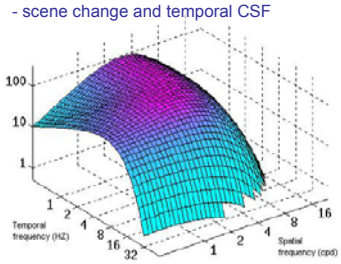
Intra-mode Indexed Non-uniform Quantization Parameter Matrices in AVC/H.264

Jing Hu, Professor Jerry D. Gibson
ViVoNets Lab

Background

- **Advanced Video Coding/ ITU-T Recommendation H.264/ ISO/IEC MPEG-4 (Part 10)**
  - Video coding layer (VCL) offers twice of coding efficiency than MPEG-2
  - Network abstraction layer (NAL) provides "network friendliness"
  - More powerful error resilience tools
- **Psychophysics of HVS**
  - Luminance Masking / Amplitude Nonlinearity / Light Adaptation
    - uniform image and background (Weber's Law: $\Delta I / I = k \approx 0.33$)
  - Baseline Contrast Masking
    - sine-waves on uniform background (Contrast Sensitivity Function (CSF))
  - Texture Masking
    - image components with similar spatial locations and frequency contents
  - Temporal Masking
    - scene change and temporal CSF



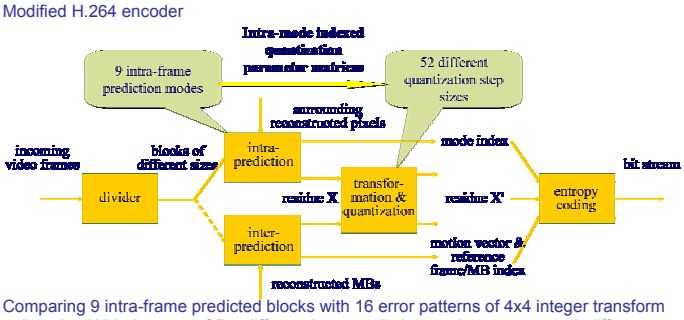
Contrast Sensitivity Function

Example of effects of texture masking

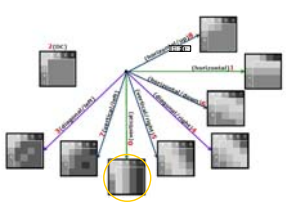
Algorithm

We investigate the visibility of quantization errors in different frequency coefficients after the integer transform, as a function of the intra-mode, local luminance level and frame rate. Two subjective experiments are designed and conducted to explore the luminance, texture and temporal masking of the human vision system (HVS) and the optimal QP matrices are calculated from the data collected in the experiments. The resulting 9 non-uniform quantization parameter (QP) matrices are indexed to the 9 intra-frame prediction modes of the 4x4 blocks and can be applied to the 16x16 macroblock(MB)s.

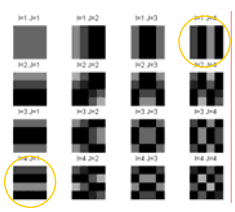
- Modified H.264 encoder



- Comparing 9 intra-frame predicted blocks with 16 error patterns of 4x4 integer transform predicts that With the same QPs, different intra-prediction modes texture-mask different error patterns to different levels.



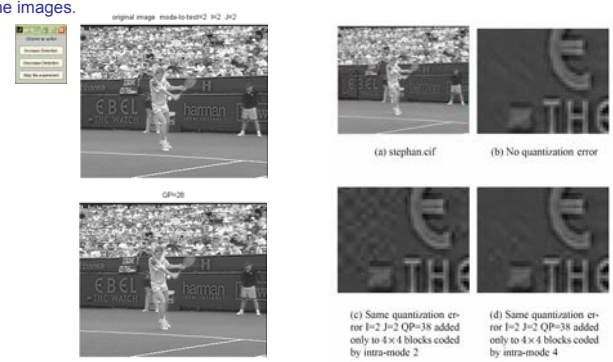
9 intra-frame prediction modes



Normalized error of 4x4 integer transform

Experiments

- **Quantization parameter matrices for the intra-coded MBs (Subjective experiment I)**
  - Only the 4x4 blocks which are intra-coded using one of the 9 intra-prediction modes in the image are chosen at one time to add only one out of the 16 distortion patterns.
  - The quantization parameter(QP)s gradually increase until the distortion becomes perceivable to the subject and the QPs at threshold are recorded.
  - Three images in CIF (352x288 pixels) and three subjects at the age of 20, 25, 31 are involved. A 17-inch monitor with resolution 1152x864 pixels is used and the average background luminance is 170cd/mm^2. The viewing distance is about 6 times the height of the images.

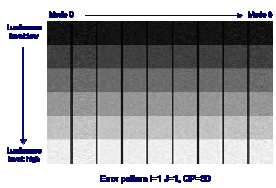


(a) stephan.cif (b) No quantization error

(c) Same quantization error I=2 J=2 QP=38 added only to 4×4 blocks coded by intra-mode 2

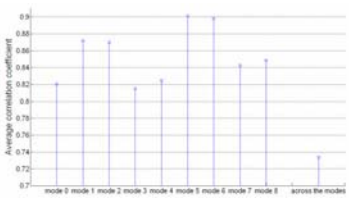
(d) Same quantization error I=2 J=2 QP=38 added only to 4×4 blocks coded by intra-mode 4

- **Quantization parameter matrices for the inter-coded MBs (Subjective experiment II)**
  - An artificial video sequence is generated for each error pattern and a certain QP value.
  - All 9 intra-prediction blocks at 6 different luminance levels are generated as the background image, same for every frame in every video sequence.
  - For each video sequence, only one error pattern at one QP value is added. But this error propagates spatially and temporally by simulating inter-frame motion estimation and compensation.

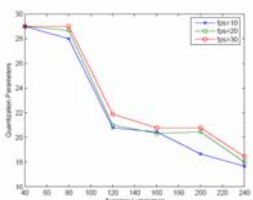


Results

- **Data generated from two subjective experiments**



Averages of correlations among all QP matrices in experiment I



The average QPs for error pattern I=1 J=1 over all the intra-coding modes in experiment II

- **Generated non-uniform QP matrices indexed to intra-prediction modes**

$$QP_0 = \begin{bmatrix} 33 & 33 & 34 & 38 \\ 32 & 34 & 35 & 40 \\ 32 & 33 & 36 & 40 \\ 34 & 36 & 38 & 43 \end{bmatrix}, QP_1 = \begin{bmatrix} 31 & 28 & 30 & 33 \\ 28 & 30 & 31 & 36 \\ 31 & 32 & 34 & 39 \\ 33 & 36 & 38 & 42 \end{bmatrix},$$

$$QP_2 = \begin{bmatrix} 31 & 30 & 31 & 36 \\ 31 & 30 & 32 & 37 \\ 32 & 32 & 35 & 39 \\ 34 & 36 & 37 & 43 \end{bmatrix}, QP_3 = \begin{bmatrix} 31 & 33 & 35 & 37 \\ 33 & 35 & 36 & 40 \\ 34 & 36 & 37 & 40 \\ 36 & 38 & 41 & 44 \end{bmatrix},$$

$$QP_4 = \begin{bmatrix} 35 & 34 & 35 & 38 \\ 36 & 36 & 38 & 42 \\ 36 & 37 & 39 & 44 \\ 38 & 40 & 42 & 46 \end{bmatrix}, QP_5 = \begin{bmatrix} 34 & 34 & 35 & 38 \\ 35 & 36 & 37 & 41 \\ 35 & 37 & 38 & 43 \\ 36 & 39 & 41 & 46 \end{bmatrix},$$

$$QP_6 = \begin{bmatrix} 34 & 34 & 35 & 38 \\ 34 & 35 & 36 & 41 \\ 35 & 36 & 38 & 41 \\ 35 & 36 & 38 & 43 \end{bmatrix}, QP_7 = \begin{bmatrix} 34 & 33 & 34 & 37 \\ 33 & 33 & 35 & 39 \\ 34 & 35 & 37 & 41 \\ 36 & 38 & 39 & 44 \end{bmatrix},$$

$$QP_8 = \begin{bmatrix} 33 & 31 & 33 & 35 \\ 32 & 34 & 34 & 38 \\ 32 & 34 & 34 & 39 \\ 34 & 37 & 39 & 44 \end{bmatrix}.$$

- **Coded video comparison**

(a) Using uniform QP (b) Using nonuniform QP matrices

Sponsors: The authors gratefully acknowledge discussions with Prof. Miguel Eckstein. This work was supported by the California Micro Program, Applied Signal Technology, Dolby Labs, Inc. and Qualcomm, Inc., by NSF Grant Nos. DGE-0221713, Nos. CCF-0429884 and CNS-0435527, and by the UC Discovery Grant Program and Nokia, Inc..