draft-wenger-avt-rtp-jvt-00.txt

RTP Payload for JVT Video

Stephan Wenger
stewe@cs.tu-berlin.de

Overview

- What is JVT Video
- VCL and NAL
- The ERPS Concept
- Open Issues
  - Semantic of the RTP Timestamp
  - Semantic of the Marker Bit
  - (Compound Packets vs. RTP Multiplexing)

What is JVT Video

- New video coding from
  - ITU-T VCEG (the guys who did H.263)
  - MPEG-Video sub-group
- Status of the work
  - Technically frozen spec hopefully in May
    - Input from AVT still possible
  - Standard by late 2002/early 2003
  - Products shortly thereafter

What is JVT Video? VCL

- VCL = Video Coding Layer
  - "Normal" Video Codec
    - 4x4 blocks, many macroblock shapes, affine motion, …, PLUS
    - Enhanced Reference Picture Selection (ERPS)
    - Data Partitioning
  - Output of the VCL is a Slice
    - Consists of one or three Partitions
  - Slices are independently decodable

What is JVT Video? NAL

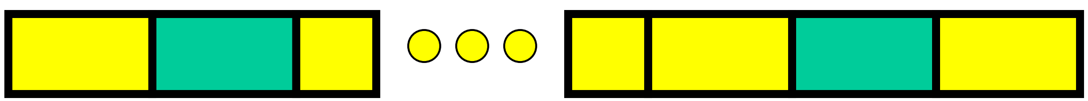
- NAL = Network Adaptation Layer
  - Maps Slices to network/mux specifics
  - Generates/consumes NALPs
    - Network Adaptation Layer Packets
  - Bit-oriented networks: start codes / framing
  - Packet networks: packetization scheme
  - See I-D for NALP types

The ERPS Concept

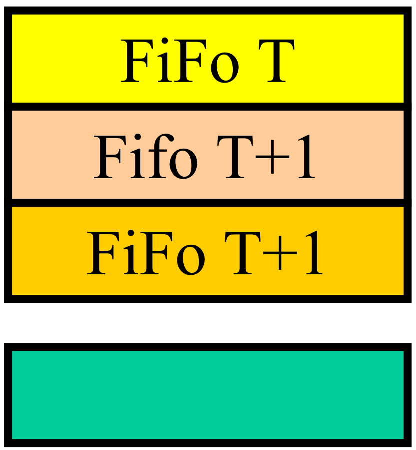
- Three-dimensional Motion Vectors
  - Coding Efficiency, Error Resilience, special
  - Bit stream addresses Reference Buffer
  - Reference Pictures are “stored” in a Reference Buffer
    - Normally FIFO
    - Explicit Addressing possible

The ERPS Concept

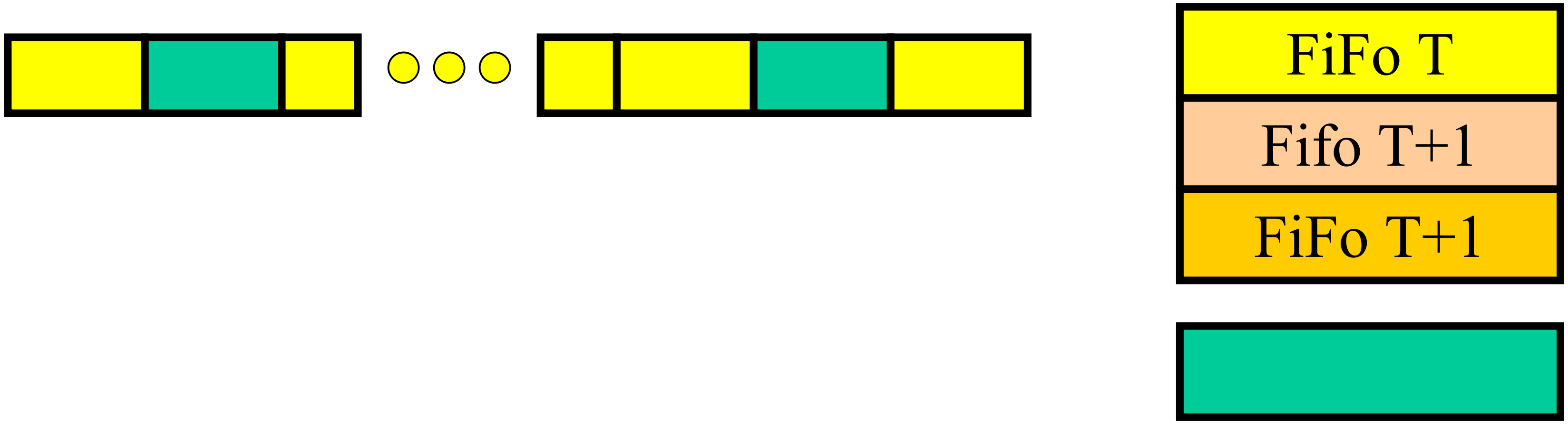
- **Explicit Addressing (write to Reference Picture)**



Reference Pics



- Packet Stream sets up green RefPic
- Unclear, when this will be used (if ever)



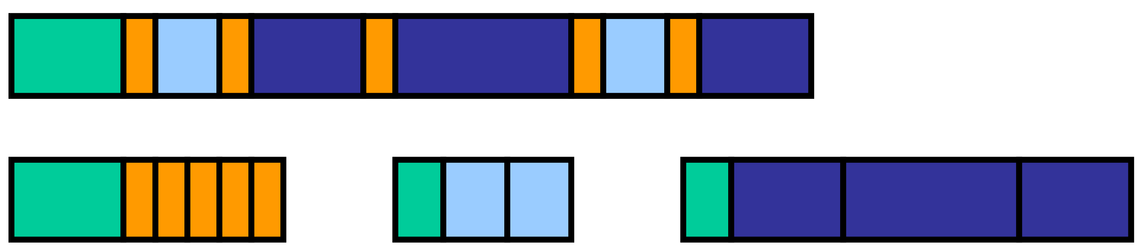
- **Yellow: Live Football Game**
- **Green: Slices of Commercial**
  - **Included in the packet stream whenever bit rate is available**
  - **To be display some (unknown) time in the future, by simply referencing it**
    - **Inter Picture w/ (0,0) MV, no coefficients**

Data Partitioning

Header Data (MBTYPE, MVs, …)	most important
Intra CBPs and Coefficients	less important
Inter CBPs and Coefficients	least important
Slice/Partition Header	

SSP NALP

DPA/DPB/DPC



Open Issues

- Semantic of the RTP Timestamp
- Semantic of the Marker Bit
- (Compound Packets)

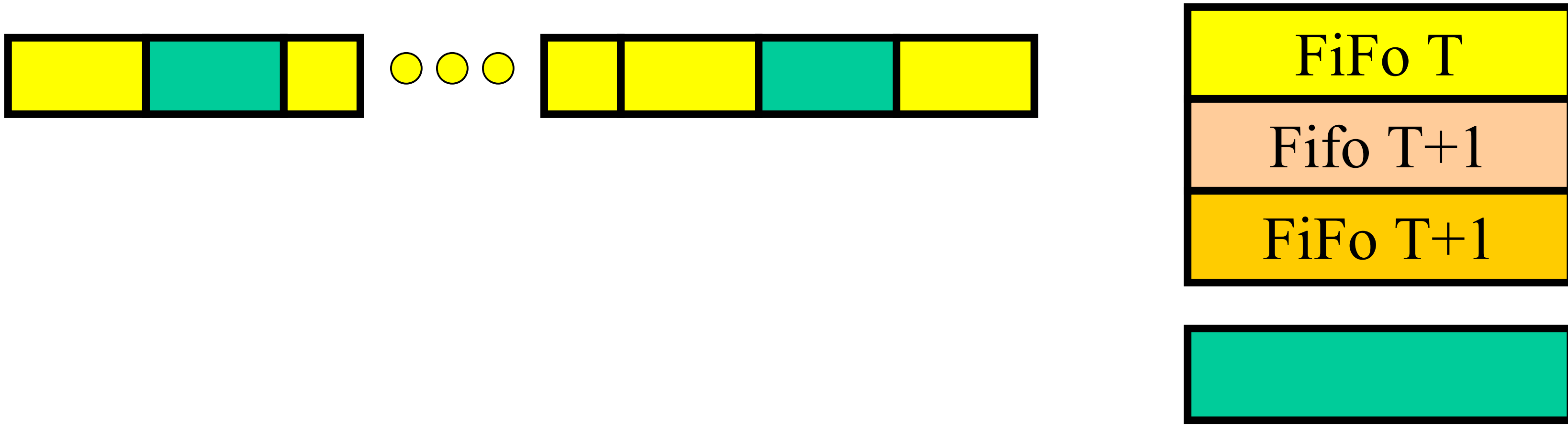
RTP Timestamp 1/3

- **Can't use Sampling Instant**
  - **Unknown for the Commercial**
- **Can't use Presentation Time**
  - **Unknown for the commercial**
  - **Jitter Buffer problematic (avt@ietf.org)**
- **Decoding Timestamp**
  - **Decode a Slice? A Picture? Tricky!**

RTP Timestamp 2/3

- **Proposal: Sending Time of RTP packet**
- **Pro: Perfect Jitter Buffer handling**
- **Contra: Media synchronization not possible**
  - **BUT: does this term make sense for JVT anyway?**
  - **Yes, for the yellow packets, no for the green packets**

RTP Timestamp 3/3



- **How do you want to "synchronize" the green packets with anything?**
- **Conceptually, JVT Video could consist ONLY of green packets**

Marker Bit

- Traditional: M bit set for last slice of a picture
- JVT: No true picture concept!
  - Could keep it as is
  - Could change it to last partition of a slice

(Compound Packet 1/2)

- **Build-in multiplex tool**
  - Not only for RTP, but
  - We could (if we really want) disallow its use over RTP
- **Low overhead (2 bytes per NALP)**
- **Intended for**
  - MTU size change (wireless/wireline gateway)
  - Combining different NALPs that "belong" together

(Compound Packet Examples)

- **Combine DPAs of a picture together to save IP/UDP/RTP header bits**
  - **And to protect them**
- **Combine 10 wireless SSPs (100 bytes each) to one 1 Kbyte RTP payload**
- **Combine the intra coefficients of 10 pictures together in one packet and protect them**