

# Instant Learning Sound Sensor for Ubiquitous Computing

根岸 佑也

名古屋大学大学院情報科学研究科

河口 信夫

名古屋大学大学院工学研究科

## 1. はじめに

近年、ユビキタス環境の実現の期待とともに、ユーザ自身の状況と取り巻くコンテキストに基づき、そのユーザに適したサービスを提供するコンテキストウェア・システムが盛んに研究されている。コンテキスト情報を取得するデバイスとしては、GPS や加速度・圧力・温度センサなどが利用される。そして、それらを集約した小型無線センサデバイス[1]も開発され、実際にスマートルームなどに組み込む実験もされている。Jianfeng Chen[2] らによって、バスルームにおける手を洗うといった行動イベントを、室内に設置したマイクロフォンを用いた音認識により取得し、健康管理に役立てるシステムが提案されている。しかしながら、そのような信号処理を伴う複雑なパターン認識処理の設計は、個別の特微量の解析などに手間がかかるため、日常空間中の任意の物音に対して、音認識処理を適用し、手軽にコンテキストウェア・システムに活用するようなことは困難である。

我々は、時系列信号を伴う各種センサ情報から、イベントとして検出したい信号パターンを、センサを設置する現場で手軽に学習できる Instant Learning Sensor を提案する。提案手法は、実演されたイベント信号を自動的に解析し、その信号パターンの認識に適した特微量抽出処理や認識アルゴリズムを自動的に選択し、組み合わせ、認識処理を試行し、認識率と誤認識率を評価する。そして、最も性能の良い認識処理を最適なものとして、ユーザに提示する。本手法を用いることにより、ユーザは信号処理プログラミングを用いることなく、信号処理を用いた高度なコンテキスト情報が取得可能になる。

上記の手法に基づく最初の Instant Learning Sensor として、生活音や環境音の認識に特化した Instant Learning Sound Sensor を開発した。まず、多様な物音に対応するために、様々な物音の特微量や認識アルゴリズムについて検討し、多様なパラメータ設定を持つ認識処理を設計した。本認識処理は、多様なパラメータを持つため、様々な物音に応じて調整可能である。しかしながら、イベント音学習時に、全パラメータを試行することは計算量の点から非現実的である。これを解決するために、本研究では、あらかじめ多種多様な物音をサンプル音データベースとして構築した。各サンプル音に対して、最適なパラメータ設定をテンプレートとして与えておき、未知のイベント音を学習する際に、いくつかのテンプレートを自動選択し、パラメータ決定に利用する。

実際に、家具や日用品に貼付けた圧電マイクによって、コーヒーカップにコーヒーを注ぐ音や、ドアノブを回す音、水道から水が流れる音などを、80%を超える認識率が得られる処理を自動生成可能なことが確認できた。本手法を用いることにより、物音認識を用いた高度なコンテキスト情報を、誰でも手軽にコンテキストウェア・システムに組み込むことができる。

## 2. Instant Learning Sound Sensor

### 2.1 システム概要

我々が提案する Instant Learning Sound Sensor は、次の3つの特徴を持つ。

(1)Instant Learning: センサ取り付け箇所にて、検出対象としたイベントを開発者が実演することにより、提案センサに対象信号パターンを即興的に学習させることが可能。

(2)Smart Sensor: 提案センサは信号処理を単体で処理可能。

(3)Simple Device: 将来的に、提案センサを dsPIC のように小型・低コスト・デバイス上に実装し、小型ネットワークセンサ化すること視野に入れている。そのために、可能な限り軽量の認識処理にて実現する。また1つのセンサに1つ、多くて2,3種類のイベント・パターン専用の信号解析処理を手軽に設定可能な、非常にシンプルなセンシング・コンポーネントを提供する。

Fig.1 にシステム構成を示す。提案システムは検出したいイベント音に最適な認識処理を自動生成する Sensor Configurator と、実際に環境に設置され、生成された認識プログラムを実行する Sound Sensor Device から構成する。

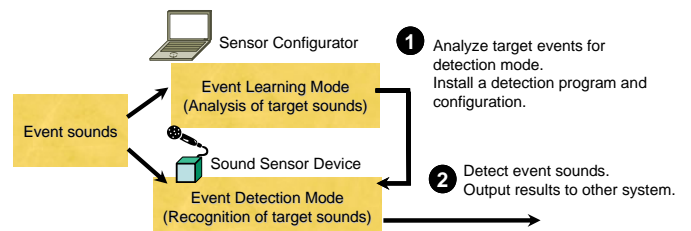


Fig.1. System Overview

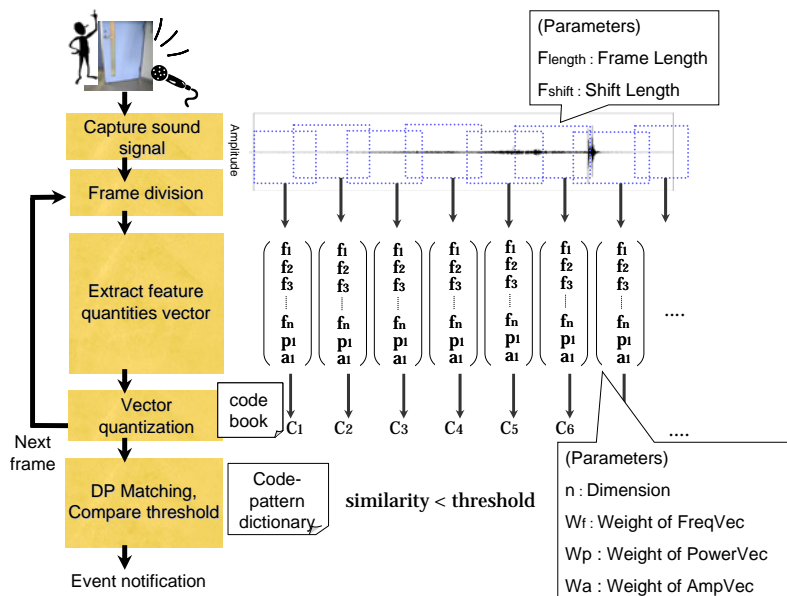


Fig.2. Basic Sound Recognition Process

## 2.2 生活音・環境音認識処理

続いて、具体的に幾つかの生活音を実際に観測・分析し、考案した音認識処理をFig.2に示す。今回用いた音認識処理は、振幅特性と周波数特性、パワーの相対的变化の特性に関して、時系列変化を時間伸縮パターンマッチング[3]により検出判定する。図中のパラメータを調整することにより、認識対象音に応じて、特徴量の重み付けなどを調整可能である。なお、省メモリ化・計算量の削減のために、ベクトルの次元数やコードブックサイズなどを小さい値に抑えることが望ましい。また、既存の環境音認識と異なり、本研究では未知のイベント音を認識するための処理を自動的に生成することを目標としているため、対象とする音によってどのような特徴量やパラメータが最適であるか不明である。そのため、Instant Learningを実現するためには、効率的な最適パラメータの決定手法が必要と考えられる。

## 2.3 未知のイベント音学習

前節の課題を踏まえ、提案システムでは、あらかじめ多様な環境音・生活音(Sample Sound)を収集し、それぞれの音に適した軽量な処理を実現するパラメータ集合(Parameter Template)を Sample Sounds DBとして構築した。Sensor Configurator は、Sample Sounds DBを用いて、未知の検出対象音に適したパラメータを推定する。手順は次の通りである。(Fig.3)

- (1)ユーザより入力された認識対象音と類似する SampleSound の上位候補を取得。
- (2)取得した Sample Sound の Parameter Template ごとに従い、コードブックやパターン辞書を生成。
- (3)ユーザよりテスト入力を受け、時間伸縮マッチングの閾値を決定。
- (4)各 Template による処理の誤認識率を、他の Sample Sound を用いて評価し、最良結果の Template を用いたイベント音認識処理をユーザに提示し、Sound Sensor Device にインストール。

このように本手法により、ユーザは対象音を数回分用意するだけで、手軽に音認識を利用できる。

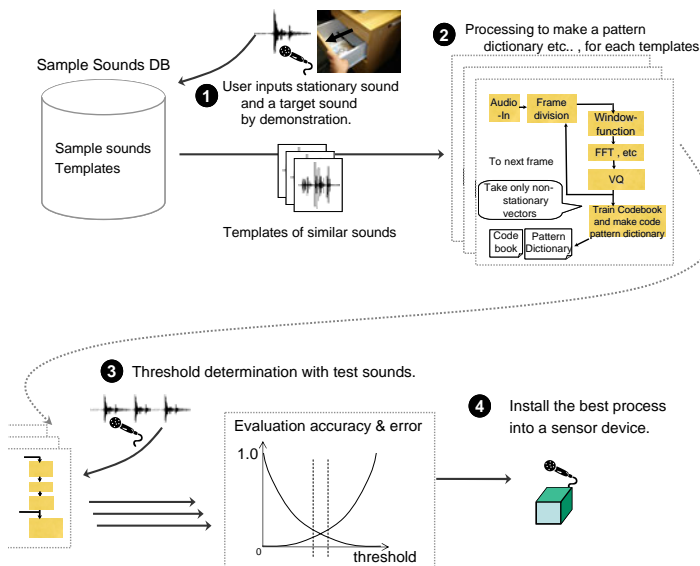


Fig.3. Instant Learning Process using Sample Sounds DB

## 3. プロトタイプ

実際に、PC 上に Sensor Configurator と Sound Sensor を実装した。スクリーンショットを Fig.4, 5 に示す。本プロトタイプでは、マイクロフォンとして圧電素子を利用した。その理由は小さく薄く、家具や日用品など、いたる箇所に貼付け可能なためである。サンプリング周波数は 16KHz、サンプリングビットは 16bit とした。

Sample Sounds DB としては、30 種類のイベント音と、各イベントに約 50 音を収集した。そして人手により最適パラメータを付与し、Parameter Template を作成した。

さらに、IP ネットワーク経由にて、Sound Sensor と通信するための Windows API ライブラリとして、Sound Recognition Toolkit を実装した。

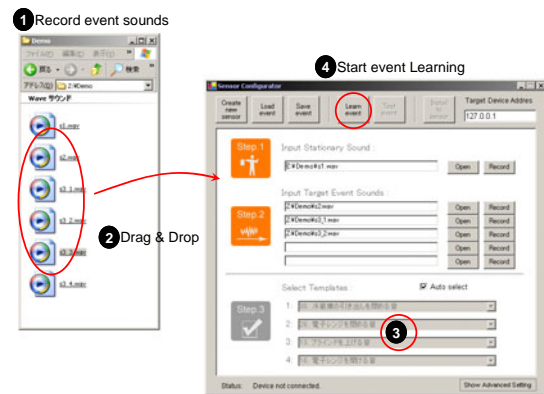


Fig.4. Sensor Configurator

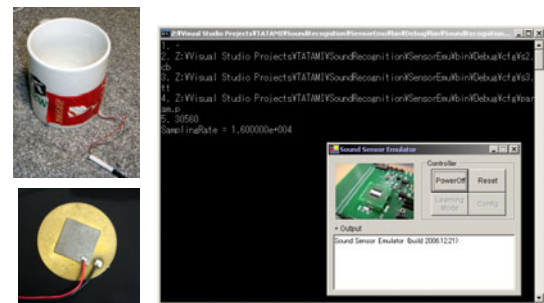


Fig.5. Sound Sensor

## 4. まとめと今後の予定

本稿では、実世界の物音をコンテキスト情報として、手軽に利用可能にする Instant Learning Sound Sensor を提案した。提案システムにより、イベント音に適した認識処理を自動的に生成することが可能である。今後の課題としては、より多くの認識アルゴリズムの追加、Sample Sounds DBの大規模化、低コスト DSP マイコンへの実装の検討が挙げられる。

## 参考文献

- [1] Crossbow Technology, MOTE, <http://www.xbow.com/Products/Wireless%20Sensor%20Networks.htm>
- [2] Jianfeng Chen, Alvin Harvey Kam, Jianmin Zhang, Ning Liu, Louis Shue, "Bathroom Activity Monitoring Based on Sound", Pervasive 2005, 2005.
- [3] 古井貞照：音声情報処理，森北出版，1998.