

経済産業大臣賞

高精度映像識別技術 Video Signature の  
研究開発

日本電気株式会社 情報・メディアプロセッシング研究所

岩元 浩太 大網 亮磨 佐藤 貴美 野村 俊之

筆者らは、ネット社会の普及とともに急速な広がりを見せる、映像・画像コンテンツ産業の健全な成長拡大に甚大な影響を及ぼす映像の不正流通を阻止するため、世界最高性能の高精度映像識別技術 Video Signature を開発した。この技術は、ネット上に埋もれている、種々の手法で改変された不正コピー映像を人間と同等以上の認識精度で自動的に瞬時に発見し、実用レベルで不正流通防止に活用できる世界初の画期的な技術である。世界レベルで横行する不正流通に対抗するにはグローバルな仕組み構築が重要であるとの認識のもと、本技術を国際標準規格化することにも挑戦し、成功した。標準化された本映像識別技術によれば、コンテンツ産業の発展に不可欠な権利と対価の健全性を先進の IT 技術で保護する基盤が構築でき、今後の日本の同産業の輸出や関連産業の拡大などにも大きな貢献が期待される。

## 1. 緒 言

日本のアニメや漫画などは海外でも大人気で、日本発のポップカルチャーとして世界的に認知されている。近年の製造業の停滞もあり、日本では新たな産業としてアニメなどのコンテンツ産業が着目され、輸出産業としての期待も高まっている。21 世紀に入り国内では「コンテンツ立国」との名のもと、同産業を拡大・保護する取り組みも盛んとなっている。その市場規模は国内 12 兆円 [1]、世界では 140 兆円と巨大かつ成長を続けている [2] (図 1)。しかし、アニメ等の輸出比率は僅か 0.15% [3] と米国、韓国など諸外国に比べ圧倒的に低く、海外進出は遅れている。そのため、経済産業省は「クール・ジャパン戦略」の一環として国家あげての取り組みを行なってきた。

輸出拡大を阻害する最大の問題の一つが、コンテンツの不正流通である。インターネット上の動画共有サイトなどの普及により、ネット上は映像・画像コンテンツの不正コピーによる無法地帯となっており、人気アニメなどは放送後すぐに字幕付きでアップロードされてしまうことさえある。このような不正流通により、関連する権利者などは適正対価を失うばかりでなく、放送や番組製作等の正規事業の拡大も阻害され社会問題化している。このような損失額は約 2400 億円 [3] と推計される。従って、日本が世界的なコンテンツ産業を育成するには、不正流通阻止は喫緊の課題であり、総務省でも国家レベルで対策を検討してきている [4]。

こうした不正流通の横行に対し、現在放送局等は多大なコストを払って人手による目視検査で被害を発見し管理者に削除依頼する、というのが実情である。しかし、ネット上の膨大なコンテンツは人手で監視しきれものではなく、通報された一部だけが削除されるのみである。一般に、不正流通コンテンツには発見されにくい工夫が施されているため発見が困難なだけでなく、一旦流通したら最後、複製され拡散してしまうことが避けられない。

我々は、以上の課題解決を目指し、インターネット規模の大規模なコンテンツデータから、機械がそのコンテンツが何であるかを瞬時に自動識別(発見)する、映像識別の自動化技術を開発した。Video Signature と呼ぶこの技術は、従来技術に比べ不正コピーの識別精度が圧倒的に高く、精度と速度の両面において実用的なレベルで不正流通を検知できる世界初の技術である。さらに我々は技術開発に留まらず、横行する不正流通に対抗する世界的な仕組み作りに取り組む重要性を認識し、世界標準として利用できるように国際標準規格 MPEG-7 Video Signature Tools として標準化することにも挑戦し、成功した。

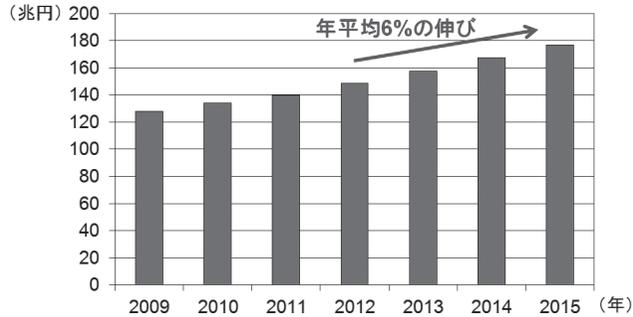


図1 コンテンツ産業の世界市場  
(PricewaterhouseCoopers「Global entertainment and media outlook 2011-2015」より)

## 2. 研究開発の背景

ネット上のコンテンツの不正流通の検知(発見)を自動化するには、コンテンツの原本(オリジナル)を登録し、流通するコンテンツが原本と同一かどうかを識別すればよい。ここで重要なのが、不正コピーの殆どには、図2のようなテロップ重畳、カメラ盗撮、符号化圧縮などの編集や、悪意のある改変が原本に加えられ、信号のレベルでは内容は大きく変化していることである。そのため、これらの行為に対し、頑健に同一性の判断ができる映像識別技術が求められる。一方で、これら行為があっても発見できるように閾値を下げると、誤検出が拡大する。不正流通検知では、以上のように本来相反する高い頑健性と低い誤検出率を両立させる識別方法が求められる。

従来の映像識別技術には情報埋め込み型と特徴抽出型がある(表1)。情報埋め込み型は、コンテンツデータに識別用情報を埋め込む手法であり、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、映像信号に見えない形で情報を埋め込む電子透かし等がある。これらは埋め込まれた情報が解読できればコンテンツの出所を追跡できる反面、流通前に予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、悪意のある改変に対しては極めて脆弱であり、容易に埋め込んだ情報が消去されるなど、同技術は不正流通防止には全く無力である。

一方特徴抽出型は、コンテンツそのものが持つ特徴で映像を識別する。その一つであるコンテンツベース検索技術では絵柄を表現する特徴量を抽出して照合し、映像の同一性を判断する。具体的には抽出された特徴量は原本コンテンツの特徴量DBと照合され、一致検索がなされる。特徴量としては、色特徴を用いる Color Histogram [5]、空間特徴を用いる Difference Block Luminance [6] や Ordinal Measure [7] などが提案されている。これら方法は事前の仕込みが不要なこと、既に流通しているコンテンツにも適用できる等の利点がある。しかし、特徴量 [5]-[7] は類似検索を目的に開発されたもののため、不正コピーに多いテロップ重畳やカメラ盗撮などの改変に対して脆弱である。よって、本技術も不正流通防止には不十分である。

そこで我々は、特徴抽出型の全く新しい映像識別技術として、各種改変・編集行為に極めて頑健であり、新たな発想による Video Signature と呼ぶ特徴量を考案し、実用レベルで不正流通に対抗できる世界初の映像識別技術の開発に取り組んだ。

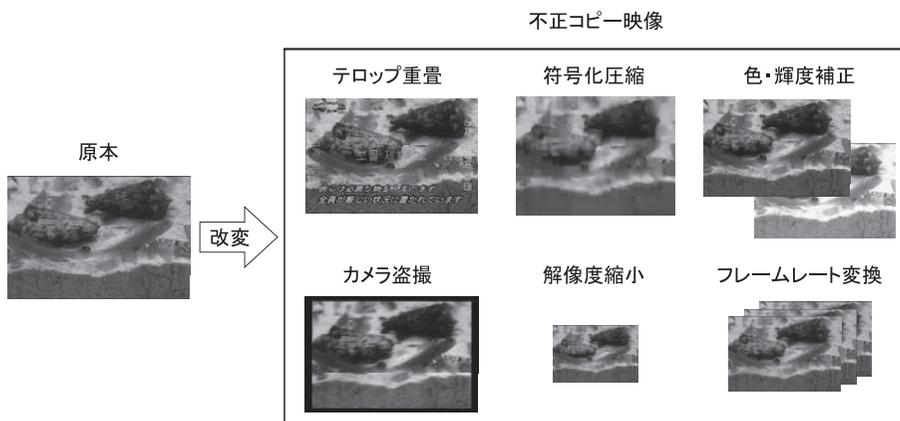


図2 不正コピー映像に施される変更・編集の例

表1 映像識別技術の優劣比較

	情報埋め込み型		特徴抽出型	
	ファイル埋め込み	電子透かし	コンテンツベース検索技術	Video Signature
事前準備	必須	必須	不要	不要
画質劣化	劣化なし	劣化あり	劣化なし	劣化なし
変更頑健性	対応不可	脆弱	脆弱	頑健

### 3. Video Signature 技術

Video Signature は、人間が画像パターンを判断する視覚認知を模した方法に基づき、絵柄を表現する固有で頑健な特性を特徴量化したものである [8]-[11]。不正流通検知を目的に、以下の技術要件をクリアできるように設計された。

[要件 1] 種々の変更・編集に頑健なこと。また同時に低い誤検出率を達成できること。(3.1 節)

[要件 2] 数秒程度に短くカット編集された一部区間でも識別できること。(3.2 節)

[要件 3] 大規模データに対応できるように特徴量の抽出・照合が超高速なこと。(3.1, 3.2 節)

#### 3.1 Video Signature 抽出

Video Signature は映像の各フレームから抽出する。[要件 1] の高い頑健性を得るために、人間の視覚認知を模した多重周波数グリッド解析というアイデアに基づいて、フレーム画像中の多様な位置・大きさ・形状の部分領域間の輝度差を特徴ととらえ、それを 3 値量子化して 76 バイトにコンパクト化して特徴量とする。また低い誤検出率を実現するために、Video Signature と併用して、絵柄の複雑度に基づく信頼度という指標を初めて導入し、誤検

出を大幅に抑制することを可能にした。

映像の各フレームから Video Signature と信頼度を抽出する処理フローを図3に示す。Video Signature の抽出には、まず予め設定された 380 種類の部分領域のペア (図4) の各々から平均輝度値を算出し、その輝度間の差分  $d_i$  ( $i=1, \dots, 380$ ) を算出する。図4に示すように、部分領域の位置・大きさ・形状に多様性を持たせて多重な周波数解析を行うことで、従来の特徴量 [6][7] で一般的であった画一的なブロック分割と比較して、様々な改変行為に対する頑健性を飛躍的に向上させている。また識別にとってより重要な情報が画像中央にくる確率が高いことを考慮し、部分領域は中央がより密にサンプリングされている。次に、輝度差分  $d_i$  をそれぞれ 3 値  $\{-1, 0, +1\}$  に量子化 (一方の輝度が顕著に大きい、もしくはほぼ同値) する。

$$x_i = \begin{cases} +1 & (\text{if } d_i > th) \\ 0 & (\text{if } |d_i| \leq th) \\ -1 & (\text{if } d_i < -th) \end{cases} \quad (1)$$

ここで閾値  $th$  は、フレームごとに輝度差分の分布を考慮して、量子化値の出現頻度が均等 (1/3 ずつ) になるように適応的に決定される。これにより、輝度変化に対する頑健性を確保すると同時に、特徴量の識別能力を最大化している。最後に、380 要素の 3 値量子化値  $x_i$  ( $i=1, \dots, 380$ ) を 5 つずつまとめて 1 バイトに符号化することで、わずか 76 バイトのコンパクトなバイトデータ  $b_j$  ( $j=1, \dots, 76$ ) に圧縮する。

$$b_j = 81 \times (x_{5j-4} + 1) + 27 \times (x_{5j-3} + 1) + 9 \times (x_{5j-2} + 1) + 3 \times (x_{5j-1} + 1) + (x_{5j} + 1) \quad (2)$$

この符号化により、3 値要素を各 2 ビットずつ独立して符号化するよりも、20% 符号量を削減できる。

信頼度はフレーム画像の絵柄の複雑さを表す数値で、抽出した 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。信頼度は算出には、Video Signature の抽出で用いた部分領域間の輝度差分の絶対値  $|d_i|$  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 Video Signature の照合で併用し、平坦な映像による誤検出を除外するのに用いられる。

以上のように、Video Signature と信頼度の抽出処理は、部分領域間の輝度差分に基づくシンプルな数値演算によるため、リアルタイム処理のさらに数十倍の処理速度を実現できる ([要件 3])。

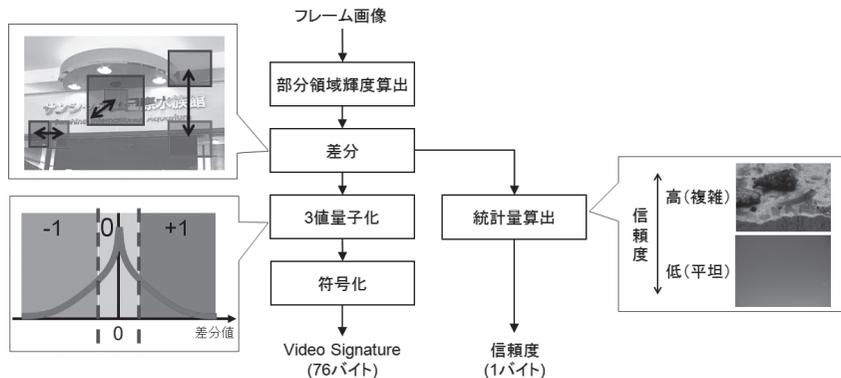


図3 Video Signature と信頼度の抽出処理

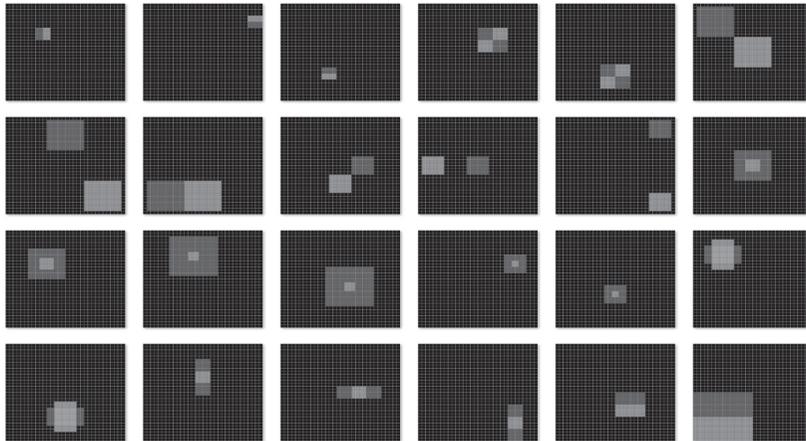


図4 部分領域ペアの例

### 3.2 Video Signature 照合

照合では、2つの Video Signature のフレーム系列間で、1対1での Video Signature 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、同一の映像区間として識別する。距離計算には L1 距離を用いるが、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。式 (2) で符号化された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 ([要件 3])。

Video Signature 系列間の照合は、具体的には以下の方法で高速・効率的な識別を行う [12](図 5)。まず、長さ  $L$  (識別したい映像区間の最長長) のウィンドウをそれぞれの系列上にスライドしていき、ウィンドウ内の対応するフレーム間で、Video Signature の距離計算を順次行う。ウィンドウの始端から終端までを距離値が既定の閾値を下回る場合は、次にウィンドウの両

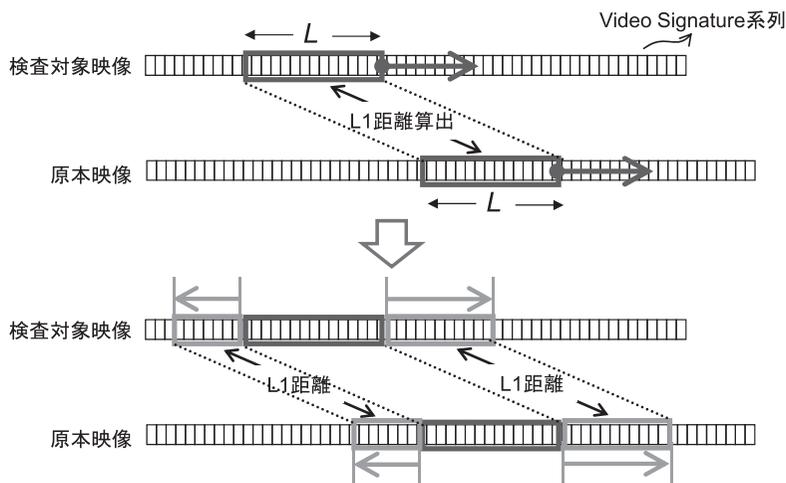


図5 Video Signature 系列の照合

端を距離値が閾値を下回るまで伸ばしていき、同一映像区間の始端と終端を確定する。こうして検出された同一映像区間に対して、双方の信頼度の平均値を算出し、これが既定の閾値より小さい場合は、平坦なシーンどうしによる誤検出と判断し、この結果を除外する。この一連の処理を、ウィンドウをずらしながら実行していく。この照合方式により、わずか2秒程度に短くカット編集された映像区間でも、大規模データから漏れなく正確にその同一区間を識別できる([要件 2])。

### 3.3 性能評価

Video Signature が、各種改変行為にどれだけ頑健に映像を識別できるのかを、後述する国際標準化団体が定めたテスト方式で評価した。このテストでは、数百時間の原本映像から切り出された2秒の短い映像区間に各種改変処理を加えて改変コピー映像を作成し、これから元の原本映像の区間を正確に識別できるかを評価する。テストする改変行為は不正コピーで多く行われる(1)テロップ重畳、(2)カメラ盗撮、(3)符号化圧縮、(4)解像度縮小、(5)アナログ録画、(6)明度変換、(7)モノクロ変換、(8)IP変換、(9)フレームレート変換の9つである。5ppm(100万分の5)という極めて低い誤検出率において、各種改変コピーから元の映像区間を識別できる識別率を測定する。従来のコンテンツベース検索技術で用いられる Difference Block Luminance[6]と Ordinal Measure[7]の特徴量と比較した。

図6に評価結果を示す。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[6][7]と比較して、全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。図7に識別結果の例を示す。従来技術では不可能であった短くカット編集された改変コピーでも識別が可能になり、[要件 1]と[要件 2]を満たすことが確認された。

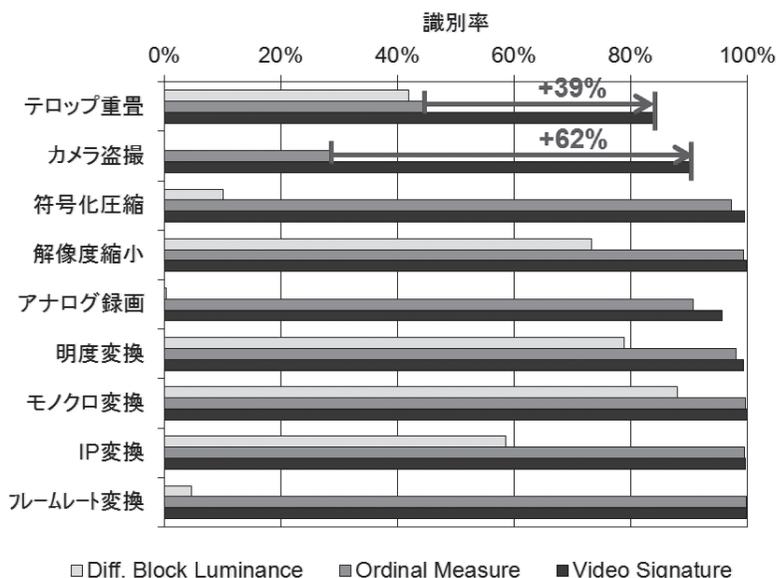


図6 性能評価結果



図7 識別結果例

次に、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 1000 時間超の DB から識別が可能であり、[要件 3] を満たすことが確認された。この高速処理により、例えばストリーミング動画のリアルタイムの映像識別が可能である。照合速度は、Video Signature がコンパクトなため、3.2 節で説明した全数探索でも十分高速であるが、インデックスを導入してフィルタリングを行うことで、さらに 2～3 桁以上の高速化も可能である。

さらに特徴データは 1 フレームあたり 76+1=77 バイトと極めて小さく、例えば 100 時間の映像であってもわずか 792MB 程度で映像自体のデータ量より 2～3 桁以上小さいため、大規模な原本 DB でもストレージ容量を圧迫しない。以上より、Video Signature は精度・速度・ストレージ容量の全てにおいて、ネット上の不正流通検知を実用レベルで実現できる技術といえる。

#### 4. 国際標準規格 MPEG-7 Video Signature Tools への採用

世界中のサイトやネットワークで横行しているコンテンツの不正流通を阻止するためには、技術開発だけでは不十分で、世界中の映像 DB を共通の仕組みで検索できるオープンな仕組み作りが必要である。そこで我々は、本技術を世界共通の標準技術として、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。我々が映像識別のための世界共通方式の重要性を訴えた結果、賛同が得られて、2007 年に MPEG-7 での映像識別用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核メンバーとして骨子策定を手掛け、技術要件定義、技術評価方法の策定など、標準規格化作業を全面的に主導した。

そして2008年に行われた技術募集に対して、開発した Video Signature を提案した。技術選定のための公平な第三者検証によるオープンな性能比較コンテストで、NECは、2位の組織に20%以上の差をつけた圧倒的な1位であった(図8)。この結果を受け、NECの Video Signature 技術は本規格のベース技術として採用された。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) [13][14]として規格が発行された。NECは標準化採用を受けて記者発表[15]を行った。不正流通検知への応用[16][17]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。

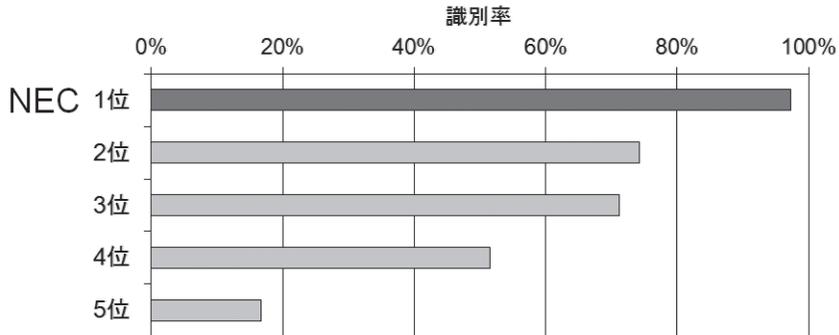


図8 MPEGでの評価テスト結果

## 5. 商用化実績

### 5.1 映像識別ソフトウェア Media-Serpla

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[18]。この製品は、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9)の機能を提供し、



図9 映像識別ソフトウェア Media-Serpla

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。不正流通検知のほか、コンテンツ利用実績の確認やアーカイブ検索など、様々な応用に活用できる汎用製品である。

図 10 は、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テムの例を示す。また、放送局などのコンテンツホルダー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ー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。

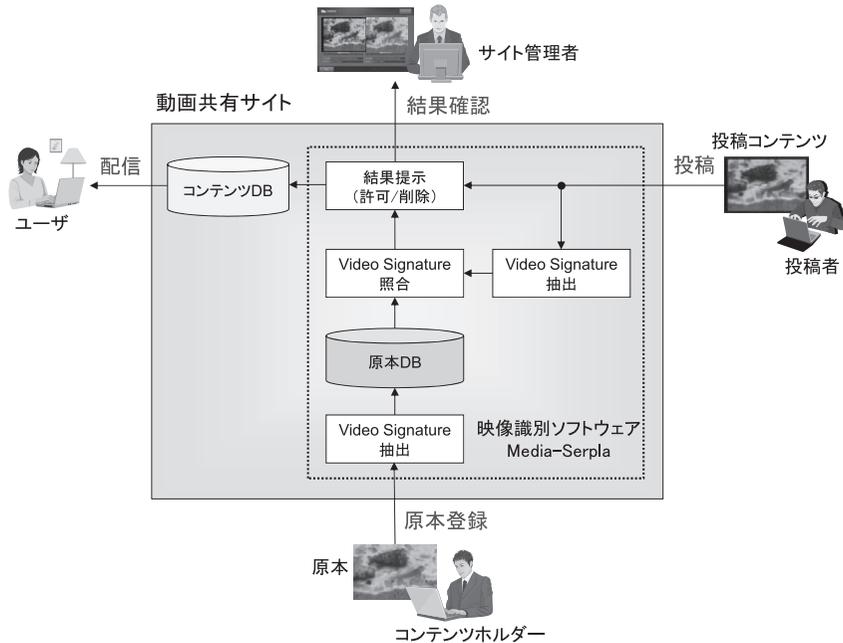


図10 動画共有サイトへの導入システム例

## 5.2 その他の製品化

Video Signature 技術の応用は不正流通検知だけに限らず、我々は映像識別を用いるその他の幅広い用途の実用化も行ってきた。まず NEC は、CM コンテンツの採集・調査するシステムを開発し、2011 年 11 月に CM コンサルティング会社に納入・稼働開始した [19]。本システムは、Video Signature を用いて一日平均約 4000 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。

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システムを開発し、放送局などの事業者にて試験導入している [20][21]。このシステムでは素材映像・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。そのリンク関係をマップやタイムライン形式で視覚化することで(図 11)、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 の利点を生かし、ショット単位の編集や、テロップ挿入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。

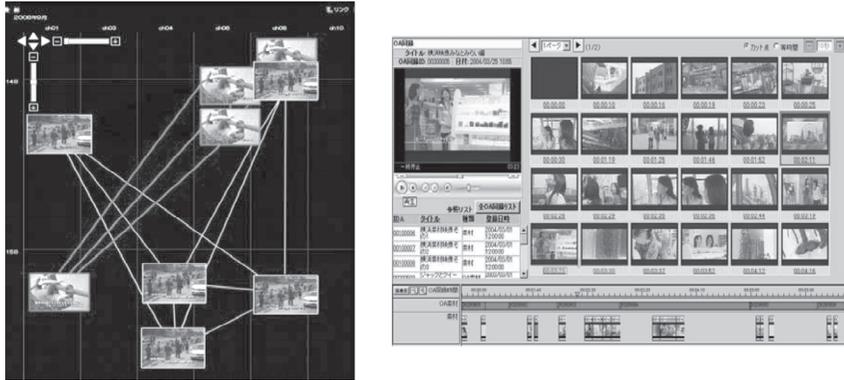


図11 映像アーカイブシステム

## 6. 結 言

日本が「クール・ジャパン戦略」のもと、アニメや漫画などの人気コンテンツを輸出産業として拡大する上で、その不正流通防止は極めて重要な課題である。筆者らが開発した Video Signature 技術は、従来では識別が不可能であった改変コピーでもネット上から瞬時に発見することを可能とし、実用レベルの精度と速度で不正流通防止に適用できる世界初のブレークスルーである。また筆者らは、単に技術開発に留まることなく、それを ISO/IEC 国際標準規格として世界共通の標準技術化することに取り組み、世界中で問題となっているコンテンツの不正流通行為に対抗する仕組み作りにも貢献した。不正流通を阻止できる実用的な技術と仕組み双方が整った結果、それを実用化した製品やシステムの開発にも成功し、現在ではそれら製品の販売やシステム運用を開始している。以上の活動は、コンテンツ産業の健全な収益構造を守り、日本の同産業の成長拡大に大きく貢献することが期待される。我々は、今後もより安全かつ効率的で快適な IT 社会の実現に向けて、映像識別を含む画像認識の技術開発で引き続き世界の最先端を突き進んでいく所存である。

## 参考文献

- [1] デジタルコンテンツ協会, “デジタルコンテンツ白書 2010”, 2010.
- [2] PricewaterhouseCoopers, “Global entertainment and media outlook 2011-2015 –Industry overview-”, 2011/6.
- [3] 経済産業省, “コンテンツ産業の現状と今後の発展の方向性”, [http://www.meti.go.jp/policy/mono\\_info\\_service/contents/downloadfiles/121226-1.pdf](http://www.meti.go.jp/policy/mono_info_service/contents/downloadfiles/121226-1.pdf), 2012/12.
- [4] 総務省, “インターネット上での放送コンテンツ等の不正流通防止に向けた取組について”, [http://www.soumu.go.jp/main\\_content/000154516.pdf](http://www.soumu.go.jp/main_content/000154516.pdf).
- [5] M. R. Naphade, M. M. Yeung, and B.-L. Yeo, “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.
- [6] J. Oostveen, T. Kalker, and J. Haitsma, “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, 2000.
- [7] X.-S. Hua, X. Chen, and H.-J. Zhang, “Robust video signature based on ordinal measure”, Proc. of ICIP2004, 2004.
- [8] K. Iwamoto, R. Oami, and T. Nomura, “MPEG-7 Video Signature for robust video identification”, Forum on Information Technology (FIT2011), 2011/9.
- [9] K. Iwamoto, T. Sato, R. Oami, and T. Nomura, “Visual Duplicate based Topic Linking using a Robust Video Signature”, Proc. of ICCE, 2013/1.
- [10] 岩元, 佐藤, 大綱, 野村, “Video Signature を用いた大規模映像データベースの構造化”, 映像情報インダストリアル, 第45巻, 第4号, p.57-63, 2013/4.
- [11] 岩元, 佐藤, 大綱, 野村, “Video Signature を用いた大規模映像データベースの構造化”, ビジョン技術の実利用ワークショップ ViEW2012, 2012/12.
- [12] E. Kasutani, R. Oami, A. Yamada, T. Sato, and K. Hirata, "Video material archive system for efficient video editing based on media identification", Proc. of ICME2004, vol.1, pp.727-730, 2004/6.
- [13] ISO/IEC 15938-3:2002/AMD 4:2010, Information Technology - Multimedia content description interface - Part 3: Visual, Amendment 4: Video signature tools.
- [14] S. Paschalakis, K. Iwamoto, P. Brasnett, N. Sprljan, R. Oami, T. Nomura, A. Yamada, and M. Bober, “The MPEG-7 Video Signature Tools for Content Identification”, IEEE Trans. on Circuits and Systems for Video Technology, vol. 22, issue 7, pp.1050-1063, 2012/7.
- [15] 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.
- [16] 山田, “違法コピー動画を高速かつロバストに検知できる映像識別技術”, コピライト 2010年10月号, p.53-57, 2010/10.
- [17] K. Iwamoto, “Stop Internet Piracy! Video signatures offer a solution”, ISO Focus+ Nov. 2010, p.38-39, 2010/11.
- [18] 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.
- [19] 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.
- [20] 金子, 小澤, 野村, 岩元, “「ビデオシグネチャ」を活用した映像識別ソリューション”, NEC 技報, Vol. 64, No. 3, 2011/9.
- [21] 野村, 佐藤, 岩元, “MPEG-7 Video Signature を用いた映像アーカイブ管理システム”, 第10回情報科学技術フォーラム (FIT2011), 2011/9.

< 関連特許 > 29件

< 新聞・雑誌掲載 > 14件