

眼球計測データおよび音響信号の特徴量抽出に基づく 「音楽の好み」の推定

米家 惇[†] リャオ シンイ[†] 古川 茂人[†] 柏野 牧夫^{†‡}

[†] 日本電信電話株式会社 NTT コミュニケーション科学基礎研究所 人間情報研究部 〒243-0198 神奈川県厚木市森の里若宮 3-1

[‡] 東京工業大学大学院 総合理工学研究科 物理情報システム専攻 〒226-0026 神奈川県横浜市緑区長津田町 4259

概要

脳計測に基づく心理推定に替わって、眼球計測に基づく心理推定を行うための手法を提案する。具体的には「楽曲の好み」の程度を推定するため、眼球計測信号と音響信号に対して独自の特徴抽出を行い、特徴量に基づいた学習を行った例を報告する。推定に用いる特徴として microsaccade および縮瞳・散瞳に着目し、制御論モデルに基づく新たな特徴量を抽出した。また、生体応答がどのような刺激によって生じたのかを推定の手がかりとするため、音響信号の各種特徴量も抽出した。眼球計測データと音響信号から抽出された特徴量と、楽曲の好みに関する主観評定の対応関係を学習することで、高い精度で音楽の好み推定が可能となった。

1. 背景

近年、脳機能計測および機械学習の高度化に伴い、脳信号に基づいてヒトの知覚体験を読み出すデコーディングの研究が発展している。しかし、脳計測には一般にノイズやロバスト性、簡便性等に問題があり、実用上多くの制約があるものと考えられる。そこで我々は、上記の問題による影響が比較的少なく、かつ「脳の一器官」とも言われるほど我々の知覚状態を反映している「眼球」に着目し、眼球計測に基づくデコーディングの手法を提案している。本発表では具体的に音楽の好みの推定に関する例を示す。

眼球運動や瞳孔径の計測データは、多チャンネルで測定する脳波と比べノイズが少なく計測が容易な反面、単位時間当たりで取り出せる情報量が少ないという一面をもつ。本研究では、microsaccade（無意識に生じる微小な眼球運動）に対して制御論モデルを適用することで、従来着目されてこなかった潜在的な特徴量を抽出し、さらに刺激（音響信号）の特徴も組み込むことで、眼球計測に基づく知覚状態のデコーディングを可能とした。

2. 実験

20-40 代の健常成人 22 名が実験に参加した。実験参加者は、クラシック/ロック/ジャズの3つのジャンルから5曲ずつ、計15曲をランダム順に聴取した。各曲の長さは90秒とした。実験参加者が音楽を聴取する際の眼球運動を、SR Research 社製 EyeLink-CL System を用い、1000Hz のサンプリングで計測した。実験中、実験参加者には正面に設置されたモニター中央の注視点に視線を向けるよう教示した。実験終了後、実験参加者は15曲の楽曲全てについて、好みの度合いを7段階評価（1から7）で主観評定（アンケート回答）を行った。

3. 推定方法

実験参加者が1曲を聴取した90sのデータを1サンプルの単位とした。各サンプルから計算された多次元特徴ベクトルと主観評定値の対応関係を学習し、交差検定を用いて推定精度の評価を行った。

3-1. 特徴抽出

眼球運動/瞳孔径変化に関する特徴量、音響信号に関する特徴量に加え、特徴量同士の相関値も計算することで、各サンプルから133次元特徴ベクトルを計算した。ただし、各次元の特徴量は90s間での平均値をとったものを用いた。

Estimation of music preference based on the feature extraction from eye-movement data and acoustic signals

Makoto Yoneya[†], Hsin-I Liao[†], Shigeto Furukawa[†], Makio Kashino^{†‡}

[†] Human Information Science Laboratory, NTT Communication Science Laboratories, Nippon Telegraph and Telephone Corporation 3-1 Morinosato Wakamiya, Atsugi-shi, Kanagawa, 243-0198 Japan

[‡] Department of Information Processing, Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology 4259 Nagatsuda-cho, Midori-ku, Yokohama-shi, Kanagawa, 226-0026 Japan

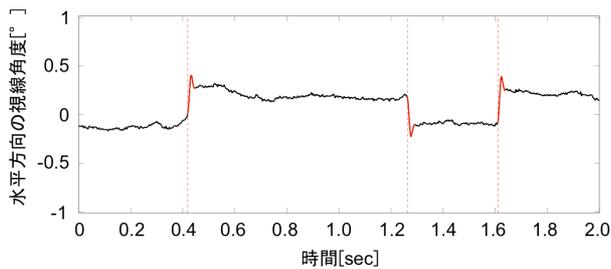


図1. 眼球計測信号中の microsaccade の例 (赤線) .

①眼球運動の特徴量

microsaccade は無意識に生じる跳躍性の微小眼球運動であり，聴覚知覚に応じた変化を示すことが報告されている^[1]．我々は microsaccade の波形が二次系のステップ応答として記述することに着目し^[2]，波形の振動性を表す減衰係数，波形の速応性を表す固有角振動数を特徴量として計算した．その他，眼球運動の特徴量として microsaccade の発生頻度，振幅，最大速度に加え，瞬目頻度等も含めた．

②瞳孔径変化の特徴量

瞳孔は自律神経系の支配を受けて驚きや不快のような感情に応じて変化を示すことが知られる．本研究では瞳孔径変化の計測データの平滑化を行った上で，一定の基準に基づいて極小点／極大点を検出することにより散瞳（瞳孔の散大）／縮瞳（瞳孔の縮小）を抽出し，散瞳／縮瞳の発生頻度，振幅，平均速度，持続時間等を瞳孔径変化の特徴量として計算した．

③音響信号の特徴量

音楽は一般に音のパターンの繰り返しと逸脱から構成される．本研究では音のパターンの時系列的な予測不可能性を定量評価することで，音響信号の「サプライズ」を特徴量として計算した．本研究が提案するサプライズモデルの詳細は口演にて説明を行う．その他，音響信号の特徴量としてスペクトログラム，ビートスペクトログラム^[3]の計算も行った．

3-2. 学習と推定

学習段階では，サンプル毎に計算された 133 次元の特徴量ベクトルに対して，アンケートで得られた好みの程度を表すスカラ値（1～7）を対応づけ，識別平面等の学習を行った．学習アルゴリズムには SVM，k-NN および単純クラスタリングの 3 通りを用いた．

推定段階では，各サンプルから計算された特徴ベクトルが 7 種類のスカラ値（好みの程度）に対応する尤度を計算し，最も尤度の高いスカラ値を推定値として出力した（最尤推定）．

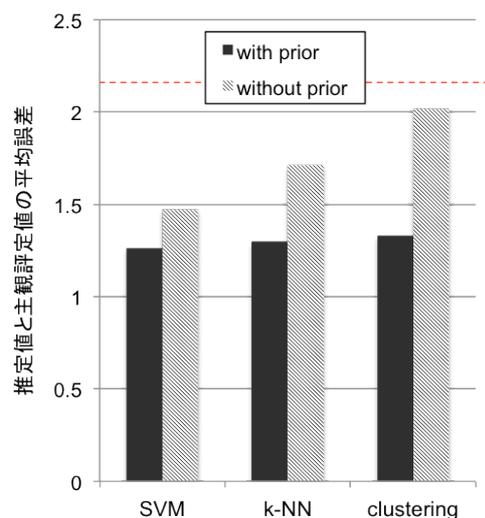


図2. 交差検定に基づく推定結果．

事前分布ありは黒色，事前分布なしは灰色で示す．点線はチャンスレベルを示す．

上記の最尤推定に加え，22 名のアンケート結果に基づき，各曲に対して主観評定がとりうる値の分布（対数正規分布で近似）を事前確率分布として計算し，尤度と組み合わせた事後確率を最大とする推定も行った（MAP 推定）．

推定精度の評価のため，330 サンプルのうち 1 サンプルを評価データとし，残り 329 サンプルを学習データとする交差検定を繰り返し行った．

4. 結果とまとめ

交差検定で計算された推定値と主観評定値の平均誤差の結果を図 2 に示す．最も推定精度が高かったのは，SVM を用いた MAP 推定で，推定値と主観評定値の平均誤差は 1.26 であった．SVM を用いた最尤推定の平均誤差は 1.47 であった．チャンスレベル（推定値を 1 から 7 までの無作為の値とした場合の主観評定値との平均誤差）は 2.18、尤度を一様分布とした場合の MAP 推定の平均誤差は 1.30 であった．

以上，提案法に基づき眼球計測信号および音響信号から特徴抽出を行い，楽曲の好みの主観評定と対応づける学習を行うことで，音楽の好みをチャンスレベルよりも高い精度で推定することができた．

参考文献

- [1] A. Widmann, R. Engbert, & E. Schröger (2014). Microsaccadic responses indicate fast categorization of sounds: a novel approach to study auditory cognition. *The Journal of Neuroscience*, 34(33), 11152-11158.
- [2] M. Yoneya, H. I. Liao, S. Kidani, S. Furukawa, & M. Kashino. (2014, February). Sounds in Sequence Modulate Dynamic Characteristics of Microsaccades. In 37th ARO MidWinter Meeting.
- [3] Foote, J., & Uchihashi, S. (2001, August). The Beat Spectrum: A New Approach To Rhythm Analysis. In ICME.