

映画の文法に基づくビデオ画像の内容検索

石井 孝和† 吉高 淳夫‡ 平川 正人‡ 市川 忠男‡

†広島大学大学院工学研究科

‡広島大学工学部

〒739 東広島市鏡山 1-4-1

映画の映像において、ある特定の意味内容を表すシーンで一般的に用いられる撮影技法及び編集技法に注目し、それらの技法を用いた結果として動画像上に現れる特徴を抽出することによって、対応した意味内容を表すシーンを検索する手法を提案する。それら技法は、特定の意味内容を表現あるいは強調し、見る側に伝えることを目的として、映画の制作者（映画監督、撮影者、編集者）によって意図的に用いられるものである。本手法では、意味内容とそれを表すシーンが満たすべき特徴とを対応付けておき、画像処理によってデータから抽出された特徴とのマッチングをとることにより、制作者によって意図された意味内容に基づく検索を行うことができる。

Content-Based Retrieval of Film Scenes based on the Grammar of the Film

Takakazu Ishii†, Atsuo Yoshitaka‡, Masahito Hirakawa‡, and Tadao Ichikawa‡

†Graduate School of Hiroshima University

‡Faculty of Engineering, Hiroshima University

1-4-1 Kagamiyama, Higashi-Hiroshima 739, Japan

In the production of films, there are many techniques of filming and editing that have been used in common by many film producers (directors, cameramen, editors) for a long time. This paper proposes a method for retrieving video scenes by extracting the features caused by using the techniques. The techniques are used by the producers intentionally in order to express or emphasize the meaning of scenes. The system maintains the features corresponding to the meaning, and compares them with the features that are extracted directly from video data. Thus, this method enables users to retrieve scenes based on the meaning intended by the producers.

1 はじめに

コンピュータのデータ処理能力の向上、記憶装置の大容量化によって、動画像データを手軽に扱える環境が整ってきた。

それに伴い、VOD(Video on Demand)等の動画像データベースシステムが研究されており、そのようなシステムの特徴を生かすためにも、膨大なデータの中からユーザが見たいシーンを効率よく検索する機能が必要とされている。

従来の検索では、動画像データの各シーンに対し、それが表す意味内容をあらかじめテキストで付加しておき、それによって検索を行うものが多い[1]。しかし、テキストデータを付加する作業は大変な労力を必要とするため、必ずしも現実的な手法であるとはいえず、人手を介さずに動画像データ自体からその内容を抽出できることが望ましいといえる。

現状では、ドメインを限定しない一般的な動画像データからそれが表す意味内容を自動的に抽出することは困難である。しかしながら、適用分野を限定し、その分野の画像特有の性質を利用することによって、内容に基づいたデータへのアクセスが可能となる場合もある。

例えば、対象をTVのニュース番組に限定した場合、あるニュース番組中でアナウンサーを捉えたショットはパターンがほんの数種類であり、かつこれらのショットでは画像の変化が画面上の特定の領域でのみ起こるといった性質がある。また、アナウンサーが話題を紹介するショットの後には、その話題に関するライブ映像等の、画面全体にわたって画像が変化するショットが1つ以上続くという性質がある。このような点に注目し、[4]ではニュース番組を各ショットに分割した上で、アナウンサーのショットとそれに続くライブ映像のショットを合わせてユーザに提示するシステムを提案している。

また、映画やTVドラマ等において、同一シーン内では類似したショットが繰り返し出現するケースが多いことに注目し、[5]では類似ショットを自動的に検出することによりそれらの出現パターンからシーンの切れ目を検出する手法を提案している。ただし、データを各シーン(あるいは、サブシーン)に分割することを目的としており、シーン内で表されている意味内容については述べられていない。

一方、映画の画像には、ある特定の意味内容を表すシーンにおいて一般的に使われる撮影技法、編集

技法がある。例えば、映画の文法[6]によると、あるシーンでの緊迫感を高める演出をするためには、ショットの長さをだんだん短くするという編集技法がよく用いられる。その結果、映しだされている映像そのものの内容とあいまって、制作者の意図した意味内容をより効果的に表現し、見る側に伝えているのである。

本稿では、対象を映画の画像とし、画像そのものに含まれる内容を抽出しようとするのではなく、いわば行間に表現される意味内容に注目した検索手法を提案する。

2 アプローチ

2.1 映画の画像の性質

検索対象として映画の画像を用いる。それは、映画の画像には一般に以下のような性質があるからである。

映画の制作においては、一般に複数の位置に置かれたカメラから同時に、あるいは別々に撮影された複数のフィルムを後に編集することによって1つのシナリオを完成させる。その撮影や編集では、制作者(映画監督、撮影者、編集者等)によって特定の意味(意図)を強調する目的で使用される技法がある。そしてそれらの中には長年にわたって多くの制作者に使用されてきた、共通の規則とも呼べる技法がある[6]。

2.2 動画像データから抽出する特徴

本研究で用いる言葉を以下のように定義する。

カット: 連続した動画像データにおいて撮影時のカメラの start, stop 操作、及び編集作業によって作られた継ぎ目

ショット: カットによって分割されてきた個々の区間。すなわち、各ショット内では画像の動きは連続している

シーン: 1つ以上のショットの連続から成り、一般には場所あるいは背景が同一であるまとまり

映画の制作者によって使用される様々な技法のうち、本研究では、一定の特徴を持ったショットとそれらの出現パターンに注目する。各ショットから抽

出する特徴は、ショットの長さ、ショット内の画像変化、及び類似ショットの組とする。そして、オブジェクトの抽出及びその動きの追跡といった個々のショット内での画像の詳細な内容を理解しようとするのではなく、複数のショットの組み合わせによって表される意味内容に注目する。

ここで、ショット内の画像変化には、オブジェクト自身の運動によるものとカメラの移動及び操作（パン、チルト、ズーム等）によるものがある。本研究ではこれらを区別せずに画像変化として扱う。また、類似ショットとは、一般に同じ（位置に置かれた）カメラから撮影され、フレーミング及びカメラに対する操作もほぼ等しいショット同士と定義する。

2.3 各ショットの特徴とその出現パターンによって表される意味内容

前節で述べた画像の特徴の出現パターンとそれによって表されるシーンの意味内容を以下に示す。

1. ショットの長さに関するパターン

一般に、短いショットは時の早い流れを、長いショットは時のゆっくりした流れを暗示する。

- (a) 「長いショット、あるいはそれが連続する」
ゆったりとしたシーン、単調なシーン
- (b) 「短いショットが連続する」
激しい動きを表すシーン、驚嘆や衝撃などを表すシーン、緊迫したシーン
- (c) 「連続したショットの長さが徐々に短くなる」
緊迫感の高まるシーン
- (d) 「長いショットと短いショットが交互に現れる」
(主に2(c)と併用して) 激しいリズムを表すシーン
- (e) 「連続した短いショットの後の比較的長いショット」
緊張感を解放するシーン、余韻を残すシーン

2. ショット内の画像の変化に関するパターン

一般に、画像変化が大きい場合はあわただしさ、興奮、激しさを、小さいあるいは全くない場合は重苦しさ、静けさ、憂うつなどの感じを生む。

- (a) 「画像変化の大きいショット、あるいはそれが連続する」

あわただしいシーン、激しいシーン

- (b) 「画像変化の小さいショット、あるいはそれが連続する」

落ち着いたシーン、ゆっくりとした時の流れを表すシーン、重苦しいシーン

- (c) 「画像変化の大きいショットと小さいショットが交互に現れる」

躍動感を表すシーン、激しさを強調するシーン

3. 類似ショットとその出現パターン

- (a) 「1-2-1'-2' (1,2はそれぞれショットを表し、同じ番号は類似したショット同士を表す)」

追跡シーン、動きの少ない（カメラの動きを伴わない）人物の対話シーン

- (b) 「1-1'-2-2'」

追跡シーン

以上に基づいて、本研究で検索対象とするシーンの意味内容とそのシーンが満たすべき画像の特徴に関する制約条件をあらかじめ定義しておき、これを検索時に用いる。詳しくは第4章で述べる。

3 動画像データに対する処理

画像処理の流れとして、まず動画像データを各シーン単位でシステムに入力する。システムは各シーンからカットを検出し、それによって分割された各ショット毎の特徴を抽出、保持する。以下に、具体的な処理方法について述べる。

3.1 カットの検出

カットの検出には、[2]で提案されている以下の手法を用いる。これは、すべての連続した2フレーム間での輝度の変化量を調べるものである。

フレーム i 内の画素 n ($n \in \{1, 2, \dots, N\}$, N : フレーム内の全画素数)における輝度を $Y_n(i)$ として、連続した2フレーム間での画素 n における輝度差を $\Delta Y_n(i) = |Y_n(i+1) - Y_n(i)|$ と定義する。ただし、本研究では画素 n におけるRGBの値をそれ

ぞれ $R_n(i), G_n(i), B_n(i)$ としたときの輝度 $Y_n(i) = 0.30R_n(i) + 0.59G_n(i) + 0.11B_n(i)$ を用いた。

$$IDarea(i) = \frac{1}{N} \sum_{n=1}^N 1 \quad \text{if } \Delta Y_n(i) > \Delta Y_{th}$$

ここで、 ΔY_{th} は $\Delta Y_n(i)$ ($0 \leq \Delta Y_n(i) \leq 255$) に対するしきい値 (本研究では、16 に設定) を表す。

この $IDarea(i)$ ($i \in \{1, 2, \dots, I-1\}$, I : 全フレーム数) に対して、Pfilter [2] を適用した結果があるしきい値を越えるフレーム間をカットと判定する。Pfilter により、カット部分の $IDarea$ の値はあまり変化させず、ショット内の画像変化等により大きな値となる $IDarea$ の値のみを小さくできるため、カットの検出率を向上させることができる。

3.2 ショットの長さ

ショット m の長さを $Len(m)$ とし、ショット m を構成するフレーム数を用いて以下のように表す。

$$Len(m) = t - s + 1$$

ただし、 s, t はそれぞれショット m 内の最初と最後のフレーム番号を表す。

3.3 ショット内の画像変化

ショット m 内での画像の変化量を $AID(m)$ とし、以下のように定義する。

$$AID(m) = \frac{1}{t-s} \sum_{i=s}^{t-1} IDarea(i)$$

ただし、 s, t はそれぞれショット m 内の最初と最後のフレーム番号を表す。

3.4 類似ショットの検出

類似ショットを検出するために、まず、同一シーン内のすべての任意のショットの組み合わせについて類似度を算出し、その結果を用いて各ショットに対しラベル付けを行う。そして、同じラベルを付けられたショット同士を類似ショットであると判断する。

3.4.1 ショット間の類似度

ショットの長さに関係なく類似度を調べるために、各ショットからそれぞれ代表フレームを作成し、そ

の代表フレーム同士で比較を行う。そしてフレームの領域を x 軸方向、 y 軸方向ともに 8 等分し (領域分割数 $J = 64$)、その対応する領域ごとにカラーヒストグラムを比較する。代表フレームのカラーヒストグラムは、分割領域毎にそのショットを構成する各フレームのカラーヒストグラムの平均を求めることによって得られる。その際、色については RGB の値をそれぞれ R, G, B としたとき、以下の式によって得られる YIQ に変換して用いる。

$$Y = 0.30R + 0.59G + 0.11B$$

$$I = 0.60R - 0.28G - 0.32B$$

$$Q = 0.21R - 0.52G + 0.31B$$

この Y, I, Q のとる値の範囲を人間に対する重要度に応じて 16, 6, 2 等分し、色数 $K = 192$ とする。

以上に基づいて、ショット間の類似度の算出方法を以下に述べる。

$$MPR(m, l, j) = 1 - \frac{1}{2 \times M} \sum_{k=1}^K |h_{m,j}(k) - h_{l,j}(k)|$$

ここで、 $h_{m,j}(k)$ はショット m の代表フレーム上の領域 j ($j \in \{1, 2, \dots, J\}$, J : 領域分割数) において色が k ($k \in \{1, 2, \dots, K\}$, K : 色数) である画素の数、 M は 1 領域当たりの画素数を表す。すなわち、 $MPR(m, l, j)$ はショット m と l の代表フレームにおいて対応する領域 j 同士を比較して色が一致している画素の割合を表す。

そして全領域のうち評価値がしきい値を越える領域の割合を求め、これを類似度とする。すなわち、ショット m, l の類似度 $DRE(m, l)$ を以下のように定義する。

$$DRE(m, l) = \frac{1}{J} \sum_{j=1}^J 1, \quad \text{if } MPR(m, l, j) > MPR_{th}$$

ここで、 MPR_{th} は MPR に対するしきい値を表す。 $MPR_{th} = 0.5$ として用いた。

なお、領域分割方法、色数、及び MPR_{th} は実験により求めた。

3.4.2 ショットのラベリング

シーン内の類似ショットの組とその出現パターンを調べる目的は、第 2.3 節の 3 で述べたような類似ショットの繰り返しのパターンがある場合に、そ

れを正しく検出することである。すなわち、必ずしもシーン内の全ショットに対して正確なラベル付けが行われていなくても、特に近傍の類似したショットとの類似度を正確に検出し、繰り返しにおいて対になるショット（例：1-2-1'-2'の1と2）を異なるショットと認識することができればよい。

以下にラベル付けのアルゴリズムを示す。

```

for i=1 to num {
  if ( ラベル [i] = null )
    ラベル [i] ← 新規ラベル
  for j=i+1 to num {
    if ( 類似度 [i,j] ≥ th ) {
      if ( 類似度 [i,j] > 値 [i] ) {
        親ショット [i] ← j
        値 [i] ← 類似度 [i,j]
      }
    }
    if ( 値 [j] > 値 [i] ) {
      ラベル [i] ← ラベル [j]
      ラベル [x] ← ラベル [j]
      (x: i を先祖とするすべてのショット)
    }
  }
  else if ( 類似度 [i,j] ≥ 値 [j] or 値 [j]=null ){
    ラベル [j] ← ラベル [i]
    親ショット [j] ← i
    値 [j] ← 類似度 [i,j]
    ラベル [y] ← ラベル [i]
    (y: j を先祖とするすべてのショット)
  }
}
}
}

```

num: シーン内のショット数

th: 類似度に対するしきい値

ラベル [i]: ショット i のラベル

類似度 [i,j]: ショット i と j の類似度

親ショット [i]: ショット i にラベル [i] を付ける原因となったショット（ショット i から見て、最も類似度が高く、かつしきい値以上であるショット）

値 [i]: ショット i と親ショット [i] との類似度

このアルゴリズムにより、結果としてあるショット a に付けられるラベル [a] は、シーン内の他のすべてのショットの中でショット a から見て類似度が最も高く、かつその類似度がしきい値以上であるショット

（親ショット [a]）と同一のラベルである、ということが保証される。

そのことにより、実際に類似したショット b と c の類似度 [b,c] が仮にしきい値未満であっても、共通の親ショットを持つ場合（親ショット [b]=親ショット [c]=d）においては、ショット b, c にはそれぞれショット d と同じラベルが付けられることから、正しくショット b と c の類似を検出できる。

また、実際には類似していないショット e と f の類似度 [e,f] が仮にしきい値以上であっても、ショット e の親ショットがショット f とは異なるラベルを持っている場合（ラベル [親ショット [e]] ≠ ラベル [f]）においては、その影響を受けずに、正しくショット e と f は類似していないと判断される。

4 システムの概要

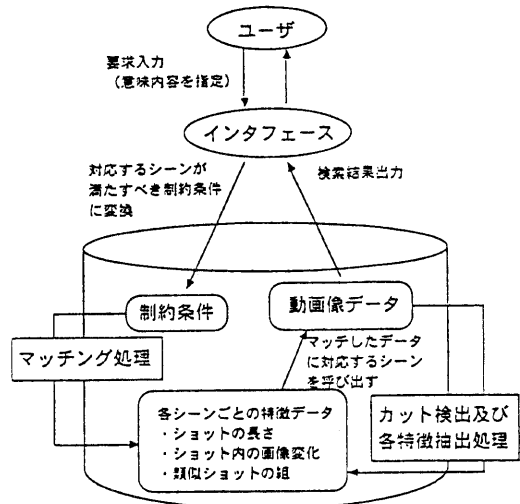


図1: システムの流れ

システムの流れを図1に示す。まず、システムにはあらかじめ検索対象とする意味内容に対応したシーンが満たすべき制約条件を定義しておく。また、動画像データに対して行う画像処理は、データをシステムに入力する際に行い、抽出された各特徴データは付加データとして保持しておく。以下、制約条件の定義及び検索の流れについて述べる。

4.1 制約条件の記述形式

以下に制約条件の記述形式を示す。

$s(l[\text{operator}]par_l, m[\text{operator}]par_m, label,$
 $lr[\text{operator}]par_lr, mr[\text{operator}]par_mr)$

$s()$: (単一の) ショットが満たすべき制約

$ns()$: $s()$ を満たすショットが n 個連続することを表す制約

l : ショットの長さ

m : ショット内の画像変化

$label$: ショットのラベル

lr : 直前のショットの長さを 1 としたときの、そのショットの長さ

mr : 直前のショット内の画像変化を 1 としたときの、そのショット内の画像変化

$[\text{operator}]$: $>$, $<$, \geq , \leq , $=$ のいずれか

$par_l, par_m, par_lr, par_mr$: 各パラメタ

ここで、 $label$ については整数値を用いて記述するが、他のショットと同じであるか異なるかのみを識別するために用いられる。

以下に具体的な記述例を示す。

例 1 「連続する 5 つのショットの長さがだんだん短くなる」という制約条件を定義する場合

$4s(\phi, \phi, \phi, lr < 1.0, \phi)$

例 2 「連続する 4 つのショットにおいて画像変化が大きい ($m > 30.0$)、小さい ($m < 5.0$) が交互に現れる」という制約条件を定義する場合

$s(\phi, m > 30.0, \phi, \phi, \phi) \wedge s(\phi, m < 5.0, \phi, \phi, \phi) \wedge$
 $s(\phi, m > 30.0, \phi, \phi, \phi) \wedge s(\phi, m < 5.0, \phi, \phi, \phi)$

4.2 制約条件の定義

検索時にユーザが指定する意味内容に対し、あらかじめ定義しておく制約条件について述べる。

第 2.3 節で述べた各特徴は大まかなものであるため、個々の条件に該当するシーンは数多く存在することが考えられる。そのため、複数の条件を同時に満たす制約条件を定義、あるいは各パラメタの設定値により制限することで、該当するシーンをより忠実に意味内容を反映したシーンに限定することができる。しかし、制約条件を厳しくすることで、本来検索されるべきシーンも同時に除外されてしまうといったことも起こり得る。

そこで、これらを考慮した上で、各意味内容に対しその厳しさのレベルが異なる制約条件 (レベルを

数値で表し、数値が大きいほど厳しい) を複数用意し、検索時の要求に応じてそれらを利用することで効率よく所望のシーンが得られるようにする。

4.3 検索の流れ

Step1: システムが提示した意味内容の一覧から、ユーザは自分の検索したいシーンの意味内容を指定することによって要求を入力する。

Step2: システムは指定された意味内容からそれに対応した画像の特徴が満たすべき制約条件を参照し、画像データの付加データに対してマッチング処理を行う。そして、その制約条件を満たす画像データを検索結果として提示する。(ここでの制約条件のレベルは、デフォルト値として設定されたものとする。)

Step3: 意図したシーンが得られれば終了。

意図したシーンを検索することができない場合、そのシーンが検索されるようにより小さいレベル (緩い制約条件) を指定する。

要求と異なるシーンを多く含み、提示されたシーン数が多過ぎる場合は、より大きいレベル (厳しい制約条件) を指定し、より忠実に意図を反映したシーンに限定する。

5 実験

「人物の対話シーン」、「追跡シーン」、「緊迫したシーン」の意味内容について、それぞれ制約条件を定義し、実際のシーンを検索する実験を行った。

5.1 実験に用いたデータ

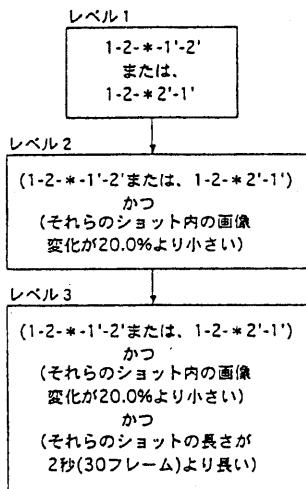
以下に示す 1 ~ 5 の映画から「人物の対話シーン」15 個、「追跡シーン」4 個、そして 1 ~ 9 の映画から「緊迫したシーン」23 個を選んだ。なお、データはすべて、フレームサイズ 160 × 120 ピクセル、色数 1677 万色中の 256 色、フレームレート 15 フレーム/秒である。また、フレーム数は 544 ~ 2284、ショット数は 6 ~ 79 である。

1. 「風と共に去りぬ (Gone with the Wind)」

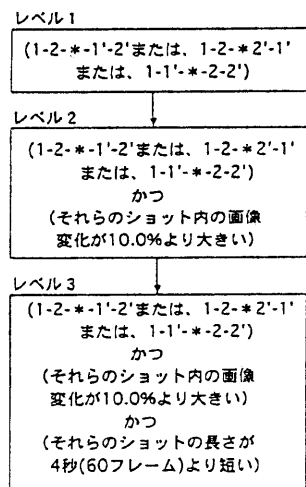
ヴィクター・フレミング監督、1939 年、アメリカ、文芸映画

2. 「北北西に進路を取れ (North by Northwest)」
アルフレッド・ヒッチコック監督、1959年、アメリカ、サスペンス映画
3. 「鳥 (The Birds)」
アルフレッド・ヒッチコック監督、1963年、アメリカ、サスペンス映画
4. 「俺たちに明日はない (Bonnie and Clyde)」
アーサー・ペン監督、1967年、アメリカ、アメリカン・ニュー・シネマ

5. 「トラ!トラ!トラ! (Tora!Tora!Tora!)」
リチャード・フライシャー、舛田利雄、深作欣二監督、1970年、アメリカ、戦争映画
6. 「JFK」
オリバー・ストーン監督、1991年、アメリカ、法廷劇
7. 「スター・ウォーズ (Star Wars)」
ジョージ・ルーカス EP、1977年、アメリカ、SF アクション映画
8. 「レイダース-失われたアーク- (Raiders of the Lost Ark)」
スティーブン・スピルバーグ監督、1984年、アメリカ、アクション映画
9. 「クリフハンガー (Cliffhanger)」
レニー・ハーリン監督、1993年、アメリカ、アクション映画



(*は、制約のない0個以上のショットを表す)
図2: 「人物の対話シーン」の制約条件



(*は、制約のない0個以上のショットを表す)
図3: 「追跡シーン」の制約条件

5.2 実験内容

実験1 「人物の対話シーン」「追跡シーン」の計19個のデータに対して、それぞれ図2, 3のように定義した制約条件を用いて検索を行った。

実験2 「緊迫したシーン」の23個のデータに対して、図4のように定義した制約条件を用いて検索を行った。

5.3 実験結果

まず実験1について述べる。本実験で用いたショット間の類似度 DRE に対するしきい値を次のように選んだ。全データの合計で、類似ショットの数 $C = 276$ のうち、類似度がしきい値未満であった数を A 、逆に類似ショットでないにもかかわらず、類似度がしきい値以上となった数を B としたとき、類似ショットの認識率 $DPF = 1 - \frac{A+B}{C}$ を最大にするようにしきい値を選んだ。その結果、しきい値が74%のとき、 $A = 53$, $B = 35$ であり、 $DPF = 74.5\%$ で最大となった。(なお、 DPF は、 A, B が結果に与える悪影響度の比を2:1として算出した。)

それに基づき、各ショットへのラベル付けを行い、その結果を用いて検索を行った。結果を表1(a)に示

す。「人物の対話シーン」「追跡シーン」について、レベル1(制約条件)でも検索できなかったシーンは、共に1つずつであり、それ以外のシーンについては検索することができた。また、「人物の対話シーン」ではレベル3、「追跡シーン」ではレベル2または3を用いることで、それぞれ要求に合ったシーンのみ限定することができた。

次に、実験2の結果を表1(b)に示す。「緊迫したシーン」について、レベル1を用いることで、すべてのシーンを検索することができた。ただし、本実験では要求を満たすシーンのみに対して、それらが制約条件を満たしているかどうかを調べたが、制約条件が緩いために、要求とは異なるシーンも検索結果に多く含まれることが考えられる。そのような場合、必要に応じてレベル2及び3を用いることで、要求に合わないシーンを除外することができる。

6 まとめと今後の課題

映画の撮影技法、編集技法によって動画データ上に現れる特徴に注目し、それを自動的に検出することにより、制作者が意図した意味内容に基づく検索を行う手法を提案した。

今後の課題として、今回実験したシーン以外の意味内容を表すシーンについても実験を行い、各意味内容に対する制約条件の定義、及びその検証を行う必要がある。そのために、実際の映画のシーンをより多く分析することによって、各制約条件における各パラメタについての検討を現在行っている。

また、現在はシステムに対する動画データの入力はシーン単位で行っているが、システムの自動化のために、今後はシーンの切れ目を検出する方法を考察していく予定である。

なお、システムは現在 MS-Windows95 上で Visual C++ を用いて構築中である。

参考文献

- [1] R.Weiss, A.Duda, and D.K.Gifford, "Composition and Search with a Video Algebra", IEEE MultiMedia, Vol.2, No.1, pp.12-25, 1995.
- [2] K.Otsuji and Y.Tonomura, "Projection Detecting Filter for Video Cut Detection", Proc., ACM Multimedia'93, pp.251-257, 1993.

- [3] A.Nagasaka and Y.Tanaka, "Automatic Video Indexing and Full-Video Search for Object Appearances", Visual Database Systems, Vol.11, pp.119-133, 1992.
- [4] H.J.Zhang, Y.Gong, S.W.Smoliar, and S.Y. Tan, "Automatic Parsing of News Video", Proc., IEEE Multimedia Computing and Systems, pp.45-54, 1994.
- [5] M.Yeung, B.L.Yeo, and B.Liu, "Extracting Story Units from Long Programs for Video Browsing and Navigation", Proc., IEEE Multimedia Computing and Systems, pp.296-305, 1996.
- [6] ダニエル・アリホン著, 岩本憲児, 出口丈人訳, "映画の文法", 紀伊國屋書店, 1980.
- [7] 浦岡敬一, "映画編集とは何か", 平凡社, 1994.

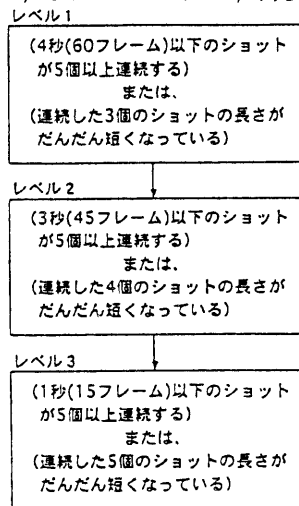


図4: 「緊迫したシーン」の制約条件

意味内容 制約条件	(a)実験1		(b)実験2
	人物の対話シーン	追跡シーン	緊迫したシーン
レベル1でも検索できない	1	1	0
レベル1での検索シーン数	14 [2]	3 [14]	23
レベル2での検索シーン数	13 [1]	3 [0]	20
レベル3での検索シーン数	10 [0]	1 [0]	14

(a)の[]内の数値は、要求とは異なるシーンが検索された数

表1: 実験結果