

音の抑揚と光量の変化を利用したスマートフォンにおける 周辺映像コンテンツ推定

A Proposal for Surrounding Video Contents Estimation using SmartPhone Sensors

齋藤 松汰†
Shota Saito

奥野 淳也‡
Zyunya Okuno

瀬崎 薫‡
Kaoru Sezaki

岩井 将行†
Masayuki Iwai

1. はじめに

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。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一切わからない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を用いて調査し、周辺に流れる映像コンテンツの推定を行う技術について提案を行う。

2. 現在の波形解析現状

2.1 近年の音楽情報処理分野

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、[1]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。

2.2 映像波形分析の目的と実証方法

周辺映像を映像コンテンツに絞る。映像コンテンツは、同コンテンツで同一音、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。映像コンテンツが流れる画面から1m, 2mと距離を変えて計測を行う。その目的は、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。

3. スマートフォンを使用した音の抑揚と光量の変化のデータの取得

3.1 スマートフォンによるデータ収集

まず、スマートフォンを活用したデータ収集システムを提案する理由を述べる。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

† 東京電機大学未来科学部情報メディア学科
Tokyo Denki University

‡ 東京大学空間情報科学研究センター
The University of Tokyo

- 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

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。

3.2 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。

3.3 スマートフォンによる収集した波形

図1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く50msのストック時間で行うこととする。このように、ストック時間を変えて波形を示し、ストック時間は50msが最適であると判断した。

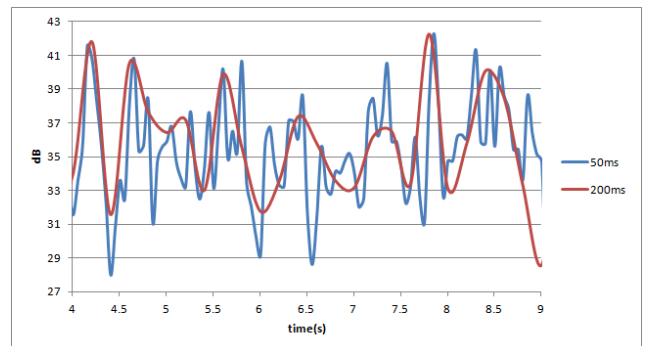


図1 異なるストック時間の波形比較

3.4 距離による音圧レベルの変化

距離と音圧レベルの関係は次式で求められる。[2]

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

$r1$: 音源と受音点の距離 [m]

$r2$: 求めたい騒音予測地点の距離 [m]

$Lp1$: 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]

したがって、今回の実験の2m地点($r2$)での減衰は

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

となり、1m 地点に比べると 2m 地点の音圧は継続して 6.0206 [dB]低くることが分かる。

図 2 は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ため、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。

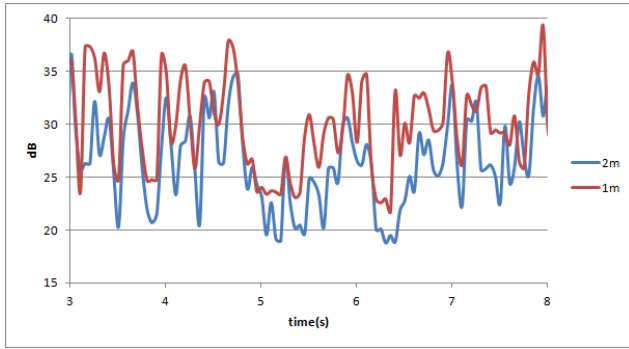


図 2 異なる距離の波形比較

3.5 各距離での音、光データの問題点

同じ距離、同一映像コンテンツにおける音の抑揚の相関係数を求めたところ、 $cor=0.3723457$ となった。また、光量の変化も、同様に相関がほとんどない値と、なってしまった。これは、波形の形状は似ているが、時間軸の差異によって低い値になってしまったと考えられる。そのため、時間軸が異なっても同一形状であれば相関値が高くなるアルゴリズムを考える必要がある。

4. 計測結果によるデータの選別、類似度計測

4.1 異なる時間系列データの関連密度

15 種類の映像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った。[3]コヒーレンスとは、二組の時系列データの関連密度をそれぞれのサンプリング周波数で値を 0 から 1 の間で返す。値が 1 に近いほどその周波数で相関が強いことを示す。図 3、図 4 はそれぞれの場合のコヒーレンス図である。同一映像コンテンツは値が高いところが多いことが分かる。

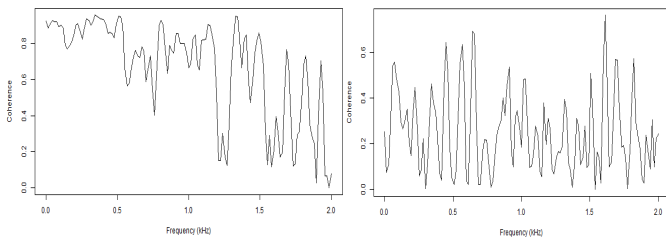


図 3 同一コンテンツ比較 図 4 異種コンテンツ比較

また、[4]Pre-multiplied Spectra(PMS)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 5、図 6 はそれぞれの場合の PMS である。rmax は時間軸をシフトして計算した相関係数が最も高い値であり、それぞれ 0.89、0.28 となっている。しかし、同様に光量の変化の値を計算したところ、両アルゴリズム共に望

む値とならなかった。スマートフォンを用いた光量の変化の取得はセンサの精度の観点から困難があると言える。

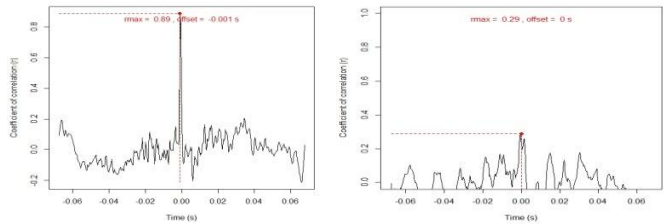


図 5 同一コンテンツ比較 図 6 異種コンテンツ比較

すべての映像コンテンツのそれぞれの場合のコヒーレンスの面積、PMS の rmax を求めた。表 1、表 2 は比べるのが同一映像コンテンツ(Same)、距離が違う(Distance)、異種映像コンテンツ(Different)の場合の平均(Average)、低周波(Low)、高周波(High)の組み合わせを図に示した。

	Average (V)	Low (Hz)	High (Hz)
Same	81.66969	50.19293	31.47676
Distance	71.413	46.53475	24.87825
Different	33.00617	18.415	14.59117

表 1 コヒーレンスのそれぞれの場合の面積

	Same	Distance	Different
rmax	0.89	0.76	0.29

表 2 PMS のそれぞれの場合の rmax の値

4.2 実験結果、評価

表 1、表 2 のように、それぞれの場合において有意差が表れる結果となった。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、遠距離において映像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。

5. おわりに

本稿では、最適なデータの取得から、データの減衰、データの類似性の検討を行った。今後、遠距離においても相関値が下がらない波形マッチングアルゴリズムを考察する。

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] 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)
- [2] “Sound Attenuation” (<http://www.silex.com/pdfs/sound%20attenuation.pdf>)
- [3] “SweaveListingUtils: Utilities for Sweave together with TeX listings package”(http://cran.r-project.org/web/packages/SweaveListingUtils/index.html)
- [4] 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)