FAST TEMPLATE MATCHING FOR INTRA PREDICTION

Gayathri Venugopal
Philipp Merkle
Detlev Marpe
Thomas Wiegand

Fraunhofer HHI

Overview

- Introduction
- Conventional Template Matching (TM) for Intra Prediction
- Computational complexity of TM
- Proposed Fast Template Matching
- Test conditions and Experimental results
- Conclusions

Introduction

- The state-of-the-art H.265/HEVC has **21.9 %** BD-rate gain compared to its predecessor H.264/AVC with All Intra main configuration.
- The intra coding tools in HEVC are more advanced than AVC.

Functionality	AVC	HEVC
Luma Prediction block sizes	4x4, 8x8 and 16x16	4x4, 8x8, 16x16 and 32x32
Number of Luma intra prediction modes	9 (4x4 and 8x8), 4 (16x16)	35
Number of Chroma intra prediction modes	4	4 + luma mode
Reference sample smoothing	8x8	8x8 and above
Boundary smoothing	N/A	For HOR, VER and DC modes
Operation when reference samples missing	Use DC mode	Reference sample substitution
Number of most probable modes in mode coding	1	3

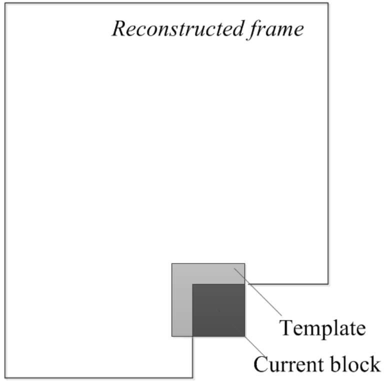
Introduction

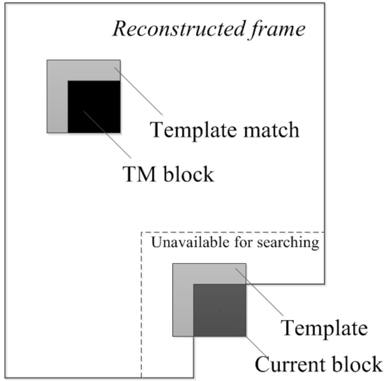
- Template Matching (TM) is a texture synthesis technique used in image processing.
- TM was proposed for H.264/AVC intra prediction with upto 11% BD-rate gain.

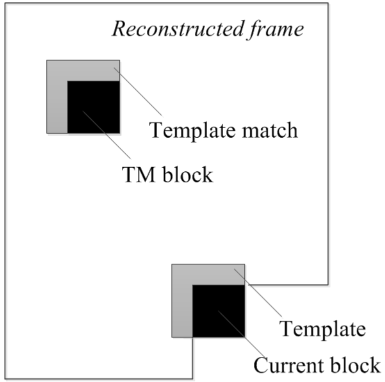
- *Problem with TM:* Substantial **computational complexity**, especially for **decoder**, from extensive search algorithm makes TM less attractive.

Conventional template matching intra mode

What is TM intra prediction mode ?







Consider the reconstructed frame. We have a block to be predicted (i.e the current block). The samples above and left of the current block are considered as the *template* (rotated L-shaped).

We search for the best template match by minimizing the error metric SSD. The corresponding block of the best template match is called the **TM block or TM predictor**.

TM block is the prediction of the current block. The decoder should also repeat the search method for finding the predictor as **no side information** is sent.

Gayathri Venugopal

Conventional template matching intra mode

Pro’s and Con’s

Displacement Vectors are not signaled to the decoder. Hence saves bits.

Increased decoder run-time due to search algorithm for finding predictors.

Conventional template matching intra mode

Averaged superposition

- What if we take the first ‘k’ template matches using the error minimizing metric?
- Then we will have ‘k’ number of predictors.
- The **average of the k TM blocks** is the final prediction in averaged superposition mode.
- Our experimental results show that k = 3 is a nearly optimal configuration.

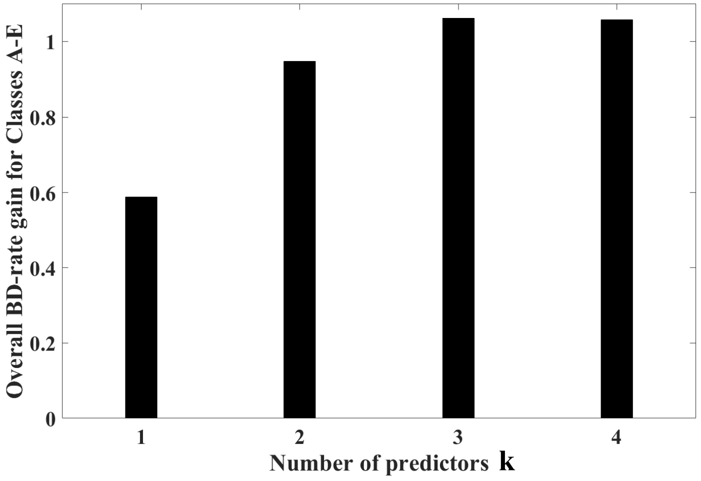


Fig.1 : Averaged Superposition

Previous researches show that averaged superposition improves accuracy.
Our research also confirms this behavior.

Computational complexity of TM

Limitation of conventional TM

- Searching for the template match leads to an increased run-time, for both encoder and decoder.
- A practical solution is using a search window. Let **(M+Δ)×(M+Δ)** be the search window size, where Δ×Δ region is not available for searching.
- What is the best value for M ?

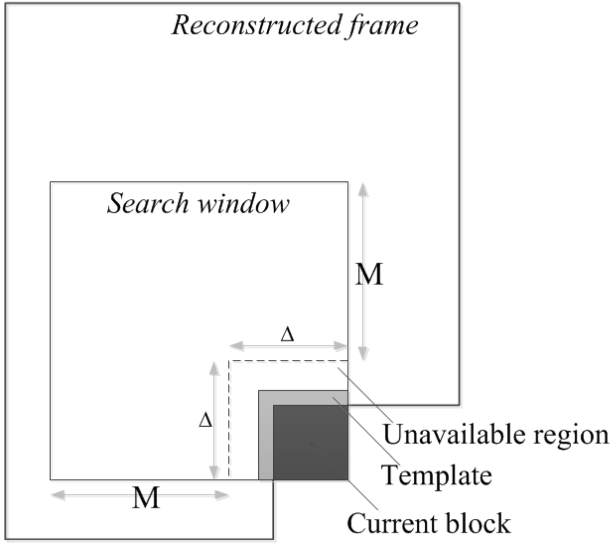


Fig.2 : Search window for Template Matching

Gayathri Venugopal

Computational complexity of TM

BD-rate gain and run-time increase for different values of ‘M’

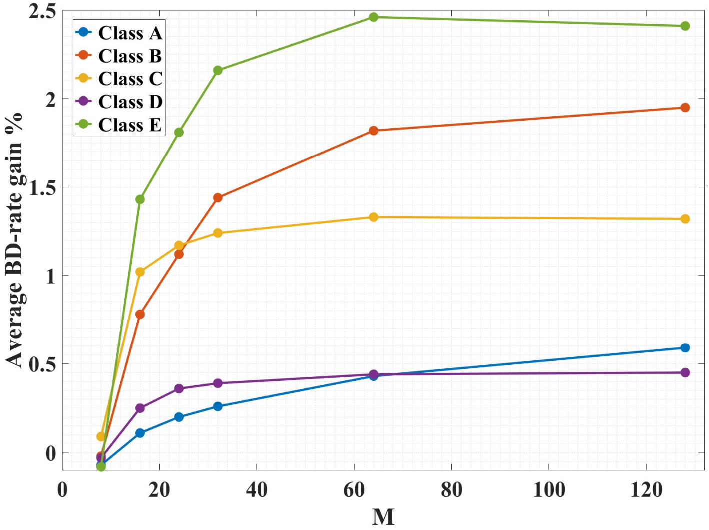


Fig.3 : Search window size ‘*M* ’ Vs coding efficiency

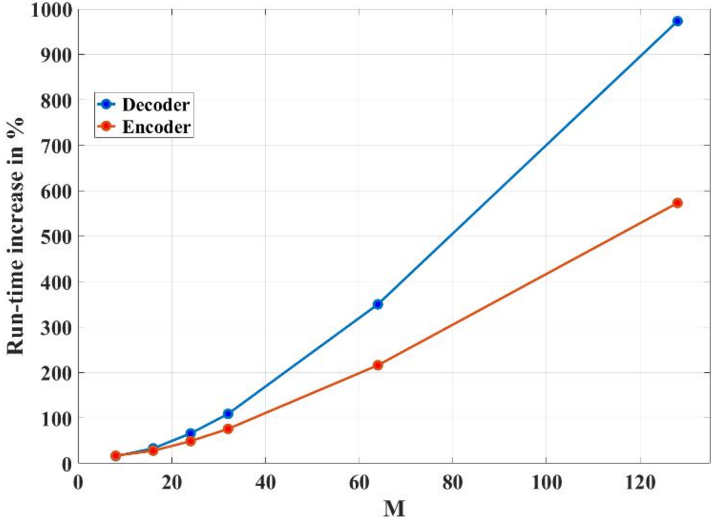


Fig.4 : Search window size ‘*M*’ Vs run-time

The results are obtained by implementing TM on HM16.6 as an additional intra mode. We have used the JCT-VC test conditions with All Intra configuration.

Gayathri Venugopal

Fraunhofer HHI

Computational complexity of TM

For M = 64,

Overall BD-rate gain = -1.26%

Run-time increase,

Decoder = 350%

Encoder = 216%

Not acceptable for practical applications

How to find a good trade-off between coding gain and run-time with TM?

Gayathri Venugopal

Analysis on conventional TM intra prediction

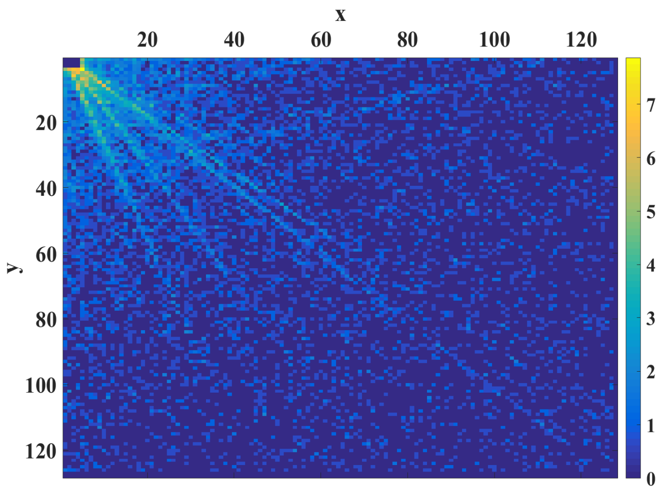


Fig.5 : Position of the first best TM block relative to the current block.
Histogram of the number of TM blocks wrt to current block inside 128x128 window (BasketBallDrill with QP = 22)

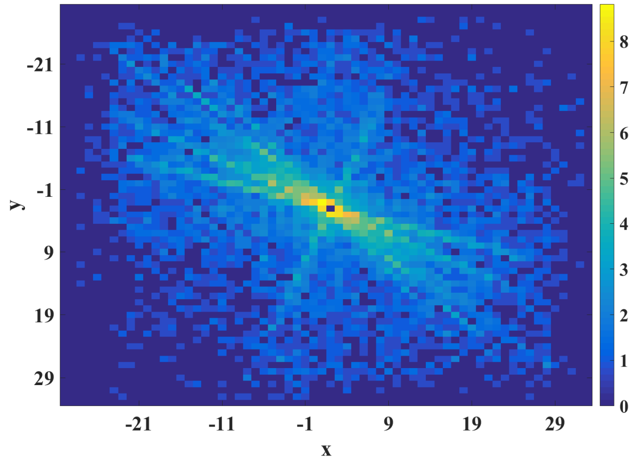


Fig.6 : Position of the second best TM block relative to the first best TM block.
Histogram of number of 2nd best TM block wrt to 1st TM block inside 64x64 window (BasketBallDrill with QP = 22)

- TM blocks are more often present in the immediate neighbourhood of the current block

- k>0 predictors are present close k=0 predictor

Gayathri Venugopal

Fraunhofer HHI

Proposed Fast Template Matching intra mode

- In our fast TM intra prediction mode, the search window is divided into **5 regions**.
- The width of each region, $n = \frac{M}{3}$

Encoder

- We find the 3 best TM blocks from each region through template matching and apply averaged superposition.
- The region that minimizes the RD cost function, ***J***, is chosen as the best region

$J = D + \lambda R$, where D is the distortion between the original and predicted blocks, and R is the number of bits for signaling the associated region

- We send the information indicating the best region to the decoder.

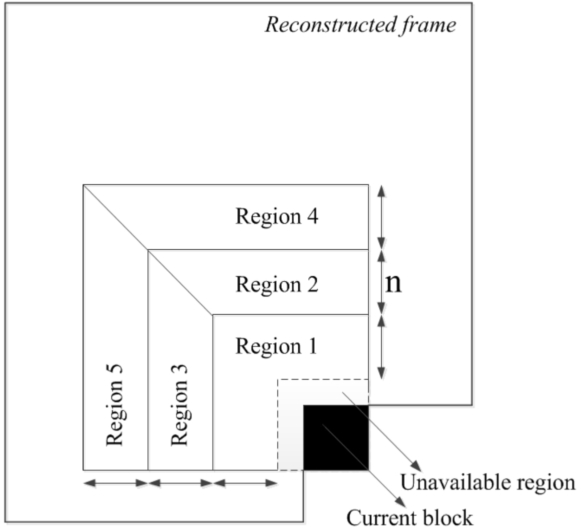


Fig.7 : Fast TM Intra mode search regions

Proposed Fast Template Matching

Signaling

MPMFlag – Most Probable Mode Flag
TMFlag – TM Mode Flag
RegionFlag – Most Probable Region Flag

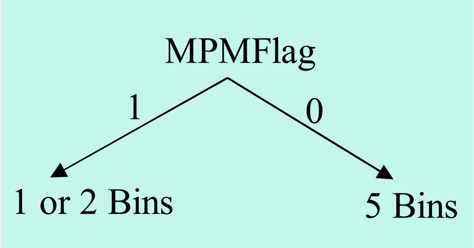


Fig.8 : HEVC Signaling

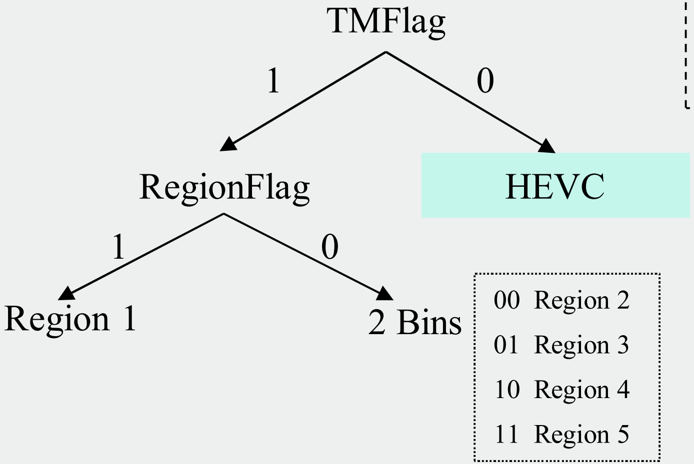


Fig.9 : HEVC + Fast TM Signaling

Decoder

- Runs search algorithm **only** in the signaled region for finding the predictors.

Proposed Fast Template Matching

Adaptive search region

- The search region size is adapted based on,

  (a) The **size of the block** to be predicted

  (b) The **frame width** of the input video

Block size, N×N	**Frame-width, w**	**Region width, n**
4, 8	$0 \leq w \leq 832$ $832 < w < 1280$ $1280 \leq w < 2560$ otherwise	A 2 x A 3 x A 4 x A
16, 32	always	2 x A

Table 1 : Adaptive search window for Fast TM

- The aspect ratio of the search window is kept constant always (1:1).
- What is the best value of A ?

Investigation on the value of ‘A’

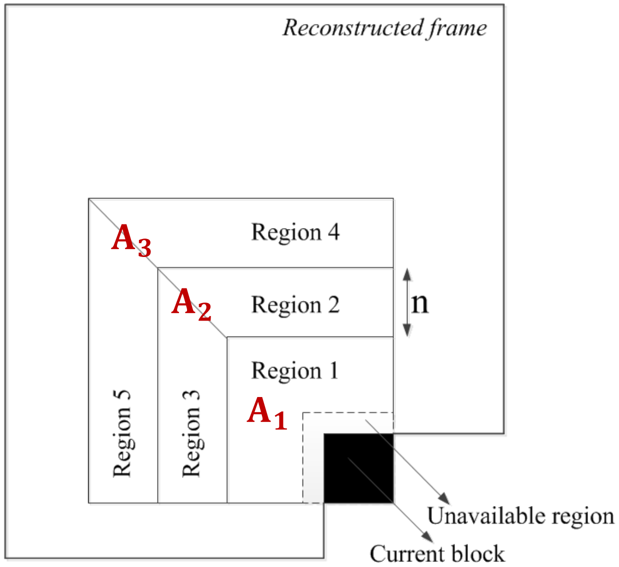


Fig.10 : ‘A’ for each region

For our experiment,

- We consider, for Region1, $A = A_1$

  for Region2 and Region3, $A = A_2$

  for Region 4 and Region5, $A = A_3$

- We assume,

  $A_1 = \{4, 6, 8\}$

  $A_2 = \{4, 6, 8\}$

  $A_3 = \{4, 6, 8\}$

Investigation on the value of 'A'

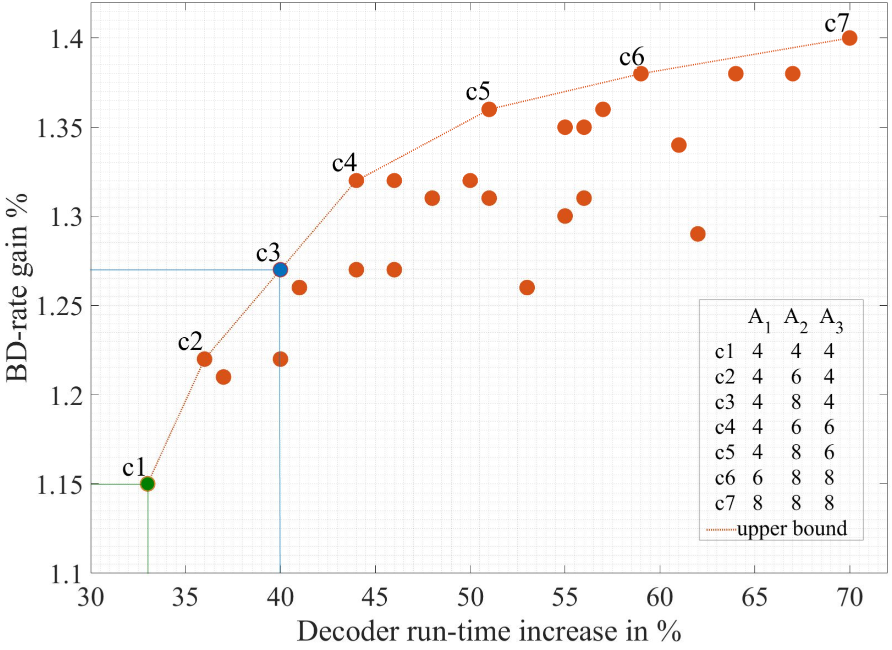


Fig.11 : Average BD-rate gain Vs Decoder complexity for different combinations of A_1, A_2 and A_3

- **c1** has the least decoder run-time increase. So, it is selected as the best choice of our proposed fast TM mode
- **c3** is the point in the upper bound of fast TM which has the closest gain as that of conventional TM for M=64. c3 is used for comparison purpose.

Conventional TM Vs Fast TM intra mode

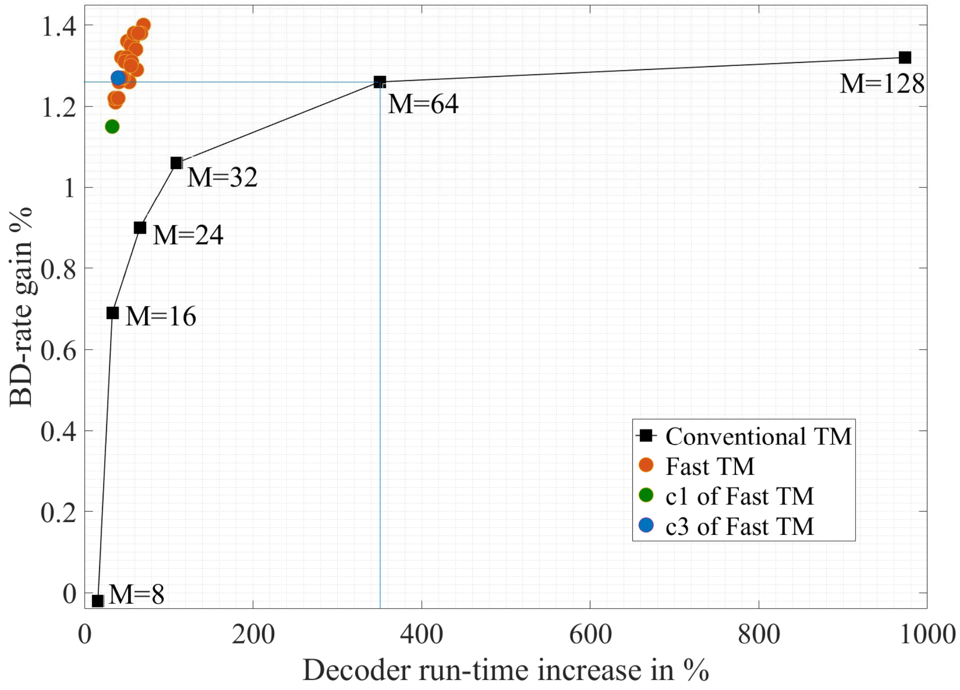


Fig.12 : Conventional TM Vs Fast TM intra mode

- Fast TM mode points are concentrated in the top-left of the plot in fig.14. i.e in the area with more gain and less complexity.
- Fast TM intra mode has much lesser decoder run-time increase compared to conventional TM
- For M=64, conventional TM achieves 1.26% coding gain with 350% increase in decoder run-time while fast TM (• c3) achieves the same gain with just 40% increase in decoder run-time (nearly 9 times less).
- • c1 is the best choice of fast TM mode, as it has the least decoder complexity.

Gayathri Venugopal

Fraunhofer HHI

Experimental results

	Test Sequences	BD-rate gain
Class A 2560x1600	Traffic	-0.49%
	PeopleOnStreet	-0.84%
	NebutaFestival	-0.10%
	SteamLocomotiveTrain	-0.16%
Class B 1920x1080	Kimono	-0.63%
	ParkScene	-0.45%
	Cactus	-2.36%
	BasketballDrive	-2.15%
	BQTerrace	-2.43%
Class C 832x480	BasketballDrill	-3.23%
	BQMall	-0.72%
	PartyScene	-0.54%
	RaceHorses	-0.12%
Class D 416x240	BasketballPass	-0.30%
	BQSquare	-0.80%
	BlowingBubbles	-0.18%
	RaceHorses	0.01%
Class E 1280x720	FourPeople	-1.28%
	Johnny	-4.60%
	KristenAndSara	-1.64%

Average BD-rate gain	**-1.15%**
Decoder Complexity	**133%**
Encoder Complexity	**180%**

- All implementations on HM16.6
- JCT-VC Common Test Conditions for All intra main configuration
- $A_1 = A_2 = A_3 = 4$ (i.e combination c1)

Gayathri Venugopal

Fraunhofer HHI

Conclusions

- We propose a fast TM intra mode for HEVC.
- Fast TM intra mode can achieve an average BD-rate gain of 1.15% with 33% decoder run-time increase (for A = 4).
- The new method can be tuned for various combinations of BD-rate gain and complexity, and thus the one suitable for the application can be chosen.
- For the same coding gain of 1.26%, conventional TM has 350% decoder run-time increase, whereas our proposed method has just 40% increase in decoder run-time (nearly 9 times less).
- Sequences with homogeneous textures are found to be favorable for TM prediction.

Thank you !

Contact:

Gayathri Venugopal
gayathri.venugopal@hhi.fraunhofer.de

Einsteinufer 37
10587 Berlin, Germany

Experimental results ctd..

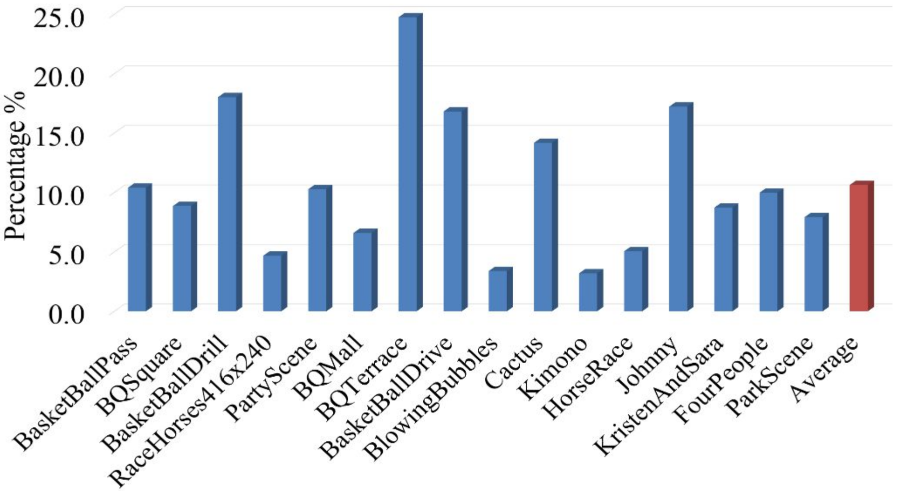


Fig.13 : Average percentage of area in a frame predicted by TM mode for different sequences

Experimental results ctd..

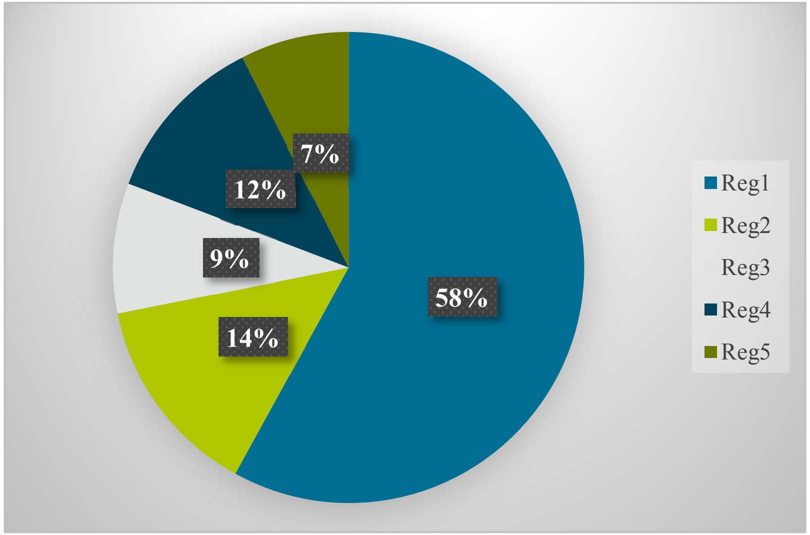


Fig.14 : Average percentage of area in a frame contributed by each region

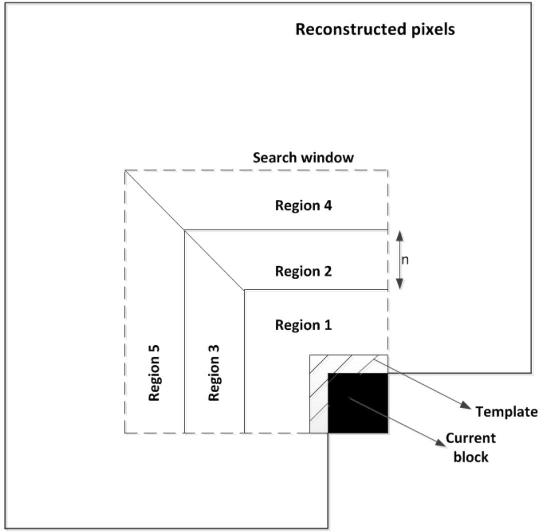


Fig.15 : Regions of Fast TM

References

[1] Vivienne Sze, Madhukar Budagavi, and Gary J. Sullivan, High Efficiency Video Coding (HEVC) Algorithms and Architectures, Springer, 2014.

[2] Gary J. Sullivan, Jens-Rainer Ohm, Woo-Jin Han, and Thomas Wiegand, “Overview of the high efficiency video coding HEVC standard,” in IEEE Transactions on Circuits and Systems for Video Technology, September 2012, vol. 22, pp. 1649 – 1668.

[3] Thiow Keng Tan, Choong Seng Boon, and Yoshinori Suzuki, “Intra prediction by template matching,” in IEEE International Conference on Image Processing (ICIP), Altlanta, GA, 2006, pp. 1693 – 1696.

[4] Thiow Keng Tan, Choong Seng Boon, and Yoshinori Suzuki, “Intra prediction by averaged template matching predictors,” in Proc. CCNC 2007, Las Vegas, NV, USA, 2007, pp. 405–109.

[5] Yi Guo, Ye-Kui Wang, and Houqiang Li, “Prioritybased template matching intra prediction,” in IEEE International Conference on Multimedia and Expo 2008, Hannover, Germany, June 2008, pp. 1117–1120.

[6] Yinfei Zheng, Peng Yin, Oscar Divorra Escoda, Xin Li, and Cristina Gomila, “Intra prediction using template matching with adaptive illumination compensation,” in Proc. IEEE International Conference on Image Processing ICIP2008, San Diego, CA, October 2008.

[7] Matthieu Moinard, Isabelle Amonou, Pierre Duhamel, and Patrice Brault, “A set of template matching predictors for intra video coding,” in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2010, March 2010.

[8] Safa Cherigui, Christine Guillemot, Dominique Thoreau, Philippe Guillotel, and Patrick Perez, “Hybrid template and block matching algorithm for image intra prediction,” in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, March 2012, pp. 781–784.

[9] Karam Naser, Vincent Ricordel, and Patrick Le Callet, “Local texture synthesis: A static texture coding algorithm fully compatible with HEVC,” in International Conference on Systems, Signals and Image Processing,IWSSIP, September 2015

Fraunhofer
HHI