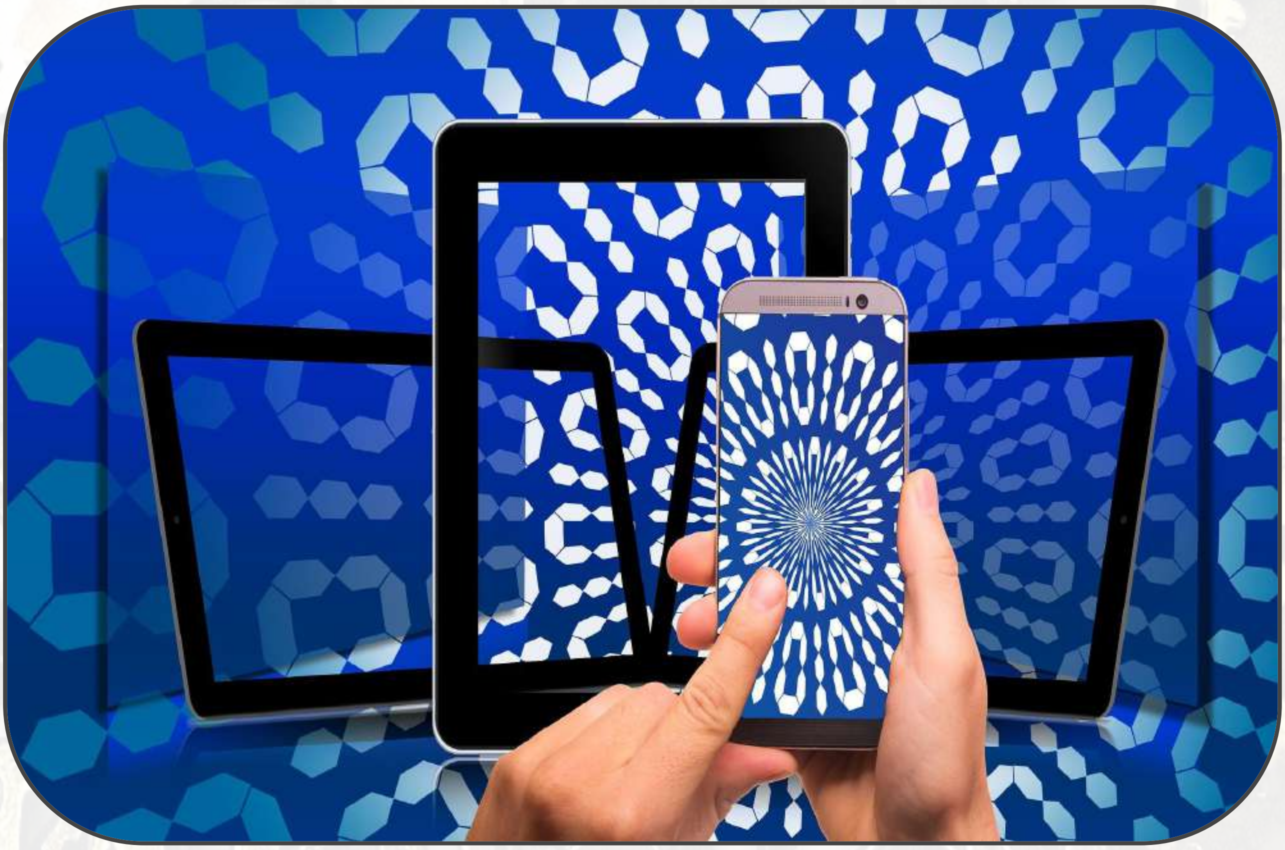
HCI

Human-Computer Interaction



Associate partner:

Co-funded by the Erasmus+ Programme of the European Union

INTRODUCTION TO THE INDUSTRIAL REVOLUTION 4.0

These didactical materials, which have been developed in the framework of the European project 'Industry 4.0 - INTRO 4.0', funded by the European Commission aims to come up with an overview of what has been done in the European Industry in terms of Industry 4.0.

The content of these didactical materials provides the most relevant and useful information on Industry 4.0 to a target group that includes: adults, educators (VET & Higher Education), teachers, trainers, coaches, employers, employees, the general public, and suppliers of innovative solutions.

This information is rooted within the report 'Current Status Of The Industry 4.0' and the report 'Summary Report of the expert interviews/questionnaires and the specific research on the field of manufacturing companies”, both developed by the partners of this project.

INDEX

THIS CONTENT MAY BE OF GREATER INTEREST TO THE COMPANIES

THIS CONTENT MAY BE OF GREATER INTEREST TO THE GENERAL PUBLIC

LEARNING OBJECTIVES

- Understand what is Human-Computer Interaction.
- Know how people and technology joins together thanks to HCI.
- Know the importance of design and implementation in HCI.
- Know about useful information about tools involving HCI.

INTRODUCTION

HCI (Human Computer Interaction) Human-Computer Interaction (HCI) is a multidisciplinary field of study focusing on the design of computer technology and the interaction between humans (the users) and computers.

Learning Objectives

- Understand what is Human-Computer Interaction.
- Know how people and technology joins together thanks to HCI.
- Know the importance of design and implementation in HCI.
- Know about useful information about tools involving HCI.

What is this for?

- Evolving technology improving design and implementation methods
- Improving communication between humans and computers
- Improving design and implementation methods on computer and devices.
- Increasing productivity through interface accesibility

SOME BENEFITS

- Better accessibility
- Evolving technology
- Improved interfaces
- Increase in productivity
- Better user experience

FUTURE APPLICATIONS

Integrated interfaces in your everyday life where the communication between humans and computers reach its limits like fully guided voices devices,virtual augmented reality implementation, etc.

WHAT IS IT?

HCI (human-computer interaction) is the study of how people interact with computers and to what extent computers are or are not developed for successful interaction with human beings.

As its name implies, HCI consists of three parts: the user, the computer itself, and the ways they work together.

User

By "user", we may mean an individual user and a group of users working together. An appreciation of the way people's sensory systems (sight, hearing, touch) relay information is vital. Also, different users form different conceptions or mental models about their interactions and have different ways of learning and keeping knowledge either and. In addition, cultural and national differences play a part.

Computer

When we talk about the computer, we're referring to any technology ranging from desktop computers, to large scale computer systems. For example, if we were discussing the design of a Website, then the Website itself would be referred to as "the computer". Devices such as mobile phones or VCRs (video cassette recorders) can also be considered to be “computers”.

Interaction

There are obvious differences between humans and machines. In spite of these, HCI attempts to ensure that they both get on with each other and interact successfully. In order to achieve a usable system, you need to apply what you know about humans and computers, and consult with likely users throughout the design process. In real systems, the schedule and the budget are important, and it is vital to find a balance between what would be ideal for the users and what is feasible in reality.

WHAT IS IT?

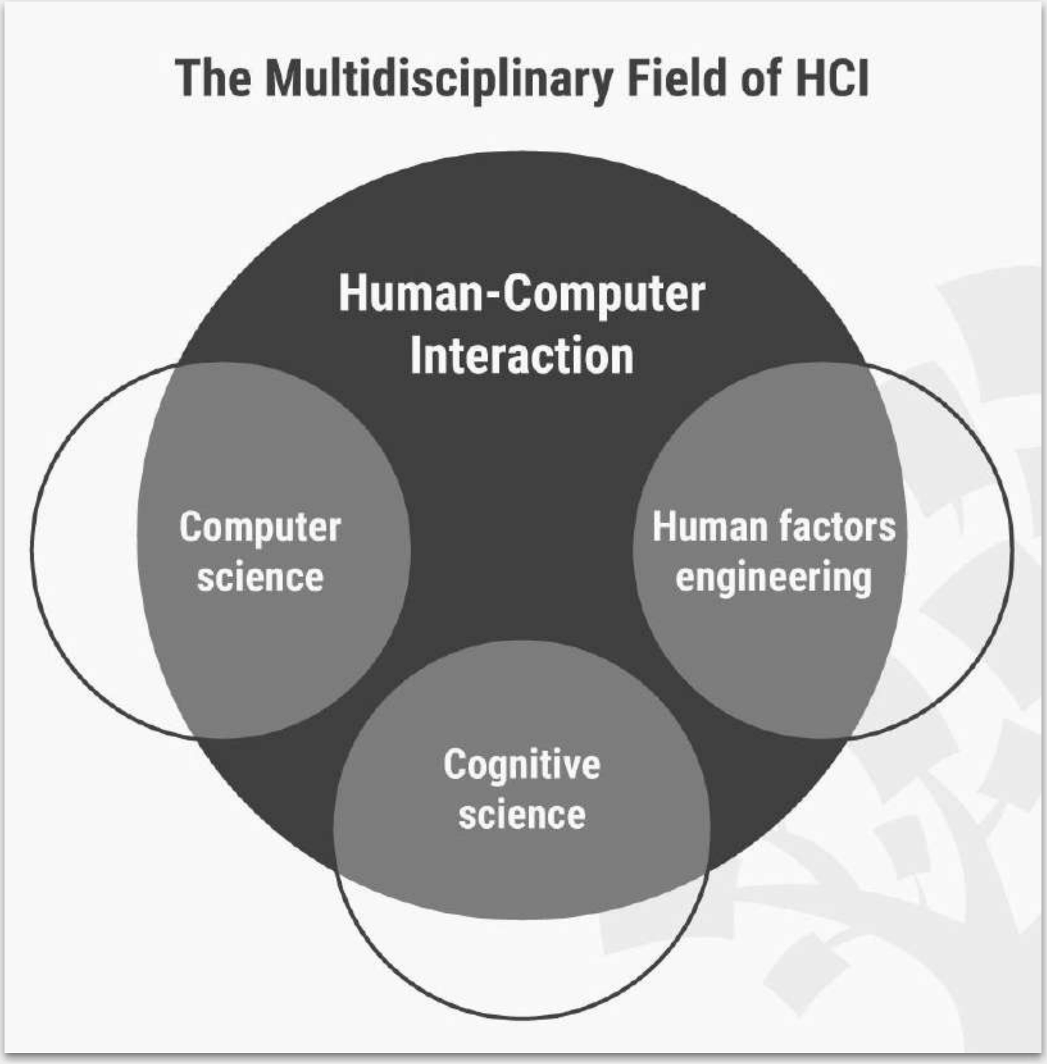


Figure 0. User interfaces in everyday devices. Source: www.interaction-design.org/

"...it no longer makes sense to regard HCI as a specialty of computer science; HCI has grown to be broader, larger and much more diverse than computer science itself. HCI expanded from its initial focus on individual and generic user behavior to include social and organizational computing, accessibility for the elderly, the cognitively and physically impaired, and for all people, and for the widest possible spectrum of human experiences and activities. It expanded from desktop office applications to include games, learning and education, commerce, health and medical applications, emergency planning and response, and systems to support collaboration and community. It expanded from early graphical user interfaces to include myriad interaction techniques and devices, multi-modal interactions, tool support for model-based user interface specification, and a host of emerging ubiquitous, handheld and context-aware interactions."

— John M. Carroll, author and a founder of the field of human-computer interaction.

WHAT IS THIS FOR?

HCI AND THE EVOLUTION OF TECHNOLOGIES

User Interface

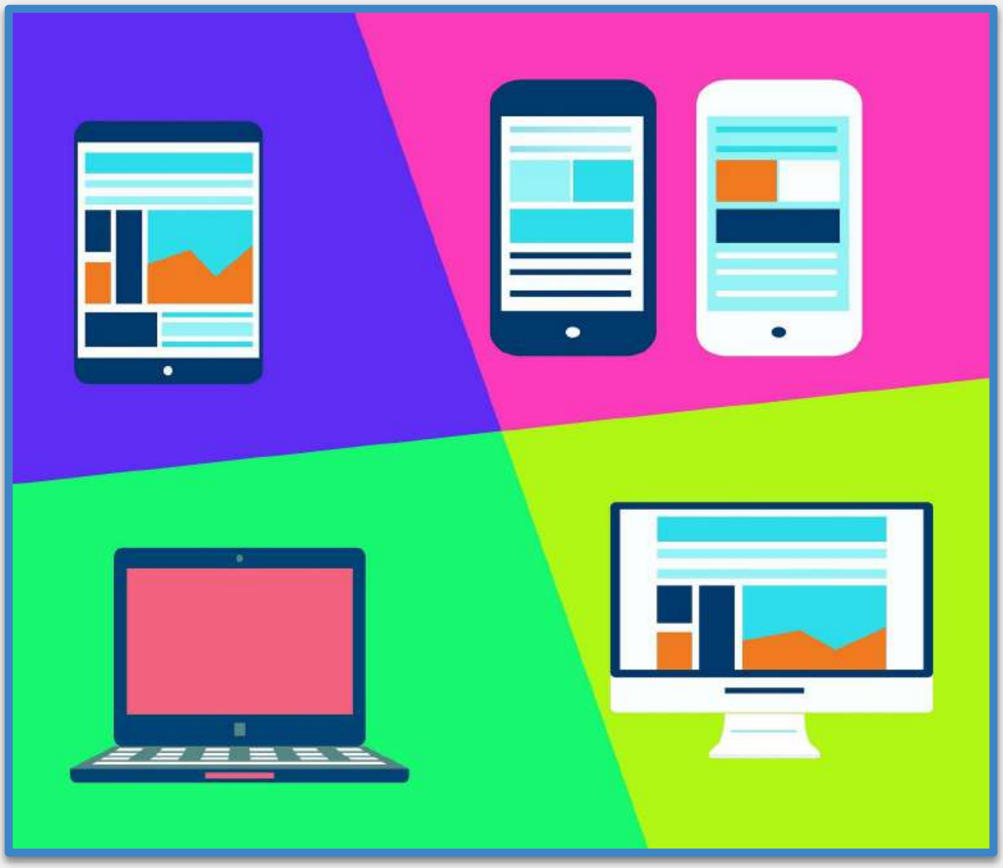


Figure 1. User interfaces in everyday devices. Source: medium.theuxblog.com

An interface sits **between you and technology**, and nearly every technology has one. Yet when we say the word interface, we inevitably think of the UI between a user and a computer, smartphone, tablet, or similar device. They can be physical devices such as keyboards, mice, touchscreens, and virtual objects such as screen icons and menus, voice-driven natural language assistants, gesture recognition devices, and more.

WHAT IS THIS FOR?

Interfaces are **intermediaries** that shield us from the underlying complexities of what we want to do from how the item needs to operate. In the computerized world, those complexities can be the underlying programs, operating systems, and networks. For example, when AOL, Prodigy, and CompuServe were popular, they all tried to shield us from some of the underlying intricacies through their GUI (Graphic User Interface). Nowadays, Firefox, Edge, Safari, and Chrome browsers have user interfaces that hide the underlying HTML, style sheets, and scripts. Web and app designers strive to create user interfaces that prevent us from making mistakes, improve productivity, ensure smooth operations, and shield us from inefficient, confusing or unusable products and systems.

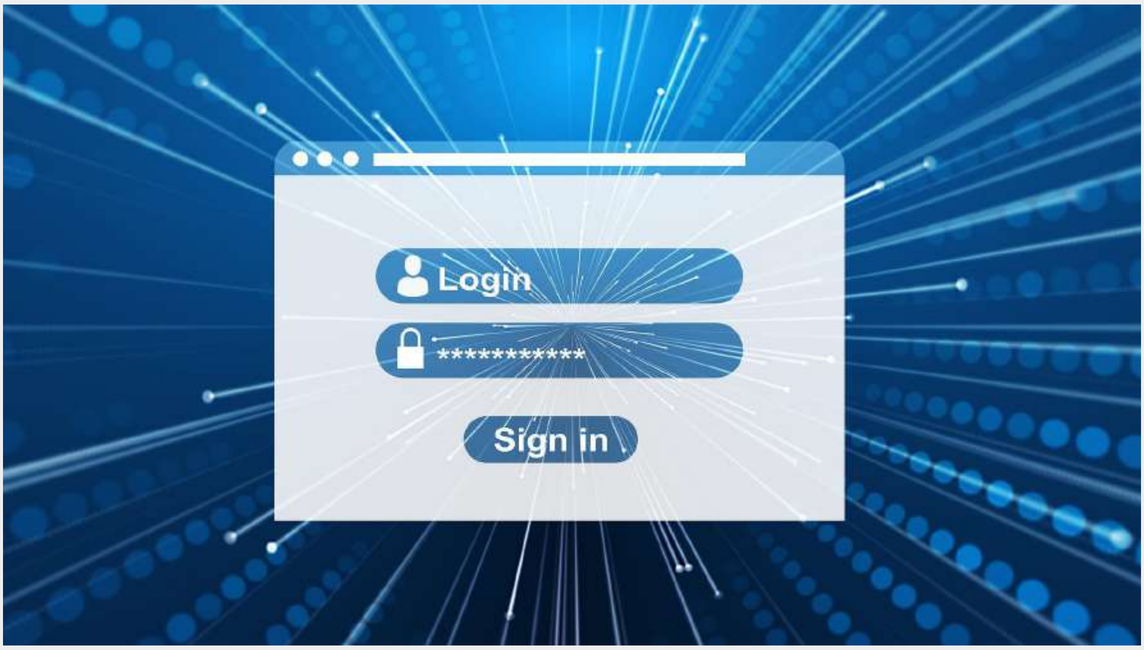


Figure 2. Interface. Source: www.pixabay.com

They employ **design and implementation techniques that fall under the scope of Human-Computer Interaction**. Many of the standards governing this field of study are contained in ISO 9241 “Ergonomics of human-system interaction”. The benefits of a good UI are tangible and measurable. In the business world, good interfaces lead to higher morale and job satisfaction, and lower training costs and staff turnover, all leading to lower operational costs.

WHAT IS THIS FOR?

From a user's perspective, **the UI is a gateway into their computer** and represents easy access to the intricacies of its underlying hardware, software, and networking. For example, when you enter an elevator, its interface works with elevator logic to close the door, get you to your floor at an acceptable speed, process multiple floor requests, adjust the floor indicator, open the door, and dispatch the car to waiting passengers.

Important elements in an user interface

Clear

A UI needs to be easy to use and obvious. For example, present the user with a sorted list if it is appropriate.

Concise

Clarity is important, but it should not be overly wordy.

Familiar

People learn new concepts based on ones they already know.

Responsive

User feedback is key. Rapidly acknowledging a GUI screen press.

Consistent

Visual interfaces in a series of screens should have the same "look and feel." For example, Microsoft Word, Excel and PowerPoint have a similar appearance.

Attractive

Users may need to use the interface daily, so it should be nice-looking.

WHAT IS THIS FOR?

Efficient

Users should interact with complex applications in the simplest way possible. For example, parse their "first" and "last" name instead of using two fields.

Forgiving

If the UI can't prevent a user error, it should allow the user to correct it.

When hardware, software, and the way we interact with computers are constantly evolving then HCI remains relevant for UI designers and engineers alike who want to investigate the "Why" behind the "How" of the interfaces they design. **The answer to that "Why?" will almost always be: To make interfaces "easy to learn, and easy to use"** an early mantra of HCI.

WHAT IS THIS FOR?

Augmented Reality

Augmented reality is **enhancing the view of reality** by supplementing virtual objects using technology.

Figure 3. Augmented Reality example. Source: forbes.com

Augmented reality is **enhancing the view of reality** by supplementing virtual objects using technology. Using AR technology the environment around a person can become much more **interactive and digital**. Apart from sense of sight **AR applies to all senses**, such as hearing, smell, and touch.

Various types of hardware components are required for the functioning of augmented reality: processors, sensors, display and input devices. Smartphones or tablets often use camera, GPS and other sensors.

Figure 4. Augmented Reality example. Source: phys.org

WHAT IS THIS FOR?

Augmented reality takes a real world scene with the help of camera on device and superimposes images, videos or sounds on the real world scene. AR works in two ways, first based on positioning of markers which is identified by the software on the device and then the content hidden in the marker is displayed and second way is to identify the location of device through GPS and displays the content according to the field of view of the device.

Systems in future that will allow human computer interaction will require technology to **interpret human gestures and movements** including complexity of motions like joint movements. New systems that can mimic human brain and systems that are capable of deep learning have been developed by researchers .These systems can even understand endless complexities of joint angles.

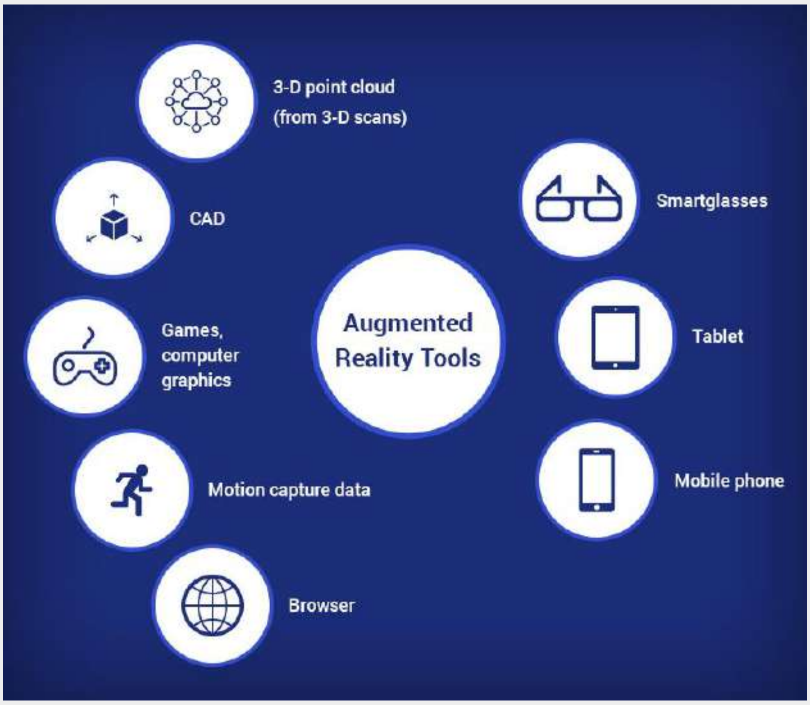


Figure 5. Augmented Reality tools. Source: mobixed.com

Augmented Reality is emerging as one of the most powerful technologies in the field of computer science. It has added a new dimension in the world of computing. With its capability of superimposition it has been contributing to entertainment, education, medical sciences, commercial, sports, military etc. With the rapid development of Human Computer Interaction and its ability to interpret three dimensional human gestures, it will lead Augmented Reality to an exceptional level.

WHAT IS THIS FOR?

Social Computing

Social computing involves **the digital systems that support online social interaction**. Some online interactions are obviously social – exchanging email with a family member, sharing photos with friends, instant messaging with coworkers. These interactions are prototypically social because they are about communicating with people we know. But other sorts of online activity also count as social – creating a web page, bidding for something on eBay™, following someone on Twitter™, making an edit to Wikipedia. These actions may not involve people we know, and may not lead to interactions, but nevertheless they are social because we do them with other people in mind: the belief that we have an audience – even if it is composed of strangers we will never meet – shapes what we do, how we do it, and why we do it.

Figure 6. Social Computing Media. Source: posachaidonut.wordpress.com

Thus when we speak of social computing we are concerned with **how digital systems go about supporting the social interaction** that is fundamental to how we live, work and play. **They do this by providing communication mechanisms** through which we can interact by talking and sharing information with one another, and by capturing, processing and displaying traces of our online actions and interactions that then serve as grist for further interaction.

WHAT IS THIS FOR?

Social Computing examples

Blogs, (derived from Web logs) are the most visible of the social computing initiatives. Started in late '90s, they have come to take the world of journalism by storm, and have extended their presence into several other domains as well. Blogs may be thought of as online journals, which may be published by an individual or a small group, through the Web interface, and focused either on a single topic or a variety of topics reflecting interests of the authors.

Wikipedia is an online open source encyclopedia built by aggregating so-called wikis, which are tools (or instances) of collaborative authoring of tagged hypertext content, with version control and user feedback features built in. Wikis allow several users to contribute their knowledge so that a structured hypertext article on a topic can build itself from grassroots. The quality control derives from user feedback, and version control allows undoing changes and reverting when necessary. Wikis are popularly used as knowledge sharing tools and for collaborative authoring in teams.

Skype, the peer-to-peer, Internet-based voice and video communication service, represents social cooperation in bandwidth usage subverting traditional telephony; millions of users on the edge collaborate to share their bandwidth and realize service quality that is competitive with circuit-switched expensive lines, and thus undermine the usage-based pricing model of traditional telephony.

LinkedIn is a social network for business professionals rapidly gained popularity. In essence, it takes "networking" online; allowing professionals to create their profiles, and invite their professional contacts to be part of their "network." Networks grow rapidly, and users help each other by "endorsing" them and by various referrals and testimonials, as well as by providing access to the networks of each.

GOOD PRACTICES

Over the past decade, the research community has made major inroads in building common HCI design patterns, resulting in a lingua franca for UI design. These patterns are increasingly important for designers as a **vehicle for mediating between HCI and other software engineering practices**, but there is still some room for improving them to maximize their utility to designers. Today's software applications introduce challenges that also call for HCI patterns capable of organizing the interaction on modern interfaces— for example, mobile platforms.

Usability functionalities are one of the best ways to tackle routine user tasks—some examples include letting users undo an action, informing them about an action's progress, enabling them to cancel a command in progress, or letting them set their own preference. However, much of this functionality goes beyond UI design: providing an undo functionality has more implications than just displaying an undo button in the UI.

Therefore, integrating usability functionalities into a software application is not straightforward. In our work, we've developed the following related best practices:

- The development team must gather information related to the usability functionalities suitable for the **target software application based on user interactivity**, the complexity of the required usability functionality, and the tradeoff between usability and any other quality attributes.

- Developers must integrate such usability functionalities into **the artifacts used in the application to describe the system's functional requirements, use cases, user stories, or any other artifact used for that aim.**

GOOD PRACTICES

- Usability functionalities imply **specific responsibilities that the software system must address**. Responsibilities for the abort functionality, for example, must include recording information to recover the system status prior to the cancelable command's execution, stopping the active command, or freeing allocated resources.

- Developers **must map all of these responsibilities to different components of the design models**—for example, to particular classes in class diagrams—which also need to be allocated to the respective layer or subsystem in the architecture according to the main architectural style.

- Finally, developers must **test usability functionalities like any other software requirement**. These practices are clear indicators of usability's impact on software design. Developers must carefully deal with usability functionalities during software construction to deploy them properly.

Beyond UI design and usability components engineering, an important issue for developers wrestling with usability is how to deal with **current and future user experiences**—that is, how users will apply, perceive, and learn the software, as well as how it will evolve and adapt to users' changing expectations.

HCI practitioners advocate a user centered design (UCD) approach, which includes a set of activities for building interactive systems with user involvement in all development stages. According to this model, if a UCD process is identified as being necessary, developers must determine who is to use the product and for what purpose, in addition to what other requirements a successful product must fulfill. Developers also need to evaluate design alternatives, create design solutions, and evaluate their usability with real users.

GOOD PRACTICES

The notion of design as it is defined in user-centered approaches refers to UI design or interaction design, not to software design as conceived from a software architecture perspective. Therefore, developers must integrate this UCD process into a particular software development process to build systems with the required quality attributes, including usability.

An important challenge in this integration is how to manage the design loop generated by the continuously evolving, ambiguous, and unclear user experience during the software development process. HCI techniques such as user observation, focus groups, or even social networks can help elicit proper user needs, paper prototyping or storyboarding can drive development and heuristic evaluations, and usability testing can help develop software applications that are more focused on real users. But an organizational change is necessary to align the software process with the design loops needed to properly address the user experience. This involves tackling the following issues:

- **Different HCI techniques require different expertise, resources, and user availability**. Laboratory usability testing and video recording are two HCI techniques proposed for assessing usability, for example, but they require applying an entire physical infrastructure. HCI experts must select the best techniques for each project based on its own idiosyncrasies.

- **Integrating UCD activities into particular software development processes** with their own particularities will produce differing results. In an agile project, for example, usability questionnaires might be replaced by a thinking aloud protocol to assess the software product's usability after each iteration.

GOOD PRACTICES

- **Developers must integrate HCI techniques incrementally and logically into a software engineering process**. We advocate the use of HCI techniques first for gathering user needs and expectations, then for designing the respective interaction, and, finally, for evaluating the resulting system's usability. It's inefficient to apply usability testing at the end of the development project without applying any other HCI technique during requirements engineering or UI design.

HCI practice	Description	Examples	Project phase	Person in charge	Benefits
HCI interface patterns	Patterns used to capture best practices for solving particular user-interface design scenarios	Color-coded divisions, titled sections, scrolling menu, icon menu, shopping carts, small groups of related things, chart or graph, map of navigable spaces, thumbnail, collapsible panels	At any time, but preferably during interface design.	The UI designer or the programmer, if no UI designer participates in the development team	Improve system interface appearance and navigation; improve access to system functionalities (such as menus)
Usability functionalities	Usability requirements that ease the use of any software system and facilitate daily user tasks	Undo/cancel commands, predict a task duration, aggregate commands, check for errors, present system state, provide good help, minimize user recovery work due to systems errors	During requirements elicitation to capture the usability functionalities as any other requirement.	The requirements analyst must be conscious about the need to incorporate usability functionalities in the system; the project manager must estimate the cost and resources needed	Enrich user interaction capabilities with new usability functionalities that the user is not aware of a priori
User centered design	Approach for building interactive systems that explicitly involves the user through all development stages; particular HCI techniques contribute to this aim	Focus groups, card sorting, scenarios of use, thinking aloud, heuristic evaluation, Wizard of Oz prototyping, expert evaluation, participatory evaluation, laboratory testing, surveys	At project startup because it affects the entire software development process	Whole organization must be involved to adapt the traditional software development process to take into account user-centered needs and HCI techniques	Helps capture user interaction with the system; reduces the level of user rejection of the software product

GOOD PRACTICES

HUMANISING AUTONOMY

We were founded on the premise of enabling a safer, more human-centered implementation of autonomous technology, and called ourselves Humanising Autonomy because that was what we wanted to do. We started the company because we realised that our position on urban mobility was not being reflected by the automotive industry's view of automated vehicles in the cities of today and tomorrow. We want to see cities where mobility systems — private vehicles, ridesharing fleets, public transportation — take into account the vulnerable road users outside the vehicle, not just focus on the interior experience. The pedestrians, cyclists, and other users of the road should be able to interact with these vehicles — and the vehicles will need to understand this nuanced communication to be able to safely and efficiently navigate complex urban environments. We discovered an opportunity to create something truly necessary, but in our view underserved by current technology, and started developing our own AI-powered technology that is able to predict the full range of pedestrian and vulnerable road user behaviour in real time for the safer, and more trustworthy mobility systems that we have in our minds.

Some leading companies:

CTRL-Labs ORBAI tobii

TRIFACTA emteq

Furhat Robotics

BENEFITS FOR THE COMPANY

The Benefits of Converged Infrastructure

- The infrastructure deploys faster than traditional data centers, particularly for cloud environments.

- HCI is easier to manage since it includes a software interface. **IT admin can keep a pulse on monitoring and troubleshooting actions by using the HCI software**.

- **HCI lowers operating and capital costs** since it relies on commodity hardware such as a white box or x86 platforms. In addition, "converged systems also provide a single resource pool for applications that make them easier to share thus improving utilization efficiency."

- HCI is more flexible, scalable, and **easy to maneuver**.

- Cisco points out that HCI "enables companies to gain the benefits of on-demand infrastructure for data-centric workloads **without placing resources in public clouds.**"

- HCI is readily **accessible** to purchase.

- HCI is capable of **increased data protection**. HCI is designed to effectively deal with this data protection problem since it already comes with included comprehensive backup and recovery capabilities allowing to meet even the strictest RTO (recovery time objective) and RPO (recovery point objective) requirements.

- HCI can operate as a virtual machine (VM) since it includes a hypervisor.

FUTURE APPLICATIONS

Intelligent and Adaptive HCI

Although the devices used by majority of public are still some kind of plain command/action setups using not very sophisticated physical apparatus, the flow of research is **directed to design of intelligent and adaptive interfaces**. The exact theoretical definition of the concept of intelligence or being smart is not known or at least not publicly agreeable. However, one can define these concepts by the apparent growth and improvement in functionality and usability of new devices in market. As mentioned before, it is **economically and technologically crucial to make HCI designs that provide easier, more pleasurable and satisfying experience for the users**. To realize this goal, the interfaces are getting more natural to use every day.

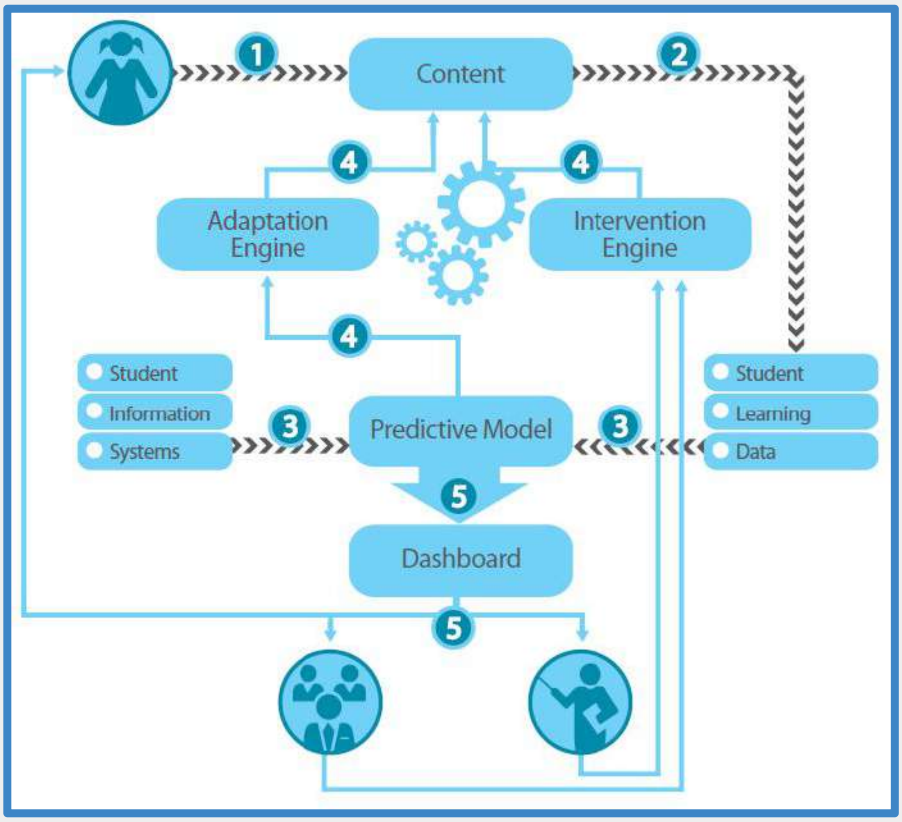


Figure 7. Adaptive learning process. source: www.dreambox.com

Evolution of interfaces in note-taking tools is a good example. First there were typewriters, then keyboards and now touch screen tablet PCs that you can write on using your own handwriting and they recognize it change it to text and if not already made, tools that transcript whatever you say automatically so you do not need to write at all. One important factor in new generation of interfaces is to differentiate between using intelligence in the making of the interface (Intelligent HCI) or in the way that the interface interacts with users (Adaptive HCI).

FUTURE APPLICATIONS

Intelligent HCI designs are interfaces that incorporate at least some kind of intelligence in perception from and/or response to users. A few examples are speech enabled interfaces that use natural language to interact with user and devices that visually track user's movements or gaze and respond accordingly. Adaptive HCI designs, on the other hand, may not use intelligence in the creation of interface but use it in the way they continue to interact with users. An adaptive HCI might be a website using regular GUI for selling various products. This website would be adaptive -to some extent- if it has the ability to recognize the user and keeps a memory of his searches and purchases and intelligently search, find, and suggest products on sale that it thinks user might need. Most of these kinds of adaptation are the ones that deal with cognitive and affective levels of user activity.

Another example that uses both intelligent and adaptive interface is a PDA (Personal Digital Assistant) or a tablet PC that has the handwriting recognition ability and it can adapt to the handwriting of the logged in user so to improve its performance by remembering the corrections that the user made to the recognised text.

Finally, another factor to be considered about intelligent interfaces is that most non-intelligent HCI design are passive in nature i.e. they only respond whenever invoked by user while ultimate intelligent and adaptive interfaces tend to be active interfaces. The example is smart billboards or advertisements that present themselves according to users' taste. In the next section, combination of different methods of HCI and how it could help towards making intelligent adaptive natural interfaces is discussed.

FUTURE APPLICATIONS

Ubiquitous Computing and Ambient Intelligence

The latest research in HCI field is unmistakably ubiquitous computing (Ubicomp). The term which often used interchangeably by ambient intelligence and pervasive computing, refers to the ultimate methods of human-computer interaction that is **the deletion of a desktop and embedding of the computer in the environment** so that it becomes invisible to humans while surrounding them everywhere hence the term ambient.

The idea of ubiquitous computing was first introduced by Mark Weiser during his tenure as chief technologist at Computer Science Lab in Xerox PARC in 1998. His idea was t**o embed computers everywhere in the environment and everyday objects so that people could interact with many computers at the same time while they are invisible to them and wirelessly communicating with each other.**

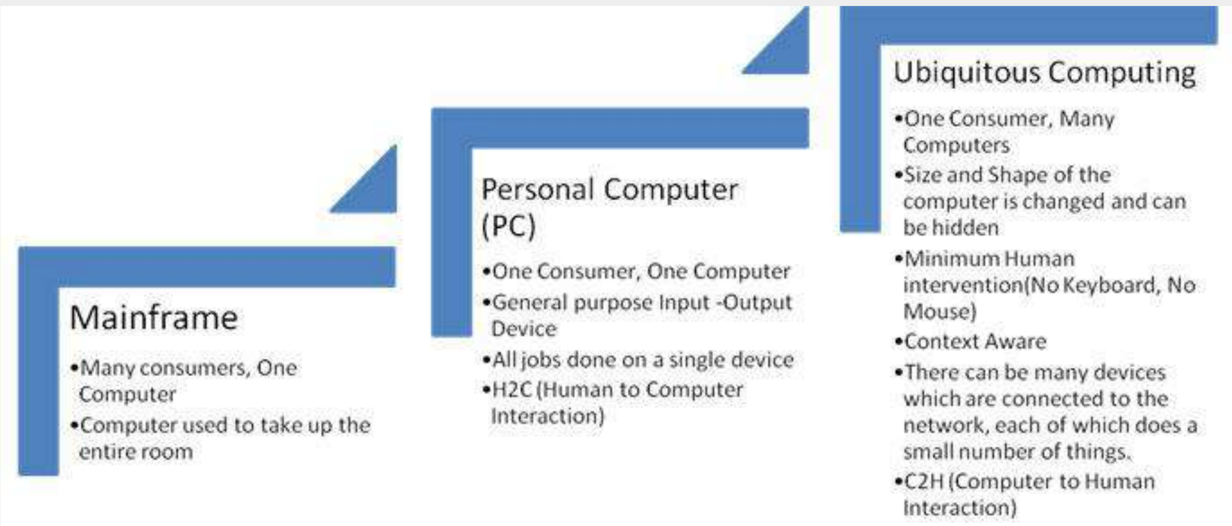


Figure 8. Computing waves. Source: www.thbs.com

Ubicomp has also been named the Third Wave of computing. The First Wave was the mainframe era, many people one computer. Then it was the Second Wave, one person one computer which was called PC era and now Ubicomp introduces many computers one person era.

FUTURE APPLICATIONS

Visual-Based HCI

The visual based human computer interaction is probably the most widespread area in HCI research. Considering the extent of applications and variety of open problems and approaches, researchers tried to tackle different aspects of human **responses which can be recognized as a visual signal**.

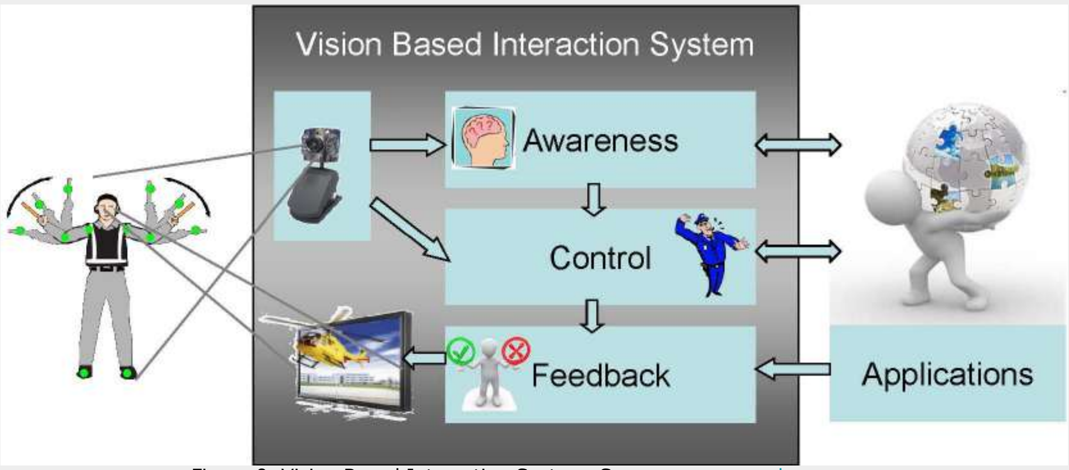


Figure 9. Vision Based Interaction System. Source: www.ganghua.org

While the goal of each area differs due to applications, a general conception of each area can be concluded. Facial expression analysis generally deals with recognition of emotions visually. Body movement tracking and gesture recognition are usually the main focus of this area and can have different purposes but they are mostly used for direct interaction of human and computer in a command and action scenario. Gaze detection is mostly an indirect form of interaction between user and machine which is mostly used for better understanding of user's attention, intent or focus in context-sensitive situations. The exception is eye tracking systems for helping disabilities in which eye tracking plays a main role in command and action scenario, e.g. pointer movement, blinking for clicking. It is notable that some researchers tried to assist or even replace other types of interactions (audio-, sensor-based) with visual approaches. For example, lip reading or lip movement tracking is known to be used as an influential aid for speech recognition error correction.

FUTURE APPLICATIONS

Audio-Based HCI

The audio based interaction between a computer and a human is another important area of HCI systems. This area **deals with information acquired by different audio signals**. While the nature of audio signals may not be as variable as visual signals but the information gathered from audio signals can be more trustable, helpful, and is some cases unique providers of information.

Historically, speech recognition and speaker recognition have been the main focus of researchers. Recent endeavors to integrate human emotions in intelligent human computer interaction initiated the efforts in analysis of emotions in audio signals. Other than the tone and pitch of speech data, typical human auditory signs such as sigh, gasp, and etc helped emotion analysis for designing more intelligent HCI system. Music generation and interaction is a very new area in HCI with applications in art industry which is studied in both audio- and visual-based HCI systems.

Figure 10. Audio based HCI Speech recognition. Source: www.medium.com

ADVANCED CONTENT

Multimodal HCI Systems

The term multimodal refers to combination of multiple modalities. In MMHCI systems, these modalities mostly refer to **the ways that the system responds to the inputs**, i.e. communication channels. The definition of these channels is inherited from **human types of communication which are basically his senses: Sight, Hearing, Touch, Smell, and Taste.** The possibilities for interaction with a machine include but are not limited to these types.

Therefore, a multimodal interface acts as a facilitator of human-computer interaction via two or more modes of input that go beyond the traditional keyboard and mouse. The exact number of supported input modes, their types and the way in which they work together may vary widely from one multimodal system to another. Multimodal interfaces incorporate different combinations of speech, gesture, gaze, facial expressions and other non-conventional modes of input. One of the most commonly supported combinations of input methods is that of gesture and speech. Although **an ideal multimodal HCI system should contain a combination of single modalities that interact correlatively, the practical boundaries and open problems in each modality oppose limitations on the fusion of different modalities**. In spite of all progress made in MMHCI, in most of existing multimodal systems, the modalities are still treated separately and only at the end, results of different modalities are combined together.

ADVANCED CONTENT

The reason is that the open problems in each area are yet to be perfected meaning that there is still work to be done to acquire a reliable tool for each sub-area. Moreover, roles of different modalities and their share in interplay are not scientifically known. "Yet, people convey multimodal communicative signals in a complementary and redundant manner. Therefore, **in order to accomplish a human-like multimodal analysis of multiple input signals acquired by different sensors, the signals cannot be considered mutually independently and cannot be combined in a context-free manner at the end of the intended analysis but, on the contrary, the input data should be processed in a joint feature space and according to a context-dependent model**. In practice, however, besides the problems of context sensing and developing context dependent models for combining multisensory information, one should cope with the size of the required joint feature space. Problems include large dimensionality, differing feature formats, and time-alignment."

An interesting aspect of multimodality is the collaboration of different modalities to assist the recognitions. For example, lip movement tracking (visual-based) can help speech recognition methods (audio-based) and speech recognition methods (audio-based) can assist command acquisition in gesture recognition (visual-based). The next section shows some of application of intelligent multimodal systems.

ADVANCED CONTENT

A classic example of a multimodal system is the "Put That There" demonstration system. **This system allowed one to move an object into a new location on a map on the screen by saying "put that there" while pointing to the object itself then pointing to the desired destination**. Multimodal interfaces have been used in a number of applications including mapbased simulations, such as the aforementioned system; information kiosks, such as AT&T's MATCHKiosk and biometric authentication systems.

Multimodal interfaces can offer a number of advantages over traditional interfaces. For one thing, **they can offer a more natural and user-friendly experience**. For instance, in a real-estate system called Real Hunter, one can point with a finger to a house of interest and speak to make queries about that particular house. Using a pointing gesture to select an object and using speech to make queries about it illustrates the type of natural experience multimodal interfaces offer to their users. Another key strength of multimodal interfaces is their ability to provide redundancy to accommodate different people and different circumstances. For instance, MATCHKiosk **allows one to use speech or handwriting to specify the type of business to search for on a map. Thus, in a noisy setting, one may provide input through handwriting rather than speech.**

Few other examples of applications of multimodal systems are: Smart Video Conferencing, Intelligent Homes/Offices, Driver Monitoring, Intelligent Games, E-Commerce, Helping People with Disabilities.

Multimodal HCI Systems Applications

People with disabilities

One good application of multimodal systems is **to address and assist people with disabilities (as persons with hands disabilities), which need other kinds of interfaces than ordinary people.** In such systems, users with disabilities can perform work on the PC by interacting with the machine using voice and head movements. Two modalities are then used: speech and head movements. Both modalities are active continuously. The head position indicates the coordinates of the cursor in current time moment on the screen. Speech, on the other hand, provides the needed information about the meaning of the action that must be performed with an object selected by the cursor.

Synchronization between the two modalities is performed by calculating the cursor position at the beginning of speech detection. This is mainly due to the fact that during the process of pronouncing the complete sentence, the cursor location can be moved by moving the head, and then the cursor can be pointing to other graphical object; moreover the command which must be fulfilled is appeared in the brain of a human in a short time before beginning of phrase input.

In spite of some decreasing of operation speed, **the multimodal assertive system allows working with computer without using standard mouse and keyboard.** Hence, such system can be successfully used for hands-free PC control for users with disabilities of their hands.

ADVANCED CONTENT

Emotion Recognition

As we move towards a world in which computers are more and more ubiquitous, **it will become more essential that machines perceive and interpret all clues**, implicit and explicit, that we may provide them regarding our intentions. **A natural human-computer interaction cannot be based solely on explicitly stated commands**. Computers will have to detect the various behavioural signals based on which to infer one's emotional state. This is a significant piece of the puzzle that one has to put together to predict accurately one's intentions and future behaviour.

People are able to make prediction about one's emotional state based on their observations about one's face, body, and voice. Studies show that if one had access to only one of these modalities, the face modality would produce the best predictions.

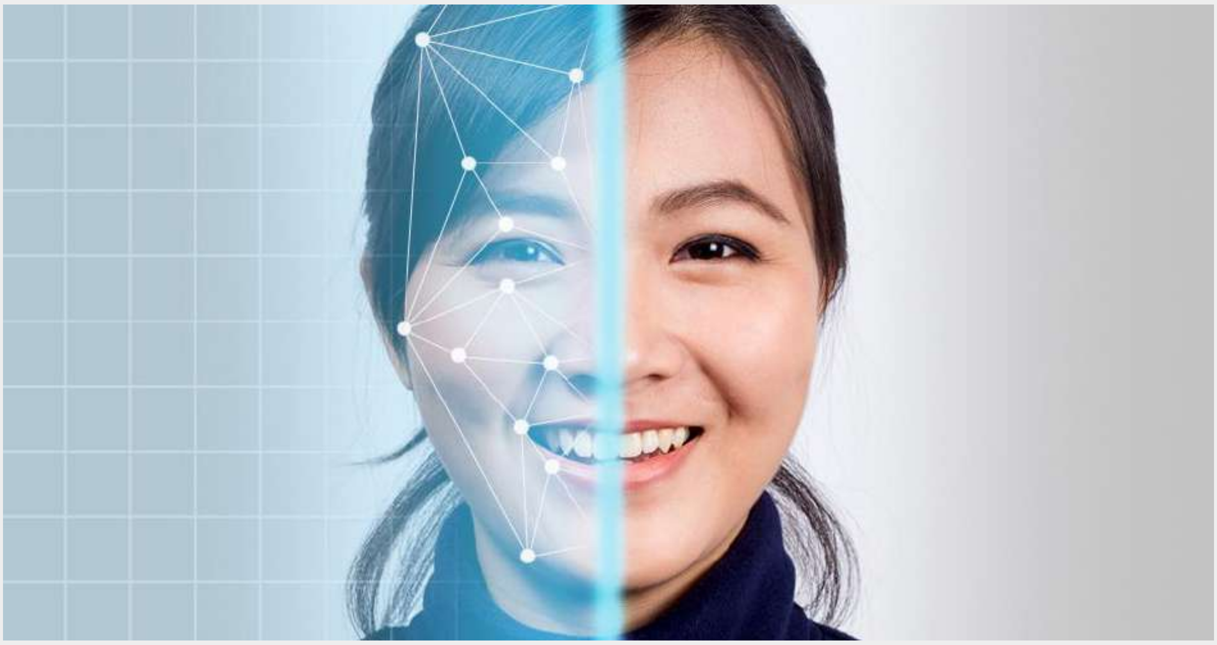


Figure 11. Facial based recognition. Source: www.acart.com

ADVANCED CONTENT

However, this accuracy can be improved by 35% when human judges are given access to both face and body modalities together. This suggests that affect recognition, which has for the most part focused on facial expressions, **can greatly benefit from multimodal fusion techniques**.

One of the few works that has attempted to integrate more than one modality for affect recognition is in which facial features and body posture features are combined to produce an indicator of one's frustration. Another work that integrated face and body modalities is in which the authors showed that, similar to humans, machine classification of emotion is better when based upon face and body data, rather than either modality alone. The authors attempted to fuse facial and voice data for affect recognition. Once again, remaining consistent with human judges, machine classification of emotion as neutral, sad, angry, or happy was most accurate when the facial and vocal data is combined.

They recorded the four emotions: "sadness, anger, happiness, and neutral state". The detailed facial motions were captured in conjunctions with simultaneous speech recordings. **Deducted experiments showed that the performance of the facial recognition based system overcame the one based on acoustic information only.** Results also show that an appropriate fusion of both modalities gave measurable improvements.

Results show that the emotion recognition system based on acoustic information only give an overall performance of 70.9%, compared to an overall performance of 85% for a recognition system based on facial expressions. This is, in fact, due to the fact that the cheek areas give important information for emotion classification.

On the other hand, for the bimodal system based on fusing the facial recognition and acoustic information, the overall performance of this classifier was 89.1%.

ADVANCED CONTENT

Different input modalities are suitable for expressing different messages. For instance, speech provides an easy and natural mechanism for expressing a query about a selected object or requesting that the object initiate a given operation. However, speech may not be ideal for tasks, such as the selection of a particular region on the screen or defining out a particular path. These types of tasks are better accommodated by hand or pen gestures. However, making queries about a given region and selecting that region are all typical tasks that should be accommodate by a map-based interface. Thus, **the natural conclusion is that map-based interfaces can greatly improve the user experience by supporting multiple modes of input, especially speech and gestures**.

Quickset is one of the more widely known and older map-based applications that make use of speech and pen gesture input. Quickset is a military-training application that allows users to use one of the two modalities or both simultaneously to express a full command. For instance, users may simply draw out with a pen a predefined symbol for platoons at a given location on the map to create a new platoon in that location. Alternatively, users could use speech to specify their intent on creating a new platoon and could specify vocally the coordinates in which to place the platoon. Lastly, users could express vocally their intent on making a new platoon while making a pointing gesture with a pen to specify the location of the new platoon.

A more recent multimodal map-based application is Real Hunter. It is a real-estate interface that expects users to select objects or regions with touch input while making queries using speech. For instance, the user can ask "How much is this?" while pointing to a house on the map.

ADVANCED CONTENT

Tour guides are another type of map-based applications that have shown great potential to benefit from multimodal interfaces. One such example is MATCHKiosk, the interactive city guide. In a similar fashion to Quickset, MATCHKiosk allows one to express certain queries using speech only, such as "Find me Indian restaurants in Washington."; using pen input only by circling a region and writing out "restaurants"; using bimodal input by saying "Indian restaurants in this area" and drawing out a circle around Alexandria. These examples illustrate MATCHKiosk's incorporation of handwriting recognition that can frequently substitute for speech input. **Although speech may be the more natural option for a user, given the imperfectness of speech, especially in noisy environments, having handwriting as a backup can reduce user frustration**.

Human-Robot Interface

Similar to some map-based interfaces, human-robot interfaces usually have to provide mechanisms for pointing to particular locations and for expressing operation-initiating requests. As discussed earlier, the former type of interaction is well accommodated by gestures, whereas the latter is better accommodate by speech. Thus, the human-robot interface built by the Naval Research Laboratory (NRL) should come as no surprise. NRL's interface allows users to point to a location while saying "Go over there". Additionally, it allows users to use a PDA screen as a third possible avenue of interaction, which could be resorted to when speech or hand gesture recognition is failing. Another multimodal human-robot interface is the one built by Interactive System Laboratories (ISL), **which allows use of speech to request the robot to do something while gestures could be used to point to objects that are referred to by the speech.** One such example is to ask the robot, "switch on the light" while pointing to the light. Additionally, in ISL's interface, **the system may ask for clarification from the user when unsure about the input**. For instance, in case that no hand gesture is recognized that is pointing to a light, the system may ask the user: "Which light?"

ADVANCED CONTENT

Medicine

By the early 1980s, surgeons were beginning to reach their limits based on traditional methods alone. The human hand was unfeasible for many tasks and greater magnification and smaller tools were needed. **Higher precision was required to localize and manipulate within small and sensitive parts of the human body**. Digital robotic neuro-surgery has come as a leading solution to these limitations and emerged fast due to the vast improvements in engineering, computer technology and neuroimaging techniques. Robotics surgery was introduced into the surgical area.

State University of Aerospace Instrumentation, University of Karlsruhe (Germany) and Harvard Medical School (USA) has been working on developing man-machine interfaces, adaptive robots and multi-agent technologies intended for neuro-surgery.

The neuro-surgical robot consists of the following main components: An arm, feedback vision sensors, controllers, a localization system and a data processing centre. Sensors provide the surgeon with feedbacks from the surgical site with real-time imaging, where the latter one updates the controller with new instructions for the robot by using the computer interface and some joysticks.

Neuro-surgical robotics provide the ability to perform surgeries on a much smaller scale with much higher accuracy and precision, giving access to small corridors which is completely important when a brain surgery is involved.

EDUCATION

MOOCS:

- Human-Centered Design: an Introduction
- UX Design: From Concept to Prototype
- Understanding User Needs
- Visual Elements of User Interface Design
- UX Design Fundamentals

EXTERNAL MANUALS FOR MORE INFORMATION:

- Human-Computer Interaction Fundamentals
- The evolution of Human-Computer Interaction
- A Missing Link in the Evolution of Human-Computer Interaction

BIBLIOGRAPHY

- Dix, A. (2019). *Human-Computer Interaction (HCI).*
  Retrieved from
  https://www.interaction-design.org/literature/topics/human-computer-interaction

- Yellin, B. (2018). *Human-Computer Interaction and the User Interface.*
  Retrieved from
  https://education.emc.com

- Grudin, J. (2009). *AI and HCI: Two Fields Divided by a Common Focus.*
  Retrieved from
  https://pdfs.semanticscholar.org

- Biseria, A., Rao, A. (2016). *Human Computer Interface-Augmented Reality.*
  Retrieved from
  http://ijesc.org/

- Whinston, A., Parameswaran, M. (2007). *Social Computing: An Overview.*
  Retrieved from
  https://pdfs.semanticscholar.org

- Moreno, A., Selfah, A., Capilla, R., Sánchez, M. (2017). *HCI practices for building usable software.*
  Retrieved from
  https://www.researchgate.net

- Karray, F., Alemzadeh, M., Abou, J., Nours, M. (). *Human-Computer Interaction: Overview on State of the Art.*
  Retrieved from
  http://s2is.org/

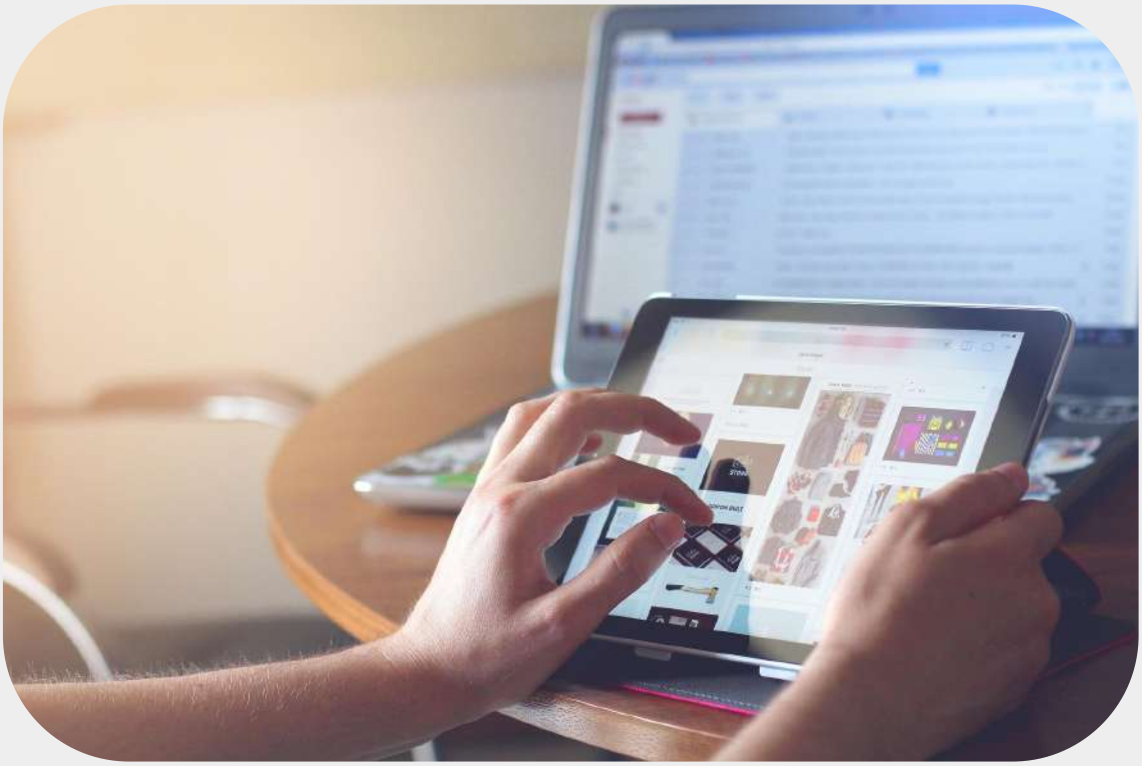
SELF APPRAISAL

★ After reading this text have I a clear idea of what Human-Computer Interaction is?

★ What tools am I used to?

★ Do I know the benefits that Human-Computer Interaction can bring to my company?

★ Can I recognize the advantages and disadvantages of its implementation for my company?



INTRODUCTION TO THE INDUSTRIAL REVOLUTION 4.0