

ウェアラブルカメラと生体情報を用いた映像自動編集システム

Automatic editing system using wearable camera and brain waves

I-93

石島 健一郎
Kenichiro Ishijima

相澤 清晴
Kiyoharu Aizawa

中山 雅哉
Masaya Nakayama

東京大学新領域創成科学研究科
School of Frontier Science,, The University of Tokyo

1 はじめに

情報機器が小型化し、ウェアラブルデバイスに関する研究 [4, 5] が盛んに行われるようになってきた。これに伴い、ウェアラブルカメラ等を用いた映像情報が大量に流通するようになると考えられる。しかし映像の取得時間の増大に伴い、記憶に残ったシーンの編集にも長時間が必要となる問題がある。そこで、我々は映像取得とともに同じくウェアラブルな脳波計による生体情報を取得し、自動編集によるシーン生成の研究を行っている。本稿では、映像編集に用いるパラメータを自動的に決定する手法を提案する。

2 映像自動編集方法

2.1 脳波を利用した映像自動編集方法の枠組み

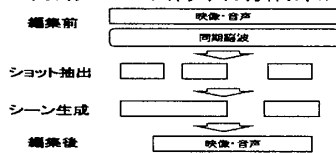


図 1: 映像自動編集の枠組み

映像と同期して記録した脳波を基に、単純な閾値処理により記憶に残った複数の映像(ショット)を抽出できることは [1, 2] で示した。また、脳波特性に基づいた後述の処理により生成される複数の映像(シーン)を結合して要約映像を生成(図 1)できることは [3] で示した。なお以下では α 波のみに着目した処理を示す。

2.2 ショット統合によるシーン生成

脳波に関する生理学的な特性を基に以下の閾値による抽出を行ってきた [3]。 α 波の振幅はつねに一定ではなく、多くの場合 1 秒ないし数秒の周期で漸増漸減を繰り返す [6]。従って、前記閾値処理によって抽出される映像は断片的で細切れの映像である。刺激の時間間隔の不規則性によっても脳波の振幅は増減を繰り返し、局所的に見れば抽出すべきでないショットでも、シーンを生成する上では必要になる場合がある。そこで、振幅閾値条件を満たす映像に挟まれる映像に時間閾値を設定し、条件を満たした場合にショットの統合を行いシーンを生成する。時間閾値は T_1 (500ms) 以下 (シーン生成処理 1) とし、前後のいずれかが抽出されている場合は T_2 (1000ms) 以下 (シーン生成処理 2) とし、前後のいずれも抽出されている場合は T_3

(2000ms) 以下 (シーン生成処理 3) とした。

2.3 抑制処理による統合

同様の刺激が繰り返されると脳波反応が抑制され、誘発電位応答の振幅が半分以下になることが知られている [8]。そこで、抽出されるシーンにおける振幅の最後の極値の R ((100+)60%) に至る時点までの映像を付加して抜き出す (抑制補完処理)。

2.4 遅延処理による統合

脳波は様々な刺激に対する誘発電位応答及び事象関連電位応答の重ね合わせであり [7]、脳波反応に伴う遅延はそれらの潜時に相当すると考えられる。抽出されるシーンの時間閾値 T_p (500ms) 前の映像を付加して抜き出す (遅延補完処理)。

3 閾値の自動決定

前章で述べた閾値は、生体情報の特性に基づいて値を決定しているが、個人差があるため、ここでは、より最適な閾値を見つけ出す手法について述べる。

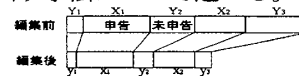


図 2: 指標の定義

3.1 指標の定義

最適な閾値を見つけるために、次に示す指標を定義する。

• 抽出率 (E)

自動編集システムが手動による編集と同じものを抽出することができれば最適である。具体的には、被験者が興味をもったと申告した対象をより多く抽出し、そうでない対象はより少なく抽出することで評価できる。またシーンの価値は本来その長さには依存しないという立場から、上図の例において次式で抽出率 (E) を定義する。なお申告対象数が変わる場合も同様に定義できる。

$$E = \frac{1}{2} \left(\frac{x_1}{X_1} + \frac{x_2}{X_2} \right)$$

• 誤抽出率 (M)

抽出率と同様にして次式で誤抽出率 (M) を定義する。

$$M = \frac{1}{3} \left(\frac{y_1}{Y_1} + \frac{y_2}{Y_2} + \frac{y_3}{Y_3} \right)$$

• 総抽出率 (T)

編集後の映像の編集前の映像に対する割合を総抽出率 (T) として次式で定義する。

$$T = \left(\frac{y_1 + x_1 + y_2 + x_2 + y_3}{Y_1 + X_1 + Y_2 + X_2 + T_3} \right)$$

● 評価関数 (F)
 抽出率 (E) と誤抽出率 (M) を用いて評価関数 (F) を次式で定義する。評価関数 (F) は自動編集システムの精度を直接的に表現する指標である。すなわち評価関数 (F) の値が大きければ大きいほど編集精度が高いものとする。なお、興味対象ではない映像もある程度残した方が映像として見やすいことから、次式における係数 k を定義している。以降においては係数 k の値として 2 を用いることにする。

$$F = kE - M (k > 1)$$

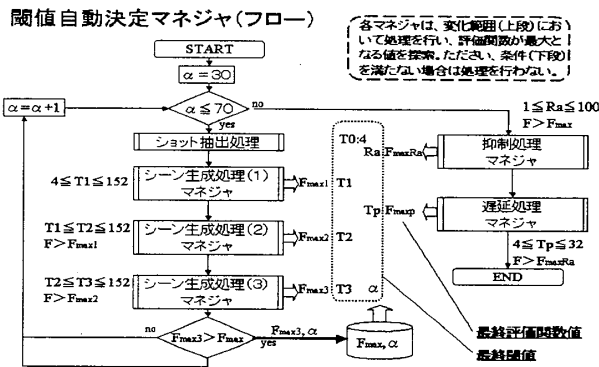


図 3: 閾値自動決定マネージャ

3.2 閾値自動決定方法

次に閾値を自動的に決定する方法について述べる。時間閾値 T_0 については、その値を変化させると T_0 が 8 フレーム (1 秒は 30 フレーム) のときに評価関数が最大になることが多いので固定値とする。

- Step1 振幅閾値 α を偏差値 30 から 70 まで変化させて Step2 から Step5 までの処理を繰り返す
- Step2 時間閾値 T_1 を 4 フレームから 4 フレーム単位で 152 フレームまで変化させてシーン生成処理 1 を行い、評価関数 F を最大にする T_1 を見つける。このときの F の値を F_{max1} とする。
- Step3 時間閾値 T_2 を Step2 で見つけた T_1 から 4 フレーム単位で 152 フレームまで変化させてシーン生成処理 2 を行い、評価関数 F が F_{max1} 以上であるという条件の下にこれを最大にする T_2 を見つける。このときの F の値を F_{max2} とする。
- Step4 時間閾値 T_3 を Step3 で見つけた T_2 から 4 フレーム単位で 152 フレームまで変化させてシーン生成処理 3 を行い、評価関数 F が F_{max2} 以上であるという条件の下にこれを最大にする T_3 を見つける。このときの F の値を F_{max3} とする。
- Step5 F_{max3} が以前の処理で得られた値よりも大きければ、 F_{max} としてそのときの振幅閾値 α 、時間閾値 T_1, T_2, T_3 とともに保存して、Step1 に戻る。

Step6 閾値 R を 1(%) から 100 まで変化させて抑制補完処理を行い、評価関数 F が F_{max} 以上であるという条件の下にこれを最大にする R を見つける。このときの F の値を F_{maxR} とする。

Step7 時間閾値 T_p を 4 フレームから 4 フレーム単位で 32 フレームまで変化させて抑制補完処理を行い、評価関数 F が F_{maxR} 以上であるという条件の下にこれを最大にする T_p を見つける。このときの F の値が評価関数 F の最大値となる。またこのときに保存されていた振幅閾値 α 、時間閾値 T_1, T_2, T_3 、閾値 R 、時間閾値 T_p が最適な閾値となる。

4 日常を素材とした実験

23 歳男性を被験者として府中競馬場における約 47 分の映像と脳波を記録し、後に興味対象を申告させ、2 節における生理学的特性を利用して決定した閾値を用いて自動編集したところ、抽出率 E は 92.2、誤抽出率 M は 30.7、評価関数 F は 153.7 であった。一方で閾値を最適化する手法を用いたところ、抽出率 E は 95.1、誤抽出率 M は 20.3、評価関数 F は 167.2 となって編集精度を向上させることができた。このときの閾値の値は、振幅閾値 α が 44、時間閾値 T_0 が 8 フレーム 時間閾値 T_1 が 140 フレーム、時間閾値 T_2 が 152 フレームであり、シーン生成処理 3 以降の処理では評価関数を大きくする閾値が存在しなかった。閾値の自動決定に要した時間は約 18 分であった。

5 まとめ

本稿では、著者らが提案する映像自動編集システムと閾値を自動的に決定する手法の有効性を示した。

参考文献

- [1] 石島、椎名、相澤:” 個人体験映像の構造化と要約—生体情報を用いた映像要約によるライフメディア—”, 信学技報, PRMU2000-48, pp.51-58, July 2000
- [2] K. Aizawa, K. Ishijima, M. Shiina:” SUMMARIZING WEARABLE VIDEO”, IICIP2001, Vol. -pp.398-401, Oct 2001
- [3] 石島、相澤:” ウェアラブルによる長時間個人体験記録の編集—脳波を利用した映像の自動編集の試み—”, 信学技報, PRMU2000-158, pp.85-92, Jan 2001
- [4] J. Healey and R. W. Picard, ” StartleCam: A Cybernetic Wearable Camera, ” In Proceedings of the Intl. Symposium on Wearable Computers, Pittsburgh, Oct. 1998.
- [5] 広瀬通孝、上岡玲子 他:” 体験記録としてのウェアラブルコンピュータの研究 ” 日本バーチャルリアリティ学会 第 6 回大会論文集, 14A3 p.149-152, Sep. 2001.
- [6] 大熊輝雄:” 臨床脳波学 第 4 版 ” 医学書院, 1991
- [7] 下地恒毅 編:” 誘発電位—基礎から臨床応用まで— ” 西村書店, 1992
- [8] 八木昭宏:” 現代心理学シリーズ 6 知覚と認知 ” 培風館, 1997