Pattern Recognition Coursework Report

Praveen Tharmarajan (CID: 01517322), Shaanuka Gunaratne (CID: 01905941)

1. Section A - Data Preparation

Time Instance Selection Whilst initially examining the time series data provided, a prominent transient stage was discovered in the PVT data. According to D Erickson [1], in the initial stage of contact, the system exhibits significant levels of randomness and noise, giving rise to the observed transience. Therefore, a time instance of 740 was selected for two reasons: firstly, it is sufficiently far from the initial transience, at a point where the system can be considered to be in steady state, and secondly, it is also sufficiently far from any transience that can occur whilst releasing the object. Figure 1 and Figure 2 show the time series PVT data and time series electrode data for a number of objects on two separate trials. The corresponding section of the submitted .m file includes the more elaborate data visualisation that was undertaken in order to be meticulous.

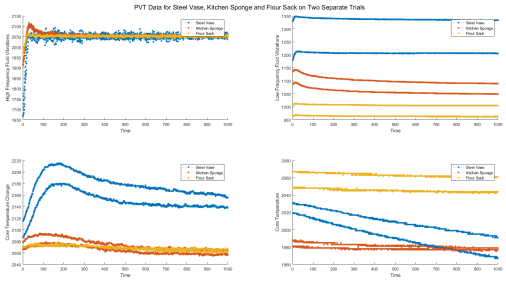

Figure 1. Time series PVT data collected from the steel vase, kitchen sponge and flour sack on trials 1 and 2.

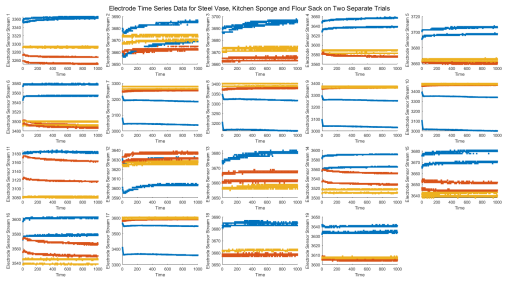

Figure 2. Time series electrode data collected from the steel vase, kitchen sponge and flour sack on trials 1 and 2.

Finger Selection: F0 The time series data was sampled at a time instance of 740 and appropriately saved in 'F0_PVT.mat' and 'F0_ElectrodeData.mat'.

3D Scatter Plot Figure 3 shows the complete contents of 'F0_PVT.mat' on a 3D scatter plot.

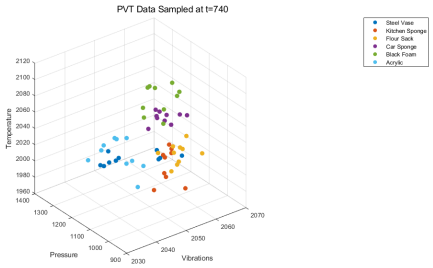

Figure 3. PVT data collected at t=740

2. Section B: Principal Component Analysis

Part 1 In order to remove the bias from different variable ranges, the PVT data was initially standardised. Then, Principal Component Analysis (PCA) was undertaken. The obtained covariance matrix, eigenvalues and eigenvectors are presented below:

$$S = \begin{bmatrix} 1.0000 & -0.1395 & -0.0017 \\ -0.1395 & 1.0000 & -0.5561 \\ -0.0017 & -0.5561 & 1.0000 \end{bmatrix} \quad (1)$$

$$eigenvalues = \begin{bmatrix} 0.4263 & 0 & 0 \\ 0 & 1.0008 & 0 \\ 0 & 0 & 1.5729 \end{bmatrix} \quad (2)$$

$$eigenvectors = \begin{bmatrix} -0.1739 & 0.9699 & 0.1702 \\ -0.7069 & -0.0026 & -0.7074 \\ -0.6856 & -0.2434 & 0.6861 \end{bmatrix} \quad (3)$$

Figure 4 shows the standardised PVT data with principal components displayed.

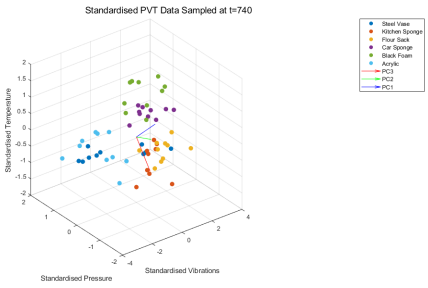

Figure 4. Standardised PVT data collected at t=740 with principal components displayed

Figure 5 shows the projection of the standardised PVT data onto the plane spanned by PC1 and PC2.

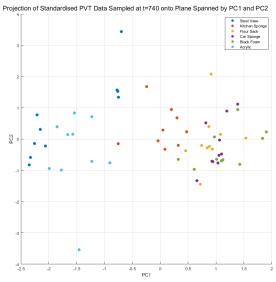

Figure 5. Projection of standardised PVT data collected at t=740 onto plane spanned by PC1 and PC2

Figure 6 shows the projection of the standardised PVT data onto PC1, PC2 and PC3.

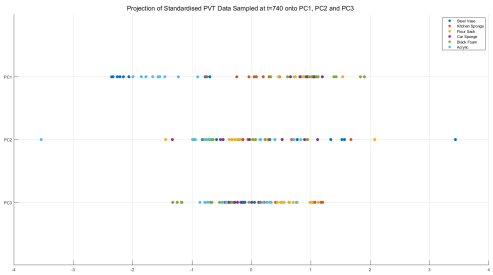

Figure 6. Projection of standardised PVT data collected at t=740 onto PC1, PC2 and PC3

As expected, there is greater potential for discrimination and clustering when the PVT data is projected onto the 2D plane spanned by PC1 and PC2, when compared to the 1D PCs themselves because a more complete feature vector is used. With regards to the 1D projections, as expected, the data along PC1 has the largest variance, followed by PC2 and finally PC3. For this reason, the data along PC1 has greater potential for discrimination and clustering, when compared to PC2 and PC3. Essentially, as the principal component number increases, the ability to group and discriminate deteriorates due to lower data variance.

Part 2 Figure 7 shows a scree plot corresponding to the standardised electrode data. It indicates that the first 3 principal components exhibit the largest variance and all other principal components are negligible in comparison. Therefore, the projection of this 19-dimensional data onto a 3D space spanned by PC1, PC2 and PC3 preserves the quality of the data for clustering and discrimination.

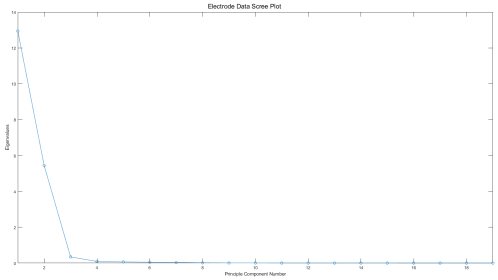

Figure 7. Scree plot corresponding to standardised electrode data

Figure 8 shows the projection of the electrode data onto a 3D space spanned by PC1, PC2 and PC3. We can observe that there is strong potential for discrimination and clustering in this 3D space, despite there being some overlap between some of the data points corresponding to the steel vase and acrylic. This overlap is likely due to common physical properties, such as the density of the two objects, which makes them hard to differentiate.

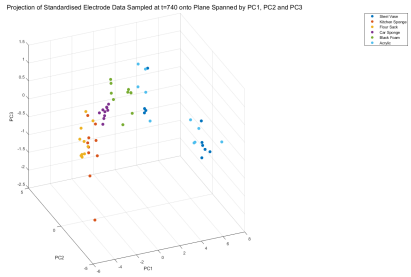

Figure 8. Projection of electrode data onto 3D space spanned by PC1, PC2 and PC3

Figure 15 shows the projection of the electrode data onto a 2D space spanned by PC1 and PC2, and Figure 16 shows the projection of the electrode data onto PC1, PC2 and PC3. Overall, we can see that there is the greatest potential for discrimination and clustering in the 3D space spanned by PC1, PC2 and PC3, followed by the 2D space spanned by PC1 and PC2, and finally, the 1D principal components themselves. This is because as the completeness of the feature vector diminishes, information pertaining to the data variance also diminishes. With regards to projections onto the 1D PC planes, the variance of the data has an inverse relationship with the principal component number so discrimination is more difficult at higher principal components.

3. Section C - Linear Discriminant Analysis

The following figures show a number of graphs, corresponding to the PVT sensor readings taken, whilst experimenting with a car sponge and black foam. Figure 9 shows a graph of the 2D PV data, along with the generated linear discriminant function and corresponding separation line. The same can be seen for the PT and TV data in Figures 10 and 11.

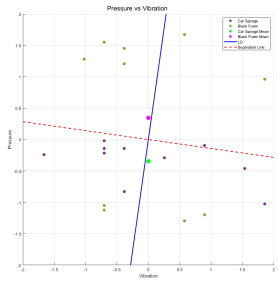

Figure 9. Graph of Pressure vs Vibrations

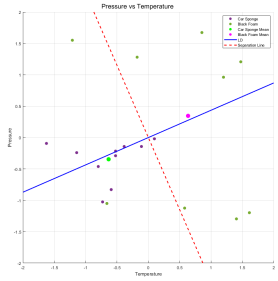

Figure 10. Graph of Pressure vs Temperature

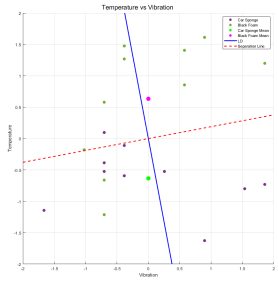

Figure 11. Graph of Temperature vs Vibrations

Figure 12 shows the PVT data, corresponding to the car sponge and black foam, highlighting the generated 2D linear discriminant hyperplane.

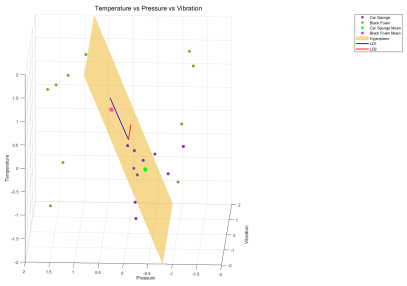

Figure 12. PVT scatter plot highlighting generated 2D linear discriminant hyperplane.

Figure 17, 18 and 19 in the Appendix shows the projection of these 2D and 3D data onto their respective linear discriminant planes. Table 1 shows how successfully LDA separates the various types of data. This shows that LDA has separated the 3D PVT data better than it has the 2D data, since 90% of the car sponge data points have fallen on one side of the separation line and 80% of the black foam data points have fallen on the complementary side of the separation line. The reason for an imperfect success rate is the presence of some aberrant sensor readings, which are likely attributed to the similar material composition of both objects, commonly polyurethane foam which is known for its lightweight, flexible and porous nature. This physical property likely contributed to overlap in the PVT sensor readings, resulting in imperfect class separation.

	Separation Success Rate (%)	
Type of Data	Car Sponge	Black Foam
2D PV Data	90	60
2D PT Data	90	70
2D TV Data	90	80
3D PVT Data	90	80

Table 1. How successfully LDA separates data

The same LDA analysis was conducted, using the PVT sensor readings taken whilst experimenting with a steel vase and acrylic. The generated graphs are in Figures 20, 21, 22, 23, 24, 25, 26 of the Appendix. We wanted to investigate whether data separation is possible when it comes to two objects with higher densities. The success rate of this analysis was essentially identical to the previous LDA analysis.

4. Section D - Clustering and Classification

4.1. Hierarchical Clustering

Euclidian Distance Having performed Hierarchical Clustering using the Euclidian distance metric to generate 6 clusters, the dendrogram in Figure 13 was obtained.

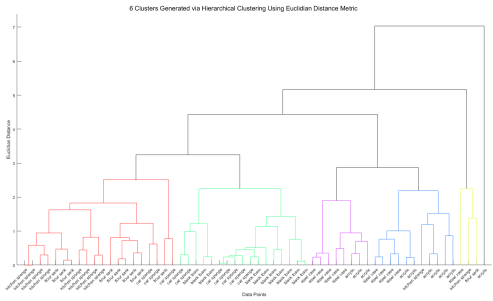

Figure 13. 6 clusters generated via hierarchical clustering using the Euclidian distance metric (dendrogram)

In the ideal scenario, each of the 6 clusters would only contain all the data points corresponding to a single object; however, this is not what is shown in the figure. The lime and black clusters have no similarities between its members, in terms of physical properties. The other four clusters do share similarities in terms of physical properties: the members of the purple and blue clusters have high density, the members of the green cluster are soft, flexible and porous, and finally, the members of the red cluster have similar external texture. The generated clusters can be visualised in Figure 27 of the Appendix.

Chebychev Distance Similarly, Figure 28 in the Appendix shows the 6 clusters, generated via Hierarchical Clustering using the Chebychev distance metric, as a dendrogram.

Using the Chebychev distance metric, a greater variance is observed in cluster size when compared to using the Euclidian distance metric. Also, there are more outliers when it comes to cluster members as there are fewer commonalities when it comes to physical properties. Therefore, the Euclidian distance metric is superior for clustering.

4.2. Bagging

30 bags This number of bags was large enough for accurate classification but not too large such that the data is overfitted. We characterised this quantity using the '*ooberror()* function and produced Figure 30. After 30 bags, there were diminishing returns; in fact, the error increased again, suggesting overfitting.

Figures 31 and 32 in the Appendix show the first and final decision trees, respectively.

Figure 14 shows the generated confusion matrix when the trained model was run on the test data. The accuracy rate is 79.2%. The model struggles to classify the denser objects, but performs very well with less dense objects. The initial PCA step was especially invaluable for electrode data analysis because it allowed us to view the 19-dimensional data in 3D, whilst preserving data variance, enabling classification and discrimination.

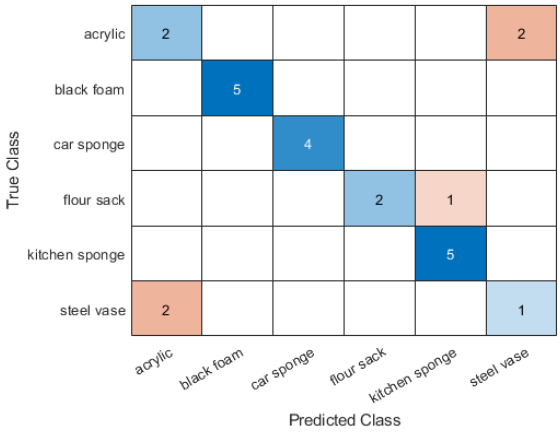

Figure 14. Confusion matrix

5. Conclusion

Summary PCA provided a means for dimensionality reduction, whilst maximising data variance, which increased the potential for clustering. Class separation was maximized for class discrimination through the application of LDA. Hierarchical Clustering allowed the data points to be grouped, based on a distance metric. Finally, bagging allowed the model to be trained so that it could classify new data based on past learning experiences.

Distinguish Clustering similar objects perfectly is possible using touch alone, but determining a specific object among similar objects, with 100% accuracy, using touch alone, isn't possible.

Critical Property With regards to the electrode channels, it was noted during PCA that only 3 of the 19 principal components had non-negligible eigenvalues, meaning the data variance was largest in those three directions. Therefore, only those channels contributing significantly to those principal components should be kept. With regards to the PVT sensors, as you can see in Figure 4, temperature plays a dominant role in all of the principal components so it would be logical to say it is the most important, but the temperature time series data indicates that the temperature readings aren't at steady state for all objects so that conclusion cannot be made. The next most dominant feature is pressure, so it is advisable to have a pressure sensor over a temperature or vibration sensor.

Alternative Methods Instead of sampling a single time instance, we can sample at a certain low frequency to allow features of higher mathematical order to be analysed and evaluated. However, due to the increased complexity, the processing speed would decrease significantly because of the so-called "curse of dimensionality".

References

[1] D. Erickson, M. Weber, and I. Sharf. Contact stiffness and damping estimation for robotic systems. *The International Journal of Robotics Research*, 22(1):41–57, 2003.

6. Appendix

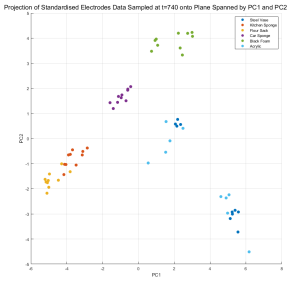

Figure 15. Projection of electrode data onto 2D space spanned by PC1 and PC2

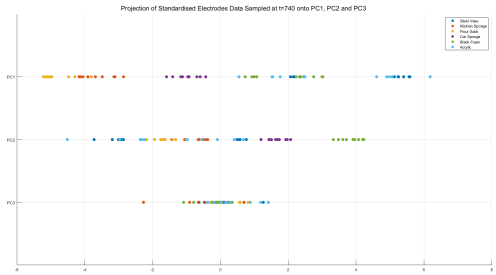

Figure 16. Projection of electrode data onto PC1, PC2 and PC3

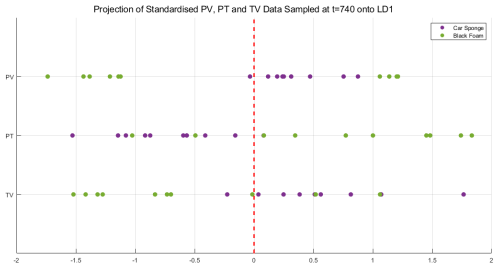

Figure 17. Projection of PV, PT, and TV data onto their respective linear discriminant

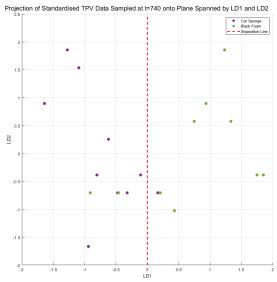

Figure 18. Projection of PVT data onto 2D linear discriminant hyperplane spanned by LD1 and LD2

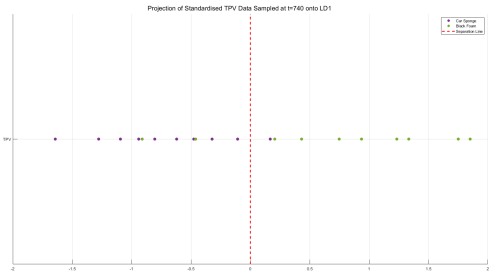

Figure 19. Projection of the PVT data onto 1D linear discriminant, LD1

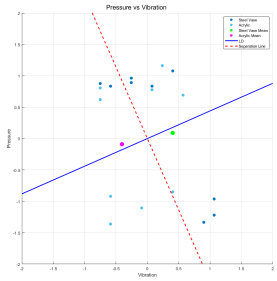

Figure 20. Graph of Pressure vs Vibrations

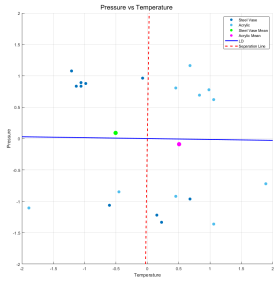

Figure 21. Graph of Pressure vs Temperature

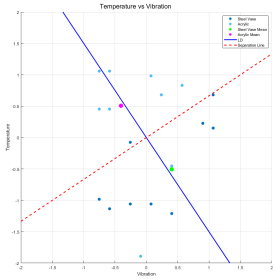

Figure 22. Graph of Temperature vs Vibrations

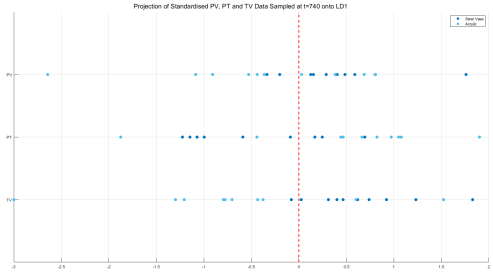

Figure 23. Projection of PV, PT, and TV data onto their respective linear discriminant

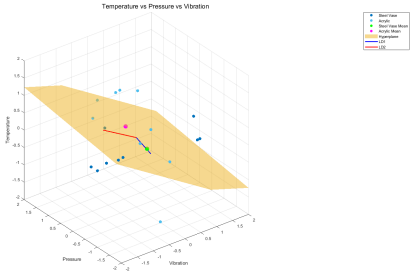

Figure 24. PVT scatter plot highlighting generated 2D linear discriminant hyperplane.

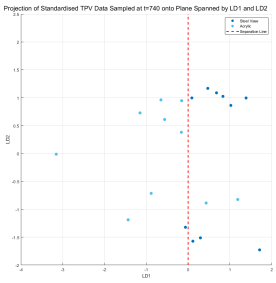

Figure 25. Projection of PVT data onto 2D linear discriminant hyperplane spanned by LD1 and LD2

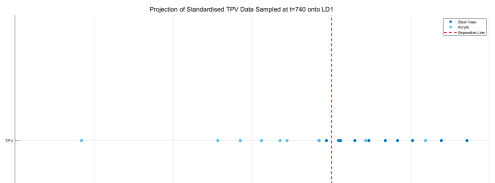

Figure 26. Projection of the PVT data onto 1D linear discriminant, LD1

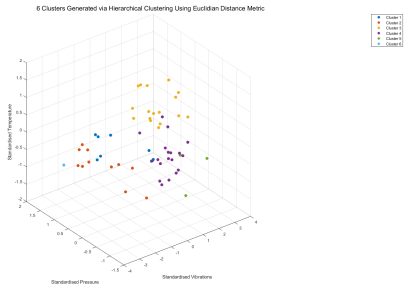

Figure 27. 6 clusters generated via hierarchical clustering using Euclidian distance (scatter plot)

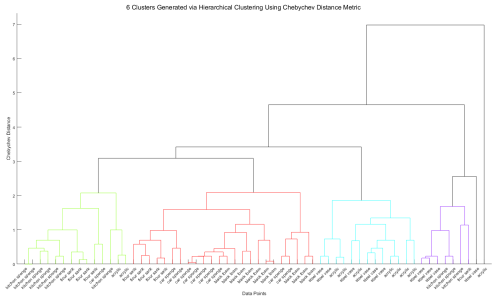

Figure 28. 6 clusters generated via hierarchical clustering using the Chebychev distance metric (dendrogram)

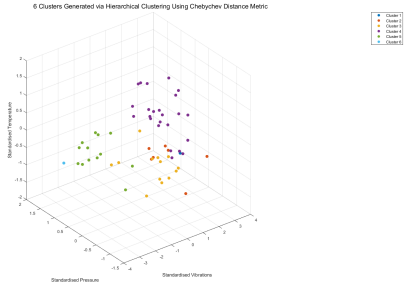

Figure 29. 6 clusters generated via hierarchical clustering using Chebychev distance (scatter plot)

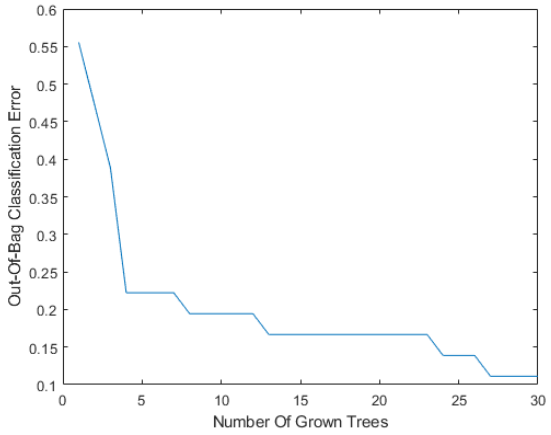

Figure 30. Out-of-Bag error plot

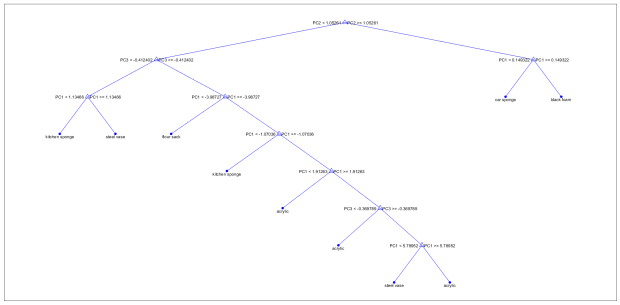

Figure 31. First decision tree

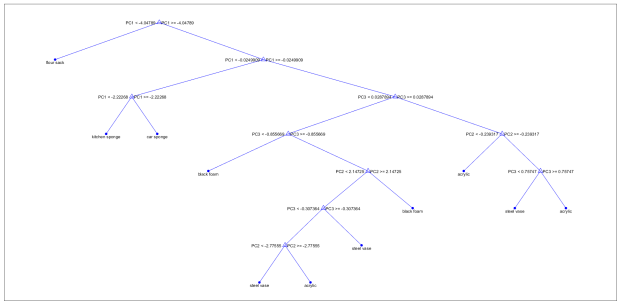

Figure 32. Final decision tree