Ambient Assisted Living Joint Programme

Project full title:

Development of a non-invasive CAPactive sensor oral MOUSE interface for the disabled elderly (CAPMOUSE, AAL-2008-1-203)

Deliverable:

D1.4: Market Analysis

AAL project number:
Project starting date: 15/06/2009
Project duration: 30 months

Coordinator: Tomas Brusell
Coordinator e-mail: tomas.brusell@brusell-dental.com
Project website: http://www.brusell-dental.com/aal

Contributors: BD, HMC
Planned delivery date: 28.02.2011
Actual delivery date: 25.04.2011

Table of Contents

1. Human Interface technologies

The user interface (also known as Human Computer Interface or Human Machine Interface (HMI)) is a combination of parts, ways and procedures by which people—the users—interact with the system—a particular machine, device, computer program or other complex tool both cognitively and physically; it is the layer between a product and a human.

To work with a system, users have to be able to control the system and assess the state of the system. For example, when driving an automobile, the driver uses the steering wheel to control the direction of the vehicle, and the accelerator pedal, brake pedal and gearstick to control the speed of the vehicle. The driver perceives the position of the vehicle by looking through the windscreen and exact speed of the vehicle by reading the speedometer. The user interface of the automobile is on the whole composed of the instruments the driver can use to accomplish the tasks of driving and maintaining the automobile.

The user interface of a mechanical system, a vehicle or an industrial installation is sometimes referred to as the Human-Machine Interface (HMI). Although the term HMI can refer to any type of device that allows a person to manipulate a machine or process, the term as used today, refers to the hardware and software that serve the specific role as the user's interface for controlling a system or a device.

The term user interface is often used in the context of computer systems and electronic devices. In the context of computer science, the user interface refers to the way a program presents itself to a user, what it looks like on the screen, the commands it puts at the user's disposal, or the level at which the user can communicate with the program. User interface design, research and development is central to the field called human-computer interaction (HCI).

Research in human-computer interaction started in the 1960s. One of the pioneering studies was presented by Douglas Engelbart of Stanford Research Institute (SRI), who introduced the now-ubiquitous computer mouse at the Fall Joint Computer Conference in San Francisco in 1968.[1] The mouse was further developed at Xerox Palo Alto Research Center (PARC). Several other important design innovations originated from SRI and PARC, including the graphical user interface (GUI) and direct manipulation.[2] These design principles were applied to the interface of Xerox Star computer, which was introduced in April 1981. The design effort in developing this particular computer was without parallel. Several hundred experiments were performed with human test subjects to validate design details. Largely, today's computer users still rely on this 30 year old design based on a graphical user interface which is mostly operated with a hand controlled peripheral pointing device such as a mouse.

During the three decades, computers have revolutionized the way we work, communicate, and participate in all forms of activities, including leisure. We have found new ways of collecting data, solving problems, and making decisions. In doing so, we have been forced to adopt work methods, which fit the requirements of computers and developers, but not necessarily the requirements of users. The adoption of computers has been a gradual process, and it may seem that we have accepted the evolution of computer interaction interfaces without much reflection or criticism.

Undoubtedly, interface design in the future will move toward designs where the right input modality for the job is selected, rather than the modality at hand being forced to fit the job. The GUI combined with its traditional input modalities is a powerful and adaptable interface form, and the software engineers and designers have tended to approach any given interface implementation with a GUI solution. However,

[1] http://sloan.stanford.edu/MouseSite/

[2] graphical user interface. (2009). In Encyclopædia Britannica. Retrieved January 19, 2009, from Encyclopædia Britannica Online: http://search.eb.com/eb/article-93005

careful analysis of the needs of the user in a given interface environment may suggest that another (nontraditional) form of user interface is required.

1.1 Hand controlled interface technologies

Hand controlled human interface technologies are today by far the most common means for human-computer or human-machine interaction. Hand controlled interface technologies can be separated into two major segments – text-entry devices and pointing devices. Manual text-entry devices are commonly recognized as computer keyboards. Pointing devices include computer mice, trackballs, joysticks, touchpads or trackpads, touchscreens and various haptic devices.

Since the introduction of graphical user interfaces for human computer interaction the main input devices have been the keyboard and pointing devices.

Computer keyboards are used for inputting alphanumeric data into a computer or other electronic device. Pointing devices are generally used as a mode of communication between users and computer interfaces. The mouse, the trackball, the touchpad and the joystick are all devices used for pointing at locations or objects on the graphical computer interface.

The computer peripherals market revenues reached $61.2 billion in 2009 and are expected to grow to $78.8 billion by 2014. The combined global market for keyboards and pointing devices was estimated to be worth $2.45 billion in 2010 – $1.51 billion for mice, trackballs and other pointing devices and $940 million for keyboards and desktops (mouse and keyboard combined). The market for computer touchpads and trackpads was estimated at $810 million in 2010. Touch screen panel market was worth $2.45 billion in 2009 and was estimated to generate revenues of $3.6 billion in 2010.

Leading computer keyboard and pointing device manufacturers are Logitech (headquartered in Romanel-sur-Morges, Switzerland) Microsoft (based in Redmond, Washington, USA) and Chicony (based in Taipei, Taiwan). Together these three companies are estimated to hold 75% of the global peripheral market. U.S. company Synaptics Inc is the market leader for integrated touchpad and trackpad sales with 70% of the global market. The touch screen industry is notoriously fragmented. There are more than 170 companies that claim to be manufacturing some parts of a touch screen. There are no clear leaders, with the largest companies – Taiwanese producer Chung Hua EELY Enterprise Group and U.S. company Elo Tyco accounting each for only around 10% of the industry.

Keyboards

Computer keyboard is a typewriter keyboard, which uses an array of buttons or keys, to act as mechanical levers or electronic switches used for direct input into computers. The design of modern computer keyboards is fairly standardized – traditional alphanumeric keyboards have square keys (side length 19.05 mm), and have a key travel of at least 3.81mm. Desktop computer keyboards, such as the 101-key US traditional keyboards or the 104-key Windows keyboards, include alphabetic characters, punctuation symbols, numbers and a variety of function keys.

Alternative keyboards such as adjustable split and split-fixed designs have been on the market for over two decades and are currently offered by leading peripheral companies such as Logitech and Microsoft.

Fixed-split keyboards have been shown to reduce awkward postures, muscle strain and overall pain and discomfort as well as improving the functional status of participants with pre-existing hand and wrist pain. Further, research has shown the benefit of the fixed-split keyboard design in reducing the incidence of new cases of carpal tunnel syndrome and other symptoms. Despite evident advantages widespread sales and use of split design keyboards have not exceeded those of conventional, straight keyboard designs.

Mice

Computer pointing devices are typified by the mouse, which has revolutionized the use of computers over the past decades. Douglas Englebart and colleagues invented the mouse in 1967 at the Stanford Research Institute. Forty years later, the mouse in relatively unchanged design persists because its properties provide a good match between human performance and the demands of graphical interfaces.

Modern computer mice make use of light-emitting diodes (LEDs) or low-power lasers and an imaging array of photodiodes to detect movement relative to the underlying surface. Traditional corded mice are today largely being replaced by cordless mice using radiofrequency technology for communicating with a computer. Today, the worldwide cordless mouse market alone has an estimated worth of over $1 billion, accounting for two thirds of total computer mice market.

Trackballs

A trackball senses the relative motion of a partially exposed ball in two degrees of freedom. Trackballs have a small working space (footprint), and afford use on an angled surface. Trackballs may require frequent clutching movements because users must lift and reposition their hand after rolling the ball a short distance. The buttons are located to the side of the ball, which can make them awkward to hold while rolling the ball. A trackball engages different muscle groups than a mouse, offering an alternative for users who experience discomfort when using a mouse.

Joysticks

A joystick is an input device consisting of a controlling stick or a handle that pivots on a base and reports its angle or direction to the device it is controlling. In consumer applications joysticks are most commonly used to control video games on personal computers and gaming consoles.

Most joysticks are classified as isometric devices. An isometric joystick is a force-sensing joystick that returns to center when released. Most isometric joysticks are stiff, offering little feedback of the joystick's displacement. The rate of cursor movement is proportional to the force exerted on the stick; as a result, users must practice in order to achieve good control of the operated virtual or physical device.

In addition to applications for computer gaming control joysticks have also become commonplace in many industrial and manufacturing applications, such as; heavy construction and field machinery, cranes, assembly lines, forestry equipment, mining trucks, and excavators.

Touchpads

Touchpads or trackpads are small, touch-sensitive tablets or surfaces often found on laptop computers. Touchpads use relative mode for cursor control because they are too small to map to an entire screen, but most touchpads also have an absolute mode to allow features such as sliding along the edge of the pad to scroll.

Touchpads support clicking by recognizing tapping or doubletapping gestures. Lastest touchpads (such as those found in Apple MacBook series of computers) also support multi-touch gestures that are performed with multiple fingers on the surface of the pad. Such actions can include zooming in and out, swiping to flip through documents or photos, rotating to adjust an image, etc. Accidental contact (or loss of contact) with pad surface can erroneously trigger such gestures. Like trackballs, the small size of touchpads necessitates frequent clutching, and touchpads can be awkward to use while holding down a button, unless the user employs his or her other hand.

Touchscreens and pen-operated devices

Touchscreens are transparent, touch-sensitive surfaces mounted on a display. Some touchscreens can only sense a bare finger; others can sense either a plastic stylus or a bare finger. Touchscreens offer user the ability to use finger or stylus to interact with devices instead of pointing devices or keyboards. Touchscreens have recently become especially widespread in mobile devices such as mobile phones, smartphones an tablet computers. The popularity of touchscreen devices has notably risen since the launch of Apple iPhone smartphone in 2007. Currently, most of the smartphone and tablet PC devices sold in the world also incorporate a touchscreens that can be controlled by bare fingers. It is estimated that worldwide touchscreen mobile device sales reached 362.7 million units in 2010

Haptic interfaces

Haptics generally describes touch feedback, which may include kinesthetic (force) and cutaneous (tactile) feedback. As a technical field, haptics research has been active for several decades. In the 1990s, haptics research expanded significantly with the availability of high fidelity, commercially available force feedback systems from companies such as SensAble Technologies, Inc. (USA), Immersion, Inc. (USA) and Force Dimension (Switzerland). Sophisticated haptic devices developed by these companies allow for up to 7 degrees of freedom (DOF) and are due to their high cost (starting from 10,000 USD) used only in professional settings.

Currently, much of the force feedback research focuses on developing practical control interfaces for application in fields such as entertainment, design, education, training, medicine and dentistry, and rehabilitation.

In the medical field, there are many application systems proposed and developed so far such as a surgery training system for inexperienced medical doctors, a surgery rehearsal system enabling surgeons to virtually experience uncommon and highly difficult cases before the real surgery, a distance medical care system opening the door for consulting and diagnosing in less-populated areas and isolated islands by remote doctors via networks and haptic devices.

In the industrial design field, 3D CAD systems used in conjunction with haptic control devices can provide intuitive 3D data entry functions and improve the effectiveness of the design

While the use of expensive haptic devices is generally limited to professional applications, some developers have recently launched fairly sophisticated devices priced suitably for consumer markets. One of the first mass-marketed consumer haptic interfaces was force-feedback PC controller Falcon launched by Novint Technologies, Inc in 2007 at a price of $249.95. Novint Falcon is a 3 DOF force feedback device. It tracks in 3 DOF (right-left, forwards-backwards, and up-down), and gives forces in those same degrees of freedom. Falcon is aimed mainly at the computer gaming market.

Limitations of hand controlled interface devices

The use of hand controlled interface devices carries some significant health risks. Use of a computer keyboard, computer mouse or other manually controlled input device can lead to persistent muscle aches, tendon inflammation, nerve compression, and sub-sequent disability that in some cases may be permanent. Computer users experience high rates of injury and disability, broadly termed repetitive strain injury (RSI). It is estimated that roughly 100 million individuals worldwide are predicted to suffer from some type of RSI or other computer associated health problems that are largely caused by the use of hand controlled computer interfaces.

The type of interaction expected when using a manually controlled hardware interface excludes a range of users who do not possess the required precision of arm and hand movements needed to operate these devices. It is estimated that 1.5-2% of the global population fall outside of the standard norm of access and require help with using traditional manually controlled computer interface devices or are unable to use computers due to lack or absence of upper limb mobility.

The range of alternative modalities available today creates an opportunity for new methods of human-computer interaction where computers adopt to human behavior, narrowing the gap between the device and the user.

1.2 Hands-free interface technologies

The past decade has seen the emergence of latest generation of commercial user interfaces that redefine how users communicate with computers. Research in human-computer interaction spans a wide range, from improving graphical user interfaces to conveying information to computers via new non-traditional interfaces such as speech recognition, gesture recognition, eye tracking and brain-computer interfaces. All of these non-traditional modalities have been already successfully implemented in some applications. Some of these modalities use rather mature technology, while others are on the cutting edge of the research domain. All of them have their strengths and limitations, mostly stemming from the level of technological advancement in the specific field.

All in all, when selecting an interface, it is important to remember that the main goal of the interface is to effectively support users in completing some required task. International Standard ISO 9241:11 specifies that usable designs should have three important attributes.[3] First, they should be effective, meaning that the user can successfully use the interface to accomplish a given goal. Second, the interface should be efficient. This means that the user not only can accomplish the goal, but can do so quickly and easily with a minimum of error or inconvenience. Finally, the interface should leave the user satisfied with the experience. This does not mean the user has to be happy, but it does mean that the user should have high confidence that the task was accomplished according to his intentions.

Table 1 Non-traditional interface technologies

Technology	Applications	Strengths	Current limitations
Eye tracking	Medical applications Marketing/advertising Ergonomics & usability research PC control	Solution for severely disabled patients Ergonomics Fast (eye movements are the fastest movements in human body)	Accuracy Requires frequent calibration "Midas touch" problem Expensive
Voice control	Consumer electronics Mobile devices PC control Automotive applications Avionics Defense/military Medical technology	Resilient as a communication side channel Allows control from distance Relatively affordable	Reliable with limited vocabulary Poor reliability in mismatched conditions (especially affected by noise)
Gesture control	PC control Consumer electronics Mobile devices Medical technology Marketing/advertising	Relatively affordable Natural communication channel Immersive (allows full body control)	"Midas touch" problem Ergonomics ("gorilla arm" effect) Accuracy & reliability
Brain-computer interface	PC control Consumer electronics Potentially all HCI applications	Translating brain activity directly into control signals, therefore potentially lowest latency	Poor information throughput Limited real-world usability

[3] http://usabilitynet.org/tools/r_international.htm#9241-11

		Solution for severely disabled patients	Requires training, minority of subjects exhibit little or no control

1.2.1. Eye tracking and eye control

Eye tracking is a general term for techniques whereby an individual's eye movements are measured so that both where a person is looking at any given time and the sequence in which their eyes are shifting from one location to another are determined. As the visual system is the primary channel for acquiring information, eye movements can provide a valuable insight into the way people receive and process different information through inner cognitive process known as visual attention. Because on its speed, accuracy and stability, it may also be advantageous to use the direction of the eye gaze as a control input.

The measurement device most often used for measuring eye movements is commonly known as an eye tracker. In general, there are two types of eye movement monitoring techniques: those that measure the position of the eye relative to the head, and those that measure the orientation of the eye in space, or the "point of regard".

There are four broad categories of eye movement measurement methodologies involving the use or measurement of: Electro-OculoGraphy (EOG), scleral contact lens/search coil, Photo-OculoGraphy (POG) or Video-OculoGraphy (VOG). These three methods are quite invasive and today the most widely applied eye movement technique, primarily used for point of regard measurements, is the video-based method that uses corneal reflection for eye tracking.

Corneal reflection systems record people's eye movement on video using a camera mounted on a headband, glasses or a standalone or integrated desktop eye tracking device. The light from the Infra- Red or Near Infra-Red Light-Emitting Diodes (NIR LEDs) creates a reflection on the user's eyeball (infrared light is used to avoid dazzling the user with visible light). As the eyeball is nearly perfect sphere, the reflection stays stable on the same position, independent of the direction of the gaze. Image recognition detects the center of the pupil and after calibration (i.e. looking at the corners of the screen), the direction of the gaze can be calculated using the vector from the corneal reflection point to the center of the pupil.

As the detection of a black pupil inside a brown iris is difficult because of low contrast, many eye-trackers use the white-pupil-method – if infrared light illuminates the eye, the eye background (retina) reflects the light and causes a white pupil to appear for the infrared camera device. To allow free head movement in front of the display, without use of a chin rest, the eye tracker is usually combined with a video-based head tracker.

Human sight has a visual field of about 120 degrees, encompassing three degrees of visual acuity: foveal, parafoveal, and peripheral vision. We primarily take in visual data from the outside world through the fovea, which is the central region of the retina responsible for the eye's sharpest vision. We move our head and eyes to focus the fovea on objects of interest that we want to see. The main measurements used in eye-tracking research are fixations, smooth pursuits and saccades:

- Saccade – A saccade is a rapid eye movements used in repositioning the fovea to a new location in the visual environment.
- Fixation – Fixations are eye movements that stabilize the retina over a stationary object of interest.
- Smooth pursuit – Pursuit movements are involved when visually tracking a moving target.

Saccades range in duration from 10 milliseconds to 100 milliseconds, which is a sufficiently short duration to render the executor effectively blind during the transition. Saccades are considered manifestations of the desire to voluntarily change the focus of attention. The world is primarily perceived visually through fixations. The brain virtually integrates the acquired visual images through successive fixations. Fixations naturally correspond to the desire to maintain one's gaze on an object of interest. Similarly, pursuits are used in the same manner for objects in smooth motion.

Eye tracking devices

Eye tracking device is an instrument that is capable of capturing data about saccadic activity, fixations and smooth pursuits of the foveal area. Commercial eye tracking setups come in a variety of configurations – they can be head-mounted, some function remotely as standalone or integrated devices and automatically track the head during motion. Some of the standalone devices give users the freedom to move their head during tracking; others require the head to be stable (for example, with a chin rest).

Head-mounted eye tracking systems are mostly used in mobile environments where full freedom of movement is needed to collect eye movements and point of gaze information during the performance of natural tasks.

The table mounted integrated eye trackers are ideal for all forms of eye tracking studies with stimuli that can be presented on a computer screen. Integrated eye tracker may appear no different from a common flat panel display, and that is intentional. Unlike a typical monitor, however, a camera and infra-red LED optics are embedded beneath the LCD flat panel.

Standalone eye trackers are most flexible point-of-gaze systems that can be used for eye tracking studies relative to any surface. These trackers are used for visual testing of physical brochures, magazines, products and shopping shelves, or of similar scenes that are best presented on a projection screen or TV. Scientific uses include psychological studies where stimuli need to be presented on a projection screen and infant studies that involve real world objects. Most standalone tracker come in form of a large web camera.

Standalone eye trackers and integrated trackers can also be used for gaze-based human computer interfaces (HCI) that allow the computer to be completely controlled by the eyes.

While most commercially offered eye trackers today work at sampling rates of 50/60 or 120 Hz, some video-based high speed eye trackers run at 250, 500 or even up to 2000 Hz, which is needed in order to capture the detail of the rapid saccadic eye movements, ranging in duration from 10 ms to 100 ms, during reading, or during medical studies.

Figure 1 Head-mounted-, standalone- and integrated eye tracking devices (EyeLink II from SR Research and Tobii X120 and Tobii T120 shown here)

Applications

A wide variety of eye tracking applications exists, which can broadly be described within two broad categories, termed here as diagnostic or interactive. In its diagnostic role, the eye tracker provides objective and quantitative evidence of the user's visual and attentional processes. As an interface modality, the eye tracker serves as a powerful input device that can be utilized by a host of visually mediated applications.

In general, in their diagnostic capacity, eye movements are simply recorded to ascertain the user's attention patterns over a given stimulus. Diagnostic applications are distinguished by the unobtrusive use of the eye tracking device. In some cases (marketing/advertising or usability research), it may even be desirable to disguise the eye tracker so that potential subjects are not aware of its presence. Furthermore, the stimulus being displayed may not need to change or react to the viewer's gaze. In this scenario, the eye tracker is simply used to record eye movements for post-trial, off-line assessment of the viewer's gaze during the experiment. In this way, eye movement data may be used to objectively corroborate the viewer's point of regard, or overt locus of attention. Diagnostic eye tracking techniques are applicable (but not limited) to the following fields and applications:

- Medical applications – psychology and research of neurological disorders. Eye tracking technology is extremely beneficial for assessing cognitive defects of people who are suffering from neurological disorders such as stroke, traumatic brain injury and blast injury (TBI), Alzheimer's disease, multiple sclerosis, amyotrophic lateral sclerosis.[4]
- Marketing/advertising – By measuring eye fixations eye tracking technology allows researchers to measure the advertisements ability to gain and hold customers attention and break through surrounding "clutter". Eye-tracking has also been used in retail settings to learn more about shoppers' viewing habits.[5]
- Ergonomics and usability research – Eye tracking has given researchers major insights into how users consume Web pages–such as where people start browsing on a page, whether they have banner and text link blindness, where users look for navigation, how they react to different text types, relative attention allocated to text vs. pictures and more.[6] Eye tracking allows to observe how people interact with different man-machine interfaces, control, navigation and entertainment

[4] http://www.tto.ohiou.edu/upload_files/file/Hallowell08018.pdf
[5] http://www.wired.com/gadgets/miscellaneous/news/2007/06/eyetracking
[6] http://www.useit.com/eyetracking/

systems and other instrumentation in such vehicles as aircraft, cars, trains, ships, as well as in other areas, such as air traffic control, power plant control, operating rooms.

Equipped with an eye tracker as an input device, an interactive system is expected to respond to, or interact with the user. Interactive applications are therefore expected to respond to the user's gaze in some manner.

The most published interactive eye tracking application is one where the user's gaze is used as a pointing device. Eye movements can be measured and used to enable an individual actually to interact with a computer interface. Users could position a cursor by simply looking at where they want it to go, or click an icon by gazing at it for a certain amount of time or by blinking. The first obvious application of this capability is for disabled users who cannot make use of their hands to control a mouse or keyboard. This type of ocular interaction can also be considered but one of a set of multimodal input strategies from the system's point of view. Using gaze to aid communication has also been explored in multiparty computer-supported collaborative work systems. Besides being used as a pointing device, knowledge of the user's gaze may be utilized to alter the display for speed-up purposes, as may be required in the rendering of complex virtual environments. Interactive eye tracking techniques are applicable (but not limited) to the following fields and applications:

- Human– computer interaction (HCI)/consumer electronics/embedded devices – Eye tracking technology can be used for mediating human computer interaction - users can control a computer and make things happen by just looking at it. Eye control can be used as sole interaction technique or combined with keyboard, mouse, physical buttons and voice. This application of eye tracking has potential uses in personal computing, the automotive industry, medical technology, education.
- Visual displays, and computer graphics – Virtual reality environments can be controlled by the use of eye movements. Eye movements seem to be the ideal tool in such a context, as moving the eyes to span long distances requires little effort compared with other control methods. Some techniques alter a display depending on the point of regard. Some large-display systems, such as flight simulators channel image processing resources to display higher quality or higher resolution images only within the range of highest visual acuity (i.e., the fovea) and decrease image processing in the visual range where detail cannot be resolved (the parafovea).[7]

Limitations

Eye tracking has several human factors and technology limitations. The human eye fixates visual targets within the fovea, which fundamentally limits the accuracy of eye gaze tracking to 0.5 degree of the field of view. The eye constantly jumps around constantly, moving rapidly in saccades between brief fixations points, so a high sampling rate and intelligent filtering is necessary to make sense of eye tracking data.

Another significant problem is accuracy of the eye tracker. Video-based eye trackers need to be fine-tuned to the particularities of each person's eye movements by a "calibration" process. Following initial calibration, eye tracker accuracy may exhibit significant drift, where the measured point of regard gradually falls off from the actual point of gaze.

If one uses eye gaze to execute commands, the so-called Midas touch problem results, because the user cannot glance at a command without activating it. That is, unlike a mouse with which a user signifies activation of an object by pressing a mouse button, with gaze pointing everything that a user looks at is potentially activated. Frequently applied solutions for the Midas touch problem are the concept of dwell time (of about 150–200 ms) where user activates functions simply by a prolonged fixation of an object and

[7] http://www.alexpoole.info/academic/Poole&Ball%20EyeTracking.pdf

blinking which assumes selection of the gazed object when the user blinks. Prolonged fixation means that the interaction is slower since the user has to sit through the dwell-time, blinking solution is obviously prone to inadvertent clicking due to inadvertent blinking by a user. Another solution for Midas problem is to use eye movements in combination with other input devices to make intentions clear. Speech commands can add extra context to users' intentions when eye movements may be vague, and vice versa.

Companies offering eye tracking solutions

Tobii Technology AB
http://www.tobii.com

Tobii Technology AB, based in Stockholm, Sweden, is a world leader in hardware and software solutions for eye tracking. The company was founded in 2001 and has continuously shown very rapid year-to-year growth. With a revenue growth of 2404 % between 2003 and 2007, the company is awarded a top ten position in this year's "Sweden Technology Fast 50" - Deloitte's annual ranking of the fastest-growing technology companies in Sweden.

Tobii's eye tracking products are used within the scientific community and in commercial market research and usability studies, as well as by disabled people as a means to communicate. Company offers eye trackers integrated into TFT monitors, standalone eye tracking units and software for analyzing eye gaze data. Tobii Technology also offers off-the-shelf or customized original equipment manufactured ("OEM") eye tracking components for integration into various devices. Tobii's eye tracking technology provides eye-gaze point, eye/head position and pupil size data.

Applied Science Laboratories (ASL)
http://www.a-s-l.com

ASL has been a pioneer in the examination of human eye movement and pupil dynamics for over 30 years. Founded by a M.I.T. scientist in 1962, ASL developed the first video based eye tracker in 1974. Company is located in Bedford, Massachusetts, Unites States.

Company's main product is the EYE-TRAC®6 Series control unit which can be configured to work with a wide range of head mounted eye tracking systems, desktop/remote mounted systems or in magnetic resonance imaging (MRI) environment. Company also offers software solutions for analyzing almost any eye fixation or eye movement parameter.

SensoMotoric Instruments GmbH (SMI)
http://www.smivision.com

SensoMotoric Instruments (SMI) is a Berlin, Germany based producer of dedicated computer vision applications, eye & gaze tracking systems and OEM solutions for a wide range of applications. Founded in 1991 as a spin-off from academic research, SMI was the first company to offer a commercial, vision-based 3D eye tracking solution. Company offers different standalone and head mounted eye tracking units and systems for the fMRI environment. SMI also produces software for conducting gaze tracking experiments and visual stimulus presentations as well as eye tracking data analysis software.

SR Research Ltd.
http://www.eyelinkinfo.com/

SR Research develops and markets head mounted and standalone EyeLink eye tracking systems, optional equipment for tracking systems and data analysis software. Standalone EyeLink 2K system provides a 2000 Hz sampling rate and is currently the fastest eye tracking device available on the market. Head mounted EyeLink II has the highest resolution (noise-limited at <0.01°) and also fastest data rate (500 samples per second) of any other head mounted video-based eye tracker currently on the market.

Company is located in Ontario, Canada.

1.2.2. Voice control

The term "voice control" is used to refer to applications in which a person uses spoken voice commands to control functions of a computer or other machine. Voice control applications make use of automatic speech recognition (ASR) technology what seeks to understand patterns or information (in form on words) in human speech which can be converted to machine-readable input. In that way speech recognition enables to design interfaces in which the input is primarily or exclusively speech.

Speech recognition technology has improved steadily, if not spectacularly, over the last few decades, but it is still not possible to accurately convert anyone's voice talking about any subject. Speech recognition and voice control can succeed for a limited vocabulary, such as speaking the name of a person from one's contact list to place a cell phone call; however, error rates increase as the vocabulary and complexity of the grammar grows, if the input signal is poor, or if users employ "out-of-vocabulary" words. In short, voice control is best suited to functions requiring selection between a discrete set of choices, rather than to selection of continuous quantities or to positional control.

In general, it is effective to use voice command applications for situations when speech can enable a task to be done more efficiently, such as when a user's hands and eyes are busy doing another task or where hands and eyes might otherwise be better occupied – repairing equipment, for instance, or sorting through inventory, or managing some other highly mobile, hands-on assignment. And not every environment is equally amenable to direct-manipulation input. Likewise, speech input is useful when no keyboard is available for text entry, or when a user has a physical disability that limits the use of his or her hands (a category which includes not just the severely disabled but the increasing legions of computer users with repetitive strain injuries), or if a user is just not comfortable typing. Some people, children for example, don't have the dexterity, or even the hand size for keyboards and pointers.

Speech is, further, very resilient as a side channel, making it the ideal mode for so called secondary-task interfaces. These are interfaces for systems or functions when the computational activity is not the primary task (for instance, supporting an installation where the user is busy handling equipment but still needs to check part numbers, follow procedures and the like). The issue of safety can become quite important for secondary-task interfaces when the primary activity is potentially dangerous, such as driving a car or piloting an aircraft.

Technologically, speech recognition by machine is deeply interdisciplinary by nature. Disciplines that have been applied to one or more speech recognition problems include signal processing, acoustics, pattern recognition, communication and information theory, linguistics, physiology, computer science and psychology.

Broadly speaking, there are three approaches to speech recognition, namely: the acoustic-phonetic approach, the pattern recognition approach and the artificial intelligence approach. The most used method today is the pattern recognition approach in which speech patterns are used directly without explicit feature determination (in the acoustic-phonetic sense) and segmentation. As in most pattern-recognition approaches, the method has two steps-namely, training of speech patterns, and recognition of patterns via pattern comparison. Speech "knowledge" is brought into the system via the training procedure. Pattern recognition approach has gained popularity in because it is relatively widely used and understood, because of its robustness and invariance to different speech vocabularies, users, feature sets, pattern comparison algorithms and decision rules and because of its high performance.

Voice control devices

On a personal computer or a handheld device (mobile phone, personal digital assistant, handheld gaming device) voice control via speech recognition software can generally be implemented without the need of a dedicated hardware setup. All that is needed is a signal input device (a microphone). Software programs that allow personal computers to be controlled via voice are readily available, for example Windows Speech Recognition, which is free and built into Microsoft Windows Vista operating system.[8] Similar voice recognition and control software is also included with Apple Mac OS X operating system.[9]

A number of companies are producing integrated circuits that allow speech recognition functionality to be added to embedded and electronic consumer devices. These chips usually integrate speech-optimized digital and analog processing blocks into a single chip solution that with an addition of a microphone can be used for speech recognition and product control.

Speech recognition in electronic devices can also be implemented as a standalone device. In this solution, a portable wireless device (the Speech Understanding Unit, or SUU) contains the speech recognition and natural language understanding capabilities. The SUU can be a separate device, or part of another device such as a mobile phone or a headset. The SUU discovers which Bluetooth enabled devices are in the vicinity, and which words can be used to control them. This solutions eliminates the need to integrate speech recognition which requires much processing power and would substantially raise the price of consumer electronic devices.[10]

Applications

Modern speech recognition technology has developed to the point, where it allows voice control to be used in a wide variety of fields and applications:

- Consumer electronics/mobile technology – Speech recognition and synthesis can help to differentiate and improve the user interface in a variety of household products. Sensory, Inc. has developed specialized hardware chips that allow the addition of speech input and output to home appliances with very little incremental cost.[11] Speech input technology has been used in a number of prototype appliances, like for example a fully voice controlled oven.[12] World's largest computer peripherals producer Logitech has announced that the future models of its Harmony universal remote controls will include voice recognition features.[13] US based startup Vlingo has released voice control application for popular iPhone and Blackberry smartphones that allows users to speak into their device and have many popular applications carry out their respective functions. This includes dialing your phone, sending an email or SMS, creating and saving a memo or task, opening a web browser and performing a web search, composing a social-networking status message and more.[14] Also Google offers speech recognition service for the iPhone, which lets users enter queries into the Google search box by speaking.[15]

---

[8] http://www.microsoft.com/enable/products/windowsvista/speech.aspx
[9] http://www.apple.com/accessibility/macosx/physical.html
[10] http://www.voxi.com/docs/Voxi_Business_Case_Intelligent_Home.pdf
[11] http://www.sensoryinc.com
[12] http://www.loquendo.com/en/news/news-loquendo-indesit-sophius.htm
[13] http://www.pcmag.com/article2/0,1895,2126061,00.asp
[14] http://www.vlingo.com/
[15] http://googleblog.blogspot.com/2008/11/now-you-can-speak-to-google-mobile-app.html

- Interfaces for the disabled – Prototype voice-controlled human-computer interfaces have been designed that enable severely handicapped individuals to operate a computer.[16] Also, voice controlled wheelchairs have been developed and tested in research environments.[17]
- Automotive applications – Speech recognition and voice command systems are widely used in user interfaces in the automotive environment.[18] Research has shown the positive impact to safety and response times when people use speech recognition to control their in-car systems. Study conducted in 2008 revealed significant benefits when drivers were able to use their voice to select music, input addresses in navigation systems and dial the phone while driving.[19]
- Avionics – Control of aircraft systems by voice commands is one of the most attractive ways of reducing the distraction the pilot inevitably suffers when gazing at the cockpit instrumentation. The F-35 Lightning II military strike fighter that will enter service in 2011 will be the first U.S. fighter aircraft to employ a speech-recognition system. The speech-recognition system will enable F-35 pilots to control communications, navigation, and other aircraft subsystems via voice commands. As a result, fighter pilots are better able to focus on flying and the combat environment around them.[20] Voice input is also used in European fighter jet Eurofighter Typhoon.[21] Research conducted by UK Army Air Corps has shown that voice recognition system enhances aircraft safety by significantly increasing pilot's 'head out' time.[22]
- Defense/military applications – In military research environment speech recognition has been successfully implemented as a viable interface technology for military command and control (C2) software applications in Air Operations Centers (AOC). Speech recognition proved extremely useful for entering data, navigating menus and locating infrequently used information. The operational assessments showed that speech recognition can significantly reduce the time required for Air Tasking Order planning. Speech Input was also superior to Manual Input for subject operators performing in a simulated Uninhabited Aerial Vehicle (UAV) control station environment. Subjects' performance was better with Speech Input, both for the flight/navigation task and data entry tasks. Additionally, their subjective ratings indicated Speech was better than Manual.[23]
- Medical technology – Voice-control systems allow users to interact to and control computers or medical equipment via speech even from several meters away. Clinical trials of such medical systems are already proving highly successful by providing clinicians with a convenient, reliable and accurate means of hands-free control that allows them greater freedom to concentrate on their patients. These hands-free sterile human-machine interfaces are especially important in medical environments, as they allow the medical staff to control information without the contamination associated with traditional interfaces (mouse, keyboard) or touch screens that must be cleaned after each surgical procedure. Hands-free operation also allows users greater freedom

---

[16] Mu-Chun Su, Ming-Tsang Chung, "Voice-controlled human-computer interface for the disabled", Computing & Control Engineering Journal Volume 12, Issue 5, Oct. 2001 Page(s):225 – 230 http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=962674

[17] Masato Nishimori, Takeshi Saitoh and Ryosuke Konishi, "Voice Controlled Intelligent Wheelchair", SICE Annual Conference 2007, http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=4421003

[18] http://reviews.cnet.com/4321-13249_7-6605649.html?tag=mncol;txt

[19] http://www.reuters.com/article/pressRelease/idUS115891+08-Jul-2008+BW20080708

[20] http://www.af.mil/news/story.asp?id=123071861

[21] http://www.eurofighter.com/et_as_vt_dv.asp

[22] http://www.gizmag.com/go/7484/

[23] David T. Williamson, Mark H. Draper, Gloria L. Calhoun and Timothy P. Barry, "Commercial Speech Recognition Technology in the Military Domain: Results of Two Recent Research Efforts", International Journal of Speech Technology 1/2005

to concentrate on their patients. Philips Healthcare and British medical devices company Smith & Nephew both offer medical devices that can be controlled via voice commands.[24][25]

Limitations

While speech seems like it might be the ideal way to communicate with a computer anytime, there are situations when it is best not to use a speech user interface. In spite of focused research in this field for the past several decades, the understanding of the acoustic– phonetic characteristics of speech, speech variability and speech perception is far from complete, and robust speech recognition with high reliability has not been achieved. The speech recognition process may work well in clean conditions but degrades significantly in speaker and channel mismatch conditions.[26] Three of the biggest challenges for speech recognition performance come in form of speech ambiguity, limited acoustic information, and noise:

- Ambiguity – The ability of speech recognition software to correctly determine the spoken word string depends on finding a better match (between measurable characteristics of speech and acoustic models) along the correct path than any other path. The recognizer's biggest enemy is similar-sounding paths because they can easily be confused. In general, as the vocabulary and grammar get larger, the potential for ambiguity increases. Therefore in the use of critical tasks voice control should be confined to a small number of distinct words.
- Limited acoustic information – In general, shorter words and phrases are harder to recognize than longer ones. Longer words and phrases provide more acoustic information that can help in differentiating paths through the speech recognition model. For example the problem of differentiating the spoken city names "Boston" and "Austin." The only acoustic information that can distinguish between these two cities is the B in Boston. In general, the realization of B's are quite short. If the utterance was 0.75 seconds long, it is quite likely that the B, if it was there, played a small role in the overall pattern matching score. Alternatively, if the user said "Austin," any slight distortion at the beginning of the word (such as a lip smack) could easily have matched the B model. Clearly, differentiating "Boston, Massachusetts" from "Austin, Texas" is far easier than differentiating "Boston" from "Austin."
- Noise – Noise and distortion can come from numerous sources, such as channel noise, and environmental noise. It is well-known that the presence of noise severely degrades the performance of speech recognition systems and much research has been devoted to the development of techniques to alleviate this effect. Noise adds a random factor to the voice feature vectors so that they no longer represent the user's actual speech as accurately. They may not then match as closely the acoustic models along the correct path. The noise may also mask important features for matching. In general, anything that changes the feature vectors so that they are less like the data used to train the acoustic models will make recognition less accurate. The problem of noisy environments has been addressed by the use of array microphones that combine audio from multiple microphones. Through the process of beamforming, the outputs of the multiple microphones of an array are combined to form a single audio signal in which all but the dominant speaker's signal has been removed. At the same time beamforming can also reveal information about the position of the speaker.

The gap between human and machine recognition of speech remains large for many practical tasks. Current speech recognition accuracy is adequate for small vocabularies; humans, on the other hand, are highly capable in difficult conditions, e.g., understanding unknown speakers in noisy environments saying

[24] http://www.apptech.philips.com/healthcare/projects/voice_assisted_healthcare_systems.html
[25] http://global.smith-nephew.com/us/DIGITAL_OR_PLATFORM_15306.htm
[26] Serajul Haque, Roberto Togneri, Anthony Zaknich, "Perceptual features for automatic speech recognition in noisy environments", Speech Communication 51 (2009) 58–75

arbitrary utterances. Speech recognition functions well in ``matched conditions", where the system has been previously trained on all: speakers who would use the system, words that may be used, and possible recording conditions. The challenge remains for speech recognition to increase accuracy for mismatched conditions: the word error rate (WER), or the percentage of added or deleted words in the original speech, for spontaneous speech remains as high as 25% in many cases, while commercially acceptable WERs are often under 2%. The market for speech recognition is growing, but its rate of growth will increase greatly only when performance approaches that of humans.

Companies offering voice control solutions

Sensory, Inc.
http://www.sensoryinc.com/

Sensory, Inc. develops and markets speech recognition, speech synthesis and other speech technologies. Sensory sells both integrated circuit (IC) and embedded software solutions. Sensory's flagship RSC line of low-cost speech chips performs speech recognition, speech and music synthesis, speaker verification, and general purpose microcontroller functions.

Sensory's FluentSoft™ software development kit allows developers to incorporate technologies including speech recognition, continuous digit dialing, animated speech, and text-to-speech on many different platforms and operating systems.

The company offers its products for use in applications, such as telephones, home automation, toys, remote controls, automotive, security, and learning aids; automotive, wireless handset, PC/PDA, set top box, and home automation applications; and portable MP3 players, internet access devices, wireless handsets, automobiles, PDA's, and medical and industrial instrumentation.

Sensory, Inc. was founded in 1994 and is based in Sunnyvale, California. The company also has offices in Portland, Oregon; Hong Kong; Tokyo, Japan; and Vienna, Austria.

SRI International, Inc.
http://www.speechatsri.com/

SRI International operates as a nonprofit research and technology development company. The company engages in conducting client-sponsored research and development for government agencies, commercial businesses, foundations, and other organizations. It conducts engineering research, systems engineering and development, and integration services in many different areas.

SRI develops and licenses speech recognition engines and tools packaged as software development toolkits, which developers incorporate into their products and services. In addition to high performance speech recognition engines, SRI develops software-based systems that enable advanced functionality for a broad range of speech-based applications. These applications include data entry and form-filling, audio mining, free-form translation, spoken address capture, and content search by voice.

Company's consolidated 2007 revenues were approximately $450 million. Its main client is The United States Department of Defense.

The company was founded in 1946 as Stanford Research Institute and changed its name to SRI International in 1977. SRI International is headquartered in Menlo Park, California.

Nuance Communications, Inc. (Nasdaq: NUAN)
http://www.nuance.com/

Nuance Communications, Inc. offers speech-based solutions for businesses and consumers worldwide. It primarily delivers a portfolio of speech-enabled customer care solutions that enhance customer communications and automate customer services and business processes in information and process-intensive vertical markets, such as telecommunications, financial services, travel and entertainment, and government.

Nuance Communications' products include Dragon NaturallySpeaking speech recognition software for personal computers, Nuance Voice Control software for controlling mobile devices, Nuance Recognizer and Call Steering for creating automated customer service solutions, and VoCon embedded speech recognition engine for automotive, entertainment, mobile communications, and consumer electronics applications.

Nuance has grown considerably through mergers and takeovers that have brought together a number of former competitors, including Lernout & Hauspie, ScanSoft, SpeechWorks, and Dictaphone. Its main rivals in speech recognition market are Microsoft and Philips, along with speech technology specialist Intervoice.

Nuance's yearly revenue rose 44 percent to $868.5 million in 2008, from $602 million in the prior year. For the full fiscal year, Nuance booked a loss of $30.1 million, or 14 cents per share, compared with a loss of $14 million, or 8 cents per share, in fiscal 2007. In mid-January 2009 Nuance's market cap stood at $2.5 billion.

Nuance Communications was founded in 1992. It was formerly known as Visioneer, Inc. and changed its name to ScanSoft, Inc. in 1999 and then to Nuance Communications, Inc. in 2005. The company is headquartered in Burlington, Massachusetts.

1.2.3. Gesture control

A gesture may be defined as the physical movement of hands, arm, face or any other part of the body with the intent to convey information or command. Gesture recognition consists the tracking of human movement and interpretation of that movement as semantically meaningful commands for the computer or a mechanical device.

The use of gesture as a means of communicating with computers and machines is attractive for several reasons. A common motivation behind the analogy between nonverbal (gestural) communication and human–computer communication is that it allows for better, more natural and intuitive interaction. In addition, gesture recognition allows interacting with and controlling computers and machines when the use of traditional input (mouse or keyboard) devices may be hindered because of a disability.

Modern gesture control technologies can be classified into two categories, camera-based and movement sensor-based. Camera-based recognition or computer vision is most suitable for stationary applications, and often requires specific camera set up and calibration. The movement sensor based approach utilizes different kinds of sensors e.g. tilt, acceleration, pressure, conductivity, capacitance, etc. to measure movement. These sensors can be integrated into mobile phones, handheld remote control devices, wearable garments (gloves, shirts), attached directly to different parts of body, or implanted into the body (e.g. tongue). In addition to the camera and movement sensor based gesture control approaches, 2D patterns, such as characters or strokes drawn on a surface with a mouse or a pen can be used as an input modality. This category of input has been referred to as character recognition and gesture recognition. Unlike the camera-based and movement sensor-based approach that allow distant interaction without having to be within reach or in contact with a physical device in order to interact, character recognition requires a physical contact with a screen or display for controlling or manipulating virtual objects.

The most interesting potential in the field of gesture control is to make accessory free and wireless gesture interfaces, such as in virtual-reality and intelligent rooms, because the use of physical or wired gadgets can make the interface and gesturing tedious and less natural. Hence, when the aim is to make gesture interfaces unobtrusive to the user, camera-based computer vision is the best way to detect gestures. Computer vision is inherently wireless, and people have become accustomed to using standalone or inbuilt web-cameras. Most of the modern computer vision applications make use of compact and affordable 3D cameras, which can sense distance in real-time between an imaging sensor and the objects in its field of view. This guarantees that any data beyond a certain depth (such as background color or objects behind the user) can be ignored and will not affect accuracy. The input from the camera is analyzed by computer algorithms that interpret video frames and recognize human's activity and natural gestures.

Computer vision algorithms often consist of three parts: segmentation that spots relevant parts in the field of view, tracking that follows the movements, and classification that finds meaningful information. Segmentation detects the body parts in the images, such as relying on shape or skin color, or motion detection in video sequences for dynamic gestures. In order to move past this step, many solutions have simplified the segmentation by using infrared reflectors or other such markings that are easy to segment and track. Tracking follows a set of features when it moves with, for example, condensation or Kalman filtering. Classification detects the actual gestures using hidden Markov models, template or model, or fuzzy or Bayesian logic. A new approach to computer vision tracks the movement of a camera phone by analyzing its images of surroundings. These phone movements are used as gestures.

Commercially available sensor-based gesture capture systems employ a variety of strategies to track the position and orientation of multiple points. The most popular and probably also the cheapest method today is accelerometer sensing that is based on either tilt or orientation. At present, sensor based gesture control is often preferred due to of better accuracy compared to camera-based solutions and because in case of the camera-based computer vision algorithms, the necessary image processing can be slow, which creates unacceptable latency for applications requiring fast-moving control.

The future will likely bring many interesting breakthroughs in this gesture recognition field. Raw computer vision and other detection systems are expected to evolve into robust classifiers with less obstructive designs. This means that there will be no devices attached to the user, so that the user can feel completely immersed in the virtual environment that blends perfectly with the real world.

Applications

Gesture interfaces are popular wherever the interface requires some freedom of movement and an immersive feeling or applicational context or user's physical features impede the usage of traditional input devices.

Gesture recognition technology probably found its first commercial applications in the video game industry. The earliest gesture gaming device to go mass market was the Power Glove controller for the Nintendo Entertainment System (NES) in 1989. Wearable glove-like Power Glove was the first peripheral interface controller to recreate human hand movements on a television or computer screen.[27] It could track motion of the glove in three-space, finger position, and has a set of buttons/switches on the top of the wrist. Gesture control using computer vision has since been implemented in two gaming products by Sony – The Sony EyeToy for PlayStation 2 in 2002 and PlayStation Eye for PlayStation 3 console.

A variety of spontaneous gestures, such as finger, hand, body and head movements, are used to convey information in interactions among people. Gestures can hence be considered a natural communication channel, which has not yet been fully utilized in human–computer interaction. This situation is changing quickly as new applications of gesture recognition are brought to the market at a faster pace than ever. Some the fields where gesture control has been recently implemented include:

[27] http://www.abc.net.au/tv/goodgame/stories/s2248843.htm

- Personal computer/consumer electronics/mobile technology – Los Angeles-based Oblong Industries has created a gesturally driven G-Speak spatial operating environment.[28] G-Speak system allows users to control everything they see on a computer screen with their hand gestures. Instead of using just one hand to control a mouse, someone with G-Speak — after slipping on special gloves — uses both, and is able to communicate with a PC intuitively. G-Speak uses special "I/O bulbs" that can detect what a user does as well as emit light. The technology is being experimented with for uses such as air-traffic control or medical imaging, which could help surgeons.

  One of the top, global, mobile phone manufacturers Sony Ericsson has products on market that allow some of the functions to be controlled by hand gestures. Users can simply wave their hand over the phone to mute the ring tone and ignore a call on the W380 and the Z555 flip style handsets. Similarly a person using the phones as an alarm clock can activate the snooze function by waving across the phone.[29]

  Another world's largest mobile phone maker Samsung has patented a similar system of cellphone and mobile device control that responds to a users gestures, not on the screen as with the multitouch Apple iPhone but as recognized in the space around the handset. Using the cellphone's front-mounted camera, the software recognizes preset motions and translates them into on-screen control. For instance, pointing at the screen and then moving the finger could control a mouse or cursor, while rotating the wrist with the hand outstretched might flip an image or layer.[30]

  At The International Consumer Electronics Show (CES) 2009 in Las Vegas, Japanese electronics manufacturer Toshiba demonstrated a conceptual computer interface that uses hand gestures for control. With simple computer vision based motion sensing technology and a software interface, Toshiba hopes to open up applications for video games and other interactive media. Toshiba said it was looking into the technology for possible use on its Cell TV, a TV set based on the powerful Cell processor. The TV is due on the market in Japan in 2009.[31] The first concept devices that enable the TV unit to be controlled by hand gestures were designed already in the beginning of 90's.[32] In 2008 Toshiba was also the first PC maker, who introduced limited computer vision based gesture control features in their laptop computers. The Qosmio G55 is the first notebook that lets users control music and video playback as well as PowerPoint presentations, using just their hands. The webcam senses movements from 3 to 10 feet away and the Intel Quad Core HD processor interprets users hand motions in real time.[33]

  Gesture control using inbuilt three-axis accelerometers has already appeared in numerous consumer electronics, e.g. Nintendo Wii remote, and mobile device, e.g. Apple iPhone. Using Wii Remote and Nunchuk, games can be controlled by natural gestures which made the console very popular, especially for sports games.[34]

  Belgian company SoftKinetic is developing camera-based gesture recognition and control technology for computer games.[35] The company makes a software development kit (SDK) and API that makes the switch to gesture recognition simple for game makers. Essentially, SoftKinetic takes the data the camera is returning and translates it into straightforward commands for the gaming people. This makes it relatively easy and cheap for game developers to add gesture

[28] http://oblong.com/
[29] http://www.sonyericsson.com/cws/products/mobilephones/overview/z555i
[30] US Patent nr: US2008089587 (A1)
[31] http://www.engadget.com/2009/01/11/video-hands-flailing-wildly-with-toshiba-spatial-motion-interfa
[32] http://www.merl.com/publications/TR1994-024/
[33] http://blog.laptopmag.com/hands-on-with-the-qosmio-g55
[34] http://www.nintendo.com/wii
[35] http://www.softkinetic.net

recognition support into their games. In addition to SoftKinetic there are several other companies working on adapting gesture recognition technology into gaming environment.[36]

- Medical technology – Researchers at Ben-Gurion University of the Negev (BGU) in Israel have developed a hand gesture recognition system, tested at a Washington, D.C. hospital, that enables doctors to manipulate digital images during medical procedures by motioning instead of touching a screen, keyboard or mouse which compromises sterility and could spread infection. The system, named "Gestix", was tested during a brain biopsy procedure. In the in vivo experiment, this interface prevented the surgeon's focus shift and change of location while achieving a rapid intuitive reaction and easy interaction.[37]
- Marketing/advertising – Computer vision based gesture recognition technology allows retailers enhance their window displays by creating interactive experiences. People can control and interact with multi-media images and special effects, play games, manipulate advertisements or immerse their video image right into an ad or display, simply by waving their hand or pointing their finger in front of the retail window. The same gesture-control functionality can be delivered on a screen, panel, floor, wall, cash counter or tabletop.[38] Gesture control is also used for creating motion activated interactive information displays and for controlling basic applications that would be running in an electronic kiosk: a map or a product browser for example.
- Interfaces for the disabled – US Virginia-based start-up company Gravitonus has developed a device called the Alternative Computer Control System (ACCS) – a hands-free system that enables people who have suffered spinal cord injuries or other forms of paralysis to dictate basic computer functions entirely with their tongue. ACCS comprises of a control-module with a sensor-type movement tracking device and up to 9 tongue controllable buttons. It is fully placed in a person's mouth cavity and communicates with the head-set through a wire running from the mouth along the right cheek.[39]

  Researchers at Georgia Tech School of Electrical and Computer Engineering have created tongue-operated assistive technology, called the Tongue Drive system that allows individuals with disabilities to operate a computer, control a powered wheelchair and interact with their environments simply by moving their tongues. The system is operated using a small magnet, the size of a grain of rice that is attached to an individual's tongue by implantation, piercing or tissue adhesive. Movement of the magnetic tracer attached to the tongue is detected by an array of magnetic field sensors mounted on a headset outside the mouth or on an orthodontic brace inside the mouth.[40]

Limitations

Humans naturally gesture and point using their hands during verbal communication, which has motivated research into freehand gestures. A major challenge is to correctly identify and interpret when a gesture, as opposed to an incidental hand movement, starts and stops. This gesture recognition problem is referred to as The Midas touch problem.

Designers can expect spontaneous gestures from users all the time if the goal is natural immersive behavior. Therefore, the gesture recognition must be very tolerant. Otherwise, users would suffer rigid constraints to their behavior while in the system. Unfortunately, mainly due to technological aspects, designers are very often forced select rigid constraints as solutions.

[36] http://venturebeat.com/2007/12/14/gesture-recognition-technology-for-games-poised-for-breakthrough/
[37] http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2410001
[38] http://www.gesturetek.com/marketuses/digitalsignage_advertising.php
[39] http://gravitonus.com/hardware/accs/
[40] http://www.gatech.edu/newsroom/release.html?id=1960

This means that one of the most difficult tasks in computer vision based gesture control is finding a feasible gesture vocabulary that is easy for the user to remember and perform. Sign language is not convenient because the gestures are rather complicated, and sign languages differ according to the underlying vocal language. Limiting the vocabulary is important, and will benefit both users and designers.

A typical approach to defining an application's gesture vocabulary is to make it easy for the computer system's recognition algorithm to recognize the gestures. The result of this approach for finding gestures can be called a technology-based gesture vocabulary. This often results in a vocabulary where there is no particular meaning to the gesture itself or the association between gesture and meaning is arbitrary. These "clear-cut" gestures are often also stressful or fatigue producing for the user and nearly impossible for some of the user to perform.

The ergonomics problem associated with gesture control is called "gorilla arm" and it was initially observed by scientists in 1980s when conducting tests with touch-menu systems. Research showed that after more than a very few selections, the users arm begins to feel sore, cramped, and oversized — the operator looks like a gorilla while using the touch screen and feels like one afterwards. This shows that humans aren't designed to hold their arms constantly in front of them and at the same time making motions.[41]

Camera-based gesture recognition may also not be the best solution for controlling detailed virtual objects due to user's imprecision of pointing at a distance. Also, computer vision-based tracking systems often suffer from poor tracking reliability and sensitivity to variations in background illumination. Tracking reliability can be enhanced by controlling the appearance of the object so that it can be tracked unambiguously.

Companies offering gesture control solutions

GestureTek Inc.
http://www.gesturetek.com/

GestureTek Inc., founded in 1986, is the largest producer of camera-enabled gesture-recognition technology for presentation and entertainment systems. In the past 20 years, GestureTek has installed approximately 4,000 interactive multi-media displays, kiosks, exhibits, digital signs and advertisements, virtual gaming systems and other interactive surface computing solutions.

Company manufactures and markets gesture recognition hardware and software for a variety of applications. Products include computer interfaces for interactive advertising displays, gesture-control solutions for the health, disability and education markets, gesture-controlled mobile gaming and navigation engine, computer interface solutions for horizontal surfaces, desktop hand-tracking unit for mouse replacement and a gaming system for full-body gesture control.

GestureTek has also licensed its patents and technology to numerous consumer electronics companies. Sony licensed GestureTek patents for use with its EyeToy products running on the PlayStation2. Hasbro licensed has licensed gesture-recognition technology for its ION Educational Gaming System. Microsoft uses GestureTek technology with the Xbox Live Vision camera for Xbox 360. NTT DoCoMo has licensed technology for use in its 904i, 905i and 705i series handsets. Reactrix has licensed patents for use in its immersive signage and display business.

Company currently owns 7 patents in United States and has further 37 patent applications are in various stages of the patent process.

[41] http://catb.org/jargon/html/G/gorilla-arm.html

In the end of 2007 GestureTek received a strategic investment from Telefonica, a telecom company based in Spain. NTT DoCoMo, Japan's largest cellular operator, has also previously placed an investment in the company.[42]

Mgestyk Technologies
http://www.mgestyk.com/

Mgestyk Technologies is developing affordable camera-based gesture recognition technology for PC and gaming environments. Company has recently demonstrated its technology at numerous trade shows in North-America.

Mgestyk combines custom software for hand-gesture language processing with an affordable 3D camera by 3DV Systems. By using the Deep Media SDK, the patent-pending technology makes it possible to capture small hand movements and translate them into useful commands for controlling practically any Windows application.

Company doesn't market the product at the moment and has not specified when the device will come on sale. It has been announced that the pricing is expected to be within the range of a high end webcam or estimated around $150.

Oblong Industries, Inc.
http://oblong.com/

Oblong Industries, Inc., licenses human gesture-based G-Speak operating system to licensee/development partners in various industries. It offers a gesture recognition engine that parses and interprets positions of elementary targets, such as the user's hand and finger positions and orientations; and a motion tracking software system, which performs glove tracking. Oblong Industries, Inc. is based in Los Angeles, California.

Oblong is run by John Underkoffler who has worked on human machine interfaces at MIT's Media Lab for over two decades.

In December 13, 2007 company announced that it had raised around $8.80 million in its series A round of venture capital funding led by Foundry Group.

3DV Systems, Ltd.
http://www.3dvsystems.com/

3DV Systems, Ltd. develops and markets video imaging technologies that enable sensing motion and recognizing shape within a defined three-dimensional space. The company's products include Z-Cam, a real time depth camera; and DeepC, a chipset incorporating depth sensing technology. 3DV Systems, Ltd. was founded in 1997 and is headquartered in Yokneam, Israel.

In December 12, 2006 3DV Systems announced that it had closed a $15 million investment round, led by new investors, Kleiner Perkins Caufield & Byers, a leading U.S venture capital fund and Pitango Venture Capital, a leading Israeli venture capital fund. Other investors in the company include Israeli technology holding-company Elron and RDC Rafael Development Corporation Ltd.

[42] http://venturebeat.com/2007/12/12/gesturetek-receives-investment-from-telefonica/

1.2.4. Brain-computer interfaces (BCI)

Human brain activity produces electrical signals that can be detected from the scalp, from the cortical surface, or within the brain. A brain-computer interface (BCI), also known as a direct brain interface (DBI) or a brain-machine interface (BMI), is a system that provides a means for people to control computers and other devices directly with brain signals. BCI systems are not able to directly interpret thoughts or perform mind reading. Instead, BCI systems monitor and measure specific aspects of a user's brain signals, looking for small but detectable differences that signal the intent of the user. The aim of BCI interfaces is typically to emulate traditional interfaces by triggering keystrokes and cursor control.

Research on brain-computer interfaces spans many disciplines, including computer science, neuroscience, psychology, and engineering. BCIs were originally conceived in the 1960s, and since the late 1970s have been studied first of all as a means of providing a communication channel for people with very severe physical disabilities.

BCI systems can be broadly classified into two categories, depending on the placement of the electrodes used to detect and measure neurons firing in the brain.

Invasive systems interact with the brain directly, i.e., with electrodes that penetrate the brain or lay on the surface of the brain, while noninvasive systems interact with the brain indirectly by transmissions through the skull, e.g., electroencephalography (EEG), functional magnetic resonance imaging (fMRI), and magnetic sensor systems. Invasive techniques, such as implanted electrodes, usually provide better control through clearer, more distinct signal acquisition. Noninvasive systems have obvious advantages, but also the skull blocks a lot of the brains electrical signals, and it distorts what does get through. This means that the signals from the scalp are difficult to interpret, and they are very susceptible to external and internal interferences. Therefore, robust signal processing methods are needed to reliably detect and classify voluntarily generated brain patterns. Although state-of-the-art signal processing methods are applied in BCI research, the output of a BCI is still unreliable, and the information transfer rates are very small compared with conventional human interaction interfaces such as keyboard and mouse.

However, although BCIs exhibit serious drawbacks relative to other interfaces, very recent research developments suggest that BCIs may soon become much more powerful, flexible, and usable tools, providing improved communication to severely disabled users and opening new applications and new user groups. Although research to date has focused mainly on controlling output from the brain, recent efforts are also focusing on input channels.

BCI devices

The most ubiquitous BCI approach is the electroencephalogram (EEG), a recording of signals representing activity over the entire surface of the brain or a large region of the brain, often incorporating the activity of millions of neurons. An EEG can be recorded noninvasively (without surgery) from electrodes placed on the scalp, or invasively (requiring surgery) from electrodes implanted inside the skull or on the surface of the brain. Brain signals can also be recorded from tiny electrodes placed directly inside the brain cortex, allowing researchers to obtain signals from individual neurons or small numbers of co-located neurons.

Nearly all BCI studies using noninvasive sensors involve the use of silver or gold disk electrodes with conducting paste that are affixed to the skull using some type of head cap configuration to facilitate the application of the EEG electrodes. Limited progress has been made in improving these devices over the last two decades to rapidly and comfortably affix them to the skull of a BCI user. Head caps have been developed that aid in the measurement and placement of 64 to 256 EEG electrodes using the "International 10–20 grid system." The process of fitting individuals with EEG electrodes with head caps is time consuming, requires testing of individual electrodes for their impedance, and results in a system that is not comfortable or practical for routine BCI use. There is a need for development of "dry electrodes," which could be used without the preparation required for the current designs. There are promising "dry-type" electrode configurations that have been under development using carbon nanotube (CNT) electrodes and other dry-type sensor designs.

Figure 2 An electroencephalogram (EEG) based noninvasive brain-computer interface from Guger Technologies OEG (http://www.gtec.at)

Applications

As the BCI field matures, considerable interest has arisen in applying BCI techniques to real-world problems. The principal goal has been to provide a communication channel for people with severe motor disabilities, but other applications may also be possible. Researchers are focusing on applications for BCI technologies in several critical areas:

- Interfaces for the disabled – As brain–computer interfaces are systems that can translate brain activity directly into signals that control external devices, they can represent the only technology for severely paralyzed patients to increase or maintain their communication and control options.

  A team of researchers led by IDIAP Research Institute in Switzerland have created a brain-computer interface for the disabled within the EU-funded Mental Augmentation through Determination of Intended Action (MAIA) project. So far, the team has carried out a series of successful trials in which users have been able to maneuver a wheelchair around obstacles and people using brainpower alone. Same technology could be applied to artificial limbs to allow quadriplegics to pick up objects or unlock a door. By using the BCI to interact with computer systems, meanwhile, they could control the lighting in their homes, surf the internet, or change the channels on the TV.[43] The EU has also funded a MAIA follow-up project TOBI that will focus on BCI technologies which have the potential to become the basis for commercial systems in near future.[44]

[43] http://cordis.europa.eu/ictresults/pdf/factsheet/INF70100_ICT_%20Results_Fact_sheet_08_May_MAIA.pdf
[44] http://www.tobi-project.org/

Non-invasive BCI systems have been successfully integrated with the control architecture of the rehabilitation robot which consists of a 7 degrees of freedom (DOF) manipulator mounted on an electrical wheelchair.[45]

Research conducted in the University of Rome allowed patients with Spinal Muscular Atrophy type II (SMA II) or Duchenne Muscular Dystrophy (DMD) to move and communicate within surrounding environment via a system based on a software application. The system was controlled by the subjects' voluntary modulations of EEG sensorimotor rhythms recorded on the scalp; this skill was learnt even though the subjects have not had control over their limbs for a long time. Study effectively showed that people with severely disabling neuromuscular or neurological disorders can acquire and maintain control over detectable aspects of brain signals, and use this control to drive output devices.[46]

- Consumer electronics – World's largest toy maker Mattel has demonstrated a Mind Flex game (due 2009 fall), which uses brain-wave activity to move a ball through a tabletop obstacle course, at the 2009 Consumer Electronics Show in Las Vegas. The aim of the game is to concentrate hard enough to generate enough energy to power a fan which in turn causes a ball to levitate and move through a series of hoops. A special wireless headset reads user's brain activity, in a simplified version of EEG tests, and the circuitry translates it to physical action.[47]

  Similar game has also been released by Lucas Licensing, the owner of the Star Wars franchise. The game, called The Force Trainer comes with a headset that uses brain waves to allow players to manipulate a sphere within a clear 10-inch-tall training tower.[48]

  Players of electronic games are likely early adopters of BCI technology. They often wear headgear, enjoy novelty and technical challenges, have money and time available for peripherals and training, and are competitive and increasingly numerous. Currently the only gaming BCI system on sale is the Neural Impulse Actuator (nia) produced by OCZ Technology.[49] The nia is not seen as a replacement for traditional input methods, but merely a potentially powerful supplement for standard keyboard and mouse or a gamepad / joystick.

Limitations

Most BCIs have been used to control computer applications such as spelling devices, simple computer games, limited environmental control, and generic cursor control applications. BCIs are not well suited for controlling more complex details of demanding applications because of two reasons: (a) complex applications increase the mental workload of the user and can thus negatively affect BCI performance, and (b) complicated tasks require a number of subtasks, which, when controlled on a low-level basis, can be time consuming, fatiguing, and frustrating.

Brain-computer interfaces are notorious for poor information throughput. Although online BCI systems have exhibited performance slightly above 60 bits per minute (which roughly translates to selecting eight characters per minute from an alphabet), such performance is not typical of most users in real-world settings. Other recent publications have described information throughput between 30 and 60 bits per minute, but these still may not reflect average performance in real-world settings. There is often considerable variation across subjects and BCI usage sessions.

[45] Diana Valbuena, Marco Cyriacks, Ola Friman, Ivan Volosyak and Axel Graser, "Brain-Computer Interface for high-level control of rehabilitation robotic systems", Proceedings of the 2007 IEEE 10th International Conference on Rehabilitation Robotics

[46] Febo Cincotti, Donatella Mattia, Fabio Aloise, Simona Bufalari, Gerwin Schalk, Giuseppe Oriolo, Andrea Cherubini, Maria Grazia Marciani, Fabio Babiloni, "Non-invasive brain–computer interface system: Towards its application as assistive technology", Brain Research Bulletin, Volume 75, Issue 6, 15 April 2008, Pages 796-803

[47] http://www.engadget.com/2009/01/09/video-mattel-mind-flex-hands-and-heads-on/

[48] http://www.usatoday.com/life/lifestyle/2009-01-06-force-trainer-toy_N.htm

[49] http://www.ocztechnology.com/products/ocz_peripherals/nia-neural_impulse_actuator

Further, a minority of subjects exhibit little or no control. The reason is not clear, but even long sustained training cannot improve performance for some of these subjects.

There are also considerable limitations to the real-world usability of BCI systems. Both the hardware and software currently available for brain-computer communication is more suitable for experiments in the lab than for practical applications in real-world environments like user homes. The hardware needed for an EEG BCI requires a trained expert to precisely position the EEG cap, scrape the skin where each electrode will go, apply gooey electrode gel, further abrade the skin, and continue this process until all electrodes (often a large number of electrodes is necessary) produce a clean signal. This process is not painful, but not exceptionally pleasant either. After each BCI usage session, the cap and the user's hair must be washed. BCIs not only require an expert to help set up the necessary hardware but also to configure and adapt key software parameters.

Companies offering BCI solutions

Guger Technologies OEG (g.tec)
http://www.gtec.at/

Guger Technologies is an Austrian medical and electrical engineering firm that has become the first company to ship a commercial BCI system called g.BCIsys for research applications or controlling computers and other devices.

g.BCIsys uses mesh of electrodes and sensors positioned on a helmet-like cap combined with Guger's own biosignal acquisition and biosignal amplifier systems.

The g.BCIsys brain interface kit can be used with a standard Windows PC, or a Windows Mobile Platform device when paired with a lightweight biosignal recording system called g.MOBIlab. In either case, EEG processing, analysis and pattern recognition can be used to perform rudimentary tasks such as spelling and message composition, or to control a computer game. Its unique head cap for EEG electrodes design allows for some of the best signal-to-noise achievable in the business from wet electrode technology. In particular, the electrode cap design requires extra time for attachment of electrodes but achieves excellent signal-to-noise characteristics. For a variety of BCI uses, g.BCIsys is considered to be the source of one of the best head caps used in the field involving wet electrode recordings. The BCI solution costs about $5,000 with a 99 - 100% level of accuracy for "trained subjects."

At the 2007 CeBIT trade show in Germany Guger Technologies won the European Information and Communication Technology (ICT) 2007 Grand Prize for g.BCIsys system.[50]

NeuroSky, Inc.
http://www.neurosky.com/

California start-up company NeuroSky develops BCI technology, which can be incorporated into everyday products made by consumer electronics manufacturers.

NeuroSky technology is based on electroencephalography, or EEG, the measurement of the brain's electrical activity through electrodes placed on the scalp. NeuroSky's "dry-active" sensors don't require gel, are the size of a thumbnail, and could be put into a headset that retails for as little as $20, according to the company. NeuroSky's technology measures a person's baseline brain-wave activity, including signals that relate to concentration, relaxation and anxiety. The technology ranks performance in each category on a scale of 1 to 100, and the numbers change as a person thinks about relaxing images, focuses intently, or gets kicked, interrupted or otherwise distracted

[50] http://www.epistep.org/documentation/The%202007%20European%20ICT%20Prize%20Nominees.pdf

NeuroSky does not market its products directly to the public. Instead company licenses its technology to other companies to incorporate into that company's products. NeuroSky BCI sensor technology is the driving force behind the first commercial BCI toys Mattel Mind Flex and The Force Trainer that are expected to start selling in 2009. Companies which have publicly announced they are working with NeuroSky also include Sega Toys, Square Enix, Nokia and Musinaut.

Emotiv Systems, Inc.
http://www.emotiv.com/

Emotiv Systems, Inc. develops brain computer interface technology. It focuses on neuro-technology solutions for the electronic games industry. The company offers Project Epoc, an interface for human computer interaction that uses a set of 18 sensors to tune into electric signals naturally produced by the brain to detect player thoughts, feelings, and expression, and connects wirelessly with game platforms from consoles to personal computers to be controlled and influenced by the player's mind. Eventually Emotive will also build equipment for clinical use.

The 30-person company hopes to begin selling a consumer headset in 2009, but it would not speculate on price. The BCI system is expected to work with gaming consoles such as the Nintendo Wii, Sony PlayStation 3 and Microsoft Xbox 360.

Emotiv Systems, Inc. was founded in 2003 and is based in San Francisco, California, United States.

2. Comparative advantage of CapMouse technology

Past few decades have seen the emergence of new human computer interface technologies that have slowly started to challenge the traditional input modalities consisting mainly of a graphical user interface and a hand directed pointing device. The main catalyst for these developments has been the need for new and more natural ways of interacting with computers or other machines.

Among the most important factors in developing these new interfaces are the ease of usage and convenience in control. Operating the control device must be easy to learn and require minimum effort on the users' part. The device should be small, unobtrusive, low cost, and non- or minimally invasive.

Most common new nontraditional interfaces that have been applied in real-world applications include eye tracking, speech recognition, gesture recognition, and also brain-computer interfaces. All these technologies have their own limitations that currently decisively hinder their applicability in real-world situations that require fair amount of robustness and flexibility.

Tongue operated input interfaces form a subtype of gesture-based technologies that currently offer a number of significant advantages over other nontraditional interfaces making them highly usable in assistive control devices.

Since the tongue and the mouth occupy an amount of sensory and motor cortex that rivals that of the fingers and the hand, they are inherently capable of sophisticated motor control and manipulation tasks.[51] This is evident in their usefulness in vocalization and ingestion. The tongue can also move very fast accurately within the mouth cavity. It is thus a highly suitable organ for manipulating assistive or other control devices. Therefore, unlike voice control interfaces that are reliable when making selections between a discrete set of choices, tongue is perfect for selection of continuous quantities or for positional control. Additionally tongue muscle is similar to the heart in that it does not fatigue easily. Therefore, a tongue operated device has a very low rate of perceived exertion.

Another way to approach the real-world usability of a control interface is to look into the ergonomics and biomechanics of using the interface to ensure that a physically stressing gesture or body posture is avoided. The tongue muscle is not afflicted by repetitive motion disorders that can arise when few exoskeletal muscles and tendons are regularly used. These problems often surface when for example using hand gesture recognition interfaces. The tongue is also not influenced by the position of the rest of the body, which may be adjusted according to need and for maximum comfort. This shows that tongue operated interfaces are ergonomically superior to many other control technologies.

The most efficient technology to hands-free control of appliances would be a direct brain-computer interface where the control signals are recorded directly from the firing neurons. However, this technology is still in the infancy of its development and there are several fundamental scientific problems to be solved so it is unclear when, if ever, this technology becomes mature to be applied for control.

From economic standpoint, tongue control interfaces are considerably more affordable than eye tracking systems.

Although using human tongue as a source for control signals has great potential, most developed tongue controlled interfaces are invasive and obtrusive in nature. Examples include inserting a trackball, magnets, joystick, plastic palate with discrete control buttons, or a “sip-and-puff” controller straw into the mouth of

[51] E.R. Kandel, J.H. Schwartz, T.M. Jessell, "Principles of neural science“, 4th ed. McGraw-Hill, 2000

an individual. These devices are extremely intrusive, irritate the mouth, may impair verbal communication, present hygiene issues, can be difficult to operate, and are not very reliable.

CapMouse technology enables the control of a cursor or pointer on an electronic viewing screen by means of the tongue and/or other parts of the mouth cavity, without having to apply any manipulating object inside the mouth cavity. CapMouse is currently the only non-obtrusive and non-invasive tongue control interface that has been shown to be highly effective in real-world applications.

The advantage of the CapMouse technology also lies in the fact that by the means of a training process the device can potentially capture an unlimited number of tongue movements, each of which can represent a user command. A set of specific tongue movements can be tailored for each individual user and mapped onto a set of customized functions based on functional needs, user's oral anatomy, personal preferences, and lifestyle. CapMouse technology provides users more natural and unobtrusive control ability than the existing tongue control devices that are mostly switch based with only a limited number of direct selections.

3. Business model for CapMouse

The business model is at the heart of an innovation. A business model describes the means by which an organization generates revenues and profits from an innovation and resultant products. A business model also defines the product (to some extent), positions the company in the value chain (product commercialization process) and describes the mechanisms of getting revenues from the product over time.

Some questions that frequently rise when defining medical device company business model are:

- Should the company focus on research and development of services or products?
- Is the company going to manufacture or partner or outsource?
- Should the company build its own sales and marketing channels or should it co-operate with resellers or commercial partners for commercializing the product or service?

The key problem is deciding at what stage to sell the product or service so that the most amount of value could be extracted from the offering at available resources.

The medical device industry can be divided into integrated companies that have the capacity to take innovations from concept to market, and those that cannot or are not interested in doing so. Most start-ups belong to the second group of companies. They are typically built around innovators who develop technologies and then sell or license them to other companies. These, so called innovator companies are hence mainly situated in the R&D part of the medical device industry value chain.

Large integrated medical equipment companies have over the past decade gone through a very significant restructuring of their business models. These big companies are now realizing that they are primarily sales and marketing and distribution organizations.

A lot of their technology is currently in-licensed or acquired and further development work has been outsourced and has been going down through the supply chain. Companies have let their suppliers undertake the kind of R&D that's necessary to get a finished product. This has shifted the medical device industry R&D budget from the major corporation down to the next level.

At the R&D end of the value chain, there are national labs, academic and clinical research centers, and emerging companies—all funded by a variety of federal, grant, and private resources. These institutions license and sell developed medical device technology and patents before entering the next part of the value chain i.e. manufacturing and marketing. Some device companies form alliances with larger, established device companies for market access. Licensing between innovators and big players is beneficial for the whole industry, as it gives new technologies and products better commercial prospects thanks to access to large development and marketing know-how.

Considering the following points:

- Large part of the value of the CapMouse technology can be extracted from the R&D stage as Oral Mouse technology belongs to the category of less risky medical devices and therefore there is no regulatory risk in the latter part of the value chain associated with it;
- CapMouse device, as a later-stage and real-world proven technology, is less risky, and therefore more valuable, than early-stage and unproven technology i.e. the licensing terms are potentially more favorable;
- CapMouse intellectual property is well protected;
- Medical device manufacturing is capital intensive and the building of strong sales and marketing capability is time-consuming and expensive…

…it can be concluded that the first appropriate business model for Brusell Dental, as an innovator company, would be to license the CapMouse technology to a larger and established market participant after the R&D objectives for the device have been met. This commercial partner would be solely responsible for the manufacturing and marketing and distribution of the CapMouse device. This business model allows Brusell Dental to bring CapMouse technology to the market in the shortest time frame by focusing its resources on the R&D part of the value chain and leaving the management of all the other stages to the prospective buyer, who has significant experience and resources for executing these downstream activities.

The second option for Brusell Dental is to outsource the manufacturing activities to a contract manufacturing company after completing the R&D phase and sell the finished product to participants who have access to the end-user markets. As CapMouse is a universal peripheral device, this business model would allow Brusell Dental to enter a large number of markets at once without needing to sign licensing agreements with manufacturers in every single of those segments. Outsourcing product manufacturing to specialty companies allows Brusell Dental to avoid the cost of constructing new facilities and the purchase of high-priced equipment, preserving capital. Outsourcing also eliminates the need to hire employees who have the requisite technical expertise.

In case of licensing agreements, the value of the property can be captured through different contract structures. By looking at similar technology transfer deals, one can expect CapMouse technology license agreement to include following financial terms:

- Technology transfer, access fees, or up-front payments – This lump sum payment, due at signing, recognizes the investment made to date by the licensor both in developing know-how and the technology itself, and also includes consideration that some licensor preparation and effort may be necessary to allow access to the technology.
- Patent prosecution and maintenance fees – The licensee can be asked to pay legal and patent office fees for maintaining the patent.
- Royalties – Royalties allow the licensor to participate in the net revenues of the out-licensed product with a certain percentage. This percentage will vary by the state of the patented technology at execution of the license agreement and will typically be based on industry or market rates prevailing at that time. A rule of thumb sometimes used by many licensors is the 25% rule – it is often accepted that a royalty that is equal to 25% of the expected pre-tax net profit is a fair rate. The royalty rate in the license agreement will then depend on the market forces of each particular product. For example, if the licensee has profit margins of 60%, the royalty paid to the licensor should be 15% (¼ of 60%) of sales revenues. If profits of 4% are expected, the royalty rate should be around 1% of net revenues. Many licensing agreements include adaptive royalty rates that change as the annual revenues reach certain thresholds.

Research has shown that the median royalty rate across all industries, as a percentage of average licensee operating profit margins is 26.7 percent (median royalty rate of 4.3%).[52] Royalty rates in medical technology licensing are expected to be considerable higher as the industry is one of the most profitable sectors of the global economy. Consequently, higher royalty rates are associated with more developed technology, as is CapMouse.
Typically, higher royalty rates are associated with license agreements that provide the licensee with exclusive rights to use the intellectual property. An exclusive right to use a keystone IP places the licensee in a superior position. If the IP provides highly desirable utility, premium prices can be demanded for the product. Competitors cannot counter with the same product without risking infringement, and the exclusive licensee will earn superior profits. Such an arrangement is worth higher royalty payments.

[52] Russell L. Parr, „Royalty Rates for Licensing Intellectual Property“, 2007 John Wiley & Sons, Inc.

4. Business environment

Demographics and income

The single greatest factor influencing demand for the medical device industry is the continued aging of populations and increasing incomes in the U.S., Europe, Japan, and other large developed and emerging economies.

The Organisation for Economic Co-operation and Development (OECD) data show that the elderly (65 years or older) currently represent 14.02% of the population of the developed countries.[53] This percentage share is expected to increase substantially over the next 40 years.

The proportion of Japan's population age 65 or older is growing faster than in the United States and EU, and stood at 20.8 percent in 2006 (compared with 17 percent in 2006 in the EU and 12.4 percent in 2006 in the United States).

By 2050 approximately 34.9 percent of Japan's population will be aged 65 or older. About 37 percent of the European population is projected to be 60 or over in 2050, up from 20 percent in 2000. Almost 30 percent is projected to be 65 or over, up from 15 percent in 2000. The U.S. Census Bureau estimates that, by 2050, the percentage of people in U.S. over the age of 65 will be 20.7%, up from 12.4% in 2000. Altogether the share of elderly people in the total population of Japan, EU and U.S. will grow at a pace of 1.1 percent a year till 2050. By then the total number of the elderly people in developed countries will reach 336.7 million compared to 164.7 million in 2006.[54]

In contrast with the slow process of population ageing experienced in the past by most countries in the more developed regions, the ageing process in most of the less developed regions is taking place in a much shorter period of time, and it is occurring on relatively larger population bases. The leading developing economies have already seen dramatic fertility decline and improved longevity over the past two decades. Prime example of these tendencies is the most populous of the emerging economies, China. The country already has about 102 million elderly (those ages 65 and over), or over one-fifth of the world's elderly population. And the percentage of elderly in China is projected to triple from 8 percent to 24 percent between 2006 and 2050, to a total number of 322 million.

Globally the population of older persons is growing at a rate of 2.6 percent per year, considerably faster than the population as a whole which is increasing at 1.1 percent annually. At least until 2050, the older population is expected to continue growing more rapidly than the population in other age groups. This increase will put enormous strain on the medical industry as a whole and represents a significant opportunity for medical equipment providers. Since older people typically require higher levels of medical treatment, the aging of developed countries relatively active and wealthy population will drive demand for medical devices and equipment over the next three or four decades.

---

[53] OECD Health Data 2008

[54] http://www.oecd.org/topicstatsportal/0,3398,en_2825_494553_1_1_1_1_1,00.html

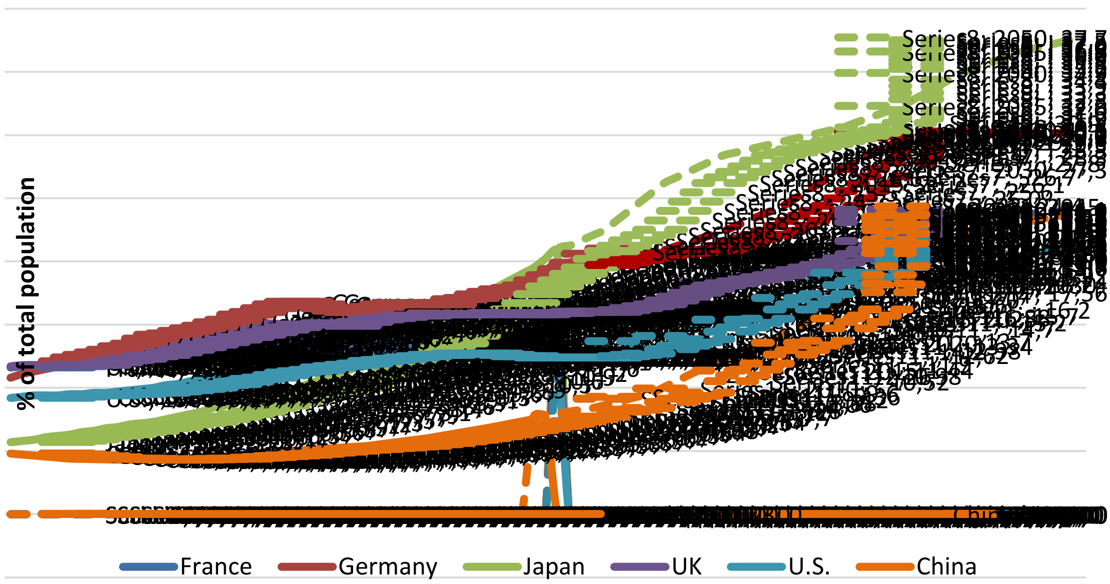

Figure 3 Population age structure, Population: 65 years old and over, % of total population

Source: OECD, UN World Population Prospects

Real household incomes have been increasing in all large developed countries through mid-1980s to mid-2000s (the only exception being Japan in the 90s). Furthermore, statistics from China reveal that self-reported average household income in country grew more than 70% from 1997 to 2006, and that the income of the typical urban Chinese household more than doubled in that period.

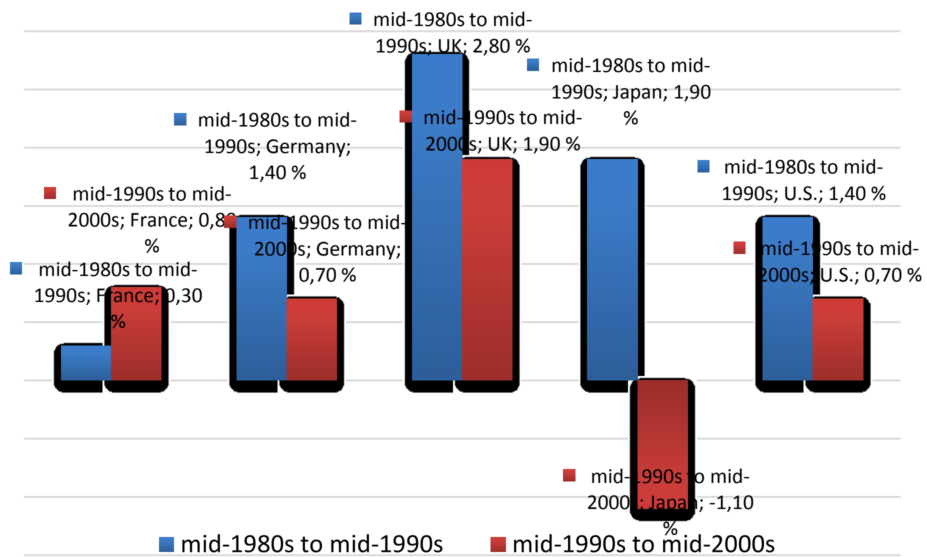

Figure 4 Trends in real household income; Average annual change mid-1980s to mid-2000s

Source: OECD

Those around retirement age – 55-75 – have seen the biggest increases in incomes over the past 20 years, and pensioner poverty has fallen very rapidly indeed in many countries, so that it is now less than the average for the OECD population as a whole.[55]

[55] Growing Unequal? Income Distribution and Poverty in OECD Countries, OECD 2008

Increasing household incomes in developed nations and especially rising middle-class incomes in large emerging economies will remain one of the driving forces behind the growing demand for medical technology.

Healthcare spending

In recent years, expenditures for healthcare as a share of gross domestic product (GDP) in the EU, Japan, and the United States have risen and are projected to continue to increase, although at a declining rate. In EU, currently an average of 8.7% of gross domestic product is spent on healthcare. Of this figure, 6.3% (i.e. 0.55% of GDP) goes to medical technology. Average spending per person on healthcare is €2,173 or $2,825 and of this, €145 or $189 is spent on medical technology.[56] In the US, expenditure on healthcare is about 15.3%of GDP with 5.5% (i.e. 0.84% of GDP) spent on medical technology. Japan is estimated to spend about 5 percent of health expenditures on medical devices, and health expenditures account for 8 percent of nations GDP.

Health expenditures have been growing very rapidly in the U.S. Since 1970, health care spending has grown at an average annual rate of 9.8%, or about 2.5 percentage points faster than the economy as measured by the nominal GDP. Annual spending on health care increased from $75 billion in 1970 to $2.0 trillion in 2005, and is estimated to reach $4 trillion in 2015. As a share of the economy, health care has almost tripled over the past 45 years, rising from 5.1% of GDP in 1960 to 15.3% of GDP in 2006, and is projected to be 20% of GDP in 2015. Health care spending per capita increased from $356 in 1970 to $6,697 in 2005, and is projected to rise to $12,320 in 2015.

Health expenditures in the countries of the European Union have also risen faster than gross domestic product since 1970, namely by 3.2 percentage points in the 1970s, 0.8 in the 1980s, 1.2 in the 1990s, and 1.6 points in 00s (up to 2006). If future health care expenditure will rise faster than GDP by 1 percentage point, the EU average will increase from over 8.7% in 2006 to 11.2% in 2030. If the increase is 1.5 percentage points higher, it will reach 12.6% in 2030.

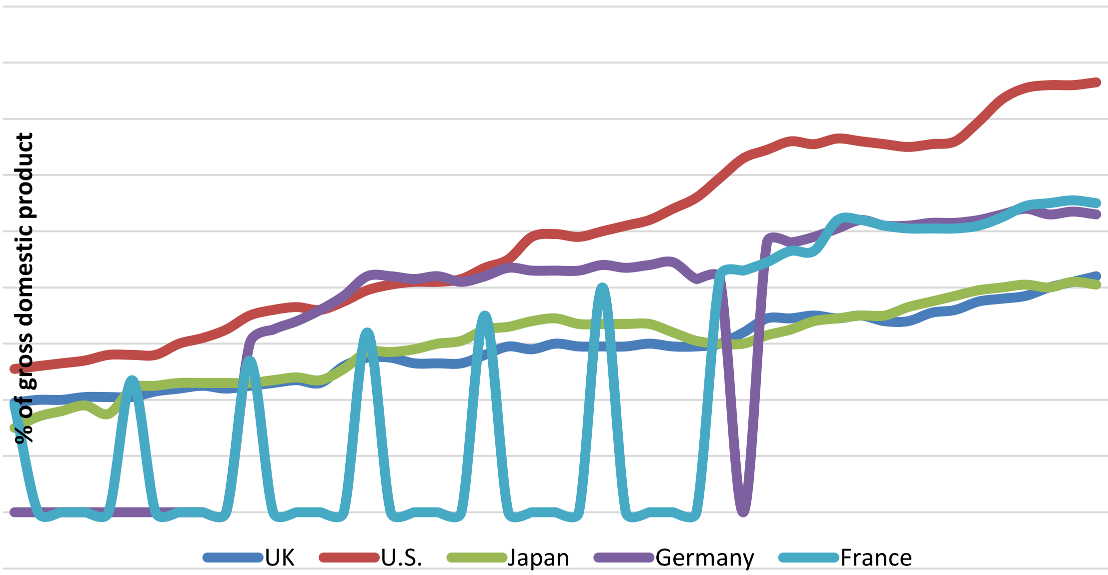

Figure 5 Total expenditure on health, % of gross domestic product

Source: OECD Health Data 2008

[56] Eucomed, http://www.eucomed.org/

The amount of money a country spends on healthcare is a key factor affecting the market for medical devices. National levels of healthcare expenditures are a strong indicator of potential success for the industry. Technological advances lead to the development of more innovative medical devices, which, in turn, enables a number of medical problems to be treated, thereby increasing demand and expenditures on medical devices. Technical change has been the largest driver of growth in healthcare spending over the past 50 years; it has resulted in annual increases in per capita spending on healthcare of approximately 2 percent, which is about half of the total real growth in healthcare spending from 1950–2000.

Reimbursement

Adequate reimbursement rates and transparent reimbursement policies are important demand factors in the medical device market. Reimbursement rates are determined by government and private payers after a medical device has been approved for use in a particular market. If medical devices are not reimbursed at a rate that enables the medical device consumer to recoup its cost or if rates are not predictable or transparent, the consumer then may not purchase the device.

Cost containment policies being adopted by healthcare systems that face growing budget deficits also put substantial monetary constraints on potential consumers of medical devices. These policies may take the form of budgetary caps (at the national, regional, or local level) or the adoption of payment systems that classify particular devices or treatments in categories for which the associated payment does not adequately cover the total cost of the device, innovation, testing, and/or marketing.

Most sector revenue comes from the sale of surgical and medical equipment to institutional purchasers rather than from the sale of expensive devices to individuals. Institutional purchasers, including public and private sector hospitals, have made efforts to trim costs by standardizing treatment protocols, making purchasing more rational, and forming group purchasing organizations (GPOs). The GPOs attempt to use their size to negotiate lower prices and possibly to counter market power on the seller's side. The increasing presence of GPOs in the United States and the EU has been met with some controversy and reported concerns about their impact on small and medium size suppliers and on the adoption of new technology.

U.S. medical device and equipment manufacturers sell the largest portion of their products to hospitals and physicians who typically bill various third-party payers, such as the Medicare and Medicaid programs, private insurance plans, and health maintenance organizations (HMOs), for the healthcare provided to their patients. As the largest single insurer in the United States, the Federal Medicare and Medicaid program has a profound influence on the healthcare market. About one-third of funding for hospitals, or approximately $125 billion a year, is from Medicare, and much of the balance is from private insurers and HMOs. Thus, the ability of customers (healthcare providers) to obtain reimbursement from third-party payers is critical to the success of medical device producers because it determines which products a customer can purchase and the price it is willing to pay.

Health care in the EU systems is either financed through general taxation or by contributions to health insurance funds.

There are three predominant systems of health care finance in the European Union. The first is public finance by general taxation. Secondly, there is public finance based on compulsory social insurance. Thirdly, there is private finance based on voluntary insurance, which covers only a small minority of EU citizens entirely, but which also operates on top of social insurance as a supplementary form of funding health care. Initiatives intended to cut costs and improve efficiency are prevalent in the EU, including a recent shift away from the global budgeting of healthcare spending to activity-based methods of payment, such as diagnosis-related groups (DRGs), and the increased use of health technology assessments (HTAs).

There are generally two ways in which a medical device can be reimbursed in Europe. Either the device is recognised as providing a health benefit in its own right, or it is recognised as part of a beneficial

procedure. For the former, reimbursement levels will be set for the device itself, but when a device is recognised as part of a procedure, payment for the device must come from within the budget set for the procedure as a whole. Reimbursement agencies across Europe have compiled lists of devices and procedures that can be reimbursed, along with the value of reimbursement. The procedures lists in Western Europe are generally based on a version of the diagnosis-related group. In this system, similar and related medical procedures are grouped together. Each group is then coded and given a value, which is the set amount of money that will be reimbursed for each procedure.
In addition to the DRGs, certain government bodies across Europe have been tasked with assessing medical devices and procedures to determine whether the average costs are justified by the patient outcome. These measurements, referred to as health technology assessments are not standardised across Europe, although they employ the same principles. An HTA will consider how well the technology works for both the provider and the patient, and it will also compare the device or procedure against alternatives, such as medication. The International Network of Agencies for Health Technology Assessment (INAHTA) defines HTA as:[57]

"*The systematic evaluation of properties, effects, and/or impacts of healthcare technology. It may address the direct, intended consequences of technologies as well as their indirect, unintended consequences. Its main purpose is to inform technology-related policymaking in healthcare. HTA is conducted by interdisciplinary groups using explicit analytical frameworks drawing from a variety of methods.*"

HTAs have a greater scope than the efficacy assessments conducted under the Medical Devices Directives for CE marking. For example, an HTA would consider cost and time implications of training, as well as running costs, such as maintenance.
As a result of HTA and reimbursement, the amount of clinical and supporting data needed to market a medical device in Europe has increased substantially.

[57] http://www.eunethta.net/HTA/

5. Market for CapMouse

In a broad sense, non-invasive oral interface can be introduced to a variety of market segments and applications such as:

- Market for disabled and elderly with muscle illnesses – the main driver for oral interface is related to its compliance to the needs of paralyzed people. These people are unable to use conveniently alternative interfaces and the CapMouse is potentially a killer application in this segment. The share of people having muscular problems is around 2.5% of population, which makes the potential clientele for the CapMouse interface in Europe reach 15 million users.
- Dental and medical markets – professionals who need to use their both hands at the same time may find the CapMouse interface as a value added component to operate computers or other machines without necessarily using hands. Speech recognition is here the main competitor. According to the FDI World Dental Federation there are approximately million dentists internationally, whereas the idea and necessity for designing the CapMouse interface as such has stemmed from the dental sector.[58]
- Market for people with Carpal Tunnel Syndrome – people who experience pain in using their hand in operating computer mouse. Oral interface may be an additional way of controlling the computer cursor. The lifetime risk for CTS is around 10% of the adult population, i.e. 72 million people in Europe alone.
- Computer game markets – one of the most dynamic markets, where oral interface may be used for additional functionality or as a novel interface for new experience. In 2008 34.7 million game consoles were sold globally, with a fast market growth during the upcoming years.[59]
- Handheld terminal market – 116 million smart phones shipped globally in 2007, which mushroomed to 171 million in 2008 and should reach 203 million in 2009. That gives smartphones 14.1% share of the total handheld market.[60]

CapMouse technology can be of great help for the disabled community, particularly disabled who suffer from muscular illnesses or problems with upper extremities. By using CapMouse interface they can easily be integrated into digital environment, as the CapMouse interface provides an efficient means of navigating and managing devices or content. The community which is likely to benefit from this innovation is globally 180 million people and over 6 million in Nordic countries.

In general, the CapMouse technology has the potential to become a universal interface for the elderly with chronic conditions, for people with different types of disabilities as well as for the able-bodied people for controlling a number of additional devices.

58 http://www.fdiworldental.org/
59 http://www.npd.com/
60 http://www.pcworld.com/businesscenter/article/158697/smart_phones_lead_market_growth.html

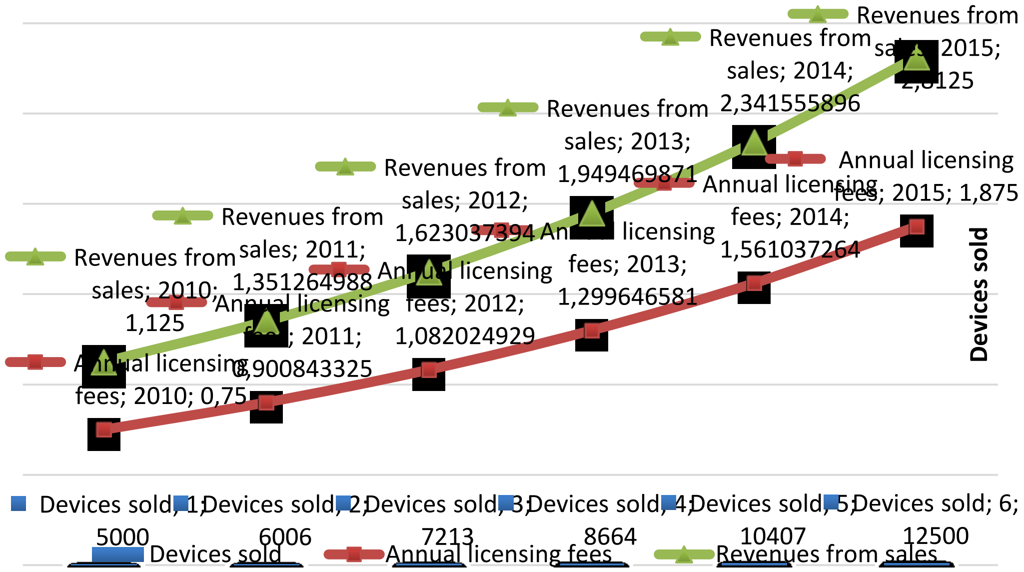

Figure 6 CapMouse device sales and 1) annual fees from royalties when licensing 2) revenues from sales when selling finished devices

The market entry for CapMouse interface is planned through the high-potential market for assistive medical devices for the disabled. Initial commercialization of the product on the highly regulated medical device market will help CapMouse technology to gain wider acceptance in the consumer marketplace. Technological solutions proven in the medical marketplace will help CapMouse consortium to adapt the interface to various consumer applications.