

モバイル端末のセンサ情報を用いた飲酒状態認識

男澤 晃陽[†] 服部 峻[†]

[†] 室蘭工業大学 〒050-8585 北海道室蘭市水元町 27-1

E-mail: [†]10024041@mmm.muroran-it.ac.jp, hattori@csse.muroran-it.ac.jp

あらまし アルコールを摂取する機会が多いが、飲酒行為には様々なリスクが伴う。これらのリスクは現在の飲酒量を把握し、これ以上の飲酒が危険であると自己認識を行うことである程度回避することができる。これらのリスクを回避するため、飲酒量認識や飲酒状態認識が使われている。しかし、従来の飲酒状態認識は酩酊して脳が麻痺し平常時とは異なる状況下にあるユーザに対して、息を吹きかける、飲酒内容を手で入力するといった自発的行動を必要としている。そこで本稿ではアルコールによって酩酊してしまったユーザを、ユーザの自発的な行動ではなく酩酊時特有の症状をモバイル端末のセンサ情報（特に加速度センサや音センサ）を監視して認識し、家族や警察に連絡や通報、周囲の人に警報を発することで、アルコールによる被害から身の安全を見守るシステムを提案する。

キーワード 飲酒, センサ情報, コンテキスト認識, モバイル端末

Drinking State Recognition Using Sensor Data of Mobile Devices

Kouyou OTOKOZAWA[†] and Shun HATTORI[†]

[†] Muroran Institute of Technology 27-1 Mizumoto-cho, Muroran, Hokkaido 050-8585, Japan

E-mail: [†]10024041@mmm.muroran-it.ac.jp, hattori@csse.muroran-it.ac.jp

Abstract We have many opportunities to drink alcohol, but it is accompanied by various risks. These risks can be avoided to some extent by getting to know about our current amount of drinking and being aware by ourselves that it is dangerous for us to drink any more. Drinking Amount Recognition and Drinking State Recognition are used to avoid these risks. However, most of conventional Drinking State Recognition systems require users, who are in an unordinary state that they are drunk and paralyzed on their brain, to do their voluntary actions such as breathing and manually inputting about their drinking. Therefore, this paper proposes a novel Drinking State Recognition system that watches for such users and protects their safety from damages by alcohol. The system does not require them to do their voluntary actions, but observes sensor (especially accelerometer and microphone) data of mobile devices, recognizes their particular symptoms when drunk, and then alerts their family, police, and/or their surrounding people.

Key words Drinking, Sensor Data, Context-Awareness, Mobile Devices

1. ま え が き

厚生労働省の調査によると、1年間に1度以上飲酒する男性は83.1%、女性は60.9%にのぼり[1]、多くの人アルコールを摂取する機会にあると言える。アルコールは飲料として歴史が長く、身近な存在であり入手が非常に容易であるが、アルコールで酔っている状態は脳が麻痺している状態であるため平常時とは違う状態であることに留意しなければならない。飲酒量が増加することで脳の麻痺も進行するが、脳の麻痺が進むと抑制が取れなくなる、運動能力が低下するなどの症状により自分で自身の安全を確保できなくなるという問題点がある。

飲酒量を推定する研究はいくつか存在する。その1つに快適な飲酒を支援するコースター型飲酒検知デバイスの開発[2]というものがある。この研究では、コースター型デバイスの上に置かれたグラス内の酒量の変化を測定し飲酒量を提示している。そのため、飲酒量の提示にコースター型デバイスを必要としている。一方で、総務省によるとスマートフォンの世帯普及率は平成25年度末時点で62.6%[3]であり、スマートフォン等のモバイル端末は少なくともコースター型デバイスよりは普及していると考えられる。

モバイル端末において飲み過ぎを警告するサービスもまた存在している。例えば、飲酒量計算アプリは自分で飲んだ酒量を

入力することで摂取したアルコール量を計算し表示する。また、モバイル端末に接続可能な呼気検知器を持ち歩き、息を吹きかけ血中アルコール濃度を測定する方法もある。これらの方法での飲酒量検知にはユーザの自発的な行動を必要としている点が共通している。しかし、脳が麻痺している状態、つまり酩酊している状態はユーザが平常通りの行動を取ることができなくなっているからこそ問題が発生し得るため、ユーザが自分で酒量入力や息の吹きかけを行わない、行えない可能性がある。

そこで本稿では、アルコールによって酔ってしまった、特に一般的に酩酊状態と呼ばれ大脳だけでなく小脳まで麻痺し始める、アルコールを4単位（1単位20g）以上摂取[4]、もしくは以下の血中アルコール濃度と酔いの条件式[5]より

$$\frac{\text{飲酒量 (ml)} \times \text{アルコール度数 (\%)}}{833 \times \text{体重 (kg)}} \geq 0.16$$

を満たすユーザを、ユーザの自発的な行動ではなく酩酊時特有の症状をモバイル端末のセンサ情報（特に加速度センサや音センサ）を監視して認識し、家族や警察に連絡、通報を行うこと、周囲の人間に警報を発することでアルコールによる被害から身の安全を見守るシステムを提案する。

2. 提案手法

2.1 システム概要

本稿における飲酒状態認識システムは前述の通り、酔った人間に発生する平常時との特有な違いを認識する。飲酒状態の認識には、平常な人間は酔った振りを行うことはできるが、酔った人間は平常な振りを行うことができないという生理的特徴と、アルコールにより脳が麻痺する時にもアルコールの量により麻痺する脳の部位が異なる、特に血中アルコール濃度が高くなった時にはじめて運動機能に影響が出るという生理的特徴[4]を利用して認識を行う。

具体的には、人間は酔うとまっすぐ歩けなくなる点と人間は酔うと声に変化してしまう点を、それぞれ加速度センサと音センサに基づき判定をする。加速度センサに基づく飲酒状態認識は、平常時に歩行した際に発生する加速度をあらかじめ学習しておき、学習した加速度より大きな加速度を検知した時に酩酊していると判定する。

酩酊時に平常時よりふらふらにならずまっすぐ歩けるようになる人は存在しないが、声は酩酊時に平常時より大きくなる人と逆に平常時より小さくなる人が居る。そこで、音センサに基づく飲酒状態認識は、モバイル端末で録音と Android 標準 API による高速フーリエ変換を行い、酩酊時特有の振幅のばらつきを検知できた時に酔っていると判定する。

2.2 加速度に基づく飲酒状態認識

本節の処理の流れを図1に示す。ユーザがモバイル端末を身につけ、歩行を行うと加速度センサが端末の加速度をその都度記録していく。まず事前に学習フェーズとして、平常時に歩行した際に発生する加速度と加速度の分散を取得する。この取得した時系列の加速度ベクトル長データ $A[i]$ の要素について、極大値のみ、つまり、 $A[i] > A[i-1]$ かつ $A[i] > A[i+1]$ を満たす加速度値のみを記録する。記録した加速度値の中から α_1 番

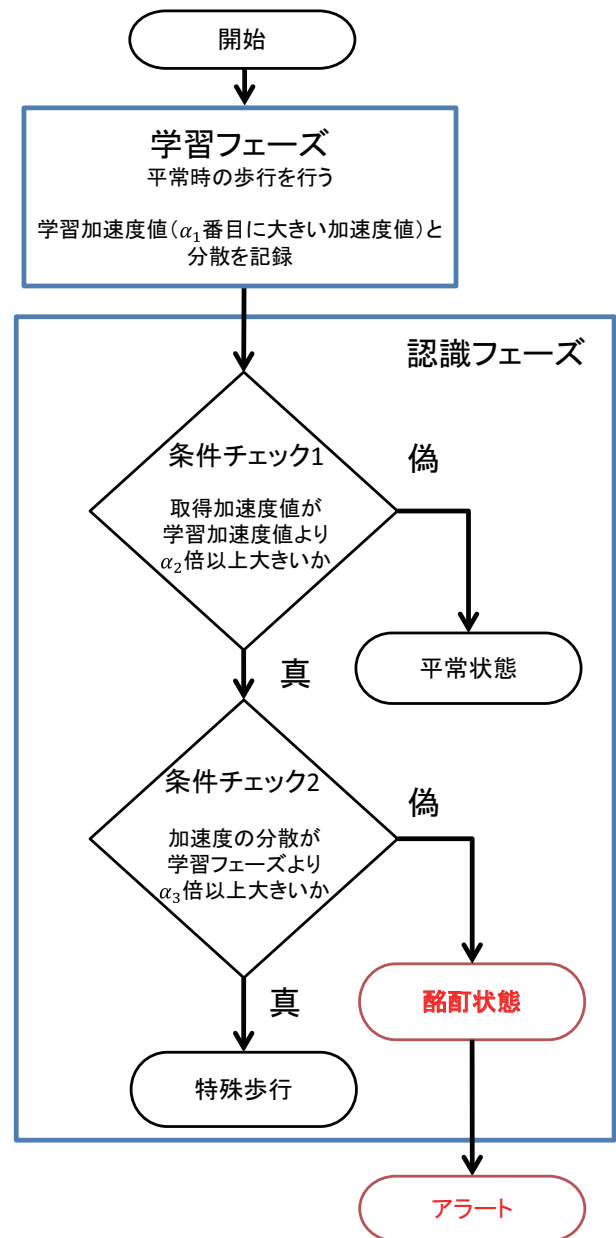


図1 加速度に基づく認識の流れ

Fig. 1 An overview of the acceleration-based recognition.

目に大きい加速度値を「学習加速度値」とする。

学習フェーズ終了後、認識フェーズに移行し、認識フェーズで本節では例として10秒分取得した加速度系列を学習フェーズと同様に $A[i] > A[i-1]$ かつ $A[i] > A[i+1]$ を満たす加速度値のみを記録する。そして、 α_1 番目に大きい加速度値を学習加速度値と比較して α_2 倍よりも小さい場合にはユーザが平常状態にあると認識し、 α_2 倍以上の場合にはユーザが平常状態ではない（酩酊状態または特殊歩行状態である）と認識する。

本節の認識フェーズの条件チェック1は加速度値の比較を行うのみであるので、ユーザが走行した場合も加速度値が非常に大きくなるため酩酊状態であると認識してしまう。そこで、走行時は平常状態歩行時、酩酊状態歩行時よりも加速度の分散が異常に大きくなることを利用する。認識フェーズの条件チェック1で平常状態ではないと認識した時に、加速度の分散が学習

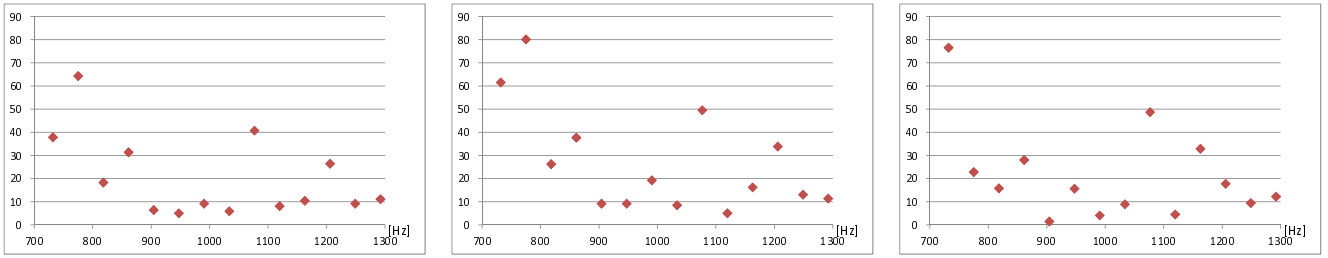


図2 Android 標準 API によって音声を高速フーリエ変換 (FFT) した結果の一部

Fig.2 A part of FFT results of voice by Android API.

フェーズ時より α_3 倍以上大きくなっていた場合、ユーザは酩酊状態ではなく特殊歩行状態であると判定する。

2.3 音声の周波数に基づく飲酒状態認識

ユーザがモバイル端末に向かって発した音声を音センサが取得する。本稿の実験では端末に向かって発した音声はすべて「あ」の単音のみに統一し、録音を行っている。録音した音声を Android 標準 API で高速フーリエ変換 (FFT) を行うと、図 2 のように 43Hz 刻みで同じ周波数毎の振幅が得られる。本節では、事前に酩酊時に声が大きくなる男性 (第 1 著者) と酩酊時に声が小さくなる別の男性の平常時と酩酊時の音声の高速フーリエ変換結果を比較した結果、特に 700Hz から 1300Hz の間で両男性共に大きな特徴が現れていたため、飲酒状態認識に用いる音声の周波数の範囲を 700Hz から 1300Hz に限定する。

図 3 は図 2 のような高速フーリエ変換の結果を 1 つにまとめたものであり、本稿ではある周波数 f について得られた、本節では例として 5 秒分の振幅データ列を昇順ソートした配列を配列 $V_i^{(f)}$ と表す。一方、酩酊時の音声 (図 4) は平常時の音声 (図 5) と比べ、音声周波数の振幅の配列 $V_i^{(f)}$ の要素間で値の差が大きくなっている周波数 f が特に 700Hz から 1300Hz の間に多く見られた。したがって、周波数 f ($700 \leq f \leq 1300$) 毎に $V_i^{(f)}$ をチェックして $V_i^{(f)} - V_{i-1}^{(f)}$ が 1 度でも閾値 β_1 より大きくなる周波数 f の数が 14 回中に β_2 回以上である音声を音センサが取得した時、ユーザが酩酊状態であると認識する。

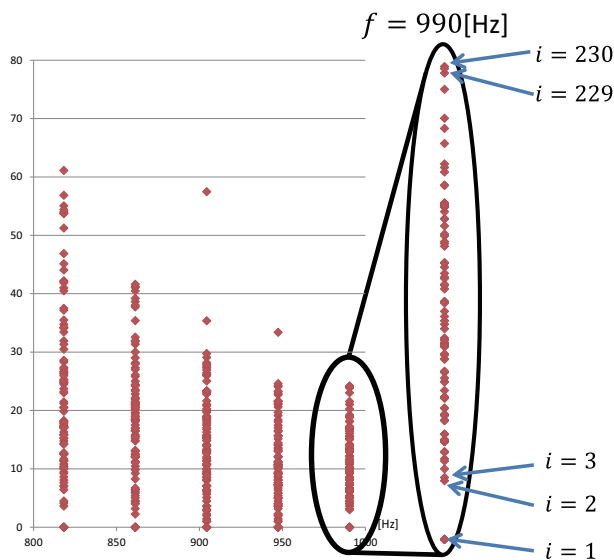


図3 音声周波数 (振幅) の配列 $V_i^{(f)}$

Fig.3 Array of voice frequency (amplitude), $V_i^{(f)}$.

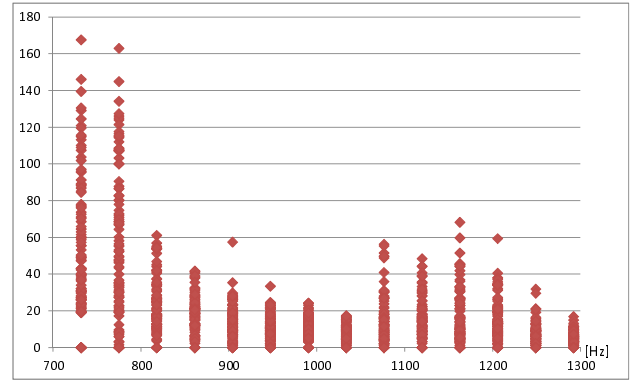


図4 酩酊時の音声周波数 (振幅)

Fig.4 Drunk voice frequency (amplitude).

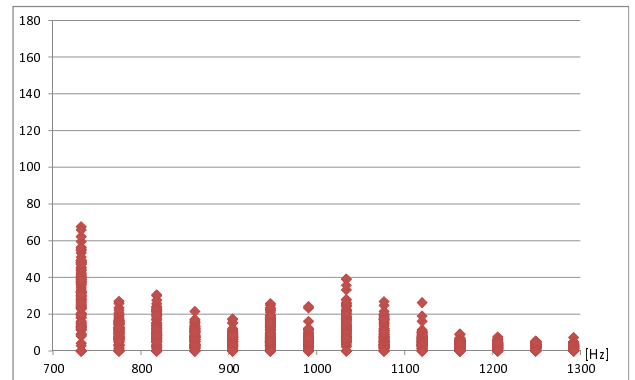


図5 平常時の音声周波数 (振幅)

Fig.5 Ordinary voice frequency (amplitude).

3. 評価実験

本章では、第 2 章で提案した加速度に基づく飲酒状態認識におけるパラメータの最適化、及び、音声の周波数に基づく飲酒状態認識におけるパラメータの最適化に関する評価実験を行う。なお、本実験において提案システムを使用する被験者は 1 名 (第 1 著者) のみであり、空腹状態で直近 30 分以内の飲酒量 (ml) が血中アルコール濃度と酔いの条件式を満たした場合のみを酩酊状態とした。評価尺度としての認識精度は、酩酊である状態の認識率 (実際に酩酊状態の時にシステムが酩酊状態と正しく認識した確率)、及び、酩酊でない状態の認識率 (実際に酩酊でない状態の時にシステムが酩酊でない状態と正しく

認識した確率)に基づき、以下の式で線形結合したものを使う。

$$\begin{aligned} \text{認識精度} &= \gamma \times \text{酩酊である状態の認識率} \\ &+ (1 - \gamma) \times \text{酩酊でない状態の認識率} \end{aligned}$$

本稿の実験では、酩酊である状態の認識率をより重視して、線形結合パラメータ γ を 0.6 に設定している。

3.1 加速度に基づく飲酒状態認識のパラメータ最適化

実験時は加速度センサが内蔵されており、かつズボンのポケットに収納可能なサイズである Android 端末（本実験では Sony 社の Xperia Z1）で、加速度の取得を行った。計測場所は直線の廊下とし、端末の x 軸方向の加速度を X 、 y 軸方向の加速度を Y 、 z 軸方向の加速度を Z とし、3 軸方向の二乗和の平方根 $\sqrt{X^2 + Y^2 + Z^2}$ を端末の加速度ベクトル長として最長でも 60 ミリ秒間隔毎に取得する。

本来ならば学習加速度値を取得するために、平常状態で 1 回だけ歩行すれば十分であるが、今回は実験用として学習加速度値のために平常状態で 50 回の歩行を行った。50 回分の学習加速度値の中央値を、本実験で使用するそれぞれの学習加速度値とした。学習加速度値とは別に平常時と酩酊時にそれぞれ 35 回歩行し、特殊歩行として 10 回歩行し、合計 80 個の加速度データを評価データとして使用した。

第 2.2 節で提案した加速度に基づく飲酒状態認識手法において、以下のような 3 種類のパラメータが存在する。

- 極大値の α_1 番目を学習加速度値とする
- 条件チェック 1 で学習加速度値の α_2 倍以上
- 条件チェック 2 で加速度の分散値が α_3 倍以上

本節では、これらのパラメータを最適化し、加速度に基づく飲酒状態認識の最良の認識精度を求める。

パラメータが取り得る範囲は、

- $1 \leq \alpha_1 \leq 10$ (1 刻み)
- $1 \leq \alpha_2 \leq 2$ (0.001 刻み)
- $1 \leq \alpha_3 \leq 10$ (0.1 刻み)

とする。パラメータを変動させた結果、 α_1 は 3 から 6 の辺りで、 α_2 は 1.1 の辺りで、 α_3 は 2.9 以上で最良の認識精度として 0.982 が得られた（酩酊である認識精度は 0.971、酩酊でない認識精度は 1.000）。なお、酩酊でない状態は平常状態の歩行か特殊歩行のどちらかである状態を指す。

α_1 が 3 から 6 における認識精度のパラメータ α_2 と α_3 への依存性を図 6 から図 9 にそれぞれ示す。また、 α_1 の最適なパラメータは複数存在するが、認識精度が高くなる α_2 を 1.1、 α_3 を 3.0 に固定して α_1 を変動させた。すると、図 10 のように α_1 が 1 から 5 の傾きは緩やかであり、6 から 10 の傾きは急である。したがって、図 10 において認識精度が最も高い α_1 は 5 と 6 であるが、5 の方が 6 より適切であると判断できる。

一方で、図 12 と図 13 のように α_1 をそれぞれ認識精度が最も高い 5 と 6 に固定して（共に α_3 は 3.0 に固定して） α_2 を変動させると、 α_1 が 6 の方が最も高い認識精度を示す α_2 の点が多かった。 α_1 が 6 の時に最良の認識精度を満たす α_2 は 57 点と、他の最良の認識精度を満たすことができる α_1 よりも多く現れた。これは、2 番目に α_2 が多く現れた α_1 が 5 の時は

α_2 が 16 点、3 番目に α_2 が多く現れた α_1 が 4 の時は α_2 が 9 点であることを考えると、 α_1 が 6 の時に最良の認識精度を満たす α_2 の点の数は非常に多いと言えるだろう。最良の認識精度を満たす α_2 の点の数が最も少ないのは α_1 が 3 の時であり、 α_2 の数が 3 点しか存在しない。しかし、図 6 から図 9 を見ると、 α_1 が 3 の時がカラーマップにおいて明るい範囲が最も広く、最良の認識精度を満たす α_1 はそれぞれにパラメータとして選択するメリットがあるように考察できる。

また、図 11 のように α_1 と α_2 を固定して α_3 を変動させたが、認識精度は α_3 の値が 2.9 以上の時常に最良値であったため、 α_3 は 2.9 以上であればパラメータ選択の自由度が高い。実際に 10 秒間加速度を端末で取得した結果、平常状態と酩酊状態で歩行した時の加速度の分散が 25 を超える加速度は 70 回の中に 1 つも存在していないが、走行した時の加速度の分散は 10 回中すべての加速度が加速度の分散が 100 以上であった。この結果からも、酩酊状態の歩行と特殊歩行の認識のためのパラメータ選択にはかなり余裕があることがわかる。

3.2 音声に基づく飲酒状態認識のパラメータ最適化

実験時は音センサが内蔵されており、かつ両手で安定して保持可能なサイズである Android 端末（本実験では Sony 社の Xperia Tablet）で音声の取得を行った。計測場所は静かな室内とし、音声はサンプリング周波数は 44100Hz、音声フォーマットは PCM16bit、wav 形式で録音を行った。実験用として平常時と酩酊時にそれぞれ 20 回録音し、合計 40 個の音声データを評価データとして使用した。

第 2.3 節で提案した音声の周波数に基づく飲酒状態認識手法において、以下のような 2 種類のパラメータが存在する。

- $V_i^{(f)} - V_{i-1}^{(f)}$ の閾値 β_1
- 閾値 β_1 を上回る回数の閾値 β_2

本節では、これらのパラメータを最適化し、音声の周波数に基づく飲酒状態認識の最良の認識精度を求める。

パラメータが取り得る範囲は、

- $1 \leq \beta_1 \leq 20$ (0.1 刻み)
- $1 \leq \beta_2 \leq 14$ (1 刻み)

とする。パラメータを変動させた結果、図 14 のように $\beta_1 = 1.3$ 、 $\beta_2 = 13$ において認識精度が 0.960 となり最も高くなった（酩酊である認識精度は 1.000、平常である認識精度は 0.900）。

最適なパラメータは閾値 β_1 がかなり小さいが、閾値 β_1 を上回る回数の閾値 β_2 は非常に条件が厳しくなった。一方、2 番目に高い認識精度は 0.950 で、パラメータの組み合わせは $\beta_1 = 18.0$ 、 $\beta_2 = 1$ となり、 β_1 が大きく、 β_2 の条件は緩いという真逆の結果となった。これは図 15 にも表れており、認識精度の高いパラメータの組み合わせが左上から右下にかけて弧を描くように集まっている。

β_1 が 1.3 と 18.0 の間に位置する 7.9 の時にも同じく 2 番目に高い認識精度 0.950 となる β_2 が存在している。その β_2 の値は 13 と 1 の中間付近ではなく、3 である。 β_1 が 7.9 の時、 β_2 が 7 や 8 であると、表 1 のように認識精度が低くなることから、酩酊時の人間の音声には特有の法則性があると考えられる。

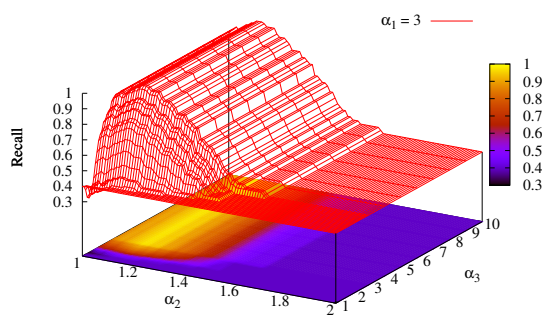


図 6 $\alpha_1 = 3$ の時の再現率
Fig. 6 Recall when $\alpha_1 = 3$.

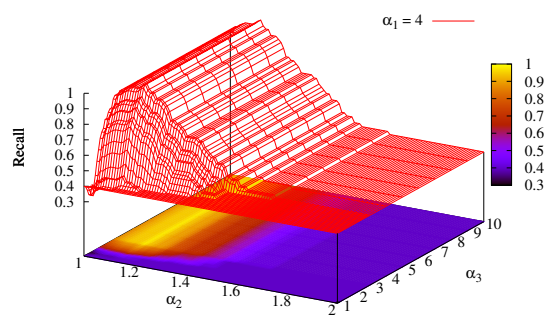


図 7 $\alpha_1 = 4$ の時の再現率
Fig. 7 Recall when $\alpha_1 = 4$.

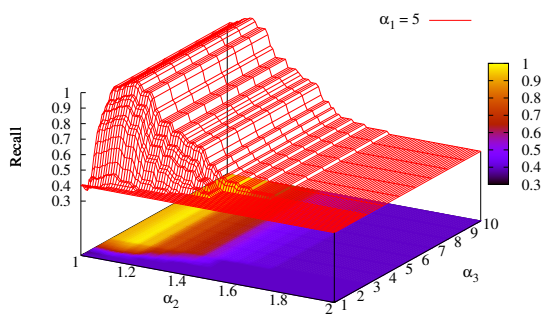


図 8 $\alpha_1 = 5$ の時の再現率
Fig. 8 Recall when $\alpha_1 = 5$.

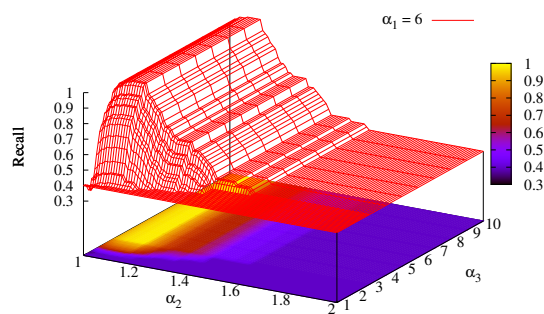


図 9 $\alpha_1 = 6$ の時の再現率
Fig. 9 Recall when $\alpha_1 = 6$.

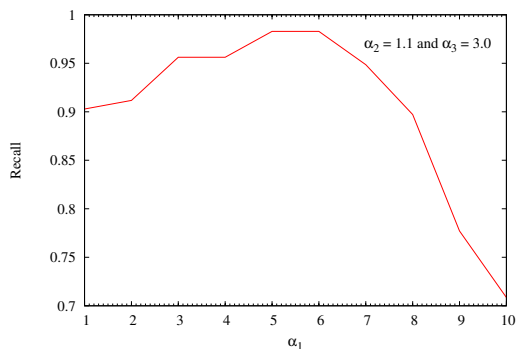


図 10 α_1 のみ変動させた時の再現率
Fig. 10 Recall in varying only α_1 .

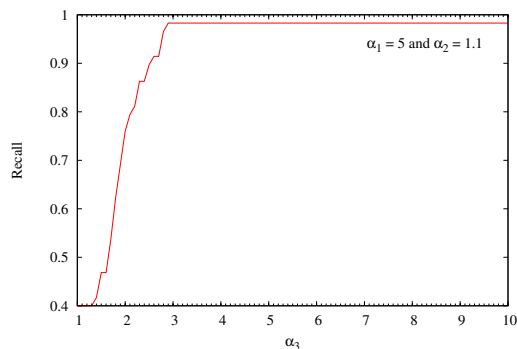


図 11 α_3 のみ変動させた時の再現率
Fig. 11 Recall in varying only α_3 .

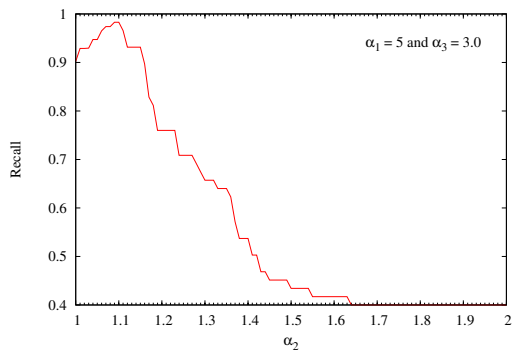


図 12 α_2 のみ変動させた時の再現率 ($\alpha_1 = 5$)
Fig. 12 Recall in varying only α_2 ($\alpha_1 = 5$).

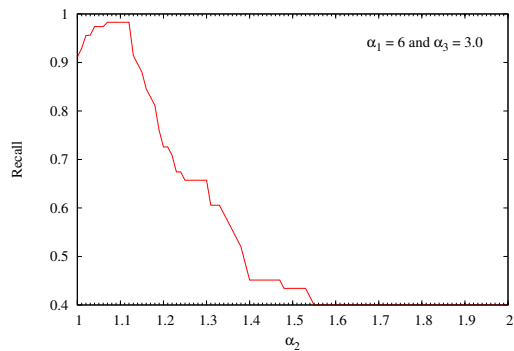


図 13 α_2 のみ変動させた時の再現率 ($\alpha_1 = 6$)
Fig. 13 Recall in varying only α_2 ($\alpha_1 = 6$).

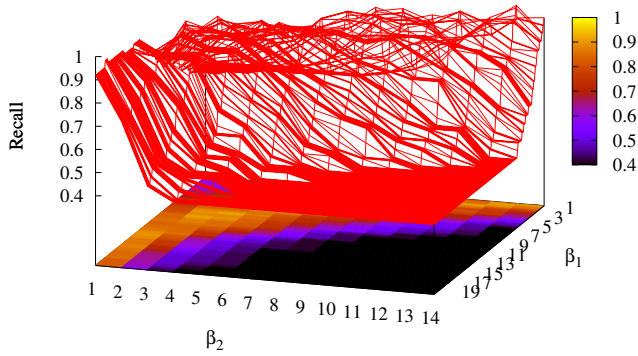


図 14 再現率の β_1 及び β_2 への依存性
Fig. 14 Dependency of recall on β_1 and β_2 .

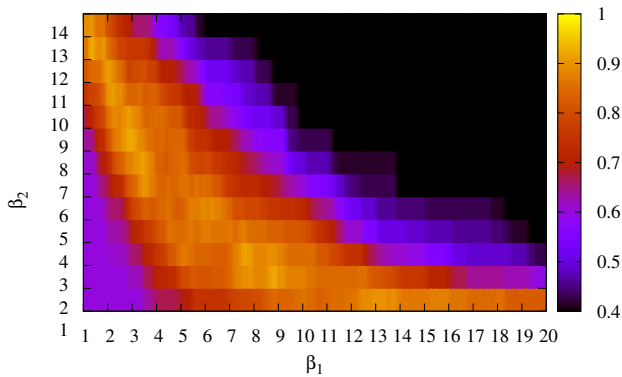


図 15 β_1 及び β_2 に依存した再現率のカラーマップ
Fig. 15 Color map of recall on β_1 and β_2 .

表 1 認識精度の順位とパラメータ β_1 及び β_2
Table 1 Ranking of recall with β_1 and β_2 .

順位	β_1	β_2	認識精度 (酩酊, 平常)
1	1.3	13	0.960 (1.000, 0.900)
2	18.0	1	0.950 (0.950, 0.950)
2	7.9	3	0.950 (0.950, 0.950)
2	1.4	13	0.950 (0.950, 0.950)
2	18.1	1	0.950 (0.950, 0.950)
2	18.2	1	0.950 (0.950, 0.950)
2	18.3	1	0.950 (0.950, 0.950)
8	2.9	9	0.940 (1.000, 0.850)
8	3.1	8	0.940 (1.000, 0.850)
8	8.9	2	0.940 (1.000, 0.850)
11	2.0	11	0.930 (0.950, 0.900)
11	7.6	3	0.930 (0.950, 0.900)
11	9.2	2	0.930 (0.950, 0.900)
⋮	⋮	⋮	⋮
907	7.9	7	0.700 (0.500, 1.000)
⋮	⋮	⋮	⋮
1066	7.9	8	0.640 (0.400, 1.000)
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮

4. まとめと今後の課題

本稿では、アルコールによる被害から身の安全を確保するため、酩酊と呼ばれる状態となってしまったユーザーに関して、ユーザーの自発的な行動ではなく酩酊時特有の症状をモバイル端末のセンサ情報、特に加速度センサと音センサを監視して加速度データと音声データを取得して飲酒状態認識を行い、酩酊状態と判定した場合には家族や警察に連絡や通報、周囲の人に警報を発するシステムを提案した。また、飲酒状態認識のためにパラメータ最適化を行い、加速度センサに基づく飲酒状態認識は 80 件の評価データの入力に対して認識精度 0.982 (酩酊である認識精度は 0.971, 酩酊でない認識精度は 1.000), 音センサに基づく飲酒状態認識は 40 件の評価データの入力に対して認識精度 0.960 (酩酊である認識精度は 1.000, 平常である認識精度は 0.900) が得られた。

今後の課題として、本稿における実験で使用した加速度と音声はすべて第 1 著者のみの加速度と音声であるため、より多くの人の加速度と音声を取得し、他の人が本システムを使用した場合にも飲酒状態を精度良く認識できるかどうかの検証を行う。また、本稿では歩行による加速度データを記録する時間の長さを 10 秒間、音声データを記録する時間の長さを 5 秒間と仮に設定したが、この時間の長さもパラメータ化し、それぞれの飲酒状態認識手法で必要とする最適な記録時間がどの程度であるかの最適化も行っていく必要がある。

加えて、本稿の音声の周波数に基づく飲酒状態認識は、加速度に基づく飲酒状態認識が酩酊状態と平常状態以外の特殊歩行状態も考慮しているのに対し、酩酊状態と平常状態の 2 値のみを想定しているため、ユーザーがわざと大声を出した、歌を歌った、周囲で大きな音がしたなどの酩酊状態と平常状態のどちらでもない特殊な音が入力された場合、正確な認識を行えないと予測されるのでアルゴリズムの改良を行う必要がある。その改良の際、モバイル端末の音センサ以外に、端末の GPS 位置情報を利用し、例えばユーザーがカラオケ屋に居るといった位置情報を得た際に、ユーザーが歌う、大声を出す状態にある、周囲で大きな音がする可能性が高いとシステムに推定させ、位置情報に応じたアルゴリズムやパラメータに切り替えさせるなどの手法が可能であるか検証を行っていく。さらに、発展課題としてモバイル端末の音声認識を利用して、ユーザーが何度も音声認識に失敗した時、アルコールが原因で呂律が回らなくなっていると推定して飲酒状態認識を行うことも検討している。

文 献

- [1] 20 代前半の飲酒率、女性が男性上回る～厚生労働省調査～, <http://news.envy-j.com/news/detail.jsp?newsno=643> (2014/10/20).
- [2] 植田将基, 久原政彦, 伊藤誠, 遠藤守, 山田雅之, 宮崎慎也, 快適な飲酒を支援するコースター型飲酒検知デバイス, 電子情報通信学会, 信学技報, MVE 研究会, vol.112, no.25, pp.43-46 (2012).
- [3] 情報通信白書 平成 26 年版, <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h26/html/nc1411100.html> (2014/10/20).
- [4] 梅田悦生, 飲酒の生理学: 大虎のメカニズム, 裳華房 (1997).
- [5] 血中アルコール濃度と酔いの症状, <http://www.kirin.co.jp/csv/arp/fundamental/blood.html> (2014/12/16).