

特集号  
招待論文

# 映像「指紋」技術 Video Signature の開発—その研究開発, 国際標準化, そして事業化の実践—

岩元 浩太<sup>†1</sup> 鍋藤 悠<sup>†2</sup>

<sup>†1</sup> 日本電気 (株) <sup>†2</sup> (株) NEC 情報システムズ

2000年代に入り、インターネット上での映像の氾濫や不正コピーの横行により、世界中に複製・拡散が進んだ映像の身元 (ID) を識別できる現実的な仕組みの必要性が叫ばれるようになった。そこで筆者らは、人間の「指紋」と同様に、映像自体から固有で頑健な特徴量を抽出することで、映像の ID 識別ができる映像「指紋」技術 Video Signature を開発した。さらに、世界中に拡散した映像を ID 識別できる世界的な仕組みを構築する重要性を認識し、本技術を国際標準化することにも挑戦し、成功した。そして、標準化された Video Signature 技術を用いて、不正流通検知、コンテンツ使用実績調査、映像アーカイブ管理などのさまざまな応用システムを開発し、事業化を進めてきた。本稿では、筆者らが Video Signature 技術を世の中に出していくために実践した研究開発、国際標準化、そして事業化の取り組みを解説する。

## 1. はじめに

2000年代に入り、インターネット上の映像配信や映像共有サイト、簡易に映像撮影ができる端末機器の普及により、世の中に出回る映像コンテンツの量は急速に拡大した。映像を容易に編集・加工できるソフトウェアの普及も相まって、コンテンツ制作における正規の引用・二次利用や不正コピーなどにより、映像コンテンツは爆発的な勢いで複製・拡散が進んでいった。一方で、世界中のサイトやシステムに渡って拡散した映像の身元 (ID) を識別できる現実的な方法が存在しなかったため、映像の管理はますます困難になった。所望の映像を探すのは難しくなり、映像の不正流通が横行するようになった。このため、映像の身元を識別する現実的な仕組みの必要性が急速に拡大していった。

映像の身元を識別する方法としては、従来からコンテンツ自体にIDを埋め込む方法が存在する。ファイルフォーマット中に任意の文字列を書き込む方法や、映像信号中に見えない形で情報を埋め込む電子透かし技術がある。これらは、埋め込まれた情報が解読できれば映像の出所を追跡できる反面、流通前にあらかじめ情報を付加しなければならないため膨大な手間とコストがかかるという問題があり、情報埋め込みせずに流通したものは識別できない。さらに、映像への編集・改変に脆弱であり、容易に埋め込んだ情報が消去されるなど、重大な欠点がある。

そこで筆者らは、インターネット規模の大規模な映像データからでも、その映像の身元 (ID) を識別でき

る映像「指紋」技術 Video Signature<sup>1), 2)</sup> を開発した。Video Signature は、人間の指紋と同様に、映像自体から抽出できる、映像の絵柄を表現する固有な特性を特徴量化したものである。この特徴量は、映像へのあらゆる編集・改変に対して頑健であるため、映像信号レベルで変化しても、元映像との同一性を判断できる。このため、映像から抽出した Video Signature を、あらかじめ DB 登録された原本映像の Video Signature と照合することで、その身元 (ID) を正確かつ容易に識別することが可能となった (図 1)。

さらに筆者らは技術開発だけでなく、インターネットを通じて世界中に拡散した映像を ID 識別できる世界的な仕組みを構築する重要性を認識し、本技術を世界標準化することにも挑戦し、国際標準規格 MPEG-7 Video Signature Tools として標準化することに成功した。そして、標準化された Video Signature 技術を用いて、不正流通検知、コンテンツ使用実績調査、映像アーカイブ管理などのさまざまなシステムを開発し、多数の顧客に対して事業化を進めてきた。

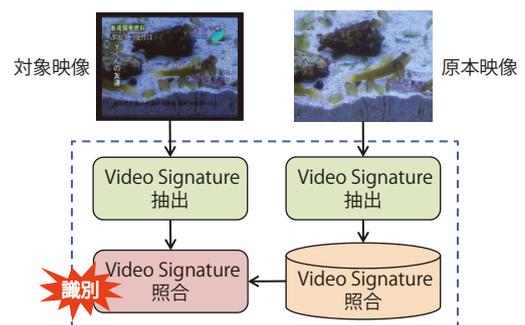


図 1 Video Signature の概要図

本稿では、筆者らがVideo Signature技術の研究開発、国際標準化、そして事業化にあたって実践した取り組みを解説する。

## 2. Video Signature 開発の全体像

### 2.1 背景と課題

映像自体から特徴量を抽出し、それを照合することで類似した映像を検索する技術は、コンテンツベース検索 (CBIR: Content-based Image Retrieval) 技術として、2000年代初頭にさまざまな研究がされていた。これらのCBIR技術では、画像からその絵柄を表す特徴量を抽出し、その特徴量を照合することで類似画像を検索する。映像 (動画) を対象する場合は、映像のフレームから特徴量を抽出し、特徴量系列を照合することで類似映像を検索する。色特徴<sup>3)</sup> や空間特徴<sup>4), 5)</sup> などを用いたさまざまな特徴量が提案されていた。NECでもColor Layout<sup>6)</sup> という名称の色特徴量を開発し、放送事業者向けの検索システムなどの構築を手掛けていた。

2000年代中盤に入り、YouTubeなどの出現によるインターネット上での映像の氾濫が起り、映像の身元 (ID) を識別できる仕組みの必要性が高まると、従来のCBIR技術がターゲットとしていた特定システム内での映像検索ではなく、インターネット規模の映像から唯一特定できる「指紋」として活用できる、まったく新しい発想に基づく強力な特徴量技術が必要となった。すなわち技術的な課題としては、CBIR技術では成し得なかった、(1) 異なる映像は異なるものと判定できる固有性 (ユニーク性) と、(2) さまざまな編集・改変 (テロップ重畳、符号化圧縮、カメラ再撮など) によって信号レベルでは変化した映像でも同一だと判定できる頑健性 (ロバスト性)、の相反する2つの要件を高いレベルで両立することである。世界最高レベルの特徴量技術を創出するとともに、大規模データでそれを立証する必要があった。

さらなる課題としては、従来のCBIR技術のように各システムが独自の特徴量技術を使ってDB構築をされたら、相互に互換性がなく、同じシステム内でしか使えないものになってしまうということである。世界中に拡散した映像を網羅的にID識別できる仕組みを構築するには、世界中で共通に使える特徴量技術として普及させる必要があった。

### 2.2 開発戦略

これらの課題に対処するために、我々は研究開発を(1) Video Signature技術の創出、(2) MPEG国際標準化、(3) システム開発・事業化の3段階で進めていくことにした (図2)。(1) Video Signature技術の創出では、これまでCBIR技術の開発で培ったノウハウを活かし、映像の「指紋」として使える新たな特徴量技術を早期に創出した。次に(2) MPEG国際標準化では、その技術を、国際コンペの場で磨きをかけて世界最高レベルの技術に昇華させるとともに、国際標準化することで相互互換性のある世界共通方式にすることに成功した。そして(3) システム開発・事業化では、標準化した技術をベースに、周辺技術と組み合わせて、さまざまな顧客ニーズに対応するシステムを開発し、早期に多数の事業化を実践した。次章からは、それぞれの活動について解説する。

## 3. Video Signature 技術の創出

我々は、それまでのCBIR技術開発で培ってきたノウハウをもとに、映像の「指紋」として使える新たなVideo Signature特徴量技術と、それを用いた照合技術の開発を、2005年頃から着手した。

### 3.1 Video Signature 特徴量技術の開発

Video Signatureは、映像の各フレームの絵柄を表す特徴量であり、映像を唯一特定できる「指紋」として使えるように、要件である固有性と頑健性を両立するように設計していった。Video Signature特徴量の設計開発の経緯を振り返ると、Version 1～3の3段階で進められていき、その都度、性能を向上していった。

Version 1は、放送局アーカイブの検索システムを想定して2005年頃に開発した方式であり、編集作業で生じるテロップ重畳に対して頑健な方式として開発した。テロップ重畳のように絵柄の一部が変化する改変に対しては、画像全体に波及する特徴ではなく、画像のローカルな特徴の集合であるほうが頑健だという知見を基に、画一的に区切ったブロックごとに、輝度変化に頑健な輝度



図2 Video Signature 開発の3段階の戦略

勾配を量子化する方式<sup>7)</sup>であった。従来の色情報を用いた特徴量方式と比較して、テロップ重畳に対する頑健性を大幅に改善した。

Version 2は、第4章で説明するMPEG国際標準化で2008年に行われた技術募集へ提案するために、Version 1のアイデアをベースに改善を加えたものである。固有性と、テロップ重畳以外のあらゆる編集・改変に対する頑健性を向上させるために、画像のローカルな輝度勾配を特徴化する方法は踏襲し、画一サイズのブロックではなく、サイズ・形状に多様性を持たせた局所領域から、多重周波数で特徴抽出を行う方式とした。サイズ・形状の多様性は、ランダムな局所領域パターンを発生させて、その中から学習によって固有性と頑健性を両立するパターンを自動選定する方法をとった。この方式により、標準化で行われた技術コンペで他組織を大幅に上回る第1位の成績を納めることができた。

Version 3は、標準化活動の中でVersion 2の技術をさらに性能改善を図っていき、最終的にMPEG国際標準技術として採用された方式である。学習の偏りや過学習の懸念を考慮し、学習に基づく局所領域パターンの選定をやめ、サイズ・形状に多様性を維持しつつ、幾何学的な規則に基づいたパターンを用いた。さらに、局所領域パターンは、画像の周辺よりも中央領域を重視し、画像の中央のほうがより密にサンプリングされるようにした。これにより、標準化での評価実験でさらに認識精度を10%改善することができた（評価実験結果については第4章に示す）。

Version 3のMPEG国際標準に採用されたVideo Signature特徴量の抽出処理を図3に示す。フレーム画像から、標準規格で規定した計380パターンの多様な位置・サイズ・形状の局所領域ペア（図4に一部の例を示す）の各々について、その輝度差分を算出し、差分値を{+1, 0, -1}の三値量子化（一方の輝度が顕著に大きい、もしくはほぼ同値）したものを特徴量とする。三値量子化では、量子化値の出現頻度が均等になるように閾値を適応的に決定することで、輝度変化に対する頑健性を確保すると同時に、特徴量の固有性（識別能力）を最大化している。380次元の三値量子化値は、5つずつまとめて1バイトで符号化することで、わずか76バイトのコンパクトなバイト列に圧縮する。

またVideo Signatureと併用して、フレーム画像の絵柄の複雑度を表す数値であり、Video Signatureがどの程度識別に有効な情報を保有しているかを示す信頼度という指標を初めて導入した。信頼度の算出には、Video

Signatureの抽出で用いた局所領域ペア間の輝度差分の絶対値から、その統計量として中央値を求め1バイトで記述する。絵柄が複雑な場合は信頼度が高く、絵柄が平坦で特徴のない場合は低くなる。信頼度は、従来から平坦で特徴のない絵柄で特徴量の照合が不安定になるという問題に対処するために導入しており、Video Signatureの照合で併用することで、平坦な映像により発生する誤検出を大幅に削減することに成功している。

### 3.2 照合技術の開発

Video Signatureの照合は、CBIR技術の照合を踏襲したシンプルな方法であり、基本的には2つの映像（動画）のVideo Signature特徴量系列間フレーム対フレーム間の距離計算を順次行い、距離値が連続的に小さい区間を同一区間と識別する。この処理を高速化するために、CBIR技術の研究で映像の高速照合方式として開発した間引きマッチング手法<sup>8)</sup>などを適用することもできる。

フレーム対フレーム間の距離計算には、Video Signatureの380次元の三値ベクトル間のL1距離を用いるが、実装上は、これにはルックアップテーブル（LUT）を用いた超高速演算ができる。符号化された76バイトデータのまま、1バイト単位でLUTを参照することで、わずか76回のLUT参照と76回の加算で算出できる。この超高速演算により、数千時間程度のDBからの瞬時の映像識別を可能にしている。

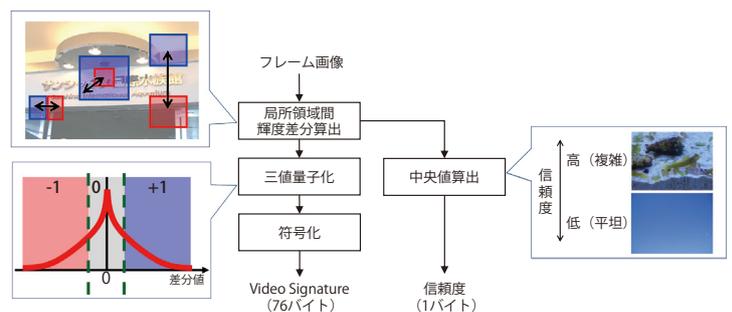


図3 Video Signature と信頼度の抽出処理

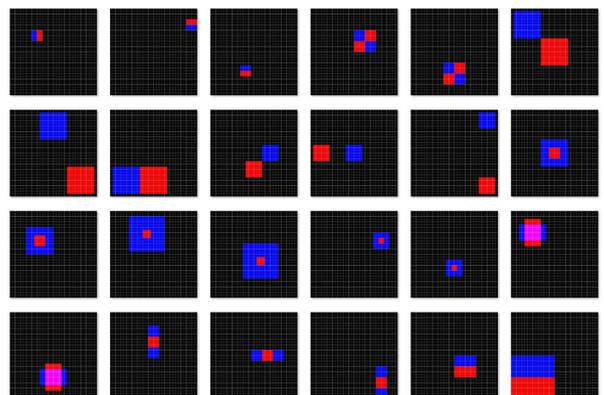


図4 標準規格で規定した380パターンの局所領域ペアの例

## 4. MPEG 国際標準化

Video Signature 技術の開発と並行して、2007年から本技術を世界共通で使える標準技術として、ISO/IEC 国際標準のMPEG-7規格（映像・画像の検索用メタデータに関する規格）の一部として規格化することに挑戦した。筆者らがMPEG標準化会合の場で、映像をID識別するための特徴量の標準化の重要性を訴えた結果、賛同が得られて、2007年にMPEG-7での標準規格化のプロジェクトが立ち上がった。筆者らは、このプロジェクト活動に世界各国から集まった20社以上の参加者の中の中核メンバとして、年に4回開催されるMPEG標準化会合での標準化作業を全面的に主導した。標準化作業は、参加者の議論と合意に基づき進めていき、(1)標準化目的・技術要件の策定、(2)技術評価方法の策定、(3)技術コンペ・技術改善の実施、(4)規格文書の作成、の順に進められた。以下では、4年ほどかけて行われたMPEG国際標準化の各作業について、活動内容を紹介する。

### 4.1 標準化目的・技術要件の策定

まず、標準化プロジェクトの立ち上がりにあたり、標準化の目的を議論した。映像の不正流通検知、コンテンツ使用実績調査、アーカイブ検索などを実現するために、映像の同一性を判断するための一意識別可能な「指紋」特徴量を標準化することを目的と規定した。そして標準化のスコープとしては、インターオペラビリティ（相互互換性）を確保する最低限の要素として、特徴量の抽出方法とその記述フォーマットを規格化する範囲と定めた。特徴量の照合方法については、各々が自由な方法・実装で実現しても良いとし、推奨方法は提示するものの標準必須外とすることにした。

次に、標準化する特徴量技術が持つべき技術要件として、以下を満たすことを規定した。

- 固有性（ユニーク性）：大量映像の中から、映像を唯一特定できること。
- 頑健性（ロバスト性）：各種編集・改変が加わっても同一性を判定できること。
- 高速性：インターネット規模の大量映像に対応できるように、特徴量の抽出・照合速度が高速であること。
- コンパクト性：特徴量のサイズが小さいこと。
- 部分照合機能：映像の一部区間が切り出されても、区間を特定できること。

### 4.2 技術評価方法の策定

MPEGでは策定した技術要件に合う技術を広く募集して、公平でオープンな技術コンペを通じて最適な技術を選定していく。そのために、技術の性能・優劣を評価する実験方法を策定する必要がある。本規格の評価実験方法では、世界最高レベルの特徴量技術を選定するために厳しい評価実験を設定した。編集・改変の加わった2～10秒程度の映像片から、大量映像の中に埋もれている元映像を識別できるかをテストとし、その識別率を測定することとした。その際に許容する誤検出率は5ppm=0.0005%という極小値として、ほぼ誤検出を許容しない設定とした。映像片に加える各種編集・改変については、世の中の実際の実施状況を網羅するように、①テロップ重畳、②符号化圧縮、③解像度変換、④カメラ盗撮、⑤アナログ録画、⑥フレームレート変換、⑦明度変換、⑧モノクロ変換、⑨IP変換、の9種類の編集・改変を定めた（図5）。それぞれについて、編集・改変の度合も低・中・高の3レベルを規定し、その中でも高レベルは一般にインターネットで見られるような編集・改変をはるかに超えるようなレベルとして、意図的に難しいテストとなるように設定した。

この評価実験を回すために、100時間以上の映像データが必要となった。当時、そのような公開データセットがなかったので、標準化活動に参加している各社が分担して映像収集することとなった。収集には、放送局に提供を依頼したり、インターネット上からフリー素材の映像を探索したり、自ら映像撮影したりなど、1年以上かかることとなった。最終的には、映画、ニュース番組、スポーツ、バラエティ、ホームビデオなど多種・多様な映像データを集めることができた。そして収集した映像から、映像片を切り出し、編集ソフトウェアを使って9種類の編集・改変処理を加えた。こうして苦労して評価実験用の映像データが用意でき、2008年に技術募集をかけることができた。

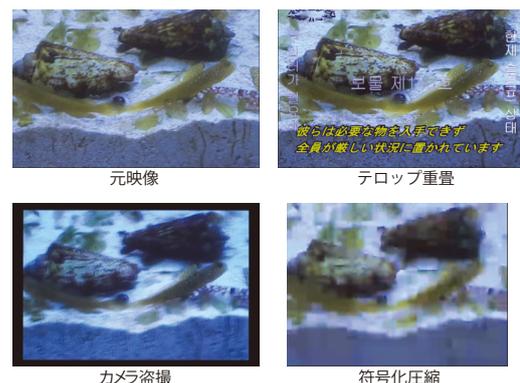


図5 評価実験での編集・改変の例

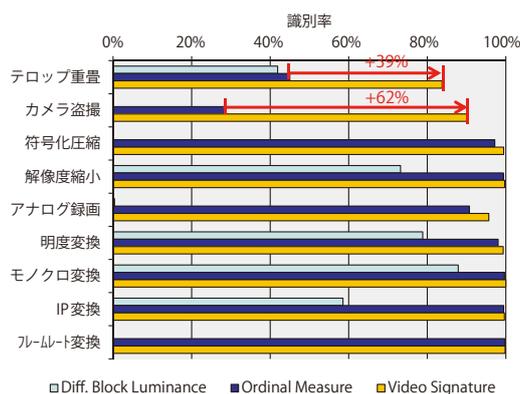


図6 評価実験結果

### 4.3 技術コンペ・技術改善の実施

技術募集に対して我々 NEC も 3.1 節で説明した Version 2 の特徴量技術を応募した。4.2 節で説明した技術評価方法で技術コンペを行った結果、NEC は 2 位の組織に平均識別率で 10% 以上の差をつけた圧倒的な 1 位の成績であった。この結果を受け、NEC の提案した特徴量技術は本規格のベース技術として採用された。それ以降は、標準化活動の中で他組織と共同で技術改善を図っていき、最終的には Version 3 の技術にまで昇華させ、これが標準技術となった。この改善プロセスにより、平均識別率をさらに 10% 向上させることができた。

最終的に標準技術となった Version 3 の Video Signature 技術を用いて、技術評価方法に基づく評価実験を行った結果（9 種類の編集・改変に対する識別率）を図 6 に示す。従来の CBIR 技術で用いられる Difference Block Luminance<sup>4)</sup> と Ordinal Measure<sup>5)</sup> の特徴量と比較した。Video Signature は、あらゆる改変行為に対して安定した性能を示しており、平均して 96% の識別率を達成した。従来の特徴量<sup>4), 5)</sup> と比較して、すべての改変に対して精度改善が認められるが、特にインターネット上の不正コピーで最もよく使われているテロップ重畳 (+39%) とカメラ盗撮 (+62%) でその改善は顕著である。

また、Video Signature の抽出・照合の処理時間を PC (CPU 3.4GHz) で計測した。抽出は約 1.4ms/フレーム、照合は PC1 台で 1 秒間に 1,000 時間超の DB から識別が可能であり、十分な高速性が実証された。

### 4.4 規格文書の作成

標準化する技術が選定されたことで、筆者らは ISO/IEC のプロジェクトエディタに任命され、規格文書の執筆も担当した。そして、2010 年 10 月に MPEG-7 Video Signature Tools (ISO/IEC 15938-3/Amd.4)<sup>9)</sup> として規格が発行された。

さらに筆者らは、標準化された Video Signature の抽出・照合方法の実装例が参照できる参照ソフトウェア、抽出された Video Signature データが規格に適合するかをテストする適合性試験、などの周辺規格も策定し、最終的にすべての Video Signature 関連の標準化作業を完了したのは 2011 年であった。プロジェクトの始動から 4 年ほどかかったが、標準化活動により世界最高レベルの技術を完成させ、映像を ID 識別する世界共通の標準技術として確立することができた。

## 5. システム開発・事業化

2010 年に、NEC は Video Signature 技術の標準化採用を受けて記者発表<sup>10)</sup> を行い、映像の不正流通検知<sup>11)</sup> などへの応用が期待され、多数の新聞や雑誌への掲載やテレビ放送がされて大きな反響を受けた。また、多数の顧客からの引き合いを受け、我々は早速システム開発・事業化に着手した。

Video Signature 技術を用いたシステムとして、(1) 不正流通検知、(2) コンテンツ使用実績調査、(3) 映像アーカイブ管理、などさまざまな顧客向けのシステムを開発し、顧客に提供していった。各システムを開発・事業化するにあたっては、それぞれ望まれている技術やシステムの要件が異なったため、すべて同じ標準化された Video Signature 技術を用いたものであっても、照合方法、データの保存方法、システム構築方法、ユーザインタフェースなどさまざまな面で異なる方法・実装をとる必要があった。以下、開発して事業化した各々のシステムについて説明する。

### 5.1 不正流通検知システム：映像識別ソフトウェア

#### Media-Serpla

Video Signature 技術の最も反響の大きかった応用方法は、大きな社会問題となっていた映像コンテンツの不正流通の検知であった。そこで、NEC は標準化された Video Signature 技術を用いた世界初の製品として、映像識別ソフトウェア Media-Serpla を 2012 年 4 月から販売開始した<sup>12)</sup>。このソフトウェアは、原本 DB の登録、検査対象映像からの Video Signature 抽出と照合、および照合結果確認 (図 7) の機能を提供し、PC サーバに導入することで簡単に映像識別システムを構築できる。図 8 に、本製品を動画共有サイトなどのサービス事業者を導入し、コンテンツの投稿時に不正コピーか否を識別することで、サイト上での不正流通を未然に防止するシス



図7 映像識別ソフトウェア Media-Serpla

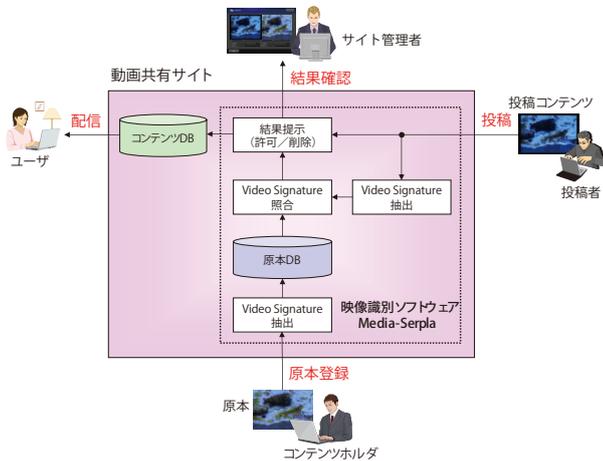


図8 動画共有サイトでの不正流通検知

テムの例を示す。また、放送局などのコンテンツホルダや警察などの第三者が、本製品とネット上の映像を巡回収集するクローラと組み合わせることで、不正流通の発見・摘発に活用することもできる。

本製品を開発するにあたり、標準化では規定されていない高速な照合方法を開発した。開発した高速照合は、フレームごとに、その後に続くフレームへの特徴量距離値をテーブルとしてメモリ上に蓄積し、フレーム対フレーム間の照合を順次行っていく際に、テーブルを参照して距離値が閾値を下回らない区間を算出して、照合するフレームをスキップしていくというものである。この方法は、多少のメモリ量増加はあるものの、照合速度を数倍から最大で10倍ほど高速化する効果があった。また、照合の後処理として、識別した同一区間の統合処理を導入した。これは、場合によって、1つの長い同一区間が複数の短い同一区間に分断されてしまうことがあり、これを防ぐ処理である。位相がほぼ同じであれば、複数の識別された同一区間を統合する。この処理はシンプルでありながら、実用上はとても効果的であった。

## 5.2 コンテンツ使用実績調査システム

Video Signature技術の応用としてもう1つ要望が大きかったのが、映像コンテンツの使用実績調査である。特にCMの放映実績の自動調査に対するニーズは大き

く、NECはCMコンテンツの採集・放映実績調査するシステムを開発し、2011年11月に事業化した<sup>13)</sup>。本システムは、Video Signatureを用いて1日平均約4,000件放映されるCMを正確かつ効率的に採集・DB登録し、CMの放映実績を集計する。コンテンツの調査・分析の業務効率化に貢献している。

本システムは、放送チャンネルごとに映像を同録し、CM区間を無音区間に基づいて切り出してDBに保存するPC（放送チャンネル分）と、切り出したCMとすでにDBに保存されているCMとをVideo Signatureを用いて照合するPCから構成される。本システムでは、CMに特化した照合を実装した。切り出されたCMは、15秒（もしくは30秒）のほぼ決まった長さであるため、それらの間の照合は、切り出しの開始点・終了点のズレだけを考慮すればよい。この特性を利用して、最大で数秒のズレまでを考慮し、探索範囲を狭めることで、3カ月分のCM（約12,000件）と照合するのに必要な時間をわずか1秒間にすることができた。

またCMの照合では、テロップがわずかに異なる同一映像のCMを異なるCMとして識別して集計するという要件があった。本来、Video Signatureはテロップ重畳に対しては頑健に設計されているので、通常のVideo Signatureの照合ではテロップ違いのCMを見分けるのは困難である。そこで、Video Signatureの380次元の三値ベクトルの各要素が画像内の特定の局所領域から算出されることを利用し、特徴ベクトル間で差のある要素について、対応する局所領域をハイライトしていくことで、テロップのわずかな差を浮き出させることができた。これにより、テロップ違いのCMの見分けも可能になった。本来のVideo Signatureの使い方とは異なる新たな使い方の発見であった。ここでの例のように、元映像を識別するとともに、一部の領域に加わった編集・変更の差も見分けたいという要望は多く、Video Signatureが多数の局所領域から算出されるという特性を利用することで、このような要望にも幅広く応えられることを学んだ。

## 5.3 放送局向けアーカイブ管理システム

またNECは、放送局などの大量の映像コンテンツを保有する事業者向けに、Video Signatureを用いた映像の効率的な管理・検索を実現する映像アーカイブ管理システムを開発し、放送局などの事業者に試験導入している<sup>14)</sup>。このシステムでは放送局が蓄積する素材映像・編集映像・完パケ映像など編集段階ごとの映像の相互



図9 映像アーカイブ管理システムのGUI

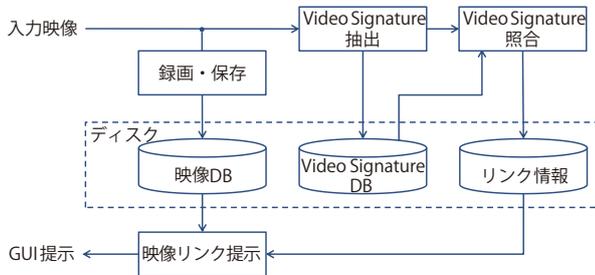


図10 映像アーカイブ管理システムの構成

引用関係（リンク）を Video Signature で自動識別する。そのリンク関係をタイムライン形式で GUI 提示することで（図9）、アーカイブ内の編集履歴、素材の引用頻度・実績を把握でき、素材探索や二次利用のための権利確認などの業務作業を効率化する。Video Signature の利点を活かし、ショット単位の編集や、テロップ重畳や符号化による映像信号の変化があっても、正確にリンクを生成できる。

映像アーカイブ管理システムは、蓄積された映像データベース内の映像リンク（同一区間）を随時検出していく、そのリンク情報を用いて映像間の構造をユーザに GUI 提示する。図10に、PCで実装した本システムの基本構成を示す。まず、システムに新たな映像が入力されると、それを蓄積（保存・録画）すると同時に、並行して Video Signature を抽出し、ディスクに保存する。Video Signature の抽出が完了すると、次に Video Signature の照合を実施し、新規の入力映像とすでに蓄積されている DB 映像との間での同一区間を検出する。検出された同一区間はクラスタリング（グルーピング）され、共通する同一区間を共有する映像リンクのグループとしてリンク情報を保存する。こうして求められた映像リンク情報を用いて、図9に示すような映像リンク構造の GUI 提示を行う。

このようなシステムで、蓄積された映像データが数万時間を超えてくると、5.1節で説明した高速照合を用いても、照合に時間がかかりすぎてしまう。そこで、超大規模映像DB向けの超高速検索として、インデックスを用いた絞り込み検索方式を開発した。この方式で

は、まずDB映像の各フレームの Video Signature をハッシュ値に変換する。ハッシュ値は、Video Signature の380次元の三値ベクトルから、あらかじめ定められた要素の値を合成して算出する。そして、フレーム系列を数秒単位の短い区間に分割し、各区間に含まれるハッシュ値で、逆インデックス（ハッシュ値を参照することで、その値を含む区間を絞り込むことのできるインデックス）を生成する。クエリとなる入力映像に対しては、同様に各フレームからハッシュ値を算出し、逆インデックスを用いてそのハッシュ値を参照することで、DB映像の候補となる区間を絞り込むことができる。こうして絞り込まれた区間に対してのみ、Video Signature による通常の照合を行う。このインデックスを用いた絞り込み方式により、2秒の短い映像区間を、36,500時間の大規模映像DBの中から、わずか1秒で検索することが実現できた。これにより、ますます拡大していく映像データ規模にも対処することが可能になったといえる。

## 6. おわりに

本稿では、筆者らが開発した映像のID識別のための映像「指紋」技術 Video Signature を世の中に出していくために実践した研究開発、国際標準化、そして事業化の取り組みについて解説した。2000年代中盤に、インターネット上での映像の複製・拡散が急速に進んだ背景の中で、筆者らは映像の「指紋」として使える世界最高レベルの特徴量技術を開発した。世界中に拡散した映像をID識別できる世界的な仕組みが必要と考え、MPEG国際標準化することに挑戦し、成功した。そして、標準化された Video Signature 技術を用いて、不正流通検知、コンテンツ使用実績調査、映像アーカイブ管理などのさまざまな応用システムを開発し、事業化を進めてきた。国際標準化や事業化を通じて、Video Signature 技術は開発当初よりも、より高性能で実用的な技術に進化していった。着実に実績は出しているものの、まだ世界レベルでの普及には至っておらず、今後も Video Signature 技術の普及に努めていく所存である。

### 参考文献

- 1) Paschalakis, S., Iwamoto, K., Brasnett, P., Sprljan, N., Oami, R., Nomura, T., Yamada, A. and Bober, M.: The MPEG-7 Video Signature Tools for Content Identification, IEEE Trans. on Circuits and Systems for Video Technology, Vol.22, Issue7, pp.1050-1063 (2012).
- 2) 岩元浩太, 大網亮磨, 佐藤貴美, 野村俊之: 大規模な映像識別のための改変にロバストな Video Signature, 精密工学会誌, Vol.79, No.11, pp.1089-1095 (2013).

- 3) Naphade, M. R., Yeung, M. M. and Yeo, B.-L. : A Novel Scheme for Fast and Efficient Video Sequence Matching Using Compact Signatures, Proc. of SPIE, Storage and Retrieval for Media Databases, Vol.3972, pp.564-572 (2000).
- 4) Oostveen, J., Kalker, T. and Haitsma, J. : Feature Extraction and a Database Strategy for Video Fingerprinting, Proc. of 5th Int'l Conf. on Recent Advances in Visual Information Systems, pp.117-128 (2002).
- 5) Hua, X.-S., Chen, X. and Zhang, H.-J. : Robust Video Signature Based on Ordinal Measure, Proc. of ICIP2004 (2004).
- 6) Kasutani, E. and Yamada, A. : The MPEG-7 Color Layout Descriptor - A Compact Image Feature Description for High-speed Image/Video Segment Retrieval, Proc. of ICIP2001 (2001).
- 7) Iwamoto, K., Kasutani, E. and Yamada, A. : Image Signature Robust to Caption Superimposition for Video Sequence Identification, Proc. of ICIP2006 (2006).
- 8) Kasutani, E., Oami, R., Yamada, A., Sato, T. and Hirata, K. : Video Material Archive System for Efficient Video Editing based on Media Identification, Proc. of ICME2004, Vol.1, pp.727-730 (2004).
- 9) ISO/IEC 15938-3:2002/AMD 4:2010, Information Technology - Multimedia Content Description Interface - Part 3 : Visual, Amendment 4 : Video Signature Tools.
- 10) NEC プレスリリース, 瞬時に違法コピー動画を発見できる映像識別技術を開発 ~国際標準規格 (MPEG-7) に採用~, <http://www.nec.co.jp/press/ja/1005/0701.html> (2010/5/7) (2017年2月1日現在)
- 11) Iwamoto, K. : Stop Internet Piracy! Video Signatures Offer a Solution, ISO Focus+ Nov.2010, pp.38-39 (2010).
- 12) NEC プレスリリース, NEC, 大量の映像データの中から同一映像を瞬時に見つけ出す「Media-Serpla 映像識別ソフトウェア」を販売開始~国際標準規格 MPEG-7 採用技術を製品化~, <http://www.nec.co.jp/press/ja/1204/1001.html> (2012/4/10) (2017年2月1日現在)
- 13) NEC プレスリリース, NEC, CM コンサルティング会社の「東京企画」へ CM 素材の同録システムを納入, <http://www.nec.co.jp/press/ja/1111/2501.html> (2011/11/25) (2017年2月1日現在)
- 14) Iwamoto, K., Sato, T., Oami, R. and Nomura, T. : Visual Duplicate based Topic Linking Using a Robust Video Signature, Proc. of ICCE2013 (2013).

岩元 浩太 (非会員) k-iwamoto@ay.jp.nec.com

2003年早稲田大学大学院理工学研究科電子・情報通信学専攻(修士課程)修了。2003年日本電気(株)に入社。現在、同社データサイエンス研究所主任研究員。画像・映像認識、画像・映像検索技術の研究開発および事業化、またISO/IEC 15938(MPEG-7)標準化のプロジェクトエディタとして国際標準化活動に従事。2012年文部科学大臣表彰若手科学者賞、情報処理学会情報処理規格調査会標準化貢献賞を受賞。2013年先端技術大賞経済産業大臣賞を受賞。2014年精密工学会高城賞を受賞。

鍋藤 悠 (非会員) y-nabeto@uf.jp.nec.com

2009年鹿児島大学大学院理工学研究科情報工学専攻(修士課程)修了。2009年(株)NEC情報システムズに入社。画像・映像認識、画像・映像検索技術のエンジン開発・高速化に従事。

採録決定：2017年2月1日

編集担当：峯松信明(東京大学)