Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN

*N. Bakir, **W. Hamidouche**, S. Fezza, K. Samrouth and O. Déforges*

CLIM 2 workshop – INRIA Rennes

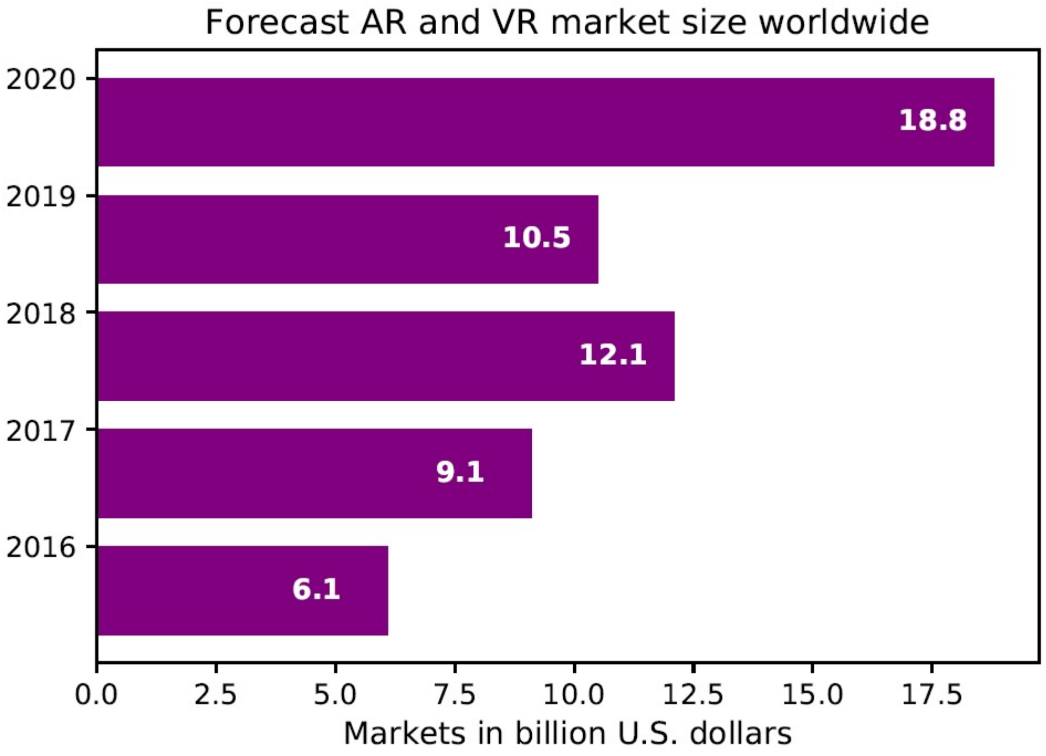
September 30th 2021

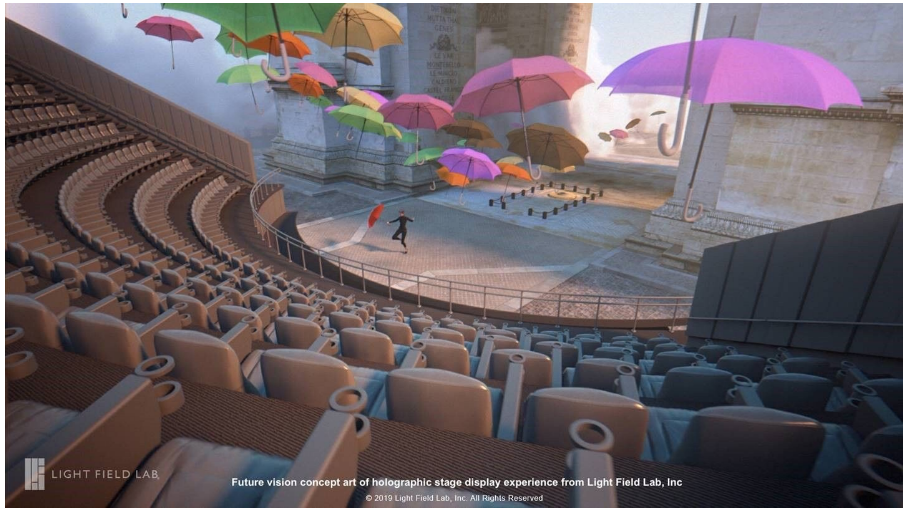
Outline

- **Context & challenges**
- Related work
- Proposed solution
- Experimental Results
- Conclusion

Context & Challenges

❑ Augmented Reality (AR) & Virtual Reality (VR) applications



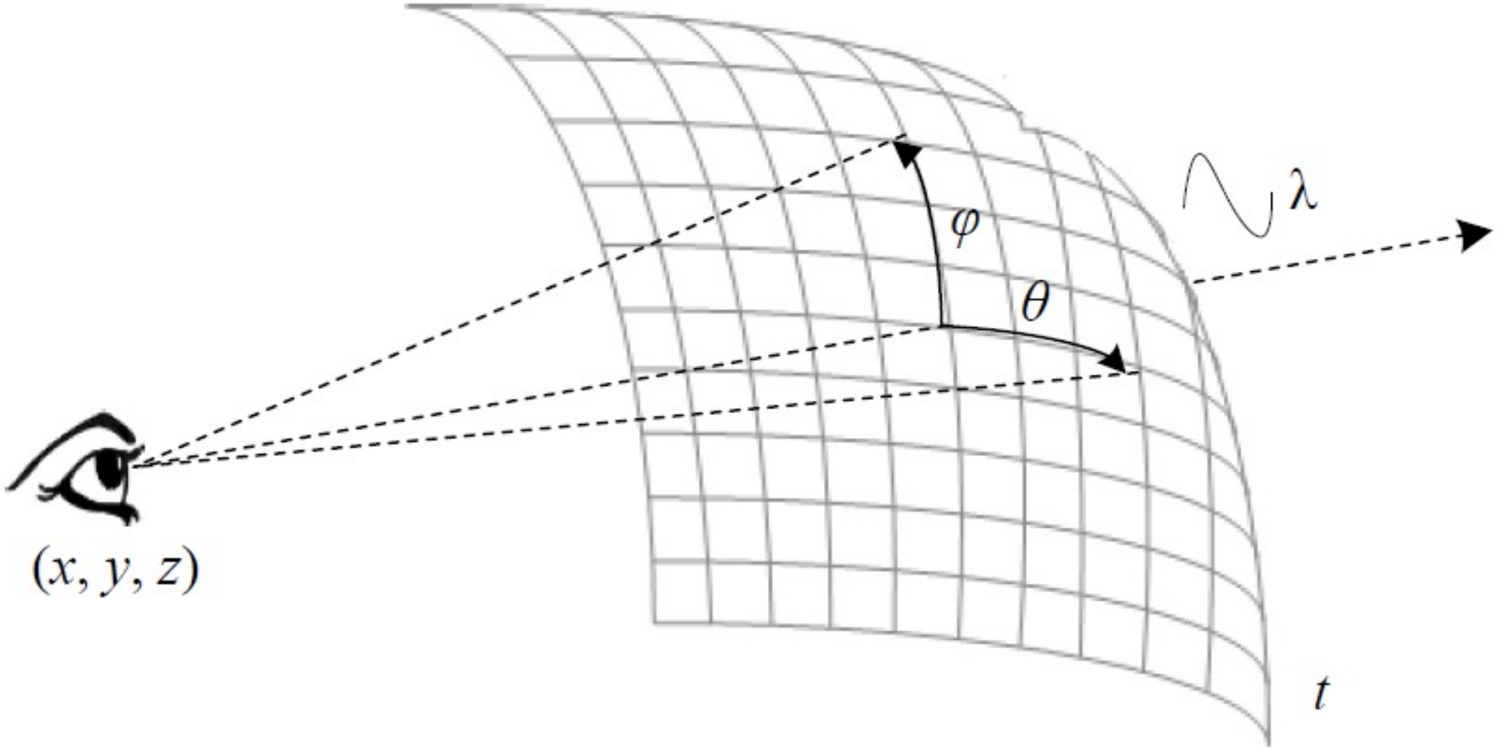


N. Bakir, **W. Hamidouche** et al., Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Context & Challenges

❑ Definition of Light Field

- Seven dimensional function $L(x, y, z, \theta, \varphi, \lambda, t)$

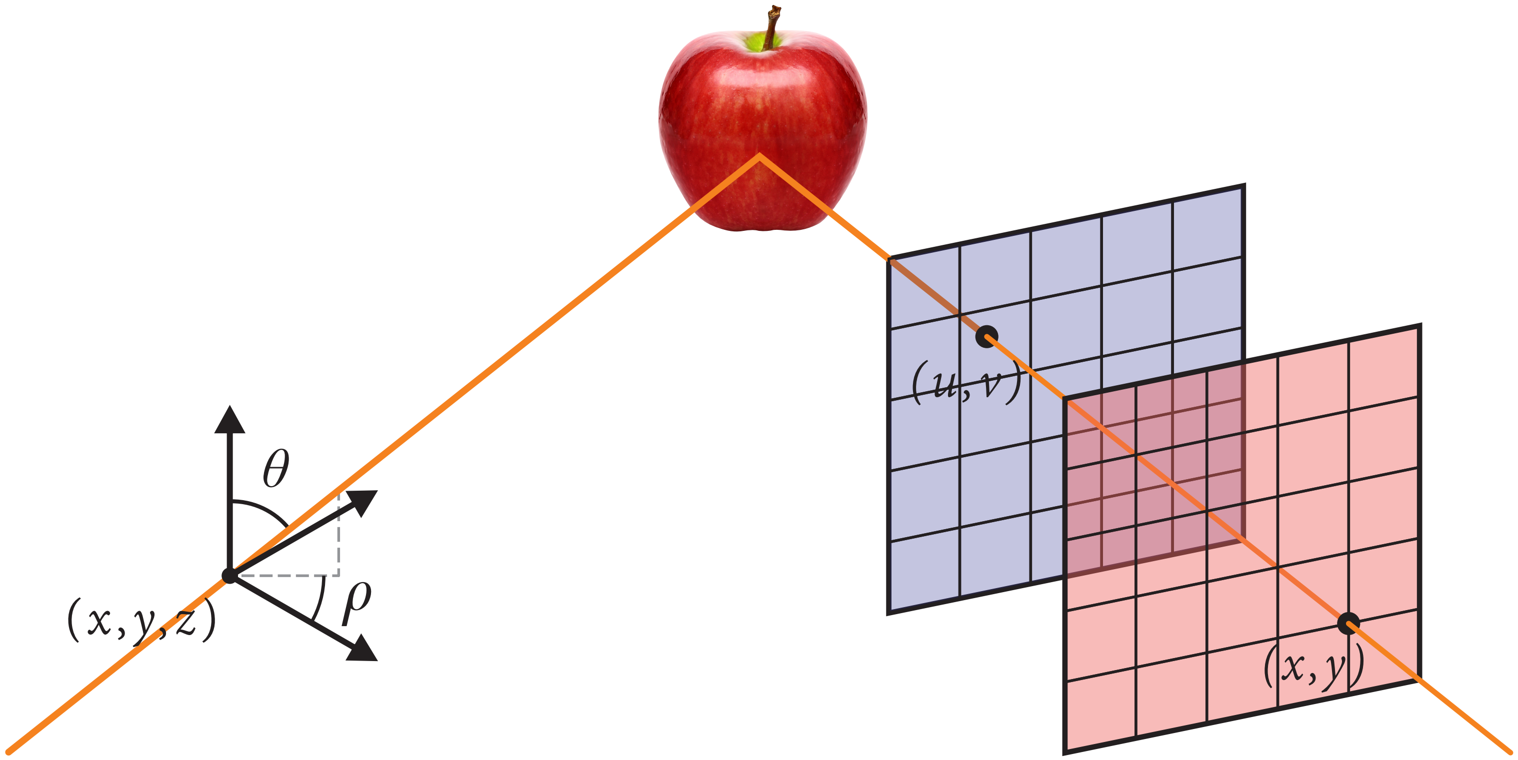


N. Bakir, **W. Hamidouche** et al., Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Context & Challenges

Definition of Light Field

- Seven dimensional function $L(x, y, z, \theta, \varphi, \lambda, t)$
- Four dimensional function with two parallel planes $L(u, v, t, s)$



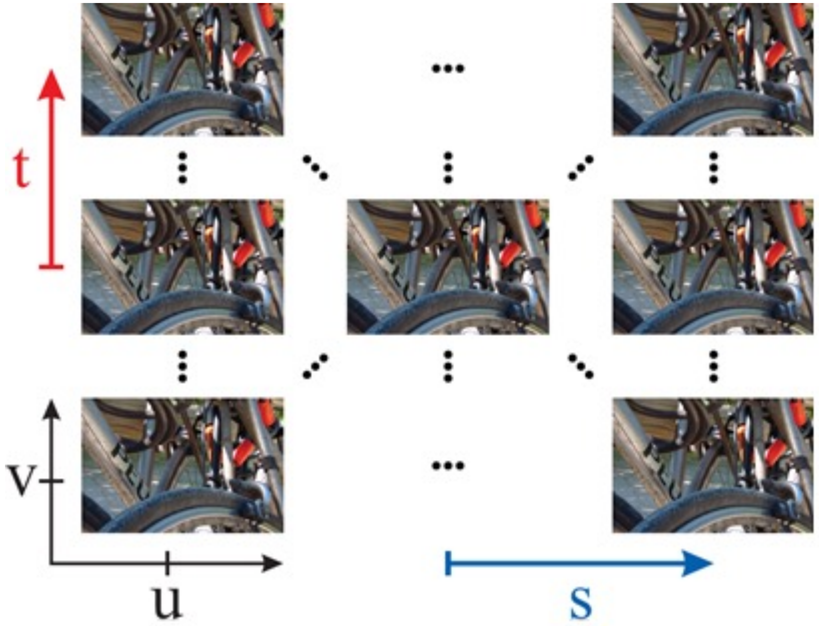
N. Bakir, **W. Hamidouche** et al., Light Field Image Coding using D2GAN and VVC Temporal Scalability, IEEE Tran. on Multimedia 2021.

Context & Challenges

❑ Light Field image acquisition and representation

Multi camera

Plenoptic camera

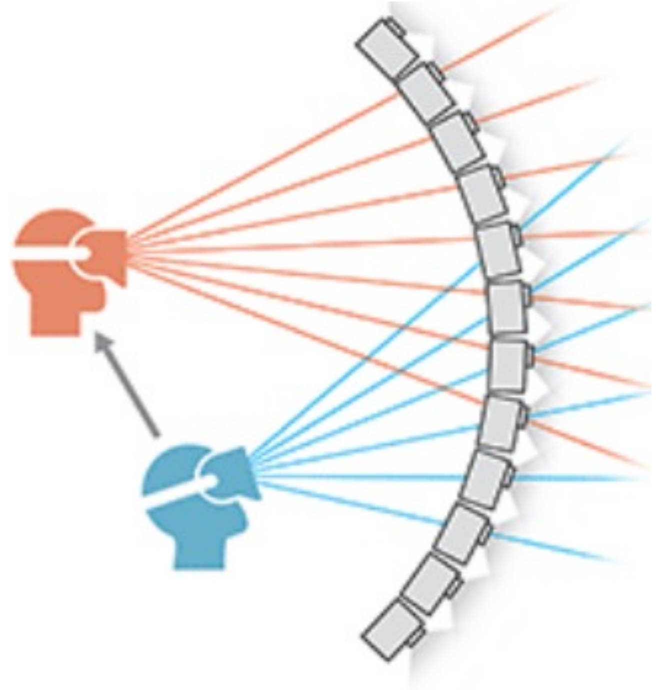


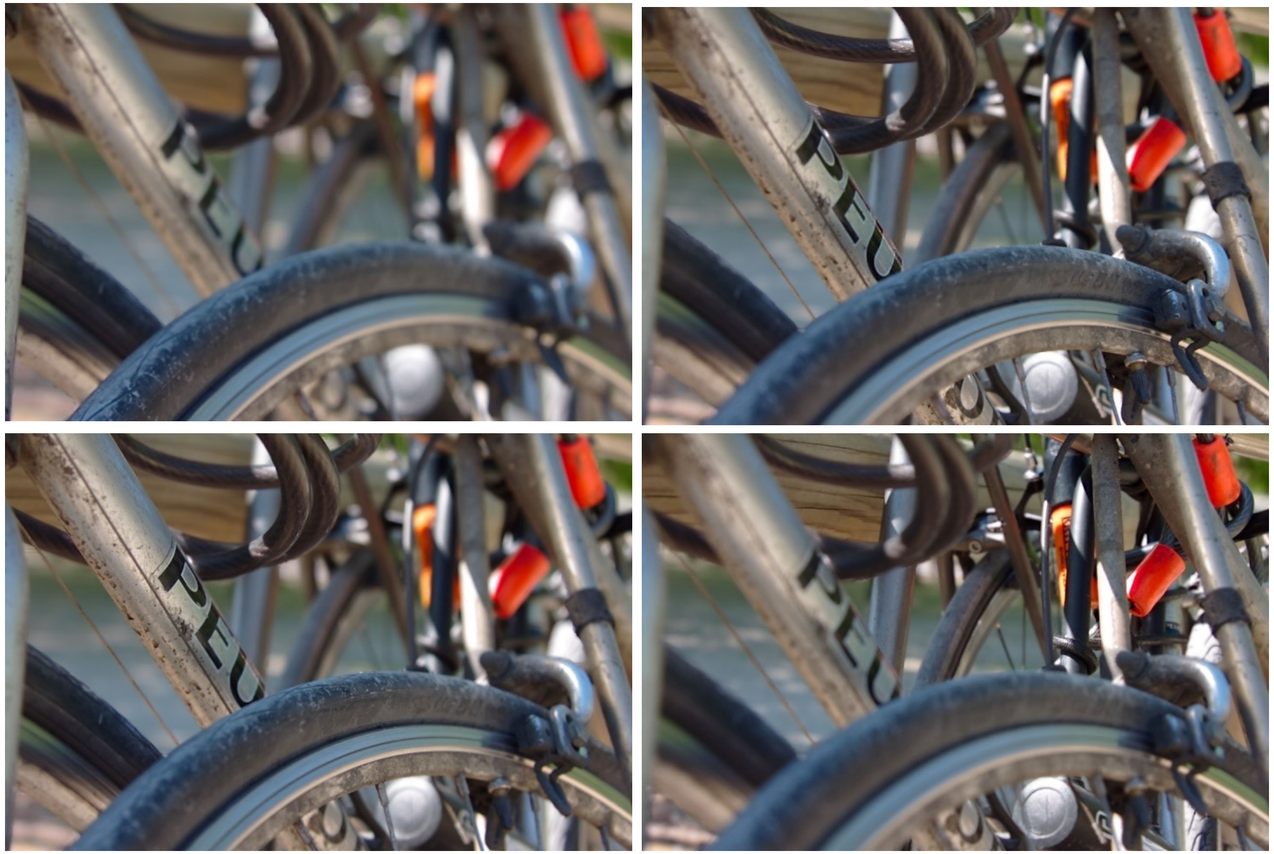
N. Bakir, **W. Hamidouche** et al., [Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN](), IEEE T. on Multimedia 2021.

Context & Challenges

❑ **Advantages of Light Field technology**

- ➢ Refocusing capability after capture
- ➢ Change of perspective (view point)
- ➢ Field display for VR





N. Bakir, ***W. Hamidouche*** *et al.*, Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Context & Challenges

VAADER
IETR research team

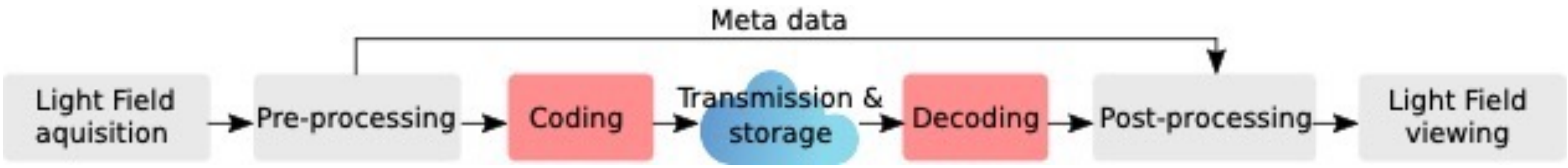
❑ Limitations of Light Field technology

- ➢ Very large amount of data for storage and transmission
- ➢ High computational processing and energy
- ➢ Highly redundant data

1 LF image Lytro Illum Camera requires 1.5 Gbits

❑ Objective

- ➢ Propose an efficient and low complexity coding solution for LF images



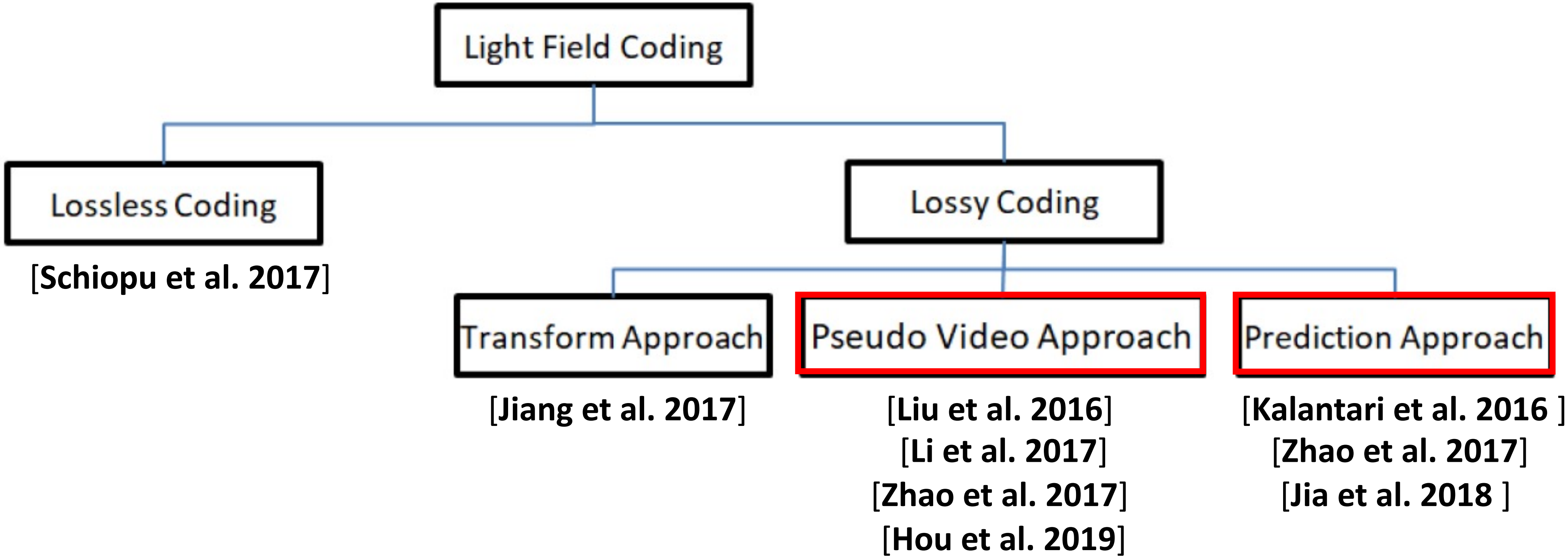
N. Bakir, **W. Hamidouche** *et al.*, Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Outline

- Context & challenges
- **Related work**
- Proposed solution
- Experimental Results
- Conclusion

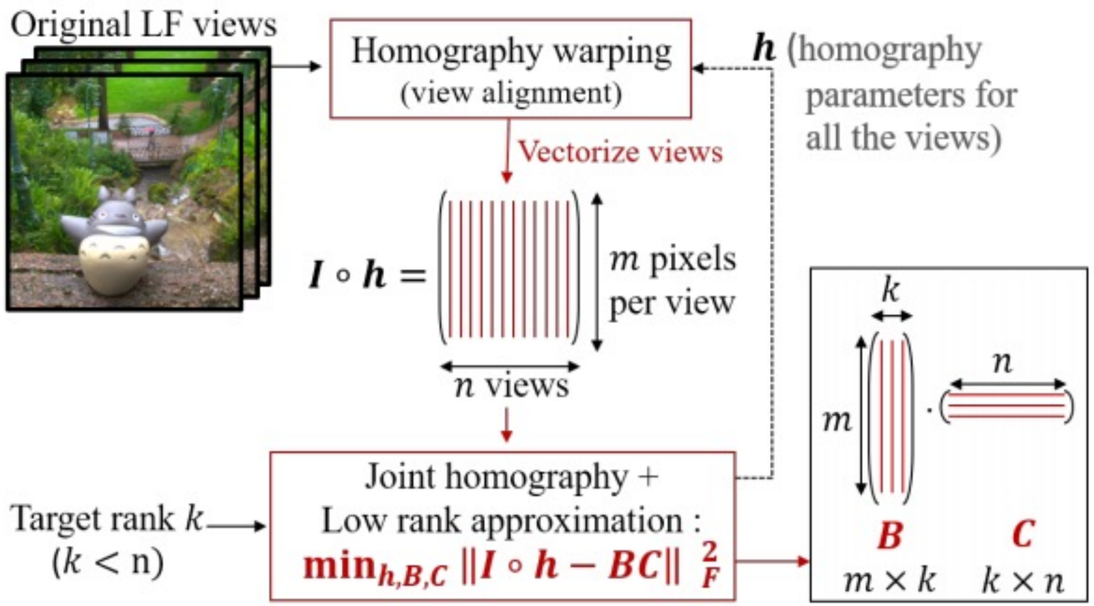
Related work

❑ State-of-the-art of Light Field Image Compression



N. Bakir, **W. Hamidouche** *et al.*, Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

- LF Compression with Homography-based Low Rank Approximation (HLRA) [Jiang et al. 2017]
  - Based joint optimization of multiple homographies and **LRA**[1]
  - Using homography in order to reduce the error produced by **LRA** between **SAI**[2]



Overview diagram of the HLRA method.

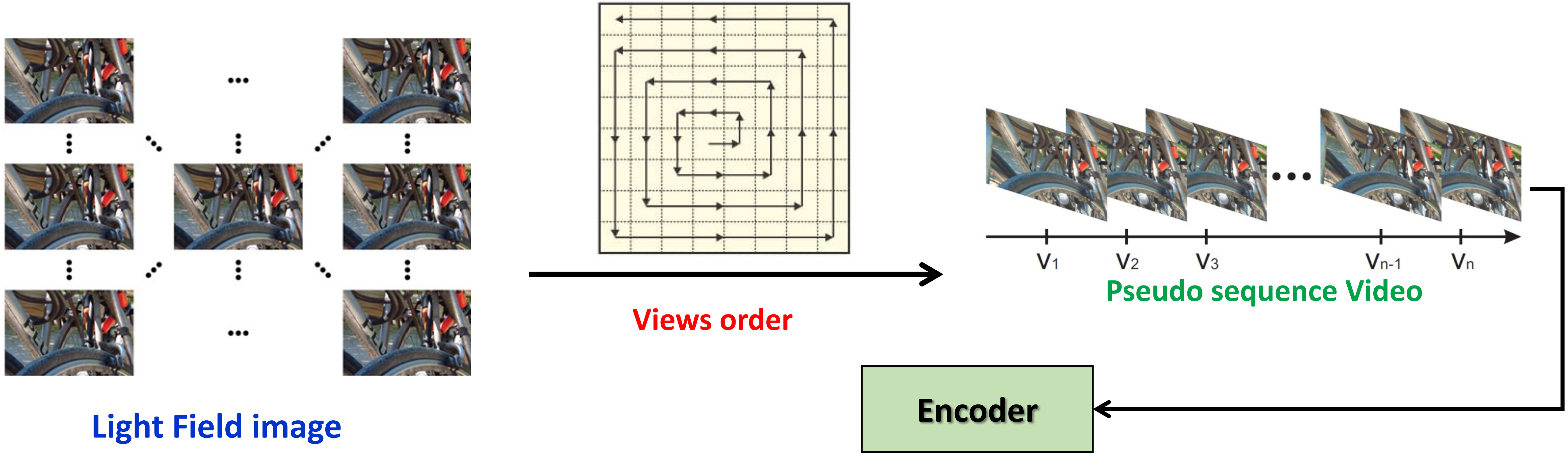
[1]: Low Rank Approximation
[2]: Sub-Aperture Images

N. Bakir, **W. Hamidouche** et al., Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Related work

❑ State-of-the-art of Light Field Image Compression [Lui et al. 2016]

- ➢ Pseudo video coding approach



N. Bakir, ***W. Hamidouche*** *et al.*, [Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN](), IEEE T. on Multimedia 2021.

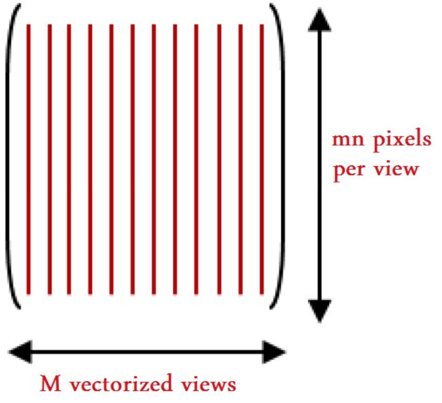
❑ LF image coding via linear approximation [Zhao et al. 2017]

$\boldsymbol{S_R} \ \epsilon\ \mathbb{R}^{mn\ x\ M}$, consists of M vectorized references views in $\boldsymbol{S_A}$

$\boldsymbol{s_b}$ is the target view to be recovered

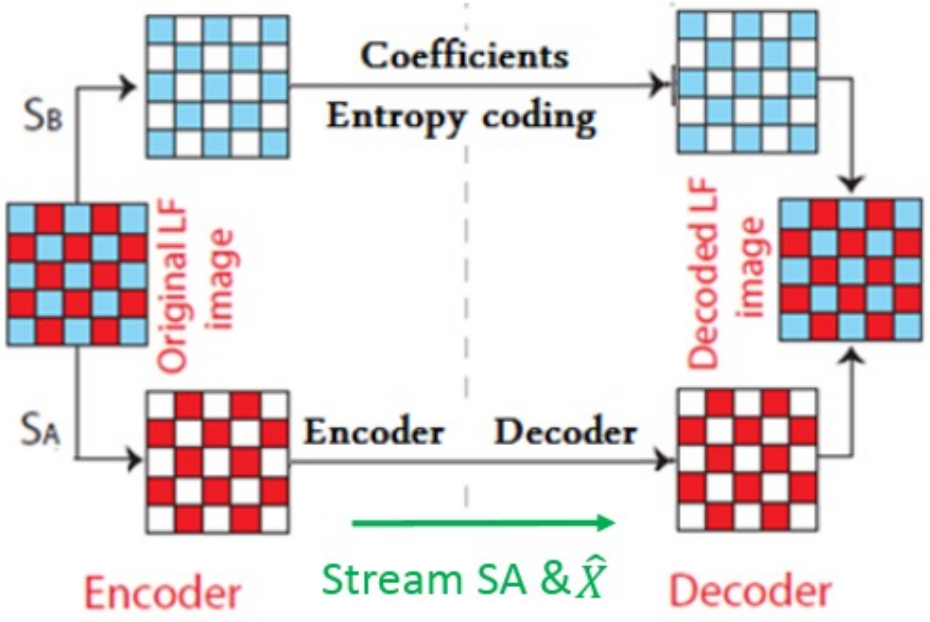
$$X_b = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \dots \\ \dots \\ a_M \end{bmatrix} \in \ \mathbb{R}^M$$

$\boldsymbol{S_R}$ =



$$\min \|X_b\|_1$$

$$\textbf{subject to} \quad \|S_R.X_b - s_b\|_2 \leq \ \varepsilon$$



Linear Approximation (**LA**) Coding Scheme [Zhao et al. 2017]

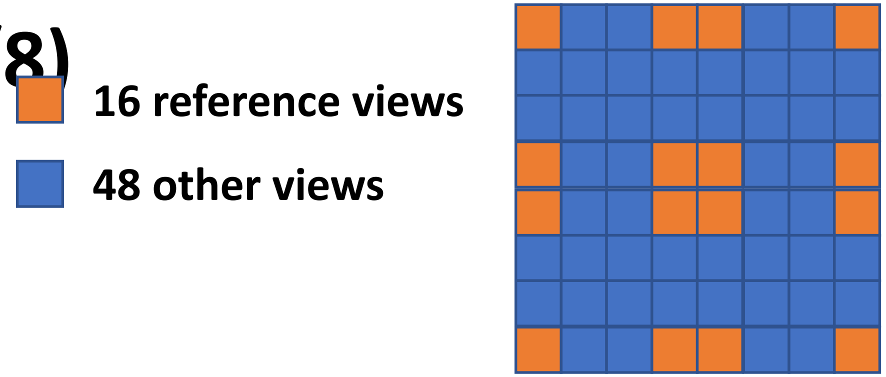
N. Bakir, ***W. Hamidouche*** *et al.*, Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

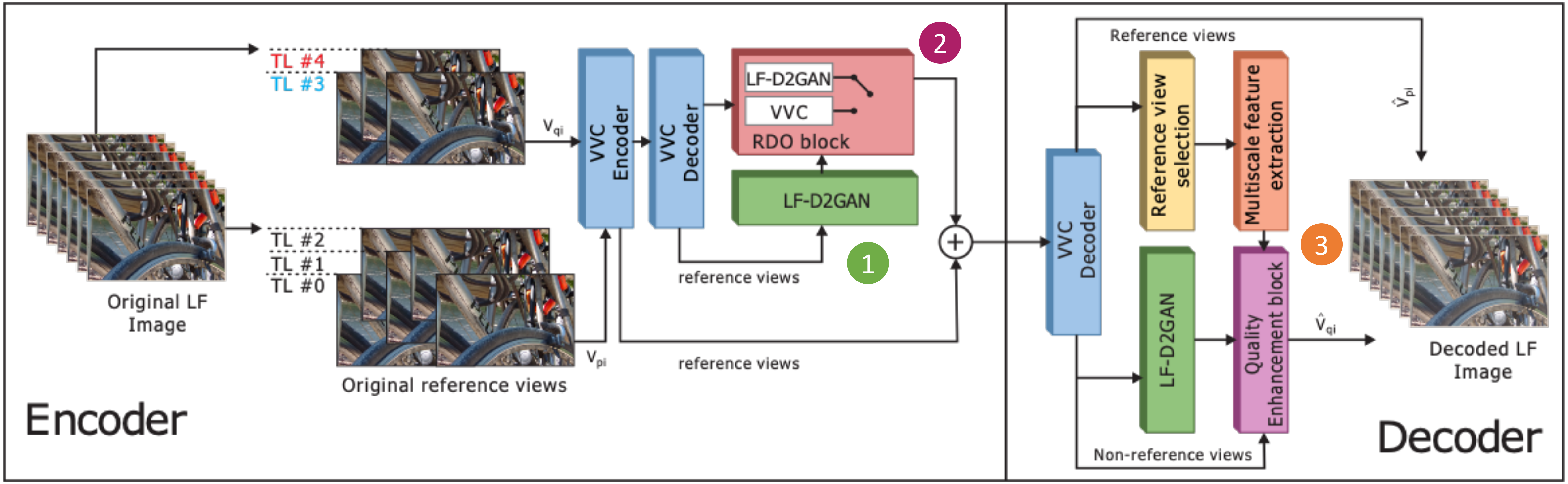
Outline

- Context & challenges
- **Proposed solution**
- Experimental Results
- Conclusion

Proposed Solution (1/8)

❑ Overall scheme of the proposed solution

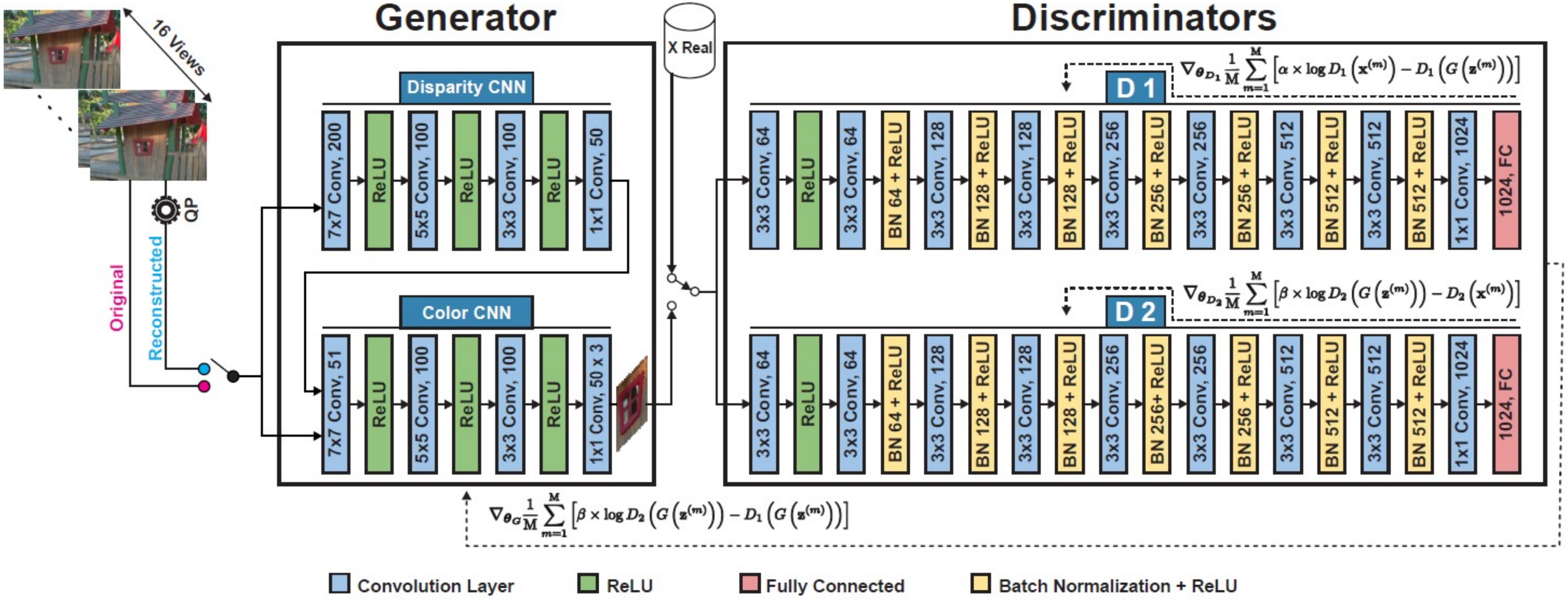




N. Bakir, **W. Hamidouche** et al., Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Proposed Solution (2/8)

❑ Architecture of the view synthesis block D2GAN



Nguyen et al., **Dual Discriminator Generative Adversarial Nets**, 31st Conference on Neural Information Processing System (NIPS 2017), USA.

N. Bakir, ***W. Hamidouche*** *et al.*, Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Proposed Solution (3/8)

❑ D2GAN Loss function formulation

$$\min_{G} \max_{D1,D2} \mathcal{L}\,(G, D1, D2)$$

❑ Three player minimax optimization game:

- Discriminators maximizing its reward $V(G, D1, D2)$
- Generator minimizing Discriminators reward (or maximize its loss)

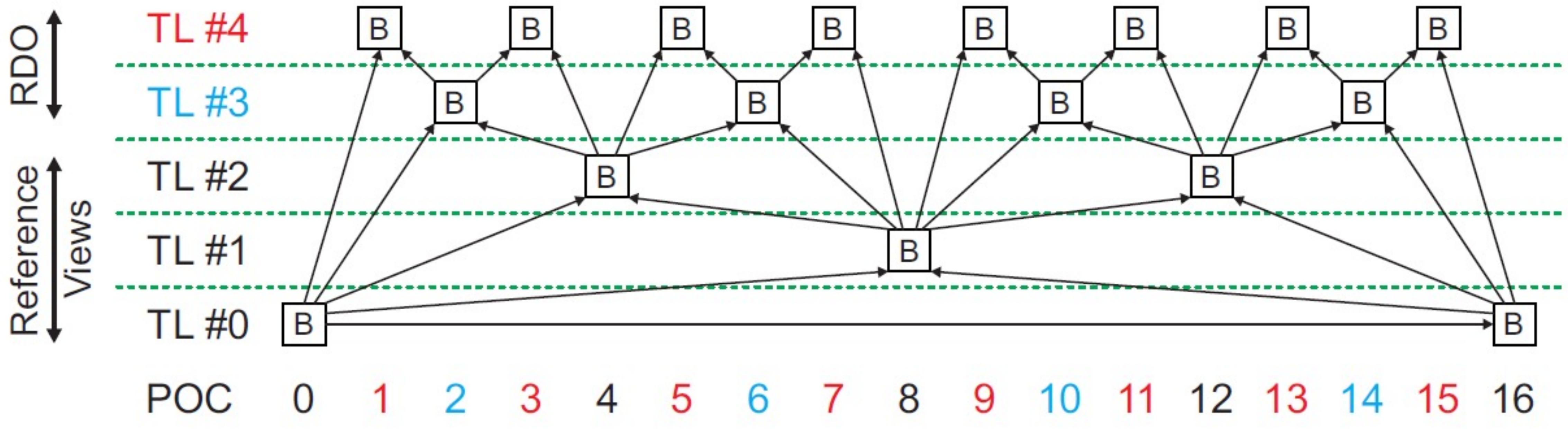
$$\mathcal{L}(G, D_1, D_2) = \alpha \mathbb{E}_{x \sim P_{data}} [\log D_1(\mathrm{x})] + \mathbb{E}_{z \sim P_z} [- D_1(\mathrm{G(z)})]$$

$$+ \mathbb{E}_{x \sim P_{data}} [- D_2(\mathrm{x})] + \beta \mathbb{E}_{z \sim P_z} [\log D_2(\mathrm{G(z)})]$$

N. Bakir, ***W. Hamidouche*** *et al.*, Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Proposed Solution (4/8)

❑ Versatile Video Coding (VVC) hierarchical coding architecture

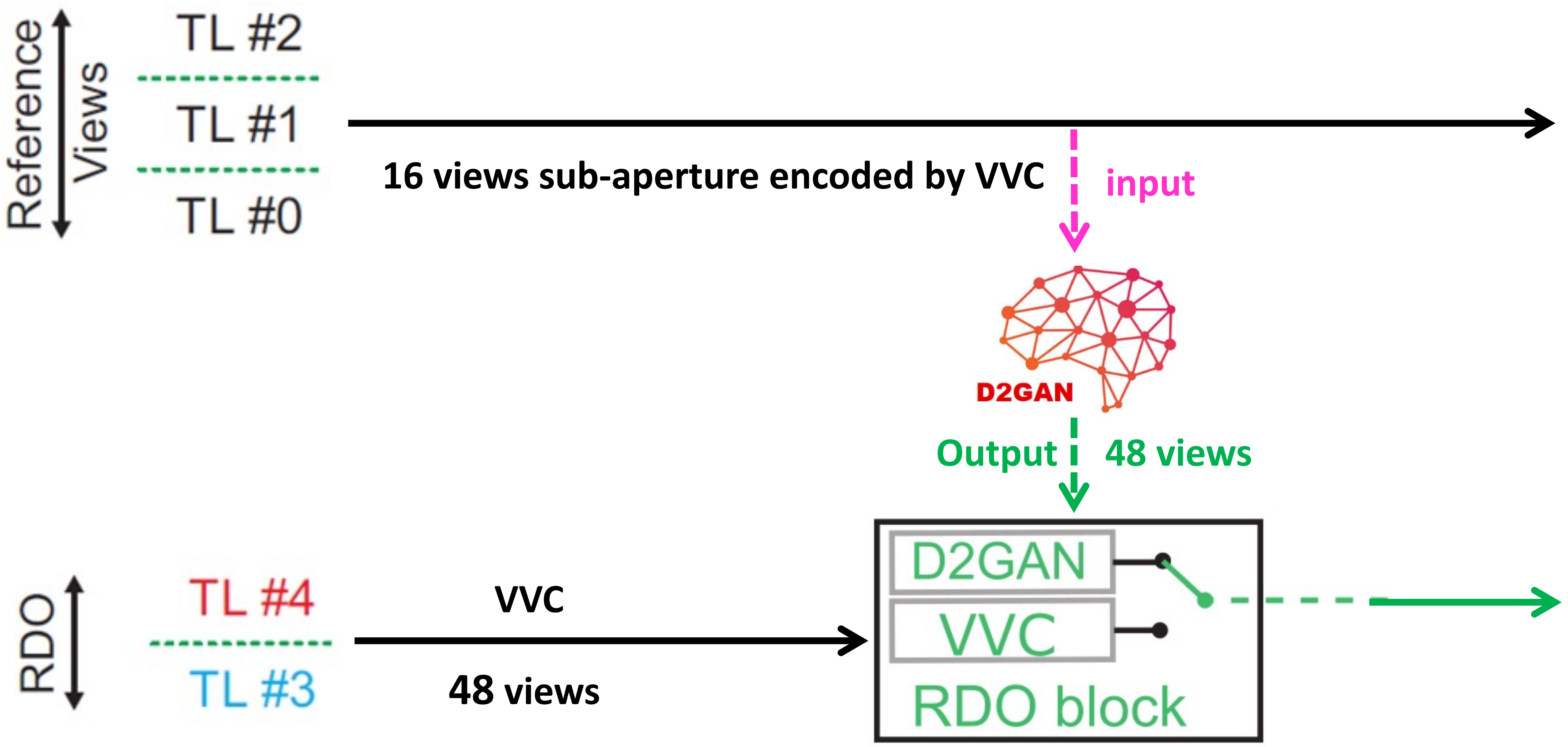


Hierarchical prediction structure in VVC - One GOP is shown.

N. Bakir, **W. Hamidouche** et al., [Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN](), IEEE T. on Multimedia 2021.

Proposed Solution (5/8)

❑ The proposed strategy



*N. Bakir, **W. Hamidouche** et al.*, Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Proposed Solution (6/8)

❑ Rate distortion optimization

Algorithm 2: Algorithm of the RDO between VVC and D2GAN

Require: $\mathcal{J} \leftarrow \{\forall\, m,\ \forall\, v \in TL\#[3\ or\ 4],\ \mathcal{J} = D + \lambda R\}$

  m: metod {VVC, D2GAN}

  for all $v \in TL\#4$ **do**

    if $\mathcal{J}(VVC) < \mathcal{J}(D2GAN)$ **then**

      Encode v by VVC

      flag(v) $\leftarrow$ false

    else

      generate v by D2GAN

      flag(v) $\leftarrow$ true

    end if

  end for

  for all $v \in TL\#3$ **do**

    if $\mathcal{J}(VVC) < \mathcal{J}(D2GAN)$ **then**

      Encode v by VVC

      flag(v) $\leftarrow$ false

    else {flag(previous(v)) and flag(next(v))}

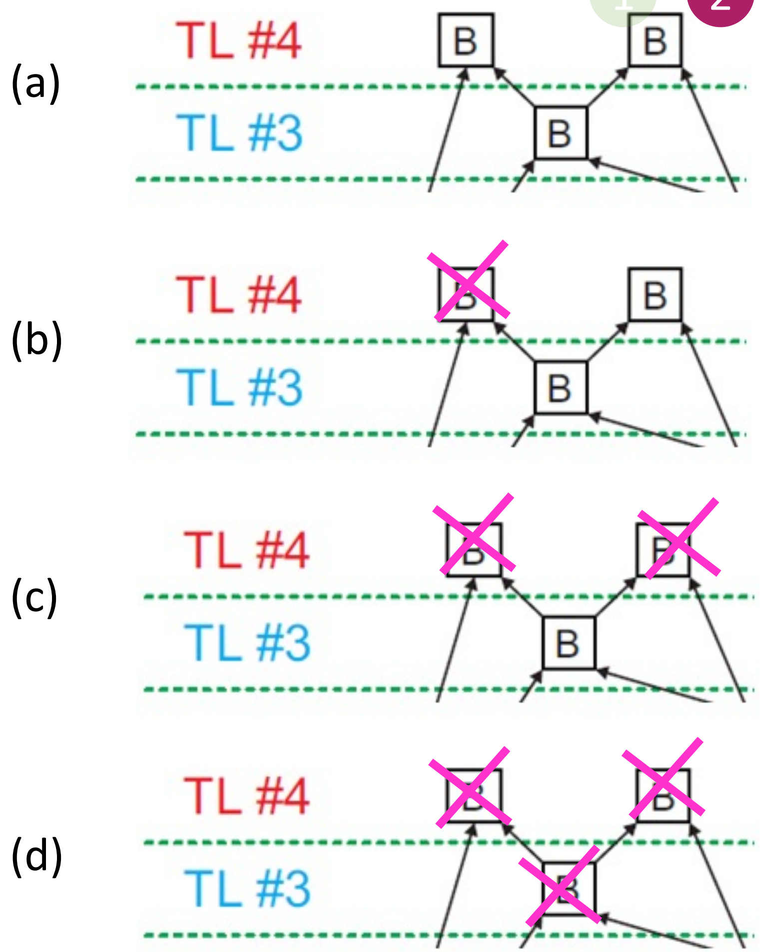
      generate v by D2GAN

      flag(v) $\leftarrow$ true

    end if

  end for

$$J = D + \lambda\, R$$

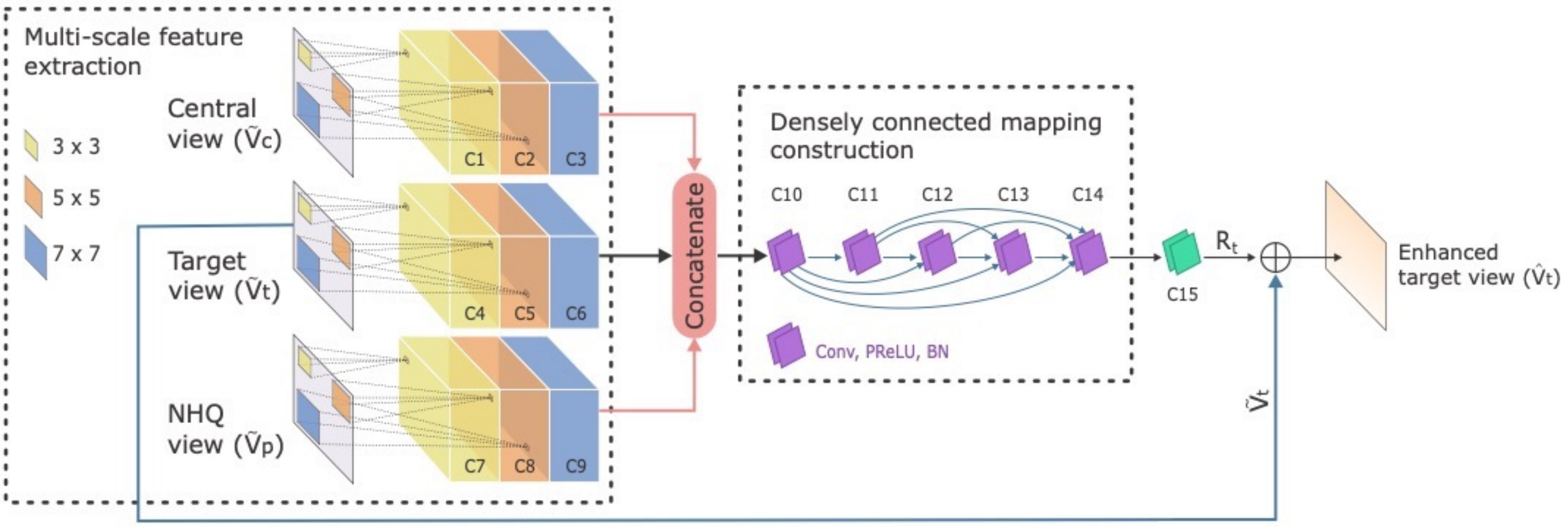


RDO cases

N. Bakir, **W. Hamidouche** *et al.*, Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Proposed Solution (7/8)

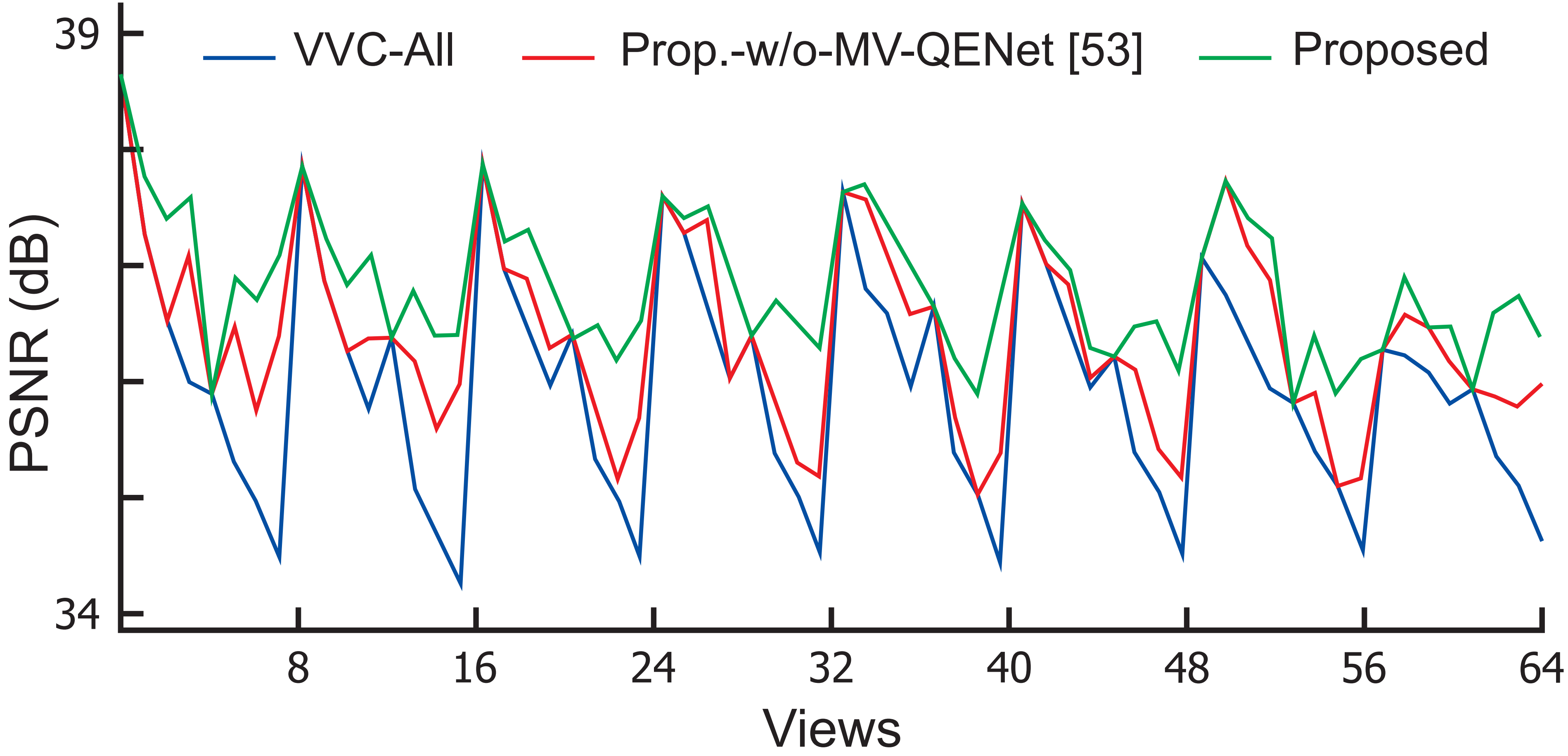
❑ Quality Enhancement Module: MV-QENet



N. Bakir, **W. Hamidouche** et al., Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Proposed Solution (8/8)

❑ MV-QENet performance



N. Bakir, ***W. Hamidouche*** *et al.*, Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Outline

- Context & challenges
- Proposed solution
- **Experimental Results**
- Conclusion

Experimental Results

❑ For the training phase, 140 LF images (8x8 views) have been selected:

- EPFL dataset (70 images), Stanford university LF dataset (50 images), HCI dataset(20 images).
- Each sub aperture are splitting into patches (60x60 pix).
- Trained with more than 150,000 patches.

❑ 2D-AN Training Parameters

Mini-batches size	10
ADAM solver	$\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\alpha_{Adam} = 0.0001$
Learning rate	0.0001
Activation function	Relu
Regularization coefficients of D1 and D2	$\alpha_{reg} = 0.2$ and $\beta_{reg} = 0.2$

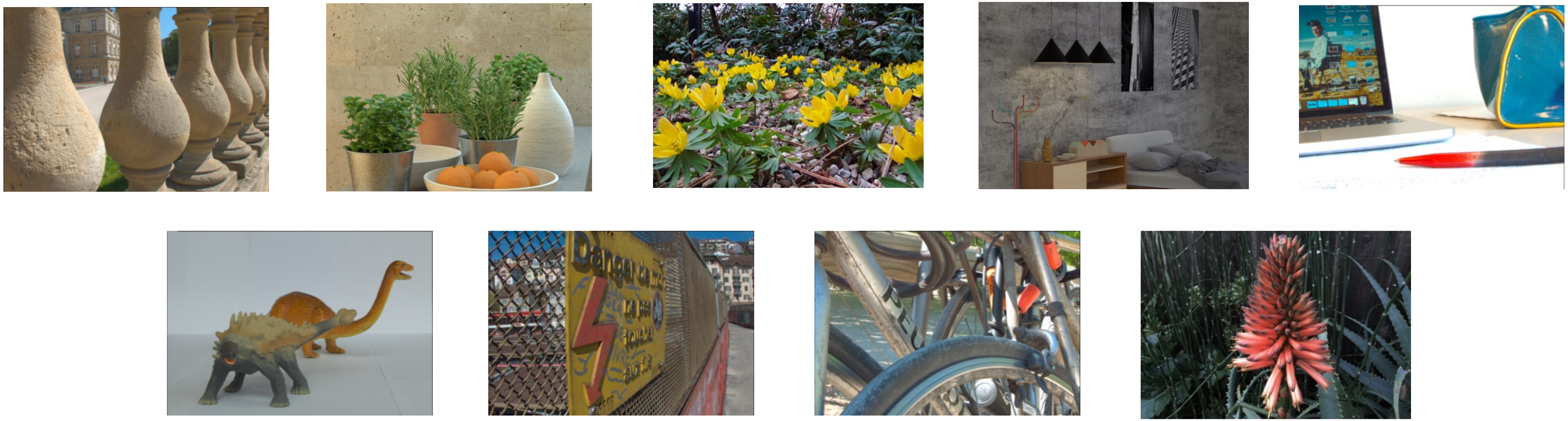
N. Bakir, ***W. Hamidouche*** *et al.*, Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Experimental Results

❑ 9 testing LF images

- ➢ EPFL dataset (6 images), Stanford university LF dataset (1 images), HCI dataset(2 images).

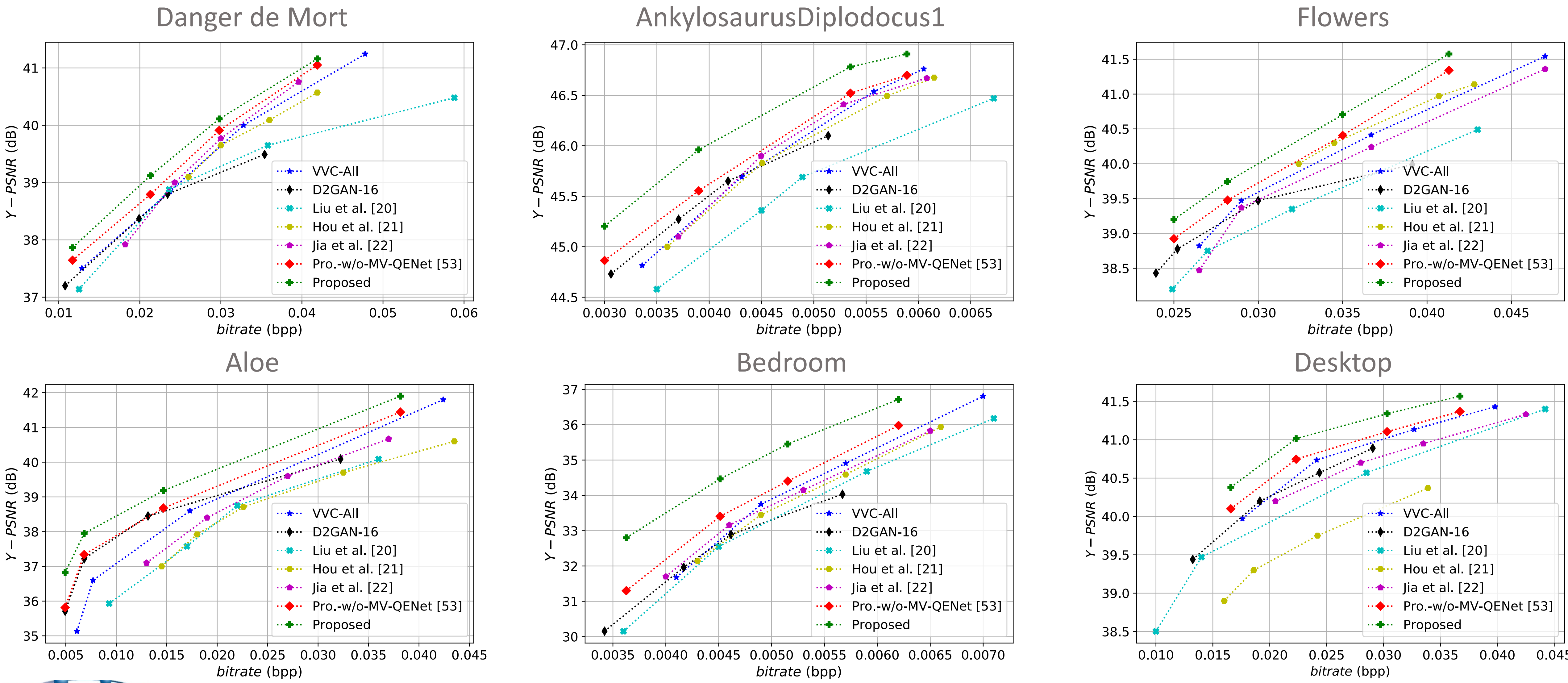
Central perspective view from each **LF** image used in the test



N. Bakir, ***W. Hamidouche*** *et al.*, [Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN](), IEEE T. on Multimedia 2021.

Experimental Results

❑ RD curves of the 5 considered solutions for the 6 LF images



N. Bakir, **W. Hamidouche** et al., Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Experimental Results

TOP-2 TOP-1

- BD-BR and BD-PSNR gains calculated against anchor method described in [Lui at al. 2016].

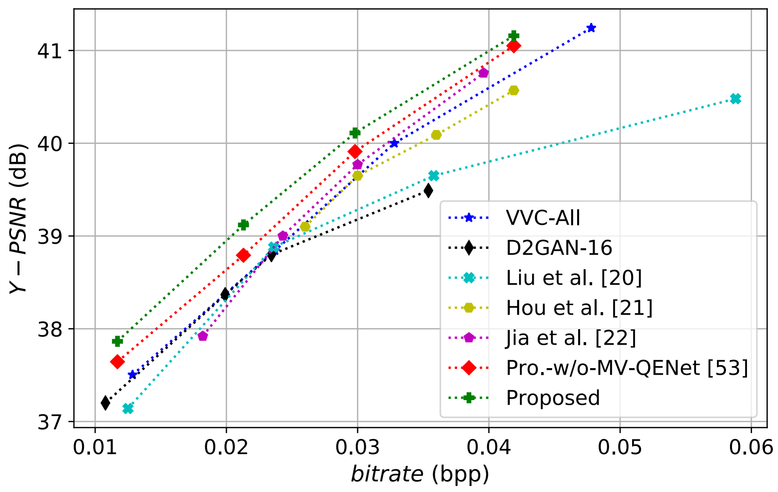
Image	BD-BR versus [Lui et al. 2016]									
	VVC-All		[Jia et al. 2018]		[Hou et al. 2019]		Ours w/o QE		Ours	
	BD-BR	BD-PSNR	BD-BR	BD-PSNR	BD-BR	BD-PSNR	BD-BR	BD-PSNR	BD-BR	BD-PSNR
Bikes	-11.7%	0.72	-6.3%	0.48	-6.9%	0.49	-22.4%	0.96	-31.56%	1.19
DangerDeMort	-7.8%	0.22	-10.8%	0.28	-8.7%	0.26	-16.5%	0.4	-25.69%	0.78
Flowers	-12.3%	0.56	-11.9%	0.54	-16.2%	0.72	-16.6%	0.74	-23.66%	1.03
Ankylosaurus Dip1	-13.2%	0.44	-14.9%	-0.72	-12.3%	0.39	-18.0%	0.57	-31.17%	1.15
Aloe	-26.4%	0.85	-9.1%	0.31	-2.46%	-0.12	-42.3%	1.23	-56.59%	1.84
StonePillarsOutside	-18.3%	0.61	-15.1%	0.52	-11.9%	0.28	-35.6%	0.98	-49.76%	1.42
Bedroom	-5.3%	0.46	-4.0%	0.32	-2.3%	0.18	-9.5%	0.85	-24.78%	2.11
Desktop	-19.6%	0.32	-7.5%	0.11	44.1%	-0.61	-26.3%	0.45	-40.58%	0.79
Herbs	-26.2%	1.14	-4.4%	-0.11	6.9%	-0.21	-29.8%	1.32	-42.25%	1.85
Average	-15.6%	0.59	-8.3%	0.35	-0.54%	0.15	-24.1%	0.83	-36.22%	1.35

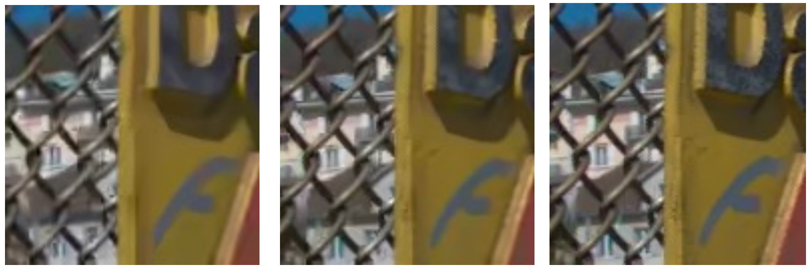
N. Bakir, ***W. Hamidouche*** *et al.*, Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Experimental Results

❑ Visual comparison LF images decoded by different codecs

Danger de Mort





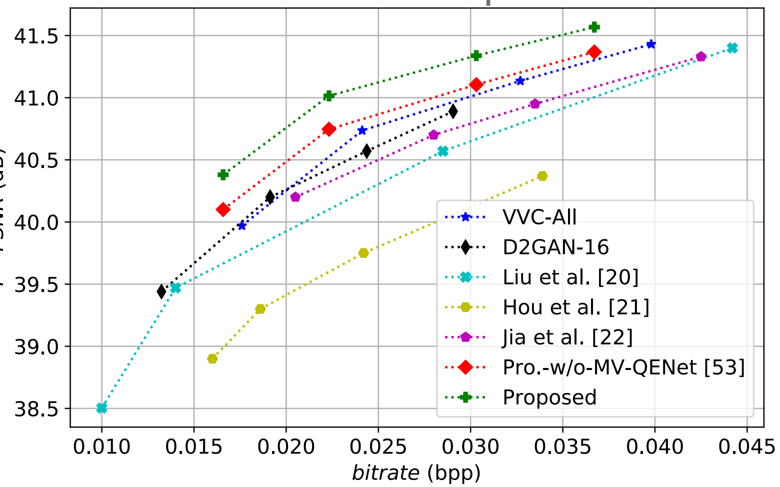
[Liu et al.]
y-psnr: 39.15

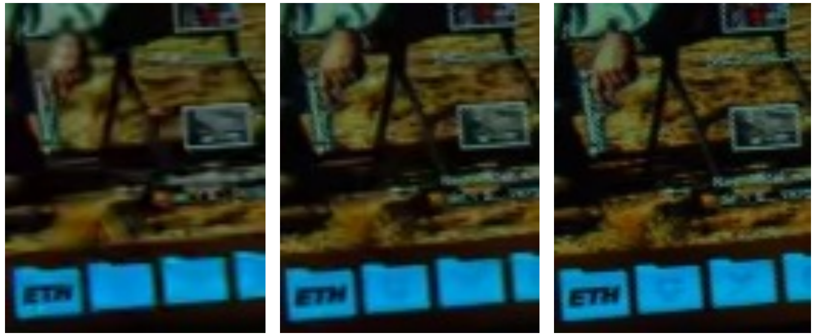
VVC-All
y-psnr: 39.86

Our w/o QE
y-psnr: 40.08

@0.032 bpp

Desktop





[Liu et al.]
y-psnr: 40.12

VVC-All
y-psnr: 40.36

Our w/o QE
y-psnr: 40.65

@0.021 bpp

*N. Bakir, **W. Hamidouche** et al.*, Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Experimental Results

- Complexity analysis : run time on CPU (Inter i9-7900X 3.3GHz) & GPU (NVDIA TITAN XP)

	Encoder run time (second)				
	VVC-All	[Jia et al. 2018]	[Hou et al. 2019]	Ours	
QP	**CPU**	**GPU**	**CPU**	**CPU**	**GPU**
22	**259**	450	6028	559	449
26	**152**	350	6028	452	342
32	**101**	220	6028	401	291
37	**66**	142	6028	366	256
Average	**145**	291	6028	445	335

	Decoder run time (second)				
Average	**4**	53	583	124 to 333	94 to 285

N. Bakir, ***W. Hamidouche*** *et al.,* Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Outline

- Context & challenges
- Proposed solution
- Experimental Results
- **Conclusion**

Conclusion

❑ Contributions

- New **D2GAN** model for **LF** image **S**ynthesis
- Efficient coding solution outperforming the state-of-the-art learning based coding approach
- **O**pen **Q**uestions
  - ➢ **E**ncode or **S**ynthetize ?
  - ➢ What about the **D**epth **I**nformation ?
  - ➢ **S**ubjective **E**valuation ?

*N. Bakir, **W. Hamidouche** et al.*, Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Visual illustration and web page

https://naderbakir79.github.io/LFD2GAN.html

N. Bakir, ***W. Hamidouche*** *et al.*, Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN, IEEE T. on Multimedia 2021.

Thank you for your attention

N. Bakir, ***W. Hamidouche*** *et al.*, [Light Field Image Coding Using VVC standard and View Synthesis based on Dual Discriminator GAN](), IEEE T. on Multimedia 2021.

❑ References

[**Schiopu et al. 2017**] Schiopu et al. **Lossless compression of subaperture images using context modeling**, 2017 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON) 2017.

[**Jiang et al. 2017**] Jiang et al. **Light Field Compression with Homography-based Low Rank Approximation**, IEEE Journal of Selected Topics in Signal Processing (J-STSP), vol. 11, No. 7, pp. 1132-1145, Oct. 2017

[**Liu et al. 2016**] Liu et al., **Pseudo-Sequence-Based Light Field Image Compression**, in 2016 IEEE International Conference on Multimedia Expo Workshops (ICMEW), July 2016, pp. 1–4.

[**Li et al. 2017**] Li et al., **Pseudo-Sequence-Based 2-D Hierarchical Coding Structure for Light-Field Image Compression**, in IEEE Journal of Selected Topics in Signal Processing, vol. 11, no. 7, pp. 1107-1119, Oct. 2017.

[**Zhao et al. 2016**] Zhao et al., **Light field image coding with hybrid scan order**. pp 1-4. VCIP. 2016.

[**Kalantari et al. 2016**] Kalantari et al., “Learning-based view synthesis for light field cameras," ACM Trans. Graph., vol. 35, no. 6, pp. 193:1-193:10, Nov. 2016.

[**Zhao et al. 2017**] Zhao et al., **Light field image coding via linear approximation prior**, in Proc. IEEE International Conference on Image Processing (ICIP), Sept 2017.

[**Jia et al. 2018**] Jia et al., **Light Field Image Compression Using Generative Adversarial Network Based View Synthesis**, IEEE Journal on Emerging and Selected Topics in Circuits and Systems, pp. 1–1, 2018.

[**Sullivan et al. 1998**] Sullivan et al., “Rate-distortion optimization for video compression”, IEEE signal processing magazine, 1998.

[**Hou et al. 2019**] Hou et al., **Light Field Image Compression Based on Bi-Level View Compensation With Rate-Distortion Optimization**, in IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 2, pp. 517-530, Feb. 2019.