applied sciences

MDPI

Review

Multimodal Natural Human–Computer Interfaces for Computer-Aided Design: A Review Paper

Hongwei Niu [1,2,3], Cees Van Leeuwen [2,3], Jia Hao [1,*], Guoxin Wang [1] and Thomas Lachmann [2,3,4]

[1] Institute of Industrial and Intelligent Systems Engineering, Beijing Institute of Technology, Beijing 100081, China; nhwfend@bit.edu.cn (H.N.); wangguoxin@bit.edu.cn (G.W.)
[2] Brain and Cognition Research Unit, KU Leuven, 3000 Leuven, Belgium; cees.vanleeuwen@kuleuven.be (C.V.L.); lachmann@rhrk.uni-kl.de (T.L.)
[3] Center for Cognitive Science, University of Kaiserslautern, 67663 Kaiserslautern, Germany
[4] Centro de Investigación Nebrija en Cognición, Universidad Nebrija, 28015 Madrid, Spain
* Correspondence: haojia632@bit.edu.cn

Abstract: Computer-aided design (CAD) systems have advanced to become a critical tool in product design. Nevertheless, they still primarily rely on the traditional mouse and keyboard interface. This limits the naturalness and intuitiveness of the 3D modeling process. Recently, a multimodal human–computer interface (HCI) has been proposed as the next-generation interaction paradigm. Widening the use of a multimodal HCI provides new opportunities for realizing natural interactions in 3D modeling. In this study, we conducted a literature review of a multimodal HCI for CAD to summarize the state-of-the-art research and establish a solid foundation for future research. We explore and categorize the requirements for natural HCIs and discuss paradigms for their implementation in CAD. Following this, factors to evaluate the system performance and user experience of a natural HCI are summarized and analyzed. We conclude by discussing challenges and key research directions for a natural HCI in product design to inspire future studies.

Keywords: human–computer interface; multimodal interface; computer-aided design; natural interaction

check for updates

Citation: Niu, H.; Van Leeuwen, C.; Hao, J.; Wang, G.; Lachmann, T. Multimodal Natural Human–Computer Interfaces for Computer-Aided Design: A Review Paper. *Appl. Sci.* **2022**, *12*, 6510. https://doi.org/10.3390/app12136510

Academic Editor: Miguel Ángel Redondo Duque

Received: 18 May 2022
Accepted: 22 June 2022
Published: 27 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 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 (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Computer-aided design (CAD) systems have traditionally been used to aid the processes of product design and manufacturing [1]. Recent advances in CAD technology have led to a greater efficiency of 3D modeling in multiple industrial domains, such as mechanical and aerospace engineering [2,3]. Yet, there are some significant obstacles preventing further advancements in the research on human–computer interfaces (HCIs) for CAD, slowing the same pace of development. The current CAD systems still heavily rely on the conventional WIMP (Windows, Icons, Menus, Pointer) interface, using keystrokes and pointer movements as input. The behavioral patterns providing the system with input in this way show a mismatch in mapping between the user's mental model and the computer interface. This mismatch forces the user to perform complex mental transformations between the 3D modeling space and 2D input devices, resulting in a high lead-time and hindered design creativity [3,4]. Moreover, novice users need to master a dense set of toolbars and menus, which takes considerable time to learn [5]. It is therefore necessary to move to user-centered natural HCI technology that can make CAD modeling more natural for users to operate and more intuitive to learn.

Efforts to mitigate this need have led to the development of multimodal interfaces that use eye tracking [6], gesture recognition [7], speech recognition [8], and brain signal recording techniques such as electroencephalograms (EEG) [9]. Standing at the crossroads of several research areas including psychology, computer vision, signal analysis, artificial intelligence, and cognitive science, the field of multimodal HCIs for CAD is in need of a solid theoretical foundation [10]. With the assistance of neighboring disciplines, researchers

have developed a variety of novel interfaces for CAD [8,11,12]. Unfortunately, new input modalities, such as gesture, gaze, and EEG signals, are currently processed separately and integrated only at the application stage. Not only does this ignore the many issues regarding the application of each separate modality but also, more importantly, the roles and functions of multiple modalities for CAD and their interplay remain to be quantified and scientifically understood. In addition, there is a one-size-fits all attitude with respect to these modalities, in that an application fails to take individual differences into account. To elaborate on these issues, we review the state-of-the-art interaction patterns in CAD systems and discuss likely future research trends.

The remainder of the paper is organized as follows: Section 2 delineates the review methodology, including the methods adopted for the literature search and literature scope. In Section 3, the requirements and effects of a natural HCI for traditional CAD are presented. Based on this, Section 4 illustrates the state-of-the-art interaction techniques to implement a natural HCI, including eye tracking, gesture recognition, speech recognition, and EEG recognition. Sections 5 and 6 summarize the measurements and devices of natural HCIs for CAD, respectively. A discussion of future research directions is provided in Section 7, and Section 8 summarizes this review and highlights its contributions.

2. Methodology

We performed a two-step literature search in this study. Firstly, the major engineering databases, including Web of Science, Engineering Village, and IEEE Xplore Digital Library, were retrieved by using specific keywords, such as “Computer aided design”, “Conceptual design”, “Human computer interface”, “Multimodal”, and “Natural interaction”, to obtain a list of publications in English dating from 1980 to 2022. To this, results from a Google Scholar search with the same keywords were added. Secondly, a manual screening was conducted by reading the titles and abstracts of these papers, so as to retain those related to the field of computer-aided design in engineering while excluding those focused on other areas such as architecture or medicine.

Ninety-five papers were identified through this search. Other papers cited in our study deal with general concepts and devices, for example, eye tracking, gesture recognition and EEG acquisitions.

It should be emphasized that this review does not claim to comprehensively and exhaustively cover all aspects of CAD systems for product design in engineering. We limit ourselves to providing an overview of the state-of-the-art studies related to multimodal interfaces in the application of CAD and to discussing future directions for the implementation of natural HCIs for CAD.

3. Natural HCI for CAD

The concept of natural HCIs was first introduced by Steve Mann [13,14], in an effort to bridge the gap between human and computer systems. As a continuously expanding field, natural HCIs are difficult to describe exactly [15]. When we think of user interfaces that are natural and easy to use, we most often think of one where the interaction is direct and consistent with our natural behaviors [4]. As such, these behaviors can be considered as media for everyday social communication through voice, gestures, and emotional expressions. Exploring the world through locomotion and the use of our senses, especially vision, is also prominent in natural behavior. Whenever we visually explore the world, we move our eyes to keep the information of interest in focus. Hence, natural HCIs visually explore virtual environments in manners analogous to how we visually explore the environment through eye movement and the manipulate virtual models in ways analogous to how we manipulate physical objects. The key focus of natural HCIs is to allow humans to interact with computer systems as much as possible in the same way as human–human communication in everyday life [14,16].

The requirements and effects of natural HCIs for CAD, as identified in this section, cover product design activities and aim to emphasize the importance of natural HCI technology for CAD development.

3.1. The Reasons Why Natural HCI Is Required for CAD

In the early 1960s, product design processes evolved from manual to computer-aided drawing, in parallel with the emergence of computer graphics display technology [17]. The initial technology based on 2D computer graphics evolved into 3D CAD systems around the 1970s, and the subsequent evolution of computer software and hardware allowed a rapid development of CAD technology [18–20]. CAD functions have incrementally been standardized and integrated, and artificial intelligence has been incorporated into their design [20,21]. However, a key component of CAD systems has lagged in development: interactive interface technology still reflects the earlier stages of CAD technology.

In most CAD applications, a text-based Command Line Interface (CLI) and WIMP-based Graphical User Interface (GUI) are still the mainly used components to date [20,22,23]. These traditional interaction modes have some obvious limitations for CAD applications as users have to create, modify, and manipulate 3D models with 2D devices (mouse and monitor), which make it difficult to meet users' needs for natural interactions.

On the one hand, users are forced to frequently toggle certain on/off modules, and traverse across hundreds of menu items and toolbar buttons in the current interface, just to execute the desired CAD command [24]. Moreover, users need to memorize complicated hotkey combinations to control specific software products. Different CAD software programs usually have a variety of ways to represent the same command functions. This becomes even more challenging when designing across different 3D modeling applications, such as FreeCAD, AutoCAD, and SolidWorks [11]. As a result, current 3D CAD systems require a "very steep learning curve" for novice users to master the complex task of using a dense set of tool bars and menus [25].

On the other hand, the visuomotor requirements of using a mouse and keyboard may interfere with the visuospatial resources of the design process. The user needs to switch between the 3D modeling environment to the 2D input and output space [3]. In this process, the cognitive load is further enhanced by planning ahead for the manipulations needed to build the model [26,27]. Thus, the design intentions of users are frequently interrupted since considerable attention must be paid to finding the right commands via different button clicks [24].

Therefore, to ease learning and increase modeling efficiency, there is a strong demand for natural HCIs that bypasses these difficulties and enables users to interact with the system in a natural manner.

Natural HCIs for CAD maximally encapsulate the characteristics of human–human interactions (HHIs) in the modeling process [14,16,28]. As the next generation of natural HCIs, capturing original signals, such as human EEG, eye movements, gestures, and voices, is the premise and foundation for understanding users' model operation intentions accurately [29,30]. Note that human–human communication and exploration behavior involve visual, auditory, and tactile processes, and thus are intrinsically multimodal. Multimodal interfaces therefore become an effective approach for realizing natural HCIs for CAD.

3.2. The Effects of Using Natural HCI in CAD

Natural HCIs will drive a gradual shift in the product design process from a traditional system-centered design to user-centered design, and provide a more intuitive model operation experience for users [31]. Thus, research on natural HCIs will become a main development trend of CAD interaction technology. While natural HCI technology is not yet commonly used in current CAD systems in industry, it has been explored extensively in academic research and areas where natural HCIs could bring significant positive effects have been identified [32,33]. The potential effects of natural HCIs in CAD are mainly reflected in three aspects described below.

1. They can shorten the learning cycle and reduce the difficulty of learning for novice users.

Since users traditionally have to spend a huge amount of energy on learning and memorizing dense software menus and complicated operation instructions in modeling, it is not easy for the novice user to learn and master a new CAD application. To overcome this shortcoming, a natural HCI provides users with a novel interactive pattern where users just need to model in the way that is most comfortable and natural for them. The users' intention of modeling will be automatically sensed and accurately recognized by the multimodal natural HCI without the user having to remember a lot of rules and instructions embedded in CAD systems. For instance, in order to manipulate a target model, users can select it just by directing their gaze at it and, at the same time, instruct the system to move the model by making a move gesture with their hands [34]. This interactive pattern is easy for novices to learn and could be quickly mastered regardless of experience level.

A natural HCI also adds cross-platform features to CAD systems that enable users to control 3D models in different software environments. As a result, users do not need to relearn a large number of new commands when faced with a new CAD application. This could significantly shorten the learning cycle and quickly turn users from novices to experts.

2. They enable a more intuitive and natural interaction process.

Traditional WIMP interfaces use 2D subspace inputs such as mouse and keyboard to create and operate 3D models, which disconnect the input space from the modeling space and make the interaction process tedious and non-intuitive [3]. A natural HCI circumvents the limitations of traditional interfaces in terms of interaction space and dimensions, and provides a new end-to-end channel for direct interaction between users and 3D models. By capturing and recognizing the users' EEG, gestures, speech, and other original signals in 3D modeling, natural HCIs can provide a more familiar and direct interaction for users, meaning that the interaction between the user and model can be performed more intuitively and naturally [35].

3. They can enhance the innovative design ability of designers.

Current CAD systems have difficulty in supporting innovative design as they have been developed with too much emphasis on graphical representation techniques, and not enough on exploring how to support the processes that are essential for creative design [36,37]. Central to innovative design are ambiguity, uncertainty, and parallel lines of thought. All of this is limited by the circumscribed thinking and premature fixation from current CAD systems [38], which are readily supported by natural HCIs.

For natural HCIs, the input and output devices that separate the user and the system should be as unobtrusive as possible or hidden, so that users do not notice them and instead pay more attention to the product design process [3]. In addition, a natural HCI gives designers the freedom to creatively express their ideas and allows them to rapidly generate a wide variety of concepts, which will increase designers' ability to innovate.

4. Ways to Implement Natural HCI for CAD

Multimodal interfaces provide a promising solution to realize natural interactions for CAD applications. In this study, we consider the following single-mode components: (1) eye tracking, (2) gesture recognition, (3) speech recognition, and (4) brain–computer interfaces (BCIs). We discuss each of them briefly before turning to the key topic of their multimodal integration in this section. An overview of core references is presented in Table 1.

Table 1. Overview of core references for natural HCIs for CAD.

Category of Interfaces	Descriptions	References
Unimodal	Eye Tracking	[39–41]
	Gesture	[5,7,25,27,42–47]
	Speech	[24,48–55]
	BCI	[9,28,56–61]
Multimodal	Gesture + Speech	[8,11,12,62–68]
	Gesture + Eye Tracking	[6,34,69]
	Gesture + BCI	[70]
	Gesture + Speech + BCI	[3]
	Others	[30,71–75]

4.1. Eye Tracking-Based Interaction for CAD

The human gaze is defined as the direction to which the eyes are pointing in space, which could indicate where the attention of the user is focused [76,77]. Eye tracking is a technique in which eye-trackers record a user's eye movements to determine where on the screen the user's attention is directed when looking at the screen [78–80]. Ever since 1879, eye tracking has been used extensively in psychological science in the field of reading and image perception, and more recently in neuroscience and computer applications, especially HCI applications [80].

In the area of HCIs, eye tracking offers an important interface for a range of input methods including pointing and selecting [81], text input via virtual keyboards [82], gaze gestures [83], and drawing [84]. By tracking the direction and position of human eyes, the user's intention could be judged and analyzed directly and rapidly [80,85,86]. For example, Connor et al. [87] developed two eyeLook applications (i.e., seeTV and seeTXT) based on a gaze-aware mobile computing platform. For seeTV, a mobile video player, gaze is used as a trigger condition such that content playback is automatically paused when the user is not looking. seeTXT, as an attentive speed-reading application, can flash words on a display, advancing the text only when the user is looking. Similarly, Nagamatsu et al. [88] integrated an eye-tracking-based interface into a handheld mobile device, through which the user is able to move the cursor and interact with a mobile device through their gaze. Furthermore, Miluzzo and Wang [89] demonstrated the capability of a smart phone to track the user's eye movements across the phone. They presented a prototype implementation of EyePhone on a Nokia phone, enabling the user to operate a phone in touch-free manner. Incorporating such a facility in the process of interaction enhances the entertainment value and practicability, and benefits the development of natural HCIs.

For CAD applications, eye tracking could be used to interpret the user's intentions for modeling and operating a 3D model effectively in a touch-free manner [34,39]. For example, model selection is one of the fundamental operations and also the initial task for the most common user interactions in CAD modeling [41]. Using eye-tracking technology, the user could select models in 2D/3D environments directly and quickly. Ryu et al. [6] and Pouke et al. [69] adopted eye tracking as an extra channel of input and employed the user's gaze to find the intersection point with the surface of the target to select 3D objects in a virtual environment.

In addition, eye tracking could also be used in certain CAD operation tasks as an independent pointing interface, such as rotation, zooming, and translation. In traditional CAD systems, free rotation is obviously difficult, requiring additional commands for setting the rotation center [34]. With an eye-tracking-based interface, the user in the CAD environment can easily choose the rotation center without much effort just by fixating their gaze on it. For the zooming task, in order to explore detailed information through zooming, the user needs to preset the point of interest to the target that the model view camera moves towards in CAD applications. To ensure intuitiveness of operation, the location of interest is fixated and recorded by the eye-tracker, while the zooming operation could be controlled by another modality, e.g., a hand gesture [34]. Furthermore, the direction of translation

operation in CAD could also be designed to be controlled by gazing. During all of the above operations, the application of eye tracking could decrease the user's effort in position error feedback and reduce the fatigue of users, since it involves moving their eyes instead of hands. Thus, an eye-tracking-based interface in CAD can be a more natural and intuitive interface than traditional mouse and keyboard-based interfaces.

In sum, eye-tracking technology presents itself as a potentially natural interface for CAD applications. However, eye tracking has rarely been used independently as an interactional modality for CAD modeling. To achieve more complex modeling tasks, researchers have focused on combining eye tracking with other modalities [32,69,79], as described in Section 4.5.

4.2. Gesture Recognition-Based Interaction for CAD

Gestures are a form of body language that offer an effective way of communicating with others. People use a variety of gestures ranging from simple ones (such as pointing and pinching) to more complex actions for expressing feelings [90]. In recent years, the maturity of gesture-recognition technology based on depth vision has promoted the development of gesture interaction technology in applications of CAD [91]. Gestural interfaces aim to offer highly intuitive and free-form modeling modes that emulate interactions with physical products [42] and allow designers to create 3D conceptual models quickly, while just requiring a minimal amount of CAD software experience [43]. Approaches and modes using gestures to execute CAD modeling manipulation tasks will be covered in this section.

According to their spatiotemporal characteristics, gestures used in CAD interactions can mainly be divided into static and dynamic ones [25]. Research on static gestures considers the position information of a gesture, while that on dynamic gestures needs to consider not only the change in spatial position of the gesture, but also that in the expression of a gesture during the time sequence. Therefore, static gestures could be mainly used to invoke CAD commands for model-creation tasks while dynamic gestures could be used for model manipulation and modification [27]. In some studies [3,34,92], static gestures and dynamic gestures are combined for CAD manipulation instructions. For example, users can translate a model by closing the left hand into a fist (i.e., static gesture) and moving the right hand in the direction they want to translate it (i.e., dynamic gesture) [92,93]. In this process, the left-hand gesture plays the role of a trigger signal for the translation task while the right-hand gesture controls the direction and distance of the translation.

In our literature review, we found that researchers designed various sets or repositories of gestures for carrying out various CAD model tasks in the past [7,8,25,44]. The gestures in these repositories could, in principle, readily be used to test some gesture recognition algorithms and also directly as resources to select gestures in the application of CAD interfaces. However, most gestures in these repositories are typically chosen by researchers rather than users, often for the sake of ease of implementation with existing technology, while ignoring users' preferences. It could be difficult for users to remember this information and, as a result, their cognitive loads are enhanced rather than reduced [94]. To advance from this situation, user-based research has been established as an effective gesture elicitation method [7,95]. Vuletic [7] and Thakur [25] developed user-based gesture vocabulary for conceptual design via evaluating the user's activities in the modeling process in order to achieve a higher adoptability of gestures by inexperienced users and reduce the need for gesture learning.

Moreover, some authors explored the use of simple prescribed, even if free-form, gestures, such as pinching or grasping with one's hand, to quickly create a variety of constrained and free-form shapes without the need for extensive gesture training [45,46,96]. Their evaluations demonstrated that it is possible to enable users to express their intents for modeling without the need for a fixed set of gestures. Therefore, a potential future gesture system for CAD could allow individual designers to use non-prescribed gestures that will support rather than inhibit their conceptual design thinking. Meanwhile, further research into both the procedural structure of CAD-based activities and the ways in which they

might change depending on the shape being created could be conducted to explore better gesture-based design patterns that adapt to users and their specific workflows as they are used [7].

4.3. Speech Recognition-Based Interaction for CAD

With the rapid development of speech synthesis and recognition technologies in recent years, speech-based HCIs have been extensively employed in a variety of household, automobile, office, and driving applications [48,97]. In particular, voice-enabled intelligent personal assistants (IPAs), such as Amazon's Alexa, Apple's Siri, Google Assistant, and Microsoft Cortana, are widely available on a number of automatic devices [97]. Using a speech-based interface, routine operations can be efficiently executed with intuitive voice commands and the ease of use of the system can be improved. As well as in the household, speech-recognition technology can be applied in the context of CAD.

CAD modeling involves the development of a virtual object, always using step-by-step commands of instantiating primitives and using operations such as move, copy, rotate, and zoom for model manipulation and assembly [11]. These CAD commands could be realized naturally and directly through a speech-based interface. Therefore, some studies tried converting the names of menus and icons into voice-controlled commands [48–53]. Most of these studies were aimed at using fixed expressions to perform modeling operations. For example, Gao et al. [98] developed a prototype system with more than 80 speech commands which were primarily used to activate different CAD operations, for instance, selecting various operations of the system or switching from one mode to another. Similar research was also carried out by [50,52,53].

However, for all these speech-based CAD systems, CAD commands can only be activated when users utter the exact corresponding words or expressions [51]. So, novice users still have to devote considerable time to familiarize themselves with these systems and remember all the fixed vocabularies or preset expressions, which limits the utility of these speech-based CAD systems [54]. In order to implement a more flexible modeling operation, X.Y. Kou and S.K. Xue [24,54] proposed integrating a semantic inference approach into a speech-based CAD system, so users will no longer be constrained by predefined commands. In such a system, instead of using one-to-one mapping from word expressions to model actions, users can, for example, employ various expressions, such as "Rectangle", "Draw a Rectangle" and "Create a Rectangle", all with the same effect, namely, to generate a rectangle. This frees users from the need to memorize a host of predefined voice commands [24]. Whereas it is plain to see how this would work for simple commands, recognition methods should be further advanced for a speech-based CAD system to deal flexibly and intelligently with complex verbal commands [24,55].

Note that, as an effective interaction modality for CAD modeling, speech has been mostly used in conjunction with other interaction interfaces, such as "speech and gesture" [62], "speech and EEG" [3], and "speech and eye tracking" [32]. For this topic, more detailed studies will be reviewed in Section 4.5.

4.4. BCI-Based Interaction for CAD

Brain–computer interfaces (BCIs) provide a novel communication and control channel from the brain to an output device without the involvement of users' peripheral nerves and muscles [99]. Typically, brain activity can be detected using a variety of approaches, such as EEG, magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), electrocorticography (ECoG), and near-infrared spectroscopy (NIRS) [100]. Among them, the EEG signal is considered as the input of choice in most BCI systems to relate patterns in brain signals to the users' thoughts and intentions [28]. The very first advancements in this respect were made by Pfurtscheller [101], who for the first time demonstrated that it is possible for users to move through a virtual street just by imagining their feet movements. Lécuyer [102] and Zhao [103] also developed some prototypes allowing users to navigate in virtual scenes and manipulate virtual objects just through EEG signals. Furthermore,

Trejo [56] and Li [104] further extended the application of BCI to cursor movement. These studies show that BCI, as a new non-muscular channel, could be used in many applications involving human cognition, such as computer games, vehicle driving, assistive appliances, and neural prostheses [99].

In CAD systems, BCI could offer a more intuitive and natural pattern of interaction between the user and CAD application. Esfahani and Sundararajan [57] conducted the first study that investigated the possibility of using BCI for geometry selection. They used the evoked potential component P300-based BCI for selecting different target surfaces of geometrical objects in the CAD systems. Some other important functions for CAD application such as creating models or manipulating models via BCI have been studied by other researchers [9,28,58,59,105]. Esfahani and Sundararajan [58] also carried out experiments to distinguish between different primitive shapes based on users' EEG activity, including cube, sphere, cylinder, pyramid and cone shapes. The EEG headset was used to collect brain signals from 14 locations on the scalp and a linear discriminant classifier was trained to discriminate between the five basic primitive objects with an average accuracy of about 44.6%, significantly above the chance level of 20%. Postelnicu [60] conducted similar studies where the user was able to create and modify geometrical models by using EEG signals. As an important task in CAD applications, model manipulation-based BCI was also realized by [9,59,60]. In another recent work, Sree [61] used EEG and electromyogram (EMG) signals from facial movement using the Emotiv headset to enable users to carry out certain tasks in the CAD environment of Google SketchUp. Meanwhile, a human-factors study assessing the usability of the EEG/EMG interface was performed with participants from different backgrounds. The results suggested that the BCI-based system could help to lower the learning curve and has high usability as a generalized medium for CAD modeling [61].

In sum, BCI shows great potential in allowing users in CAD applications to create, modify and manipulate models directly. However, there are still some common limitations on the practical applicability of BCI resulting from the difficulty of brain signal acquisition and the limited accuracy of instruction discrimination. Future, lightweight EEG acquisition equipment and high-precision classification algorithms will become the new frontier for developing BCI-based CAD applications. Additionally, detecting the emotional state and user's satisfaction from EEG signals could be used in CAD systems to correct for errors and to strengthen proper classifications, which will make the system more reliable [106].

4.5. Multimodal HCI for CAD

Multimodal interfaces aim to construct interactive systems that leverage natural human capabilities to communicate through different modalities such as speech, gesture, gaze, and others, bringing more accurate and robust intention recognition methods to human–computer interaction [29]. Thanks to advances in the development of pattern-recognition techniques (in natural language processing, machine vision, etc.) and hardware technologies of input and output (cameras and sensors, etc.), there has been a significant increase in multimodal HCI research [107]. Turk [29], Jaímes [33], and Dumas [108] performed valuable surveys on the current status of multimodal HCI research. From these studies, we can easily see that the main goal of research on multimodal interactions is to explore a more transparent, flexible, efficient, and natural interface that can remove existing constraints on what is possible in the field of HCIs, and progress towards the full use of human communication and interaction capabilities. The realization of this aim depends on collaboration among researchers from different disciplines, such as computer scientists, mechanical engineers, cognitive scientists, and other experts [29].

Meanwhile, multimodal interfaces are considered to offer an improved user experience and better control than a unimodal input in the application of CAD modeling. The pioneering study by Bolt [63] in creating a "Put That There" system showed the potential and advantages of combining gesture and speech inputs for the selection and displacement of virtual models. This system proposed the use of commands such as "Move that to the

right of the green square" and "Delete that", allowing users to employ vague language to activate CAD functions while disambiguating them with gestures. It is noteworthy that none of these commands can be interpreted properly from just a single input modality—at least two were required—but this multimodal interface created simple and expressive commands that are natural to users.

The work by Bolt [63] led to a rapid development of many multimodal interfaces that typically used gesture, speech, eye-tracking, or BCI-based input modals for creating and manipulating CAD models, as is shown in Figure 1. From the survey results, "Gesture and Speech" is the most widely used interface in the application of CAD and drove the majority of multimodal interface research. Alternative formulations have also been followed, bringing new multimodal interfaces such as "Gesture and Eye Tracking" [34,69], "Gesture and BCI" [70], "Gesture, Speech and BCI" [3] and others [30,71–75]. For these multimodal interface-based CAD applications, there are three main types of combined modalities, as shown below.

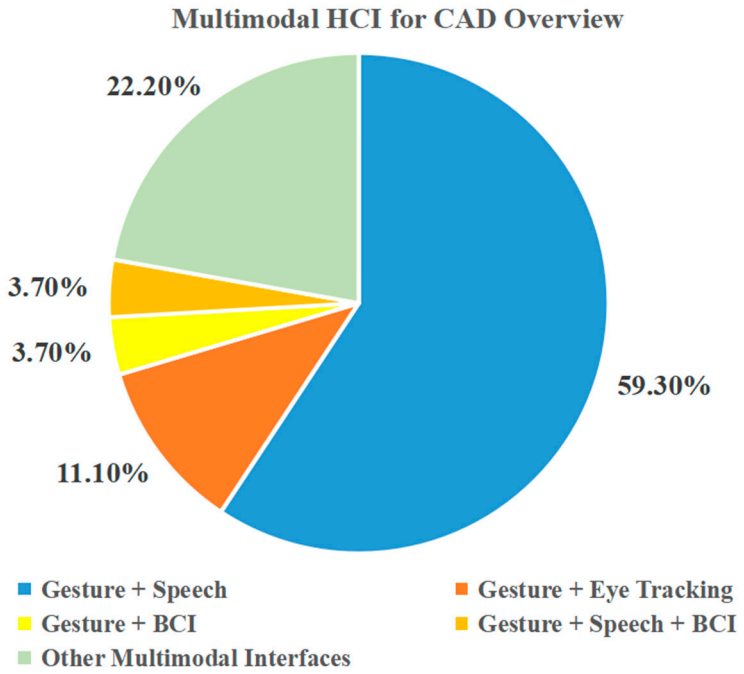


Figure 1. Survey on multimodal interfaces for CAD.

1. Different modalities can be used to execute different CAD commands. For example, in [69], eye tracking is used to select CAD models and gesture is used to activate different types of manipulation commands.
2. Different modalities can be used to execute the same CAD commands at the same time. In another words, there is overlap in the set of CAD commands that can be implemented through different modalities. For example, the task of creating models can be completed both by BCI-enabled commands and speech commands [3]. In this case, users can choose different interfaces according to their preferences. Moreover, to a certain extent, this kind of overlap in commands improves the robustness and flexibility of the interfaces.
3. Different modalities can be combined to execute a CAD command synergistically. In this respect, the advantages of different interfaces can be leveraged to work together on a model manipulation task. For example, in most CAD prototype systems combined with speech and gesture [8,12,64], speech is used to activate specific model manipulation commands (translation, rotation, zooming, etc.), while gestures are used to control the magnitude and direction of specific manipulations.

Most studies on multimodal interaction have concluded that using eye tracking, gesture, speech, and BCI for conceptual CAD modeling is easy to learn and use [3,8,12]. However, they mainly focused on the feasibility of the multimodal interface developed, considering aspects such as the accuracy and efficiency of user intent recognition, rather than user experience [3]. Notably, one of the most important goals of multimodal interfaces is to provide a natural interactive experience for CAD users. Hence, it is essential to carry out human factor studies for multimodal systems. Song [34] tested the intuitiveness and comfort levels of a multimodal system, i.e., GaFinC, by means of user interviews. Participants reported that it was very natural to pick the point of interest using their own gaze. However, they felt more uncomfortable and fatigued using two hands as opposed to one during the process of CAD modeling. In addition, in most gesture-based multimodal systems, the gestures used are always prescribed in advance and users have to remember complex gesture vocabulary before modeling, which significantly increases the additional cognitive load imposed on users [7]. All the issues mentioned above will benefit from the development of future multimodal-based CAD systems.

4.6. Other Interfaces for CAD

In addition to the interfaces listed in the previous sections, a short overview of other interfaces developed is given in this section.

Touch-based interfaces are those in which designers can input design information on a touchscreen with their fingers. Sharma et al. [71] developed a multimodal interface, "MozArt", which combines speech and touch for creating and interacting with CAD models. This system features a minimalist user interface and uses a touch table whose orientation can be changed depending on the needs of the user. While MozArt provides users with efficient results, the main problem with this interface is that the touch-based system is subject to ambiguity since it is two dimensional in nature. To resolve these ambiguities, EEG is considered here as the main method to resolve these ambiguities. Therefore, a novel system combining touch and BCI was proposed by Bhat [30] in order to reduce the ambiguity associated with 2D touch to perform 3D model manipulation. The experimental results have shown that BCI could help to identify the true intention of the users while performing the modeling tasks by touch [30].

Virtual reality (VR) is providing a new perspective for human interaction with CAD models, allowing for the stereoscopic visualization of virtual objects on one-to-one scale within the working space of users [74]. In a VR-based system, the user can directly create, modify, and manipulate models without learning complex protocols and procedures [109]. Haptic interface, as a promising 3D HCI for various applications, could allow the user to feel the contact force in the process of free-form design and more realistically convey virtual reality to humans so as to aid the design process [110–112]. To construct more immersive design environments for users, Damien Chamaret [113] and Simon Kind [114] combined haptic and VR in multimodal systems and found that haptic feedback can provide a more accurate manipulation of virtual objects and significantly improve the working efficiency of CAD systems. On this basis, a novel multimodal application, "VRSolid", was proposed by Mogan [73], further extending the input modalities within the VR environment. The VRSolid system consists of five modalities, including two for output (haptic feedback and 3D visual) and three for input (gesture, speech, and motion tracking). Following this, Bourdot [72] and Toma [35] constructed some similar multimodal systems for CAD in a VR environment. All applications mentioned above are designed to provide an intuitiveness and natural environment for interaction with CAD models, while creating a common multimodal interface that can be integrated with the existing commercial CAD software, such as SolidWorks, Creo, and Parasolid [35,72,73].

Additionally, Stark et al. [74] tested "pen and VR" in an application called VR-cave, which has a physical pen for free-hand sketching in 2D and 3D immersive modeling environments. Xiangshi et al. [75] investigated the performances of four modes of CAD applications, such as, "mouse", "pen", "pen and speech", and "pen, speech and mouse".

They found that "pen, speech and mouse" was the best multimodal combination in terms of drawing time, modification time, and subjective preferences among these four modes [75].

All the studies reviewed in this section further emphasize the diversity of CAD interface development research and the difference in the outcome of the interface evaluations. Further exploring more natural interfaces and integration methods for CAD application will be an important issue for future studies.

5. Evaluation of Natural HCI for CAD

To evaluate the performance and usability of natural interfaces for CAD, we need to establish whether the interface is able to carry out tasks effectively and efficiently and if it is also easy to understand and use for novices in the field of CAD [35]. Generally, each novel interaction system needs to pass several assessment and modification cycles during its development so as to reach a high quality and ensure user satisfaction [115]. Consequently, the evaluation of the natural HCI is one of the indispensable parts during system development.

From the literature review, the evaluation factors can be broadly categorized into two groups, from the system performance and user experience perspectives [115,116]. For the system performance, Effectiveness and Efficiency are the two main evaluation metrics. For the user experience, evaluation metrics include Learnability, User preference, Cognitive load, and Physical fatigue. The specific description about these metrics is as follows:

1. Effectiveness: This metric refers to the recognition accuracy (or error rate) and completeness of modeling tasks while using the novel CAD interaction system. To compute the accuracy parameter, Song et al. [34] carried out a manipulation test for the developed multimodal interface and the errors of translation, rotation, and zooming were recorded. The test results showed that the user performed the zooming task with higher accuracy than the translation task in the novel interaction system.
2. Efficiency: This metric is mainly established in terms of the task completion time and includes the system response time. In the view of [117], the time period, when the user is planning how to operate a generate/modified task through the user interface in a CAD system, should also be included as a component of task completion time in addition to the operation execution time. From the time test results, Chu et al. [117] found that using the novel multimodal interface results in a speedup of 1.5 to 2.0 times over the traditional interface (mouse and keyboard) in the process of building simple models, such as block, cylinder, sphere, etc. For more complex models, modeling with a traditional interface takes three to four times longer than with the novel interface [117]. The main factor contributing to these differences is that the number of operations required for traditional interface increases significantly relative to the novel interface with the increase in model complexity. Additionally, the total task completion time is proportional to the total number of operations for modeling. It is worth noting that only users with the same level of experience for CAD were considered in the study of Chu et al. [117]. Future research is needed to conduct a comparative experimental analysis of system performance using users with different experience levels.
3. Learnability: This metric describes how easy it is for novice users to interact with CAD systems and accomplish modeling tasks. For multimodal HCIs, learnability is a kind of quality that allows novice users to quickly become familiar with these interfaces and make good use of all their features and capabilities. Generally, the use of user questionnaire is a common method for evaluating the learnability of interaction systems. In the studies of [11,64], users were asked to rate the task on a five-point Likert scale in terms of the ease of performing tasks and satisfaction with the modeling results after the completion of tasks. The results showed that for a multimodal interface-based CAD system, users' perceptions of the ease of performing tasks were high for all tasks, except for rotation and modeling tasks [64]. On the other hand, some quantitative parameters can also reflect the learnability of a system.

For example, the improvement of user performance associated with task repetitions was used to evaluate learnability in [113], including the task completion time and number of collisions. Meanwhile, Chu et al. [117] focused on the number of design steps for modeling and the idle time period when the user knows exactly what to do and is planning how the interaction approach may be performed to achieve the design tasks. In the study of [117], compared to traditional interfaces, a VR-based multimodal system required fewer design steps and less idle time in the modeling process, which indicates that the user could quickly understand what operations need to be performed and complete the modeling tasks.

4. User preference: This metric refers to the subjective evaluation of the user's pleasure in using the novel interface based on the user characteristics, including static characteristics (e.g., age, gender, and language) and dynamical characteristics (e.g., motivation, emotional status, experience, and domain knowledge). This metric is also an important reflection of user-centered natural interaction design. In terms of multimodal systems, it can offer a set of modalities for user input and output depending on the interactional environment and functional requirements [115]. User experience and preference should be fully taken account in the process of the choice of different modalities and the design of a unimodal interface. In the studies of [3,118,119], participants were questioned about their preferred interaction interface and most of them appreciated the use of an immersive multimodal interface for CAD modeling, which can offer a more intuitive interactive environment than traditional interfaces where the user does not have to navigate through a series of windows and menus in order to achieve a desired action. Additionally, the unimodal interfaces, such as gesture-based interfaces [7] and speech-based interfaces [24], should be designed taking into account user preference, not just based on the expert opinion.
5. Cognitive load: This metric refers to the amount of mental activity imposed on working memory in the process of modeling tasks, including the necessary information processing capacity and resources. In terms of cognitive psychology [120], the product designer is considered as a cognitive system utilizing CAD systems for modeling efficiently [35]. The cognitive interaction between the designer and system is performed by various types of interfaces. In general, the use of multimodal interfaces could make better use of the human cognitive resources and will thus result in a lower cognitive load on the user during the process of modeling [115,121]. To verify this point, Sharma et al. [71] collected subjective task load data with the NASA TLX method after each task, and evaluated the cognitive load of users in modeling with multitouch and multimodal systems. Following this, the qualitative self-reported data and EEG signals were used to analyze the user's cognitive load during the process of modeling in the multimodal interface-based system [26]. The results showed that compared with the results of expert users, novice users' alpha band activity increased when using a multimodal interface, which indicates that their cognitive load decreased [26]. This result was confirmed in the questionnaire responses. Consequently, it was easy for novice users to use a new set of multimodal inputs for modeling compared to expert users. Additionally, some performance features (e.g., reaction time and error rate) and psycho-physiological parameters (e.g., blink rate and pupil diameter) can also be used to evaluate the cognitive load of user [122].
6. Physical fatigue: This metric refers to the physical effort needed for users while interacting with the CAD system in order to perform modeling tasks. Even if the multimodal interface puts users in the loop of an intuitive and immersive interactive environment, the repeated execution of body movements (e.g., arm, hand, eye, etc.) for a long time can lead to fatigue and stressful factors for users [35]. Therefore, it is necessary to evaluate multimodal interface-based CAD systems assessing the body motion in the performance of modeling tasks so as to reduce body fatigue and increase user satisfaction in the design process. Administering a questionnaire to users is a mainstream method adopted to evaluate the physical fatigue caused by using

multimodal interfaces [6,64]. Meanwhile, the overall body posture, movement pattern, and hand movement distance can also be used as analytical factors for physical fatigue. Toma [35] used the distance covered by hand movement in the performance of a task to compare and evaluate the VR-CAD system with a traditional desktop workspace, and found that multimodal interfaces lead to higher physical fatigue, with the hand movement distance being on average 1.6 times greater than the desktop interface for the modeling process.

The presented metrics are not yet exhausted, and some other characteristics can still be considered to evaluate a multimodal interface-based CAD system, such as the number and type of interactive modalities; the ability to use modes in parallel, serially, or both; the size and type of recognition vocabularies; and methods for integrating multiple sensors and the types of applications supported [29].

As the modeling requirements become more complex, evaluating the system must become more formalized. Evaluating metrics will become more quantitative and evaluating methods will become more objective. The ultimate goal of analyzing these metrics is to create a prototype system that users like and can be used to perform modeling tasks in a natural and intuitive approach [123]. After conducting a system evaluation, it is important to record what was observed in addition to why such a behavior occurred and modify the interfaces according to the results. It is noteworthy that an effective evaluation may not generate a direct solution to the problems, but can provide modified design guidelines for continued testing [124].

6. Devices for Natural HCI

Although these interfaces are unlikely to completely replace traditional WIMP-based interfaces, the importance of multimodal natural HCIs is growing due to advances in hardware and software, the benefits that they can provide to users, and the natural fit with the increasingly ubiquitous mobile computing environment [125].

In the process of CAD modeling, signals that can represent the design intention should first be obtained to identify the user intention. Thus, signal acquisition and identification devices are important for achieving a natural HCI. In this section, the devices used for eye tracking, gesture recognition, speech recognition, and BCI are introduced in detail, as presented in Table 2. In additional, some other devices related to natural HCIs are also briefly described.

Table 2. Overview of devices for natural HCIs.

Signal Modalities	Categories	Devices
Eye tracking	Head-mounted	Dikablis Glass 3.0 [126], Tobii Glass 2 [127], Pupil Labs Glasses [128]
	Tabletop	Gazepoint GP3 [129], Tobii Pro Spectrum 150 [130], EyeLink 1000 Plus [131]
Gesture	Sensor-based	MoCap Pro Glove [132], Cyber Glove [133], Vrtrix Glove [134]
	Vision-based	Kinect [135], Leap Motion [44,136]
Speech	DP-100 Connected Speech Recognition System (CSRS) [63], Microsoft Speech API (SAPI) [54], CMUSphinx [137]	
EEG	Saline electrode	Emotiv Epoc+ [3,58]
	Wet electrode	Neuroscan [56]
	Dry electrode	Neurosky MindWave [138], InteraXon Muse [139]
Others	Haptic	SensAble PHANTom [110], SPIDAR [113]
	VR/AR	HTC Vive [140], Oculus Rift [141], HoloLens 2 [142]

6.1. Devices for Eye Tracking

Eye tracking technology is used to track the eye movement state of the user and recognize where the user is looking on a computer screen through some advanced optical recognition methods, such as the pupil–corneal reflection method [143], pupillary-canthus method [144], HN method [145], and so on. The first eye-tracker that could measure eye movements quantitatively was developed by Dodge and Cline in 1901 [146]. With the dramatic improvement in eye-tracking technology, more advanced and easy-to-use hardware has been developed on the market in recent years [147]. Currently, there are two main categories of eye-trackers available as interactive input devices, including head-mounted devices and tabletop devices.

The head-mounted eye-tracker, which usually needs to be attached to the eyes as special contact lenses or headsets, is composed of a scene camera, which records the user's first-person view, and an eye camera, which continuously records the changes of the sight line by using the infrared ray of cornea and pupil reflection [148]. Some common commercial head-mounted eye-trackers can be used for HCIs, as shown in Figure 2. Obviously, these head-mounted devices are inconvenient and involve some physical load for users. Therefore, the lightweight design of head-mounted eye-trackers is needed to make them more widely used in natural HCIs.



(**a**) (**b**) (**c**)

Figure 2. Commercial head-mounted eye-trackers: (**a**) Dikablis Glass 3.0 [126]; (**b**) Tobii Glass 2 [127]; (**c**) Pupil Labs Glasses [128].

Compared to the head-mounted tracer, the tabletop eye-tracker is more comfortable and flexible for users as it can be integrated with a monitor and does not require the user to wear any equipment, as shown in Figure 3. However, most of these commercial eye-trackers are designed to target the PC environment, with a short distance of 50 to 80 cm between users and eye-trackers, which limits their application in the field of HCIs [32]. Moreover, complex calibration between the eye-tracker and display coordinates is required and the user's freedom of head movement is limited owing to the fixed camera. All the mentioned drawbacks need to be improved and optimized to promote the future development of tabletop eye-trackers.



(**a**) (**b**) (**c**)

Figure 3. Commercial tabletop eye-trackers: (**a**) Gazepoint GP3 [129]; (**b**) Tobii Pro Spectrum 150 [130]; (**c**) EyeLink 1000 Plus [131].

6.2. Devices for Gesture Recognition

Gesture recognition involves algorithms to detect and recognize the movement of fingers, palms, arms, or the entire body so as to interpret the user's interaction intent. Currently, the devices for gesture recognition can be broadly divided into sensor-based devices, such as data gloves, and vision-based devices, such as normal video cameras or depth-aware cameras.

Sensor-based devices can directly capture the motion and position of hand gesture by using data gloves. Commonly used sensors mainly include EMG sensors [149], bending sensors [150], pressure sensors [151], and acceleration sensors [152]. By using these glove sensors, we can easily obtain various gesture signals and accurately identify the hand pose and movement. Representative products in data gloves for gesture recognition are MoCap Pro Glove, Cyber Glove, and Vrtrix Glove, as shown in Figure 4. However, these data gloves are too expensive and their wired sensors restrict natural hand movement [47]. To overcome these limitations, the vision-based approach came into existence [93].



(**a**) (**b**) (**c**)

Figure 4. Commercial data gloves for gesture recognition: (**a**) MoCap Pro Glove [132]; (**b**) Cyber Glove [133]; (**c**) Vrtrix Glove [134].

The vision-based device is less constrained and moves more naturally than the sensor-based one, and does not require users to wear anything over the hands. This device is relatively cheap, simple, natural, and convenient to use in the application of CAD. At present, representative vision-based products used in the process of CAD are Kinect and Leap Motion, as shown in Figure 5. Vinayak [45] developed a novel gesture-based 3D modeling system and used Kinect to complete the recognition and classification of human skeletal and hand posture in the process of modeling. However, the recognition performance of vision-based devices is sensitive and easily affected by environmental factors such as lighting conditions and cluttered backgrounds [153]. Therefore, the illumination change, multiuser cluster, partial or full occlusions are important challenges for vision-based gesture input devices to be addressed in the future study [93].



(**a**) (**b**)

Figure 5. Commercial vision-based devices for gesture recognition: (**a**) Kinect [135]; (**b**) Leap Motion [136].

6.3. Devices for Speech Recognition

Speech recognition-based interfaces provide a natural method to operate and modify the CAD models for users. Along with the development of signal input devices and natural

language-processing technology, the accuracy and effectiveness of speech recognition has been significantly improved, which further promotes the application of speech-based interfaces for CAD.

Generally, the devices for speech recognition consist of two main modules, i.e., an input microphone and speech recognition module. After converting the analog signal with speech information into a digital signal through the input microphone, the speech recognition module integrated with natural language-processing algorithms completes the final recognition of semantics [154]. Obviously, the speech recognition module is a cord part of this system as an internal algorithm processor. Additionally, many companies have developed a variety of speech-recognition modules. For example, NEC (Nippon Electric Company) developed a system called DP-100 Connected Speech Recognition System (CSRS) [63] which is able to recognize a limited number of connect voices without pause. Microsoft also introduced its own voice-recognition interface for computers, Microsoft Speech API (SAPI) [54]. Carnegie Mellon University developed a speech-recognition toolbox, CMUSphinx, that can be used for microphone speech recognition [137].

In addition, machine learning and deep learning algorithms such as the Hidden Markov Model (HMM) [155], Convolutional Neural Network (CNN) [156] and Long Short-term Memory (LSTM) [157] have been widely used in the core part of speech recognition modules to improve the accuracy of speech recognition.

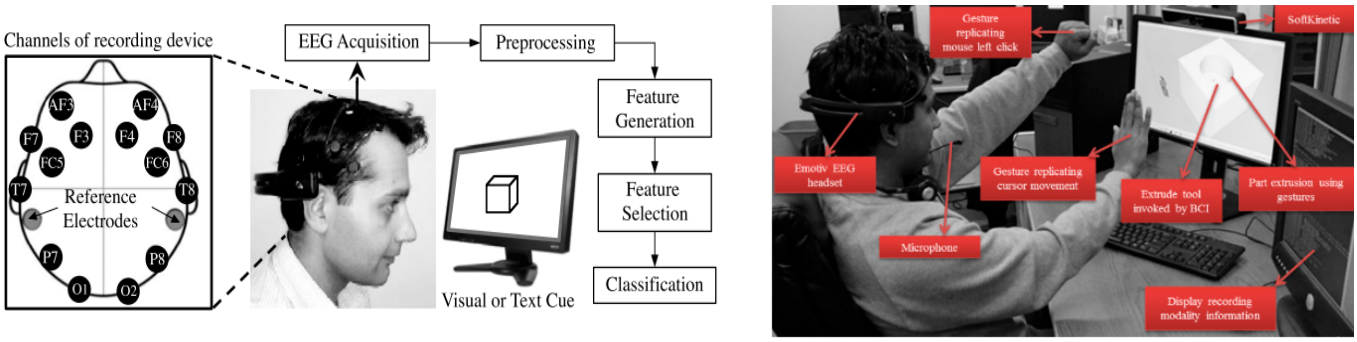
6.4. Devices for EEG Acquisition

The EEG acquisition device is a piece of equipment that can record the electrical activity of the brain through some specific electrodes placed on a user's scalp [158]. In 1924, the first EEG detection device in the world was developed by Berger in Jena, Germany [159]. Following the development of signal processing and microelectronics technology, the structure and function of EEG-acquisition devices have gradually matured. There are numerous EEG devices from different companies that are able to cater to the specific needs of EEG users, such as Neuroscan, Emotiv, and Neurosky MindWave.

Generally, EEG-acquisition devices can be divided into invasive, semi-invasive, and non-invasive categories according to the connection with the brain [160]. Among them, a non-invasive EEG device is mainly used in the area of natural HCIs due to the damage of the scalp caused by the invasive device and the semi-invasive device. For non-invasive EEG devices, the types of electrodes used can be categorized into saline electrode, wet electrode, and dry electrode [160]. Saline electrodes use salt water as a conductive medium, which is easy to carry and low in cost [161]. Wet electrodes usually need to use gel as a conductive medium and can collect high-quality EEG signals. Dry electrodes can directly contact the user's scalp to achieve conduction without any conductive medium which is more natural and convenient to the user, but the signal acquisition accuracy is limited [162].

In current studies in the literature, Emotiv Epoc+ is commonly used as an EEG-acquisition device in the application of CAD [3,57–59], as shown in Figure 6. Emotiv Epoc+ is a small and low-cost device used to record and recognize EEG and facial muscles' movements (EMG) with 14 saline electrodes. This device can communicate wirelessly with computer via a USB dongle and provide the Emotiv API, a C++ API, which can be used to easily obtain the EEG/EMG data and transform the user-specific action command into an easy-to-user structure [3].

Additionally, some other lightweight and low-cost EEG devices, such as Neurosky MindWave and InteraXon Muse, may also be used to build future BCI for CAD.



(**a**) (**b**)

Figure 6. Emotiv used for BCI in CAD applications: (**a**) constructing 3D models from BCI. Reproduced with permission from [58], Elsvier, 2012; (**b**) multimodal interface-based CAD system. Reproduced with permission from [3], ASME, 2013.

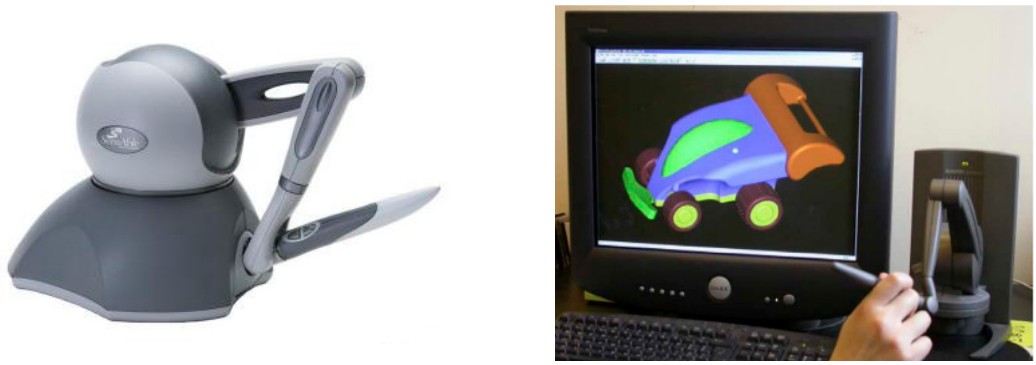
6.5. Other Devices for Natural HCI

In addition to the main devices mentioned above, some other devices for natural HCIs are briefly introduced in this part.

A haptic device is an interaction device which can deliver force feedback to a user, while interacting with the virtual environment (VE) or real environment (RE) [114]. This device could provide a more realistic interaction experience by making users feel the movement and touch the interaction models. The SensAble PHANTom [110,163,164] is one of the most common haptic devices used in CAD applications, as shown in Figure 7, which consists of a mechanical robotic arm with joints and a pen-shaped end-effector as manipulator. The arm follows the movement of the manipulator and is capable of generating a force equivalent to that applied to the manipulator [10]. Additionally, the wearable haptic device, SPIDAR, is also used to interact with virtual mock-ups [113], as shown in Figure 8. Compared to SensAble PHANTom, SPIDAR has a more flexible workspace and allows the user to grasp the virtual models with natural experience.

From the studies of [10,73,114], haptic devices are always applied in a virtual environment (VR) for CAD. In another word, VR devices are also needed and important for natural HCIs of CAD. Currently, the most popular VR/AR devices on the market are HTC Vive, Oculus Rift, HoloLens, and so on, as shown in Figure 9.

Additionally, some other interactive devices such as touchscreen and EMG-based devices, are not listed here due to their limited application in the area of CAD.



(**a**) (**b**)

Figure 7. Haptic device for CAD: (**a**) SensAble PHANTom; (**b**) haptic-based device for modeling system. Reproduced with permission from [110], Elsvier, 2005.

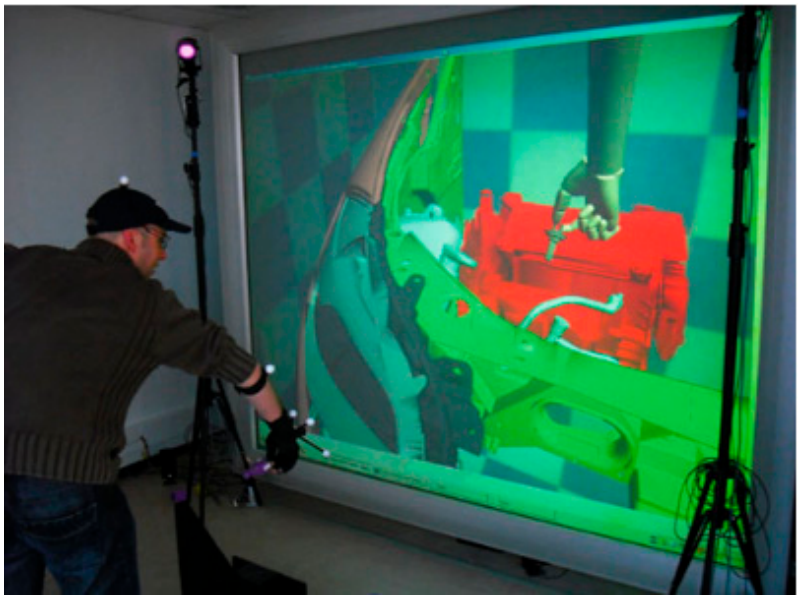


Figure 8. SPIDAR-based interaction system. Reproduced with permission from [113], Springer Nature, 2010.



(**a**) (**b**) (**c**)

Figure 9. VR/AR devices: (**a**) HTC Vive [140]; (**b**) Oculus Rift [141]; (**c**) HoloLens 2 [142].

7. The Future of Natural HCI for CAD

Multimodal natural HCIs for CAD constitute a growing research area and lots of research efforts are expected. To develop more natural and intuitive multimodal systems for CAD with better performance characteristics than traditional interfaces, many fundamental scientific issues and multidisciplinary research challenges still remain to be addressed. In this section, the challenges and potential research directions of multimodal natural interfaces for CAD are discussed.

7.1. Changing from a Large Scale and Independence to Lightweight, Miniaturization and Integration for Interaction Devices

Interaction devices that can accurately capture physiological signals such as the EEG, gaze, and gestures of the user are the technological basis for building natural HCIs in the application of CAD. Most existing interaction devices in this review are designed to be wearable and independently recognize the user's signals. However, these wearable devices' excessive size and weight can make the user easily fatigued in the process of using. Therefore, lightweight and miniaturized designs will be a mainstream trend in the future, which can enhance the user's comfort and decrease their physical fatigue by reducing the size and weight of interaction devices. For instance, Figure 10 shows the evolution of the EEG headset from conventional to in-ear EEG. Furthermore, the need for compact wearables also means that actuators and sensors should be nimble enough to fit into tight crevices on different bodies, especially for gloves, glasses, and EEG headsets [165]. With such requirements, fitting the actuation for signal acquisition of different modes could also

be a challenge. In other words, instead of employing different devices to capture distinct signals, researchers should focus more on integrating multiple acquisition functions in one device. This trend is exacerbated by the popularization of multimodal interfaces [10].

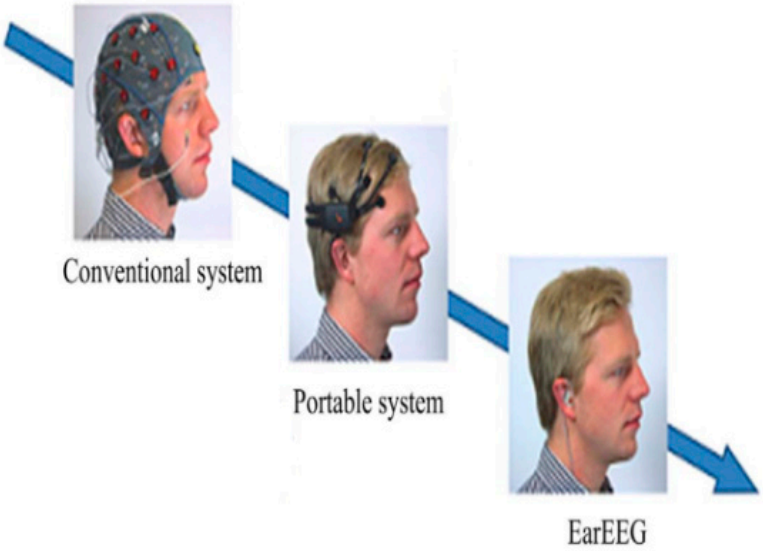


Figure 10. Evolution of EEG from conventional device to in-ear EEG. Reproduced with permission from [166], Elsvier, 2017.

7.2. Changing from System-Centered to User-Centered Interaction Patterns

Historically, the development of human–computer interfaces for CAD application has been a system-centered pattern, driven by the functional requirements of the system and the current technological capability enabling their implementation. This pattern mostly focuses on organizing the functionality of the system and builds the product in the designer's own interpretation and through the implementation of systems thinking, rather than the user's. This means that the system users have to memorize a large number of instructions in order to interact in a manner that matches system processing rules, which is not typically user friendly and contrary to the original intention of natural HCIs.

To build a natural HCI, the user's preferences should be fully considered in the design process, thus forming a user-centered interaction pattern. A user-centered pattern is mainly based on a user's abilities and real preferences, and focusing on the user rather than forcing the users to change their behavior to accommodate the interface. In the future, further empirical work is needed to build the user-centered interaction pattern for CAD, such as (a) developing natural and efficient unimodal interface for the CAD applications driven by what the users would naturally do instead of defined by the researchers developing the studies, just due to ease of application or alignment with the technology used, so that users do not need to learn too many particular commands or procedures and can focus more on the product design process, (b) giving users free control over the selection or combined use of input modes in a multimodal interface according to their preference to further reduce the cognitive load and improve their comfort, and (c) proposing a blended interaction style that combines a passive input mode (e.g., gaze tracking) and active input mode (e.g., BCI or gesture) toward supporting the recognition of users' natural activities in context [31,167].

7.3. Changing from Single Discriminant to Multimodal Fusion Analysis for Interaction Algorithms

One of the most fundamental issues for multimodal HCIs is to integrate input information from different sensing modalities during the process of modeling. However, different input modalities yield disparate data forms and rates, which makes the integration of such signals a difficult and challenging task. For most current systems, multimodal data are typically processed separately and only combined at the end without taking full advantage of the complementary and redundant information from multimodal signals [168]. There-

fore, to achieve the desired naturalness and robustness of the multimodal interface, more efficient and effective technologies of data fusion are needed to study.

Firstly, the temporal dimension is an important aspect in the process of multimodal fusion because different modalities may have different temporal constraints and semantic features. Some modal integrations are intended to be interpreted in parallel, which others may typically be offered sequentially [169]. In addition, not all multimodal systems are needed to fuse information to form a joint interpretation, which is more closely tied with the environment context and modeling task context [31]. Therefore, the questions about when to integrate and which needs to be combined should be solved according to a context-dependent model, including the temporal features and modeling tasks so as to provide better recognition results. Furthermore, based on the type of input modalities in a multimodal interface, different levels of integration methods should be defined, mainly including feature-level fusion and decision-level fusion [170]. For example, EEG and eye tracking have a strong correlation in the process of model selection, which can be fused at the feature level to increase the robustness and accuracy of user intent discrimination. Depending on the chosen level of fusion, the future integration can be performed using varies methods, ranging from simple feature concatenation to complex deep fusion algorithms. In particular, future work needs to continue to study the adaptive multimodal fusion architectures for realizing error handling and coupling multiple modalities in a natural manner [170,171].

7.4. Changing from Single-Person Design to Multiuser Collaborative Design for Interaction

Most of current multimodal interfaces for CAD have been developed as standalone single-user solutions, which unfortunately, are rarely designed to meet complex collaboration requirements for multiple people. Now more than ever, there is a growing need for collaborative CAD solutions, essentially moving away from the high-performance single user solutions of the past and towards more multiuser systems [172]. Therefore, in future interactive environments of product design, one important research goal will be allowing multiple users to interact together in larger physical spaces and ubiquitous forms, such as pointing to a shared display surface, gesturing concurrently, and talking in a natural manner [167]. For this target, some specific technological issues will need to be addressed because the current interfaces generally assume that a particular device provides input from a single user at a given moment, for example, how to capture data from multiple users at the same time and identify the contribution of an individual user to the overall multiuser interactive environment, as well as switching between different interactive subjects. The amount or size of multimodal devices that a user has to wear or carries will also need to be minimized.

In addition, due to the rapid development of network and VR/AR technologies, distributed collaborative design will be another important trend in the future. Technology for data sharing and communication allows users to share a design resource efficiently in co-located or distributed situations [173]. Multimodal interface-based AR/VR should also be designed and developed to support solid modeling and collaborative design activities in the distributed environment, where multiple users can simultaneously view and interact with virtual/real objects [174,175].

In sum, in order to meet the future requirements of the multiuser collaborative design of products, the multimodal natural HCI still faces an array of long-term research challenges.

8. Conclusions

To better understand the state-of-the-art research of multimodal natural HCIs for CAD, this paper summarizes the related research works and advanced techniques in the literature. The method of literature retrieval and research scope were determined. By analyzing the process of traditional product concept design from the retrieved papers, the requirements and effects of natural HCIs for CAD are presented. Then, the mainstream

techniques for implementing natural HCIs, evaluation indicators and supported devices are summarized. Finally, challenges and potential directions for natural HCIs are identified for further discussion.

From the studies reviewed in this paper, we observed that although many issues are being addressed in the field of multimodal natural HCIs for CAD, such as novel interface designs, interface performance evaluations, and device development for natural HCIs, the field is still young and needs further research to establish reliable multimodal natural HCI and mature applications for CAD. In particular, the current novel interfaces for CAD are mainly used in the conceptual design phase of products and it is difficult to meet the interactive functional requirements of the other design process, such as detailed design, simulation, and analysis. Therefore, research of new multimodal natural HCIs needs to be considered to support multiple phases of product design, not just the conceptual design.

Finally, this paper contributes to the development of multimodal natural HCIs for CAD in two main aspects: summarizing the up-to-date software- and hardware-supported techniques for natural HCIs in the application of CAD in many disciplines, and proposing key future research directions to be addressed in the field. Hopefully, this paper can provide an insightful informative reference to researchers for the development of multimodal natural HCIs of CAD and inspire future studies on natural HCIs.

Author Contributions: Conceptualization, H.N.; methodology, H.N. and J.H.; analysis, H.N.; writing—original draft preparation, H.N. and C.V.L.; writing—review and editing, C.V.L., G.W. and T.L.; supervision, J.H., C.V.L. and T.L.; funding acquisition, J.H. All authors have read and agreed to the published version of the manuscript.

Funding: This study is supported by National Ministry Projects of China (No. JCKY2018204B053).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Allen, J.; Kouppas, P. Computer Aided Design: Past, Present, Future. In *Design and Designing: A Critical Introduction*; BERG: London, UK, 2012; pp. 97–111.
2. Bilgin, M.S.; Baytaroğlu, E.N.; Erdem, A.; Dilber, E. A Review of Computer-Aided Design/Computer-Aided Manufacture Techniques for Removable Denture Fabrication. *Eur. J. Dent.* **2016**, *10*, 286–291. [CrossRef]
3. Nanjundaswamy, V.G.; Kulkarni, A.; Chen, Z.; Jaiswal, P.; Verma, A.; Rai, R. Intuitive 3D Computer-Aided Design (CAD) System with Multimodal Interfaces. In Proceedings of the 33rd Computers and Information in Engineering Conference, Portland, OR, USA, 4–7 August 2013; ASME: New York, NY, USA, 2013; Volume 2A, p. V02AT02A037. [CrossRef]
4. Wu, K.C.; Fernando, T. Novel Interface for Future CAD Environments. In Proceedings of the 2004 IEEE International Conference on Systems, Man and Cybernetics, The Hague, The Netherlands, 10–13 October 2004; Volume 7, pp. 6286–6290. [CrossRef]
5. Huang, J.; Rai, R. Conceptual Three-Dimensional Modeling Using Intuitive Gesture-Based Midair Three-Dimensional Sketching Technique. *J. Comput. Inf. Sci. Eng.* **2018**, *18*, 041014. [CrossRef]
6. Ryu, K.; Lee, J.J.; Park, J.M. GG Interaction: A Gaze–Grasp Pose Interaction for 3D Virtual Object Selection. *J. Multimodal User Interfaces* **2019**, *13*, 383–393. [CrossRef]
7. Vuletic, T.; Duffy, A.; McTeague, C.; Hay, L.; Brisco, R.; Campbell, G.; Grealy, M. A Novel User-Based Gesture Vocabulary for Conceptual Design. *Int. J. Hum. Comput. Stud.* **2021**, *150*, 102609. [CrossRef]
8. Friedrich, M.; Langer, S.; Frey, F. Combining Gesture and Voice Control for Mid-Air Manipulation of CAD Models in VR Environments. In Proceedings of the VISIGRAPP 2021—16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, Online, 8–10 February 2021; Volume 2, pp. 119–127. [CrossRef]
9. Huang, Y.-C.; Chen, K.-L. Brain-Computer Interfaces (BCI) Based 3D Computer-Aided Design (CAD): To Improve the Efficiency of 3D Modeling for New Users. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Schmorrow, D.D., Fidopiastis, C.M., Eds.; Springer International Publishing: Cham, Switzerland, 2017; Volume 10285, pp. 333–344. [CrossRef]
10. Toma, M.-I.; Postelnicu, C.; Antonya, C. Multi-Modal Interaction for 3D Modeling. *Bull. Transilv. Univ. Brasov. Eng. Sci.* **2010**, *3*, 137–144.

11. Khan, S.; Tunçer, B.; Subramanian, R.; Blessing, L. 3D CAD Modeling Using Gestures and Speech: Investigating CAD Legacy and Non-Legacy Procedures. In Proceedings of the "Hello, Culture"—18th International Conference on Computer-Aided Architectural Design Future (CAAD Future 2019), Daejeon, Korea, 26–28 June 2019; pp. 624–643.
12. Khan, S.; Tunçer, B. Gesture and Speech Elicitation for 3D CAD Modeling in Conceptual Design. *Autom. Constr.* **2019**, *106*, 102847. [CrossRef]
13. Câmara, A. Natural User Interfaces. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2011; Volume 6946, p. 1. [CrossRef]
14. Shao, L.; Shan, C.; Luo, J.; Etoh, M. *Multimedia Interaction and Intelligent User Interfaces: Principles, Methods and Applications*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2010.
15. Chu, M.; Begole, B. Natural and Implicit Information-Seeking Cues in Responsive Technology. In *Human-Centric Interfaces for Ambient Intelligence*; Elsevier: Amsterdam, The Netherlands, 2010; pp. 415–452.
16. Karray, F.; Alemzadeh, M.; Abou Saleh, J.; Nours Arab, M. Human-Computer Interaction: Overview on State of the Art. *Int. J. Smart Sens. Intell. Syst.* **2008**, *1*, 137–159. [CrossRef]
17. Beeby, W. The Future of Integrated CAD/CAM Systems: The Boeing Perspective. *IEEE Comput. Graph. Appl.* **1982**, *2*, 51–56. [CrossRef]
18. Fetter, W. A Computer Graphic Human Figure System Applicable to Kineseology. In Proceedings of the 15th Design Automation Conference, Las Vegas, GA, USA, 19–21 June 1978; p. 297.
19. Tay, F.E.H.; Roy, A. CyberCAD: A Collaborative Approach in 3D-CAD Technology in a Multimedia-Supported Environment. *Comput. Ind.* **2003**, *52*, 127–145. [CrossRef]
20. Tornincasa, S.; Di Monaco, F. The Future and the Evolution of CAD. In Proceedings of the 14th International Research/Expert Conference: Trends in the Development of Machinery and Associated Technology, Mediterranean Cruise, 11–18 September 2010; Volume 1, pp. 11–18.
21. Miyazaki, T.; Hotta, Y.; Kunii, J.; Kuriyama, S.; Tamaki, Y. A Review of Dental CAD/CAM: Current Status and Future Perspectives from 20 Years of Experience. *Dent. Mater. J.* **2009**, *28*, 44–56. [CrossRef]
22. Matta, A.K.; Raju, D.R.; Suman, K.N.S. The Integration of CAD/CAM and Rapid Prototyping in Product Development: A Review. *Mater. Today Proc.* **2015**, *2*, 3438–3445. [CrossRef]
23. Lichten, L. The Emerging Technology of CAD/CAM. In Proceedings of the 1984 Annual Conference of the ACM on The Fifth Generation Challenge, ACM '84, San Francisco, CA, USA, 8–14 October 1984; Association for Computing Machinery: New York, NY, USA, 1984; pp. 236–241. [CrossRef]
24. Kou, X.Y.; Xue, S.K.; Tan, S.T. Knowledge-Guided Inference for Voice-Enabled CAD. *Comput. Des.* **2010**, *42*, 545–557. [CrossRef]
25. Thakur, A.; Rai, R. User Study of Hand Gestures for Gesture Based 3D CAD Modeling. In Proceedings of the ASME Design Engineering Technical Conference, Boston, MA, USA, 2–5 August 2015; Volume 1B-2015, pp. 1–14. [CrossRef]
26. Baig, M.Z.; Kavakli, M. Analyzing Novice and Expert User's Cognitive Load in Using a Multi-Modal Interface System. In Proceedings of the 2018 26th International Conference on Systems Engineering (ICSEng), Sydney, Australia, 18–20 December 2019; pp. 1–7. [CrossRef]
27. Esquivel, J.C.R.; Viveros, A.M.; Perry, N. Gestures for Interaction between the Software CATIA and the Human via Microsoft Kinect. In *International Conference on Human-Computer Interaction*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 457–462.
28. Sree Shankar, S.; Rai, R. Human Factors Study on the Usage of BCI Headset for 3D CAD Modeling. *Comput. Des.* **2014**, *54*, 51–55. [CrossRef]
29. Turk, M. Multimodal Interaction: A Review. *Pattern Recognit. Lett.* **2014**, *36*, 189–195. [CrossRef]
30. Bhat, R.; Deshpande, A.; Rai, R.; Esfahani, E.T. BCI-Touch Based System: A Multimodal CAD Interface for Object Manipulation. In *Volume 12: Systems and Design, Proceedings of the ASME 2013 International Mechanical Engineering Congress and Exposition, San Diego, CA, USA, 15–21 November 2013*; American Society of Mechanical Engineers: New York, NY, USA, 2013; p. V012T13A015. [CrossRef]
31. Oviatt, S. User-Centered Modeling and Evaluation of Multimodal Interfaces. *Proc. IEEE* **2003**, *91*, 1457–1468. [CrossRef]
32. Lee, H.; Lim, S.Y.; Lee, I.; Cha, J.; Cho, D.-C.; Cho, S. Multi-Modal User Interaction Method Based on Gaze Tracking and Gesture Recognition. *Signal Process. Image Commun.* **2013**, *28*, 114–126. [CrossRef]
33. Jaímes, A.; Sebe, N. Multimodal Human Computer Interaction: A Survey. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Figure 1; Springer: Berlin/Heidelberg, Germany, 2005; Volume 3766, pp. 1–15. [CrossRef]
34. Song, J.; Cho, S.; Baek, S.Y.; Lee, K.; Bang, H. GaFinC: Gaze and Finger Control Interface for 3D Model Manipulation in CAD Application. *CAD Comput. Aided Des.* **2014**, *46*, 239–245. [CrossRef]
35. Toma, M.I.; Gîrbacia, F.; Antonya, C. A Comparative Evaluation of Human Interaction for Design and Assembly of 3D CAD Models in Desktop and Immersive Environments. *Int. J. Interact. Des. Manuf.* **2012**, *6*, 179–193. [CrossRef]
36. Lawson, B.; Loke, S.M. Computers, Words and Pictures. *Des. Stud.* **1997**, *18*, 171–183. [CrossRef]
37. Goel, A.K.; Vattam, S.; Wiltgen, B.; Helms, M. Cognitive, Collaborative, Conceptual and Creative—Four Characteristics of the next Generation of Knowledge-Based CAD Systems: A Study in Biologically Inspired Design. *CAD Comput. Aided Des.* **2012**, *44*, 879–900. [CrossRef]

38. Robertson, B.F.; Radcliffe, D.F. Impact of CAD Tools on Creative Problem Solving in Engineering Design. *Comput. Des.* **2009**, *41*, 136–146. [CrossRef]
39. Jowers, I.; Prats, M.; McKay, A.; Garner, S. Evaluating an Eye Tracking Interface for a Two-Dimensional Sketch Editor. *Comput. Des.* **2013**, *45*, 923–936. [CrossRef]
40. Yoon, S.M.; Graf, H. Eye Tracking Based Interaction with 3d Reconstructed Objects. In Proceeding of the 16th ACM International Conference on Multimedia—MM '08, Vancouver, BC, Canada, 27–31 October 2008; ACM Press: New York, NY, USA, 2008; p. 841. [CrossRef]
41. Argelaguet, F.; Andujar, C. A Survey of 3D Object Selection Techniques for Virtual Environments. *Comput. Graph.* **2013**, *37*, 121–136. [CrossRef]
42. Dave, D.; Chowriappa, A.; Kesavadas, T. Gesture Interface for 3d Cad Modeling Using Kinect. *Comput. Aided. Des. Appl.* **2013**, *10*, 663–669. [CrossRef]
43. Zhong, K.; Kang, J.; Qin, S.; Wang, H. Rapid 3D Conceptual Design Based on Hand Gesture. In Proceedings of the 2011 3rd International Conference on Advanced Computer Control, Harbin, China, 18–20 January 2011; pp. 192–197. [CrossRef]
44. Xiao, Y.; Peng, Q. A Hand Gesture-Based Interface for Design Review Using Leap Motion Controller. In Proceedings of the 21st International Conference on Engineering Design (ICED 17), Vancouver, BC, Canada, 21–25 August 2017; Volume 8, pp. 239–248.
45. Vinayak; Murugappan, S.; Liu, H.; Ramani, K. Shape-It-Up: Hand Gesture Based Creative Expression of 3D Shapes Using Intelligent Generalized Cylinders. *CAD Comput. Aided Des.* **2013**, *45*, 277–287. [CrossRef]
46. Vinayak; Ramani, K. A Gesture-Free Geometric Approach for Mid-Air Expression of Design Intent in 3D Virtual Pottery. *CAD Comput. Aided Des.* **2015**, *69*, 11–24. [CrossRef]
47. Huang, J.; Rai, R. Hand Gesture Based Intuitive CAD Interface. In Proceedings of the ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Buffalo, NY, USA, 17–20 August 2014. [CrossRef]
48. Kou, X.Y.; Liu, X.C.; Tan, S.T. Quadtree Based Mouse Trajectory Analysis for Efficacy Evaluation of Voice-Enabled CAD. In Proceedings of the 2009 IEEE International Conference on Virtual Environments, Human-Computer Interfaces and Measurements Systems, Hong Kong, China, 1–13 May 2009; pp. 196–201. [CrossRef]
49. Samad, T.; Director, S.W. Towards a Natural Language Interface for CAD. In Proceedings of the 22nd ACM/IEEE Design Automation Conference, Las Vegas, NV, USA, 23–26 June 1985; pp. 2–8. [CrossRef]
50. Salisbury, M.W.; Hendrickson, J.H.; Lammers, T.L.; Fu, C.; Moody, S.A. Talk and Draw: Bundling Speech and Graphics. *Computer* **1990**, *23*, 59–65. [CrossRef]
51. Kou, X.Y.; Tan, S.T. Design by Talking with Computers. *Comput. Aided. Des. Appl.* **2008**, *5*, 266–277. [CrossRef]
52. Menegotto, J.L. A Framework for Speech-Oriented CAD and BIM Systems. In *Communications in Computer and Information Science*; Springer: Berlin/Heidelberg, Germany, 2015; Volume 527, pp. 329–347. [CrossRef]
53. Behera, A.K.; McKay, A. Designs That Talk and Listen: Integrating Functional Information Using Voice-Enabled CAD Systems. In Proceedings of the 25th European Signal Processing Conference, Kos, Greek, 28 August–2 September 2017.
54. Xue, S.; Kou, X.Y.; Tan, S.T. Natural Voice-Enabled CAD: Modeling via Natural Discourse. *Comput. Aided. Des. Appl.* **2009**, *6*, 125–136. [CrossRef]
55. Plumed, R.; González-Lluch, C.; Pérez-López, D.; Contero, M.; Camba, J.D. A Voice-Based Annotation System for Collaborative Computer-Aided Design. *J. Comput. Des. Eng.* **2021**, *8*, 536–546. [CrossRef]
56. Trejo, L.J.; Rosipal, R.; Matthews, B. Brain-Computer Interfaces for 1-D and 2-D Cursor Control: Designs Using Volitional Control of the EEG Spectrum or Steady-State Visual Evoked Potentials. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2006**, *14*, 225–229. [CrossRef]
57. Esfahani, E.T.; Sundararajan, V. Using Brain Computer Interfaces for Geometry Selection in CAD Systems: P300 Detection Approach. In Proceedings of the 31st Computers and Information in Engineering Conference, Parts A and B, Washington, DC, USA, 28–31 August 2011; Volume 2, pp. 1575–1580. [CrossRef]
58. Esfahani, E.T.; Sundararajan, V. Classification of Primitive Shapes Using Brain-Computer Interfaces. *CAD Comput. Aided Des.* **2012**, *44*, 1011–1019. [CrossRef]
59. Huang, Y.C.; Chen, K.L.; Wu, M.Y.; Tu, Y.W.; Huang, S.C.C. Brain-Computer Interface Approach to Computer-Aided Design: Rotate and Zoom in/out in 3ds Max via Imagination. In Proceedings of the International Conferences Interfaces and Human Computer Interaction 2015 (IHCI 2015), Los Angeles, CA, USA, 2–7 August 2015; Game and Entertainment Technologies 2015 (GET 2015), Las Palmas de Gran Canaria, Spain, 22–24 July 2015, Computer Graphics, Visualization, Computer Vision and Image Processing 2015 (CGVCVIP 2015), Las Palmas de Gran Canaria, Spain, 22–24 July 2015. pp. 319–322.
60. Postelnicu, C.; Duguleana, M.; Garbacia, F.; Talaba, D. Towards P300 Based Brain Computer Interface for Computer Aided Design. In Proceedings of the 11th EuroVR 2014, Bremen, Germany, 8–10 December 2014; pp. 2–6. [CrossRef]
61. Verma, A.; Rai, R. Creating by Imagining: Use of Natural and Intuitive BCI in 3D CAD Modeling. In Proceedings of the 33rd Computers and Information in Engineering Conference, Portland, OR, USA, 4–7 August 2013; American Society of Mechanical Engineers: New York, NY, USA, 2013; Volume 2A. [CrossRef]
62. Moustakas, K.; Tzovaras, D. MASTER-PIECE: A Multimodal (Gesture + Speech) Interface for 3D Model Search and Retrieval Integrated in a Virtual Assembly Application. In Proceedings of the eINTERFACE'05-Summer Workshop on Multimodal Interfaces, Mons, Belgium, 12–18 August 2005; pp. 1–14.

63. Bolt, R.A. Put-That-There. In Proceedings of the 7th Annual Conference on Computer Graphics and Interactive Techniques—SIGGRAPH '80, Seattle, WA, USA, 14–18 July 1980; ACM Press: New York, NY, USA, 1980; pp. 262–270. [CrossRef]
64. Khan, S.; Rajapakse, H.; Zhang, H.; Nanayakkara, S.; Tuncer, B.; Blessing, L. GesCAD: An Intuitive Interface for Conceptual Architectural Design. In Proceedings of the 29th Australian Conference on Computer-Human Interaction, Brisbane, Australia, 28 November–1 December 2017; ACM: New York, NY, USA, 2017; pp. 402–406. [CrossRef]
65. Hauptmann, A.G. Speech and Gestures for Graphic Image Manipulation. *ACM SIGCHI Bull.* **1989**, *20*, 241–245. [CrossRef]
66. Arangarasan, R.; Gadh, R. Geometric Modeling and Collaborative Design in a Multi-Modal Multi-Sensory Virtual Environment. In Proceedings of the ASME 2000 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Baltimore, MD, USA, 10–13 September 2000; pp. 10–13.
67. Chu, C.-C.P.; Dani, T.H.; Gadh, R. Multi-Sensory User Interface for a Virtual-Reality-Based Computeraided Design System. *Comput. Des.* **1997**, *29*, 709–725. [CrossRef]
68. Chu, C.-C.P.; Dani, T.H.; Gadh, R. Multimodal Interface for a Virtual Reality Based Computer Aided Design System. In Proceedings of the International Conference on Robotics and Automation, Albuquerque, NM, USA, 20–25 April 1997; Volume 2, pp. 1329–1334. [CrossRef]
69. Pouke, M.; Karhu, A.; Hickey, S.; Arhippainen, L. Gaze Tracking and Non-Touch Gesture Based Interaction Method for Mobile 3D Virtual Spaces. In Proceedings of the 24th Australian Computer-Human Interaction Conference on OzCHI '12, Melbourne, Australia, 26–30 November 2012; ACM Press: New York, NY, USA, 2012; pp. 505–512. [CrossRef]
70. Shafiei, S.B.; Esfahani, E.T. Aligning Brain Activity and Sketch in Multi-Modal CAD Interface. In Proceedings of the ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Buffalo, NY, USA, 17–20 August 2014; Volume 1A, pp. 1–7. [CrossRef]
71. Sharma, A.; Madhvanath, S. MozArt: A Multimodal Interface for Conceptual 3D Modeling. In Proceedings of the 13th International Conference on Multimodal Interfaces, Alicante, Spain, 14–18 November 2011; pp. 3–6. [CrossRef]
72. Bourdot, P.; Convard, T.; Picon, F.; Ammi, M.; Touraine, D.; Vézien, J.-M. VR–CAD Integration: Multimodal Immersive Interaction and Advanced Haptic Paradigms for Implicit Edition of CAD Models. *Comput. Des.* **2010**, *42*, 445–461. [CrossRef]
73. Mogan, G.; Talaba, D.; Girbacia, F.; Butnaru, T.; Sisca, S.; Aron, C. A Generic Multimodal Interface for Design and Manufacturing Applications. In Proceedings of the 2nd International Workshop Virtual Manufacturing (VirMan08), Torino, Italy, 6–8 October 2008.
74. Stark, R.; Israel, J.H.; Wöhler, T. Towards Hybrid Modelling Environments—Merging Desktop-CAD and Virtual Reality-Technologies. *CIRP Ann.-Manuf. Technol.* **2010**, *59*, 179–182. [CrossRef]
75. Ren, X.; Zhang, G.; Dai, G. An Experimental Study of Input Modes for Multimodal Human-Computer Interaction. In Proceedings of the International Conference on Multimodal Interfaces, Beijing, China, 14–16 October 2000; Volume 1, pp. 49–56. [CrossRef]
76. Jaimes, A.; Sebe, N. Multimodal Human–Computer Interaction: A Survey. *Comput. Vis. Image Underst.* **2007**, *108*, 116–134. [CrossRef]
77. Hutchinson, T.E. Eye-Gaze Computer Interfaces: Computers That Sense Eye Position on the Display. *Computer* **1993**, *26*, 65–66. [CrossRef]
78. Sharma, C.; Dubey, S.K. Analysis of Eye Tracking Techniques in Usability and HCI Perspective. In Proceedings of the 2014 International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 5–7 March 2014; pp. 607–612. [CrossRef]
79. Brooks, R.; Meltzoff, A.N. The Development of Gaze Following and Its Relation to Language. *Dev. Sci.* **2005**, *8*, 535–543. [CrossRef]
80. Duchowski, A.T. A Breadth-First Survey of Eye-Tracking Applications. *Behav. Res. Methods Instrum. Comput.* **2002**, *34*, 455–470. [CrossRef]
81. Kumar, M.; Paepcke, A.; Winograd, T. Eyepoint: Practical Pointing and Selection Using Gaze and Keyboard. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 28 April–3 May 2007; pp. 421–430.
82. Majaranta, P.; Räihä, K.-J. Twenty Years of Eye Typing: Systems and Design Issues. In Proceedings of the 2002 Symposium on Eye Tracking Research & Applications, New Orleans, LA, USA, 25–27 March 2002; pp. 15–22.
83. Wobbrock, J.O.; Rubinstein, J.; Sawyer, M.W.; Duchowski, A.T. Longitudinal Evaluation of Discrete Consecutive Gaze Gestures for Text Entry. In Proceedings of the 2008 Symposium on Eye Tracking Research & Applications, Savannah, GA, USA, 26–28 March 2008; pp. 11–18.
84. Hornof, A.J.; Cavender, A. EyeDraw: Enabling Children with Severe Motor Impairments to Draw with Their Eyes. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Portland, OR, USA, 2–7 April 2005; pp. 161–170.
85. Wu, C.-I. HCI and Eye Tracking Technology for Learning Effect. *Procedia-Soc. Behav. Sci.* **2012**, *64*, 626–632. [CrossRef]
86. Hekele, F.; Spilski, J.; Bender, S.; Lachmann, T. Remote Vocational Learning Opportunities—A Comparative Eye-Tracking Investigation of Educational 2D Videos versus 360° Videos for Car Mechanics. *Br. J. Educ. Technol.* **2022**, *53*, 248–268. [CrossRef]
87. Dickie, C.; Vertegaal, R.; Sohn, C.; Cheng, D. EyeLook: Using Attention to Facilitate Mobile Media Consumption. In Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology, Seattle, WA, USA, 23–26 October 2005; pp. 103–106.
88. Nagamatsu, T.; Yamamoto, M.; Sato, H. MobiGaze: Development of a Gaze Interface for Handheld Mobile Devices. In Proceedings of the Human Factors in Computing Systems, New York, NY, USA, 10–15 April 2010; pp. 3349–3354. [CrossRef]

89. Miluzzo, E.; Wang, T.; Campbell, A.T. Eyephone: Activating Mobile Phones with Your Eyes. In Proceedings of the Second ACM SIGCOMM Workshop on Networking, Systems, and Applications on Mobile Handhelds, New Delhi, India, 30 August 2010; pp. 15–20.
90. Franslin, N.M.F.; Ng, G.W. Vision-Based Dynamic Hand Gesture Recognition Techniques and Applications: A Review. In Proceedings of the 8th International Conference on Computational Science and Technology, Labuan, Malaysia, 28–29 August 2021; Springer: Berlin/Heidelberg, Germany, 2022; pp. 125–138.
91. Tumkor, S.; Esche, S.K.; Chassapis, C. Hand Gestures in CAD Systems. In Proceedings of the Asme International Mechanical Engineering Congress and Exposition, San Diego, CA, USA, 15–21 November 2013; American Society of Mechanical Engineers: New York, NY, USA, 2013; Volume 56413, p. V012T13A008.
92. Florin, G.; Butnariu, S. Design Review of Cad Models Using a NUI Leap Motion Sensor. *J. Ind. Des. Eng. Graph.* **2015**, *10*, 21–24.
93. Kaur, H.; Rani, J. A Review: Study of Various Techniques of Hand Gesture Recognition. In Proceedings of the 1st IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems, ICPEICES 2016, Delhi, India, 4–6 July 2016; IEEE: Piscataway, NJ, USA, 2017; pp. 1–5. [CrossRef]
94. Fuge, M.; Yumer, M.E.; Orbay, G.; Kara, L.B. Conceptual Design and Modification of Freeform Surfaces Using Dual Shape Representations in Augmented Reality Environments. *CAD Comput. Aided Des.* **2012**, *44*, 1020–1032. [CrossRef]
95. Pareek, S.; Sharma, V.; Esfahani, E.T. Human Factor Study in Gesture Based Cad Environment. In Proceedings of the International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Boston, MA, USA, 2–5 August 2015; Volume 1B-2015, pp. 1–6. [CrossRef]
96. Huang, J.; Jaiswal, P.; Rai, R. Gesture-Based System for next Generation Natural and Intuitive Interfaces. In *Artificial Intelligence for Engineering Design, Analysis and Manufacturing-AIEDAM*; Cambridge University Press: Cambridge, UK, 2019; Volume 33, pp. 54–68. [CrossRef]
97. Clark, L.; Doyle, P.; Garaialde, D.; Gilmartin, E.; Schlögl, S.; Edlund, J.; Aylett, M.; Cabral, J.; Munteanu, C.; Edwards, J.; et al. The State of Speech in HCI: Trends, Themes and Challenges. *Interact. Comput.* **2019**, *31*, 349–371. [CrossRef]
98. Gao, S.; Wan, H.; Peng, Q. Approach to Solid Modeling in a Semi-Immersive Virtual Environment. *Comput. Graph.* **2000**, *24*, 191–202. [CrossRef]
99. Rezeika, A.; Benda, M.; Stawicki, P.; Gembler, F.; Saboor, A.; Volosyak, I. Brain–Computer Interface Spellers: A Review. *Brain Sci.* **2018**, *8*, 57. [CrossRef]
100. Amiri, S.; Fazel-Rezai, R.; Asadpour, V. A Review of Hybrid Brain-Computer Interface Systems. *Adv. Hum.-Comput. Interact.* **2013**, *2013*, 187024. [CrossRef]
101. Pfurtscheller, G.; Leeb, R.; Keinrath, C.; Friedman, D.; Neuper, C.; Guger, C.; Slater, M. Walking from Thought. *Brain Res.* **2006**, *1071*, 145–152. [CrossRef]
102. Lécuyer, A.; Lotte, F.; Reilly, R.B.; Leeb, R.; Hirose, M.; Slater, M. Brain-Computer Interfaces, Virtual Reality, and Videogames. *Computer* **2008**, *41*, 66–72. [CrossRef]
103. Zhao, Q.; Zhang, L.; Cichocki, A. EEG-Based Asynchronous BCI Control of a Car in 3D Virtual Reality Environments. *Chin. Sci. Bull.* **2009**, *54*, 78–87. [CrossRef]
104. Li, Y.; Wang, C.; Zhang, H.; Guan, C. An EEG-Based BCI System for 2D Cursor Control. In Proceedings of the 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), Hong Kong, China, 1–6 June 2008; pp. 2214–2219.
105. Chun, J.; Bae, B.; Jo, S. BCI Based Hybrid Interface for 3D Object Control in Virtual Reality. In Proceedings of the 4th International Winter Conference on Brain-Computer Interface—BCI 2016, Gangwon, Korea, 22–24 February 2016; pp. 19–22. [CrossRef]
106. Esfahani, E.T.; Sundararajan, V. Using Brain–Computer Interfaces to Detect Human Satisfaction in Human–Robot Interaction. *Int. J. Hum. Robot.* **2011**, *8*, 87–101. [CrossRef]
107. Sebe, N. Multimodal Interfaces: Challenges and Perspectives. *J. Ambient Intell. Smart Environ.* **2009**, *1*, 23–30. [CrossRef]
108. Dumas, B.; Lalanne, D.; Oviatt, S. *Multimodal Interfaces: A Survey of Principles, Models and Frameworks*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 3–26. [CrossRef]
109. Ye, J.; Campbell, R.I.; Page, T.; Badni, K.S. An Investigation into the Implementation of Virtual Reality Technologies in Support of Conceptual Design. *Des. Stud.* **2006**, *27*, 77–97. [CrossRef]
110. Liu, X.; Dodds, G.; McCartney, J.; Hinds, B.K. Manipulation of CAD Surface Models with Haptics Based on Shape Control Functions. *CAD Comput. Aided Des.* **2005**, *37*, 1447–1458. [CrossRef]
111. Zhu, W. A Methodology for Building up an Infrastructure of Haptically Enhanced Computer-Aided Design Systems. *J. Comput. Inf. Sci. Eng.* **2008**, *8*, 0410041–04100411. [CrossRef]
112. Picon, F.; Ammi, M.; Bourdot, P. Case Study of Haptic Methods for Selection on CAD Models. In Proceedings of the 2008 IEEE Virtual Reality Conference, Reno, NV, USA, 8–12 March 2008; pp. 209–212. [CrossRef]
113. Chamaret, D.; Ullah, S.; Richard, P.; Naud, M. Integration and Evaluation of Haptic Feedbacks: From CAD Models to Virtual Prototyping. *Int. J. Interact. Des. Manuf.* **2010**, *4*, 87–94. [CrossRef]
114. Kind, S.; Geiger, A.; Kießling, N.; Schmitz, M.; Stark, R. Haptic Interaction in Virtual Reality Environments for Manual Assembly Validation. *Procedia CIRP* **2020**, *91*, 802–807. [CrossRef]
115. Wechsung, I.; Engelbrecht, K.-P.; Kühnel, C.; Möller, S.; Weiss, B. Measuring the Quality of Service and Quality of Experience of Multimodal Human–Machine Interaction. *J. Multimodal User Interfaces* **2012**, *6*, 73–85. [CrossRef]

116. Zeeshan Baig, M.; Kavakli, M. A Survey on Psycho-Physiological Analysis & Measurement Methods in Multimodal Systems. *Multimodal Technol. Interact.* **2019**, *3*, 37. [CrossRef]
117. Chu, C.C.; Mo, J.; Gadh, R. A Quantitative Analysis on Virtual Reality-Based Computer Aided Design System Interfaces. *J. Comput. Inf. Sci. Eng.* **2002**, *2*, 216–223. [CrossRef]
118. Gîrbacia, F. Evaluation of CAD Model Manipulation in Desktop and Multimodal Immersive Interface. *Appl. Mech. Mater.* **2013**, *327*, 289–293. [CrossRef]
119. Feeman, S.M.; Wright, L.B.; Salmon, J.L. Exploration and Evaluation of CAD Modeling in Virtual Reality. *Comput. Aided. Des. Appl.* **2018**, *15*, 892–904. [CrossRef]
120. Community, I.; Davide, F. 5 Perception and Cognition in Immersive Virtual Reality. 2003, pp. 71–86. Available online: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.433.9220&rep=rep1&type=pdf (accessed on 10 May 2020).
121. Wickens, C.D. Multiple Resources and Mental Workload. *Hum. Factors* **2008**, *50*, 449–455. [CrossRef]
122. Zhang, L.; Wade, J.; Bian, D.; Fan, J.; Swanson, A.; Weitlauf, A.; Warren, Z.; Sarkar, N. Multimodal Fusion for Cognitive Load Measurement in an Adaptive Virtual Reality Driving Task for Autism Intervention. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2015; Volume 9177, pp. 709–720. [CrossRef]
123. Wickens, C.D.; Gordon, S.E.; Liu, Y.; Lee, J. *An Introduction to Human Factors Engineering*; Pearson Prentice Hall: Upper Saddle River, NJ, USA, 2004; Volume 2.
124. Goodman, E.; Kuniavsky, M.; Moed, A. Observing the User Experience: A Practitioner's Guide to User Research. *IEEE Trans. Prof. Commun.* **2013**, *56*, 260–261. [CrossRef]
125. Cutugno, F.; Leano, V.A.; Rinaldi, R.; Mignini, G. Multimodal Framework for Mobile Interaction. In Proceedings of the International Working Conference on Advanced Visual Interfaces, Capri Island, Italy, 21–25 May 2012; pp. 197–203.
126. Ergoneers Dikablis Glasses 3 Eye Tracker. Available online: https://www.jalimedical.com/dikablis-glasses-eye-tracker.php (accessed on 10 May 2020).
127. Tobii Pro Glasses 2—Discontinued. Available online: https://www.tobiipro.com/product-listing/tobii-pro-glasses-2/ (accessed on 10 May 2020).
128. Pupil Labs Core Hardware Specifications. Available online: https://imotions.com/hardware/pupil-labs-glasses/ (accessed on 10 May 2020).
129. GP3 Eye Tracker. Available online: https://www.gazept.com/product/gazepoint-gp3-eye-tracker/ (accessed on 10 May 2020).
130. Tobii Pro Spectrum. Available online: https://www.tobiipro.com/product-listing/tobii-pro-spectrum/ (accessed on 10 May 2020).
131. EyeLink 1000 Plus. Available online: https://www.sr-research.com/eyelink-1000-plus/ (accessed on 10 May 2020).
132. MoCap Pro Gloves. Available online: https://stretchsense.com/solution/gloves/ (accessed on 10 May 2020).
133. CyberGlove Systems. Available online: http://www.cyberglovesystems.com/ (accessed on 10 May 2020).
134. VRTRIX. Available online: http://www.vrtrix.com/ (accessed on 10 May 2020).
135. Kinect from Wikipedia. Available online: https://en.wikipedia.org/wiki/Kinect (accessed on 10 May 2020).
136. Leap Motion Controller. Available online: https://www.ultraleap.com/product/leap-motion-controller/ (accessed on 10 May 2020).
137. Satori, H.; Harti, M.; Chenfour, N. Introduction to Arabic Speech Recognition Using CMUSphinx System. *arXiv* **2007**, arXiv0704.2083.
138. NeuroSky. Available online: https://store.neurosky.com/ (accessed on 10 May 2020).
139. Muse (Headband) from Wikipedia. Available online: https://en.wikipedia.org/wiki/Muse_(headband) (accessed on 10 May 2020).
140. VIVE Pro 2. Available online: https://www.vive.com/us/product/vive-pro2-full-kit/overview/ (accessed on 10 May 2020).
141. Meta Quest. Available online: https://www.oculus.com/rift-s/ (accessed on 10 May 2020).
142. Microsoft HoloLens 2. Available online: https://www.microsoft.com/en-us/hololens (accessed on 10 May 2020).
143. Sigut, J.; Sidha, S.-A. Iris Center Corneal Reflection Method for Gaze Tracking Using Visible Light. *IEEE Trans. Biomed. Eng.* **2010**, *58*, 411–419. [CrossRef]
144. Chennamma, H.R.; Yuan, X. A Survey on Eye-Gaze Tracking Techniques. *Indian J. Comput. Sci. Eng.* **2013**, *4*, 388–393.
145. Ma, C.; Baek, S.-J.; Choi, K.-A.; Ko, S.-J. Improved Remote Gaze Estimation Using Corneal Reflection-Adaptive Geometric Transforms. *Opt. Eng.* **2014**, *53*, 53112. [CrossRef]
146. Dodge, R.; Cline, T.S. The Angle Velocity of Eye Movements. *Psychol. Rev.* **1901**, *8*, 145. [CrossRef]
147. Stuart, S.; Hickey, A.; Vitorio, R.; Welman, K.; Foo, S.; Keen, D.; Godfrey, A. Eye-Tracker Algorithms to Detect Saccades during Static and Dynamic Tasks: A Structured Review. *Physiol. Meas.* **2019**, *40*, 02TR01. [CrossRef] [PubMed]
148. Chen, C.H.; Monroy, C.; Houston, D.M.; Yu, C. *Using Head-Mounted Eye-Trackers to Study Sensory-Motor Dynamics of Coordinated Attention*, 1st ed.; Elsevier: Amsterdam, The Netherlands, 2020; Volume 254. [CrossRef]
149. Lake, S.; Bailey, M.; Grant, A. Method and Apparatus for Analyzing Capacitive Emg and Imu Sensor Signals for Gesture Control. U.S. Patents 9,299,248, 29 March 2016.
150. Wang, Z.; Cao, J.; Liu, J.; Zhao, Z. Design of Human-Computer Interaction Control System Based on Hand-Gesture Recognition. In Proceedings of the 2017 32nd Youth Academic Annual Conference of Chinese Association of Automation (YAC), Hefei, China, 19–21 May 2017; pp. 143–147.
151. Grosshauser, T. Low Force Pressure Measurement: Pressure Sensor Matrices for Gesture Analysis, Stiffness Recognition and Augmented Instruments. In Proceedings of the NIME 2008, Genova, Italy, 5–7 June 2008; pp. 97–102.

152. Sawada, H.; Hashimoto, S. Gesture Recognition Using an Acceleration Sensor and Its Application to Musical Performance Control. *Electron. Commun. Jpn.* **1997**, *80*, 9–17. [CrossRef]
153. Suarez, J.; Murphy, R.R. Hand Gesture Recognition with Depth Images: A Review. In Proceedings of the 2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication, Paris, France, 9–13 September 2012; Volume 2012, pp. 411–417. [CrossRef]
154. Abu Shariah, M.A.M.; Ainon, R.N.; Zainuddin, R.; Khalifa, O.O. Human Computer Interaction Using Isolated-Words Speech Recognition Technology. In Proceedings of the ICIAS 2007: International Conference on Intelligent and Advanced Systems, Kuala Lumpur, Malaysia, 25–28 November 2007; pp. 1173–1178. [CrossRef]
155. Zhang, J.; Zhang, M. A Speech Recognition System Based Improved Algorithm of Dual-Template HMM. *Procedia Eng.* **2011**, *15*, 2286–2290. [CrossRef]
156. Chang, S.-Y.; Morgan, N. Robust CNN-Based Speech Recognition with Gabor Filter Kernels. In Proceedings of the Fifteenth Annual Conference of the International Speech Communication Association, Singapore, 14–18 September 2014.
157. Soltau, H.; Liao, H.; Sak, H. Neural Speech Recognizer: Acoustic-to-Word LSTM Model for Large Vocabulary Speech Recognition. *arXiv* **2016**, arXiv1610.09975.
158. Tyagi, A. A Review of Eeg Sensors Used for Data Acquisition. *Electron. Instrum.* **2012**, 13–18. Available online: https://www.researchgate.net/profile/Sunil-Semwal/publication/308259085_A_Review_of_Eeg_Sensors_used_for_Data_Acquisition/links/57df321408ae72d72eac238e/A-Review-of-Eeg-Sensors-used-for-Data-Acquisition.pdf (accessed on 10 May 2020).
159. Pinegger, A.; Wriessnegger, S.C.; Faller, J.; Müller-Putz, G.R. Evaluation of Different EEG Acquisition Systems Concerning Their Suitability for Building a Brain–Computer Interface: Case Studies. *Front. Neurosci.* **2016**, *10*, 441. [CrossRef]
160. Zerafa, R.; Camilleri, T.; Falzon, O.; Camilleri, K.P. A Comparison of a Broad Range of EEG Acquisition Devices–Is There Any Difference for SSVEP BCIs? *Brain-Comput. Interfaces* **2018**, *5*, 121–131. [CrossRef]
161. Liu, Y.; Jiang, X.; Cao, T.; Wan, F.; Mak, P.U.; Mak, P.-I.; Vai, M.I. Implementation of SSVEP Based BCI with Emotiv EPOC. In Proceedings of the 2012 IEEE International Conference on Virtual Environments Human-Computer Interfaces and Measurement Systems (VECIMS), Tianjin, China, 2–4 July 2012; pp. 34–37.
162. Guger, C.; Krausz, G.; Allison, B.Z.; Edlinger, G. Comparison of Dry and Gel Based Electrodes for P300 Brain–Computer Interfaces. *Front. Neurosci.* **2012**, *6*, 60. [CrossRef]
163. Hirose, S. A VR Three-Dimensional Pottery Design System Using PHANTOM Haptic Devices. In Proceedings of the 4th PHANToM Users Group Workshop, Dedham, MA, USA, 9–12 October 1999.
164. Nikolakis, G.; Tzovaras, D.; Moustakidis, S.; Strintzis, M.G. CyberGrasp and PHANTOM Integration: Enhanced Haptic Access for Visually Impaired Users. In Proceedings of the SPECOM'2004 9th Conference Speech and Computer, Saint-Petersburg, Russia, 20–22 September 2004; pp. x1–x7.
165. Wee, C.; Yap, K.M.; Lim, W.N. Haptic Interfaces for Virtual Reality: Challenges and Research Directions. *IEEE Access* **2021**, *9*, 112145–112162. [CrossRef]
166. Kumari, P.; Mathew, L.; Syal, P. Increasing Trend of Wearables and Multimodal Interface for Human Activity Monitoring: A Review. *Biosens. Bioelectron.* **2017**, *90*, 298–307. [CrossRef]
167. Oviatt, S.; Cohen, P.; Wu, L.; Duncan, L.; Suhm, B.; Bers, J.; Holzman, T.; Winograd, T.; Landay, J.; Larson, J.; et al. Designing the User Interface for Multimodal Speech and Pen-Based Gesture Applications: State-of-the-Art Systems and Future Research Directions. *Hum.-Comput. Interact.* **2000**, *15*, 263–322. [CrossRef]
168. Sharma, R.; Pavlovic, V.I.; Huang, T.S. Toward Multimodal Human-Computer Interface. *Proc. IEEE* **1998**, *86*, 853–869. [CrossRef]
169. Sunar, A.W.I.M.S. Multimodal Fusion- Gesture and Speech Input in Augmented Reality Environment. *Adv. Intell. Syst. Comput.* **2015**, *331*, 255–264. [CrossRef]
170. Poh, N.; Kittler, J. Multimodal Information Fusion. In *Multimodal Signal Processing*; Elsevier: Amsterdam, The Netherlands, 2010; pp. 153–169. [CrossRef]
171. Lalanne, D.; Nigay, L.; Palanque, P.; Robinson, P.; Vanderdonckt, J.; Ladry, J.-F. Fusion Engines for Multimodal Input. In Proceedings of the 2009 International Conference on Multimodal Interfaces, Cambridge, MA, USA, 2–4 November 2009; p. 153. [CrossRef]
172. Zissis, D.; Lekkas, D.; Azariadis, P.; Papanikos, P.; Xidias, E. Collaborative CAD/CAE as a Cloud Service. *Int. J. Syst. Sci. Oper. Logist.* **2017**, *4*, 339–355. [CrossRef]
173. Nam, T.J.; Wright, D. The Development and Evaluation of Syco3D: A Real-Time Collaborative 3D CAD System. *Des. Stud.* **2001**, *22*, 557–582. [CrossRef]
174. Kim, M.J.; Maher, M.L. The Impact of Tangible User Interfaces on Spatial Cognition during Collaborative Design. *Des. Stud.* **2008**, *29*, 222–253. [CrossRef]
175. Shen, Y.; Ong, S.K.; Nee, A.Y.C. Augmented Reality for Collaborative Product Design and Development. *Des. Stud.* **2010**, *31*, 118–145. [CrossRef]