

動画像からのキーフレーム抽出に関する検討

大串 亮平[†] 竹内 一樹[†] 朱 青^{††} 小館 亮之^{††} 富永 英義^{††}[†] 早稲田大学 理工学部

〒169-8555 東京都新宿区大久保 3-4-1 55 号館 -N0602

^{††} 早稲田大学 国際情報通信研究センター

〒169-0051 東京都新宿区西早稲田 1-3-10 29-7 号館

ryohei@tom.comm.waseda.ac.jp

動画像の効率的ブラウジングのための支援として、本研究では撮影者が有意と考えて撮影したと思われるフレームを映像から自動で抽出するシステムについての検討を行う。従来の研究では対象映像を編集済み映像としている研究が多いが、本研究では素材映像を対象映像とする。素材映像の特徴としてはショットが長いことが挙げられ、ショットの中の時間的変化を知ることが重要となる。そこで、ショットを複数の区間に分割し、その区間に対してクラスタリングを行う。そして、要素数の多いクラスタから優先的にそのクラスタの重心に最も近いフレームをキーフレームとして抽出する。キーフレームを抽出することで、動画へのランダムアクセス、動画のインデキシングにかかる処理コストの低減も可能となる。

キーワード

キーフレーム、映像要約、素材映像、動画像編集、ヒストグラム、映像検索、クラスタリング

Extraction of Keyframes from Video Sequence

Ryohei OGUSHI[†] Kazuki TAKEUCHI[†]Qing ZHU^{††} Akihisa KODATE^{††} Hideyoshi TOMINAGA^{††}[†] Dept. of Science and Engineering, Waseda University

3-4-1 Ohkubo, Bldg.55-N0602, Shinjuku-ku, 169-8555, Tokyo JAPAN

^{††} Global Information and Telecommunication Institute, Waseda University

1-3-10 Nishi-Waseda, Bldg.29-7, Shinjuku-ku, 169-0051, Tokyo JAPAN

ryohei@tom.comm.waseda.ac.jp

The spread of digital video cameras are enabling the general users to get the video sequence easily. However, efficient access to desired video scenes and quick browsing over the video technology are not sufficient. Therefore this paper proposes keyframes extraction from shot to settle these problems. As the first step, we divide the shot into sections. In the second place, we select keyframes from sections using a clustering based approach. Keyframes also greatly reduce the amount of data required in video indexing and provides an organizational framework for dealing with video content.

Keywords

Keyframe, Video Summarization, Non-edited Video, Video Editing, Video Indexing, Histogram

1. はじめに

近年、デジタルビデオカメラの普及、映像の圧縮技術の高度化、記憶媒体の大容量化、通信速度の向上などにより、一般ユーザは大量の映像にアクセスすることが可能になってきた。一方で、大量の映像にアクセスすることが可能となっても、その中から必要な情報を取り出すためには、動画の中身を見る必要があり、多大な労力と時間を要するのが現状である。そうした背景から

- (a) ユーザの欲しい情報が含まれている部分を検索し、提示する技術
- (b) ユーザが短時間で映像の内容を正確に理解できる形で映像を表現する技術

への必要性が高まりつつあり、映像検索、映像のダイジェスト表現に関する研究が盛んに行われている。しかし、これらの研究で問題となることは、動画像の表す意味内容には様々な認識が可能であるということである。そのため、映像をテキスト情報で表すことは大変困難であり、また、人間によって映像のインデキシングを行うためには映像ソースにおける知識はもちろん、多大な時間、労力、手間を要してしまう。したがって、映像を解析することによって、自動で動画の意味内容を抽出したり、インデキシングを行うことが望まれており、現在様々な研究がなされている⁽¹⁾。本稿では、特に家庭用デジタルビデオカメラによって撮影される映像(以下、素材映像)を対象に自動でキーフレームを抽出することで、従来の早送りによるブラウジングよりも効率的な映像の内容情報を取得するシステムの構築を目指す。ここで、一般に素材映像の特徴としては、編集済み映像に比べ、カット点¹の頻度が少ないことで、編集済み映像である実際の映画の場合、15分に300のカット点が存在するという報告もある⁽²⁾。また、映像をショット²単位でセグメンテーションするカット点検出に関してはこれまで様々な研究がなされており、かなりの精度でカット点の検出が可能となっている⁽³⁾⁽⁴⁾。そこで、本稿では映像がショット単位にセグメンテーションされているものとしてショットからキーフレームを抽出することを考える。

2. 従来の関連研究

キーフレームとは本来ならば、ショット中の意味のあるフレームであることが理想であるが、どのフレームが有意であるかは個人の主観に依存し、現在の技術ではそのようなフレームを自動で抽出することは大変困難である。そのため、画像特徴量などから独自にキーフレームを定義している研究がほとんどである。文献⁽⁵⁾⁽⁶⁾では、ショットの切り替わるフレームを検出し、そのフレームをユーザに提示することで、ユーザが求める画像にランダムアクセスできる手法を提案している。しかし、素材映像ではショットの長さが編集済み映像に比べ長いことを考慮すると、ショットの切り替わるフレームだけではショット内での時間的変化を把握することができない。動作認識に関する重要な情報を多く含んだフレームをキーフレームとしている手法もある⁽⁷⁾。しかし、カメラが固定であることを前提としており、素材映像にカメラワークが多く含まれている場合もあることを考えるとロバスト性があまりない。文献⁽⁸⁾⁽⁹⁾では映像の構造化がある程度モデル化でき、カメラワークもある程度決まっているスポーツ映像、ニュース映像といった映像対象を限定させ、意味内容ごとに特徴的なフレームを設定し、そのフレームをキーフレームとしている。これらの手法では、対象画像毎に画像に現れる意味内容を考慮し、その意味内容に基づいた特徴抽出を行っているために、対象画像固有の意味内容が表現できる反面、他の映像への応用が困難である。動画によらずにキーフ

レームを抽出するものとして、HSV空間でHSのヒストグラムを特徴量として、教師なしクラスタリングを行うことで、類似しているフレームを集め、そのクラスターの重心に最も近いフレームをキーフレームとする研究もある⁽¹⁰⁾。この手法では時間について考慮していないのでビジュアル的に類似した余分なフレームを抽出する可能性を低減できるが、ヒストグラムは類似しているが、まったく別のものが同じクラスターに含まれてしまう可能性もある。文献⁽¹¹⁾ではYUV空間でYUVのヒストグラムを特徴量として、フレーム間差分では検出できないような、ゆるやかなショット内での場面変化も検出する手法を提案している。しかし、変化があったと判断したフレームをキーフレームとしており、その変化があったフレームが必ずしもキーフレームであるとは限らない。

また、動画像検索において、検索処理に必要なフレームをあらかじめキーフレームに限定しておけば、動画像のインデキシングなどの処理コストの低減、効率の良い検索が可能となる⁽¹²⁾。

3. 提案手法

3.1 有意と考える物の撮影方法

撮影者が有意と考える物を撮影するときの意図を反映していると思われるものとして以下のものが挙げられる。

- (a) カメラワーク
ズームをして撮影しているところなどは重要である。
- (b) 撮影時間
撮影者が重要と考えるものは、そうでないものよりも撮影時間が長くなる。
- (c) フレームの中央
撮影者は重要なもの、フォーカスがフレームの中央になるように撮影する。

本稿では(b),(c)を前提として、動画像中のショットからキーフレームを抽出することを考える。

3.2 類似度を定義する画像特徴量

画像内容検索では、類似度を求めるために用いられる特徴として、色情報、テクスチャ、形状、オブジェクトの位置などの特徴量が挙げられる⁽¹³⁾。一方、動画像は時間軸に沿って順次フレームが再生されていくデータであると言えるので、同一ショット内では連続するいくつかのフレーム間では、画像の構図、撮影されている物の形状はほぼ同一であると考えられ、特徴量として色情報だけを考慮すれば十分であると思われる。なお、色情報を表すカラーモデル⁽¹⁴⁾はいくつかあるが、ここでは、MPEG-2で圧縮された映像に適用することも考慮し、YCbCrカラーモデルを用いる。以上を踏まえて、本稿では、画像の類似性を表す特徴量として輝度ヒストグラム H 、YCbCr空間でのユークリッド距離 D 、YCbCr空間でのフレームの輝度、色差平均($Y_{avr}, Cb_{avr}, Cr_{avr}$)を用いる。

(a) ヒストグラム H

ヒストグラムは画像内での色の空間分布の情報(画像の構図情報、画像中のオブジェクトの形)などが失われてしまうので、ヒストグラムが類似しているからと言って、画像が類似しているとは限らない。しかし、動画では連続しているいくつかのフレームは共通の部分が撮影されているので画像の構図情報、画像中のオブジェクトはある程度類似していると仮定できる。その仮定を前提とするとヒストグラムを用いた類似度の信頼性は高いと思われる。また、連続するフレームを比較する特徴量としてフレーム間差分がよく用いられるが、フレーム間差分はカメラ操作によって発生する画面全体に渡る動きに対し

¹1つのショットから次のショットに切り替わる変わり目

²一台のカメラの電源のオン/オフの間に連続的に撮影された映像区間

て人間の視覚的類似度をよく反映できないことに対して、ヒストグラムでは人間の視覚的類似度をよく反映することができると考えられる。その点を踏まえ、本稿では、輝度情報の類似度を表すものとしてヒストグラムを用いる。ここで、カメラ操作による画像全体に渡る変化を表した図を図1に示す。



- フレーム間差分値を類似度を求める特徴量とすると、この2つの画像の類似度はあまり高くない
- ヒストグラムを類似度を求める特徴量とすると、この2つの画像の類似度は比較的高い。

人間の視覚的類似度では、このフレームは非常に似ていると判定できる

→ ヒストグラムは人間の視覚的類似度を比較的反映することができると思われる。

図 1: カメラ操作による画像の画面全体に渡る変化

なお、フレーム i とフレーム j の類似度 m はそれらのヒストグラム $H_i(s), H_j(s)[s; \text{輝度値}]$ が重なった部分の画素数とすると、以下の式で表すことができる。

$$m = \sum_{s=0}^{255} \min(H_i(s), H_j(s)) \quad (1)$$

- (b) $YCbCr$ 空間でのユークリッド距離 D
- フレーム i の $YCbCr$ 空間の値を (Y, Cb, Cr) 、フレーム j の $YCbCr$ 空間の値を (Y', Cb', Cr') とすると、フレーム i とフレーム j のユークリッド距離 D は以下のように表すことができる。

$$D = \sqrt{(Y - Y')^2 + (Cb - Cb')^2 + (Cr - Cr')^2} \quad (2)$$

- (c) $YCbCr$ 空間でのフレームの中心の輝度値の平均 Y_{ave} 、色差平均 Cb_{ave}, Cr_{ave}
- 3.1(c) の前提より、フレームの中心に重要なもの、フォーカスがあると仮定し、フレームの中心の特徴量を用いる。また、フレームの中央に着目することで、図2のような全体では類似したフレームであっても、そのフォーカスによって区別することが可能となる。ここで、フレームの中心とはフレームを 5×5 分割したものの中央 3×3 の部分とし、輝度、色差情報として $YCbCr$ 空間での Y 成分、 Cb 成分、 Cr 成分の輝度、色差平均を用いた。 Y 成分、 Cb 成分、 Cr 成分の輝度、色差平均はフレーム画像中の座標 (x, y) の画素の Y 成分の輝度を $Y(x, y)$ 、 Cb 成分の色差を $Cb(x, y)$ 、 Cr 成分の色差を $Cr(x, y)$ で表すと次式で表される。

$$Y_{avr} = \frac{1}{\frac{3w}{5} \times \frac{3h}{5}} \sum_{x=\frac{w}{5}}^{\frac{4w}{5}} \sum_{y=\frac{h}{5}}^{\frac{4h}{5}} Y(x, y) \quad (3)$$

$$Cb_{avr} = \frac{1}{\frac{3w}{5} \times \frac{3h}{5}} \sum_{x=\frac{w}{5}}^{\frac{4w}{5}} \sum_{y=\frac{h}{5}}^{\frac{4h}{5}} Cb(x, y) \quad (4)$$

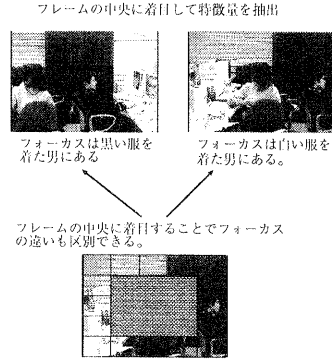


図 2: フォーカスの違い

$$Cr_{avr} = \frac{1}{\frac{3w}{5} \times \frac{3h}{5}} \sum_{x=\frac{w}{5}}^{\frac{4w}{5}} \sum_{y=\frac{h}{5}}^{\frac{4h}{5}} Cr(x, y) \quad (5)$$

ここで、 w は画像の幅、 h は画像の高さの画素数である。

3.3 ショットの区間への分割処理

3.3.1 目的

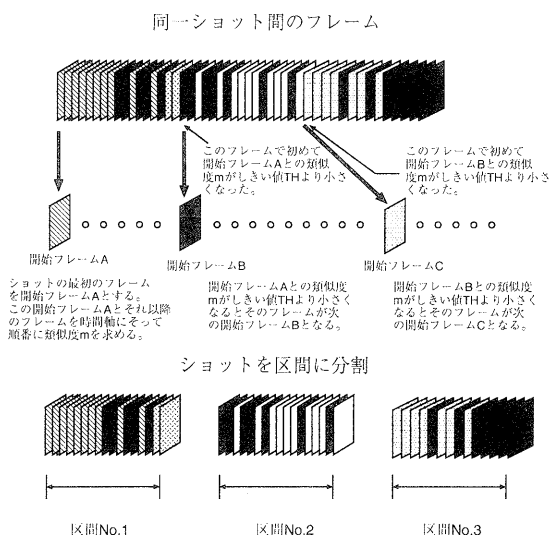
3.4 でクラスタリングを行うときに時間軸を考慮しないと、フレームの中心の特徴量は類似しているが、実際はまったく異なるものが同じクラスに属してしまう可能性が高くなる。例えば、ある人の顔を撮影してから、パンを行って別の人の顔を撮影したとする。このとき、この二人の人物が写っているフレームの中心の輝度、色差は人間の肌の色が似ていることを考慮すると類似していると判定され、この二人の人物を撮影したフレーム群は同じクラスに属する可能性が高くなる。なお、カメラ操作がある映像、ショットのフレーム数が多い映像では、以上のような問題が起こりやすくなる。このように、実際はまったく異なる人物が同じクラスに属してしまうことを防ぐために、最初の人を撮影している時の区間、その次の人を撮影している区間といったような区間に分割し、その区間をクラスタリングすることで以上の問題を低減できる。つまり、あらかじめショットを区間に分割し、その分割した区間でクラスタリングを行えば、輝度平均、色差平均は類似しているがまったく異なるフレームが同じクラスに属してしまう可能性を低減できる。また、動画画像をブラウジングするときに、その内容を把握するときにも時間という流れは重要であると思われる。

3.3.2 ショットの区間への分割処理手順

ショットを以下の手順によって区間に分割し、その概略図を図3に示す。

- ショット中の最初のフレームを区間1の開始フレームとする。
- 区間の開始フレームと、それ以降のフレームとの間の類似度 m を時間軸にそって計算する。
- (b) の類似度がある閾値 Th よりも小さくなれば、そのときのフレームを次の区間の開始フレームとする。
- 上記 (b)-(c) を繰り返すことによって、ショットを各区間に分割する。

本来であればフレーム全体の変化に着目しているのに、(c)では類似度を表す特徴量としてフレーム全体のヒストグラムだけを用いるのが理想であるが、背景が大きな割合を占める映像においては、全体のヒストグラムが背景の影響を大きく受けてしまい、本来、区間に分割すべきところを分割できない。一方、フレームの中心のヒストグラムでは映像中のオブジェクトが背景に占める割合は全体のヒストグラムの場合に比べ大きくなる。よって、フレームの中心のヒストグラムも合わせて考慮することで、より背景の影響を小さくすることが可能となる。また、フラッシュなどの影響で一瞬フレームの輝度値が大きく変化をすると、本来同一区間であるべき区間にもかわらず、区間の切れ目として誤検出してしまう場合も考えられる。その場合は、開始フレームとの類似度が2フレーム連続であるしきい値 Th よりも小さくなったとき、区間に分割することにし、その時の開始フレームを2フレームのうち、先に比較を行ったフレームとすればよい。



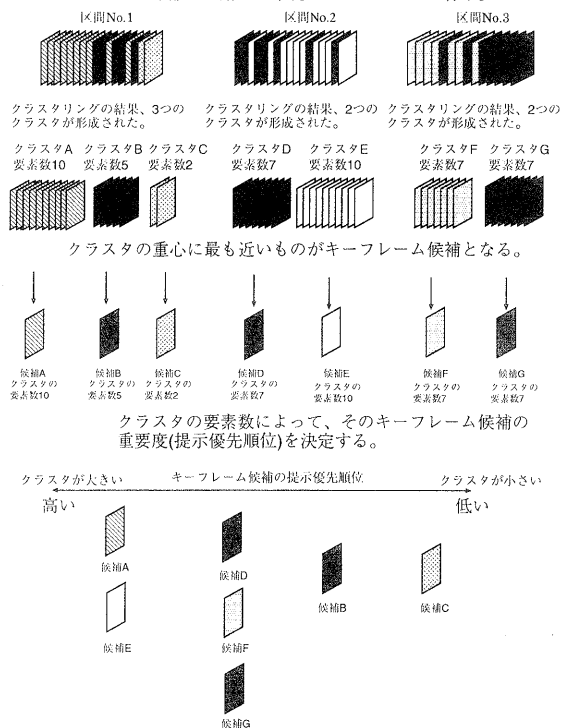
3.4 分割された区間のクラスタリング

次に、分割された各区間から以下の手順でキーフレームを抽出し、その概略図を図4に示す。

- フレームの中心の $YCbCr$ 空間での輝度、色差情報を用いて単純クラスタリングを行なう。
- 生成された複数のクラスタからその各クラスタの重心に最も近いものをそれぞれキーフレーム候補として抽出する。
- 自分の求めている詳細さによってキーフレームを抽出するフレームの数を決める。このとき、より要素数の多いクラスタから抽出したキーフレーム候補は、その提示優先順位を高くする。これは、3.1(b)の前提より、要素数の多いクラスタのほうが要素数の少ないクラスタよりも有意と考えられるためである。

なお、単純クラスタリングを行なう時に、区間内での時間軸を考慮するとビジュアル的に類似した余分なフレームを抽出

各区間内でフレームの中央に着目して、 $YCbCr$ 空間でのユークリッド距離 D を用いて単純クラスタリングを行う。



してしまう可能性が大きくなるのに対し、区間内での時間軸を考慮しないことで余分なフレームを抽出する可能性を低減できる。これは、ビデオ撮影者が一度撮影したところを、再び撮影するケースが考えられるからである。ここで、 n 個の

特徴ベクトル $(X_1, X_2, \dots, X_n)$ を単純クラスタリングする

手順を簡単にしめす。

- 任意のベクトル X_i をとり、これを第一クラスタ C_1 の中心 $Y_1 (Y_1 = X_i)$ とする。

- 次に、ベクトル X_j をとり、 Y_1 と X_j の距離 $D_{1,j}$ を求め、

$$D_{1,j} > T \quad (6)$$

であれば、 X_j を第二のクラスタの中心 $Y_2 (Y_1 = X_i)$ とする。もし、

$$D_{1,j} \leq T \quad (7)$$

ならば、 $X_j \in C_1$ とし、クラスタの中心 Y_1 を更新する。

- X_k をとり、 Y_1 および Y_2 との距離 $D_{1,j}, D_{2,j}$ を求め

$$D_{1,k} > T, D_{2,k} > T \quad (8)$$

であれば, X_k を第三クラス C_3 の中心 $Y_3(Y_3 = X_i)$ とする.

$$D_{1,k} \text{ or } D_{2,k} \leq T \quad (9)$$

ならば, X_k は中心との距離の短い方のクラスに所属するものとする.

- (d) すべてのベクトルについて, 上記 (a)-(c) の処理を行ない, クラスタリングを完了する. なお, クラスタの中心は新しい要素が増えるごとに更新する.

本稿の場合, 新しい要素が増える前のクラスタの中心を $(Y_{cen}, Cb_{cen}, Cr_{cen})$, 要素数を num , 新しくクラスタに加わる要素を $(Y_{new}, Cb_{new}, Cr_{new})$ であるとき, 新しいクラスタの中心は

$$Y'_{cen} = (num \times Y_{cen} + Y_{new}) / (num + 1) \quad (10)$$

$$Cb'_{cen} = (num \times Cb_{cen} + Cb_{new}) / (num + 1) \quad (11)$$

$$Cr'_{cen} = (num \times Cr_{cen} + Cr_{new}) / (num + 1) \quad (12)$$

とする. クラスタリングの結果, クラスタの数が多過ぎるか, または少な過ぎた場合, T の値を変えて再びクラスタリングを行なう.

4. 実験, 評価

4.1 実験

本システムにより, キーフレーム抽出処理の実験を行った. 実験で用いた動画は一般家庭用デジタルカメラ (SONY 製, DCR-PC1000) によって撮影された映像で, 実験映像は撮影者が自分の仲間達を撮影した映像 (720×240, 30 フレーム/秒, ppm 形式, 撮影時間=31 秒) である. しきい値は, フレームの全体のヒストグラムの類似度 m を求める時のしきい値には $Th_{all}=121000$, フレームの中央のヒストグラムに着目した場合には $Th_{center} = 32000$, そしてユークリッド距離のしきい値として $D=30$ を用いた. ここで, 素材映像が分割された区間の区間番号とその区間から生成されたクラスタを代表するキーフレーム候補番号, クラスタの要素数を区間の範囲が 5 フレーム以上である区間に限定して表 1 に, フレームの中央の輝度平均, 色差平均のグラフを図 5 に類似度 m のグラフを図 6 示す.

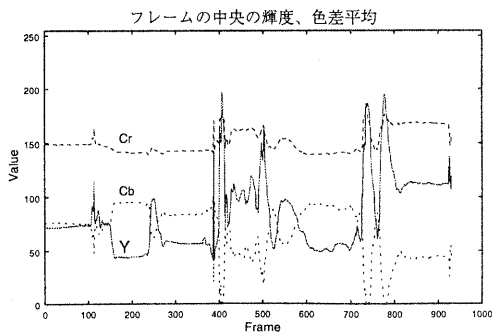


図 5: フレーム中央の輝度, 色差平均

表 1: 分割された区間番号とキーフレーム候補

区間番号	区間の範囲	キーフレーム候補番号 (クラスタの要素数)
1	0-123	84(121), 113(2)
2	124-247	155(118), 244(5)
3	248-255	252(7)
4	256-389	300(130)
18	417-423	423(6)
20	428-455	450(27)
21	456-515	457(51)
22	516-538	532(22)
23	539-723	713(184)
28	737-742	741(5)
32	751-762	756(11)
34	767-774	770(7)
44	796-801	799(5)
47	810-815	812(5)
48	816-824	821(8)
49	825-924	857(99)
50	925-930	930(5)

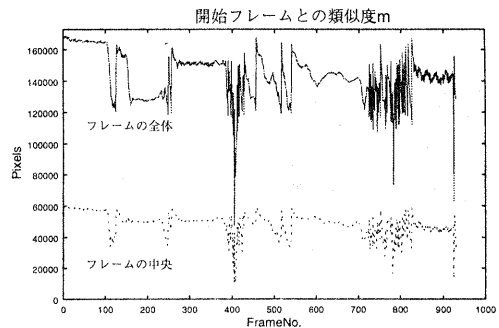


図 6: 類似度 m

4.2 評価

ここで, システムの評価をするにあたり, 撮影者に撮影の意図の観点からどのフレームが実験映像のダイジェストを知るうえで必要となるか調査を行ったところ図 7 のような結果となった. なお, 撮影者以外の一般のユーザにも調査を行ったところ同じ結果となった. また, 本システムから得られたキーフレーム候補のうち, キーフレーム提示の優先順位が 9 位までのフレームをキーフレームとして図 8 に示す.

以上の結果より, 過剰検出になるようにキーフレームを抽出すれば撮影者の必要とするフレームを提示することが可能となることが分かった. しかし, 他の実験映像でショットを区間に分割するとき過分割が起きてしまい本来のキーフレームとして提示する優先順位よりも低くなってしまった場合もあった. また, 撮影者が必要とするフレームとは, 目的の

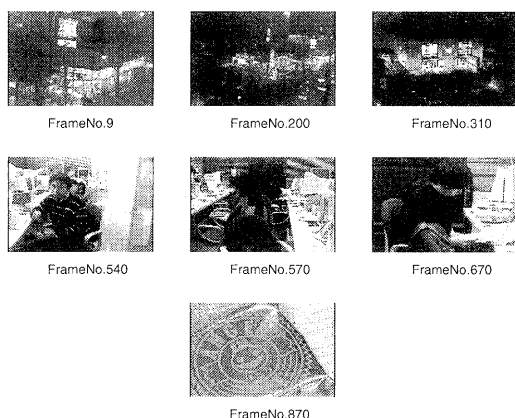


図 7: 撮影者が必要と考えたフレーム

本システムによるキーフレーム(提示枚数9枚)



フレームの並び方はフレーム番号順(時間順)である。

図 8: システムが提示したフレーム

物が写っているフレームであれば基本的には何でも良いと幅が広く、できればその中でもぼけていないはっきりしたフレーム、人間が笑顔で写っているフレームがよいなどの細かい要望もあった。

5. むすび

本稿では、カメラ操作によって発生する画面全体に渡る動きに対して、人間の視覚的類似度を反映することができるようするために、類似度を定義する画像特徴量としてヒストグラムを用いた。そして、撮影者の意図を含むものとして、フレームの中央に撮影者が有意と考えるものがある、撮影者が有意と考えるものほど撮影時間が長いという前提を元にキーフレームを抽出した。また、撮影者の意図を含むものとしてカメラワークも考えられ、動画からカメラワーク、手ぶれ

を検出する研究も行われている⁽¹⁵⁾。キーフレームを抽出するときにカメラワーク、手ぶれなどの情報も用いれば、さらに効果的なキーフレームの提示ができると考えられる。また、ショットを区間に分割する時に本システムでは輝度情報しか用いていないが、色差成分なども考慮するとより適切な区間に分割することが可能になると考えられる。なお、本稿のシステムの課題として、動画像では映像中の物体の動きや、変化がその映像の意味内容に関わっている場合も考えられるが、現システムでは動画像中の動作に関してはほとんど考慮していないということが挙げられる。また、現システムでは3つのしきい値をあらかじめ決めておいたが、映像の種類によって最適なしきい値も変化すると考えられるので、しきい値も映像によって変化させる必要がある。

参考文献

- (1) 粕谷英司: "映像の内容検索とその記述方式に関する研究" 早稲田大学理工学研究科修士論文, Feb.1998
- (2) Yeung, M.M., Yeo, B.L. and Liu, B.: Extracting Story Units from Long Programs for Video Browsing and Navigation, Proc. IEEE Intl. Conf. on Multimedia Computing and Networking, Vol. 2417, pp. 389-398 (1995)
- (3) 大辻清太, 外村佳伸, 大場有二: "動画カット検出" 信学技報, IE91-116, 1991
- (4) 菅野勝, 中島康之: "部分復号を用いた MPEG データからのカット点検出" 画像ラボ Apr. 1999
- (5) 田中聡, 脇本浩司, 神田準史郎: "シーン検出による動画情報の自動要約・閲覧技術の開発" 信学技報 IE99-20, PRMU99-44, MVE99-40 (1999-07)
- (6) 谷口他: "映像ショット切替え検出法とその映像アクセスインターフェースへの応用", 信学論, Vol. J79-D-2, No. 4 pp. 538-546, 1996
- (7) 末永丈士, 上原英昭, 他: "特徴的フレームの抽出とそれに基づく時系列画像の簡潔な表現" 信学技報 PRU94-113 (1995-01)
- (8) 越後富夫: "視覚的な手がかりによる動画検索", IPSJ Magazine Vol. 41 No. 4 2000
- (9) 有木康雄: "社会を支える映像情報メディア" 映像情報メディア学会 Vol. 53, No. 1, pp. 34-40 (1999)
- (10) ZHUANG Y, RUI Y, HUANG T S, HUANG T S, MEHROTRA S: "Adaptive Key Frame Extraction Using Unsupervised Clustering" Proc Int Conf Image Process 1998 Vol. 1 (USA) PAGE. 866-870
- (11) GUENSEL B, TEKALP A M: "Content-Based Video Abstraction" Proc Int Conf Image Process 1998 Vol. 3 (USA) PAGE. 128-132
- (12) 真鍋佳図士, 他: "動画データベースにおける内容検索のためのキーフレームの検出" 信学技報 PRMU98-262 (1999-03)
- (13) 串間和彦, 赤間浩樹, 紺谷精一, 山室雅司: "色や形状等の表層的特徴量にもとづく画像内容検索技術" 情報処理学会論文誌 Vol. 40 No. SIG3 (TOD1)
- (14) 小谷一孔: "コンピュータグラフィックにおける色再現" テレビジョン学会誌 Vol. 47, No. 9, pp. 1216-1224 (1993)
- (15) 土橋健太郎, 小館亮之, 西塔隆二, 富永英義: "MPEG-2 動きベクトルを用いたカメラワークの検出の検討" PCSJ2000 p. 113-114