2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Comparison of various linear discriminant analysis techniques for fault diagnosis of Re-usable Launch Vehicle

A. Akilez Krishnamurthy, Madhu N. Belur and Debraj Chakraborty

***Abstract*— In this paper, we use Linear Discriminant Analysis (LDA) techniques to diagnose Reaction Control System (RCS) thruster faults in a Re-usable Launch Vehicle (RLV) upon re-enrty. An RCS thruster operates in binary mode i.e. either ON or OFF. A mode is a particular combination of thruster ON/OFF values which is commanded by the controller. Different Linear Discriminant Analysis (LDA) techniques like CLDA (Classical LDA), FSLDA (Foley- Sammon LDA), ULDA (Uncorrelated LDA) are implemented in Matlab and used here to estimate the mode in which the vehicle lies based on the double derivative of pitch, roll and yaw angles. If the estimated mode is not same as the commanded mode then it implies a fault. Misclassification percentage of each of the LDA techniques with respect to percentage of training samples used, number of loading vectors, number of nearest modes and number of instances dropped after a mode change has occurred are evaluated and are compared to decide the one that suits the application. A thorough comparison of the three LDA techniques brings out a contrasting conclusion: unlike reported in LDA literature, CLDA performs better than FSLDA and ULDA for this RCS fault application, though FSLDA and ULDA are advanced variants of the CLDA.**

***Keywords:* Linear Discriminant Analysis (LDA), Fault diagnosis, Re-usable Launch Vehicle (RLV), Misclassification percentage.**

I. INTRODUCTION

Re-usable Launch Vehicle (RLV) are those launch vehicles which have the capability to be used more than once. After disengagement from the payload, they return back to earth and during their re-entry into earth's atmosphere there comes a stage in which their motion is controlled by Reaction Control System (RCS) thrusters. These thrusters are ON-OFF kind of devices that give torque to control pitch, yaw and roll angles of the vehicle. The performance index puts a bound on pitch, yaw and roll angles and whenever the vehicle crosses the bound, the controller sets a particular combination of RCS thrusters to bring the vehicle back to desired region. During this re-entry stage, the RLV is said to reside in different modes where each mode is a particular binary combination of RCS thrusters.

In this paper, we use Linear Discriminant Analysis (LDA) techniques to diagnose Reaction Control System (RCS) thruster faults. LDA algorithm is used to find new projection directions such that, when data samples are projected on those directions, the within mode distance decreases while between mode distance increases. Thus it helps in the classification process. Three different Linear Discriminant Analysis (LDA) techniques are used here. They are

- CLDA (Classical LDA)
- FSLDA (Foley-Sammon LDA)
- ULDA (Uncorrelated LDA)

These 3 LDA techniques are commonly used for classification purpose [1][7][11] and hence they are chosen here for analysis and comparison. These LDA techniques, from double differentiated values of pitch, yaw and roll angles, are used to estimate the mode in which the RLV lies. Controller commands a mode to RCS thrusters whenever the RLV crosses its set boundary limits on pitch, yaw and roll angles. When this estimated mode is not same as commanded mode then it implies a fault in RCS thrusters on the assumption that sensors are not faulty. The following is done for each of the LDA techniques and their misclassification percentage is compared to decide the one that best suits for this fault diagnosis of RLV application.

1) Percentage of training samples used is varied and the corresponding misclassification percentage is found out.
2) The number of samples dropped after each mode change is varied and corresponding misclassification percentage is found out to observe the effect of inertia and differentiation [1] of successive samples.
3) The number of loading vectors used is varied and misclassification percentage is found out to know the optimal number of loading vectors (and hence the computational intensity) to be used for classification purpose.
4) Number of nearest modes/cluster to which the new test sample might belong to is varied and misclassification percentage is found out. This is because the test data, when projected onto the direction of loading vectors, need not necessarily belong to first nearest mode/cluster but it may belong to any of the few nearest modes/clusters.

A thorough comparison of the three LDA techniques brings out a contrasting conclusion: unlike reported in LDA literature[1][10], CLDA performs better than FSLDA and

This research was supported in part by the Bharti Centre for Communication, IIT Bombay and the Indian Space Research Organisation: Grant ISRO/RES/STC-IITB/08-09.

A. Akilez Krishnamurthy, Madhu N. Belur and Debraj Chakraborty are with the Department of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India. {akilez.krish@gmail.com}, {belur, dc}@ee.iitb.ac.in

[1]Strictly speaking, we are dealing with samples of a continuous variable. By differentiation we mean successive differences of samples using backward difference method. We use the word 'differentiation' to mean this for the rest of the paper.

ULDA for this RCS fault application, though FSLDA and ULDA are advanced variants of the CLDA[12].

LDA is also usually used for image processing applications like face recognition[7]. It is also used for object tracking applications[8]. Fault diagnosis is mostly done using artificial neural networks [9]. Techniques like support vector machines are also used for fault diagnosis applications[6].

The paper is organized as follows. Section II discusses the assumptions made, input data processing done and fault detection procedure used. Section III explains about the three different LDA techniques i.e. CLDA, FSLDA and ULDA. Section IV gives the results and discussion while section V serves as the conclusion to the paper.

II. Model, assumptions, input data processing and fault detection

A. Model

The RLV simulator model, a software simulated one, is governed by its six degree of freedom equations. These equations basically describe the RLV dynamics. They give the dependence of rate of change pitch, yaw and roll angles on the pitch, yaw and roll angles, moment of inertia about body axis of the vehicle and pitch, yaw and roll moments due to RCS thrusters. There are also equations governing the rate of change of side slip angle, rate of change of angle of attack and rate of change of vehicle velocity.

The model gives time, ON/OFF combination of thrusters, pitch, yaw and roll angles, their corresponding rate of change and their double differentiated values (with respect to time) as its output. There is a bound specified on these pitch, yaw and roll angles whose values is based on the performance index required. The pitch, yaw and roll angles from the output of the model is fedback and compared with these bounds. Necessary action (firing of appropriate thrusters) is then done to ensure that the pitch, yaw and roll angles stay within these bounds. The double differentiated values of pitch, yaw and roll angles, thruster ON/OFF combination and the corresponding time obtained from the model is used in this paper for fault diagnosis in the RLV.

B. Assumptions

Actual vehicle and the model has 12 thrusters out of which 4 are used to control pitch, 4 to control yaw and 4 to control roll angle. Each of these 4 act in pairs i.e. 2 are for rotation along clockwise direction and 2 for rotation along anticlockwise direction. Here in this paper it is assumed that only 6 thrusters are there by combining a pair of thrusters into one. Hence there are only 2 thrusters to control pitch (one for each direction), 2 to control yaw and 2 to control roll angle. Further, for control of a particular angle (pitch or roll or yaw) it is assumed that there are only 3 combinations namely both the thrusters are ‘OFF’ or one of them is ‘ON’ and other one is ‘OFF’ as the effect both the thrusters are ‘ON’ is same as both the thrusters ‘OFF’. Figure 1 shows the pitch, yaw and roll control of a flight along with two thrusters (marked circle) to control roll angle.

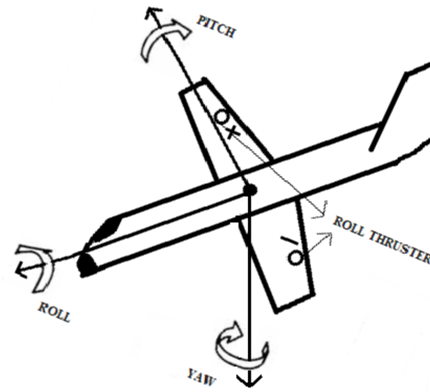

Fig. 1. Schematic figure showing roll control thrusters and pitch, yaw and roll angles of a flight

C. Input data processing

Input data available from the model (RLV simulator) are

- Binary combinations of RCS thrusters that gives the mode in which the RLV lies.
- Corresponding double differentiated values of pitch, yaw and roll angles.
- Corresponding time values.

These double differentiated values are differentiated further to check their effect on classification. Here the number of modes are 3^3 (3 combinations for each of 3 angles). The data is separated according to their mode. Hence now each mode consists of 6 columns first three corresponding to triple differentiated values of pitch, yaw and roll angles and next three columns corresponding to double differentiated values of pitch, yaw and roll angles. These six column vectors are fed into the different classification algorithm (CLDA, FSLDA, ULDA) for training and testing purpose. Input data processing, implementation of different LDA algorithms and comparison of their performance measures to find the best one that suits the application are done in Matlab. The binary combination of RCS thruster and their corresponding mode number is given in table I. Here 0 and 1 correspond to OFF and ON of RCS thruster respectively.

The reason for using double and triple differentiation of pitch, yaw and roll is as follows. Double differentiation of angle is proportional to torque, while impulse imparted is proportional to triple differentiation of angle. The effect of including them is discussed in section IV-F.

D. Fault diagnosis

After feeding the training samples to LDA classification algorithms, a set of loading vectors are obtained. The loading vectors are such that when the training samples are projected onto the direction of loading vectors, samples of any particular mode are close together while the individual modes are well separated from the other modes. The training samples are now projected onto the direction of loading vectors where each mode gets clustered appropriately. When a new test data comes in, it is projected onto the direction of loading vectors

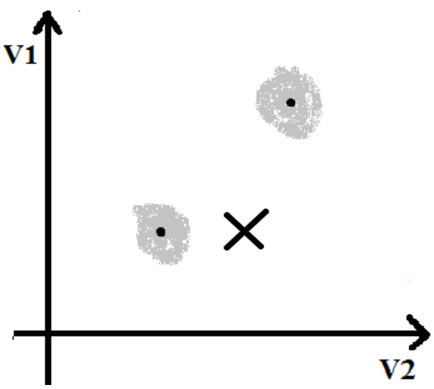

Fig. 2. Clusters of two modes in new coordinates $V1$ and $V2$. $\bullet \equiv$ mean of cluster and $\times \equiv$ new data point

and its distance with mean of each of the clustered modes is found out. The new data is estimated to belong to the mode to which it is closest. This is shown in figure 2. This estimated mode may not be same as the commanded mode given by the controller to the RCS thrusters. This mismatch between estimated mode and commanded mode implies that a fault has occurred. For example, if the estimated mode is 1 while the commanded mode is 2, then from table I it is possible to conclude that thruster 6(Roll -) is faulty. By using a similar analysis, any number of thruster faults can be diagnosed explicitly using this algorithm.

The block diagram of the overall procedure is given figure 3. The lower part of the block diagram i.e. input data processing, LDA blocks and comparison of estimated and commanded mode to detect fault are the one that is explained in this paper. Controller and sensor values are simulated using the model (RLV simulator). Upper part of the block diagram consist of the real hardware i.e. the plant (RLV) along with RCS thrusters. RCS thruster ON or OFF values produce forces that acts on the plant to change its pitch, yaw and roll angle which is then sensed by sensors. Further any variation between the estimated mode and actual mode implies a misclassification due to the LDA algorithm.

III. Linear Discriminant Analysis techniques

Three main types of Linear Discriminant Analysis (LDA) techniques are used here. They are

1) Classical Linear Discriminant Analysis (CLDA)
2) Foley- Sammon Linear Discriminant Analysis (FSLDA)
3) Uncorrelated Linear Discriminant Analysis (ULDA)

All these LDA techniques focus on finding the loading vectors that maximizes the Fisher criteria function (J). Maximizing the Fisher criteria implies maximizing the between class distance while minimizing the within class distance[4]. Let n be the number of observations, m be the number of measurements and p be the number of classes/modes. The Fisher criteria function (J) is given by the determinants,

$$J(W) = \frac{|W^T S_B W|}{|W^T S_W W|} \tag{1}$$

TABLE I

Mode number and binary combination of thrusters

Mode number	[Pitch+	Pitch-	Yaw+	Yaw-	Roll+	Roll-]
1	[0	0	0	0	0	0]
2	[0	0	0	0	0	1]
3	[0	0	0	0	1	0]
⋮			⋮			
27	[1	0	1	0	1	0]

where,
$W \in \mathbb{R}^{m\times m}$ is the loading vector matrix
$S_B \in \mathbb{R}^{m\times m}$ is the between scatter matrix
$S_W \in \mathbb{R}^{m\times m}$ is the within scatter matrix.

The loading vector matrix on which the data set is projected is specific for the particular LDA technique used. Here m is six as we have six measurements to be fed into the algorithm namely three corresponding to double derivative of pitch, yaw and roll angles and three corresponding to derivative of double derivative of pitch, yaw and roll angles. The procedure to find the scatter matrices is common for all the three methods [2]. The procedure to find the loading vectors by the three different methods is given below.

A. *Classical Linear Discriminant Analysis (CLDA)*

This is also called as Fisher Discriminant Analysis (FDA). This technique focuses on finding the vector w that maximizes the Fisher criteria (J) given by,

$$J = \frac{w^T S_B w}{w^T S_W w} \tag{2}$$

Hence the objective of the first FDA vector is to maximize the scatter between classes while minimizing the scatter within classes. The second FDA vector does a similar job in a direction other than the direction of first FDA vector. It turns out that the FDA loading vectors are equal to the eigenvectors w_k of the generalized eigenvalue problem

$$S_B w_k = \lambda_k S_W w_k \tag{3}$$

where the eigenvalues λ_k indicate the degree of overall separability among the classes when projecting the data onto w_k [2]. Thus the first loading vector is the eigenvector associated with the largest eigenvalue, the second loading vector is the eigenvector associated with the second largest eigenvalue and so on. A large eigenvalue indicates that when the data in the classes are projected onto the associated eigenvector then there is overall a large separation of class means relative to class variances and consequently, a large degree of separation among the classes along the eigenvector direction.

B. *Foley-Sammon Linear Discriminant Analysis (FSLDA)*

FSLDA can get perpendicular loading vectors [11][5]. FSLDA has the same first loading vector as CLDA. That is, its first loading vector maximizes the Fisher criterion given by Equation 1. Then the next loading vector that maximizes Fisher criterion is obtained by imposing a constraint that it

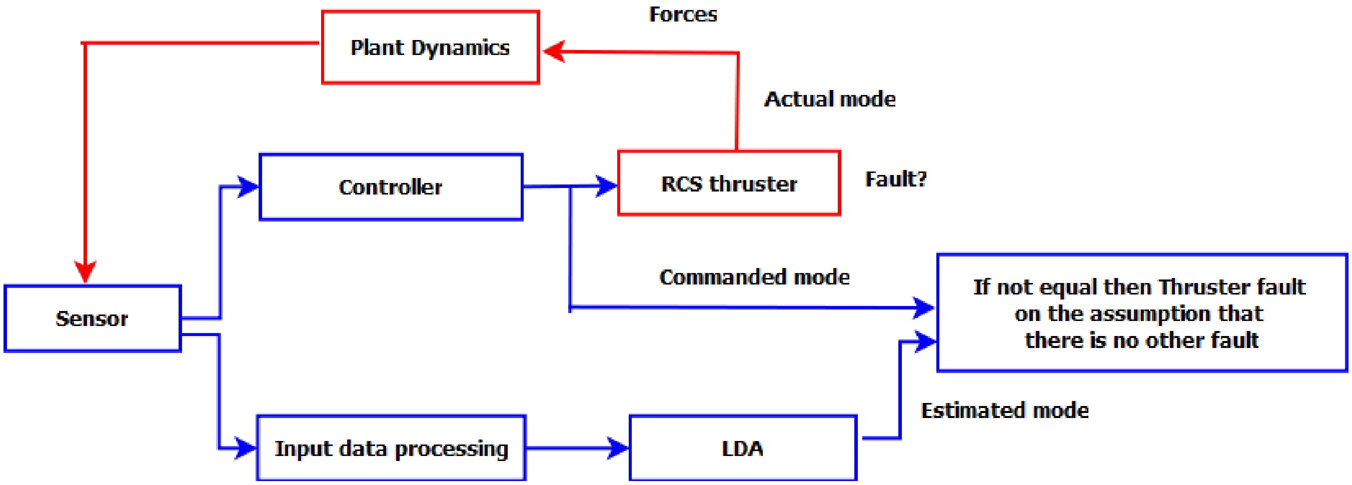

Fig. 3. Block diagram of the overall process

should be perpendicular to the first loading vector. Thus in general if $f_1 ..... f_j$ are the loading vectors then the constraint is

$$f_{j+1}{}^T f_k = 0, k = 1,2...j \tag{4}$$

It turns out that every $(j+1)_{th}$ loading vector is the eigenvector corresponding to the largest eigenvalue of the generalized eigenequation

$$MS_B f = \lambda S_W f \tag{5}$$

where

$$M = I - D(D^T S_W{}^{-1} D)^{-1} D^T S_W{}^{-1}, D = [f_1, f_2 .... f_j], (j > 1) \tag{6}$$

and I is the identity matrix sharing the size of S_W.

C. Uncorrelated Linear Discriminant Analysis (ULDA)

ULDA ensures that the variables in the projected space are uncorrelated [11][7]. ULDA has the same first loading vector as CLDA. That is, its first loading vector(u_1) maximizes the Fisher criterion given by Equation 17.

After obtaining the first j loading vectors u_1,u_2,..... u_j the $(j+1)^{th}$ loading vector is obtained by maximizing the Fisher criterion along with the following uncorrelation constraint

$$u_{j+1}{}^T S_T u_i = 0, (i = 1,2....j) \tag{7}$$

It turns out that the j^{th} loading vector is the eigen vector corresponding to the maximum eigenvalue of the generalized eigenequation

$$U_j S_B u = \lambda S_W u \tag{8}$$

where

$$U_1 = I_N$$

$$U_j = I_N - S_T D_j{}^T (D_j S_T S_W{}^{-1} S_T D_j{}^T)^{-1} D_j S_T S_W{}^{-1},$$

$$D = [u_1, u_2 .... u_j], (j > 1) \tag{9}$$

where I_N is the identity matrix sharing the size of S_W.

ULDA is same as CLDA [11] if

1) Rank(S_B) = $c-1$, where c is the number of classes.
2) S_W is non-singular.
3) All $c-1$ non zero eigenvalues of matrix $S_W^{-1} S_B$ are distinct.

IV. Results and Discussion

Misclassification is said to have occurred when the mode to which the new test data point gets classified as (estimated mode) is not same as the mode to which it should belong to (actual mode). Estimated mode and actual mode are those mentioned in figure 3. For example if the test data belonging to mode 2 is classified as mode 3 then misclassification is said to have occurred. How a mode classification takes place was already explained in the section II.D. The LDA techniques mentioned in section III are evaluated based on four different criteria and their effect on misclassification percentage is found out. The following subsections explain this.

A. Effect of percentage of training samples

Here the training data percentage is varied and fed into the 3 classification algorithms. The remaining data is used for testing. All the 6 loading vectors obtained are used here for classification. Further the data samples are such that after a mode change has occurred 2 data samples are dropped.

From figure 4 it can be observed that the misclassification percentage is almost constant for CLDA while for FSLDA and ULDA they seems to decrease with increase in training sample percentage beyond 70. The spread of data across different time ranges gets accounted automatically when the percentage of training data is increased.

B. Effect of number of instances dropped after mode change

Whenever a mode change occurs by change in the thruster combination that are ON/OFF, the vehicle does not respond immediately due to its inertia. Hence immediately after the mode change, the double differentiated[2] values of pitch, yaw

[2]Note that difference between consecutive samples is used to approximate the derivative (see footnote 1)

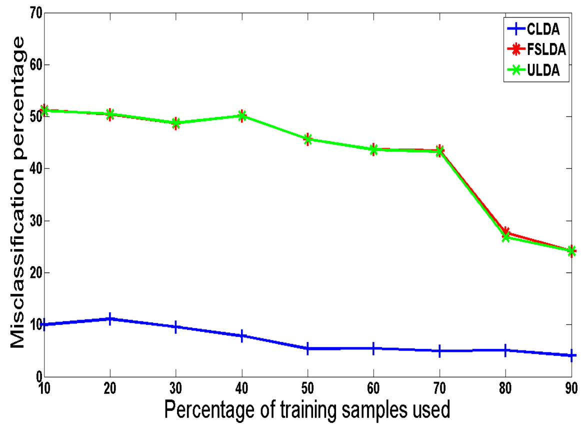

Fig. 4. Misclassification percentage versus percentage of training samples used

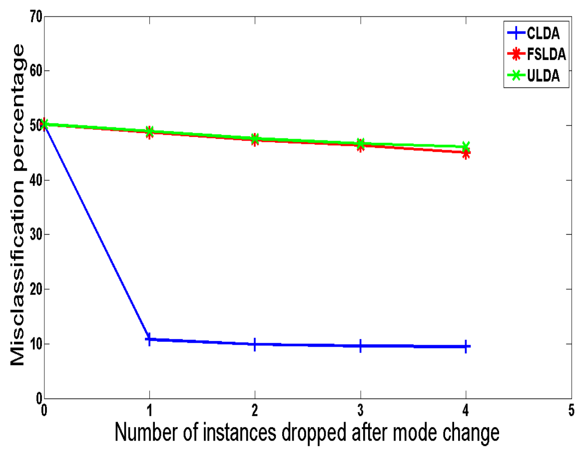

Fig. 5. Misclassification percentage versus number of instances dropped after mode change

and roll angles may neither correspond to the present mode nor to the previous mode. Further, when there is a mode change, the derivative calculated from successive samples to obtain triple differentiated pitch, yaw and roll angles may vary from that of those values in the present and previous mode.

Hence, here the number of samples dropped after mode change is varied from 0 to 4 and misclassification percentage is found for each of the three LDA algorithms. 30% training sample is used and rest is used for testing. All the six loading vectors are used for classification. From figure 5 it is clear that CLDA misclassification percentage is very less when compared with FSLDA and ULDA.

C. Effect of number of loading vectors

Here the number of loading vectors used for classification is varied from 1 to 6 in the order starting from the one corresponding to maximum Fisher criteria. 30% training sample is used and the rest is used for testing. Further the data samples are such that after a mode change has occurred 2 data samples are dropped.

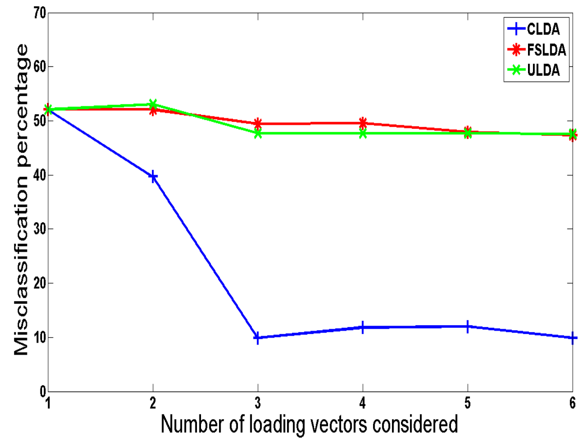

Fig. 6. Misclassification percentage versus number of loading vectors considered

It is evident from figure 6 that CLDA has lower misclassification percentage when compared with FSLDA and ULDA. Also the misclassification percentage decreases until first three loading vectors are considered and after that it is almost constant. Hence the loading vectors corresponding top three eigenvalues are good enough for classification. The top three eigenvalues are 2.17×10^5, 28.12 and 9.67. This says that the first loading vector is able to separate different modes and reduce the variance within each mode much better than the others[2].

D. Effect of number of nearest modes considered

In the above three variants, misclassification is said to have occurred if the nearest mode to which the new test data point gets classified as is not same as the actual mode to which the test data should belong to. Assuming that the test data point could belong to any of the few nearest modes/clusters instead of only the first nearest mode/cluster (for example one among the top 3 nearest modes from the test data point), the number of nearest modes is varied from 1 to 5 and misclassification percentage is plotted in figure 7. It was also observed that whenever the second nearest mode is classified as the correct mode, the first nearest mode remains the same corresponding to the second mode. For example when the second nearest mode is 19 the first nearest mode is 21, for 11 it is 10, for 6 it is 4 etc.,.

E. Misclassification with respect to time

The actual mode to which the data belong to and the estimated mode are compared with respect to time and graph is plotted. This is shown in figure 8. CLDA is the method used here to find the estimated mode. X- axis shows time in seconds and Y-axis shows the mode number (1 to 27). Further, the misclassification persist for utmost 2 seconds after a mode change.

F. Effect of a particular variable in classification

The relation between highest magnitude in the loading vector and dominant variables of classification remains to be

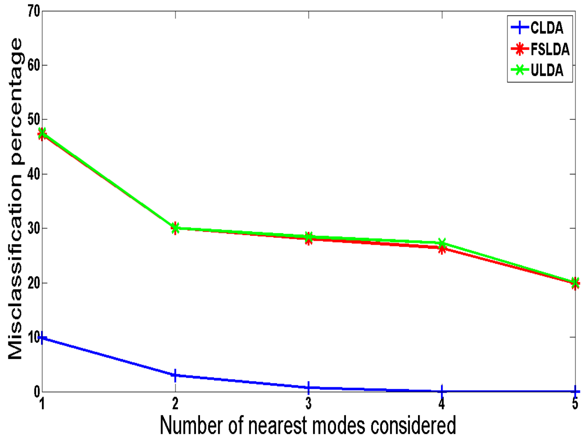

Fig. 7. Misclassification percentage versus number of nearest modes considered

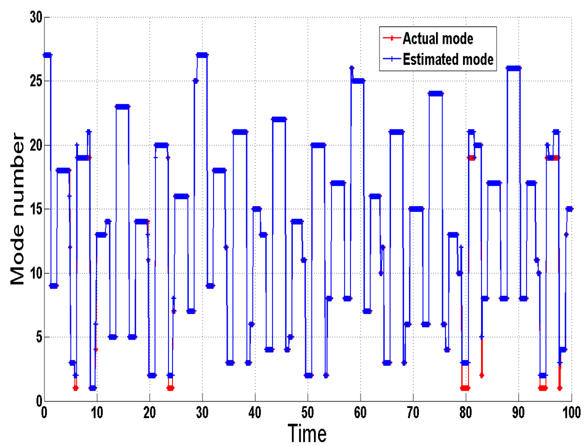

Fig. 8. Misclassification with respect to time

studied more thoroughly though some of the observations made are presented here. The top three loading vectors corresponding to top three eigenvalues of CLDA are

$$\begin{bmatrix} -16.91 & -55.31 & -14.34 \\ 4.99 & -4.47 & -9.19 \\ -7.22 & -6.29 & 0.15 \\ -207.19 & 0.04 & 0.03 \\ -0.09 & 35.90 & -3.65 \\ 0.06 & -6.97 & -29.80 \end{bmatrix} \tag{10}$$

The first loading vector (first column) corresponds to the largest eigenvalue and it gives more weight (Fourth row) to the value corresponding to double differentiation of pitch (Fourth column in data matrix fed to LDA algorithm). This value seems to dominate the classification in the direction of first loading vector. The second vector has relatively more weight for double differentiation of yaw and triple differentiation of pitch angle. Third vector gives more weight to double differentiation of roll angle. Hence it can be said that the double differentiated angles dominate the classification much more than the triple differentiated angles. Normalization of loading vectors does not affect the classification as the loading vectors gives only the direction [2]. RCS thrusters provide torque that changes pitch, yaw and roll angle and torque is proportional to double derivative of angle (angular acceleration). This is also evident from the loading vectors given in 10 as the double derivative of pitch, yaw and roll angles dominate the classification .

V. CONCLUSIONS

A. Conclusions

CLDA does not impose any constraint to find its loading vectors that maximizes the Fisher criteria. As observed from the previous section CLDA seems to have lower misclassification percentage, for this fault diagnosis of RLV thrusters application, compared to FSLDA and ULDA techniques (advanced variants of CLDA) with respect to the following criteria.

1) Percentage of training samples
2) Number of instances dropped after mode change
3) Number of loading vectors
4) Number of nearest modes considered

By knowing the commanded mode and the mode estimated from the classification it is possible to say which RCS thruster is faulty on assumption that there is no sensors are faulty.

VI. ACKNOWLEDGMENTS

We thank Mani Bhushan for providing reference materials. We also thank Kushal Chatterjee and Deepak Patil for providing the input data, used in these classification algorithms, from their simulation model.

REFERENCES

[1] X. Cao, J.Y. Yang, When Uncorrelated Linear Discriminant Analysis Are Combined with Wavelets, *Intelligent Computing in Signal Processing and Pattern Recognition*, 2006, Volume 345, 556-565.
[2] L.H. Chiang, E.L. Russell, R.D. Braatz, *Fault Detection and Diagnosis in Industrial Systems*, Springer- Verlag London Limited 2001.
[3] R.O.Duda, P.E.Hart, D.G.Stork, *Pattern Classification*, Second Edition, Wiley student edition 2006.
[4] R.A. Fisher, The use of multiple measurements in taxonomic problems, *Annals of Eugenics*, Volume 7, 1936, 178-188.
[5] D.H. Foley, J.W. Sammon, An optimal set of discriminant vectors, *IEEE Transactions on Computers*, 1975, Volume 24, Issue 3, 281289.
[6] M. Ge, R. Du, G. Zhang, Y.Xu, Fault diagnosis using support vector machine with an application in sheet metal stamping operations, *Mechanical Systems and Signal Processing*, 2004, Volume 18, Issue 1, 143-159.
[7] Z. Jin, J.Y.Yang, Z.S. Hu, Z. Lu, Face recognition based on uncorrelated discriminant transformation, *Pattern Recognition*, 2001, Volume 34, Issue 7, 1405-1416.
[8] G. Li, D. Liang, Q. Huang, S. Jiang, W. Gao, Object tracking using incremental 2D-LDA learning and Bayes inference, *15th IEEE International Conference on Image Processing*, 2008, 1567-1571.
[9] B. Samanta, K. R. Al-Balushi, Artificial neural network based fault diagnostics of rolling element bearings using time-domain features, *Mechanical Systems and Signal Processing*, 2003, Volume 17, Issue 2, 317-328.
[10] F. Song, S. Liu, J.Y Yang, Orthogonalized Fisher discriminant, *Pattern Recognition*, 2005, Volume 38, Issue 2, 311-313.
[11] J. Wang, Y, Xu, D. Zhang, J. You, An efficient method for computing orthogonal discriminant vectors, *Neurocomputing*, 2010, Volume 73, Issues 10-12, 2168-2176.
[12] Y. Xu, J.Y. Yang, Z. Jin, Theory analysis on FSLDA and ULDA, *Pattern Recognition*, 2003, Volume 36, Issue 12, 3031-3033.