

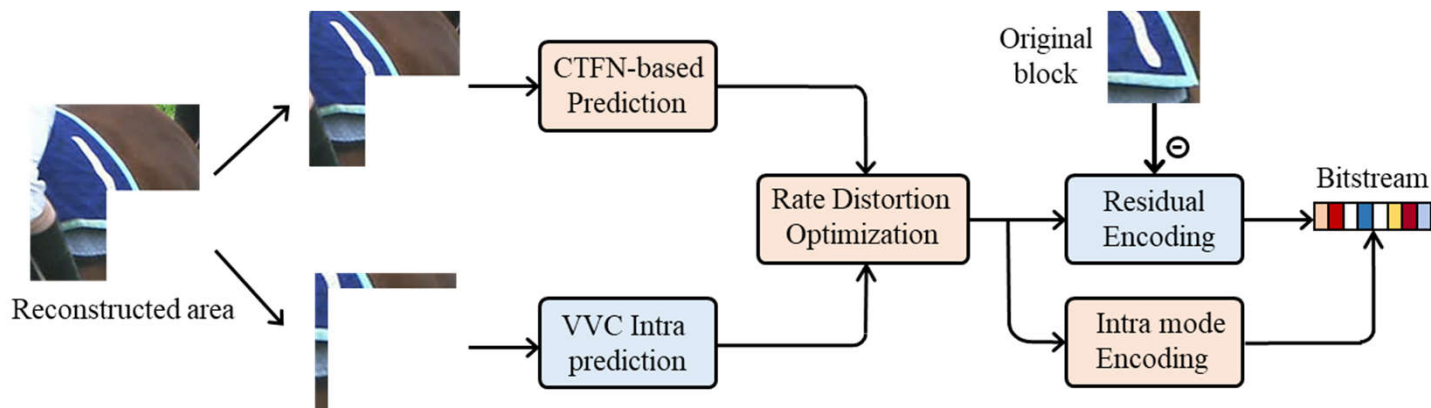
Coarse-to-fine Prediction With Local and Nonlocal Correlations for Intra Coding

Meng Lei, Xuewei Meng, Chuanmin Jia, Shanshe Wang,
Zhipeng Cheng and Siwei Ma

Virtual Conference
DCC 2022

Integration of the CTFN in VVC codec

- Serve as a PU-level intra prediction mode
 - Compete with all intra prediction modes in VVC.
- A larger reconstructed region as input(local→nonlocal)
 - Compared to four reference lines in VVC.



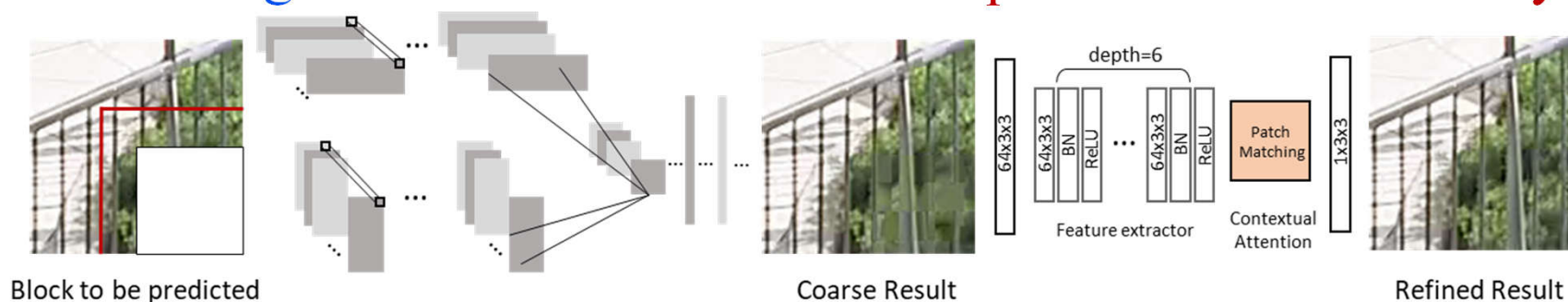
Architecture of the two-stage CTFN

□ The coarse prediction stage

- composed of fully-connected networks and convolutional networks.

□ The fine prediction stage

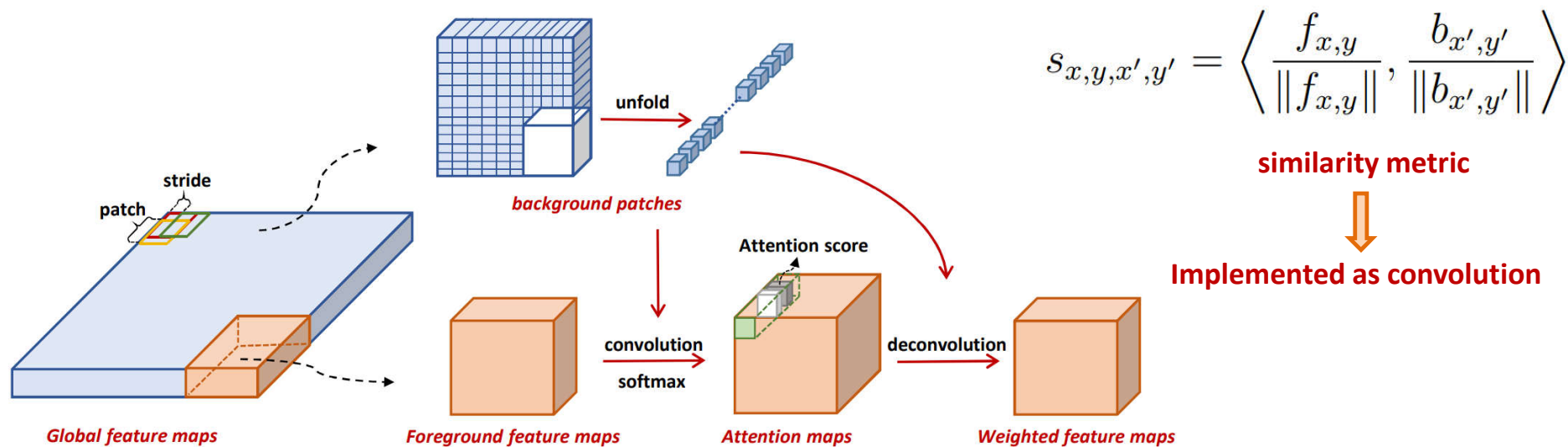
- consists of a global feature extractor and a patch-wise attention layer.



Layer number	8×8/8×16/16×8		16×16			
	Layer type	Node	Layer type	Kernel	Filter/Node	Stride
1	fully-connected	1216	convolution	3×3	32	(2, 2)
2	fully-connected	1216	convolution	3×3	64	(2, 2)
3	fully-connected	64/128	convolution	3×3	128	(1, 2)/(2, 1)
4	-	-	convolution	3×3	128	(1, 2)/(2, 1)
5	-	-	fully-connected	-	512	-
6	-	-	fully-connected	-	256	-

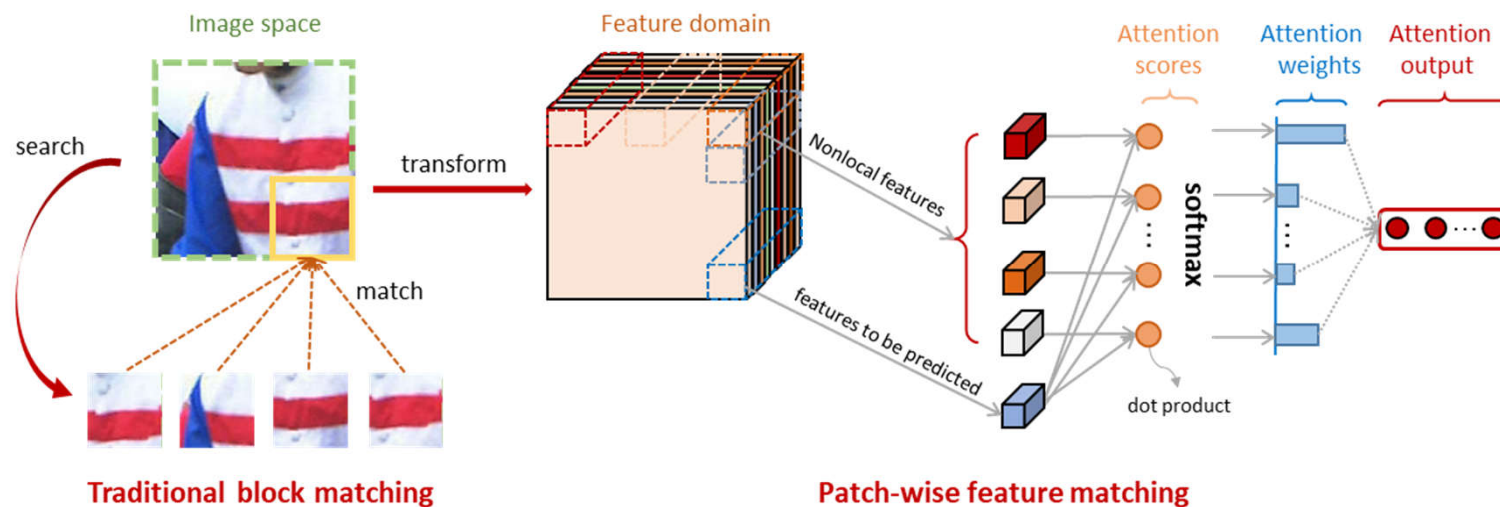
Patch-wise attention layer

- Attention based matching process(similarities → attention scores)
 - The coarse prediction region plus the L-shape reconstructed area with a width of 2 are determined as template.
 - Matching and weighting are implemented as convolutions.



Patch-wise attention layer

- Extracting effective nonlocal information in feature domain
 - Spatial-domain block matching → Patch-wise feature matching.
 - Assigns weights to each nonlocal feature with attention mechanism.



Experimental results

Objective performance

Anchor: VTM11.0-NNVC

Test: Proposed CTFN, JVET-W0081

CTC: **-1.92%**, SCC: **-1.48%**

GPU type	GPU: RTX-3080Ti-12GB	Framework	Pytorch v1.8.1
Epoch	50	Batch size	64
Parameter Precision	32-bit float	Optimizer	ADAM

Class	Sequence	JVET-W0081 [12]		Proposed-v1		Proposed-v2	
		BD-BR	UR	BD-BR	UR	BD-BR	UR
F	BasketballDrillText	-0.73%	18.2%	-1.11%	20.3%	-1.55%	22.4%
	ArenaOfValor	-1.17%	21.8%	-1.23%	21.9%	-1.46%	22.5%
	SlideEditing	-0.10%	4.10%	-1.21%	8.80%	-1.20%	9.00%
	SlideShow	-0.27%	8.10%	-0.24%	9.70%	-0.61%	10.7%
TGM	FlyingGraphics	0.21%	2.60%	-3.31%	14.3%	-3.60%	17.6%
	Desktop	0.06%	1.00%	-1.56%	9.00%	-1.78%	10.9%
	Console	0.12%	1.00%	-0.87%	8.00%	-1.05%	9.50%
	ChineseEditing	0.13%	1.40%	-0.25%	5.00%	-0.41%	5.90%
	Map	-0.44%	6.40%	-0.97%	10.0%	-1.39%	12.9%
	Browsing	-0.23%	3.50%	-1.44%	8.70%	-1.97%	10.8%
	Robot	-1.03%	17.3%	-1.82%	19.0%	-1.70%	19.3%
	Programming	-0.24%	2.40%	-0.73%	6.90%	-1.09%	7.70%
Average		-0.31%	7.32%	-1.23%	11.8%	-1.48%	13.3%

Class	Sequence	JVET-W0081 [12]		Proposed-v1		Proposed-v2	
		BD-BR	UR	BD-BR	UR	BD-BR	UR
A1	Tango2	-3.65%	60.6%	-3.72%	62.3%	-3.71%	61.7%
	FoodMarket4	-4.14%	61.4%	-4.17%	61.8%	-4.16%	61.2%
	Campfire	-1.23%	33.0%	-1.41%	34.3%	-1.43%	34.7%
A2	CatRobot1	-1.86%	40.6%	-2.47%	45.5%	-3.05%	48.6%
	DaylightRoad2	-1.45%	37.9%	-1.88%	43.5%	-2.14%	43.1%
	ParkRunning3	-1.08%	54.0%	-1.09%	54.7%	-1.16%	55.3%
B	MarketPlace	-1.52%	49.8%	-1.66%	51.6%	-1.70%	52.7%
	RitualDance	-2.20%	34.6%	-2.56%	38.3%	-2.63%	37.6%
	Cactus	-1.31%	29.3%	-1.82%	33.5%	-1.93%	33.8%
	BasketballDrive	-1.95%	29.5%	-2.23%	35.6%	-2.55%	36.7%
	BQTerrace	-0.75%	22.1%	-1.34%	26.2%	-1.67%	29.8%
C	BasketballDrill	-0.42%	24.2%	-0.77%	24.9%	-0.89%	25.4%
	BQMall	-0.83%	13.5%	-0.85%	14.2%	-0.93%	14.4%
	PartyScene	-0.43%	10.1%	-0.54%	11.1%	-0.50%	11.8%
	RaceHorseC	-1.24%	17.9%	-1.32%	21.3%	-1.14%	24.0%
D	BasketballPass	-1.04%	13.8%	-1.42%	16.5%	-1.92%	18.6%
	BQSquare	0.08%	6.20%	-0.43%	7.50%	-0.47%	7.90%
	BlowingBubbles	-0.69%	14.8%	-0.84%	16.7%	-0.62%	16.0%
	RaceHorses	-0.93%	17.7%	-0.90%	18.7%	-1.15%	19.5%
E	FourPeople	-1.85%	26.9%	-2.12%	28.9%	-2.21%	29.3%
	Johnny	-1.78%	27.5%	-3.23%	33.1%	-3.72%	35.1%
	KristenAndSara	-2.09%	26.4%	-2.66%	28.4%	-2.73%	27.9%
Average		-1.47%	29.6%	-1.79%	32.1%	-1.92%	33.0%

Thanks!



mail: lemony@pku.edu.cn

