

視聴覚心理生理ベースの信号処理

高木, 英行
九州大学大学院芸術工学研究院

<https://hdl.handle.net/2324/4482770>

出版情報 : 電子情報通信学会総合大会 基礎・境界講演論文集. 2008, pp.62-63, 2008-03-18. The Institute of Electronics, Information and Communication Engineers(IEICE)

バージョン :

権利関係 : Copyright(C) IEICE



視聴覚心理生理ベースの信号処理

Signal Processing based on Audio-Visual Psychology and Physiology

高木英行

Hideyuki Takagi

九州大学大学院芸術工学研究院

Faculty of Design, Kyushu University

1. はじめに

信号処理理論は対象信号の性質と信号処理知識を基に数値ベースの処理を行う。処理する信号が主役であり、視聴する人間は脇役に置かれる。これに対し本論文では、視聴者の心理生理反応に基づいて信号処理をする方法論について述べる。従来の制御理論が制御対象のモデル化を目指したのに対し、ルールベースの制御は制御者のモデル化を目指したことに類似した関係ともいえる。

この方法論では、インタラクティブ進化計算 (IEC) [1] の枠組みを用い、視聴者の視聴覚判断や生理反応に基づいて処理された音や画像を評価し、視聴者にとって最適な信号処理フィルタを設計するものである。

この手法が必要になってきた理由の 1 つは、信号処理応用の拡大に伴い、信号処理ユーザが信号処理の専門家でない場合が増えているからである。医療画像診断の専門家や補聴器ユーザは信号処理の専門家でないが、彼らの視覚や聴覚が彼らにとって最適な医療画像強調フィルタ設計や補聴器フィッティングの手掛かりであり、彼らの視聴覚に基づいて信号処理フィルタを自動的に最適化し信号処理する技術が必要になってくる。

以下、このアプローチの具体的な応用例を紹介する。紙面制約の関係で全参考文献を紹介できないので、文献番号なしで紹介している信号処理事例については、これらを引用紹介している IEC 解説論文 [1] を参照されたい。

2. 聴覚ベースの音響信号処理

低域が抑圧され第 1, 第 2 フォルマント (F1, F2) に影響を受けた歪み音を用い、聴感的に歪みを軽減する 8 次 FIR フィルタの係数を IEC で決定する歪み補正フィルタ設計問題を行う。3 名の IEC ユーザが 10, 20, 40 世代目で得たフィルタで処理した音声と歪み原音声と、32 名の被験者が Sheffé の一対比較法で比較した結果、3 名の 10 世代目以降の処理音声はすべて危険率 1% で原音声より有意に聴感的な歪みの減少が認められた [渡辺ら 1995]。図 1 に処理前後の音声パワースペクトルを示す。聴感的に設計したフィルタであるが、特に F1 の回復が顕著である。

音声には /a/ や /i/ などの音素を区別する音韻情報と、個人

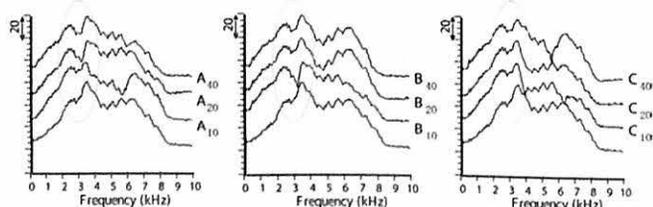


図 1 原歪み音声と、IEC ユーザ A, B, C が設計した 10, 20, 40 世代目の 8 次 FIR フィルタで処理した音声のパワースペクトル。各スペクトルは 20dB ずらして表示している。

性や表情を表す韻律情報があり、韻律情報は、音の強さ、長さ、高さで制御できる。そこで、悲しい声、怒った声のようなターゲット表情の声になるよう、聴覚的評価に基づいて IEC が予め用意した音声のピッチ、強度、長さの 3 物理量を決定する応用がある [佐藤 1996]。

補聴器ユーザの聞こえを計測することは不可能なので、補聴器フィッティングは IEC に適した応用事例である [大崎ら 1998]。

現在の補聴器フィッティングは、静音環境下で、純音、あるいは帯域雑音を補聴器ユーザに提示して計測した可聴特性に基づいて補聴器の初期化を行う。次に、その補聴器を装着したユーザの主観的な聞こえを言語的に評価させる。フィッティングはユーザの聞こえを解釈して、経験と勘に基づいて補聴器パラメータを調整する。

この方法ではデジタル補聴器のように調整するパラメータが増えるとフィッティングが困難になる。また、補聴器フィッティングでは非日常音を使うため、実使用の環境音で最適になる保証がない。そもそも、他人の聞こえは誰にも分からないわけなので、原理的に限界があるフィッティング手法といえる。IEC fitting は、ユーザの聞こえの判断に基づいて最適化手法がフィッティングするので、この原理的な限界をブレイクスルーする手法と言える [2]。

国内トップクラスのフィッティングが調整した商用アナログ補聴器と IEC で調整した同補聴器とを聴覚障害者が実生活で長期間利用して評価した結果を以下に示す。

- | | |
|-----------------------|----------------------|
| (1) 5 名が 1~2 週間使用後に評価 | |
| 単音節認識率: | 両補聴器は同等 |
| 音質評価: | 両補聴器は同等 |
| フィッティング時間: | IEC fitting の方が有意に早い |
| (2) 4 名が半年間使用後に評価 | |
| 音質, APHAB 評価: | 両補聴器は同等 |

人工内耳のフィッティングでは、多くの臨床応用があり、フィッティングの大きな改善効果が示されている [3]。興味深いことは、人工内耳の帯域チャンネル電極の最大-最小強度幅 (C-T 値) が IEC fitting によって狭くなったにも関わらずユーザの聞こえが悪くない事例が見つかったことである。これは人工内耳フィッティングの常識と思われていたことに反する結果であり、IEC fitting によってこれまでの仮説を覆す知見が得られる可能性がある。

3. 視覚ベースの画像処理

フィルタ設計には、前節のように構造が決まっているフィルタの係数を最適化する方法と、遺伝的プログラミング (GP) で生成した数式をフィルタとする方法とがある。

例えば、入力画素濃度(x)から出力画素濃度変(y)への濃度変換フィルタ設計を考えてみよう。 $y=x$ であれば入力と同じ画像が得られ、 $y=-x$ であればポジネガ反転画像が得られる。 $y=f(x)$ になれば、濃度階調が少ないため判別しにくい部分の濃度差異を拡張して見やすくなることもあり、画像強調フィルタとして使える。この関数を線分の組合せで表現し、各線分座標を決定する場合は前者の設計方法になり、GP で生成した関数をそのまま使う場合は後者の設計方法になる。

図 2 は前者の方法で、医療画像診断の専門家が IEC ユーザとして視覚的に診断が可能になるまで繰り返して得られた医療画像強調フィルタの処理結果である。原画像で判別が困難であったリンパ節内の濃度階調が拡張され、濃淡分布の判別が容易になった。

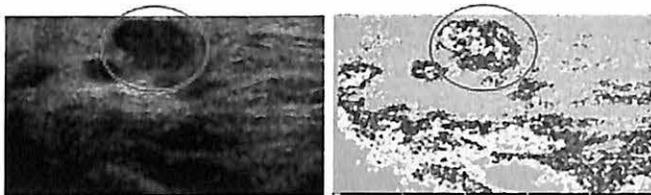


図 2 超音波原画像のリンパ節(左図円内)と IEC12 世代目で得られた画像強調フィルタによる処理後のリンパ節画像(右図円内)

この濃度変換を RGB それぞれに設計することで、グレイ画像をカラー化する色フィルタが設計できる。医療画像に応用すれば、医師が注目する箇所の視覚判断を助けるカラー化ができる可能性がある。

後者の応用例では、超音波画像の動画中の 2 枚の画像を入力し、差異を強調する画像を出力するフィルタ設計がある[Poli ら 1997]。心臓の伸縮画像を用いた例で、例えば、

$$y = 0.55 - g_1 g_2 - g_1^2 g_2 + 0.21 g_1 - 0.75 g_2$$

の関数が GP で得られたとし、 g_1 と g_2 を入力画像 1 と入力画像 2 の同じ位置の画素値、 y を出力画像の同じ位置の画素値とすることで、2 枚の入力画像から 1 枚の出力画像が生成される。IEC ユーザは視覚的に差異が最も分かり易い画像出力に高い評価を与えることで、差異を強調した合成画像が生成できる。

4. 生理ベースの AV 信号処理

映画や音楽の視聴者は、興奮、安らぎ、恐怖などの生理心理的な影響を受ける。多くの場合、コンテンツの文脈が心身に影響を与える主因であるが、同じ内容であっても、音量、色、輝度、時間的変化などの物理特徴の変化によって視聴者の心理生理反応が異なる。1997 年のポケモン事件の場合、赤青 12Hz 点滅という物理特徴が光感受性発作を引き起こすほど大きな生理的影響を与えた。

逆に考えれば、物理特徴を制御することで生理反応を間接的に制御できる可能性があることを示している。よりリラックスした生理状態、よりワクワクドキドキした生理状態(ゴールとなる理想生理状態)に近づける物理特徴を最適手法で見出すことができれば、感動強調フィルタ設計ができる他、デジタルコンテンツ制作上有益になることが期待できる。

この物理特徴の最適化には、図 3 の拡張 IEC[4]の枠組みを用い、fitness には理想生理状態と計測生理状態との差

を用いる。

多種類の物理特徴の多数の特徴量毎に多種類の生理反応を計測することは、時間的に現実的ではない。そこで最初に、生理反応に影響を与える可能性のある物理特徴を心理実験での絞込み、次に生理実験で確認するという、以下の研究ステップを採る。

- (1) メディア物理特徴の抽出と心理実験による絞込み
- (2) 絞り込まれた物理特徴の変化と生理反応の変化の相関計測による、生理反応に影響を与える物理特徴の確定
- (3) 拡張 IEC によるターゲット生理状態に近づける物理特徴の最適化

第 1 ステップは、20 秒の映画 60 シーン各々から、映像特徴 78 種類を抽出する。次に、(緊張-弛緩)軸と(爽快-鬱屈)軸からなる心理空間(情動平面)を用意し、被験者 30 名に各映画シーンの印象を空間座標で与えさせる。60 映像の情動平面上の分布から、情動平面座標に有意に相関がある映像特徴を生理反応に影響を与える可能性のある特徴候補として抽出する。

生理反応に関係があると思われる 2 軸で心理空間である情動平面を構成する点が特徴候補を絞り込むポイントで、この結果、(緊張-弛緩)軸と(爽快-鬱屈)軸に危険率 1% で有意に相関がある映像特徴が 26 種類と 2 種類が各々得られた。現在は、第 2 ステップの生理実験を行っている途中である。

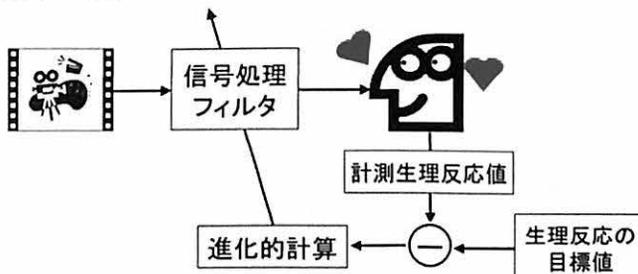


図 3 拡張 IEC を用いた感動強調フィルタ設計

5. まとめ

IEC の枠組みを用い、視聴者の心理生理状態に基づいて進化的計算が信号処理フィルタを設計する、従来の信号処理理論とは異なる方法論とその応用について述べた。

謝辞 本研究は一部文科省 21 世紀 COE プログラム、および、科研費補助金(基盤研究 B, No.17300073)の補助を受けた。

参考文献

- [1] Takagi, H., "Interactive evolutionary computation: fusion of the capabilities of EC optimization and human evaluation," Proc. of the IEEE, 89(9), 1275-1296 (2001).
- [2] Takagi, H. and Ohsaki, M. "Interactive evolutionary computation-based hearing-aid fitting," IEEE Trans. on Evolutionary Computation, 11(23), 414-427 (2007).
- [3] Legrand, P., Bourgeois-Republique, C., et. al., "Interactive evolution for cochlear implants fitting," Genetic Programming and Evolvable Machines, 8(4), 319-354 (2007).
- [4] Takagi, H., Wang, S., and Nakano, S., "Proposal for a framework for optimizing artificial environments based on physiological feedback," J. of Physiological Anthropology and Applied Human Science, 24(1), 77-80 (2005).