

A camera based human computer interaction through virtual keyboard assistant

M Uma^{*1,4}, T Sheela², Vivek³

¹ Assistant Professor, Department of Software Engineering, SRM Institute of Science and Technology, Chennai, India

² Professor, Department of Information Technology, Sri Sairam Engineering College, Chennai

³ Assistant Professor, Department of Software Engineering, SRM Institute of Science and Technology, Chennai, India

⁴ Research Scholar, Bharathiar University, Coimbatore, India

* Corresponding author: umaprabhu78@gmail.com

Abstract: Human Computer Interaction has evolved with time and exploited almost all the possible relationships that human can have with computers. The present scenario of using the computer is often inaccessible to people with severe motor impairment, which cannot inhabit the traditional keyboard and mouse, and require another specialized input device. A Camera mouse Human-Computer Interface (HCI) based virtual assistant is designed and planned to be evaluated with 15 subjects and hoped to work for individuals with advanced Amyotrophic Lateral Sclerosis (ALS), all of whom were given tasks to execute through the virtual assistant. In the proposed system, this includes primarily, a virtual keyboard with an embedded Internet browser. The virtual keyboard allows users with movement impairments, basically hand movement issues, to directly interact with the computer. The keyboard has varied functions, apart from just typing; namely, searching anything on google, login to social media and checking emails, text to speech facilities and has a database that holds all the basic helper commands for the person. The camera mouse is software that allows people to interact with the computer by registering any part of their body and then using that part as the pointing device to work on the computer system. The participants of the evaluation will be both, subjects that have knowledge about the HCI interface and those completely new to it. We define additional criteria for unrestricted internet access for evaluation of the presented and future internet browsers, Camera based browser provides unrestricted access and enables free web surfing for individuals with motor impairments. This system improves the subject's quality of life and provides uninterrupted access for the subject to use the computer.

Keywords: Camera, Human computer Interaction, Virtual keyboard, Graphical User Interface



1. Introduction

Communication and the capability to cooperate with the atmosphere are elementary needs for human associations. For those, who suffer from unadorned physical incapacities or palsy, the ability to fulfill this need is partial or at times impossible. In contrast to reduced motor activity, the sensual and intellectual functions are habitually almost intact (a locked-in state). The proposed paper therefore, forecasts a method in which those with severe impairments of physical movement are able to interact with the computer with a systematic ease [1-3]. Patients who are suffering from ALS have learned to work an electronic spelling device [1]. With the same aim, this paper hopes that the interaction of people with the basic functionalities of the computer, like browsing the internet, or storing some information that needs to be converted to speech can be easily done through this virtual assistant.

As far as the design of the virtual keyboard is concerned, the design of proposed in this system is not the traditional qwerty keyboard, but a standard A-Z input keyboard in the order of the letters, as in English language. The reason for this non-conventional keyboard design is the fact that the targeted audiences for this keyboard are people who have very less or absolutely no exposure to the computer till date. The keyboard would also allow them with extravagance functions like keeping a music playlist and playing music or audio at their choice and a predictive word loader that would auto-complete words based on a the user's helper database and the user will not have to type the entire word every time.

1.1. Previous Work

The severe motor impairment caused due to brain stroke, cerebral palsy and other neurological problem. People who cannot speak and cannot use their hands to operate a traditional input devices such as mouse, keyboard are very difficult to communication with outside environment. EOG is one of the solution to communicate with external world such as robotic aids, rehabilitation aids (Barea et al., 2002), computer control (Betke et al., 2000), etc. are being achieved using EOG signal. Subjects with severe paralysis or motor impairment have challenging one to controls the traditional input devices such as moving the cursor to desire position [4]. The GUI interface has been developed for the user to perform different tasks through their eyes movements [5].

Camera mouse was the one of the alternative input devices and customized assistive software helps the disabled to communicate with family and caregivers [6, 7]. Severe motor disorder may have a choice of eye movement as only way to control the computer. Example head mounted cameras with infrared signal or EOG. Severe paralysis may leave the eyes as the only muscles that a person can control. Such an individual may be able to control the computer via interfaces that detect eye movements, for example via head-mounted cameras, which typically process infrared signals, [8] or EOG. Al-Haddad et al., Blinks and gaze angle are measured and taken as inputs for the controlling method. The new auto navigation method was done by microcontroller which knows the goal point direction and distance by calculating the gaze angle that the user is gazing at. Future work to integrate Human Computer Interface (HCI), wheelchair and EOG system together [9]. Wajeetha Akram, et al., introduced an enhanced computer vision based strategy where features, for example on a user's face, viewed through an inexpensive USB camera, are tracked and translated to pointer movement to accessing desktop files and folders for disabilities[10]. Eye blinks and lateral eye movements are prominent in EEG signals which are obtained by placing electrodes in the frontal region of the brain, a machine learning approach to

detect eye movements and blinks from EEG data and map them as intents to control external devices like a computer desktop or a wheel chair [11].

2. Methodology

The camera mouse is open-source software available with the camera mouse official web page. The concept of the application is to allow fully access to the pointer of the computer by registering any body part with the device. The device opens with the computer's camera and on the camera, image the user is supposed to click on the body part that he wants to register as the new pointing device. The system tracks the computer user's movements with a video camera and converts them into the movements of the mouse pointer on the screen. The visual tracking algorithm is based on cropping an online template of the tracked feature from the current image frame and testing where this template correlates in the subsequent frame. The user can interact with the computer by moving their cursor through eyes [13]. Electrical discharge machining (EDM), once considered a non-traditional machining process, has been replacing conventional machining operations and is now a well-established the preferred machining option in many manufacturing industries throughout the world. Modern EDMs are capable of machining geometrically complex and/or hard material components that need to be precise and are difficult-to-machine, such as heat treated tool steels, composites, super alloys, ceramics, etc.

Table 1. Camera-based interface system for user with severe motion impairment

System	Action Mechanism	References
Camera Mouse	Move Body Part. Eg: Foot, Head and Finger.	Betke, Gips, and Fleming [14] Betke [6]
Blink Link	Control over the blink length	Grauman, Betke, Lombardi, Gips, Bradski [15]
Eyebrow Raiser	Raise Eyebrows	Lombardi and Betke [16]
Eye Keys	Control horizontal gaze direction	Magee, Betke, Gips, Scott, and Waber [17]
Finger Counter	Move fingers	Crampton and Betke [18]
Head Tilt Detector	Tilt head left or right	Waber, Magee, and Betke [19]
Symbol Design	Move Body part	Betke, Gusyatn, and Urinson [20]

From the above comparisons we draw a conclusion that the Camera Mouse is a more simpler way of using the pointing device because there is no limitation as to which part of the body needs to be registered. The following figure illustrates the registration of an eye as the pointing device.

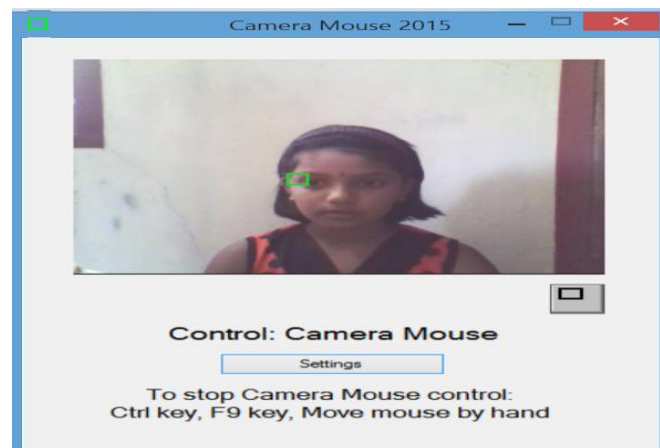


Figure 1. The Camera Mouse Interface [12]

2.1 Tracking Algorithm

When the user originally clicks on the desired feature to be selected as the pointer, a square is strained around that portion and the sub-image within that square is collected out of that image frame. The cropped sub-image is then used as “template” to know the site of the feature in the coming next frame of image. To find this site, the tracking system uses this template to search for the same feature in a “search window” that is positioned at the location of the feature in previous selected frame. The template shifts through this window and is correlated with the original sub-images. The window is well-defined to contain the centers of all previously selected sub-images tested. Since a novel frame gets received in a thirtieth of second, the prototype is typically very alike to the intensity pattern of that feature in the new acquired frame, which can be searched for the best correlated sub-image. The supposition that corresponding intensity lines in succeeding frames are constant, the “constant brightness assumption,” is frequently made when scheming algorithms for gesture analysis in such images [21-22]. The sub-image with the maximum correlation coefficient amongst all the sub-images in the available search window is resolute. It serves two distinct purposes; first, its focus coordinates are moved to the user computer to be understood as pointer coordinates. Second, the image is collected from the present frame and converts the current template that is used to search for the best match in the sub-image in the next and upcoming frames. The process reiterates and the pattern is updated for every image frame. This apprising of the template ensures that a strong match can be produced in every image frame and that the gesture of the initially selected feature is then tracked. If at any point of time a low correlation is got and the template no longer nearly resembles the template’s desired feature, the overseeing subject is free to re-click on that particular feature on the screen, thus informing the pattern manually.

2.2 System Architecture and Algorithm:

```

Start
Capture (Camera Video Feed)
Breakpoint;
Register the feature -> new Image template;
Template -> Mouse Action
Tracking Algorithm ( );
Select (Desired Module);

```

Stop.

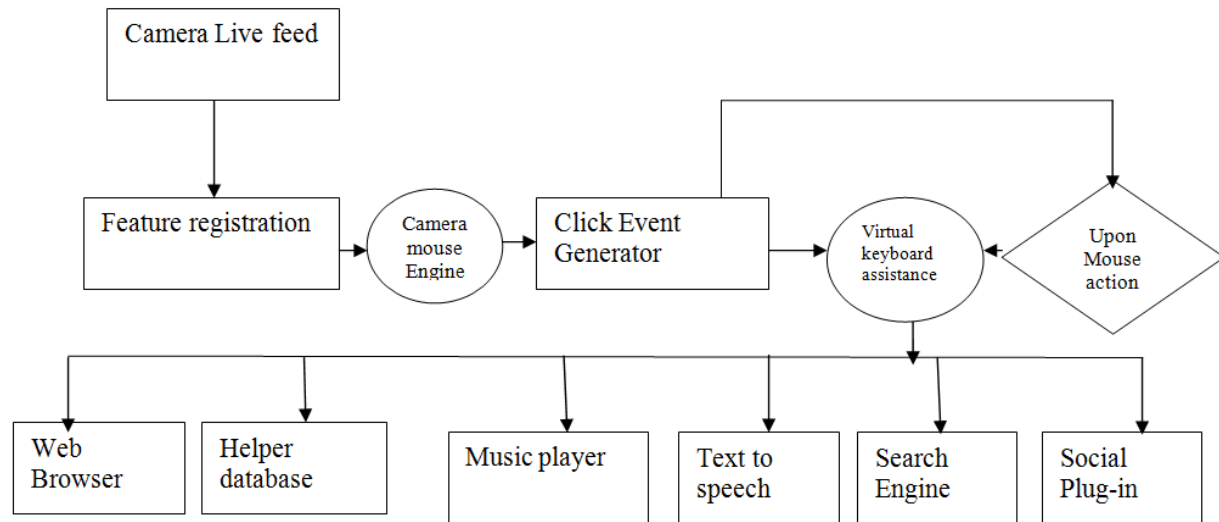


Figure 2. System Architecture of Virtual Keyboard

3. Result and Discussions

3.1 Architecture of the Keyboard Assistant

The complete software has been designed and developed in the Visual Studio IDE using vb.net as the programming language. Most of the elements of the software use Windows Forms for their layout attached with Microsoft Office Access Database Files.

- A. *Keyboard Layout*: The layout of the keyboard or the first form of the software shows the outline from where all the other elements of the keyboard can be accessed. The design is not the traditional one, with a qwerty model with only one reason being the user is expected to be using it for the first time, and would use it very less often. Apart from the alphabets and numbers there are eight functional keys on the top.
 - a) *Music Player*: The music player can be used for both audio and video files, and also provides a feature of creating the user's personal favorites' playlist [B].
 - b) *Web Browser*: The web browser is integration into the keyboard itself, and functions similar to the other browsers and is simple to operate and understand [C].
 - c) *Help*: The self-help option is a button that is a one-stop help option for the user, and can be used to convey messages or demands to other people around the user asking for help, or simply displaying the needs of the user [D].
 - d) *Social Networking (Gmail / Face book)*: Social Networking has become an important and significant part of our lives, and therefore this keyboard allows the user to access their social networking accounts and actively participate in them [E].

- e) *Text-To-Speech*: This is an important feature for those who have difficulty in speaking and can use the camera mouse interface to type their needs, and then use this feature to speak it out, so that, they can allow people around them to realize that they are need of something.
- f) *Google Search*: The google search works by making the textbox onto which the user types anything into the google search bar. Upon typing, what needs searching, and clicking the Google Search button, the software opens up the system's default browser and runs a google search on the typed keywords.
- g) *Delete/Backspace*: The delete and backspace keys function in the same way as in any other keyboard, used for erasing text from both the directions.

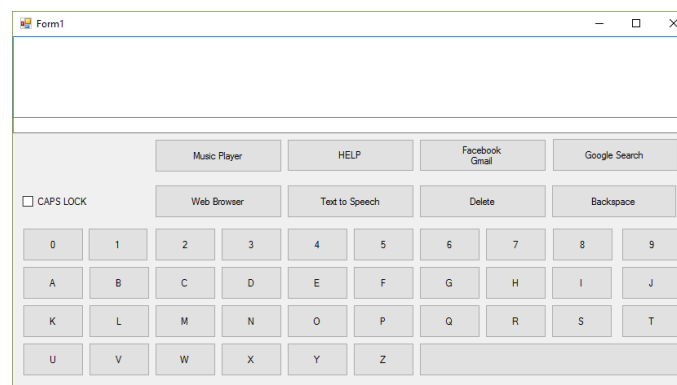


Figure 3. Layout of the Keyboard Virtual Assistant

3.2 Music Player

The music player is an extension that is meant to help the user create and edit songs or video files of choice and preference. It is a media player plug-in attached to the Windows Form Application and connected to user dependent database, that would hold the songs added to the playlist.

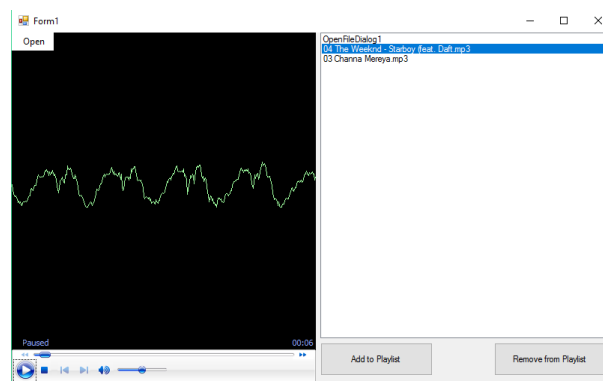


Figure 4. Music Player GUI in the Virtual Assistant

3.3 Web Browser

The web browser functions as any other web browser, although the proposed design is simple and does not have any functions other than reload, forward and backward. The design is made by integrating a web browser tool into the form and then encoding the buttons to do the respective functions of going a page back or forward, reloading the page. The home page by default is set to open the Google Search page.

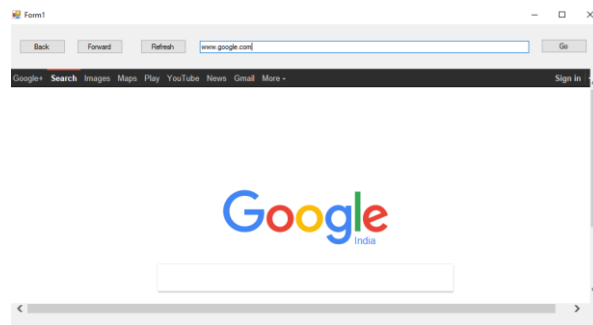


Figure 5. Web Browser GUI in the Virtual Assistant

3.4 Help

The help form of the project helps the user with communication imparities to be able to communicate properly with their care-takers. The help module has commands that are pre-installed into the system, and at the same time can also be added manually by the user. These commands can later be used to help the user by converting them into speech, so that the care-taker or anyone around the person could come to their help.

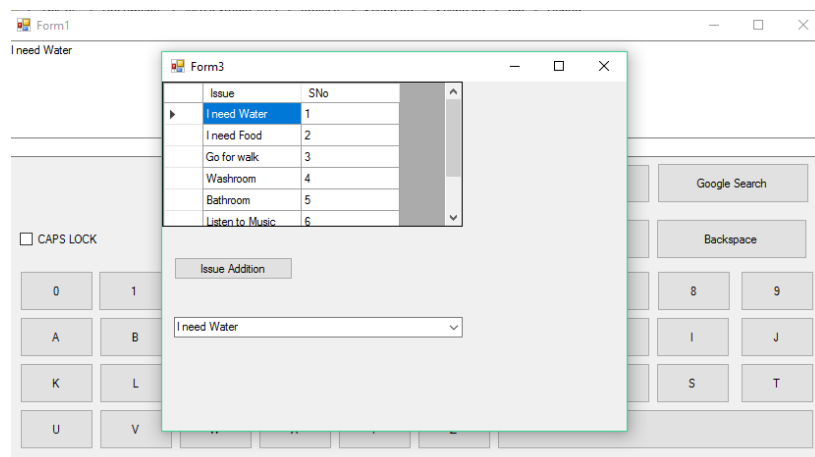


Figure 6. Helper Database

3.5 Social Plug-in

The social plug-in has two separate modules, Gmail and Face book. The user can use the virtual assistant to log onto any of these sites, and save their login information. They can easily then, access these websites through the social plug-in form and use these websites like any other browser would show it to them.

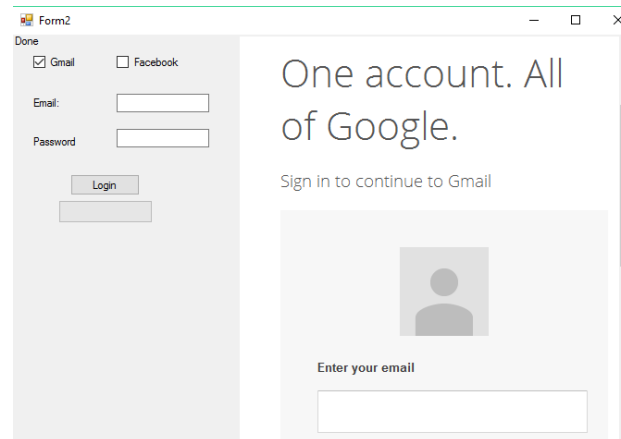


Figure 7. Gmail and Face book Social Plug-in

3.6 Predictive Drop Down

Predictive drop down is a feature useful for both, accessing the Helper toolbox and also for the regular typing on the main screen. The drop down functions based on a database that contains words that the user has been using frequently and actions that need to be performed at a regular basis. This would help, save time for the user and he would not have to type the entire word again and again.

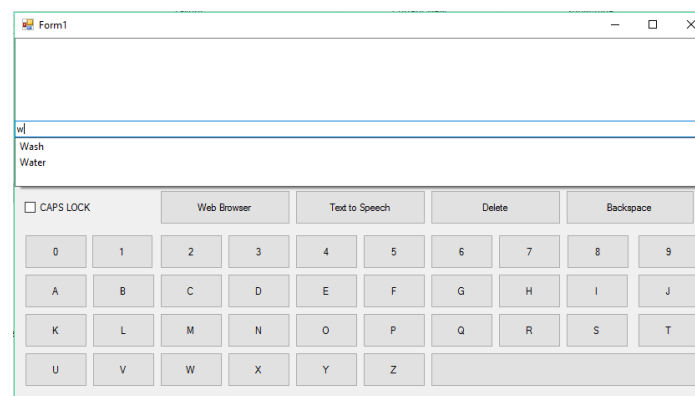


Figure 8. Predictive drop down window

3.7 Experimentation and Results

The Virtual keyboard assistant was experimented on 15 subjects in 5 separate modules, taking 3 subjects at a time. Records were maintained for subjects being able to use the tool prior to any training and later with training and slight differences were noted.

Table 2. Combined Experiment results of first 3 modules without training the subjects.

Task	Average Time Taken (seconds)	Accuracy (%)
Type keyword Water	34.281	83
Login to Gmail	47.232	81
Add music to playlist	52.415	79

Search on Google	31.152	89
Add helper command	63.224	91

Another reading was taken on the same subjects in 3 separate modules, after giving them proper time and training with the Camera Mouse software and knowledge about the Virtual Assistant tool. Accuracy calculation for both the training sets were based on a threshold value set for each task. Number of readings performed nearest to that threshold value would make the subject's task doing, more accurate.

Table 3. Experiment results after giving the subjects training on using the software

Task	Average Time Taken (seconds)	Accuracy
Type keyword Water	31.471	91%
Login to Gmail	44.961	87%
Add music to playlist	49.346	88%
Search on Google	27.864	91%
Add helper command	60.124	92%

4. Conclusions

1. The virtual assistant is an application that would make lives easier for people suffering from physical imparities. The proposed usage mining based personalized assistive technology provides ease of access of the external world.
2. The system is flexible, user-friendly and easy to operate. The system control is provided by using the eye tracking movements, or any other body part movement in real time. The user can access the internet, play music or ask for help without spending much time and clicks. The real-time experiment is carried out by allowing healthy subjects to use the system. The result shows that they can easily access the system without any selection issues and learn the operation of the system in a short time period.
3. Music and video players are developed based on the user's interest. The user can listen to music and watch videos without any outside help. One of the most important things is the emergency panel which allows the user to cast emergency responses and ask for help to anyone around the user. He can also show messages to other people sharing his interests.

5. Future Work

The current virtual keyboard assistant is developed using the camera mouse system that allows the user to register a particular body part as the pointing device to the computer and then use the camera mouse to perform the regular functions of any pointing device. The future plan would be to use eye blink and use the EEG signals captured by the emotive device and function according to the emotional conscience of the user. The system would completely be controlled by the mind and the signals that the emotive software would capture.

6. References

- [1] Scherer R, Muller GR, Neuper C, Graimann B and Pfurtscheller G 2004 An Asynchronously Controlled EEG-Based Virtual Keyboard: Improvement of the Spelling Rate, *IEEE Trans Biomed Eng.* **51**(6), pp.979-84.
- [2] Birbaumer N, Kubler A, Ghanayim N, Hinterberger T, J Perelmouter, J Kaiser, I Iversen, B Kotchoubey, N Neumann, and H Flor 2000 The thought translation device (TTD) for completely paralyzed patients, *IEEE Trans. Rehab. Eng.* **8**, pp.190–193.
- [3] Neuper C, Müller G, Kubler A, Birbaumer N, and Pfurtscheller G, 2003 Clinical application of an EEG-based brain-computer interface: A case study in a patient with severe motor impairment, *Clin. Neurophysiol.* **114**(3), 399–409.
- [4] Keates S 2009 Motor impairments and universal access, In: *Stephanidis Constantine (ed.), The Universal Access Handbook. CRC Press, Cleveland.* 5-12.
- [5] Anwesha Banerjee, et al 2015 Voluntary eye movement controlled electrooculogram based multitasking graphical user interface, **18**(3), 254-271.
- [6] Margrit Betke 2009 Intelligent interfaces to empower people with disabilities, *Handbook of Ambient Intelligence and Smart Environments*, DOI 10.1007/978-0-387-93808-0_15 Springer, New York.
- [7] Mcmurrough C, Ferdous S, Papangelis A, Boisselle A and Makedon F 2012 A survey of assistive computing devices for cerebral palsy patients, *5th ACM International Conference on Pervasive Technologies Related to Assistive Environments (PETRA 2011), Greece.* F1:1–F1:8.
- [8] *Eye Tracking System*, Applied Science Laboratories 2008 Bedford, MA, USA. <http://www.a-s-l.com>
- [9] Al-Haddad, Sudirman and R Omar 2011 Guiding Wheelchair Motion based on EOG Signals using Tangent Bug Algorithm *Third International Conference on Computational Intelligence, Modelling & Simulation.* 695-702.
- [10] Wajeeha Akram, et al 2007 A Customizable Camera-Based Human Computer Interaction System Allowing People with Disabilities Autonomous Hands-Free Navigation of Multiple Computing Tasks, *Universal Access in Ambient Intelligence Environments.* 28-42.
- [11] Soumya sen gupta, et al 2012 Detecting Eye Movements in EEG for Controlling Devices *Computational Intelligence and Cybernetics, 2012 IEEE International Conference.* ISBN: 978-1-4673-0892-2.
- [12] Epstein S, Missimer E, and Betke M 2014 Using kernels for a video-based mouse-replacement interface, *Personal and Ubiquitous Computing.* **18**(1), 47-60.
- [13] Margrit Betke, Member, and Peter Fleming, 2002 Features to Provide Computer Access for People With Severe Disabilities, *IEEE Transactions on neural systems and rehabilitation engineering.* **10**(1).
- [14] Betke M, Gips J, Fleming P 2002 The Camera Mouse: Visual tracking of body features to provide computer access for people with severe disabilities, *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* **10**(1), 1–10.
- [15] Grauman K, Betke M, Lombardi J, Gips J, Bradski GR 2003 Communication via eye blinks and eyebrow raises: Video-based human-computer interfaces, *International Journal Universal Access in the Information Society* **2**(4), 359–373
- [16] Lombardi J, Betke M 2002 A camera-based eyebrow tracker for hands free computer control via a binary switch. In: *7th ERCIM Workshop on User Interfaces for All, Paris, France*
- [17] Magee, J.J., Betke, M., Gips, J., Scott, M.R., Waber, B.N. 2008 A human-computer interface using symmetry between eyes to detect gaze direction, *IEEE Transactions on Systems Man & Cybernetics,* **38**(6), 1-14.

- [18] Crampton S, Betke M 2003 *Counting fingers in real time: A webcam-based human-computer interface with game applications*, In: Proceedings of the Conference on Universal Access in Human-Computer Interaction (UA-HCI), Crete, Greece, 1357–1361.
- [19] Waber BN, Magee JJ, Betke M 2005 Fast head tilt detection for human computer interaction. In: Sebe N, Lew M, Huang T (eds) *Computer Vision in Human-Computer Interaction*, ICCV 2005 Workshop on HCI, Beijing, China, October 21, 2005, Proceedings. LNCS, Vol. 3766, Springer-Verlag, 90–99.
- [20] Betke M, Gusyatin O, Urinson M 2006 Symbol Design: A user-centered method to design pen-based interfaces and extend the functionality of pointer input devices. *Universal Access in the Information Society*, **4**(3), 223–236.
- [21] B. K. P. Horn, *Robot Vision* Cambridge, MA: MIT Press, 1986.
- [22] M. Betke, E. Haritaoglu, and L. S. Davis 2000 Real-time multiple vehicle detection and tracking from a moving vehicle, *Mach. Vis. Appl.*, **vol. 12, no. 2**, 69–83.