

入力モード推定に基づく 作業環境音最適化アンビエントインテリジェンスシステム

三井健史^{1,a)} 濱川 礼²

概要: 本研究では, ユーザの作業用 PC に振動マイクと振動スピーカーを取り付けるだけで集中度推定し個々の状態において最適な環境演出を行うシステムについて述べる. 我々は PC 作業者のキーロギング・姿勢・視線情報からユーザの状態を推定し, 環境音再生やコーヒーやメントールのアロマオイルを噴霧することで作業環境を最適化するシステムを検討してきた. しかし従来法の作業状態推定手法では入力項目が多くセットアップが煩雑であるといった問題があった. そこで本研究では, アクティブ音響センシングを参考にした入力モード推定を行うことでユーザの作業用 PC に振動マイクと振動スピーカーを取り付け, 作業者の手が PC とどのように接触しているかを取得し集中度推定を行う試みを提案する. さらに, 集中度推定結果から作業用 BGM の音量を最適化することで, 作業効率向上を図る.

BGM Volume Optimize Ambient Intelligence System Based on Estimating Input Mode

TAKEFUMI MITSUI^{1,a)} REI HAMAKAWA²

Abstract: This paper describe a BGM volume optimize system based on estimating user's concentration using vibration microphone and vibration speaker. We have examined the system to optimize the working environment by playing the environmental sound playback and spraying the coffee and menthol aroma oil based on estimating user's concentration. However, The conventional method is a problem that complicated to setup. Therefore, in this paper, we propose attempt to estimate user's concentration by performing an input mode estimation with the active acoustic sensing. Moreover, By optimizing the BGM volume from the degree of concentration estimation results, to achieve a working efficiency.

1. はじめに

人間の行動状態やその空間をセンシング・解析・推定し, 個々の状況に応じて環境を変化させたり, 情報提示を行うことで人間の活動を支援する試みを総称してアンビエントインテリジェンス (Ambient intelligence) [1] と呼ぶ. 人間の行動状態とは, 頭・手の動きや姿勢・作業内容・デバイス使用頻度などの人間の行動に関する状態を指したり, 集中度, 眠気などの人間の心理に関する状態を指し, それら

を様々なセンサデバイスを用いてセンシング・解析・推定することで, ユーザに対して適応的な環境演出を行い, 作業効率の向上を目的としたシステムが考案されて来ている. 作業者の状態から作業効率向上を目的としたシステムとして, 我々は PC 作業者のキーロギング・姿勢・視線情報からユーザの状態を推定し, 環境音再生やコーヒーやメントールのアロマオイルを噴霧することで作業環境を最適化する試みを検討してきた [2][3]. しかし, システム実用上, 作業前にシステム動作に必要なアプリケーションを起動し, 視線追跡の為にキャリブレーション処理を行い, 姿勢推定のために事前に指定アプリケーションを起動させたスマートフォンを胸ポケットに入るなど, 集中度推定のための入力項目が多く, 煩雑であるという課題があった.

そこで本研究では, ユーザの作業用ノート PC に振動

¹ 中京大学 情報科学研究科
Graduate School of Information Science and Technology,
Chukyo University

² 中京大学 工学部
School of Engineering, Chukyo University

^{a)} h11516m@st.chukyo-u.ac.jp

マイクと振動スピーカーを取り付けるだけで集中度推定し個々の状態において最適な環境演出を行うシステムについて述べる。

2. 関連研究

集中度推定には様々なアプローチが考案されている。昨今では各種センサデバイスが非常に安価で手に入る為、脳波計などの大掛かりで非常に高価なセンサを使用せずにユーザの集中度を推定することが可能となっている。以下に安価なセンサデバイスを用いた集中度推定手法を幾つか述べる。

小林ら [4] は Web カメラを用いた集中度推定手法を提案している。この研究ではユーザの顔を Web カメラで撮影し、顔中心座標位置の x 軸・ y 軸・ z 軸の変化量をそれぞれ集計し、その移動特徴から集中度推定を行うものである。評価実験ではタイピング作業時の進捗状況の変化を集中状態の変化と位置づけて行っており、非集中時の作業時には図 1 の円で囲まれた例のような特徴的な波形となる事を示している。また、脳波計などを用いる集中度推定方法と異なり、非接触センサデバイスを用いることでユーザへの圧迫感を抑えるなどしている。

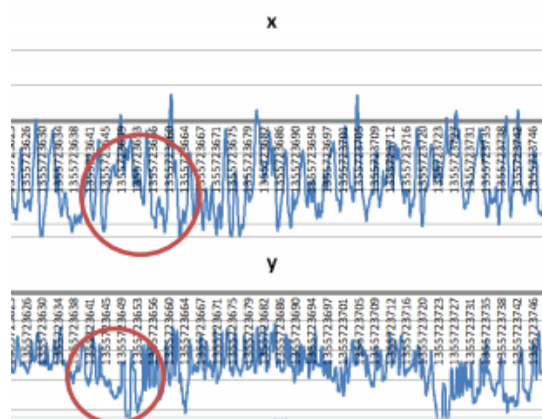


図 1 顔座標位置プロット

濱口ら [5] の研究では、教育支援システムのために作業者の机背面部に設置した振動センサから、筆記作業時の卓上振動音をアンビエントにセンシングし、筆記作業時と思考時にリアルタイムに分類する手法を提案している (図 2)。さらにユーザの心理に影響を与える主要因をタスクの難易度と仮定し、難易度の異なる記述問題を課す被験者実験を行い、その結果から卓上振動音のうち非筆記音の占める割合からタスクに対する主観的難易度の推定の可能性まで見出している。

2.1 関連研究との関連

センシングデータから集中度推定を行う研究は上記に挙げる通りいくつか存在し、本稿の集中度推定手法は排他的



図 2 筆記作業のセンシング環境

なものではなく並列に処理することが可能である。

一方、作業者の集中度推定結果に基づいた環境演出まで含めアンビエントインテリジェンスを構築している研究例は少ない [2][3][6]。これは、オフィスなどの複数人が介在する環境下でのさまざまな作業を想定した場合、フロア全体の騒音や作業者会話の有無など、作業への集中度を推定する上で考慮すべき要因が数多くある為、集中度推定手法の提案のみの研究が多い為と考えられる。

そこで本研究では、問題をシンプルにするため、個人利用かつ PC 作業を前提とする。こういった状況下での作業者の集中度推定結果に基づき、環境演出として再生されている BGM の音量を最適化し、さらにこの処理サイクルを単位時間毎に継続して行うことで作業効率の向上を目標とする。

3. 提案手法

本研究では前述の通り個人のノート PC 作業者を前提として行う。

従来 PC 作業時の集中度推定にはキーロギングなどを用いて行ってきた [2][4][6]。しかし、アルゴリズム思案時など考え事をしている時にはキーボードの打鍵が行われなない為、キーロギング情報からユーザの行動を見たとき、何の作業も行っていない状態との識別は困難であった。

そこで本研究では作業時に作業者の手がどのように PC と接触しているかに着目し、物体に音響信号を流しその伝搬応答から物体の状態を認識する「アクティブ音響センシング」[7] 技術を参考に問題を解決する手法を用いるアプローチを取る。

図 3 に示す通り、作業者の PC に小型の振動センサーと振動マイクを設置する。設置した振動スピーカーから正弦波を流し、振動マイクから振動をセンシングすることで PC から作業者が手を離している状態 (none モード) か、手のひらが PC のキーボードに触れている状態 (home モード) か、タイピング作業が行われている状態 (typing モード) の三つの入力モードをマイクから取得した振動情報の周波数解析を高速フーリエ変換 (Fast Fourier Transform: FFT) を用いて行った結果を SVM でクラス分類した後、モード分類結果に基づいて単位時間に占める各モードの割

合から作業者の集中度推定を行う．最終的に環境演出として作業時にユーザが再生している BGM の音量を集中度推定の結果に基づき最適化する．

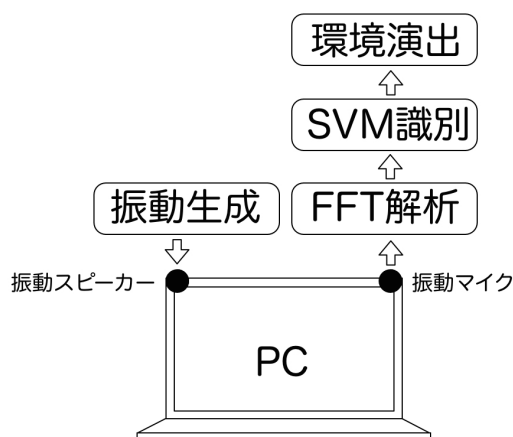


図3 システム構成

本手法では単純なキーロギングと比べ、ユーザのキー入力がないような（キーボードに手を添えて考え事を行っている等）思考状態も検知する事が可能である．さらにキーロギング方式ではユーザが作業している PC に専用のソフトウェアを導入する必要があるのに対し、提案手法ではその必要がないなどの利点が挙げられる．

そして、モード分類結果に基づいて作業者の集中度推定を行い、環境演出として作業時にユーザが再生している BGM の音量を最適化する．

4. 実装

提案手法のプロトタイプを実装した．振動スピーカとしてコスモテクノ社の VIBROY CSP-V101 [8] を、振動マイクにはアーテック社のピエゾピックアップ A1-OSJ [9] を使用した．実際の設置例を図4に示す．



図4 実際のセンサ設置例

振動マイクの入力データを PC を転送するための USB オーディオインターフェースとして Roland 社の DUO-CAPTURE mk2 UA-11-MK2 [10] を使用した．

5. システム詳細

提案手法は入力モード推定処理部・集中度推定処理部・作業用 BGM 音量調整処理部の三つの処理部が連動し動作している．

5.1 入力モード推定処理部

入力モード推定処理部では作業者が PC から作業者が手を離している状態（none モード）か、手のひらが PC のキーボードに触れている状態（home モード）か、タイピング作業が行われている状態（typing モード）かの三つの入力モードの分類を行う．

振動マイクから取得した振動データをもとに、フレーム単位に FFT を行い、周波数ごとのパワースペクトルを算出し、それを各モードの特徴とする（図5）．

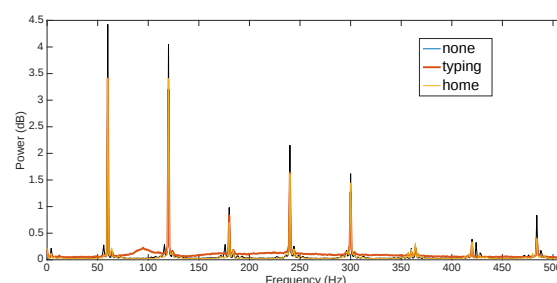


図5 各入力モードの周波数パワースペクトル

この周波数ごとのパワースペクトルから入力モード推定を行うための機械学習手法としては、決定木、ランダムフォレスト、サポートベクターマシン（Support Vector Machine: SVM）、ニューラルネットワークなどさまざまな方法が検討されるが、本研究では SVM を利用し入力モード推定処理を行う．

SVM は、学習サンプル $x = (x_1, x_2, \dots, x_n)$ とその教師信号 y 、重みベクトル w 、バイアス項を b とした時、二つのクラスに分離されるような超平面 $\langle w, x \rangle + b = 0$ を、その周辺のマージン領域が最大となるように選択することで、汎化性能がより高くなるような識別境界を求める手法である．ここで先の処理で得られた各入力モードごとの周波数スペクトルデータ学習データとし SVM の識別器を学習させる．

また、今回は分類を 3 つの状態に分類するため、typing モードを判定する識別器と、home モードを判定する識別器の 2 つを組み合わせ処理を行う．

5.2 集中度推定処理部

集中度推定処理には、入力モード推定処理より得られた三つの入力モードが、単位時間に占める平均割合を算出し、状態推定を行う作業者の各入力モードの単位時間に占める

割合との差が小さい方の状態を現在の作業者の状態として判定する処理を行う。故に、出力結果は作業者が単位時間内で集中状態であるか、非集中状態であるかの二値結果が得られる処理を行う。

5.3 作業用 BGM 音量調整処理部

作業者が集中状態にあるとき、作業用 BGM の音量を単位時間内で徐々に下げる処理を行う。

6. 評価

提案手法の可能性を検討するため、ユーザの入力モードの推定を行う実験を行った。作業者からホームポジションに手を添えた状態 (home モード) と、タイピング作業を行う状態 (typing モード) と、作業者が PC に何も触れていない状態 (none モード) の三つの入力モードのデータ各 100 秒間のデータを学習データとし入力モード推定精度の調査を行った。

PC 設置位置と振動センサ設置位置を固定した学習データから、同様の設置環境での未知データに対して精度評価を行った実験結果を図 6 に示す。各行は正解データを示し、列は判別結果を示す。

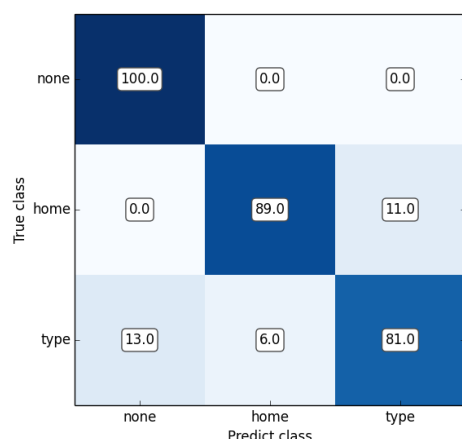


図 6 入力モード推定結果の混交行列

実験結果から入力モード推定精度は非常に高い精度で推定可能なことがわかった。しかし、学習データ取得時での PC 設置位置と振動センサ設置位置から、異なる状況に設置された未知データに対して精度評価を行ったところ、こちらでは入力モード推定精度は大幅に落ちることも判明した。これは振動センサの微妙な位置変化や卓上の状態によっても伝搬応答に変化があるためであると考えられる。このことから、提案手法では PC 設置毎にキャリブレーション処理を行う必要がある事がわかった。

7. 展望

今回は入力モード推定処理の妥当性を評価する段階まで

進めたが、今後は home モードの判定状態が作業者の集中とどのような関連があるかの実験を行い、集中度推定処理の評価や主音量調整機能がユーザの集中度にどのような影響をあたえるかの調査を継続して行う。

さらに入力モード推定処理について、現状の手法では PC 設置毎にキャリブレーション処理が必要となるが、判定方法を工夫することで解決を目指す。

また、不確定な時系列データに対して統計的なアプローチからモデル化する手法として隠れマルコフモデル (Hidden Markov Model: HMM) がある。現状集中度推定処理は今後は単位時間内でのユーザの行動のみのデータを処理するのではなく、HMM などの時系列を考慮したモデルを適用することで時系列でのユーザの状態を加味した集中度推定手法をシステムに組み込んでいく。

参考文献

- [1] Diane J. Cook, Juan C. Augusto, Vikramaditya R. Jakkula: "Ambient intelligence: Technologies, applications, and opportunities": Pervasive and Mobile Computing 5(4), pp.277 - 298 (2009)
- [2] 三井健史, 大橋省吾, 小林拳人, 志津野之也, 濱川礼: "個々のユーザの集中度に応じた室内環境を演出するシステム ~ ヴァーチャル・スタバ ~": 第 77 回全国大会講演論文集 2015(1), pp.653 - 655 (2015)
- [3] 三井 健史, 濱川 礼: " ~ ヴァーチャル・スタバ ~ 作業状態に応じた室内環境演出システムの開発と評価": マルチメディア, 分散, 協調とモバイル (DICOMO2015) シンポジウム, pp.1424 - 1435 (2015)
- [4] 小林竜司, 中田豊久: "頭部の動きに着目した非接触型の集中度測定方法の開発": 第 75 回全国大会講演論文集 2013(1), pp.469 - 470 (2013)
- [5] Nana Hamaguchi, Keiko Yamamoto, Daisuke Iwai, Kosuke Sato: "Subjective Difficulty Estimation for Interactive Learning by Sensing Vibration Sound on Desk Panel": Lecture Notes in Computer Science, Springer 6439 pp.138 - 147 2010
- [6] HONDA Shinkuro, TOMIOKA Hironari, KIMURA Takaaki, OOSAWA Takeharu, OKADA Kenichi, MATSUSHITA YUTAKA: "A Home Office Environment Based on the Concentration Degrees of Workers -A Virtual Office System Valentine(Special Issue on Next Generation Human Interface and Interaction)": Information Processing Society of Japan (IPSJ) pp.1472 - 1483 (1998)
- [7] Ono Makoto, Shizuki Buntarou, Tanaka Jiro: "Touch & Activate: Adding Interactivity to Existing Objects Using Active Acoustic Sensing": Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology, pp.31 - 40 (2013)
- [8] パイプレーションスピーカー VIBROY CSP-V101: <http://www.cosmotek.co.jp/sp/seihin-joho-CSP-VI01.html>
- [9] ARTEC(アーテック)/A1-OSJ: <http://www.soundhouse.co.jp/products/detail/item/201191/>
- [10] DUO-CAPTURE mk2 UA-11-MK2: http://www.roland.co.jp/products/duo-capture_mk2/