Video Coding Standards

- Introduction
- Video Coding Standards
- MPEG-4
- H.264
- HEVC
Introduction

• Why have standards?
  - To ensure interoperability amongst different manufacturers and to encourage interworking, competition and increased choice.

• What is a video coding standard?
  - It defines a coded representation (or syntax) that describes visual data in a compressed form, and a method of decoding the syntax to reconstruct the visual information.
  - It aims to ensure that compliant encoders and decoders can successfully interwork with each other, whilst allowing manufacturers the freedom to develop competitive and innovative products.
Introduction

- The standards do not define the encoder; rather they define the output that an encoder should produce.
- Each standard defines a decoding method but the manufacturers are free to develop alternative decoders as long as they can decode the syntax, and produce the same result as that in the standard.
Video Coding Standards

- **ITU-T Recommendation H.261**
  - Video codec for audiovisual services at $p \times 64$ Kbits/s

- **ISO/IEC 11172 (MPEG-1)**
  - Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbits/s

- **ITU-T Recommendation H.262 or ISO/IEC 13818 (MPEG-2)**
  - Generic coding of moving pictures and associated audio information – Part 2: Video
Video Coding Standards

- **ITU-T Recommendation H.263**
  - Video coding for low bit rate communication

- **ISO/IEC 14496 (MPEG-4)**
  - Coding of audiovisual objects – Part 2: Visual

- **ITU-T Recommendation H.264**
  - MPEG-4 Part 10: Advanced video coding

- **ISO/IEC 23008 (MPEG-H)**
  - High efficiency coding and media delivery in heterogeneous environments
  - Part 2: High Efficiency Video Coding (HEVC)
Related Standards

- **AVS** – Audio Video coding Standard
  - China’s national audiovisual coding standard

- **AVS-M**
  - Audio video coding standard for mobile phones

- **ISO/IEC 15938 (MPEG-7)**
  - Multimedia content description interface

- **ISO/IEC 18034 (MPEG-21)**
  - Multimedia framework
Features of MPEG-4

• Efficient compression of progressive and interlaced natural (rectangular shape) video sequences. The core compression tools are based on the ITU-T H.263 standard.

• Coding of arbitrarily-shaped video objects for object-based video applications in which a video scene may be composed of a number of distinct video objects, each independently coded.

• Support for effective transmission over practical networks using error resilient and scalable coding tools.
Composition of Audiovisual Objects (AVOs)

• MPEG-4 provides a standardized way to describe a scene, allowing the user to:
  - place AVOs anywhere in a given coordinate system;
  - apply transforms to change the geometrical or acoustical appearance of a AVO;
  - group primitive AVOs in order to form compound media objects;
  - apply streamed data to AVOs, in order to modify their attributes;
  - change interactively the user’s viewing and listening points anywhere in the scene.

• With reference to Fig. 4.1, for example, one can replace the person with a different person, changes her dress or hairstyle; group the desk and the globe to form a compound AVO since they are static; or change the background using a different sprite.
Fig. 4.1
An MPEG-4 audiovisual scene.
Video Objects

- MPEG-4 treats a video sequence as a collection of video objects.
- A video object (VO) is an area of video scene that may occupy an arbitrary-shaped region and may exist for an arbitrary length of time.
- An instance of a VO at a particular point in time is a video object plane (VOP).
- In the traditional video coding sense, a rectangular video frame is a VOP and a video sequence is a VO.
Examples of Video Object Plane

VOP1

VOP2

VOP3

VOP4
Fig. 4.2 VOP formation.
MPEG-4 Encoder

Fig. 4.3
MPEG-4 encoder.
Fig. 4.4 VOP prediction.
Macorblock Structure

Fig. 4.5  Macroblock structure (4:2:0).
## MPEG-4 Profiles

<table>
<thead>
<tr>
<th>Profile</th>
<th>Coding Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Simple Profile</strong></td>
<td>I-VOP, P-VOP, 4MV, UMV, Intra prediction, Video packets, Data Partitioning, RVLCs, Short Header (compatible with H.263 Baseline)</td>
</tr>
<tr>
<td>Coding of rectangular video frames</td>
<td></td>
</tr>
<tr>
<td><strong>Core Profile</strong></td>
<td>Simple coding tools, B-VOP, Alternate quantizer, Binary shape, Shape adaptive DCT</td>
</tr>
<tr>
<td>Coding of arbitrary-shaped video objects</td>
<td></td>
</tr>
<tr>
<td><strong>Scalable Profile</strong></td>
<td>Simple coding tools, Temporal scalability, Spatial scalability, Fine granular scalability, Object-based spatial scalability</td>
</tr>
<tr>
<td>Scalable coding of rectangular video frames or video objects</td>
<td></td>
</tr>
</tbody>
</table>