Signal Dependent Transform Based on SVD for HEVC Intracoding

Tao Zhang, Haoming Chen, Ming-Ting Sun, *Fellow, IEEE*, Debin Zhao, *Member, IEEE*, and Wen Gao, *Fellow, IEEE*

***Abstract*—Transform is used to compact the energy of the blocks into a small number of coefficients and is widely used in recent image/video coding standards. In the latest video coding standard high efficiency video coding (HEVC), a combination of discrete cosine transform (DCT) and discrete sine transform (DST) is adopted to transform the residuals from intra prediction. Since the DCT and DST are the fixed transforms that are derived from the Gauss–Markov model, some of residual blocks may not be compacted well by the DCT/DST. In this paper, we propose a signal dependent transform based on singular value decomposition (SVD) for HEVC intracoding. The proposed transform (SDT-SVD) is derived by performing SVD on the synthetic block and applied to the residual block considering the structural similarity between them. Furthermore, we extend SDT-SVD to template matching prediction (TMP) to further improve the intracoding performance. Experimental results show that the proposed transform on angular intra prediction (AIP) outperforms the latest HEVC reference software with a bit rate reduction of 1.0% on average and it can be up to 2.1%. When the proposed transform is extended to TMP-based intracoding, the overall bit rate reduction is 2.7% on average and can be up to 5.8%.**

***Index Terms*—Transform, HEVC, Intra coding, SVD, DCT.**

I. Introduction

WITH the development of the Internet and hardware, various multimedia applications related to image/video coding have been developed recently, such as video communication, video streaming, cloud-based image storage/transmission [1] and cloud computing and screen sharing [2] based on screen video coding [3]–[6]. Intra coding [7] plays an important role in recent video coding standards. The latest HEVC [8] video coding standard employs a quadtree structure in which the size of the coding unit (CU) can be from 8×8 to 64×64. Transform units (TU) are the basic units in the transform which are generated by the quadtree partition of the CU, and the size of TU can be from 4×4 to 32×32. The quadtree partitions for the CU make the block-size more flexible and adaptive to the characteristics of the content which improves the coding efficiency significantly. The quadtree partitions for the TU make the transform block-size adaptive to the prediction residual distribution so that the residual energy can be compacted more effectively. Moreover, compared to only 9 modes in H.264/AVC [9], there are 35 prediction modes [10] in HEVC intra coding which significantly improves the prediction accuracy. In this paper, we focus on the transform in intra coding which is an important part for improving the intra coding efficiency.

In H.264 or HEVC intra coding, the transform is used to remove the redundancy remained in the prediction residual block through compacting these residuals into a small number of transform coefficients. Karhunen–Loeve transform (KLT) is the optimal transform for decorrelation if only one fixed transform is used for all the coding blocks. In [11], DCT has been proved to be close to KLT in terms of energy compaction for coding blocks without any prediction [12]–[14]. However, with the highly accurate prediction, the residual block after prediction does not contain much correlation compared to the original coding block. In this case, DCT may not be optimal any more.

To deal with this problem, many new transforms have been proposed. Mode-dependent directional transform (MDDT) for H.264/AVC intra coding was proposed in [15]. In MDDT, nine transforms which are trained from an off-line dataset are used for nine different intra prediction modes. In [16] and [17], a set of transforms are trained offline and the best transform is selected by rate-distortion optimization (RDO). In [18], rate distortion based rotational transform (ROT) was proposed for HEVC intra coding. ROT is implemented as a secondary transform applied after the primary DCT. The encoder tries every rotational transform from the dictionary and then selects the best one. In [19]–[21], DST was proposed to transform prediction residuals in intra coding. In these methods, a combination of DCT and DST is derived from the Gauss-Markov model for the transforms of the rows and columns of the residual block. From statistical analysis, the residual block has different features compared to the original block. The pixels close to the reference boundary can be predicted better than those far away

Manuscript received March 21, 2016; revised March 3, 2017 and April 18, 2017; accepted April 27, 2017. Date of publication May 10, 2017; date of current version October 13, 2017. This work was supported in part by the Major State Basic Research Development Program of China (973 Program) under Grant 2015CB351804. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Balakrishnan Prabhakaran. *(Corresponding author: Debin Zhao.)*

T. Zhang is with the Department of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China. He is now with Tencent, Beijing 518057, China (e-mail: taozhang.hit@hotmail.com).

H. Chen is with the Google, Inc., Kirkland, WA 98033 USA (e-mail: haominghoward@gmail.com).

M.-T. Sun is with the Department of Electrical Engineering, University of Washington, Seattle, WA 98105 USA (e-mail: sun@uw.edu).

D. Zhao is with the Department of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China (e-mail: dbzhao@hit.edu.cn).

W. Gao is with the Department of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China (e-mail: wgao@pku.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMM.2017.2703114

from the reference boundary, which leads to small residuals close to reference boundary while large residuals far away from the boundary. The combination of DCT and DST can handle these residuals well.

The above mentioned transforms are fixed transforms which are not adaptive to the coding content. In order to overcome the drawback of fixed transforms, several signal dependent transforms [22]–[27] were proposed to improve the coding performance. As the signal dependent transform is derived for each coding block during the coding process, the energy can be better concentrated. In this paper, we propose a signal dependent transform based on SVD (SDT-SVD) for HEVC intra coding. Considering the structural similarity between the residual block and its synthetic block, SDT-SVD is derived by performing SVD on its synthetic block which is available both in the encoder and decoder, therefore avoiding the transmission of the transforms to the decoder.

The rest of the paper is organized as follows. Section II gives the related work of the proposed transform. Section III presents the proposed signal dependent transform based on SVD. HEVC Intra coding with the proposed SDT-SVD is described in detail in Section IV. The simulation results of the proposed method are presented in Section V. Concluding remarks are given in Section VI.

II. Related Work

In this section, the related work of the proposed SDT-SVD transform is given. First, the existing signal dependent transforms are briefly reviewed. Then, transforms in HEVC intra coding are presented in detail.

A. *Signal Dependent Transforms*

Several signal dependent transforms were proposed to improve the coding performance. In [22], an image-coding algorithm which combines DCT and SVD was proposed. SVD transform was calculated from the coding block and the transform coefficients which are the eigenvalues can also be derived by SVD. Then, a small number of eigenvalues and the derived eigenvectors are encoded and transmitted. Since the bits cost on the eigenvectors is significant, the coding performance is limited. In [23], a content adaptive transform was proposed for H.264/AVC intra coding, in which a residual block was downsampled into four equal-sized sub-blocks, the first sub-block was transformed by DCT and the other three sub-blocks were transformed by the transforms derived by performing SVD on the reconstructed first sub-block. Another adaptive transform based on SVD was proposed in [24] for HEVC inter coding. The transform for the current residual was directly derived from the prediction block in the reference frame considering that there has similarity between the residual block and the original block because of the imperfect motion estimation/compensation and the original block can be approximated by the prediction block [25]. A signal dependent transform based on KLT was proposed in [26], [27], in which lots of blocks similar to the current coding block are searched by template matching and then used to train the KLT transform for the block predicted by template matching prediction [28] in intra coding.

B. *Transforms in HEVC Intra Coding*

Under the Gauss-Markov model when the correlation coefficient between pixels is close to 1 which means there is large redundancy in the coding block, it has been demonstrated that DCT is close to KLT [11]. Since the prediction in HEVC intra coding can reduce major redundancy, the correlations between residual pixels are not as strong as those in the original pixels. Theoretically, DST has been demonstrated to be close to the optimal KLT for the residual block after the intra prediction in some cases [19]–[21]. For the prediction modes which are close to the vertical direction, DST is close to KLT as the vertical transform, and DCT is close to KLT as the horizontal transform. Similarly, for the prediction modes which are close to the horizontal direction, DCT is close to KLT as the vertical transform and DST is close to KLT as the horizontal transform. According to [21], different combinations of DCT and DST should be used for different intra prediction modes.

In the HEVC video coding standard, in order to achieve a good balance between the complexity and the coding efficiency, it was proposed to always use DST as the transforms for the luma 4×4 blocks in intra coding, and use DCT in other cases [30]. This simplification brings little coding gain loss compared to [21] while removing the switching logic to choose between DCT and DST.

III. Proposed Signal Dependent Transform Based on SVD

In this section, intra prediction residual analysis is provided firstly. Then, the derivation of the proposed transform is presented.

A. *Intra Prediction Residual Analysis*

DCT/DST transforms are the fixed transforms, which are not adaptive to the prediction residuals. According to our observations, some residual blocks have similar structural information as their prediction blocks in HEVC intra coding. For example, some blocks have horizontal textures and can be predicted by the horizontal modes in the intra prediction. After intra prediction, the residuals often still have some horizontal textures.

This structural similarity between the residual and the prediction mainly comes from three aspects. The first one is that only 33 angular modes (shown in Fig. 1) in HEVC intra coding cannot cover all the directions in a coding block. Thus, only the one which is closest to the ideal direction of the block is selected in the prediction. The second one is that the pixel values along one direction may vary for the coding block, which cannot be predicted well by using one fixed prediction value. The third one is that the reference pixels may be distorted by the quantization resulting in inaccurate prediction. Because of such imperfect prediction, the residual block may still remain some similar structures as the original block or prediction blocks.

Fig. 2(a) and 2(b) show one residual frame and the corresponding prediction frame generated by HEVC intra coding for sequence *BasketballPass*, respectively. It can be seen that there exists structural similarity between the residual frame and the

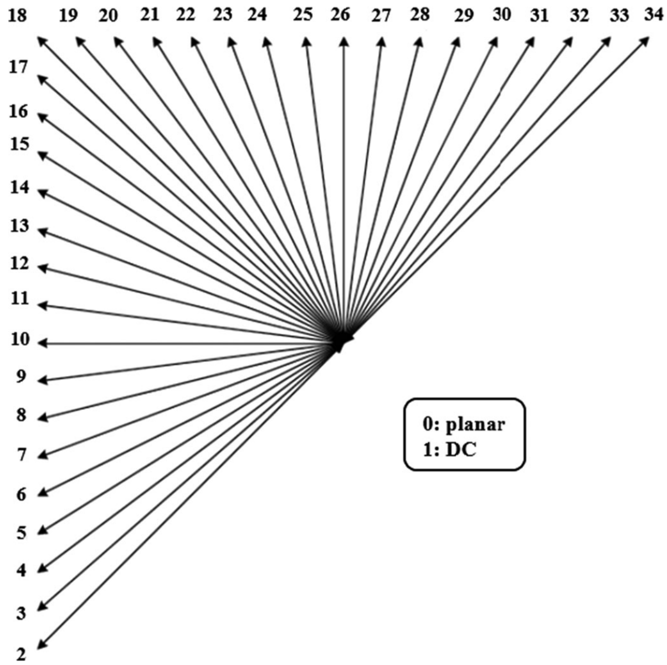


Fig. 1. Intra prediction modes in HEVC.

(a) (b)

Fig. 2. (a) One residual frame (shown with enhanced contrast and best viewed when zoom in) and (b) the corresponding prediction frame from sequence *BasketballPass* by HEVC intra coding when QP = 22. (a) Residual frame. (b) Prediction frame.

133	132	132	133
131	131	133	130
129	131	133	133
118	119	118	117

(a)

137	137	137	137
132	132	132	132
128	128	128	128
117	117	117	117

(b)

-4	-5	-5	-4
-1	-1	1	-2
1	3	5	5
1	2	1	0

(c)

Fig. 3. Example of a 4 × 4 block predicted by the horizontal mode (mode 10). (a) Original block. (b) Prediction block. (c) Residual block.

prediction frame in some regions, such as the boundaries of ceiling, some textures of the floor, and the outline of the players. Fig. 3 shows an example of a 4 × 4 block from the frame in Fig. 2 predicted by the horizontal mode. We can see that the residual block still has some horizontal textures which are similar to its prediction block. In some cases, transforms adaptive to the structures of residual can compact the energy better than DCT/DST.

In order to statistically show that there has structural similarity between the prediction block P and the residual block R, we calculate the Pearson correlation coefficient (PCC) between P and R as

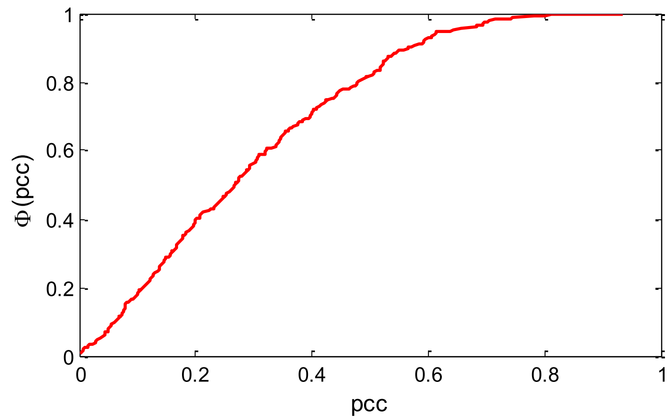


Fig. 4. Distribution of PCCs for CUs coded with the size of 8 × 8. Φ is the cumulative distribution function.

$$\mathrm{PCC} = \frac{\sum_{i=1}^{n} (R_i - \bar{R})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^{n} (R_i - \bar{R})^2}\sqrt{\sum_{i=1}^{n} (P_i - \bar{P})^2}} \tag{1}$$

where R_i is residual value and P_i is the prediction value. $\bar{R}$ and $\bar{P}$ are the mean value of the residual block and the prediction block. n is the number of pixels in the block. The distribution of PCCs for CUs coded with the size of 8 × 8 is given in Fig. 4 for sequence *BasketballPass*. In Fig. 4, Φ is the cumulative distribution function for PCC. It can be seen that P and R have some structural similarity. For example, there is about 30% of the 8 × 8 CUs with PCC > 0.4.

B. Derivation of SDT-SVD

In video coding, the SVD based transform on residual block R can be represented as

$$S_R = {U_R}^T R V_R \tag{2}$$

where S_R is the transform coefficient matrix which is a diagonal matrix with singular values along the main diagonal. U_R and V_R are the eigenvector matrices derived by performing SVD on R as

$$[U_R, S_R, V_R] = SVD(R) \tag{3}$$

Since the SVD transform can diagonalize the residual block, the redundancy of the residual block is fully removed. In [22], a hybrid DCT-SVD transform for image coding was proposed. In this hybrid transform, the eigenvectors U and V need to be coded and transmitted to the decoder. It is obvious that the bits cost on U and V takes a large part of the total bits.

Reducing the coding bits on transform matrices U and V is crucial to SVD transform based image/video coding. However, previous works [31], [32] show that it is difficult to find an efficient compression method for matrices U and V. Another solution is to compromise the decorrelation capability of the SVD transform in order to decrease the bits cost on coding transform matrices, such as the works in [23], [24].

In the previous analysis, we have shown that some residual blocks in HEVC intra coding still remain some structural

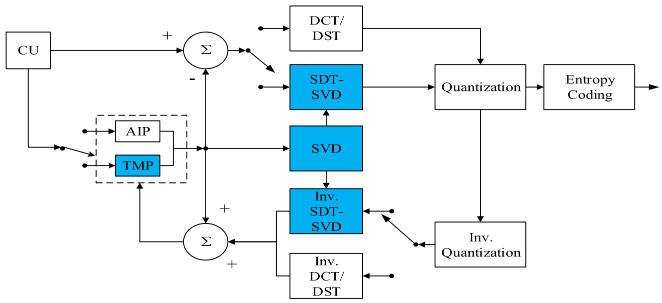


Fig. 5. Framework of intra coding with the proposed SDT-SVD. The blue boxes are new added parts for the proposed coding method.

information as the prediction blocks. Therefore, the transform derived by performing SVD on the prediction block P can be used to transform the residual block.

The SVD transform derived from P is represented by

$$[U_P, S_P, V_P] = SVD\,(P) \tag{4}$$

where $\mathrm{U_P}$ and $\mathrm{V_P}$ are the eigenvector matrices for P, $\mathrm{S_P}$ is a diagonal matrix with singular values along the main diagonal and $\mathrm{S_P} = \mathrm{U_P^T P V_P}$. $\mathrm{U_P^T}$ and $\mathrm{V_P}$ can be used as the transform for residual block R, named as SDT-SVD.

Based on above analysis, we can replace the DCT/DST in HEVC intra coding with the proposed SDT-SVD in some cases.

IV. HEVC Intra Coding With SDT-SVD

In this section, we first introduce the framework of HEVC intra coding with SDT-DCT, then we give the details of the proposed SDT-SVD on AIP. Finally, we extend the proposed SDT-SVD to TMP based intra coding.

A. The Framework of HEVC Intra Coding with SDT-SVD

Fig. 5 shows the framework of intra coding with the proposed SDT-SVD. The blue boxes are the changes compared to the original HEVC intra coding. In Fig. 5, a coding unit (CU) is the input. AIP represents intra prediction in HEVC. TMP indicates template matching prediction which is a new prediction mode and will be described in the next subsection.

The transform matrices U and V for SDT-SVD can be directly derived from the prediction block P of the coding block O by (4), and then applied to the residual block R in order to get transform coefficients C by

$$C = {U_P}^T R V_P. \tag{5}$$

The corresponding inverse transform is represented by

$$R = U_P C {V_P}^T. \tag{6}$$

In order to avoid the floating-point arithmetic computation and also be consistent with the design of transforms [37] in HEVC, the elements of derived SVD transform matrices $\mathrm{U_P}$ and $\mathrm{V_P}$ are multiplied by $2^{(6+log(N)/2)}$ and then rounded to the closest integer. $\mathrm{N \times N}$ is the size of the transform. In this case, there is no drift problem.

As mentioned in Section II, there has some structural similarity between the residual block and the prediction block. If the structural similarity is higher, the derived transform will be more efficient to compact the energy of the residual. In order to get higher structural similarity, a synthetic prediction block P' is generated by combining the prediction block and the two prediction blocks from its two neighboring intra modes as follows:

$$P' = (w_L \cdot P_L + w_C \cdot P + w_R \cdot P_R)\,/w \tag{7}$$

where P is the prediction for current block, $\mathrm{P_L}$ and $\mathrm{P_R}$ are the prediction blocks generated from the left and right intra modes of the current mode, respectively. w_{L}, w_{C} and w_{R} are weighting factors for these three predictions and $w = w_{\mathrm{L}} + w_{\mathrm{C}} + w_{\mathrm{R}}$. In practical realizations, w_{L}, w_{C} and w_{R} are integers and w can be represented by $w = 2^{\mathrm{n}}$ to support division using low-complexity bit-shifting. For example, if the current mode is mode 10, then the left mode is mode 9 and the right mode is mode 11. Table I shows the derivation of the left mode (modeL) and the right mode (modeR) for block with modeC.

TABLE I
Derivation of the Left Mode and the Right Mode

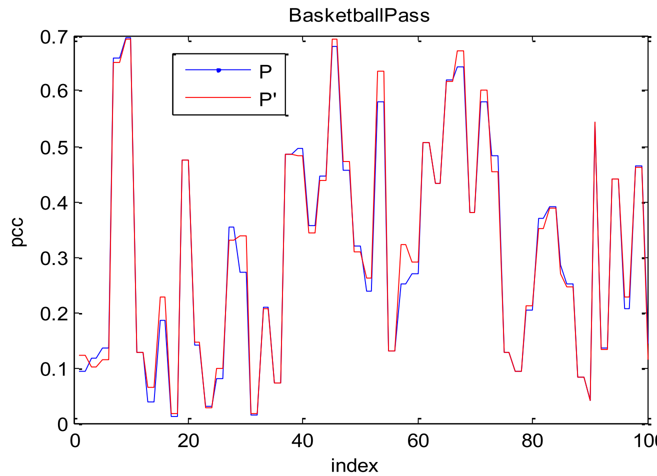
modeC	modeL	modeR
2	33	3
34	3	33
3∼33	modeC − 1	modeC + 1

To statistically show that the synthetic prediction P' is more similar to R than that of P, the PCCs between P' and R, and the PCCs between P and R for a group of coding blocks (100 blocks with the size of 8×8) are compared, as shown in Fig. 6. It can be observed that the PCCs between P' and R are higher than the PCCs between P and R. It means that the synthetic prediction P' contains higher structural similarity to the residual R than that of P. Thus, P' can be used to derive the SDT-SVD transform as follows:

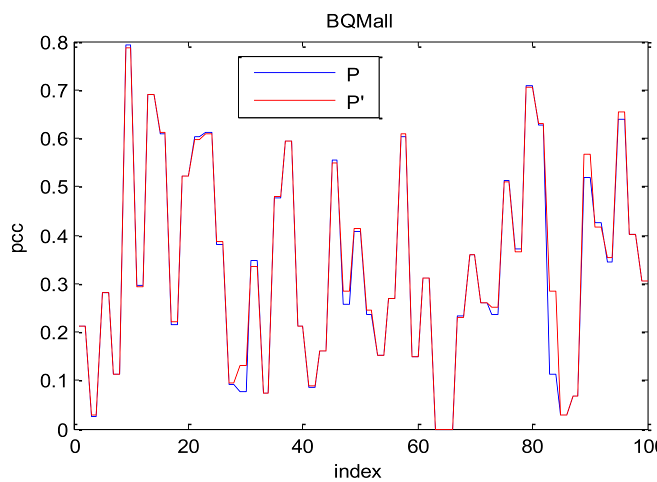
$$[U'_P, S'_P, V'_P] = SVD\,(P')\,. \tag{8}$$

From the analysis in the previous section, we know that for most of the residual blocks which fit the Gauss-Markov model well, DST/DCT will be more efficient; while for some other residual blocks which contain similar structures as the prediction blocks, the proposed SDT-SVD transform will be more efficient. In this paper, we propose to use the RDO process to decide the best transform for the coding block. To reduce the overhead for signaling the best transform type, we implement the proposed SDT-SVD scheme on CU level. It means all the TUs in the CU will choose the same transform.

In HEVC intra coding, several selected intra prediction modes from rough mode decision (RMD) will go through the RDO process, and the best intra prediction mode and TU partition are derived by minimizing the RD cost. In the proposed coding method, for each of the selected intra mode, both the original DST/DCT and the proposed SDT-SVD are tested in the RDO process, and the best transform is determined for each intra mode. Finally, the best combination of intra mode, transform type and TU partition are obtained based on the RD costs.



(a)



(b)

Fig. 6. PCCs between P (P') and R for a group of SDT-SVD transformed blocks with index from 1 to 100. (a) *BasketballPass.* (b) *BQMall.*

B. *Extending to Template Matching Prediction*

Angular intra prediction is efficient for predicting blocks with strong directional information which can be handled well by different angular modes. However, it is unable to predict blocks with complex textures or with multiple directions. Several new intra prediction methods have been proposed to exploit redundancy within an image. A block based prediction was proposed in [33] by using motion estimation and motion compensation techniques which are classically used in inter coding. In this method, the reconstructed part of the frame is used as reference, and the prediction for current block is searched in the reference by block matching. The derived block vector which indicates the best match for current block should be encoded. This method was also called intra block copy (IBC) [34], [35] and was adopted in the HEVC Range Extension for screen content coding because of its high coding performance.

However, IBC is not efficient for natural video since the large overhead for signaling the block vector [36]. A similar method based on block searching called template matching prediction (TMP) [29] was proposed as an extra prediction mode for H.264/AVC. In TMP, the template (encoded pixels close to the block) is used to find blocks with similar templates and these blocks are used as the predictions for the coding block, as shown in Fig. 7. In the encoder side, the index of the template with the best prediction block needs to be encoded. The decoder can perform the same template matching process to find a set of candidates and get the best prediction by the decoded index.

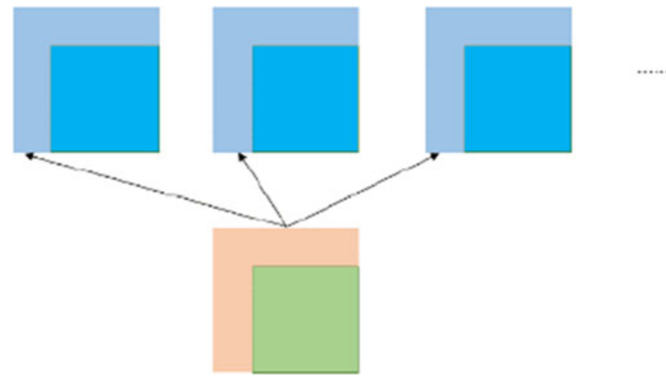


Fig. 7. Template matching process. Inverse-L area is the template (in light blue) used in the matching process. All the deep blue blocks are the prediction candidates for the current block (in green).

Similar to AIP, we observed the residual block from TMP contains some structural information (e.g., boundaries and edges) as the predictions generated by TMP. This structural similarity between residual block and original block (or prediction block) mainly comes from the imperfect block matching. Although similar blocks can be found by TMP, some parts of the coding block especially object boundaries and edges may not be predicted well by the matched blocks. In this case, the SDT-SVD transform can be efficient to remove the redundancy in the residual blocks since it can capture the structural information from the prediction blocks. Thus, besides DCT, SDT-SVD can be used as an alternative transform for TMP.

Based on the above analysis, for TMP, the SDT-SVD transforms can be directly derived from the prediction block. To get a synthetic block with more similar structures as the residual block, we can obtain several residual blocks by subtracting the prediction candidates from the best prediction block which is selected by RDO process when DCT is applied. These calculated residual blocks may contain more similar structures as the residual block R. The synthetic residual block R' used to derive the SVD transform for TMP is calculated by averaging these residual blocks as

$$\mathrm{R}' = \frac{1}{\mathrm{M}} \sum_{\mathrm{i}=1}^{\mathrm{M}} (\mathrm{P} - \mathrm{P}_i) = \mathrm{P} - \frac{1}{\mathrm{M}} \sum_{\mathrm{i}=1}^{\mathrm{M}} \mathrm{P}_i \tag{9}$$

where P is the best prediction generated by TMP, $\mathrm{P_i}$ ($\mathrm{i} = 1, 2, \ldots, \mathrm{M}$) are the prediction candidates with least SAD costs for the templates. M is the number of candidates used for calculating R'.

Similar to the analysis for P' in previous subsection, to statistically show that the synthetic residual R' is more similar to R than that of P, the PCCs between R' and R, and the PCCs between P and R for a group of coding blocks with the size of 8×8 in the first frame of each sequence are compared, as shown

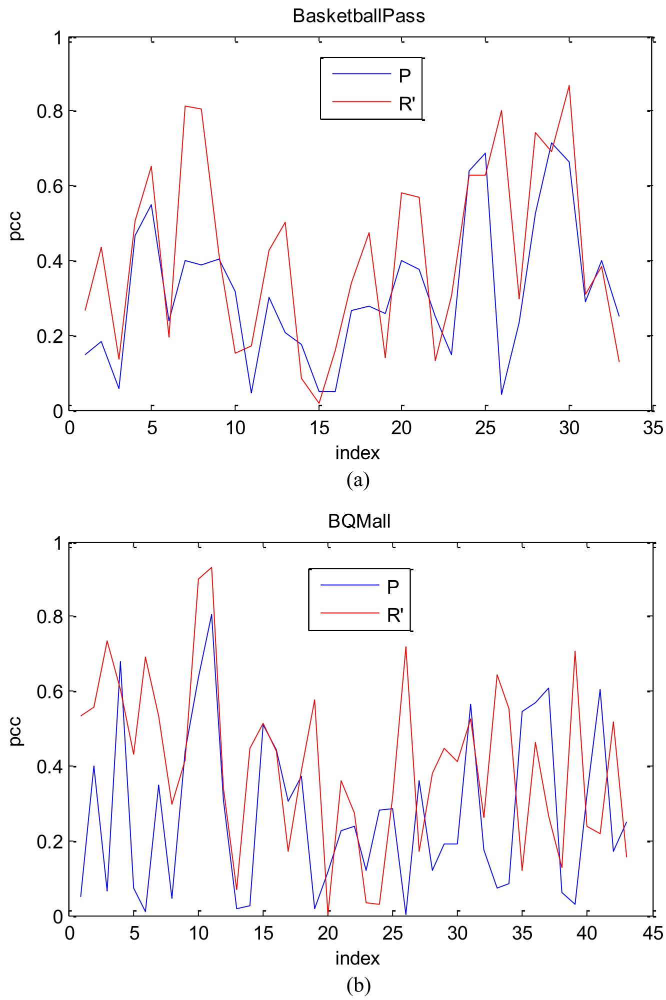


Fig. 8. PCCs between P (R') and R for a group of SDT-SVD transformed blocks in the first frame of each sequence. (a) *BasketballPass.* (b) *BQMall.*

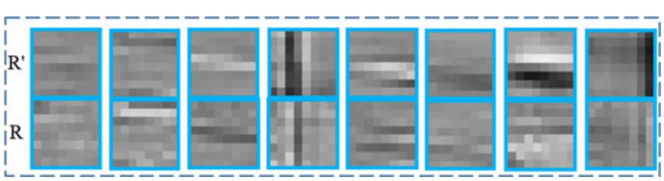


Fig. 9. Examples of residual block R by TMP and its corresponding R' calculated by (9) for 8 TUs. The first row shows R' and the second row shows the corresponding R. (a) *BasketballPass.* (b) *BQMall.*

in Fig. 8. It can be observed that the PCCs between R' and R are higher than that between P and R. It means that the synthetic residual R' contains higher structural similarity to the residual R than that of P. Thus, R' can be used to derive the SDT-SVD transform for TMP prediction residuals as follows:

$$[U'_R, S'_R, V'_R] = SVD\,(R')\,. \qquad (10)$$

To illustrate the similarity between R' and R visually, Fig. 9 gives examples of the residual R from TMP and the corresponding R' calculated by (9) for 8×8 TUs transformed by SDT-SVD. It can be seen some structural similarity exists between R and its corresponding R'.

The proposed template matching prediction with SDT-SVD is only implemented on 8×8 and 16×16 CUs considering the balance between coding efficiency and complexity. Firstly, we choose N candidates which have the smallest SAD values for the templates. Then, a two-step search algorithm is employed to get the best combination of the prediction candidate and transform type in TMP coding. In the first step, a rough candidate selection is conducted to get a small number of candidates which have the smallest SADs between the coding block and the candidate blocks. In the second step, the RDO process is conducted to obtain the best prediction candidate when DCT is used as the transform from these candidates. In the proposed TMP, the number of candidates after rough selection is much smaller than the allowed maximum number of candidate $(\mathrm{N_t})$. In this paper, $\mathrm{N_t}$ is set to 32, and the candidate numbers after rough selection are set to 8 for 8×8 CU and 3 for 16×16 CU, respectively. After obtaining the best prediction candidate, the SDT-SVD is applied to the residual block from this prediction. The best transform is then selected between DCT and SDT-SVD by comparing their RD costs. The index of the best prediction candidate is coded with fixed length coding using $\log_2(\mathrm{N_t})$ bits and a 1-bit flag for signaling the best transform type is entropy coded.

V. Experimental Results

This section presents experimental results of the proposed SDT-SVD transform based intra coding compared to intra coding in HEVC reference software in terms of coding gain and computational complexity.

A. *Experimental Settings*

To evaluate the performance of the proposed method, we implemented the proposed method on the recent HEVC reference software (HM 14.0)[1]. The simulation results are provided using test sequences from JCTVC [38] encoding with the All-Intra main configuration. QP values of 22, 27, 32, and 37 are used for the evaluation. We measure the coding performance by the Bjontegaard-Delta Bit Rate (BD-BR) [39]. The performance of the proposed method is compared to the anchor which is the intra coding in HM 14.0 with default settings. Three parts of the proposed method are compared separately to the anchor. The three parts are: 1) AIP with SDT-SVD; 2) TMP with SDT-SVD; 3) AIP with SDT-SVD + TMP with SDT-SVD (Proposed overall);

In the proposed TMP method, the width of template is two pixels adjacent to the coding block and the search range used for searching the template candidates is limited to not exceed left LCU and above LCU. For deriving transforms for TMP, M in (9) is set to 8. For deriving transforms for AIP, w_{L}, w_{C} and w_{R} in (7) are set to 1, 6 and 1, respectively. To verify the efficiency of synthetic prediction P' for deriving transforms by (8) for AIP, we also test the method when deriving transforms from P by (4).

In addition, TMP without SDT-SVD (only DCT are applied to TMP) is also tested in order to know the additional gain of the proposed SDT-SVD on TMP. Furthermore, the proposed overall

[1][Online]. Available: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-14.0.

TABLE II
CODING PERFORMANCE OF THE PROPOSED METHOD AND [27]

Class	Sequences	Resolution	Bit depth	AIP w/ SDT-SVD (4)	AIP w/ SDT-SVD (8)	TMP w/o SDT-SVD	TMP w/ SDT-SVD	Proposed overall	[27]
A	Traffic	2560 × 1600	8	−0.3%	−0.3%	−1.1%	−1.6%	**−1.9%**	−1.4%
	PeopleOnStreet	2560 × 1600	8	−0.3%	−0.4%	−1.4%	−2.0%	**−2.3%**	−1.2%
B	Kimono	1920 × 1080	8	0.0%	−0.1%	−0.2%	−0.3%	**−0.3%**	−0.3%
	ParkScene	1920 × 1080	8	−0.2%	−0.2%	−0.4%	−0.7%	**−0.8%**	−0.5%
	Cactus	1920 × 1080	8	−0.7%	−0.8%	−2.1%	−2.9%	**−3.4%**	−2.7%
	BasketballDrive	1920 × 1080	8	−1.6%	−2.1%	−1.7%	−2.6%	**−4.2%**	−3.5%
	BQTerrace	1920 × 1080	8	−0.7%	−0.8%	−3.5%	−4.5%	−5.0%	**−5.5%**
C	BasketabllDrill	832 × 480	8	−0.6%	−0.6%	−3.4%	−5.1%	−5.5%	**−6.2%**
	BQMall	832 × 480	8	−1.6%	−1.9%	−0.7%	−1.0%	**−2.7%**	−2.0%
	PartyScene	832 × 480	8	−1.4%	−1.5%	−0.6%	−1.0%	**−2.1%**	−0.8%
	RaceHorses	832 × 480	8	−0.4%	−0.5%	−0.2%	−0.4%	**−0.8%**	−0.4%
D	BasketballPass	416 × 240	8	−1.9%	−2.1%	−0.9%	−1.3%	**−3.2%**	−1.2%
	BQSquare	416 × 240	8	−1.2%	−1.3%	−0.6%	−1.0%	**−2.2%**	−0.3%
	BlowingBubbles	416 × 240	8	−0.9%	−0.9%	0.0%	−0.1%	**−1.0%**	−0.2%
	RaceHorses	416 × 240	8	−0.5%	−0.5%	0.0%	−0.2%	**−0.7%**	−0.1%
E	FourPeople	1280 × 720	8	−0.6%	−0.8%	−1.3%	−1.9%	**−2.5%**	−1.8%
	Johnny	1280 × 720	8	−0.5%	−0.9%	−3.7%	−5.0%	**−5.8%**	−5.2%
	KristenAndSara	1280 × 720	8	−0.9%	−1.1%	−2.3%	−3.4%	**−4.0%**	−3.2%
	Average			−0.8%	−1.0%	−1.3%	−2.0%	**−2.7%**	−2.0%
	Enc. time ratio			2.39	2.45	2.51	3.04	4.48	19.41
	Dec. time ratio			1.14	1.17	4.40	5.88	5.76	13.22

method is also compared with SDT [26] [27] in which TMP is also used for prediction and KLT is trained from a set of TMP matched candidates. The KLT is then used to transform residual generated by TMP.

B. Rate-Distortion Performance

The results for the proposed methods are shown in Table II. From this table, we can see that the average BD-rate reduction for AIP with SDT-SVD derived by (4) is 0.8%. When the transforms are derived by (8), the average gain is increased to 1.0%. It also can be seen that the improvement compared to the method using (7) can be up to 0.5% for sequence *BasketballDrive*. In the TMP based intra coding, the pure TMP (TMP w/o SDT-SVD) provides 1.3% gain compared to the anchor, while the proposed SDT-SVD on TMP further improves TMP by 0.7%. The average gain of the proposed overall method is 2.7% and the gain can be up to 5.8% for sequence *Johnny*, while the BD-rate reduction for SDT [27] is 2.0% on average.

Fig. 10 plots the rate distortion curves for sequence *BasketballPass* and *BQMall* for AIP w/ SDT-SVD. We can observe that the proposed SDT-SVD for AIP outperforms HEVC at different quantization levels, which verifies that the derived SDT-SVD works well under different QPs.

In order to further verify the efficiency of the proposed method, we also test the proposed overall method on JVET sequences [40]. Compared to JCTVC sequences, JVET sequences includes additional 8 UHD sequences (the resolutions are 4096x2160 and 3840x2160) with bit depth of 10. These UHD sequences are classified as Class A1 and Class A2. The coding gains of the proposed overall method on JVET sequences are given in Table III.

According to Table III, the proposed method can achieve 2.7% BD-rate reduction on average which is the same as the

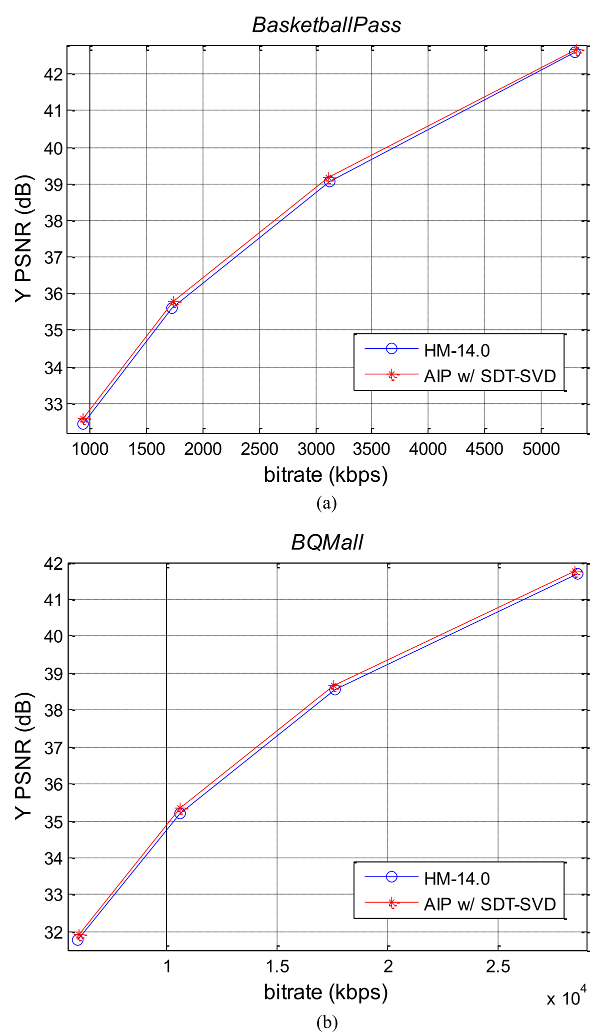


Fig. 10. Rate distortion curves for (a) *BasketballPass*, (b) *BQMall*.

TABLE III
PERFORMANCE OF THE PROPOSED METHOD ON JVET SEQUENCES

Class	Sequences	Resolution	Bit depth	Proposed overall
A1	Tango	4096 × 2160	10	−1.0%
	Drums100	3840 × 2160	10	−0.9%
	CampfireParty	3840 × 2160	10	−2.5%
	ToddlerFountain	4096 × 2160	10	−0.5%
A2	CatRobot	3840 × 2160	10	−5.4%
	TrafficFlow	3840 × 2160	10	−4.8%
	DaylightRoad	3840 × 2160	10	−5.1%
	Rollercoaster	4096 × 2160	10	−1.5%
	Average			−2.7%

gain for JCTVC sequences. Therefore, the proposed method is also efficient for UHD sequences.

C. Computational Complexity

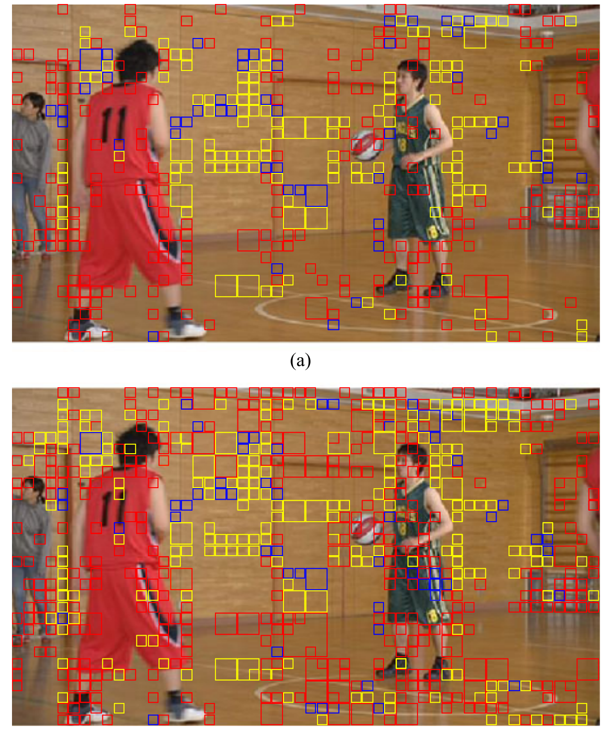
The coding complexity of the proposed method is also shown in Table II. The encoding/decoding time ratios between the proposed methods and anchor are given in this table. For the encoder, the complexity is 2.45 times for the proposed SDT-SVD on AIP and the encoding complexity for the proposed SDT-SVD on TMP is 3.04 times. For the proposed overall method, the encoding complexity is 4.48 times. The complexity of the proposed overall method is mainly from the increased RDO processes, the template matching process and the singular value decomposition for the derivation of SVD transforms. The complexity ratios for SDT [27] are also shown in Table II. It can be observed that the encoding complexity of SDT is 19.41 times which is much larger than the proposed method. The complexity for SDT [27] mainly comes from the template matching and the derivation of KLT kernels.

For the decoder, the complexity of the proposed AIP with SDT-SVD is 1.17 times and the complexity is 5.88 times for TMP with SDT-SVD. The complexity for the proposed overall method is 5.76 times. The decoding complexity mainly comes from the derivation of SVD and template matching. For SDT [27], since both the derivation of KLT and template matching need intensive computation, the decoding complexity is also much larger than the proposed method which is 13.22 times the one of HEVC.

D. Analysis of the Proposed Method

For the analysis of the proposed method, we get CU partitions of the first frame in sequence *BasketballPass* coded as intra at QP = 22 and 32 which are shown in Fig. 11. In Fig. 11, blocks marked with red boxes are coded by AIP with SDT-SVD, blocks marked with yellow boxes are coded by TMP without SDT-SVD and blocks marked with blue boxes are coded by TMP with SDT-SVD. Other regions without colored boxes are coded by original HEVC intra coding. We can see a large amount of blocks with strong edges are coded by the proposed SDT-SVD transform. This is because blocks with strong edges are more likely to generate residuals having similar structures as the corresponding synthetic blocks.

The usage rates of SDT-SVD under different QPs are shown in Table IV, respectively. From Table IV, the average usage rates of SDT-SVD in AIP w/ SDT-SVD are 20.0%, 20.0%, 19.7% and 16.8% when QP is 22, 27, 32 and 37, respectively. For TMP, Table V shows the average usage rates of SDT-SVD for TMP in TMP w/ SDT-SVD are 25.0%, 23.2%, 20.4% and 16.2% when QP is 22, 27, 32 and 37.



(a)

(b)

Fig. 11. Modes distributions for the first frame in sequence *BasketballPass*. The blocks marked with red boxes are coded by AIP with SDT-SVD, the blocks marked with yellow boxes are coded by TMP without SDT-SVD, and the blocks marked with blue boxes are coded by TMP with SDT-SVD. Other regions except these blocks with colored boxes are coded by original intra coding in HEVC. (a) QP = 22. (b) QP = 32.

TABLE IV
USAGE RATES OF SDT-SVD TRANSFORM IN AIP W/ SDT-SVD

Class	Sequences	QP = 22	QP = 27	QP = 32	QP = 37
A	Traffic	9.6%	9.1%	7.6%	5.4%
	PeopleOnStreet	10.6%	10.9%	10.3%	6.9%
B	Kimono	7.6%	5.2%	4.6%	2.9%
	ParkScene	14.8%	12.1%	9.4%	6.6%
	Cactus	27.4%	18.3%	14.4%	9.1%
	BasketballDrive	31.4%	26.7%	25.7%	25.1%
	BQTerrace	13.7%	23.7%	22.6%	18.1%
C	BasketabllDrill	21.0%	14.2%	11.0%	9.3%
	BQMall	31.5%	31.6%	31.0%	27.0%
	PartyScene	35.1%	37.6%	41.4%	41.5%
	RaceHorses	14.6%	17.1%	18.7%	17.6%
D	BasketballPass	27.8%	35.1%	36.7%	27.4%
	BQSquare	33.1%	36.6%	39.2%	37.3%
	BlowingBubbles	28.4%	26.5%	25.8%	22.6%
	RaceHorses	14.6%	17.1%	18.7%	17.6%
E	FourPeople	10.6%	10.9%	10.3%	6.9%
	Johnny	15.9%	14.6%	14.1%	9.9%
	KristenAndSara	12.7%	13.3%	13.5%	11.4%
	Average	20.0%	20.0%	19.7%	16.8%

TABLE V
USAGE RATES OF SDT-SVD TRANSFORM IN TMP W/ SDT-SVD

Class	Sequences	QP = 22	QP = 27	QP = 32	QP = 37
A	Traffic	13.0%	14.6%	13.2%	9.8%
	PeopleOnStreet	23.1%	20.8%	16.5%	8.1%
B	Kimono	8.5%	8.9%	5.7%	3.4%
	ParkScene	13.7%	14.5%	13.0%	9.6%
	Cactus	39.4%	23.9%	20.1%	13.1%
	BasketballDrive	29.9%	26.0%	15.1%	10.8%
	BQTerrace	25.6%	31.2%	29.4%	20.1%
C	BasketablIDrill	24.1%	17.5%	9.1%	5.3%
	BQMall	31.7%	32.3%	32.4%	27.7%
	PartyScene	39.0%	40.2%	38.9%	34.0%
	RaceHorses	19.4%	16.9%	17.5%	17.7%
D	BasketballPass	26.1%	21.9%	15.3%	10.5%
	BQSquare	35.6%	21.8%	30.9%	34.1%
	BlowingBubbles	34.3%	24.7%	24.2%	19.9%
	RaceHorses	13.1%	17.1%	17.7%	12.1%
E	FourPeople	18.8%	25.1%	25.1%	20.9%
	Johnny	30.0%	28.5%	11.6%	8.7%
	KristenAndSara	24.1%	31.1%	30.6%	26.1%
	Average	25.0%	23.2%	20.4%	16.2%

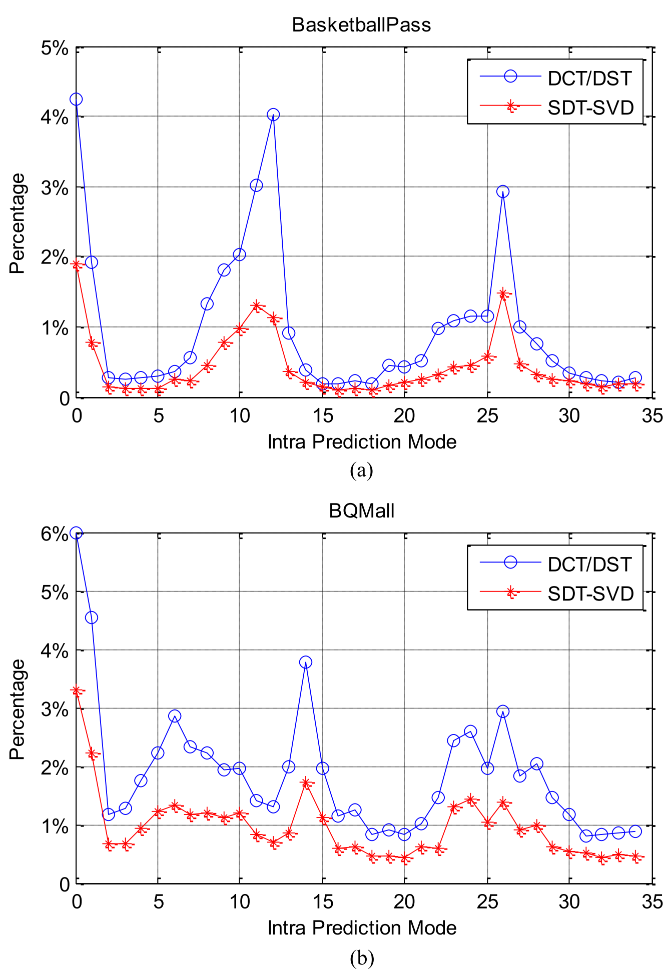


Fig. 12. Percentage of blocks transformed by DCT/DST and SDT-SVD in AIP w/ SDT-SVD for different intra prediction modes, when QP = 22. (a) *BasketballPass*. (b) *BQMall*.

In order to understand the efficiency of the proposed SDT-SVD for different intra prediction modes, the distributions of the modes for blocks coded by SDT-SVD and DCT/DST in AIP w/ SDT-SVD are given in Fig. 12 for *BasketballPass* and *BQMall*. According to this figure, the shapes of modes distributions for blocks transformed by SDT-SVD and DCT/DST are similar. This indicates that SDT-SVD is efficient for all the intra prediction modes.

VI. CONCLUSION

This paper investigated the residuals characteristics of intra prediction and the deficiency of DCT/DST for residual blocks. Based on the residual analysis, a signal dependent transform based on SVD (SDT-SVD) which is derived from the synthetic block is proposed to compact the residuals from angular intra prediction. The proposed SDT-SVD is also extended to template matching prediction. Experimental results suggest that the proposed coding method with SDT-SVD for AIP outperforms the latest HEVC reference software with a bit rate reduction of 1.0% on average and it can be up to 2.1%. When SDT-SVD is extended to TMP based intra coding, the overall bit rate reduction is 2.7% on average and can be up to 5.8%. In the future, fast algorithms, such as early skip for SDT-SVD transform and fast template searching method, can be exploited to reduce the complexity of the proposed method. To further improve the efficiency of the SDT-SVD transform, more work can be done to design an efficient block with more similar structures as the residual block for deriving SVD transforms.

REFERENCES

[1] H. Yue, X. Sun, J. Yang, and F. Wu, "Cloud-based image coding for mobile devices-toward thousands to one compression," *IEEE Trans. Multimedia*, vol. 15, no. 4, pp. 845–857, Jun. 2013.
[2] Y. Lu, S. Li, and H. Shen, "Virtualized screen: A third element for cloud mobile convergence," *IEEE Multimedia Mag.*, vol. 18, no. 2, pp. 4–11, Apr. 2011.
[3] W. Zhu, W. Ding, J. Xu, Y. Shi, and B. Yin, "Hash-based block matching for screen content coding," *IEEE Trans. Multimedia*, vol. 17, no. 7, pp. 935–944, Jul. 2015.
[4] W. Zhu, W. Ding, J. Xu, Y. Shi, and B. Yin, "Screen content coding based on HEVC framework," *IEEE Trans. Multimedia*, vol. 16, no. 5, pp. 1316–1326, Aug. 2014.
[5] L. Zhao, T. Lin, K. Zhou, S. Wang, and X. Chen, "Pseudo 2D string matching technique for high efficiency screen content coding," *IEEE Trans. Multimedia*, vol. 18, no. 3, pp. 339–350, Mar. 2016.
[6] H. Chen, A. Saxena, and F. Fernandez, "Nearest-neighboring intra prediction for screen content video coding," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2014, pp. 3151–3155.
[7] H. Chen, T. Zhang, M.-T. Sun, A. Saxena, and M. Budagavi, "Improving Intra prediction in high efficiency video coding," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3671–3682, Aug. 2016.
[8] G. J. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.
[9] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.
[10] J. Lainema, F. Bossen, W.-J. Han, J. Min, and K. Ugur, "Intra coding of the HEVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1792–1801, Dec. 2012.
[11] K. R. Rao and P. Yip, *Discrete Cosine Transform-Algorithms, Advantages and Applications*, New York, NY, USA: Academic, 1990.
[12] H. Chen and B. Zeng, "New transforms tightly bounded by DCT and KLT," *IEEE Signal Process. Lett.*, vol. 19, no. 6, pp. 344–347, Jun. 2012.

[13] H. Chen and B. Zeng, "Design of low-complexity, non-separable 2-D transforms based on butterfly structures," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2012, pp. 2921–2924.

[14] H. Chen, S. Zhu, and B. Zeng, "Design of non-separable transforms for directional 2-D sources," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2011, pp. 3697–3700.

[15] Y. Ye and M. Karczewicz, "Improved H.264 intra coding based on bi-directional intra prediction, directional transform, and adaptive coefficient scanning," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2008, pp. 2116–2119.

[16] X. Zhao, L. Zhang, S. W. Ma, and W. Gao, "Video coding with rate-distortion optimized transform," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 1, pp. 138–151, Jan. 2012.

[17] F. Zou *et al.*, "Rate-distortion optimized transforms based on the Lloyd-type algorithm for intra block coding," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 6, pp. 1072–1083, Dec. 2013.

[18] E. Alshina, A. Alshina, and F. Fernandes, "Rotational transform for image and video compression," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2011, pp. 3689–3692.

[19] C. Yeo, Y. H. Tan, Z. Li, and S. Rahardja, "Mode-dependent transforms for coding directional intra prediction residuals," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 4, pp. 545–554, Apr. 2012.

[20] J. Han, A. Saxena, V. Melkote, and K. Rose, "Jointly optimized spatial prediction and block transform for video and image coding," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1874–1884, Apr. 2012.

[21] A. Saxena and F. C. Fernandes, "DCT/DST-based transform coding for intra prediction in image/video coding," *IEEE Trans. Image Process.*, vol. 22, no. 10, pp. 3974–3981, Oct. 2013.

[22] A. Dapena and S. Ahalt, "A hybrid DCT-SVD image-coding algorithm," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 2, pp. 114–121, Feb. 2002.

[23] M. Wang, K. N. Ngan, and L. Xu, "Efficient H.264/AVC video coding with adaptive transforms," *IEEE Trans. Multimedia.*, vol. 16, no. 4, pp. 933–946, Jun. 2014.

[24] X. Cao and Y. He, "Singular vector decomposition based adaptive transform for motion compensation residuals," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2014, pp. 4127–4131.

[25] T. Zhang *et al.*, "Adaptive transform with HEVC intra coding," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf.*, Dec. 2015, pp. 388–391.

[26] C. Lan, J. Xu, G. Shi, and F. Wu, "Exploiting non-local correlation via signal-dependent transform (SDT)," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 7, pp. 1298–1308, Jul. 2011.

[27] C. Lan, J. Xu, and F. Wu, *Enhancement of HEVC Using Signal Dependent Transform (SDT)*, VCEG-AZ08, Warsaw, Poland, Jun. 2015.

[28] T. Zhang, H. Chen, M.-T. Sun, D. Zhao, and W. Gao, "Hybrid angular intra/template matching prediction for HEVC intra coding," in *Proc. IEEE Int. Conf. Visual Commun. Image Process.*, Dec. 2015, pp. 1–4.

[29] T. K. Tan, C. S. Boon, and Y. Suzuki, "Intra prediction by template matching," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2006, pp. 1693–1696.

[30] K. Ugur and O. Bici, *Performance Evaluation of DST in Intra Prediction*, ITU-T and ISO/IEC, Doc. JCTVC I0582, Geneva, Switzerland, May 2012.

[31] H. C. Andrews and C. L. Patterson, "Singular value decomposition (SVD) image coding," *IEEE Trans. Commun.*, vol. 24, no. 4, pp. 425–432, Apr. 1976.

[32] P. Waldemar and T. A. Ramstad, "Image compression using singular value decomposition with bit allocation and scalar quantization," in *Proc. NORSlG Conf.*, 1996, pp. 83–86.

[33] S.-L. Yu and C. Chrysafis, *New Intra Prediction Using Intra Macroblock Motion Compensation,* 3rd Meet. Joint Video Team, Doc. JVT-C151, May 2002.

[34] J. Xu, R. Joshi, and R. A. Cohen, "Overview of the emerging HEVC screen content coding extension," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 1, pp. 50–62, Jan. 2016.

[35] H. Chen, Y. Chen, M. T. Sun, A. Saxena, and M. Budagavi, "Improvements on intra block copy in natural content video coding," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2015, pp. 2772–2775.

[36] J. Xu, A. Tabatabai, and O. Nakagami, *On Intra Block Copying in RExt*, ITU-T and ISO/IEC, Doc. JCTVC-O0232, Geneva, Switzerland, Oct. 2013.

[37] M. Budagavi, A. Fuldseth, G. Bjontegaard, V. Sze, and M. Sadafale, "Core transform design in the high efficiency video coding (HEVC) standard," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 6, pp. 1029–1041, Dec. 2013.

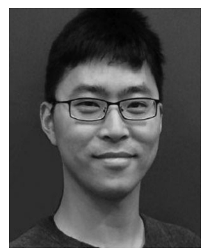
[38] F. Bossen, *Common Test Conditions and Software Reference Configurations*, ITU-T and ISO/IEC, Doc. JCTVC-B300, Geneva, Switzerland, Jul. 2010.

[39] G. Bjøntegaard, *Calculation of Average PSNR Differences Between RD-Curves*, ITU-T Q.6/SG16, Doc. VCEG-M33, 2001.

[40] K. Suehring and X. Li, *JVET Common Test Conditions and Software Reference Configurations*, 2nd Meet. Joint Video Exploration Team ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-B1010, Feb. 2016.

Tao Zhang received the B.S. degree from Xidian University, Xi'an, China, in 2010, and the M.S. degree and Ph.D. degree from the Harbin Institute of Technology, Harbin, China, in 2012 and 2017, respectively, all in computer science.

He is currently with Tencent, Shenzhen, China. From September 2012 to September 2013, he was with Microsoft Research Asia, Beijing, China, as an Intern. From September 2014 to September 2015, he was a visiting student in the Information Processing Laboratory, University of Washington, Seattle, WA, USA. His research interests include video coding and image processing.



Haoming Chen received the B.S. degree from Nanjing University, Nanjing, China, in 2010, the M.Phil. degree from the Hong Kong University of Science and Technology, Hong Kong, China, in 2012, and the Ph.D. degree from the University of Washington, Seattle, WA, USA, in 2016, all in electrical engineering.

He is currently with Google, Kirkland, WA, USA. He has interned with Samsung Research America, Richardson, TX, USA, in 2013, and Apple, Cupertino, CA, USA, in 2015. His research interests include video coding, multimedia communication, and image processing.

Dr. Chen was the recipient of the Top 10% Paper Award at IEEE ICIP 2014, and was the Student Best Paper Award Finalist at the IEEE International Symposium on Circuits and Systems 2012.

Ming-Ting Sun (S'79–M'81–SM'89–F'96) received the B.S. degree from National Taiwan University, Taipei, Taiwan, in 1976, and the Ph.D. degree from the University of California, Los Angeles, Los Angeles, CA, USA, in 1985, both in electrical engineering.

In 1996, he joined the University of Washington, Seattle, WA, USA, where he is a Professor. Previously, he was the Director of Video Signal Processing Research at Bellcore, Red Bank, NJ, USA. He has been a Chaired/Visiting Professor with Tsinghua University, Beijing, China; Tokyo University, Bunkyō, Japan; National Taiwan University, Taipei, Taiwan; National Cheng Kung University, Tainan, Taiwan; National Chung Cheng University, Chiayi County, Taiwan; National Sun Yat-Sen University, Kaohsiung, Taiwan; Hong Kong University of Science and Technology, Hong Kong, China; and National Tsing Hua University, Hsinchu, Taiwan. He holds 13 patents and has authored or coauthored more than 200 technical papers, including 17 book chapters in the area of video and multimedia technologies. He co-edited the book *Compressed Video Over Networks* (CRC Press, 2000). His research interests include video and multimedia signal processing, and transport of video over networks.

Dr. Sun was a guest editor for 12 Special Issues for various journals and given keynotes for several international conferences. He has served in many Awards Committees. He was a Technical Program Co-Chair for several conferences including the International Conference on Multimedia and Expo (ICME) 2010. He was the Editor-in-Chief of the IEEE TRANSACTIONS ON MULTIMEDIA and a Distinguished Lecturer of the Circuits and Systems Society from 2000 to 2001. He was the General Co-Chair of the Visual Communications and Image Processing 2000 Conference. He was the Editor-in-Chief of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY (TCSVT) from 1995 to 1997. He served as the Chair of the Visual Signal Processing and Communications Technical Committee of the IEEE Circuits and Systems (CAS) Society from 1993 to 1994. From 1988 to 1991, he was the Chairman of the IEEE CAS Standards Committee and established the IEEE Inverse Discrete Cosine Transform Standard. He is a General Co-Chair of the IEEE ICME 2016, Seattle, WA, USA. He was the recipient of an Award of Excellence from Bellcore for his work on the digital subscriber line in 1987, the TCSVT Best Paper Award in 1993, and the IEEE CASS Golden Jubilee Medal in 2000.

Debin Zhao (M'11) received the B.S., M.S., and Ph.D. degrees in computer science from the Harbin Institute of Technology (HIT), Harbin, China, in 1985, 1988, and 1998, respectively.

He is currently a Professor with the Department of Computer Science, HIT. He has authored or coauthored more than 200 technical articles in refereed journals and conference proceedings. His research interests include the areas of image and video coding, video processing, video streaming and transmission, and pattern recognition.

Wen Gao (S'87–M'88–SM'05–F'09) received the Ph.D. degree in electronics engineering from the University of Tokyo, Tokyo, Japan, in 1991.

He is currently a Professor of computer science with Peking University, Beijing, China. Before joining Peking University, he was a Professor of computer science with the Harbin Institute of Technology, Harbin, China, from 1991 to 1995, and a Professor in the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China. He has authored or coauthored 5 books and more than 600 technical articles in refereed journals and conference proceedings. His research interests include the areas of image processing, video coding and communication, pattern recognition, multimedia information retrieval, multimodal interface, and bioinformatics.

Prof. Gao has served or currently serves on the Editorial Board for several journals, such as the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE TRANSACTIONS ON AUTONOMOUS MENTAL DEVELOPMENT, the *EURASIP Journal of Image Communications*, and the *Journal of Visual Communication and Image Representation*. He chaired a number of prestigious international conferences on multimedia and video signal processing, such as the IEEE ICME and ACM Multimedia, and also served on the advisory and technical committees of numerous professional organizations.