Smelling in Chemically Complex Environments: An Optofluidic Bragg Fiber Array for Differentiation of Methanol Adulterated Beverages

Adem Yildirim,[1,2,†] *Fahri Emre Ozturk,*[1,2,†] *and Mehmet Bayindir*[1,2,3]

[1]UNAM-National Nanotechnology Research Center, [2]Institute of Materials Science and Nanotechnology, [3]Department of Physics, Bilkent University, 06800 Ankara, Turkey

[†]These authors contributed equally to this work

We have provided the details of HCA analysis, scheme of the experimental set up, figures showing the response of fiber 2 to ternary mixtures and dendrogram plots of the response of each single fiber, and tables containing I/I_0 values of all measurements.

S1. Hierarchical cluster analysis

Hierarchical cluster analysis techniques are used for classifying data under groups with similar properties, i.e. clusters. The dendrogram represents formed clusters and their similarities (the proximity of the clusters) graphically.

The main parameters of a HCA technique are the linkage method which determines how the clusters are formed and the distance metric which is used to calculate the distance between data points. In this study we used 'squared euclidean distance metric' for all dendrograms and 'single-link linkage method' for the dendrogram at Figure 3c , 'Ward's linkage method' for all the remaining dendrograms. These are so called agglomerative methods, in which each data point is registered as a cluster initially and two of the closest clusters are merged at every step. Basic steps utilized by built in HCA algorithms of the commercial software used are provided below.[1,2]

Algorithm 1. Basic clustering algorithm with squared euclidean distance metric and single-link linkage method.

1: Calculate the distance matrix, i.e. the matrix that gives the squared euclidean distances between each dual data points P_n , P_m.

$$P_n(x_n, y_n, z_n)\ , P_m(x_m, y_m, z_m)$$

$$d_{nm} = (x_n - x_m)^2 + (y_n - y_m)^2 + (z_n - z_m)^2$$

2: Merge the two closest data points as a cluster.

3: Register each single data point as a cluster. Update the distance matrix by calculating the distances between the new cluster and original clusters. Calculate the inter clusteral distances as d_{nm} where P_n and P_m are the closest points from the two different clusters C_1 and C_2.

$$d_{nm} = min\{dist(P_n, P_m) :\ P_n \in C_1\ , P_m \in C_2\}$$

4: Merge the two closest clusters.

5: If the number of remaining clusters is larger than 1; go to step 3.

6: End.

Algorithm 2. Basic clustering algorithm with squared euclidean distance metric and Ward's linkage method.

1: Register each data point as a cluster.

2: Repeat steps 3-5 for each possible set of clusters which is to be formed by merging two of the original clusters.

3: Calculate a mean 'P_M' for each cluster:

$$P_M(x_M, y_M, z_M)$$

$$x_M = \frac{\sum_{i=1}^{N} x_i}{N}, y_M = \frac{\sum_{i=1}^{N} y_i}{N}, z_M = \frac{\sum_{i=1}^{N} z_i}{N}$$

where 'N' is the number of data points in the cluster.

4: Calculate the squared euclidean distances between each cluster mean 'P_M' and the data points of that cluster 'd_{iM}', where subscript 'i' denotes the distance for each data point.

$$d_{iM} = (x_i - x_M)^2 + (y_i - y_M)^2 + (z_i - z_M)^2$$

5: Add the calculated squared euclidean distances for each cluster:

$$D = \sum_{i=1}^{N} d_{iM}$$

where 'N' is the number of data points in the cluster.

6: Add the obtained 'D' values to get the sum-of-squares index 'E':

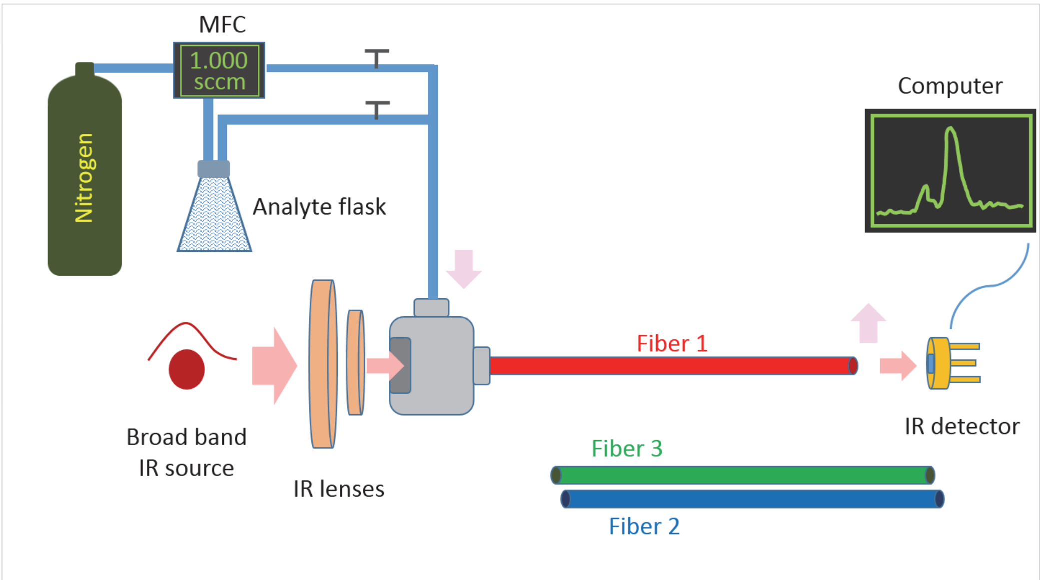
$$E = \sum_{i=1}^{K} D_i$$

where 'K' is the number of clusters in the set.

7: Merge the two clusters that form the cluster set which provides minimum value of 'E'.

8: If the number of remaining clusters is larger than 1; go to step 2.

9: End.



Scheme S1. Optoelectronic nose concept measurement system. Analyte vapors were introduced into fibers with nitrogen as carrier gas using a mass flow controller (MFC). IR light from broad band source of an FTIR was coupled to the fiber. Measurements were taken from each fiber by rounds. IR detector of the FTIR (uncooled DLaTGS) was utilized for measuring total transmitted intensities of the fibers.

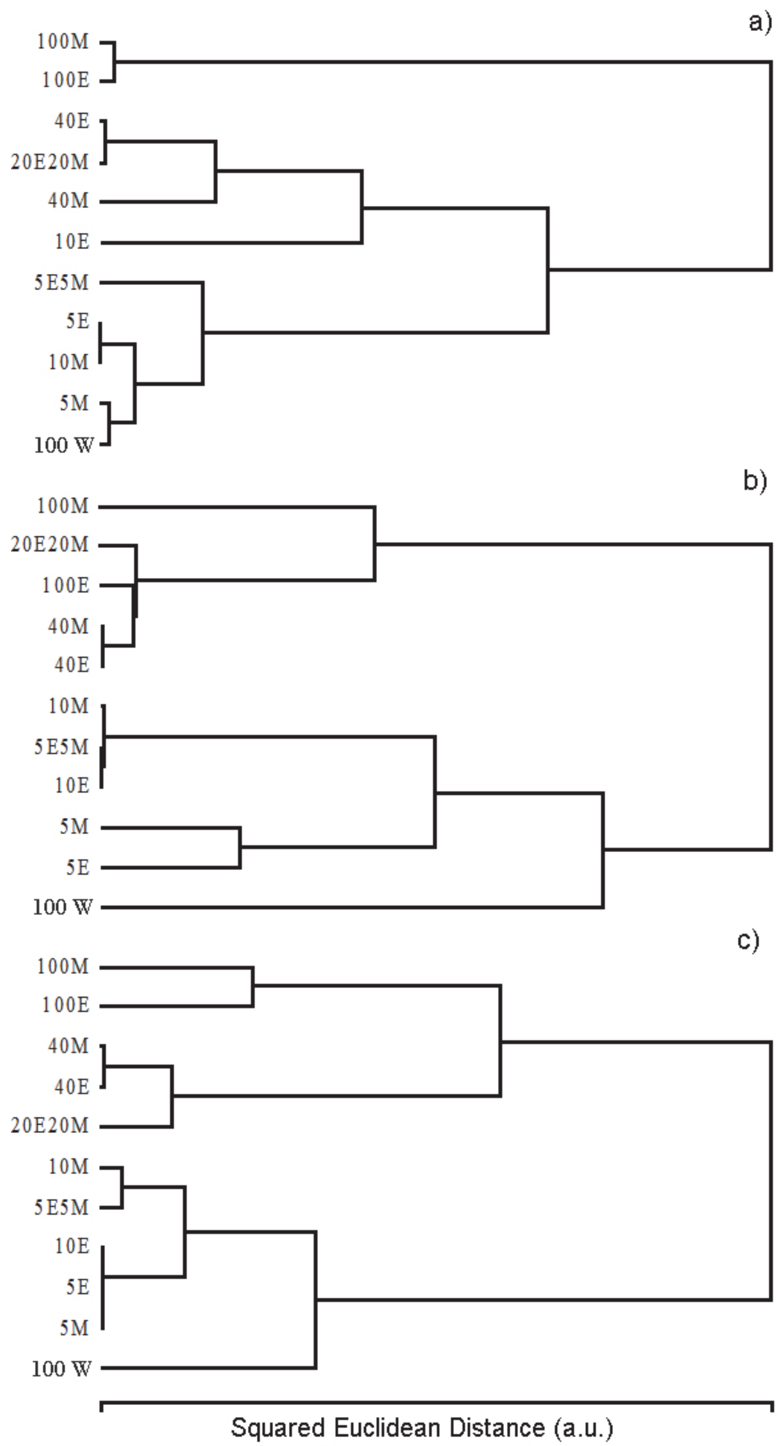


Figure S1. Hierarchical cluster analysis of alcohol mixture measurements for each fiber. a) HCA of sensor response of fiber1, b) fiber 2 and c) fiber 3. Contrary to the analysis made with using all three fibers (shown in Figure 3c of the manuscript), a meaningful, reliable differentiation is not possible with using only one fiber.

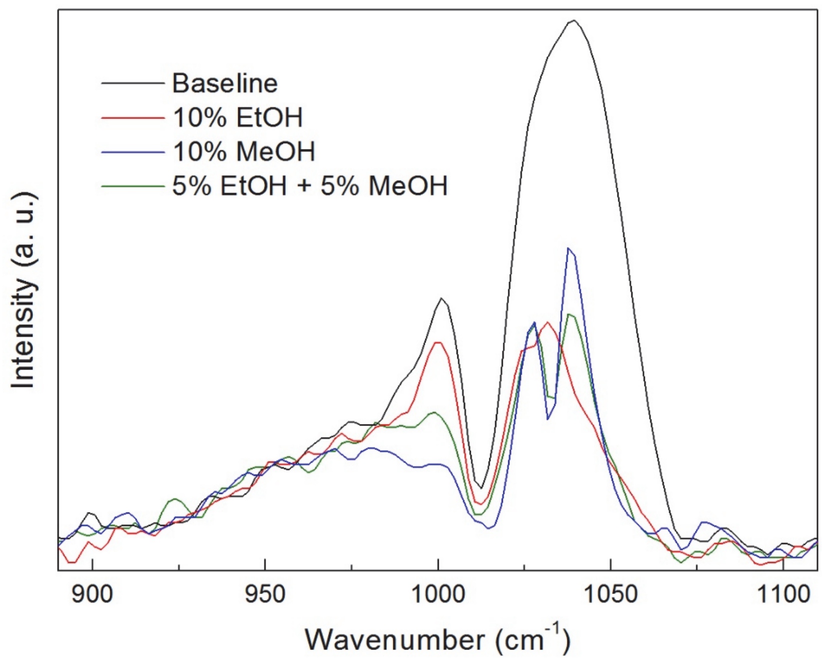


Figure S2. Response of fiber 2 to mixtures with same total alcohol concentrations but varying in percentages of ethanol and methanol. 5% Ethanol + 5% methanol ternary solution quenches the transmission of the fiber in a way which can be predicted as a combination of responses of binary mixtures.

Table S1. (I/I_0) between total transmitted intensity before (I_0) and after (I) interacting with the analyte for ethanol (EtOH) methanol (MeOH) mixtures for concentrations varying in the range of 5% to 40%.

Analyte	**(I/Io) x 100 (Fiber 1)**	**(I/Io) x 100 (Fiber 2)**	**(I/Io) x 100 (Fiber 3)**
Water	97.64	97.87	98.63
EtOH 100%	77.24	59.33	73.75
EtOH 40%	82.81	62.66	79.94
EtOH 10%	88.99	74.98	93.58
EtOH 5%	95.71	88.10	94.03
MeOH 100%	76.44	49.5	70.25
MeOH 40%	85.51	61.84	79.41
MeOH 10%	95.81	73.94	89.67
MeOH 5%	97.08	82.95	94.47
EtOH+MeOH 20%	83.19	56.72	82.34
EtOH+MeOH 5%	93.54	74.76	90.99

Table S2. Averaged I/I_0 values and standard errors of several alcohol mixtures calculated from at least three seperate measurements.

Analyte	**Average (I/Io) x 100 (Fiber 1)**	**Standard error (Fiber 1)**	**Average (I/Io) x 100 (Fiber 2)**	**Standard error (Fiber 2)**	**Average (I/Io) x 100 (Fiber 3)**	**Standard error (Fiber 3)**
Water	98.92	1.01	99.65	1.74	99.84	1.12
EtOH 10%	n/a	n/a	77.09	1.51	n/a	n/a
EtOH 5%	95.39	0.78	88.42	0.52	94.98	0.76
MeOH 10%	n/a	n/a	75.09	0.98	n/a	n/a
MeOH 5%	96.8	0.22	82.97	0.46	94.92	1.18

Table S3. (I/I_0) between total transmitted intensity before (I_0) and after (I) interacting with the analyte for repeated measurements of 5% ethanol (EtOH), 5% methanol (MeOH) solutions and water.

Analyte	(I/Io) x 100 (Fiber 1)	(I/Io) x 100 (Fiber 2)	(I/Io) x 100 (Fiber 3)
EtOH 5% #1	95.71	88.10	94.03
EtOH 5% #2	95.07	88.78	95.81
EtOH 5% #3	95.96	87.84	94.69
EtOH 5% #4	96.02	88.24	94.64
EtOH 5% #5	94.17	89.12	95.72
MeOH 5% #1	97.08	82.95	94.47
MeOH 5% #2	96.95	82.80	93.68
MeOH 5% #3	96.80	82.42	96.72
MeOH 5% #4	96.58	83.69	95.40
MeOH 5% #5	96.58	83.01	94.31
Water #1	97.64	97.87	98.63
Water #2	100.25	99.14	100.95
Water #3	98.97	100.70	100.71
Water #4	99.45	98.40	100.24
Water #5	98.29	102.12	98.65

Table S4. (I/I_0) between total transmitted intensity before (I_0) and after (I) interacting with the analyte for repeated measurements of 5% ethanol (EtOH), 5% methanol (MeOH) adulterated beer mixtures and unalcoholic beer.

Analyte	(I/Io) x 100 (Fiber 1)	(I/Io) x 100 (Fiber 2)	(I/Io) x 100 (Fiber 3)
Beer #1	99.20	98.50	100.64
Beer #2	98.32	97.93	98.45
Beer #3	98.22	96.56	100.05
Beer #4	98.48	97.38	98.99
Beer #5	98.72	97.94	99.97
EtOH 5% #1	96.54	85.33	95.05
EtOH 5% #2	95.51	85.35	94.31
EtOH 5% #3	96.14	86.50	96.65
EtOH 5% #4	96.90	86.42	95.25
EtOH 5% #5	96.24	84.22	95.22
MeOH 5% #1	93.72	81.38	95.06
MeOH 5% #2	94.73	81.45	96.52
MeOH 5% #3	95.10	81.09	94.80
MeOH 5% #4	94.36	79.61	91.96
MeOH 5% #5	96.01	80.04	92.40

Table S5. (I/I_0) between total transmitted intensity before (I_0) and after (I) interacting with the analyte for repeated measurements of 5% ethanol (EtOH), 5% methanol (MeOH) adulterated mixed fruit juice and unadulterated mixed fruit juice (FJ).

Analyte	(I/Io) x 100 (Fiber 1)	(I/Io) x 100 (Fiber 2)	(I/Io) x 100 (Fiber 3)
Fruit Juice #1	99.73	100.66	99.88
Fruit Juice #2	99.05	97.45	101.34
Fruit Juice #3	99.36	99.77	101.83
Fruit Juice #4	100.18	101.29	102.34
Fruit Juice #5	99.75	100.92	100.00
EtOH 5% #1	98.40	87.01	96.00
EtOH 5% #2	97.20	87.24	94.28
EtOH 5% #3	96.08	87.61	94.32
EtOH 5% #4	95.44	89.25	95.67
EtOH 5% #5	95.16	88.85	96.82
MetOH 5% #1	96.74	83.99	94.07
MetOH 5% #2	97.24	83.82	96.52
MetOH 5% #3	98.16	82.84	94.39
MetOH 5% #4	96.08	82.22	95.98
MetOH 5% #5	97.38	83.55	94.12

References

(1) Tan, P.N.; Steinbach, M.; Kumar, V. Introduction to Data Mining; Addison-Wesley, 2006.

(2) Romesburg, C.H.; Cluster Analysis for Researchers; Lulu Press: North Carolina, 2004.