Available online at www.sciencedirect.com
SCIENCE DIRECT®
Sensors and Actuators B xxx (2003) xxx–xxx
SENSORS AND ACTUATORS B CHEMICAL
www.elsevier.com/locate/snb

zNose™ technology for the classification of honey based on rapid aroma profiling

Jeroen Lammertyn [a,b], Els A. Veraverbeke [a,b], Joseph Irudayaraj [a,*]

[a] *Department of Agricultural and Biological Engineering, 227 Agricultural Engineering Building, The Pennsylvania State University, University Park, PA 16802, USA*
[b] *Laboratory of Postharvest Technology, Katholieke Universiteit Leuven, W. de Croylaan 42, Leuven B-3001, Belgium*

Received 29 August 2003; received in revised form 29 August 2003; accepted 11 September 2003

Abstract

Rapid aroma profiling of food products is a first step towards at-line flavor quality control and off-flavor assessment. In this paper, the potential of the zNose™ was tested for the first time to address this application. Honey was chosen as the food product because of its characteristic aroma. Both a chromatogram and a spectral approach to the interpretation of the zNose™ signal were established. In the chromatogram approach, the signal was treated as a traditional chromatogram and relative peak areas were calculated and compared, while the whole aroma spectrum was considered in the spectral approach. Shifts in GC-column retention times initially led to misinterpretation of the results in the spectral approach. A data processing algorithm was, hence, developed to correct for these shifts. Data were analyzed with principal component analysis (PCA), and canonical discriminant analysis (CDA). With both relative peak areas and corrected spectra, the aroma of six different honey varieties and two types of sugar solutions were successfully discriminated. A classification model was built and validated externally, which resulted in a correct classification of 15 out of 16 honey aroma profiles (94%).

Keywords: zNose™; Aroma profiling; Honey; Classification; PCA; CDA

1. Introduction

Food quality is a complex concept referring to multiple characteristics that make a food product acceptable or more desirable to eat. Important food quality aspects are safety, nutritional value, functionality, and aesthetics (color, texture, flavor, appearance). While the first three are subjectively quantifiable, the last has an even more important subjective component, which makes it more difficult to describe and/or quantify. Flavor, as the combination of aroma and taste, is a very important component of this subjective quality [1].

Traditional analytical and quantitative techniques for flavor analysis include HPLC, GC with headspace sampling and GC-MS analysis with solid phase microextraction [2–5]. Numerous reports exist on the flavor analysis of a wide range of food products with these techniques and they have proven to give very precise and reliable results. These techniques, however, involve a lot of sample preparation, are time consuming and can only be carried out in a specially equipped laboratory environment by well trained operators. Next to a chemical characterization, flavor analysis often also includes a sensory evaluation by both trained taste panels and consumer panels [4,6]. This type of evaluation is important in classifying flavor characteristics according to human perception and consumer behavior. Evidently, this is a very subjective and variable evaluation, which involves a very costly and time consuming procedure.

New techniques that allow a faster, objective flavor characterization without the need for special equipment or skills offer value to industries attempting an on-line or at-line flavor evaluation. The best known of these new techniques is probably the electronic nose (E-nose) [7–10]. The E-nose has been introduced as a fast, non-destructive and at-line alternative for aroma analysis measuring the change in piezo-electric properties of a sensor array in the presence of aroma components in the sample headspace. Aroma analysis with the E-nose has not always been very successful. It is very sensitive to drift and lacks the possibility for identification of the different aroma compounds causing the signal change. Recently, the mass spectrometry based E-nose (MSE-nose) has been introduced as a fast and sen-

* Corresponding author. Tel.: +1-814-865-2807; fax: +1-814-863-1031.
E-mail address: josephi@psu.edu (J. Irudayaraj).

0925-4005/$ – see front matter © 2003 Published by Elsevier B.V.
doi:10.1016/j.snb.2003.09.012

sitive, but also an expensive, alternative to fast aroma finger printing [11].

Since a few years, another fast, non-destructive, low-cost, and sensitive alternative sensor for food aroma analysis is commercially available: the zNose™. The zNose™ is a fast GC technique, which allows identification and finger printing of aroma as with regular GC but at the same time operates at the speed of the E-nose [12]. The zNose™ has not yet been as widely used as the E-nose. Reports exist on the use of zNose™ for the evaluation of the aroma of black tea [13], the detection of off-flavors in wine [14] and the analysis of plant volatiles [15].

The objective of this paper is to evaluate the potential of the zNose™ as an aroma finger printing tool. Honey was chosen as the product under study because of its specific aroma, which depends on factors such as the botanical and geographical origin of honey. Also sugar solutions were included in the analysis because of their resemblance to honey. The zNose™ was used to discriminate between honey varieties of different botanical origins, between pure honey varieties and sugar solutions and in between different sugar solutions.

2. Experimental

2.1. Honey and adulterant samples

All honeys were provided directly by US honey producers. In experiment 1, samples of three different honey varieties (buckwheat, clover, orange blossom) were used. For experiment 2, samples of six different honey varieties (buckwheat, clover, orange blossom, black locust, mint, carrot) from different geographical origin compared to the honeys considered in the first experiment were used. Liquid medium invert cane and beet sugars were purchased from the Imperial Sugar Company (Sugarland, TX, USA).

In experiment 1, 10 independent samples were analyzed for each honey variety and sugar solution. In experiment 2, 10 independent samples per honey variety and sugar solution were measured, of which eight were used for calibration purposes and two for external validation. The measurement protocol is described in the next paragraph.

2.2. zNose™ measurements

The zNose™ (7100/4100 vapor analysis system, Electronic Sensor Technology, USA) used for this work has a surface acoustic wave sensor (SAW) with a parts per billion sensitivity. The SAW detector is a small miniature vapor chemical sensor used to detect volatile organic compounds (VOCs). The base material of a SAW device is an uncoated piezo-electric quartz crystal. This crystal is in contact with a thermoelectric element, which controls the temperature for cooling during vapor adsorption and for heating during cleaning of the crystal. The crystal operates by maintaining highly focused and resonant surface acoustic waves of 500 MHz on its surface. When volatiles adsorb on the surface of the sensor the frequency of the surface acoustic wave will be altered, which will in turn affect the detection signal and allow identification of the contaminants [12].

For the zNose™ measurements, 8 g of pure honey or pure sugar solution was transferred into a vial of 40 ml (98 mm length and 28 mm outer diameter) sealed with a screw cap containing a septum. The vials were then transferred into a waterbath at 50 °C where the samples were allowed to equilibrate with the headspace in the vial for a minimum of 120 min. To prevent any leakage during this equilibration period the screw cap with septum was covered with an extra plastic cap. The analysis temperature of 50 °C was chosen after an initial set of experiments whereby the profiles of all pure honeys were compared over five repetitions at room temperature, 50 and 70 °C (results not shown). At room temperature the profiles were less concentrated and more sensitive to changes in ambient temperature. At 70 °C the profiles were very intense but more noise susceptible, possibly due to reactions occurring in the honey at high temperature. At 50 °C the profiles were eventually both intense and stable and equilibration of the headspace was relatively fast. After equilibration the samples were measured one by one with the zNose™.

The zNose™ was provided with a 5 cm needle at the inlet, which was used for sampling through the septa of the vials. The sampling mode was set to 5 s after which the system switched to a 10 s data acquisition mode. During this time period the gas sample was released from the trap inside the system and carried over the column (DB-5) in a helium flow of 3.00 cm^3. On the column the different chemical components in the gas sample were separated and sequentially detected by the SAW detector through a deviation from its set frequency change. Data were collected every 0.02 s. The inlet temperature was 150 °C, the valve temperature was 120 °C, and the initial column temperature was 70°C. During analysis the column temperature was ramped at the rate of 10 °C per second to a final column temperature of 100 °C. The SAW sensor was operated at a temperature of 40 °C.

After each data sampling period the system needed a 15 s baking period, in which the sensor was shortly heated to 125 °C and after which the temperature conditions of the inlet, column, and sensor were reset to the initial conditions. In between each sample measurement at least one blank was run to ensure cleaning of the system and a stable baseline.

2.3. Data analysis approach

As the zNose™ is a combination of a sensor based detection and a regular GC analysis, the data resulting from the zNose™ measurements were thus approached in two different ways.

First, a regular GC data analysis approach with the comparison of different peaks and peak areas was attempted. This was possible through the software of the instrument,

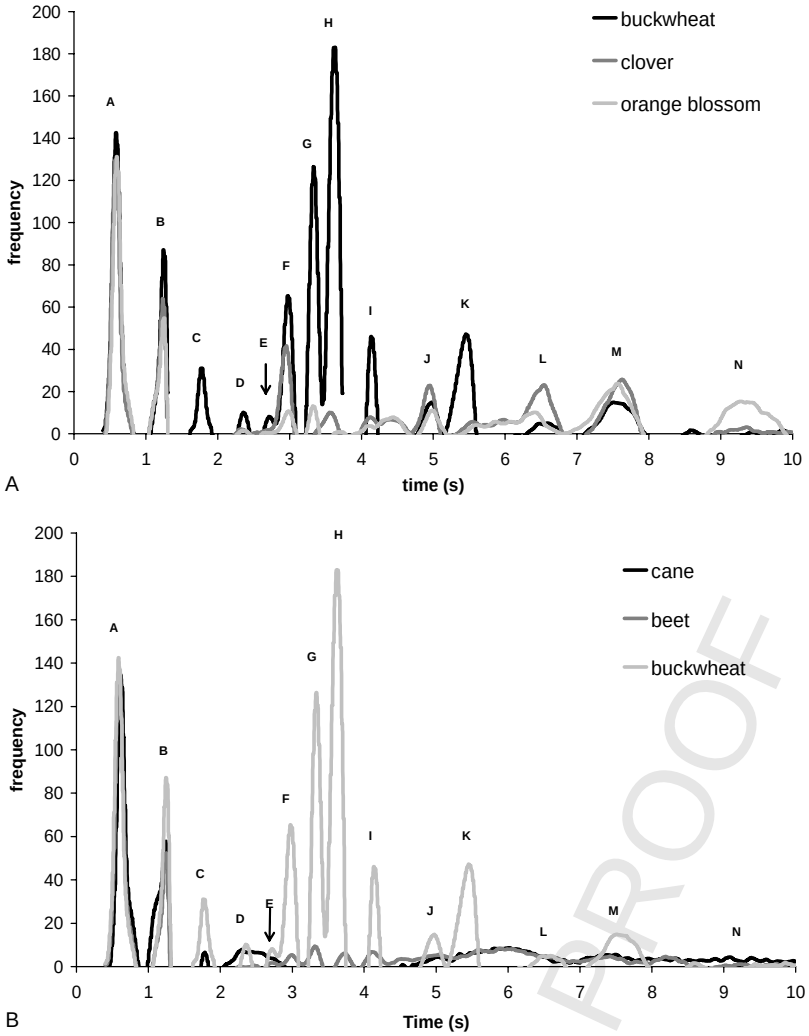

Fig. 1. Chromatograms of (A): buckwheat, clover, and orange blossom honey and (B): buckwheat honey, beet, and cane invert sugar. Letters indicate peaks considered for the PCA analysis.

which automatically transforms the frequency profile that is read from the SAW sensor to its first derivative. When only the positive values of this first derivative plot were considered, a chromatogram, which is similar to a regular GC chromatogram resulted (Fig. 1). Each peak found in this derivative plot corresponded to a specific volatile compound and had a retention time specific for the column and analysis temperature. The area under the peak was correlated to the compound concentration and was expressed in counts (cts).

For the chromatogram approach, 14 corresponding peaks in all chromatograms of all different products were selected and their relative areas compared. Relative peak areas were calculated as the absolute peak area (in counts) of each peak divided by the sum of all peak areas. When a peak was not present in a certain chromatogram its relative area was set to zero.

In a second approach, the full first derivative profile (positive and negative values) was considered and treated as spectral data (Fig. 2). In this case the full frequency spectrum of every sample was analyzed. Vertical baseline shifts in the frequency profiles were automatically filtered out by taking the first derivative. Next to the vertical shifts also horizontal shifts are a very common phenomenon in all types of chromatography. Small fluctuations in injection time, temperature profile, and data processing of the system cause the different components to be released and detected at slightly different retention times or within a 'time window'. In normal chromatographic analysis this does not generally result in problems since only a limited number of selected peaks are compared, each within its own window. In the case when full spectra are compared this shift leads to misinterpretation, however, as important peak information is compared with noise when two spectra are not perfectly aligned. To correct for horizontal shifts, an algorithm was developed in MATLAB version 6.1 (The Mathworks, Inc.).

Assume the recorded zNose^{TM} aroma spectrum consists of n datapoints. Every datapoint i, where $1 = i = n$, consists of one frequency value and one time point at which

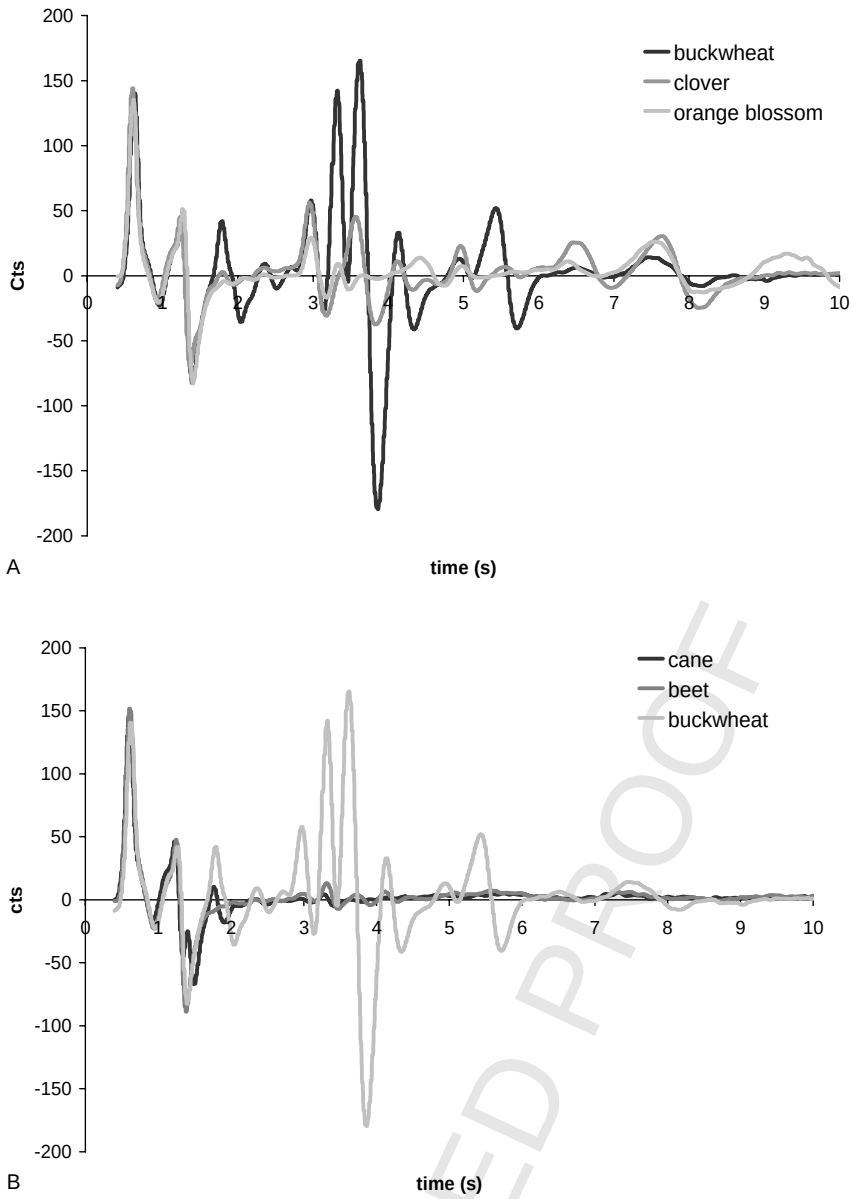

Fig. 2. Spectra of (A): buckwheat, clover, and orange blossom honey and (B): buckwheat honey, beet, and cane invert sugar.

this frequency reading was made. For a baseline correction, a constant value is subtracted or added to all the frequency readings of the spectrum. The whole spectrum shifts parallel in the vertical direction. In case of a horizontal shift correction of the spectrum, the values of the frequency readings stay constant, but the time at which they occur is adjusted according to the following formula:

$$t_{\text{new},i} = a + bt_{\text{old},i} + ct_{\text{old},i}^2$$

where $t_{\text{new},i}$ is the corrected time which is assigned to the ith frequency reading, $t_{\text{old},i}$ is the original time for the ith frequency reading, a, b, and c are the regression coefficients applied to transform the old time value into a new one. For $a = 0$, $b = 1$ and $c = 0$, no horizontal shift correction is carried out. For $a \neq 0$, $b \neq 1$ and $c \neq 0$, the spectrum shifts over a constant value a. For a positive and negative value of a, the shift will be to the right and left, respectively. For $a \neq 0$, $b \neq 1$ and $c \neq 0$, the new time value is a linear function of the old time value. The spectrum is stretched linearly over time. Frequency readings with a large original time value will be shifted over a larger time interval than values with a low original time value. For $a \neq 0$, $b \neq 1$ and $c \neq 0$ the original spectrum is stretched non-linearly over time, with the largest shifts for the frequency points with the largest original time values.

For the developed algorithm a spectrum of buckwheat honey was selected as reference spectrum as buckwheat was the product with the most complex aroma profile. All other spectra were shifted horizontally to have the best overlap with this reference spectrum. In the algorithm the

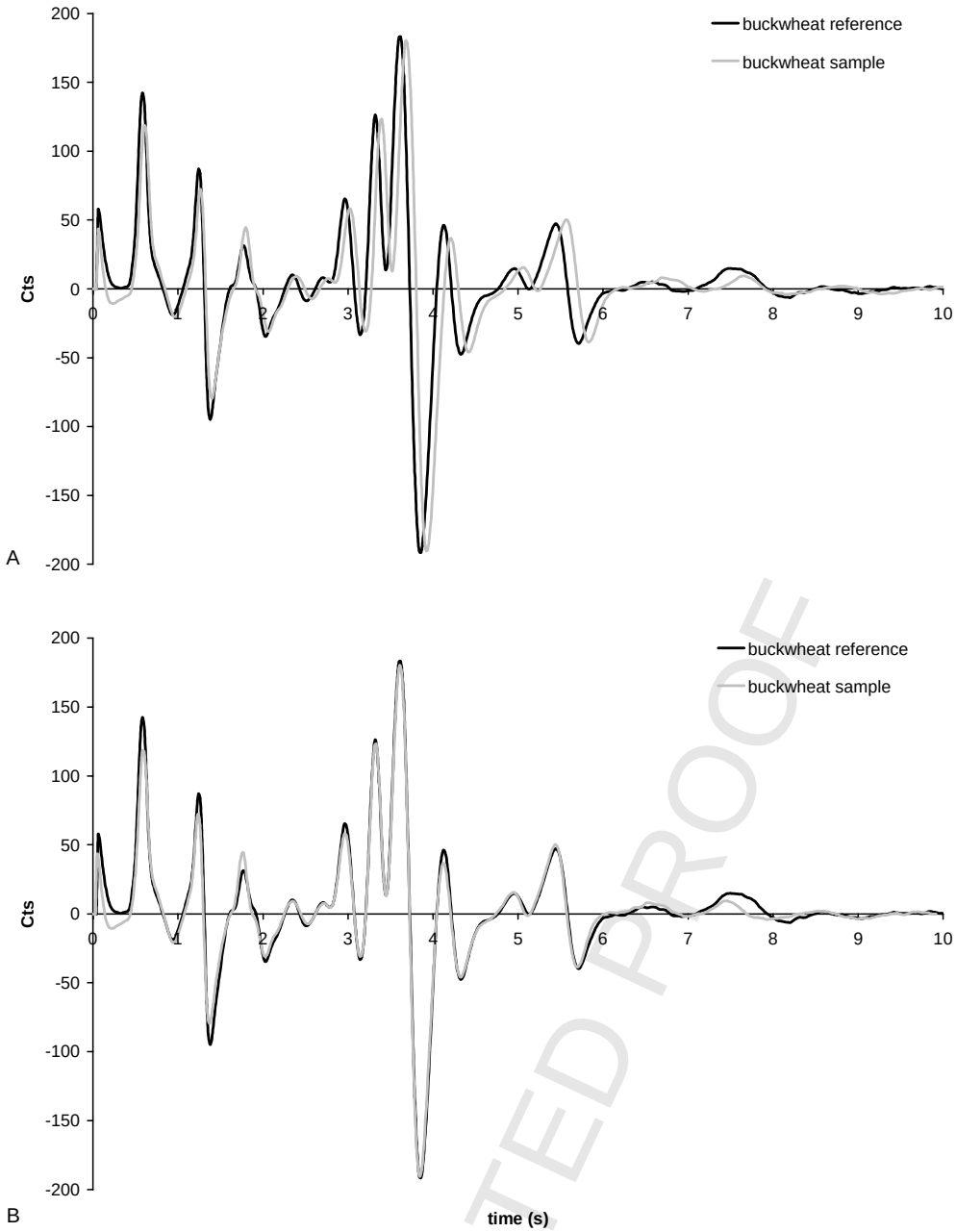

Fig. 3. Illustration of correction algorithm applied to zNose^{TM} honey aroma spectra. A buckwheat honey spectrum was used as reference spectrum. The spectra of all other honey samples were corrected with respect to this spectrum. (A): reference spectrum and raw spectra of a buckwheat honey sample (B): horizontal shift correction algorithm applied to the buckwheat honey sample.

three parameters were adjusted manually in order to shift and stretch the spectra linearly or non-linearly, depending on the needs. In Fig. 3 this corrective algorithm is illustrated for another buckwheat honey sample. The horizontal shift in chromatogram or spectrum is clearly illustrated in Fig. 3A. For small time values, the peaks of the buckwheat sample overlay those of the buckwheat reference, but as the time increases, the maxima of corresponding peaks lay further apart, suggesting the need for a horizontal shift correction. The corrected buckwheat spectrum is shown in Fig. 3B. The corresponding peak maxima of the reference and measured honey sample now occur at the same time. The corrected spectra are now ready for further statistical analysis.

2.4. Statistical analysis

The data were processed with principal component analysis (PCA) using "The Unscrambler" software version 6.11 (CAMO AS, Trondheim, Norway) and with canonical discriminant analysis (CDA) using SAS/STAT software version 8.2 (SAS Institute, Cary, NC, USA) [16].

3. Results and discussion

3.1. Experiment 1: approach to zNose™ data analysis

In a first experiment the zNose™ was tested for its ability to provide individual fingerprints of the aroma of different pure honeys and sugar solutions. In this experiment, the fingerprints obtained were not used for full identification of the different honey aroma compounds nor for classification of different honeys. Instead, two different data analysis approaches: the chromatogram and spectral approach, were developed for optimal extraction of information contained in the raw data. To not overload the graphs only three different honeys and two sugar solutions were included at this point. Sugar solutions were included in the test because of their resemblance with honey and because they are often mentioned as mimics for honey in adulteration practices.

With the chromatogram approach, PCA analysis on the total dataset of all relative peak areas of three honey types (buckwheat, clover and orange blossom) and two sugar solutions (beet invert and cane invert) resulted in a separation between all products with PC1 and PC2 explaining 90% of the total variance (Fig. 4). The clearest separation was among pure honeys and pure sugar solutions along the PC1 axis. The corresponding loading plot revealed this separation to be related mainly to components in the first part of the chromatogram (Fig. 1A and B), which were more prominent for the sugars than for the pure honeys. A further separation was found among the individual pure honeys. This separation was dominated to a much larger extent by PC2. This reflected the very distinct aroma of buckwheat honey compared to other honeys. Especially in the middle part of the chromatogram (2.5–5 s) a few very pronounced peaks were found for buckwheat, which were not present in the other honey varieties. This was confirmed in the loading plot of PC2 where the highest loadings were also appointed to these peaks in the middle (not shown).

In the spectral approach, both the negative and positive values in first derivative plots of all honeys and sugar solutions were included in the PCA analysis. This did not at first lead to a good separation (Fig. 5A). Only the aroma fingerprint of buckwheat appeared specific enough to be separated from the rest in a PCA plot with PC1 and PC2 explaining 80% of the total variance.

Closer examination of the spectra, however, revealed that a horizontal shift between the different spectra was causing this poor PCA separation. This horizontal shift was attributed to fluctuations in retention time of the chemical components on the chromatographic column. An algorithm was developed to correct the spectra for this horizontal shift. This is described in Section 2.4 of the materials and methods.

A PCA analysis on the corrected spectra did result in a much better separation with PC1 and PC2 explaining 89% of the total variance (Fig. 5B). Buckwheat honey did again form an isolated group, this time separated from the rest along the PC1 axis. Separation among honey varieties and sugars was recorded along the PC2 axis. Evaluation of the corresponding loading plots of PC1 and PC2 showed again that the separation of buckwheat, which was dominated by PC1, was explained mainly by the contribution of the middle part of the chromatogram. At this point in the spectrum the loadings for PC1 were the highest. The separation among honeys and sugars was in turn determined by the beginning (0–2.5 s) and end (5–10 s) parts of the spectrum with the loading plot of PC2 carrying the highest loadings at these positions.

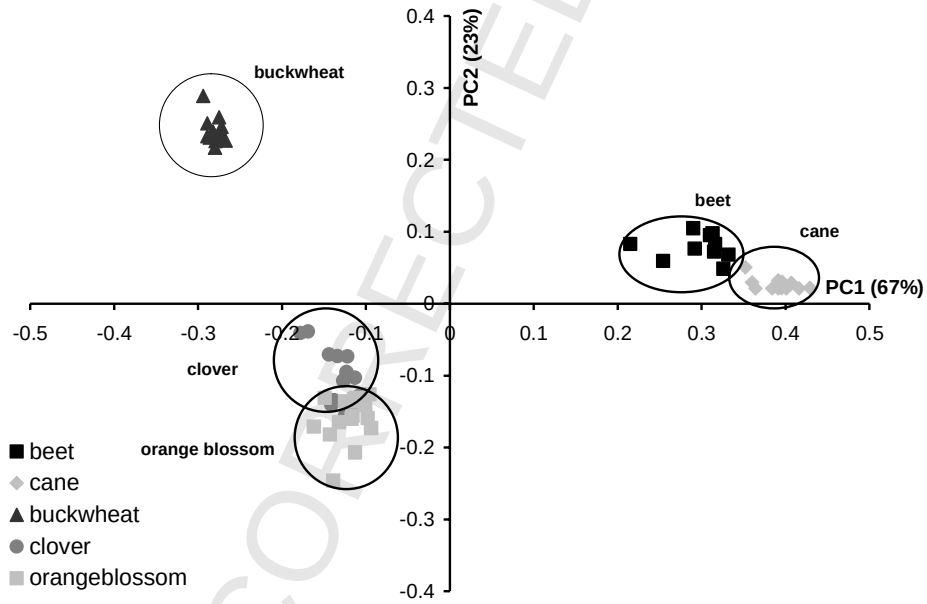

Fig. 4. PCA score plot of buckwheat honey, clover honey, orange blossom honey, beet invert sugar, and cane invert sugar, based on zNose™ chromatograms.

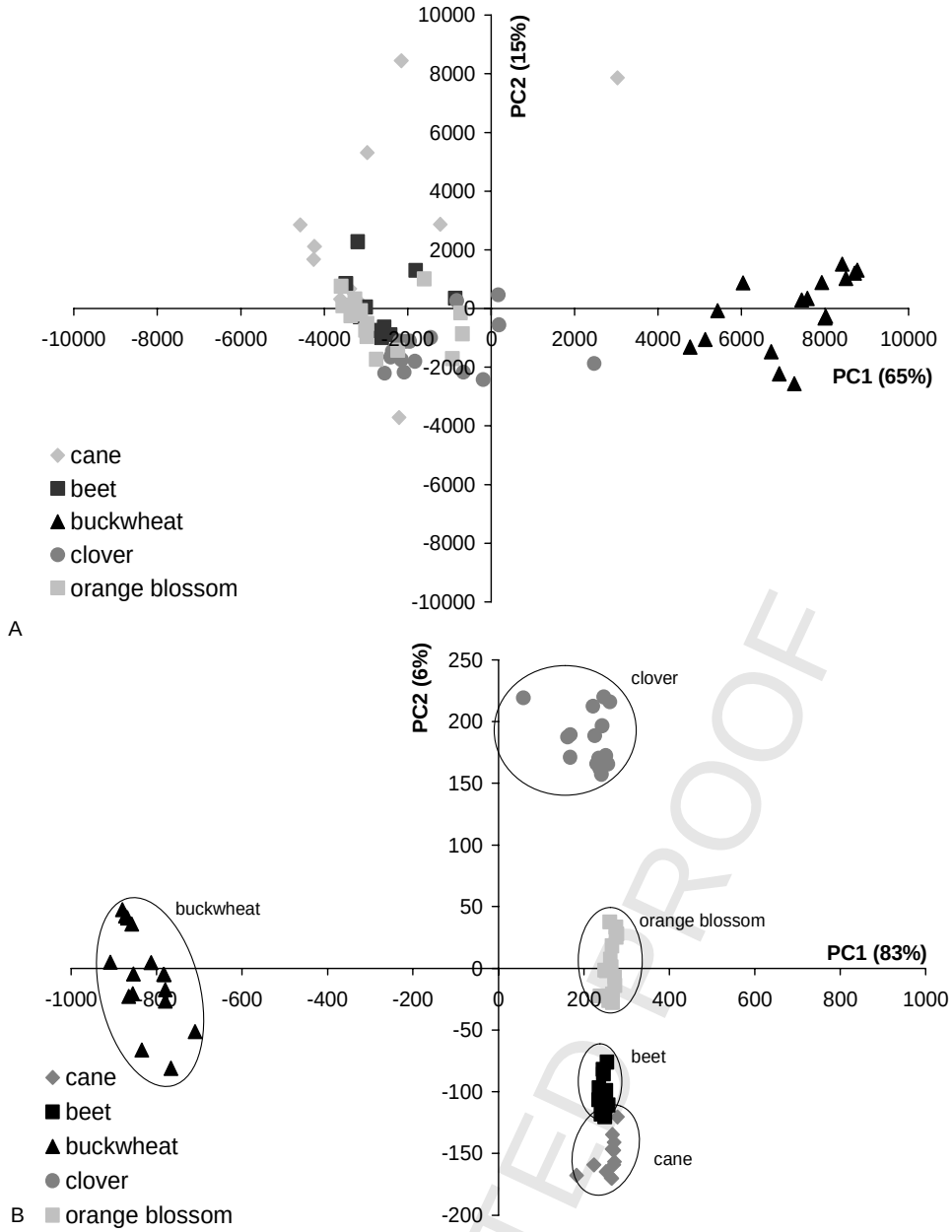

Fig. 5. PCA score plot of buckwheat honey, clover honey, orange blossom honey, beet invert sugar, and cane invert sugar, based on zNose™ (A): uncorrected and (B): corrected full spectra.

From this analysis it appeared that both the chromatogram and spectral approach hold potential to classify different types, which will be illustrated in the experiment described in the next paragraph.

3.2. Experiment 2: classification of different honeys and sugar solutions with zNose™

In this experiment six honeys of different botanical and/or geographical origins as in the previous experiment and two sugar solutions were measured. The PCA analysis in experiment 1 has indicated that both the chromatograms and the corrected spectral data are potentially valuable to build classification models for honey. In this experiment classification models are built for both data types. All models were calibrated on 64 measurements. Two independent measurements per honey and sugar solution were used as an external validation set.

In the chromatogram approach, the 12 most abundant honey volatiles were selected and used directly as explanatory variables in the discriminant analysis. This model showed a good classification performance. All but one of the 16 external validation samples were classified correctly. Only one sample from the carrot honey was classified in the group of the clover honey.

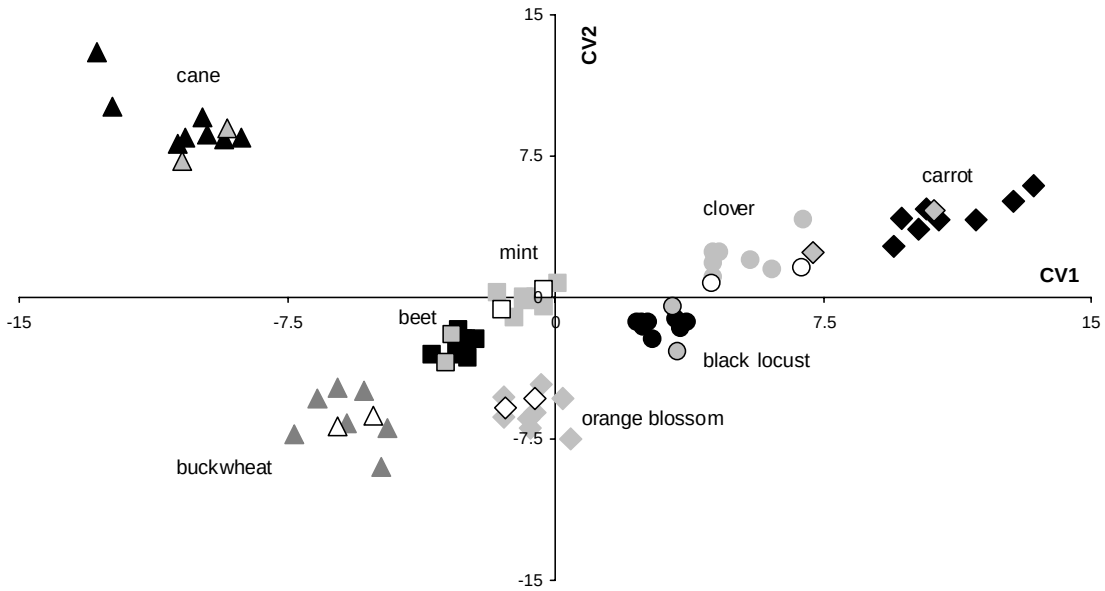

Fig. 6. Two dimensional CDA plot of buckwheat, clover, orange blossom, black locust, mint, and carrot honey, and beet and cane invert sugar based on zNose^{TM} corrected full spectra. Calibration measurements are depicted by filled symbols. Validation measurements are indicated by lighst colored and dark bordered corresponding symbols.

In case of the corrected spectral data, it was not possible to use the full spectra directly to perform the discriminant analysis. The number of variables (480) exceeded the number of observations (80) considerably, resulting in an overfit. To resolve this problem data reduction techniques such as principal components or canonical variable analysis had to be applied first. A discriminant analysis based on eight PC's resulted in a good classification model, in which, again, all but one of the validation samples were correctly classified. The PCA data reduction has the disadvantage that linear combinations of the original variables are constructed to describe the total variance in the data structure rather than accentuating the sometimes very small differences in spectral information between honeys. Canonical discriminant analysis offers a good alternative to overcome this. In CDA, canonical variables (cv) are calculated, which are also linear combinations of the original variables, but which maximize the ratio of between-groups variance over within-groups variance. Applying discriminant analysis on these canonical variables results in a discriminant function, which enables classification of any future measurement depending on mahalanobis distance to group means.

Twelve time points on the corrected spectra were visually selected to calculate five canonical variables on which the discriminant analysis was conducted. This resulted again in a good classification performance as illustrated in Fig. 6. In a two dimensional canonical variate plot, all honeys and sugar solutions can be visually discriminated. The calibration observations are depicted in plain symbols and the external validation observations are indicated with a lighter and bordered symbol similar to the corresponding honey calibration observations. Results obtained indicated that the validation measurements coincide with the corresponding calibration measurements, except for one carrot honey validation observation, which is classified among the clover honey observations. This corresponds to a 94% correctly classified external validation samples.

The zNose^{TM} can, therefore, be considered sufficiently sensitive to discriminate among the aroma of the different honey varieties examined. In addition, the aroma fingerprints of adulterant sugars can also be discriminated from those of pure honeys and also among different adulterants the fingerprints are sufficiently unique to separate them from each other.

4. Conclusion

In this work, the zNose^{TM} was introduced as a new potential at-line technique to analyze the aroma of honey. Aroma fingerprints of 6 different honeys were sufficiently specific to discriminate these honeys based on their aroma composition. With CDA, pure honey, and pure sugar solutions could be discriminated from each other, whether the data were approached as chromatograms with relative peak areas or as full spectra, which were corrected for horizontal shifts. Validation of the discriminant models was done externally with an independent sample set.

This work clearly shows the potential of the zNose^{TM} as a fast aroma finger printing technique. With some future work on optimization of the experimental conditions and extension to a broader range of honey types and origin the zNose^{TM} has the potential for practical implementation.

Acknowledgements

The National Honey board is acknowledged for supporting this research. Jeroen Lammertyn is postdoctoral fellow with the Flanders Fund for scientific research (FWO–Vlaanderen). Els Veraverbeke is postdoctoral fellow with the Catholic University of Leuven (PDM 02/158) and benefits form a NATO grant (N1/5-CVW.D7.842) to conduct research at The Pennsylvania State University (USA).

References

[1] J. Lammertyn, B.E. Verlinden, B.M. Nicolaï, in: W. Jongen (Ed.), Applying Advanced Instrumental Methods: mealiness in fruit, in Fruit and Vegetable Processing: maximising quality. Woodhead Publishing Ltd., Cambridge, 2002, pp. 170–187.

[2] R.F.A. Moreira, L.C. Trugo, M. Pietroluougo, C.A.B. De Maria, Flavor composition of cashew (*Anacardium occidentale*) and marmeleiro (*Croton species*) honeys, J. Agr. Food. Chem. 50 (2002) 7616–7621.

[3] R.A. Perez, C. Sanchez-Brunete, R.M. Calvo, J.L. Tadeo, Analysis of volatiles from Spanish honeys by solid-phase microextraction and gas chromatography–mass spectrometry, J. Agr. Food. Chem. 50 (2002) 2633–2637.

[4] Q.X. Zhou, C.L. Wintersteen, K.R. Cadwallader, Identification and quantification of aroma-active components that contribute to the distinct malty flavor of buckwheat honey, J. Agr. Food Chem. 50 (2002) 2016–2021.

[5] J.S. Bonvehi, F.V. Coll, Flavor index and aroma profiles of fresh and processed honeys, J. Sci. Food Agric. 83 (2003) 275–282.

[6] H.R. Muskowitz, Correlating sensory ratings with objective measurements of foods, in: Product Testing and Sensory Evaluation of Foods, Marketing, and R&D approaches, Food and Nutrition Press (1983) pp. 524–566.

[7] T. Eklöv, G. Johansson, F. Winquist, I. Lundström, Monitoring sausage fermentation using an electronic nose, J. Sci. Food Agric. 76 (1998) 525–532.

[8] D.D.H. Boothe, J.W. Arnold, Electronic nose analysis of volatile compounds from poultry meat samples, J. Sci. Food Agric. 82 (2002) 315–322.

[9] F. Sinesio, C. Di Natale, G.B. Quaglia, F.M. Bucarelli, E. Moneta, A. Macagnano, R. Paolesse, A. D'Amico, Use of the electronic nose and trained sensory panels in the evaluation of tomato quality, J. Sci. Food Agric. 80 (2000) 63–71.

[10] J.R. Stetter, S. Strathmann, C. McEntegart, M. Decastro, W.R. Penrose, New sensor arrays and sampling systems for a modular electronic noses, Sens. Actuators B Chem. 69 (2000) 410–419.

[11] S. Saevels, J. Lammertyn, A. Berna, E.A. Veraverbeke, C. DiNatale, B.M. Nicolaï, Evaluating apple quality after storage: assessment of the aroma by means of an E-nose and MSE-nose, Postharvest Biol. Tec. (2003), in press.

[12] E. Staples, The zNose™, a new electronic nose technology for analytical exploration of the chemical world which surrounds us, in: proceedings of the on-site conference (Geo–Odyssey Meeting), Boston, MA, November 8, (2001).

[13] E. Staples, The chemistry of black tea aroma, EST Internal Report, April 2002. (2002).

[14] E. Staples, Detecting 2, 4, 6 TCA in corks and wine using the zNose™, http://www.estcal.com/TechPapers/TCA_in_Wine.doc, (2000).

[15] M. Kunert, A. Biedermann, T. Koch, W. Boland, Ultrafast sampling and analysis of plant volatiles by a handheld miniaturised GC with pre-concentration unit, J. Sep. Sci. 25 (2002) 677–684.

[16] SAS Institute Inc., SAS/STAT® 1999, User's Guide, Version 8, SAS Institute Inc., Cary, NC, 1999.

Biographies

Jeroen Lammertyn is MSc in Biological Engineering and Ms Sci in Biostatistics. He obtained his PhD in applied biological engineering at the Catholic University of Leuven, Belgium in 2001. In 2002–2003 he was research associate at the Pennsylvania State University, USA. Since October 2003 he is Assistant Professor at the Catholic University of Leuven. His main research interests involve (bio-)sensor technology for quality control of the biological production chain and biostatistics.

Els Veraverbeke is MSc in Biological Engineering. She obtained her PhD in applied biological engineering at the Catholic University of Leuven in 2001. In 2002–2003 she spent one year as research associate at the Pennsylvania State University, USA. Since October 2003 she is research associate at the Catholic University of Leuven. Her main research interests involve postharvest technology, food quality control, and safety.

Joseph Irudayaraj has a PhD in agricultural and biological engineering in the area of Numerical Modeling of Food systems. His present research focus on sensor methodology development addresses the monitoring of food quality and safety attributes. Techniques examined are spectroscopy and zNose™ for chemical finger printing and optical biosensors for microorganism detection. His teaching responsibility includes instruction in the areas of instrumentation and measurements, biosensors, and food engineering.