



## 生体信号情報の情報通信工学への適用 — 生体信号によるユーザ主観評価の客観的推定 —



早稲田大学 基幹理工学部情報通信学科 教授 かめやま わたる  
亀山 渉

### 1. はじめに

情報通信の分野においては、最終的にサービスを享受するのはユーザであることから、サービスに対するユーザの主観評価を得て、サービス及びネットワークの管理や制御に利用する必要がある。ユーザの主観評価を得る具体的な手法として、従来、アンケートによるユーザ主観評価が広く用いられており、例えば、ITU-R BT.500<sup>[1]</sup>によるMOS評価手法は映像品質に対する主観評価を得るものとして有名である。しかしながら、アンケート等による主観評価には後述する種々の難しさがあるため、サービスに対する様々なユーザの主観評価を精度よく確実に取得できるとは限らない。

筆者はこの点に対し、サービスを受けているユーザの生体信号を観測し分析することによって、サービスに対するユーザの主観的な評価を客観指標として取得し、推定する手法について研究を行っている。本稿では、生体信号情報を情報通信工学へ適用する一つの事例として、生体信号によるユーザ主観評価の客観的推定手法について、筆者の研究内容を交えて論じる。

### 2. ユーザ主観評価取得における従来手法の問題点

例えば、オーディオビジュアルコンテンツ（以下、「コンテンツ」と呼ぶ）の推薦システムを構築する場合、従来手法では、あるユーザのコンテンツの視聴傾向を視聴済みコンテンツに付随するメタデータを用いて分析し、新しい推薦情報を生成するアプローチが一般的である。しかしながら、推薦が受け入れられる割合（推薦精度）はそれほど高くないことが問題となっている。これは、例えば、ジャンル等のメタデータはユーザごとに受け止め方が異なるため、言語で記述されたメタデータからの推定には限界があると考えられる。具体的には、ある人がSF映画と感じているものを別の人はアクション映画と感じている事例、また、ある人がロックと感じている音楽はJポップというジャンルに分類されている事例等が挙げられる。この問題を解決するため、コンテンツ中の歌の歌詞やテキストで書かれたあらすじに機械学習を適用することによりコンテンツが持つムードやジャンルを抽出する試み、また、音楽解析及び映像解析によってそれらを抽出する手法が研究されている。しかしながら、コンテンツのムードやジャンル等は、非常

に個人的な要因に由来してユーザに理解されるものであり、同じコンテンツを視聴しても、ユーザの反応、あるいは、ユーザが生起する情動は、個々のユーザの過去の経験や知識等によって異なる。極端な話、ある人が面白いと思ったコンテンツを別の人は悲しいと感じている可能性がないわけではない。

一方、ネットワーク制御の分野では、QoS (Quality of Service) に代わるものとして、実際にユーザがサービスを体感する品質であるQoE (Quality of Experience) を利用する研究が盛んに行われている<sup>[2, 3]</sup>。ここでの仮定は、QoSが等しい場合であってもユーザが置かれた状況等によってQoEは異なるというものである。ここで、ユーザが置かれた状況とは、単にユーザの置かれた物理的な環境条件だけの問題ではなく、先にも述べたように、ユーザの経験や知識等を含むものと解釈すべきである。実際、ITU-T Rec. P10/G.100<sup>[4]</sup>では、QoEは「アプリケーションあるいはサービスに対するユーザの喜びや苛立ちの程度」と定義されており、QoEに影響を与えるものとして、「アプリケーションあるいはサービスのタイプや特性、使用コンテキスト、ユーザの期待や満足感、ユーザの文化的背景、社会的及び経済的観点、心理的側面、ユーザの情動」等が挙げられている。

このような要因を受けるユーザ主観評価をアンケートで正確に得ることは困難である。特に、ユーザの経験や知識を基にして生起される情動は、次章に述べる点もあり、アンケートのみで取得することは難しい。

### 3. アンケートの難しさ

アンケートの難しさは、以下の3点にまとめられる。

(1) 妥当性を担保することが難しい。

アンケートで知りたいことは、外的刺激によって人が理解する「概念」である。しかしながら、概念は直接測定できない。そのため、概念は「行動や動作」となって表れることを利用してそれらを測定、あるいは観察するが、概念から行動や動作へのマッピングの妥当性、即ち、アンケートの設問事項への妥当なマッピングが問題となる。例えば、「このシステムは使いやすいですか」という設問は簡単で一見よさそうな設問に思える。しかし、「使いやすい」という表現に関する人



の理解は一通りではないため、妥当性が問われる。より妥当な設問としては、「従来に比べてデータ入力する手間は減りましたか」、「従来に比べて入力時間は短縮されましたか」、「従来に比べて出力されたデータの理解はしやすいですか」等が考えられる。

#### (2) 人の特性を考慮することが難しい。

人はしばしば、気を遣い、自分をよく見せようとし、経験に影響し、疲れ、手を抜く。このような人の特性を考慮せずにアンケートを取ると、結果に重大な瑕疵を招く。例えば、「このシステムを日常生活で使いたいですか」という何気ない設問も、上司から部下、研究室内の大学院生から卒論生に向けられると、聞かれた側には気遣い、あるいは、忖度の気持ちが生ずるかもしれない。また、「この映像は怖かったですか」という設問に対しても、この程度で怖がっていると思われたら恥ずかしいという気持ちが働くかもしれない。

#### (3) リアルタイムに答えてもらうのが難しい。

ITU-R BT.500<sup>[1]</sup>にはスライダを利用してリアルタイムに主観評価を計測する方法が記述されてはいる。しかし、実際に実験を行うと、スライダの操作を忘れる、スライダの操作ばかりが気になって評価に集中できない、といった状況が発生しがちである。

## 4. 種々の生体信号情報とその特徴

前章で述べた点を解決する一手法として、生体信号情報の利用は有効であると考えられる。特に、興味、飽き、集中、

喜び、悲しみ、驚きといった人の情動反応を客観的な指標として取得できる可能性がある。情報通信工学分野に利用できると考えられる主な生体信号情報を表1にまとめる。

生体信号に表れる人の反応は個人間で一定の傾向が認められるものの、実際にはかなりの個人差が存在する。そのため、観測された信号から個人差を考慮したデータ分析をする必要がある。これは生体信号を扱う際の難しい点の一つである。例えば、表情解析の研究は近年急速に進歩した分野であるが、個人差を吸収できる分析手法等が検討されている<sup>[7]</sup>。また、アンケートによる回答をグラントゥールズとして利用した教師あり機械学習を適用する場合には、前章で述べた難しさがあるため、注意深く設問内容を検討する等、常に妥当性の検証が求められる。

## 5. 研究事例

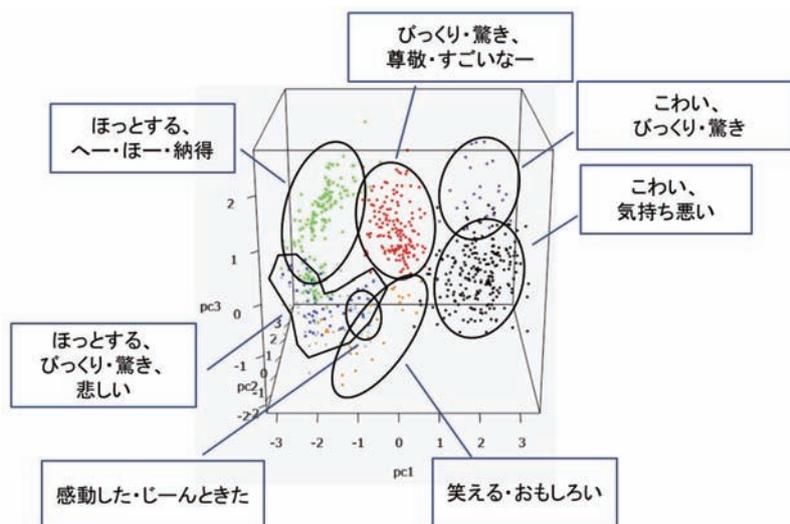
以下、筆者の研究室で行っている2つの研究事例を紹介する。

### 5.1 映像視聴者の情動推定

より精度の高いコンテンツ推薦システムを目指し、映像視聴者がコンテンツから受ける情動を推定する手法を研究している。文献 [8、9、10、11] では、被験者に、驚き、癒し、楽しい、感動、悲しい、怖い、面白い (2種類)、退屈といった情動が生起されると想定する9種類の映像を視聴してもらい、その際の瞳孔径、基礎律動 (背景脳波)、RRI (RR

■表1. 主な生体信号情報とその特徴

生体信号	特徴
視線及び瞳孔径等	<ul style="list-style-type: none"> <li>興味<sup>[5]</sup>、集中、注目点、瞬目等の反応が取得可能。</li> <li>アイトラッカによる非拘束の計測が可能。</li> <li>興味度を測定するためには、環境光による瞳孔径変化の影響を取り除く「対光反射補正」<sup>[6]</sup>が必要。</li> </ul>
脳波	<ul style="list-style-type: none"> <li>様々な情動反応を計測可能。</li> <li>簡易脳波計を利用すれば比較的手軽に測定可能。</li> <li>完全な非拘束性はないが、将来はウェアラブルデバイス等に搭載される可能性も。</li> </ul>
心拍及び血圧等	<ul style="list-style-type: none"> <li>緊張及びリラックス等の情動反応が計測可能。</li> <li>簡易心拍計では脈拍程度の情報が得られないため、心電計、特にR波を計測できるものが望ましい。そのため、装置装着に若干手間がかかる。</li> <li>スポーツ選手の身体状態の測定等でよく利用されている。</li> </ul>
顔面温度	<ul style="list-style-type: none"> <li>特に鼻梁周辺には動脈吻合と呼ばれる特殊な血管があり、鼻梁付近の皮膚温度を得ることでストレス度合いを計測可能。</li> <li>赤外線カメラを利用した非拘束な計測が可能。</li> </ul>
唾液アミラーゼ	<ul style="list-style-type: none"> <li>ストレス度合いを計測可能。</li> <li>計測用のチップを口中に入れる必要がある。</li> </ul>
発汗	<ul style="list-style-type: none"> <li>主として指や手のひらの精神発汗を電気伝導によって測定。</li> <li>緊張度等を計測可能。</li> <li>非拘束には計測できない。</li> </ul>
表情	<ul style="list-style-type: none"> <li>様々な情動反応の測定が可能。</li> <li>カメラを用いた非拘束な手法が利用可能。</li> </ul>



■図. ある被験者の生体信号データとアンケート結果との対応関係（筆者らの文献 [11] の図1より）

間隔、心電波形におけるR波とR波の間隔)を測定する実験を行った。測定したデータに対し、正規化、主成分分析、多段階クラスタリングを施したところ、主成分空間中で図のようなデータ分布とアンケートとの対応関係が得られた。

図から、生体信号データとアンケート結果は非常によく対応していることが分かり、また、主成分空間中では線形分離できない形状に分布しているクラスタの存在も認められる。実験を行った26名の被験者のほとんどで、クラスタ分布や形状は異なるものの、かなりはっきりとした対応結果が得られた。これを基に未知データに対してk-NNによるクラスタ推定を行ったところ、多くの場合に妥当な推定結果が得られた。

一方、線形分離できないという点を考慮し、スパースコーディングによるデータ解析を行った<sup>[12]</sup>。ここでは、上述の生体信号に加え、顔特徴点も併せて解析を行った。得られたスパース係数をクラスタリングによって分析したところ、情動別にまとまって分布するクラスタが存在するという傾向が見られた。また、スパース係数と基底ベクトルに着目した分析からは、特徴的なクラスタは少数の特徴的な基底にのみ反応しているという結果が得られた。加えて、スパース係数分析の結果から、例えば、怖いという情動は生体信号によって大きく特徴付けられ、面白いという情動は顔特徴点によって大きく特徴付けられることが分かった。

以上より、生体信号による映像視聴者の情動推定が可能であることが示唆された。今後の課題としては、他の機械学習手法の適用による推定精度の向上、個人間での類似度分析等が挙げられる。

## 5.2 動画視聴時におけるユーザのQoE推定

QoEによるネットワーク制御等の実現を目指し、動画再生時に品質劣化が生じる場合のQoEを推定する研究を行っている。文献 [13] では、再生中に動画が一時停止する遅延を含んだHD品質の動画像を被験者に視聴してもらい、その際の被験者の瞳孔径、視線移動量、基礎律動及びRRIを測定した。幾つかの異なった遅延条件下で複数の動画を視聴してもらい、1つの動画視聴終了ごとに主観品質評価と動画像に対する興味度をアンケートで取得した。取得したデータに対し、主観品質評価を目的変数とし、瞳孔径、視線移動量、基礎律動、RRI及び動画像に対する興味度を説明変数として、取得したデータの6割をランダムフォレストで被験者ごとに学習させ、残りの4割で被験者ごとの主観品質評価を推定できるかどうかの実験を行った。表2及び表3に被験者10名の実験結果を示す。

主観品質評価値の推定は、興味度を含まない場合は約5割、

■表2. 興味度を説明変数に含めない場合の分類結果（筆者らの文献 [13] の表7より）

		実際の主観品質評価値					適合率
		1	2	3	4	5	
分類結果	1	298	170	54	22	85	47%
	2	180	609	126	30	150	56%
	3	53	99	223	28	60	48%
	4	22	22	15	75	11	52%
	5	74	119	66	13	252	48%
再現率		48%	60%	46%	45%	45%	51%

右下のセルは正答率



■表3 興味度を説明変数に含んだ場合の分類結果  
(筆者らの文献 [13] の表8より)

		実際の主観品質評価値					適合率
		1	2	3	4	5	
分類結果	1	419	103	53	12	106	60%
	2	119	721	137	32	143	63%
	3	40	79	231	21	41	56%
	4	10	32	23	97	2	59%
	5	39	84	40	6	266	61%
再現率		67%	71%	48%	58%	48%	61%

右下のセルは正答率

興味度を含んだ場合は約6割の精度であり、いずれもチャンスレベルである2割を超え、比較的高い精度で主観品質評価値を推定できることが分かった。特に、興味度を含めた場合には精度が10ポイント上がるという結果を得た。また、説明変数の重要度を調べたところ、興味度、瞳孔径、視線移動量の重要度が高かった。

以上から、「興味のある動画では品質劣化が気になるが、興味のないものでは気にならない」、並びに、「興味のある動画では品質劣化が気にならないが、興味のないものでは気になる」という2つのグループが存在する可能性が示唆された。今後の課題としては、より多くの種類の動画と妨害要因の条件下で、より多くの被験者の実験結果を確認して検証する必要がある。

## 6. おわりに

本稿では、生体信号情報を情報通信工学へ適用する事例として、生体信号によるユーザ主観評価の客観的推定手法について、筆者の研究事例を交えて述べた。映像サービスやIoTサービスが更に広まるにつれ、それらのサービスが与える快適度や満足度の評価と分析が重要になってくると考えられる。安価で装着が簡単な生体信号測定装置も利用できるようになってきていることもあり、様々な情報通信工学分野での生体信号情報の有効利用が、今後ますます重要となるであろう。

(2018年2月27日 情報通信研究会より)

## 参考文献

- [1] ITU-R BT.500-13, “Methodology for the subjective assessment of the quality of television pictures”, 2012年
- [2] 黒川 章、江崎 修司、平松 淳、堀越 博文, “次世代ネットワーク (NGN) を支えるネットワーク基盤技術”, 電子情報通信学会、通信ソサイエティマガジン、No.13 [夏号]、pp.10-21、2010年
- [3] 江口 真人、三好 匠、矢守 恭子、山崎 達也, “QoE 状況依存性に係る同一ユーザー群の評価レンジのモデル化”, 情報処理学会論文誌、Vol.54、No.12、pp.2451-2460、2013年
- [4] ITU-T Rec. P.10/G.100, “Vocabulary for performance, quality of service and quality of experience”, 2017年
- [5] Hess E. H., Polt J. M., “Pupil size as related to interest value of visual stimuli”, Science, 132, pp.349-350、1960年
- [6] 加藤 敦士、菅沼 陸、亀山 渉, “ディープラーニングによる瞳孔径の対光反射補正方式に関する検討”, 電子情報通信学会、2017年総合大会、H-2-12、2017年
- [7] 佐々木 康輔、渡邊 健斗、橋本 学、長田 典子, “顔キーポイントの移動方向コードに基づく個人差の影響を受けにくい表情認識”, 電気学会論文誌C、Vol.138、No.5、pp.611-618、2018年
- [8] 傅 櫻、菅沼 陸、亀山 渉、サイモン クリピングデル, “マルチモーダル生体情報による映像視聴時の感情分類に関する検討”, 電子情報通信学会、2016年総合大会、H-2-14、2016年
- [9] 傅 櫻、橋本 稜平、田上 結衣、菅沼 陸、亀山 渉、サイモン クリピングデル, “生体情報による映像視聴者の感情分類分析に関する検討”, 電子情報通信学会、2017年総合大会、H-2-5、2017年
- [10] 傅 櫻、菅沼 陸、亀山 渉、サイモン クリピングデル, “生体信号による映像視聴者の情動分類とアンケート回答との対応関係に関する考察”, 電子情報通信学会、HCGシンポジウム、C-1-5、2017
- [11] 傅 櫻、橋本 稜平、田上 結衣、菅沼 陸、亀山 渉、サイモン クリピングデル, “アンケート回答と生体信号の対応関係を用いたk-NNによる映像視聴者の情動推定に関する検討”, 電子情報通信学会、2018年総合大会、H-2-1、2018年
- [12] 田上 結衣、菅沼 陸、亀山 渉、サイモン クリピングデル, “生体信号と顔特徴点のスパースコーディングによる映像視聴時情動分類に関する一考察”, 電子情報通信学会、HCGシンポジウム、C-1-4、2017年
- [13] 河内 瞭彦、菅沼 陸、亀山 渉, “ランダムフォレストを用いた複数の生体情報等による動画遅延時のQoE推定の検討”, 電子情報通信学会、信学技報CQ2017-124、2018年