109th MPEG Sapporo, Japan, 7 - 11 July 2014, Meeting Report

Panos Kudumakis
qMedia, Queen Mary University of London

Contents

1 Publish/Subscribe Application Format (PSAF)

The PSAF Group issued an updated requirements document mainly introducing a 'matching' information service between publishers and subscribers. This affected both Contract Expression Language (CEL) and Media Contract Ontology (MCO) standards by issuing amendments in order to facilitate PSAF requirements. PSAF (ISO/IEC 23000-16) standard reached Working Draft (WD) status at 109th MPEG meeting. A summary of both of these documents is given in the next paragraphs.

Publish/Subscribe (PubSub) is an established communication paradigm where senders do not communicate information directly to intended receivers but rely instead on a service that mediates the relationship between senders and receivers. In Publish/Subscribe model senders (called Publishers) post information on and receivers (called Subscribers) declare their interest in a certain type of information – before or after a publication – to a service.

A typical workflow of a content distribution context that can benefit from the Publish/Subscribe modality is given by Table 1.1 where the need for 4 information objects is highlighted in italic.

Table 1.1: Steps in multimedia Publish/Subscribe

	Step	Information type required	Acron.
1	Creator stores resource	Resource ID or locator	
2	Creator stores information on resource	*Resource Information*	RI
3	Publisher publishes information on resource	*Publication Information*	PI
4	Subscriber subscribes to a class of resources	*Subscription Information*	SI
5	Service matches subscription with publication	No information object	
6	Service issues notification(s)	*Notification Information*	NI
7	Subscriber opens notification, requests/plays resource	No information object	

Note that steps 3 and 4 may also happen in reverse order. The steps of the Publish/Subscribe mechanism can be expressed by the following walkthrough and graphically depicted in Figure 1.1.

1. User 1
    a. Creates SI
    b. Subscribes SI
2. User 2
    a. Creates RI
    b. Stores RI
    c. Stores Resource
    d. Creates PI
    e. Publishes PI
3. Match service
    a. Finds Match between Peer1's SI and Peer2's PI
    b. Notifies Peer1 and Peer2
4. User 1
    a. Retrieves RI
    b. Retrieves Resource

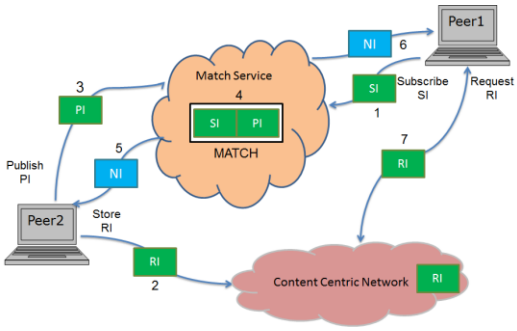


Figure 1.1: Implementation of Publish/Subscribe for multimedia applications

Updated requirements

Publish/Subscribe users for multimedia applications shall be able to:

1. Define users to be/not to be notified of an event in the form of
    a. Enumerations
    b. Groups of users
    c. Conditions that users should satisfy
2. Define a standard information package (Resource Information) containing at least the following information elements
    a. Descriptions of Resource
    b. Permissions, obligations and prohibitions for the use of the Resource by the Publisher or by the Subscriber
    c. List of users to be notified/not to be notified that a specific use of a Resource has been made
3. Define a standard information package (Publication Information) containing at least the following information elements
    a. Metadata related to the Resource Information
    b. Permissions, obligations and prohibitions for the use of Metadata related to the Resource Information by the Match Service Provider
    c. List of users to be notified/not to be notified that a match between this publication and subscriptions has been found
4. Define a standard information package (Subscription Information) containing at least the following information elements
    a. Query related to Metadata related to the Resource Information
    b. Permissions, obligations and prohibitions for the use of the Query by the Match Service Provider
    c. List of users to be notified/not to be notified that a match between this subscription and publications has been found
5. Define a standard information package (Notification Information) containing at least the following information elements
    a. Resource Information ID
    b. Publication Information ID
    c. Subscription Information ID
6. Select Match Service Provider(s)
7. Prescribe
    a. Which Publications shall be/shall not be considered in computing matches with a Subscription
    b. Which Subscriptions shall be/shall not be considered in computing matches with a Publication
8. Prescribe to Match Service Provider which users are to be/not to be notified that a match has been found
9. Define the validity of Publication/Subscription (i.e. the period of time defined by a start and end time during which the Match Service Provider shall notify matches)
10. Guarantee authenticity of Resource, Publication, Subscription and Notification Information
11. Update Publications/Subscriptions

Relationships between Publish/Subscribe users

The relationships between the four users identified in Table 1.2 are depicted in Figure 1.2.

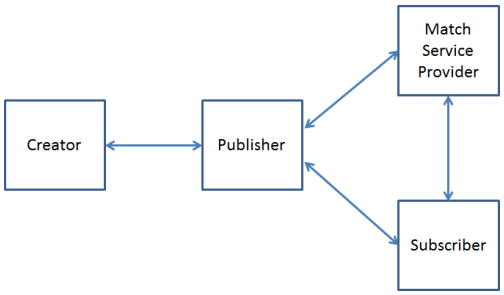


Figure 1.2: Relationships between Publish/Subscribe users

Table 1.2: Relationships between Publish/Subscribe users

Party 1	Act	Party 2	Act
Creator	Grants right to distribute	Publisher	E.g. shares revenues
Publisher	Requests match of pub with subs	MSP	E.g computes statistics
Subscriber	Requests match of sub with pubs	MSP	E.g computes statistics
Publisher	Grants right to use (e.g. play) resource	Subscriber	E.g. pays

The standard specifies the formats used as payloads by a Publish/Subscribe mechanism for media applications and specifically the Information objects of Table 1.3. The payloads are exchanged by the users defined in Table 1.4.

Table 1.3: Information objects needed in PSAF

Information Objects	Definition
Resource Information	Information that is represented by a Digital Item (DI). This includes Resource ID, Metadata, Usage Rights and Conditions and Event Report Request
Publication Information	Information that is sent to the selected MSP(s) in order to make known certain Resources. This is represented by a DI that includes Metadata related to RI, Usage Rights and Conditions and Event Report Request. PI also contains the identifiers of (groups of) users whose subscriptions are eligible for a match with the publication and the identifiers of (groups of) users who should be notified of a match
Subscription Information	Sent by subscriber, speculative to a PI, where descriptive metadata are replaced by one or more than one queries
Notification Information	Information that, as a minimum, enables a subscriber to retrieve RI

Table 1.4: Users in multimedia Publish/Subscribe

Creator	Creates Resource Information
Publisher	Sends Publications Information to MSPs
Subscriber	Sends Subscriptions Information to MSPs
Match Service Provider (MSP)	Performs matches of Subscriptions with Publications on the request of Pub-Sub users Sends Notification Information to users listed in Publication Information and Subscription Information
Pub-Sub user	Is a Publisher or a Subscriber

Technologies

The following MPEG technologies are used to specify the PSAF

1. Digital Item
2. Digital Item Identifier
3. Simple Metadata Profile
4. MPEG Query Format
5. Rights Expression Language
6. Event Reporting

Outside MPEG the W3C Signature Recommendation.

Output documents
N14556 - Requirements for Publish/Subscribe Application Format (PSAF)
N14657 - WD of Publish Subscribe Application Format

2 Audio Synchronization (aka audio fingerprinting)

ISO/IEC 14496-3:2009/DAM5 describes the Audio Synchronization algorithm. An example of the applications using the audio synchronization scheme is a "second screen" application where the 2nd screen content is automatically synchronized to the 1st screen content. In this scenario, no common clock covering the 1st and 2nd screen devices is required, nor an exchange of time-stamps between the devices. Synchronization of the contents between the devices is done by using audio features extracted from the 1st screen content.

For example, the 1st screen content is distributed over existing broadcast system, and the 2nd screen content is distributed over IP network. The audio feature stream of the 1st screen content is sent to the 2nd screen together with the 2nd screen audio/video content over the IP network. In the 2nd screen device, the audio of the 1st screen content is also captured by a microphone and its feature is extracted. The extracted feature from the microphone input and received feature from IP network is compared and the time difference is computed. This time difference is used to align the 2nd screen audio/video content to the 1st screen content. One of the greatest benefits of this approach is that there is no need to modify the transmitter/receiver system of main media stream (for 1st screen).

Figure 2.1 shows the overview of an Audio Synchronization system describing how the system synchronizes two input audio signals. Audio Signal #1 is to be broadcasted as the 1st screen content and Audio Signal #2 is an audio of the 1st screen content captured by a microphone of the 2nd screen device. The system consists of an Audio Feature Extraction tool and an Audio Feature Similarity Calculation tool. The Audio Feature Extraction tool generates audio feature for synchronization from a time domain audio signal. The Audio Feature Similarity Calculation tool compares two audio feature streams to find time difference between the audio signals.

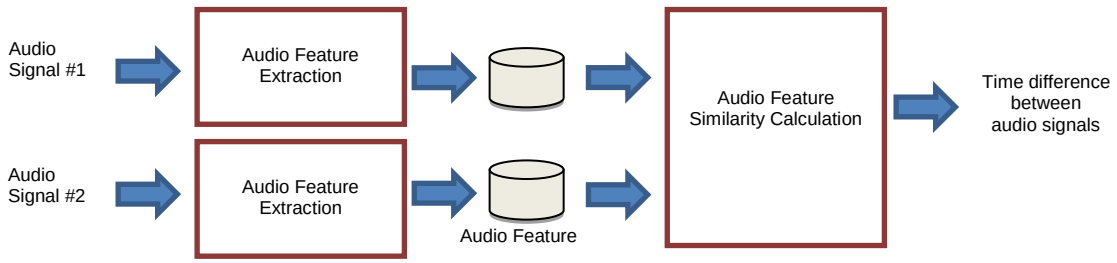


Figure 2.1: Audio Synchronization System

Audio Feature Extraction Tool (Normative)

The block diagram of Audio Feature Extraction tool of Figure 2.2 shows how the audio feature is extracted from a time domain audio signal.

First of all, the sampling rate of the input audio signal is converted to 8kHz and divided into audio frames in time domain. For each audio frame, pre-emphasis filter is applied to emphasis high frequency, then band pass filtering is applied in order to split the audio signals into 5 equally spaced frequency bands in log frequency domain.

Then, auto-correlation within each sub-band is calculated and each of the auto-correlation is normalized by maximum peak of the auto-correlation within the sub-band. The normalized auto-correlations obtained from

the sub-bands with strong pitch component are summed together to obtain a single integrated auto-correlation values for each time frame.

The lag values which give peaks in the integrated autocorrelation values are detected. The integrated auto-correlation values are converted to audio features represented with binary data based on the detected peak position. The rate of the binary data is converted to that for transmission. The series of the above process is repeated while the input audio signal is available.

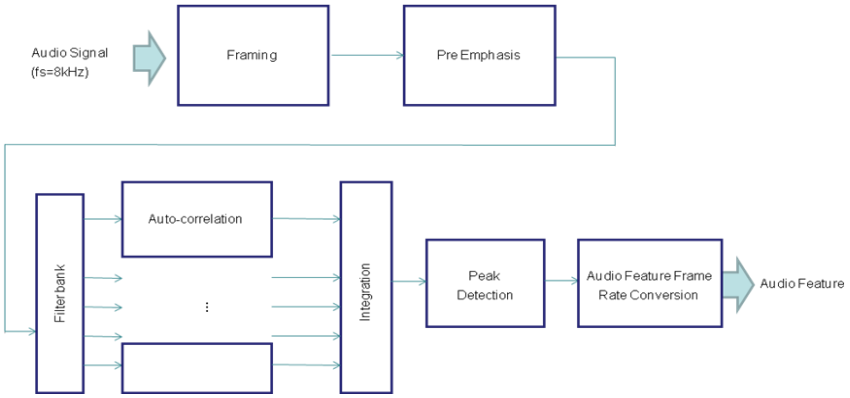


Figure 2.1: Block Diagram of Audio Feature Extraction Tool

Audio Feature Similarity Calculation Tool (Informative)

The block diagram of the Audio Feature Similarity Calculation tool of Figure 2.3 shows how the time difference between audio signals is calculated from audio feature streams obtained by Audio Feature Extraction tool. The calculation is performed to measure the similarity of two "blocks" consisting of audio feature vectors extracted from audio feature sequences.

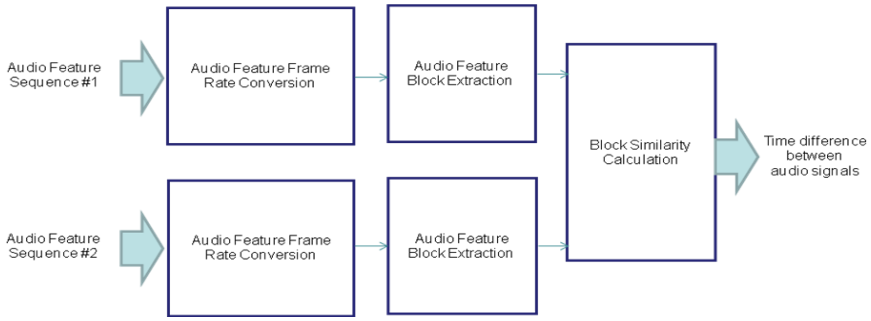


Figure 2.3: Block Diagram of Audio Feature Similarity Calculation tool

Output documents

N14733 - Text of ISO/IEC 14496-3:2009/DAM 5, Support for Dynamic Range Control, New Levels for ALS Simple Profile, and Audio Synchronization

3 Sample Variants in ISOBMFF (aka forensic watermarking)

Sample Variants in ISOBMFF (ISO/IEC 23001-12) standard reached Committee Draft (CD) status at 109th MPEG meeting. This is about signalling non-standardised “transaction” watermarking decoders in ISOBMFF and consequently determine media players' behaviour with respect to authorised/non-authorised content. A summary is given below.

This standard (ISO/IEC 23001-12) adds support for a general framework for sample “variants” in ISO/IEC 14496-12 ISO base media file format (ISOBMFF). This would be used by a forensic “watermarking” system

to modify the base sample, but is independent of the "watermarking" algorithm. Variants are sample data that may be used by a decoder and Digital Rights Management (DRM) system to ultimately output video or audio that is marked in a way that can be unique to individual decoders or decoder product models. The application of the variants during the decode process is under control of the DRM system (and ultimately the content provider).

As ISOBMFF (14496-12) and Common Encryption (CENC) (23001-7) becomes increasingly utilized for the distribution of high-value content, there is a growing need to support coding tools to aid with the identification of players from which content has been copied without authorization.

A common identification capability is known generically as "forensic marking". A forensic marking system utilizes "variances" in the sample data that in aggregate are unique to the player. "Forensic marking" enables efficient technical content protection response to the discovery of unauthorized copies to more promptly prevent future occurrences. The ability to better respond to such events encourages publishers to continue to release high value content using ISOBMFF.

There is today no standardized way to deploy "forensic marking" with ISOBMFF files. Having a standardized framework enables:

- industry adoption of "forensic marking" for better interoperability
- support of ISOBMFF standard by services which require "forensic marking"
- reduced complexity of analysis tools.

This is NOT about standardizing any specific "forensic marking" technique.

Requirements

The proposed technical approach to support a forensic marking framework in ISOBMFF:

- maintains file format compliance with ISOBMFF and CENC
- supports the provision of one or more media data variants for any sample in the ISOBMFF file
- enables cryptographic control over which media data variants are made available to a particular player
- is compatible with, and agnostic to, any particular sample and sub-sample based forensic mark system.
- is resilient to scrubbing attacks (the removal of variant sample data from rendered output)
- minimizes the encoder and player overhead of adding variant data to the file
- enables independent verification of ISOBMFF files - verification should occur with no dependency on knowledge of the authoring tool or authoring approach
- maintains support for download, progressive download and super-distribution use cases
- maintains compatibility with CENC, and thus is independent of any specific DRM system
- provides security robustness:
  - sample keys are of utmost value and thus must not be mixed with other keys since they may be subject to different compliance and robustness rules, for example used within trusted hardware (in contrast to keys encrypting metadata which may be in software)
  - variant samples can be double-encrypted with both a media key and another key
  - variant metadata must be encrypted to obscure the variant samples that are actually used by a decoder

In meeting the above requirements, there are some disadvantages to standard editing tools such as not having offset values in the clear. Editing would have to be performed before the variant metadata was encrypted. Within a file, there are a number of samples that are encoded with variant samples. This is shown in Figure 3.1 below. The Figure shows 3 samples in a series left to right, the middle of which has variants. The top row shows what's encoded in the ISOBMFF and the bottom row is what is output to the decrypter/decoder after the variant processing.

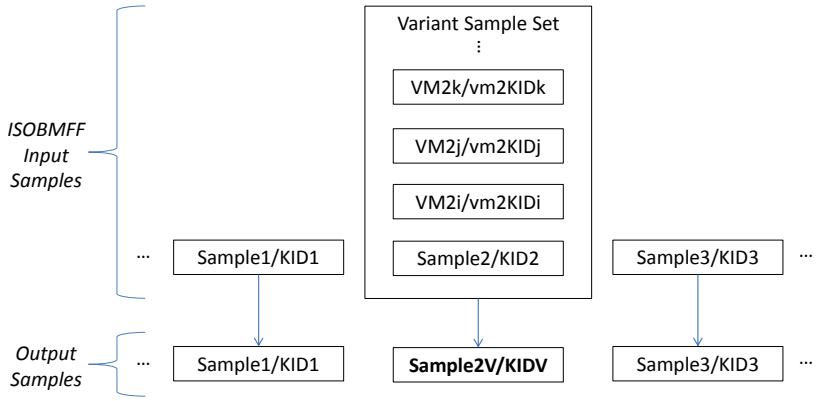


Figure 3.1: Variant sample structure

The control point for the use of the proposed framework is the content publisher:

- the content publisher will encode encrypted, compressed variant data into the ISOBMFF file and ensure that each set of variant sample data for a given sample time is encrypted with a different key.
- the content publisher will work with the DRM to the release of keys such that the playback path (the actual sample and variant data utilized during playback) is controlled and the player can only decrypt and render the data that it has been authorized to render.

The decoder model for the processing of the file is shown in Figure 3.2. Critical to the variant decoding process is a DRM system that controls if and how the variant samples are processed. Note that the decrypt and decode steps are standard operations as they would be for any CENC-enabled decoder.

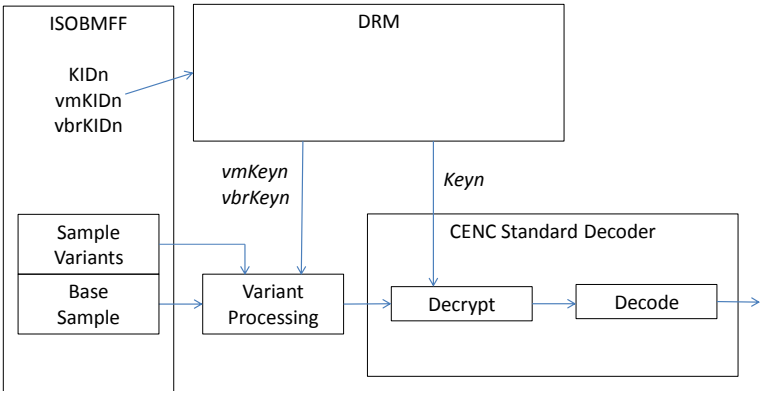


Figure 3.2: Variant Decoder Model

By operating in the encrypted/compressed domain, secure baseband link operation (e.g. dedicated, secure video pathways) is preserved and is intended to be fully compatible with CENC.
The proposed approach is intended to support player model identification via DRM client identification, such as a unique decoder model number.

Example
Consider a variant constructor consisting of three byte range groups:

- The first byte range group has one variant byte range S1, which is unencrypted.
- The second byte range group has one variant byte range S2, which is encrypted.
- The third byte range group has two variant byte ranges, S3 and S4, each of which are encrypted.

At encryption time:

- The variant media associated with variant byte range S1 is not encrypted, resulting in unencrypted variant media data M1.
- The variant media associated with variant byte ranges {S2, S3,S4} are each encrypted with Media Key K1 (KID KID1), resulting in encrypted variant media data {M2*, M3*, M4*}.
- The encrypted media data M3 is further encrypted with byte range key K3 (KID KID3) and encrypted media data M4 is encrypted with byte range key M4 (KID KID4), resulting in doubly encrypted media data M3** and M4**.

The resulting variant constructor will have four byte ranges and is structured as [| S1 | S2 | S3 S4], where the symbol "|" indicates the start of a byte range group. The underlying media data is stored as {M1, M2*, M3**, M4**}.

If the decoder has access to KID1 and KID3 only, per the Decoder Model it will do the following:

1. Process S1, establish it as unencrypted and consequently add M1 to the sample assembly and identify it as unencrypted.
2. Process S2, match KID1 and consequently add M2* to the sample assembly and identify it as encrypted.
3. Process S3, match KID3 and consequently decrypt M3** using K3, then add the resulting M3* to the sample assembly and identify it as encrypted.
4. Process S4, not recognize KID4 and consequently skip M4**.
5. Decrypt the sample assembly [M1 M2* M3*] by skipping M1 and using the Media Key K1 to decrypt M2* and M3*, resulting in unencrypted variant media data[M1 M2 M3].

Output documents
N14599 - Text of ISO/IEC CD 23001-12 Sample Variants in ISOBMFF

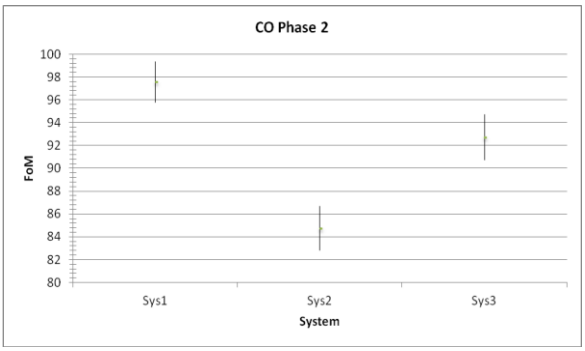
4 Selection of 3D Audio Phase 2 Technology

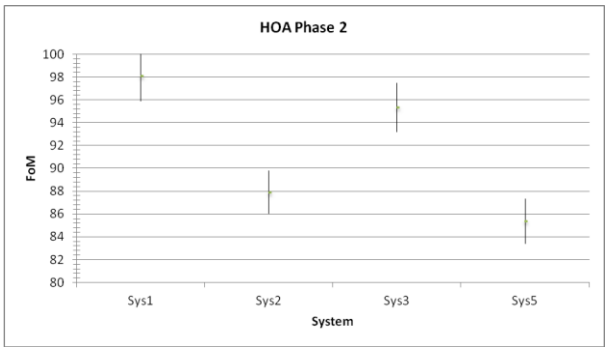
Proponent submissions for 3D Audio Phase 2 Technologies were due May 16, 2014, and in the period leading up to the 109th MPEG meeting a listening test was conducted to evaluate the performance of the submitted technology. Test scores were available as a contribution to the 109th meeting held in Sapporo, JP in July 2014, and submissions were evaluated at the 109th MPEG. This document reports on the results of the listening test and on the selection of technology for 3D Audio Phase 2.

The listening test was conducted for two signal sets: Channel/Object (CO) and Higher Order Ambisonics (HOA). Submitted technology was evaluated at the following bit rates for each signal set: 48 kb/s, 64 kb/s, 96 kb/s and 128 kb/s. The listening tests used the MUSHRA test methodology. Tests were conducted at the test sites shown in the following table. The number of listeners for each test (after post-screening) is shown in the rightmost column.

Test and Signal Set	Bitrate	ETRI	IDMT	IIS	MGL	ORG	QUAL	SONY	TECH	N
Test2-1-CO	128 kb/s	X	X	X	X			X		42
	96 kb/s	X	X	X	X					47
	64 kb/s	X	X	X	X			X		36
	48 kb/s	X		X	X			X		51
Test2-1-HOA	128 kb/s			X	X	X	X		X	43
	96 kb/s				X	X	X		X	32
	64 kb/s				X		X		X	23
	48 kb/s			X	X		X		X	41

Furthermore, the Listening Test Logistics for 3D Audio Phase 2 defined a Figure of Merit (FoM) for the listening tests. The FoM plots for the submitted technology are shown here:



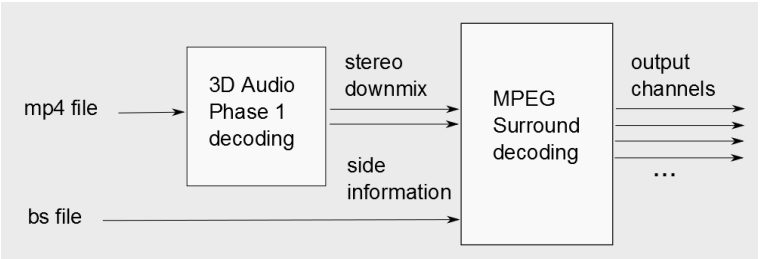


Selected Technology

Audio experts discussed the architecture of the submissions, particularly whether there was opportunity to combine technology from the best submissions.

CO Signal Set

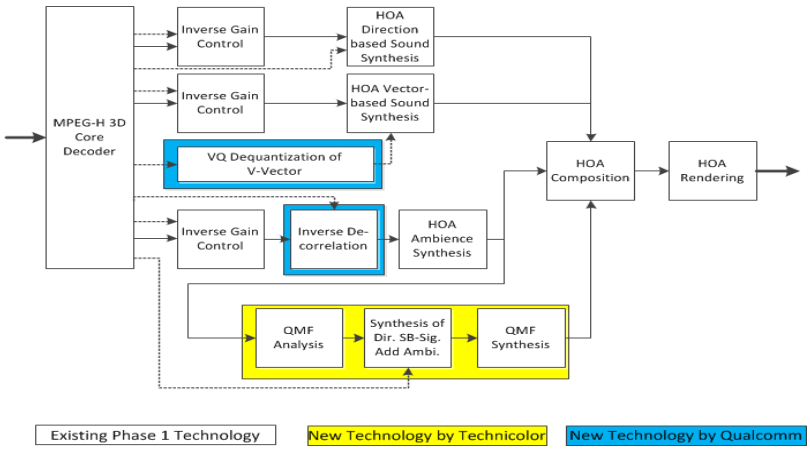
For the CO signal set, it was the consensus of the Audio subgroup to select Sys1 (FhG-IIS) as Phase 2 technology. This system had the highest Figure of Merit. It was acknowledged that Sys3 (ETRI) had a very similar architecture so that a first collaborative step for Phase 2 is to explore which parts of ETRI technology could be brought into the Sys1 architecture by means of Core Experiments such that it provides demonstrated benefit.



3D Audio Phase 2: C/O Decoder

HOA Signal Set

For the HOA signal set, it was acknowledged that Sys1 (Qualcomm) had the highest Figure of Merit. However, Sys3 (Technicolor) had a complementary architecture and produced better results for some items at some bit rates. Hence it was the consensus of the Audio subgroup to merge Sys3 into Sys1 and that this would be Phase 2 technology. A first collaborative step for Phase 2 is to merge the Qualcomm and Technicolor technology into a single architecture.



3D Audio Phase 2: HOA Decoder

Timeline for Standardization
Phase 2 WD text and RM decoder reference software will be available prior to the 110th MPEG meeting, October 2014. At this point Core Experiments will begin. The envisioned timeline for standardization is:

Status	Meeting	Date
WD	110	October, 2014
CD	111	February, 2015
DIS	112	June, 2015
IS	114	February, 2016

Output documents
N14748 - Report on Selection of Technology for 3D Audio Phase 2

5 Internet of Things (IoT)

The Requirements subgroup recognizes that MPEG-V provides technology that is applicable in the area of Internet of Things. The Requirements subgroup would like to ask MPEG members to encourage relevant external organisations to share their views on this subject directing them to the MPEG-V white paper.

6 Exploration - Use Cases for Processing and Sharing of Media under User Control

In the actual use of MPEG technologies, there are many situations that require the media of its components to be private, with processing sharing under user control. Examples of this are:

1. searching an encrypted audio visual database with an encrypted query
2. identifying a spoken keyword in a private conversation, e.g. encrypted audio
3. removal of identification clues from multimedia content such as media, audio or speech
4. sharing multimedia content in a limited context, e.g. making a picture available to a limited list of persons, for a limited time, or for a specified purpose

MPEG has been developing very successful standards that process audio and video information based on its significant expertise in this domain MPEG is now identifying application domains and extract requirements to achieve both application of existing standards, and use of its expertise to develop new standards, supporting the private processing and sharing of media.

Use Case: Privacy-Preserving Media Search & Analysis
In this use case, Alice would like to search a media database that is owned by Bob. Alice encrypts her query so that what she is looking for is kept private and not revealed to an un-trusted party. The media in the database is also encrypted to keep it secure. A secure protocol is executed between Alice and Bob to perform matching without decrypting either the query or the media stored in the database. The best matching results are then returned to Alice (or a third party) without revealing the query to Bob, or the media in the database to Alice.

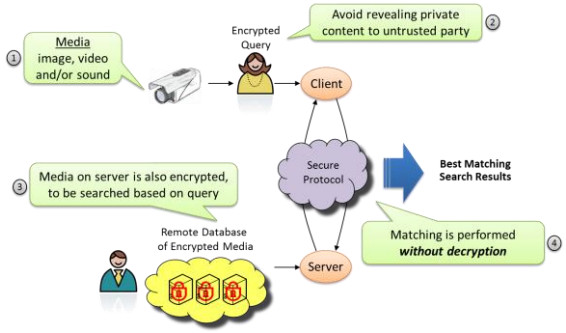


Figure 6.1: Use Case – Search a Media database

Rather than searching for media, it is also possible to consider a privacy-preserving analysis of the media. For example, consider the analysis of a speech signal from a conversation. The speech signal is encrypted to keep the conversation private. The encrypted speech signal is then sent to an un-trusted server for analysis, where keywords from the speech could be detected using a secure protocol.

The general problem of finding nearest neighbours in a privacy-preserving manner is applicable to a broad range of problems, many of which involve media such as images, video and audio signals. This problem can be broken down into two steps. The first step requires a method for privately computing distances, and the second step requires a method for privately finding the minimum distance to determine the nearest neighbour.

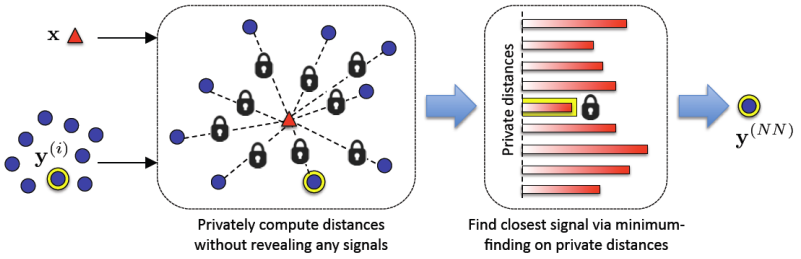


Figure 6.2: Secure Fingerprinting

While the above example scenarios refer to the use of cryptographic primitives, there also exist methods to perform these operations using a form of secure fingerprints.

Use Case: Context aware multimedia privacy protection in video surveillance and social networks
In this use case, recorded video surveillance footage, or social networks share content that may become available to a large audience. Sharing of such content and the context of sharing such information is an essential issue.

The context could be a select group of people, purpose of sharing, time and date, metadata and the likes. Applications on smart devices tend to interact with platforms mutually without informing to the user the privacy and security concerns they may have.

In video surveillance, a number of consortia have been exploring various implications and problems that sharing of video surveillance footage can arise from ethical as well as legal points of views. Examples include EC funded network of excellence VideoSense (http://www.videosense.eu).

Likewise, several examples of applications and services exist that allow sharing of video in social networks, such as Socialcam, a mobile application which allows users to share with their friends as well as with public, various video.

Currently, the usage control of such content is either trivial, or complex, and inefficient. For example, Spotify and Yahoo News automatically publish songs and news one has listened to or read on the profile of users.

Privacy is not static nor deterministic but rather a dynamic and stochastic parameter which needs to take into account both the context in which the protection took place and the context in which a user accesses content and information. It is therefore important to devise a framework in any solution for user control, which explicitly takes into account not only the relationship between content and privacy, but also the context in which this relationship occurs.

Use Case: Privacy-Preserving Video Transcoding
For applications like dynamic adaptive streaming and screen content sharing, video content generated at its source needs to be transcoded to fit a large variety of client devices and dynamic network environments.

When video contains privacy sensitive information and content and transcoding is conducted by an un-trusted party, it is necessary to keep privacy undisclosed while allowing transcoding transformations.

One possibility is to encrypt video content, or a segment (for a period of time, e.g., a particular scene) or a portion of it (for a region in the content, e.g., a person's face), in a transcoding friendly manner. This requires a good understanding of common transcoding transformations used in practice and what encryption schemes can be used to permit conducting these transformations while video content is encrypted using any of those schemes.

Media Usage Control Policies
In order to provide user control over processing and sharing of multimedia content, a flexible, effective and scalable mechanism is to provide users a way to express their control desires in a form that can be processed and monitored systematically, consistently and persistently throughout the lifecycle of the multimedia content. These control desires can be considered as multimedia usage-control policies, much like data usage-control policies many websites (e.g., Facebook) use to protect user identity and data usage-control. The difference here is that, in the user control case, it is the users, not the websites, who specifies how their multimedia content should be processed and shared.

In order to express users' desire on controlling usage of their multimedia content, the following should be supported in a usage-control policy expression language such as a profile of the MPEG-21 CEL:

- Specification of ownership of the multimedia content and their components, with optional authentication checking mechanisms.
- What kinds of processing, or types of operations, can be applied to the multimedia content, and their deontic characteristics such as permissions, obligations and prohibitions (i.e., rights, duties and bans in the MPEG-21 CEL terms).
- How to share, with whom, for how long and under what conditions, also along with their deontic characteristics such as permissions, obligations and prohibitions (i.e., rights, duties and bans in the MPEG-21 CEL terms).

While MPEG-21 CEL provides a generic mechanism, specific information related to the features above will come from applications; a source for this type of information can come from existing user data usage-control policies, such as the ones from Facebook and the usage-control policy template (www.contractstore.com).

Output documents
N14550 - Use Cases for Processing and Sharing of Media under User Control

7 Exploration – Uniform Signalling for timeline alignment

MPEG has started an exploration activity on media stream synchronization in heterogeneous delivery environments, transported using existing MPEG systems technologies. The Systems subgroup is planning a seminar at the 110th MPEG meeting in Strasbourg to advertise this activity and collect more requirements/needs from the industry.

In a number of scenarios, ancillary timed content can be made available to enhance the experience of consuming some primary content. Examples include accessibility, sub-titling or captions, different audio tracks but also content that enhances the user's experience in other ways.

To enable a wide range of interoperable services, it is desirable that the availability and the alignment of ancillary media for the primary media are signalled using standard techniques and terms, in a uniform way, independently from formats and transport protocols.

MPEG has started an exploration activity on discovery and synchronization of auxiliary media streams for its Systems technologies (MPEG-2 Transport Stream, ISO Base Media File Format, MPEG-DASH, MPEG Media Transport). This exploration focuses on the problem of playing one content that is time-aligned to some other content, and possibly delivered using a different transport, physical network and/or encapsulated in different containers. Both contents may be played on the same device, or on different devices sharing synchronization information.

There are several use cases where it may be interesting for a service provider to deliver part of its content on one transport channel, for example free-to-air terrestrial or satellite broadcast, and part of it on another such as broadband IP connection. In order to depict a complete overview of the use cases they can be categorized and distinguished against several characteristics. Here is a non-exhaustive list of such criteria:

- Number of different contents
- Type of the content, e.g., audio, video, etc.
- Format of the content
- Number of devices playing-out at least one of the content
- Physical locations of the devices
- Physical networks the devices are connected to
- Requirement on synchronization accuracy (e.g., frame accurate or not)

Proposed Topics:

- Hybrid Broadcast - Broadband distribution for UHD deployments: a use cases perspective
- Hybrid Broadcast - Internet applications on companion screens
- HEVC and Layered HEVC for UHD deployments (potentially HDR ?)
- Inter Destination Media Synchronization – Social TV use cases
- Audio Fingerprinting based Synchronization

Round table discussion on issues on UHD deployment, second-screen apps, hybrid broadcast-broadband, HEVC, MPEG systems, ..., and demos.

The Systems group welcomes feedback on the above topics and suggestions of demonstrations, which are deemed of interest for MPEG activities.

Output documents
N14644 - Uniform Timeline Alignment
N14645 - Plan of Seminar on Hybrid Delivery at the 110th MPEG Meeting

8 Digital Media Project 3rd Phase: Hybrid-Delivery Media Services (HDMS)

The Digital Media Project (DMP), founded in 2003. In its 1st Phase work programme developed specifications for an Interoperable DRM Platform (IDP) designed to respond to the needs of all players in the digital media value chain. In its 2nd Phase work programme DMP developed specifications and a commercial grade platform on Open Connected TV (OCTV). The source code is available to DMP members for commercial exploitation.

Recently, after two brainstorming meetings held in San Jose, CA, USA, 11 Jan. 2014 and Valencia, Spain, 29 Mar. 2014, DMP announces the specification areas to be covered by its 3rd Phase work programme. In particular, the latter aims to define specifications for a comprehensive system view of media services that are based on a main broadcasting service supplemented by interactive and personalised services delivered via the internet. The specification will include a general model of Hybrid-Delivery Media Services (HDMS) and their stakeholders, HDMS requirements, system-wide interfaces and user application interfaces.

In Sapporo (GA38) substantial input was received in response to the DMP Call for Proposals issued at the 37th DMP General Assembly in Valencia. However, DMP members have been asked to bring more contributions

to the next DMP General Assembly (GA39) in Strasbourg, FR, 25 Oct. 2014, in order to comply with the time line of the HDMS work plan of 24 months with a major milestone after 12 months.