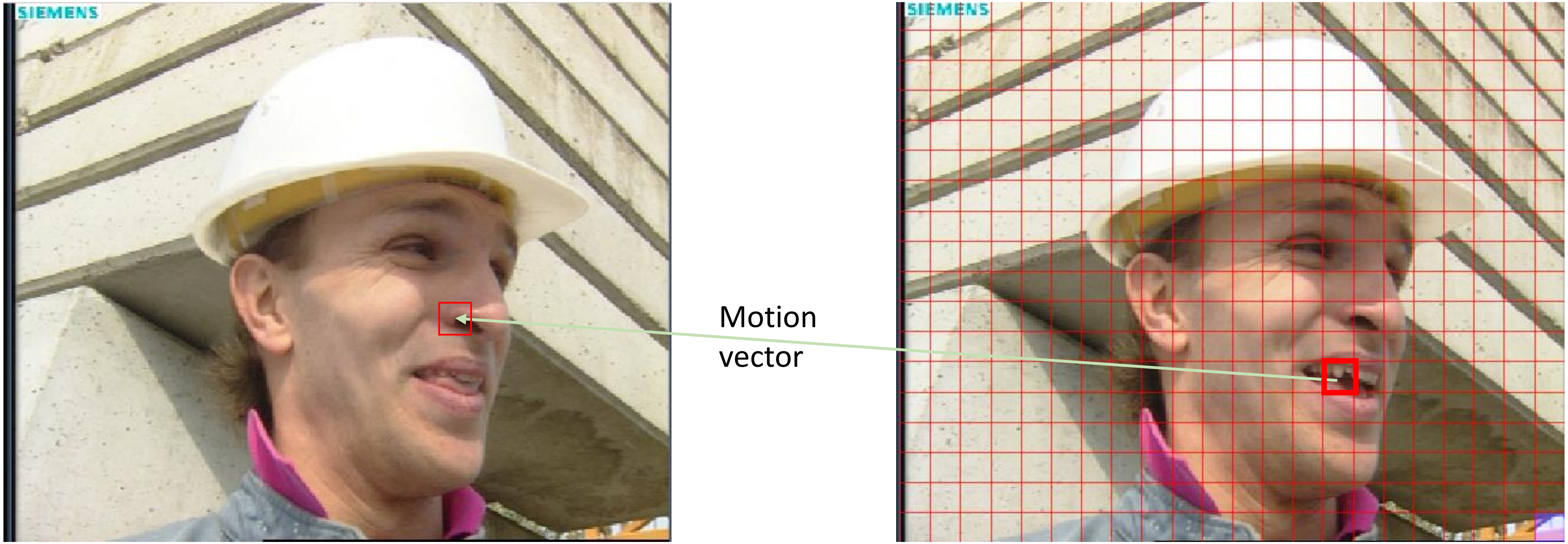
Adaptive interpolated motion compensated prediction

Wei-Ting Lin, Tejaswi Nanundaswamy, Kenneth Rose

Signal Compression Lab

University of California Santa Barbara

Block-based motion compensation



Frame: n-1

Frame: n

Block-based motion compensation

Target

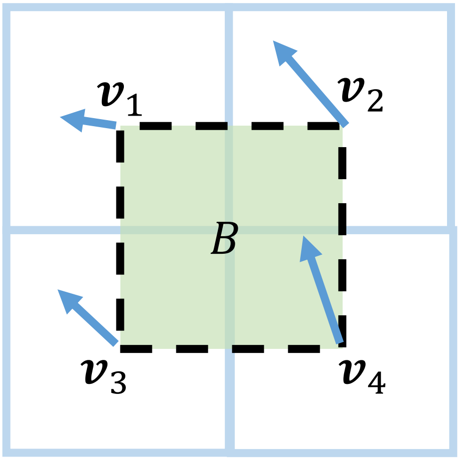
Prediction

Motivation

- Pixel domain block matching to optimize RD cost Pixels in a block move uniformly
  - Motion vector do not necessarily represent the actual motions
  - The influence of a motion vector is restricted within a rectangular block
- We propose to break free the restriction by explicitly treating motion vectors as pointers to multiple estimation sources
  - Using multiple observations to estimate a pixel (based on the nearby motion vectors)
  - Deriving K sets of optimal linear estimation coefficents for predicting each pixel

Prediction Model

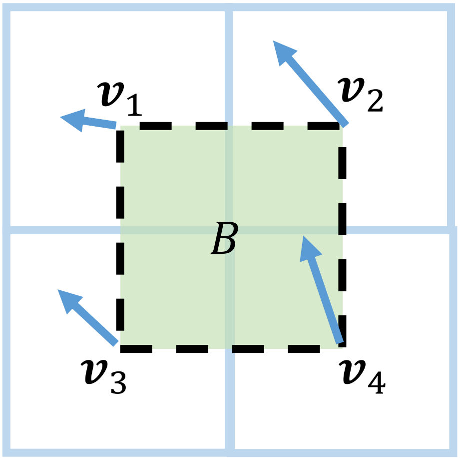
- Use the four closest motion vectors to obtain multiple estimations for each block
- Form the best linear predictor based on the estimations



Grid of motion vectors

Prediction Model

- Use the four closest motion vectors to obtain multiple estimations for each block
- Form the best linear predictor based on the estimations



Grid of motion vectors

$$\tilde{x}_k(\boldsymbol{s}) = \sum_{m=1}^{4} w_m^q(\boldsymbol{s})\hat{x}_{k-1}(\boldsymbol{s}-\boldsymbol{v}_m)$$

$$= \boldsymbol{w}^q(\boldsymbol{s})^T \hat{\boldsymbol{x}}_{k-1}(\boldsymbol{s})$$

$\tilde{x}_k(\boldsymbol{s})$: Prediction of the pixel $\boldsymbol{s}$ in the frame k
$\hat{x}_k(\boldsymbol{s})$: Reconstructed pixel $\boldsymbol{s}$ in the frame k

$w_m^q(\boldsymbol{s})$: The qth set of prediction coefficients for prediction corresponding to the pixel $\boldsymbol{s}$ and the motion vector

Basic Prediction Coefficient Design

Initialize K sets of coefficients

Classify blocks into K clusters in which prediction error is minimized

$$q = \arg \min_{r \in 0,\ldots,K-1} \sum_{s \in B} \left(x_k(\boldsymbol{s}) - \boldsymbol{w}^r(\boldsymbol{s})^T \widehat{\boldsymbol{x}}_{k-1}(\boldsymbol{s}) \right)^2$$

Optimize the coefficients for each cluster C_q to minimize prediction error

$$\boldsymbol{w}^q(\boldsymbol{s}) = \arg \min \sum_{B_{i,j} \in C_q} \sum_{s \in B_{i,j}} \left(x_k(\boldsymbol{s}) - \boldsymbol{w}^q(\boldsymbol{s})^T \widehat{\boldsymbol{x}}_{k-1}(\boldsymbol{s}) \right)^2$$

Refined Prediction Coefficients Design

- Designing coefficients to minimize prediction does not guarantee better reconstruction due to quantization
- We need to design the prediction coefficients accounting for the reconstruction error. The target function can be written as

$$J = \sum_{B_n \in C_q} \sum_{s \in B_n} \left(x_n(\boldsymbol{s}) - \tilde{x}_n^{(t)}(s) - \hat{e}_n^{(t)}(s) \right)^2$$

$\hat{e}_k(s)$: quantized prediction residual

Given the discrete nature of quantization, the above function is piecewise continuous in the prediction coefficients. Sufficiently small changes in the coefficient values will only affect the reconstructed value through the prediction term

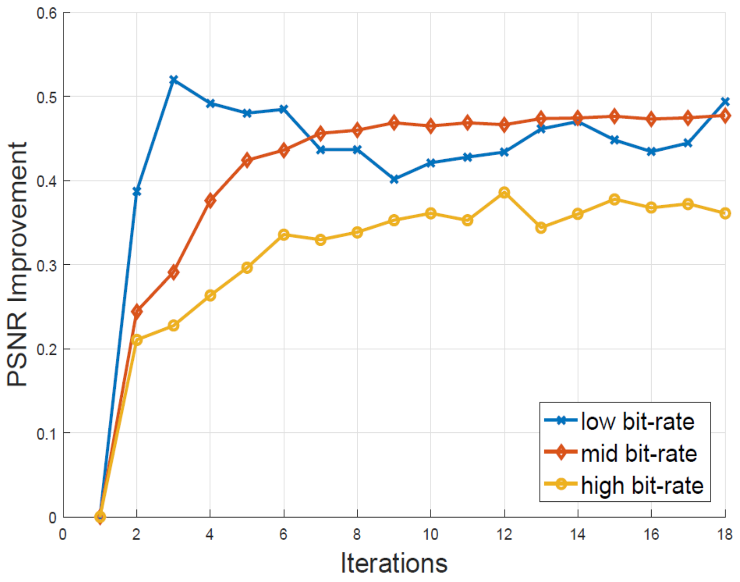
Prediction Coefficient Update Algorithm

Given the set of coefficients at iteration t - 1, a training set of reconstructions $\{\hat{x}_0^{(t)}, \hat{x}_1^{(t)}, \ldots, \hat{x}_N^{(t)}\}$ and quantized prediction error $\{\hat{e}_0^{(t)}, \hat{e}_1^{(t)}, \ldots, \hat{e}_N^{(t)}\}$

Given the training set, minimize

$$J = \sum_{B_n \in C_q} \sum_{s \in B_n} \left(x_n(\mathbf{s}) - \tilde{x}_n^{(t)}(s) - \hat{e}_n^{(t)}(s) \right)^2$$

for each cluster C_q

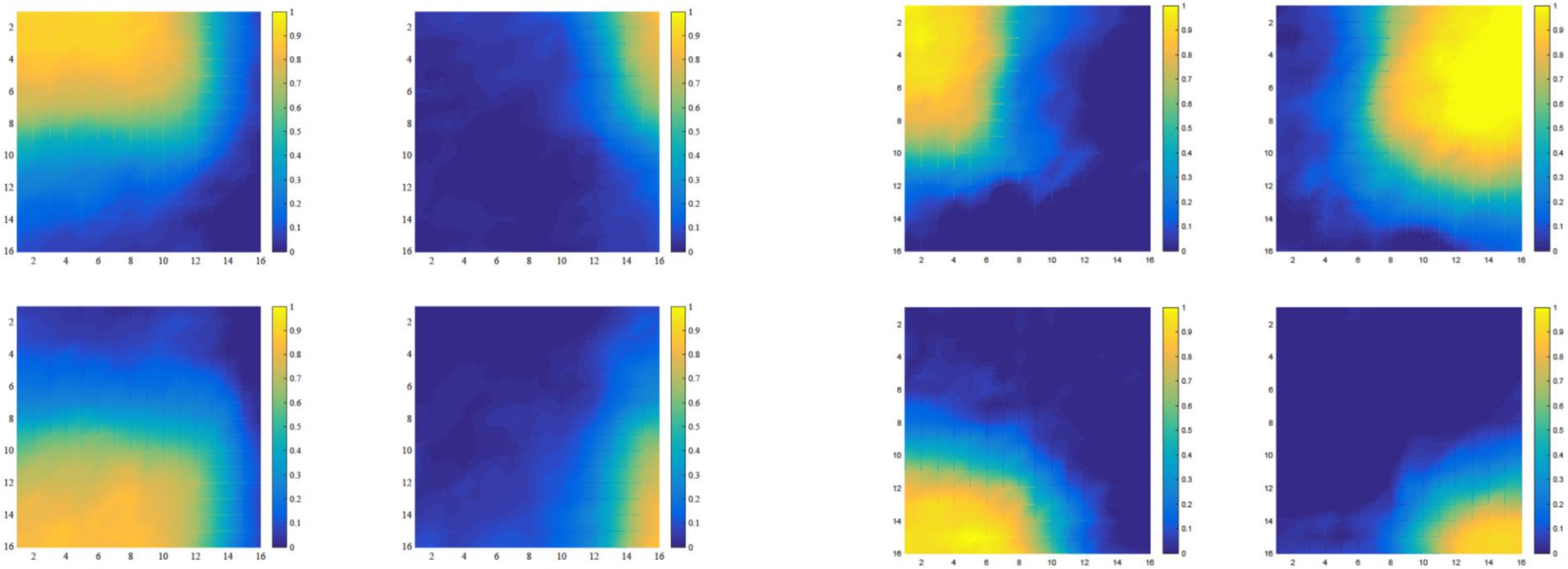


PSNR improvement (in dB) versus iterations of the proposed K-mode clustering algorithm for different target bit rate regions.

Example Sets of Coefficients

The values of coefficients increase when approaching the motion vectors

Allow us to capture the significance of predictions due to neighboring motion vectors



Motion Refinement for Interpolated Prediction

- Motion vectors are optimized independently without considering the interpolation
- Motion update:

1. Calculate the optimal mode for each block $B_{i,j}$ given the motion vectors

2. Fix the modes and $B_{i,j}$'s neighboring block's motion vectors; run motion search to minimize the RD cost

Experimental Results

Prediction quality improvement

Original Prediction

Interpolated Prediction

Experimental Results

Target:

- Side by Side Comparison

Original Prediction

Interpolated Prediction

Experimental Results

- Codebase: VP9
- Coding Structure:
  - IPPP with only the previous frame allowed as reference for inter prediction
  - No intra blocks in inter frame
  - 16x16 fixed block

Sequence	BD-rate Reduction	
	Without motion refinement	With motion refinement
Foreman	11.174	11.316
Bus	13.783	14.455
Ice	6.213	6.863
HighWay	9.500	9.969
BQMall	7.804	7.891
Vidyo4	3.973	4.011
CrowdRun	9.068	9.266
BasketBallDrive	7.746	7.937
Average	8.658	8.964