

スマートフォンを用いた音量，光量による 視聴コンテンツ判定手法

齋藤 松汰¹ 瀬崎 薫² 岩井 将行¹

概要：スマートフォンでは検索履歴，位置情報から状態，趣味を予測して適切な情報を選択し表示している。このようにユーザの状態や趣味に応じたプロファイリングを行い，より適切な情報を表示することが求められている。[1] 予測するための元となる情報は，収集されている情報のみとなっている。これに対して，ユーザが音楽コンテンツ，PC で閲覧している動画コンテンツなどの提供者から不明確な状態の情報を加えることで，より最適な情報の選択表示が可能となる。本研究では，音の抑揚と光量の変化をスマートフォンのマイク及び光量センサーを用いて調査し，視聴コンテンツ判定の方法についての提案と実装を行う。

キーワード：スマートフォン，マーケティング，音声分析

The viewing and listening contents judging technique by sound volume and light volume using SmartPhone Sensor

SHOTA SAITO¹ KAORU SEZAKI² MASAYUKI IWAI¹

Abstract: In the smart phone, a state and a hobby are predicted from a search history and position information, and suitable information is chosen and displayed. Thus, profiling according to the user's state and hobby is performed, and displaying more suitable information is called for. The information which becomes the origin for predicting is only the information currently collected. On the other hand, selection presenting of the more nearly optimal information of a user is attained from donors, such as a music content and animation contents currently perused with PC, by adding the information on an indefinite state. In this research, the intonation of sound and change of light volume are investigated using the microphone and light volume sensor of a smart phone, and proposal about the method of a viewing-and-listening contents judging and mounting are performed.

Keywords: smartphone, marketing, speech analysis

1. 序論

1.1 外的要因とマーケティングとの関係性

人の行動の 90 % は無意識に行われていると言われてい
る [2]。商品選択には大きく分けて内的要因と外的要因が
あり，内的要因とは，自分自身の欲求を基準に選択に優劣
をつけるものであり，外的要因とはその名の通り自分自身

以外の外部から影響を受けることである。外的要因とは，
例えば，スーパーに行き，周りの影響により買う予定では
なかったものを買ってしまうということである。これは非
計画購買と呼ばれ，これらには店頭の POP や得値商品，
エンド棚などのオススメ商品，試食品などの外的要因と密
接な関係がある。購入物選択は，自分の中の「内的要因」
だけでなく「外的要因」にも影響を受けているのでマーケ
ティングを考える際に必要がある。よって，スマートフォ
ンのバナー広告の表示選択において外的要因を加えること
により，より最適な情報選択表示となる。

¹ 東京電機大学
Tokyo Denki University

² 東京電学
The University of Tokyo

1.2 スマートフォンを用いたセンシングについて

まず、スマートフォンを活用したデータ収集システムの利点、欠点を下記に示す。

- 利点
 - － さまざまな場所の現場データの収集が可能
 - － 事前情報入力により再入力が必要
 - － 多様なセンサによる計測の冗長性
- 欠点
 - － デバイス差による計測値 f 4 誤差の存在
 - － 情報漏えいなどのセキュリティ問題
 - － バッテリーの制約

スマートフォンには様々なセンシング機能が備わっている。加速度、気圧、方位、音、光、GPS、傾き、標高など多くのデータを計測することが可能である。その機能を利用したアプリなどが開発されている。ユーザは、様々な場所にいるため、多くの場所におけるデータの収集が可能であり、多様な種類のデータが得られる。[3]

1.3 関連研究

音楽、歌声などの音波形に関する研究が従来から盛んに行われており、学術的な観点からだけでなく、産業応用的な観点からも注目を集めている。例えば、歌声は通常の音声よりもはるかに変動が大きく、歌声と相互に関連し合う伴奏音も大きな音量で含まれている。歌声は、[4] オーバーシュート、ビブラート、プレパレーション、微細変動の4種類に歌唱者に依存せず含まれており歌声の基本周波数の軌跡により分割することが可能である。この技術を応用し、歌唱力向上支援、音声合成システムの研究が行われている。

外的要因による最適な情報選択表示では、ターゲティングTV広告というのがある。[5] それは、視聴者の見ている番組や年齢、年収などから、最も興味を得られると推測される内容をテレビCMに反映させるものである。複数の視聴者が同じテレビドラマを視聴していても、テレビCMに切り替わると、Aには高級車、Bにはファミリーカー、Cには軽自動車といった具合に、CMの出し分けが行われる。

2. スマートフォンによる音量、光量の取得

取得したデータから音圧を計算し、[6] その音圧の変化を音の抑揚とし、その特徴分析を行う。同様に、光量の特徴分析も行う。視聴コンテンツにした理由は、同一コンテンツで同一な音、映像であり同データ、違データとの相関が明確に表すことができると予想したためである。

実証方法は、目的となるマスターデータと同一コンテンツのテストデータのみ相関値が高くなる方法を考える。15種類の視聴コンテンツの音の抑揚と光量の変化のデータをスマートフォンにより計測し比較を行う。視聴コンテンツが流れる画面から1 m, 2mと距離を変えて計測を行う。そ

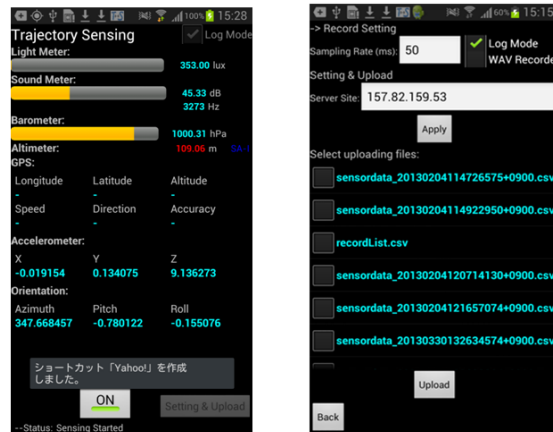


図 1 Trajectory Sensing の画面
Fig. 1 Screen of Trajectory Sensing

の目的は、距離により情報の欠落がどの程度あるかを判断し、マッチングの精度を推定するためである。

2.1 データ取得に用いたアプリ

Trajectory Sensing を使用した。図 1 これは、東京大学の劉 広文が作成したものである。各種センサから取得したデータを表示する図 1 左と同時に、データを CSV ファイルに保存する。データの取得する周期を設定することができ図 1 右、データの精密度を変更できる。

2.2 スマートフォンによるデータ取得方法

Android アプリケーション開発の API[7] には、照度センサと音の周波数を計測するクラスが用意されており、それを使用することでデータの取得が可能である。光量は、計測値が変わるたびに書き換えられるのでデータの取得が容易である一方、音の周波数は、データが変わる毎に書き換えられるのではなく、ある程度の周波数をストックしておき、それを出力するので、プログラミングの書き方に注意が必要である。よって、周波数をストックしておく最適な時間を次節で考察する。

2.3 スマートフォンによる収集した波形

図 2 のように、200ms のストック時間で周波数を計測した場合は、50ms に比べて凹凸がなく滑らかになっている。本研究では波形の類似性を求めるため、特徴となる周波数の数が多いほどよい。しかし、ストック時間を短くしすぎると乱雑になり波形のマッチングが困難になる。ストック時間を変えて波形を示し、目視によりストック時間は 50ms が最適であると判断した。

2.4 スマートフォンによる取得データ

15 種類の視聴コンテンツの種類、距離で取得した音の抑揚と光量の変化のデータのグラフ化を行い、目視でのデー

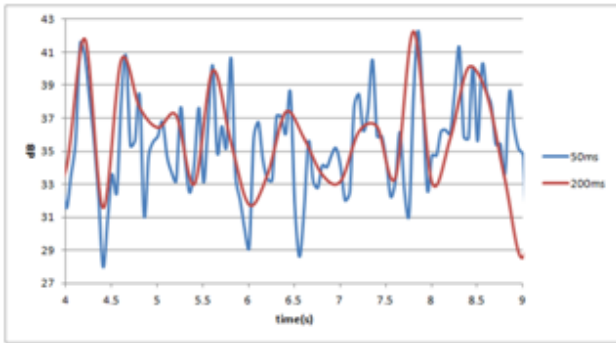


図 2 異なるストック時間の波形比較
Fig. 2 Waveform comparison
of different stock time

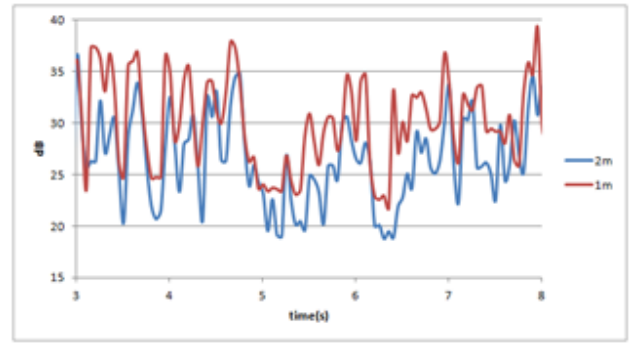


図 3 異なる距離の波形比較
Fig. 3 Waveform comparison
of a different distance

タの類似度は高いので、スマートフォンによる視聴コンテンツ判定は可能であると判断した。

同一視聴コンテンツにおける同一距離、異なる距離での音の抑揚の相関係数 [8] を求めたところ、 $cor=0.3723457$ の低い相関値となった。また、光量の変化も、同様に計算したところ相関値が低い値となった。これは、波形の形状は似ているが、距離による減衰、時間軸の差異により低い値になったと考えられる。そのため、時間軸が異なっても類似している形状であれば相関値が高くなるアルゴリズムにする必要がある。

2.5 視聴コンテンツ判定のためのアルゴリズム推測

距離による音圧レベルの減衰は次式で求められる。 [9]

$$Lp2 = Lp1 - 20 \log_{10} \frac{r2}{r1} \quad (1)$$

$Lp1$: 観測地点 $r1$ の音圧レベル [dB]

$Lp2$: 求めたい地点 $r2$ の音圧レベル [dB]

$r1$: 音源と受音点の距離 [m]

$r2$: 求めたい観測地点の距離 [m]

したがって、1m 地点 ($r1$) の音圧レベルに比べて、2m 地点 ($r2$) での減衰は

$$Lp2 = Lp1 - 6.0206[\text{dB}] \quad (2)$$

となり、1m 地点に比べ 2m 地点の音圧は継続して 6.0206 [dB] 低くなる事が分かる。

図 3 は同種コンテンツのそれぞれの距離での波形である。距離が近いほうが継続して音圧レベルが高くなっている。しかし、波形の形状はそれほど変化していない。

また、距離による光の減衰は次式で求められる。 [11]

$$\frac{i_1}{i_2} = \frac{d_2^2}{d_1^2} \quad (3)$$

i_1 : 観測地点 $r1$ の光量 [lux]

i_2 : 求めたい地点 $r2$ の光量 [lux]

d_1 : 光源と受光点の距離 [m]

d_2 : 求めたい観測地点の距離 [m]

したがって、1m 地点 (d_1) と 2m (d_2) での光量の比較は

$$i_2 = \frac{1}{4} i_1 [\text{lux}] \quad (4)$$

となり、2m 地点の光量は継続して 1m 地点の $\frac{1}{4}$ の値になることが分かる。1m 地点と 2m 地点との同種コンテンツの波形をグラフ化したところ音量と同様に継続して値の差があったが、波形の形状はそれほど変化していない。

以上のことから、異なる距離での類似性を計算することは可能と考えられる。また、音量、光量の両データは距離によって値が減衰されるので、それを考慮した判定のためアルゴリズムを考えるべきである。なので、波形の特徴の類似性により相関値を決めるアルゴリズムで分析を行う。

3. 視聴コンテンツ判定のための手法

3.1 コヒーレンスを用いた判定のための手法

コヒーレンスとは、波形同士の干渉のしやすさを表す。干渉とは、複数の波を重ね合わせるとき、波が打ち消し合ったり強め合ったりすることである。例えば、太平洋赤道上の 2 点で計測された流速の時系列データを比べるとき、コヒーレンスを用いると流速の時系列データは長周期の変動では相関が高く、短周期の変動では相関が低いことを知ることができ、流速の特徴分析ができる。よって、コヒーレンスはそれぞれの周波数の波形の変動がどれくらい相関かを定量的に計ることができる。一般式は、フーリエ周波数成分の相関係数に相当し以下の式で表され $0 \leq coh^2(\omega) \leq 1$ を満たす。

$$coh^2(\omega) = \frac{|S_{xy}|^2}{S_{xx}(\omega)S_{yy}(\omega)} \quad (5)$$

つまり、コヒーレンスは、二組の時系列データの関連密度

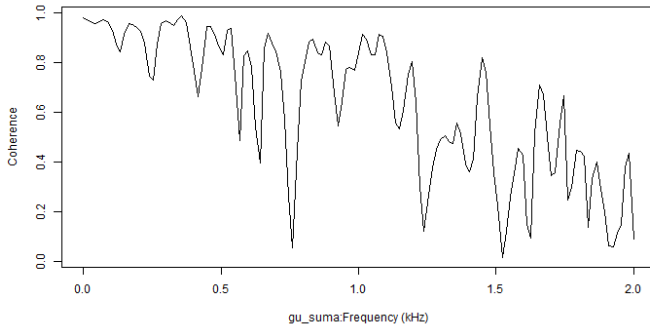


図 4 同種コンテンツの音量のコヒーレンス図

Fig. 4 Figure of coherence of the volume of similar contents

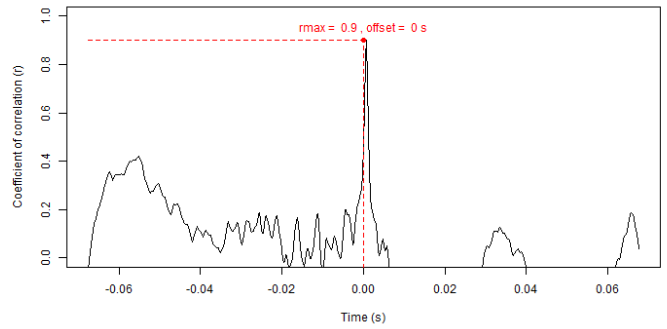


図 6 同種コンテンツの音量の PMS 図

Fig. 6 Figure of PMS of the volume of similar contents

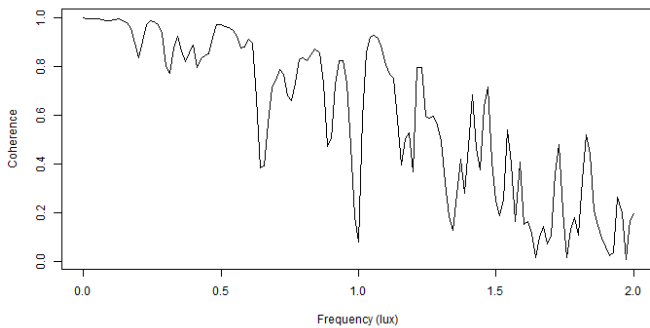


図 5 同種コンテンツの光量のコヒーレンス図

Fig. 5 Figure of coherence of the light quantity of similar contents

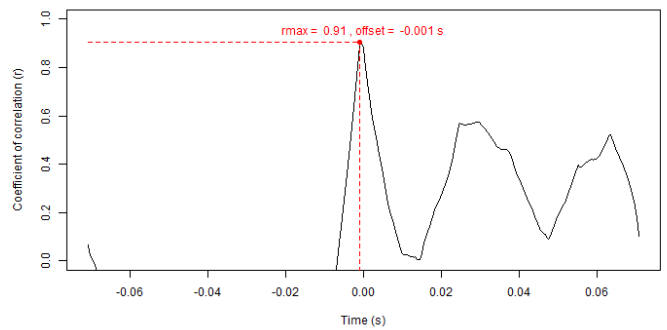


図 7 同種コンテンツの光量の PMS 図

Fig. 7 Figure of PMS of the light quantity of similar contents

をそれぞれのサンプリング周波数で値を 0 から 1 の間で返し、値が 1 に近いほどその周波数で相関が強いことを示す。スマートフォンにより取得した音量をコヒーレンスにより計算したところ図 4 になった。同種視聴コンテンツ比較は、短周波数の相関が高く平均 0.8 を超えてくる、一方長周波数は相関が低い。同様に、光量の計算を行い図 5 のようになった。光も短周波数の相関が高く、長周波数は相関が低い。なので、音量、光量ともに長周波の雑音に対しては強く、マッチングに影響を及ぼさないことが分かる。

3.2 PMS を用いた判定のための手法

Pre-multiplied Spectra(PMS)[12] は、スペクトル強度に波数を乗算量を表す。横軸を波数または波長の対数に対してプロットしたときにエネルギー含有量を示す指標となる。

図 6 は、音量の PMS であり、図 7 は光量の PMS である。rmax は時間軸をシフトして計算した相関係数であり、0 から 1 の間で返し、1 に近いほど相関が強いことを示す。rmax が高い値の場合は、類似度の高い波形同士ということになる。

3.3 コヒーレンス、PMS を用いた音量、光量の評価

15 種類の映像コンテンツの音量、光量それぞれのコヒーレンス線分の積分の平均値を求めたところ表 1 のようになった。音量、光量の両データで同一コンテンツ、異種コンテンツにおいて有意差が表れる結果となった。なので、コヒーレンス、PMS を用いて視聴コンテンツ判定が可能である。異なる距離同士も求めたところ表 2 のようになった。光量の方が距離による減衰が少っている。

同様に、それぞれの PMS の rmax の平均を取りまとめたところ表 3 のようになり、異なる距離同士も求めたところ表 4 のようになった。しかし、両アルゴリズムにおいて距離の変化により減衰がみられるため、遠距離において視聴コンテンツのマッチング精度が著しく下がると考えられる。

3.4 音量、光量データによる視聴コンテンツ判定方法提案

視聴コンテンツ判定に音量、光量それぞれのデータでマッチングが可能であるとわかった。視聴コンテンツ判定の値をそれぞれの音量、光量データの coherence 積分値または、PMS の rmax の値のユークリッド距離とする。なぜなら、その時の周りの状況によりデータの信頼性が変化し、不適切なシチュエーションでは有意差が出ない場合があるからである。ユークリッド距離の判定値ならば片方の

表 1 視聴コンテンツそれぞれの種類による積分値

Table 1 Integral calculus level by the kind of the each seeing and hearing contents.

	同種コンテンツ	異種コンテンツ
音量	84.66969	31.47676
光量	82.08225	38.22343

表 2 視聴コンテンツそれぞれの距離による積分値

Table 2 The integration value by the distance of each viewing-and-listening contents.

	同一距離	異なる距離
音量	84.66969	61.59588
光量	82.08225	73.29339

表 3 視聴コンテンツそれぞれの種類による rmax の値

Table 3 Value of rmax by the kind of the each seeing and hearing contents.

	同種コンテンツ	異種コンテンツ
音量	0.89	0.29
光量	0.88	0.27

表 4 視聴コンテンツそれぞれの距離による rmax の値

Table 4 Value of rmax by the distance of the each seeing and hearing contents.

	同一距離	異なる距離
音量	0.89	0.76
光量	0.88	0.74

計測値のマッチングが高ければ、判定値は高くなる。

4. 結論

本稿では、視聴コンテンツ判定を音量、光量に着目し、時系列データ間の関連を表すコヒーレンス、PMS を用いて、視聴コンテンツの判定手法を行った。提案手法を評価するため、コヒーレンス線分の積分または、PMS の rmax の値によりそれぞれの場合の値の算出をしまとめた。その結果、それぞれの手法により視聴コンテンツの判定に音量、光量の両データが使用可能ということがわかった。

5. おわりに

今後の展望として、マーケティングと結び付けられているアプリの開発を行っていく。現在プロトタイプとして、CM が流れている同時刻にその CM の内容のハイパーリンク作成や広告を表示するアプリの開発を行っている。また、様々なシチュエーションによる判定値の変動を実験により調査し、より最適な方法がないか模索している。

謝辞 本研究は H25 科研費若手研究 (A)(代表者:岩井将行, 課題番号:25700007) の一部により行われている。

参考文献

- [1] “精度の高いターゲティング” MicroAd (<http://www.microad.co.jp/service/blade/merit.php/sec-targeting>)
- [2] “行動に影響を与える内的要因と外的要因” MARKE LABO(<http://www.markelabo.com/base/yoin/>)
- [3] 笠崎 伸一郎: “スマートフォンを活用したデータ収集システムについて”
- [4] 後藤真孝, 平 田圭二: “音楽情報処理の最近の研究” 日本音響学会誌 60 巻 11 号 pp.675-681 (2004)
- [5] kenji sakaue: “グレースノート、テレビ CM のターゲティングができる新技術を日本初披露” (<http://1topi.jp/curator/CarGuyTimes/1304/19/178073>)
- [6] “Sound Pressure” The Engineering ToolBox(http://www.engineeringtoolbox.com/sound-pressure-d_711.html)
- [7] “Sensor” Developers (<http://developer.android.com/reference/android/hardware/Sensor>)
- [8] “Correlation Coefficients: Find Pearson’s Correlation Coefficient” Statistics How To (<http://www.statisticshowto.com/how-to-compute-pearsons-correlation-coefficients/>)
- [9] Beranek: “Sound Attenuation” (<http://www.silex.com/pdfs/sound/20attenuation.pdf>)
- [10] Fletcher Dunn, Lan Parberry: “ゲーム 3 D 数学”
- [11] 辻 義之, 田中 宏彦, 大野 哲靖: 相関とスペクトル解析
- [12] 河村 洋: “平行平板間乱流の大規模 DNS の展開と壁面及び中央部に現れる大規模構造” ながれ 22 pp.467-476 (2003)

齋藤 松汰

1991 年生。2010 年東京電機大学未来科学部情報メディア学科入学。2014 年東京電機大学未来科学部情報メディア学科卒業予定

瀬崎 薫

1989 年 3 月東京大学大学院工学系研究科電気工学専攻博士課程修了 (工学博士)。2001 年 4 月東京大学空間情報科学研究センター空間情報システム研究部門助教 (現職)

岩井 将行

2002 年 3 月慶應義塾大学大学院政策・メディア研究科修士課程修了。2004 年 9 月慶應義塾大学大学院政策・メディア研究科博士号取得。2013 年 4 月東京電機大学未来科学部情報メディア学科 准教授