

Xing-guo Zhu, Lu Yu, 2015. A reversibility-gain model for integer Karhunen-Loève transform design in video coding. *Frontiers of Information Technology & Electronic Engineering*, **16**(10):883-891. [doi:10.1631/FITEE.1500071]

A reversibility-gain model for integer Karhunen-Loève transform design in video coding

Key words: Integer transform, Karhunen-Loève transform (KLT), Integer-KLT, Transform coding, Video coding

Contact: Xing-guo Zhu

E-mail: 05xxgcsqx@zju.edu.cn

 ORCID: <http://orcid.org/0000-0002-6319-2471>

Introduction

- Karhunen-Loève transform (KLT)
 - The optimal transform that minimizes distortion at a given bit allocation for a Gaussian source
 - Contains non-integers: need integerization
- Integer-KLT design
$$\mathbf{A} = \text{int} (\mathbf{T} \cdot \alpha)$$
 - \mathbf{T} is the original KLT with non-integers
 - α is a scalar factor
 - 'int' denotes integerization which could be rounding, floor, ceiling or other operations
 - \mathbf{A} is the resulted integer-KLT

Traditional method of integer-KLT design

- Integer-KLTs with mathematical rounding

$$A(i, j) = \text{round}(T(i, j) \cdot \alpha).$$

- Integer-KLTs with TCG

$$\mathbf{A} = \arg \max_{\mathbf{C} \in \Phi} \{ \text{TCG}(\mathbf{C}) \}$$

- Φ is the candidate set of integer-KLTs and \mathbf{C} is a sample in the set
- $\text{TCG}(\mathbf{C})$ is the transform coding gain of \mathbf{C}

Integer-KLT design with R-G model

- Reversibility of a Transform \mathbf{A}

$$\begin{cases} R(\mathbf{A}) = \|\mathbf{U}\| \\ \mathbf{U} = \frac{\alpha^2 \mathbf{I} - \mathbf{A}^T \mathbf{A}}{\alpha^2} \end{cases}$$

- TCG loss ratio of a transform

$$\text{LR}(\mathbf{A}) = \frac{\text{TCG}(\mathbf{T}) - \text{TCG}(\mathbf{A})}{\text{TCG}(\mathbf{T})} \times 100\%.$$

Integer-KLT design with R-G model

- Reversibility-gain (R-G) model

$$\mathbf{A} = \arg \min_{\mathbf{C} \in \Phi} \{R(\mathbf{C})\} \quad \text{s.t.} \quad 0 \leq \text{LR}(\mathbf{C}) \leq \eta.$$

- η is a threshold to control the TCG of \mathbf{A} close to that of KLT.
- The equation seems to be an optimization problem. We use an enumeration method to solve it since it is not convex.

Performance

- Given 16 KLT matrices, we generate three groups of 16 integer-KLTs by different integerization methods.
 - \mathbf{A}_r : Integerization with rounding
 - \mathbf{A}_g : Integerization with maximal TCG
 - \mathbf{A}_{R-G} : Integerization with R-G model

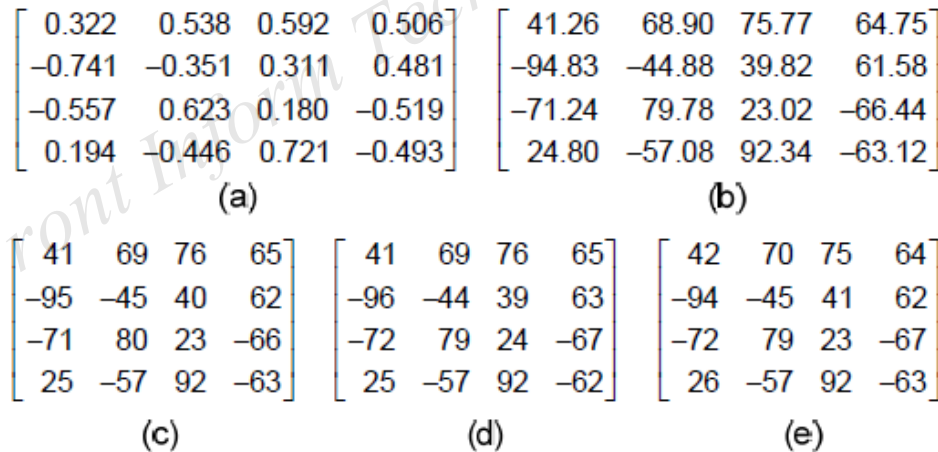


Fig. 3 Row integer-KLTs for vertical prediction mode
 (a) Float KLT matrix; (b) Scaled KLT matrix with $\alpha=128$;
 (c) \mathbf{A}_r ; (d) \mathbf{A}_g ; (e) \mathbf{A}_{R-G}

Performance

- Performance comparison on video coding
 - BD-bitrate: the bitrate increases of one method when compared with the anchor under the same PSNR
 - Anchor: video coding with A_r
 - Experimental platform: KTA-1.9
- Results
 - A_{R-G} has the best rate distortion performance in video coding

Table 5 BD-rates of A_g and A_{R-G} when compared with A_r

Sequence number	High bitrate (%)		Median bitrate (%)	
	A_g	A_{R-G}	A_g	A_{R-G}
0	2.74	-0.78	-0.06	0.03
1	3.60	-0.92	0.02	-0.11
2	1.98	-0.35	0.03	-0.08
3	1.80	-0.77	0.08	-0.05
4	2.30	-1.03	0.29	-0.11
5	2.19	-0.32	0.11	0.06
6	1.00	-0.25	0.00	0.02
7	2.23	-0.72	0.05	0.04
8	3.05	-0.91	0.05	0.01
9	0.83	-0.25	-0.06	-0.07
Average	2.17	-0.63	0.05	-0.02

Conclusions

- In this paper, we discuss integerization of KLT and focus on how to maintain useful properties (orthonormality and high TCG) of KLT for video compression.
- An R-G model that jointly considers the TCG and reversibility of a transform is proposed for integer-KLT design. The experimental results show effectiveness of R-G model.