

Review

Perspective and Evolution of Gesture Recognition for Sign Language: A Review

Jesús Galván-Ruiz ^{1,2}, Carlos M. Travieso-González ^{1,2,*} , Acaymo Tejera-Fetmilch ¹,
Alejandro Pinan-Roescher ¹, Luis Esteban-Hernández ³  and Luis Domínguez-Quintana ² 

¹ IDeTIC, Universidad de Las Palmas de Gran Canaria, 35017 Las Palmas de G.C., Spain; jesus.galvan101@alu.ulpgc.es (J.G.-R.); acaymo.tejera101@alu.ulpgc.es (A.T.-F.); alejandro.pinan101@alu.ulpgc.es (A.P.-R.)

² Signals and Communications Department, Universidad de Las Palmas de G.C. (ULPGC), 35017 Las Palmas de G.C., Spain; luis.dominguezquintana@ulpgc.es

³ Aries Research Center, Universidad Antonio Nebrija, 28015 Madrid, Spain; lesteban@nebrija.es

* Correspondence: carlos.travieso@ulpgc.es

Received: 18 May 2020; Accepted: 22 June 2020; Published: 24 June 2020



Abstract: This review analyses the different gesture recognition systems through a timeline, showing the different types of technology, and specifying which are the most important features and their achieved recognition rates. At the end of the review, Leap Motion sensor possibilities are described in detail, in order to consider its application on the field of sign language. This device has many positive characteristics that make it a good option for sign language. One of the most important conclusions is the ability of the Leap Motion sensor to provide 3D information from the hands for due identification.

Keywords: gesture recognition; algorithms; Leap Motion; pattern recognition; EMG; RFID; gloves; Wi-Fi

1. Introduction

Gestures are a component of non-vocal communication, in which body language sends specific messages instead of or in addition to speech. They are made through the movement of different body parts, i.e., typically hands, arms and face.

Hand gestures are the most expressive and most frequently used, as they seem to be more natural and intuitive gestures for people in general. Their impact on the message of speech is extremely high. For example, within the context of emotional messages—according to researcher Albert Mehrabian—both signals and gestures can carry up to a 55% of the message impact [1].

Humans have been using this type of nonverbal communication for over a million years, and even some of these basic signals have become universal gestures: moving one's head to affirm or deny something, frowning to express annoyance, shrugging one's shoulders to indicate that something is not being duly understood, or even those gestures inherited from the animal kingdom, such as showing the teeth to express aggressiveness. Nonverbal communication has developed in such a way (see Figure 1), that some languages have been created from a series of gestures. Some examples are the sign languages used among deaf people or the universal signs used for aviation, diving or first aid.

Nowadays, technologies form a natural part of daily human life, thanks to a rising number of applications. The growing capacity in terms of computing power and speed for processors and sensors (Table A1), enables the existence of increasingly complex applications (Table A2) (Figure A1). Hence, the need for a more efficient human-machine interaction (HMI) is also proportionally increasing.

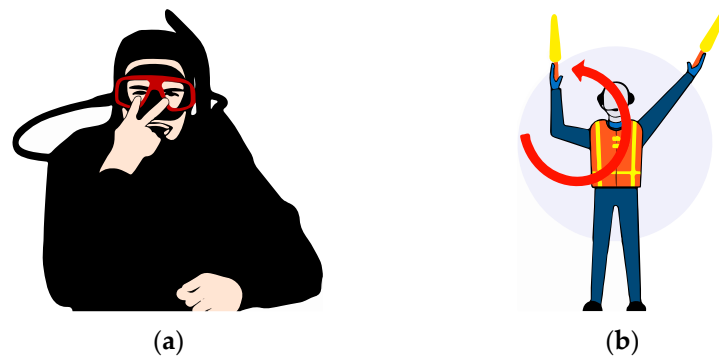


Figure 1. Sign languages for different uses. (a) Diving Sign. (b) Aviation Sign.

Within this context, on one side, the use of touch screens and voice interfaces is currently becoming consolidated, whereas on the other side, the keyboard-mouse exchange in favour of more natural and intuitive interfaces is still far from being a reality. This search for more direct interfaces has become an interesting option in those situations where the use of hands can represent some kind of inconvenience or difficulty when it comes to manipulate an input device directly. For example, when a driver wants to check the navigation system of his/her vehicle without taking his/her hands off the steering wheel, or when a person in the middle of a meeting wants to act discreetly with a computing device. Moreover, direct physical manipulation might spread microbes. Hence contact-free input mechanisms might become crucial in those spaces that cannot be contaminated in any way whatsoever, for instance, surgery rooms.

In this sense, gesture recognition may be considered useful in this path of increasing the interaction between machines and humans even more (Table A2). Since the rise of digital video capture technologies, there have been some attempts to recognize dynamic gestures for different purposes. Moreover, the update of new technologies such as depth sensors or high-resolution cameras enables the development of several ways to detect different movements and act in real time.

These types of data captures can be applied in multiple areas, such as patient monitoring, virtual and augmented reality navigation and manipulation, home automation, robotics, vehicle interfaces, PC interfaces, and lexicon translation of sign languages, among others (Table A2).

By definition, all these interactions consist of making gestures understandable for computers. The position of the human body, its configuration (angles and rotations) and its movements (speed or acceleration) must therefore be adequately detected. To perform this task, there are two main different approaches: the non-optical approach consists of using sensor devices such as gloves or bracelets, whereas the optical approach relies on the use of camera vision techniques. Both methodologies present a range of variations in terms of accuracy, calibration complexity, resolution, latency, range of motion, user comfort and costs.

Furthermore, it is necessary to make a clear distinction between gestures and poses. To start with, a pose corresponds to a static gesture, for example, a hand with no movement at all. In contrast, the gesture is a dynamic act formed by a sequence of connected poses during a short period of time, for example waving goodbye with a hand. Hence, gesture recognition might be challenged in two levels: a lower one demanding an adequate pose detection and a higher level regarding adequate gesture recognition.

Apart from this differentiation, interpreting both poses and gestures is not a simple task for various reasons. Firstly, there are usually several concepts associated with one same gesture and vice versa, and secondly, gesture meanings can vary between different languages, cultures and contexts. For example, sign languages used by deaf people vary between countries (sometimes even within the same country), these languages mix gestures and poses, and logically they might also vary depending on the conversation context.

2. Evolution of Gesture Recognition Devices

As mentioned above, in terms of human-machine interaction, gestural recognition is considered as the most intuitive and natural, so its development is constantly evolving, depending on the improvement of sensors used to capture gestures (Figure A1). In this sense, gesture recognition has essentially evolved from an intuitive recognition, to a more formal recognition based on the improvements from experiments on sensors used for this purpose. Initially, during the 1960s, researchers started using tablets and special pencils that captured writing (sketchpad [2]), touch-sensitive interfaces or pointing devices (see Figure 2). Later, in 1969, engineer Myron Krueger started working with virtual reality prototypes, and claimed that a future was possible without screens, in which people would interact directly with their environment [3]. This approach would later be referred to as natural interaction, and these interaction interfaces as natural user interfaces (NUI) [4].

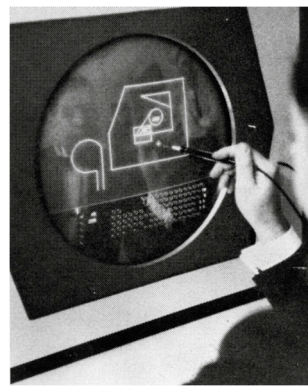


Figure 2. Written on sketchpad in 1963.

It would not be until the 1980s when movement capture by wearing gloves with sensors for flexion and position would start. In the 1990s, works on identifying gestures in images and video with computer vision methods increased, something that has been improving until today where real-time and people tracking systems are available [5,6]. Apart from computer vision, there are also electromagnetic systems, which locate an object position by measuring the electromagnetic fields generated by a transmitter, for example, radiofrequency.

It would not be until the 1980s when movement capture by wearing gloves with flexion and position sensors started. In the 1990s, work on identifying gestures in images and video with computer vision methods was increased, and has been improving until today, where real-time and people-tracking systems are already available [5,6]. Apart from computer vision, there are also electromagnetic systems, which can locate an object position by measuring the electromagnetic fields generated by a transmitter, for example, radiofrequency.

2.1. Data Gloves

These devices are worn on the hands in order to measure their position and their movements. They usually have either tactile or other types of sensors incorporated, capable of producing an exact mapping of the movements performed by all different phalange and wrist joints. This has the great advantage that no data processing step for obtaining descriptors is needed, as would happen in the case of an image obtained with a camera, for instance. Its disadvantage is that they are expensive and uncomfortable for the user.

They can be divided into two distinct categories: actives and passives [7] (Figure 3). On one hand, active systems include all such gloves that include some type of sensor or accelerometer. Formerly, these types of data gloves are connected to the computer by cables, a feature that has already disappeared thanks to wireless technologies. On the other hand, passive or non-invasive gloves refer to non-electronic devices, which include certain colour markers in order to facilitate image identification.



Figure 3. Passive gloves to help differentiate finger position.

The first type to appear was the Sayre Glove in 1977 and based its operation on flexible tubes (instead of optic fibre) to measure finger flexion by emitting a beam of light at one end, while a sensor (photoelectric diode) detected the intensity of that emitted beam according to the flexion and position of the fingers.

In 1980, MIT Media Lab started using gloves similar to Sayre's, only they were more interested in developing a movement capture device, rather than a control instrument. Unfortunately, the technology was not yet developed enough to obtain a truly effective input device, so the use of this glove shortly came to an end [8].

In 1983, the first glove that actually recognized hand positions with the intention of creating alphanumeric characters was patented, and the first virtual reality glove was marketed only four years later [9]. Physically, this DataGlove consisted of a lightweight glove with optical fibres attached to the back of the fingers. In this model, the flexion of the fingers bends the fibres, thus attenuating their transmitting light. The signal intensity for each fibre is duly sent to a processor, which then determines the angles of the joints based on the pre-calibrations for each user.

Later, in 1989 and 1995, the Power Glove and Super Glove appeared, and they both used materials of variable electric resistance in order to measure the finger flexions. These were cheaper but at the same time less accurate [10]. Once the use of magnetic sensors and accelerometers were added, numerous examples of these devices started to appear, both for commercial and scholar purposes. Currently, the CyberGlove [11] (Figure 4) is the most acknowledged model, and has been particularly developed for the graphic animation and film industries. In addition, the MIT AcceleGlove [12] is being widely used for virtual reality frameworks.



Figure 4. CyberGlove III mainly used to control robots.

The evolution in experiments on data gloves has enabled their use in different fields, ranging from classic applications such as sign understanding, or entertainment and virtual reality, to other fields such as medicine [13], robotics or industrial production [14].

2.2. Electromyography (EMG) Electrodes

Exactly the same way as active gloves, bracelets also try to avoid environmental interferences that voice interaction or computer vision systems might suffer. However, its operation is not based on accelerometers or optical fibres; it relies on electromyography (EMG) sensors. This means that it seeks to measure the electrical potentials produced by muscle cell activity.

Among these recognition technologies, active research is being conducted on studies based on bio-signal recognition technology: electrooculography signal (EOG: electrical activity of the eyeball movements or “gaze position” recorded around the eyes), electroencephalography signal (EEG: electrical brain activity recorded from the scalp), electrocardiogram signal (ECG: electrical activity of the heart), EMG and others.

Electromyography is based on the study of the neuromuscular system, so it is responsible for detecting, analysing and processing the electrical signals produced by muscles and nerves, by using electrodes.

Muscles are composed of a set of specialized cells capable of contracting and relaxing, so they are responsible for functions such as producing movement, transporting substances through the body or providing adequate stabilization and temperature [15]. In addition, three types of muscle tissues can be differentiated according to their structure, contractile properties and control mechanisms: (a) skeletal muscles, (b) smooth muscles and (c) heart muscles. Skeletal muscles—according to its own name—are attached to the skeleton and facilitate body movement, while smooth muscles are inside the intestines and are responsible for the transport of substances, and finally, heart muscles are in charge of producing heartbeat. The EMG signal is obtained from skeletal muscles [16].

It must be noted that bioelectric potentials generated by the organism are actually ionic potentials produced by flows of ionic currents. The efficient measurement of these ionic potentials requires that they must be therefore converted into electric potentials, before they can be measured with conventional methods. Devices responsible for this conversion are the electrodes obtaining intramuscular or surface voltages [17]. EMG electrodes are capable of recording potentials from all the muscles within their reach. This means that the potentials of nearby large muscles can interfere during attempts to measure the EMG of small muscles, even when the electrodes are placed directly on such small muscles. Therefore, in these cases, the obtained signals are usually the result of a sum of all these individual potentials created by the fibres that form up the muscle or muscles that are being measured. When this might cause a problem, needle electrodes directly inserted into the muscle are required, as these can access muscle fibres individually. Consequently, intramuscular or needle electromyography is used in order to study the physiology or pathologies of motor units, whereas the surface type (SEMG) is more suitable for studies on types of muscular behaviour, patterns of temporal activity or muscular fatigue [18].

Factors such as the distance between electrodes, the area of muscular activity, the skin properties, the signal processing and the contact that occurs between the skin and the electrode are essential for analysing EMG signals, since amplitude and its properties depend directly upon all these.

Typically, signals obtained by the electrodes must be amplified and processed (Figure 5). At the same time, signal amplitude and bandwidth might vary depending on the size, type of electrodes, and the spacing between them [19].

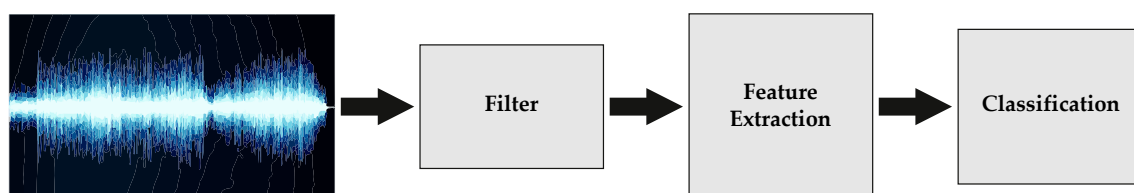


Figure 5. Different phases for the treatment of electromyography (EMG) signals.

Amplitude of an EMG signal ranges from 10 μV to 5 mV and its bandwidth from 10 Hz to 10 KHz, depending on the electrode configuration. Commercially, it uses silver–silver chloride electrodes (Ag–Ag Cl), due to its stability and noise reduction.

As for the number of electrodes needed, a minimum of three is required. Two of them comprise a differential input pair and the third one corresponds to the grounding. Sometimes, active or activated electrodes are added in order to create a feedback control circuit between the sensor and the body [20], thereby eliminating the need of using conductive gel between electrodes and the skin (its use decreases the electrical impedance).

Traditionally, EMG technology has been employed for medical diagnosis [21], prosthetic control [22] and medical rehabilitation technologies [23]. Within the human–computer interaction (HCI) context, a series of studies have emerged focusing on the use of EMG inputs as an alternative channel for users with physical disabilities, given that the activity in one or more muscles can be interpreted in a Graphical User Interface (GUI) both on a continuous basis (as a substitute for the mouse) or on a static basis by gesture recognition [24–27]. Other examples of EMG applications for HCI include robotic control [28–30], non-voice speech recognition [31], emotional state recognition [32,33], musical expression interfaces [34–36] and generic gestures recognition, using these for either controlling a music player [37], as a GUI pointer [38] or as a numeric keypad [39,40].

Regarding existing products in the market, due to recent developments in wireless communications and integrated computing, we can find widespread portable devices used to obtain EMG data using bracelets [41,42]. These include multiple EMG sensors radially positioned around the circumference of a flexible band, enabling a relatively comfortable fitting. An outstanding example might be the Myo bracelet produced by Thalmic Labs company, which has been used in many applications and experiments [43–47].

2.3. Ultrasound

Two techniques can be distinguished for gesture capturing. One of them uses ultrasound images (sonomyography), and the other that takes advantage of the Doppler Effect, profiting from the ultrasound radiation in any room. This technique, opting for the use of ultrasound images, enables the possibility of the observation of the muscles inside the body on a real-time basis.

This is achieved thanks to the fact that our tissues bear different acoustic impedances, so when a sound wave passes from one to another with a different acoustic impedance, different amounts of energy are reflected, thus forming an ultrasound image. It will unlikely provide greater accuracy than other systems measuring directly the moving parts of the body, such as cameras or data gloves, but instead, as in EMG techniques, potential benefits concerning information losses due to occlusion can be obtained, since in this case, any part of the body does not remain hidden by others. Additionally, and in contrast with EMG, sonomyography does not experiment with the difficulty of differentiating between individual muscles (for non-invasive techniques), and those lying deeper in the forearm, which limit the soundness of EMG techniques for capturing a variety of wide hand movements [48].

Various studies have emerged in this regard: from Hodges in 2003, measuring muscle contraction of several muscles [49], to Zheng and Chen in 2006, 2008 and 2010 [50–52], or Mujibiya, which in 2013 put forward a band of ultrasonic transducers around the forearm and fingers [53], Hettiarachchi which differentiated six gestures in 2015 [54] and McIntosh in 2017, which produced a system that detects and monitors a series of finger gestures using a probe fitted on the forearm [55].

Another technique using ultrasonic frequency signals for gesture capturing is based on the management of the Doppler Effect. It essentially consists of a transmitter that emits ultrasonic continuous tones, which bounce off of objects (arms and hands in gesture recognition) within the detection field. Reflected signals returning to the sensor are then captured for their due analysis. In a static environment, the returning signal bears the same frequency, but has a different phase and amplitude. However, if an object moves, its echoes shift their frequency (Doppler shift), thus creating components in other frequencies which are proportional to their speed in relation to the sensor.

The Doppler shift caused by reflection from a moving target is approximately $fd = 2vf/(cv)$, where fd is the frequency offset, v stands for the target's velocity (in the direction of the sensor), f shows the emitted frequency, and c represents the speed of sound [56].

In the case of movements performed by several objects with different speeds, the return signal will contain multiple frequencies, one for each moving object.

This is the same principle of radar operation used by the police for speed controls. However, ultrasound is used here instead of RF, and the ultrasonic sensor observes the full velocity profile in the field of view. In comparison, police radars can only detect the speed of a single object [57].

Although ultrasonic ranges have been considerably exploited in the market [58], research on this technique applied towards ultrasound gesture recognition has not been widely reported.

The first accounts in this area date back to 2009, when hand recognition was developed and was considered an important issue according to [59], in particular, when it is based on the signal power being reflected into three receivers (3D). A year later, a simple ultrasonic gesture recognition system for music control [60] was reported. In 2011, Kreczmer reported gesture recognition using the conventional ultrasonic range finder in a mobile robot [61]. Shortly thereafter, Microsoft Research and the University of Washington developed a nearby gesture ultrasonic recognition system for laptops [62]. That same year, a Berkeley group designed an ultrasonic gesture recognition system that uses a matrix with a large number of capacitive ultrasonic transducers on a single chip [63,64]. This research line continued developing, looking for a more compressed hardware with less spaced sensors, like the works in 2014 by Booji [65], or in 2017 by [66] or Yu Sang in 2018 [67].

2.4. WiFi

With the aim of re-using existing Wi-Fi infrastructures, which would save the corresponding employment of a hardware and with the possibility of an easy large-scale deployment of systems, the search for gesture recognition through Wi-Fi networks (Figure 6) began. An additional advantage of this technique, which shall be furtherly detailed in the cameras section, consists of its capacity for providing gesture recognition even under non line of sight (NLOS) scenarios (without direct vision).

Moreover, different indicators are used on these type of systems to obtain all the information necessary for gestural recognition. For instance, some research studies are based on the received signal strength indicator (RSSI) [68], on the signal flight time indicator (ToF) [69] or on Doppler shifts [70]. However, these systems require either specialized devices [69,71], or need modifying by already-existing commercial devices [70], and some of them are excessively susceptible to interference [68]. Consequently, research has currently focused more on the observation of the channel status information (CSI). This information provides more accurate indications as it comes from a single signal flow, due to the orthogonal frequency division (OFDM). As a result, CSI has proven to be a more robust indicator even with indoor interferences [72].

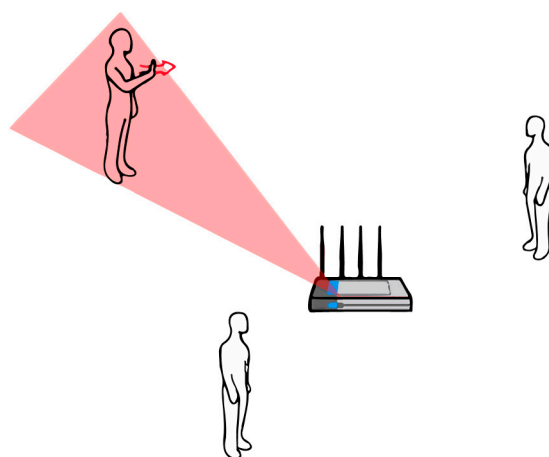


Figure 6. WFI recognition of different hand positions.

Based on the RRSI use, Abdelnasser et al., [73] created a gesture recognition system using Wi-Fi (WiGest). This system can identify several hand gestures and map them for monitoring various actions, with accuracy rates ranging between 87.5% and 96%. For this task, it uses a single simple access point and three aerial ones. Other researchers like Pu et al. [70] introduced the WiSee system, which can recognize nine body gestures that interact with Wi-Fi-connected home devices, by taking advantage of Doppler changes in wireless signals. On the same line, Adib et al. introduced WiVi [71], Witrack [69] and WiTrack0 [74]. These systems can track human movement through the walls and classify simple hand gestures. Concerning CSI, Nandakumar et al. [75] presented a four-gesture recognition system with a 91% success in LOS scenarios and with 89% in NLOS. In contrast, El WiG of He et al. [76] classify the same number of gestures but with a 92% and an 88% success respectively. Some other research works include more complex gestures, such as the WiFinger by Li et al. [77] which identifies considerably subtle finger gestures with a success rate of 93%.

2.5. Radio Frequency Identification (RFID)

RFID systems generally consist of ultra-high frequency (UHF) commercial readers, which are capable of detecting labels within a range of even a few dozen meters, depending on the transmitted power. Identification tags can either be carried by a person or not, depending on the RFID application, and they can work with or without a battery (active and passive). Normally, RFID tags have been used for supply chain management and automatic object identification [78,79]. They also have been used to increase the digital information environment [80], to monitor indoor human activities [81], or to detect human interactions with RFID-tagged objects [82]. Just like in the case of Wi-Fi signals, relevant information such as phase changes, RSSI and Doppler shifts [83] can be collected from UHF signals, thus providing potential possibilities for gesture recognition.

An advantage of using passive RFID technology is its inexpensive costs but—logically, since they do not count with their own energy supply—detection ranges are reduced to centimetres. Due to this inherent setback, related gesture detection research is very focused towards detecting interactions between humans and objects or tactile interfaces [84–87]. Active tags, due to their long range, are normally used in traceability, location and identification applications.

2.6. RGB and RGB-d Cameras

Nowadays the most widely used devices for gesture captures are definitely RGB and RGB-d cameras [7]. Current digital cameras capture the light in three main channels of red, green and blue light (hence the acronym red (R), green (G) and blue (B)). The resulting image is organized in an array of pixels, each one bearing its three RGB values. Therefore, any attempt to identify an object within

the image will result in a computer-vision algorithm, in which, essentially, the order and intensity of colour of the different pixels will provide a meaning.

Depth cameras, also known as RGB-d, are a variation of conventional cameras. These have a fourth channel for showing the depth information, i.e., the distance between the focus and the captured object. In order to obtain this depth information, and therefore produce 3D images, three clearly differentiated techniques are used.

2.6.1. Stereoscopic Vision

This is achieved by placing two cameras at par values to obtain two different views of the same scene. This process is similar to human binocular vision procedures. In order, however, to obtain such depth, a series of calculations must be performed on the obtained pair of images. Since two pixels actually correspond to one same point of the scene, its depth can be measured by knowing the system calibration conditions of such parameters, like the focal length or the separation between both camera centres [88].

2.6.2. Structured Light

This consists of an active illumination of the scene by projecting a pattern that varies spatially in the x and y coordinates. Therefore, it is considered a modification of a stereoscopic vision, which replaces the use of a second camera in favour of a light source [89].

2.6.3. Time of Flight (TOF)

These cameras can estimate the 3D structure directly, without the help of any artificial vision algorithms. Its operation is based on the emission of a modulated light source, usually within the infrared (IR) spectrum, which illuminates the scene. Distance is obtained by measuring the flight time of the signal that travels from the light source, collides at the scene and returns to the sensor element [90].

The problem of designing a system capable of recognizing gestures for controlling applications usually arise in the following stages (Figure 7): data collection and processing (camera capturing, processing and segmentation), modelling of the captured object (selection and management of object features), definition of gestures (static and dynamic), their detection (posture separations and motion detection) and transition control (state machines).

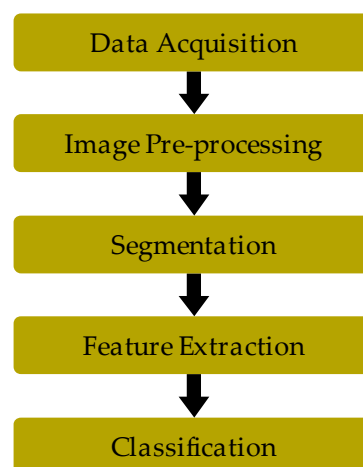


Figure 7. Stages of vision-based systems.

Each of the previous sections produces a range of research studies, depending on the processing modality, chosen segmentation and classification, and the possible use of multiple cameras to deal with the problem of objects occlusion (hands and arms in the case of gesture capture) [91,92].

Some outstanding examples are Cohen and Li [93], which classify body postures with the support vector machine (SVM) technique of a 3D visual helmet, built from a set of data entry. This system returns classified positions of the human body as miniature images. Moreover, ChengMo et al. [94] have pointed out a human behaviour analysis by using a system that recognizes human postures such as walking, bending down or sitting. Moreover, Corradini et al. [95] could analyse and recognize positions by using hybrid neural networks. Furthermore, Raptis et al. [96] for instance, presented a classification system which recognizes dance gestures based on the movement of six joints, collected on a real time basis by a Kinect camera with a 96.9% accuracy.

Although these 3D cameras have been used for gesture recognition systems, traditionally they have been expensive [97]. Since the appearance of the Kinect Microsoft device, costs have been significantly reduced, thus becoming a very attractive option for many researchers [98]. The system returns the classified positions of the human body as miniature images [99].

Recently, Microsoft released a new version of the Kinect device, called Azure Kinect DK. Although the previous version was mainly focused on games; this new device is focused on professional use, covering markets such as logistics, robotics, and healthcare, among others. Azure Kinect DK is a developer kit and a PC peripheral that uses advanced artificial intelligence (AI), sensors for sophisticated computer vision and speech models. It combines a depth sensor and spatial microphone array with a 4K RGB video camera and an inertial measurement unit. The whole device is applied with multiple modes, options, SDKs, and Azure cloud services support.

Microsoft has also provided a new skeleton tracking kit that detects the 3D coordinates of human body joints. The body-tracking features provides body segmentation, an anatomically skeleton for each partial or full body in the field of view, a unique identity for each body, and can track bodies over time.

It is a device with a high cost that requires fairly considerably minimal requirements. Additionally, it is a recent device and no scientific publications related to sign language have been found, but rather some proprietary solutions from private companies.

2.7. Leap Motion

The Leap Motion [100] (Figure 8) is a compact and affordable recognition device, which is capable of 3D tracking forearms, hands and fingers in real time. Its configuration is oriented to interactive software applications.



Figure 8. Leap Motion operating mode.

It contains two infrared cameras (Figure 9), with an angle of 120° and three infrared LEDs, so it would be classified like a ToF camera, according to the previous section on camera types. These sensors work at a wavelength of 850 nanometres, i.e., they work within a non-visible spectrum for the human

eye, and use sample speeds of up to 200 fps (frames per second), adapting their lighting to detected light, clarifying the flood detection zone and ensuring a constant image resolution.

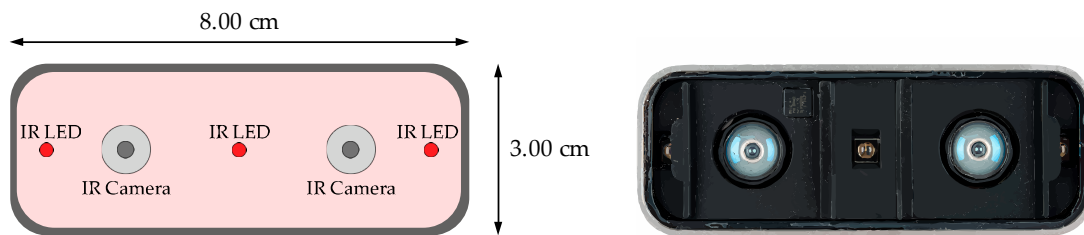


Figure 9. Different components of the Leap Motion hardware system.

As can be observed, small plastic barriers are placed between these LEDs and the cameras, in order to avoid interference and saturation, and to achieve uniform illumination throughout all covered areas.

In addition, sensors are CMOS type, which means that no external electronics are required, as the digitization of each pixel occurs within each cell. All this results in a faster capture speed, using less hardware space [101].

Moreover, the microcontroller (Macronix MX25L3206E [102]) performs tasks such as lighting regulation, obtaining the information collected by the sensors and sending it to the USB controller.

Additionally, the USB controller controls the data transfer to the computer to which the device is connected. It is high speed and can support USB 3.0. Moreover, sent data are received by the computer controller through two serial ports, UART_Rx and UART_Tx.

The interaction area (Figure 10a), forms a semi-hemisphere with a viewing angle of 150° and with an effective range that extends from 25 mm to 600 mm. All these parameters delimit the interaction scope (interaction box). The interaction height range (see Figure 10b) can be configured between 70 mm and 250 mm.

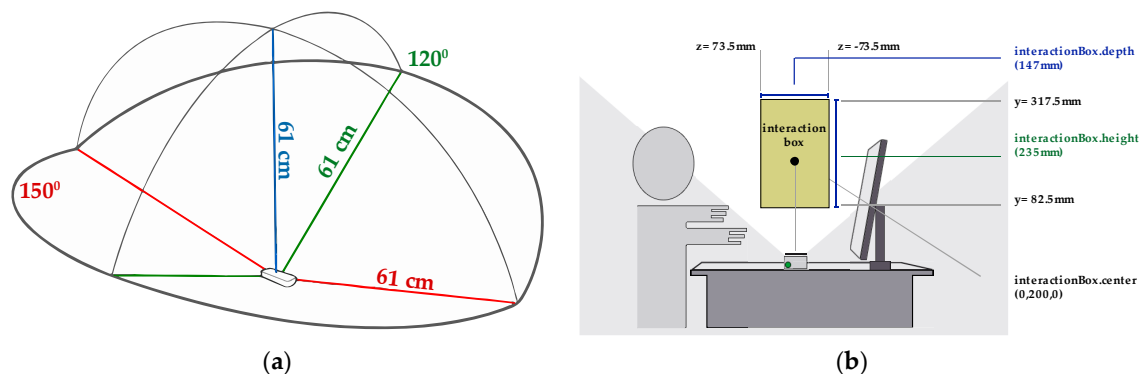


Figure 10. (a) Interaction area of Leap Motion. (b) Interaction box of Leap Motion.

This area can be influenced by several factors such as the camera viewing angle or the maximum intensity that the USB can deliver. It should be noted that the viewing angle (1) is also affected by the focal length and sensor size, in the following way: “ d ” is the sensor diagonal and “ f ” stands for focal length.

$$\alpha = 2 * \tan^{-1}\left(\frac{d}{2f}\right) \quad (1)$$

It has a USB connection and is compatible with Windows, Linux and Mac operating systems. Furthermore, the company provides APIs for developers to program this device in Python, Java, C++, C#, Objective-C and JavaScript, finally evolving to Leap C, a C style to access tracking data of the Leap Motion service. It also distributes plugins for the Unreal Engine and Unity graphics engines.

Here, the controller analyses each image looking for hands and fingers, and sends a set of frames to the computer via USB. Each frame object (Figure 11) contains Hand-type classes with its correspondent Arm subclasses (a bone-like object that provides the orientation, length, width and endpoints of an arm) as well as Finger subclasses, which contain data concerning identifying types, direction position, orientation, speed, length and width, angle between them and even percentage of coincidence with the internal model contained by this device. Besides all of these parameters, it can also recognize certain predefined dynamic gestures (Gesture) and tools (Tool) [103].

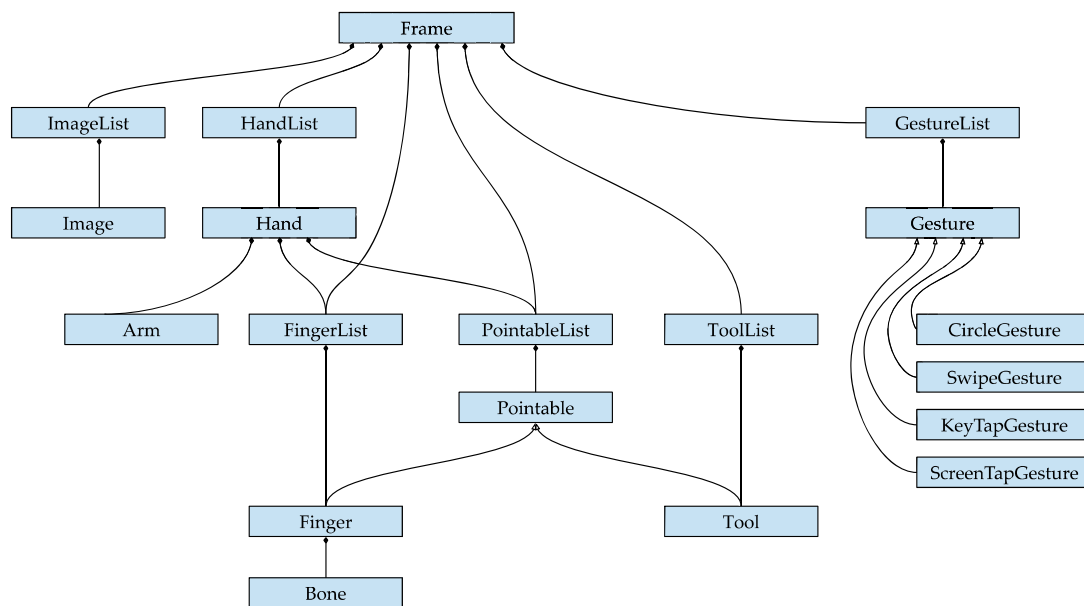


Figure 11. Frame Object of Leap Motion Source.

The internal model of a human hand is used to provide a predictive tracking even when parts of a hand are not visible (due to occlusion). The software uses the visible parts of the hand, its internal model, as well as past observations to calculate the most probable positions of non-visible parts at any given time. A *Hand.confidence()* rating indicates how observed data fit the internal model [104].

Besides the hand skeletal tracking model (Figure 12a), Leap Motion also provides a screen, where the recorded sequences of a scene happening in front of the device can be observed in real time (Figure 12b). As mentioned earlier, this occurs thanks to an infrared radiation (IR) of the scene, where its reflection is captured by biconvex cameras. Data obtained by the sensors are stored on a digitized image.

As images are obtained, a distortion produced by the lenses can be observed (see Figure 13) on these optical images, thus deforming them. This distortion produced by the Leap Motion is known as complex distortion and is a mixture between barrel distortion and cushion distortion [105].

This distortion can be calibrated (Figure 14) by superimposing a mesh map of calibration points on the captured image. As pairs of image data are being sent to the microcontroller, each value is accompanied by its distortion information. Finally, the correct brightness value can also be achieved by using a calibration map [106].

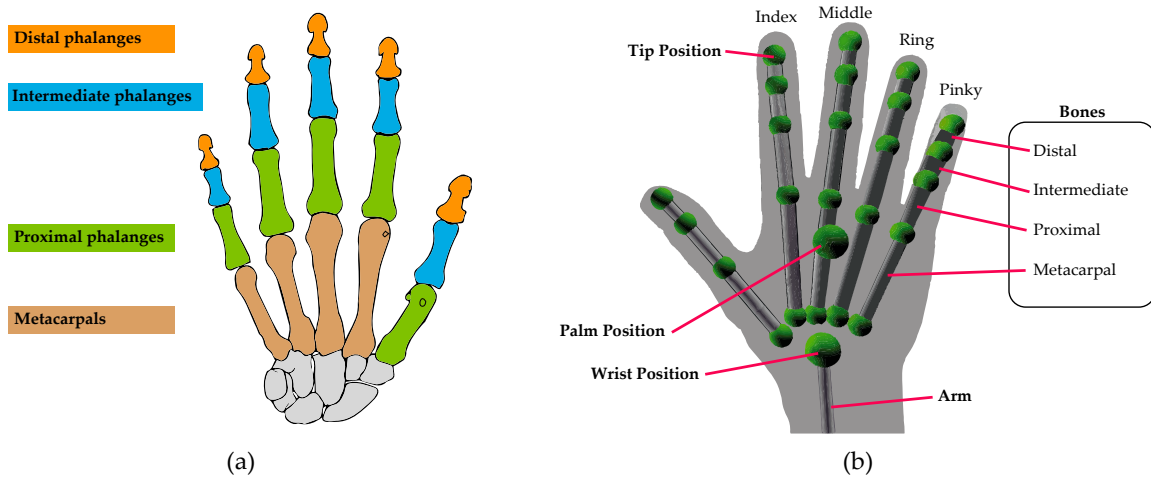


Figure 12. Tracking of the hands skeletal model (a) Object Bone. (b) Object Hand.

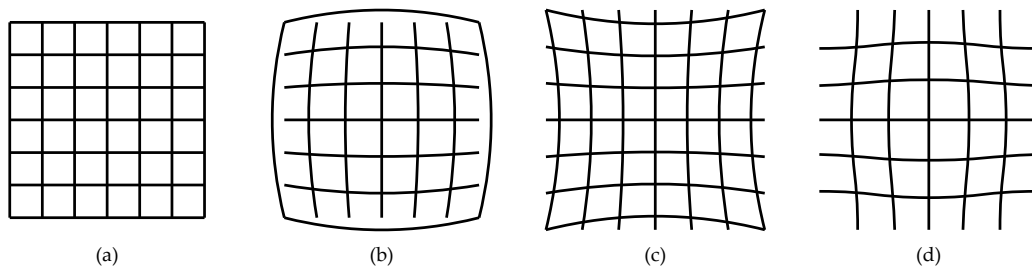


Figure 13. Distortions that lenses can produce (a) Without distortion. (b) Barrel distortion. (c) Cushion distortion. (d) Complex distortion.

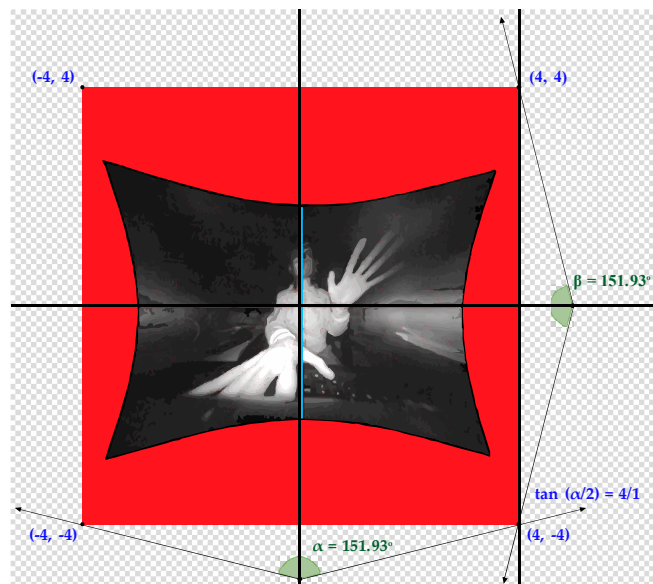


Figure 14. Image with the distortion corrected before reaching the microcontroller.

Once the corrected images have been obtained and the microcontroller has identified the hands and fingers in the interaction zone, their position in the coordinate system is determined through stereoscopic vision techniques [107].

There are different studies that have checked this device accuracy and its robustness [105,108–110]. It has been evidenced that it obtains better results than other contemporary devices of the Kinect or Myo types. Additionally, both static and dynamic measurements have also been performed to analyse

systematically the controller sensory space and define the spatial dependence of its accuracy and reliability. During static experiments, when objects move away from the device and are located at the right or left ends, a significant increase in standard deviation is observed. Concerning dynamic experiments, a significant drop in sample accuracy was observed when these were taken over 250 mm above the controller.

Due to the Leap Motion design, mainly oriented to applications such as virtual reality (VR), gaming or more natural interfaces use (NUI), we can find many publications in this regard [111–114].

It is not hard to imagine that the features of this device are suitable for research outside of the VR and gaming fields. This fact is evidenced by the large number of publications concerning both static and dynamic gesture recognition using Leap Motion. For example, in 2014, Mohandes et al. [115] managed to capture 28 letters of the Arabic sign alphabet with a 99% success rate, by using multilayer neural networks (MLP) with the Nave Bayes classifier. The following year, Chen et al. [116] using an SVM algorithm, captured 36 gestures (numbers and alphabet) with a success rate of 80%. That same year, McCartney et al. [117] achieved an accuracy rate of 92.4% in a database of over 100 participants, using convolutional neural networks and hidden Markov models. A year later, Wei Lu et al. [118] managed to recognize a set of dynamic data with a success rate of 89.5% by using neural networks as a classifier (HCNF).

Logically, in the same way as other technologies explained above, the use of Leap Motion has extended to other areas such as robotics [119], medical rehabilitation [120], home automation [121], identification and authentication [122], music [123] or education [124].

Currently, the same company that produces Leap Motion has released a new device called UltraLeap. This device is more accurate and powerful than the Leap Motion. It started to be commercialized a few months ago with a very high price. There are no scholarly papers yet concerning this device, nor any scientific application for it.

2.8. Virtual Reality (VR)

Nowadays, the advances based on resolution, which is included on the latest VR-based devices, help to create interesting interactive spaces for different fields, which need to be added to finger tracking, sensations or custom environments. This benefits applications that are aimed at medical rehabilitation, education, professional training and virtual practices of the architectural type, virtual test benches for industrial design and engineering, teleoperation of robots [125], etc. Efforts have been focused on creating spatial awareness, as real and accurate as possible, to create a full user experience. Therefore, users want to achieve the head-mounted display (HMD), a mounted screen that leads to the highest possible degree of presence. In the case that the HMD achieves this objective, it is understood that the performance of the trace is improved. This is explained in [126], where it proves that due to the Oculus Rift's support for Leap Motion, the user performance is significantly better in virtual reality settings. It requires combinations of sensor hardware and processing software, based on different detection technologies such as depth cameras, radio triangulation, ultrasound, IR reflectance of mobile screens, polarized light, or marker-based tracking with an IR camera. The tracked movements of the user are transformed into arbitrary coordinate systems and data formats; therefore, the VR environment serves as a kind of meta-layer to integrate and simulate different virtual tracking systems. [127].

The infinity of possibilities that VR offers has prompted companies like HTC, Oculus, Valve, HP, Microsoft and Sony to manufacture their own glasses. Striking examples are the alliance of HP, Valve and Microsoft to launch the HP Reverb G2 [128] to achieve 90 Hz resolution, which is four times higher than Oculus Rift [129] and HTC Vive [130]. Additionally, we highlight the Valve Index glasses [131] with their experimental 144 Hz mode, which increases the field of view 20° more than HTC.

3. Discussion

Since the Sketchpad in the 1960s to the present, there has been significant research and strong efforts regarding the production of devices that are capable of recognizing gestures. The most important

evolution has been observed since the 1980s, when improvements in semiconductor technologies enabled the creation of sensors recognizing different hand movements by transferring them to control systems. A fairly widespread example has been the use of different types of gloves, (DataGlove, PowerGlove, SuperGlove, CyberGlove, AcceleGlove). As devices, it should be noticed that these examples started being invasive and not very versatile, although it is true that later, their quality and accuracy improved, over the years. This improvement has been achieved not only due to the type of sensors that have been used, but also due to software development regarding information management.

Moreover, many video processing techniques have tried to recognize movement. Their main hindrance compared to other techniques is the fact that these always work in two dimensions and the recognition of gestures requires 3D systems.

In the second decade of the twenty-first century, the Kinect was released. This milestone represented an important revolution in the recognition of 3D gestures. The Kinect sensor belongs to the group of ToF cameras, and is able to detect the movements of a person and transfer—in real time—the contour of that person along with his/her movements. This device was not specially designed for the detection of hand movements; therefore, it is not precisely the most ideal device to be used with sign language. In addition, their size reduces their versatility considerably as a transportable device.

Since this mentioned decade, other types of components have appeared, such as the Leap Motion, included in the ToF component category, capable of recognizing the movements of the palm of the hand in 3D. This device was launched into the market with the aim of handling any type of environment simply by using hand movements. Research performed on Leap Motion tested the power of this device concerning the detection of different types of static gestures with percentages of up to 99% accuracy. Its dynamic use has also produced very high recognition percentages in different research studies. These properties of Leap Motion make it an ideal device for the recognition of sign language.

The rest of the technologies used for the recognition of gestures, for example, those related to the different types of sound waves (Wi-Fi, Ultrasound and RFID), are not able to detect the hand and finger movements with accuracy. In addition, there are no specific devices for these types of other technologies. Research in these fields mainly focuses on how the Doppler effect can be managed.

4. Conclusions

The evolution of different devices and their application for sign language purposes has been an active matter since the 1960s, and continues to be on a constant basis up until today. It is an open knowledge area where technological solutions are being increasingly approached, but it needs some more thorough consideration and development. Interest in this field is certainly significant, as well as the amount of studies regarding this research line.

There are many devices that are capable of recognizing gestures. Some of them have excessive complexity and size, and they are subordinate to using different active elements. An example may be the use of both active and passive gloves. Other devices have complex data post-management and some of them even have low recognition rates, also.

In contrast, the Leap Motion has many advantages over other devices. In the first place, this technology is very cheap, its usage and transport are easy too, it clearly recognizes your hands and you can work with it on multiple platforms. It can also be integrated into electronic devices and, most importantly, no active element is required. The manufacturer also provides all the necessary software to process—in a simple way—all the generated information.

Finally, their interaction area within their reach is more than sufficient, so that they can recognize, for example, sign language.

Funding: This research was granted and funded by “Fundación Indra” and “Fundación Universia”, under the award “Ayudas a Proyectos Inclusivos” in its 2018 Edition.

Acknowledgments: Authors show their acknowledgment to “ASOR Las Palmas” (Las Palmas de Gran Canaria, Spain), for its support, their shared knowledge and help provided by its staff.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Referenced Devices.

Ref.	Device	Authors	Year	Application	Focus	Results
[2]	Sketchpad	Sutherland	1963	Touch recognition	First touch interaction device	Opens up a new area of man–machine communication
[3]	Videoplace	Krueger et al.	1969	Virtual reality	Search for an alternative to the traditional human computer interaction (HCI)	Combines a participant’s live video image with a computer graphic world
[9]	DataGlove	Zimmerman	1983	Gesture recognition	The first invention of a glove with optical flex sensor	Signals from these sensors can be processed for application in kinesiology, physical therapy, computer animation, remote control and man to machine interface
[10]	PowerGlove	Mattel	1989	VR Gesture recognition	Device for Nintendo games	A low-cost alternative was created for researchers in virtual reality and hand and gesture posture recognition
[10]	SuperGlove	Nissho Electronics	1997	Gesture recognition	Pioneer use in sensors with a special resistive ink	Measures flexion of both the metacarpophalangeal and proximal interphalangeal joints for all four fingers and the thumb
[11]	CyberGlove	CyberGlove Systems	1990 (Currently on sale)	VR Robotic control Medical rehabilitation Gesture recognition	Gloves that use magnetic sensors and accelerometers	Programmable gloves with the greatest commercial impact
[12]	AcceleGlove	AnthroTronix	1999 (Currently on sale)	VR Robotic control Medical rehabilitation Gesture recognition	Gloves that use magnetic sensors and accelerometers	Programmable gloves with the greatest commercial impact
[43–47]	Myo	Thalmic Labs	2014 (Currently on sale)	VR Gesture recognition Home automation	Bracelet that reads and collects the electrical activity of muscles to control, wirelessly, other devices with gestures and arm movements	Myoelectric bracelet with greater commercial impact
[104,105,117]	Kinect	Microsoft	2010 (Currently on sale V2.0)	Gaming Gesture recognition Robotic control Medical rehabilitation Education Identification Authentication	Allows users to control and interact with the console without having physical contact with a traditional video game controller	More than 10 million sensors sold
[106,114–117, 120–131]	Leap Motion	Leap Motion Inc.	2012 (Currently on sale)	VR Gesture recognition Robotic control Medical rehabilitation Education Identification Authentication	Low cost device to interact with computer through gestures	Startup whose valuation reached 306 million in 2013 and was acquired for 30 million dollars

Table A2. Devices that work to recognize static and dynamic gestures.

Ref.	Device Type	Focus	Segmentation	Feature Extraction	Classification	Results and Observations
[4–7,10]		An overview of Gesture Based Interaction	Not applied	Not applied	Not applied	Advantages and disadvantages of various techniques that have emerged over time and in progress
[8]	DataGlove	A survey of glove-based input	Not applied	Not applied	Not applied	Achieving the goal of the “contactless” natural user interface (NUI) requires progress in many areas, including improving speed and accuracy of tracking devices and reducing manufacturing costs
[13]	DataGlove	Data-glove for medical use	Not applied	Not applied	Not applied	The use of sensorized surgical gloves in Hand-assisted laparoscopic surgery (HALS) surgery provides the surgeon with information additional tissues and organs
[14]	DataGlove	Design of a haptic control glove for teleoperated robotic manipulator	Not applied	Not applied	Not applied	Efficiency greater than 90% in opening and closing movements
[33]	Electromyogram data (EMG)	Cybernetic control of wearable computers	Not applied	Skin conductivity signals	Skin conductance sensor with a matched filter	The startle detection algorithm is robust in several users, and allows the wearable to respond automatically to events of potential interest
[34–37]	Electromyogram data (EMG)	An interface based on active EMG electrodes	Not applied	Root Mean Square (RMS) power calculations filtering	Not applied	Control of synthesizers, sequencers, drum machines or any another musical instrument digital interface (MDI) device [34,35,37]. Expanded to a multimodal interaction [36].
[38]	Electromyogram data (EMG)	A biological feedback GUI pointer (electromyograms)	Not applied	Not applied	Neural network	Detects the electromyograms of four of the muscles used to move the wrist, with performance similar to a standard mouse
[39,40]	Electromyogram data (EMG)	Capture of electromyographic signals as input of joysticks and virtual keyboards	Not applied	Not applied	Not applied	Low-cost brain-computer interface for the disabled people and the elderly
[41]	Electromyogram data (EMG)	Interpretation of forearm electromyography and finger gesture classification	31 milliseconds temporary segments	Root Mean Square (RMS). Fourier transform (FFT). Phase Coherence	SVM (Support Vector Machines)	Precision of finger gestures and arm postures above 80%
[42]	Electromyogram data (EMG)	Myoelectric control prosthetics and therapeutic games using Myo Armband	Not applied	Not applied	Not applied	Helps to improve the low myoelectric prosthesis adoption rate
[43]	Electromyogram data (EMG)	Hand posture and gesture recognition using Myo Armband	Ordered subspace clustering (OSC)	Fourier transform (FFT)	Collaborative representation classification (CRC)	Precision greater than 97% in six hand gestures and various postures

Table A2. Cont.

Ref.	Device Type	Focus	Segmentation	Feature Extraction	Classification	Results and Observations
[44]	Electromyogram data (EMG)	Provide guidelines on the use of Myo Armband for physiotherapy analysis by doctors and patients	Not applied	Not applied	Not applied	It is shown that the MYO device has the potential to be used to understand arm movements during the physiotherapy stage
[45]	Electromyogram data (EMG)	Use of electromyogram data (EMG) provided by the Myo bracelet to identify letters from Brazilian sign language (LIBRAS)	0.25 s samples	Calculation of absolute and average value	Support vector machines (SVM)	It is very difficult to perceive in real time fine finger gestures only with EMG data
[46,47]	Electromyogram data (EMG)	Real-time hand gesture recognition with Myo bracelet	200 sample windows	Signal rectification (abs) and application of a Butterworth low pass digital filter	Neighbor and dynamic temporal deformation algorithms closest to k	In [46] the model works better (86% accuracy) than the Myo system (83%) for five kinds of gestures. In [47] a muscle activity detector is included that speeds up processing time and improves the accuracy of recognition at 89.5%
[48,50,52]	Ultrasound	Ultrasound imaging based on control strategy for upper arm prosthetics	Ultrasound image sequences	Image processing algorithm. Calculation of cross-correlation coefficients between sequences. Root mean squares (RMS)	K-Nearest neighbour algorithm (k-NN) [48]. Not applied [50,52]	SMG signal has potential to be an alternative method for prosthetic control
[53]	Ultrasound	Gesture detection based on the propagation of low frequency transdermal ultrasound.	35, 40, 45, and 50 kHz of sinusoidal wave	Average amplitude, standard deviation, linear and log average of Fourier transform (FFT) points	Support Vector Machines (SVM)	Solid classification of tactile gestures made in different locations in the forearm
[54]	Ultrasound	Radial Muscle Activity Detection System with ultrasound transducers	1100 images per gesture	Rectification and application of the Hilbert transform	NN applying Combined cross-correlation (C-CC)	72% success in recognition of five flexion/extension gestures
[55]	Ultrasound	Comparison of the performance of different mounting positions of a portable ultrasound device	Splitting the video into push-ups or extensions from the neutral position	Farneback's algorithm	Optical flow. Multilayer perceptron regressor.	Flexion classification and extension of 10 discrete hand gestures with an accuracy greater than 98%
[57]	Ultrasound	A human gait classification method Based on radar Doppler spectrograms Radar Doppler Spectrograms	20 s samples	Short-Time Fourier Transform (STFT)	Support Vector Machines (SVM)	With seven directional filters, the classification performance is increased above 98%
[58]	Ultrasound	Ultrasonic sound waves to measure the distance to a subject	Not applied	Not applied	Not applied	Convert this information into lens rotation to obtain the correct focus
[59]	Ultrasound	Ultrasonic device to recognize simple hand gestures	Frames of 32 ms	Fourier transform (FFT) and discrete cosine transform (DCT)	Gaussian mixture model (GMM)	88.42% success in recognition of eight gestures
[60]	Ultrasound	Gesture-controlled sound synthesis with triangulation method using ultrasonic	Not applied	Low-pass filter	Not applied	Success in the detection and monitoring of 1 m ³ space movement

Table A2. Cont.

Ref.	Device Type	Focus	Segmentation	Feature Extraction	Classification	Results and Observations
[61]	Ultrasound	Gesture recognition based on the application of simple ultrasonic rangefinders in a mobile robot	Not applied	The speed and direction of the hand allow define a function of a speed ration	Not applied	Sensors readily available but the set of recognized gestures is limited
[62,65]	Ultrasound	Technique that takes advantage of the speaker and microphone already integrated in mobile device to detect gestures in the air	Not applied	Velocity, direction, proximity, size of target and time variation	Not applied	Robustness above 90% on different devices, users and environments [62]
[63,64]	Ultrasound	2D ultrasonic depth sensor that measures the range and direction of the targets in the air [63]. 3D applications [64]	Not applied	Signal-to-noise ratio (SNR) and incident angle	Not applied	Ultrasonic solutions are ideal for simple gesture recognition for smartphones and other devices with power limitation
[66]	Ultrasound	Ultrasonic hand gesture recognition for mobile devices	Not applied	Round trip time (RTD) Reflected signal strength (RSS)	Support vector machines (SVM)	96% success for seven types of gestures
[67]	Ultrasound	Micro hand gesture recognition system and methods using ultrasonic active sensing	256 FFT points	Range-Doppler pulsed radar signal processing method, shift and velocity	Convolutional neural network (CNN) and long short-term memory (LSTM)	96.32% of success for 11 hand gestures
[68]	Wi-Fi	Operation-oriented gesture recognition system in a range of low consumption computing devices through existing wireless signals	Not applied	Envelope detector to remove the carrier frequency and extract amplitude information	Dynamic time Warping (DTW)	Achieves the classification in 97% of success in a set of eight gestures
[70]	Wi-Fi	Gesture recognition extracting Doppler shifts from wireless signals	Combination of segments with positive and negative Doppler shifts	Computation of the frequency–time energy distribution of the narrow band signal	Coding of gestures. Doppler effects	94% success for nine gestures of a person in a room. With more people can go down to 60%
[71]	Wi-Fi	Capturing moving objects and their gestures behind a wall	Gesture decoding	Cumulative distribution functions (CDF) of Signal-to-noise ratio (SNR)	Not applied	87.5% success in detecting three gestures on concrete walls
[73]	Wi-Fi	Wi-Fi-based gesture recognition system identifying different signal change primitives	Detects changes in the raw Received signal strength indicator (RSSI)	Discrete wavelet transform (DWT)	Not applied	96% accuracy using three access point (AP)
[75]	Wi-Fi	Wi-Fi gesture recognition on existing devices	Received signal strength indicator (RSSI) and channel state information (CSI) streams	Low-pass filter Changes in received signal strength indicator (RSSI) and in channel state information (CSI)	Height and timing information to perform classification	91% accuracy while classifying four gestures across six participants

Table A2. Cont.

Ref.	Device Type	Focus	Segmentation	Feature Extraction	Classification	Results and Observations
[76]	Wi-Fi	Wi-Fi gesture recognition on existing devices	180 groups of channel state information (CSI) data from each packet	Mean value, standard deviation, median absolute deviation and maximum value of the anomaly patterns.	Support vector machines (SVM)	Average recognition accuracy of 92% in line of sight scenario and 88% without line of sight
[77]	Wi-Fi	Finger gesture recognition as continuous text input on standard Wi-Fi devices	Channel state information (CSI) stream	Discrete wavelet transformation (DWT)	Dynamic time Warping (DTW)	Average classification accuracy of up to 90.4% to recognize nine digits finger gestures of American sign language (ASL)
[87]	RFID	Device-free gesture recognition system based on phase information output by commercial off-the-shelf (COTS) RFID devices	Modified Varri Method	Savitzky–Golay filter	Dynamic time warping (DTW)	Achieves an average recognition accuracy of 96.5% and 92.8% in the scenario of identical positions and diverse positions respectively
[96]	Vision	Real-time classification of dance gestures from skeleton animation using Kinect	Video frames	Skeletal tracking algorithm (STA)	Cascaded Classifier	96.9% accuracy average for approximately four seconds skeletal motion recordings
[97]	Vision	A review of manual gesture recognition algorithms based on vision	Not applied	Not applied	Not applied	The methods using RGB and RGB-D cameras are reviewed with quantitative and qualitative algorithm comparisons
[98]	Vision	Robust gesture recognition with Kinect through a comparison between dynamic time warping (DTW) and hidden Markov model (HMM)	Video frames	Not applied	Dynamic time warping (DTW) and hidden Markov model (HMM)	Both dynamic time warping (DTW) and hidden Markov model (HMM) approaches give a very good classification rate of around 90%
[113]	Leap Motion	Segmentation and recognition of text written in 3D using Leap Motion Interface	Partial differentiation of the signal within a predefined window	Heuristics is applied to determine word boundaries	Hidden Markov model (HMM)	The proposed heuristic-based word segmentation algorithm works with an accuracy as high as 80.3% and an accuracy of 77.6% has been recorded by HMM-based words
[115]	Leap Motion	Arabic Sign Language Recognition using the Leap Motion Controller	Sample of 10 frames	Mean value	Nave Bayes classifier and multilayer perceptron neural networks (MLP)	98% classification accuracy with the Nave Bayes classifier and more than 99% using the MLP
[116–118]	Leap Motion	Recognition of dynamic hand gestures using Leap Motion	Data sample	Palm direction, palm normal, fingertips positions, and palm centre position. High-frequency noise filtering.	Support vector machines (SVM) and hidden Markov model (HMM) [116] Convolutional neural network (CNN) [117] Hidden conditional neural field (HCNF) [118]	With 80% of sample data, efficiency is achieved above 93.3% [116] Despite its good performance, it cannot provide online gesture recognition in real time [117] 95.0% recognition for the Handicraft-Gesture data set and 89.5% for the LeapMotion-Gesture3D [118]

Table A2. Cont.

Ref.	Device Type	Focus	Segmentation	Feature Extraction	Classification	Results and Observations
[120]	Leap Motion	Hand gesture recognition for post-stroke rehabilitation using leap motion	Data sample	17 features are extracted including Euclidean distances between the fingertips, pitch angle of palm, yaw angle of palm, roll angle of palm, and angles between fingers	Multi-class support vector machines (SVM). K-nearest neighbour (k-NN)-Neural Network	Seven gestures for the residential rehabilitation of post stroke patients are monitored with 97.71% accuracy
[122]	Leap Motion	Leap Motion Controller for authentication via hand geometry and gestures	Data sample	Equal error rate (EER), false acceptance (FAR) and false reject rate (FRR)	Waikato environment for knowledge analysis (WEKA)	IT shows that the Leap Motion can indeed be used successfully to both authenticate users at login as well as while performing continuous activities. Authentication accuracy reaches 98%
[110]	Leap Motion and Kinect	Hand gesture recognition with leap motion and Kinect devices	Data sample	Extract from the characteristics of the devices	Multi-class support vector machines (SVM)	81% accuracy with LPM, 65% accuracy with Kinect and when combined, increases to 91.3% for ten static gestures
[1]		Degree of impact of the gesture in a conversation	Not applied	Not applied	Not applied	55% impact on emotional conversations
[15–18,21–32]	Electromyogram data (EMG)	Techniques of EMG signal analysis	EMG decomposition using the lowest nonlinear mean square optimization (LMS) of higher order accumulators	Wavelet transform (WT). Autoregressive model (AR). Artificial intelligence Linear and stochastic models; Nonlinear modeling	Euclidean distance between the motor unit action potentials (MUAP) waveforms [21]. Motor unit number estimation (MUNE). Hardware models [26,29]	EMG signal analysis techniques for application in clinical diagnosis, biomedicine, research, hardware implementation and end user
[19]		Scientific basis of the body and its movements	Not applied	Not applied	Not applied	Global organization of system elements neuromuscular, neuroreceptors and instrumentation
[20]		Concepts of medical instrumentation	Not applied	Not applied	Not applied	A view of main medical sensors
[49,51]	Electromyogram data (EMG)	Relationships between the variables measured with sonography and EMG	Not applied	The relationship between each of the ultrasound and EMG parameters was described with nonlinear regression [49]. Root mean squares. (RMS) [51]	Not applied	Sonomyography has great potential for be an alternative method to assess muscle function
[56]	Wi-Fi	Identification of multiple human subjects outdoors with micro-Doppler signals	Not applied	Doppler filtering to eliminate clutter, torso extraction from the spectrogram, torso filtering to reduce noise, and peak period extraction using a Fourier transform.	Future work to compare classifiers.	Accuracy of more than 80% in the recognition of long-range subjects and frontal view. With angle variation it is significantly reduced below 40% in the worst case.

Table A2. Cont.

Ref.	Device Type	Focus	Segmentation	Feature Extraction	Classification	Results and Observations
[69,74]		System that tracks a user's 3D movement from the radio signals reflected in their body	Round trip distance spectrogram from each receive antenna	Kalman filter	Not applied	Its accuracy exceeds the current location of RF systems, which require the user to have a transceiver [69]. The second version [74] can localize up to five users simultaneously, even if they are perfectly static. Accuracy of 95% or higher
[72]	Electromyogram data (EMG)	Wi-Fi-based indoor positioning system that takes advantage of channel status information (CSI)	Not applied	Channel status information (CSI) processing	Not applied	Significantly improve location accuracy compared to the RSSI (Received Signal Strength Indicator) approach
[78–83]	RFID	RFID supply chain applications	Not applied	Not applied	Not applied	Management guidelines to proactively implement RFID applications
[86]	RFID	RFID-enabled fluid container without battery for recognize individual instances of liquid consumption	The time series is segmented as a single label reading	Mean, median, mode, standard deviation and range with respect to received signal strength indicator (RSSI) and Phase	Naïve Bayes (NB). Support vector machines (SVM). Random forest (RF). Linear conditional random fields (LCRF).	87% of success to recognize episodes of alcohol consumption in ten volunteers
[88]	Vision	Introductory techniques for 3D computer vision	Not applied	Not applied	Not applied	Set of computational techniques for 3D images
[89]	Vision	3D model generation by structured light	Not applied	Not applied	Not applied	Creation of a 3D scanner model based on structured light
[90]	Vision	High quality sampling in depth maps captured from a camera 3D-time of flight (ToF) coupled with a high-resolution RGB camera	Not applied	Not applied	Not applied	The new method surpasses the existing ones approaches to 3D-ToF upstream sampling
[91]		Model-based human posture estimation for gesture analysis in an opportunistic fusion smart camera network	Video frames	Expectation maximization (EM) algorithm	Not applied	Human posture estimation is described incorporating the concept of an opportunistic fusion framework
[92–95]		2D Human Features Model [94,95] 3D model-based tracking and recognition of human movement from real images [92,93]	Video frames	3_DOF model [92] Shape descriptors [93] Star skeleton. Skin color processing	Dynamic time Warping (DTW) [92]. Support vector machines (SVM) [93,94]. Neural Classifiers [95].	The experiment manages to obtain a result of 3D pose-recovery and movement classification [92,93]. Results with 2D models [94,95]
[105,108]	Leap Motion	An evaluation of the performance of the Leap Motion Controller [105] The analysis was performed in accordance with ISO 9283 [108]	Not applied	Not applied	Not applied	In the static scenario, it was shown that the standard deviation was less than 0.5 mm at all times. However, there is a significant drop in the accuracy of samples taken more than 250 mm above the controller [105]. With an industrial robot with a reference pen, an accuracy of 0.2 mm is achieved [108].

Table A2. Cont.

Ref.	Device Type	Focus	Segmentation	Feature Extraction	Classification	Results and Observations
[107]	Vision	Stereoscopic vision techniques to determine the three-dimensional structure of the scene	Not applied	Not applied	Not applied	Study on the effectiveness of a series of stereoscopic correspondence methods
[109]	Leap Motion	Validation of the Leap Motion Controller using markers motion capture technology	Not applied	Not applied	Not applied	The LMC is unable to provide data that is clinically significant for wrist flexion/extension, and perhaps wrist deviation
[111]	Vision	Potential of multimodal and multiuser interaction with virtual holography	Not applied	Not applied	Not applied	Multi-user interactions are described, which support local and distributed computers, using a variety of screens
[112]		Operating Virtual Panels with Hand Gestures in Immersive VR Games	Not applied	Not applied	Not applied	Gesture-based interaction for virtual reality games using the Unity game engine, the Leap Motion sensor, smartphone and VR headset
[114]	Leap Motion	Exploration of the feasibility of adapting the LMC, developed for video games, to the neurorehabilitation of the elderly with subacute stroke	Data sample	Not applied	Not applied	The rehabilitation with CML was performed with a high level of active participation, without adverse effects, and contributed to increase the recovery of manual skills
[119]	Leap Motion	Robotic arm manipulation using the Leap Motion Controller	Data sample	Mapping Algorithm	Not applied	Robotic arm manipulation scheme is proposed to allow the incorporation of robotic systems in the home environment
[121]		The system that remotely controls different appliances in your home or office through natural user interfaces HCI	Not applied	Not applied	Not applied	A low-cost and real-time environment is built capable of help people with disabilities
[123]	Leap Motion	Implementations of the Leap Motion device in sound synthesis and interactive live performance	Not applied	Not applied	Not applied	It seeks to empower disabled patients with musical expression using motion tracking technology and portable sensors
[124–127]	Leap Motion and VR glasses	VR glasses and Leap Motion trends in education, robotic, etc.	Not applied	Not applied	Not applied	Technology continues to update and promote innovation. VR and sensors such as Leap Motion are intended to have a great impact on it

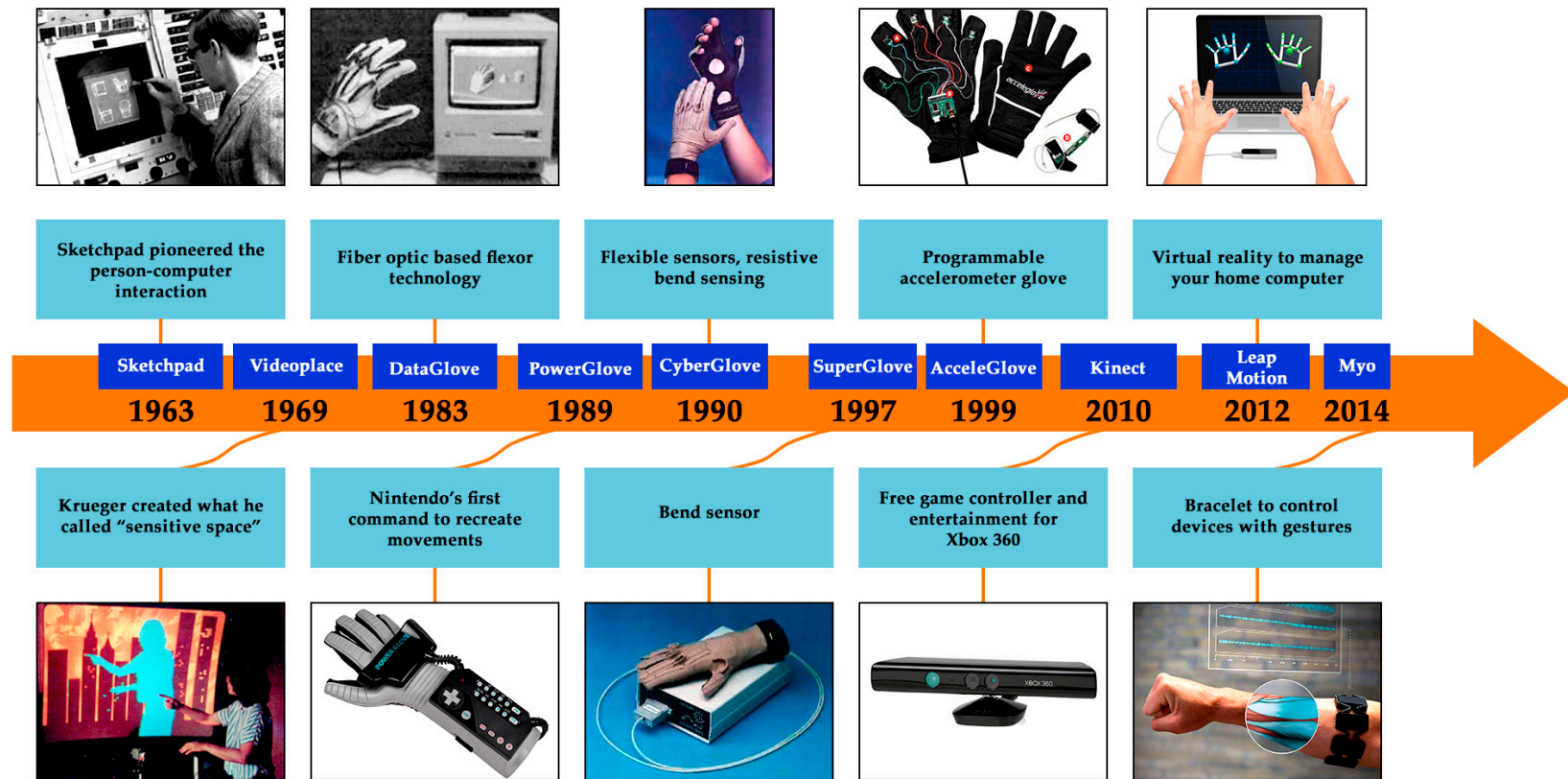


Figure A1. Timeline evolution of gesture recognition devices.

References

1. Mehrabian, A. *Nonverbal Communication*; Routledge: London, UK, 1972.
2. Sutherland, E.I. *Sketchpad: A Man-Machine Graphical Communication System*; University of Cambridge Computer Laboratory: Cambridge, UK, 2003.
3. Krueger, W.M.; Gionfriddo, T.; Hinrichsen, K. Videoplace an artificial reality. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, San Francisco, CA, USA, 10–13 April 1985; pp. 35–40. [[CrossRef](#)]
4. Kaushik, M.; Jain, R. Gesture Based Interaction NUI: An Overview. *Int. J. Eng. Trends Technol.* **2014**, *9*, 633–636. [[CrossRef](#)]
5. Myers, A.B. A brief history of human-computer interaction technology. *Interactions* **1998**, *5*, 44–54. [[CrossRef](#)]
6. Sharma, P.K.; Sharma, S.R. Evolution of Hand Gesture Recognition: A Review. *Int. J. Eng. Comput. Sci.* **2015**, *4*, 4.
7. Premaratne, P.; Nguyen, Q.; Premaratne, M. Human Computer Interaction Using Hand Gestures. In *Advanced Intelligent Computing Theories and Applications*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 381–386. [[CrossRef](#)]
8. Sturman, J.D.; Zeltzer, D. A survey of glove-based input. *IEEE Comput. Graph. Appl.* **1994**, *14*, 30–39. [[CrossRef](#)]
9. Zimmerman, G.T. Optical flex sensor. U.S. Patent 4542291A, 17 September 1985.
10. La Viola, J.J., Jr. *A Survey of Hand Posture and Gesture Recognition Techniques and Technology*; Department of Computer Science, Brown University: Providence, RI, USA, 1999.
11. CyberGlove III, CyberGlove Systems LLC. Available online: <http://www.cyberglovesystems.com/cyberglove-iii> (accessed on 17 January 2020).
12. Hernandez-Rebollar, J.L.; Kyriakopoulos, N.; Lindeman, R.W. The AcceleGlove: A whole-hand input device for virtual reality. In Proceedings of the ACM SIGGRAPH 2002 Conference Abstracts and Applications, San Antonio, TX, USA, 21–26 July 2002; p. 259. [[CrossRef](#)]
13. Santos, L.; González, J.L.; Turiel, J.P.; Fraile, J.C.; de la Fuente, E. Guante de datos sensorizado para uso en cirugía laparoscópica asistida por la mano (HALS). Presented at the Jornadas de Automática, Bilbao, Spain, 2–4 September 2015.
14. Moreno, R.; Cely, C.; Pinzón Arenas, J.O. Diseño de un guante háptico de control para manipulador robótico teleoperado. *Ingenium* **2013**, *7*, 19. [[CrossRef](#)]
15. Reaz, M.B.I.; Hussain, M.S.; Mohd-Yasin, F. Techniques of EMG signal analysis: Detection, processing, classification and applications. *Biol. Proced. Online* **2006**, *8*, 11–35. [[CrossRef](#)] [[PubMed](#)]
16. Valentinuzzi, M.E. Bioelectrical signal processing in cardiac and neurological applications and electromyography: Physiology, engineering, and noninvasive applications. *Biomed. Engl. OnLine* **2007**, *6*, 27. [[CrossRef](#)]
17. Farina, D.; Stegeman, D.F.; Merletti, R. Biophysics of the generation of EMG signals. In *Surface Electromyography: Physiology, Engineering and Applications*; Wiley Online Library: Hoboken, NJ, USA, 2016; pp. 1–24. [[CrossRef](#)]
18. Merletti, R.; Parker, P. *Electromyography: Physiology, Engineering, and Noninvasive Applications*; Wiley-IEEE Press: New York, NY, USA, 2004.
19. Gowitzke, B.A.; Milner, M. *El Cuerpo y sus Movimientos. Bases Científicas*; Paidotribo: Buenos Aires, Argentina, 1999.
20. Webster, J.G. *Medical Instrumentation: Application and Design*; John Wiley & Sons: New York, NY, USA, 2009.
21. Luca, C.J.D. The Use of Surface Electromyography in Biomechanics. *J. Appl. Biomech.* **1997**, *13*, 135–163. [[CrossRef](#)]
22. Heftner, G.; Zucchini, W.; Jaros, G.G. The electromyogram (EMG) as a control signal for functional neuromuscular stimulation. I. Autoregressive modeling as a means of EMG signature discrimination. *IEEE Trans. Biomed. Eng.* **1988**, *35*, 230–237. [[CrossRef](#)]
23. Merletti, R.; Farina, D. *Surface Electromyography: Physiology, Engineering, and Applications*; John Wiley & Sons: New York, NY, USA, 2016.

24. Barreto, A.; Scargle, S.; Adjouadi, M. Hands-off human-computer interfaces for individuals with severe motor disabilities. In Proceedings of the HCI International '99 (the 8th International Conference on Human-Computer Interaction), Munich, Germany, 22–26 August 1999; Volume 2, pp. 970–974.
25. Coleman, K. *Electromyography Based Human-Computer-Interface to Induce Movement in Elderly Persons with Movement Impairments*; ACM Press: New York, NY, USA, 2001; p. 4.
26. Guerreiro, T.; Jorge, J. EMG as a daily wearable interface. Presented at the International Conference on Computer Graphics Theory and Applications, Lisbon, Portugal, 25–28 February 2006; pp. 216–223.
27. Ahsan, R.; Ibrahimy, M.; Khalifa, O. EMG Signal Classification for Human Computer Interaction A Review. *Eur. J. Sci. Res.* **2009**, *33*, 480–501.
28. Crawford, B.; Miller, K.; Shenoy, P.; Rao, R. Real-Time Classification of Electromyographic Signals for Robotic Control. Presented at the AAAI-Association for the Advancement of Artificial Intelligence, Pittsburgh, PA, USA, 9–13 July 2005; p. 6.
29. Kiguchi, K.; Hayashi, Y. An EMG-Based Control for an Upper-Limb Power-Assist Exoskeleton Robot. *IEEE Trans. Syst. Man Cybern. Part. B Cybern.* **2012**, *42*, 1064–1071. [[CrossRef](#)] [[PubMed](#)]
30. Joshi, C.D.; Lahiri, U.; Thakor, N.V. Classification of Gait Phases from Lower Limb EMG: Application to Exoskeleton Orthosis. In Proceedings of the 2013 IEEE Point-of-Care Healthcare Technologies (PHT), Bangalore, India, 16–18 January 2013; pp. 228–231. [[CrossRef](#)]
31. Manabe, H.; Hiraiwa, A.; Sugimura, T. Unvoiced speech recognition using EMG–mime speech recognition. In Proceedings of the CHI '03 Extended Abstracts on Human Factors in Computing Systems, Ft. Lauderdale, FL, USA, 5–10 April 2003; pp. 794–795. [[CrossRef](#)]
32. Benedek, J.; Hazlett, R.L. Incorporating Facial EMG Emotion Measures as Feedback in the Software Design Process. Presented at the HCIC Consortium, Granby, CO, USA, 2–6 February 2005; pp. 1–9.
33. Healey, J.; Picard, R.W. StartleCam: A cybernetic wearable camera. In Proceedings of the Digest of Papers. Second International Symposium on Wearable Computers (Cat. No.98EX215), Pittsburgh, PA, USA, 19–20 October 1998; pp. 42–49. [[CrossRef](#)]
34. Dubost, G.; Tanaka, A. A Wireless, Network-Based Biosensor Interface for Music. In Proceedings of the 2002 International Computer Music Conference (ICMC), Gothenburg, Sweden, 16–21 September 2002.
35. Knapp, R.B.; Lusted, H.S. A Bioelectric Controller for Computer Music Applications. *Comput. Music J.* **1990**, *14*, 42–47. [[CrossRef](#)]
36. Tanaka, A.; Knapp, R.B. Multimodal Interaction in Music Using the Electromyogram and Relative Position Sensing. In *A NIME Reader*; Jensenius, A.R., Lyons, M.J., Eds.; Springer International Publishing: Cham, Vietnam, 2002; Volume 3, pp. 45–58.
37. Fistre, J.; Tanaka, A. Real time EMG gesture recognition for consumer electronics device control. Presented at the Sony CSL Paris Open House, Paris, France, October 2002.
38. Rosenberg, R. The biofeedback pointer: EMG control of a two dimensional pointer. In Proceedings of the Digest of Papers Second International Symposium on Wearable Computers (Cat. No.98EX215), Pittsburgh, PA, USA, 19–20 October 1998; pp. 162–163. [[CrossRef](#)]
39. Wheeler, K.R.; Jorgensen, C.C. Gestures as input: Neuroelectric joysticks and keyboards. *IEEE Pervasive Comput.* **2003**, *2*, 56–61. [[CrossRef](#)]
40. Dhillon, H.S.; Singla, R.; Rekhi, N.S.; Jha, R. EOG and EMG based virtual keyboard: A brain-computer interface. In Proceedings of the 2009 2nd IEEE International Conference on Computer Science and Information Technology, Beijing, China, 8–11 August 2009; pp. 259–262. [[CrossRef](#)]
41. Saponas, T.S.; Tan, D.S.; Morris, D.; Balakrishnan, R.; Turner, J.; Landay, J.A. Enabling always-available input with muscle-computer interfaces. In Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology, Victoria, BC, Canada, 4–7 October 2009; pp. 167–176. [[CrossRef](#)]
42. Tabor, A.; Bateman, S.; Scheme, E. Game-Based Myoelectric Training. In Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts, Austin, TX, USA, 16–19 October 2016; pp. 299–306. [[CrossRef](#)]
43. Boyali, A.; Hashimoto, N.; Matsumoto, O. Hand posture and gesture recognition using MYO armband and spectral collaborative representation based classification. In Proceedings of the 2015 IEEE 4th Global Conference on Consumer Electronics (GCCE), Osaka, Japan, 27–30 October 2015; pp. 200–201. [[CrossRef](#)]

44. Sathiyarayanan, M.; Rajan, S. MYO Armband for physiotherapy healthcare: A case study using gesture recognition application. In Proceedings of the 2016 8th International Conference on Communication Systems and Networks (COMSNETS), Bangalore, India, 5–10 January 2016; pp. 1–6. [\[CrossRef\]](#)
45. Abreu, J.G.; Teixeira, J.M.; Figueiredo, L.S.; Teichrieb, V. Evaluating Sign Language Recognition Using the Myo Armband. In Proceedings of the 2016 XVIII Symposium on Virtual and Augmented Reality (SVR), Gramado, Brazil, 21–24 June 2016; pp. 64–70. [\[CrossRef\]](#)
46. Benalcázar, M.E.; Jaramillo, A.G.; Jonathan, A.; Zea Páez, A.; Andaluz, V.H. Hand gesture recognition using machine learning and the Myo armband. In Proceedings of the 2017 25th European Signal Processing Conference (EUSIPCO), Kos, Greece, 28 August–2 September 2017; pp. 1040–1044. [\[CrossRef\]](#)
47. Benalcázar, M.E.; Motoche, C.; Zea, J.A.; Jaramillo, A.G.; Anchundia, C.E.; Zambrano, P.; Pérez, M. Real-time hand gesture recognition using the Myo armband and muscle activity detection. In Proceedings of the 2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM), Salinas, Ecuador, 16–20 October 2017; pp. 1–6. [\[CrossRef\]](#)
48. Sikdar, S.; Rangwala, H.; Eastlake, E.B.; Hunt, I.A.; Nelson, A.J.; Devanathan, J.; Pancrazio, J.J. Novel Method for Predicting Dexterous Individual Finger Movements by Imaging Muscle Activity Using a Wearable Ultrasonic System. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2014**, *22*, 69–76. [\[CrossRef\]](#)
49. Hodges, P.W.; Pengel, L.H.M.; Herbert, R.D.; Gandevia, S.C. Measurement of muscle contraction with ultrasound imaging. *Muscle Nerve* **2003**, *27*, 682–692. [\[CrossRef\]](#)
50. Zheng, Y.P.; Chan, M.M.F.; Shi, J.; Chen, X.; Huang, Q.H. Sonomyography: Monitoring morphological changes of forearm muscles in actions with the feasibility for the control of powered prosthesis. *Med. Eng. Phys.* **2006**, *28*, 405–415. [\[CrossRef\]](#)
51. Guo, J.-Y.; Zheng, Y.-P.; Huang, Q.-H.; Chen, X.; He, J.-F. Comparison of sonomyography and electromyography of forearm muscles in the guided wrist extension. In Proceedings of the 2008 5th International Summer School and Symposium on Medical Devices and Biosensors, Hong Kong, China, 1–3 June 2008; pp. 235–238. [\[CrossRef\]](#)
52. Chen, X.; Zheng, Y.P.; Guo, J.Y.; Shi, J. Sonomyography (SMG) Control for Powered Prosthetic Hand: A Study with Normal Subjects. *Ultrasound Med. Biol.* **2010**, *36*, 1076–1088. [\[CrossRef\]](#)
53. Mujibiyah, A.; Cao, X.; Tan, D.S.; Morris, D.; Patel, S.N.; Rekimoto, J. The sound of touch: On-body touch and gesture sensing based on transdermal ultrasound propagation. In Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces, St. Andrews, UK, 6 October 2013; pp. 189–198. [\[CrossRef\]](#)
54. Hettiarachchi, N.; Ju, Z.; Liu, H. A New Wearable Ultrasound Muscle Activity Sensing System for Dexterous Prosthetic Control. In Proceedings of the 2015 IEEE International Conference on Systems, Man, and Cybernetics, Kowloon, China, 9–12 October 2015; pp. 1415–1420. [\[CrossRef\]](#)
55. McIntosh, J.; Marzo, A.; Fraser, M.; Phillips, C. EchoFlex: Hand Gesture Recognition using Ultrasound Imaging. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; pp. 1923–1934. [\[CrossRef\]](#)
56. Tahmouh, D.; Silvius, J. Radar micro-doppler for long range front-view gait recognition. In Proceedings of the 2009 IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems, Washington, DC, USA, 28–30 September 2009; pp. 1–6. [\[CrossRef\]](#)
57. Tivive, F.H.C.; Bouzerdoum, A.; Amin, M.G. A human gait classification method based on radar Doppler spectrograms. *EURASIP J. Adv. Signal. Process.* **2010**, *2010*, 1–12. [\[CrossRef\]](#)
58. Biber, C.; Ellin, S.; Shenk, E.; Stempeck, J. The Polaroid Ultrasonic Ranging System. Presented at the Audio Engineering Society Convention 67, New York, NY, USA, 31 October–3 November 1980.
59. Kalgaonkar, K.; Raj, B. One-handed gesture recognition using ultrasonic Doppler sonar. In Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, Taiwan, 19–24 April 2009; pp. 1889–1892. [\[CrossRef\]](#)
60. Fléty, E. 3D Gesture Acquisition Using Ultrasonic Sensors. In *Trends in Gestural Control of Music*; Wanderley, M., Battier, M., Eds.; Institut de Recherche et Coordination Acoustique Musique -Centre Pompidou: Paris, France, 2000; pp. 193–207.
61. Kreczmer, B. Gestures recognition by using ultrasonic range-finders. In Proceedings of the 2011 16th International Conference on Methods Models in Automation Robotics, Miedzyzdroje, Poland, 22–25 August 2011; pp. 363–368. [\[CrossRef\]](#)

62. Gupta, S.; Morris, D.; Patel, S.; Tan, D. SoundWave: Using the doppler effect to sense gestures. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, TX, USA, 5–10 May 2012; pp. 1911–1914. [[CrossRef](#)]
63. Przybyla, R.J.; Shelton, S.E.; Guedes, A.; Krigel, R.; Horsley, D.A.; Boser, B.E. In-air ultrasonic rangefinding and angle estimation using an array of ALN micromachined transducers. In Proceedings of the 2012 Solid-State, Actuators, and Microsystems Workshop Technical Digest, Hilton Head, SC, USA, 3–7 June 2012; pp. 50–53. [[CrossRef](#)]
64. Przybyla, R.J.; Tang, H.-Y.; Shelton, S.E.; Horsley, D.A.; Boser, B.E. 3D ultrasonic gesture recognition. In Proceedings of the 2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC), San Francisco, CA, USA, 9–13 February 2014; pp. 210–211. [[CrossRef](#)]
65. Booij, W.E.; Welle, K.O. Ultrasound Detectors. U.S. Patent 8792305B2, 29 July 2014.
66. Saad, M.; Bleakley, C.J.; Nigram, V.; Kettle, P. Ultrasonic hand gesture recognition for mobile devices. *J. Multimodal User Interfaces* **2018**, *12*, 31–39. [[CrossRef](#)]
67. Sang, Y.; Shi, L.; Liu, Y. Micro Hand Gesture Recognition System Using Ultrasonic Active Sensing. *IEEE Access* **2018**, *6*, 49339–49347. [[CrossRef](#)]
68. Kellogg, B.; Talla, V.; Gollakota, S. Bringing Gesture Recognition to All Devices. Presented at the 11th (USENIX) Symposium on Networked Systems Design and Implementation (NSDI 14), Seattle, WA, USA, 2–4 April 2014; pp. 303–316.
69. Adib, F.; Kabelac, Z.; Katabi, D.; Miller, R.C. 3D Tracking via Body Radio Reflections. Presented at the 11th (USENIX) Symposium on Networked Systems Design and Implementation (NSDI 14), Seattle, WA, USA, 2–4 April 2014; pp. 317–329.
70. Pu, Q.; Gupta, S.; Gollakota, S.; Patel, S. Whole-home gesture recognition using wireless signals. In Proceedings of the 19th Annual International Conference on Mobile Computing & Networking, Miami, FL, USA, 30 September–4 October 2013; pp. 27–38. [[CrossRef](#)]
71. Adib, F.; Katabi, D. See through walls with Wi-Fi! In Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM, Hong Kong, China, 12–16 August 2013; pp. 75–86. [[CrossRef](#)]
72. Wu, K.; Xiao, J.; Yi, Y.; Gao, M.; Ni, L.M. FILA: Fine-grained indoor localization. In Proceedings of the 2012 Proceedings IEEE INFOCOM, Orlando, FL, USA, 25–30 March 2012; pp. 2210–2218. [[CrossRef](#)]
73. Abdelnasser, H.; Youssef, M.; Harras, K.A. WiGest: A ubiquitous Wi-Fi-based gesture recognition system. In Proceedings of the 2015 IEEE Conference on Computer Communications (INFOCOM), Hong Kong, China, 26 April–1 May 2015; pp. 1472–1480. [[CrossRef](#)]
74. Adib, F.; Kabelac, Z.; Katabi, D. Multi-Person Localization via RF Body Reflections. In Proceedings of the 12th USENIX, etworked Systems Design and Implementation, Oakland, CA, USA, 4–6 May 2015.
75. Nandakumar, R.; Kellogg, B.; Gollakota, S. Wi-Fi Gesture Recognition on Existing Devices. *arXiv* **2014**, arXiv:1411.5394. Available online: <https://arxiv.org/abs/1411.5394> (accessed on 23 June 2020).
76. He, W.; Wu, K.; Zou, Y.; Ming, Z. WiG: Wi-Fi-Based Gesture Recognition System. In Proceedings of the 2015 24th International Conference on Computer Communication and Networks (ICCCN), Kowloon, Hong Kong, 26 April–1 May 2015; pp. 1–7. [[CrossRef](#)]
77. Li, H.; Yang, W.; Wang, J.; Xu, Y.; Huang, L. WiFinger: Talk to your smart devices with finger-grained gesture. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Heidelberg, Germany, 12–16 September 2016; pp. 250–261. [[CrossRef](#)]
78. Angeles, R. RFID Technologies: Supply-Chain Applications and Implementation Issues. *Inf. Syst. Manag.* **2006**, *51*–65. [[CrossRef](#)]
79. Borriello, G.; Brunette, W.; Hall, M.; Hartung, C.; Tangney, C. Reminding About Tagged Objects Using Passive RFIDs. In *UbiComp 2004: Ubiquitous Computing*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 36–53. [[CrossRef](#)]
80. Want, R.; Fishkin, K.P.; Gujar, A.; Harrison, B.L. Bridging physical and virtual worlds with electronic tags. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Pittsburgh, PA, USA, 15–20 May 1999; pp. 370–377. [[CrossRef](#)]
81. Buettner, M.; Prasad, R.; Philipose, M.; Wetherall, D. Recognizing daily activities with RFID-based sensors. In Proceedings of the 11th International Conference on Ubiquitous Computing, Orlando, FL, USA, 30 September–3 October 2009; pp. 51–60. [[CrossRef](#)]

82. Fishkin, K.P.; Jiang, B.; Philipose, M.; Roy, S. I Sense a Disturbance in the Force: Unobtrusive Detection of Interactions with RFID-tagged Objects. In *UbiComp 2004: Ubiquitous Computing*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 268–282. [CrossRef]
83. Yang, L.; Chen, Y.; Li, X.-Y.; Xiao, C.; Li, M.; Liu, Y. Tagoram: Real-time tracking of mobile RFID tags to high precision using COTS devices. In Proceedings of the 20th Annual International Conference on Mobile Computing and Networking, Maui, HA, USA, 7–11 September 2014; pp. 237–248. [CrossRef]
84. Asadzadeh, P.; Kulik, L.; Tanin, E. Real-time Gesture Recognition Using RFID Technology. *J. Pers. Ubiquitous Comput.* **2012**, *16*, 225–234. [CrossRef]
85. Bouchard, K.; Bouzouane, A.; Bouchard, B. Gesture recognition in smart home using passive RFID technology. In Proceedings of the 7th International Conference on Pervasive Technologies Related to Assistive Environments, Rhodes, Greece, 27–30 May 2014; pp. 1–8. [CrossRef]
86. Jayatilaka, A.; Ranasinghe, D.C. Real-time fluid intake gesture recognition based on batteryless UHF RFID technology. *Pervasive Mob. Comput.* **2017**, *34*, 146–156. [CrossRef]
87. Zou, Y.; Xiao, J.; Han, J.; Wu, K.; Li, Y.; Ni, L.M. GRfid: A Device-Free RFID-Based Gesture Recognition System. *IEEE Trans. Mob. Comput.* **2017**, *16*, 381–393. [CrossRef]
88. Trucco, E.; Verri, A. *Introductory Techniques for 3-D Computer Vision*, 1st ed.; Prentice Hall: Upper Saddle River, NJ, USA, 1998; pp. 178–194.
89. Manteca, F. *Generación de Modelos 3D Mediante luz Estructurada*; Universidad de Cantabria: Cantabria, Spain, 2018.
90. Park, J.; Kim, H.; Tai, Y.W.; Brown, M.S.; Kweon, I. High quality depth map upsampling for 3D-TOF cameras. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 1623–1630. [CrossRef]
91. Wu, C.; Aghajan, H. Model-based human posture estimation for gesture analysis in an opportunistic fusion smart camera network. In Proceedings of the 2007 IEEE Conference on Advanced Video and Signal Based Surveillance, London, UK, 5–7 September 2007; pp. 453–458. [CrossRef]
92. Gavrilu, D.M.; Davis, L.S. Towards 3-d model-based tracking and recognition of human movement: A multi-view approach. In Proceedings of the International Workshop on Automatic Face-and Gesture-Recognition, Zurich, Switzerland, 26–28 June 1995; pp. 272–277.
93. Cohen, I.; Li, H. Inference of human postures by classification of 3D human body shape. In Proceedings of the 2003 IEEE International SOI Conference, Proceedings (Cat. No.03CH37443), Nice, France, 17 October 2003; pp. 74–81. [CrossRef]
94. Mo, H.-C.; Leou, J.-J.; Lin, C.-S. Human Behavior Analysis Using Multiple 2D Features and Multicategory Support Vector Machine in MVA. In Proceedings of the MVA2009 IAPR Conference on Machine Vision Applications, Yokohama, Japan, 20–22 May 2009; pp. 46–49.
95. Corradini, A.; Böhme, H.-J.; Gross, H.-M. Visual-based posture recognition using hybrid neural networks. Presented at the European Symposium on Artificial Neural Networks, Bruges, Belgium, 21–23 April 1999; pp. 81–86.
96. Raptis, M.; Kirovski, D.; Hoppe, H. Real-time classification of dance gestures from skeleton animation. In Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, Vancouver, BC, Canada, 5–7 August 2011; pp. 147–156. [CrossRef]
97. Pisharady, P.K.; Saerbeck, M. Recent methods and databases in vision-based hand gesture recognition: A review. *Comput. Vis. Image Underst.* **2015**, *141*, 52–165. [CrossRef]
98. Raheja, J.L.; Minhas, M.; Prashanth, D.; Shah, T.; Chaudhary, A. Robust gesture recognition using Kinect: A comparison between DTW and HMM. *Optik* **2015**, *126*, 1098–1104. [CrossRef]
99. Kinect: Desarrollo de Aplicaciones de Windows. Available online: <https://developer.microsoft.com/es-es/windows/kinect> (accessed on 18 January 2020).
100. Digital Worlds that Feel Human | Ultraleap. Available online: <https://www.ultraleap.com/> (accessed on 18 January 2020).
101. Leap Motion (I): Características técnicas. Available online: <https://www.showleap.com/2015/04/20/leap-motion-caracteristicas-tecnicas/> (accessed on 18 January 2020).
102. MX25L3206E DATASHEET. Available online: <https://docs.rs-online.com/5c85/0900766b814ac6f9.pdf> (accessed on 23 June 2020).

103. API Overview—Leap Motion JavaScript SDK v2.3 Documentation. Available online: https://developer-archive.leapmotion.com/documentation/v2/javascript/devguide/Leap_Overview.html (accessed on 18 January 2020).
104. API Overview—Leap Motion Java SDK v3.2 Beta Documentation. Available online: https://developer-archive.leapmotion.com/documentation/java/devguide/Leap_Overview.html?proglang=java (accessed on 27 January 2020).
105. Guna, J.; Jakus, G.; Pogačnik, M.; Tomažič, S.; Sodnik, J. An Analysis of the Precision and Reliability of the Leap Motion Sensor and Its Suitability for Static and Dynamic Tracking. *Sensors* **2014**, *14*, 702–3720. [[CrossRef](#)]
106. Leap Motion (II): Principio de funcionamiento. Available online: <https://www.showleap.com/2015/05/04/leap-motion-ii-principio-de-funcionamiento/> (accessed on 18 January 2020).
107. Montalvo Martínez, M. Técnicas de Visión Estereoscópica Para Determinar la Estructura Tridimensional de la Escena. 2010. Available online: <https://eprints.ucm.es/11350/> (accessed on 18 January 2020).
108. Weichert, F.; Bachmann, D.; Rudak, B.; Fisseler, D. Analysis of the Accuracy and Robustness of the Leap Motion Controller. *Sensors* **2013**, *13*, 6380–6393. [[CrossRef](#)]
109. Smeragliuolo, A.H.; Hill, N.J.; Disla, L.; Putrino, D. Validation of the Leap Motion Controller using marked motion capture technology. *J. Biomech.* **2016**, *49*, 1742–1750. [[CrossRef](#)]
110. Marin, G.; Dominio, F.; Zanuttigh, P. Hand gesture recognition with leap motion and kinect devices. In Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014; pp. 1565–1569. [[CrossRef](#)]
111. Noor, A.K.; Aras, R. Potential of multimodal and multiuser interaction with virtual holography. *Adv. Eng. Softw.* **2015**, *81*, 1–6. [[CrossRef](#)]
112. Zhang, Y.; Meruvia-Pastor, O. Operating Virtual Panels with Hand Gestures in Immersive VR Games. In *Augmented Reality, Virtual Reality, and Computer Graphics*; De Paolis, L., Bourdot, P., Mongelli, A., Eds.; Springer: Cham, Switzerland; pp. 299–308. [[CrossRef](#)]
113. Agarwal, C.; Dogra, D.P.; Saini, R.; Roy, P.P. Segmentation and recognition of text written in 3D using Leap motion interface. In Proceedings of the 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), Kuala Lumpur, Malaysia, 3–6 November 2015; pp. 539–543. [[CrossRef](#)]
114. Iosa, M.; Morone, G.; Fusco, A.; Castagnoli, M.; Fusco, F.R.; Pratesi, L.; Paolucci, S. Leap motion controlled videogame-based therapy for rehabilitation of elderly patients with subacute stroke: A feasibility pilot study. *Top. Stroke Rehabil.* **2015**, *22*, 306–316. [[CrossRef](#)] [[PubMed](#)]
115. Mohandes, M.; Aliyu, S.; Deriche, M. Arabic sign language recognition using the leap motion controller. In Proceedings of the 2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE), Istanbul, Turkey, 1–4 June 2014; pp. 960–965. [[CrossRef](#)]
116. Chen, Y.; Ding, Z.; Chen, Y.-L.; Wu, X. Rapid recognition of dynamic hand gestures using leap motion. In Proceedings of the 2015 IEEE International Conference on Information and Automation, Lijiang, China, 8–10 August 2015; pp. 1419–1424. [[CrossRef](#)]
117. McCartney, R.; Yuan, J.; Bischof, H.-P. Gesture Recognition with the Leap Motion Controller. In Proceedings of the 19th International Conference on Image Processing, Computer Vision, & Pattern Recognition: IPCV'15, Las Vegas, NV, USA, 27–30 July 2015.
118. Lu, W.; Tong, Z.; Chu, J. Dynamic Hand Gesture Recognition with Leap Motion Controller. *IEEE Signal. Process. Lett.* **2016**, *23*, 1188–1192. [[CrossRef](#)]
119. Bassily, D.; Georgoulas, C.; Guettler, J.; Linner, T.; Bock, T. Intuitive and Adaptive Robotic Arm Manipulation using the Leap Motion Controller. In Proceedings of the ISR/Robotik 2014, 41st International Symposium on Robotics, München, Germany, 2–3 June 2014; pp. 1–7.
120. Li, W.-J.; Hsieh, C.-Y.; Lin, L.-F.; Chu, W.-C. Hand gesture recognition for post-stroke rehabilitation using leap motion. In Proceedings of the 2017 International Conference on Applied System Innovation (ICASI), Sapporo, Japan, 13–17 May 2017; pp. 386–388. [[CrossRef](#)]
121. Gándara, C.V.; Bauza, C.G. IntelliHome: A framework for the development of ambient assisted living applications based in low-cost technology. In Proceedings of the Latin American Conference on Human Computer Interaction, Córdoba, Argentina, 18–21 November 2015; pp. 1–4. [[CrossRef](#)]

122. Chan, A.; Halevi, T.; Memon, N. Leap Motion Controller for Authentication via Hand Geometry and Gestures. Presented at the International Conference on Human Aspects of Information Security, Privacy, and Trust, Los Angeles, CA, USA, 2–7 August 2015; pp. 13–22. [[CrossRef](#)]
123. Hantrakul, L.; Kaczmarek, K. Implementations of the Leap Motion device in sound synthesis and interactive live performance. In Proceedings of the 2014 International Workshop on Movement and Computing, Paris, France, 16–17 June 2014; pp. 142–145. [[CrossRef](#)]
124. Ling, H.; Rui, L. VR glasses and leap motion trends in education. In Proceedings of the 2016 11th International Conference on Computer Science Education (ICCSE), Nagoya, Japan, 23–25 August 2016; pp. 917–920. [[CrossRef](#)]
125. Cichon, T.; Rossmann, J. *Simulation-Based User Interfaces for Digital Twins: Pre-, in-, or Post-Operational Analysis and Exploration of Virtual Testbeds*; Institute for Man-Machine-Interaction RWTH Aachen University Ahornstrasse: Aachen, Germany, 2017; p. 6.
126. Vosinakis, S.; Koutsabasis, P. Evaluation of visual feedback techniques for virtual grasping with bare hands using Leap Motion and Oculus Rift. *Virtual Real.* **2018**, *22*, 47–62. [[CrossRef](#)]
127. Jetter, H.-C.; Rädle, R.; Feuchtner, T.; Anthes, C.; Friedl, J.; Klokmose, C. In VR, everything is possible!: Sketching and Simulating Spatially-Aware Interactive Spaces in Virtual Reality. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020. [[CrossRef](#)]
128. HP Reverb G2 VR Headset. Available online: <https://www.hp.com/us/en/vr/reverb-g2-vr-headset.html> (accessed on 13 June 2020).
129. Oculus Rift | Oculus. Available online: https://www.oculus.com/rift/?locale=es_LA#oui-csl-rift-games=images-tale (accessed on 13 June 2020).
130. Virtual Reality | HTC España. Available online: <https://www.htc.com/es/virtual-reality/> (accessed on 13 June 2020).
131. Mejora tu Experiencia-Valve Corporation. Available online: <https://www.valvesoftware.com/es/index> (accessed on 13 June 2020).



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).