

複数の生体信号を利用した要約映像作成に関する一考察

On Consideration of Video Summarization Based On Multi-Bio Signals

大野雄也†
Yuya ONOオン コックメン†
Kok-Meng Ong亀山渉†
Wataru KAMEYAMA

1. まえがき

近年、映像コンテンツの流通量が増加し、それらの内容をコンパクトに把握したいというニーズが高まっており、映像をコンパクトにするために様々な短縮の方法が提案されている。映像の音声情報や輝度などの特徴量を利用して映像を短縮する方法や、映像技法、セリフ、BGMなどを利用して映像を短縮する方法などがある。

しかし近年ではこれらのコンテンツそのものに関する研究だけではなく、コンテンツと我々人間がどのように関係し影響しているかという観点での研究の重要性が増してきている。そこで本研究では映像コンテンツを、視聴者から得られる情報を基に映像を要約する手法について検討する。本稿では心拍動と瞳孔径の二つの生体信号を利用した映像要約についての考察を示す。

2. 映像と生体信号に関する従来の研究

映像視聴中における生体信号に関する研究として、自律神経系反応の変化を計測した研究がある。杉田らはジェットコースターのような強い情動反応を示す映像を提示した時に心拍のパワースペクトルに有意な変化が現れることを示した[1]。また鎌田らは運動体験を示す映像により、発汗や心拍動の変化が計測されることを示した[2]。映像と生体信号の関係性として明確な反応式は見出されていないが、これらの研究に見られるように、映像によって生体信号に反応が現れることは明らかとされている。

これらの生体信号を利用して映像を短縮する研究として、ユーザーの心拍動を利用した手法が豊沢らによって提案されている[3]。ここでは視聴者の心拍動低下と高周波成分を利用して視聴者が覚醒的になった時点のショットを集めて短縮映像を作成する手法を提案している。

また瞳孔径に関して、視聴中の瞳孔径と覚醒度には相関関係があることが先行研究で明らかにされており[4]、映像視聴時における覚醒箇所の特定に利用することができる。

3. 映像要約方法

本研究では心拍動と瞳孔径の二つの生体信号を用いて視聴者の覚醒度や興奮度といった情報を総合的にまた補完的に判断することで精確な映像要約が可能であると考えている。本研究では要約された映像において、ストーリーの再現性に重点を置き作成することとしている。これは作成された要約映像は、元の映像のストーリーを思い出しやすいものであるべきと考えるためである。提案方式では視聴者が覚醒・興奮した箇所について両指標から別々に特定し、それらをつなぎ合わせた要約映像を作成し評価する(図1)。

† 早稲田大学大学院 国際情報通信研究科

† Graduate School of Global Information and Telecommunication Studies, Waseda University

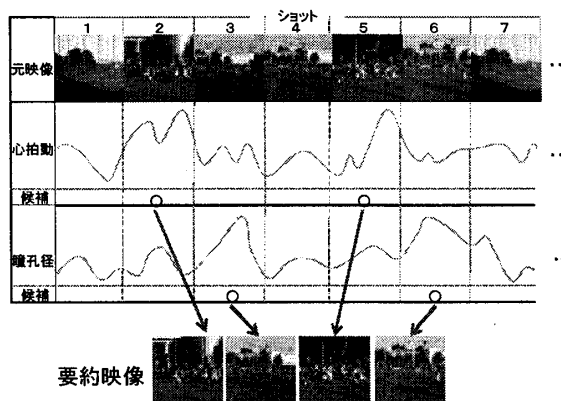


図1. 心拍動と瞳孔径を利用したショットの特定

3. 1 心拍動による抜粋箇所候補

心拍動はどきどきなどの緊張な場合は心拍数が上昇し、リラックスするなどの弛緩な場合は心拍数が下降する。映像視聴中にみられる心拍動の変化は映像の内容による影響があるため、映像視聴中の心拍数の上昇箇所は視聴者の緊張を示し、覚醒度・興奮度の高い箇所であると予想される。このような箇所を心拍動による要約映像への抜粋箇所候補とする。

抜粋箇所候補の特定には以下のような基準を設ける。

- ① 6秒間以上継続して心拍数が上昇した箇所
- ② 視聴中の平均心拍数を上回っている箇所
- ③ 条件①で挙げた箇所の二つの間隔が3秒以内の場合は連続した上昇箇所とみなす

以上の基準をすべて満たす箇所を含むショットを抜粋箇所候補とする。このとき要約映像への抜粋はショット単位で行うが、ストーリーの再現性に重点を置く場合には部分的な映像が飛散的に抜粋されるよりも、ショットというある程度のまとまりで抜粋されるほうがストーリーを理解しやすいと思われる。そのため、得られた心拍データに指数平滑移動平均フィルタをかけデータを平滑化する。これはある時点での心拍動は過去の数秒または数十秒のストーリーによって影響された結果であるとの仮説による。また、平滑化はデータ計測中のノイズを除去するという目的もある。ショットの途中から心拍の上昇開始がみられた場合、ストーリーの再現性を考えるとそのショットも連続して抜粋したほうがよいと思われるので抜粋箇所候補とした。

3. 2 瞳孔径による抜粋箇所候補

瞳孔径は映像の輝度が多分に影響するため、本研究ではニューラルネットワークを介し輝度による影響を削除し瞳孔径として計算している。刺激に対する瞳孔径の反応は0.2

～1.0秒の遅延があるためそれを考慮したうえでCPR (Cumulative Pupil Responce)を求めた[5].

ショット毎にCPR平均値を求め、CPR平均値が全体の上位15%以上であるショットを瞳孔径による抜粋箇所候補とした。しかしショット毎のCPR平均値だけでは、映像の中で特にインパクトが強いと思われるような一瞬の強い瞳孔反応があった箇所を見逃す可能性があるため、CPR値の特に高い箇所を含むショットも抜粋箇所の候補とした。

これら二つの指標ごとに要約映像へ抜粋するショットの候補を決め、どちらかの指標において抜粋すべきショットであると特定されたものを補完的に繋ぎ合わせて要約映像を作成した。

4. 評価手法

提案方式によって作成された要約映像について、次の2種類の評価手法を実施した。

① 要約映像の比較

本方式で作成された要約映像と、等間隔でショットを抜粋した要約映像を被験者に提示し、どちらのほうが元のストーリーがわかりやすいかを答えてもらう。

② スイッチ入力

映像視聴後に再度映像を視聴してもらい、その際に“映像を要約する際に残しておきたい箇所”についてスイッチによってリアルタイム入力してもらった。本方式によって抜粋されたショットと、スイッチ入力されたショットとの適合率と再現率について評価する。

5 実験

5.1 実験環境

被験者は20～30代の男子学生2名でいずれも事前に心拍動・瞳孔径の生体信号を計測し研究として利用することに了承を得てから実験を開始した。実験に使用した映像の長さは44分45秒で、アメリカドラマの第一話でジャンルはサスペンスのものを提示した。

5.2 データの計測

データの計測には以下の機器を利用した。

- ・心拍動：POLAR RS800 (POLAR社製)
- ・瞳孔径：VIS (VIS総研社製)

この二つの機器は無侵襲で視聴中においても装着感を意識することなく計測することができる。

5.3 実験手順

本映像視聴前に瞳孔径を計測するためのキャリブレーションと心拍動を安定させるために安静時間を設けた。

その後被験者は本映像を視聴し、休憩をはさんで評価②で利用するためのスイッチ入力をする。その後実験実施者が作成した要約映像を視聴し、評価をした。(図2)

キャリブレーション	安静 3分	映像 45分	休憩 15分	スイッチ入力 45分	休憩 15分	要約映像視聴	比較評価
-----------	----------	-----------	-----------	---------------	-----------	--------	------

図2. 被験者に対する実験手順

6. 結果

被験者のうちの一人に関する心拍動と瞳孔径の反応箇所とそれらを基にして抜粋候補として特定されたショット数、

要約映像に抜粋されたショット数を表1に示す。本映像全体の長さは44分45秒、ショット数は880ショットである。

表1. 抜粋候補と抜粋ショット

	心拍動	瞳孔径	抜粋ショット	全体
長さ	9m02s	5m21s	13m28s	44m45s
ショット数	158	131	266	880

心拍動からは、3.1の条件により合計6分25秒検出されこの反応箇所を包含するショットとして9分2秒158ショットが抜粋候補として挙げられた。

一方、瞳孔径からは3.2の条件により合計5分21秒、131ショットが挙げられた。

この二つの生体信号ごとに抜粋候補として挙げられたショットを補完的に選択し、本方式で作成された要約映像は266ショット、全体で13分28秒となり元映像の約30%の長さとなった。

7. 評価

①要約映像の比較

本方式により作成された要約映像と等間隔で抜粋された要約映像を被験者に比較してもらったところ、両映像に“わかりやすさ”の差はほとんどないという評価であった。

②スイッチ入力

スイッチによる“被験者が要約映像のために残したい箇所”と本方式での要約映像の適合率と再現率を表2にまとめた。被験者がスイッチ入力したショットは653ショットでそのうち190ショットが要約映像中に含まれており、適合率は71.4%、再現率は29.1%であった。

表2. スイッチ入力との比較

スイッチ入力されたショットの数	653
抜粋したショットの数	266
マッチしたショットの数	190
適合率	71.4%
再現率	29.1%

8. まとめと今後の課題

本稿では生体信号を利用した映像要約に関して、心拍動と瞳孔径を補完的に用いることでその精度を上げ、ストーリーの再現性の高い映像を作成することを試みた。現段階ではその精度は高くないが、今後は心拍動と瞳孔径の相関についても検討し、映像に対する心拍動と瞳孔径の反応速度等についても実験を行っていく予定である。

参考文献：

- [1] 杉田典大,吉沢誠,田中明,阿部健一,山家智之,仁田新一:血圧・心拍数間の最大相互相関係数を用いた映像刺激の生体影響評価,ヒューマンインターフェース学会論文誌,Vol.4,No.4,pp227-233,(2002)
- [2] 鎌田幹夫,坂東武彦,黒岩義之:映像コンテンツ視聴時の生体信号計測 評価,2001,ソニー(株)インターネット研究所 <http://acordo.jp/lido/uploads/paper/JMAC2001KAMADAfinal.pdf>
- [3] 豊沢聡,河合隆:視聴者の心拍活動を用了映像短縮方法とその評価,映像情報メディア学会誌,vol.63,No.1,pp86-94,2009
- [4] T.Partara,V.surakka, *Pupil size variation as an indication of affective processing*,Int.J.Human-ComputerStudies,volume59, issue1-2,pp185-198(2003)
- [5] Kok-meng ong, w.kameyama: Classification of Video shots Based on Human Affect,映像情報メディア学会誌,Vol.63,No.6,pp847-856(2009)