

Figure 4 generated as in the Lewis paper (using raw data with 1151 features):

- the data is clustered to 3 clusters
- the mean is found for each cluster and plotted

Figure 2 generated using fastmapped data of 4 features. The values of each feature are plotted.

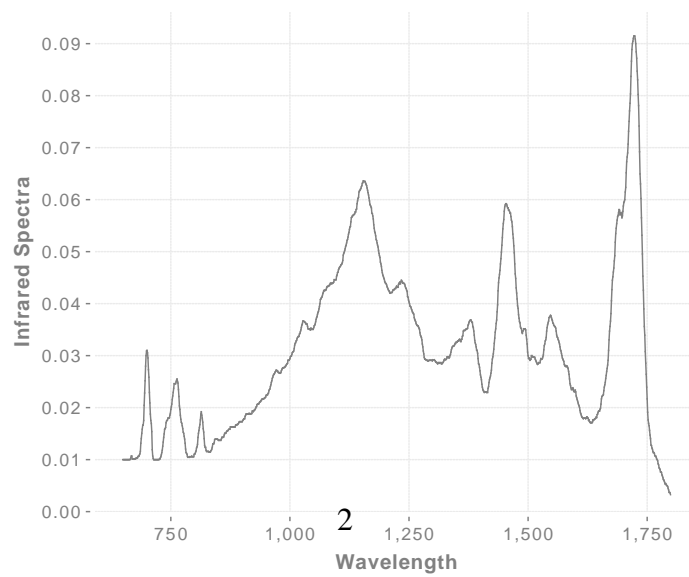
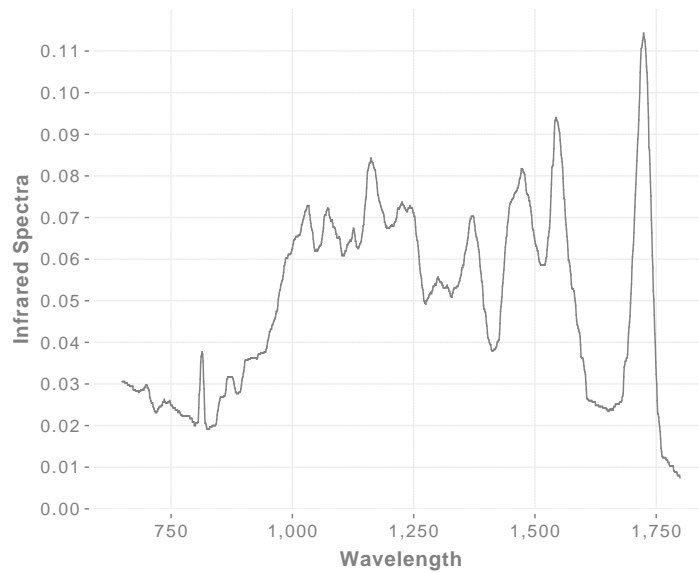
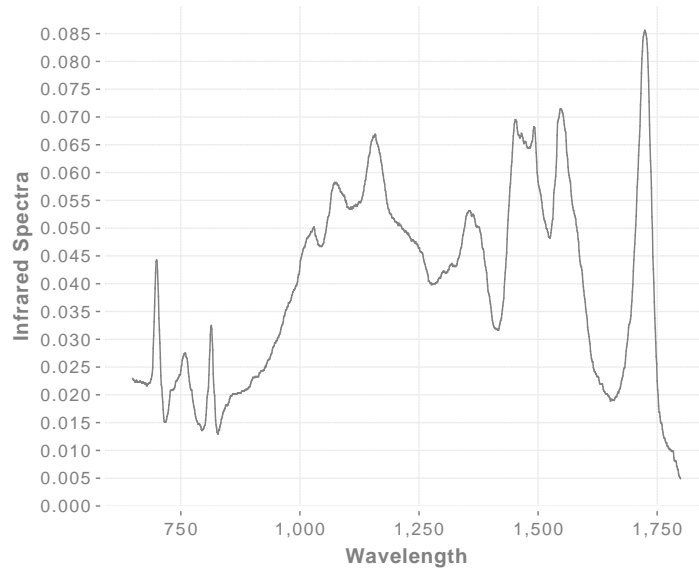


Figure 1: Infrared spectra of the mean of classes produced by kmeans clustering on 185 infrared spectra from 37 vehicles.

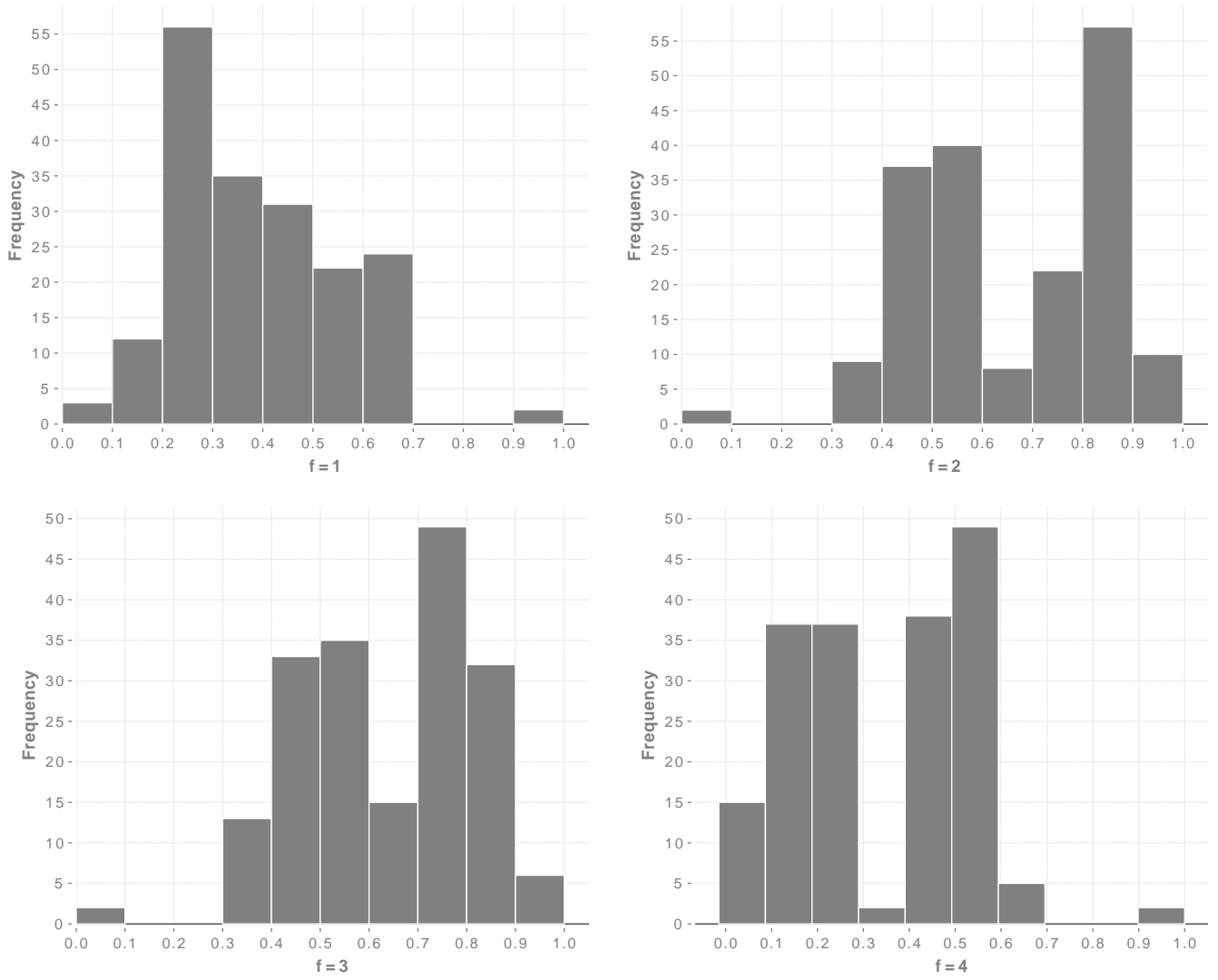


Figure 2: FastMapped infrared spectra for 4 features.

Clusters	n	Final Instances	Reduction%
bc (286, 10, 2)	n=1	272	95
	n=2	272	95
	n=4	200	70
	n=7	58	20
	n=8	8	3
	n=16	0	0
heart (297, 14, 5)	n=1	249	84
	n=2	200	67
	n=4	176	59
	n=8	64	22
	n=9	35	12
	n=16	0	0
lym (148, 19, 4)	n=1	120	81
	n=2	118	80
	n=4	115	78
	n=8	115	78
	n=16	32	22
	n=17	20	14
pima (768, 9, 2)	n=1	375	49
	n=2	95	12
	n=4	13	2
	n=8	0	0
	n=16	0	0
tumor (339, 18, 21)	n=1	270	80
	n=2	255	75
	n=4	255	75
	n=8	255	75
	n=16	193	57
	n=18	126	37

Figure 3: Instance selection using the CLIFF selector. The Reduction% column shows the percentage of the original data set left after selection.

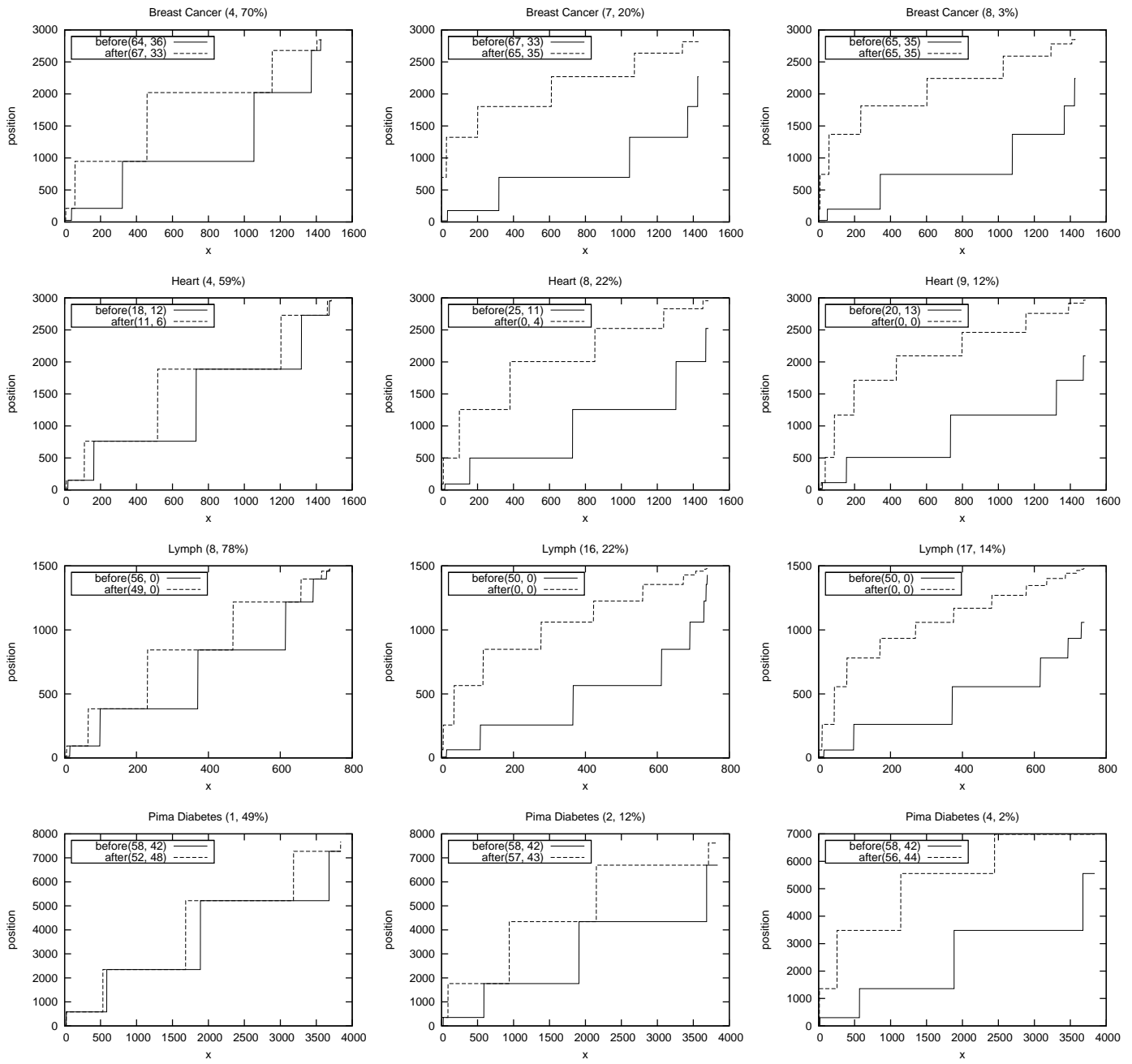


Figure 4: Breast Cancer, Heart, Lymph, Pima.