스케일러블 비디오 부호화에 대한 계층적 보호 기법

*헨드리 *김문철

**함상진 **이근식 **박근수

*한국정보통신대학교 (ICU)

**한국방성 (KBS)

*{hendry, mkim}@icu.ac.kr **{cashy, kslee22, kspark}@kbs.co.kr

A Layered Protection Scheme for Scalable Video Coding

*Hendry *Munchurl Kim

**Sangjin Hahm **Keunsik Lee **Keunsoo Park

*School of Engineering, Information and Communications University (ICU)
**Korean Broadcasting System (KBS)

Abstract

Protection to the multimedia contents is inevitable to ensure that only authorized users be able to access the protected contents for consumption. Since protection mechanisms need to be designed efficiently by exploiting the type of the contents, we propose a protection scheme for the video bitstream encoded by Scalable Video Coding (SVC) technique. Our scheme exploits the property of SVC in which a video is encoded into spatial, temporal, and quality scalability layers. By applying our proposed protection scheme to the appropriate scalability layers we can effectively control the SVC contents completely or partially. Each layer can be flexibly protected with different encryption keys or even with different encryption algorithms. The algorithms that are used to protect each layer are described by the standardized protection description tool, which is the MPEG-21 Intellectual Property Management and Protection (IPMP) Components. In this paper, we present the design of the proposed layered SVC protection scheme, its implementation and experimental results. The experiment result shows that the proposed layered SVC protection scheme is very effective and can easily be applied.

1. Introduction

JVT (Joint Video Team), a joint group of MPEG (Moving Picture Experts Group) and VCEG (Video Coding Experts Group), is working on the standardization of the scalable video coding (SVC) which is the scalable extension to H.264/AVC. Within the SVC standard, a video can be encoded into several scalability layers with respects to the spatial, temporal, and/or quality scalabilities. The first scalability layer of the encoded video is called the base layer while the rest of the scalability layers are called the enhancement layers. The benefit of encoding the video in this way is that we can adjust the video consumption to delivery network conditions/characteristics, device capabilities, etc. Hence the SVC allows for producing encoded video bitstream for one-source-multiple-use.

One of the interesting properties in SVC is that the SVC scalability layers are nested. One scalability layer

depends on its lower scalability layers. For example, the second enhancement layer cannot be reconstructed without reconstructing the data from the first enhancement layer which in turn requires reconstruction of frames from the base layer. Considering this property, in this paper we propose a protection mechanism for the SVC contents by exploiting the nesting property of SVC. Depending on which layer we protect, we can get different effects of protection to the whole content. For example, protecting the base layer only is equivalent to protecting the whole content since all the enhancement layers cannot be reconstructed without the reconstruction at the base layer. In this regard, the protection of an enhancement layer will allows for an effective protection of its higher enhancement layers.

To make our protection scheme flexible, we do not mandate any specific encryption algorithm (i.e., DES, AES, XOR-ing, etc). Rather, it can accommodate various algorithms so that the selection is left to users (protectors). The information about the algorithm applied to the content is then described into standardized protection description by which appropriate un-protection algorithm can be invoked at user terminals. By utilizing this protection description, we can have a flexible control on protection of SVC contents at each scalability layer. Each layer can be protected by different algorithms and the description is used to signal the protection information about the algorithm. At the time of decoding, appropriate decryption algorithm can be invoked to un-protect the content. This layered protection approach gives much flexibility in various business models for the SVC contents services. One interesting example is that the base layer of SVC content may be left unprotected for free view but the enhancement layers can be protected for pay-per-view.

In this paper, we utilize the MPEG-21 Intellectual Property Management and Protection (IPMP) Components [4] for the description of the protection information. The description tools of the MPEG-21 IPMP Components are defined in the form of XML schema to convey information such as the list of protection algorithms, the location of the content where the tools are applied, and additional information to execute the tool to un-protect the protected part of the content.

This paper is organized as follow: in Section 2, we briefly describe the fundamental concept of SVC in Section 3, we present a brief explanation about MPEG-21 IPMP Components; in Section 4,we describe the proposed layered SVC protection scheme; in Section 5, we present the experimental result and lastly, in Section 6 we conclude our paper.

2. Scalable Video Coding

Scalability in video coding is defined as a functionality that allows for removal of parts of the bit-stream while achieving a reasonable coding efficiency of the decoded video at reduced temporal, quality, or spatial resolution [1]. It means that the bit-stream itself may contains several scalability parts (later we denote them as scalability layers). Depending on the delivery network and/or terminal capabilities the video resolution and/or quality can be increased by additionally sending the bitstream at the enchancement layer.

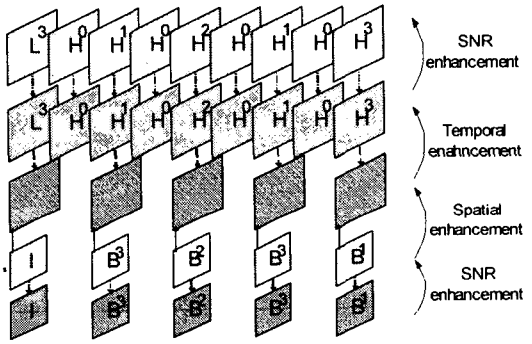

Figure 1. SVC bitstream illustration

Figure 1 illustrates an example of scalability levels that a SVC bitstream can convey. The SVC bitstream can contain base layer only, or the base layer plus an SNR enhancement layer, or the base layer plus an SNR enhancement and a one spatial enhancement. The full resolution that is illustrated in Figure 1 conveys the base layer with its SNR enhancement plus spatial enhancement layer and a temporal enhancement layer with its SNR enhancement layer.

To cope with diverse transport environments, the SVC bitstream is encapsulated into Network

Abstraction Layer (NAL) Unit. Later, it will be adapted according to the transport protocols such as streaming approach or packetized approach. Fig. 2 shows a simple illustration of NAL. Each NAL has its header (1 ~ 3 bytes header depending on its type) and payload (output of encoding).

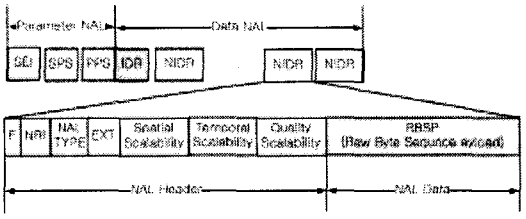

Figure 2. NAL structure illustration in SVC bitstream [1]

3. MPEG-21 Intellectual Property Management and Protection (IPMP)

MPEG-21 IPMP Components (ISO/IEC 21000-4) defines syntax and semantics for IPMP tool applied to Digital Items to facilitate the exchange of governed content between Peers. It is designed to provide a flexible signaling of the protection information applied to the content (or Digital Items). The signaling description is model by XML schema.

MPEG-21 IPMP Components contains two major parts: firstly, the protection descriptions to the Digital Item declaration, *IPMP_DIDL*, in which by using this, we are able to provide a protection to the Digital Item model itself the second part is the governance description to describe the applied protection tools which comprises *IPMP Info* and *IPMP General Info descriptors*. They carry information about the protection tools that are used to describe which part of content they are applied, how to use them, and what are the configuration and/or initialization data for the applied tools. They also provide a container where the rights information can be placed so that the protection and governance may come in one package.

While *IPMP_DIDL* describes the protected version of the Digital Item description, *IPMP Info* contains metadata for *General Info descriptor* and *the Info descriptor*. *General Info descriptor* carries the aggregation information that is valid for the entire content while *Info descriptor* is used to carry information about the protection at the location of the protected data. In brief, *General Info descriptor* carries the list of tools and the list of rights while *Info descriptor* carries information about how to use those tools for the protected parts of content.

4. Protection to Scalable Video Coding: A Layered Approach

4.1 Concept

SVC content is encoded in such a way that a higher layer depends on its lower layers. Hence protecting a layer will have protection effect to all its higher layers. For example, protecting the base layer only is equivalent to protecting the whole SVC encoded video with all scalability levels.

The proposed layered protection accommodates various business models. For example, video on demand (VOD) content may be encoded in three levels of temporal scalability and be protected at each level with different encryption key. Depending on the membership types, the users are provided with different decryption keys to access their designated layers of the SVC video. For unregistered users, they can get only a key to consume the base layer but the registered users can get the key to unprotect the second layer which enables them to consume the base layer and the first enhancement layer. Finally the premium userscan get the key to unprotect up to the second enhancement layer which allows them to access the full resolution of the content.

4.2 Implementation Design

Figure 3 shows a high level illustration where our proposed scheme is incorporated with the encoder and decoder parts. The protection module is not integrated to encoder; rather, we apply the protection to the NAL units which are outputted by the encoder. In this way, the design of the protection is loosely coupled with the encoder and decoder. For example, protecting the base layer only is equivalent to protecting the whole SVC encoded video with all scalability levels.

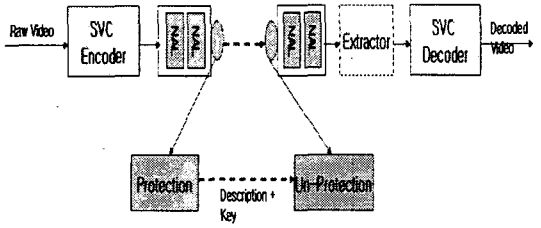

Figure 3. Protection module in encoder and decoder

As shown in Figure 3, the protection module needs to transfer two types of information which arethe protection description and the key(s) (the key(s) is considered to be delivered through "out of band"mechanism which is not the scope of this paper) to the un-protection module in order for it to be able to un-protect the NAL prior to sending it to the SVC decoder.

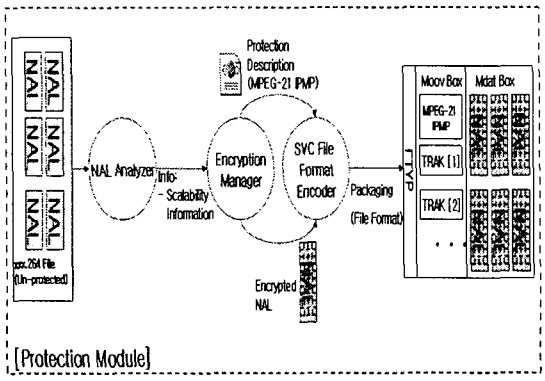

Figure 4. Protection module building block diagram

Figure 4 shows more detailed building blocks of the protection module. The protection module takes the sequence of NAL units from the SVC encoder output file and produces the encrypted NAL units. The protected NAL units together with the description of what protection algorithms are used may be packaged together in the form of MPEG SVC file format. Prior to encrypting the NAL units, related information carried by the NAL units is analyzed. The important information is:

- Number of layers and their corresponding identifier (from the Supplemental Enhancement Information NAL unit)
- Layer ID
- Quality level
- Temporal level

This information is used by the encryption manager to choose which NAL units are to be encrypted and which are left un-encrypted.

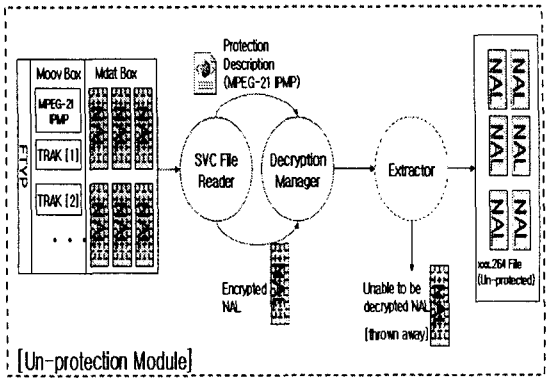

Figure 5. Un-protection module building block

The un-protection module shown in Figure 5 is basically the reversed version of the protection module with the extractor. In case that the un-protection module cannot perform the un-protection due to unavailable key (because the user is not authorized), only un-protected NAL units are sent to the decoder. This will be shown in section 5.

4.3 Packaging the Protected Content

The protected NAL units and description of how it is protected are packaged in a single file following the MPEG SVC file format [3] which is derived from ISO-Base file format [6]. ISO-Base file format use the concept of box in carrying the information. This is meant to ease the parsing and interpretation of the file. Within ISO-Base file format, the major boxes are defined as follow [6]:

- FTYP box. It describes the type of the file type. In case of SVC content, the value of FTYP box is 'svc1'.
- MOOV box. It describes the representation of the data and also the structure of the data in the file.
- METAbox. A container to carry the metadata representation.
- MDAT box. A container to carry the binary representation of the data. The data carried inside the MDAT box are in the form of access unit.

Figure 6 shows an illustration of ISO-Base file

format's boxes for SVC content and how they are protected by the IPMP description. As aformentioned, supposed that we have SVC content with two protected layers of scalability: the base layer and the enhancement layer 1. The NAL units of each protected layer are stored in the MDAT box. One or more NAL units are further grouped together into access units (samples).

The MOOV box contains important descriptions about the samples stored in the MDAT box. Firstly, it carries TRAK box(es). TRAK box provides information about the structure of the samples that belong to the same group/layer. Through its children boxes, TRAK box provides grouping information that associates samples with layers (SampleGroup Box). Secondly, MOOV box may also contain META box. This META box provides a place holder for XML description. The MPEG-21 IPMP description is located in the META box.

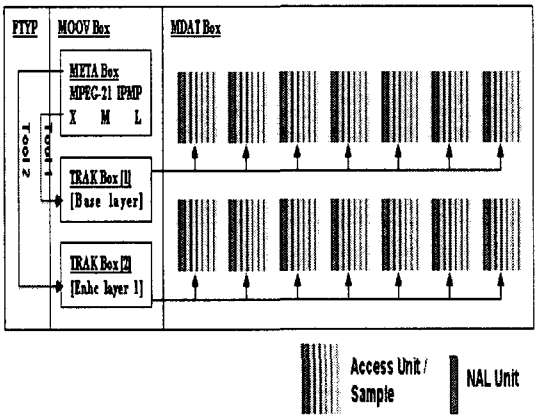

Figure 6. Protection with MPEG-21 IPMP in SVC file

The tools' structure in the MPEG-21 is arranged to be associated with each TRAK. As shown in Fig. 6, Tool1 by its structure is associated to the TRAK 1 while Tool2 is associated to the TRAK 2.

5. Experimental Result

We test our implementation with the MPEG SVC reference software JSVM version 5.0 [2] to verify our proposed layered SVC protection scheme. The scenario for the experiment is as follow: we use FOOTBALL video sequences (one of MPEG SVC test video sequence) and encode its 8 first frames into the base layer and one fine grain scalability (FGS) layer as an enhancement layer. Then we protect only the enhancement layer with our simple encryption tool (XOR-ing the byte with the key). We simulate the situation where two possible situations occur: first the user is the un-authorized user who does not have the key to un-protect the enhancement layer and the second situation is assumed that the user is the authorized one.

Figure 7 shows the console output of JSVM decoder when it is run by the un-authorized user. Without acquiring the decryption key, the user's un-protection module cannot decrypt the protected enhancement layer's NAL units. Hence they are all filtered out by the extractor before sent to the decoder.

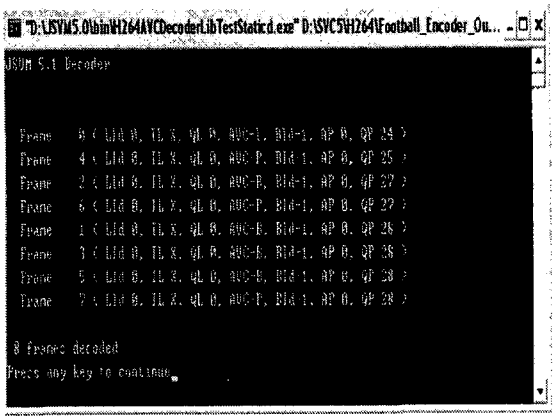

Figure 7. Output of JSVM decoder at the un-authorized user terminal. Only NAL units from base layer are successfully decoded

Figure 8 shows the console output of JSVM decoder when it is run by the authorized user. Since he has the key, all the protected NAL units are successfully un-protected so that there is no NAL unit that is filtered out by the extractor before. As show in Figure 8, the frames (numbered from 0 to 7) are duplicated meaning that they all consist of one base layer frame and one enhancement layer frame.

Figure 8. Output of JSVM decoder at the authorized user terminal. Each decoded frame consists of one base layer frame (AVC compatible) and one enhancement layer frame

6. Conclusions

In this paper, we present a layered SVC protection scheme. The layered protection allows an effective protection mechanism so that bitstreams can be efficiently protected at different scalability layers without protecting the while bitstream which usually causes large computational load at both encoder and decoder side. We utilize a standard description scheme of MPEG-21 IPMP Components for describing protection information for the layered protected bitstream.

We implement the proposed scheme and test it with JSVM reference software. Our implementation and experiment result exhibit that this protection scheme can be easily implemented and effectively works well.

7. References

1. MPEG Video Subgroup, ISO/IEC JTC 1/SC 29/ WG 11/N7795: Text of ISO/IEC 14496-10-2006/PDAM3 Scalable Video Coding, Thailand, January, 2006

2. MPEG Video Subgroup, ISO/IEC JTC 1/SC 29/ WG 11/N7796: Joint Scalable Video Model (JSVM) 5, Thailand, January, 2006

3. MPEG System Subgroup, ISO/IEC JTC 1/SC 29/ WG 11/N7906: WD 3.0 of ISO/IEC 14496-15/PDAM2 (SVC File Format), Thailand, January, 2006

4. MPEG MDS Subgroup, ISO/IEC JTC 1/SC 29/ WG 11/N7717: Text of ISO/IEC 21000-4 IPMP Components FDIS, France, October, 2005

5. Ohm, Jens-Rainer., Advances in Scalable Video Coding, Proceedings of the IEEE, Vol. 93, No. 1, January, 2005

6. MPEG System Subgroup, ISO/IEC JTC 1/SC 29/ WG 11/N6596-B: ISO Based Media File Format incorporating amendment 1, USA, June, 2004