Personalized Gesture Recognition with Hidden Markove Model and Dynamic Time Warping

Shubhangi S. Marudkar[1] Dr. Prof. Prashant R. Deshmukh[2]
[1]M.E. Student [2]Assistant Professor
[1,2]Department of Computer Science & Engineering
[1,2]Sipna College of Engineering and Technology Amravati, India

Abstract— Spinal cord injury, brain injury and strokes can cause upper extremity mobility impairment. Because of these the person suffering from becomes paralyzed and unable to do their daily routines. There are some existing assisting technology solutions to provide access as user input but mostly they are intrusive and expensive and some requires physical contact which can have dangerous effects such as skin friction injury for the paralyzed selection users. Our system allows user for selection of gestures, we are using Hidden Markove model and Dynamic Time Warping to convert raw capacitance values to alphanumeric gestures.

Key words: Upper Extremity Impairments, Gesture Recognition

I. INTRODUCTION

Given a source of data it has found that there are millions of people hospitalized every year because of strokes, brain injuries and spinal cord injuries. These patients become paralyzed and weak even after completion of their treatment and they become dependent on the assistive care devices for the environmental control physical therapy, such assistive devices can help them for maximum independence and prove supplement for direct care. The assistive devices are generally based on the gesture recognition- based environmental control systems, there are many techniques are available such as use of inertial sensors, vision systems, eye tracking systems, voice recognition systems etc. These systems present a set of fundamental challenges such as the recognition system fail to address[3]. Many individuals performs gestures in different ways, Infact the same user may perform a gesture differently depends on the time of the day, medication and fatigue. Some eye tracking systems uses mounted cameras while EEG electrode-based systems which can cause skin irritation and abrasion that can have hazardous effects. In order to overcome there challenges, we designed and implemented the wearable sensors system built from the textile capacitive sensor array using the Hidden Markove Model and Dynamic Time Warping[2].

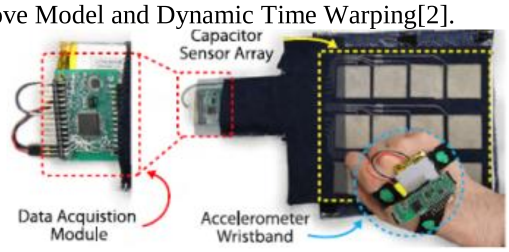


Fig. 1: Demonstrates a 4 _ 3 capacitive sensor array sewn into the denim fabric using conductive wires. It also demonstrates the accelerometer wristband used to find the orientation of the capacitor sensor array with respect to the hand. The data from the Sensors is analyzed using our custom-designed wireless module which uses capacitance measurement ICs, an MSP430 micro-controller, and Bluetooth wireless module.

Fig. 1 illustrates a prototype system built using the capacitive plates and conductive threads sewn into the denim fabric and the wrist band with an accelerometer[2]. These sensors are proximity sensors and obviate the need for touch-based gesture recognition. An array of these sensors can be effectively used to track body positions and movements that can be consequently used to recognize gestures.

II. WORKING PRINCIPLE

In gesture recognition system the spotting tasks is more challenging. The recognition of gestures has been studied for many years. The work of gesture recognition presented in this paper is part of large effort towards reliable spotting of complex activities using simple on body sensors. The primary aim of this is to support complex activities of human gestures. The contribution of this paper is to present gesture spotting method based on body to worn motion sensors[4]. We are using HMM to reduce the classification errors. This paper also describes the verification on two scenarios that comprise thousands of relevant gestures. First is to interact with everyday different objects which are part of wide range of wearable system applications. The second one is nutrition intake is highly specialized applications motivated by the needs a large industry dominated health monitoring project when there is large variations in the way of gestures the HMM performs well, but the DTW requires minimal training[5].

A. *Hidden Markove Model*

The Hidden Markove Model is defined as set of states and an alphabet of output symbols. The each state of the HMM is characterized by the two probability distributions:

- Transition distribution over state and
- Emission distribution over the output symbols[3].

The algorithm is illustrated in Algorithm 1 Formally, it is defined by 5-tuple $\Omega=(S,\Sigma,\Pi,\delta,\lambda)$ where S represents a set of hidden states that are not directly observable. Transitions between the states are denoted by a transition probability matrix, δ. Π is the set of initial probabilities corresponding to the states in S. Every state has a set of possible emissions Σ and continuous probabilities λ for these emissions. The emissions can be observed by giving some information about the most likely underlying hidden state sequence which led to a particular sequence of observations. The HMM model assumes that the underlying physical process is Markovian that means any prediction of future behavior can be optimally calculated by knowing the present state without any history. We map such a Hidden Markov Model to our problem of decoding a gesture based on a time series of hand position values. This definition of states constraints the set of gestures for which the Markovian assumption is valid. Our evaluation, however, demonstrates that this state definition works well across a large set of gestures [2].

B. *Algorithm 1: HMM Feed Forward (O, Φ, Σ, π, δ, λ, k)*

1) *Input:*

Observations (O):{$(x_1,y_1),\ldots,(x_n, \mathbf{y}_n)$} for the gesture, Hidden states (S): {1,….k}, Initial Probabilities: {$\pi_1,\ldots, \pi_k$}, Transition Probabilities: $\delta(i, j),(1 \le i , j \le k)$, Emission probability distribution: $\lambda_1,\ldots\lambda_k$ (Gaussian (μ_k,σ)) for every state k ϵ [1,…len(S)]

2) *Output:*

p (Model Probability given the Observations)
for i: =1 to len(O) do
for k: =1 to len(S) do
Emission = $(1/(\text{distance}((x_i,y_i), \mu_k)/\sigma)^2)+3$
for j: =1 to len(S) do
p(i ,k) = p(i ,k)+p(i – 1, j) x Emission
end for
end for
end for
return Σ_k p(|O| - 1, k)
Asymptotic Running time complexity: O(|O| x k^2)

The transition probability matrix, δ is determined during a training phase when the subject performs a set of gestures. While calculating the probability of transitions between states, we make the following key adaptation to fit our problem domain. We have found that if the sampling frequency of our sensors is large compared to the transition rate, the transition probability matrix begins to resemble an identity matrix, and the importance of transitions between states is diminished.

The emission probabilities λ is calculated accordingly Gaussian function. Each state is assigned a center position. The distance (D) from the estimated position and the center is used as the input for the Gaussian. The Gaussian width (σω) is approximately 1/3 of the center-to-center plate distance. This width is determined and used throughout the experiments. For computational efficiency, we use the following approximation ((D/ σω) 2+ 3)-1 for the Gaussian. Applying the above modifications to the HMM we generate a model per gesture. the gesture to be classified is passed through each of these models to create a vector of posterior probabilities relating to each Underlying model, and the model with the highest probability is selected as the classification for the gesture.

C. *Dynamic Time Warping*

The Dynamic time Warping is the generation of classical algorithms for comparing the desecrate sequences. This model is trained on a large set of gestures[6]. This model is alternative to the statistical model such as HMM. The DTW performs entire gesture with a single speed factor, therefore avoids requiring a reference gesture for every possible timing variations. Dynamic temporal distortion contrasts an approach using a uniform time scaling that assumes the entire gesture is performed with a single speed factor[2]. This avoids requiring a reference gesture for every possible timing variation. Variations in the positional path are handled by providing a representative set of gesture variations in the codebook. This set is sufficient enough to classify gestures using unsupervised clustering algorithms such as a single nearest neighbor or a k-nearest neighbor. An optimal choice of dynamic time warping and minimum error is based on a Euclidean distance metric.

D. *Algorithm 2: Dynamic Time Warping(O,M)*

1) *Input:*

O = [$(x_1 ,y_1),\ldots,(x_n , y_n)$] (positions for the gesture)
M = [$(x'_1, y'_1),\ldots, (x'_n , y'_n)$] (model positions for the gesture)

2) *Output:*

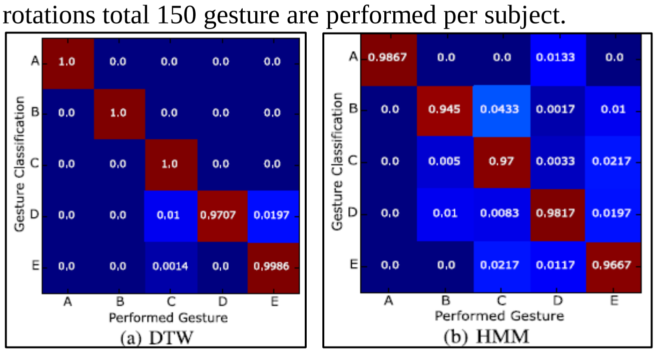
d (Warped Distance),
d(0,0) = distance(M(1), *O*(1))
for i:=1 to len(*O*) do
d(i ,1) = d(i -1, 1)+ distance(*M*(1), *O*(i))
end for
for j: =1 to len(*M*)do
d(1 ,j) = d(1,j -1)+ distance(*M*(j), *O*(1))
end for
for j: =1 to len(*M*)do
for i:=1 to len(G)do
d(1,j) = $\min_{i,j}$ d[d(i-1, j-1),d(i-1, j),d(i, j-1)] + distance(M(j), O(i))
end for
end for
return d(n,m)
Complexity: *O*(n x m)

III. SYSTEM EVOLUTION

The goal of this system is to provide highly accurate gesture recognition so that users can adapts to changes in the position and orientation of the sensors[2]. The sensor array move and rotate according to the user hand. The HMM and DTW model would ideally need to be retrained on the gestures for every rotation, to overcome this problem a system is designed that augments the capacitive sensor array with wrist worn accelerometer band, this band could be replaced by smart watch.

A. *Case Study*

In order to discuss the implementation of this approach, consider the everyday life gestures in continuous data stream from body worn sensors. We tested this system on five subjects, who has C6 spinal cord injury. For each session subject were asked to perform the five gestures, during the experiment ten of each type of gesture at different rotations total 150 gesture are performed per subject.



(a) DTW (b) HMM

Fig. 2(a): Confusion matrix illustrating the accuracy of recognizing the five gestures performed by all the subjects when our system used Dynamic Time Warping. (b) Confusion matrix illustrating the accuracy of recognizing the gestures performed by the subjects when our system used Hidden Markov Model. For both figures, the confusion matrix is calculated using 150 _ 5 = 750 gestures. Our system has an average accuracy of 99 and 97 percent for the DTW and HMM models respectively.

The figure 2 illustrates the classification of accuracy of gesture recognition algorithm. The figure generated using data from all gestures performed by five subjects when sensor array rotated different degrees with respect to hand. The average accuracy is 99 and 97 percent of DTW and HMM respectively. If there is a large variation in the way of gestures performed then HMM model would perform best while the DTW would perform poorly.

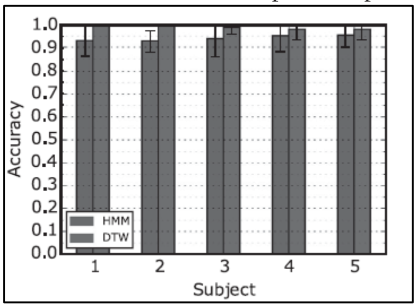


Fig. 3: Accuracy per subject for five gestures when the classification algorithm uses a DTW model and an HMM classification.

The figure 3 illustrates the accuracy of the DTW and HMM based gesture recognition per subject. The height bar shows the accuracy across five gestures while the error bar shows the standard deviation in the accuracy across gestures per subject. The focus of this experiment on the subjects who has C6 spinal cord injury. The subject perform some set of gestures and it shows that the accuracy of DTW algorithm on the gestures is 99 percent , Hence we can say that this system works very well even though the subject use limited mobility.

IV. Conclusion

In this paper we present a system which recognizes the gestures of the users with limited mobility. We have designed an adaptive personalized signal processing system which can convert capacitance data from sensor array to alphanumeric gestures. This gestures then can be used to control appliances in the home. we have tested our system using DTW and HMM based classification on five C6 spinal cord patients, we have found that the DTW is performed wee and it is more convenient option.

References

[1] L. Zhao and E. M. Yeatman, “Micro capacitive tilt sensor for human body movement detection,” in Proc. 4th Int. Workshop Wearable Implantable Body Sensor Netw., pp. 195–200.

[2] Alexander Nelson, Student Member, IEEE, Gurashish Singh, Member, IEEE,Ryan Robucci, Member, IEEE, Chintan Patel, Member, IEEE, and Nilanjan Banerjee “Adaptive and Personalized Gesture Recognition Using Textile Capacitive Sensor Arrays”.

[3] Tim oates, laura Firoiu and paul R. cohen “Clustering Time series with Hidden markove Models and Dynamic Time Wrping”.

[4] Holger Junkera, Oliver Amfta,∗, Paul Lukowiczb, Gerhard Tröstera “Gesture spotting with body-worn inertial sensors to detect user activities”. aWearable Computing Lab., ETH Zurich, Gloriastrasse 35, 8092 Zurich, Switzerland Embedded Systems Group, University of Passau, Innstrasse 43, 94032 Passau, Germany Received 10 January 2007; received in revised form 15 November 2007; accepted 19 November 2007.

[5] “From Dynamic Time Warping (DTW) to Hidden Markov Model (HMM)”. Final project report for ECE742 Stochastic Decision Chunsheng Fang University of Cincinnati, 2009/3/19.

[6] “Everything you know about Dynamic Time Warping is Wrong “. Chotirat ann Ratanmahatana Eamonn Keogh University of California, Riverside C 92521.

[7] Ali Erol a,*, George Bebis a, Mircea Nicolescu a, Richard D. Boyle b, Xander Twombly b” Vision-based hand pose estimation: A review” Received 13 September 2005; accepted 13 October 2006 Available online 19 January 2007 Communicated by Mathias Kolsch.