A hand-based visual intent recognition algorithm for wheelchair motion

T. Luhandjula[1,2], K. Djouani[1], Y. Hamam[1], B.J. van Wyk[1], Q. Williams[2]

1. French South African Technical Institute in Electronics at the Tshwane University of Technology, Pretoria, RSA
2. Meraka Institute at the Council for Scientific and Industrial Research, Pretoria, RSA

tluhandjula@gmail.com, vanwyk@gmail.com, hamama@tut.ac.za

Abstract. **This paper describes an algorithm for a visual human-machine interface that infers a person's intention from the motion of the hand. Work in progress shows a proof of concept tested on static images. The context for which this solution is intended is that of wheelchair bound individuals whose intentions are the direction and speed variation of the wheelchair. Results show that the symmetry property of the hand in motion can serve as an intent indicator.**

Keywords**: Human-machine interface, intention, intention curves, intention detection, symmetry curves.**

I. INTRODUCTION

ONE of the challenges facing the task of realising an enabled environment where people with disabilities and the aged are independent and can therefore be active and contribute in society, is to develop systems that can assist them in performing the tasks they wish to carry out without other people's assistance [1], [2]. Good performance in a team/society environment is heavily conditioned by the awareness of people's intention within society [3] and therefore human-machine interaction where the machine has a support role, requires that the intention of the user is well understood by the machine. This intention awareness capability is important for Human-Machine Interaction (HMI) and for the more specific area of the enabled environment.

Plan recognition is the term generally given to the process of inferring intentions from actions and is therefore an important component of HMI. The literature shows that the plan recognition community has spent some interest in probabilistic network based approaches [4]. Although plan recognition is a well-known feature of human collaboration, it has proven difficult to incorporate into practical human-computer collaboration systems due to its inherent intractability in the general case. Lesh et al. [5] describe a plan recognition algorithm which is tractable by virtue of exploiting properties of the collaborative setting, namely: the focus of attention, the use of partially elaborated hierarchical plans, and the possibility of asking for clarification. It has been shown that plan recognition can allow more efficient and natural communication between collaborators, and can do so with relatively modest computational effort. These are important results as Human-computer collaboration provides a practical and useful application for plan recognition techniques.

One frequent human-machine interaction may be found in the context of a person with a physical disability whose mobility is constrained by a wheelchair. There are some solutions found in the literature such as [6], where a new human-machine interface for controlling a wheelchair by head movements is presented. The position of the head is determined by the use of infrared sensors. The placements of the infrared sensors are behind the head of the user so that the field of view is not limited. Jia and Hu [7] propose an integrated approach to real time detection, tracking and direction recognition of human faces, which is intended to be used as a human-robot interaction interface for the intelligent wheelchair. It is implemented using Adaboost face detection and a canonical template matching to tell the nose position, therefore giving an indication of the position of the head and the direction the wheelchair must take. However these solutions are specifically dedicated to Head motion detection and are not suitable for people who would rather use their hands but not a joystick.

Many other platforms have already been devised to help people in their daily manoeuvring tasks: OMNI, Bremen autonomous wheelchair, RobChair, Senario, Drive Assistant, VAHM, Tin man, Wheelesley (stereo-vision guided), and Navchair (sonar guided) [8]. These systems are based on "shared control" where the control of the wheelchair or any other assistive device is shared between the user and the device. Often the developed architectures consist of different algorithms that each realise specific assistance behaviour, "such as drive through door", "follow corridor" or "avoid collision". The presence of multiple operating modes creates the need to choose from them, and therefore makes the user responsible for selecting the appropriate mode, which might in some instances be an inconvenience.

In this paper, an alternative visual solution is proposed that infers the intention a subject using his hand in motion from the dorsal (footer) view as object of interest. The application intended for this solution is that of wheelchair bound individuals where the intentions are the direction and the speed variation intended by the subject for the wheelchair. This hand-based solution is proposed as an alternative to systems using joysticks, pneumatic switches and the motion of the head [1], [2]. The solution is non-intrusive and does not require the multiplicity of operating modes. This paper provides a proof of concept as a contribution to the task of realising a Human System Interaction solution for the enabled environment allowing people with disabilities and the elderly to be more independent and as a result more active in society.

II. METHODS

The type of data used for intention inference is visual: A sequence of images is captured by a CCD camera with a hand in motion from its dorsal view as object of interest. The pre-processing step of detecting the hand in the field of view is not part of this work, and therefore it is assumed that the hand is already detected. No visual aid or marker is provided on the hand to analyse the motion in the sequence. The hand performs two types of motion: Rotation and Vertical Motion to indicate an intention in direction and speed variation of the wheelchair respectively. These intentions of interest become the commands for the motion of the wheelchair as described in Table 1 below:

TABLE 1: MAP OF HAND MOTION TO INFERRED INTENTION

Motion of Hand	**Inferred Intention**
Direction	
Rotation to the Right	Move to the Right
Rotation to the Left	Move to the Left
No Rotation (Centred hand)	Move Straight
Speed Variation	
Vertical motion Down	Increase speed
Vertical motion Up	Decrease speed
No Vertical motion (Centred hand)	Move at constant speed

For direction classification the method consists in extracting a symmetry curve from the input image with the hand as object of interest, and a classifier (in this work the classification task has been performed using a neural network and a support vector machine) is used to distinguish between the symmetry curves associated the different positions of the hand. A sequence of these positions is used to classify between the different directions namely “going straight”, “going left” and “going right”.

For speed variation classification the method consists in extracting a symmetry curve from the input image with the hand as object of interest and the centre of gravity of the symmetry curve is calculated. A sequence of these centres of gravity is used to classify between the different motions namely “going at the same speed”, “going faster” and “going slower”.

The rest of this section describes the approach used for intention recognition. This approach assumes that the hand is the only object in the camera’s field of view and therefore no pre-processing steps (such as detection and tracking) are described in this work. The approach consists of a symmetry-based approach [1], [2] that extracts symmetry curves of the hands assuming that different positions of the hand gives different symmetry curves as it did for the face in previous work in [1] and [2].

A. Symmetry-based approach

In previous work [1], [2] the merit of using a symmetry-based approach for intent recognition has been established. Though hands are not as symmetrical as faces, the underlying assumption is that a human hand from its dorsal view displays different symmetry properties as it moves vertically or rotates. These properties can be used to detect the motions (rotation and vertical motion) undertook by the hand: Given a $X \times Y$ greyscale image *I*, the symmetry is calculated using the following expression:

$$f(y) = \sum_{\omega=1}^{k} \sum_{x=1}^{X} | I(x, y-\omega) - I(x, y+\omega) | \qquad (1)$$
$$\forall y \in [k+1 \;\; Y-k]$$

The symmetry-value *f(y)* of each pixel-row in the image is evaluated by taking the sum of the differences of two pixels at a variable distance ω = [1 k] from it on both sides making the pixel-row the centre of symmetry. This process is repeated for each column and the resulting symmetry-value is the summation of these differences. The symmetry curve is composed of these symmetry values calculated for all the pixel-row in interval [*k+1 Y-k*]. In earlier work [1], [2], it has been shown that the distance that gives a more discriminative symmetry curve among the different positions is given by *k* = 35 for both images of size 240x200 and 145x250 of the hand in rotation and vertical motion respectively. The reason for this size difference is that the vertical motion of the hand requires more vertical space for the hand to remain in the field of view than in the case of the hand in rotation. This affects the position of the camera as well as the size of the field of view required for recognition.

B. Vertical motion: Classification of individual positions of the hand

The symmetry curves’ centre of gravity (COG) may be used to classify the different positions of the hand in vertical motion. The centre of gravity is calculated as the point in the curve at which all the values of the curve can be considered centred:

$$C = \frac{y_1 f(y_1) + y_2 f(y_2) + ... + y_n f(y_n)}{f(y_1) + f(y_2) + ... + f(y_n)} \qquad (2)$$

Here the symmetry curve is defined by the function $f : y \rightarrow f(y)$ with *f(y)* given by (1) and *I* is a 145x250 greyscale image frame. Fig. 1 shows the position of the centre of gravity on the symmetry curve for different positions of the hand in vertical motion as an indication of the position of the hand.

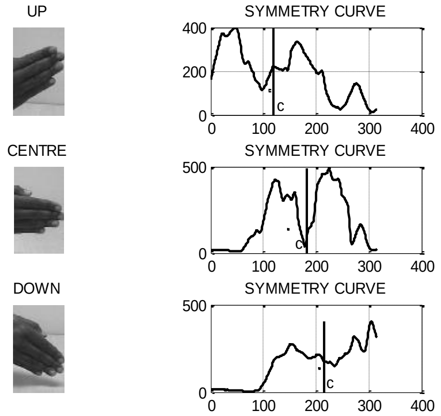

Fig. 1. Three positions of the hand in vertical motion with associated symmetry curves. The positions of the symmetry curves' COG are indicated by the vertical lines.

Two approaches have been used to classify these different positions into three categories (centre, up and down): The difference of means of the COGs, and the mean and standard deviation of the COGs in Gaussian distributions'. This is justified by the different positions of the centres of gravity for the different positions of the hand as shown in Fig. 1.

1) Difference of means

In this approach, the mean of the centres of gravity is calculated for the training set of each category. The difference between the centre of gravity to be classified and the mean of each class is calculated, and the class corresponding to the mean where the difference is the smallest is chosen:

TABLE 2: DIFFERENCE OF MEANS APPROACH FOR COG CLASSIFICATION

Let $\mu_{centre}, \mu_{up}, \mu_{down}$ be the means of the different classes and C the centre of gravity corresponding to the symmetry curve and therefore a given image to be classified:

$$d_1 = | C - \mu_{centre} |,\ d_2 = | C - \mu_{up} |,\ d_3 = | C - \mu_{down} |$$

$$d = \min([d_1, d_2, d_3])$$

$$\begin{array}{ll} if\ d == d_1 & class 1 \\ elseif\ d == d_2 & class 2 \\ else & class 3 \end{array}$$

2) Statistics in a Gaussian distribution

The mean and the standard deviation of the centres of gravity are calculated for the training sets of the three different categories. They are associated to Gaussian distributions along with the given centre of gravity to be classified. The resulting highest probability measure among the three cases corresponds to the class the given centre of gravity belongs to.

TABLE 3: MEAN AND STANDARD DEVIATION IN A GAUSSIAN DISTRIBUTION APPROACH FOR COG CLASSIFICATION

Calculate:

$P_1 = \mathcal{N}(\mu_{Centre}, \sigma_{Centre})$,

$P_2 = \mathcal{N}(\mu_{up}, \sigma_{up})$,

$P_3 = \mathcal{N}(\mu_{down}, \sigma_{down})$

where $P_i = \mathcal{N}(\mu_{class}, \sigma_{class})$

$$= \frac{1}{\sqrt{2 \times \pi}\sigma_{class}} \exp\{-\frac{(C - \mu_{class})^2}{2\sigma_{class}^2}\},$$

C is the centre of gravity, class = {Centre, Right, Left}, μ_{Class} and σ_{Class} *are the (means and standard deviations of the centres of gravity in the training set:*

$$P = \max([P_1, P_2, P_3])$$

$$\begin{array}{ll} if\ P == P_1 & class\ 1 \\ elseif\ P == P_2 & class\ 2 \\ else & class\ 3 \end{array}$$

C. Vertical motion: Intention detection for speed variation

The task of intent recognition involves the detection of the direction the subject intends to take and the speed variation he wishes to perform by looking at the motion of the hand. This section describes the recognition of the hand's vertical motion indicating intent of variation in speed (increase and decrease for a down and up motion respectively). The time sequences of the symmetry curves' centre of gravity give 15-elements vectors referred to in this work as "intention curves" and are used to recognize the different possible intentions namely: constant speed, decrease and increase in speed.

Let E = {I_i : I_i is the ith frame in a sequence of N = 15 frames}. $\forall\ I_i \in E$, C_i is the centre of gravity of the symmetry curve (1) associated to I_i. The resulting intention curve designated by the vector V = {C_i: i = 1…15} is shown on Fig. 2 and 3 for each scenario. In Fig. 2 both up and down vertical motions are captured from the centre and in Fig. 3 the up vertical motion is captured from down to the centre while the down vertical motion is captured from up to the centre. These three types of motion (the hand remaining centred, the vertical motion of the hand up and the vertical motion of the hand down over time) exhibit different patterns and can therefore be easily classified.

Given the level of clarity on the difference between the intention curves associated to the three different classes of motion as shown in Fig. 2 and 3, decision rules as described in Table 4 may be used for classification. They are also based on 'difference of means' and 'mean and standard deviation in Gaussian distribution' approaches (refer to Tables 2 and 3). Tables 6 and 8 summarize the results for individual position and intent classification respectively.

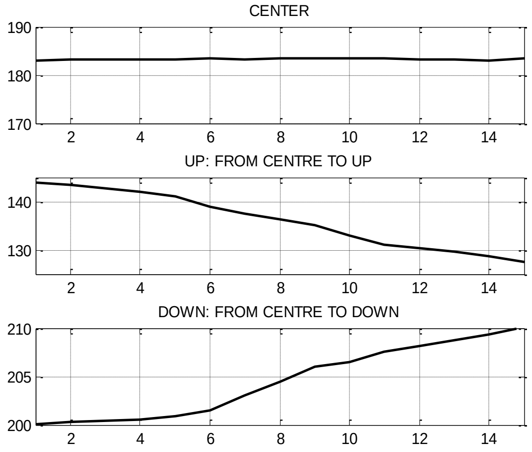

Fig. 2. Time sequence of the symmetry curves' COGs for hands in vertical motion. Motion captured from centred position.

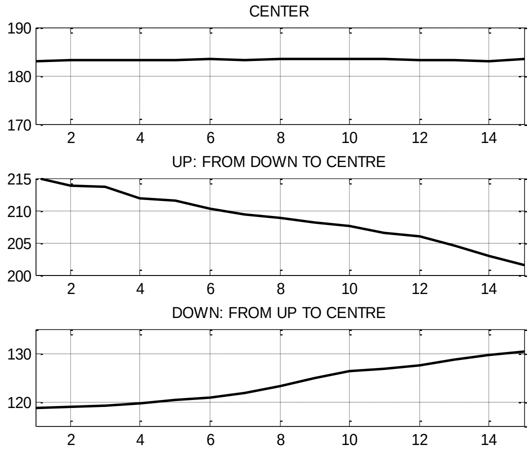

Fig. 3. Time sequence of the symmetry curves' COGs for hands in vertical motion. Motion captured from down and up positions for up and down motions respectively.

TABLE 4: DECISION RULE FOR CLASSIFICATION OF INTENTION CURVES. IT CAN BE IMPLEMENTED USING EITHER DIFFERENCE OF MEANS OR MEANS AND STANDARD DEVIATION IN A GAUSSIAN DISTRIBUTION.

Let V be the intention curve to be classified:

Initialisation

A = 0; B = 0; (Initialise A notifying a decrease and B notifying an increase)

$\forall i \in \{x : x \geq 1 \text{ and } x \leq length(V) - 1\}$, $D = V(i) - V(i+1)$

If $D > 0$ $\quad A = A + |V(i) - V(i+1)|$ *(notifying a decrease in value of C, by adding the extent to which there Is a decrease to the value of A)*

If $D < 0$ $\quad B = B + |V(i) - V(i+1)|$ *(notifying an increase in value of V, by adding the extent to which there is a decrease to the value of B)*

Classification

Let μ_{Class}, σ_{Class} *be the statistics (means and standard deviations) of the difference between A and B in a training set for each class:*

Class = {Centre, Up, Down}; *and* $n = \{1, 2, 3\}$.

Difference of means:

$d_n = |(A - B) - \mu_{Class}|$, $d = \min([d_1\ d_2\ d_3])$

If $(A > B$ *and* $d = d_1)$ *or* $(A < B$ *and* $d = d_1)$
Intention = Going Straight

If $A > B$ *and* $d = d_2$ *Intention = Going Right*

If $A < B$ *and* $d = d_3$ *Intention = Going Left*

Statistics (mean and standard deviation) in Gaussian distribution:

Calculate: $P_n = \dfrac{1}{\sqrt{2 \times \pi}\,\sigma_{class}} \exp\{-\dfrac{((A-B)-\mu_{class})^2}{2\sigma^2_{class}}\}$

$P = \max([P_1\ P_2\ P_3])$

If $(A > B$ *and* $P = P_1)$ *or* $(A < B$ *and* $P = P_1)$
Intention = Going Straight

If $A > B$ *and* $P = P_2$ *Intention = Going Right*

If $A < B$ *and* $P = P_3$ *Intention = Going Left*

D. Rotation: Classification of individual positions of the hand

The symmetry curves associated to images with the hand in rotation as object of interest do not display the same discriminative property as the hand in vertical motion, as well as the face in rotation and vertical motion as described in previous work [1], [2]. However, as shown in Fig. 4 they still exhibit different patterns for the different positions.

The approach proposed consists in calculating the statistics (means and standard deviation) of the symmetry curves for the different classes. Fig. 5 shows points corresponding to the statistics of the three different classes namely left, right and centre on a feature space made of means as x-axis and standard deviation as y-axis. Given Fig. 5, known non-linear classification methods may be used. A Multi Layer Perceptron Neural network and a Support Vector Machine are used in this study.

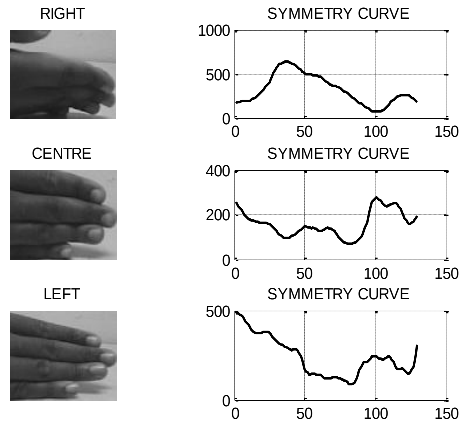

Fig. 4. Symmetry curves corresponding to three different positions of the hand in rotation

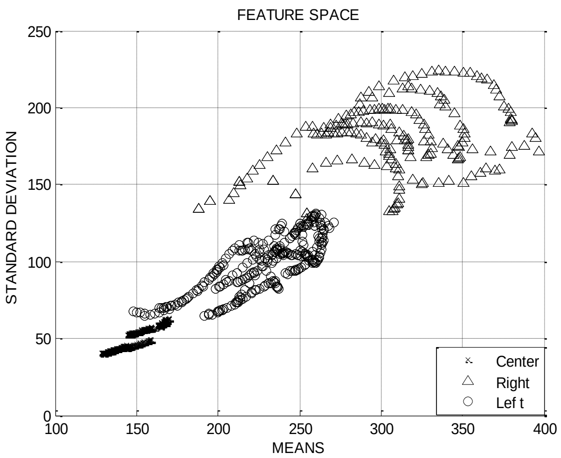

Fig. 5. Feature space showing points representing hands in the centre right or left position. Each point is given by (μ_c, σ_c), and c = {centre, right, left}

1) Neural Network: Multilayer Perceptron (MLP)

As a powerful data modelling tool, the neural network's ability to learn non-linear relationships [9] from data such as those shown in Fig. 5 is used. From empirical study [1] the topology of the multilayer perceptron (MLP) is chosen to consist of a two neuron input layer, a 10 neuron hidden layer and the output. The training is performed using a backpropagation algorithm: Given a labelled training set consisting of a set of data points $x_c = (\mu_c, \sigma_c)$ with their accompanying labels T_c, c = {centre, right, left}, the output is given by

$$Y = f(\sum_{i=1}^{N} \omega^i x_i + b), \quad (3)$$

where N is the number input neurons x_i from the previous layer ω_i is the weight associated to x_i, b is the offset from the origin of the feature space and f is the activation function chosen to be the sigmoid function $f(x) = \frac{1}{1+e^{-x}}$.

The weights are updated using

$$\omega_{jk}^{i+1} = \omega_{jk}^{i} + \Delta\omega_{jk}^{i}, \quad (4)$$

where $\Delta\omega_{jk}^{i} = -\eta \frac{\partial E^i}{\partial \omega_{jk}^{i}}$, $E^i = \frac{1}{2}\sum_{o=1}^{N_O}(T_o - Y_o)^2$ (T_o and Y_o are the target and actual output of the network respectively).

2) Support Vector machine

Support vector machines have become increasingly popular tools in data mining tasks such as regression, novelty detection and classification [10] and can therefore be used for the classification problem at hand: Given a labelled training set consisting of a set of data points $x_c = (\mu_c, \sigma_c)$ with their accompanying labels T_c and c = {centre, right, left}, the hyperplane expression is given by

$$y = \langle x, \omega \rangle + b, \quad (5)$$

where ω and b are the weights (giving the shape of the hyperplane) and offset from the origin respectively, and x is the data. The value of ω and b that maximizes the margin between the hyperplane and the support vectors is obtained using

$$\arg\min_{w,b}\{\frac{1}{2}\|\omega\|^2 - \sum_{i=1}^{N}\alpha_i[y_i(\omega . x_i - b) - 1]\}, \quad (6)$$

yielding

$$\omega = \sum_{i=1}^{N}\alpha_i y_i x_i \quad (7)$$

$$b = \frac{1}{N_{SV}}\sum_{i=1}^{N_{SV}}(\omega . x_i - y_i), \quad (8)$$

where α_i is the i^{th} Langrange multiplier and N_{SV} are the numbers of support vectors. Since this is a non-linear problem (refer to Fig. 5), the kernel trick is used to construct the hyperplane. The main idea behind the kernel trick is to map the data into a different space, and to construct a linear classifier in that space [10]. The polynomial kernel k is used, and the solution becomes:

$$F(x) = \sum_{i=1}^{N}\alpha_i y_i k(x_i, x) \quad (9)$$

For this 'three class' problem, a one against one decomposition of the binary classifiers is used. Table 7 summarizes the results for individual classification, using the MLP and SVM.

E. Rotation: Intention detection for direction

Let E = {I_i : I_i be the i^{th} frame in a sequence of N = 15 frames}: $\forall\ I_i \in E,\ M_i = \frac{1}{L}\sum_{k=1}^{L} f_i(y)$, where f is the symmetry curve associated to I_i. The resulting vector $V_2 = \{M_i : i = 1 \ldots 15\}$ is shown on Fig. 6 and 7 for each scenario. It may be observed that rotation from the centre to either side exhibits the same pattern while rotation from either side to the centre also exhibits the same pattern. It is therefore possible to distinguish between rotations from the centre and those from either side. However, insufficient information is provided in V_2 to distinguish between rotation to the left and rotation to the right. To address this problem, a preliminary step is implemented that consists in getting another 15 elements vector V_1 made of the outputs of the MLP or the SVM. For a 'centre' scenario, 15 consecutive 1s are expected, while 15 consecutive 2s and 3s are expected for right and left scenarios respectively. The Euclidean distance is therefore calculated between the given vector V_1 and the three 15-points vectors made of ones, twos and threes respectively. The smallest distance indicates which class should be chosen. Table 5 summarizes the resulting decision rule used for classification and Table 9 summarizes the results.

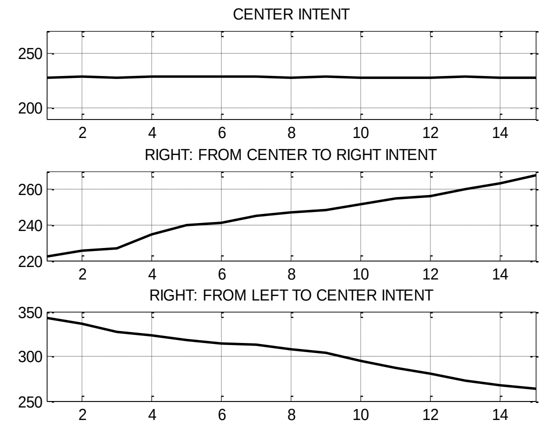

Fig. 6. Time sequence of the symmetry curves' means for hands in rotation to the right

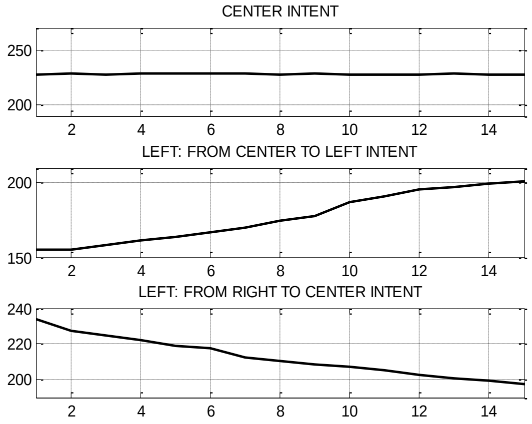

Fig. 7. Time sequence of the symmetry curves' means for hands in rotation to the left

TABLE 5: DECISION RULE FOR INTENT RECOGNITION FOR THE HAND IN ROTATION. THE DISTANCE USED IS THE EUCLIDEAN DISTANCE.

Get the output of the MLP/SVM for 15 consecutive frames resulting in the symmetry curve: vector V_1.
$d_1 = dist(V_1, ones(1,15))$
$d_2 = dist(V_1, 2\times ones(1,15))$
$d_3 = dist(V_1, 3\times ones(1,15))$
$d = min(d_1, d_2, d_3)$;
Use the decision rule on Table 4 on V_2 to classify between flat, ascending and descending V_2.
if $d == d_1$ && V_2 flat
Centred motion
Else if $d == d_2$
Motion to the right
if V_2 ascending
From centre to the right
Else if V_2 descending
From left to the centre
Else if $d == d_3$
Motion to the Left
if V_2 ascending
From centre to the left
Else if V_2 descending
From right to the centre

III. RESULTS

The experimental results have been obtained by collecting video sequences of five different subjects with three sequences each. The dorsal side of the right hand in rotation and in vertical motion are the objects of interest. Two sets of results are given below namely the classification rates of individual positions of the hand and that of the hand in motion indicating the intention.

A. Results for individual position classification

For classification of individual positions, 450 examples are used for training and the validation set is made of 900 image frames. From Tables 6 and 7 it may be observed that the up/down/centre classification rate is better than the left/right/centre classification. This is justified by the fact that the symmetry curves display more explicit changes for the vertical motion than for the rotation. However, the first requires a bigger vertical region of interest than the later as the vertical motion scans a bigger area than the rotation. For speed variation recognition (Table 6) the results are consistent in both methods. The Centre class has the best classification rate and the Down classification displays the worst rate. From Table 6 it can also be observed that the "Statistics of COG in Gaussian distribution" approach (98.3333%) performs slightly better than the "Difference of means of COG" approach (98.1852%) again because of the added information provided by the standard deviation in the first method. From Table 7 it can be observed that for both MLP and SVM, the worst classification is that of the left class. The reason is that left and centre hands are visually close to each other (Fig. 4 and Fig. 5) resulting in left hands misclassified as centred hands. This is more pronounced in the SVM giving a worse overall result than the MLP. Note also that the SVM gives 100% classification rate for the centred class as opposed to the MLP that displays 93.7778% because the optimal hyper plane places all the centre examples in the correct class at the cost of left examples being left in the centred class as errors.

B. Results for intention detection

For direction detection the decision rule described in Table 5 is used where a combination of the sequence of symmetry curves' means (intention curves V_2) and the sequence of output from the MLP or SVM classifiers (intention curves V_1) constitute the input. For speed variation the intention curve is simply made of a sequence 15 consecutive centres of gravity. The training sets are made of 400 examples of intention curves and 600 intention curves are used for validation. The size of the intention curve being 15 means that theoretically, for a 25 frames per second frame grabber, the proposed solution can perform recognition in 0.6 second. The intent recognition results are summarized in Tables 8 and 9.

For direction detection the MLP classification yields better results than the SVM and the combination with the statistics in Gaussian distribution approach gives the better classification rate. This is due to the added information of the standard deviation and the fact that given the data in Fig. 5, the MLP performs better than the SVM.

TABLE 6: RESULTS ON POSITION CLASSIFICATION FOR THE HAND IN VERTICAL MOTION (SPEED VARIATION) ON INDIVIDUAL FRAMES

Methods	**Class**	**Training set**	**Testing set**	**Correct classification**	**Incorrect classification**	**Classification rate**
Difference of means of COG	Centre:	450	900	900	0	100%
	Up:	450	900	885	15	98.3333%
	Down:	450	900	866	34	96.2222%
Total:		1350	2700	2651	49	98.1852%
Statistics of COG in Gaussian distribution	Centre:	450	900	900	0	100%
	Up:	450	900	893	7	99.2222%
	Down:	450	900	862	38	95.7778%
Total:		1350	2700	2655	45	98.3333%

TABLE 7: RESULTS ON POSITION CLASSIFICATION FOR THE HAND IN ROTATION (FOR DIRECTION) ON INDIVIDUAL FRAMES

Methods	**Class**	**Training set**	**Testing set**	**Correct classification**	**Incorrect classification**	**Classification rate**
MLP	Centre:	450	900	844	56	93.7778%
	Right:	450	900	855	45	95%
	Left:	450	900	823	77	91.4444%
Total:		1350	2700	2522	178	93.4047%
SVM	Centre:	450	900	900	0	100%
	Right:	450	900	818	82	90.8889%
	Left:	450	900	757	143	84.1111%
Total:		1350	2700	2475	225	91.6667%

TABLE 8: RESULTS ON SPEED VARIATION RECOGNITION

Methods	**Class**	**Training set**	**Testing set**	**Correct classification**	**Incorrect classification**	**Classification rate**
Difference of means of COG	Centre:	400	600	531	69	88.5%
	Up:	400	600	564	36	94%
	Up (back):	400	600	489	111	81.5%
	Down:	400	600	507	93	84.5%
	Down (back):	400	600	565	35	94.1667%
Total:		2000	3000	2656	344	88.5333%
Statistics of COG in Gaussian	Centre:	400	600	522	78	87%
	Up:	400	600	580	20	96.6667%
	Up (back):	400	600	489	111	81.5%
	Down:	400	600	517	93	86.1667%
	Down (back):	400	600	581	19	96.8333%
Total:		2000	3000	2689	321	89.6333%

TABLE 9: RESULTS ON DIRECTION RECOGNITION

Methods	**Class**	**Training set**	**Testing set**	**Correct classification**	**Incorrect classification**	**Classification rate**
MLP + difference of Means	Centre:	400	600	573	27	95.5%
	Right:	400	600	560	40	93.3333%
	Right (back):	400	600	568	32	94.6667%
	Left:	400	600	554	46	92.3333%
	Left (back):	400	600	534	66	89%
Total:		2000	3000	2789	211	92.9667%
MLP + Statistics	Centre:	400	600	573	27	95.5%

in Gaussian	Right:	400	600	588	12	98%
distribution	Right (back):	400	600	564	36	94%
	Left:	400	600	550	50	91.6667%
	Left (back):	400	600	570	30	95%
Total:		2000	3000	2845	155	94.8333%
SVM +	Centre:	400	600	548	52	91.3333%
difference of	Right:	400	600	552	48	92%
Means	Right (back):	400	600	522	78	87%
	Left:	400	600	569	31	94.8333%
	Left (back):	400	600	530	70	88.3333%
Total:		2000	3000	2721	279	90.7%
SVM + Statistics	Centre:	400	600	548	52	91.3333%
in Gaussian	Right:	400	600	586	14	97.6667%
distribution	Right (back):	400	600	526	74	87.6667%
	Left:	400	600	563	37	93.8333%
	Left (back):	400	600	564	36	94%
Total:		2000	3000	2787	213	92.9%

IV. CONCLUSION

This paper proposes a visual interface to infer a subject's intentions. The preliminary results were obtained using static images and show promise as an alternative human-machine interaction solution in the context of wheelchair mobility. Further work is ongoing to test the robustness of the method in different lighting conditions and with more realistic hands of people with disabilities. A real time implementation is also part of ongoing work.

REFERENCES

[1] Luhandjula T, Hamam Y, van Wyk B.J et al (2009) Symmetry-based head pose estimation for intention detection. In: Proc Twentieth Annual Symposium of the Pattern Recognition Association of South Africa, Stellenbosch, South Africa, pp 93-98

[2] Luhandjula T, Monacelli E, Hamam Y, et al (2009). Visual Intention Detection for Wheelchair Motion. In: Proc Springer-Verlag (ed) 5th International symposium on visual computing 2009, Las Vegas, USA, pp 407-416

[3] Kanno T, Nakata K, Furuta K, (2003) Method for team intention inference. Human-Computer Studies 58:393-413

[4] Geib C.W, (2002) Problems with intent Recognition for Elder Care. In: Proc Association for the Advancement of Artificial Intelligence Workshop on automation as Caregiver. Menlo Park, CA, USA pp 13-17

[5] Lesh N, Rich C, Sidner C.L, (1999) Using Plan Recognition in Human-Computer Collaboration. In: Proc Seventh International Conference on User Modeling. Banff, Canada

[6] Christensen H.V, Garcia J.C, (2003) Infrared Non-Contact Head Sensor, for Control of Wheelchair Movements. Assistive Technology: From Virtuality to Reality. A. Pruski and H. Knops (ed) IOS Press pp 336-340

[7] Jia P, Hu H, (2005) Head Gesture based Control of an Intelligent Wheelchair. In: Proc Annual Conference of Chinese Automation and Computing Society. UK, Sheffield

[8] Demeester E, Nuttin M, Vanhooydonck D, et al (2003) Assessing the User's Intent Using Bayes' Rule: Application to Wheelchair Control. In: Proc of the first International Workshop on Advanced in Service Robotics. Bardolino, Italy, pp 117-124

[9] Bishop C, (1995) Neural Networks for Pattern Recognition. Oxford University Press

[10] Cristianini N, Shawe-Taylor J, (2000) An introduction to support Vector Machines: and other kernel-based learning methods. Cambridge University Press, New York