# Rearrangement Pixel Granularity Template Matching for Lossy Screen Content Picture Intra Coding

Zheng Wang[12], Pin Tao[12], Lixin Feng[12]

[1]*Department of Computer Science and Technology, Tsinghua University*
*Beijing 100084, China*

[2]*Tsinghua National Laboratory for Information Science and Technology*
*Beijing 100084, China*

[12]wangzheng14@mails.tsinghua.edu.cn
[12]taopin@tsinghua.edu.cn
[12]fenglixin@gmail.com

***Abstract*— Screen Content (SC) video coding is becoming more important in screen sharing and screen broadcasting applications. There are many distinctly different characteristics between the SC video and the camera-captured video. In this paper, we propose a rearrangement pixel granularity template matching method for lossy SC intra picture coding with higher compression ratio. Firstly, the original picture will be rearranged for lossy compression and parallel processing purpose. Then every pixel is predicted with template matching method. The template matching method can achieve high prediction accuracy on SC pictures. Our proposed method improves about 4.5dB and saves 11.7% bitrate in comparison with the HEVC range extension software HM12.0+RExt4.1.**

***Index Terms*— Screen Content Coding, Template Matching, Lossy Compression, High Efficiency Video Coding**

## I. Introduction

Screen Content (SC) video coding has become more important as a result of the rapid advances in mobile computing and cloud computing. Users demand to share the screen visuals with different devices in many applications, such as gaming broadcasting, wireless display, screen mirroring, and remote assistant, etc. The screen content picture has some significant differences compared with the camera captured picture, such as generated by computer program without any signal noise, a lot of discontinuous textures, and small colour palette for one picture, etc. The state-of-the-art video coding standards, such as AVC/H.264 and High Efficiency Video Coding (HEVC)/H.265, were mainly targeted on the camera-captured video sequences. These commonly-used transform-based video codecs always bring the artifacts after encoding, and the coding performance for high-fidelity screen picture is not satisfied. Joint Collaborative Team on Video Coding (JCT-VC), which is the standardization working group of HEVC/H.265 [11], has established an ad-hoc group about screen content coding since 2010, and a new software extension named Range Extensions (RExt) [12] for HEVC has been developed since January 2014. Recently, the Video Electronics Standards Association also completed a Display Stream Compression (DSC) [13] standard for next-generation display interfaces.

As the screen content video contains many exact-same patches in one picture, template matching method was proposed to discover the similarities. At first, template matching was adopted by Kazuo Sugimoto et al. [1] for inter picture prediction in 2004, and then for intra picture prediction by Thiow Keng Tan et al. [2] in 2006. By changing candidates of template matching, several works [3, 4, 5, 6, 7, 8] further improved the compression ratio. [3] researched on the inter frame coding for screen content video. However, most researches were focused on the screen content intra coding as the intra pictures play the most important role in the whole sequence. Normally, the target block size of template matching is 2x2 at least, commonly 4x4 or larger. The 1x1 size, which means pixel by pixel, can achieve higher matching accuracy, but cannot do the transform for lossy compression. Besides the target block, the template size as well as shape also plays an important role in template matching. Dynamic template can improve the coding efficiency [11] although most previous research works used fixed template. Furthermore, strategies to choose the prediction value among candidates deserves studying, such as choosing averaged value[5] as predicted value, choosing the best value [4] or using linear model [6].

In this paper, we propose a rearrangement pixel granularity template matching method for screen content intra coding, the picture rearrangement of our proposal gives the codec ability to do the lossy compression, and the pixel granularity operation can achieve more accurate template matching. Section 2 presents the proposed template matching scheme in detail. Section 3 gives the experiment results and section 4 draws conclusions and describes possible future work.

## II. Proposed Framework

Our proposed Pixel Granularity Template Matching &Rearrangement (PGR) method framework is shown in Fig. 1. The framework is only for SC intra coding. For one SC picture, all pixels of the picture should do rearrangement firstly. The 64x64 block after rearrangement is called Virtual LCU (VLCU). The rearrangement, which will be introduced in detail later, enables lossy compression and parallel processing. Then for each VLCU, all pixels will be predicted with the pixel granularity template matching method and other prediction methods as well. The VLCU will be divided into sub-CUs and each CU can select different prediction method according to rate-distortion optimization (RDO) algorithm. After prediction, most of the prediction residuals are zeros, and the other few non-zero residuals can be quantized. The final step is entropy coding to generate the output bit stream. The inverse quantization can reconstruct the VLCU buffer for the future template matching operations on the same intra picture.

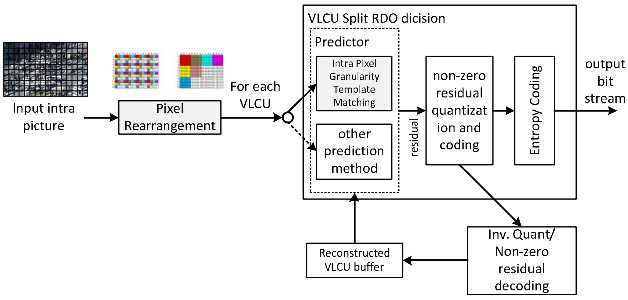


Fig.1 the framework of PGR

In our proposed framework, the most important two steps are picture rearrangement and pixel granularity template matching. In order to explain why we design the framework like this, we should explain the potential problems of the traditional intra coding. The commonly-used prediction method for intra picture was block based approaches. Currently, the state-of-the-art video standard, HEVC, uses angular intra prediction which has up to 35 prediction directions as shown in Fig. 2. The intra prediction procedure is based on linear interpolation of pixel samples from the left and upper neighbouring Prediction Unit (PU) [10]. It is clearly to see that the left and upper neighbour pixels is hard to predict the right bottom pixels of current CU accurately, especially when the size of current CU is 8x8 or larger, or the current CU contains complex textures.

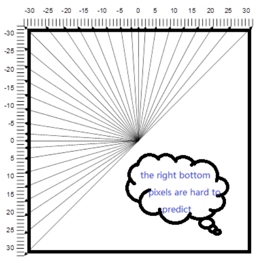


Fig. 2 angular intra prediction is hard to predict the right bottom pixels of current CU

The template matching method seems to have the potential to find the best similar copy for the current CU, but it faces the dilemma problem. If the target block size of template matching is large, for example the size is 8x8, wrong predictions may occur as shown in Fig.3. The red block is the current coding block. By using the surrounding blue block as the template, it will find the block containing "VLC" which is supposed to be the best match template. Then the green area will be used as the best pre-diction. But the "s" after "U" in the green area is the wrong prediction as the corresponding area of the red current block is "i". And it is easy to find that the smaller the size of target block, the prediction performance is better. In the extreme case, the 1x1 target size is best. That means the pixel granularity template matching is the best choice.

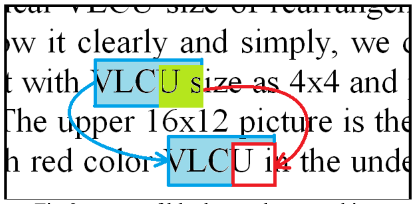

Fig.3 wrong of block template matching

However pixel granularity template matching is not fit for previous intra coding suite. As shown in Fig.4, if the codec predict pixel by pixel, it is okay for lossless coding but will cause the serious error for lossy coding, because the previous pixels in the same CU have not been reconstructed, the prediction values of them cannot be used as the part of the template of the next pixel. These slight differences caused by quantization and inverse-quantization may strongly impact the template matching results and issue the prediction error propagation in the encoder and decoder side.

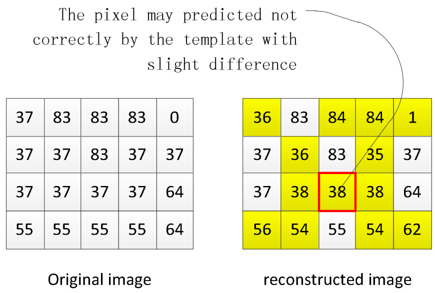


Fig. 4 error caused by pixel granularity template matching in lossy coding

### *A. Rearrangement of intra picture*

In order to solve the above dilemma problem, we propose the Pixel Granularity Template Matching & Rearrangement (PGR) method. After rearrangement, the pixel granularity template matching becomes reasonable which could improve the performance of prediction greatly.

The practical VLCU size of rearrangement is 64x64. In order to show it clearly and simply, we demonstrate the rearrangement with VLCU size as 4x4 and picture size as 16x12 in Fig.5. The upper 16x12 picture is the original picture. All pixels with red colour denoted as "a" in the upper original picture will be grouped together into a new 4x4 VLCU in the rearrangement picture below. So do other pixels with grey colour denoted as "b", cyan colour denoted as "c" and so on. After rearrangement procedure, the original picture is changed into a new picture, and for each VLCU in the new picture, all pixels in one VLCU can be predicted with template matching individually and simultaneously. The individually predicting means the template matching procedure of each pixel in a VLCU is independent. It will not cause the reconstruction error and will not cause the error propagation. The simultaneously predicting means all pixels in one VLCU can be predicted in parallel. It allows our proposed method can be speedup in the multicore processor or GPU in the future.

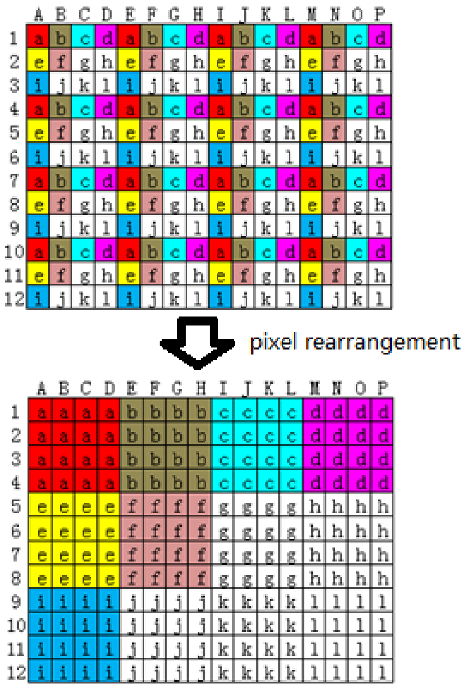


Fig. 5 rearrangement in VLCU

Generally, if the picture size is N*M where N and M is not exactly the multiples of LCU size T, the rearranged picture size will be ((r(N/T)+1)*T) * ((r (M/T)+1)*T), and r(x) denote as the round down operation. For example, the picture size is 1080*720 and T is 64, then the rearranged picture size is 1088*768.

For each pixel $P_{p,q}$ in rearranged picture, it actually comes from the $P_{i,j}$ of the original picture, we get that pixel in original picture by calculating i and j with formula (1)-(6):

$$i = (p\%T)*(r(M/T)+1) + r(p/T) \quad \text{when } (p\%T) < (M\%T) \tag{1}$$

$$i = (M\%T)*(r(M/T)+1) + ((p\%T)-(M\%T))*r(M/T) + r(p/T) \quad \text{when } (p\%T) >= (M\%T) \text{ and } p<M \tag{2}$$

$$i \text{ is not valid, and } R_{p,q}=0 \quad \text{when } p>=M \tag{3}$$

$$j = (q\%T)*(r(N/T)+1) + r(q/T) \quad \text{when } (q\%T) < (N\%T) \tag{4}$$

$$j = (N\%T)*(r(N/T)+1) + ((q\%T)-(N\%T))*r(N/T) + r(q/T) \quad \text{when } (q\%T) >= (N\%T) \text{ and } q<N \tag{5}$$

$$j \text{ is not valid, and } R_{p,q}=0 \quad \text{when } q>=N \tag{6}$$

## B. *Pixel granularity template matching*

For each pixel in a VLCU in the rearrangement picture, we can derive its position in the original picture by formula (1)-(6). The template matching method is applied in original picture pixel order, and pixel reconstruction is applied in rearrangement picture pixel order. In this way, before predicting a certain pixel value, the left and above pixels in the original picture of this pixel which compose its template have already been reconstructed. As shown in Fig. 6, the red pixel denoted as “x” of current VLCU which needs to be predicted, some left and above blue pixels act as its template. Template matching procedure will find the most similar area of this template, and use the value of pixel with corresponding position as the prediction of pixel “x”.

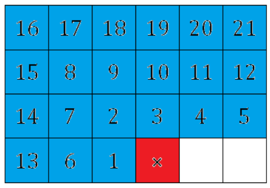


Fig. 6 pixel and its template in the original picture

In this paper, we design the template with 21 pixels as shown in Fig.6. For rapid matching, all 21 pixels are divided into 8 groups from $G_1$ to $G_8$,

| | | | |
|---|---|---|---|
| $G_1$ | 1 | $G_5$ | 6, 7, 13, 14 |
| $G_2$ | 2 | $G_6$ | 8, 15, 16, 17 |
| $G_3$ | 3 | $G_7$ | 9, 10, 18, 19 |
| $G_4$ | 4, 5 | $G_8$ | 11, 12, 20, 21 |

We calculate the average value of each group, and only the highest 3 bits of the average value of each group are used. All extracted bits of 8 groups compose a 3x8=24-bits vector to represent the template as a hash value. The hash function can speed up the matching process.

For a 64x64 VLCU, we should notice that all 4096 pixels can be predicted in parallel. This feature can be used to speed up the template matching algorithm on multicore processor or GPU platform.

Besides predicting with template matching, the codec can also use other prediction method, e.g. predicting with the left pixel, predicting with the upper pixel, predicting with the average value of nearby three pixels, etc. Then the 64x64 VLCU can be divided into four 32x32 sub CUs, and 32x32 CU can be divided into more smaller sub CUs based on the rate distortion optimization. Each sub CU can use the different prediction method, including the template matching prediction method. For SC picture, most pixels can be predicted exact correctly and most residuals are zero. Then the non-zero residuals will be quantized by dividing a quantization parameter. After quantization, the position and value of non-zero residuals will be coded with the entropy encoder.

## III. Experiment Results

We implemented the proposed method and the results are compared with the HEVC reference software HM12.0+RExt4.1. RExt4.1 is the special extension of HEVC for screen content picture. The WebBrowsing sequence shown in Fig. 7 is used as the test sequence which is a combination of typical screen content and the camera capture image. The test condition of HM12.0+RExt4.1 is all-intra coding configuration, and flag TransquantBypassEnableFlag and CUTransquantBypassFlagForce are both set to 1. Four QPs are 17, 22, 27, 32. And the lossless compressing is also tested. The PGR algorithm proposed in this paper is implemented as follows. The First VLCU is coding with the standard HEVC codec as all pixels in it cannot do template matching prediction. Then following VLCUs can be predicted with following four methods, the pixel granularity template matching, predicting with the left pixel, predicting with the upper pixel and predicting with the average of nearby three pixels. These four methods are applied on each level of sub CU and choose the best partition and best prediction method based on rate-distorting optimization. The simple quantization parameter is selected from 0, 2, 4, 8, to 16.

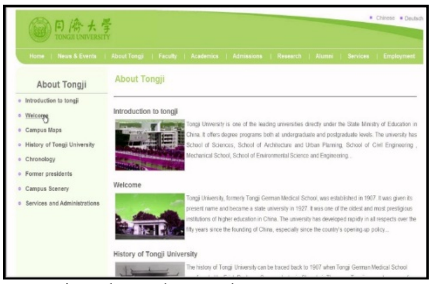

Fig.7 the WebBrowsing test sequence

The lossless compression experiment results were shown in table I. Our proposed method saves about 12% bitrate compared with HM12.0+RExt4.1.

TABLE II
EXPERIMENT RESULT OF LOSSLESS COMPRESSION

| HM12.0+RExt4.1 | | Proposed PGR | | Bits saving |
|---|---|---|---|---|
| Quant Param | Bits | Quant Param | Bits | |
| 0 | 1217347 | 0 | 1069472 | 12% |

TABLE Ⅱ
EXPERIMENT RESULT OF LOSSY COMPRESSION

| HM12.0+RExt4.1 | | | Proposed PGR | | | BD-Bit rate | BD-PSNR |
|---|---|---|---|---|---|---|---|
| Quant Param | Bits | Y-PSNR | Simple Quant Param | Bits | Y-PSNR | | |
| 22 | 584832 | 54.50 | 2 | 641740 | 59.52 | -11.7% | 4.5dB |
| 27 | 425146 | 42.84 | 4 | 505725 | 52.40 | | |
| 32 | 412539 | 41.25 | 8 | 402956 | 44.75 | | |
| 37 | 317957 | 33.01 | 16 | 332185 | 39.20 | | |

In lossless compression, over 95% pixels are predicted with pixel granularity template matching method based on RDO, and over 86% prediction residuals are zero. In lossy compression, when the simple quantization parameter is 2 or 4, the percentage of pixels adopted with template matching and the percentage of zero prediction residual have no obvious changes compared with the lossless compression. The experiment results are shown in table Ⅱ. The BDPSNR improvement is 4.5dB and the BD-Bitrate saving is 11.7%. The improvement of our proposed PGR method is strong for high quality intra coding.

## IV. Conclusions

In this paper, we proposed the rearrangement pixel granularity template matching method for lossy screen content intra coding. By taking advantages of rearrangement and pixel by pixel template matching, our proposed method can achieve about 4.5dB and saving 11.7% bitrate compared with the HEVC range extension software. And the proposed method is parallelization friendly and has future potential to get the better performance by improving the template design and better entropy encoder.

## Acknowledgment

This work was supported by National Significant Science and Technology Projects of China under Grant No. 2013ZX01039-001-002-003, and thanks to Microsoft Research Asia for the supporting on this research work.

## References

[1] K. Sugimoto, M. Kobayashi, Y. Suzuki, Sadaatsu Kato, and Boon Choong Seng, “Inter frame coding with template matching spatio-temporal prediction,” in Image Processing, 2004. ICIP ’04. 2004 International Conference on, vol. 1, pp. 465–468 Vol. 1.

[2] T. K. Tan, C. S. Boon, and Y. Suzuki, “Intra prediction by template matching,” in Image Processing, 2006 IEEE International Conference on, pp. 1693–1696.

[3] Lan Cuiling, Xu Jizheng, Wu Feng, and Shi Guangming, “Intra frame coding with template matching prediction and adaptive transform,” in Image Processing (ICIP), 2010 17th IEEE International Conference on, pp. 1221– 1224.

[4] M. Moinard, I. Amonou, P. Duhamel, and P. Brault, “A set of template matching predictors for intra video coding,” in Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on, pp. 1422–1425.

[5] Tan Thiow Keng, Boon Choong Seng, and Y. Suzuki, “Intra prediction by averaged template matching predictors,” in Consumer Communications and Networking Conference, 2007. CCNC 2007. 4th IEEE, pp. 405–409.

[6] Zheng Yunfei, Yin Peng, O. D. Escoda, Li Xin, and C. Gomila, “Intra prediction using template matching with adaptive illumination compensation,” in Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on, pp. 125–128.

[7] Hui Su, Jingning Han, and Yaowu Xu, “An optimized template matching approach to intra coding in video/image compression,” in Visual Information Processing and Communication V, February 5-6 2014, vol. 9029.

[8] E. Wige, G. Yammine, P. Amon, A. Hutter, and A. Kaup, “Sample-based weighted prediction with directional template matching for hevc lossless coding,” in Picture Coding Symposium (PCS), 2013, pp. 305–308.

[9] Lixin Feng, Pin Tao, Daxin Guang, Jiangtao Wen, and ShiQiang Yang, Pixel Granularity Template Matching Method for Screen Content

Lossless Intra Picture, vol. 8879 of Lecture Notes in Computer Science, book section 15, pp. 143–152, Springer International Publishing, 2014.
[10] Zhou Minhua, Gao Wen, Jiang Minqiang, and Yu Haoping, “Hevc lossless coding and improvements,” Circuits and Systems for Video Technology, IEEE Transactions on, vol. 22, no. 12, pp. 1839–1843, 2012.
[11] W. S. Lin, Gao Yang, and K. J. R. Liu, “Template matching for image prediction: A game-theoretical approach,” in Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on, pp. 821–824Sullivan G J, Ohm J R, Han W J, et al. Overview of the high efficiency video coding (HEVC) standard[J]. Circuits and Systems for Video Technology, IEEE Transactions on, 2012, 22(12): 1649-1668.
[12] Sullivan G J, Boyce J M, Chen Y, et al. Standardized extensions of high efficiency video coding (HEVC)[J]. Selected Topics in Signal Processing, IEEE Journal of, 2013, 7(6): 1001-1016.
[13] Kerofsky L J. Display stream compression: U.S. Patent Application 13/914,314[P]. 2013-6-10.