_____

A Persepective of Human – Computer Interaction: An Explaratory Study

[1]Ishrat Begum , [2]Dr. Meena Chaudhary

[1]Department of Computer Science Engineering, Mangalayatan University, Beswan, Aligarh, UP, India

[2]Mangalayatan University, Beswan, Aligarh, UP, India

Abstract

The rapid development of technology in the field of artificial intelligence has resulted in the widespread implementation of virtual persons in a variety of contexts, including as in the provision of virtual anchoring services, "virtual customer service, and online education". In the interaction between humans and computers, a digital humanoid anthropomorphic is employed. This artificial intelligence can interact with humans right away and enhance the user environment. Simultaneously, one of the application directions for virtual humans is multimodal human–computer interaction. The primary objective of this project is to carry out in-depth research on the interaction between humans and computers. In this particular study, a method of research known as qualitative research is used. The results of this research indicate that human–computer interaction is a technique that can be used in the creation of artificial humans in the near future. In contrast to information output that only goes in one direction, "digital humans require multimodal information such as natural language", human-like facial expressions, and "natural human-like hand gestures".

Keywords: Human; Computer; Interaction; User Experience; Digital Human.

Introduction

"The rapid development of technology in the field of artificial intelligence" has resulted in the widespread implementation of virtual persons in a variety of contexts, including as in the provision of virtual anchoring services, "virtual customer service, and online education". There is an anthropomorphic digital human that is used in human–computer interaction (Xu, 2019). This artificial intelligence can interact with humans right away and enhance the consumer's interaction. Multimodal communication between people and computers is one of the potential application paths for artificial beings at the same time. The system utilises deep learning models to create interactive components with innate characteristics, incorporating text-to-speech, voice recognition, dialogue systems, and simulated human video reproduction. There are several subcategories within virtual humanity video creation, the most well-known being 2D/3D facial the rebuilding process, talking-head another, bodily motions, and human movement in general. (Caraban et al., 2019).

_____
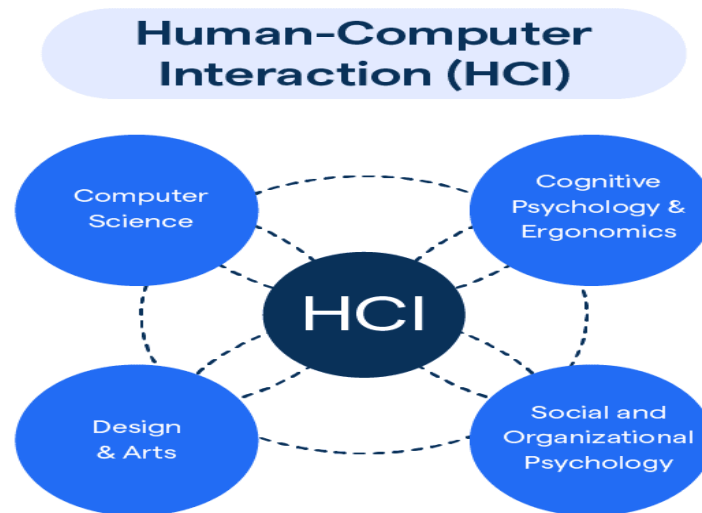
Figure 1: Key Principles of HCI [1]

In the field of study pertaining to the development of talking heads, "audio-driven lip synthesis is a common research direction". This research direction involves inputting the correct audio together with "any mesh vertex", facial picture, or video in order to generate a "lip-synced talking-head video". To put it another way, the model creates a dynamic mapping between the lower-dimensional text or speech input and the higher-dimensional video signal (Stephanidis et al., 2019). The following section elaborates the past literatures related to this field.

Literature Review

Table 1: Past literatures in Human – Computer Interactions

AUTHORS AND YEARS	METHODOLOGY	FIINDINGS
Ren and Bao (2020)	This study emphasised human interaction's ubiquitous technologies of hearing, speaking, reading, writing, and other senses.	The authors anticipated this study will assist field researchers gain the information and technologies to perform more sophisticated research.
Alnuaim et al., (2022)	The 1D CNN classification model beats classic machine learning methods. Our approach uses multiple language data sets, unlike most previous studies on emotions.	This method obtained 97.09%, 96.44%, and 83.33% accuracy for the BAVED, ANAD, and SAVEE "data sets using the most discriminating features and data augmentation".
Nicolescu & Tudorache (2022)	A sample of 40 empirical research is analysed using systematic literature review (SLR).	Empirical studies showed that response relevance and problem resolution are the most important factors in chatbot customer service, resulting in customer satisfaction,

_____

[1] https://botpenguin.com/glossary/human-computer-interaction

_____

		chatbot usage continuation, product purchases, and product recommendations.
Alnuaim et al., (2022)	Artificial intelligence is used to recognise human vocal emotions in this study.	The commonly used supervised learning algorithm multilayer perceptron (MLP) classifier was preferred for classification. The suggested model classified eight emotions with 81% accuracy on the RAVDESS dataset.
Zhen et al., (2023)	This study designed the "human–computer interaction system" framework for speech recognition, "text-to-speech, dialogue, and virtual person generation". Next, classify the talking-head video generating model using virtual human deep generation.	Methodically reviewed five years of talking-head video generating technology and trends, highlighting key works and summarising the dataset.

The interface between people and computers is crucial for enabling the different ways that people and computers communicate, according to earlier research. Today's extensively utilised graphical user interfaces (GUIs) can be found in computer kiosks, internet browsers, desktop applications, and mobile devices. Thus, this study's primary objective is to carry out an extensive investigation on many aspects of human-computer interaction.

Methodology

In this study, qualitative research methodology was utilized. Qualitative approaches are used to understand people's perspectives, circumstances, attitudes, behaviour, and exchanges. It generates "non-numerical data". The incorporation of qualitative techniques into study of intervention is becoming a more prominent area of attention in research across domains. The present study employed the Explanatory research method also, which investigates the reasons behind events when there is a scarcity of knowledge. It can aid in improving your comprehension of a certain subject, determining the cause or mechanism of a specific phenomenon, and making predictions about what will happen in the future.

RESULTS AND DISCUSSIONS

The phrase "Human-Computer Interaction/Interfacing", most referred to as "Man-Machine Interaction or Man-Machine Interfacing", was automatically portrayed with the development of the computer or, more generally speaking, the machine itself. HCI is an abbreviation for Human-Computer Interaction/Interfacing. Actually, the solution is self-evident: the vast majority of sophisticated technologies serve no purpose unless humans can operate them effectively. Both usability and functionality are offered as the primary concepts that ought to be taken into account when designing "human–computer interaction (HCI)". This argument is presented in a basic manner (Ren and Bao, 2022).

Based on artificial intelligence technology, the system aims to achieve high-fidelity, low-latency, multimodal interaction with anthropomorphic virtual humans using voice, image, and natural language processing. Figure 2 illustrates how the system is mostly made up of four modules: First, "the automatic speech recognition (ASR)" module of the system uses the voice information that the user inputs to convert it into text information. Next, "the dialogue system (DS)" receives the "text information that the ASR module produces as input". Third, "the text-to-speech (TTS)" module turns the text that the DS produces into "realistic speech information". Finally, the "talking-head generation module preprocesses the picture, video, or blendshape as the model input to extract its facial features". The model then uses the TTS module to map the lower-dimensional voice signal to the "higher-

_____

dimensional video signals", which include the "mouth, expression, motion", and so on. Ultimately, the model fuses the "features and multimodal" output video on the rendering system before displaying it on the user's side.
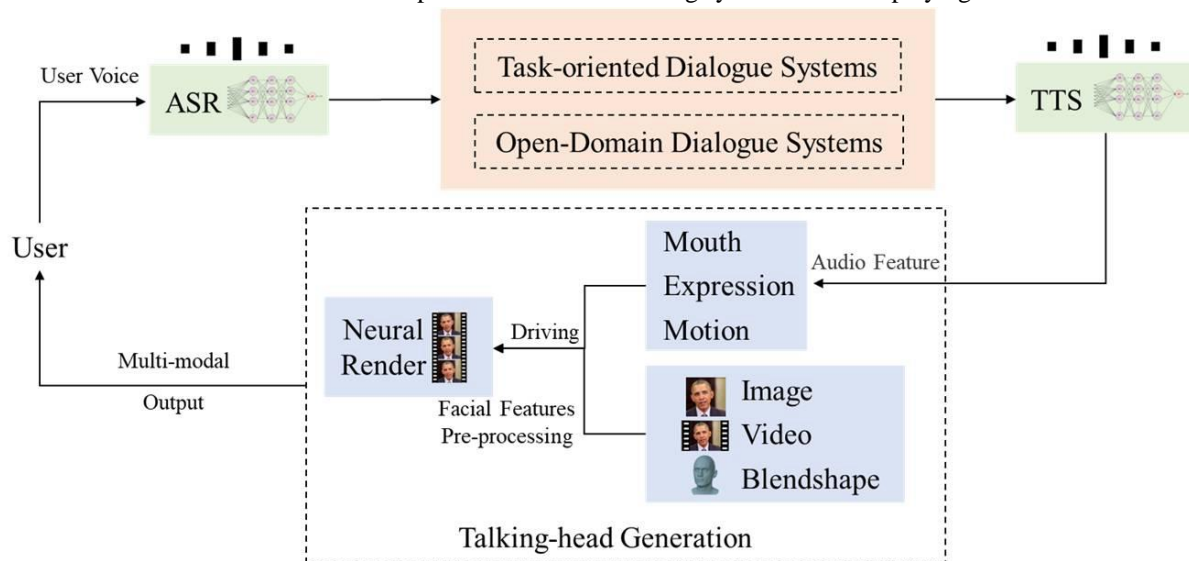

Figure 2: "The system architecture of multimodal human–computer interaction" (Zhen et al., 2023).

Even while most people still use straightforward command/action systems with somewhat basic hardware, research is heading towards creating intelligent and adaptable interfaces. What constitutes intelligence or being smart has not been determined or, at the very least, is not commonly recognised. On the other hand, these concepts can be defined by observing how evidently the utility and functionality of newly available technologies have increased and improved. The development of HCI solutions that provide users with a more straightforward, enjoyable, and satisfying experience is crucial from an economic and mechanical standpoint. The panels are becoming more user-friendly each passing day to help achieve this. One important feature of the current generation of interfaces is the ability to distinguish between using intelligence in the construction of the interaction ("Intelligent HCI") and in the way the "interface interacts with users (Adaptive HCI)". User interfaces with at least a minimal degree of intelligence in user response and/or perception are known as smart HCI designs. Examples include tools that can recognise a user's eye or movement and respond appropriately, as well as voice-activated user interfaces that speak to users in natural language.

Conclusions

In conclusion, "human-computer interaction" is an essential part of "systems design". The system's quality will be decided by how users represent and utilize it. Thus, there has been a lot of focus on creating better HCI designs. The goal of the current study direction is to replace conventional regular touch processes with multimodal, adaptable, intelligent, and genuine ones. The Third Wave, sometimes referred to as context-awareness or ubiquitous computing, seeks to make technology seem more inconspicuous and natural by integrating it into the environment. Virtual reality is another emerging field in HCI that could eventually lead to a standard protocol.

References

[1] Xu, W. (2019). Toward human-centered AI: a perspective from human-computer interaction. *interactions*, *26*(4), 42-46.

[2] Caraban, A., Karapanos, E., Gonçalves, D., & Campos, P. (2019, May). 23 ways to nudge: A review of technology-mediated nudging in human-computer interaction. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1-15).

[3] Stephanidis, C., Salvendy, G., Antona, M., Chen, J. Y., Dong, J., Duffy, V. G., ... & Zhou, J. (2019). Seven HCI grand challenges. *International Journal of Human–Computer Interaction*, *35*(14), 1229-1269.

_____

[4] Ren, F., & Bao, Y. (2020). A review on human-computer interaction and intelligent robots. *International Journal of Information Technology & Decision Making*, *19*(01), 5-47.

[5] Alnuaim, A. A., Zakariah, M., Alhadlaq, A., Shashidhar, C., Hatamleh, W. A., Tarazi, H., ... & Ratna, R. (2022). Human-computer interaction with detection of speaker emotions using convolution neural networks. *Computational Intelligence and Neuroscience*, *2022*.

[6] Nicolescu, L., & Tudorache, M. T. (2022). Human-computer interaction in customer service: the experience with AI chatbots—a systematic literature review. *Electronics*, *11*(10), 1579.

[7] Alnuaim, A. A., Zakariah, M., Shukla, P. K., Alhadlaq, A., Hatamleh, W. A., Tarazi, H., ... & Ratna, R. (2022). Human-computer interaction for recognizing speech emotions using multilayer perceptron classifier. *Journal of Healthcare Engineering*, *2022*.

[8] Zhen, R., Song, W., He, Q., Cao, J., Shi, L., & Luo, J. (2023). Human-computer interaction system: A survey of talking-head generation. *Electronics*, *12*(1), 218.