Chapter 2
Review of Related Works and MPEG-4 Standard

2.1 Review of Techniques for Visible Watermarking in Video

Lots of techniques related to visible watermarking of images [2-4] and videos [1] for copyright protection have been proposed in recent years. Mohanty et al. [3] proposed a DCT-domain visible watermarking technique for images. In their method, a mathematical model developed by exploiting the texture sensitivity of the human visual system (HVS) is utilized to embed visible watermarks in the DCT coefficients. Meng et al. [1] proposed a content-based method for embedding visible watermarks in the compressed domain without complete decoding of the compressed video bitstreams. A unique contribution of their work is to adjust the watermark strength dynamically depending on the local features acquired in the DCT domain. For instance, a brighter or cluttered image block needs a stronger watermark in order to keep a consistent visibility of the watermark.

2.2 Review of Techniques for Active Agent

An active agent is an executable program for performing particular tasks *actively* by itself without receiving any command. Yu et al. [13] proposed an active data hiding technique by hiding an active data stream, such as an applet or an executable file, as an active agent into a music file for copyright protection. But this method costs a lot of information hiding capacity, and an additional extraction program is necessary for extracting and executing the active agent. Another form of active agent proposed by Chuang and Tsai [15] is a program using a script language, such as Actionscript for flash files and Javascript for html files. This kind of active agent is capable of performing some tasks, such as copyright protection of the file, actively without installing an extra extraction program. Besides, it does not need to consider the problem of limited information hiding capacity because these script languages are supported by these files. In this study, we design a particular video player as an active agent in different form for claiming the ownership of the video actively by visible watermarking techniques. Although active agents are proposed with various forms, all of them have the same purpose for vitalizing the cover media.

2.3 Review of Techniques for Video Authentication

The purpose of video authentication techniques is to check whether a video has not been tampered with or not. Numerous methods have been proposed for authenticating videos in recent years. Two representative approaches are the digital signature technique [6] and the digital watermarking technique [7-11]. Schneider et al. [6] proposed an idea using the image content to generate a digital signature for protecting the authenticity of images and surviving acceptable compression. Moreover, an extension of the authentication system to videos with digital signatures is proposed by a two-level hashing technique. He et al. [7] proposed an object based watermarking solution for MPEG-4 video authentication. A watermark is embedded into sets of discrete Fourier transform (DFT) coefficients in the low-middle frequency by comparing the energy relationship between two divided sub-groups in each set. Besides the two approaches mentioned above, Dittmann et al. [12] proposed another approach which combines content-based digital signatures and digital watermarks to verify the integrity of video data and to search for manipulated images.

2.4 Review of Techniques for Video Data Hiding

Techniques for video data hiding are developed for embedding data into a video. In this way, data can be transmitted covertly to receiver sites or the video can be indexed with specific information. There are lots of approaches to hiding data into a video [16-19]. Hartung et al. [17] proposed methods for embedding additive watermark signals into uncompressed and compressed video sequences. In the method of embedding signals into uncompressed video sequences, the basic idea is borrowed from spread spectrum communications: a narrow-band signal (the watermark) has to be transmitted via a wide-band channel with interference (the image or video signal). In some other methods, the non-zero DCT coefficients are utilized for embedding data. Giannoula et al. [18] proposed a compressive data hiding techniques to hide both the chrominance components and the DCT coefficients of certain segmented audio signals into the luminance component on a frame-by-frame basis.

2.5 Review of MPEG-4 Standard

In this study, all the proposed methods employ MPEG-4 videos as cover media for hiding data. We give a brief review of the MPEG-4 standard in this section. In Section 2.4.1, a brief description of the MPEG-4 standard is presented. In Section 2.4.2, the structure of an MPEG-4 video bitstream is described.

2.5.1 Brief Descriptions of MPEG-4 Standard

MPEG-4 is an ISO/IEC standard developed by the MPEG (Moving Picture Experts Group), the committee that also developed the standards known as MPEG-1 and MPEG-2. Unlike previous standards, MPEG-4 provides a novel concept of object description. An MPEG-4 audiovisual scene may be composed of several interactive media objects, such as still images, video objects and audio objects. Moreover, MPEG-4 also provides a set of tools, including shape coding, motion estimation and compensation, texture coding, sprite coding and scalability for supporting and enhancing applications with different requirements. However, not all of these tools are implemented in most applications due to the associate overhead or the difficulty of generating video objects. Therefore, the conventional rectangular video frame which represents a simple case of an arbitrary-shaped object can be coded as a video object in the MPEG-4 visual scene.

MPEG-4 consists of many parts, where each part standardizes various entities related to multimedia, such as video (part 2 and part 10), audio (part 3), file formats (part 12, 14 and 15). MPEG-4 Part 2 (Visual), which is simply referred as MPEG-4 video in many documents, is a video compression technology and essentially still utilizes two important techniques, which are adopted in previous standards such as MPEG-1 and MPEG-2, to improve the compression effect. The first is to adopt a block-based DCT method for reducing spatial redundancy. The second is to utilize motion compensation to reduce temporal redundancy. In order to produce or display the MPEG-4 video bitstream, a *codec* for MPEG-4 videos is needed. The codec may be a software or hardware capable of performing encoding and decoding process on a digital data stream. In this study, we adopt the video codec named XviD, which is commonly used for compression of MPEG-4 videos nowadays, to implement the proposed methods.

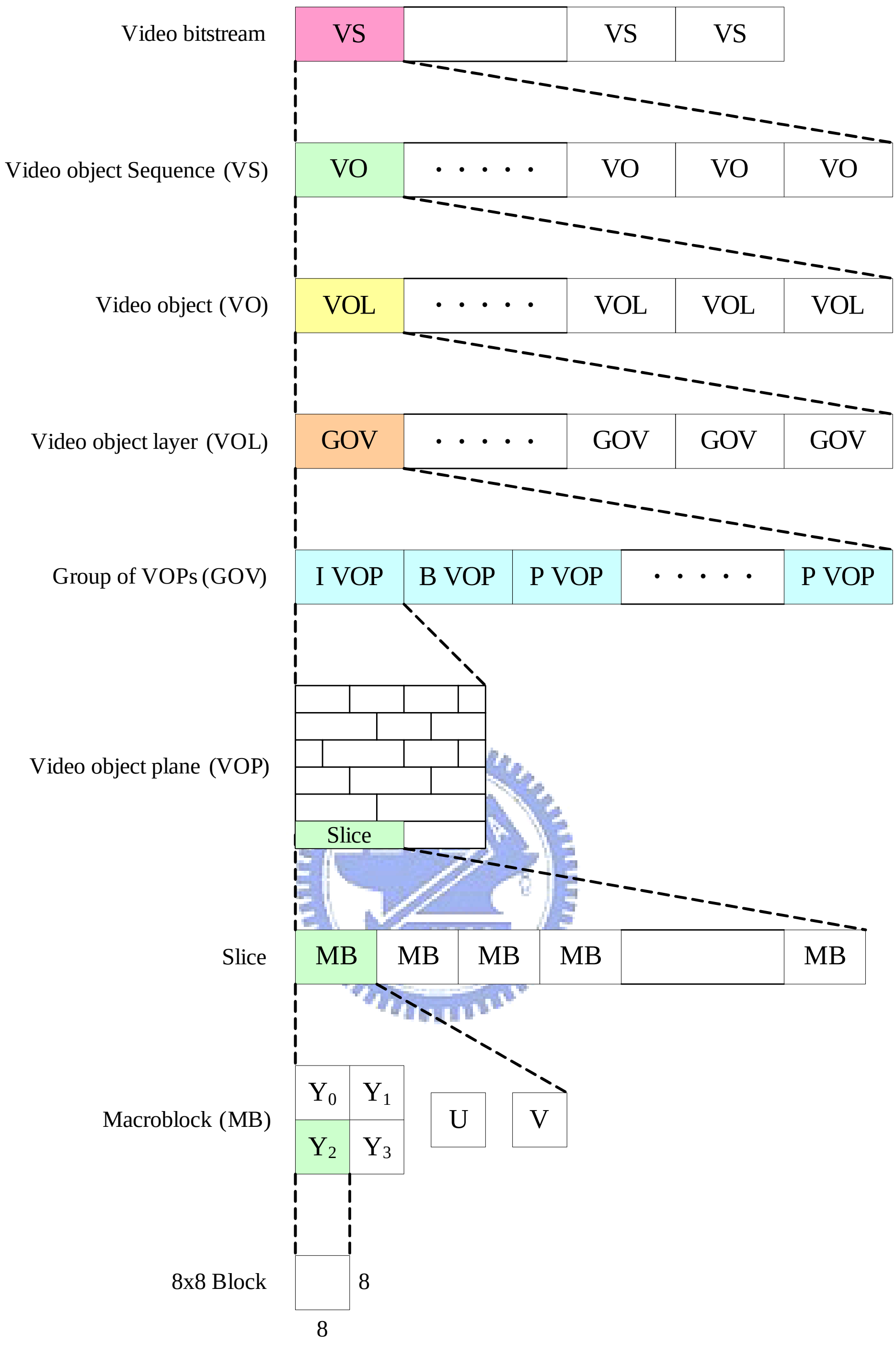

Figure 2.1 Hierarchical structure of MPEG-4 video bitstream.

2.5.2 Structure of MPEG-4 Video Bitstream

An MPEG-4 visual bitstream provides a hierarchical structure as illustrated in Figure 2.1. Each layer of the structure can be accessed in the bitstream by particular code values called start codes. A video bitstream is composed of a series of video object sequences (VS). A VS consists of one or more video objects (VO), and each

VO can be encoded in scalable and non-scalable form, depending on the application, represented by the video object layer (VOL). Each VO is sampled in time, each time sample of a VO is a video object plane (VOP). VOPs can be encoded independently of each other, or dependent on each other by using motion compensation. In addition, they can be grouped together to form a group of video object planes (GOV), which is optional in the standard. A GOV consists of a set of different types of VOPs, including intra-coded VOP (I VOP), predictive-coded VOP (P VOP) and bi-directionally predictive-coded VOP (B VOP). Because a VOP with a rectangular shape can represent a conventional video frame, we can treat these three types of VOPs respectively as I frames, P frames and B frames. A VOP is divided into a series slices, and a slices comprise a number of macroblocks. Finally, a block layer is formed by the luminance and chrominance blocks of a macroblock.