The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

A DCT-Gaussian Classification Scheme for Human-Robot Interface

Srinivas Kota, Michael Mace, Lalit Gupta, and Ravi Vaidyanathan

Abstract—**The ultimate success of a human-robot-interface system depends on how accurately user control signals are classified. This paper is aimed at developing and testing a strategy to accurately classify human-robot control signals. The primary focus is on overcoming the dimensionality problem frequently encountered in the design of Gaussian multivariate signal classifiers. The dimensionality problem is overcome by selecting, using two different ranking criteria, a small set of linear combinations of the input signal space generated by the discrete cosine transform (DCT). The application of the resulting DCT-Gaussian signal classification strategy is demonstrated by classifying tongue-movement ear-pressure (TMEP) bioacoustic signals that have been proposed for control of an assistive robotic arm. Classification results show that the DCT-Gaussian classifiers outperform classifiers described in a previous study. Most noteworthy is the fact that the Gaussian multivariate control signal classifiers developed in this paper can be designed without having to collect a prohibitively large number of training signals in order to satisfy the dimensionality conditions. Consequently, the classification strategies will be especially beneficial for designing personalized assistive interfaces for individuals from whom only a limited number of training signals can reliably be collected due to severe disabilities.**

I. Introduction

The goal of this paper is to develop a strategy to accurately classify signals that are used to control assistive human-robot (or human-machine) interfaces (HMIs). In general, devices for human-robot interface involve detecting an input from a user and converting that signal into a command signal, which in turn causes a desired event to occur. However, nearly all of the mechanisms designed for human control of peripheral devices require the user to generate an input signal through bodily movements, most often with their hands, arms, legs, or feet. Such devices predispose that the user is not utilizing their appendages for any other activity and clearly exclude individuals with impairments that cause painful or limited extremity control. Spinal chord injuries (SCI), loss of motor control due to stroke, central nervous system (CNS) disorders, amputations, paraplegia, severe arthritis, and repetitive strain disorders (RSD) all represent examples of these impairments.

A very large number of researchers have attempted to address this issue through the design of interfaces based on mapping input signals generated directly from the body. Examples of such interface control signals include speech [1]-[3], ear pressure signals [3, 4], electromyographic signals [5, 6], electroencephalogram (EEG) and event-related potential (ERP) brainwaveforms [6-9], gestures and motion tracking [10-12], and tracking of eye movement [13] (a brief survey is presented in [3]). In such signal classification problems, the dimension of the input signal space is typically quite large. It is often not possible to collect enough reliable signals to exceed the dimension of the signal space. This is particularly true for interfaces designed to aid individuals with severe disabilities to communicate and/or to control devices. For such cases, the implementation of multivariate parametric classifiers can be facilitated by decreasing the dimension of the signal space through feature selection [14]. Fusing a small set of features into a multivariate feature vector will definitely overcome the dimensionality problem; however, the performance of the resulting multivariate classifier will be dictated by the choice and discriminatory quality of the selected features.

In this paper, we investigate the use of the well-known the discrete cosine transform (DCT) as means to decrease the dimension of the signal space for multi-class multivariate classifier design and evaluation. Although the popular principal component transform (PCT) is optimal in the sense that it minimizes the mean-square error between the data reconstructed and original data, the drawback with the PCT is that its basis vectors are data dependent. The DCT, on the other hand, is a data-independent sinusoidal transform and also has the information packing ability close to that of the optimal principal component transform (PCT). The information packing property and the ease of computing the DCT makes it an ideal candidate transform for dimensionality reduction. We, therefore, investigate issues related to selecting DCT coefficients for the purpose of dimensionality reduction in multi-class signal classifier design. We specifically focus on the design of Gaussian signal classifiers because one of the most often made assumptions is that the class-conditional density function of the feature vector is multivariate Gaussian [15, 16].

A. . Organization of paper

We begin by presenting an overview of assistive robotic interface architectures in Section II. Section III focuses on the need for dimensionality reduction and on issues pertaining to the estimation of the parameters of the Gaussian discriminant functions. The use of the DCT to decrease dimensionality by forming features which are linear combinations of the discrete input signal space for single-channel signals are described in Section IV. The training and testing phases of the resulting DCT-Gaussian classification strategy is summarized in Section V while sections VI and VII map the implementation of the system to control a robotic arm through tongue movements. Conclusions are summarized in Section VIII.

Manuscript submitted March 1, 2009; supported by UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/F01869x/1

Srinivas Kota and Lalit Gupta are with Southern Illinois University, Carbondale, IL, USA ([lgupta; skota]@siu.edu)
Michael Mace is with the Department of Mechanical Engineering at the University of Bristol, Bristol, BS8 1TR, UK, (m.mace@bristol.ac.uk)
Ravi Vaidyanathan is with the Department of Mechanical Engineering at the University of Bristol, Bristol, BS8 1TR, UK * and the US Naval Postgraduate School, Monterey, CA, USA (r.vaidyanathan@bristol.ac.uk)

II. HUMAN ROBOT INTERFACES

Almost any system for human control of a robotic device will consist of three fundamental components: 1) the remote robot; 2) the communication link; and 3) the human-robot interface. Applications can range from the control of wheelchairs using body movements, to direction of robotic assist mechanisms, to the direct communication with computers using brain waveforms.

Control signal acquisition using a suitable sensor is the first step in the process and the following preprocessing operation typically includes filtering the acquired control signal. Detection involves determining whether a control signal has been initiated by the user. After a control signal has been detected, the start and end points of the control signal are estimated through segmentation. In the next step, features that distinguish different control signals are extracted. Based on the features, a classifier is designed to determine the class of a control signal during operation. The classifier output is used to generate a command signal to either communicate or to control a device. The whole process, therefore, could be regarded as a mapping of an interface control signal to an robotic or peripheral command signal.

III. DIMENSIONALITY ISSUES

For a C-class classification problem, the most general form of the Gaussian discriminant functions can be written as [15]

$$g_c(Z) = -(1/2)(Z-\mu_c)^T \Psi_c^{-1}(Z-\mu_c) - (1/2)\ln|\Psi_c| + \ln P_c; c=1,2,\ldots,C \quad (1)$$

where Z is a D-dimensional column feature vector, μ_c is the D-dimensional mean column vector of class c, Ψ_c is the (DxD) covariance matrix of class c, and P_c is the *a priori* probability of class c. It is often assumed that the covariance matrices are equal for all classes, that is, $\Psi_c = \Psi$, $c = 1,2,..C$.. For this case, Equation (1) can be written as

$$g_c(Z) = -(1/2)(Z-\mu_c)^T \Psi^{-1}(Z-\mu_c) + \ln P_c; \quad c = 1,2,\ldots,C \quad (2)$$

In practice, ensembles of each signal class are collected and used to estimate the classifier parameters and to test the classifier. The most commonly used technique to generate training and test sets is to randomly partition each ensemble into 2 equal-size sets. Another technique that is often used is the 'leave-one-out' or 'jackknife' method [16]. The parameters of the classifier are estimated using the feature vectors in the training set. Let the number of feature vectors in the training set of class c be N_c and the D-dimensional feature vector of pattern class c be Z_c. The parameters needed to implement the Gaussian discriminant function of Equation (1) for class c can be estimated as

$$\hat{\mu}_c = (1/N_c)\sum_{j=1}^{N_c} Z_{jc}; \quad c = 1,2,\ldots,C \quad (3)$$

$$\hat{\Psi}_c = (1/N_c)\sum_{j=1}^{N_c}(Z_{jc} - \hat{\mu}_c)(Z_{jc} - \hat{\mu}_c)^T; \quad c = 1,2,\ldots,C \quad (4)$$

where Z_{jc} is the jth feature vector in the training set of class c. The covariance matrix in Equation (3) can be estimated as

$$\hat{\Psi} = (1/N)\sum_{j=1}^{N} Z_j Z_j^T \quad (5)$$

where $N = N_1 + N_2 + \ldots + N_c$ and Z_j is the jth feature vector of the training set formed by pooling the zero-mean feature vectors of the C classes. That is, the set used to estimate the common covariance matrix is given by

$$\{Z_1 - \mu_1\} \cup \{Z_2 - \mu_2\} \cup \ldots \cup \{Z_C - \mu_C\} \quad (6)$$

where the symbol $\cup$ is used to represent the pooling operation and $\{Z_c - u_c\}$ represents the zero-mean training set ensemble of class c. The estimate of the covariance matrix must be non-singular in order to determine the inverses. For the estimate to be non-singular, the dimensionality conditions $N_c > D$, $c = 1,2,..C$ has to be satisfied in Equation (4) and the dimensionality condition $N > D$ has to be satisfied in (5).

It is often not be practical to collect enough training signals for $N > D$, especially from physically challenged individuals, to satisfy the dimensionality conditions [14]. Consequently, estimates of the covariance matrices will be singular and it will not be possible to determine the discriminant functions in Equations (1) and (3). The next section presents the DCT-based dimensionality reduction strategy.

IV. DISCRETE COSINE TRANSFORM (DCT) DIMENSIONALITY REDUCTION

Because the DCT efficiently packs information into a few coefficients, the dimension of the input feature space can be decreased by discarding a large number of DCT coefficients that do not contain useful information. The one-dimensional discrete cosine transform (DCT) of a sequence $Z=z(j)$; $j=0,1,\ldots,(D-1)$ is given by

$$Y(u) = \alpha(u)\sum_{j=0}^{D-1} z(j)\cos\left[\frac{(2j+1)u\pi}{2D}\right], \; u = 0,1,\ldots,(D-1) \quad (7)$$

The inverse DCT is given by

$$z(j) = \sum_{u=0}^{D-1}\alpha(u)Y(u)\cos\left[\frac{(2j+1)u\pi}{2D}\right], \; j = 0,1,\ldots,(D-1) \quad (8)$$

Where:

$$\alpha(u) = \begin{cases} \sqrt{1/D}, u = 0 \\ \sqrt{2/D}, u = 1,2,\ldots,(D-1) \end{cases} \quad (9)$$

The one-dimensional DCT may be written as the matrix operation

$$\tilde{Z} = \Phi Z \quad (10)$$

where Φ is the $(D \; x \; D)$ basis matrix whose elements are given by

$$\phi(u,j) = \alpha(u)\cos\left[\frac{(2j+1)u\pi}{2D}\right], u = 0,1,\ldots,(D-1); \; j = 0,1,\ldots,(D-1) \quad (11)$$

The $(D \; x \; D)$ linear transformation matrix Φ is fixed and does not have to be determined using a training set. The dimension of the transformed feature vector $\tilde{Z}$ is the same as that of Z, that is $(D \; x \; 1)$. The dimension of the transformed feature vector can be decreased to a specified d, $d < D$ by selecting d rows of Φ according to a defined selection criterion. The selection of rows (basis vectors) can in turn be based on selecting DCT coefficients. The two criteria most often to select DCT coefficients are based on the magnitudes and variances of the coefficients. Because each vector Z is transformed independently, the magnitude of each coefficient will very likely vary within the training vectors of each class (intra-class variations) and across the classes (inter-class variations). To accommodate these variations, the intra-class rank of each coefficient is determined from the rank of the intra-class rank-sum of the coefficient. The intra-class rank of each coefficient is summed to obtain the inter-class rank-sum and the final interclass ranking is obtained by from the rank of the inter-class rank-sum. In a similar fashion, the final ranking of each coefficient is determined from the rank of the inter-class rank-sum based on the intra-class variance for the variance criterion.

V. Classification Strategy

The complete multivariate DCT-Gaussian classification strategy based is summarized in Figures 1 and 2. Figure 1(a) shows the steps to determine the initial *(d x D)* dimensionality reduction matrix Φ_d. The training sets of all *C* classes are grouped into $\{Z_c\}$. The estimation of the mean vectors and the covariance matrices needed to implement the discriminant functions of the multivariate Gaussian classifier are shown in Figure 1(b). The procedure for classifying a test signal Z_c into a class c^* is summarized in Figure 2.

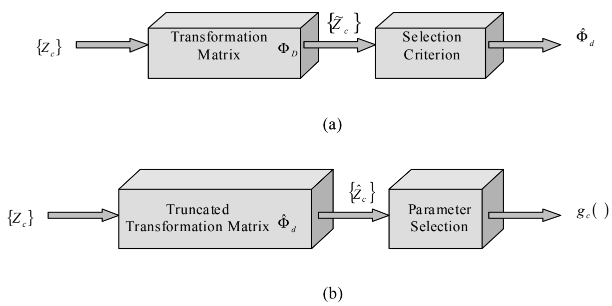

Figure 1: Training phase of classification strategy

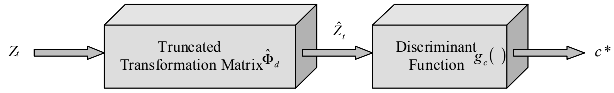

Figure 2: Testing phase of classification strategy

In order to demonstrate the application of the DCT-Gaussian multivariate classifiers summarized in Figures 2 and 3, the next section describes a HMI signal classification problem in which the dimensionality problem is known to frequently occur. That is, the number of vectors in the training set is less than the dimension of input feature vectors.

VI. Classification Example for System Performance

A. Tongue Movement Ear Pressure Signal Interface System

In a previous study, we introduced a new tongue-movement based concept for hands-free communication and control in robotic control applications [4]. It was shown that tongue movements within the human oral cavity create unique, subtle pressure signals in the ear that can be processed to produce command signals in response to that movement. The pressure changes that occur in the ear canal due to tongue movement were detected using a microphone inserted into the ear-canal. It was demonstrated that the ear-pressure signals are distinct for each tongue movement and the signals can be classified accurately. Consequently, tongue movements can be mapped, via ear pressure signals, into HMI command signals without inserting any device in the oral cavity. While many systems for severely disabled individuals have leveraged the oral cavity for control input (mouth joysticks, sip-and-puff tubes) this is the first system we are aware of capable of mapping tongue movements into command signals without inserting any device in the oral cavity and has resulted in the first available wheelchair control system that does not require bodily movement or insertion of a device in the oral cavity [17]. We have also demonstrated that spoken words can be recognized from the ear pressure signals [3]. Therefore, speech control signals can also be mapped via ear pressure signals into HMI command signals. Figure 3 shows the microphone-earpiece housing, a test subject wearing the device, and a graphic of its insertion in the ear, respectively.

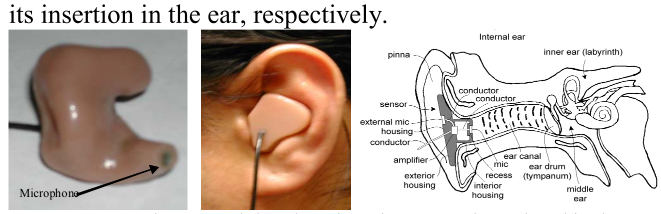

Figure 3: Earpiece containing the microphone, Earpiece placed in the ear of a subject, Graphic of earpiece inserted in the ear canal

B. Data Collection

Data from 8 test subjects were used for application and evaluation of the DCT classification strategies developed in this paper and to compare the performances of these strategies with the results reported in our previous studies. The data consisted of tongue-movement ear-pressure (TMEP) signals of 2 female and 6 male (healthy) subjects corresponding to 4 tongue movement classes (Up, Down, Left, and Right tongue movements [4]). The signals were sampled at 2 kHz. Each movement was repeated at least 100 times so that 100 tongue movements could be randomly selected to represent each tongue movement class. The signals were filtered (passband of 150Hz) and segmented using the techniques developed in [4]. The durations of the TMEP segments were 200 msec.

We now assume an additive signal plus noise model in which the ear pressure signal due to the tongue movement is deterministic and the background noise is random. Furthermore, we assume that the signal and noise are uncorrelated and that the background noise is statistically equivalent across the tongue-movement classes. Therefore, the covariance matrix is the same for all classes and can be estimated using the mixture training vectors of all 4 classes according to Equation (5). It is also assumed that the TMEP signals are equally likely for the 4 classes. For each subject, the training set of each class contained 50 ear-pressure signals. Therefore, the training set to estimate the covariance matrix of the discriminant function consisted of 200 signals. Consequently, the maximum dimension *d* of the feature vector for the classifier had to be less than 200.

C. Recognition Accuracy

For each subject, we selected the 199 highest ranked DCT coefficients out of the 400 DCT coefficients. The performance was then evaluated systematically to determine the L coefficients (out of 199), that gave the best classification accuracy. For each case, the dimension of the resulting Gaussian multivariate classifier was equal to the number of features selected. When 1 coefficient was selected, the classifier was a univariate classifier. The classification accuracies for the eight subjects are shown in Table I. The best result for each subject is shown in boldface. Also included is the value of L, in parentheses, that gave the best result reported in the table. Each result in the table was averaged over 200 trials using random sampling to form the training and sets. The table also shows the results of the best classifier (a decision fusion classifier) reported in the previous study for exactly the same data [4]. Furthermore, for each subject, the Wilcoxon rank-sum test was performed on the classification accuracies obtained from the best DCT-Gaussian classifiers and the classification accuracies of the best classifiers from the previous study. When more than one classification accuracy was 100%, any classifier yielding 100% was selected as the best. The differences between the new and previous best results are statistically significant. It should be noted that though the mean differences are quite small, the standard deviations are also quite small, especially when the average classification accuracies are very high.

Figure 4 shows examples of the classification accuracies as a function of the number of DCT coefficients selected using the maximum variance selection criterion; these criterion may vary based on the importance for each tongue movement. In this work they were selected to give the maximum overall accuracy; they were not task dependent. For this particular case (Subject S_3), the 46 highest ranked DCT coefficients gave the best results. What is interesting to note is that the classification accuracies do not follow the trend of reconstruction errors. That is, instead of increasing as more coefficients are added, the classification accuracies peak and then decrease. The optimal amount may be found experimentally, however we not that the variance is quite low over a wide range of coefficients, which allows for reasonable and quick estimation. These figures demonstrate another important difference between the selection of transform coefficients for compression and classification applications. Also interesting to note is that only a small number of linear combinations of the input are needed. This is due to the high compaction of information into a small set of transform coefficients. Therefore, a large number of DCT coefficients can be dropped to decrease the dimensionality.

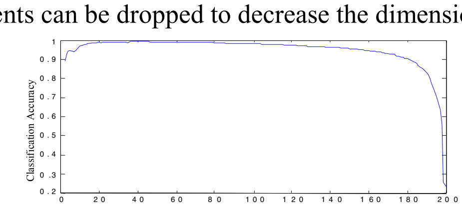

Figure 4: Classification accuracies as function of the DCT coefficients

VII. SIMULATION CASE STUDY: CONTROL OF A ROBOT ASSISTED MANIPULATOR (PROSTHETIC) ARM

In our past work, we have demonstrated the utility of TMEP signals to control mobile robots [3] and power wheelchair systems [17]. The more powerful recognition accuracies lent by the new classifier, however, open a new breadth of potential applications. To demonstrate the performance of this new classification strategy for use in robotic assistive technology, we introduce its integration for simulated control of a 4 degree of freedom robotic manipulator. The design of the manipulator arm (Figure 5) was based upon anthropomorphic dimensions including a shoulder, elbow and wrist joint with attached gripper to allow for arm and object elevation in a two-dimensional programming environment. Physical constrains were applied to the maximum and minimum joint flexure in the same manner a human arm is constrained enabling the arm to be envisaged more as a prosthetic aid then simply a robotic manipulator [18]. The task outlined for this simulation was designed to be challenging to ensure the robustness of the classifier was exercised allowing the potential of this classification strategy within a real time assistive system to be forecast. The control scheme shown in Table 2 was chosen to map the tongue movements to the kinematics of the arm with no sensory feedback involved, so that the completion of the task would rely completely on the classification process and follow the desired trajectory blindly. The arm moved in discrete time steps, thus did nothing in the absence of control input. Experiments were performed in simulation to allow several trials to be conducted rapidly for statistical analysis.

Subject	**Table 1:** DCT Classification accuracies			
	MAG Acc (%)	**p-value**	**VAR Acc(%)**	**p-value**
S1	**99.82 (20)**	9.6496e-027	99.76 (28)	1.2e-023
S2	**99.73 (23)**	1.1477e-048	99.62 (35)	7.0454e-046
S3	**99.33 (48)**	1.7448e-020	99.27(46)	3.2134e-021
S4	99.76 (26)	0.0001	**99.77 (27)**	0.0005
S5	**100.0 (15)**	1.2714e-075	**100.0 (19)**	7.0989e-075
S6	**99.96 (35)**	2.3309e-072	**99.96 (35)**	2.3309e-072
S7	**100.0 (16)**	9.5903e-028	**100.0 (19)**	5.3083e-029
S8	**99.91 (18)**	3.3801e-049	99.85 (17)	8.5068e-045
Ave	**99.81**		99.78	

In addition to utilizing the new classification system, we also implemented a control system for the manipulator utilizing data from our previous research using our decision fusion classification algorithm. The accuracy of this algorithm [4] makes it an ideal benchmark with which to compare the new DCT classification scheme.

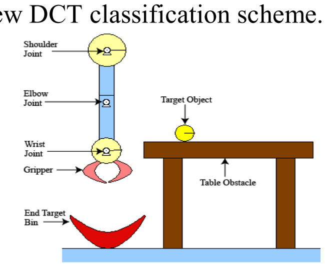

Figure 5: 2D simulation of robotic manipulator arm (prosthetic arm)

Trigger/TMEP	Response	Target
1 (L)	Incrementally rotate by 5 degrees	Shoulder
2 (D)		Elbow
3 (R)		Wrist
4 (U)	Change direction	Rotation
4,4 (Compound)	Open/Close	Gripper

Table 2: Control scheme mapping kinematics of robotic manipulator arm to tongue actions

Subject	Average Number Actions Correctly Classified to Complete Task	No. of Collisions per Run	No. of Times Task Completed
1	120.2	1	15
2	120.7	0	15
3	121.9	0	15
4	120	0	15
5	120	0	15
6	120	0	15
7	120	0	15
8	120.1	0	15
Avg	**120.4**	**0.125**	**15**

Table 3: Task results for eight test subjects over fifteen test runs

The task selected involved a complex act of controlling the dextrous manipulator to pick up a small round object from a surface, and shift it to a container. The task itself involved the arm starting in full downward extension; the arm is then rotated clockwise and the gripper brought up using the elbow joint. The arm can then move anticlockwise pass the table top surface where the gripper can be opened using a compound (double) movement of the fourth (up) action, and then brought down to the ball (object) using a combination of wrist and elbow movements. Once the ball has been gripped, the motions of the arm are reversed allowing it to be brought back towards its starting position where the ball can be dropped into the target bin. Snapshots of the arm whilst performing the task are shown in Figure 6 with the target number of tongue movements for a perfect run being 120 (note that TMEP signals can be produced and recognized in approximately 300 ms) . The task itself was chosen to mimic an everyday task that could be performed by a normal person such as picking up an object from the table and dropping it into a bin and targets future applications for control of prosthetic limbs. If a mistake is made then the control system accounts for this and has to correct the mistake using specific combinations of the appropriate movements to bring the arm back to the correct orientation. There are certain critical points associated with the task (highlighted by red circles in Figure 6) where if an error is made by the classifier then it can lead directly to a collision or even complete task failure. For instance when trying to grip the object, a compound movement of signal four is required, which significantly increases the probability of misclassification; if either of these actions is wrongly classified then the ball and arm could easily collide or the arm be knocked into the table resulting in complete task failure.

Each subject would be judged on overall accuracy, total tongue movements needed, number of collisions, and whether the task was completed. Training data was collected generically, without targeting this task. Data was collected in an open lab (uncontrolled) environment. Results for eight subjects performing the task 15 times are presented Table 3.

The results of the simulation show this classification strategy to be extremely robust, even when performing in an environment where misclassifications can be critical. It also helps to highlight the possibility of its integration into a real-time assistive technology control system with a 100% (15 out of 15) completion of the arduous task set out and a total of only one collision across the whole results set. With such a simple controller implemented and with zero feedback within the system both internally and externally (visually), it meant that the output of the task relied completely on the classification obtaining an average accuracy of 99.85% (with a standard deviation of less then 0.3%) across the eight subjects tested.

Table 4 presents results for the same task with the same test subjects simulation using the results obtained previously based on our past decision fusion classification (DFC) patter recognition algorithm (note that data for Subject 5 was

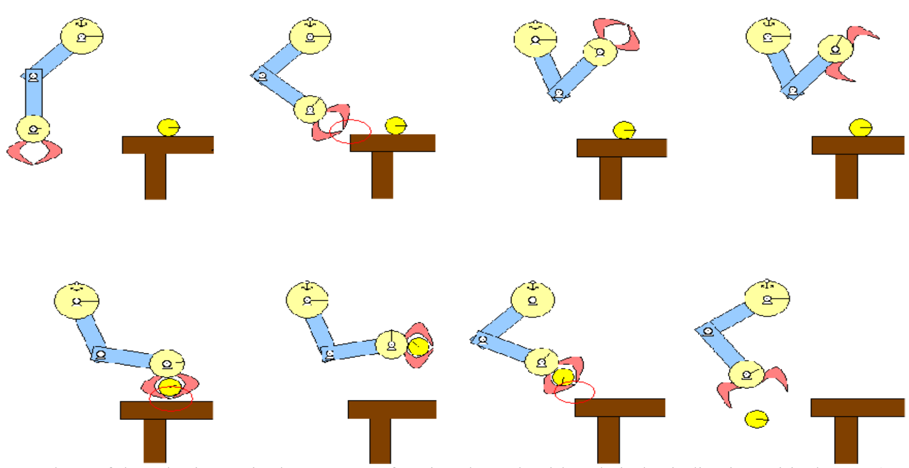

Figure 6: Snapshots of the robotic manipulator arm performing the task with red circles indicating critical areas (start point - top left, end point - bottom right)

unavailable for this test due to data corruption during collection). This algorithm had shown the highest ever previously recorded results for TMEP recognition. As can be seen from Table 4, the classifier is capable of completing the prescribed task 80% of the time compared with 100% with the DCT algorithm. Furthermore, it requires significantly more tongue actions and results in a higher number of collisions.

Subject	Average Number Actions Correctly Classified	No. of Collisions per Run	No. of Times Task completed
1	140.2	10	9
2	121	0	15
3	232	1	14
4	136.3	0	15
6	139.4	0	15
7	143.4	14	8
8	150	11	5
Avg	**152.1**	**4.2**	**12**

Table 4: Decision Fusion Classification (DFC) results for manipulator control

VIII. Conclusions

Implementing Gaussian multivariate classifiers for control signal classification is quite straight-forward when there is a sufficient number of training vectors to ensure that the estimate of the covariance matrix is non-singular. However, when the number of training vectors used to estimate the covariance matrix is less than the dimension of the training vectors, the discriminant functions cannot be implemented because the estimates of the covariance matrices are singular. In such cases, which frequently occur in the design of communication and control systems for individuals with severe disabilities, the most promising solution is to decrease the dimension of the signal space. This work has shown that the DCT basis vectors selected using the class-dependent ranking criteria are quite effective in decreasing the dimension for multivariate classifier development. In terms of implementation, the DCT based classifiers are quite simple because the DCT transformation matrices are fixed. It is important to note that the classifiers can be specifically developed for an individual without having to collect a prohibitively large number of training signals simply to satisfy the dimensionality conditions. Therefore, the Gaussian multivariate signal classification strategies based on the DCT dimensionality reduction techniques developed in this paper should be especially beneficial for designing assistive robotic interfaces for individuals from whom only a limited number of training signals can be collected due to severe disabilities. Moreover, the formulations of the strategies are quite general and can, therefore, be applied to a wide range of problems involving the classification of multivariate signals. In this work, we have specifically demonstrated the algorithm's utility for a complex task of controlling a robotic arm through tongue-movement ear pressure (TMEP) signals. We envision this work to form the basis of a new control system for assistive robotic mechanisms, including potential application to prosthetic arms.

Acknowledgments

The authors gratefully acknowledge Think-A-Move, Ltd. of Cleveland, OH, USA for their commercial research in this area and Hearing Technologies International of St. Petersburg Florida, USA. This research was supported by the UK Engineering and Physical Sciences Research Council grant EP/F01869X.

References

[1] H. Medicherla, and A. Sekmen, "Human-robot interaction via voice-controllable intelligent user interface," *Robotica*, vol. 25, no. 5, pp. 521-527, Sep. 2007.

[2] H. G. Nik, G. M. Gutt, and N. Peixoto, "Voice recognition algorithm for portable assistive devices," *Sensors*, pp. 997-1000, 28-31 Oct. 2007.

[3] R. Vaidyanathan, M. Fargues, R.S. Kurkan, L. Gupta, S. Kota, R.D. Quinn, & D. Lin, "A dual-mode human-robot teleoperation interface based on airflow in the aural cavity," *International Journal of Robotics Research*, Special Issue, 1205-1223, November 2007.

[4] R. Vaidyanathan, B. Chung, L. Gupta, H. Kook, S. Kota, & J. West, "A tongue-movement communication and control strategy for hands-free human-machine interfaces," *IEEE Transactions on Systems, Man, & Cybernetics* – A, vol. 37, No. 4, 533-546, July 2007.

[5] O. Fukuda, T. Tsuji, M. Kaneki, and A. Otsuka, "A human-assisting manipulator teleoperated by EMG signals and arm motions," *IEEE Transactions on Robotics and Automation*, vol. 19, no. 2, pp. 210-222, April. 2003.

[6] A. Ferreira, W. C. Celeste, F. A. Cheein, T. F. Bastos-Filho, M. Sarcinelli-Filho, and R. Carelli, "Human-machine interface based on EMG and EEG applied to robotic systems," *Journal of NeuroEngineering and Rehabilitation*, vol. 5, March. 2008.

[7] K. Tanaka, K. Matsunaga, and H. Wang, "Electroencephalogram-based control of an electric wheelchair." *IEEE Transactions on Robotics*, vol. 21, no. 4, pp. 762-766, Aug. 2005.

[8] J. D. R. Millán, F. Renkens, J. Mouriño, and W. Gerstner, "Noninvasive brain-actuated control of a mobile robot by human EEG." *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1026-1033, June. 2004.

[9] L.G. Cohen and N. Birbaumer, "The physiology of brain-computer interfaces," Journal of Physiology, Vol. 579(3), 570, 2007.

[10] L. Gupta and S. Ma, "Gesture-based interaction and communication: automated classification of hand gesture contours," *IEEE Transactions on Systems, Man, & Cybernetics – C*, vol. 31, No. 1, 114-120, 2001.

[11] M. Urban, and P. Bajcsy, "Fusion of voice, gesture, and human-computer interface controls for remotely operated robot." *Proceedings of the 8th International Conference on Information Fusion*, vol. 2, July 25-July 28, 2005.

[12] X. Yun and E. R. Bachmann, "Design, Implementation, and Experimental Results of a Quaternion-Based Kalman Filter for Human Body Motion Tracking," *IEEE Transactions on Robotics and Automation,* vol. 22, pp. 1216-1227, Dec 2006

[13] Y. Chen and W. S. Newman, "A human-robot interface based on electrooculography," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2004.

[14] Gupta L., Kota S., Murali S., Molfese D., Vaidyanathan R., "Dimensionality Reduction Strategies for the Design of Human Machine Interface Signal Classifiers", *IEEE Int. Conference on Systems, Man, and Cybernetics*, 2432-2436, Singapore, Oct 2008

[15] G. J. McLachlan, Disciminant Analysis and statistical pattern recognition, John Wiley, New York, 1992.

[16] R. O. Duda, P. E. Hart, D. G. Stork, "Pattern Classification", 2nd ed., John Wiley & Sons. Inc, New York, 2001

[17] http://www.think-a-move.com/, Think-A-Move, Ltd, Cleveland, USA

[18] Wang, L., Xie, M., Zhong, Z. W., Yang, H. J. & Li, J. (2008) Design of dexterous arm-hand for human-assisted manipulation. *Intelligent Robotics and Applications,* 53**,** 1233-1240