1999 7th IEEE International Conference on

Emerging Technologies and Factory Automation

Proceedings

ETFA'99

October 18-21, 1999

UPC, Barcelona, Catalonia, Spain

Sponsored by

IEEE Industrial Electronics Society

IEEE Robotics and Automation Society

Universitat Politècnica de Catalunya

Edited by: J.M. Fuertes

IEEE UPC ies CENTRECIM CIMNE SIEMENS

f a Wheelchair for Handicapped People by Face Tracking

Bergasa, M. Mazo, A. Gardel, J.C. García, A. Ortuño, A.E. Mendez

Departamento de Electrónica. Escuela Politécnica. Universidad de Alcalá
Campus Universitario s/n. 28805 Alcalá de Henares. MADRID. Spain
T:+34 91 885 6569-40 Fax: +34 91 885 6591
E-mail:bergasa@depeca.alcala.es http://www.depeca.alcala.es

This paper shows a guidance system for an electrical wheelchair for handicapped people by head A color face tracking system has been developed in order to compute head movements of the depending on them, some commands are generated to drive the wheelchair. The system is non- it allows visibility and freedom of head movements. It is able to learn the face features of the automatic initial setup, working even for people of different races. It is adaptive and, therefore, and background changes in inside environments. It has been tested with several users and are given.

1. INTRODUCTION

machine interaction which is very the direction in which a person is information can be required for several tic focus [1], teleconferencing with sation [2], face identification in [3], gaze driven panorama image viewer systems [4], lips readers [5], and mobility of disabled people [6], etc.

Department of University of Alcala has more than 6 years, on artificial means y of handicapped people. Nowadays, is being developed, within SIAMO System for Assisted Mobility) [7], in lti-functional wheelchair for disabled (Figure 1). This project includes an , by head movements, for cases of At present, a wheelchair prototype is guidance method. A 3D simulator has well, to help the users in adapting system in a safer way.

the global system architecture and the methodology followed in its design. experimental results are given and, clusions about its performance have

STEM ARCHITECTURE

e general system architecture. Through camera, placed in front of the user, uired. These images are digitized by d loaded in a PC Pentium memory. To the image, an original skin color thm has been used, called UASGM (Unsupervised and Adaptive Skin Gaussian Model) [8]. Then, a face tracking is applied to the skin blob and, depending on it, a high control state machine generates wheelchair's linear and angular speed (V_{cmd}, Ω_{cmd}).

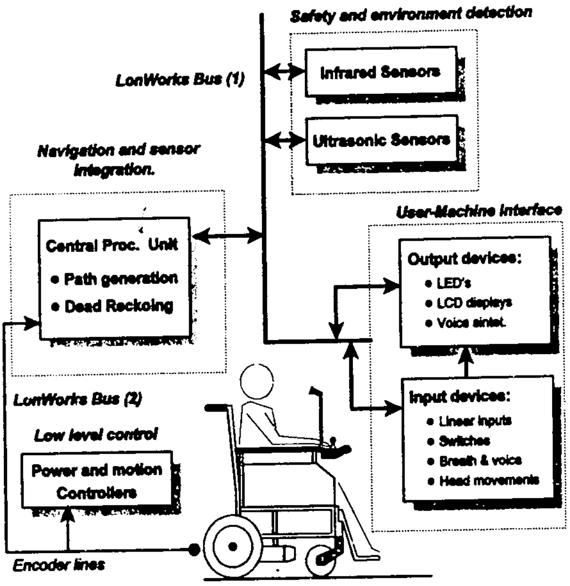


Figure 1. SIAMO Project

Applying the Kinematic model, linear and angular speed become angular speeds for every wheel ($\omega_{r,cmd}$ $\omega_{l,cmd}$) and they are sent to the low level control. In this level a PI controller has been designed to control the velocity of each wheel.

Next, the main parts of the system architecture will be explain.

2.1. Skin Segmentation

The UASGM method segmentates any person's skin, even of different races, under changing light conditions and

$10.00 ©1999 IEEE

random backgrounds. To do this, a stochastic adaptive model of skin colors in a normalised RG (Red, Green) color space has been used.

The model is initialized by a clustering process. This divides the chromaticities of an image in a number of classes (k) between one and a maximum value (k_{max}). At each step, the k cluster centers are estimated using an approximate color histogram. After that, these centers are adjusted employing a competitive learning strategy (Vector Qantization algorithm [9]) in a closest center sense. Finally a clustering quality factor is computed for each topology. This factor for k classes is given by (1), where tr[·] is the trace of a matrix and S_w and S_B denotes, respectively, the within-cluster and between-cluster scatter matrices.

$$F_k = tr[S_W^{-1} S_B] \qquad 1 \le k \le k_{max} \tag{1}$$

$$S_B = \frac{1}{kM} \sum_{l=1}^{k} M_l [m_l - m_0]^T [m_l - m_0] \tag{2}$$

$$S_W = \frac{1}{k} \sum_{l=1}^{k} \frac{1}{M_l} \sum_{i=1}^{M_l} [x_i - m_l]^T [x_i - m_l] \tag{3}$$

$$m_0 = \frac{1}{M} \sum_{i=1}^{M} x_i \quad ; \quad m_l = \frac{1}{M_l} \sum_{i=1}^{M_l} x_i \tag{4}$$

In these equations, k is the number of clusters, M_l is the number of pixels in the l^{th} cluster, x_i is a color pixel in the l^{th} cluster, m_l is the mean of the l^{th} cluster, m_0 is the mean of all of the feature vectors and M [illegible] number of pixels to be clustered.

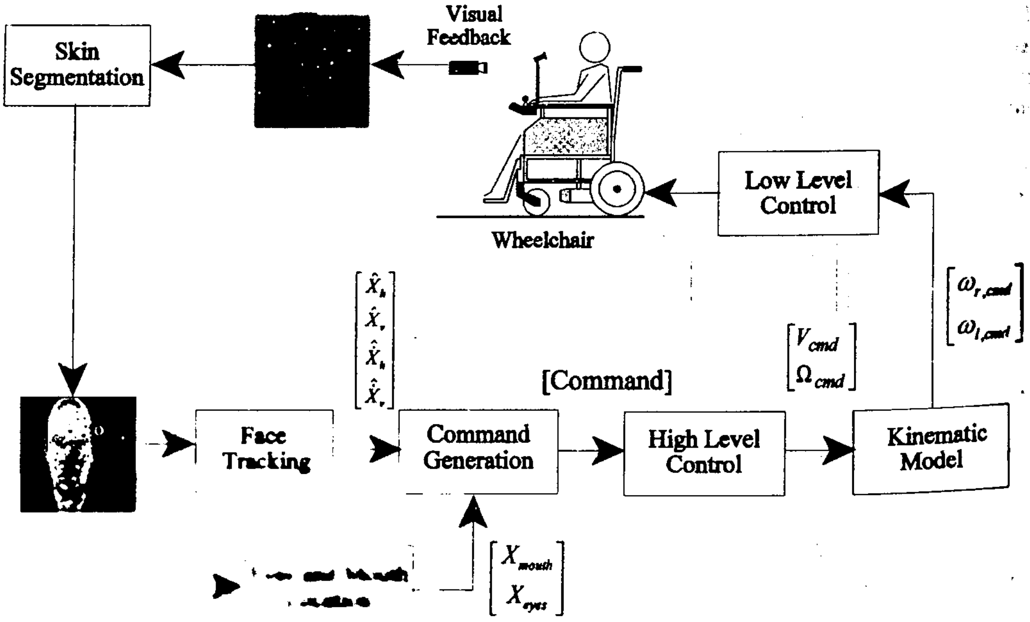
The process is repeated adding a new [illegible] each step until the maximum number of [illegible] maximum quality factor gives the number [illegible] best fits the histogram distribution. With [illegible] classes the skin cluster is located [illegible] distance between the center of the clusters [illegible] skin color position. Then, the skin class is [illegible] Gaussian function $N(m_S, C_S)$ and the [illegible] model are estimated by a linear combinatio[illegible] ones using the maximum likelihood criteri[illegible]

The estimated mean vector, $\hat{m}_S$, and [illegible] covariance matrix, $\hat{C}_S$, will be calculated [illegible]

$$\hat{m}_S = \sum_{l=1}^{v} \alpha_l R_l$$

$$\hat{C}_S = \sum_{l=1}^{v} \beta_l R_l$$

where, R_l are the previous mean vectors[illegible] previous covariance matrix; α_l are the coeffic[illegible] mean prediction and β_l are the coeffici[illegible] covariance matriz prediction. To calculate the [illegible] coefficients for the prediction, we have used the [illegible] proposed by Anderson [10]. The basic [illegible] algorithm is iteratively estimate the coeffic[illegible] coeficients $\beta_l^{(i)}$ independently, where the [illegible] denotes the ith iteration.The iteration [illegible]



[illegible] cture of the guidance by head movements

s in the order of $\alpha_l, \hat{m}_S$, C, $\beta_l, \hat{C}_S$,
ion is terminated if max($|\beta_l^{(i)} - \beta_l^{(i-1)}|$,
is an error parameter defined by the

the ith iteration will be:

$$)^{-1} R_s\Big)^{-1} R_{jS}^T\left(\hat{C}_S^{(i-1)}\right)^{-1} m_S \quad j = 1,\dots,v \qquad (7)$$

$$^{(i)} = \sum_{l=1}^{v} a_l^{(i)} R_l \qquad (8)$$

$$m_S)(x_l - m_S)^T + (x_l - \hat{m}_S^{(i)})(x_l - \hat{m}_S^{(i)})^T \qquad (9)$$

$$)^{-1} S_l\left(\hat{C}_S^{(i-1)}\right)^{-1} S_l \beta_l^{(i)} = \qquad (10)$$
$$)^{-1} S_l\left(\hat{C}^{(i-1)}\right)^{-1} C_S^{(i)} \quad j = 1,\dots,v$$

$$\hat{C}_S^{(i)} = \sum_{l=1}^{v} \beta_l^{(i)} S_l \qquad (11)$$

mber of skin pixels.

the image, the Gaussian skin model is
s value is higher than a threshold, it can
the pixel belongs to the skin class. An
d proportional to the trace of the
ce matrix has been applied in the way

$$Th = K_{Th} tr\left[\hat{C}_S\right] \qquad (12)$$

has been shown that the results,
UASGM, respecting to the optimum
in the clustering process, improve the
the following methods: FHV (Fuzzy
Evidence density and MDL (Minimum
gth), and equalize these others: MML
age Length) and GMM (Gaussian
), reducing the computing time in every
ethods are based on EM (Expectation-
algorithm. An explanation of such
data sets used in the comparison can be
rence[11].

d, with UASGM method, we have
ults than applying the GLVQ-F (Fuzzy
ing Vector Quantization) one [12], that
wn FCM (Fuzzy C-Means) algorithm.
the stochastic model shown in [4] as
lization of the model suitable for the

2.2. Face Tracking

The use of the estimation theory for tracking in computer vision is well known [13]. On the skin blob some parameters are calculated to track the face: center of gravity (x,y), horizontal (h) and vertical (v) size of the face, being able to obtain the face position and orientation.

A zero-th order Kalman filter is used to estimate two independent state vectors: one of them for the horizontal variation (X_h =(x,h)) and the other one for the vertical variation (X_v=(y,v)). Two independent state vectors have been utilized because, in our application, users can do only horizontal and vertical head rotation movements. Then can be taken into account that horizontal and vertical movements are independent: horizontal size depends on the 'x' center position and vertical size on 'y' center. Horizontal and the vertical size of the face are calculated as two parameters, because the aspect ratio of the face can change with rotations. The state vectors (X_h, X_v) and their covariance matrices (C_{Xh}, C_{Xv}) are estimated in a recursive process composed by three phases: predict, match and update.

2.3. Commands generation

Estimated state vectors $(\hat{X}_h, \hat{X}_v)$ and their derivatives $(\dot{\hat{X}}_h, \dot{\hat{X}}_v)$ are inputs to a command generation state machine.. Each state codifies one of the following commands: *turn right, turn left, increase speed, decrease speed and no command*. State transitions of the machine are achieved analysing the activation of some fuzzy conditions of input variables, based on thresholds. These are computed in the initial setup. Likewise, using the information given by eyes and mouth position (X_{eyes}, X_{mouth}) the special commands *on/off* and *forward/backward* will be obtained. These positions are calculated by the analysis of the hollows on the skin blob and imposing some geometrical restrictions. Hollows appear in the blob because in the face there are some features like: eyes, mouth, eyebrowes, etc, that have different colors related with the skin.

Figure 3 shows the actions that generates the commands. For that, we have followed a criterion of simplicity in the fulfilment of these actions and in the robustness of their detection. Therefore, if user turns the head to the right the wheelchair will turn to the same direction. This happens as well if he turns his head to the left. Head rising involves the increment of wheelchair speed, and when the user bows it, this will decrease the speed.

Every time he winks an eye the wheelchair changes its state *on/off*; when lips are hidden it changes again to the command *forward/backward*. [illegible] special actions for, at least, two [illegible] the ‘special command’ activation. Doing this, wrong commands are avoided as a consecuence of normal blinks. *On/off* commands allows to user activate or deactivate the system by themselves, so when it is in the off state, it can make any kind of movement being secure on the fact that no other command will be activated.

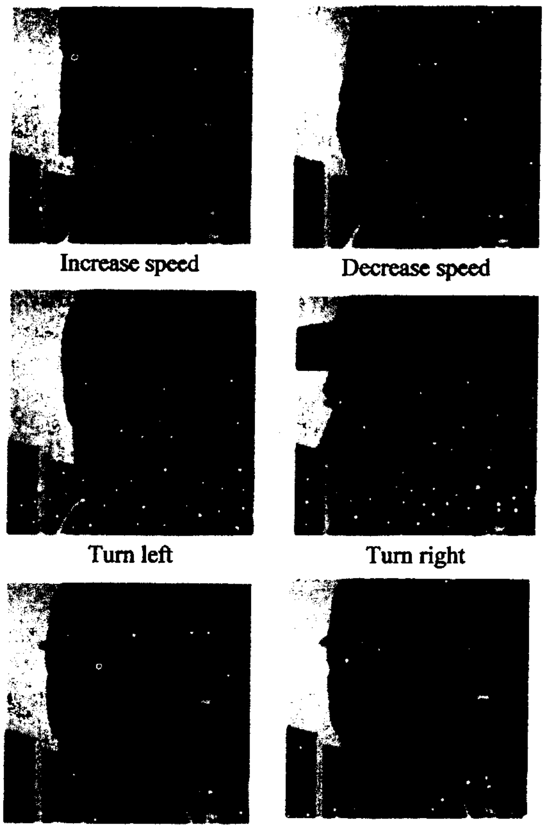

Increase speed
Decrease speed
Turn left
Turn right
On/Off
Forward/Backward

Figure 3. Actions that generate the commands

Table 1 presents the way in which the system recognizes the action that user has made, paying atention to some fuzzy conditions in the evolution of the state variables. Analysing the activation of these fuzzy conditions and their derivatives the commands are generated.

In figure 4 we can see an example of the right command activation for a sequence of right progressive rotations made by the user (30°, 60°, 90°). Here, the evolution of the state variables and their derivatives is shown.

Then, the values of the state variables for the “no command” position can be regarded as (x_R, y_R; h_R, v_R). These last variables are used as well as reference positions in the calculation of the fuzzy conditions to do the [illegible] transitions.

[illegible] the system is able to detect all the [illegible] the 30° one, in the 250 samples, as the [illegible] made very slowly. This happens because, working with the derivatives of the state the actions speed are taken into account. It is not to do one action, but it must be executed quickly. way some unvoluntary wrong actions are elimin

I: Increase, D:Decrease, E:Equal

	State Variables			
Actions with the head	h	x	v	y
No Command	E	E	E	E
Turn Right	I	D	D	I
Turn Left	I	I	D	I
Rising	E	E	I	I/D
Bows	E	E	D	I

Table 1. Fuzzy conditions in the evolution of the state variables for the different actions.

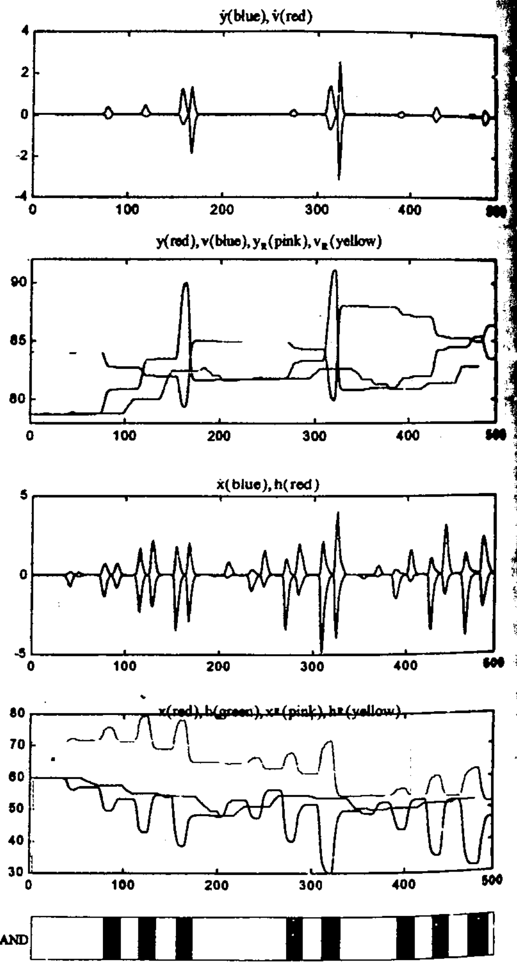
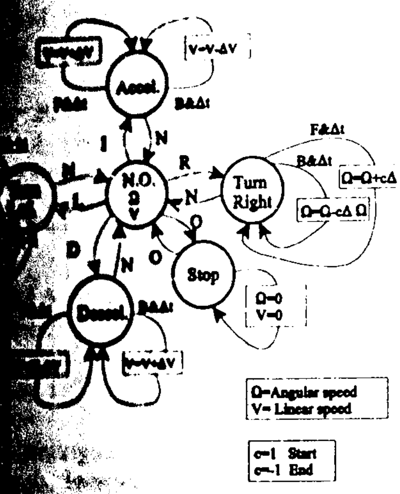


Figure 4.Examples of right command generation

2.4. High level Control

Commands are sent to another state machine whic implements high level control and will generate the line mand to the wheelchair (V_{cmd}, of time. It works as it's shown in has six states: "Stop", "No ", "Decelerate", "Turn right" and

stop position (V_{cmd} =0, Ω_{cmd}=0) "Forward", "Backward", "Left", speed a predetermined initial value a final value, positive or negative, . Then "Increase speed" and mands increase or decrease linear ction of time and depending on the ward/backward), up to certain pre- e command *on/off* allows to stop the the process.



level control state machine

YSTEM LAYOUT

that there is a visual feedback loop, reacts according to the current stance, if the system detects a right wheelchair will turn to the right until

d on board the wheelchair presents own in figure 6 [14]. A LonWorks network has been developed, where s can easily be added because of e of its design.

ven by a pair of DC-motors, one in Neuron-Chip drives an H-bridge ecessary to achieve the intended el. It also provides the support twork funtions:

des by reading the encoders

- Programmable in Neuron-C to design different control actuations. A PI controller is currently being used.

- It implements the seven layers of ISO comunication.

The driver nodes receive the commanded angular speeds coming from the high level control driven by the vision process. In this node the vision algorithm is processed on a PC Pentium II 400 Mhz NT 4.0. This node employs an interface card to Lon net, called PCLTA, also based in a Neuron-Chip. LonWorks network access is rather simple, Neuron C issues network variables data types allowing simple programming for information exchange through the bus.

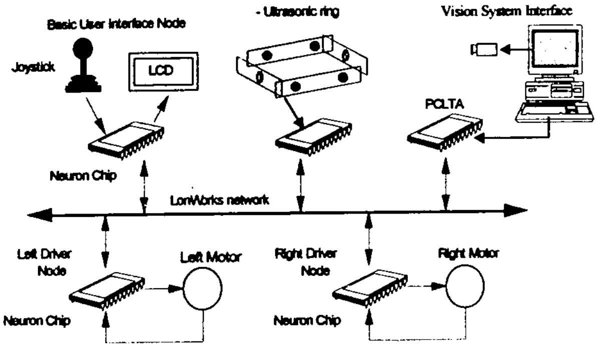


Figure 6. System layout

We have introduced, as well, advisory sounds when the commands are activated for easy control of the user through acustic feedback.

The wheelchair prototype is also provided with an ultrasonic ring to increase safety during navigation and a user interface composed by a joystick and a LCD.

4. EXPERIMENTAL RESULTS

The vision system is able to process up to 10 images per second, with a resolution of 128x128 pixels. During the testing stage it great robustness exhibited in the velocity commands. However, some details are still to be fixed on the special commands. In order to increase the system controllability two switches have been included. One of them is used to activate the on/off command, while the second is intended to activate the forward/backward command. 10 commands per second are issued to the low level controller. The maximum wheelchair velocity was set to 1 m/s.

The system was tested for 5 different users in the Electronics Department's labs after some training on the simulator. In figure 7 we show the image of one of the users during a test [illegible] performed by that user during one of the tests [illegible] in figure 8. The evolution [illegible] wheelchair's velocities [illegible] test lasted 100 seconds, taking 5 samples per second. In this figure the accelerations of the wheelchair during straight sections can be clearly apreciated. Also, the wheelchair decreases its speed before performing a curve. In this case, six left turns were executed, yielding great angular velocity peaks in the figure.

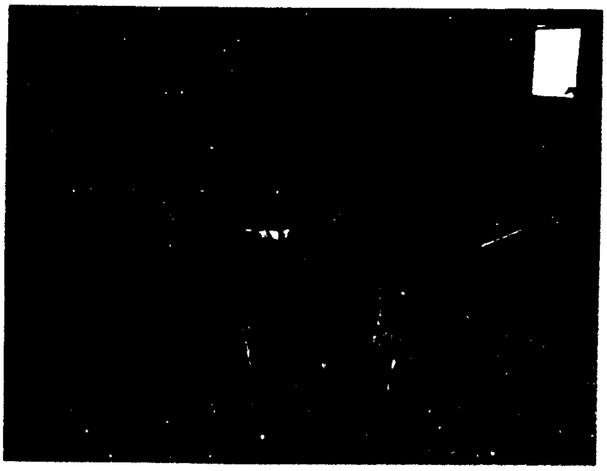


Figure 7. User doing the test

5. CONCLUSIONS AND FUTURE WORK

The conclusions obtained by the users after performing the test on the navigation system are presented bellow:

- It is non-intrusive, as it is passive, and there is no need for additional elements.

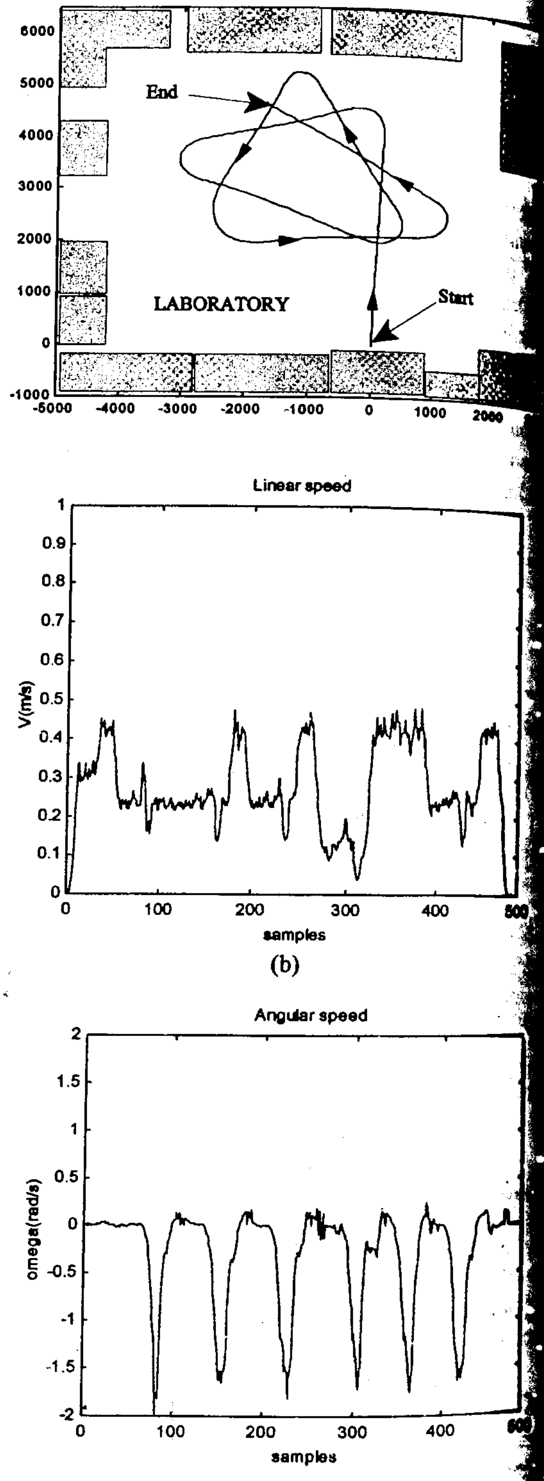
- Guidance complexity is decreased as more training is perfomed. We must also take into consideration that the camera is 80 cms in front of the user and, therefore, the system requires certain space for safe manoeubring.

- The simple commands set and the wheelchair response allow for easy controllability in environments with not too many obstacles.

- Audible feedback is included to ensure proper command acknowledgement.

- It works well in indoor environments, where suitable ilumination can be provided, decreasing the performance as light conditions get poorer. In outdoor environments there is no uniform ilumination (shadows, direct sun light, etc), decreasing also the system capabilities.

Some future guidelines are to embedded the system on a hardware platform and to develop a 3D model of the head in order to robustly obtain the gaze direction (not depending on the particular user) and perform the wheelchair control according to it.



(b)

(c)

Figure 8. Example of the guidance

6. ACKNOWLEDGEMENTS

This work has been financed by the CICYT (Sp Interministerial Science and Technology Comm through the SIAMO project TER96-1957-C03-01.

7. REFERENCES

[1] Test of Canon UC-X1Hi, HIFI electronik, 1995.

[2] De Silva, K. Aizawa, M. Hatori. "Detection and tracking of facial features by using edge pixel counting and deformable circular template matching. IEICE TRANS INF&SYST. VOL.E78-D, No 9, pp.1195-1207, September 1995.

[3] Kah-Kay Sun and Tomaso Poggio. "Example-based learning for view-based human face detection". IEEE Transactions on Pattern analisis and machine intelligence, Vol 220, n°1 , January 1998.

[4] Stiefelhagen, J. Yang and A. Waibel "A model based gaze tracking system". IEEE International Joint Symposia on Intelligence and Systems-Image, Speech&Natural Language Systems. 1997

[5] R. Meier, R. Stiefelhagen and J. Yang "Preprocessing of visual speech under real word conditions".Proceedings of European Tutorial & Research Work Shop on Audio-Visual Speach Processing, 1997

[illegible] and A. Zelinsky, "Robust real-time face tracking and gesture recognition". Proceedings of IJCAI'97, International Joint Conference on Artificial Intelligence, August 1997.

[illegible] M. Mazo, F. J. Rodríguez, J. Ureña, J. C. García, J. L. [illegible] R. García."Integral System for Assisted [illegible] 2nd International Workshop on Intelligent [illegible] 98)JCIS'98 Proceedings. Editor: Paul P. [illegible]61-364. Durham (USA), 1998.

[illegible], A. Gardel, M.Mazo, M. A. Sotelo. "[illegible] using an Adaptive Skin Color Model". [illegible]tional ICSC Symposia on Intelligent industrial automation (IIA'99) ans Soft Computing (SOCO'99). Genova. Italy. 1999. (Accepted)

[9] T. Kohonen, "Self-Organizing Maps", Springer-Verlag, 1997

[10] T.W. Anderson, "Asymptotically efficient estimation of covariance matrices with linear structure". The Annals of Statistics, Vol 1, N° 1, 135-141, 1973.

[11]S.J. Roberts, Dirk Husmeier, Iead Rezek and William Penny. "Bayesian Approaches to Gaussian Mixture Modeling". IEEE Transactions on Image Processing, Vol. 20, N°11, pp.1133-1142,November, 1998

[12]N.B. Karayiannis, J.C. Bezdek, N.R. Pal, R.J. Hathaway and P. Pai,"Repairs to GLVQ: A New Family of Competitive Learning Schemes", IEEE Transactions on Neural Networks, Vol. 7, N° 5, pp: 1062- 1071, September 1996.

[13] Crowley J.L, Coutaz J. "Vision for Man Machine Interaction" EHCI'95. Grand Targhee, August 95.

[14] J.C. García, M. Marrón, J.A. García, M.A. Sotelo, J. Ureña, J.L. Lázaro, F.J. Rodríguez, M. Mazo, M. Escudero. "An autonomous wheelchair with a LonWorks network based distributed control system".In Proc. of II Field and Services Robotics (FSR'97). Canberra. Australia. December 1997.