Chapter 12
MPEG Video Coding II
— MPEG-4, 7 and Beyond

Li & Drew ©Prentice Hall 2003

12.1 Overview of MPEG-4

- **MPEG-4**: a newer standard. Besides compression, pays great attention to issues about user interactivities.

- MPEG-4 departs from its predecessors in adopting a new **object-based coding**:

  - Offering higher compression ratio, also beneficial for digital video composition, manipulation, indexing, and retrieval.

  - Figure 12.1 illustrates how MPEG-4 videos can be composed and manipulated by simple operations on the visual objects.

- The bit-rate for MPEG-4 video now covers a large range between 5 kbps to 10 Mbps.
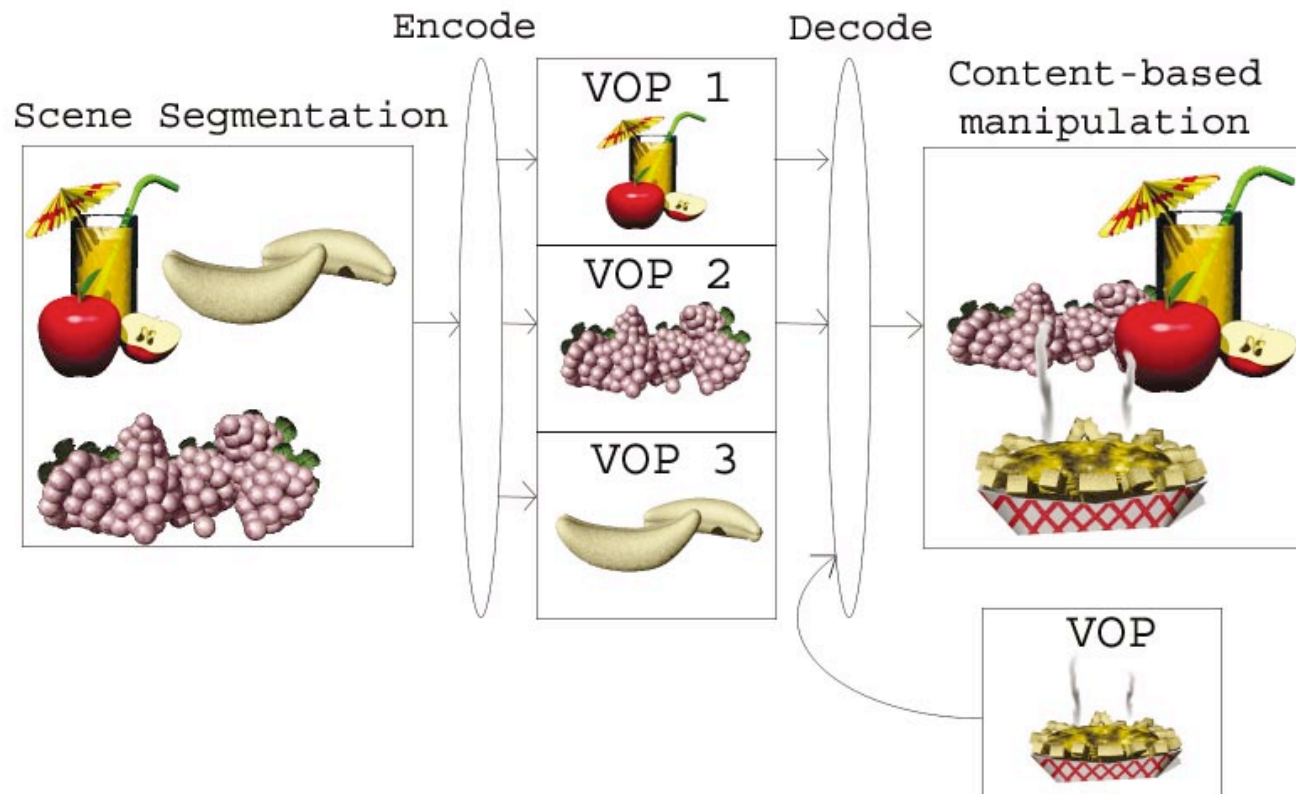
Fig. 12.1: Composition and Manipulation of MPEG-4 Videos.

Li & Drew ©Prentice Hall 2003

Overview of MPEG-4 (Cont'd)

- MPEG-4 (Fig. 12.2(b)) is an entirely new standard for:

(a) Composing media objects to create desirable audiovisual scenes.

(b) Multiplexing and synchronizing the bitstreams for these media data entities so that they can be transmitted with guaranteed Quality of Service (QoS).

(c) Interacting with the audiovisual scene at the receiving end — provides a toolbox of advanced coding modules and algorithms for audio and video compressions.

Li & Drew ©Prentice Hall 2003

Overview of MPEG-4 (Cont'd)

- The hierarchical structure of MPEG-4 visual bitstreams is very different from that of MPEG-1 and -2, it is very much video object-oriented.
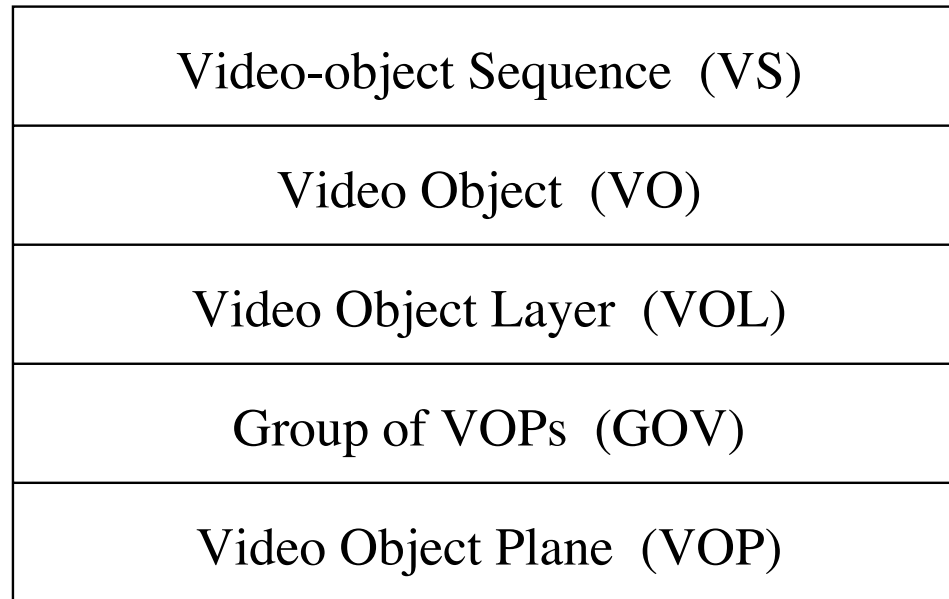
Video-object Sequence (VS)
Video Object (VO)
Video Object Layer (VOL)
Group of VOPs (GOV)
Video Object Plane (VOP)

Fig. 12.3: Video Object Oriented Hierarchical Description of a Scene in MPEG-4 Visual Bitstreams.

Overview of MPEG-4 (Cont'd)

1. **Video-object Sequence** (VS) — delivers the complete MPEG-4 visual scene, which may contain 2-D or 3-D natural or synthetic objects.

2. **Video Object** (VO) — a particular object in the scene, which can be of arbitrary (non-rectangular) shape corresponding to an object or background of the scene.

3. **Video Object Layer** (VOL) — facilitates a way to support (multi-layered) scalable coding. A VO can have multiple VOLs under scalable coding, or have a single VOL under non-scalable coding.

4. **Group of Video Object Planes** (GOV) — groups Video Object Planes together (optional level).

5. **Video Object Plane** (VOP) — a snapshot of a VO at a particular moment.

12.2 Object-based Visual Coding in MPEG-4

VOP-based vs. Frame-based Coding

- MPEG-1 and -2 do not support the VOP concept, and hence their coding method is referred to as **frame-based** (also known as **Block-based coding**).

- Fig. 12.4 (c) illustrates a possible example in which both potential matches yield small prediction errors for block-based coding.

- Fig. 12.4 (d) shows that each VOP is of arbitrary shape and ideally will obtain a unique motion vector consistent with the actual object motion.
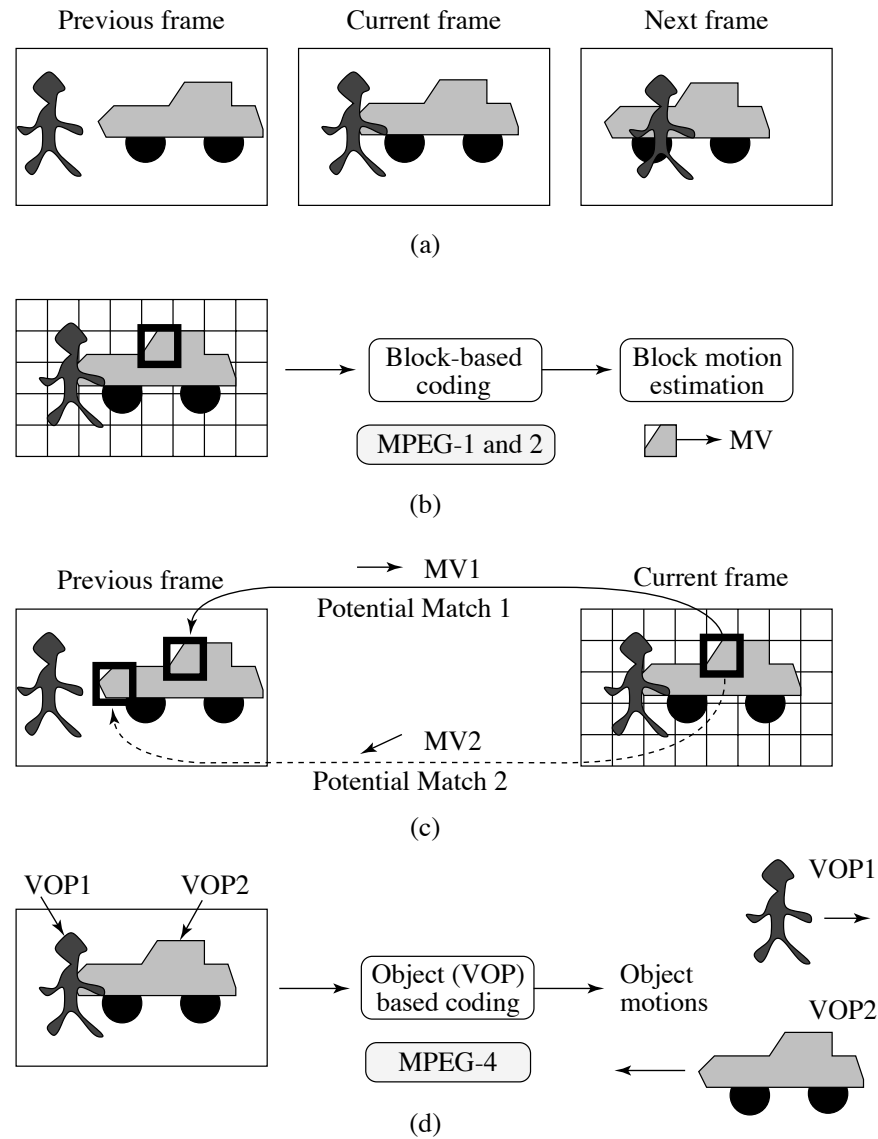
Fig. 12.4: Comparison between Block-based Coding and Object-based Coding.

VOP-based Coding

- MPEG-4 VOP-based coding also employs the Motion Compensation technique:

  - An Intra-frame coded VOP is called an **I-VOP**.

  - The Inter-frame coded VOPs are called *P-VOP*s if only forward prediction is employed, or *B-VOP*s if bi-directional predictions are employed.

- The new difficulty for VOPs: may have arbitrary shapes, shape information must be coded in addition to the texture of the VOP.

  Note: *texture* here actually refers to the visual content, that is the gray-level (or chroma) values of the pixels in the VOP.

12.6 MPEG-7

- The main objective of MPEG-7 is to serve the need of audiovisual content-based retrieval (or audiovisual object retrieval) in applications such as digital libraries.

- Nevertheless, it is also applicable to any multimedia applications involving the generation (*content creation*) and usage (*content consumption*) of multimedia data.

- MPEG-7 became an International Standard in September 2001 — with the formal name **Multimedia Content Description Interface**.

Applications Supported by MPEG-7

- MPEG-7 supports a variety of multimedia applications. Its data may include still pictures, graphics, 3D models, audio, speech, video, and composition information (how to combine these elements).

- These MPEG-7 data elements can be represented in textual format, or binary format, or both.

- Fig. 12.17 illustrates some possible applications that will benefit from the MPEG-7 standard.