Workshop and Exhibition on MPEG-4 2002

AVC/H.264 Generalized B Pictures

Markus Flierl
Telecommunications Laboratory
University of Erlangen-Nuremberg
mflierl@stanford.edu

Bernd Girod
Information Systems Laboratory
Stanford University
bgirod@stanford.edu
Classic B Pictures

- Linear combination of 2 motion-compensated signals
- Chosen from 2 fix reference pictures
- B pictures are not used as reference pictures
Multiframe Motion Compensation

- Select one prediction signal from the reference frame buffer
- Reference frame buffer contains previously decoded pictures
Multihypothesis Multiframe Motion Compensation

- Select *more than one* prediction signal from the reference frame buffer
- Average the prediction signals
- Reference frame buffer contains previously decoded pictures
- Multihypothesis pictures are used as reference pictures
Overview

- AVC/H.264 Motion Compensation
- Bi-directional prediction mode
- Multihypothesis prediction mode
- Two combined forward prediction signals
- AVC/H.264 Entropy coding
- Current joint committee draft
AVC/H.264 Motion Compensation

⇒ Picture reference parameter per macroblock

⇒ Block size data

\[
\begin{array}{c|c|c|c}
& 16 \times 16 & 8 \times 16 & 16 \times 8 & 8 \times 8 \\
0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 \\
2 & 2 & 1 & 2 & 2 \\
3 & 3 & 2 & 3 & 3 \\
4 & 4 & 3 & 4 & 4 \\
5 & 5 & 4 & 5 & 5 \\
6 & 6 & 5 & 6 & 6 \\
7 & 7 & 6 & 7 & 7 \\
\end{array}
\]

⇒ Motion vector data
A bi-directional prediction mode allows a linear combination of one past and one subsequent macroblock prediction signal.
The multihypothesis mode allows additionally a linear combination of two past macroblock prediction signals.
Motion Estimation for Bi-Directional Prediction

- **Method 1: Independent Estimation**
  - Multiframe motion estimation for each prediction signal independent of the complementary signal
  - Low computational complexity

- **Method 2: Joint Estimation**
  - Iterative algorithm with conditional multiframe motion estimation
  - Improved prediction performance
Multihypothesis Motion Estimation

Iterative Algorithm:

- Conditional rate-constrained motion estimation

- Determines conditional optimal hypothesis (= picture reference parameter, block size type, and motion vector data)

- Converges practically after 2 iterations

Mobile & Calendar

CIF, 30 fps, \( M = 5 \)
**Performance Comparison I**

B picture performance for the CIF sequence *Mobile & Calendar*, 30 fps

Multihypothesis mode outperforms bi-directional mode

**IBBPBBP...**

\[ QP_B = QP_P \]
**Performance Comparison II**

B picture performance for the CIF sequence *Mobile & Calendar*, 30 fps

Multihypothesis mode is more general than bi-directional mode

IBBPBBP...

\[ QP_B = QP_P \]
• Generalized B pictures with forward-only prediction

• Use generalized B pictures like P pictures and allow only forward prediction for the multihypothesis mode.
Avg. bit-rate at 35 dB PSNR for the CIF sequence *Mobile & Calendar*, 30 fps
A large reference frame buffer improves performance of generalized B pictures.
Draft standard permits two entropy coding schemes:

- Universal Variable Length Coding (UVLC)
- Context-Based Adaptive Binary Arithmetic Coding (CABAC)

For UVLC, the motion vectors are encoded independently.

CABAC allows a context model which considers the motion vector of the complementary prediction signal.

Multihypothesis prediction mode and CABAC can be successfully combined.
Efficient entropy coding improves performance of generalized B pictures.

B picture performance for the CIF sequence *Flowergarden*, 30 fps

$Q_{P_B} = Q_{P_P} + 2$

Reference frames:
- 5 past
- 3 future
Generalized B pictures in AVC/H.264 main profile

B pictures can be reference for other B pictures in decoding-sequence-order

Two reference frame buffers that are conceptually separate

Bi-predictive mode uses the two buffers such that each prediction signal is chosen from its reference frame buffer

The two reference frame buffer may share the same pictures (dependent on the reference frame buffer control) and consequently permit multihypothesis prediction
Summary and Conclusions

- Generalized B pictures differentiate between
  - picture reference selection and
  - linearly combined prediction signals

- Bi-predictive mode is more general than the bi-directional mode

- Two reference buffers allow multiframe prediction for each motion-compensated signal

- Multihypothesis prediction benefits from efficient entropy coding with context-based binary arithmetic coding (CABAC)