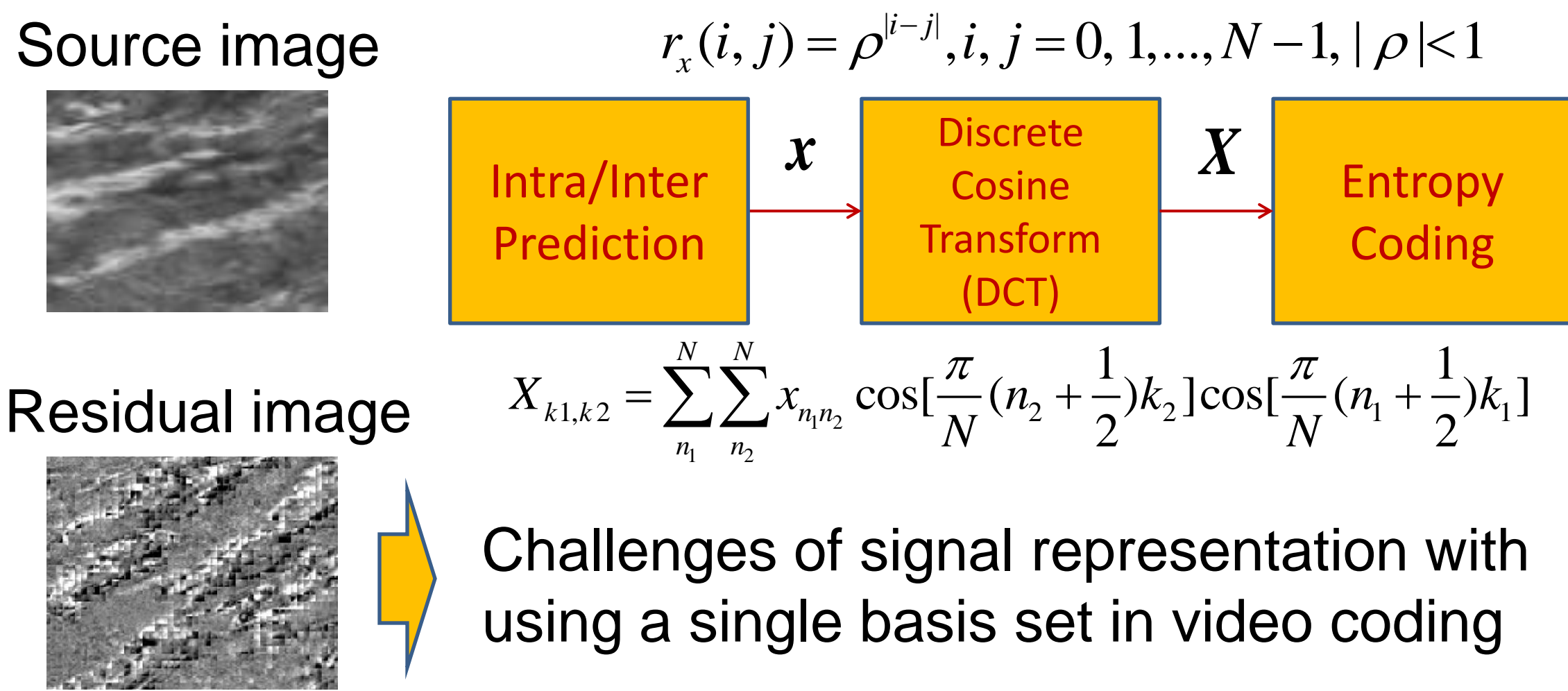


Two Layered Transform with Sparse Representation for Video Coding

Je-won Kang
EE/Multimedia Comm. Lab

Motivation



Main Idea

-Traditional signal representation, e.g. DCT

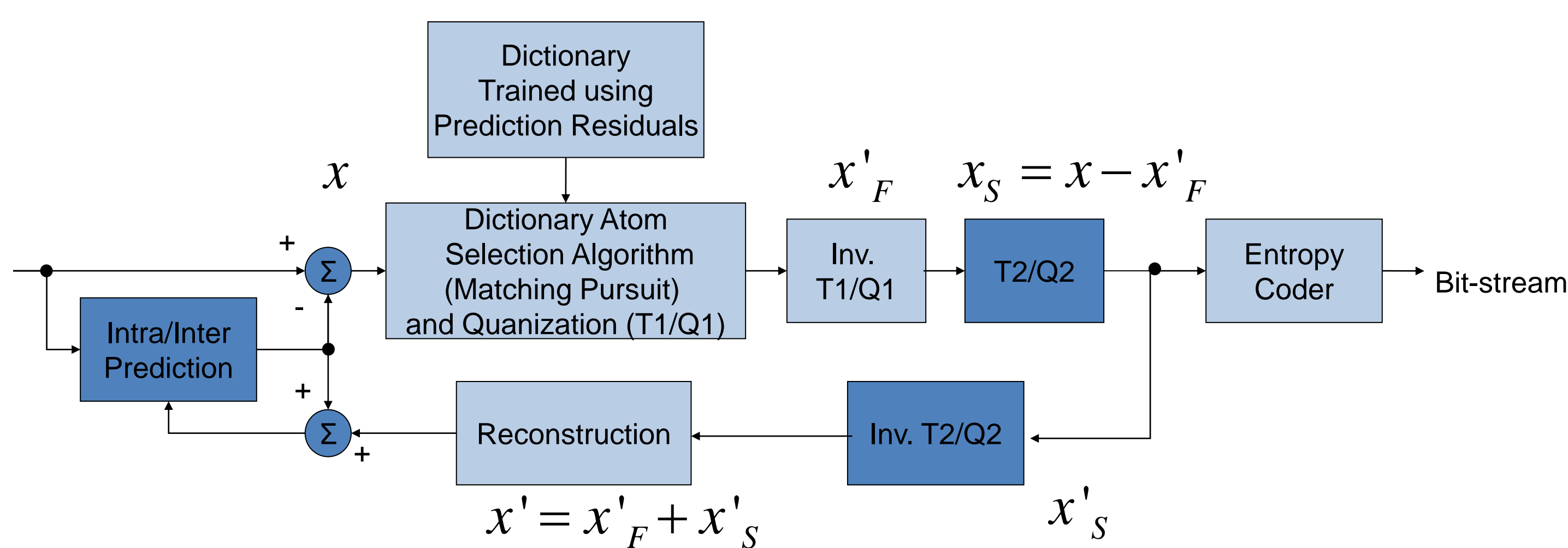
$$x = \sum_{k=0}^{K-1} \beta_k \phi_k \rightarrow \begin{matrix} \square & \square & \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square & \square & \square \end{matrix} \quad (2\text{-D DCT bases})$$

-Proposed scheme using multiple bases

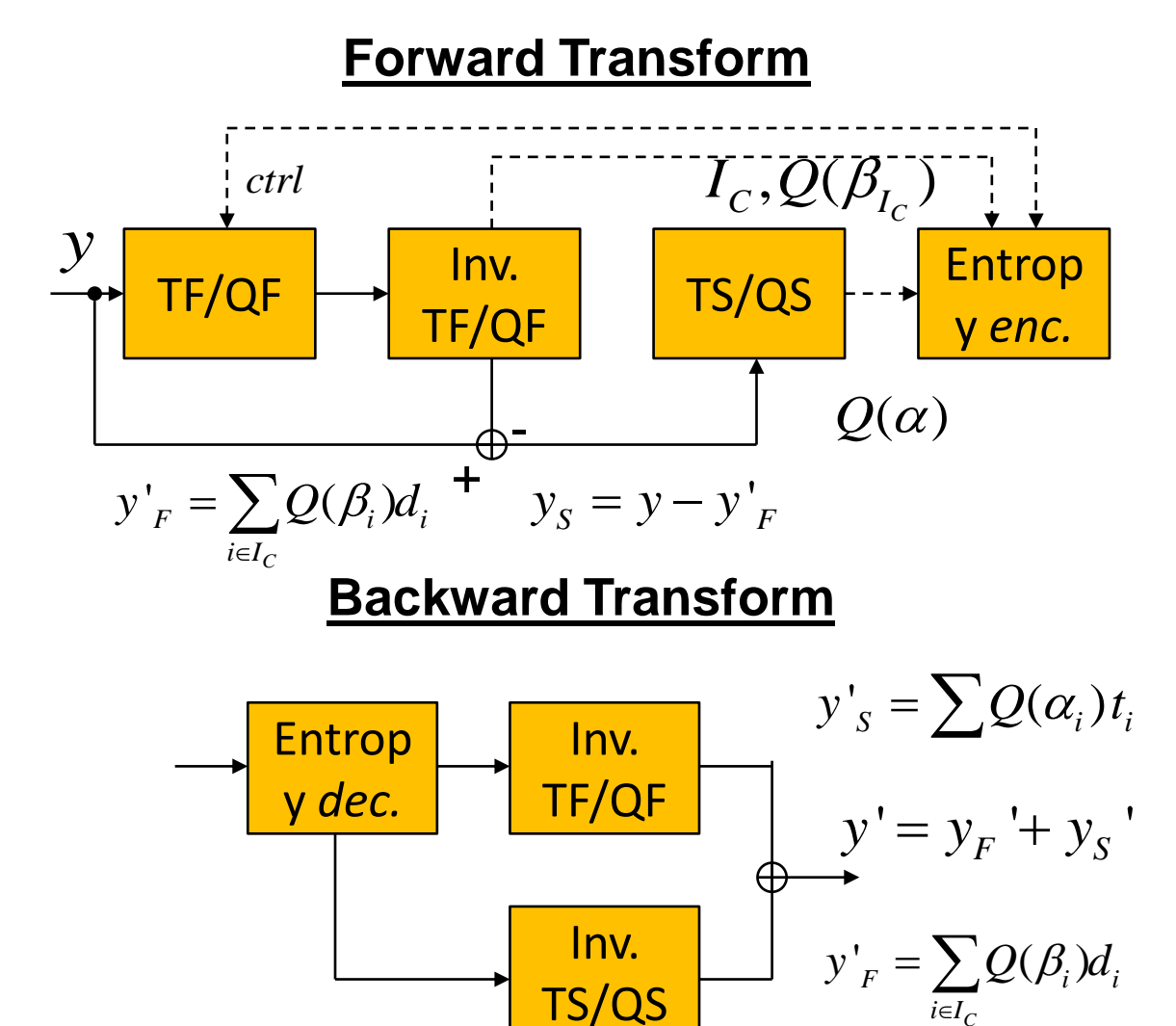
$$x = \sum_{i=0}^{I-1} \alpha_i \phi_i + \sum_{k=0}^{K-1} \beta_k \phi_k$$

Learned dictionary for sparse representation

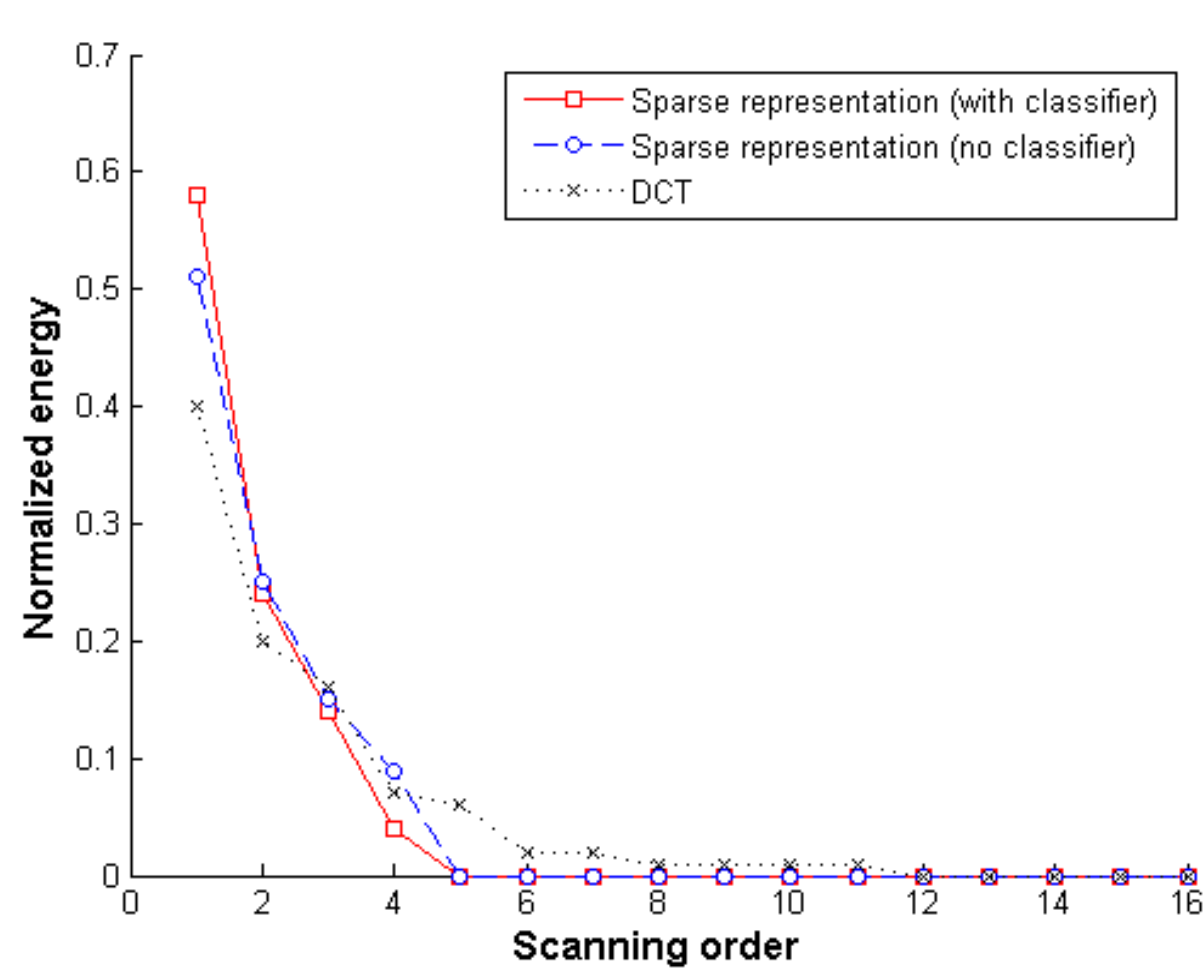
Block Diagram of the Proposed Video Coding Method



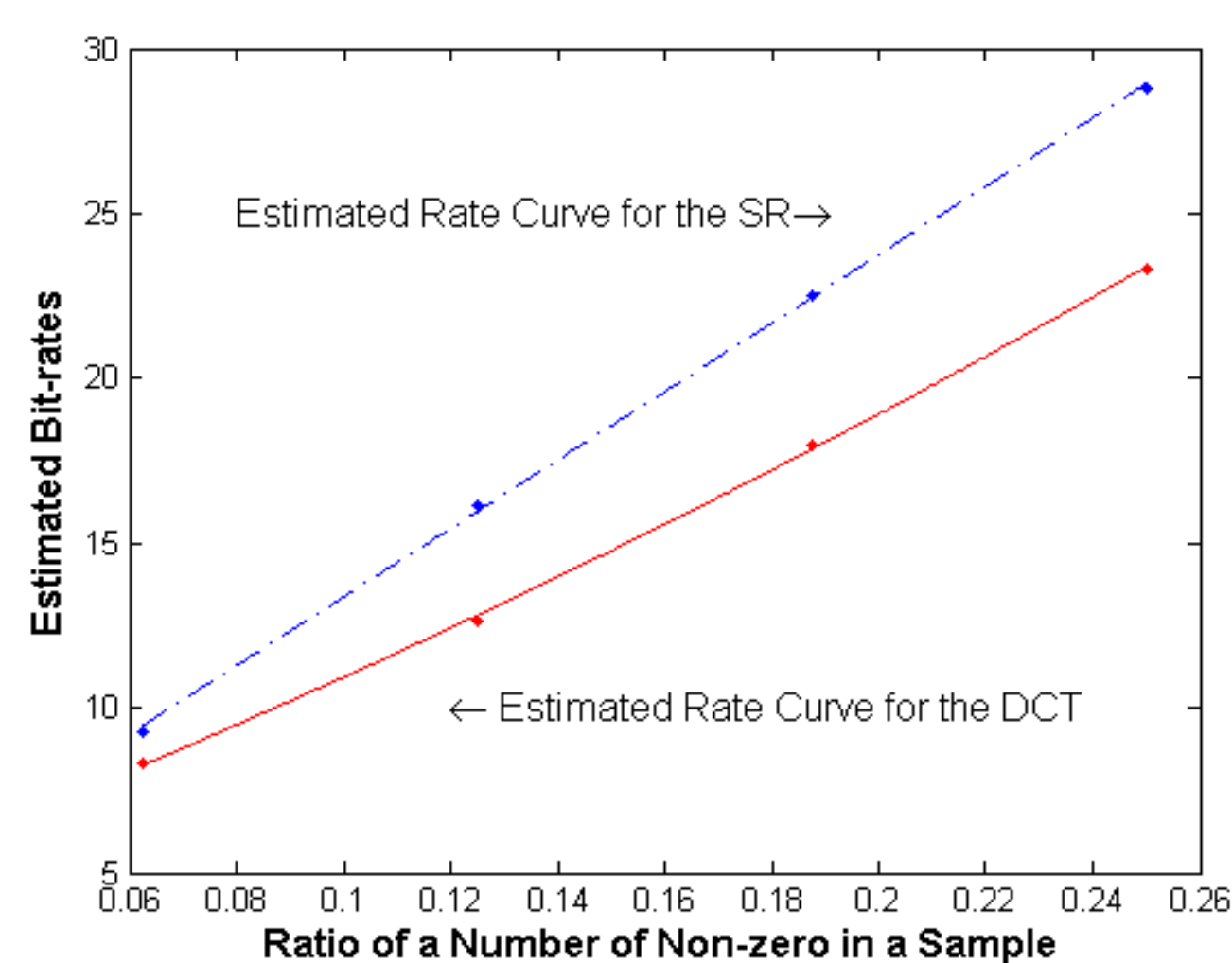
Transform



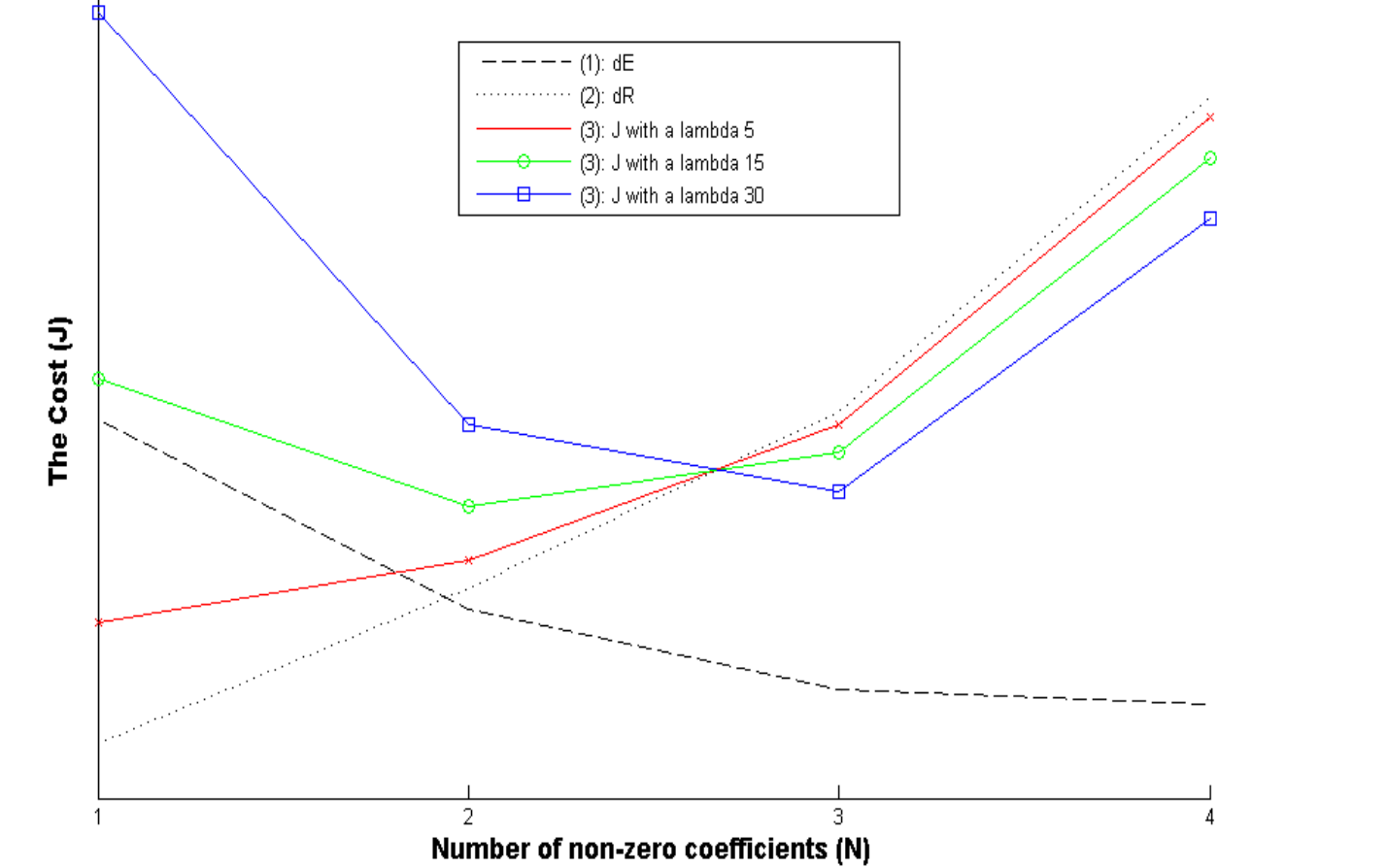
Rate-Distortion Analysis



Energy compaction capability between the two transforms



Bit-rate models of the two transforms



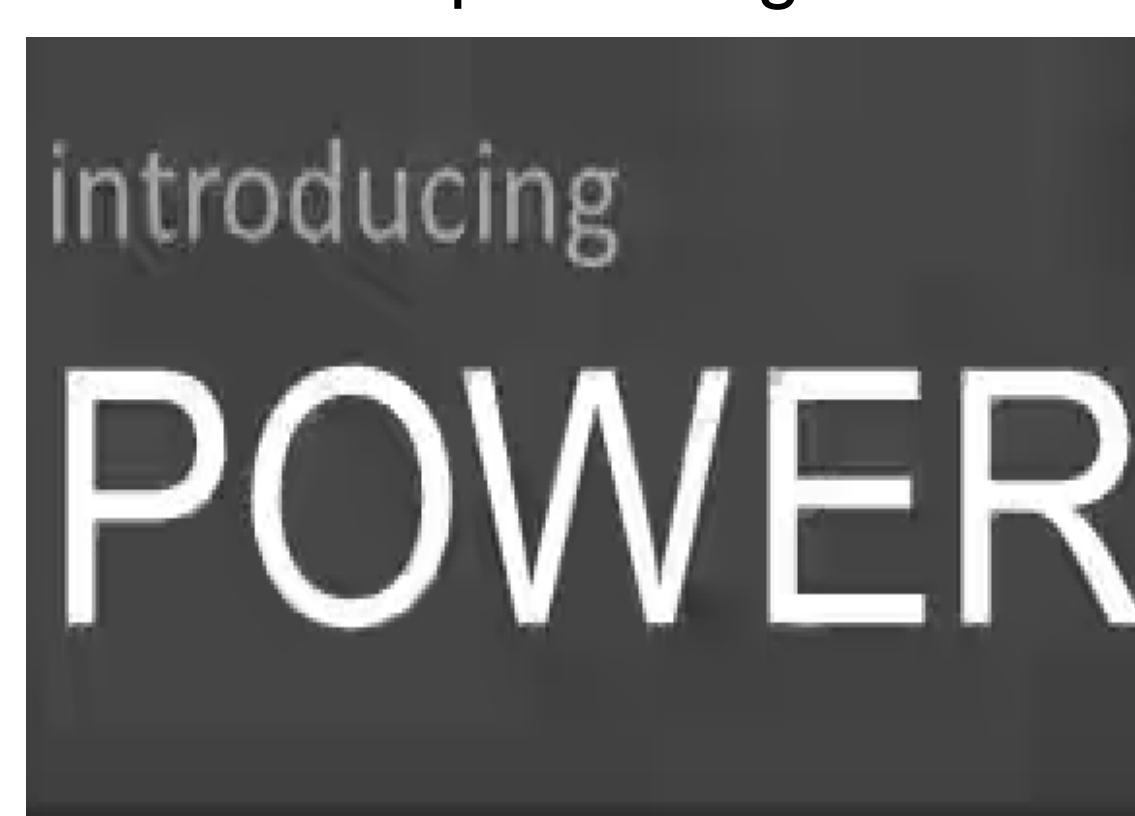
Switch the transforms by minimizing the total cost, controlled by an encoder

Experimental Results

BD-rate comparison

Types	Sequences	RA HE	RA LC	LDB HE	LDB LC
Class B (1080p)	Kimono	-0.0%	-0.1%	-0.1%	-0.2%
	ParkScene	-0.1%	-0.2%	-0.2%	-0.4%
	Cactus	-0.2%	-0.4%	-0.3%	-0.5%
	BasketballDrive	-0.4%	-0.7%	-0.6%	-1.3%
	BQTerrace	-0.5%	-0.7%	-0.7%	-1.3%
Average (B)		-0.24%	-0.42%	-0.38%	-0.74%
Class C (WVGA)	BasketballDrill	-0.5%	-1.0%	-0.9%	-1.6%
	BQMall	-0.7%	-1.1%	-1.3%	-2.4%
	PartyScene	-0.5%	-0.9%	-0.9%	-1.9%
	RaceHorses	-0.2%	-0.3%	-0.3%	-0.8%
Average (C)		-0.48%	-0.83%	-0.85%	-1.70%
Class D (WQVGA)	BasketballPass	-0.7%	-1.5%	-0.9%	-1.7%
	BQSquare	-1.1%	-1.9%	-1.4%	-2.7%
	BlowingBubbles	-0.9%	-1.3%	-1.1%	-2.2%
	RaceHorses	-0.2%	-0.4%	-0.3%	-0.7%
Average (D)		-0.73%	-1.28%	-0.93%	-1.83%
Total Average		-0.46%	-0.81%	-0.69%	-1.37%
Encoding Time		203%	191%	236%	221%

The Proposed Algorithm



37.42 dB

Conventional DCT



37.35 dB