

# Symmetry-based head pose estimation for intention detection

*T. Luhandjula<sup>1,2</sup>, Y. Hamam<sup>1</sup>, B.J. van Wyk<sup>1</sup>, Q. Williams<sup>2</sup>*

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 a visual interface that infers the intention of a person to travel in a desired direction from head movements. A rotation of the head indicates the intent to change direction. The context for which this solution is intended is that of wheelchair bound individuals. This paper describes work in progress that provides a proof of concept tested on static images where head pose changes can be detected in a sequence of frames and where the pose of an individual is remarkably detected for 30°, 60° and 90° angles. Results show that the symmetry property of the head can be used to detect a change in its position and can therefore serve as a visual intent indicator. The solution described in this paper, focusing on the specific task of head pose estimation, intends to provide a contribution to the task of intention detection applied to assistive living for the support of the elderly and people with disabilities.

## 1 Introduction

One of the challenges facing the task of realizing an enabled environment where people with disabilities and the aged are independent and can therefore be active by contributing in society is to develop solutions that can assist them in performing the tasks they wish to carry out without other people's assistance. Good performance in a team environment is heavily conditioned by the awareness of people's intention within the team [2] and therefore a human-machine team where the machine has a support role, requires that the intention of the user is well understood by the machine. The need for this intention awareness capability requires that much attention is given to this problem.

In addition to contributing to the task of robust face recognition for multiview analysis which is still a difficult task under pose variation [3], head pose estimation can be considered as a sub problem of the general area of intention detection because it is useful for the inference of nonverbal signals related to attention and intention. Estimating the head pose is crucial since it usually coincides with the gaze direction. Furthermore, head pose estimation is also essential for analyzing complex meaningful gestures [4].

Existing head pose estimation methods can be categorized as appearance-based and model-based methods. Appearance-based techniques use the whole sub image containing the face while model-based approaches use a geometric model [4]. Another set of approaches includes the application of eigenspace techniques to directly recognize the pose of a specific user [5]. According to [6] the principal advantage of global (appearance-based) approaches is that only the face needs to be located and that no facial landmarks or face models are required, making them appropriate for very low resolution images of the face. This is useful for video

surveillance, intelligent environments and human interaction modelling. Template matching is a popular method to estimate head pose where the best template can be found via a nearest-neighbour algorithm and where the pose associated with this template is selected as the best pose. Advanced template matching can be performed using Gabor Wavelets and Principle Components Analysis (PCA) or Support Vector Machines, but these approaches tend to be sensitive to alignment and are dependent on the identity of the person [6].

This paper provides an alternative head pose estimation solution for intention detection. Work in progress is described, that provides a proof of concept tested on static images where head pose changes can be detected over time and where the pose of an individual can be remarkably detected for angles 30°, 60° and 90°. The context for which this solution is intended is that of wheelchair bound individuals whose intention of interest is the direction they wish the wheelchair to follow. The motivation behind this choice is the availability and flexibility of the head for a wide range of disabilities. The proposed head pose estimation scheme can be classified as belonging to the appearance-based category [4] making use of the symmetry property of the head as a whole and therefore has the advantages of accommodating low resolution images and not being dependent on the identity of the person. The aim of this study is to make a contribution to the general problem of intention detection applied to assistive living for the support of the elderly and people with disabilities.

## 2 Methods

The type of data used is visual: A sequence of images is captured from a CCD camera and a '25 frames per second' frame grabber. No visual markers are added to the images. This work assumes that the head is already detected and does not deal with the pre-processing step of detecting and tracking the head.

The approach in this paper extracts high level information (referred to in this work as the intention curve) from the acquired images, resulting in direction intent recognition. Specific angles of the head pose are also estimated for one individual.

A symmetry-based approach [1] is used to extract symmetry curves associated with frontal views of the head. Note that in this paper, the use of the symmetry-based approach differs from that of [1] in that the latter uses the method for face detection while the same method is used in this work for head pose estimation. The assumption is that different positions of the head give different symmetry curves. The classification of the position of the head uses the pattern displayed by the centres of gravity (COG) of the symmetry curves and is implemented using two different methods namely the 'difference of means' approach and the

‘mean and standard deviation in a Gaussian distribution’ approach. Intention detection in the context of this work entails the classification of these COG sequences representing intention curves, using the two classification methods mentioned above.

## 2.1 Symmetry-based approach

The underlying assumption is that human heads viewed from the front are symmetric; and when moved from their initial position (centred position) the symmetry they display breaks down, giving the indication of a motion from the initial centred position to a new position (Left or Right). The indication of a new direction to the right or to the left is given by the head moving to the right or the left respectively. The symmetry curve, based on the work of [1], is given by

$$C(x) = \sum_{\omega=1}^k \sum_{y=1}^Y |I(x-\omega, y) - I(x+\omega, y)|, \quad (1)$$

$$\forall x \in [k+1 \ X-k]$$

The symmetry-value  $C(x)$  of each pixel-column in the image is evaluated by taking the sum of the differences of two pixels at a variable distance  $\omega = \{1, \dots, k\}$  from it on both sides making the pixel-column the centre of symmetry. This process is done for each row and the resulting symmetry-value is the summation of these differences. The symmetry curve is made of all the symmetry values calculated for all the columns in interval  $[k+1 \ X-k]$ .  $X$  and  $Y$  designate the width and the length of the image respectively,  $x$  and  $y$  represent a given column and row of the image respectively, and the value of the constant  $k$  as been set to 7 as a result of an empirical study. As shown in Figure 1, a centred position of the head yields a symmetry curve made of two peaks and one valley showing high level of symmetry while a right or left position of the head flatten those peaks showing a lower level of symmetry and indicate the direction of the head by the way in which the values of the curves continuously increase. Figures 1 and 2 show the symmetry curves for different (at angles 30°, 60° and 90°) positions of the head when rotated to the left and right respectively.

## 2.2 Classification for individual image frames

### 2.2.1 Centre of gravity of the Symmetry curve

The centre of gravity is calculated as a point in the symmetry curve at which all the values of the curve can be considered centred. The symmetry curves shown in Figures 1 and 2 have different looks and therefore give different centres of gravity. This measure can be used to classify the curves corresponding to different positions of the head. The centre of gravity of the symmetry curve is calculated as follow:

$$COG = \frac{x_1 C(x_1) + x_2 C(x_2) + \dots + x_n C(x_n)}{C(x_1) + C(x_2) + \dots + C(x_n)} \quad (2)$$

Where the symmetry curve is defined by the function:  $C: x \rightarrow C(x)$  where  $C(x)$  is given by the expression in Equation 1 and  $I$  is a 120x140 greyscale image frame. The position of the symmetry curve’s centre of gravity gives an indication on the position of the head. Figures 1 and 2 show the position of the centre of gravity on the symmetry curve for different positions of the head when rotated to the left and to the right respectively, at angles 30°, 60°, and 90°.

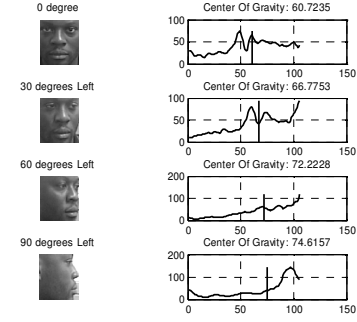


Figure 1: Symmetry curves corresponding to four different positions of the head in rotation to the left and position of the centre of gravity indicated by the vertical line on the symmetry curves.

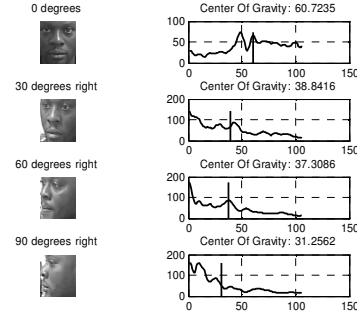


Figure 2: Symmetry curves corresponding to four different positions of the head in rotation to the right and position of the centre of gravity indicated by the vertical line on the symmetry curves.

Two approaches are used to classify these different positions into three classes (centre, right and left) and into the positions associated to the specific angles 30°, 60° and 90°:

### 2.2.2 Difference of means

The mean of the centres of gravity is calculated for the training set of each category. The difference between the centre of gravity of a curve 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 1: Difference of means approach

Let  $\mu_{centre}$ ,  $\mu_{right}$ ,  $\mu_{left}$  be the means of three classes.

Let  $COG$  be the centre of gravity of the symmetry curve and therefore a given image frame to be classified:

$$d_1 = |COG - \mu_{centre}|, d_2 = |COG - \mu_{right}|, d_3 = |COG - \mu_{left}|$$

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

If  $d == d_1$  class 1

else if  $d == d_2$  class 2

else class 3

### 2.2.3 Statistics in a Gaussian distribution

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

Table 2: Statistics (mean and standard deviation) in a Gaussian distribution approach

Calculate

$$P_1 = \mathcal{N}(\mu_{Center}, \sigma_{Center}), P_2 = \mathcal{N}(\mu_{Right}, \sigma_{Right}),$$

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

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

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

Class = {Centre, Right, Left},  $\mu_{Class}$  and  $\sigma_{Class}$  are the statistics (means and standard deviations) of the centres of gravity in a training set for each class and COG is the new centre of gravity of the curve to be classified.

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

if  $P = P_1$  class 1

elseif  $P = P_2$  class 2

else class 3

### 2.2.4 Classification of different angles

The methods described above are also used for classification of head pose at the specific angles  $30^\circ$ ,  $60^\circ$  and  $90^\circ$ . It has been tested on a single individual. A two layer classifier implemented by the above mentioned methods is used first to classify between Left, right and centre positions of the head and the second to classify by the angles as illustrated in Table 3:

Table 3: classification of individual head positions for angles  $30^\circ$ ,  $60^\circ$  and  $90^\circ$

---

Input: Acquire a new image frame  
 Layer 1: Classify between Left Right and Centre  
   If centre  
     Go to Input  
   Else if Left  
 Layer 21: Classify between  $30^\circ$  left,  $60^\circ$  left and  $90^\circ$  left  
   Go to Input  
   Else if Right  
 Layer 22: Classify between  $30^\circ$  right,  $60^\circ$  right and  $90^\circ$  right  
   Go to Input

---

### 2.3 Intention detection

The task of intention detection in the context of this work involves the recognition of the direction the subject intends to take, indicated by the motion of the head. The problem of monitoring the time sequence of individual head poses is addressed using the sequence of the symmetry curves'

centres of gravity for successive image frames. The assumption underlying the use of these indications is that the value variations in the symmetry curves' centres of gravity are different in the three following scenarios: Head in rotation to the left, head in rotation to the right, head centred.

#### 2.3.1 Sequence of centre of gravity

Let E be a set equal to  $\{I_i : I_i \text{ is the } i^{\text{th}} \text{ frame in a sequence of } N \text{ frames}\}$  where  $N = 15$ . For each image frame in E the symmetry curve and the centre of gravity of the symmetry curve is obtained using equations 1 and 2 respectively. Figure 3 shows the plots (referred to in this work as intention curves) of the given centres of gravity in a time sequence for the three different types of motion. It shows that these three types of motion (head rotation to the left, head rotation to the right and centred head) exhibit each a different pattern.

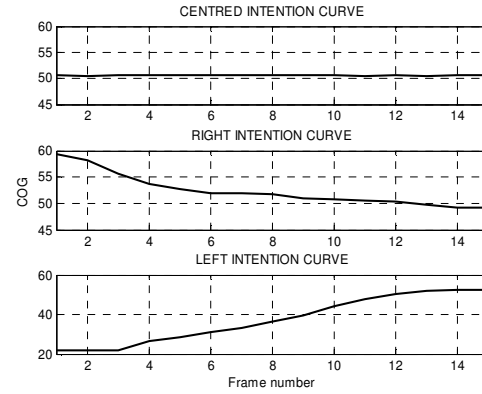


Figure 3: Time sequence of the symmetry curves' centres of gravity for heads in rotation.

#### 2.3.2 Head motion classification

Given the level of clarity on the higher level information given by the intention curve extracted in Figure 3, a simple decision rule is implemented for classification. Table 4 describes a pre-processing step resulting in an indication of the increase or decrease of the intention curve (and therefore of the COGs values) through time. This information is then used in the decision rule described in Table 5.

Table 4: A and B are the counters keeping track of the increase or decrease in values of the intention curves through time

---

Let  $Cu$  be the given intention curve,  $A = 0$ ;  $B = 0$ ;  
 (Initialize counter A notifying a decrease and counter B notifying an increase)  
 $\forall i \in \{x : x \geq 1 \text{ and } x \leq \text{length}(Cu) - 1\}$ ,  
 $D = Cu(i) - Cu(i+1)$   
 If  $D > 0$      $A = A + |Cu(i) - Cu(i+1)|$  (notifying a decrease in value of C, by adding the extend to which there is a decrease to the value of A)  
 If  $D < 0$      $B = B + |Cu(i) - Cu(i+1)|$  (notifying a decrease in value of Cu, by adding the extend to which there is a decrease to the value of B)  
 Else            Do nothing

---

From Figures 3 and given the algorithm in Table 4, it can be visually established that in case of a centred face, A and B should be approximately the same in case of a “going left” scenario, B should be much higher than A, and in the case of a “going right” scenario, A should be much higher than B. The following decision rule can be used for classification:

Table 5: Decision rule for classification of the high level information for intention detection: A and B are obtained using the algorithm in Table 4

---

Calculate:

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

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

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

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

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

$\mu_{class}, \sigma_{class}$  are the mean and standard deviation of the difference between A and B in a training set for each class and  $I = \{1, 2, 3\}$ , and class = {Centre, Right, Left},

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

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

If  $A > B \text{ and } P = P_2$  *Intention = Going Right*

If  $A < B \text{ and } P = P_3$  *Intention = Going Left*

Otherwise *Inconclusive*

---

### 2.3.3 Summary

In summary, images with a frontal view of the head as objects of interest are used as input to the proposed solution. A symmetry-based approach is used to evaluate the symmetry of the face measured by a symmetry curve, and for individual head pose classification, the curve’s centre of gravity is extracted for further pattern recognition tasks namely the difference of means of the Symmetry curve’s COG, the mean and standard deviation of the Symmetry curve’s COG in a Gaussian distribution.

Table 6: Results on position classification on individual frames

Classification Method	Class	Training set	Testing set	Correct classification	Classification rate
<b>Difference of means using Centre of gravity of Symmetry curve</b>	Centre:	100	325	320	98.4615%
	Right:	100	325	286	88%
	Left:	100	325	325	100%
					Total classification Rate: 95.4872%
<b>Means and standard deviations of Centre of gravity of Symmetry curve in Gaussian distribution</b>	Centre:	100	325	320	98.4615%
	Right:	100	325	290	90.7692%
	Left:	100	325	290	90.7692%
					Total classification Rate: 93.3333%

The intention detection task uses the same methods on intention curves made of sequences of centres of gravity from 15 consecutive frames. Three sets of results have been given below namely, the classification rate of individual positions of the heads, the classification rate of individual positions of the head at angles

## 3 Results

Three sets of results have been given below namely, the classification rate of individual positions of the heads, the classification rate of individual positions of the head at angles 30°, 60°, 90° and the classification rate of a particular head motion giving us an indication of the intention. The experimental results for the first and third scenarios have been obtained by collecting video sequences of five different subjects with three sequences each with the frontal face as the object of interest. The results for the classification at specific angles have been obtained with a single subject. For classification of individual positions, 100 frames are used for training and 325 testing frames are used to acquire the performance rate. For classification between the different angles, 100 examples are used for training and 100 for validation for each class. For classification of head motions, the training sets are made of 150 examples of intention curves each obtained using 15 frames and 350 intention curves are used for performance testing. This means that theoretically, for a 25 frames per second frame grabber, the proposed solution can perform recognition in 0.6 second.

An observation worth mentioning is the performance disproportion exhibited in the difference of means approach (84% and 97.1429% in Table 8) between ‘right’ and ‘left’ for intention curve classification. The disproportion is inexistent when using the mean and standard deviation of the Gaussian distribution approach (98.2857% and 97.1429%). This can be explained by the added information provided in the second method by the standard deviation of the Centres of Gravity. In some cases this added information is redundant (refer to Tables 6 and 7) when the standard deviations associated to the classes are close to each other. But in this particular case it adds relevant information for classification because the difference between the standard deviations associated with each class is significant.

Table 7: Results on position classification between centre, left and right position for angles 30°, 60° and 90°

	Classification Method	Class	Training set	Testing set	Correct classification	Classification rate	
Layer 1: Classification between Left and Right and Centre	<b>Difference of means using Centre of gravity of Symmetry curve</b>	Centre:	200	200	200	100%	
		Left:	200	200	186	93%	
		Right:	200	200	200	100%	
							Total classification Rate: 97.6667%
	<b>Means and standard deviations of Centre of gravity of Symmetry curve in Gaussian distribution</b>	Centre:	200	200	198	99%	
		Left:	200	200	200	100%	
		Right:	200	200	200	100%	
							Total classification Rate: 99.6667%
	Layer 21: Classification between left positions of the head at angles 30°, 60° and 90°	<b>Difference of means using Centre of gravity of Symmetry curve</b>	30° :	100	100	100	100%
			60° :	100	100	100	100%
90° :			100	100	100	100%	
						Total classification Rate: 100%	
<b>Means and standard deviations of Centre of gravity of Symmetry curve in Gaussian distribution</b>		30° :	100	100	100	100%	
		60° :	100	100	100	100%	
		90° :	100	100	100	100%	
						Total classification Rate: 100%	
Layer 22: Classification between right positions of the head at angles 30°, 60° and 90°		<b>Difference of means using Centre of gravity of Symmetry curve</b>	30° :	100	100	90	90%
			60° :	100	100	89	89%
	90° :		100	100	100	100%	
							Total classification Rate: 93%
	<b>Means and standard deviations of Centre of gravity of Symmetry curve in Gaussian distribution</b>	30° :	100	100	88	88%	
		60° :	100	100	91	91%	
		90° :	100	100	100	100%	
							Total classification Rate: 93%

Table 8: Head motion recognition with a motion made of 15 consecutive image frames

Classification method	Class	Training set	Testing set	Correct classification	Classification rate
<b>Difference of means using Centre of gravity of Intention curve</b>	Centred Head motion:	150	350	350	100%
	Head motion to the Right:	150	350	294	84%
	Head motion to the Left:	150	350	340	97.1429%
<b>Means and standard deviations of Centre of gravity of Intention curve in Gaussian distribution</b>	Centred Head motion:	150	350	345	98.5714%
	Head motion to the Right:	150	350	344	98.2857%
	Head motion to the Left:	150	350	340	97.1429%

## 4 Conclusion

This paper describes a visual interface that infers the intention of a person to travel in a desired direction from head movements. A symmetry-based head-pose estimation solution has been developed for that purpose. The context for which this solution is intended is that of wheelchair bound individuals whose intention of interest is the direction of the motion they wish the wheelchair to follow. The instruments used for data collection are a CCD camera, and a 25 frames/s frame grabber.

The main contribution this work consists in the use of a symmetry-based approach that is combined with different classification schemes for head pose estimation, indicative of the intent of a person in terms of direction.

This work has shown that the symmetry property of a person's face can be used to detect any change in its position and can therefore be a visual motion indicator. This work has established a proof of concept that uses this symmetry property yielding very good results. Given the quality of these results, this solution shows promise in making a contribution to the general problem of intention detection applied to assistive leaving (for support of the elderly and people with disabilities).

The results for the classification of specific angles are obtained for a single individual making the classification very optimistic. This solution is however intended to be independent of the identity of the person, and given the above description of symmetry measure (refer to Eq. 1), not every face displays the same level of symmetry. This will result in more realistic results.

The next step is to implement the method proposed in this paper in real time and where angles can be continuously detected along with the direction of the motion, and where the safety of this solution will also be assessed. Note that this solution is only interested in the direction the person intend the wheelchair to follow. Further work is ongoing to address the task of speed variation as well as stopping. The images were taken within an office environment without accounting for any change in illumination. More tests are ongoing to obtain results in more controlled lighting conditions.

## 5 References

1. Luhandjula, K.T., van Wyk, B.J., Kith, K., van Wyk, M.A., 2006. Eye detection for fatigue assessment. *In: Proceeding of the Seventeenth International Symposium of the Pattern Recognition Society of South Africa*. Parys, South Africa.
2. Kanno, T., Nakata, K., Furuta, K. 2003. Method for team intention inference. *Human-Computer Studies*, 58:393-413.
3. Ma, B., Zhang, W., Shan, S., Chen, X., Gao, W. 2006. Robust Head Pose Estimation using LGBP. *In: Proceedings of the 18<sup>th</sup> International Conference on Pattern Recognition*. 2:512-515.
4. Vatahska, T., Bennewitz, M., Behnke, S. 2007. Feature-based Head Pose Estimation from Images. *In: Proceedings of the IEEE-RAS 7th International Conference on Humanoid Robots (Humanoids)*. Pittsburgh, USA.
5. Fitzpatrick, P. 2001. Head pose estimation without manual initialization. *Term Paper for MIT Course*. MIT, Cambridge, USA.
6. Gourier, N., Maisonnasse, J., Hall, D., Crowley, J.L. 2006. Head Pose Estimation on Low Resolution Images. *Perception, recognition and integration for interactive environments*.
7. Tu, J., Fu, Y., Hu, Y., Huang, T. 2007. Evaluation of Head Pose Estimation for Studio Data. *Multimodal Technologies for Perception of Humans*. Heidelberg: Springer Berlin.
8. Christensen, H.V., Garcia, J.C. 2003. Book Infrared Non-Contact Head Sensor, for Control of Wheelchair Movements. *Assistive Technology: From Virtuality to Reality*. A. Pruski and H. Knops (Eds) IOS Press. 336-340.
9. Kuno, Y., Shimada, N., Shirai, Y. 2003. Look where you are going [robotic wheelchair]. *Robotics & Automation Magazine, IEEE*. 10(1): 26-34.
10. Matsumoto, Y., Ino, T., Ogasawara, T. 2001. Development of intelligent wheelchair system with face and gaze based interface. *In: Proceedings of the 10th IEEE International Workshop on Robot and Human Interactive Communication*. Paris, France, 262-267.