社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

制約と評価関数に基づいた映像編集モデル

尾形　涼†　尾関　基行†　中村　裕一†　大田　友一†

† 筑波大学 機能工学系　〒 305–8573 茨城県つくば市天王台 1–1–1
E-mail: †{ogata,ozeki,yuichi,ohta}@image.esys.tsukuba.ac.jp

あらまし　本稿では，映像編集の新しい計算モデルを提案する．このモデルでは，映像を 0.5 か 1 秒単位に区切られた箱の並びであると考え，映像編集の問題を「連続的に並んでいる箱に各ショットを埋める」手続きだと定義する．すなわち，映像構成のための規則を個別の制約や評価関数の集まりとして実装し，制約を満たす組合わせの中から評価の良いものを選ぶ組み合わせ最適化問題として編集パターンを選び出す．この計算モデルにより，種々の編集パターンを生成したり，既存の映像編集テクニックを説明することを本研究の目的とする．我々は実際に，この効果をシステマティックに確かめることのできる環境を用意し，複数のカメラを使って撮影した映像を編集することによって，その有用性を確認した．

キーワード　メディア処理，映像編集，編集の計算モデル，制約充足，組合わせ最適化

Computational Video Editing Model based on Optimization with Constraints Satisfaction

Ryo OGATA†, Motoyuki OZEKI†, Yuichi NAKAMURA†, and Yuichi OHTA†

† IEMS, University of Tsukuba, Tennoudai 1–1–1, Tsukuba, Ibaraki, 305–8573, JAPAN
E-mail: †{ogata,ozeki,yuichi,ohta}@image.esys.tsukuba.ac.jp

Abstract　This paper presents a computational video editing model. In this model, a video is considered as a sequence of small boxes each of which has a length of 0.5 or 1 second, and the editing is defined as the problem of filling each box with an appropriate shot. Each editing rule is implemented independently of the others, as an evaluation function or a constraint for choosing shots and arranging shots. This formalism enables easy and systematic editing investigation by including or excluding editing rules. Based on this model, our objective is to support for the generation of a variety of editing patterns or a simulation of typical editing in movies and TV programs. We implemented this model in an actual video editing system, and we verified that the model works for multi-angle videos, and that our set of rules produce good results.

Key words　media processing, video editing, computational editing model, constraints satisfaction, combinational optimization

1. はじめに

映像製作の専門家でない人が，映像を撮影したり編集したりする機会が増えてきている．例えば，テレビ会議，遠隔講義などでも，その場でカメラを切り替えながら相手側に適切な映像を送ることが行われている．しかし，実際に我々が映画やテレビなどで目にするような質の良い映像を個人が製作するのは難しい．その理由としては，映像製作のための機材の問題だけでなく，映像を編集するための知識や経験が不足しているという問題が大きい．映像製作の現場には映像文法とも呼ばれるような種々の規則や手本となるパターンがあり，プロのカメラマンや熟練したディレクターなどは，それらの高度な技術や経験に基づいて編集を行っているからである．以上のように，映像を利用するためには目的に沿うような編集が不可欠であるにもかかわらず，編集は非専門家にとってかなり扱いにくい問題となっている．

この問題に対し，本研究では，映像編集をサポートする環境や知的システムを実現することを目的とし，そのために自動編集を行うための計算モデルを提案する．このモデルでは，映像を 0.5 か 1 秒程度の単位に区切られた箱だと考え，連続的に並んでいる箱に適切なショットを埋めるために，制約と評価関数を用いた組合わせ最適化問題として映像編集問題を扱う．この

モデルでは，映像構成のための規則を個別の制約や評価関数の集まりとして考えることができるため，各々の効果をシステマティックに確かめることが可能となる．本研究では，これにより，種々の編集パターンを生成したり，既存の映像編集テクニックを説明することを目標とする．

我々は実際に，このモデルの機能をシステマティックに確かめることのできる環境を構築し，複数のカメラを使って撮影した映像を編集することによって，その有用性を確認した．

本論文の構成は以下のようになっている．まず，2 章で映像編集とカメラ切り替えの問題について述べ，3 章で編集モデルについて説明する．4 章，5 章で，イベント，評価関数，制約等の重要な項目について説明し，6 章で実験例を示す．

2. 映像編集とカメラ切り替え

映像編集の目的は単純なものではなく，わかりやすい，見辛くない，楽しい等，種々の要因を考え，それらの要求が満たされるように映像の構成を考える必要がある．現在，その簡単な方法論はなく，種々のテクニックがその場に応じて用いられている．本研究ではそれを少しずつ計算可能なモデルとして明らかにすることを目的としている．

研究対象として，本稿ではカメラ切り替えによる編集を扱う．カメラ切り替えによる編集は，時間的に連続するショットのみを扱うため，編集の問題の一部分でしかないが，一般的な編集に共通する問題を多く含んでいるため，得られた知見が一般的な編集のために利用できる可能性が高い．そのため，本研究では，まずカメラ切り替えの問題を詳細に考え，得られた知見を一般的な編集の問題に利用することを目標としている．

これまでの研究におけるカメラ切り替えでは，シーン中で発生したイベントに連動してカメラを切り替えるイベント駆動型のアルゴリズムが使われてきた [3], [4]．しかし，このような逐次的編集では，編集目的，編集規則やアルゴリズム，結果の各々の間の関係が非常にわかりにくい．イベントの生起時刻の微妙な違いやその順番の変化が結果に大きな影響を及ぼすからである．また，相反する要求を一つのアルゴリズムに取り入れて，そのトレードオフを探ることが難しい．例えば，退屈な映像にならないように適度な変化を与えることと，映像としての見苦しさを解消するための工夫を両立させることは難しい．

それに対し，本研究では，映像編集を組合わせ最適化問題としてとらえることにより，種々の編集パターンをシステマティックに洗い出すための計算モデルを提案する．このモデルでは，映像を単位時間毎に区切られた箱の集まりとして考え，全ての箱に適切なショットを割り当てた後，ショット列としての良さを評価する．この問題を単純に扱うと，爆発的に組合わせの数が増えるが，制約充足型の計算を取り入れることにより，ある程度の自由度を持たせながら，最終的に可能性を絞り込むことを可能にする．ただし，強い制約がかけられない場合には，扱える時間が短くなってしまうため，常に計算可能な範囲となるように調整することは今後の課題となっている．

3. 最適化と制約充足に基づいた編集モデル

3.1 計算モデル

本研究の映像編集モデルでは，まず，映像を単位時間毎に区切られたショットの集まりとして扱う．現在は，この単位時間を 0.5 秒から 1 秒程度と設定しているが，それは以下の考察による．

- 古い映画では，1 秒程度の短時間のショット切り替えがほとんど用いられていないが，十分に映像として成立している．
- 短時間のショット切り替えが連続すると，視聴の負荷が増大するため，最近の映像でも，目的なしに短時間のショット切り替えを連続させることはない．また，はっきりした目的がある場合には，複数のショットをまとめ，特殊効果を狙った一つのショット群としてマクロ的に扱うことができる可能性がある．
- ショット切り替え (シーンチェンジ) の 0.5 秒程度の進み/遅れが致命的な問題となることは少ない．つまり，ショット切り替えを時間的に進めるか遅らせるかのどちらかによって適切な映像になる場合が多い．

次に，この計算モデルを以下のように，5 つの要素で定義する．

$$\text{Editing} = \{V, S, E, O, C\} \tag{1}$$

各要素は次のような意味を持つ．

ショット集合 (S):　ショット (s_i) の集合，$S = \{s_1, \ldots, s_n\}$ を表す．時刻によってその種類は変化しないとする．

ビデオシーケンス (V):　単位時間の長さを持つビデオセグメント $v(t)$，$(t = 0 \sim t_{max})$ の並びからなるビデオシーケンスを表わす．

イベント集合 (E):　対象シーン中で起きているイベント (e_i) の集合である．イベントとしては，見る者に伝えたい出来事，もしくはカメラ切り替えのきっかけになりそうな出来事が用いられる．イベントが生起したことを 1，生起していないことを 0 として表現し，生起の程度によってその間の値を使用することもある．例えば，時刻 t にイベント e_1 が 0.9 の程度 (又は確率) で起った場合は，$e_1(t) = 0.9$ と表わす．

評価関数 (O):　各ショット (またはショット列) の良さを計算するための評価関数 (o_i) の集合を表す．用いられる形態により，2 種類に分けられる．一つは各時刻における各ショットの良さを前後のショットに関係なく単独に評価する**前評価関数**で，これにより与えられたスコアを基に，制約充足系を用いて候補の探索が行われる．二つ目はショットの組み合わせや長さを評価する**後評価関数**で，ショット列として $v(t)$ が全て即値化された後に初めて計算可能になる．

制約 (C):　ショットの並び方やつなぎ方に関する制約を記述する．良いショット列を選ぶための評価という点では，評価関数との意味的な違いはない．そのため，できるだけ多くの項目を評価関数ではなく制約として実装することにより解候補を減らすことが好ましい．

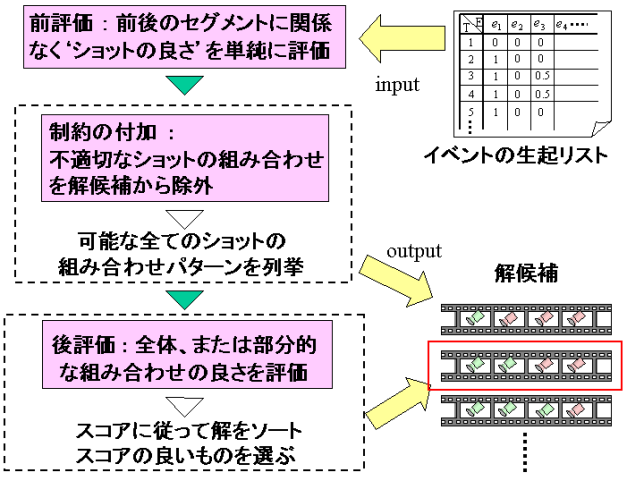


図 1　計算の流れ

これらの設定で，以下の評価 G を最大にするビデオシーケンス V を求めること，つまり，各ビデオセグメント $v(t)$ へのショット s_i の割り当てを求めることを編集の目的とする．

$$G = \sum_{t=0}^{t_{max}} \sum_{i=0}^{N_o} o_i(t) \qquad (2)$$

ただし N_0 は評価関数の数を表す．

3.2　計算の流れ

計算の流れを図 1 に示す．処理は大きく分けて，前評価，制約充足による解候補の探索，後評価の 3 段階となっている．既に述べてきたように，最初から全ての解候補を一つ一つ評価することができないため，前評価の結果を使いながら制約充足により解候補を減らし，その後，後評価を加えて，最もスコアの良かった一つまたは数個を採用する．

前評価の段階では，各時刻におけるイベントを基にして，各ショットの良さを単純に評価する．ただし，これは前後のビデオセグメントに割り当てられたショットに依存せず，生起したイベントによってのみ決まる評価のみを使用する．各時刻における実際のショットが即値化されていないため，ショット間の関係などを評価することができないからである．

制約充足による解候補の探索では，前評価のスコアを用いながら制約を加え，その後，解候補を全数探索する．実験では，100 万個程度の解候補まで問題なく扱うことができているが，映像の長さが長くなるにつれて指数関数的に解候補が増えるため，解候補の数 (推定数) に合わせて制約を適応的に変えるなどの枠組みが必要となる．これは今後の課題となっている．

前の段階で得られた各解候補各々に対して後評価関数を適用し，得られた評価の高いものを選ぶ．この後評価の段階では，上記の解候補の全てのビデオセグメントの値が即値化されているため，連続するショットの長さや異なるショット間の関係などの評価を行うことができる．

4.　イベントの種類と検出

イベントとしては，映像編集に関わる全ての出来事を考える必要がある．例えば，会話シーンでは，発話，動作，うなずき，表情の変化，物体の動き，その他が考えられる．現在は，その中で，表 1 に示したものをイベントの候補として考えている．現在のモデルでは，その種類や数に制限はないが，その役割 (映像編集との関係) が良くわからないものや，将来的にも自動認識に難があるものは省いている．

表 1　イベントの例

発話:　発話の有無や，発話の開始・終了は対話シーンを理解するための情報の中でも最も重要な要素の一つである．またその発話内容も，提示が必要な特定のものや場所を知るための手ががりになる．
ジェスチャー:　シーン中の人物の指示動作やその指示内容．または特別に大きな動作をした場合にはそこに特別な情報が含まれることが多い．
作業の発生:　机上作業を伴った対話シーンでは，作業が起っていることや，物体や作業の内容を参照していることが重要な情報となる．
表情の変化や頭部の動き:　頷きや表情，振り向き，視線の移動などには話し手やの聞き手の態度や興味が反映されるため，伝えるべき情報，または，その場所を教える重要な情報となる．
人どうしの体の接触:　握手や肩をたたいたりなどといった動作もその場の状況を伝えるための重要な情報となっている．

表 2　イベントの生起表 (上段がイベントの種類，下段がその生起を表す)

$e_1(t)$	人物 A の発話
$e_2(t)$	人物 B の発話
$e_3(t)$	机上作業を表現するキーワードの発話
$e_4(t)$	人物 A のうなずき
$e_5(t)$	人物 B のうなずき

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
e_1	0	0	0	1	1	1	1	1	1	0	0	0	1	1	1
e_2	1	1	1	1	1	0	1	0	1	1	1	1	0	1	0
e_3	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
e_4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
e_5	0	0	0	0	0	0	1	0	1	0	0	0	0	1	0

本稿の実験では，これらのイベントが起った時間を全て人手で記録し，編集システムに与えている．編集モデルの良さを評価するために，認識誤判等の要因を排除するためである．表 2 にその例を示す．実際には，発話の開始，キーワードの発話等は，我々の現在のシステム [1] [2] でも比較的精度良く認識できる．音声認識アプリケーションにより，発話があったこと自体は音声認識出力から何らかの出力があること，キーワードについては，単語 (読み) を認識出力から探すことが可能である．したがって，6 章の実験を自動化することもある程度可能であるが，これは今後の課題としたい．顔の向きやうなずき等も，現在の画像処理技術で比較的簡単に検出できることが期待される．また，微妙な表情，口調，複雑なジェスチャー等の検出を行うためには，まだ十分な技術が整っていないが，将来的には，認識系から出力される情報を利用する予定である．

5.　制約と評価関数

5.1　制約と評価関数の種類

制約と評価関数は次のような役割を持っている．

注目すべき情報を伝える:　視聴者の注意を正しい対象に向けることを目的とする．発話者や行為者が注目を要求している場所，編集者が視聴者に見せたり気付かせたい場所等がある．例

表 3　編集によって誤った情報を与える例

180 **度線の無視**:　180 度線 (imaginary line) を無視した映像提示をを行うと，視聴者がシーンの空間的構成を誤認知したり，動きの方向を誤って認識することにつながり，視聴者を混乱させる．
不適切なシーンチェンジ:　不適切な場所に視聴者の注意を向けたり，誤った情報や不要な情報を想起させる．
構図の似たショットを不用意につなぐこと:　仮現運動 (apparent motion) 等の視覚的な錯覚を引き起こしたり，意味のないシーンチェンジで視聴者を混乱させる．
不適切なアングル:　観察者 (視聴者) の動きを暗示したり，空間的な構成に対する誤認知を与える．

表 4　編集によって不適切な認知的負荷を与える例

ショット切り替えの不足:　刺激が不足するために，視聴者を飽きさせ，退屈な印象を与える．
短い時間内でのショット切り替えの多発:　視聴者に大きな認知的負荷を与える．また，シーンの詳細な理解が困難になる．
クロースショットの連続:　クロースアップ等の多用は視聴者を疲れさせる．
不適切なタイミングでのショット切り替え:　重要な点への注意を妨げ，心理的な不快感を視聴者に与える．

えば，対話シーンであれば，発話者の表情，聞き手の態度，その場の雰囲気等が必ず伝えるべき情報となる．ドラマなどでは，登場人物の感情の動きなど，直接観測しづらいものも重要となるが，本研究では，シナリオに指定してある場合，または，直接観測できる場合のどちらかを扱う．

誤った情報を想像させない/無駄な情報を伝達しない:　モンタージュ理論に代表されるように，ショットの構成が視聴者に様々な情報を伝えたり，想像させる．そのため，不用意に編集を行うと，間違った情報を伝えることになる．また，動きや方向の整合性を保たなければ，視聴者を混乱させることになる．その例を表 3 に示す．例えば，典型的な規則として 180 度線があげられるが(注1)，これを破ると，被写体の位置関係に対する誤解，認知的な負荷，イベント生起の誤解等を与えることが多い．

適度な認知/生理的刺激や負荷を与える:　見易い映像とするためには，視聴者の認知的負荷を適度な状態に保つことが必要となる．短いショットの連続や動きの激しいカメラワークは視聴者を疲れさせたり，不快感をもたらす．逆に，変化のない単調なショットが続くと，それを見続けるための認知負荷が高くなる．そのため，適度な刺激を与えることも重要な要素となる．不適切な認知/生理的負荷を与える幾つかの例を表 4 に挙げる．

5.2　制約と評価の記述と計算

上述したように，前評価関数，制約，後評価関数の 3 種類を映像編集の規則として用いる．

前評価関数は，時刻 t における各 s の好ましさを評価するだけで良いので，イベントの生起表 (表 2 参照) を基に，各時刻の各ショットにスコアを与えていく形とする．実際の例を表 5 の上段に，対応するショット群の例を図 2 に示す．o_{1A}〜o_7 はベクトルの形式で与えられているが，これは，各ショット (s_1〜s_6) に対するスコアを同時に表わすための便宜上の表記である．o_{1A} は e_1 (人物 A の発話) に比例したスコアを s_3 (人物 A のバストショット) に与える非常に単純なものである．また，イベントとショットの時刻が同一である必要はなく,「発話開始の 1 秒前の顔を見せたい」場合には，o_3 のように，時刻 $t+1$ のイベント (この例では，$e_1(t+1)$) を参照すればよい．

後評価関数の例を表 5 の下段に示す．後評価関数は o_7, o_8 のようにショットの組み合わせ等を評価するため，前評価関数のように簡単な式で表現できない場合が多い．

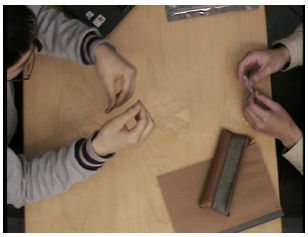
また，これらの評価関数がとる値として，便宜上，$-\infty$ を意味する大きな数も許す．本来，これは絶対的な “不採用” を表

(注1)：小津安二郎の映画では，この規則が破られている箇所が何度も見られる．しかし，前後のショットやショット中のイベント等がうまく構成されているため，不自然な印象を受けない．つまり，有用なルールである反面，絶対に守らなければならないわけではない．

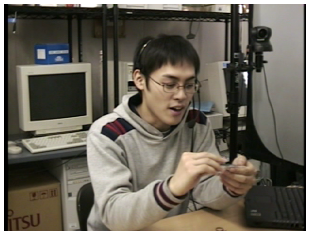
表 5　評価関数の例 (上段は前評価関数，下段は後評価関数)

$o_{1A}(t)$	$[0,0,15e_1(t),0,0,0]$
$o_{1B}(t)$	$[10e_1(t),0,15e_1(t),0,0,0]$
$o_{2A}(t)$	$[0,0,0,0,15e_2(t),0]$
$o_{2B}(t)$	$[10e_2(t),0,0,0,15e_2(t),0]$
$o_3(t)$	$[0,0,15e_1(t+1),0,0,0]$
$o_4(t)$	$[0,0,0,0,15e_2(t+1),0]$
$o_{5A}(t)$	$[0,30e_3(t),0,0,0,0]$
$o_{5B}(t)$	$[0,50e_3(t),0,0,0,0]$
$o_6(t)$	$[0,0,0,50(e_2(t)\wedge e_4(t)),0,0]$
$o_7(t)$	$[0,0,0,0,0,50(e_1(t)\wedge e_5(t))]$
o_8	s_3 と s_5, s_4 と s_6 が連続したら 10 点の減点
o_9	隣接する人物 A と B のバストショット (s_3 と s_4) の長さが同じである場合に 15 点の加点

s_1:シーンの全景



s_2:机上のショット



s_3:人物 A のバストショット

s_4:人物 B のバストショット

s_5:人物 A を人物 B の肩越しに撮ったショット

s_6:人物 B を人物 A の肩越しに撮ったショット

図 2　ショットの例 (実験で用いたもの)

表 6　制 約 の 例

c_1	同一種類のショットが最低でも 3 セグメント連続する
c_2	連続 3 セグメント以上スコアが 0 のショットが使われない
c_3	最初の 10 セグメント以内に s_1 が必ず 1 度は使われる

表 7　編集に使った制約と評価関数の組み合わせ

	制約と評価関数	解の候補数
edit1	o_{1A},o_{2A},o_{5A},c_1	1,002,156
edit2	o_{1A},o_{2A},o_{5B},c_1	1,002,156
edit3	o_{1A},o_{2A},o_{5B},c_1,c_2	596
edit4	o_{1A},o_{2A},o_{5B},o_6,o_7,c_1,c_2	1,752
edit5	o_{1A},o_{2A},o_{5B},o_6,o_7,o_8,c_1,c_2	1,752
edit6	o_{1A},o_{2A},o_{5B},c_1,c_3	459,816
edit7	o_3,o_4,o_{5B},c_1,c_2	610
edit8	o_{1B},o_{2B},o_{5B},c_1,c_3	459,816
edit9	o_{1B},o_{2B},o_{5B}	約 3×10^{15}

すため，評価関数よりも制約に組み込まれるべきものであるが，制約ライブラリの制限によって扱えないような条件が必要となる場合に用いる．

制約の例を表 6 に示す．c_1 は短いショットが連続することを禁止する制約であり，これによって，解候補の数も大幅に減る．また，ショットに前評価関数によってスコアが付加された後に適用されるので，このスコアも考慮に入れることができる．c_2 はスコアが低いショットが連続しないようにする制約である．システムで扱える範囲まで解候補の数を減らす必要があるため，これらの制約には効果的な (強い) ものを利用する必要がある．

5.3　パラメータの設定

本研究は，まだ計算モデルが実際に動作することを確認した段階であり，評価関数や制約に用いられているパラメータの値の妥当性を十分にチェックできていない．現在のパラメータ設定はイベントの重要度に応じて，-100〜+100 程度の間で 5 点刻みの点数を割り振っている．50 点以上のスコアは，優先して特定のショットを選ぶ必要がある場合に使っている．これらの値に唯一の解というものはないため，今後，状況に応じて試行錯誤しながら決めていく必要がある．また，これらのパラメータ値を実際の編集例から学習することも考えられるが，これについては現在検討中である．

6. 実　験　例

実験として，二人が机に向かい合って座り，会話をしながら机の上で作業をしているシーンを撮影し，本研究で提案したモデルを適用して編集を行った．ショット，イベント，評価関数，制約は，各々，図 2，表 2，表 5，表 6 にあげたものを利用した．

実験に用いた評価関数と制約の組み合わせのパターンを表 7 に，編集例を図 3 に示す．図 3 の edit1 と edit2 は机上を写したショットを評価するパラメータのみを変えて編集したものである．edit1 ではイベント e_3 が起きても，前後の e_1 や e_2 の影響により，結果として s_2 は選ばれなかった．しかし，edit2 では，イベント e_3 が生起したときに s_2 に付加されるスコアを edit1 より増加させた o_{5B} を使用したため，s_2 が選ばれている．

edit3 は，edit2 に対して，制約を一つ付け加えた結果である．その結果，選ばれた編集結果は edit2 と同じだったが，その解の候補数が 1,002,156 個から 596 個まで，大きく減った．つまり，低スコアのショットが連続した部分を持たないようにする制約 c_2 を追加したため，edit2 と同じ解を得ながら，不要な組み合わせを解候補から除外できたことになる．このように比較的強い制約を利用することによって解候補の数を減らすことが効果的である．

edit4，edit5 は，聞き手の様子を伝えることを重視した例である．edit4 は聞き手の様子を伝えるために，o_6, o_7 を用いることによって，聞き手がうなずいた時点での肩越しのショットのスコアを高くしている．これにより，6 秒から 9 秒までの間に s_4 が挿入されている．しかし，s_3 と s_5，s_4 と s_6 のような撮影アングルの似たショットをつなげるのは，特別な効果を狙った場合以外は避けるべきだと言われている(注2)．そのため，そのような特定のショットが連続した部分に対して評価値を下げる後評価関数 o_8 を使った編集例を edit5 に示す．これによって edit4 で現れていた s_4 と s_6 が連続する部分を含む解のスコアが下がり，edit5 のように s_4 と s_6 が連続部分を含まないものが最もスコアの高い解となった．

edit6 は場の雰囲気を見る者に伝えるため(注3)に，場の全景を写したショット s_1 をビデオの開始 10 秒以内に使う制約 c_3 を使用して編集したものである．このショット自体に良いスコアが与えられているわけではないが，制約によって，他のショットが大きなスコアを得ていない部分で強制的に挿入されている．

edit7 は話し始めを重視した例である．対話シーンの映像では，話者が話し始めた時に，その始まりに少し先行して様子を伝えることによって視聴者の理解を助けることがある．そこで，発話する人を写したショットを発話の少し前から使うための関数 o_3, o_4 を使った例が edit7 である．

edit8 では，o_{1A} と o_{2A} の代わりに，発話者のクロースショットにつけるスコアより少しだけ少ないスコアを，全景を写したショットにつける機能を付随した前評価関数 o_{1B}，o_{2B} を使っている．これによって二人が同時に発話したときは二人を同時に映したショットがスコアが各々のクロースショットのスコアを上回るようになった．編集結果からも分かる通り，s_1 が選ばれる回数が増えた．

edit9 は edit8 と同じ評価関数を利用し，制約を一つも使わずに編集を行った(注4)例でり，単純に各時刻で最も評価値の高いショットが選ばれている．その結果，一秒毎にせわしなくショットが切り替わる，非常に見にくい映像になった．またこの例から，edit8 では実際にショットの評価をしていない制約によってうまくショットが選ばれていることが逆にわかる．

以上の実験例で示したように，本稿で提案した編集モデルでは，種々の評価や制約を利用したり，また逆に無効にすることによって，種々の映像編集パターンを生成できる．また，上記の実験では実装していない映像編集規則でも，比較的単純なプ

(注2)：30 度程度 (またはそれ以内) のアングル差しかないショットが対象となる [6]．これらをつなげると，視聴者が仮現運動を感じるなどの弊害がある．

(注3)：このような目的で，最初の方に挿入されるショットを establishing shot (状況設定ショット) と呼ぶ．

(注4)：実際に本研究のモデルを利用したのではなく単純に各時刻で最もスコアの高いショットを選んだ

図 3　編　集　例

ログラムとして実現することができるものがまだ数多くあると考えられる．そのため，映像編集の問題をシステマティックに確かめる道具として，高い能力を持っているといえる．

しかし，一般的な映像編集に対してまだ十分に対応できない点もある．例えば，これまで述べてきたように，解候補の数は編集対象となる時間に対して指数関数的に増えるため，長い時間の編集ができない．一定時間の時間的窓を考え，時間方向にシフトしながら計算すればそのような制限は無くなるが、先に計算した部分が後に計算した部分の影響を受けなくなってしまう等，一般性を失う現象が起る．この点については，今後検討していく予定である．さらに，本稿では，編集規則となる評価関数や制約のパラメータについて，まだ十分な検討を行っていない．これらを実際の編集例や主観評価実験などを通して探ることも重要であり，これから重点的な研究を行う予定である．

7. ま と め

本稿では映像編集のためのモデルとして，評価と制約に基づいた最適化を行う手法を提案し，簡単な実験を行った．この編集モデルでは，種々の編集規則を評価や制約として表現し，それらを有効/無効にすることによって，種々の映像編集パターンを生成できる．それによって，映像編集の問題をシステマティックに確かめる計算モデルとなっていることを実験を通して示した．

今後はこのモデルを基にして,「良い」映像を生成するための編集規則の実装やその効果の確認を行っていく予定である．また，実用的な映像コンテンツの生成やアプリケーションの開発を目指している．

文　　献

[1] M.Ozeki, Y.Nakamura, Y.Ohta: Camerawork for Intelligent Video Production —-Capturing Desktop Manipulations, Proc. Int. Conf. on Multimedia and Expo, pp.41–44, CD–ROM TA1.5, 2001

[2] M. Ozeki, Y. Nakamura, and Y. Ohta. “Human behavior recognition for an intelligent video production system,” IEEE Proc. Pacific-Rim Conference on Multimedia, pp.1153–1160, 2002.

[3] 宮崎英明, 亀田能成, 美濃導彦, 複数のカメラを用いた複数ユーザに対する講義の実時間映像化法, 信学論, Vol.J82-D2, No.10, pp.1598-1605, 1999

[4] 大西正輝，村上昌史，福永邦雄, 状況理解と映像評価に基づく講義の知的自動撮影, 信学論, Vol.J85-D2 No.4 pp.594-603, 2002

[5] D.Bordwell K.Thompson 著, Film Art (Fifth Edition), The McGraw-Hll, 1997

[6] Daniel Arijon 著，岩本憲児 , 出口丈人 訳,「映画の文法」，紀伊国屋書店，1980.

[7] ISAC, Inc. ホームページ，http://www.isac.co.jp