

# Video Identification Solution Using a “Video Signature”

KANEKO Hiroshi, OZAWA Takato, NOMURA Toshiyuki, IWAMOTO Kota

## Abstract

NEC’s video identification technology enables instant video content identification by extracting a unique descriptor called a “video signature” from video content. This technology is approved as the MPEG-7 Video Signature Tool; an international standard of interoperable descriptors used for video identification. The video signature is an extremely robust tool for identifying videos with alternations and editing effects. It facilitates search of very short video scenes. It also has a compact design, making ultrafast searches possible via a compact system. In this paper, we propose video identification solutions for the mass media industries. They adopt video signatures as metadata descriptions to enable efficient video registration operations and the visualization of video content relationships.

## Keywords

video identification technology, video signature, metadata, content distribution, archives

## 1. Introduction

The data volume of the mass media industries has grown enormously due to the file-based / tapeless workflows and the transition to HDTV operations. On the other hand, the infrastructure of multi-platform content delivery has also been advancing via telecommunications and the Internet in accordance with the acceleration of transmitting speeds and increased volumes. In these conditions, we generally manage the video contents using metadata (keywords, thumbnail images, preview videos, etc.). However, the increase in the data volume and the circulation of video content provide a significant volume of metadata search results. This has made it difficult to identify video content or to search a specific scene manually by visual inspection. In order to resolve this issue, we have developed a technology that identifies video content automatically and efficiently by managing it as metadata. In this paper, we describe the video signatures technology and propose video identification solutions.

## 2. Video Identification Technology

### 2.1 Functions of the Video Identification Technology

NEC’s video identification technology analyzes each frame of the video content and extracts a unique descriptor called the

“video signature” to identify identical video scenes. The video signature enables high-precision and high-speed video identification without embedding ID information in the content. Due to its high performance the technology has been approved as part of the international standard ISO/IEC 15938-3/ Amd.4 - MPEG-7 Video Signature Tools <sup>1)</sup>, which standardizes an interoperable descriptor for video identification.

#### (1) Extraction of the video signature

Each video signature extracted from a video frame is composed of two components. These are the frame signature (76 bytes per frame), which is a descriptor representing the image structure, and its confidence score (1 byte per frame). **Fig. 1** illustrates the extraction procedure.

The frame signature is a 380-dimensional feature vector that describes the intensity differences between the various sub-regions of an image. As shown in **Fig. 2**, pairs of sub-regions are pre-defined for each dimension, from which the intensity difference is calculated. These sub-regions are configured at various scales, shapes and locations to provide uniqueness and robustness to the descriptor. In order to extract a frame signature, the difference between the average

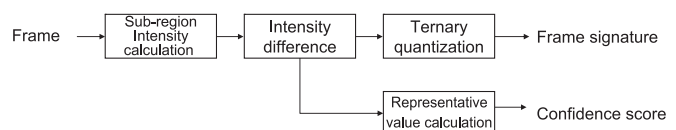


Fig. 1 Video signature extraction procedure.

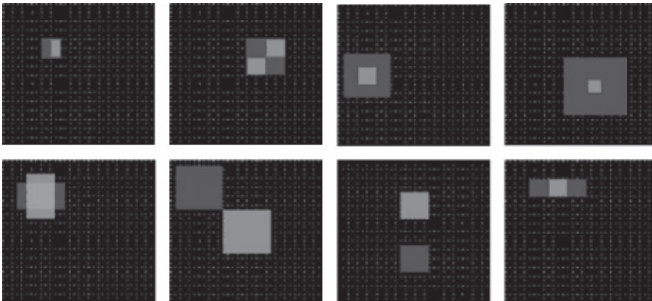


Fig. 2 Samples of sub-region pairs.

intensities of each sub-region pair is calculated. Then, the intensity differences are quantized into ternary values (-1, 0, +1), which represent the intensity relationship between the two sub-regions. The quantization threshold is determined adaptively for each frame by considering the distribution of intensity differences, so that the frequency of the quantized values becomes uniform. Finally, the extracted 380-dimensional ternary vector is encoded by representing each group of 5 consecutive dimensions into one byte value. This results in a very compact signature size of 76 bytes per frame. The confidence score shows the reliability of the frame signature, which is extracted by analyzing the complexity of the image structure. The confidence score is extracted by calculating the representative value of the intensity differences calculated at the frame signature extraction, and by converting it to 1 byte values. The confidence scores are low for the frames containing flat and featureless images.

### (2) Matching of video signatures

Matching segments between two video contents can be detected by comparing each frame signature sequentially. The L1 distance between the frame signatures is calculated and the frames are judged to be matching frames if the L1 distance is smaller than a pre-set threshold. A frame interval that contains continuous matching frames is extracted as a matching segment. False matches caused by scenes with flat and featureless images can be filtered out significantly by excluding matching segments with low confidence scores.

## 2.2 Advantages of the Video Identification Technology

The video identification technology using video signatures has the following three advantages.

### 1) Robust identification of altered or edited video content

The Video signature is an extremely robust tool for identifying videos with all kinds of alternations and editing effects.<sup>\*1</sup> It can accurately identify videos with editing effects such as: caption overlays, compression (with distortion, block noise, etc.), color compensation, resolution reduction, frame rate conversion, analog copying and camera capturing. It does this with almost no false matches. Thus, the technology can be used to automatically detect linkages between video contents before and after editing. This is a very useful aid in the video production process.

### 2) Detection of short video scenes

The technology can accurately detect not only entire video copies, but also video scenes as short as 2 seconds (approximately 1 shot). Therefore, it enables a search of a specific short scene within a longer video.

### 3) Large volume searches by a compact system

Due to a compact signature size, an in-memory matching process of long-hour video content is possible. The technology can match approximately 1,000 hours of video in 1 second using the processing speed of a typical home-class PC. This technology is suitable for carrying out searches of large-size data archives.

The video signature has been accepted in the international standard as a universal descriptor for video identification. It will facilitate interoperable management between different video systems. By using the video signature, mutual searching across video archives provided by different vendors now becomes possible.

## 3. Video Identification Solutions

### 3.1 Confirmation of Whether or not Identical Video Data is Registered

Using this technology will facilitate checking if any same video content has already been stored in the database when a video content holder tries to register a new video content to their video database. This means video content checking operations that are usually carried out manually will become significantly more efficient.

For example, when constructing a database of TV adver-

<sup>\*1</sup> The technology achieved an average detection rate of 96% for modified video content at a very low false alarm rate of 5 ppm (parts per million) in tests conducted by the international standardization organization. This result shows a 20% improvement on average compared to that of the conventional method. A maximum improvement of 65% is achieved for some types of content modification.

## Video Identification Solution Using a “Video Signature”

tisements, newly broadcast advertisements should be searched and stored in the database. Even though approximately 4,000 advertisements are broadcast every day, their inspection and storage operations are usually done manually. Our video identification technology will reduce the load of manual inspection and provide an efficient inspection operation. It extracts video signatures from an advertisement just broadcast and from all advertisements broadcast in the past. It then compares these video signatures to find out if the advertisement just broadcast matches any advertisements already stored in the database. According to the result, an advertisement that matches any of the others stored already in the database will not be registered. Only new advertisements will be registered into the database, so that operators must perform a final check of those that the system has judged to be “New content.” This will eliminate most of the operator workload that has hitherto been required to manually check the video content. So, the time consumed for this work will also be greatly reduced. Moreover, by combining this technology with an audio identification technology, a newly broadcast advertisement that contains identical images with a previous advertisement that contains different audio data can be detected and assessed. In this way, an efficient video content data checking operation can be performed when registering a new video content into an existing database. Identical video content cannot be registered in the same database so that the database memory capacity can also be maintained efficiently.

### 3.2 Detection of Illegally Distributed Copies

Due to the proliferation of the Internet and video viewer client, a significant volume of video content is available on the market. Video identification technology allows content holders to efficiently check whether distributed content is legal or illegal. This means that a large volume of video content can be checked without the need to resort to manual inspection.

For example, someone uploads video content to the Internet without acquiring permission of the copyright holder. In such a case the original video content may be subsequently modified contrary to the intention of the holder. Such illegal actions have been continuously increasing and they have now become a crucial issue for content holders as well as video hosting service providers. Video hosting service providers need to detect illegal content and prevent it from being uploaded to the Internet in order to maintain their service reliability. Hitherto all video content uploaded to the Internet has been checked by manual inspection. Our video identification tech-

nology will reduce the workload and provide efficient inspection of illegal video content.

The following explains details of the operational process of this case. First, a video content holder specifies a content to be protected from illegal distribution. The system extracts video signatures from the content in order to make a contents list. It then extracts video signatures from the content uploaded by an Internet user. Video signatures of both contents are checked against the contents list to find out if the content uploaded by an Internet user is legal or illegal. The result may also be utilized by video hosting service providers when they decide whether or not a video content should be uploaded to their website. Moreover, in many cases, video content uploaded to video hosting service sites is altered from the original content such as by including foreign language overlays, resolution modifications and camera captured copies, etc. Even so, this new technology features robust countermeasures to deal with such content alterations, so that it can detect altered video content quickly and accurately. Even in a case that manual inspection needs to be provided, operators only have to check the matched sections instead of checking the entire contents, thanks to the frame-by-frame identification technology. Illegal video content distributed via video content service provider websites can easily and quickly detect and suppress such illegal video content distributions.

### 3.3 Checking the Usage Conditions of Registered Video Content

#### (1) Auto generation of link information

By utilizing the advantage of the video signature technology, an efficient and comprehensive visualization system can be constructed to check the use of the video content stored in large-capacity archives.

**Fig. 3** shows the basic block-diagram of this system. The system consists of the following blocks: video recording, video analysis, link information analysis, and video/link information presentation. In the video recording block, video content is recorded in the video source database (hereinafter referred to as DB). Subsequently, video signatures are extracted from the video frames and stored in the video feature DB. At the same time, thumbnail images and a list of editing points are generated and stored in the metadata DB. In the link information analysis block, the link information of the video content is automatically analyzed by mutually matching the video signatures stored in the video feature DB. The analyzed link information will then be registered in the

metadata DB. The video/link information presentation block visualizes an overhead view of the link information and scene structure of the video content stored in the metadata DB. This enables users to smoothly access the related video content.

An example of the link information generated by this system is shown in Fig. 4. The link information of a shared topic broadcast by different broadcasting channels is illustrated. Broadcasting channels are located on the horizontal axis and time on the vertical axis. A topic with many links among

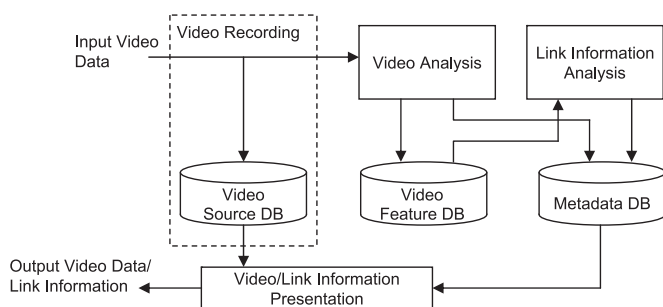


Fig. 3 Basic block-diagram of the system for checking usage conditions.

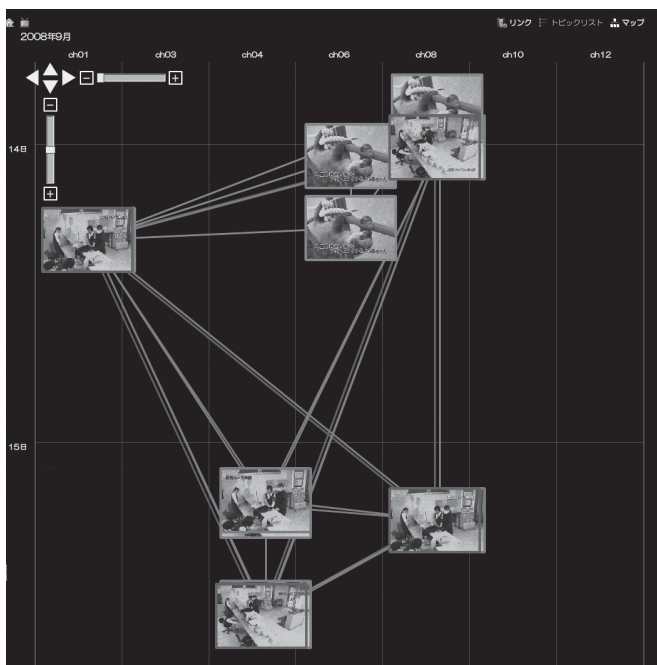


Fig. 4 Example of linked information presentation.

different broadcasting channels is usually a popular and important one. Moreover, by generating the link information of the broadcast advertisement, the most popular advertisement can be visualized for each of the broadcasting channels and times. Key topics or scenes of a TV program can be easily grasped by generating the link information between an entertainment program and its summary version, or that between a drama and its preview.

**(2) Relationships between video program materials and on-air ready video content**

Video content service providers such as TV stations usually hold a large volume of video content stored in their archives. They store on-air ready program data as well as various related video program material data including interviews, etc. Generally, metadata is attached to such video data, which is used to search a target video content.

However, a search sometimes fails due to failure to notice a fault or by mistakes while inputting keywords.

Video signature technology uses video images as a search key for searching content. It can thereby accurately detect identical video contents. Even if a caption overlay or frame rate conversion is applied, video signature technology can identify which is the video material to be used for on-air ready video content. It can also show which video material is used in which installment of the on-air ready video content.

**4. Conclusion**

In this paper, we conclude that the video identification technology can enable an innovative video content management operation and become a core technology of related operational systems. We will continue to develop and provide solutions using this technology and will thereby contribute to the continuing growth of the mass media industries.

**Reference**

- 1) ISO/IEC 15938-3:2002/AMD 4:2010, "Information Technology - Multimedia content description interface - Part 3: Visual, Amendment 4: Video signature tools," 2010.

## Video Identification Solution Using a “Video Signature”

### Authors' Profiles

**KANEKO Hiroshi**

Manager  
Mass Media Solutions Division  
Carrier and Media Solutions Operations Unit

**OZAWA Takato**

Mass Media Solutions Division  
Carrier and Media Solutions Operations Unit

**NOMURA Toshiyuki**

Principal Researcher  
Information and Media Processing Laboratories

**IWAMOTO Kota**

Assistant Manager  
Information and Media Processing Laboratories

---

The details about this paper can be seen at the following.

**Related URL:**

**<http://www.nec.co.jp/press/en/1005/0701.html>**

---

# Information about the NEC Technical Journal

---

Thank you for reading the paper.

If you are interested in the NEC Technical Journal, you can also read other papers on our website.

## Link to NEC Technical Journal website

Japanese

English

---

## Vol.6 No.3 Imaging and Recognition Solutions

Remarks for Special Issue on Imaging and Recognition Solutions

NEC's Pursuit of Imaging and Recognition Technologies

### ◇ Papers for Special Issue

#### Image recognition/analysis

Flow Line Analysis Technology for "Visualizing" Human Behavior and Utilization Examples

Video Identification Solution Using a "Video Signature"

#### Image accumulation/processing

Evolution of File-Based Image Archiving System

Broadcasting Service Platform Solution of the Next Generation

Total Nonlinear Editing Solution that Supports News Production Workflow

Rich Graphics Solution for Embedded Device - GA88 Series IWAYAG -

Development of Ultra-low Latency Codec

#### Image distribution

Wearable Unified Communication for Remote Tour Guide and Interpretation Services

Trends in Digital Signage Solutions

Next Generation Communication with a "Telecommunication Robot"

### ◇ General Papers

Development of a High-Intensity Projector Using LED Light Source

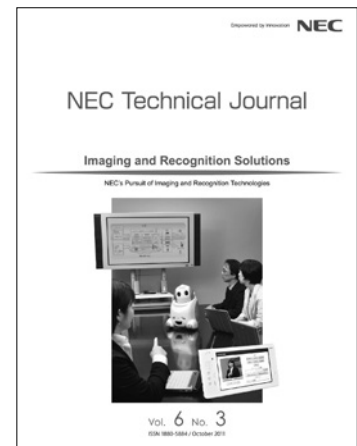
Development of an Environmentally Conscious LCD Projector

Improved Projector Functions Based on System Linkage with PC

The MultiSync PA Series of Professional Display Offers Both Accurate Color Reproduction and High Usability

Development of a Video Wall Display System Using Ultrathin-Bezel LCD Panels

"Office Cool EX Series" Featuring Unprecedented Weight/Size Reductions



Vol.6 No.3

October, 2011

Special Issue TOP