VOL. 11, NO. 10, MAY 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

FISH QUALITY STUDY USING ODOR-PROFILE CASE-BASED REASONING (CBR) CLASSIFICATION TECHNIQUE

Muhammad Sharfi Najib, Nurul Hafizah Zamberan, Nurdiyana Zahed, Fathimah Abdul Halim, Muhammad Faruqi Zahari, Wan Muhamad Azmi Mamat and Hadi Manap
Faculty of Electrical and Electronics, Universiti Malaysia Pahang, Pekan, Malaysia
E-Mail: sharfi@ump.edu.my

ABSTRACT

Fish has high quality protein and other essential nutrients and are an important part of a healthful diet. It is important to make sure the quality of fish to avoid food poisoning. There are three methods to know the quality of fish which is sensory, microbiology and chemical methods. Nowadays, some fish monger use formalin to make fish looks fresh and good. Formalin is a colourless strong-smelling chemical substance usually used in industry of textiles, plastics, papers, paint, construction, and well known to preserve human corpse. It is derived from formaldehyde gas dissolved in water. Exposure to the gas and vapour can make irritation to the eyes, nose and respiratory tract. There is quite difficult for consumers to differentiate fresh fish without formalin and fish with formalin. This is because, they look fresh and good but the different is the odor which is the fresh fish still have a fishy smell while fish with formalin do not smelly. Therefore, we use electronic nose (E-nose) to know fresh fish and formalin-based preserved fish. E-nose consists of an array of conductometric chemical sensors which change resistance when exposed to vapour. The odor-profile of the fish samples were collected based on designated experimental procedure. The measured raw data was then stored in Microsoft Excel data and converted into MATLAB. Then they were normalized and their unique features were extracted using statistical tools. The input features were than inserted into Case-based Reasoning (CBR) library and intelligently classified using CBR method and validated based specific performance measure. The results have shown that the CBR classified with 100.00% rate of accuracy.

Keywords: fish, electronic nose, formalin, case-based reasoning.

INTRODUCTION

Food is an important needs in human life to survive and get energy. It is importance for us to choose a healthy food to avoid food poisonous [1]. The health of all people in the world is needed to get peace and security. So that, everyone has rights to get the highest standard of health without distinction of religion, race, political belief, economics or social conditions. World Health Organization (WHO) is established as specialized agency to promote and protect the health of people in all states [2]. Fish has high quality protein and other essential nutrients and are an important part of a healthful diet [3]. A well-balanced diet that includes a variety of fish can contribute to heart health [4] and aid in children's proper growth and development [3]. As with any type of food, it is important to handle fish safely in order to reduce the risk of food borne illness or food poisoning [5]. There are two types of fish which is freshwater fish and saltwater fish [6]. Fresh fish means fish is not spoil and do not contains any harmful chemicals as preservative [3].

Nowadays, fisherman uses preservative to increase the shelf life of fish to avoid it from spoilage. In many studies, it was proved that most imported fishes are contaminated with formalin. The use of formalin in the control of parasites on fish is dangerous to human being and environment [7]. Formalin is an aqueous solution of formaldehyde gas and contains of not less than 37% by weight of water and 6 to 13% (12%) methanol [8]. It is recommended to control and prevention [9] of external protozoan parasites on fish [10]. The concentration of formalin solution is 250ppm for a contact period of up to one hour for the control [7]. Formalin is solution from formaldehyde gas. Formaldehyde is danger to human body because it gives effect to our health [11].

There are methods to know the freshness of fish which is sense of human [12], microbiology method [13], chemical method [14] and electronic devices [15]. Sense of human or sensory method can be defined as the scientific discipline used to analyze the characteristics of food through the sense of touch, smell, sight, hear and taste. Sensory evaluation is taken by assessing the appearance, texture and odor. Microbiological method is used to estimate the number of bacterial in the fish. It is the analysis of fish which is involved testing the presence or absence of pathogens. Chemical method is involved by determination of the concentration of a specific chemical in food. Electronic devices that will be used in this study is electronic nose (E-nose).

An electronic nose is constructed with virtual instrument to detect and quantify the concentration of volatile compounds (bio-amines) in fish, which occur at during spoilage [16]. Electronic nose can be used to identify odor variety, aroma intensity and degree of freshness [17]. The E-nose consists of an array of conduct metric chemical sensors which change resistance when exposed to vapours. Each odor generates a characteristic pattern and smell print of certain compounds [18]. The process of smell print begins by collecting data based on the respond of sensor. The chemical reaction that the sensor detected will convert into electrical signal [18]. The degree of sensitivity and type of odor detected are depends to type of sensor and number of sensor in sensor array.

The data processing and pattern recognition technique can be applied to the sensor signals to differentiate substances based on collection [19][20].

Several intelligent classifications have been reported such as artificial neural network (ANN) [21], k-nearest neighbours (k-NN) [22] and case-based reasoning (CBR)[23] . One of the new method used in E-nose is CBR that used to solve the problem is Case Based Reasoning (CBR) method [24][23]. Case based reasoning is adapt and combine old solution to solve new problem [25]. It is also a process of remembering and adapting or remembering and comparing. People use cases because of difficulties to be understood and to solve the problems. In the general course of reasoning, we tend to interleave these processes according to what is needed at that time. We cannot understand the situation so that we evaluate our solutions by projecting their results (old problem) using assessment method [26].

Hence, this research is a necessity study to classify quality of fish freshness based on the presence of formalin. Therefore, this paper proposes a new technique, employing E-nose and CBR, for classifying different fish quality from its odor-profile.

METHODOLOGY

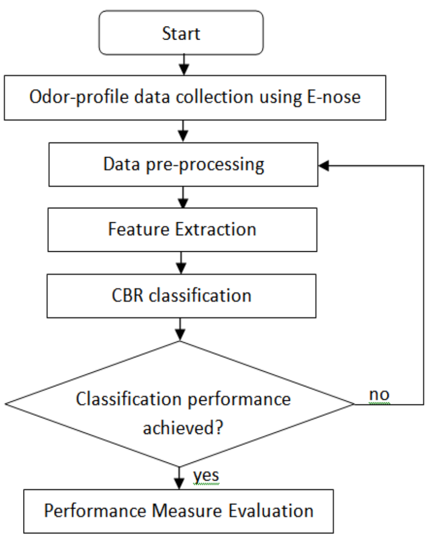

Figure-1. The overall research flow.

Figure-1 shows the overall research methodology of fish freshness odor-profile classification. It was initiated with collecting sample data of fish odor-profile using E-nose as illustrated in Figure-2. There are two samples of fish; fresh fish (FF) and formalin-based preserved (FPF) fish. Each sample was divided into 4 portions and placed into sample vials. Hence, the total of sample vials of fish that have been measured are 4 portions from both FF and FPF samples giving (140 measurements x 4 experiments x 4 sensor arrays) 2240 raw measured data. Then, the measured raw data were pre-processed using normalization technique and statistical approach. The significant features were identified statistically and were extracted to get significant variance input features. The significant input features that have been extracted and converted in cases of CBR library where the total of 8 cases were established (4 cases from FF and the rest were from FPF sample). The input features were then weighted with vector to increase the variance. After the feature extraction has been completed, the CBR classification initiated with CBR computation, voting and finally the performance of CBR obtained was measured. CBR is a method use of the reasoning; adapting and combining old solutions to solve a new problem. In this work, the new problem was identified as a new case. The goal is to build a solution to a new case based on the adaptation of solution to past cases. CBR involves the following steps which is retrieving relevant cases from the case memory. After that, selecting and reuse a set of best cases and deriving the solution. A proposed solution is revised in order to make sure that poor solutions are not repeated. Then, the confirmed solution is retained and stores the newly solved case in the case memory. In this work, the average of normalize data was stored in the CBR template and the formula is applied. One of the store cases is selected to be the current case. The value at first store case was same with the value at current case so that the similarities should be as higher as possible within the group and as lower as possible across the group. The step was repeated by changing the current case from data of different sample and the similarities will show at the template.

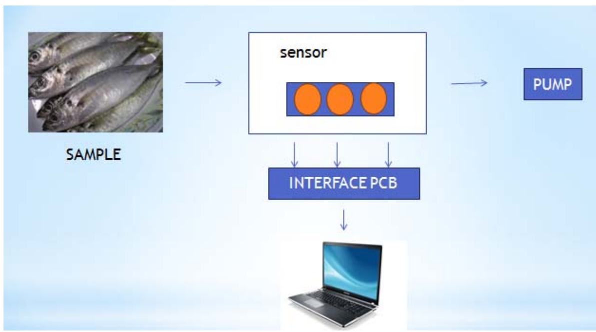

Figure-2. Illustration of the e-nose chamber.

CBR performance

By using CBR method, the classification of accuracy, sensitivity and specificity is finding from the performance of data result in the template. The similarity of data result is get from computation when all store case is selected to be current case. The result of similarity is copied to CBR performance template. The three highest value from each case is selected and assign as K=1, K=2 and K=3. The selected value is observed check if there is from the same sample or not. Then, the accuracy, sensitivity and specificity were identified from data performance by applying formula below:

VOL. 11, NO. 10, MAY 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

Accuracy = (true case/sum of case) x 100%
Sensitivity = (true sample case 1 + false sample case 2) x 100%
Specificity = [true sample case 2/ (false sample case 1 + true sample case 2)] x 100%

RESULT AND DISCUSSION

Figure-3 shows a sample data of fish odor-profile from fresh fish (FF) and formalin-based preserved fish (FPF) without normalization. The chart shows the response of resistance value produced from sensor array of the E-nose versus time in second. This chart shows the pattern of the E-nose reading in time domain. There are 4 sensors (sensor S1, S2, S3 and S4) with different color-coded are plotted. From this figure it can be shown that there is one highest resistance response towards the volatile compound the fresh fish sample. Sensor S2 is the highest while sensor S4 is the lowest response.

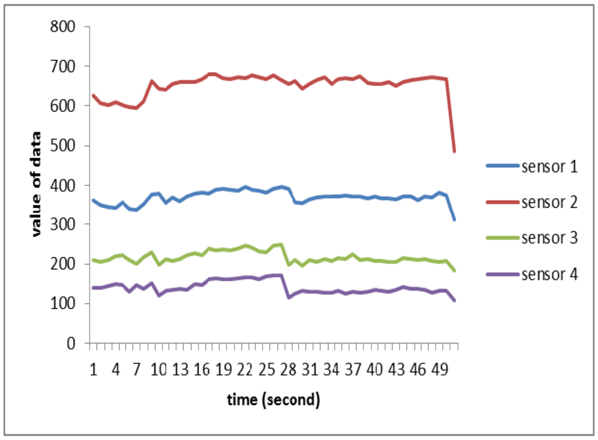

Figure-3. The data from FF sample.

As the continuity of the another data representation, Figure-4 and 5 depict the pattern that have been recognised based on frequency domain. From Figure-4 the general pattern of the odor-profile of FPF raw measured data sample have been detected. However, at this stage, the graphical representation data does not show any significant difference between FF and FPF sample. This graphical representation data were then normalised and plot in Figure-5. Figure-5 shows normalized measured data samples of E-nose reading from FF and FPF respectively. After normalization has been performed, the features of the data based on the average value of each of the sensors (S1, S2, S3 and S4) were computed. Based on pattern recognition it shows no difference between FF and FPF sample. It can be seen that Sensor S4 is very responsive towards the fish sample odor-profile. However

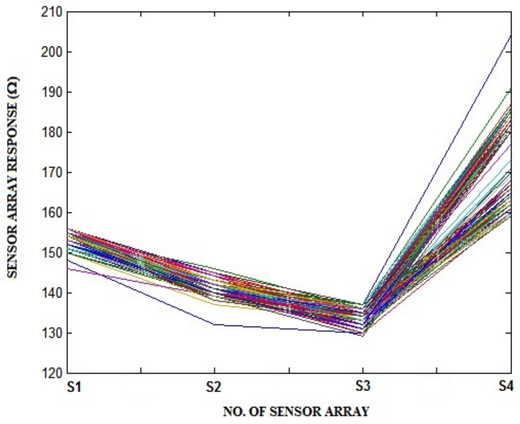

Figure-4. A raw measured data from FPF sample.

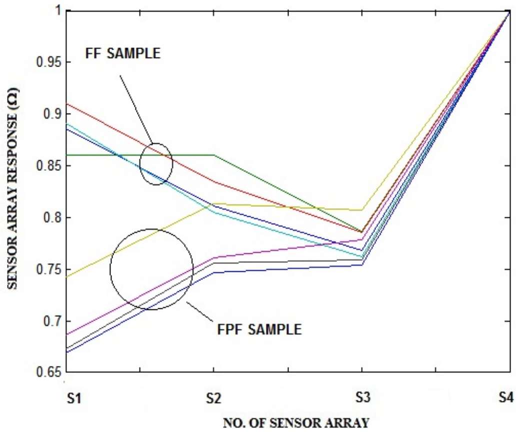

Figure-5. A normalized average measured data from FPF sample.

Case based reasoning (CBR) computation

The most critical step in a CBR system is the Retrieval step and many CBR only contain a retrieval step and one retain step (storing new cases in the case library) leaving reuse and revision for human experts, e.g. if more than one case is very close the solution may be a combination of the most similar cases. Revision is needed to ensure that the proposed solution can be mapped to the original new case that has to be classified, e.g. if there the original case of fish freshness odor-profile to classify is from fresh fish but all similar cases are from market and there is a quality difference between FF and FPF fish then an expert may need to adapt the suggested solution according to established experience.

In this paper the focus will be on the retrieval step where its main function is to calculate the similarity of two cases. One popular way to the retrieve most similar cases is that the retrieval algorithm computes the similarity value for all the cases in a case library and retrieves the most similar cases against a current problem. The similarity value between cases is usually represented as 0 to 1 where "0" means no match and "1" means a perfect

ARPN Journal of Engineering and Applied Sciences

match. The retrieval methods is based on the matching of a weighted sum of the features. For a feature vector, local similarity is computed by comparing each feature value and a global similarity value is obtained as a weighted calculation of the local similarities. A standard equation for the nearest-neighbor calculation is illustrated in Equation 1.

$$Similarity\ (T,S) = \frac{\sum_{i=1}^{n} f(T_i, S_i) \times w_i}{\sum_{i=1}^{n} wi} \qquad (1)$$

In equation 1:

T is the target case

S is the source case

n is the number of attributes in each case

i is an individual attribute from *1 to n*

f is a similarity function for attribute *i* in cases *T* and *S*

w is the importance for weighing of attribute *i*.

Execution time in a CBR system is sensitive to how much calculation is needed when comparing two features and determining how similar they are and on how many features a case has. In this work, the E-nose number of sensors have been fixed and the selection of features were focused on statistical features.

Table-1 is a sample of similarity computation of CBR system adapted from Equation 1. The average data of normalized data is inserted in the CBR template. The data from 1st trial of FF sample is stored at store case and current case. After that, the step is repeated with the other data sample.

Table-1. Similarity calculation.

Features	Source	Target	Sim	weight	norm_w	sim*norm_w
S1	0.8913	0.6686	0.2227	1.0000	0.2500	0.0557
S2	0.8046	0.7465	0.2227	1.0000	0.2500	0.0557
S3	0.7616	0.7543	0.0073	1.0000	0.2500	0.0018
S4	1.0000	1.0000	0.0073	1.0000	0.2500	0.0018
Total or global similarity between two cases						0.1150

In the above table, the similarity calculation of two cases (FF and FPF) are presented where target is a new case and source is a classified case stored in the case-library. There are 4 features (S1, S2, S3 and S4) are used for the both cases and the column 'Sim' represent the local similarity by computing the absolute difference of two features. The column 'Weight' represent the importance of each features (chemical array sensors) which is further normalized by using formula 2.

$$w_f = \frac{lw_f}{\sum_{f=1}^{n} lw_f} \qquad (2)$$

Table-2 shows the assignment of weight vectors to particular attributes. The weight has been set as 1.

Table-2. Weight vector assignment.

Weight Vector	Attributes (E-Nose Sensors)
W1=1	S1
W2=1	S2
W3=1	S3
W4=1	S4

Table-3 shows the performance evaluation of CBR classification for FF and FPF sample. After employing CBR computation, voting of true and false samples based on FF and FPF sample odor-profile; accuracy, specificity and sensitivity obtained in this works are 87.5%, 75% and 75% for the highest voting ranking (K=1). However, the performance measure in terms of accuracy, specificity and sensitivity have been improved to 100% for the medium (K=2) and lowest (K=3) voting ranking. In this initial finding, it can be said that, the sample of fresh and commercialized fish have been 100% classified. This finding shows that the characteristics of both odor-profile from FF and FPF fish contain almost similar volatile organic compound.

Table-3: Performance evaluation of CBR classification for FF and FPF.

Performance Evaluation	K=1	K=2	K=3
Criteria/ Indices	Values	Values	Values
Total Cases	4	4	4
FF case (P)	4	4	4
FPF case (N)	4	4	4
True positive (TP)	3	4	4
False positive (FP)	1	0	0
True Negative (TN)	3	4	4
False negative (FN)	1	0	0
Sensitivity=TP/(TP+FN)	0.750	1.00	1.00
Specificity=TN/(FP+TN)	0.750	1.00	1.00
Accuracy=(TP+TN)/(P+N)	0.875	1.00	1.00

CONCLUSION AND RECOMMENDATION

In conclusion, this works has shown that classification of fish freshness odor-profile is totally dependent on volatile organic compound using E-nose as an instrument. The data was successfully normalized and the features of the two samples of fish has been extracted and become input attribute based on four sensor array employed into CBR retrieval system. The CBR classification has classified two different fish sample samples with 100% accuracy, sensibility and specificity. In short, the classification using CBR as intelligent classifiers has successfully high rate of accuracy. In this work, one recommended action plan is to increase various samples of fish sample that will be identified by the fish experts specifically from Lembaga Kemajuan Ikan Malaysia (LKIM) especially fish samples that are preserved with allowable formalin and the also the forbidden one. One of the unique characteristics of CBR employed into E-Nose is the reliable parameter advised and guided by experts in this domain.

REFERENCES

[1] S. Henson and J. Caswell. 1999. Food safety regulation: An overview of contemporary issues. Food Policy. pp. 589–603.

[2] WHO. 1948. WHO definition of Health. Official records of the World Health Organization.

[3] S. J. Kaushik and I. Seiliez. 2010. Protein and amino acid nutrition and metabolism in fish: Current knowledge and future needs. Aquaculture Research pp. 322–332.

[4] M. Chalamaiah, B. Dinesh Kumar, R. Hemalatha, and T. Jyothirmayi. 2012. Fish protein hydrolysates: Proximate composition, amino acid composition, antioxidant activities and applications: A review. Food Chemistry. pp. 3020–3038.

[5] L. Ababouch. 2006. Assuring fish safety and quality in international fish trade. Mar. Pollut. Bull. pp. 561–568.

A. Askary Sary and M. Mohammadi. 2012. Mercury concentrations in commercial fish from freshwater and saltwater. Bull. Environ. Contam. Toxicol. pp. 162–165.

[6] K. C. Li, D. C. Powell, R. J. Aulerich, R. D. Walker, J. A. Render, R. K. Maes, and S. J. Bursian. 1999. Effects of formalin on bacterial growth in mink feed, feed consumption and reproductive performance of adult mink, and growth of mink kits.Vet. Hum. Toxicol. pp. 225–232.

[7] M. Ott, H. Schoenmakers, and H. Hasse. 2005. Distillation of formaldehyde containing mixtures: Laboratory experiments, equilibrium stage modeling and simulation. Chem. Eng. Process. Process Intensif. pp. 687–694.

[8] S. J. Rowland, C. Mifsud, M. Nixon, P. Read, and M. Landos. 2008. Use of formalin and copper to control ichthyophthiriosis in the Australian freshwater fish silver perch (Bidyanus bidyanus Mitchell). Aquac. Res. pp. 44–54.

[9] S. M. McLaughlin. 1996. Fish diseases and disorders. Volume 1: Protozoan and metazoan infections. Annual Review of Fish Diseases. pp. 207.

[10] D. E. Hun, J. A. Siegel, M. T. Morandi, and R. L. Corsi. 2009. Effect of ventilation rates on indoor formaldehyde concentrations in residences Buildings. pp. 2–5.

[11] H. Stone and J. L. Sidel. 2004. Sensory Evaluation Practices.

[12] G. Ólafsdóttir, E. Martinsdóttir, J. Oehlenschläger, P. Dalgaard, B. Jensen, I. Undeland, I. M. Mackie, G. Henehan, J. Nielsen, and H. Nilsen. 1997. Methods to evaluate fish freshness in research and industry. Trends in Food Science and Technology. pp. 258–265.

[13] S. B. Patange, M. K. Mukundan, and K. A. Kumar. 2005. A simple and rapid method for colorimetric determination of histamine in fish flesh. Food Control. pp. 465–472.

[14] Morsi. 2008. Electronic Nose System and Artificial Intelligent Techniques for Gases Identification. pp. 175–201.

[15] M. Delgado-Rodríguez, M. Ruiz-Montoya, I. Giraldez, R. López, E. Madejón, and M. J. Díaz. 2012. Use of electronic nose and GC-MS in detection and monitoring some VOC. Atmos. Environ. pp. 278–285.

[16] H. Gholam Hosseini, D. Luo, H. Liu, and G. Xu. 2007. Intelligent processing of e-nose information for fish freshness assessment. International Conference on Intelligent Sensors, Sensor Networks and Information Processing. pp. 173–177.

[17] D. James, S. M. Scott, Z. Ali, and W. T. O'Hare. 2005. Chemical sensors for electronic nose systems. Microchimica Acta. pp. 1–17.

[18] F. A. Halim, M. S. Najib, K. H. Ghazali, A. M. Azoddein and M. F. Zahari. 2015. ISERD - 2nd International Conference on Modern Trends in Engineering and Technology (ICMTET).

A. Loutfi. 2006. Odour Recognition using Electronic Noses in Robotic and Intelligent Systems. p. 164.

[19] H. Singh, V. B. Raj, J. Kumar, U. Mittal, M. Mishra, A. T. Nimal, M. U. Sharma, and V. Gupta. 2014. Metal oxide SAW E-nose employing PCA and ANN for the identification of binary mixture of DMMP and methanol. Sensors Actuators, B Chem. pp. 147–156.

[20] S. Güney and A. Atasoy. 2011. Classification of n-butanol concentrations with k-NN algorithm and ANN in electronic nose. International Symposium on Innovations in Intelligent Systems and Applications pp. 138–142.

ARPN Journal of Engineering and Applied Sciences

[21] M. S. Najib, M. U. Ahmad, P. Funk, M. N. Taib, and N. A. M. Ali. 2012. Agarwood classification: A case-based reasoning approach based on E-nose. IEEE 8th International Colloquium on Signal Processing and Its Applications, CSPA. pp. 120–126.

[22] M. F. Zahari, J. Thommy Anak, F. A. Halim, M. S. Najib, K. H. Ghazali, and A. A. M. Azoddein. 2015. Intelligent Classification Hazardous Gas Using Sensor Array. J. Adv. Inf. Technology. pp. 233–237.

[23] R. Lopez and E. Plaza. 1997. Case-Based Reasoning : An Overview. AI Commun. pp. 21–29.

[24] L. Kolodner. 1992. An introduction to case-based reasoning. Artif. Intell. Rev. pp. 3–34.