TRANSFORM DOMAIN TEMPORAL PREDICTION WITH EXTENDED BLOCKS

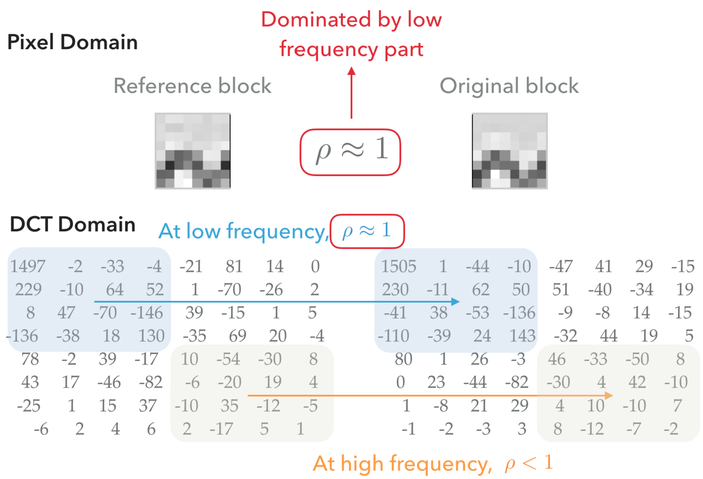
SHUNYAO LI, TEJASWI NANJUNDASWAMY and KENNETH ROSE
SIGNAL COMPRESSION LAB, ECE DEPARTMENT, UNIVERSITY OF CALIFORNIA, SANTA BARBARA

CONTRIBUTION

- We propose the EB-TDTP method to fully account for the TDTP interference with sub-pixel interpolation filter, and spatial correlations outside the reference block boundary.

BACKGROUND: TDTP

- Traditional inter prediction copies pixels one-by-one, which is suboptimal because it ignores spatial correlation
- Transform Domain Temporal Prediction (TDTP): DCT (largely) achieves spatial decorrelation, enabling optimal one-to-one prediction in DCT domain



- Temporal correlation ρ in pixel domain is dominated by the low frequencies ($\rho \approx 1$), inspiring the traditional pixel copying prediction
- TDTP: Accounts for variation in temporal correlation across frequency, which is hidden in pixel domain
- The optimal TDTP predictor for each transform domain coefficient is given by,

$$\tilde{x}_n = \rho \hat{x}_{n-1}$$

- The optimal prediction coefficient is the temporal correlation coefficient, assuming $\hat{x}_n \approx x_n$

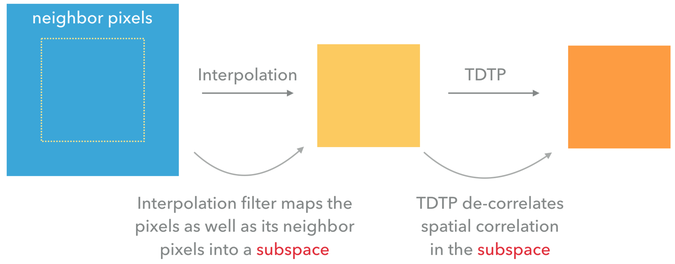
$$\rho = \frac{E(x_n \hat{x}_{n-1})}{E(\hat{x}_{n-1}^2)}$$

CHALLENGE

- TDTP: scales the DCT coefficients by ρ, where at low frequencies $\rho \approx 1$, and at high frequencies $\rho < 1$
- Interferes with the low-pass sub-pixel interpolation filters, which has similar frequency response

EB-TDTP

- TDTP did not completely dis-entangle spatial and temporal correlation, and ignored the correlation outside the block.

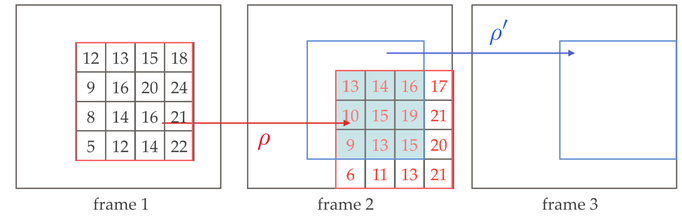


- Proposed: Extended blocks TDTP (EB-TDTP)
- First get the extended reference block with all the neighbor information, de-correlate the spatial correlation using DCT, design the prediction coefficients for each DCT coefficient, then follow the standard sub-pixel interpolation procedure to get the final prediction.
- The prediction coefficients $\mathbf{P}_{B_2}$ is designed to optimize the final prediction error, which can be converted to a linear optimization problem.

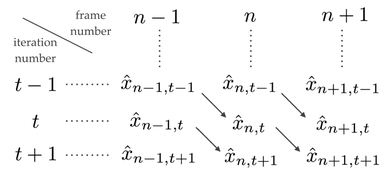
$$\tilde{\mathbf{Y}} = \mathbf{F}_1 \mathbf{D}'_{B_2} \underbrace{(\mathbf{D}_{B_2} \mathbf{X} \mathbf{D}'_{B_2})}_{\text{DCT}} \circ \underbrace{\mathbf{P}_{B_2}}_{\text{EB-TDTP}} \underbrace{\mathbf{D}_{B_2} \mathbf{F}_2}_{\text{Back to pixel domain interpolation}}, \quad min\|\mathbf{Y} - \tilde{\mathbf{Y}}\|^2$$

Asymptotic Closed-Loop (ACL) Design

- Introducing a new predictor results in different statistics, which will propagate and grow over frames through the closed loop video coding operation.
- This deviation in statistics between design and operation makes an effective off-line training impossible.

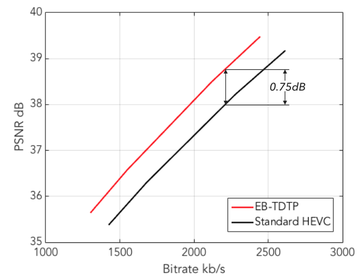


- Thus we use the asymptotic closed-loop (ACL) design approach for training the predictor.



EXPERIMENTAL RESULTS

- The proposed approach was implemented in HM 14.0. Both prediction size and transform size are restricted to 8x8, and the motion search is at half-pixel precision.



RD curve for *coastguard* at *CIF* resolution

Sequence	TDTP	EB-TDTP
Coastguard	8.69%	10.61%
Mobile	11.91%	12.91%
Highway	3.05%	5.44%
Waterfall	9.88%	10.16%
Bus	5.95%	6.70%
Tempete	5.78%	6.19%
RaceHorse	3.48%	4.32%

Reduction in BD rate over HEVC by employing TDTP and EB-TDTP

This work was supported in part by a gift from LG Electronics Inc.
http://www.scl.ece.ucsb.edu/
{shunyao li, tejaswi, rose}@ece.ucsb.edu