

映像からのアクション性とドラマ性の自動抽出とメタデータとしての活用

田中 壮詩[†] 平嶋 宗[†] 吉高 淳夫^{††}

[†] 広島大学大学院 ^{††} 北陸先端科学技術大学院大学

映像の文法によれば、アクション性はショットの長さによって、ドラマ性はカメラワークによって強調されるとされている。本稿では映画を対象として、映像の構造のみから自動でアクション性とドラマ性を抽出する手法を提案し、「ショット率とカメラワーク率」と「人によるアクション性とドラマ性の判断」の関係についての調査について述べる。さらに、この関係を踏まえた上で、ショット率とカメラワーク率から映画のアクション性とドラマ性を推定する方法を提案する。また、映画の選出に役立つようなメタデータの活用方法についても考察する。

Automatic Estimation of Action and Drama Degree in Movies

Takeshi TANAKA[†] Tsukasa HIRASHIMA[†] Atsuo YOSHITAKA^{††}
[†] Hiroshima University ^{††} Japan Advanced Institute of Science and Technology

According to film grammar, physical movement is emphasized by shortness of shot and mental movement is emphasized by use of camera work. In this text, for a movie, the method of automatic extraction of action and mental movements from the structure of the movie is proposed, and the investigation about relationship between “the rate of shot and the rate of camera work” and “judgment of action and mental movements by human” is described. In addition, the method of presuming “action and mental movements” from “the rate of shot and the rate of camera work” is proposed. And method of using useful metadata for election of movie is considered.

1. はじめに

近年、膨大な量のマルチメディアコンテンツを蓄積することが可能になり、大量のコンテンツの中からユーザが自分の嗜好に合ったコンテンツを選出することは困難になっている。これを踏まえ、コンテンツに対しメタデータを付与し、コンテンツを効率的、また柔軟に扱う技術に注目が集められている。

映画の場合はメタデータとして、ジャンルごとに分類されることが多く、同一系統に映画をグループ分けする研究[1]が行われている。

服部ら[1]はカット点の情報やカメラパラメータ、色情報など複数の特徴を抽出し、抽出した特徴を基準とする映像のグループ分けを行っている。

しかしながら、単に類似した特徴の映像をグループ分けしているだけであるため、映画を特徴付けるグループ分けの基準を考慮していない。

そこで、映画を特徴付ける基準として、映画を撮影する際に、視聴者を惹きつけるために念頭に置くような映画にとって重要な特性に着目した。視聴者を惹きつける映画にとって重要な特性を抽出することで、視聴者の映画に対する評判と同様の傾向を持つ情報を抽出

出できるのではないかと考えた。

映画にとって重要な特性はアクション性とドラマ性であることが、文献[3]や研究[2]で述べられている。

映画はアクション映画やドラマ映画といったジャンルに分ける方法は既にあるが、ジャンル分けは大量の映画の中からユーザが自分の嗜好に合った映画を選出する上では大雑把な方法であるといえる。なぜなら、同じジャンルの映画でもアクション性とドラマ性の度合いは異なり、また、それらは必ずしも相反するとは限らないからである。そのため、個人の嗜好に合った映画を選出するには、各映画のアクション性やドラマ性の度合いといった、より詳細な情報を付与することが望ましいが、手動で付与することは困難である。効率的に映画を選出するためには、視聴者の映画に対する評判となり得るアクション性とドラマ性を自動で抽出し、メタデータとして活用することが必要である。

映画の文法などの映画の撮影・編集規則が記されている文献[3]-[6]によると、アクション性はショットの長さにより、ドラマ性はカメラワークにより強調されると述べられている。本研究では映像の構造のみから自動で抽出する手法を提案し、「ショットの長さを表すシ

ショット率と、カメラワークが用いられている数を表すカメラワーク率」と「人によるアクション性とドラマ性の判断」の関係についての調査について述べる。さらに、この関係を踏まえた上で、ショット率とカメラワーク率から映画のアクション性とドラマ性を推定する方法を提案し、映画の選出に役立つようなメタデータの活用方法についても考察する。

本提案手法の実現により、大量化する映像アーカイブの効率的な管理が可能となる。これにより、ユーザーの希望するコンテンツの選出を容易にし、個人の嗜好に応じた情報配信につながる事が期待できる。

本稿では、2章でアクション性とドラマ性について説明する。3章では撮影・編集技法について述べる。4章では、映像の構造とアクション性・ドラマ性の関係について説明する。5章で「ショット率とカメラワーク率」と「人によるアクション性とドラマ性の判断」に相関性があるのかを調査した実験とその結果について述べ、6章でまとめとする。

2. アクション性とドラマ性

映画において重要な特性はアクション性とドラマ性であることが、文献[3]や研究[2]で述べられている。映画の文法[3]によると、映画の各場面はアクションとドラマに分類可能であると述べており、映画の構造表現に関する研究である高橋らの研究[2]によると、演劇評論家の検討から映画はアクション性とドラマ性により評価可能であると述べている。

[2][3]を参考にアクション性とドラマ性を定義した。

(1) アクション性

物理的な動きの情報

(2) ドラマ性

心理的な動きの情報

映画の文法などの映画の撮影・編集規則が記されている文献[3]-[6]によると、ショットの長さやカメラワークなどの撮影・編集技法により、アクション性とドラマ性は強調されると述べられている。

3. 撮影・編集技法

計算機で検出可能な撮影・編集技法が主に三つある。

(1) ショットの長さ

単一のカメラで撮影された映像の長さ

(2) カメラワーク

ズームインなどのカメラ操作

(3) BGM

背景として流す音楽

(1)-(3)の撮影・編集技法は先行研究により検出可能である。[7]-[9]

[3]-[6]によると、ショットの長さを短縮することにより映像上の激しさを強調することから、ショットの短さを表すショット率によりアクション性を抽出できるのではないかと仮定した。

また、カメラワークにより人物の心理面を強調することから、カメラワークが用いられている数を表すカメラワーク率によりドラマ性を抽出できるのではないかと仮定した。

ショット率 S_R は合計ショット数 S_n を映像の長さ V_T [分]で割った値を表す。

$$S_R = \frac{S_n}{V_T}$$

カメラワーク率 C_R は合計ショット数 S_n の内で、カメラワークが存在するショット数 C_n の率で表す。

$$C_R = \frac{C_n}{S_n} * 100$$

関連研究として菅野ら[10]は、ショットの長さやBGMの音量等の撮影・編集技法に基づき、映画の各ショットをアクションとドラマに分類しており、アクション性とドラマ性の度合いについては考慮していない。

しかしながら、効率的に映画を選出するためには、映画の各ショットの分類結果ではなく、アクション性とドラマ性の度合いをメタデータとして用いることが有効であると考えられる。

本研究では撮影・編集技法に基づき、アクション性とドラマ性の度合いを抽出することにより、個人の嗜好に合った映画の選出に有効なメタデータを付与することを旨とする。

4. 映像の構造とアクション性・ドラマ性の関係

映像の構造から得られるショット率・カメラワーク率と、映画の重要な特性であるアクション性・ドラマ性の関係を調査した。

調査する際、アクション映画はアクション性が高く、ドラマ映画はドラマ性が高いと考えられるので、インターネットで利用できる映画データベースとして最も知られているThe Internet Movie Database[11]の既存の分類を参考に、アクション映画20本、ドラマ映画20本の計40本の映画を用いた。

4.1 ショット率・カメラワーク率とアクション性・ドラマ性との関係

ショット率とカメラワーク率に相関性があるのかを調べた。図1に調査した結果を示す。

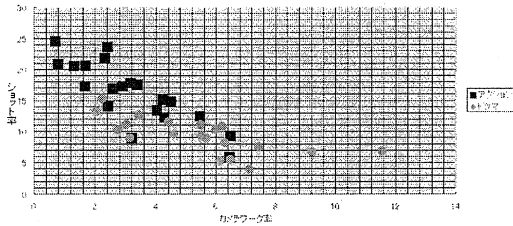


図1 ショット率とカメラワーク率の関係

アクション映画において、ショット率とカメラワーク率との相関係数は-0.83 となり、ショット率とカメラワーク率の間に強い負の相関があるといえる。

ドラマ映画において、ショット率とカメラワーク率との相関係数は-0.71 となり、同様にショット率とカメラワーク率の間に強い負の相関があるといえる。

またショット率とカメラワーク率との回帰係数は、アクション映画において-2.47 となり、ドラマ映画において-0.85 となった。

これらの結果と図1 から、ショット率が高い映画はアクション映画であり、カメラワーク率が高い映画はドラマ映画であり、ショット率とカメラワーク率の両方とも高くも低くもない映画にはアクション映画とドラマ映画が混在している。

次節で、アクション映画ではショット率が高くカメラワーク率が低くなり、ドラマ映画ではショット率が低くカメラワーク率が高くなるのかを調べるために、ショット率とカメラワーク率の傾向の違いを比較する。

4.2 ショット率・カメラワーク率の傾向の違い

アクション映画とドラマ映画での、ショット率とカメラワーク率の傾向の違いを比較した。ショット率の傾向の違いを表1に、カメラワーク率の傾向の違いを表2に示す。

表1 ショット率の平均値と標準誤差

	アクション	ドラマ
平均値	16.3	9.3
標準誤差	1.1	0.6

表2 カメラワーク率の平均値と標準誤差

	アクション	ドラマ
平均値	3.2	5.4
標準誤差	0.4	0.5

アクション映画とドラマ映画のショット率の平均値において、t検定による両側確率が 0.4×10^{-5} となり、アクション映画とドラマ映画のショット率の平均値には

有意差があるといえる。

アクション映画とドラマ映画のカメラワーク率の平均値において、t検定による両側確率が 0.1×10^{-2} となり、アクション映画とドラマ映画のカメラワーク率の平均値には有意差があるといえる。

ショット率・カメラワーク率以外で映像の構造から得られる BGM 率では、アクション映画とドラマ映画で傾向が違ってくるかを調査した結果を表3に示す。

BGM 率 B_R は合計ショット数 S_n の内で、BGM が流れるショット数 B_n の率で表す。

$$B_R = \frac{B_n}{S_n} * 100$$

表3 BGM 率の平均値と平均誤差

	アクション	ドラマ
平均値	30.1	34.6
標準誤差	2.5	2.8

アクション映画とドラマ映画の BGM 率の平均値において、t検定による両側確率が0.3 となり、アクション映画とドラマ映画の BGM 率の平均値には有意差があるといえず、標準誤差も大きい傾向となった。BGM を細かく分析すればアクション映画とドラマ映画の違いが出る可能性はあるが、BGM の量だけではアクション性・ドラマ性を抽出は困難であることから、本研究ではBGM を用いない。

表1と表2よりアクション映画では、ショット率が高くカメラワーク率が低くなり、ドラマ映画ではショット率が低くカメラワーク率が高くなることから、本研究ではショット率によりアクション性を抽出し、カメラワーク率によりドラマ性を抽出する。

しかしながら、ショット率・カメラワーク率が、高くも低くもない映画はどのような映画なのかが不明瞭になるため(図2)、映画全体を分割した範囲でのショット率・カメラワーク率も提示することで、よりコンテンツ管理に有効なメタデータとして利用できるのではないかと仮定した。

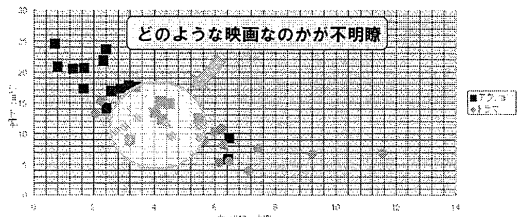


図2 ショット率とカメラワーク率の関係

映画を分割する際、2分割では少なすぎると考え、まずは3分割し、ショット率・カメラワーク率の傾向に違いが出るのかを調査した。閾値を設け、映画を分割した範囲内でショット率・カメラワーク率がそれぞれ閾値以上かを見ることで、傾向を比較した。ショット率・カメラワーク率のそれぞれの閾値は、40本の映画（アクション映画20本、ドラマ映画20本）における平均値に設定した。ショット率の閾値は12.8、カメラワーク率の閾値は4.3となった。アクション映画を用いてショット率の傾向を比較した結果の一例を表4に、カメラワーク率の傾向を比較した結果の一例を表5に示す。ドラマ映画を用いてショット率の傾向を比較した結果の一例を表6に、カメラワーク率の傾向を比較した結果の一例を表7に示す。

表中で、映画全体のショット率を S_R の列に、カメラワーク率を C_R の列に、3分割した範囲内で閾値以上の範囲を○で示した。

表4 アクション映画でのショット率の傾向

	S_R	序盤	中盤	終盤
Back to the Future	12.2			○
The Last Samurai	20.6	○	○	○

表5 アクション映画でのカメラワーク率の傾向

	C_R	序盤	中盤	終盤
Back to the Future	4.3	○	○	
The Last Samurai	1.7			

表6 ドラマ映画でのショット率の傾向

	S_R	序盤	中盤	終盤
Cube	11.6	○		
Forrest Gump	6.8			

表7 ドラマ映画でのカメラワーク率の傾向

	C_R	序盤	中盤	終盤
Cube	4.4		○	○
Forrest Gump	11.6	○	○	○

表4-7のように、映画によって映画を分割した範囲内でのショット率とカメラワーク率が変化する傾向に様々なパターンがあることが分かった。映画全体のショット率が高い映画は序盤から終盤にかけてショット率が高くなり、映画全体のカメラワーク率が高い映画は序盤から終盤にかけてカメラワーク率が高くなり、映画全体のショット率とカメラワーク率の両方とも高くも低くもない映画は序盤から終盤の分割した範囲ではショット率とカメラワーク率はどちらかに偏りのあ

るような変化をした。

5. 相関性を調査する実験と考察

「ショット率とカメラワーク率」と「人によるアクション性とドラマ性の判断」に相関性があるのかを調査する実験を行った。

実験で使用した映像データの形式は、フレームサイズ720×480[pixel]、フレームレート29.97[frames/sec.]、24ビットカラーである。

映画1本の被験者数を3名とし、20本の映画を大学生、大学院生の19名に視聴してもらい、視聴した映画の序盤や中盤、終盤、映画全体のアクション性・ドラマ性を評価してもらい。評価は5段階でしてもらい、アクション性とドラマ性がそれぞれまったくないと感じた場合1を、ほとんどないと感じた場合2、ある程度あると感じた場合3、強いと感じた場合4、非常に強いと感じた場合5をつけてもらった。

ショット率と人によるアクション性の判断の評価結果を図3と図4に、カメラワーク率と人によるドラマ性の判断の評価結果を図5と図6に示す。図中の番号が表す映画のタイトルを表8に示す。

表8 図2-5での番号が表す映画タイトル

図中の番号	タイトル
1	12 angry men
2	A Beautiful Mind
3	A Clockwork Orange
4	Awakenings
5	Back to the Future
6	Back to the Future Part II
7	Back to the Future Part III
8	Beaches
9	Bicentennial Man
10	Breakfast at Tiffany's
11	Buffalo '66
12	Close Encounters of the Third Kind
13	Firewall
14	Forrest Gump
15	Frequency
16	Face / Off
17	I, Robot
18	Independence Day
19	Jumanji
20	The Rock

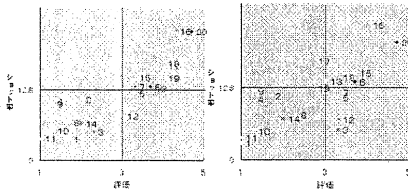


図3 映画全体(左)と序盤(右)での
ショット率とアクション性の評価

映画全体でのショット率と人によるアクション性の判断との相関係数は0.62となり、ショット率と人によるアクション性の判断の間に強い正の相関があるといえる。

また、ショット率が平均値より大きい映画の中で、人によるアクション性の評価結果が3より大きい映画が9本あり、ショット率が平均値以下の映画の中で、人によるアクション性の評価結果が3以下の映画が10本あった。ショット率は大きいアクション性の評価は小さい映画はなかったが、ショット率は小さいアクション性の評価は大きい映画が1本あった。

序盤でのショット率と人によるアクション性の判断との相関係数は0.70となり、ショット率と人によるアクション性の判断の間に強い正の相関があるといえる。

また、ショット率と評価結果が大きい映画が6本あり、ショット率と評価結果が小さい映画が9本あった。ショット率は大きい評価結果は小さい映画が1本あり、ショット率は小さい評価結果は大きい映画が4本あった。

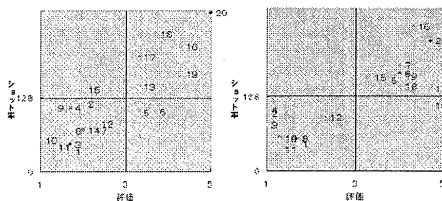


図4 中盤(左)と終盤(右)での
ショット率とアクション性の評価

中盤でのショット率と人によるアクション性の判断との相関係数は0.65となり、ショット率と人によるアクション性の判断の間に強い正の相関があるといえる。

また、ショット率と評価結果が大きい映画が6本あり、ショット率と評価結果が小さい映画が11本あった。ショット率は大きい評価結果は小さい映画が1本あり、ショット率は小さい評価結果は大きい映画が2本あった。

終盤でのショット率と人によるアクション性の判断との相関係数は0.82となり、ショット率と人によるア

クション性の判断の間に強い正の相関があるといえる。

また、ショット率と評価結果が大きい映画が9本あり、ショット率と評価結果が小さい映画が10本あった。ショット率は大きいアクション性の評価はなかったが、ショット率は小さいアクション性の評価は大きい映画が1本あった。

映画全体、序盤、中盤、終盤の結果をまとめると、ショット率と評価結果が大きい映画が30本あり、ショット率と評価結果が小さい映画が40本あった。ショット率は大きい評価結果は小さい映画が2本あり、ショット率は小さい評価結果は大きい映画は8本あった。アクション性があると判断された映画を正解とした場合、適合率は0.79、再現率は0.94となり、これらのことから、ショット率と人によるアクション性の判断はほぼ一致しているといえる。

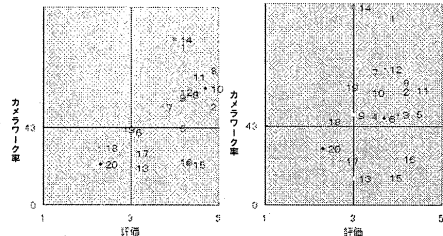


図5 映画全体(左)と序盤(右)での
カメラワーク率とドラマ性の評価

映画全体でのカメラワーク率と人によるドラマ性の判断との相関係数は0.24となり、カメラワーク率と人によるドラマ性の判断の間にやや正の相関があるといえる。

また、カメラワーク率が平均値より大きい映画の中で、人によるドラマ性の評価結果が3より大きい映画が11本あり、カメラワーク率が平均値以下の映画の中で、人によるドラマ性の評価結果が3以下の映画が6本あった。カメラワーク率は大きい評価結果は小さい映画はなかったが、カメラワーク率は小さい評価結果は大きい映画が3本あった。

序盤でのカメラワーク率と人によるドラマ性の判断との相関係数は0.26となり、カメラワーク率と人によるドラマ性の判断の間にやや正の相関があるといえる。

また、カメラワーク率と評価結果が大きい映画が13本あり、カメラワーク率と評価結果が小さい映画が3本あった。カメラワーク率は大きい評価結果は小さい映画が2本あり、カメラワーク率は小さい評価結果は大きい映画が2本あった。

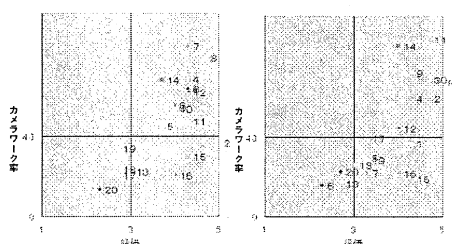


図6 中盤(左)と終盤(右)でのカメラワーク率とドラマ性の評価

中盤でのカメラワーク率と人によるドラマ性の判断との相関係数は0.28となり、カメラワーク率と人によるドラマ性の判断の間にやや正の相関があるといえる。

また、カメラワーク率と評価結果が大きい映画が11本あり、カメラワーク率と評価結果が小さい映画が5本あった。カメラワーク率は大きい評価結果は小さい映画はなかったが、カメラワーク率は小さい評価結果は大きい映画が3本あった。

終盤でのカメラワーク率と人によるドラマ性の判断との相関係数は0.57となり、カメラワーク率と人によるドラマ性の判断の間に正の相関があるといえる。

また、カメラワーク率と評価結果が大きい映画が9本あり、カメラワーク率と評価結果が小さい映画が4本あった。カメラワーク率は大きい評価結果は小さい映画はなかったが、カメラワーク率は小さい評価結果は大きい映画が7本あった。

映画全体、序盤、中盤、終盤の結果をまとめると、カメラワーク率と評価結果が大きい映画が43本あり、カメラワーク率と評価結果が小さい映画が19本あった。カメラワーク率は大きい評価結果は小さい映画が2本あり、カメラワーク率は小さい評価結果は大きい映画が15本あった。ドラマ性があると判断された映画を正解とした場合、適合率は0.74、再現率は0.96となり、これらのことから、カメラワーク率と人によるドラマ性の判断はほぼ一致しているといえる。

これらを踏まえると、ショット率とカメラワーク率は、アクション性とドラマ性の抽出に有効な指標であるといえる。しかしながら、カメラワーク率と人によるドラマ性の判断の相関係数が小さいので、ドラマ性の抽出方法を改良する必要がある。

6. まとめと今後の課題

本研究では、映像の構造だけにに基づき人間の感覚に応じた情報を抽出する手法を提案した。また映画を3

分割することにより、映画によってショット率・カメラワーク率は様々な変化をすることが分かった。そして実験により、ショット率とカメラワーク率は、アクション性とドラマ性を抽出する指標として有効であることを示した。

今後の課題として、実験が不十分なので引き続き実験を続けることや、ショット率・カメラワーク率を高いと判断する閾値の設定をすること、またドラマ性の抽出方法を改良することや、ショット率・カメラワーク率を提示することで、個人の嗜好に応じた情報配信につなげる具体的な方法を提案することが挙げられる。

参考文献

- [1] 服部しのぶ, 亀山渉, 富永英義, “映像特徴空間に基づく同系映像の分類”, 電子情報通信学会技術研究報告, Vol. 103, pp. 55-60, 2003
- [2] 高橋靖, 長谷川桂介, 杉山和雄, 渡辺誠, “セマンティックスコア法を用いた映画の構造表現”, デザイン学研究, Vol. 46, No. 6, pp. 57-66, 2000
- [3] ダニエル・アリホン著, 岩本憲児, 出口丈人訳, “映画の文法”, 紀伊国屋書店, 1980.
- [4] ジェレミー・ヴィンヤード著, 吉田俊太郎訳, “傑作から学ぶ 映画技法完全レファレンス”, フィルムアート社, 2002
- [5] ルイス・ジアネッティ著, 堤和子, 増田珠子, 堤龍一郎訳, “映画技法のリテラシー I 映像の法則”, フィルムアート社, 2003
- [6] 今泉容子著, “映画の文法—日本映画のショット分析”, 彩流社, 2004
- [7] 石井孝和, 吉高淳夫, 平川正人, 市川忠男, “映画の文法に基づくビデオ画像の内容検索”, 情報処理学会研究報告, Vol. 97, No. 7, pp. 65-72, 1997
- [8] 松井亮治, 吉高淳夫, 平嶋宗, “カメラワークを利用した演出効果の抽出”, 電子情報通信学会技術研究報告, PRMU2004-163-178, Vol. 104, No. 573, pp. 25-30, 2005
- [9] 濱村正治, 吉高淳夫, 平川正人, 市川忠男, “映画における音楽、効果音(SE)の印象評価”, 電子情報通信学会技術研究報告, Vol. 99, No. 723, pp. 69-74, 2000
- [10] M. Sugano, R. Isaksson, Y. Nakajima, H. Yanagihara, “Shot Genre Classification Using Compressed Audio-Visual Features”, IEEE ICIP 2003, Vol.2, pp.17-20, 2003
- [11] “The Internet Movie Database (IMDb)”, <http://www.imdb.com>