Application Note 10

Rapid Identification of Contaminated Water Bottles Returned for Refilling

By Steve Hobbs **October 2000**

1. Introduction

Five-gallon polycarbonate water bottles are returned to water bottling plants for refilling. During use by the consumer, the water bottles can be exposed to a wide range of contaminates. For example, consumers sometimes use the bottles as containers for fuels or liquid chemical solutions. The customer would like to identify water bottles that have been contaminated before the wash stage in their refilling process. When water bottles are not identified before the wash stage, the washing system and a large number of bottles can become contaminated. Water that is shipped in contaminated bottles can prompt customer complaints. The present system for rejecting contaminated water bottles is by a human nose. For sniffing bottles each person is limited to a 20-30 minute shift before being replaced by another.

2. Experimental

Sample preparation:
One bad and one good bottle were used for samples. A new bottle containing about 50 mL water was used as a reference. The “bad” bottle contained one drop of gasoline and approximately 50 mL water. The good bottle contained approximately 50 mL of water only.

A four inch, 14 gauge, blunt, stainless steel needle was used for sampling. A Teflon-lined lid with a hole in it was secured to the hub of the needle. The lid was used to cap the mouth of the bottles while sampling. Between sampling, the bottles were not capped. A second, 1 mm hole in the lid provided a vent to the bottle. A new bottle was used as a reference by employing a 20” sampling tube fed through a Teflon-lined cap. The tube was connected to the purge inlet of the Cyranose 320 with Teflon-lined tubing.
Testing Conditions:
A Cyranose 320 with a 32-sensor array was used to test the bottles. Method settings are shown in Table 1. Note that the total sampling time was only 8 seconds and digital filtering was off. The training set was obtained by sampling the bottles in the order shown in Table 2. To simulate the rapid sampling that will occur on a process line, *.csv data for the training and prediction sets were collected in identify mode and a repeat count of 40.
Data handling:
The sensor responses were calculated and $(R_{max}-R_{min})/R_{min}$ where R_{min} is the minimum of the resistance reading during the baseline purge and R_{max} is the maximum resistance

reading during the vapor exposure. The 32 sensor responses were then autoscaled and normalized. The KNN algorithm was used for prediction.

3. Results

The PCA plot for the training set is shown in Figure 1. The two classes are clearly separated in PCA space. Cross-validation using KNN and autoscaling correctly predicted all training set exposures and all of the prediction set as shown in Table 3.

4. Conclusions

Using a good bottle as a reference the Cyranose 320 correctly predicted clean and gasoline-contaminated bottles within an 8 second cycle time. The Cyranose 320 shows considerable promise for rapid process control monitoring. Additional bad bottles or a new algorithm could be required to accurately define the “bad” class in real-world conditions.

Table 1. Method setting used in the experiments.

Method name	FastBottle	
Class 1	Good	
Class 2	Bad	
Baseline purge	3s	High
Sample draw	0s	High
Sample draw 2	3s	High
Snout removal	0s	
1st sample gas purge	0s	High
1st air intake purge	1s	High
2nd sample gas purge	1s	High
2nd air intake purge	0s	High
Digital filtering	Off	
Substrate heater	On	42
Training repeat count	1	
Identifying repeat count	1	
Active sensors	All 32	
Algorithm	KNN	
Preprocessing	Autoscaling	
Normalization	Normalization 1	
Minimum confidence level		1

Table 2. Sampling sequence used in the training set and prediction set. The first 10 of each class were used for training and the remaining were used for predictions.

Order	Class	Order	Class	Order	Class	Order	Class
1	1	11	2	21	1	31	1
2	1	12	2	22	2	32	2
3	2	13	2	23	2	33	1
4	1	14	1	24	2	34	1
5	1	15	1	25	1	35	1
6	2	16	2	26	1	36	2
7	2	17	1	27	2	37	1
8	2	18	2	28	1	38	2
9	2	19	2	29	2	39	1
10	1	20	1	30	1	40	2

Table 3. All samples were correctly identified with an excellent match (***** grade). CYRANOSE 320 was used with KNN, Autoscaling, PCA, and a medium Identification Quality.

Sample	Result	Grade	Correct?
Sample 1	Class 2	(*****)	YES
Sample 2	Class 2	(*****)	YES
Sample 3	Class 2	(*****)	YES
Sample 4	Class 2	(*****)	YES
Sample 5	Class 1	(*****)	YES
Sample 6	Class 1	(*****)	YES
Sample 7	Class 2	(*****)	YES
Sample 8	Class 1	(*****)	YES
Sample 9	Class 2	(*****)	YES
Sample 10	Class 1	(*****)	YES
Sample 11	Class 1	(*****)	YES
Sample 12	Class 2	(*****)	YES
Sample 13	Class 1	(*****)	YES
Sample 14	Class 1	(*****)	YES
Sample 15	Class 1	(*****)	YES
Sample 16	Class 2	(*****)	YES
Sample 17	Class 1	(*****)	YES
Sample 18	Class 2	(*****)	YES
Sample 19	Class 1	(*****)	YES
Sample 20	Class 2	(*****)	YES

Figure 1 Plot of Principal Component Analysis with Autoscaling and Normalization 1. All sensors were selected.

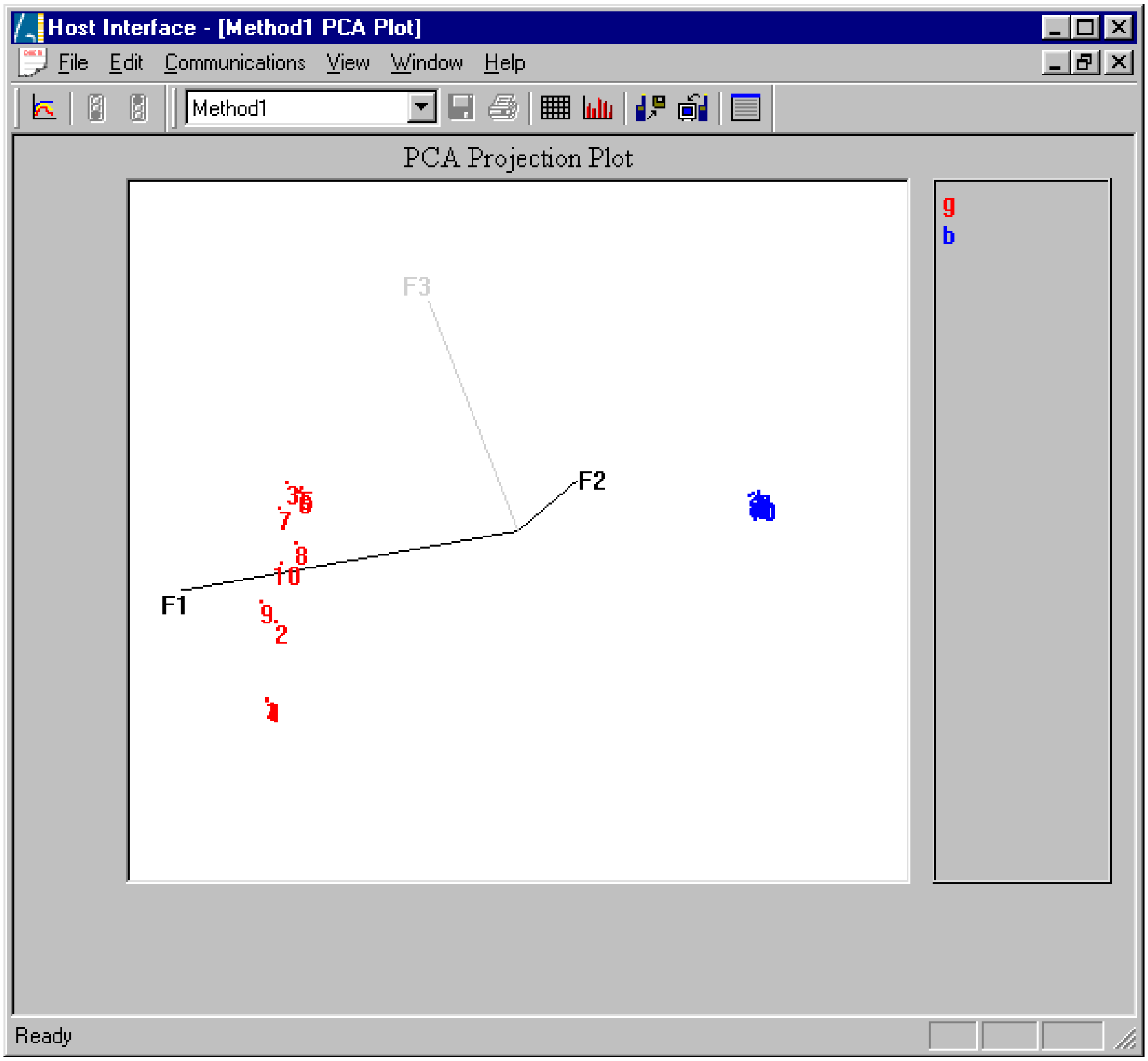

Figure 2. Canonical projection plot of training set. Interclass M-distance was 77 and cross-validation was 100% correct.

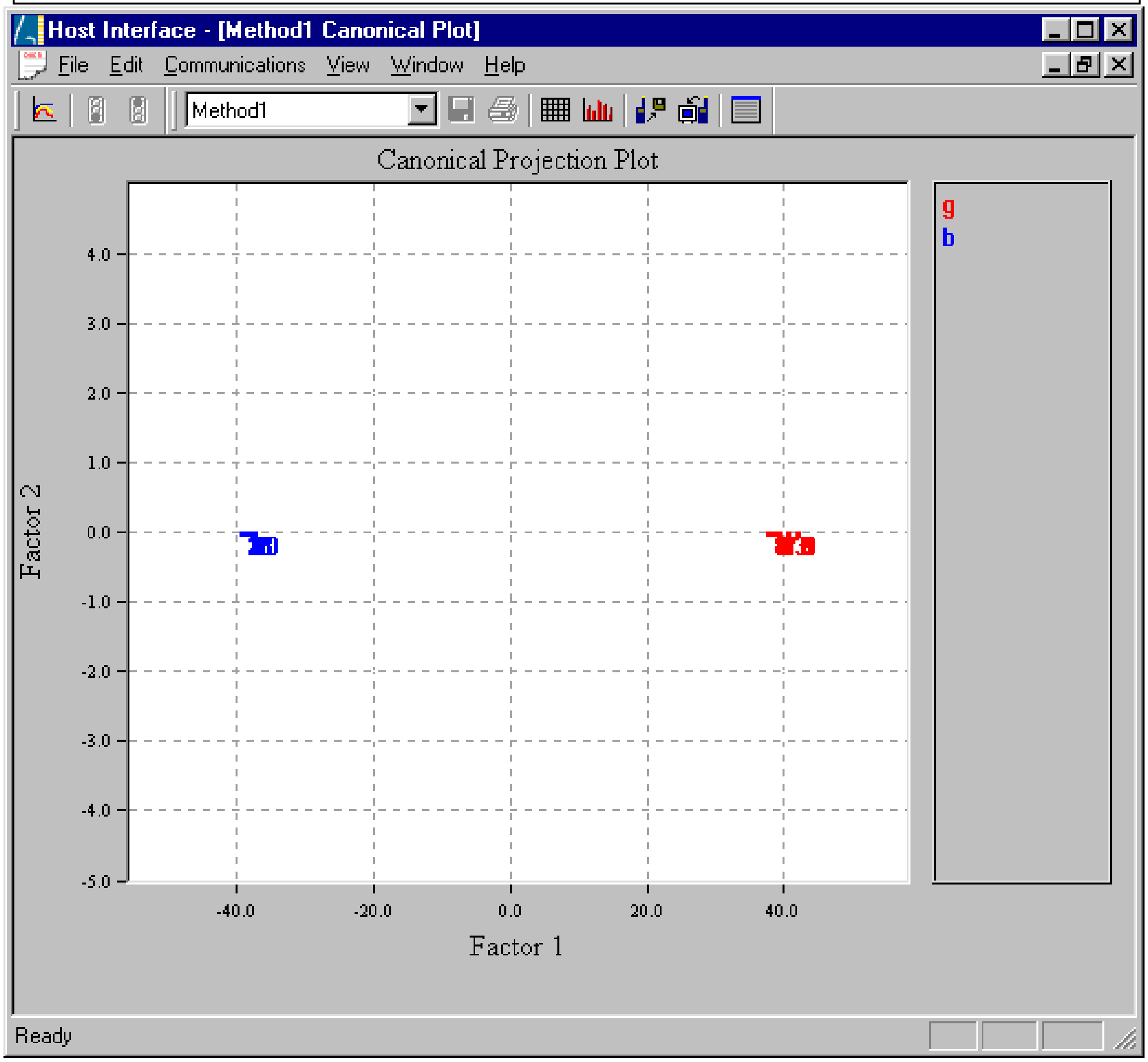

Figure 3. Strip chart plot of raw data. Total cycle time was 8 seconds. Experimental order is in Table 2.