UNIVERSITI PUTRA MALAYSIA

CUMULANT-BASED BASIS SELECTION AND FEATURE EXTRACTION TO IMPROVE HEART SOUND CLASSIFICATION

FATEMEH SAFARA

FSKTM 2013 9

CUMULANT-BASED BASIS SELECTION AND FEATURE EXTRACTION TO IMPROVE HEART SOUND CLASSIFICATION

By

FATEMEH SAFARA

Thesis submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfillment of the Requirement for the Degree of Doctor of Philosophy

February 2014

COPYRIGHT

DEDICATION

To my beloved husband,
Mohammad,
who has been a constant source of support and encouragement during this study,

and to our wonderful children,
Zahra and Mohsen,
who are indeed treasures from Allah

and to my respectable parents,
who are my pillars of inspiration, strength and compassion.

Abstract of thesis presented to the senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

CUMULANT-BASED BASIS SELECTION AND FEATURE EXTRACTION TO IMPROVE HEART SOUND CLASSIFICATION

By

FATEMEH SAFARA

February 2014

Chairperson: Shyamala A/P Doraisamy, Ph.D
Faculty: Computer Science and Information Technology

Cardiac auscultation, the direct hearing and interpreting the heart sounds, is a fundamental clinical skill that requires years to develop and refine. The interpretations are commonly prone to variations resulting in highly subjective diagnosis. Alternative technologies such as magnetic resonance imaging (MRI) and echocardiography are on the rise. However, these are expensive, and instead technologies to support or automate cardiac auscultation are becoming important and are currently widely being researched. The accuracy of cardiac auscultation could be improved through extracting objective information from phonocardiography (PCG) signals to be used for automated heart sound classification.

This study focuses on the classification of new features extracted from PCG signals represented by wavelet packet transform (WPT). A wavelet packet tree is constructed for each PCG signal, and higher-order cumulants (HOC) of the wavelet packet coefficients (WPC) are extracted and used as features, named hoc_WPC features. With the features, merits of time-frequency analysis of WPT and statistical analysis of HOC are exploited. PCG signals have been classified successfully using hoc_WPC features. An improvement of 3.02% sensitivity and 0.19% specificity have been achieved in differentiating normal heart sounds and regurgitations. The hoc_WPC features are further capable to classify heart sounds into normal, mitral regurgitation, aortic regurgitation, and aortic stenosis, with 96.95% accuracy.

Basis selection is another issue in analysis signals by WPT. For basis selection, an approach is proposed to reduce the initial search space from the entire tree to a trapezoidal sub-tree of it, and then four basis selection methods are proposed: i) multi-level basis selection (MLBS); ii) cumulant-based trapezoidal multi-level basis

selection (CT_MBS); iii) cumulant-based trapezoidal best basis selection (CT_BBS); and iv) cumulant-based trapezoidal local discriminant basis (CT_LDB).

With MLBS an energy-based information measure is used to select the best nodes of the three bottom levels of a wavelet packet tree for feature extraction. With cumulant-based trapezoidal basis selection methods, HOC are used to define information measure. This is based on the feature extraction experiment whereby the ability of HOC to represent the information laid throughout a wavelet packet tree has been shown. CT_MBS is an extension of the MLBS, whereby cumulant measure is used to prune the wavelet packet tree instead of energy. CT_BBS and CT_LDB are the extensions of the commonly used basis selection methods, which are best basis selection (BBS) and local discriminant analysis (LDB). The best classification accuracy of 98.17% was achieved by CT_LDB in classifying different types of heart sounds of this study.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi kepeluan untuk ijazah Doktor Falsafah

PEMILIHAN ASAS DAN PENYARIAN SIFAT BERDASARKAN CUMULANT BAGI PEMPERBAIKI KLASIFIKASI BUNYI JANTUNG

Oleh

FATEMEH SAFARA

February 2014

Pengerusi: Shyamala A/P Doaraisamy, Ph.D
Fakulti: Sains Komputer dan Teknologi Maklumat

Auskultasi kardium, iaitu pendengaran terus dan penterjemahan bunyi jantung, adalah kemahiran asas klinikal yang memerlukan beberapa tahun untuk berkembang dan memperhaluskan. Terdapat kemungkinan besar bagi variasi dalam penterjemahan tersebut yang dapat mengakibatkan diagnosis yang sangat subjektif. Teknologi alternatif seperti Magnetic Resonance Imaging (MRI) dan echocardiography pesat berkembang. Walaubagaimanapun, teknologi tersebut adalah mahal, dan sebaliknya teknologi-teknologi bagi menyokong atau automasikan auskultasi kardium kini menjadi penting dan dikaji secara meluas. Ketepatan auskultasi kardium boleh diperbaiki melalui pengekstrakan maklumat objektif dari isyarat phonocardiography (PCG) untuk digunakan bagi automasi klasifikasi bunyi jantung.

Kajian ini bertumpukan kepada klasifikasi baru sifat yang diekstrak dari isyarat PCG diwakili oleh wavelet packet transform (WPT). Suatu pepohon wavelet packet dibina bagi setiap isyarat PCG, dan higher-order cumulants (HOC) koefisien wavelet packet diekstrak dan diguna sebagai sifat, dinamakan sifat hoc_WPC. Dengan sifat tersebut, kelebihan penganalisaan WPT dana penganalisaan statistik HOC adalah dieksploitasi. Isyarat PCG dikasifikasikan dengan kejayaan menggunakan sifat hoc_WPC. Peningkatan sebanyak 3.02% bagi sensitivity dan 0.19% specificity telah dicapai bagi membezakan bunyi jantung biasa dan regurgitations. Sifat hoc_WPC juga berkebolehan klasifikasikan bunyi jantung kepada biasa, mitral regurgitation, aortic regurgitation dan aortic stenosis dengan 96.95% ketepatan.

Pemilihan asas, adalah satu lagi isu dalam isyarat analisis oleh WPT. Bagi pemilihan asas, suatu pendekatan dicadangkan untuk mengurangkan ruang pencarian permulaan daripada keseluruhan pepohon kepada sub-pepohon trapezoidalnya, dan empat kaedah

pemilihan asas diperkenalkan: i) multi-level basis selection (MLBS), ii) cumulant-based trapezoidal multi-level basis selection (CT_MBS), iii) cumulant-based trapezoidal best basis selection (CT_BBS), and iv) cumulant-based trapezoidal local discriminant basis (CT_LDB).

Dengan MLBS suatu ukuran berasaskan energi diguna bagi memilih nod terbaik daripada tiga peringkat terbawah suatu wavelet packet tree bagi penyarian sifat. Dengan kaedah pemilihan asas cumulant-based trapezoidal, HOC diguna bagi menakrifkan ukuran maklumat. Ini berdasarkan eksperimen penyarian sifat dimana kebolehan HOC mewakili maklumat dikeseluruhan wavelet packet tree telah ditunjuk. CT_MBS adalah suatu tambahan kepada MLBS, di mana ukuran cumulant diguna bagi memangkas wavelet packet tree. CT-BBS dan CT_LDB adalah tambahan kepada kaedah pemilihan asas yang biasa diguna iaitu best basis selection (BBS) dan local discriminant analysis (LDB). Klasifikasi terbaik sebanyak 98.17% telah dicapai oleh CT_LDB dalam mengkasifikasikan jenis-jenis dalam kajian ini.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest praise and admiration to Allah that has given me the strength, faith, confidence, and patience to complete this project despite all the challenges.

I owe my sincere gratitude to my supervisor, Dr. Shyamala Doraisamy, for giving me an opportunity to start this project. Through the course of my study, I have had great fortune to get to know and interact with her. Her comments and suggestions for further development as well as her assistance during writing this thesis are invaluable to me. Her talent, diverse background, interest, teaching and research style has provided me an exceptional opportunity to learn.

I would also like to express my sincere thanks to the supervisory committee members, Dr. Azreen Azman, Dr. Azrul Hazri Jantan, and Dr. Asri Ranga Abdullah Ramaiah for their valuable suggestions and advises throughout this work.

I acknowledge also the kind help of Dr Abdul Kahar bin Abd Ghafar, the head of Cardiology department of the Serdang Hospital, and Mr. Khairul Affandy Omar, senior cardiovascular technologist, and the staff of cardiology department, for their kind support providing access to their facilities of recording heart sounds.

I am truly indebted and thankful to Dr. Abbas Atyabi for his valuable technical supports, in particular for the long discussions that helped me sort out the technical details of my work.

The deepest gratitude to my husband, my children, my parents, my sisters, and my brothers for their unconditional supports. This thesis would not have been possible without their encouragements and patience during my PhD study.

I certify that a Thesis Examination Committee has met on 18th February 2014 to conduct the final examination of Fatemeh Safara on her thesis entitled "Cumulant-Based Basis Selection and Feature Extraction to Improve Heart Sound Classification" in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U. (A) 106] 15 March 1998. The committee recommends that the student be awarded the Doctor of Philosophy.

Members of the Examination Committee are as follows:

Marzanah A. Jabbar, Ph.D
Associate Professor
Faculty of Computer Science and Information Technology
University Putra Malaysia
(Chairman)

Fatimah Khalid, Ph.D
Associate Professor
Faculty of Computer Science and Information Technology
University Putra Malaysia
(Internal Examiner I)

Rahmita Wirza O.K. Rahmat, Ph.D
Associate Professor
Faculty of Computer Science and Information Technology
University Putra Malaysia
(Internal Examiner II)

Dimitrios Hatzinakos, Ph.D
Professor
Department of Electrical and Computer Engineering
University of Toronto
Canada
(External Examiner)

ZULKARNAIN ZAINAL, PhD
Associate Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 18th February 2014

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Shyamala Doraisamy, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Azreen Azman, PhD
Senior Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Azrul Hazri Jantan, PhD
Senior Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Asri Ranga Abdullah Ramaiah
Cardiologist
Cardiology Department
Serdang Hospital
(Member)

BUJANG BIN KIM HUAT, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

DECLARATION

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: _____ Date: _____

Name and Matric No.: _____

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: _____	Signature: _____
Name of Chairman of Supervisory Committee: _____	Name of Member of Supervisory Committee: _____
Signature: _____	Signature: _____
Name of Member of Supervisory Committee: _____	Name of Member of Supervisory Committee: _____

TABLE OF CONTENTS