

映像文法に基づいた映像編集支援システムのための使用可能ショット区間自動抽出と索引付け

Automatic Usefull Shot Extraction and Indexing for a Video Editing Support System
I-47
Based on Video Grammar熊野 雅仁†
Masahito Kumano有木 康雄†
Yasuo Ariki

1. はじめに

デジタル放送時代を迎え、放送の多チャンネル化やインタラクティブTVなどの出現により、映像コンテンツの不足や作業量の増大が問題となっている。我々は、この問題に対応するアプローチとして、映像文法に基づいた映像編集支援システムの開発に取り組んできた [1]。本稿では、映像編集支援システムにおいて索引付けを行うために必要となる処理について報告する。

2. 映像編集支援システムと映像文法

映像編集支援システムでは、カメラマンが撮影した素材映像を受け取った時点から、映像を完成させるまでの作業、つまりこれまで編集者が行ってきた作業を、以下のような3つのフェーズに分けている。図1は映像編集支援システムの中で最低限の機能を有するショットサイズを基盤としたシステムの全体像である。

フェーズ1:	素材映像の解析と索引付け
フェーズ2:	クリップの切り出し
フェーズ3:	映像編集支援部

我々のアプローチでは、フェーズ1、2の完全自動化を目指し、フェーズ3では、接続可能なClipの候補だけを編集者に一覧表示する。これにより、編集者は、非効率なフェーズ1、2の作業から解放され、フェーズ3内で、映像表現の要となるClipの接続と、時間長の調整のみに作業を集中することが可能となる。

ここで、接続可能なClipとは、放送用に使用可能な映像の区間を切り出した物理的な断片であり、表1に示す映像文法中、rule(6)に従って、FIX(カメラ固定)のみ(図2-Clip1,2,4)とPANやZOOMなどのカメラワーク前後にFIXがある図2-Clip3のような2つのタイプが考えられる。

表1: 映像文法抜粋

rule(1):	シーンの冒頭はマスターショット(通常LS)で始める
rule(2):	MSとMSは接続できない
rule(3):	TSとTSは接続できる
rule(4):	同じ被写体で同じショットサイズは接続できない
rule(5):	ショットサイズの急激な変化(LS-TS)は避ける
rule(6):	カメラワークの前後には1秒以上のフィックスが必要
rule(7):	カメラワークは安定していること
rule(8):	被写体が静止しているフィックスは15秒まで
rule(9):	LSは6秒、MSは4秒、TSは2.5秒程度の長さとする
rule(10):	シーン中最初と最後のショットは2秒増やす

ショットとは、構図や撮影対象に依存した時間方向に連続した区間であり、論理的な映像の最小単位である。また、ショットサイズとは、被写体の大きさ、見方を変えれば被写体とカメラ(映像を見る側)との距離を表しており、あるショットより被写体に近寄ったショットをタイトショット(TS)、引いたショットをルーズショット

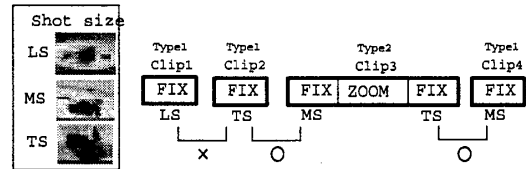


図2: ショットサイズとクリップの接続

(LS), 両者の中間となるショットをミドルショット(MS)という(図2左)相対的なショットの分類指標である。

ショットサイズは、FIX区間に対して付与される索引であり、区間内で変化しない。また、表1-rule(9)に従って、Clipの時間長を決定する。このショットサイズを元にしたClip列によって時空間の抽象化が行われる。映像文法に基づいた編集とは、このClipを接続する際に接点となるFIX部どうしを表1-rule(1)~(5)のような規則に従って行う作業である。

Phase1では、放送用に使用可能なClipを自動抽出するために、カット点やカメラワーク、ショットサイズなどの索引情報を付与する処理部であるが、素材映像中には放送用として使用できない手ぶれ、取り直しに関わるカメラワークの不安定な区間(表1-rule(7)), もしくは、カメラパラメータの調整や、意図した撮影区間の前後にあるカメラの撮影方向調整区間などの使用不能区間がある。また、撮影開始の合図を行う際のカメラマンの手が入った区間や、撮影の行われていない区間(Blank), さらに映像文法を用いてClipを切り出す文法が定義できないFollow(手持ちカメラで動く被写体を追いかける)区間などもある。Phase1を自動化するためには、素材映像の使用不能区間を除去し、使用可能区間から映像文法に沿ったClipを抽出する必要がある。

以下に、自動化されたフェーズ1を実現する方法について述べ、使用可能・不能区間の推定法とショットサイズ判定法について述べる。

3. 使用可能区間推定

素材映像中のカット点はカメラのON・OFFに対応している。素材映像に対してまずカット点検出を行い、カット点で挟まれた区間(Cut)へ分割する。現在、バッファリング法を用いた方法[1]で、実時間処理が可能であり、正解率94.3%、適合率90.9%を達成している。カメラワーク解析法については、現在、輝度投影法[2]を用い、実時間でPAN・ZOOMの変化量を求めている。この変化量を用いれば、表2の特徴により、それぞれの区間推定が可能である。

これらは、時間軸上、まとまりのある区間ごとにカメラワーク変化量の密度と短区間窓の分散値積和量との積

†龍谷大学理工学部

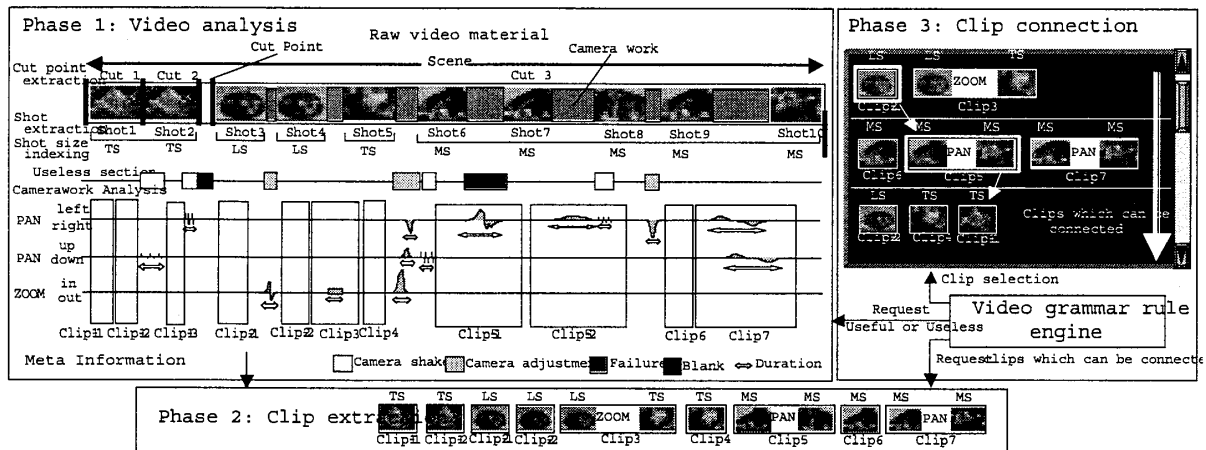


図 1: 映像編集支援システム

表 2: カメラワークの変化量と区間の分割特徴

	連続性	値・変化
Follow	不連続	断続的に大きな変化が続く
手ぶれ	不連続	値は小さい（正負にふれる）
急峻	連続	短い区間で変化がある
安定	連続	短い区間で変化が少ない

により判定を行うことができる [3]。まず、この指標を用いてカットごとに Follow 区間か否かを判定する。次に手ぶれ区間か否かを判定し、最後に急峻であるか安定であるかを判定する。ただし、Blank 区間検出については投影量で判定が行えるため、カメラワーク解析内で判定処理を行っている。表 3 に実験結果を示す。実験映像は、素材映像 1 本、320x240,30fps の MPEG1 映像を対象にフレーム精度で評価を行った。処理時間は 1 分程度である。括弧内の数値は誤検出した Follow 区間推定が正しかった場合を想定した結果である。これらの使用可能・不能なカメラワークの区間以外が FIX 区間となる。

表 3: 使用可能・不能区間推定の実験結果

	Blank	Camera work				Fix
		Follow	手ぶれ	急峻	安定	
正解検出数	6	4	22	7	21	78
未検出数	0	1	5	1	4(2)	8
湧き出し	0	0	18(14)	0	2	5
正解率 (%)	100	80	81	88	84(91)	91
適合率 (%)	100	100	55(85)	100	91	94

4. ショットサイズ判定

素材映像中 1 つのシーンを構成するショットは連続して撮影され、カメラマンは 1 つのシーン内で LS, MS, TS それぞれのショットを撮影する慣習がある。シーン内のショットは同じ被写体を撮影する可能性が高く、同じ被写体の LS, MS, TS が撮影される可能性が高いため、それぞれのショットは包含関係が成立している可能性も高い。ショットサイズは、FIX 区間内で変化しないため、FIX 区間の代表フレームを取り出し、ヒューリスティックに決定した 5 つの画像サイズを用いて相互の包含関係を求める。これをシーン内で総当りで判定し、ショットサイズを付与する [1]。図 4 は、その方法の概略を示している。

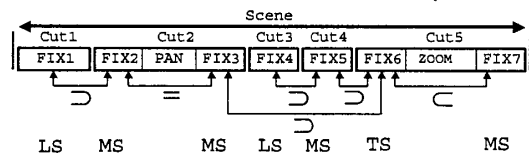


図 3: ショットサイズ付与法

表 4: ショットサイズ自動判定の結果

Scene 番号	Scene 内 FIX 部数	Shot Size 正解判定数	自動判定 正解率 (%)	Shot Size 未付与数
1	2	0	0.0	0
2	14	11	78.6	0
3	7	2	28.6	0
4	11	3	27.2	0
5	4	4	100	0
6	3	3	100	0
7	16	16	100	0
8	5	3	60.0	0
9	2	2	100	0
10	22	16	72.7	0
11	21	16	76.2	0
12	12	8	66.7	1
合計	119	84	70.6	1

この包含関係を求める処理にアクティブ探索 [4] を採用し、実験を行った。アクティブ探索は画像の解像度が低いと精度が落ちるため、640x480,30fps,MotionJPEG の素材映像 1 本を用いた。表 4 に結果を示す。

5. 今後の課題

今後は Phase2,3 の実現化と、接続規則としてショットサイズ以外の映像文法にも対応していく予定である。

参考文献

- [1] 熊野雅仁, 有木康雄, 上原 邦昭, 下條 真司, 春藤 憲司, 塚田 清志: “映像編集支援システムのためのショットサイズ自動付与”, 電子情報通信学会論文誌, Vol.J85-D-I, No.7, pp.592-602, 2002.
- [2] 長坂晃朗, 宮武孝文: “時間変化領域の画像相関に着目した実時間ビデオモザイク”, 電子情報通信学会論文誌 D-II, Vol.J82, No.10, pp.1572-1580(1999).
- [3] 熊野 雅仁, 有木 康雄: “映像編集支援システムのための使用可能ショット自動抽出”, 電子情報通信学会技術研究報告, パターン認識とメディア理解研究会, PRMU2002-31, pp.1-8, 2002.6.
- [4] 村瀬洋, V.V.Vinod: “局所色情報を用いた高速物体探索—アクティブ探索法—”, 電子情報通信学会論文誌 D-II, Vol.J81, No.9, pp.2035-2042, 1998.