

Low-Complex Environmental Sound Recognition Algorithms for Power-Aware Wireless Sensor Networks

September 2012

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy in Engineering



Keio University

Graduate School of Science and Technology
School of Integrated Design Engineering

Zhan, Yi

Thesis Summary

The past decade witnessed rapid development in the basic Internet, communications theories and in some newly emerging technologies, such as wireless sensor networks (WSNs), wearable sensing and computation. With the rapid development of these technologies, understanding individual's activities, social interaction, and group dynamics of a certain society becomes possible and plays an important role for creation a ubiquitous information society around us. This will inevitably enrich our life's content and improve our society's efficiency.

Environmental background sound is a rich information source for identifying individual and social behaviors. Therefore, many power-aware wearable devices in the WSNs system with sound recognition function are widely used to understand and trace human activities. Design of these sound recognition algorithms has two major challenges: limited computation resources and a strict power consumption requirement. These motivate us to develop a new method for recognizing environmental background sounds upon our power-aware wearable sensor node. Therefore, we address to develop a new sound recognition algorithm which can achieve high recognition accuracy while still meeting the wearable sensor's power requirement in the dissertation.

In Chapter 1, the motivation and challenge of this study are introduced. Related work is also surveyed.

In Chapter 2, the hardware architecture of our power-aware wearable sensor node and software-level sound recognition flow are introduced. Upon this resource limited platform, the assumptions and special constraints of this research are discussed. Basic approaches to tradeoff the system's accuracy and power consumption problem are proposed.

In Chapter 3, the experimental setup and process are presented. Comprehensively considering the system's accuracy and power consumption as the proposed sound recognition algorithms' performance evaluation criteria is also discussed.

In Chapter 4, sound feature extraction Mel-frequency cepstral coefficients (MFCC) and vector quantization (VQ) classification Linde-Buzo-Gray (LBG) algorithm is applied for recognizing the environmental background sounds. Applying this algorithm to 20 typical daily activity sounds, average recognition accuracy of 93.8% can be achieved. In this algorithm, how the three parameters (i.e., Mel filters number, frame-to-frame overlap and LBG codebook cluster number) affect the system's calculation burden and accuracy is also investigated. Based on the performance evaluation method in Chapter 3, the comprehensive performance of proposed MFCC+LBG algorithm is evaluated.

In Chapter 5, a new low-complex Haar-like sound feature with hidden Markov model (HMM) classification algorithm is proposed and applied to recognize the environmental sounds. Average recognition accuracy 96.3% of twenty typical daily activity sounds by the proposed algorithm can be achieved, which outperforms normal personal hearing capacity 82% accuracy. At the same time, it also satisfies the amount of calculation cost decided by the wearable sensor node's energy resource. Through experimental comparison, the proposed method outperforms other normally utilized sound recognition algorithms as the recognition accuracy and calculation cost two evaluation parameters concerned.

In Chapter 6, summary of this study is concluded. Overview of the future work is also mentioned.

要旨

無線センサーネットワークを用いて世界をリアルタイムに理解する研究が始まっている。例えば環境音をモニターして個人の活動や周囲の状況を自動認識する研究が報告されている。しかしながら、世界中のセンサー端末からクラウドにセンシング情報がそのまま転送されると、データ通信量は爆発的に増大する。そこで、センサー端末がデータを認識して転送すべきデータを圧縮することが求められる。センサー端末を環境に埋め込むためには、端末が小さく、電池の交換が不要でなければならない。つまり計算や電源の資源が極めて限られた場合でも、ある程度高精度な認識ができる高効率なアルゴリズムが求められる。そこで本研究では、省電力な無線センサーネットワークに適する計算量の少ない環境音認識アルゴリズムを検討している。

第 1 章は序論である。無線センサーネットワークと環境音認識アルゴリズムに関するこれまでの研究と課題を概説し、本研究の目的と意義を述べている。

第 2 章では、研究対象とするシステムの全体像とハードウェア構成を示し、考慮すべき仮定や制約を議論している。

第 3 章では、20 種類の日常行動に伴う環境音をフィールドで採取して、これを用いてアルゴリズムを評価する手法を検討している。また、従来の研究成果を鑑みて、性能目標を認識精度 82% 以上で電力消費 0.75mW 以下に設定している。

第 4 章と第 5 章では、特徴量の選択と識別器構成のトレーニングアルゴリズムについて性能を評価している。第 4 章では、特徴量に Mel-scale Frequency Cepstral Coefficients (MFCC) を用い、コードブックのトレーニングアルゴリズムに Linde-Buzo-Gray (LBG) や Dynamic Time Wrapping (DTW)、Gaussian Mixture Model (GMM) を用いた場合を評価している。比較検討の結果、DTW や GMM に比べて LBG は、認識精度を劣化させることなく計算量を最も低減できることを見出している。また、メルフィルタの数とフレームの重ね代量が計算量と認識精度に与える影響を詳細に調べている。メルフィルタの数は 24 個にして、フレームの重ね代を 256 ポイントからゼロにすることで、認識精度をあまり劣化させずに、計算量を大幅に低減できることを見出している。認識精度は 92.5% を達成しているが、電力消費は 0.88mW であり、目標には未達である。

第 5 章では、特徴量に Haar-like feature を用い、識別器構成のトレーニングアルゴリズムに Hidden Markov Model (HMM) や LBG を用いる方式を提案し、その性能を評価している。その結果、HMM を用いると認識率 96.3% で電力消費 0.66mW を達成でき、LBG を用いると認識率 88.6% で電力消費 0.51mW を達成できることを実証している。

第 6 章は結論である。各章で得られた知見を総括し、今後の展望を述べている。