Bio-Inspired Sensor Data Fusion for Herbal Tea Flavour Assessment

by

Nur Zawatil Isqi Binti Zakaria
(1130610576)

A thesis submitted in fulfilment of the requirements for the degree of Master of Science (Mechatronic Engineering)

School of Mechatronic Engineering
UNIVERSITI MALAYSIA PERLIS

2017

ACKNOWLEDGEMENT

First of all, *Alhamdulilah*, all praises to Almighty God. Finally, I managed to complete my master thesis after all the hardship and tribulations during this adventurous journey with Allah blessing and guidance. In addition, I would like to take this opportunity to acknowledge and appreciate the efforts of the people in the table below that helped me during my research and the documentation of this thesis. May Allah grant them the best rank in this world and hereafter.

No.		Names	Remarks
1.	Supervisor and Co- supervisor	Prof. Dr. Ali Yeon Bin. Md. Shakaff and Dr. Ammar Bin Zakaria	Guidance, motivations and patience
2.	Department	Ministry of Education, CEASTech, School of Mechatronics Engineering, Agrotech UniMAP and Centre of graduate study	Finance, instruments, samples, guidance and knowledges sharing
3.	Parents	Zakaria Ismail and Zauyah Md Zain	Support, love, patience and finance assisted
4.	Siblings and their family	Nur Zatul-Iffah, Zulkhairi and Mohamad Zulfadzli	Supports and finance assisted
5.	Sifu	Dr. Maz Jamilah Masnan, Assoc. Prof Dr. Abu Hasan Abdullah, Abdul Halis, Rohani Farook, Azian Subari, and Syahida Sulaiman	Knowledge, patience and lunch treats
6.	Inspirations	Prof. Dr. Mohd. Noor Ahmad and Pengkalan Asam team (Wani, Siti, Mubaraq, Farhanah, Kak Aza, Kak Dayah, Dayah.and Zul)	Provide wonderful environment for my thesis writing and knowledge sharing.
7.	Special people	Dr. Latifah Munirah, Nurul Maisairah, Syahida, Hamizah, Nor Hanani, Syamimi Wahida, Nurul Aini and Nurlisa	Visiting me in hospital due to car accident and assist in my research work.
8.	Al-Fateh Educational Instittute	Pak Teh Zaki Muin (Principal), Tc.Rabiha and all teachers.	Give me special leave from teaching in order to do thesis corrections.

©This item is protected by original copyright

TABLE OF CONTENTS

© This item is protected by original copyright

©This item is protected by original copyright

CHAPTER 3 METHODOLOGY

©This item is protected by original copyright

©This item is protected by original copyright

©This item is protected by original copyright

©This item is protected by original copyright

©This item is protected by original copyright

LIST OF TABLES

© This item is protected by original copyright

©This item is protected by original copyright

©This item is protected by original copyright

© This item is protected by original copyright

©This item is protected by original copyright

©This item is protected by original copyright

LIST OF FIGURES

© This item is protected by original copyright

© This item is protected by original copyright

©This item is protected by original copyright

LIST OF ABBREVIATIONS

AD	Anderson-Darling
BP	Back propagation
CDA	Canonical discriminant analysis
CDF	Computable document format
D.O.E	Design of experiment
E.I	Electron-ionization
EEG	Electroencephologram
E-nose	Electronic nose
E-tongue	Electronic tongue
FE	Features extraction
FID	Flame ionisation detector
FDA	Linear Discriminant Analysis with fisher criterion
FS	Features selection
FTIR	Fourier transform IR spectroscopy
GC	Gas Chromatography
GC-O	GC-Olfactometry
GDA	Generalized Discriminant Analysis
GRNN	General regression neural network
HLDF	High Level Data Fusion
HPLC	High performance liquid chromatography
HS	Headspace
ILDF	Intermediate Level Data Fusion
KNN	K-nearest neighbour

© This item is protected by original copyright

KS	Kolmogrov-Smirnov
LDA	Linear discriminant analysis
LF	Lilliefors
LLDF	Low Level Data Fusion
LS	Least Squares
MLP	Multi-Layer Perceptron
MOS	Metal oxide semiconductor
MS	Mass Spectrometry
MSDF	Multi sensor data fusion
MSE	Mean squared error
MVA	Multivariate analysis
NMR	Nuclear magnetic resonance
NN	Neural network
OAA	One against all
OAO	One against one
PCA	Principal component analysis
PEN3	Portable electronic nose
PH	Pureherb
PLS	Partial least square
PNN	Probabilistic neural network
POL	Polens
QP	Quadratic programming
RBF	Radial basis fuction
RH	Rainhill
RTD	Ready To Drink

©This item is protected by original copyright

SIM	Selective ion monitoring
SMO	Sequential Minimal Optimization
SOM	Self-organizing map
SPME	Solid phase microextraction
SVM	Support vector machine
SW	Shapiro-Wilk
TIC	Total Ion Chromatography
V.F.C	Volatile flavour compound

©This item is protected by original copyright

LIST OF SYMBOLS

$\mu_1\mu_2$	Mean vector
C	Box contraint
C_1C_2	Covariance matrices
G/G0	Ratio of conductance
M/Z	Mass/charge
V	Voltage
α_i	Inequalities
β	Linear model coefficient
ξ_i	Slack variable in SVM

©This item is protected by original copyright

Bio-Inspirasi Gabungan Data Sensor Untuk Penilaian Perasa Teh Herba

ABSTRAK

Produk-produk berasaskan herba menjadi amalan pengeluaran meluas di kalangan pengeluar untuk pasaran antarabangsa dan tempatan. Memandangkan bertambahnya pengeluaran bagi memenuhi permintaan pasaran, adalah sangat penting bagi pengeluar supaya memastikan produk mereka telah memenuhi kriteria dan kualiti tertentu yang telah ditetapkan oleh pengawal kualiti. Salah satu produk berasaskan herba yang terkenal ialah teh herba. Tesis ini mengkaji penilaian-penilaian rasa berdasarkan inspirasi bio dalam konteks gabungan data melibatkan e-hidung dan e-telinga. Objektif kajian ini adalah untuk mendapatkan pengelasan yang tepat bagi pelbagai jenis dan jenama teh herba, pengelasan beberapa agen 'masking' rasa dan yang terakhir pengelasan berdasarkan perbezaan kepekatan teh herba. Dalam penyelidikan ini, dua tahap gabungan data dimanfaatkan iaitu gabungan data tahap rendah (LLDF) dan gabungan data tahap pertengahan (ILDF). Empat teknik pengelasan; 'Fisher Linear Data Analysis (FDA)', 'Support Vector Machine (SVM)', 'k-Nearest Neighbour (KNN)' dan 'Probability Neural Network (PNN)' telah diuji dalam mencari pengelas terbaik bagi mencapai objektif penyelidikan. Dalam menilai prestasi pengelas, penganggar ralat berdasarkan pengesahan silang 'k-fold' dan 'leave-one-out' (LOO) telah digunakan. Pengelasan berdasarkan data GC/MS TIC turut disertakan sebagai satu perbandingan kepada prestasi klasifikasi menggunakan pendekatan-pendekatan gabungan. Secara umumnya, melalui gabungan data tahap rendah dan gabunagan data tahap pertengahan, KNN mengatasi teknik pengelasan yang lain untuk tiga penilaian rasa. Bagaimanapun, keputusan-keputusan pengelasan berdasarkan data GC/MS TIC adalah berubah-ubah bagi aplikasi yang berbeza. Memandangkan KNN dapat memberikan keupayaan pengkelasan yang tinggi, sistem automatik pengredan dibina berdasarkan algoritma teknik tersebut.

©This item is protected by original copyright

Bio-Inspired Sensor Data Fusion for Herbal Tea Flavour Assessment

ABSTRACT

Herbal-based products are becoming a widespread production trend among manufacturers for the domestic and international markets. As the production increases to meet the market demand, it is very crucial for the manufacturer to ensure that their products have met specific criteria and fulfil the intended quality determined by the quality controller. One of famous herbal-based product is herbal tea. This thesis investigates bio-inspired flavour assessments in a data fusion framework involving an E-nose and E-tongue. The objectives are to attain good classification of different types and brands of herbal tea, classification of different flavour masking effects and finally classification of different concentrations of herbal tea. Two data fusion levels were employed in this research, low level data fusion (LLDF) and intermediate level data fusion (ILDF). Four classification approaches; Fisher Linear Data Analysis (FDA), Support Vector Machine (SVM), k-Nearest Neighbour (KNN) and Probability Neural Network (PNN) were examined in search of the best classifier in order to achieve the research objectives. In order to evaluate the classifiers' performance, an error estimator based on k-fold cross validation and leave-one-out were applied. Classification based on GC/MS TIC data was also included as a comparison to the classification performance using fusion approaches. Generally, KNN outperformed the other classification techniques for the three flavour assessments in the low level data fusion and intermediate level data fusion. However, the classification results based on GC/MS TIC data varies in different application. Since KNN provide the highest classification performance, automatic grading system was developed based on this technique.

©This item is protected by original copyright