

平成 22 年 5 月 12 日現在

研究種目：若手研究 (B)

研究期間：2008～2009

課題番号：20700074

研究課題名 (和文) 差分量子化によるビデオの自動インデキシングシステムの研究

研究課題名 (英文) Research on Automatic Video indexing System Using Adjacent Pixel Intensity Difference Quantization

研究代表者

李 菲菲 (LEE FEIFEI)

東北大学・未来科学技術共同研究センター・助教

研究者番号：70451549

研究成果の概要 (和文)：

ストレージの大容量化やネットワークのブロードバンド化により、大量のビデオデータを容易にやり取りできるようになった。膨大なビデオデータベースを利用するには、人手を介さずに自動的にビデオのインデキシングを行うことが望まれている。本研究では、ビデオシーン切り替え検出し、人物が含まれるフレームに対して、高速かつ高精度で顔追跡をし、自動ビデオインデキシング手法を試みた。差分量子化によるヒストグラム特徴を用いて、時系列アクティブ探索法と組み合わせて、映像信号の高速探索法を提案し、従来の高速探索手法により高いロバスト性を実現した。

研究成果の概要 (英文)：

Video indexing has become an active area of research in recent years because video content becomes commonplace on the web and the size of video database quickly increases due to rapid developments of internet connection and disk storage technology. In this research, we sought to develop an automatic video indexing system for large video database. The adjacent pixel intensity difference quantization (APIDQ) histogram is used as the feature vector in this algorithm. Experimental results show the proposed algorithm can detect the similar video clip more accurately and robust against Gaussian noise than conventional fast video search algorithm.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	1,500,000	450,000	1,950,000
2009年度	1,000,000	300,000	1,300,000
年度			
年度			
年度			
総計	2,500,000	750,000	3,250,000

研究分野：総合領域

科研費の分科・細目：情報学・メディア情報学・データベース

キーワード：マルチメディア、データベース、ビデオインデキシング、差分量子化、顔認識、高速探索

1. 研究開始当初の背景

情報化社会の進歩につれ、さまざまな分野におけるデジタル化技術の急速な発展により、複数のメディアが統合された高度マルチメディア情報処理環境の構築に対する要望が強い。一方、ストレージの大容量化やネットワークのブロードバンド化により、大量のビデオデータを容易にやり取りできるようになった。しかし、現状のマルチメディア環境では、未整理で構造化されていないデータが溢れている。例えば、電子博物館や図書館などで保存されているデータは一般的に組織化されているが、テレビで放送されているニュースやアニメ、ドラマ、映画などのビデオ情報は、検索の対象としにくい未整理で構造化されていないデータである。ビデオデータなどを蓄積し、放送、スポーツ、警備システムなど、多岐にわたる分野での共用・再利用が期待されている。膨大なビデオデータベースを利用するには、ビデオのインデキシングが不可欠である。ここでは、映像メディアを検索する際に有効な目印を付与することをビデオインデキシング (Video Indexing) と呼び、この目印をインデックス (Index) と呼ぶ。大量のビット情報を対象とするので、人手を介せずに自動的にビデオのインデキシングを行うことが望まれている。

現状では、ビデオのインデキシングについては、映像セグメンテーションを中心として研究されてきた。一般的に、シーン切り替え検出という手法が取られている。従来よく使われてきたシーン切り替え検出手法では、たとえば、輝度ヒストグラム、尤度比、予測誤差などが使われている。ビデオ編集用などに実用化し商品化されたものも出来ている。しかし、映像を時間軸上で何らかの規範に従って分割することは、映像メディアの構造化への第1歩にしかすぎない。膨大なビデオデータベースを知的な構造を持たせるには、意味内容に基づく検索 (CBR: Content Based Retrieval) 手法を確立することが不可欠である。

一般的に撮影する映像には撮影対象が人物である場合が多く、人物インデックスの有用性は特に高いと言える。登場人物や人数、人物の動きなどのインデックス情報を得ることができれば、映像を素早く検索することが可能になり、映像閲覧、編集を効率化が図れる。

申請者らはここまで「差分量子化を用いた画像認識」という新しい画像処理手法を提案してきた。これは、差分量子化により、画像情報から差分ベクトルの頻度分布という低次元の特徴量を抽出して用いる画像認識手法である。この原理に基づいた簡便かつ高い認識率を有する顔認識手法について研究を行ってきた。実環境下での認識実験において、顔領域分割処理と複数フィルタの組み合わせ

せにより、25日までの時間経過に対して100%の認識率を実現した。一枚の顔画像に対する全認識処理時間は、一般的なパーソナルコンピュータを用いた場合で31 msであり、ビデオレートの認識速度を実現した。

また、差分量子化手法のビデオシーン切り替え検出への適用に関して研究し、差分量子化を用いることにより、高精度にビデオシーン切り替え検出が可能であることを示した。さらに、動的しきい値設定手法を確立し、シーン切り替えの自動検出を実現した。また、現在標準的な動画圧縮フォーマットである MPEG データフォーマットの DC 画像に差分量子化を適用することで、完全な MPEG デコードを実行することなく圧縮データから直接シーン切り替え検出が可能であることを示した。

これらの研究成果は、上述ビデオの自動インデキシング技術実現に向けてその基礎となる重要な成果である。

2. 研究の目的

本研究では、①申請者らが開発したロバストなビデオシーンの切り替え検出技術、②申請者らが開発した実環境でロバストな顔検出技術、③これまでの処理手法とまったく異なる、申請者らが提案した新しい概念の差分量子化による人物認識技術、④申請者らが提案した顔画像を 31ms のビデオレートで認識できる高速認識手法に基づき、ビデオシーン切り替え検出し、人物が含まれるフレームに対して、高速かつ高精度で顔追跡をし、自動ビデオインデキシング技術を実現する。

現在行われている一般的にシーン切り替え検出手法において、輝度情報を用いた切り替え検出手法は照明と色の小さい変動に弱く、尤度比と予測誤差を用いた検出手法は計算量が比較的に大きいなどデメリットがあるため、われわれは①に基づき、ビデオシーケンスに対して、差分量子化処理することにより高速でそれぞれの画像フレームの頻度分布を生成し、隣の画像フレーム同士頻度分布の変化を分析し、ディゾルブ・フェードなどを含むさまざまなタイプのビデオシーン切り替えを検出する。

その後、②に基づき、同じシーンのシーケンスの中では、人物が存在するかどうかを判断し、人物の数、位置およびサイズなどを検出する。また、検出された顔の解像度が十分であれば、③による人物認識技術および④による高速認識手法にベースにして、差分量子化処理することにより顔の特徴ベクトルを生成し、高速かつ高精度で顔認識をする。その後、ビデオの時間的連続性があることを利用し、得られた顔領域の中心位置を次の画像フレームにおける検索スタート位置とし、顔領域の差分量子化ヒストグラムをテンプレート

として使い、前のフレーム中の顔領域周辺を重点に置き、効率よくシーケンスの中で高速で顔追跡し続ける。シーンのシーケンスに自動的に顔インデキシング情報を付与することにより、自動ビデオインデキシング技術を確立することができる。

3. 研究の方法

(1) 隣接画素差分量子化ヒストグラム法

図1は隣接画素差分量子化の概要を示している。手順としては、まず、入力画像を用い、フィルタ処理を行う。これは移動平均フィルタを使う。それから、隣接画素差分計算をする。

具体的には、入力顔画像の画面上任意の一点に対して、 x 方向の差分 dIx はこの点の輝度値と右側隣接点の輝度値の差で、 y 方向の差分 dIy は下側隣接点の輝度値の差により計算できる。計算された dIx 、 dIy のペアは $dIx-dIy$ 平面のひとつの輝度変化ベクトル点として表現できる。差分ベクトル点の分布は、原点を中心として集中し、差分値が大きくなると、だんだん少なくなる傾向が見られる。これは人間顔面の輝度変化は大半緩やかであることから説明できる。 $dIx-dIy$ 平面分布の密度と形状は良く顔の特徴を表している。この空間を量子化することにより、顔特徴ベクトルが得られる。 dIx 、 dIy がある半径 r と角度 θ に変換し、それは実際輝度変化量と変化方向に対応する。

$dIx-dIy$ 平面分布の特徴を注目し、全領域はそれぞれの半径 r を持つ円とそれぞれの角度によって、扇形領域に分割される。ベクトル点は半径の小さい領域に集中しているので、半径 r の方向に非線形ステップを用いた。

われわれは APIDQ (Adjacent Pixel Intensity Difference Quantization) と呼ぶ。これにより、ヒストグラムが生成される。これを顔の特徴ベクトルとして利用する。認識システムとして、登録の段階では、この特徴ベクトルが個人情報としてデータベースに登録される。認識の段階では、入力された顔画像より特徴ベクトル(ヒストグラム)が生成され、あらかじめ登録された個人毎の特徴ベクトルと比較され、最も距離の近いものが認識結果として出力される。公開の AT&T 顔画像データベースを利用した認識実験において、95.7%という高い平均認識率が得られている。しかも、テーブル参照法により、特徴ベクトルの生成時間は 1msec に実現している。

この手法の本質というのは、差分量子化により、画像ブロック内の輝度変化方向と変化量を検出しているのである。よって、APIDQ で生成された画像の特徴ベクトル(ヒストグラム)は画像の特徴を表せる。

(2) 差分量子化による高速探索法

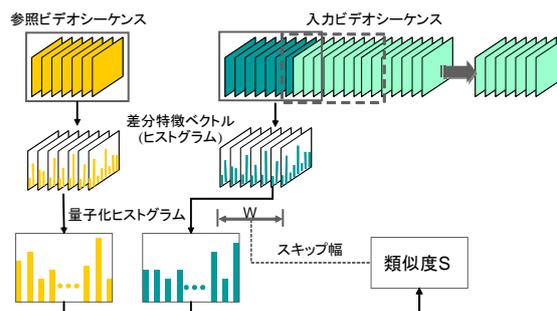


図1 差分量子化による高速探索法の概要図

図1は時系列アクティブ探索法処理の概要図である。まず差分量子化により、参照ビデオ信号(短い信号)と入力ビデオ信号(長時間信号)からそれぞれ特徴ベクトル(差分量子化によりヒストグラム特徴)を抽出する。

つぎに、参照信号と入力信号の両方に対して同じ長さの時間窓をかけて、窓内の特徴ベクトルを量子化して、各量子化符号の出現回数を数えてヒストグラムを作る。

そして、ヒストグラム同士の類似度が、あらかじめ設定した閾値を超えるかどうかで、参照信号のあり/なしを判断する。類似度の値と閾値から、探索を時間方向にスキップできる時間幅を求めることができるので、その分だけ入力信号に対する窓をずらして探索を進める。

4. 研究成果

本研究の研究成果は以下ようになる。

自動ビデオインデキシング技術を実現するため、前処理としてロバスト性のビデオシーン切り替え検出手法が不可欠である。現在標準的な動画圧縮フォーマットであるMPEGデータフォーマットのDC画像に差分量子化を適用することで、完全なMPEGデコードを実行することなく圧縮データから直接シーン切り替え検出を行った。公開のビデオデータベースからいくつかのMPEGビデオを用いて、従来手法に比べ、かなり良い結果が得られた。直接MPEGビデオのシーン切り替え検出が可能であることを示した。

また、差分量子化によるヒストグラム特徴を用いて、時系列アクティブ探索法と組み合わせ、映像信号の高速探索法を提案した。汎用のPCを用いて、6時間分の映像から15秒の参照信号を従来の高速探索手法560ms掛かった1/7の時間80msの探索時間で正しく検出することを実現した。また、白色ガウス雑音に対しても、従来の高速探索手法によりロバストであることが分かった。SN比が30dB以上であれば、提案手法のほうは99.5%以上の探索精度が得られる。

また、MPEGビデオデータベースの適用を試

みた。現在標準的な動画圧縮フォーマットである MPEG データフォーマットのビデオデータは完全な MPEG デコードを実行することなく圧縮データから部分復号し、DC 画像を生成することにより、ヒストグラム特徴を抽出できる。汎用の PC を用いて、6 時間分の MPEG ビデオデータから 15 秒の参照信号 200 本を探索する結果、従来の高速探索手法によりロバスト性の高い 3% の ERR を実現した。

さらに、差分量子化による顔認識手法を改良した。顔認識パーツ分割処理を加えることにより、位置情報とヒストグラム情報を一緒に取り込み、総合判断することで顔認識率が向上することを明らかにした。現在最も標準的な FERET データベースで評価した結果、表情、姿勢の変化を含むタスクにおいて、Top1 認識率 97.6% を得た。

また、差分量子化による顔認識手法について量子化方法を検討した。これまで、隣接画素の差分処理したあとに、 x 方向と y 方向の差分 dI_x , dI_y を極座標に変換し差分ベクトルの半径 r 、角度 θ で量子化をするが、 dI_x - dI_y の平面を直接量子化する方法を検討し、 dI_x - dI_y 平面の直交座標を長方形の量子化領域を分割し量子化する手法を試みた。公開の AT&T 顔データベースで評価した結果、直交座標での量子化方法は 97.2% の最大平均認識率が得られた。オリジナル差分量子化手法の結果 95.7% により大きく向上させた。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 4 件)

- (1) Feifei Lee, Koji Kotani, Qiu Chen, Tadahiro Ohmi, “Fast Video Search Algorithm for Large Video Database Using Adjacent Pixel Intensity Difference Quantization Histogram Feature”, *International Journal of Computer Science & Network Security*, Vol. 9, No. 9, pp. 214-220, 2009. (査読有)
- (2) Feifei Lee, Koji Kotani, Qiu Chen, Tadahiro Ohmi, “Face Recognition Using Adjacent Pixel Intensity Difference Quantization Histogram”, *International Journal of Computer Science & Network Security*, Vol. 9, No. 8, pp. 147-154, 2009. (査読有)
- (3) Feifei Lee, Koji Kotani, Qiu Chen, Tadahiro Ohmi, “Shot change detection using a novel histogram feature in compressed video”, *IEICE Electronics Express*, Vol. 5, No. 14, pp.503-509, 2008. (査読有)

[学会発表] (計 10 件)

- (1) Feifei Lee, Koji Kotani, Qiu Chen, and Tadahiro Ohmi, “Fast and Efficient Search for MPEG-4 Video Using Adjacent Pixel Intensity Difference Quantization Histogram Feature”, *2010 International Conference on Digital Image Processing (ICDIP 2010)*, edited by K. Jusoff, Y. Xie, *Proceeding of SPIE*, Vol. 7546, 75460K, Singapore, Feb. 26-28, 2010. (査読有)
- (2) Feifei Lee, Koji Kotani, Qiu Chen, and Tadahiro Ohmi, “A Modified Adjacent Pixel Intensity Difference Quantization Method for Face Recognition”, *2009 Int'l Conf. on Future Information Technology and Management Engineering (FITME 2009)*, pp. 533-536, Sanya, China, Dec. 13-14, 2009. (査読有)
- (3) Feifei Lee, Koji Kotani, Qiu Chen, Tadahiro Ohmi, “Fast Search for MPEG Video Clips Using Adjacent Pixel Intensity Difference Quantization Histogram Feature,” *Proc. of the International Conference on Image and Vision Computing (ICIVC 2009)*, pp. 777-780, Paris, France, Jun. 24-26, 2009. (査読有)
- (4) Feifei Lee, Koji Kotani, Qiu Chen, Tadahiro Ohmi, “Fast Search Algorithm for Short Video Clips from Large Video Database Using a Novel Histogram Feature,” *Proc. of the Int'l Conf. on Computational Intelligence for Modelling, Control and Automation (CIMCA2008)*, Vienna, Austria, pp. 1223-1227, Dec. 10-12, 2008. (査読有)
- (5) Feifei Lee, Koji Kotani, Qiu Chen, Tadahiro Ohmi, “Face Recognition Based on the Combination of Histogram Features and Rough Location Information of Facial Parts,” *Proc. of 9th Int'l Conf. on Signal Processing (ICSP2008)*, pp.2065-2069, Beijing, Oct. 26-29, 2008. (査読有)
- (6) Feifei Lee, Koji Kotani, Qiu Chen, Tadahiro Ohmi, “An improved fast face recognition algorithm based on adjacent pixel intensity difference quantization histogram,” *Proc. of the 2008 Int'l Conf. on Wavelet Analysis and Pattern Recognition*, pp.316-320, Hong Kong, Aug. 30-31, 2008. (査読有)
- (7) Feifei Lee, Koji Kotani, Qiu Chen, Tadahiro Ohmi, “Fast Search for Large Video Database Using Adjacent Pixel Intensity Difference Quantization Histogram”, *Proceeding of the 10th IASTED Int'l Conf.*

on *Signal and Image Processing (SIP 2008)*,
623-181, pp.114-117, Kailua-Kona, HI,
USA, Aug. 18-20, 2008. (査読有)

6. 研究組織

(1) 研究代表者

李 菲菲 (LEE FEIFEI)

東北大学・未来科学技術共同研究センター・助教

研究者番号：70451549

(2) 研究分担者

()

研究者番号：

(3) 連携研究者

()

研究者番号：