# Head Gesture Recognition for Hands-free Control of an Intelligent Wheelchair

P. Jia[1], H. Hu[1], T. Lu[2] and K. Yuan[2]

[1]Department of Computer Science, University of Essex, Colchester CO4 3SQ, U.K.

[2]Institute of Automation, Chinese Academy of Sciences, Zhongguanchun Road, Beijing, China

**Abstract:** Traditional electric-powered wheelchairs are normally controlled by users via joysticks, which cannot satisfy the needs of elderly and disabled users who have restricted limb movements. This paper presents a novel hands-free control system for intelligent wheelchairs. Both Adaboost face detection and Camshift object tracking algorithms are combined in our system to achieve accurate face detection, tracking and gesture recognition in real time. It is intended to be used as the human-friendly interface for elderly and disabled people to operate our intelligent wheelchair using their head gestures rahter than their hands. Experimental results are given to demonstrate the feasibility and performance of the proposed hands-free based control strategy.

**Keywords**: Head gesture recognition, Intelligent wheelchair, Face detection, Head gesture interface, hands-free control

## 1. Introduction

To improve quality of life for the elderly and disabled people, electric-powered wheelchairs (EPWs) have been rapidly deployed over the last 20 years (Ding and Cooper, 2005) (Simpson *et al.*, 2004). Up to now, most of these EPWs are controlled by users' hands via joysticks, and are very difficult for elderly and disabled users who have restricted limb movements caused by parkingson diseases and quadriplegics. As cheap computers and sensors are embedded into EPWs, they becomes more intelligent, and are named as intelligent wheelchairs (IWs). Various research and developments on IWs have been carried out in the last decade, such as CALL Smart Wheelchair (CALL Centre, 1994), Wheelesley (Yanco *et al.*, 1995), OMNI (Hoyer and Borgolte, 1996), NavChair (Levine *et al.*, 1999), TAO projects (T. Gomi and A. Griffith, 1998), Rolland (Röfer and Lankenau, 1998), Maid (Prassler *et al.*, 1998), UPenn Smart Wheelchair (Rao *et al.*, 2002), SIAMO (Mazo *et al.*, 2002), *etc.*

The successful deployment of IWs requires high performance and low cost. Like all the other intelligent service robots, the main performance of IWs includes: (i) The autonomous navigation capability for good safety, flexibility, mobility, obstacle avoidance, *etc.* (ii) The intelligent interface between users and IWs, including hand-based control (joystick, keyboard, mouse, touch screen), voice-based control (audio), vision-based control (cameras), and other sensor-based control (infrared sensors, sonar sensors, pressure sensors, *etc.*). As an intuitive vision based interaction interface, head gesture has already been applied in some existing IWs, such as NLPRWheelchair (Wei, 2004), WATSON (Matsumoto *et al.*, 1999), (Matsumoto *et al.*, 2001), OSAKA wheelchair (Nakanishi *et al.*, 1999), (Kuno *et al.*, 2001), *etc*. However, these systems are not robust enough to be deployed in the real world and much improvements are necessary. The new generation of head gesture based control of wheelchairs should be able to deal with the following uncertainty in the practical applications of IWs:

- The face color may change dramatically in varying illumination conditions;
- The user head may move around for looking rather than moving;
- The user may have different facial appearances at different time, such as mustache and glasses; and
- The background may be cluttered and dynamically changing when IWs move in the real world.

In this paper, a novel head gesture based interface (HGI) is developed for our IW, namely RoboChair, based on the organic integration of the Adaboost face detection algorithm (Bradski, 1998) and Camshift object tracking algorithm (Viola and Jones, 2004), and is intended to be used as the human-robot interface for our IW, *RoboChair*, to solve the problems listed above.

The rest of the paper is organised as follows. Section 2 presents the system hardware structure of our RoboChair. In Section 3, we propose a new head gesture based interface (HGI) for the control of our RoboChair, which is based on the combination of both Adaboost and Camshift algorithms. Experimental results are presented in Section 4 to show the feasibility and performance of our new algorithm. Finally, a brief conclusion and future wrok are presented in Section 5.

## 2. System Hardware Structure

Figure 1 shows the picture of our RoboChair that was built in 2004 and has the following components:

- 6 ultrasonic sensors at a height of 50cm for obstacle avoidance (4 at the front and 2 at the back);
- DSP TMS320LF2407-based controller for motion control of two differentially-driven wheels;
- A local joystick controller to connect to an A/D converter of the DSP-based controller;
- A Logitech 4000 Pro Webcam for recognising the user's head gesture; and
- Inter Pentium-M 1.6G Centrino laptop with WindowsXP installed to analyze the head gesture.

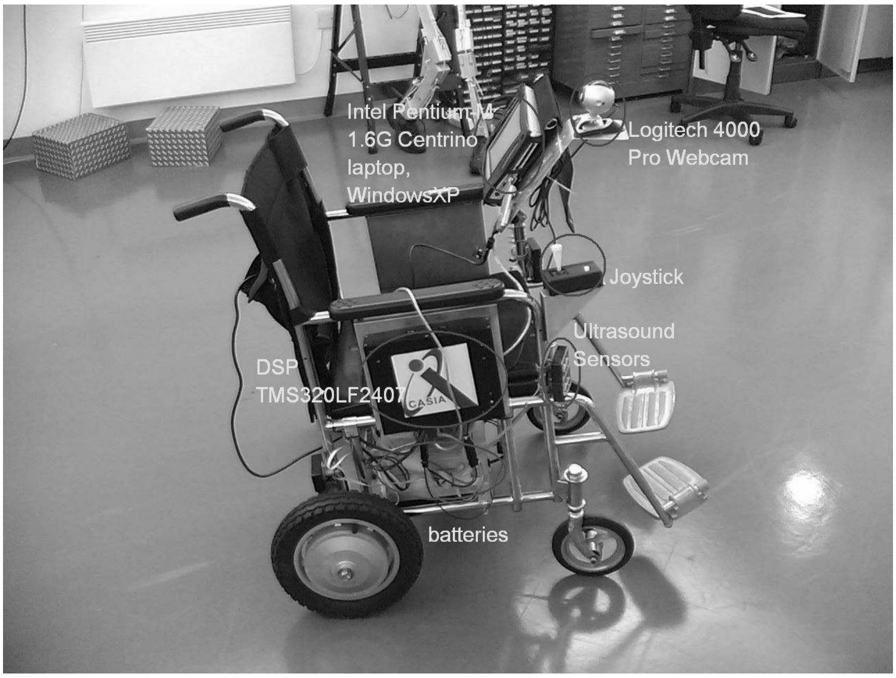



**Figure 1** Prototype of the RoboChair used in this research

The control architecture of our RoboChair is shown in Figure 2. The TI DSP chip TMS320LF2407 is used as the core processor of the motion control module. It offers excellent processing capabilities (30MIPS) and a compact peripheral integration, so that the control system is able to obtain both real-time signal processing and high performance driving control. With an obstacle avoidance module embedded in the DSP motion controller, our RoboChair has two control modes, namely intelligent control and manual control.

- Intelligent control mode: This is currently being developed in this project. Under this control mode, our RoboChair is controlled by the HGI. A Logitech webcam is used to acquire the facial images of the user. After the image data is sent to the laptop, the head gesture analysis and the decision making are implemented. Finally, the laptop sends control decision to the DSP motion controller that actuates the two motor motors.
- Manual control mode: This is already built in our RoboChair. Under this control mode, our RoboChair is controlled by the joystick that connected to the DSP motion controller with the obstacle avoidance module embedded.

It should be noticed that no matter which control mode RoboChair is working under, in order to deal with the uncertainties in the real world, the sonar readings are directly sent to the DSP motion controller to avoid the obstacles and handle the emergency in real time.

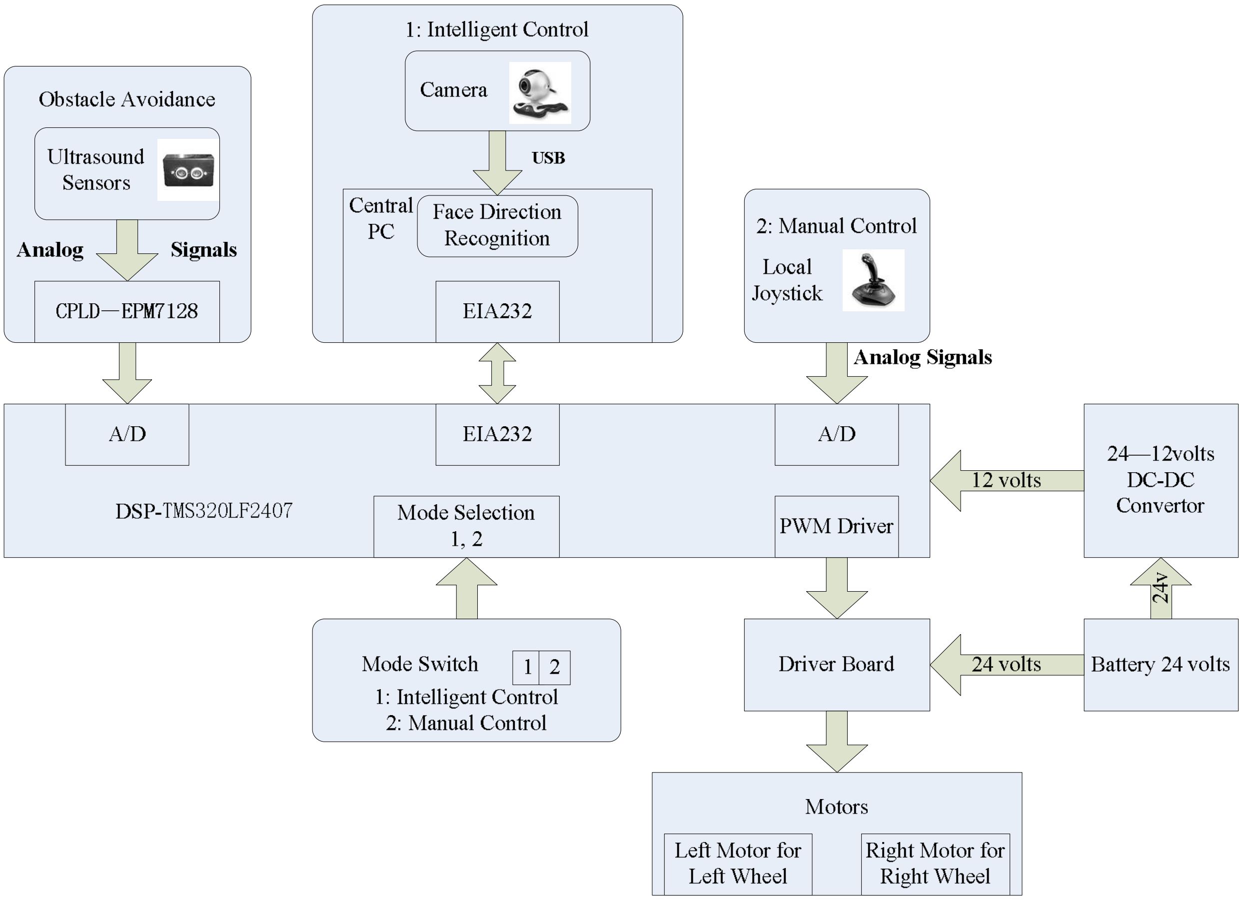



**Figure 2** Block diagram of the control architecture for our RoboChair

## 3. HGI for RoboChair

### 3.1 Adaboost and Camshift Algorithms

Adaboost is the most recent face detection method with both high accuracy and fast speed (Viola and Jone, 2004). It extracts the Haar-like features of the image, which contain the image frequency information only by integer calculation that is fast. Then a set of key features are selected from all the extracted features. After being sorted according to the importance, this set of features can be used as a cascade of classifiers that are very robust and able to detect various faces under vary illumination conditions and different face colours. Also, Adaboost is able to detect profile face detection.

Figure 3 shows the block diagram of Adaboost face detection algorithm. It consists of a sequence of stages: (i) data acquisition (image capturing); (ii) pre-processing (filtering); (iii) feature extraction (rectangular features); and (iv) a parallel stage: classifiers design (boosted cascade design) and classification. In the classifier design, supervised learning is adopted to select Ababoost features.

Camshift is a very efficient colour tracking method and is a classical optimisation algorithm. It is a fast object tracking method based on image hue (Bradski, 1998). Camshift uses a robust non-parametric technique for claiming density gradients to find the mode (peak) of probability distribution called the mean shift algorithm. Each iteration, Camshift aims to find the mean window centre using a fixed window size. If either the window centre or the window size is unstable, both values need to be adjusted accordingly until convergence.

Figure 4 shows the flowchart of the Camshift object tracking process, which consists of four stages: (i) initialization; (ii) window size adjustment (control); (iii) target search; and (iv) solutions. Note that Camshift has some limitations: (i) it cannot accurately track the face when the illimunation condition changes; (ii) it cannot work well under the clutered background.

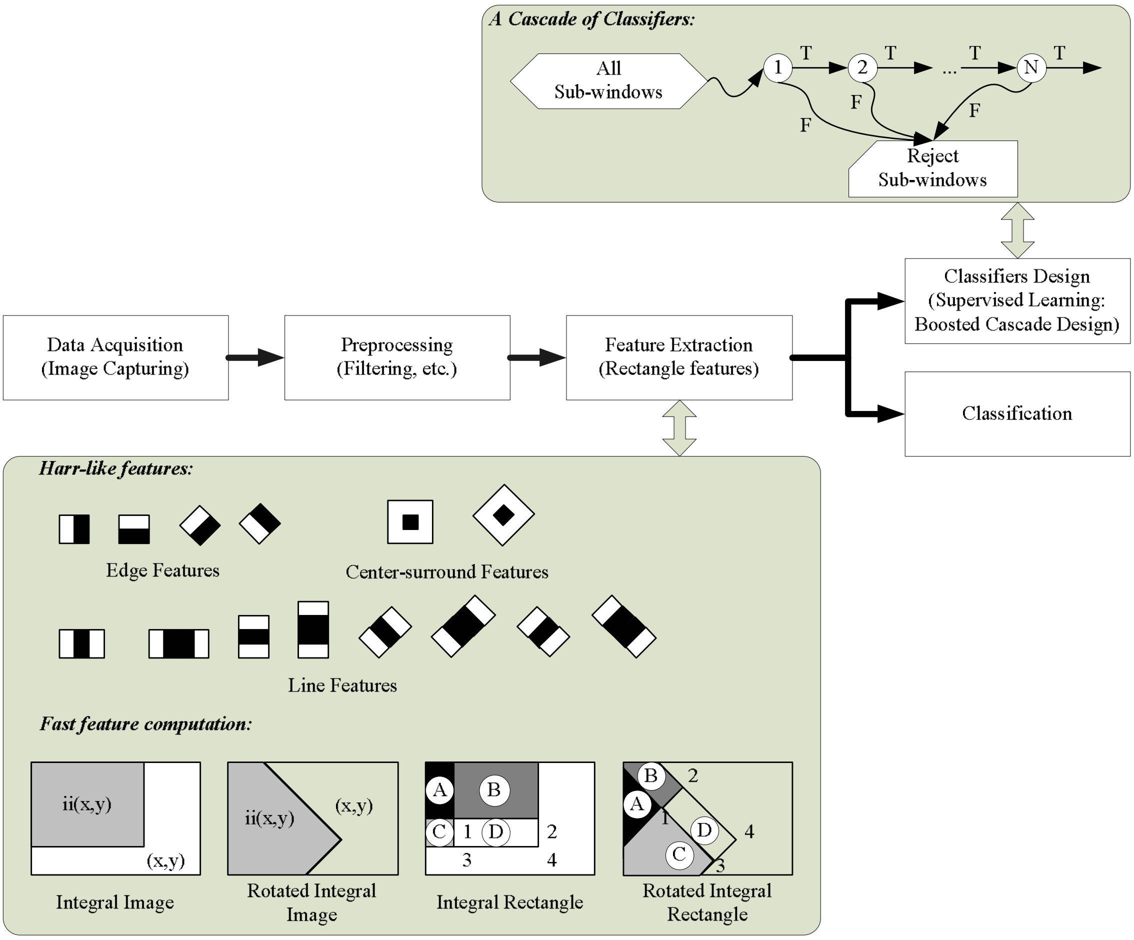



**Figure 3** Block diagram of theAdaboost face detection process

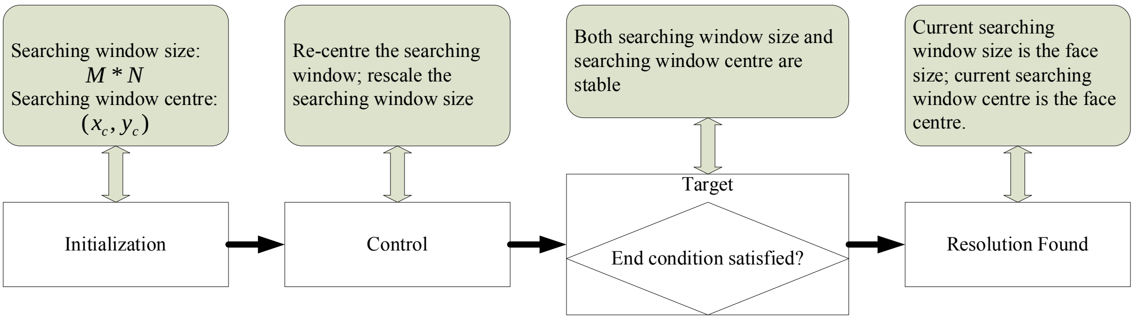



**Figure 4** Flowchart of the Camshift Object Tracking Process

### 3.2 Integration of Adaboost and Camshift Algorithms

For a low-cost IW system, the speed of the Adaboost face detection algorithm is still not fast enough to satisfy a multi-task real-time system. On the other hand, the camshift face tracking algorithm runs very fast, but is not robust to varying illumination conditions and noisy backgrounds. In order to obtain both fast speed and high accuracy, it is necessary to integrate both algorithms in a unified framework as shown in Fig. 5 (JIA and Hu, 2005).

Every frame, the system is trying to keep tracking the user’s face which is always in front of the webcam in our Robochair application.

- If the user's face is tracked successfully, Adaboost face detection is applied in the comparatively small Camshift tracking window so that the face position and size, as well as head gesture (including frontal, left profile and right profile) can be obtained rapidly. Further direction judgement is needed when frontal face is detected. Here, a simple scaled nose template matching is used to calculate the nose position in the frontal face window, through which the face directions can be determined very accurately.
- If the user’s face is lost from camera view, Adaboost face detection is applied in the whole captured image frame so that the user’s face will be tracked once again.

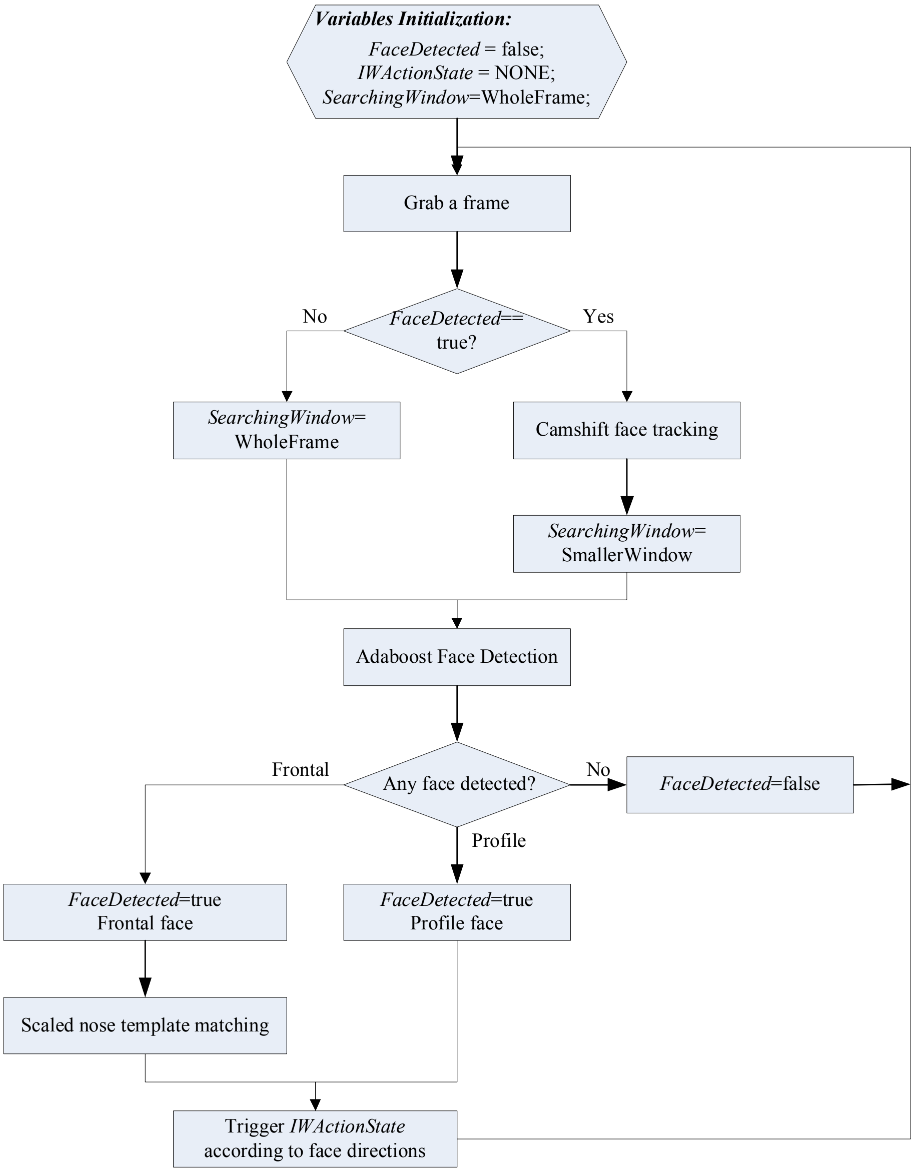



**Figure 5** Flowchart of the HGI (Integration of Adaboost and Camshift algorithms)

### 3.3 Head Gesture Recognition

Intel OpenCV has already contained the trained frontal and right profile face classifiers. After simply flipping the captured image, the left profile face can be detected as well. In order to tell the head gesture under any special situations with satisfying robustness, Adaboost frontal, left profile and right profile head gesture classifiers are applied in our research.

Because Adaboost is an appearance-based face detection method and the left face appearance is quite similar to the right face appearance. Occasionally, the face may be detected as the left profile and the right profile at the same time. A simple strategy is used to solve this situation: the profile face with bigger detection window dominates the head gesture. When both left profile and right profile detection windows are of the same size, the wheelchair will keep the status of the last cycle.

If the profile face is detected, it is for sure that the IW is going to turn left or right. But, if the frontal face is detected, further left frontal/right frontal/up frontal/down frontal /center frontal head gesture is to be recognized. Because of the varying distance from the face to the webcam, the detected face windows in different frames are not of the same size. Thus, a straight mind is to scale the detected face windows to a standard size. Here, we use a size of 100*100 as the standard face window. Then, the classical template matching method is applied in this small window to calculate the precise nose position, which will tell the exact frontal face head gestures.

### 3.4 Motion Control Commands

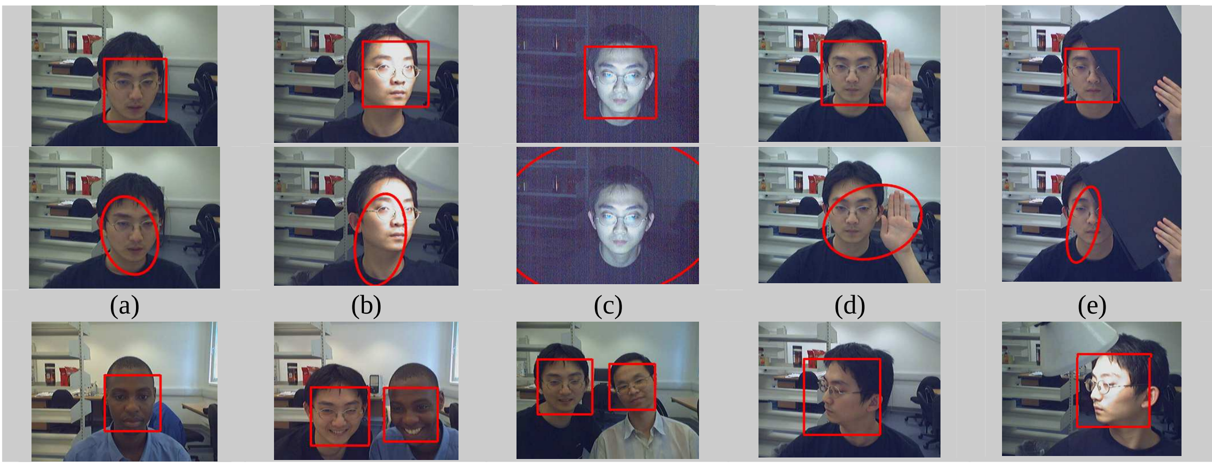
Due to a big diference between the frontal face and the profile face, Adaboost detects frontal and profile head gestures separately. Also, when the user just moves his/her head to look at something, but does not want to move the wheelchair, our HGI should be able to distinguash this situation and avoid generating any unnecessary action. To achieve this, we restrict the on-board camera to focus on the face of the user who sits right in the centre of the wheelchair. If the user's head and face are outside of the central position, our HGI will treat this situation as the user has no intention to control the wheelchair using head gestures. Also, we assume that useful head gestures should have a range of 45 degrees turning angles on each side (up, down, left and right). If the turning angles of head gestures being detected are out of the range, the HGI will notify this and no motion control commands will be generated. If the head gestures are detected within the specified range, RoboChair will act according to the following rules:

- Speed up — if frontal face up is recognised.
- Slow down until stop — if frontal face down is recognised.
- Turn left — if left frontal face or left profile is recognised.
- Turn right — if right frontal face or right profile is recognised.
- Keep speed — if face centre is recognisted.

## 4. Experimental Results

### 4.1 Adaboost Face Detection vs. Camshift Face Tracking

In this experiment, we tested the performance of Adaboost and Camshift algorithms seperatedly under 10 special conditions: (a) Normal (b) Lighting (c) Darkness (d) Face color noise (e) Occlusion (f) Different face color (g) Multi faces with different colors (h) Multi faces with similar colors (i) Profile (j) Profile lighting.



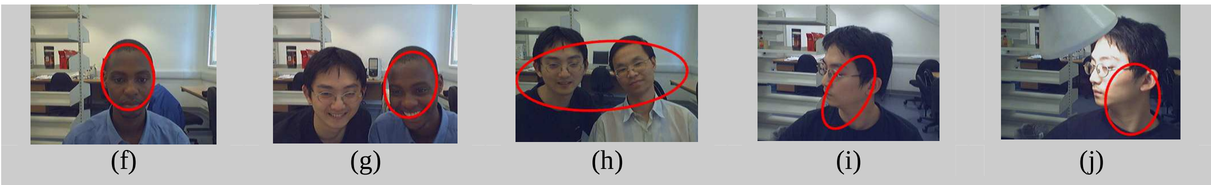



**Figure 6** Robustness of Adaboost Face Detection (row 1, 3) and Inaccuracy of Camshift Face Tracking (row 2, 4)

As shown in Figure 6, the upper image in each condition shows the performance of Adaboost face detection algorithm, and the lower image shows the performance of Camshift face tracking algorithm. It is clear that the performance of Adaboost is very robust under different lighting and noise conditions and faces can be detected very accurately. However, the performance of Camshift is not accurate and robust. In some cases, faces cannot be detected accurately although it is very fast, such as in Figure 6(c), (d), (g) and (j).

### 4.2 Speed Issues for Adaboost and Camshift

Table 1 shows that, under the same condition, the proposed method is much faster than applying Adaboost alone, without taking the USB frame capturing time into consideration. The experiments are finished in WindowXP on the Inter Pentium-M 1.6G Centrino laptop, with the frame resolution 640*480 and the minimum face size 20*20 or 40*40 . The results are statistically calculated during a period of five minutes.

**Table 1** Speed of Adaboost and Camshift

| Method | Minimum Face Size | Time Cost Per Frame |
|---|---|---|
| **Adaboost** | 20*20 | 0.67s |
| | 40*40 | 0.29s |
| **Integration** | 20*20 | 0.22s |
| | 40*40 | 0.11s |

### 4.3 Nose Template Matching for Recognition of Frontal Face Posture

Unlike the left profile and right profile head gestures, further investigation is conducted to recognize the frontal head gestures. Figure 7 shows clearly that our proposed head gesture recognition method is fairly feasible. There are five frontal head gestures to be recognised, namely (a) Center frontal; (b) Up frontal; (c) Down frontal; (d) Left frontal; (e) Right frontal. Each head gesture has two images: the upper image shows the detected face and the lower image shows the recognised head gesture. As can be seen, the small rectangular indicates the face posture based on its relation with the big rectangular box.

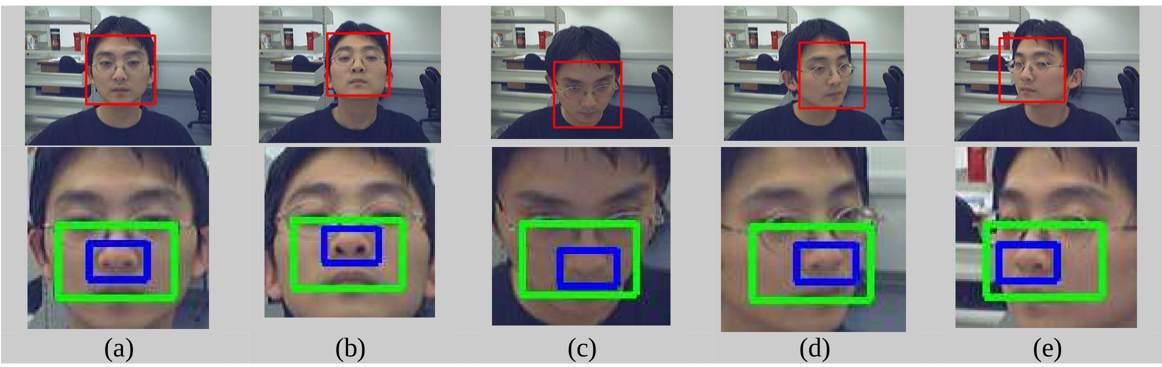



**Figure 7** Frontal Face Posture Recognition by Nose Template Matching

### 4.4 RoboChair Demonstration

Finally, RoboChair demonstrations are presented here to show the feasibility and performance of the proposed HGI framework. Figure 8 presents a sequence of images that our RoboChair is entering and pasing a gate: (a) Start forwarding; (b) Turn right; (c) Entering the gate; (d) At the gate; (e) Passing the gate; (f) Turn right again. It is clear that the proposed HGI is able to recognise the user's head gesture for the hands-free control of the RoboChair. No manual operation is needed, which will make the user life much easier and confortable.

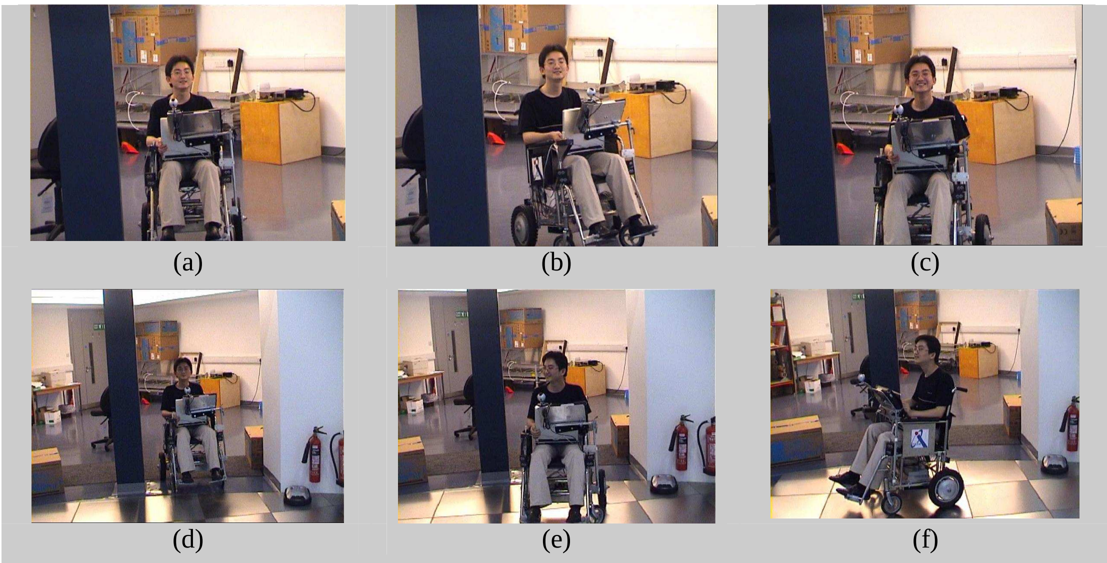



**Figure 8** An Image Sequence to show that the RoboChair is passing through a doorway

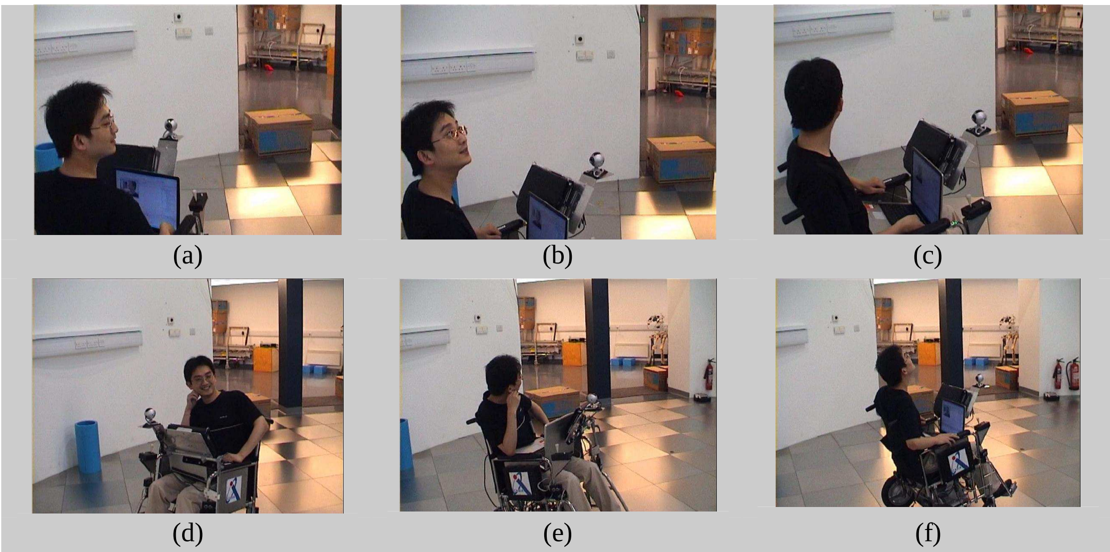



**Figure 9** A image sequence to show that the RoboChair functions well even when some uncertains were added

Figure 9 shows a sequence of images demonstrated by our RoboChair under head gesture control: (a) Turn right (b) Continue turning right with head posture "right up" (c) Turn left; (d) Turn left with hand color noise; (e) Continue turning left with hand color noise; (f) Forward. In this demonstration, some uncertainty such as hand color noise has been delibratly added into the images. When the head gestures are not in normal postures, i.e. the non-vertical head gestures in Figure 9 (d) and (e), the HGI is still able to identify the user's head gesture and control the RoboChair very well, which is very robust. However, the HGI will ignor the head gestures if the user's head is not located in the centre of images or is looking around the surroundings without intention of moving.

## 5. Conclusion and Future Work

This paper describes the design and implementation of a novel hands-free control system for intelligent wheelchairs. The developed system provides enhanced mobility for the elderly and disabled people who have very restricted limb movements or severe handicap. A robust head gesture based interface, HGI, is designed for vision-based head gesture recognition of the RoboChair user. The recognised gestures are used to generate motion control commands to the low-level DSP motion controller so that it can control the motion of the RoboChair according to the user's intention. To avoid unnecessary movements caused by user looking around randomly, our HGI is focused on the central position of the wheelchair and identify useful head gestures.

Our future research will be focused on the extensive experiments and evaluation in both indoor and outdoor environments where uncertainties such as clutered backgrounds, changing lighting conditions, sunshine and shadows may bring the complication to the head gesture recognition.

## Acknowledgement

This research project has been jointly funded by the Royal Society in the UK and the Chinese Academy of Sciences in China.

## References

G. Bradski (1998), "Real-time Face and Object Tracking as a Component of a Perceptual User Interface", Proceedings of the 4th IEEE Workshop on Applications of Computer vision, Princeton, NJ, USA, October, pp. 214-219.

CALL Centre (1994), "Learning through Smart Wheelchairs", University of Edinburgh, U.K., May.

D. Ding and R. A. Cooper (1995), "Electric powered wheelchairs", IEEE Control Systems Magazine, Vol. 25, Issue 2, April, pp. 22-34.

T. Gomi and A. Griffith (1998), "Developing Intelligent Wheelchairs for the Handicapped", Lecture Notes In Computer Science; Vol. 1458, Assistive Technology and Artificial Intelligence, Applications in Robotics, User Interfaces and Natural Language Processing, Springer-Verlag, pp. 150-178.

H. Hoyer and U. Borgolte (1996), "Office Wheelchair with High Manoeuvrability and Navigational Intelligence for People with Severe Handicap", December

Y. Kuno, Y. Murakami and N. Shimada (2001), "User and Social Interfaces by Observing Human Faces for Intelligent Wheelchairs", ACM Int. Conf. Proceeding Series archive, Proc. of the 2001 workshop on Perceptive user interfaces, Orlando, Florida, USA, , pp. 1-4.

S. P. Levine, D. A. Bell, L. A. Jaros, R. C. Simpson, Y. Koren and J. Borenstein (1999), "The NavChair Assistive Wheelchair Navigation System", IEEE Transcations on Rehabilitation Engineering, Vol. 7, No. 4, December, pp. 443-451.

Y. Matsumoto, T. Ino and T. Ogasawara (2001), "Development of Intelligent Wheelchair System with Face and Gaze Based Interface", Proceedings of the 10th IEEE International Workshop on Robot and Human Communication (ROMAN 2001), pp.262-267.

Y. Matsumoto and A. Zelinsky (1999), "Real-time Face Tracking System for Human-Robot Interaction", Proceedings of the 1999 IEEE International Conference on Systems, Man and Cybernetics (SMC99), Vol. 2, pp. 830-835.

M. Mazo, J. C. Garcia, F. J. Rodriguez, J. Urena, J. L. Lazaro and F. Espinosa (2002), "Experiences in assisted mobility: the SIAMO project", Proceedings of IEEE International Conference on Control Applications, Vol. 2, Sept. 18-20, pp. 766-771.

S. Nakanishi, Y. Kuno, N. Shimada and Y. Shirai (1999), "Robotic Wheelchair based on Observations of both User and Environment", Proceedings of IEEE/RSJ Conference on Intelligent Robots and Systems (IROS 1999), Kyongju, Korea, pp. 912-917.

Texas Instruments (2003), “TMS320LF2407, TMS320LC2406, TMS320LF2402 DSP CONTROLLERS”, SPRS0941-APRIL 1999-REVISED September

P. Jia and H. Hu (2005), “Head Gesture Based Control of an Intelligent Wheelchair”, The 11th Annual Conference of the Chinese Automation and Computing Society in the UK (CACSUK05), Sheffield, UK, 10 September, pp. 85-90.

E. Prassler, J. Scholz, M. Strobel and P. Fiorini (1998), “MAid: A Robotic Wheelchair Operating in Public Environments”, Lecture Notes In Computer Science; Vol. 1724, Springer-Verlag, pp. 68-95.

R. S. Rao, K. Conn, S. H. Jung, J. Katupitiya, T. Kientz, V. Kumar, J. Ostrowski, S. Patel and C. J. Taylor (2002), “Human Robot Interaction: Application to Smart Wheelchairs”, Proceedings of IEEE International Conference on Robotics and Automation (ICRA 2002), Washington, D.C., USA, May, pp. 3583-3588.

T. Röfer and A. Lankenau (1998), “Architecture and Applications of the Bremen Autonomous Wheelchair”, In P. P. Wang (Ed.), Proceedings of the 4th Joint Conference on Information Systems, Vol. 1, Association for Intelligent Machinery, pp. 365–368.

R. Simpson, E. LoPresti, S. Hayashi, I. Nourbakhsh and D. Miller (2004), “The Smart Wheelchair Component System”, Journal of Rehabilitation Research and Development (JRRD), Vol. 41, No. 3B, May/June, pp. 429-442.

P. Viola and M.J. Jones (2004), “Robust Real-time Face Detection”, International Journal of Computer Vision, Vol. 57, No. 2, May, pp. 137-154.

Y. Wei (2004), "Vision-based Human-robot Interaction and Navigation of Intelligent Service Robots", PhD Dissertation, Institute of Automation, Chinese Academic of Sciences, Beijing, China.

H. A. Yanco, A. Hazel, A. Peacock, S. Smith and H. Wintermute (1995), “Initial Report on Wheelesley: A Robotic Wheelchair System”, In Proceedings of the Workshop on Developing AI Applications for the Disabled, International Joint Conference on Artificial Intelligence, Montreal, Canada, August.

M.H. Yang, D.J. Kriegman and N. Ahuja (2002),“Detecting Faces in Images: A Survey”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 1, Jan., pp. 34-58