

映像同定技術 “Video Signature” と その応用技術

工藤大樹*
鶴崎真理子*
阿倍博信**

Outline and Application of Video Identification Technology "Video Signature"

Daiki Kudo, Mariko Tsurusaki, Hironobu Abe

要 旨

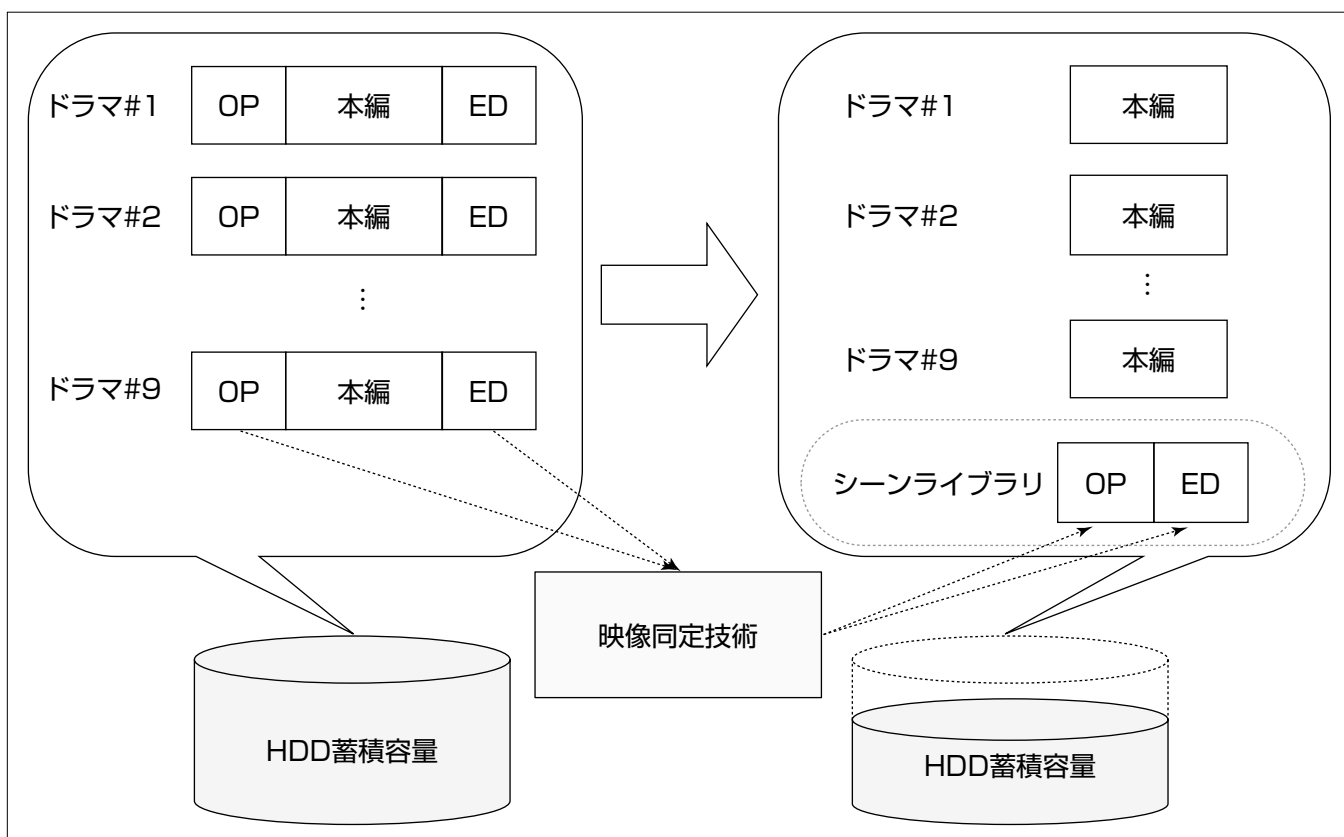
大量の画像から所望の画像を検索する“画像検索技術”には、用途・目的に応じて様々な方式が存在する。その一つに同一の映像を検索する“映像同定技術”がある。既に、各種の加工に対しても頑健に同一性の検証が可能な手法が確立しているため、インターネット上に不正に共有された映像の発見などに利用できる。本稿では、この映像同定技術の新たな応用として、HDD(Hard Disk Drive)レコーダ向けの蓄積効率向上について述べる。

一般に、レコーダに蓄積された映像には、ドラマのオープニングなど同一映像が複数記録されることが多い。これを、映像同定技術によって検出し、シーンライブラリとしてまとめる。元の映像からは、検出された同一映像は削除し、再生方法を示したメタデータを付与し、再生時はこの

メタデータを利用して、シーンライブラリから映像を読み出し、再生する。これによって効率的な映像蓄積が実現できる。

この方式では、処理時間と、映像の編集処理が問題となる。そこで、前者に対しては、音声情報を利用し、処理区間の候補を絞り込むことで高速化を行った。後者に対しては、映像符号化の単位(Group Of Picture : GOP)に基づく検出処理を行うことで、映像の削除、接続のための付加的な復号・再符号化処理を不要としている。

今後、同一映像だけでなく、類似映像を利用した方式についても検討し、単に効率的な蓄積にとどまらず、映像情報の構造化・情報活用にも応用する予定である。



映像同一部分を利用したHDD蓄積効率向上

同一ドラマをシリーズ(#1~#9)で録画したHDDレコーダを想定している。ドラマのオープニング(OP)とエンディング(ED)は同一映像であることから、これらをシーンライブラリとして蓄積し、各ドラマは本編だけを蓄積している。これによって、HDD全体での蓄積効率向上する。再生時はメタデータを利用することで、オープニングとエンディングを含めて再生する。

1. ま え が き

撮像から蓄積・配信にいたるまで、画像(静止画)・映像(動画)処理技術は急速に発達し、ユーザーは大量の画像・映像を保持し、気軽に楽しめるようになった。

これに伴い、所望の画像を探す画像検索技術の研究も盛んに行われている。画像検索には、幅広い用途・目的に合わせた様々な技術が存在する。例えば、キーワードを使ったインターネット上の画像検索⁽¹⁾のほか、顔や人物の検出⁽²⁾、構図を利用した映像検索技術⁽³⁾、同一の画像・映像を検出する画像同定技術⁽⁴⁾等が挙げられる。

ユーザーが保持している画像に目を向けると、デジタルカメラやスマートフォンで自ら撮影したものと、他者が制作したコンテンツの2つに大別できる。このうち、後者の代表的なものとして、家庭用HDDレコーダ(Hard Disk Driveを用いる映像蓄積装置、以下“HDDレコーダ”という。)内の録画番組が挙げられる。

HDDレコーダは、当初、アナログ放送を録画するものから始まり、その後、高画質のデジタル放送サポートや搭載チューナー数の増加による複数番組同時録画などによって、録画データ量は飛躍的に増大している。大容量のHDDを持つ機器でさえ、蓄積容量が不足する場合も多く、外部HDDを接続し利用するユーザーも少なくない。今後も、より高解像度の4K/8K放送が予定され、HDD容量はますます逼迫(ひっばく)するものと思われる。

この解決方法の一つが、放送映像をより高効率の符号化方式(例えば、H.264)に変換、圧縮する方法があり、既にHDDレコーダでの実用化もされている。これは、録画番組単体での蓄積効率向上を図る方法である。

これに対し、本稿で述べる技術は、HDDに蓄積される各映像の特徴量を抽出し、その映像間の同一性を利用して蓄積効率向上を図る手法である。

2. 画像・映像同定技術

2.1 概 要

画像・映像の同一性を判定する技術が研究開発され、標準化が行われた。これが、画像・映像同定技術(Visual Signature)⁽⁴⁾であり、静止画向けのImage Signature⁽⁵⁾と、映像向けのVideo Signature⁽⁶⁾の2種類からなる。

画像・映像同定について、静止画像の場合を例に述べる(図1)。女性の画像(図中の“検索キー画像”)に対して、“類似”の意味を、“女性の画像”とするのであれば、図中の全ての画像が検索結果として出力されるべきである。しかし、画像同定は、画像コンテンツの同一性を判定することが目的であり、回転・モノクロ化・左右反転等の画像処理を施し、画素のパターンとしては全く異なる画像でも、同一として、検索する技術である。

映像データに対する同定を目的としたVideo Signatureは、この画像同定の技術を動画に応用した技術で、図2で示しているように、フレームごと、すなわち個々の静止画の特徴量と、フレーム90枚で構成される一定時間長(セグメント)ごとの特徴量から同一性を検出する手法である。個々のフレーム特徴量に対しては詳細なマッチングを、セグメント特徴量に対しては、大まかなマッチングを行うことで、フレームの加工やフレームレートの変化等、映像加工に対しても誤りが少なく、頑健性の高い同定が可能である⁽⁴⁾。



図1. 同一の画像

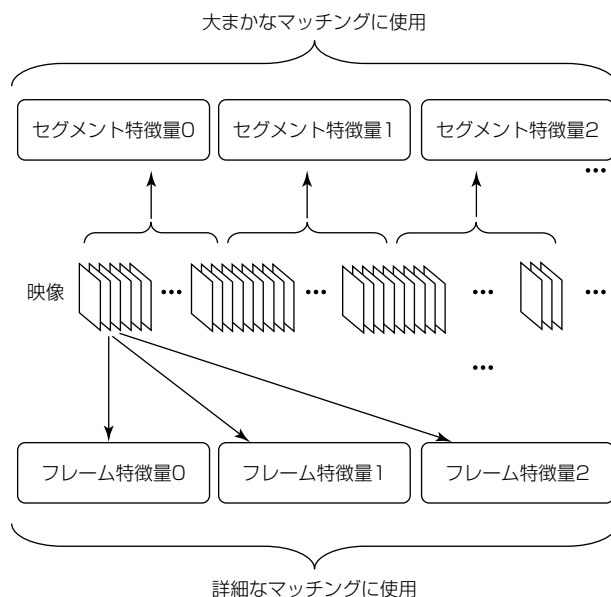


図2. Video Signatureで使用する特徴量

2.2 Video Signatureの応用

Video Signatureは、インターネット上の動画共有サイトなどでの映像コンテンツの不正流通の検出を目的に考案されたものであり、その有効性が確認されている。

3. 映像同定を応用した蓄積効率向上

3.1 概要

HDDレコーダに蓄積されている番組は、ユーザーの嗜好(しこう)に基づき、シリーズ番組(ドラマ・アニメ等)をまとめて記録しているケースが少なくない。これらシリーズ番組のオープニングとエンディングや、スポーツ番組におけるリプレイ等、レコーダ内には同一のシーンが多数存在する。

映像同定技術を利用すれば、レコーダ内で同一の映像部分を特定することが可能になり、レコーダ全体の蓄積容量を削減することが可能になる。

図3に例を示す。映像A/Bから同一映像部分(①, ②)を特定する。同一映像部分は、“シーンライブラリ”として別途保存し、元の映像からは削除する。これをレコーダ内に存在する全ての映像に対して行うことによって、レコーダ内に複数存在する同一部分はシーンライブラリに保存される一つを除き削除でき、蓄積容量の削減ができる。

なお、編集済み映像には、再生用のメタデータを付与する。この記述に従い、シーンライブラリを参照、再生することで、元の映像と同じ映像の再生が可能である。

3.2 手順

3.1節で述べた蓄積効率向上の方法を実現するため、具体的には次の手順で行う。

- (1) レコーダ内に記録されている映像のVideo Signature特徴量を抽出する。
- (2) 抽出した特徴量を利用して、レコーダ内の各映像の同一部分を特定する。
- (3) 特定した同一映像部分を抜き出してシーンライブラリを作成する。
- (4) 各映像から同一映像部分を削除して編集済み映像を作成する。

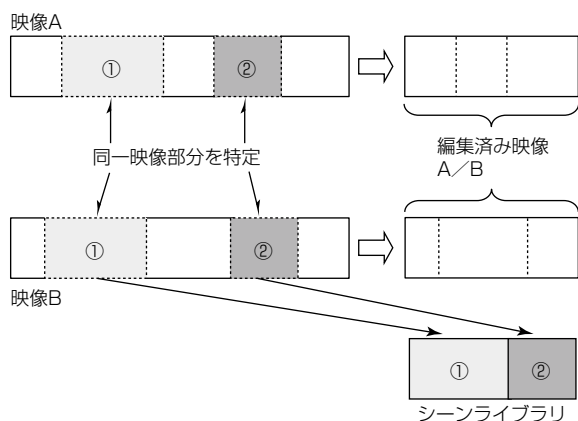


図3. 同一映像部分を利用した蓄積容量削減

成する。同時にそれぞれの再生用メタデータを作成する。

この手順は、大きくは前半(1), (2)の特徴量抽出・同一映像部分特定と、後半(3), (4)のシーンライブラリ及び編集済み映像の作成(編集)に分かれる。それぞれの処理高速化について述べる。

3.3 特徴量抽出速度の改善

Video Signatureを使用した映像同定を行う場合、映像中の全フレームについて特徴量を抽出するため、処理量が大きなものとなる。そこで、音声情報を利用し、同一映像区間候補を限定し、処理量を削減する方式を考える。

図4で示すように、映像コンテンツでは、シーンが切り替わる箇所に無音区間が存在することが多い。そこで無音区間を判定し、その位置から映像コンテンツ内の同一シーンを含む区間の候補を抽出する。特徴量抽出は、この区間に対してだけ行う。これによって、特徴量抽出だけではなく、同一性の検証のためのマッチング処理についても大きな処理量削減が可能となる。

実際のテレビ番組を用い評価実験を行った。音声情報による特徴量抽出区間の限定を行ったところ、特徴量抽出処理時間は通常のVideo Signatureを使用したときの38%に削減することができた。さらに、マッチング処理を含めた全体の処理時間は、45%に削減となった(図5)。

今後の実用化に向けては、更なる処理量削減が必要である。これは、同一映像区間候補の抽出精度を高めることで実現できる。声のパワーの一定しきい値判断による現状の無音区間判定は誤検出も多く、処理量削減効果が十分ではない。この精度向上に加え、シーンの切り替わりを特定する無音以外の情報の活用も可能である。例えば、EPG(電子番組表)の時刻情報を用いれば、番組自体の開始・終了時刻を特定できる。今後これらを検討するとともに、サーバで精度よく同一映像区間を検出、配信するクラウドサービスの可能性も検討する予定である。

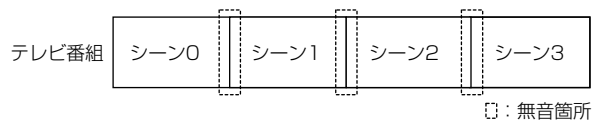


図4. 無音区間利用法

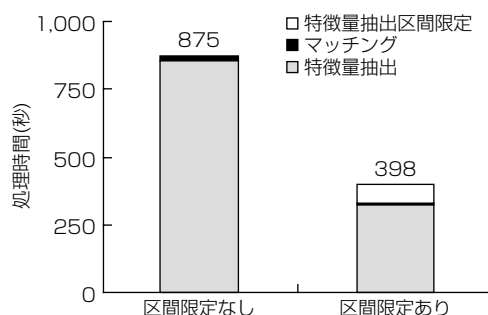


図5. 特徴量抽出処理時間の削減

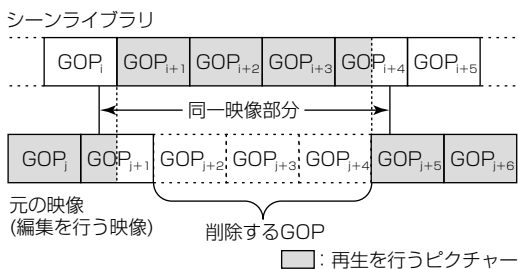


図 6. GOP単位での編集

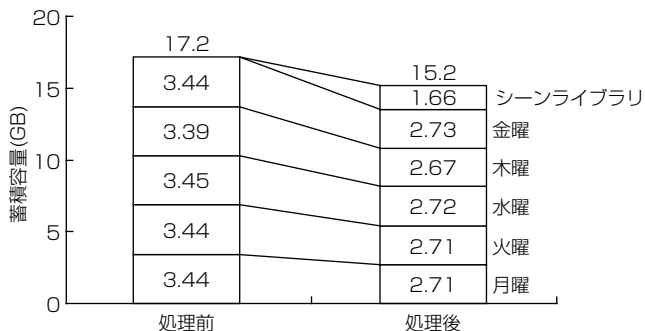


図 7. 蓄積容量の変化

3.4 編集方法

同一映像部分の削除，シーンライブラリから読み出した映像との接続では，映像符号化技術との整合性が問題となる。H.264に代表される映像符号化技術では，GOPと呼ばれる単位で，画像間の相関を利用した圧縮・符号化が行われる。そのため，映像符号は，前後のフレームに依存した情報であり，編集処理を行うには，一度復号し，編集を行い，再符号化を行う必要がある。この処理量も大きなものとなる。

そこで，符号化方式に着目し，復号・再符号化を不要とするGOP単位での編集手法を提案する。

図 6 では，シーンライブラリと元の映像の同一映像部分はGOPの区切りと一致していない。シーンライブラリではGOP_iの途中～GOP_{i+4}の途中まで，一方で編集対象の映像では，GOP_{j+1}の途中～GOP_{j+5}の途中までが同一映像部分である。

このとき，元の映像はGOP_{j+2}～GOP_{j+4}までを削除し，同一シーンの開始GOPであるGOP_{j+1}と終了GOPであるGOP_{j+5}を残す。つまり，同一映像部分に完全に内包されるGOPだけを削除し，シーンライブラリ側は，同一映像部分を完全に外包するGOPで構成する。これによって，編集によってGOPが変化することなく，映像の再符号化処理は不要となる。ピクチャー単位で編集する場合に比べて，蓄積する映像にはオーバーラップ区間が生じるが，蓄積容量は2 GOP増えるだけであり，大きな影響はない。

なお，再生用のメタデータには，ピクチャー単位の再生情報が含まれる。

3.5 削減効果

この手法を平日 5 日分の同一番組（合計 5 時間，平均 8.0MbpsのHD映像）に対して適用し，効果を検証した。

その結果，図 7 に示すように11.6%（17.2GBから15.2GB）の蓄積容量削減効果が得られた⁽⁸⁾。

この手法は，レコーダ内に多くの映像を録画しているほど同一シーンも増加し，効果が大きくなると考えられる。より長期間，番組を録画することで，更なる圧縮効率向上が期待できる。

4. むすび

映像検索技術の一つである映像同定技術“Video Signature”を利用したHDDレコーダにおける蓄積容量削減方法を述べ，11.6%の削減効果があったことを示した。今後，EPGなどの情報を活用すれば，更に効率的な検出も可能となる。

我々は，同一映像部分ではなく，類似の映像シーンを活用する蓄積効率向上についても検討している⁽⁸⁾。さらに，昨年，次世代符号化規格であり圧縮効率が高いH.265が標準化された⁽¹⁰⁾。これらを組み合わせ，HDDを最大限に活用可能な方式を開発する予定である。この技術は，単に蓄積容量の削減にとどまらず，映像情報に対する付帯情報の付与，構造化という観点でも期待が高い。監視映像への適用などへの適用も検討し，映像情報の有効活用を図る。

参考文献

- (1) Google画像検索
http://www.google.co.jp/imghp
- (2) 山内悠嗣，ほか：画像からの統計的学習法に基づく人検出，電子情報通信学会論文誌.D，J96-D，No.9，2017～2040（2013）
- (3) 望月貴裕：画像解析技術を利用した映像検索，NHK技研 R&D，No.128，18～25（2011）
- (4) 工藤大樹，ほか：画像同定技術：Video Signature，三菱電機技報，85，No.11，657～660（2011）
- (5) ISO/IEC 15938-3:2002/Amd 3：2009
- (6) ISO/IEC 15938-3:2002/Amd 4：2010
- (7) M.Tsurusaki，et al.：An Efficient Recording Method for Video Recorders Using the Video Signature Tools，Proc. of 2014，528～529（2014）
- (8) 工藤大樹，ほか：映像同定を利用したレコーダ蓄積容量削減と再生方法，P2-09，PCSJ2014
- (9) T. Semitsu，et al.：Video Re-compression Using Inter-scene Correlation among Multiple Video Streams，Proc. of IWAIT2014，231～236（2014）
- (10) ISO/IEC 23008-2