move the object to match the position of an already present virtual object in a virtual 3D space.

8 CONCLUSIONS

Computer vision based real-time monitoring is currently a key component in two highly significant aids for persons suffering from serious sensual impairments. Blind people and individuals with other serious vision disabilities can have some functional vision restored by bionic eye. Real time sign language translation provides a significant extension of serious hearing impaired persons' capabilities to communicate with other persons. Real time sign language translation is a real time gesture recognition application, where hand movements and facial expressions implement gestures. Other real time gesture recognition applications are detection of a person's emotions and alertness. Autistic persons and persons living with similar/other communication disabilities cannot clearly recognize other persons/individuals/people's feelings. Real time detection of other persons' emotional states is therefore a valuable interaction aid. Emotion detection therefore extends the disabled persons' capabilities to interact with other persons. Real time alertness detection is an important safety issue for example for a car driver alone in the car. Tracking and surveillance implementations in rehabilitation therapies are usually computer vision based real-time monitoring. A new practice today is the marker-less approach that is totally self-going and non-invasive.

Extension of the current bionic vision capabilities is an important research direction. Close cooperation with neuro-logical research on the signalling from the retina in the eye glove to the visual cortex in the brain is a necessity. A tough research challenge in sign language translation research is recognition of all words and phrases is sign language dictionaries for different natural languages. Responses to this research challenge require very big training databases for different sign languages and deployment of advanced machine learning methods

REFERENCES

[1] A.B. Albu, "Vision-Based User Interfaces for Health Applications: A Survey," *Advances in Visual Computing*, G. Bebis et al. , eds., ISVC 2006, Lecture Notes in Computer Science, vol. 4291, Berlin: Springer-Verlag, pp. 771-782, 2006.

[2] S.R. Yoshida, *Computer Vision*, Nova Science, New York, 2011.

[3] A.A. Abed, S.A. Rahman, "Computer vision for object recognition and tracking based on Raspberry Pi," *Shaping the Future of ICT: Trends in Information Technology, Communications Engineering, and Management*, I.M.M. El Emary and A. Brzozowska, eds., USA: CRC Press Taylor & Francis Group, pp. 3-14, 2017.

[4] J. Seo, S, Han, S. Lee, and H. Kim, "Computer vision techniques for construction safety and health monitoring," *Advanced Engineering Informatics*, vol. 29, no. 2, pp. 239-251, 2015.

[5] M. Pediaditis, M. Tsiknakis, and N. Leitgeb, "Vision-based motion detection, analysis and recognition of epileptic seizures—a systematic review," *Computer methods and programs in biomedicine*, vol. 108, no. 3, pp. 1133-1148, 2012.

[6] J. Peddie, K. Akeley, P. Debevec, E. Fonseka, M. Mangan, and M. Raphael, "A vision for computer vision: Emerging technologies," *ACM SIGGRAPH 2016 Panels (SIGGRAPH '16)*, Article no. 2, New York, NY, USA: ACM, 2016, doi:10.1145/2927383.2933233.

[7] J. Gao, Y. Yang, P. Lin, and D.S. Park, "Computer Vision in Healthcare Applications," *J. Healthcare Engineering*, vol. 2018, 5157020, 2018, doi:10.1155/2018/5157020.

[8] S. Sathyanarayana, R.K. Satzoda, S. Sathyanarayana, and S. Thambipillai, "Vision-based patient monitoring: a comprehensive review of algorithms and technologies," J. Ambient Intelligence and Humanized Computing, vol. 9, no. 2, pp. 225-251, 2018.

[9] Md.A.R. Ahad, S. Kobashi, and J.M.R.S. Tavares, “Advancements of Image Processing and Vision in Healthcare,” J. Healthcare Engineering, vol. 2018, 8458024, 2018, doi:10.1155/2018/8458024.

[10] M. Leo, G. Medioni, M. Trivedi, T. Kanade, and G.M. Farinella, “Computer vision for assistive technologies,” *Computer Vision and Image Understanding*, vol. 154, no. C, pp. 1-15, 2017, doi:10.1016/j.cviu.2016.09.001.

[11] *Face Detection Software SHORE®: Fast, Reliable and Real-time Capable*, Fraunhofer Institute for Integrated Circuits IIS, https://www.iis.fraunhofer.de/en/ff/sse/ils/tech/shore-facedetection.html. 2019.

[12] D. Thompson, “Google Glass Helps Kids with Autism Navigate Emotions of Others,” *HealthDay News for Healthier Living*, https://consumer.healthday.com/cognitive-health-information-26/autism-news-51/google-glass-helps-kids-with-autism-navigate-emotions-of-others-736413.html. 2018.

[13] *Driver Fatigue and Road Accidents Factsheet*, The Royal Society for the Prevention of Accidents ROSPA, https://www.rospa.com/rospaweb/docs/advice-services/road-safety/drivers/driver-fatigue-factsheet.pdf. 2017.

[14] A. Chellappa and R. Ezhilarasie, “Fatigue Detection Techniques: A Review,” *Int´l J. Pure and Applied Mathematics*, vol. 117, no. 16, pp. 503-510, 2017.

[15] A. Sayeed and S.A. Sadim, “Driver Drowsiness Detection using Face Monitoring and Pressure Measurement,” *J. Embedded System & Applications*, vol. 5, no. 3, pp. 12-18, 2017.

[16] A.S. Kulkarni and S.B. Shinde, “Monitoring Driver Distraction in Real Time Using Computer Vision System,” *Int´l J. Computer Sciences and Engineering*, vol. 5, no. 6, pp. 121-128, 2017.

[17] P. Viola and M. Jones, “Robust Real-time Object Detection,” *Int´l J. Computer Vision*, 2001.

[18] *British Sign Language Website*, https://www.british-sign.co.uk/. 2019.

[19] *American Sign Language Website*, https://www.handspeak.com/. 2019.

[20] A. Kuznetsova, L. Leal-Taixe, and B. Rosenhahn, “Real-time sign language recognition using a consumer depth camera,” *Proc. 2013 IEEE Int´l Conf. on Computer Vision Workshops*, pp. 83-90, 2013.

[21] C. Keskin, F. Kirac, Y.E. Kara, and L. Akarun, “Hand pose estimation and hand shape classification using multi-layered randomized decision forests,” *Proc. 12th European Conf. Computer Vision - Volume Part VI, ECCV’12*, pp. 852–863, 2012.

[22] P. Mekala, Y. Gao, J. Fan, and A. Davari, “Real-time Sign Language Recognition based on Neural Network Architecture,” *Proc. IEEE 43rd Southeastern Symp. System Theory*, pp. 195-199, 2011.

[23] B. Garcia and S. Viesca, “Real-time American Sign Language Recognition with Convolutional Neural Networks,” *Reports*, Stanford University, Stanford, CA, USA, http://cs231n.stanford.edu/reports/2016/pdfs/214_Report.pdf. 2016.

[24] V. Bheda and N.D. Radpour, “Using Deep Convolutional Networks for Gesture Recognition in American Sign Language,” *arXiv:1710.06836 [cs.CV]*, https://arxiv.org/abs/1710.06836. 2017.

[25] M. Yellapu, G. Anitha, and M. Anburajan, “Vision-Based Sign Language Translation Device,” *Proc. 2013 Int´l Conf. Information Communication and Embedded Systems (ICICES)*, pp. 565-568, 2013.

[26] *NI Vision Assistant Tutorial*, National Instruments, http://www.ni.com/pdf/manuals/372228m.pdf. 2011.

[27] *UNI - The world's first real time translation technology that converts sign language to grammatically correct spoken language*, http://enableneeds.com/2018/07/23/uni/. 2018.

[28] K. Tsuboi, “Sign-language translator uses gesture-sensing technology,” https://www.cnet.com/news/sign-language-translator-uses-gesture-sensing-technology/. 2014.

[29] *KinTrans, Hands Can Talk*, http://www.kintrans.com/. 2019.

[30] I. Haritaoglu, D. Harwood, and L.S. Davis, “Ghost: A Human Body Part Labeling System Using Silhouettes,” *Proc. 14th Int´l Conf. Pattern Recognition*, vol. 1, pp. 77-82, 1998.

[31] A.H. Nasution and S. Emmanuel, “Intelligent Video Surveillance for Monitoring Elderly in Home Environments,” *Proc. IEEE 9th Workshop on Multimedia Signal Processing*, pp. 203-206, 2007.

[32] T. Kroputaponchai and N. Suvonvorn, “Vision-based Fall Detection and Alert System Suitable for the Elderly and Disable Peoples,” *Proc. 24th Japanese Conf. Advancement of Assistive and Rehabilitation Technology (JCAART)*, pp. 217-218, 2009.

[33] *K Nearest Neighbors – Classification*, https://www.saedsayad.com/k_nearest_neighbors.htm. 2019.

[34] C. Stauffer and W.E.L. Grimson, “Adaptive Background Mixture Models for Real-time Tracking,” *Proc. 1999 IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, vol. 2, pp. 246-252, 1999.

[35] A.M. Sodagar and K. Najafi, “Wireless Interfaces for Implantable Biomedical Microsystems,” *Proc. 49th IEEE Int´l Midwest Symp. Circuits and Systems (MWSCAS)*, pp. 265-269, 2006.

[36] M.H. Maghami, A.M. Sodagar, A. Lashay, H. Riazi-Esfahani, and M. Riazi-Esfahani, “Visual Prostheses: The Enabling Technology to Give Sight to the Blind,” *J. Ophthalmic and Vision Research*, vol. 9, no. 4, pp. 494-505, 2014.
[37] P.M. Lewisa, H.M. Acklanda, A.J Lowery, and J.V. Rosenfeld, “Restoration of vision in blind individuals using bionic devices: A review with a focus on cortical visual prostheses,” *Brain Research*, vol. 1595, pp. 51-73, 2015.
[38] A. Barriga-Rivera, L. Bareket, J. Goding, U.A Aregueta-Robles, and G.J. Suaning, “Visual Prosthesis: Interfacing Stimulating Electrodes with Retinal Neurons to Restore Vision,” *Frontiers in Neuroscience*, vol. 11, 620, 2017, doi:10.3389/fnins.2017.00620.
[39] *Bionic Vision Technologies*, http://bionicvis.com/. 2019.
[40] *Pixium Vision*, http://www.pixium-vision.com/. 2019.
[41] A. Reddy, “Bionic eye,” *Encyclopedia Britannica*, https://www.britannica.com/topic/bionic-eye. Dec. 2018.
[42] P.J. Choi, R.J. Oskouian, and R.S. Tubbs, ”Telesurgery: Past, Present, and Future,” *Cureus*, vol. 10, no. 5, e2716, 2018, doi:10.7759/cureus.2716
[43] M. Eto and S. Naito, “Robotic Surgery Assisted by the ZEUS System,” *Endourooncology*, H. Kumon, M. Murai, and S. Baba, eds., Recent Advances in Endourology, vol 6, pp. 39-48, Tokyo: Springer-Verlag, 2005
[44] M. Shenai, R.S. Tubbs, B. Guthrie, and A- Cohen-Gadol, “Virtual interactive presence for real-time, long-distance surgical collaboration during complex microsurgical procedures,” Technical note, *J. Neurosurgery*, vol. 121, pp. 277-284, 2014.
[45] *Endoscopy*, https://www.nhs.uk/conditions/endoscopy/. 2019.
[46] *Laparoscopy*, HealthLine, https://www.healthline.com/health/laparoscopy. 2019.
[47] J. Allen, “Photoplethysmography and its application in clinical physiological measurement,” *Physiological Measurement*, vol. 28, no. 3, pp. R1-R39, 2007.
[48] S. Cheriyedath, “Photoplethysmography (PPG),” https://www.news-medical.net/health/Photoplethysmography-(PPG).aspx. 2019.
[49] *Imaginis. Mammography*, https://www.imaginis.com/mammography. 2019.
[50] S.L Colyer, M. Evans, D.P. Cosker, and A.I.T. Salo, “A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system,” *Sports Medicine-Open*, vol. 4, no. 1, p. 24, 2018, DOI: 10.1186/s40798-018-0139-y.
[51] I. Ar and Y.S. Akgul, “A computerized recognition system for the home-based physiotherapy exercises using an RGBD camera,” *IEEE Trans. Neural Systems and Rehabilitation Engineering*, vol. 22, no. 6, pp.1160-1171, 2014.
[52] M.E. Tinetti, M. Speechley, and S.F. Ginter, “Risk factors for falls among elderly persons living in the community,” *New England J. Medicine*, vol. 319, no. 26, pp.1701-1707, 1988.
[53] A. Nalci, A. Khodamoradi, O. Balkan, F. Nahab, and H. Garudadri, “August. A computer vision based candidate for functional balance test” *Proc. 37th Ann. Int´l Conf. IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3504-3508, 2015.
[54] M.E. van Baar, J. Dekker, R.A.B. Oostendorp, D. Bijl, T.B. Voorn, and J.W.J. Bijlsma, “Effectiveness of exercise in patients with osteoarthritis of hip or knee: nine months' follow up,” *Annals of the Rheumatic Diseases*, vol. 60, no. 12, pp.1123-1130, 2001.
[55] A. Wijekoon, “Reasoning with multi-modal sensor streams for m-health applications,” *Workshop Proc. 26th International conference on case-based reasoning (ICCBR 2018)*, pp. 234-238, http://iccbr18.com/wp-c...BR-2018-V3.pdf#page=234. 2018.
[56] Y. Rybarczyk, J.L.P. Medina, L. Leconte, K. Jimenes, M. González, and D. Esparza, “Implementation and Assessment of an Intelligent Motor Tele-Rehabilitation Platform,” *Electronics*, vol. 8, no. 1, p. 58, 2019.
[57] J. Dorado, X. del Toro, M.J. Santofimia, A. Parreño, R. Cantarero, A. Rubio, and J.C. Lopez, “A computer-vision-based system for at-home rheumatoid arthritis rehabilitation,” *Int´l J. Distributed Sensor Networks*, vol. 15, no. 9, p. 1550147719875649, 2019.
[58] J.P. Salisbury, R. Liu, L.M. Minahan, H.Y. Shin, S.V.P. Karnati, S.E. Duffy, N.U. Keshav, N.T. Sahin, “Patient Engagement Platform for Remote Monitoring of Vestibular Rehabilitation with Applications in Concussion Management and Elderly Fall Prevention,” *Proc. IEEE Int´l Conf. Healthcare Informatics (ICHI)*, pp. 422-423, 2018.
[59] *Vojta Therapy*, Internationale Vojta Gesellschaft e.V., https://www.vojta.com/en/the-vojta-principle/vojta-therapy. 2019.

[60] M.H. Khan, J. Helsper, M.S. Farid, and M. Grzegorzek, “A computer vision-based system for monitoring Vojta therapy,” *Int´l J. Medical Informatics*, vol. 113, pp. 85-95, 2018.
[61] Y.-L. Chen, C.-H. Liu, C.-W. Yu, P. Lee, and Y.-W. Kuo, “An Upper Extremity Rehabilitation System Using Efficient Vision-Based Action Identification Techniques,” *Applied Sciences*, vol.8, no. 7, 1161, 2018, doi:10.3390/app8071161.
[62] F. Stroppa, M.S. Stroppa, S. Marcheschi, C. Loconsole, E. Sotgiu, M. Solazzi, D. Buongiorno, and A. Frisoli, “Real-Time 3D Tracker in Robot-Based Neurorehabilitation,” *Computer Vision and Pattern Recognition, Computer Vision for Assistive Healthcare*, M. Leo and G.M. Farinella, eds., Academic Press, pp. 75-104, 2018, doi:10.1016/B978-0-12-813445-0.00003-4.
[63] P. Peer, A. Jaklic, and L. Sajn, “A computer vision based system for a rehabilitation of a human hand,” *Periodicum Biologorum*, vol. 115, no. 4, pp. 535-544, 2013.