Digital Multimedia Lab

Light Field:
A Quest for the Perfect Picture

“The Holodeck”

Dr. Panos Nasiopoulos

Electrical and Computer Engineering

IARIA

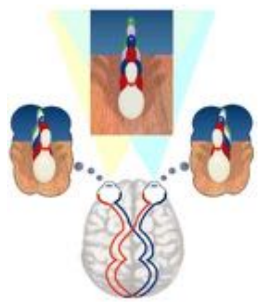
March 2023

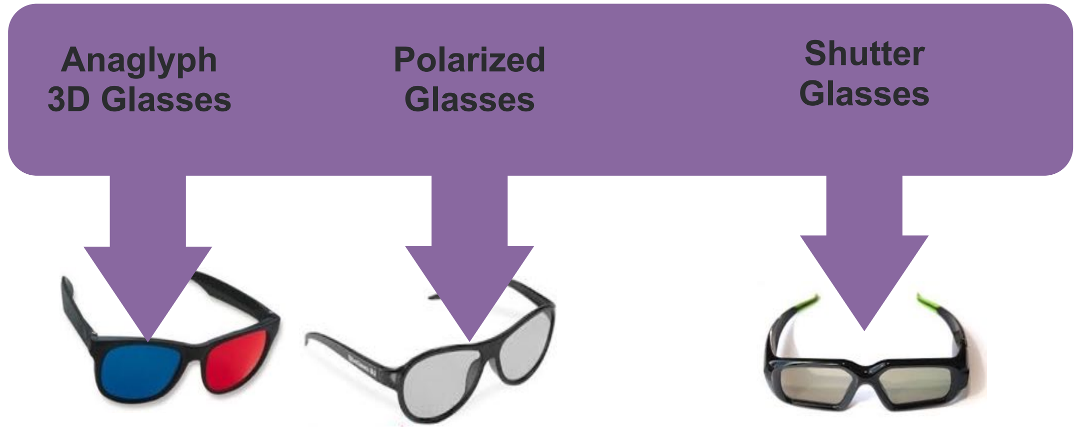
Hypervsn – holographic solutions

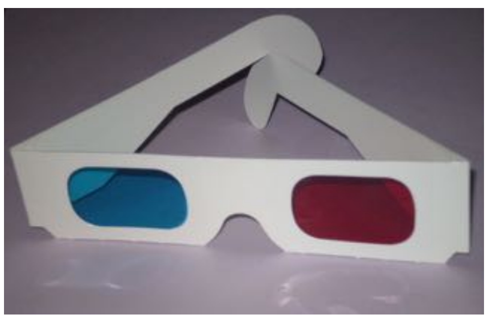
https://www.youtube.com/watch?v=4a0_Vb0eIek

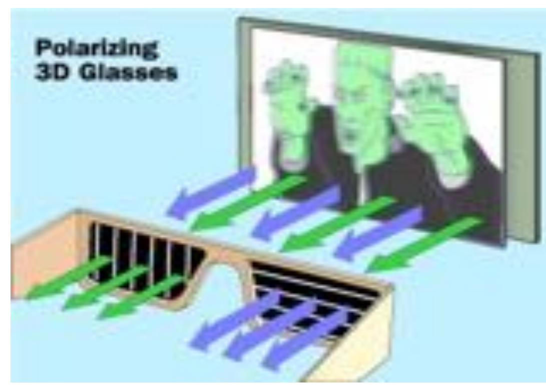

© Star Trek

Capture and display what the human eye can see









Glasses

Capture and display what the human eye can see

Multiview displays

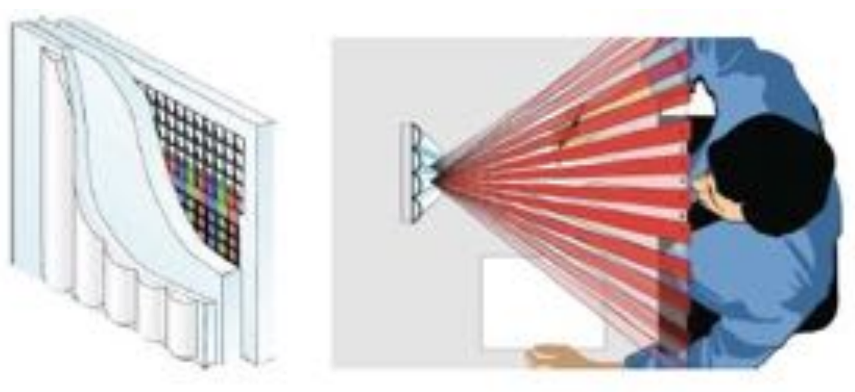


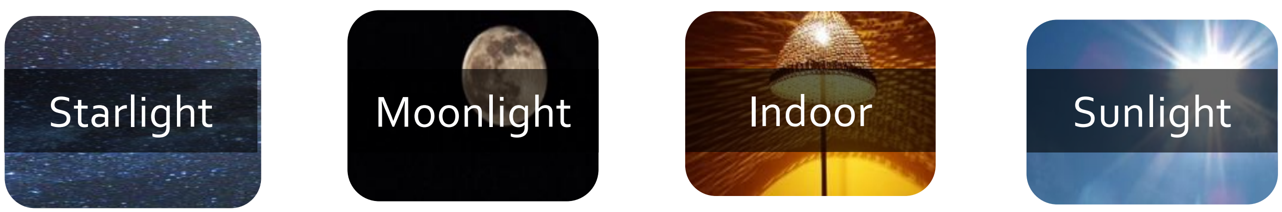
Image source: DIMENCO

Image source: Alioscopy

FOCUS & CONVERGENCE

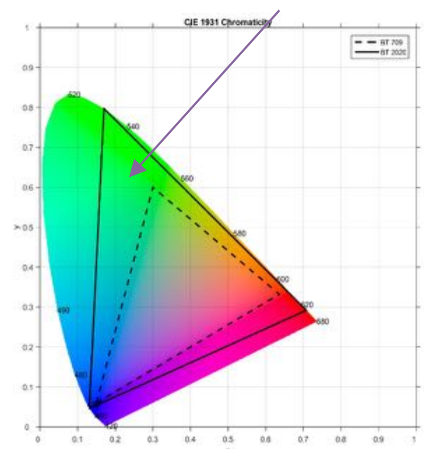
Capture and display what the human eye can see

4K: improved resolution
HDR: High Dynamic Range



10^{-6} 10^{-4} 10^{-2} 10^{0} 10^{2} 10^{4} 10^{6}

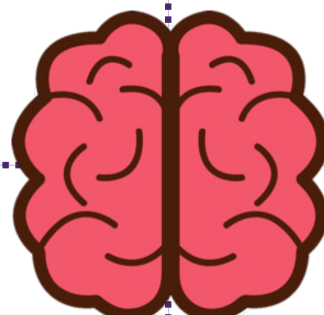
HDR



A quest for the perfect picture

We live in a Visual World

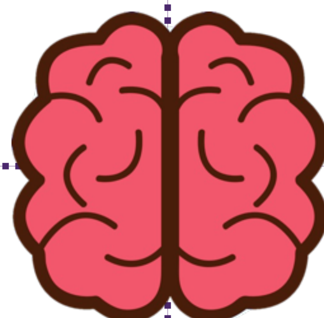
90% of information we process is visual



We live in a Visual World

90% of information we process is visual

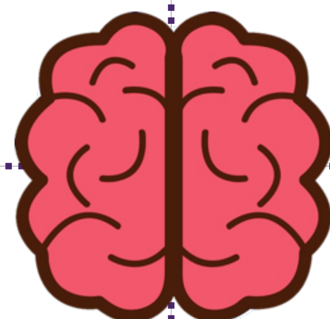
50% of our brain actively processes visual content



We live in a Visual World

90% of information we process is visual

50% of our brain actively processes visual content

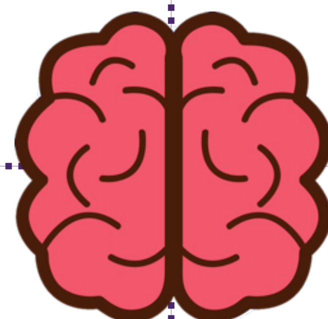


We process visuals **60,000x** faster than text or audio

We live in a Visual World

90% of information we process is visual

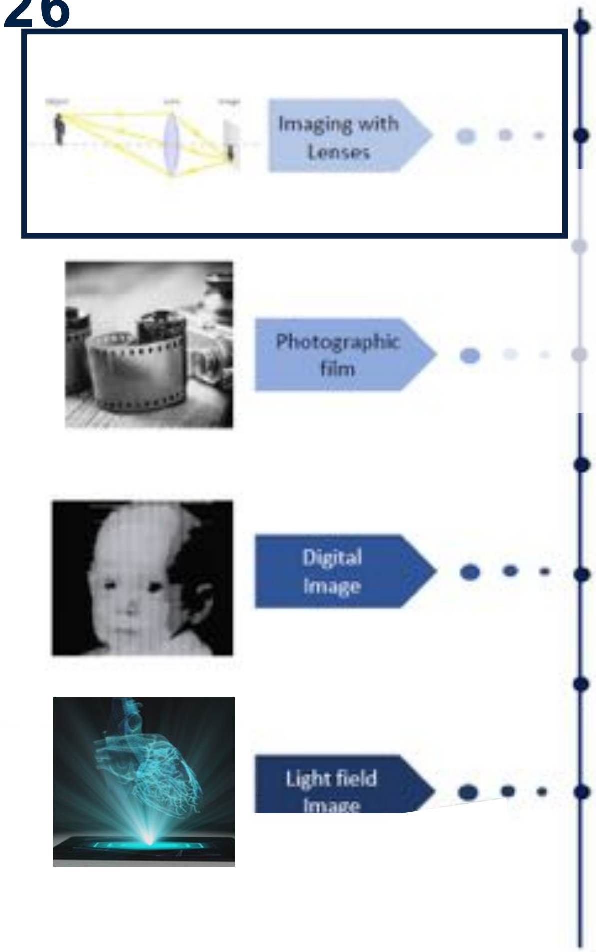
50% of our brain actively processes visual content



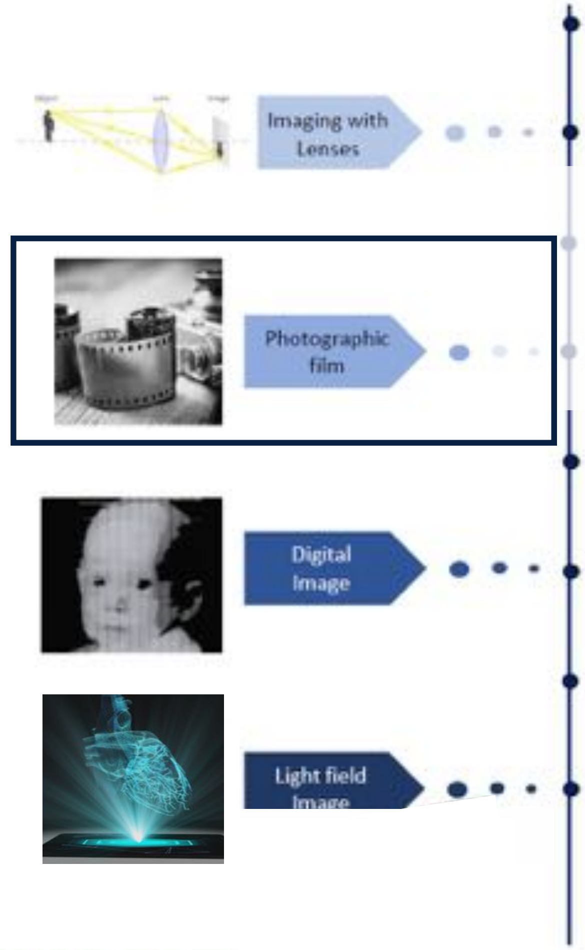
We process visuals **60,000x** faster than text or audio

We retain **80%** of what we see

FIRST PHOTO TAKEN BY A CAMERA - 1826

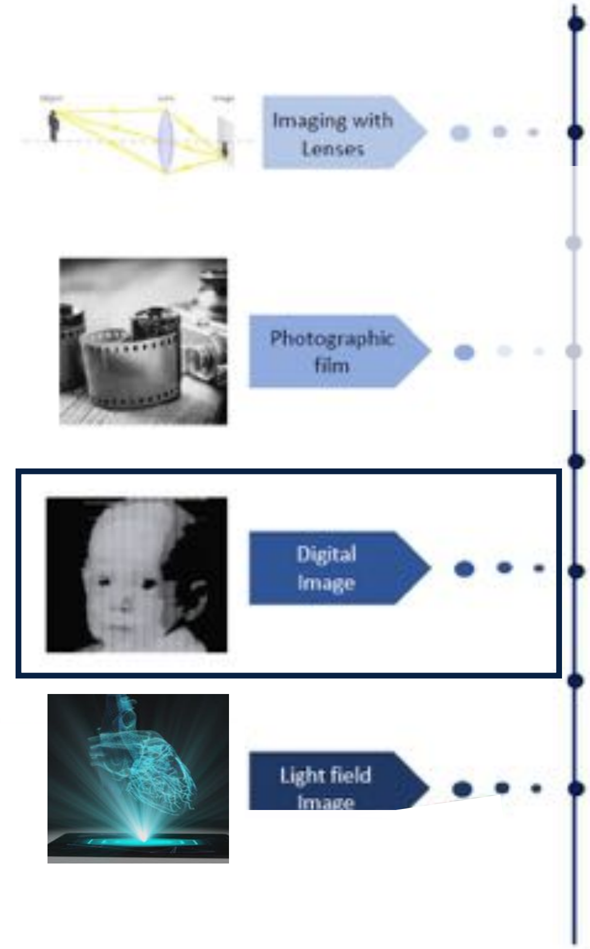


FIRST COLOUR PHOTOGRAPH - 1861

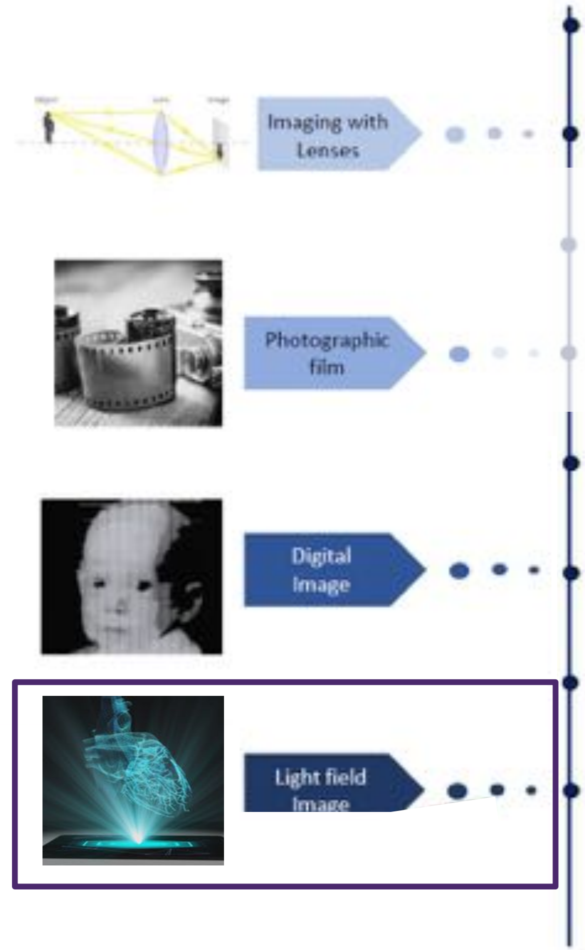


FIRST DIGITAL PHOTOGRAPH - 1957

176 x 176



HOLODECK – 20??


Imaging with Lenses
Photographic film
Digital Image
Light field Image

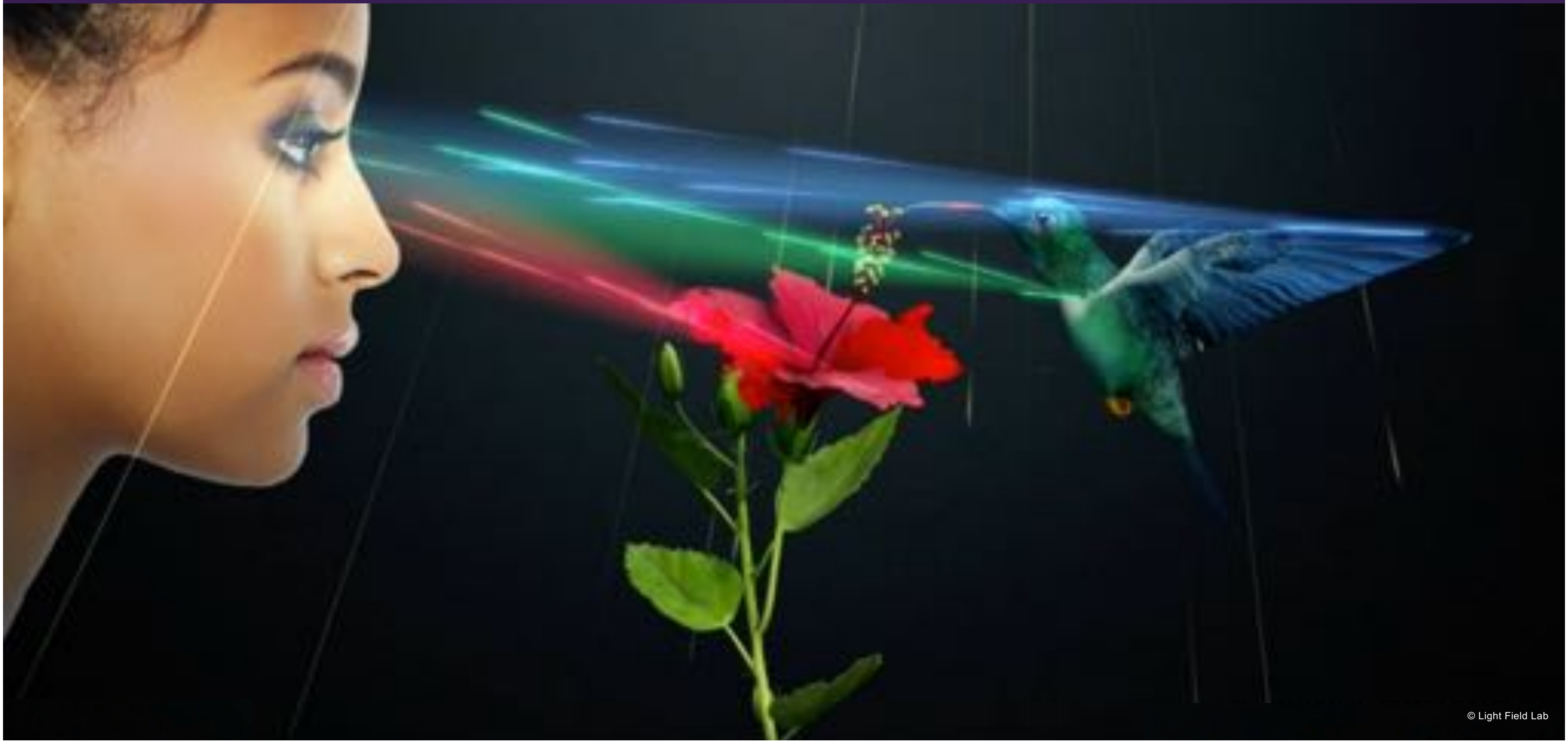
WE SEE THE WORLD BECAUSE OF LIGHT

© Light Field Lab

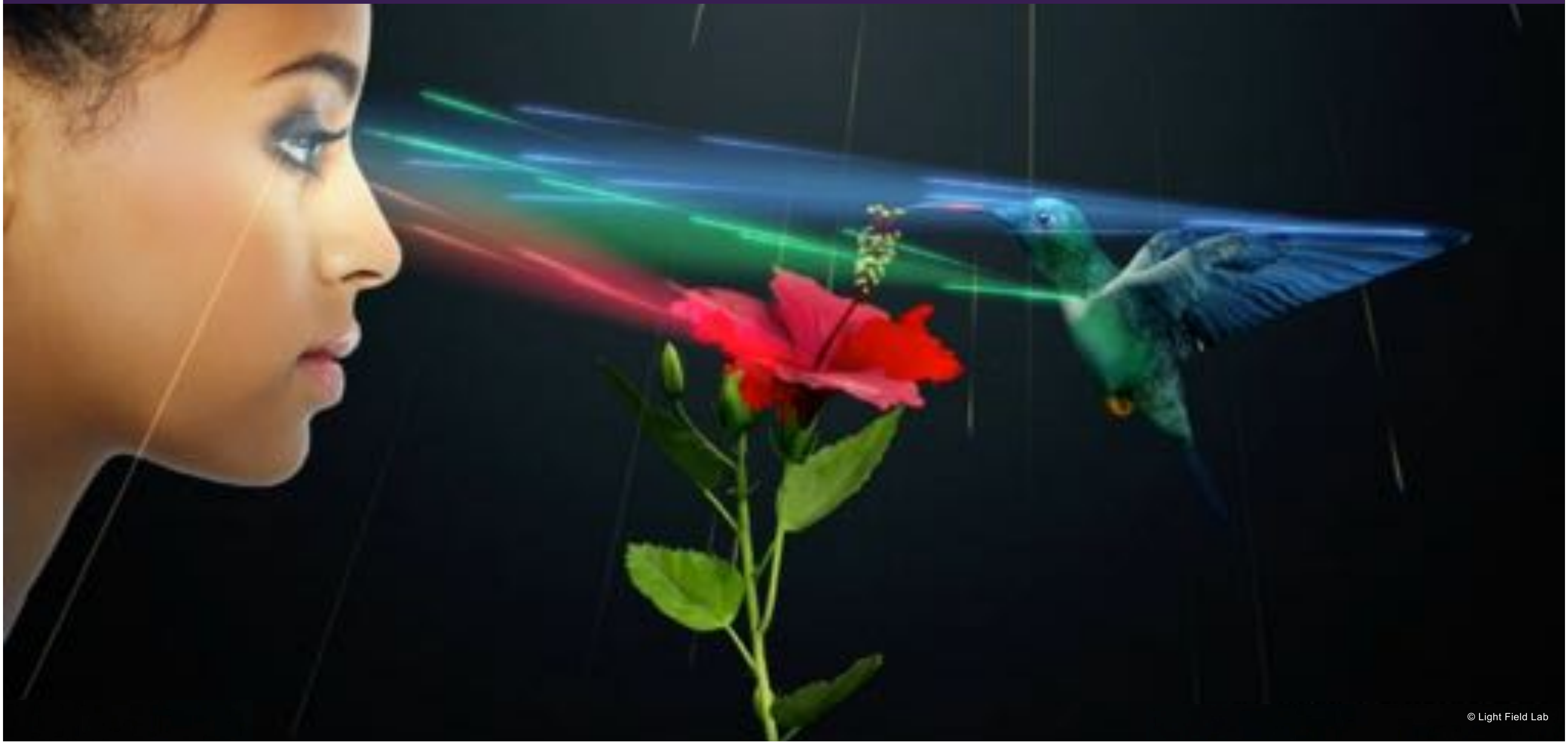
WE SEE THE WORLD BECAUSE OF LIGHT

WE SEE THE WORLD BECAUSE OF LIGHT

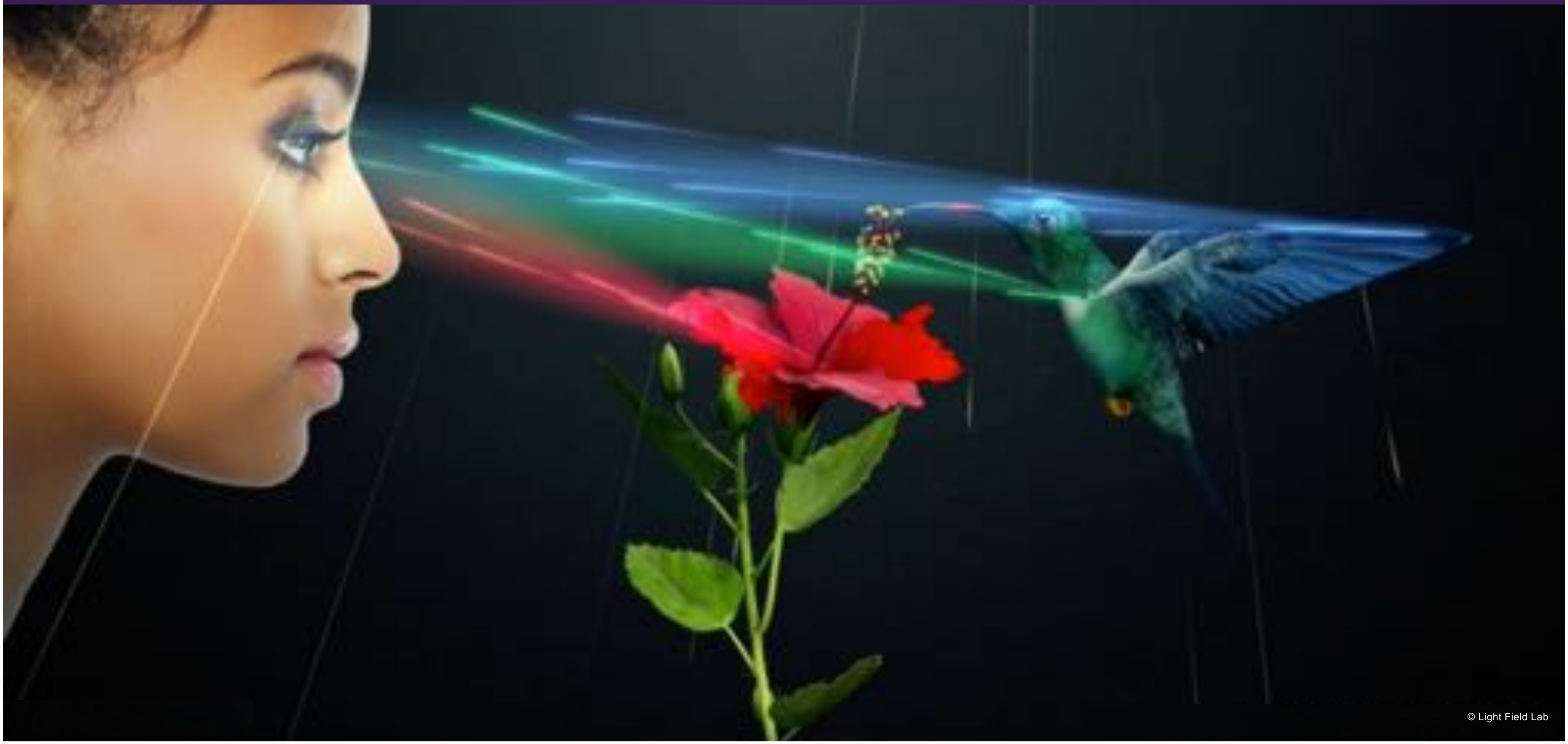
WE FOCUS ON CONVERGING RAYS



WE FOCUS ON CONVERGING RAYS



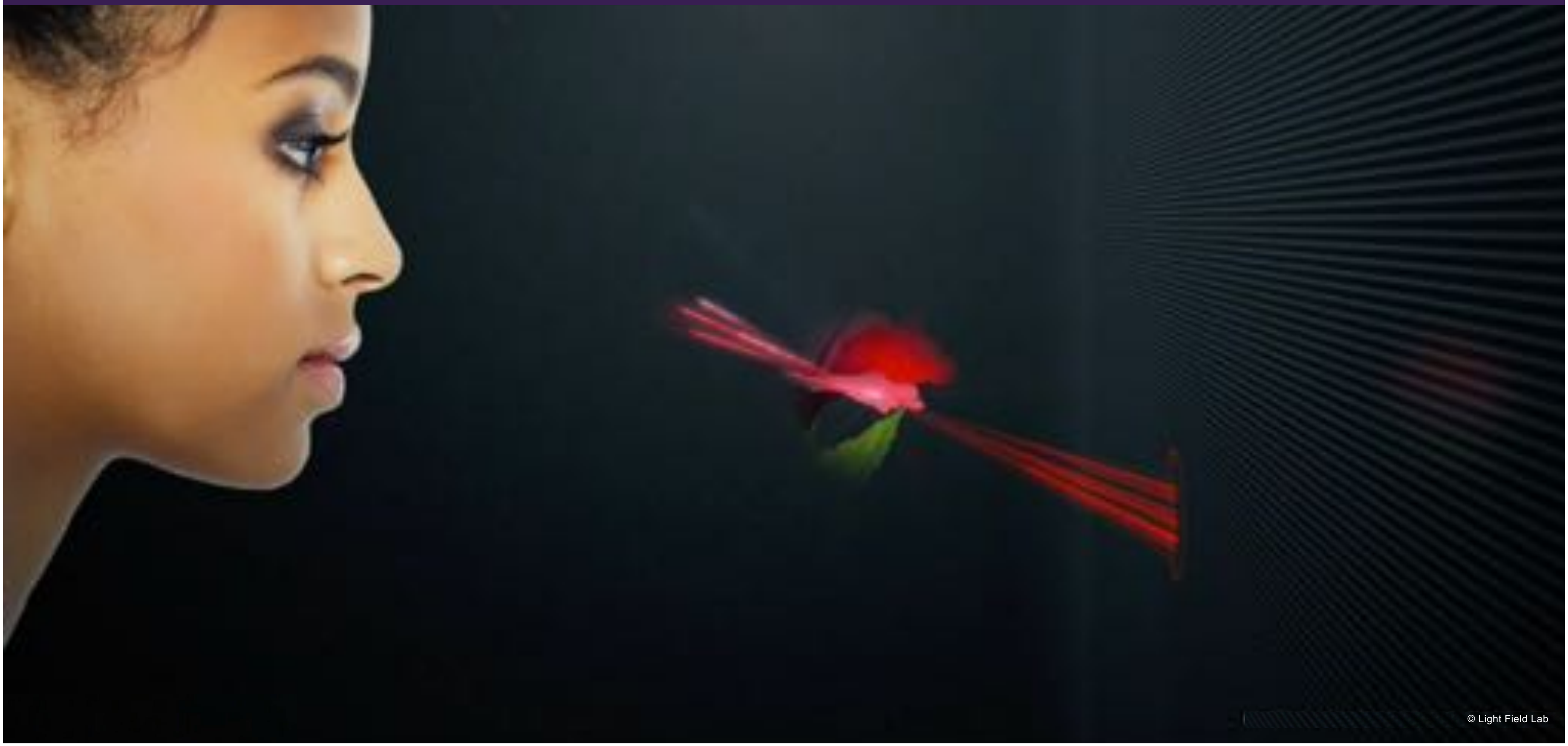
WE FOCUS ON CONVERGING RAYS



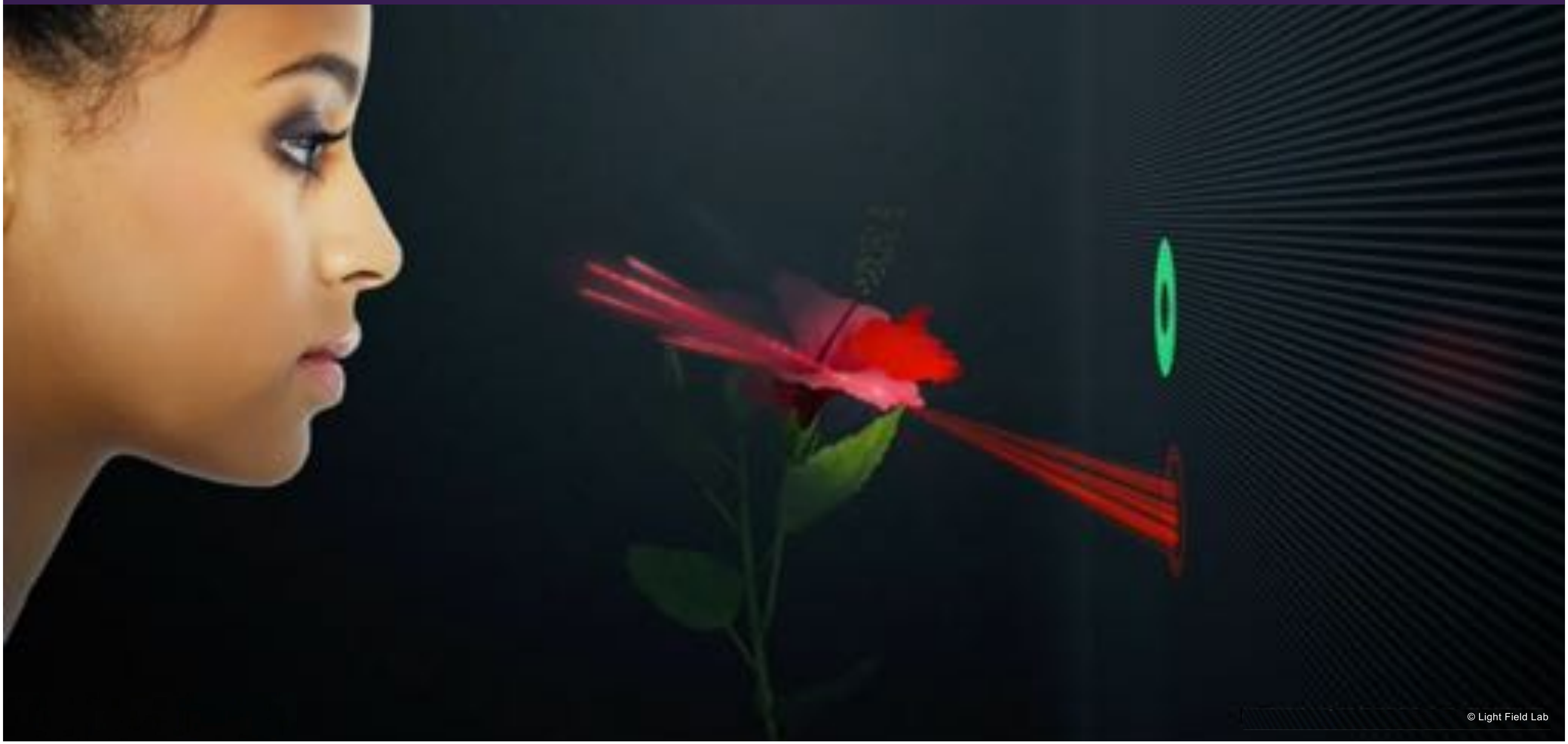
HOW TO RECREATE THE SCENE USING LIGHT FIELD?

WE CREATE INVERSE REFLECTION OF LIGHT

WE CREATE INVERSE REFLECTION OF LIGHT

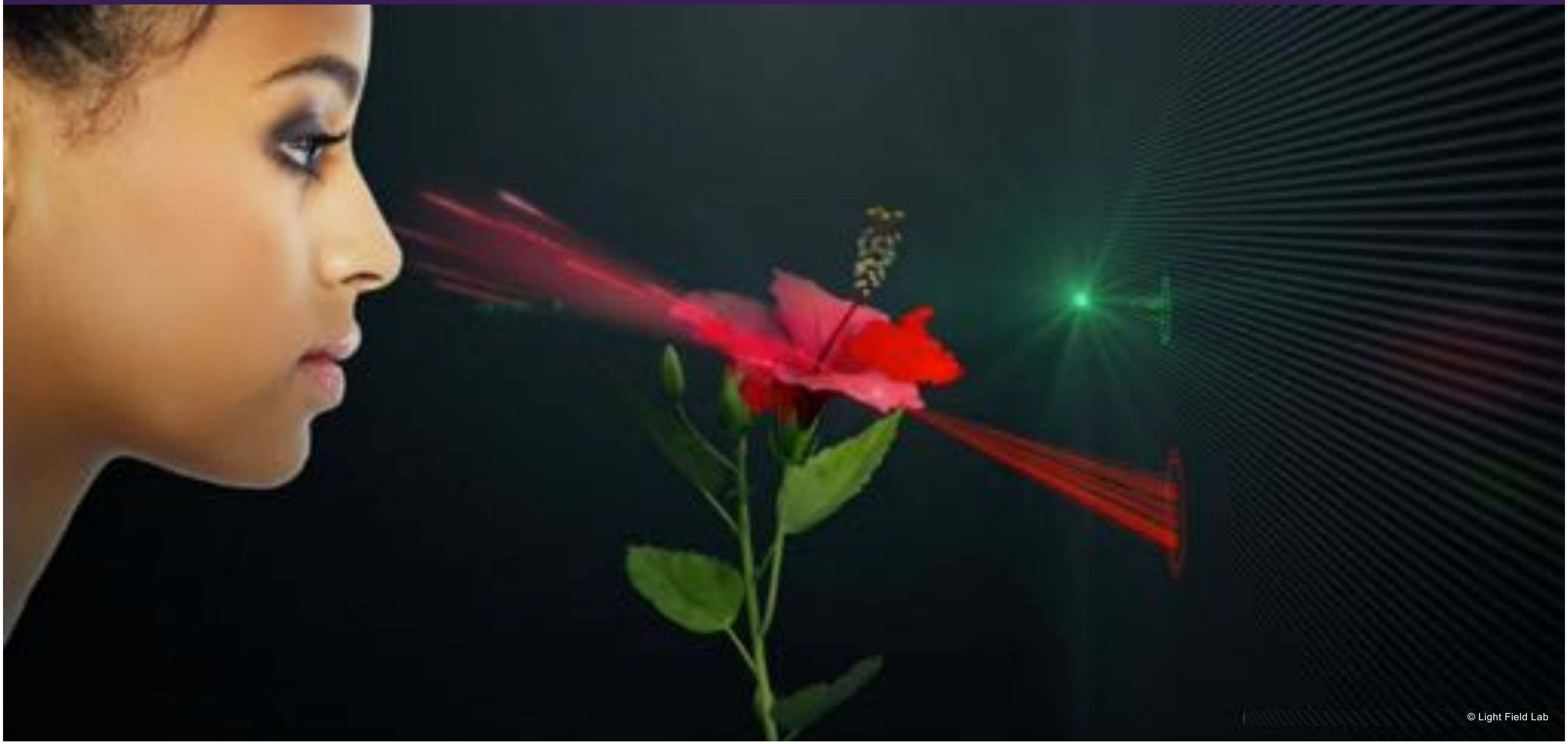


WE CREATE INVERSE REFLECTION OF LIGHT

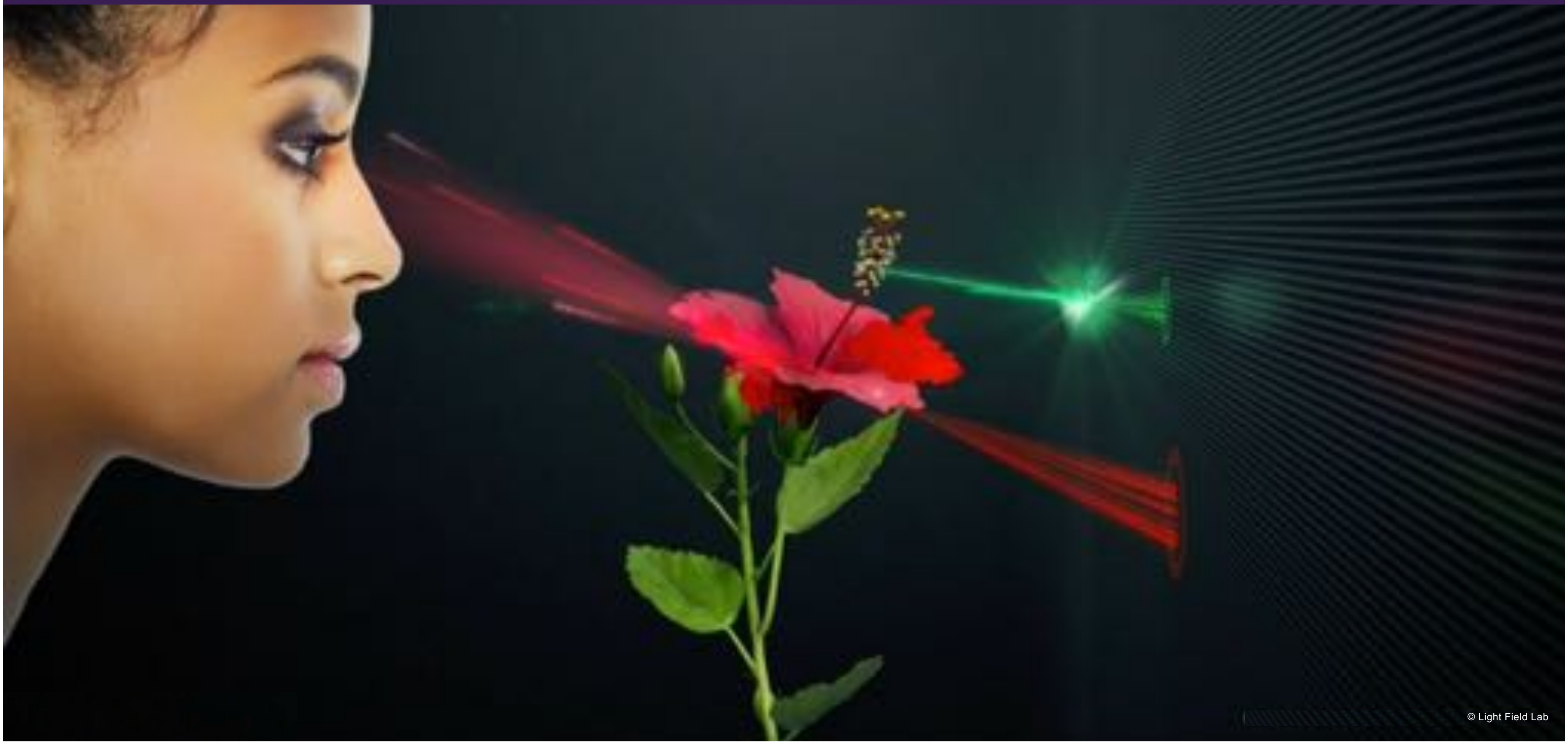


WE CREATE INVERSE REFLECTION OF LIGHT

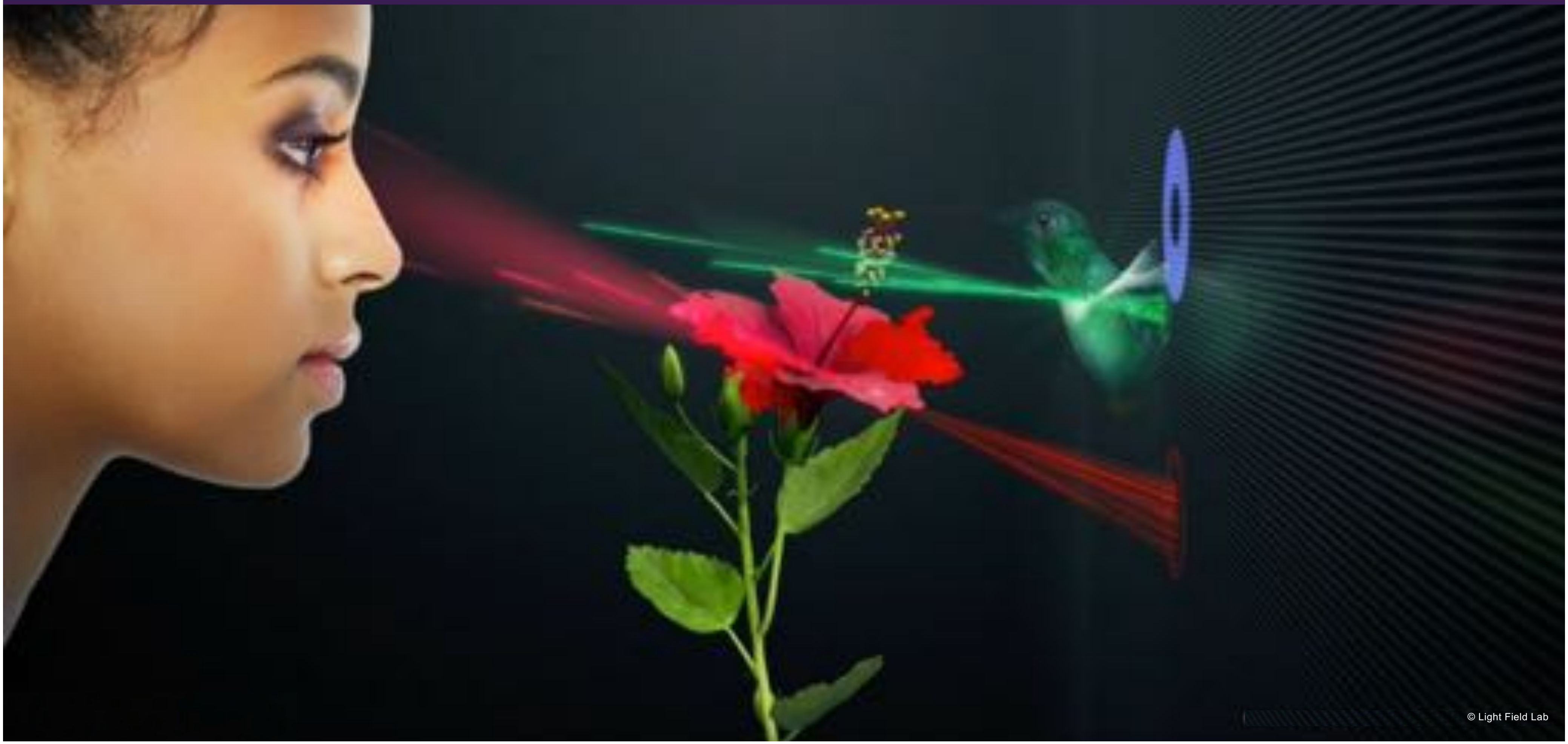
WE CREATE INVERSE REFLECTION OF LIGHT



WE CREATE INVERSE REFLECTION OF LIGHT

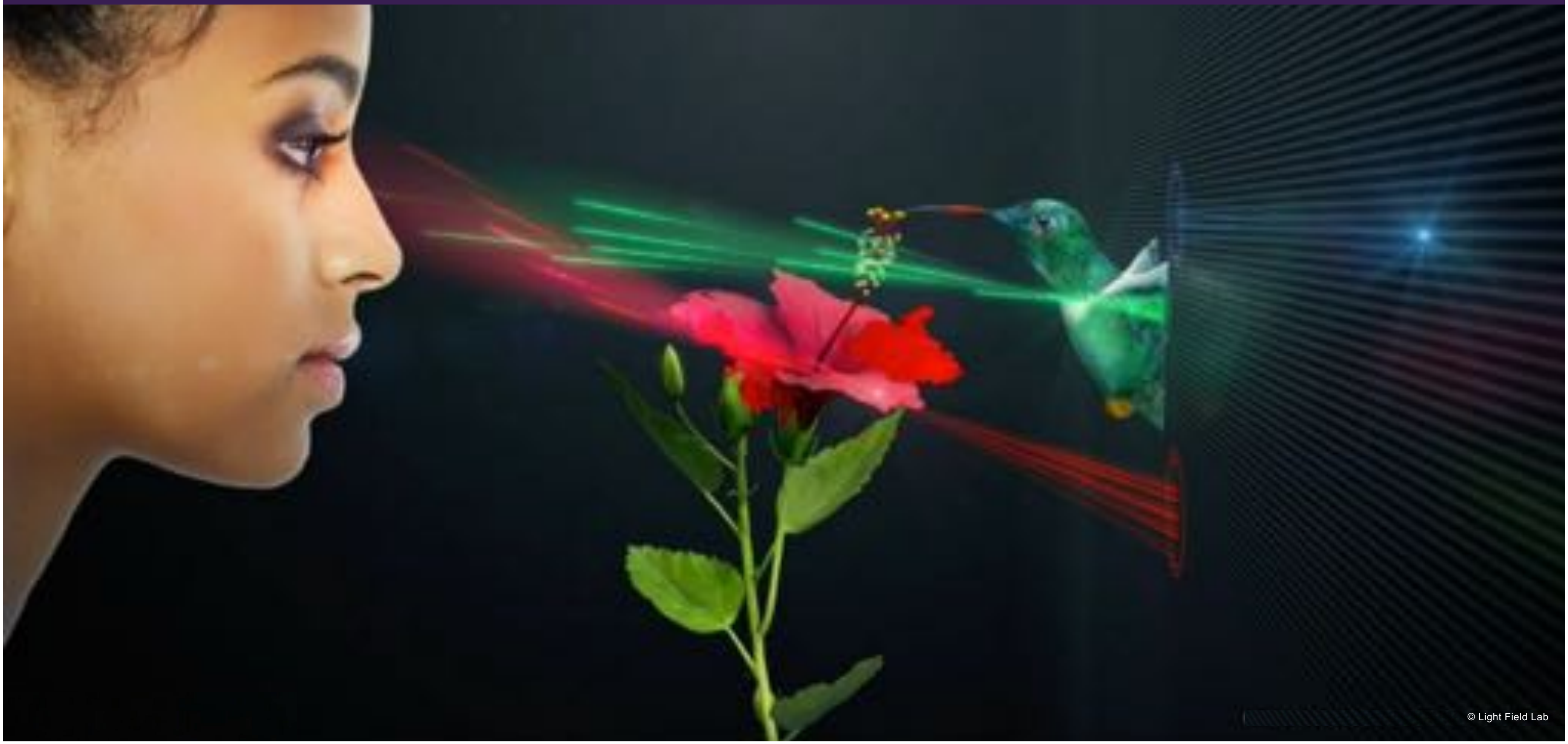


WE CREATE INVERSE REFLECTION OF LIGHT

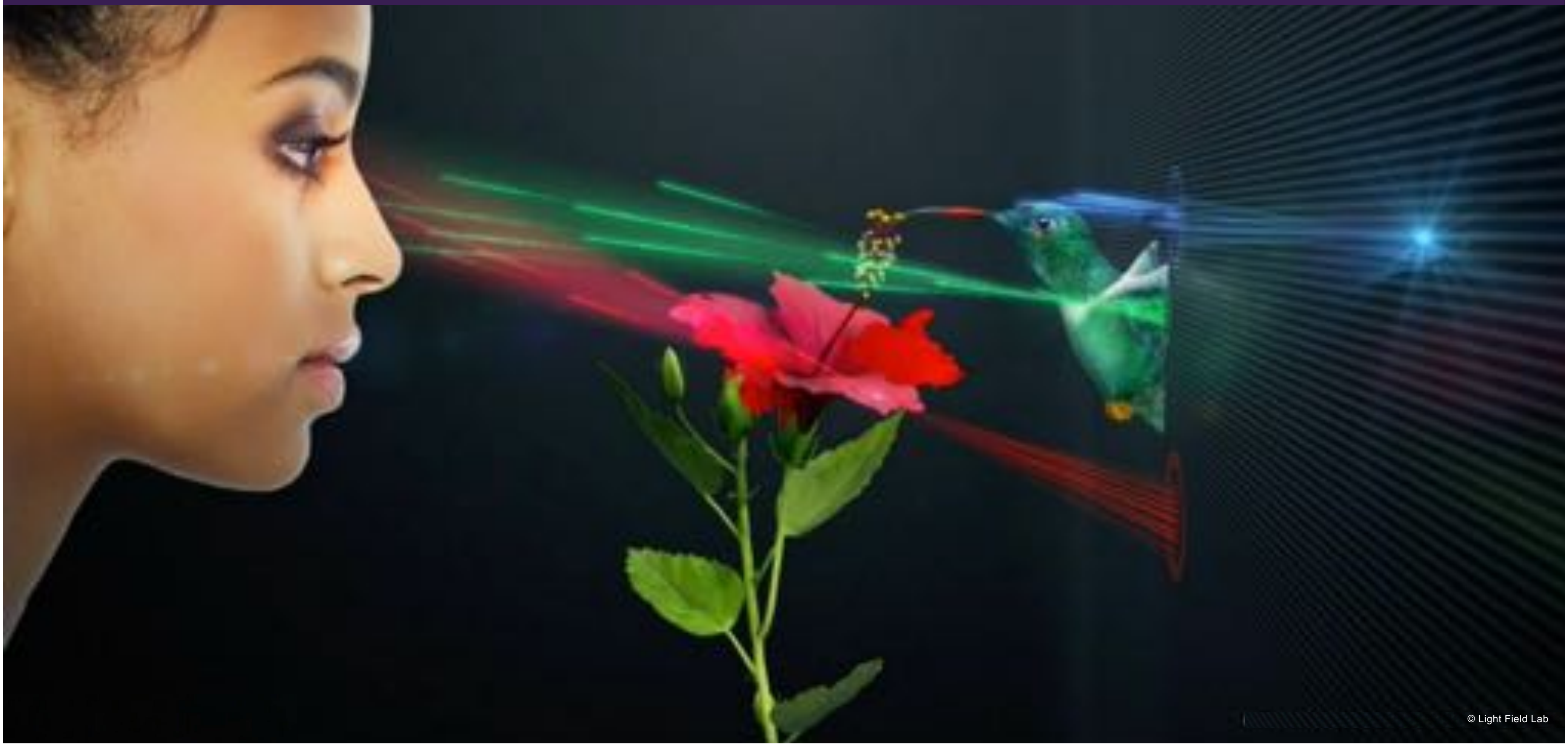


WE CREATE INVERSE REFLECTION OF LIGHT

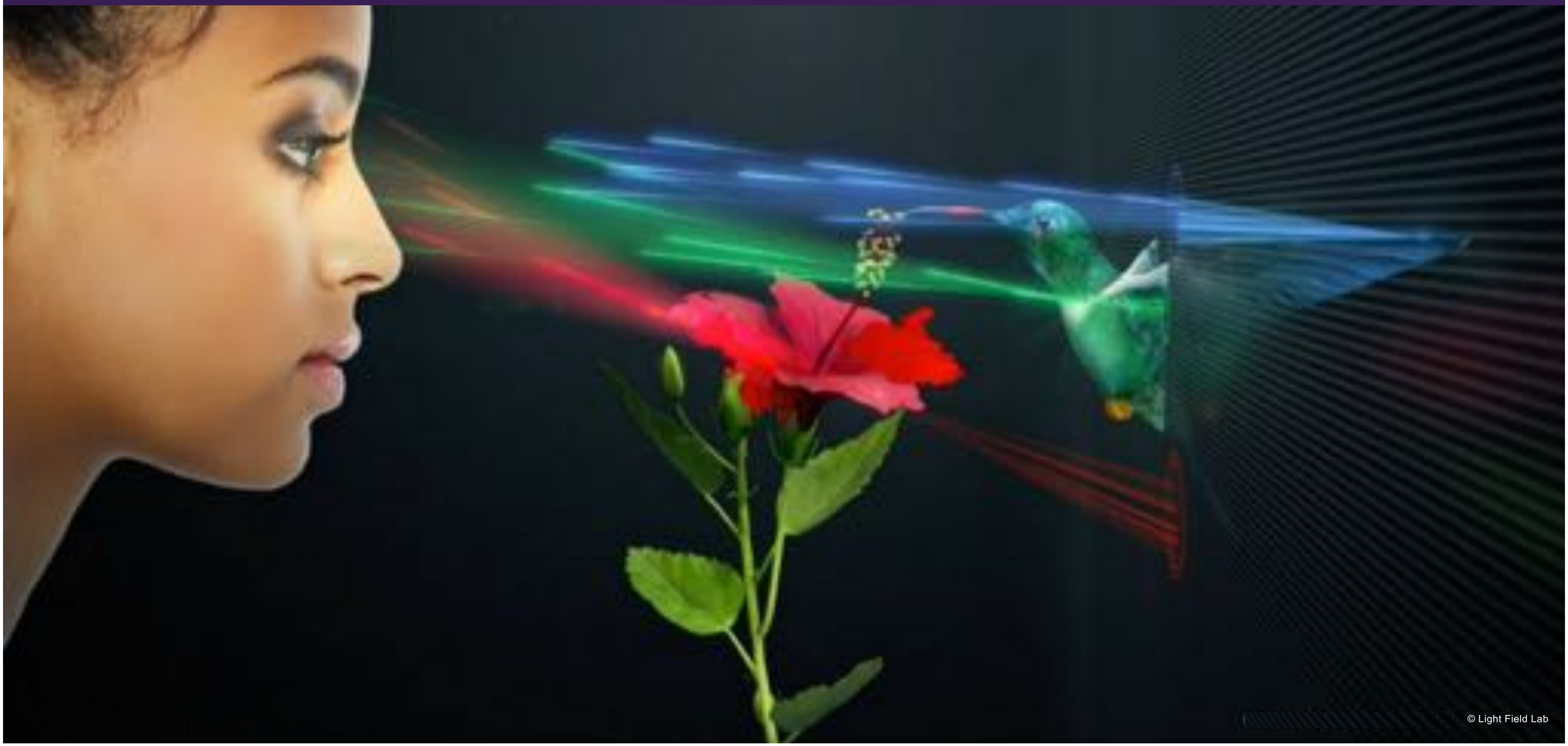
WE CREATE INVERSE REFLECTION OF LIGHT



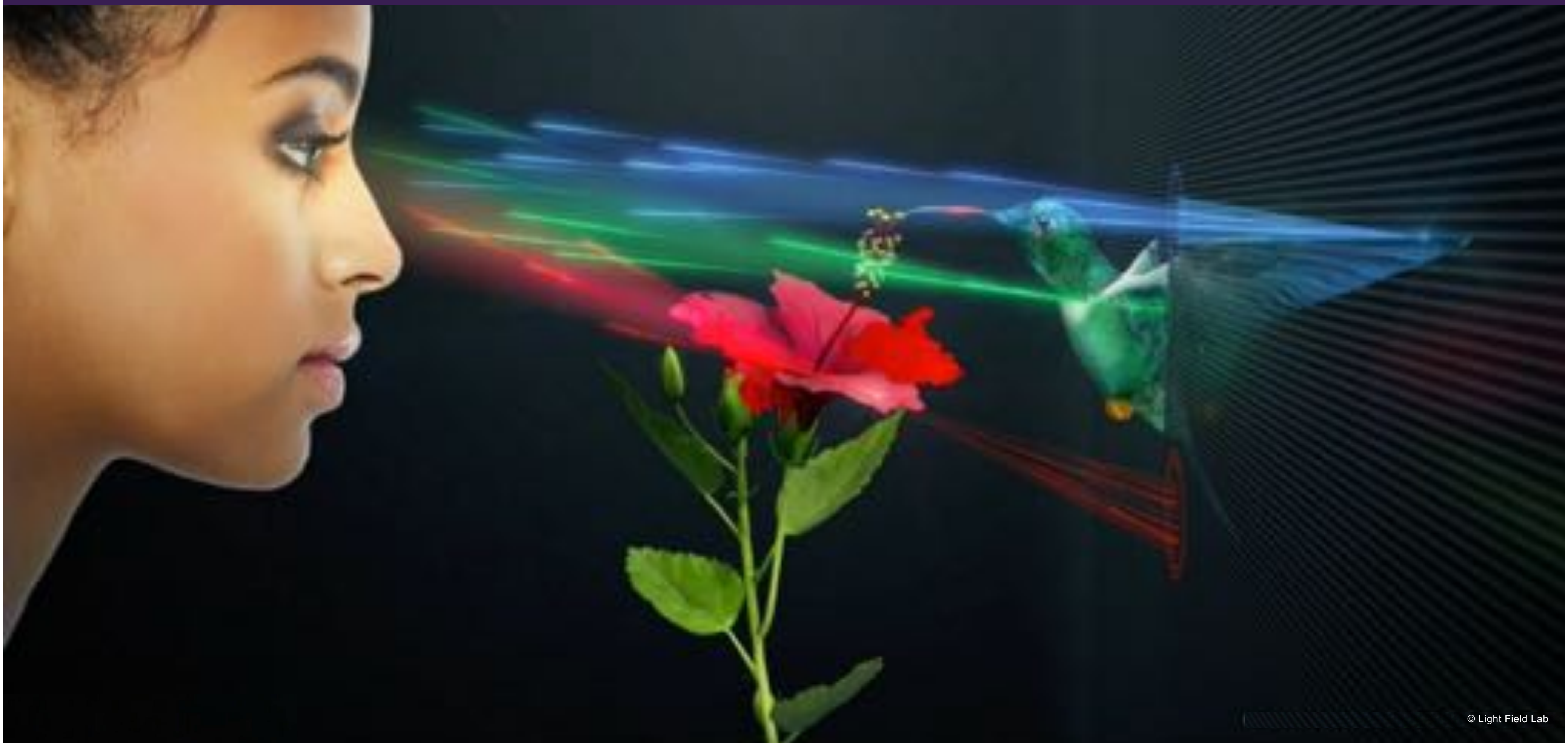
WE CREATE INVERSE REFLECTION OF LIGHT



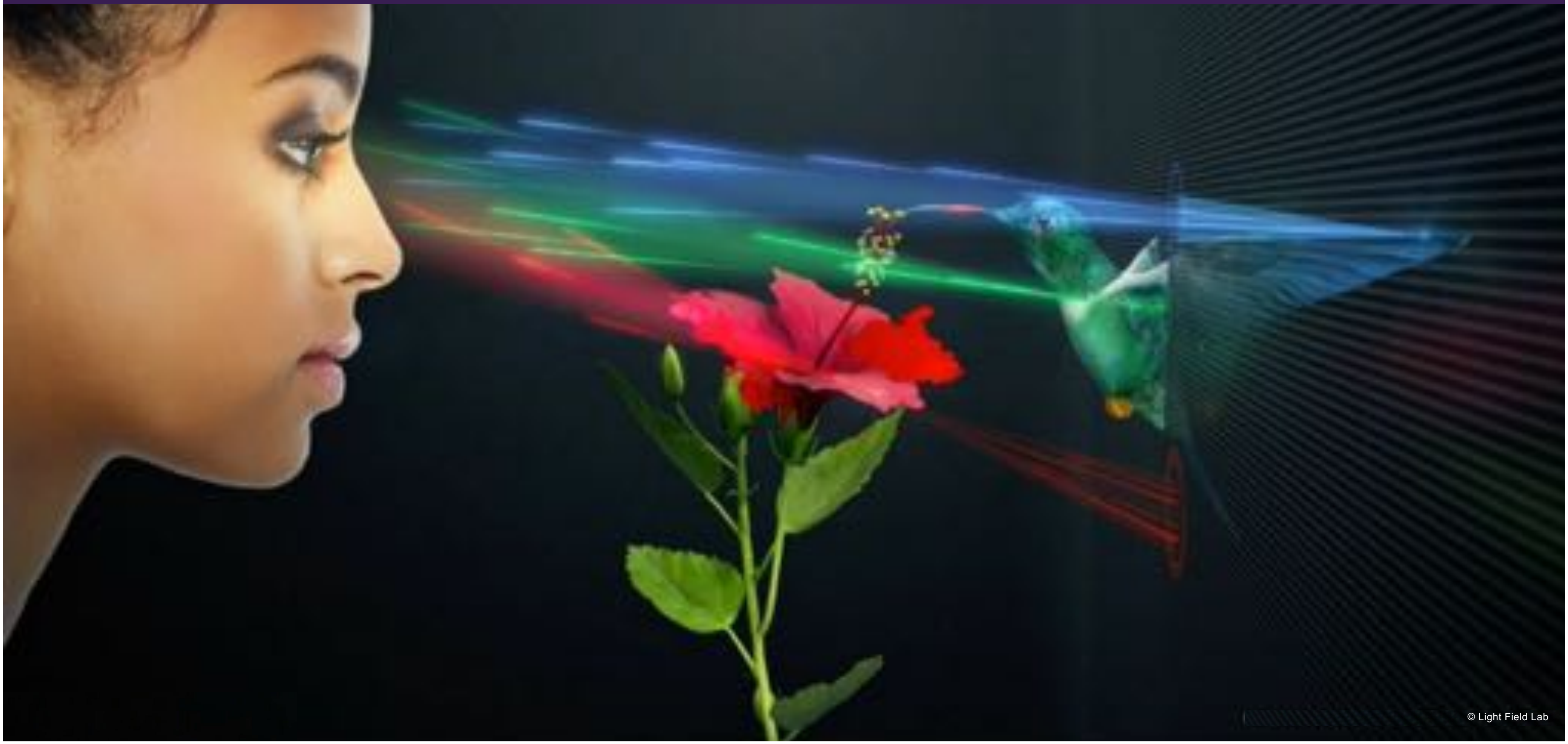
WE CREATE INVERSE REFLECTION OF LIGHT



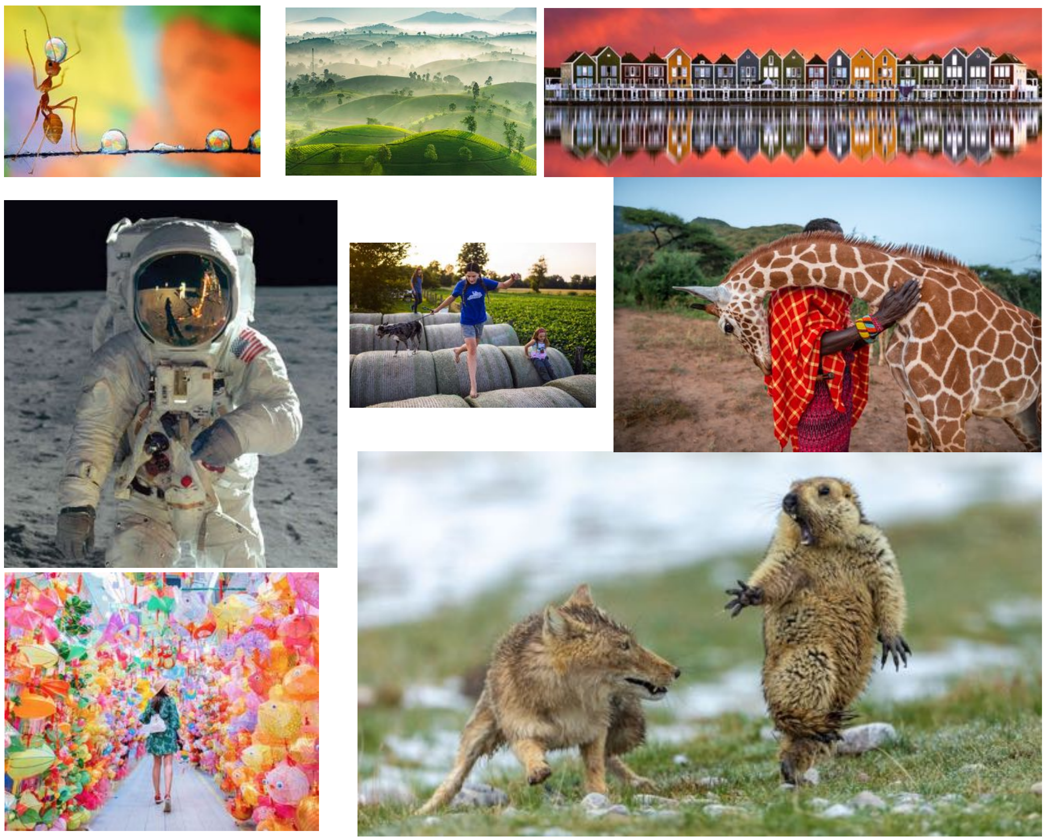
WE CREATE INVERSE REFLECTION OF LIGHT

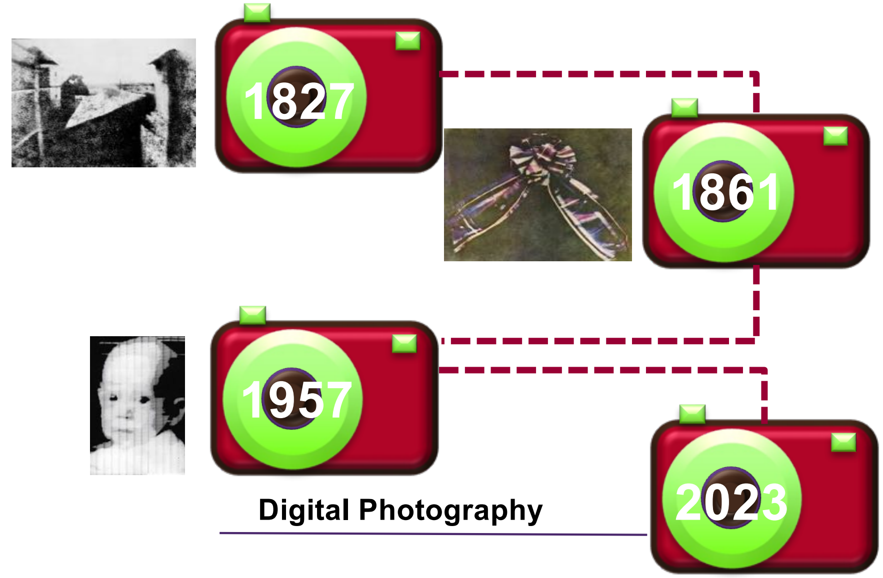


WE CREATE INVERSE REFLECTION OF LIGHT

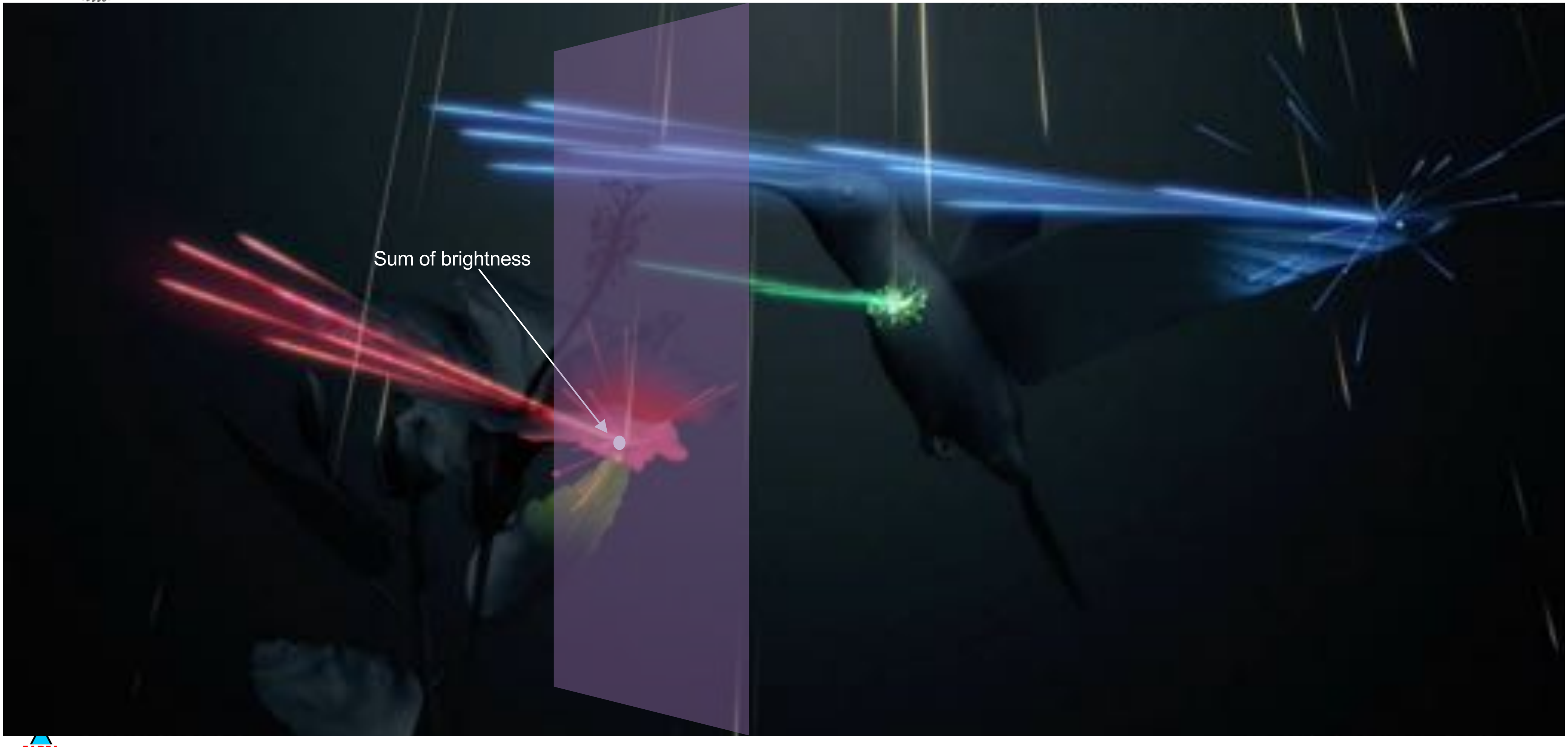


A picture is worth a 1000 words but …





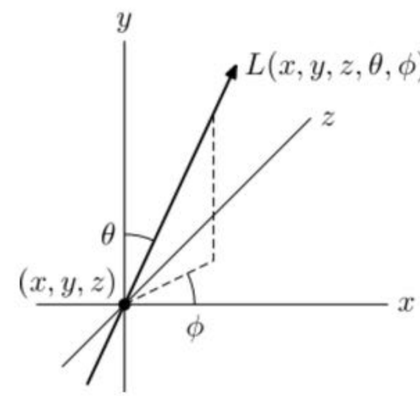
A picture is worth a 1000 words but …



Ideally ….

Concept: every light ray from every point in space

- Extremely complex
- Difficult to achieve



Plenoptic function

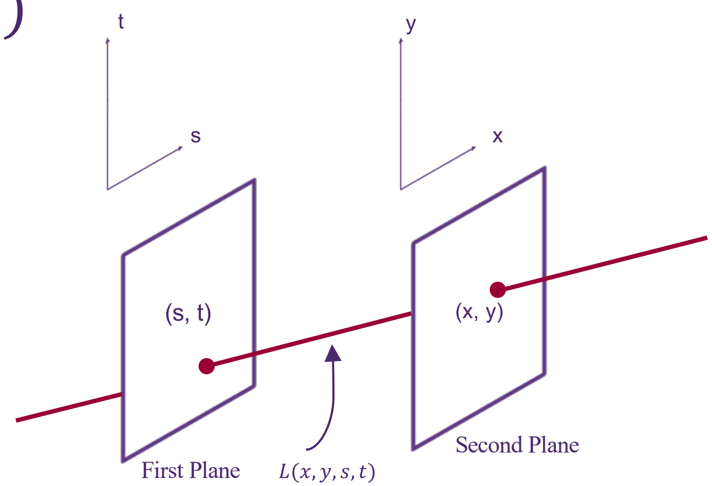
Direction time

$$L(x, y, z, \theta, \varphi, \lambda, t)$$

3D spatial position wavelength

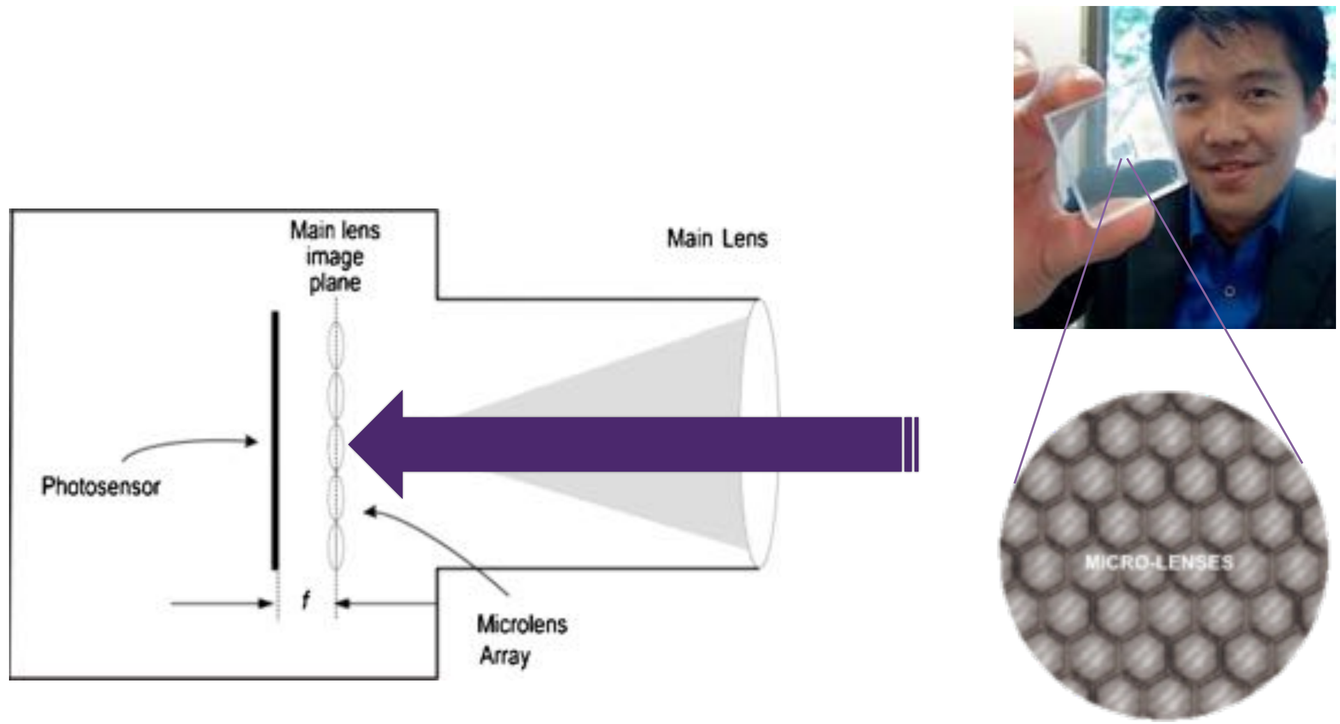
5D: $L(x, y, z, \theta, \varphi)$

4D: $L(x, y, s, t)$



© Adelson, E.H., Bergen, J.R. (1991).

Light Field Technology



2006, Ren Ng

Current Light Field Capture Systems

Microlens

Raytrix

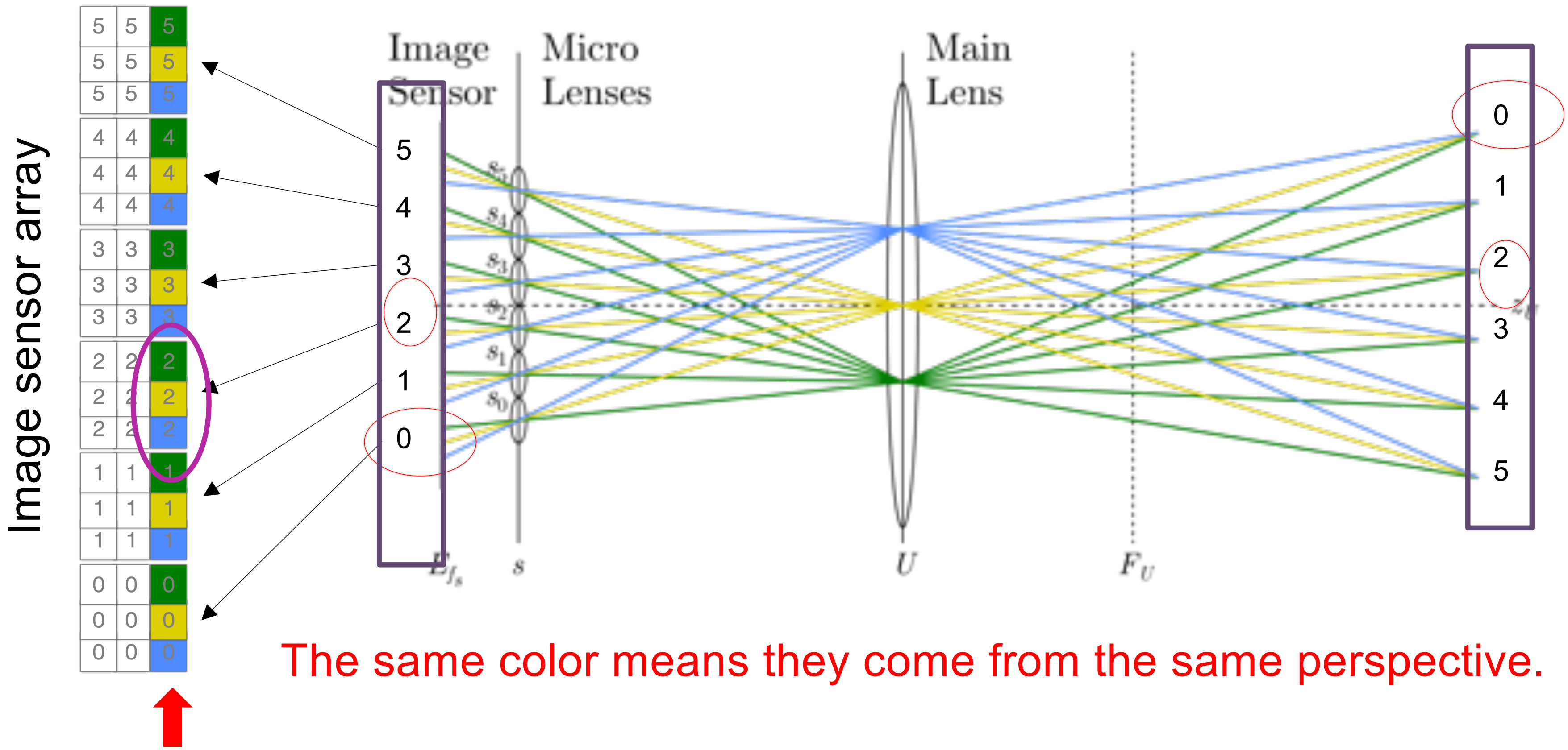
2008

Lytro

Who is involved?

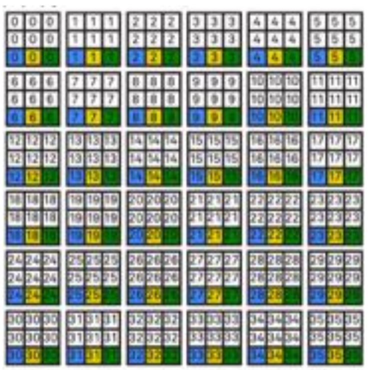
Light Field Camera – Concept of Sub-Aperture Images

Simplified model (Constant)



The same color means they come from the same perspective.

Light Field Camera – Concept of Sub-Aperture Images

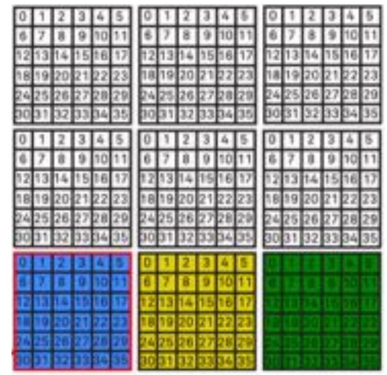


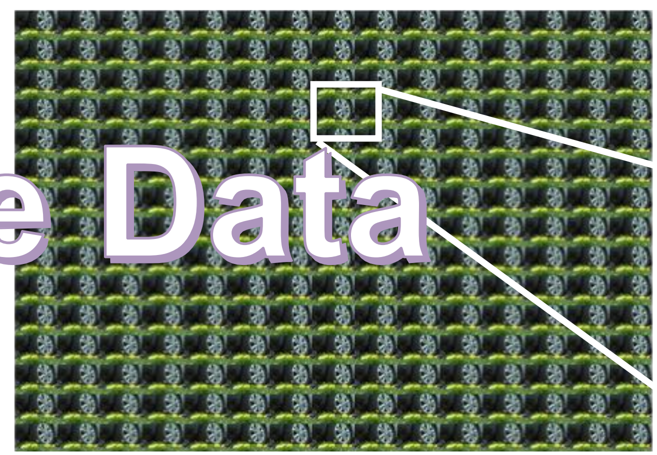
6510 (15x434)

9375 (15x625)

↑ RAW light field image captured by sensors

Massive Data





6510

9375

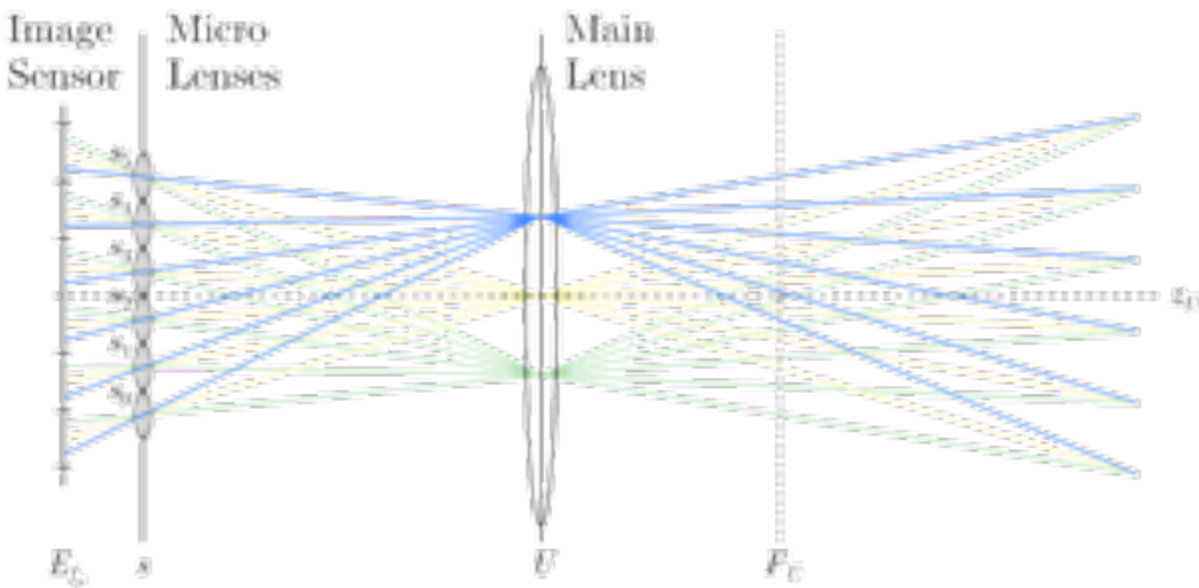
↑ Processed sub-aperture images

Single sub-aperture image

434 px

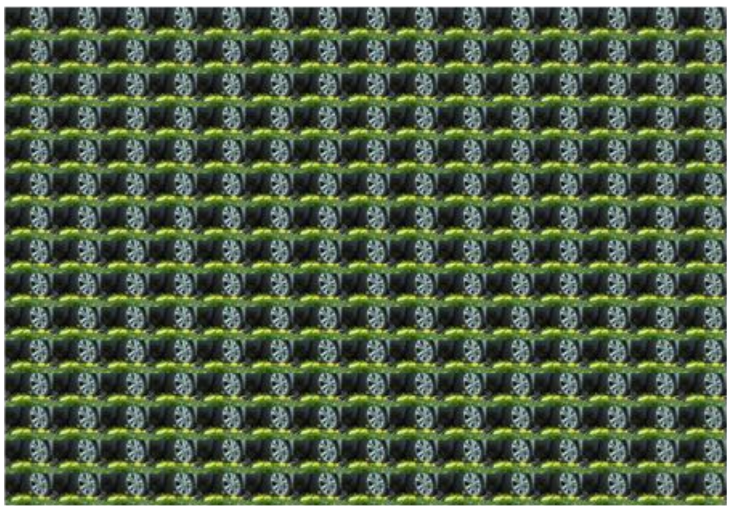
625 px

Light Field Camera – Perspective Feature



Viewpoint
Image

Light Field Camera – Perspective Feature



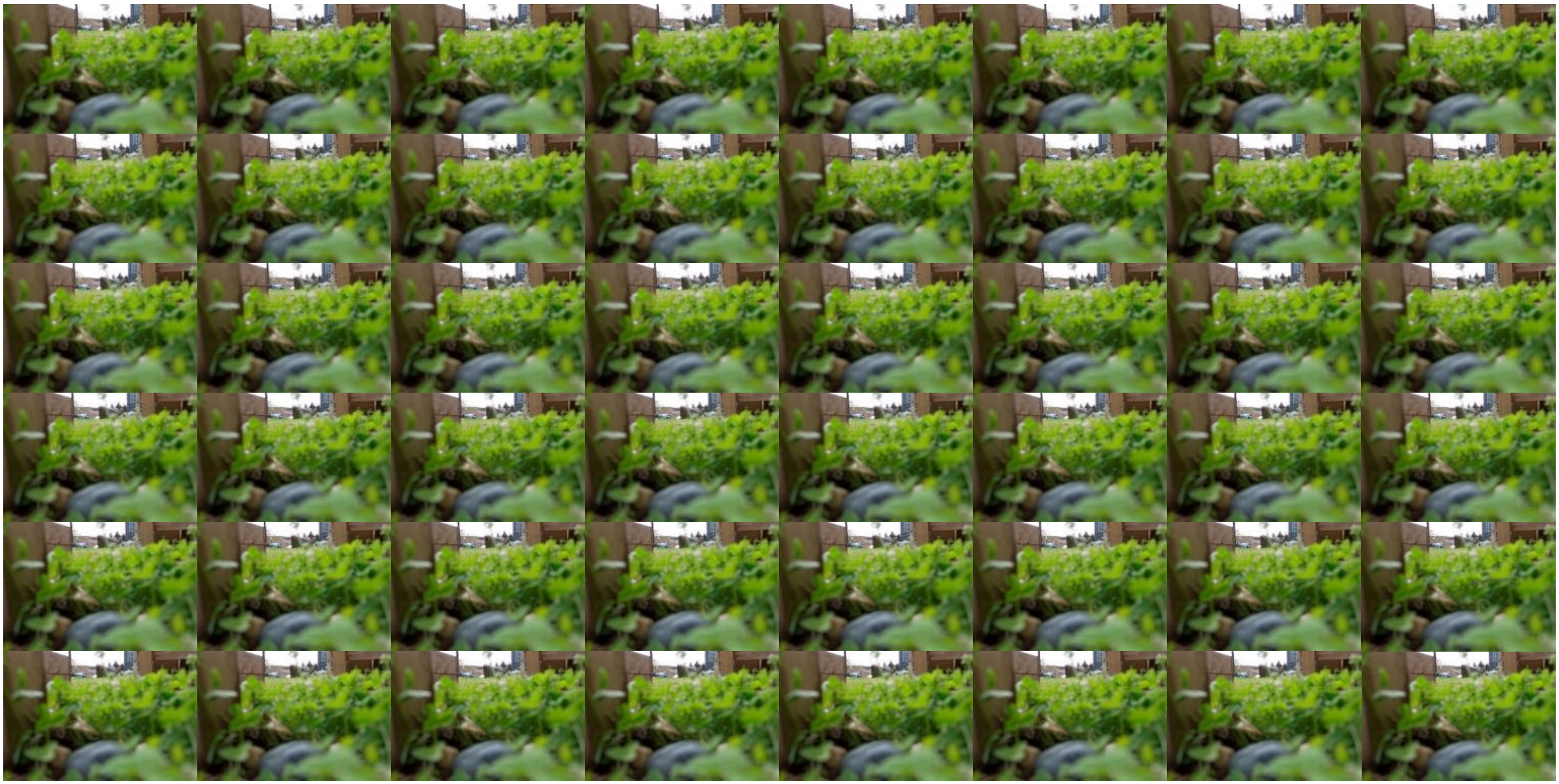
15X15 different perspectives

Freely changing perspectives

Light Field Camera – Perspective Feature

Light Field Frame

Sub-Aperture Images



More details revealed

- occlusion removal
- depth plane extraction

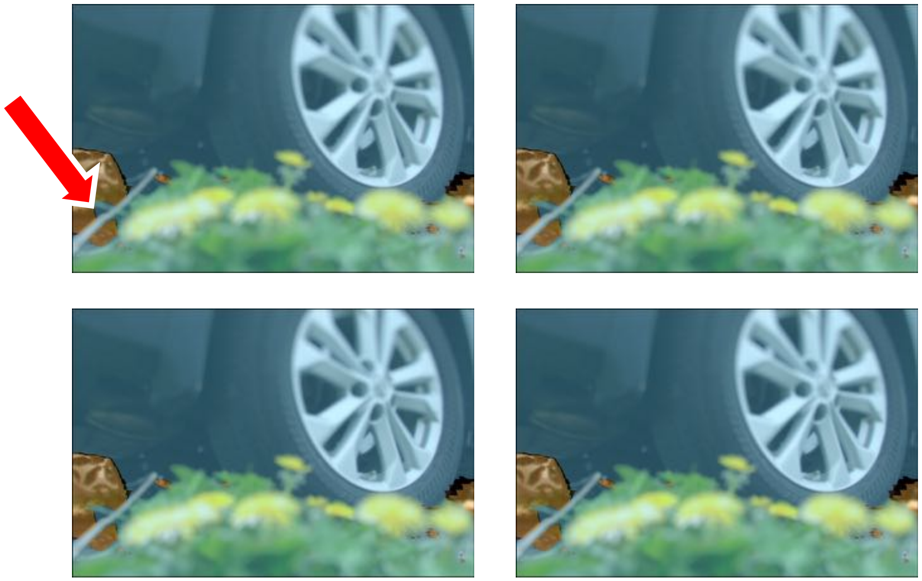
Fram

ook

Better Object Detection

Refocusing: Shift & Sum

Identify the depth layer in all the views



…

…

…

…

Refocusing: Shift & Sum

Shift each view to match the base view

Sum all the views to create the refocused view

Base view

Focused on the “Tire”

Focused on the “Small Flower”

Focused on the “Flowers”

Focused on the “Leafs”

Focused on the “other tire”

Microlens depth limitation

Current Light Field Capture Systems

Camera Array

© Stanford

Raytrix

Microlens

Lytro

REFOCUSING – DOES NOT WORK FOR CAMERA ARRAY CONTENT

REFOCUSING – DOES NOT WORK FOR CAMERA ARRAY CONTENT

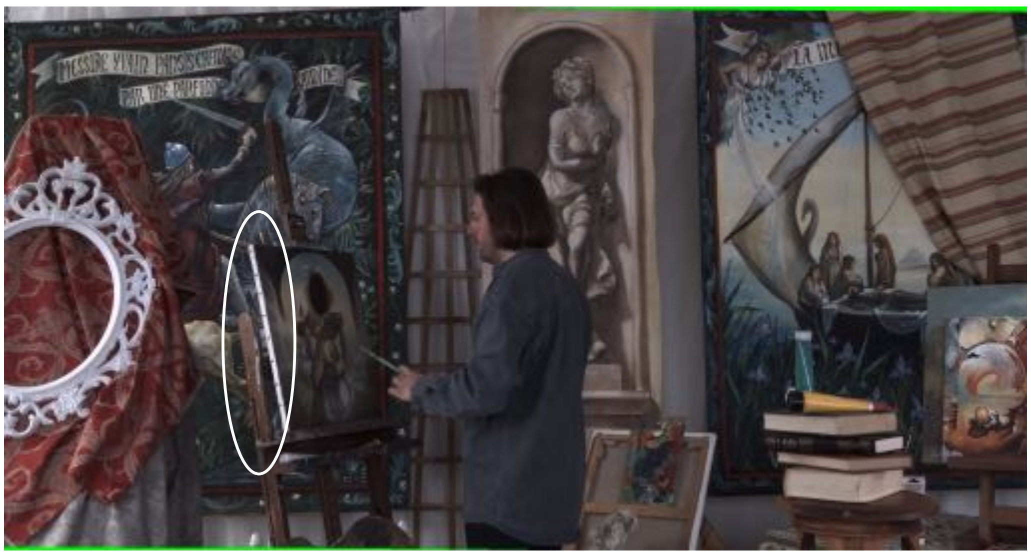
Original

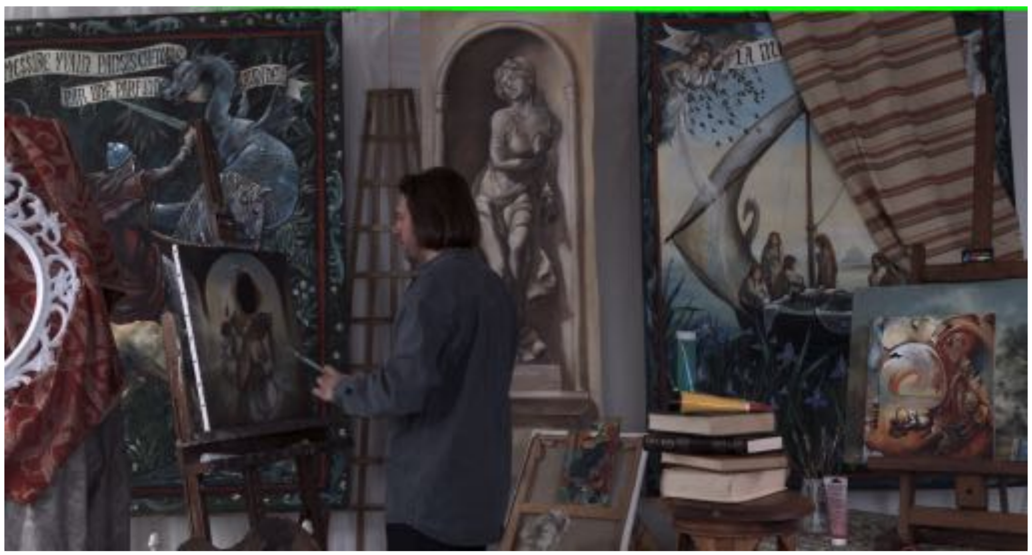
Refocused

REFOCUSING – DOES NOT WORK FOR CAMERA ARRAY CONTENT

Due to the longer baseline

Shifting only aligns to the sharpest part





- Occluded objects/ new details creates ghosting

REFOCUSING & PERSPECTIVE VIEWING IN ENTERTAINMENT

Applications:

- Live sports
- Interactive Movies
- Games

Pros:

- Every Sub-aperture image is captured from slightly different vantage point
- Better depth estimation than 3D
- Many focus points
- Better Object separation/recognition

ENTERTAINMENT – LIGHT FIELD

LIGHT FIELD IN AUGMENTED REALITY

Changing the AR/VR Landscape

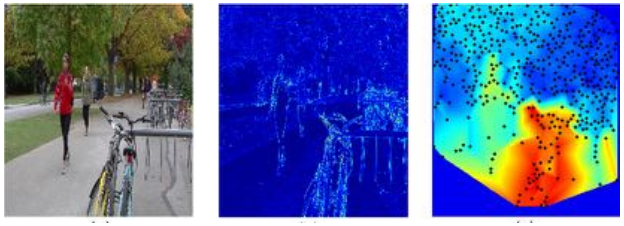
- Accurate Object Recognition
- Accurate Depth Estimation
- Accurate Overlay

Magic Leap has raised more than $2.7 billion in funding in 2 years
(AT&T is one of the investors)

LIGHT FIELD IN AUGMENTED REALITY

- Existing AR: Challenging precise overlay of synthetic augmentations on real-life content

Segmentation, object identification, depth estimation have been extremely difficult challenges for real-time applications



LIGHT FIELD IN AUGMENTED REALITY

Combine eye trackers with light field AR

Eye trackers can check where the viewer looks (heat map)

Visual information from light field camera is analyzed to identify area

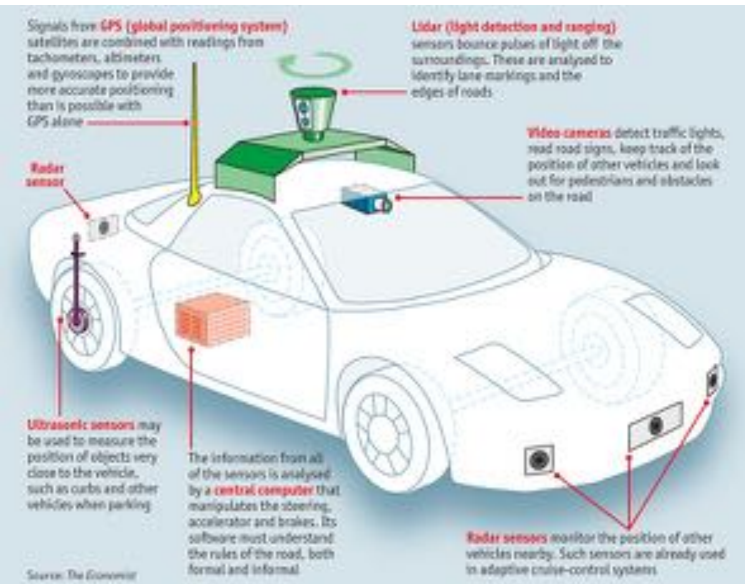
Light field properties are used to identify object of interest and focus only on that plane and location

LIGHT FIELD IN AUGMENTED REALITY

LIGHT FIELD IN AUGMENTED REALITY

LIGHT FIELD IN AUTONOMOUS DRIVING

Understanding Environment and Actions is the basis for success



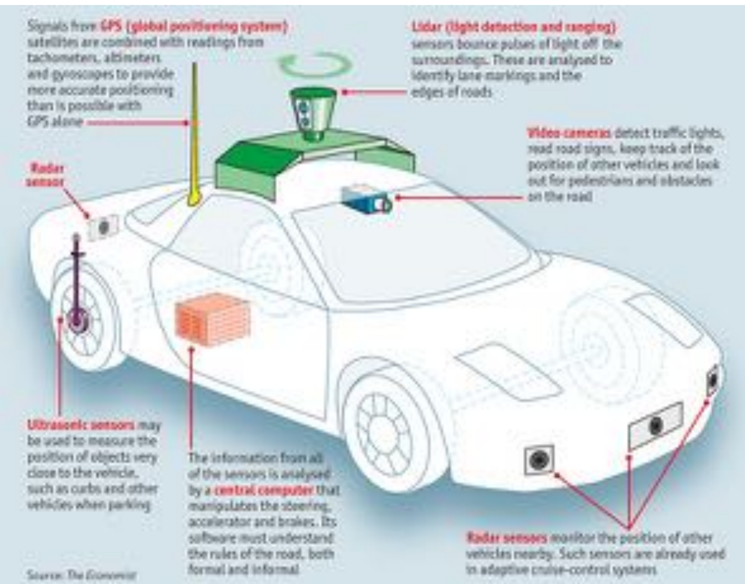
Cameras
- Detect traffic lights, read road signs, keep track of the position of other vehicles, look for pedestrians and obstacles.

Lidar
- Bounces pulses of light off the surroundings. They are analyses to identify lane markings and the edges of roads

Radar sensors
- Monitor position of other vehicles nearby. Already used in adaptive cruise control systems.

LIGHT FIELD IN AUTONOMOUS DRIVING

Understanding Environment and Actions is the basis for success

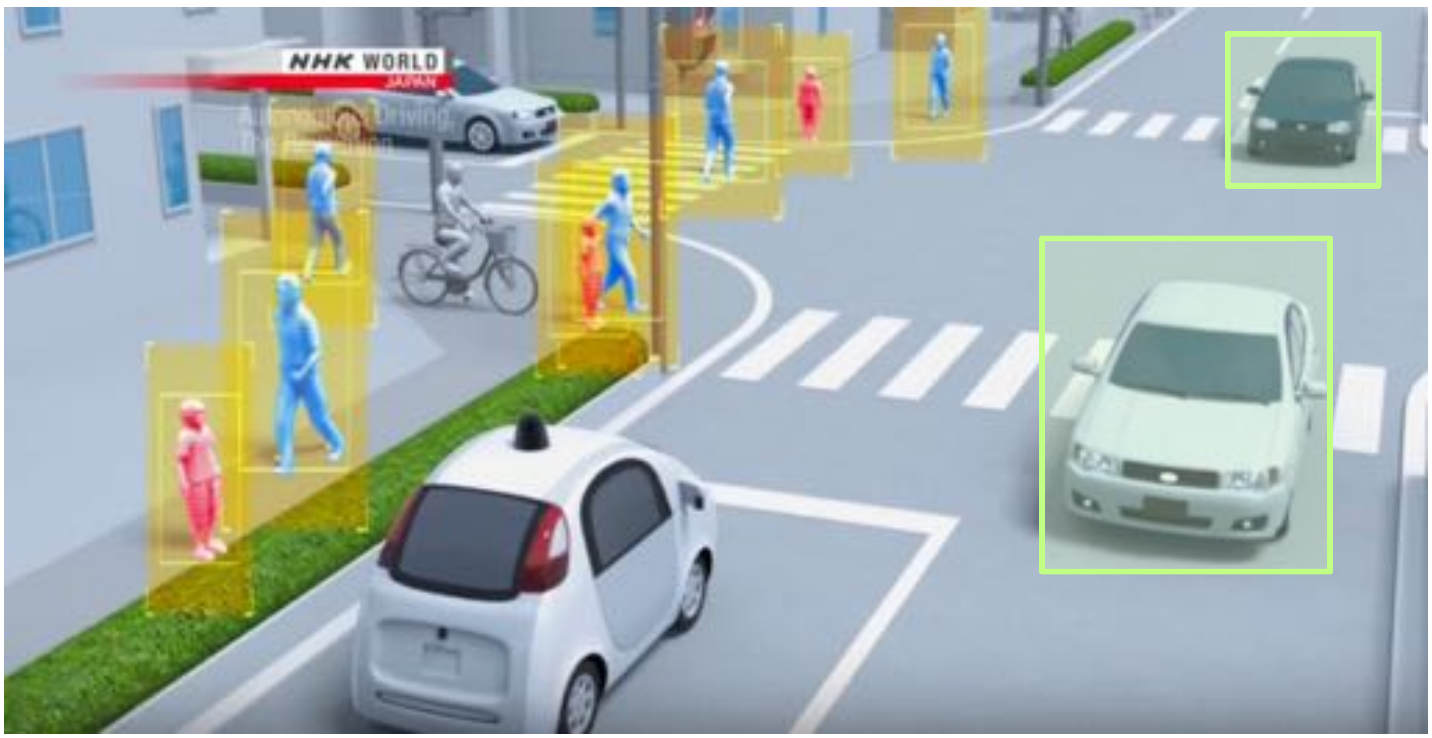


For now, autonomous vehicles/driving are limited to some places, isolated streets, specific controlled environment (depth information may suffice)

Why? Limited Visual Information

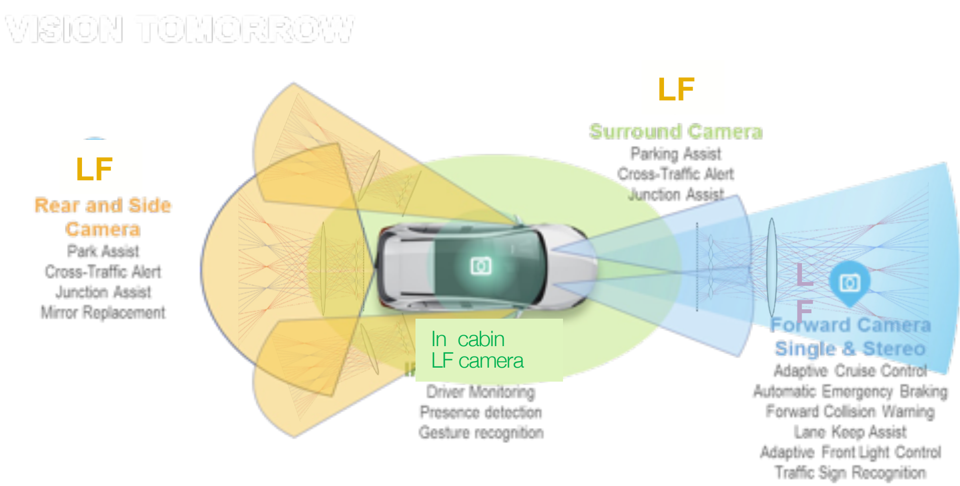
LIGHT FIELD IN AUTONOMOUS DRIVING

But the long term goal, which is 5 years and beyond, will expand these to a much broader application which will need our visual solutions to step up…



LIGHT FIELD IN AUTONOMOUS DRIVING

But the long term goal, which is 5 years and beyond, will expand these to a much broader application which will need our visual solutions to step up…

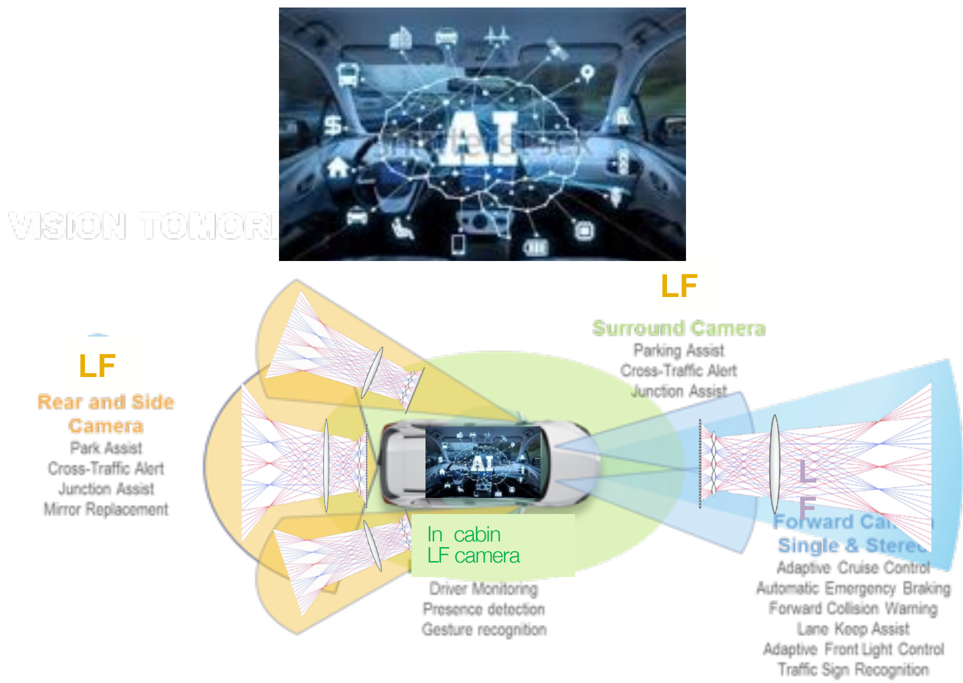


Capture

- light intensity and directional information,
- focus and depth properties, and
- visual cues.

LIGHT FIELD IN AUTONOMOUS DRIVING

The rich, immersive video information provided by Light Field will help AI "act" as close as possible to a human



Capture

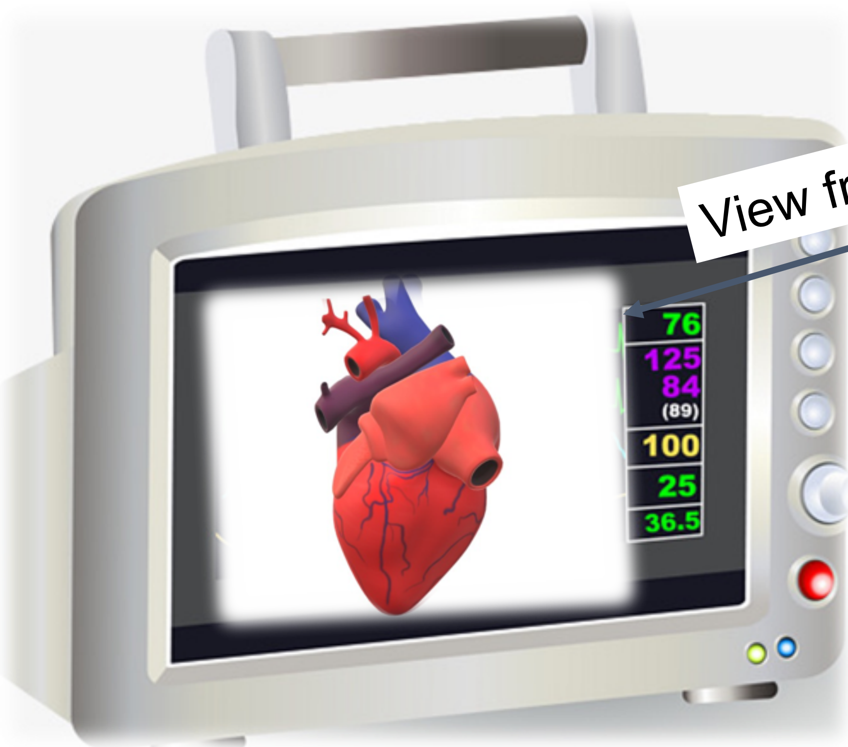
- light intensity and directional information,
- focus and depth properties, and
- visual cues.

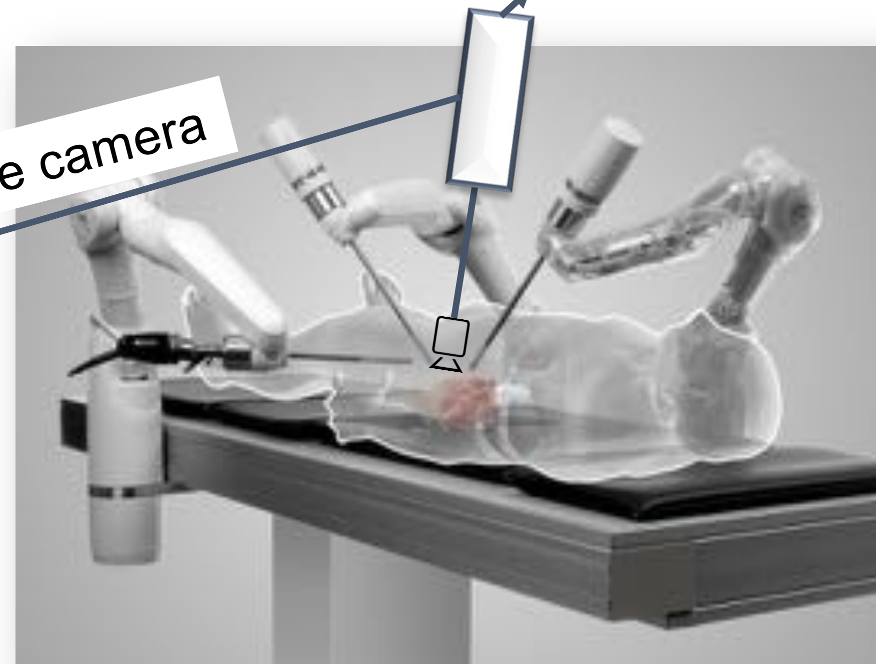
LIGHT FIELD IN DIGITAL HEALTH

Accurate Less Intrusive Endoscopy

A regular surgical camera

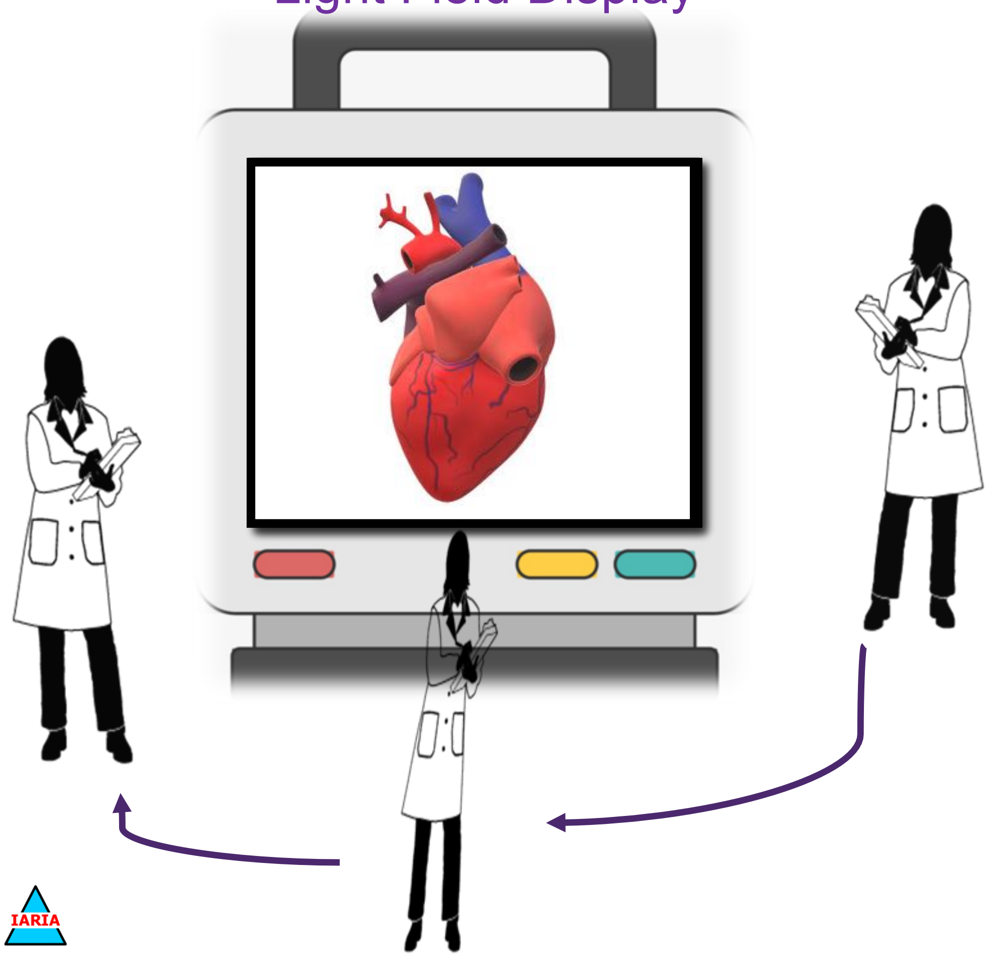
View from the camera



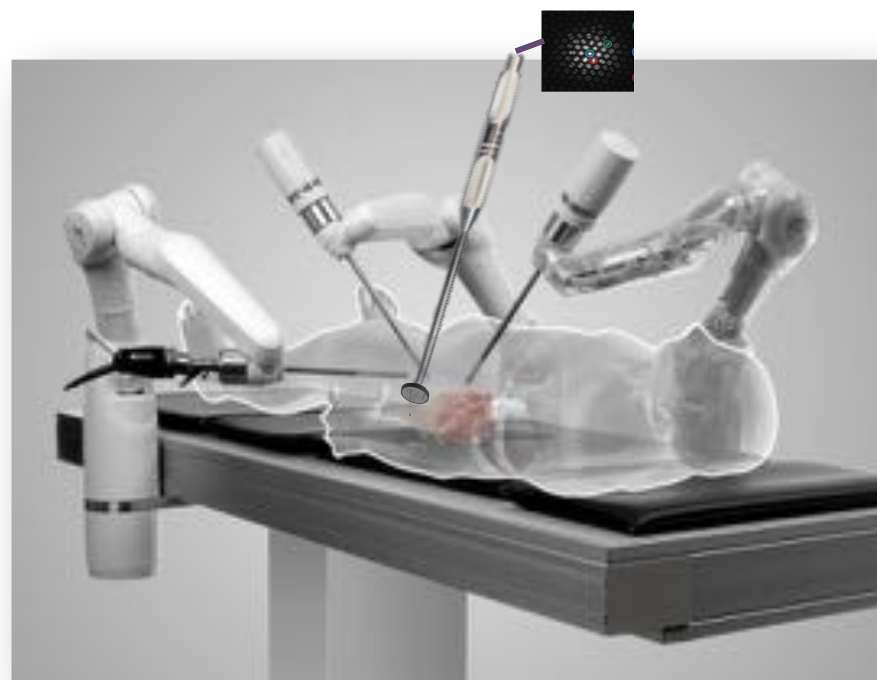


Accurate Less Intrusive Endoscopy

Light Field Display



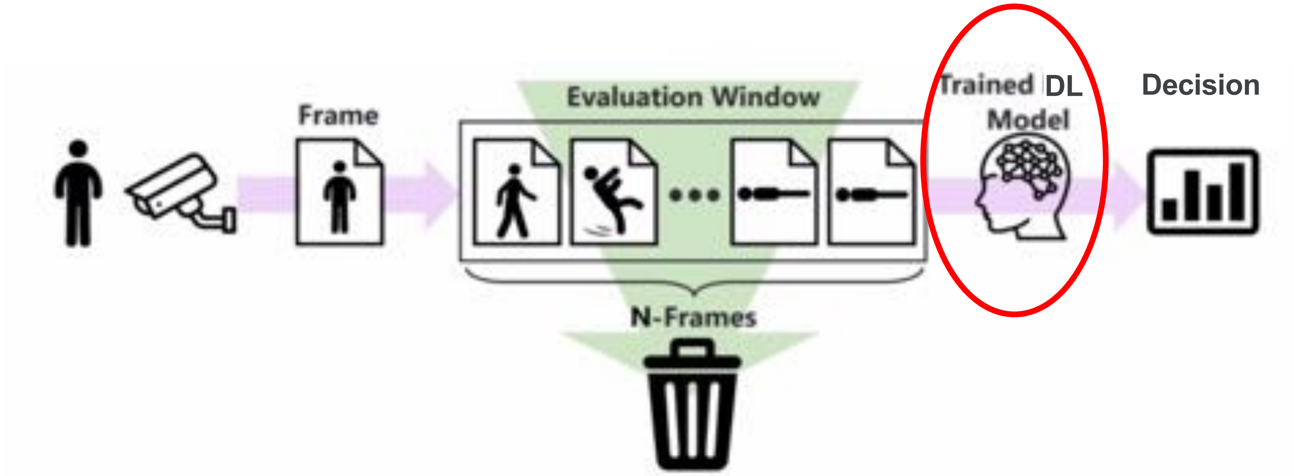
Light Field Camera



5G enables real-time remote surgery

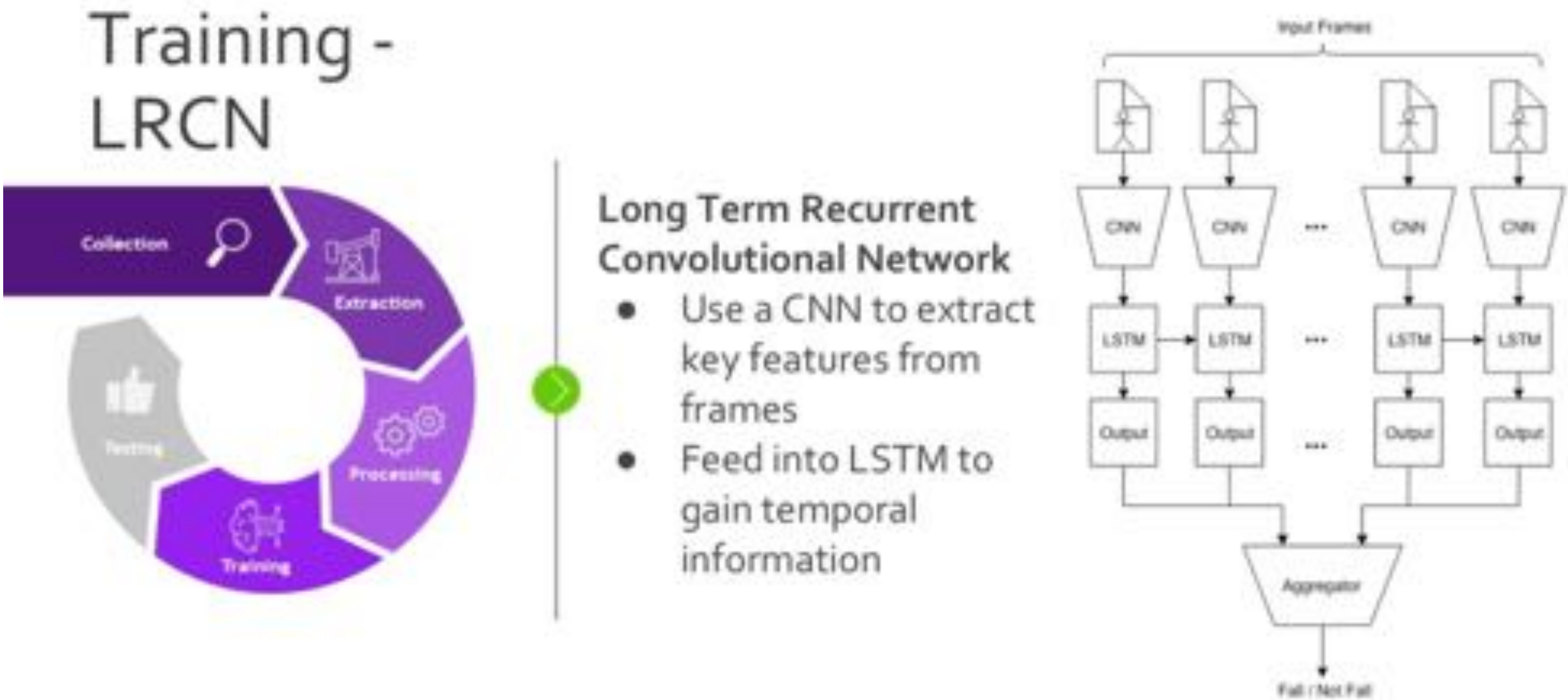
© TransEnterix

Non-Intrusive Monitoring



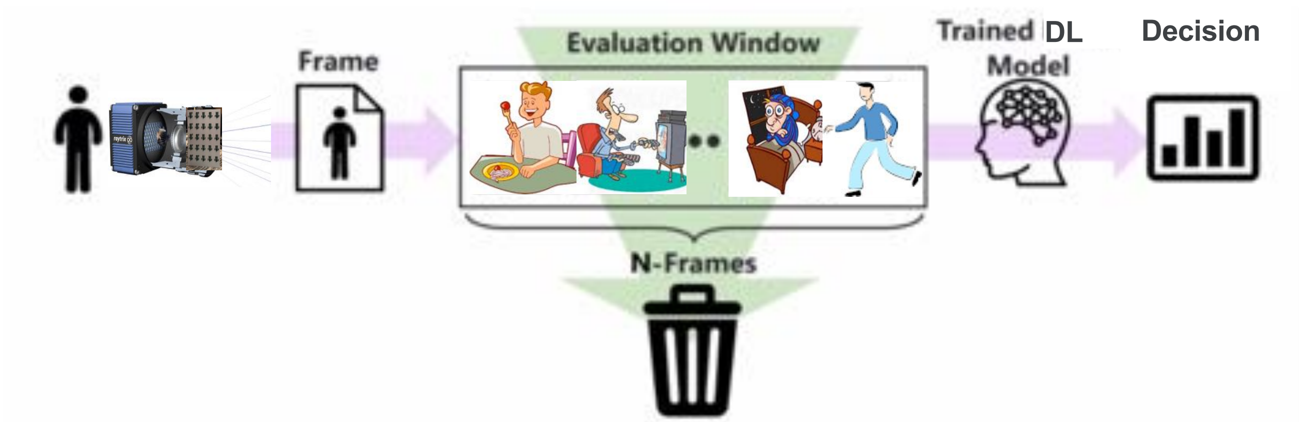
LIGHT FIELD IN DIGITAL HEALTH

LIGHT FIELD IN DIGITAL HEALTH



LIGHT FIELD IN DIGITAL HEALTH

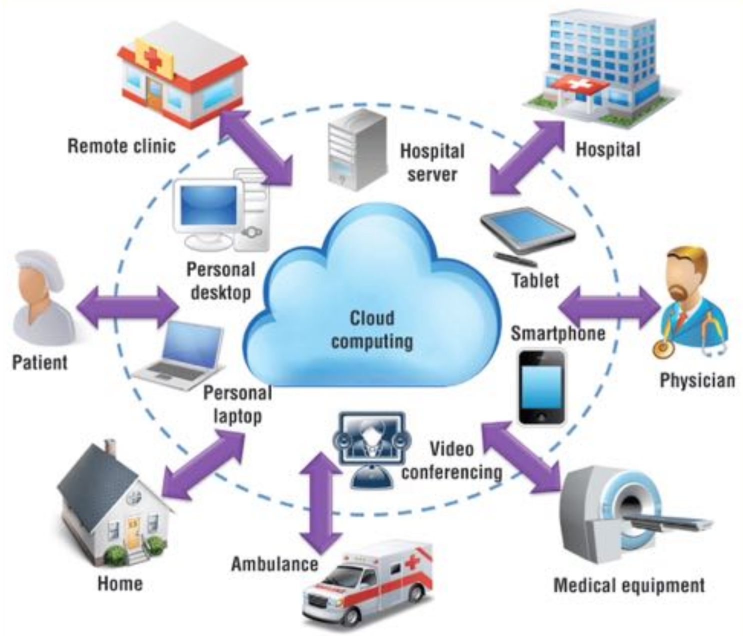
Non-Intrusive Monitoring

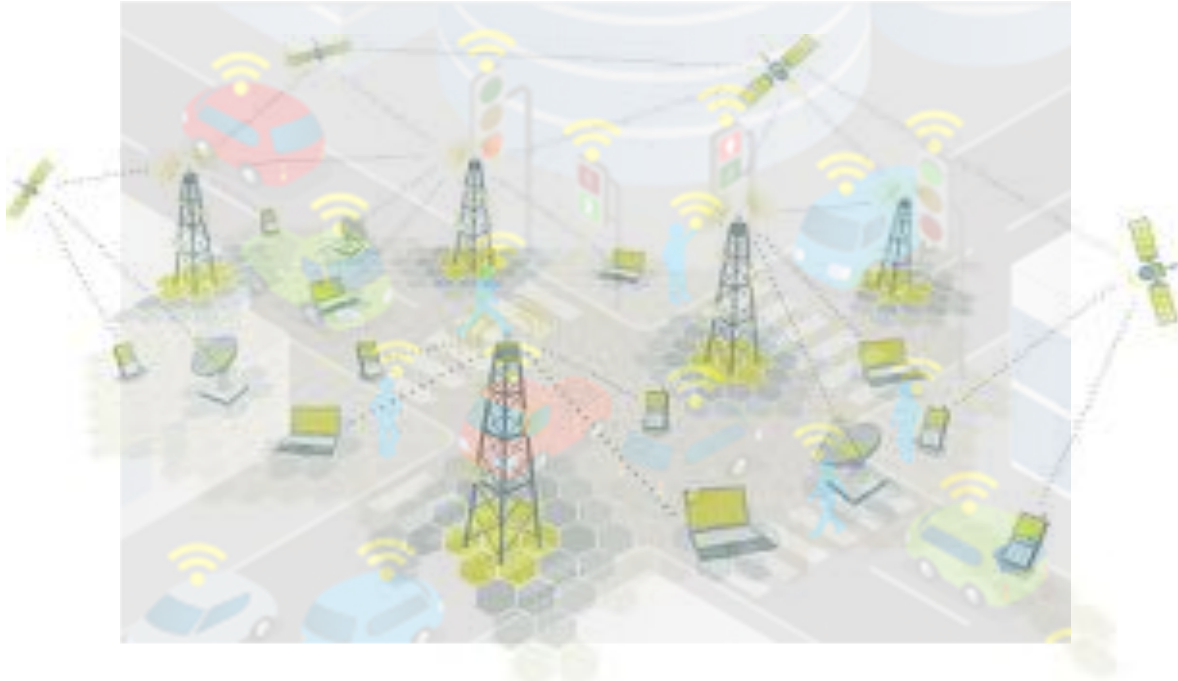


Light Field information will improve this accuracy, since richer visual information will be processed by the system.

This type of visual information will allow us to accurately track behaviour and determine changes that may be associated with various medical conditions, for early diagnosis and prevention.

DIGITAL HEALTH & COMMUNICATIONS- Challenges





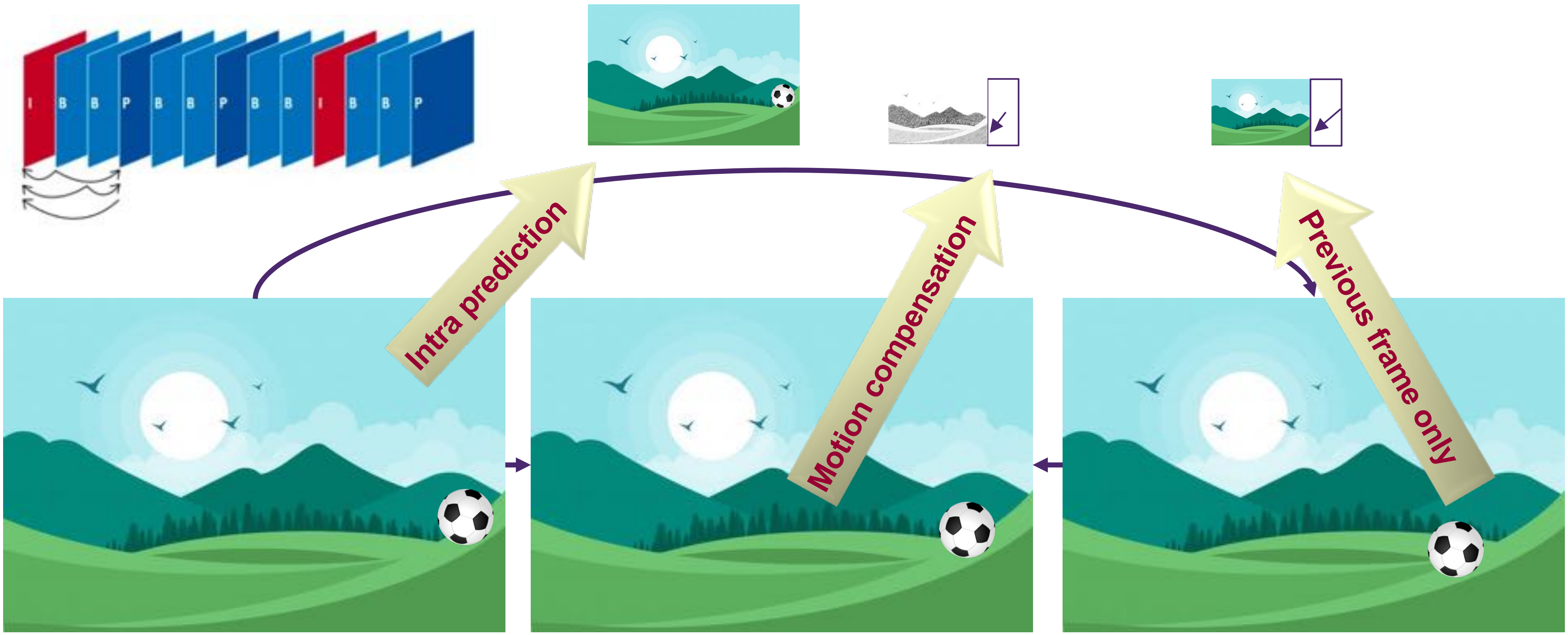
- Bandwidth
- Real-time streaming
- Security

Compression

Practical Light Field Camera arrays

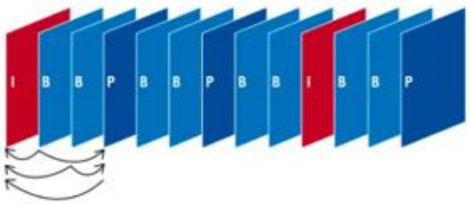
Compression IBP frames

Trade-off between overhead and residuals

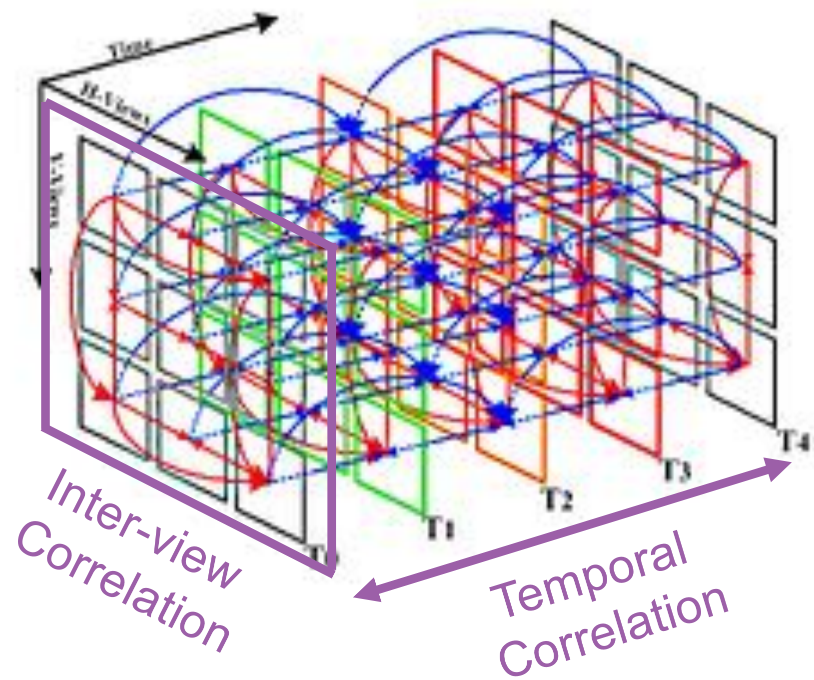


LIGHT FIELD – MASSIVE DATA: COMPRESSION

Conventional Video



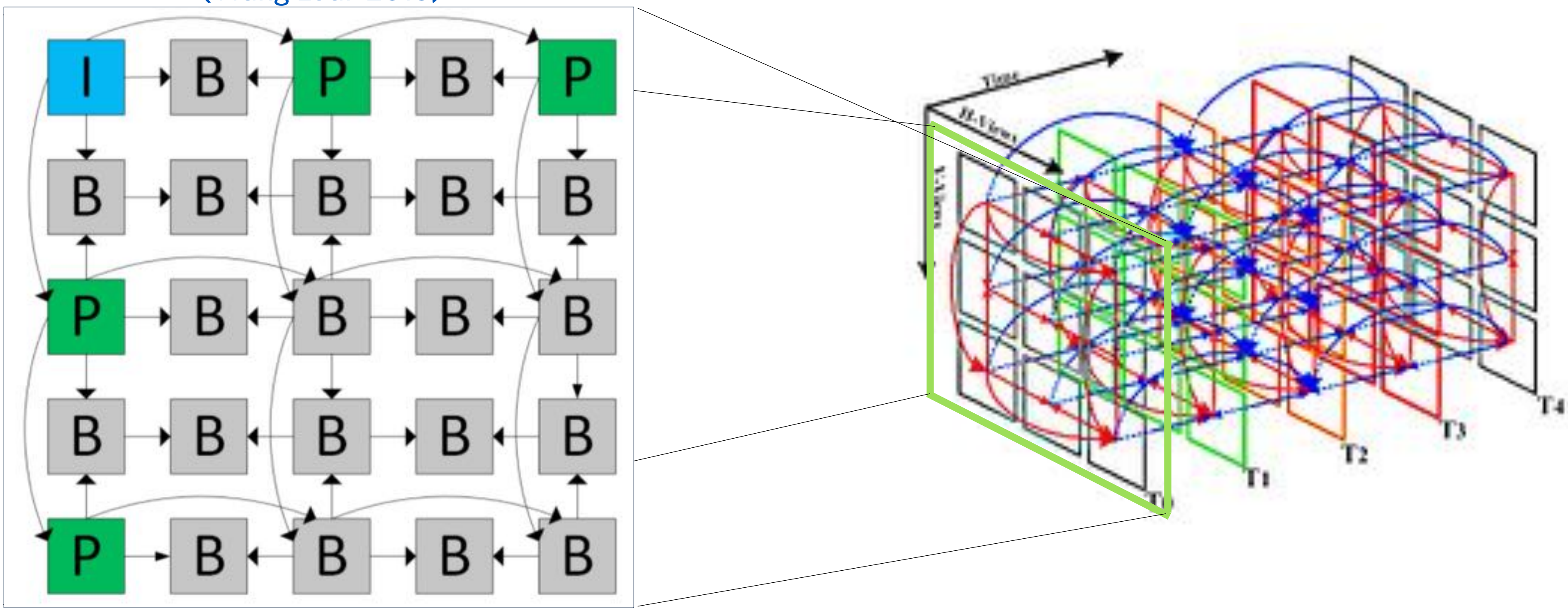
Light Field Video



LIGHT FIELD – Existing Compression Methods

LF-MVC

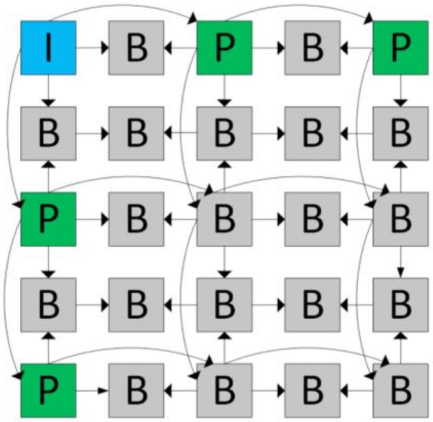
(Wang Et al. 2016)



LIGHT FIELD – Existing Compression Methods

LF-MVC

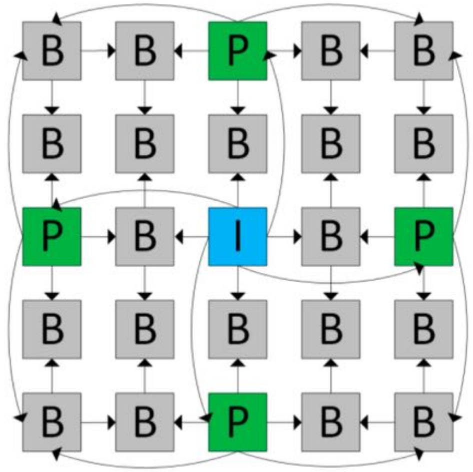
(Wang Et al. 2016)



Is Top Left the Best Place to get started?

Khoury's Method

(Khoury Et al. 2018)



Encode the central frame as I frame ✓

Again only vertical or horizontal references ×

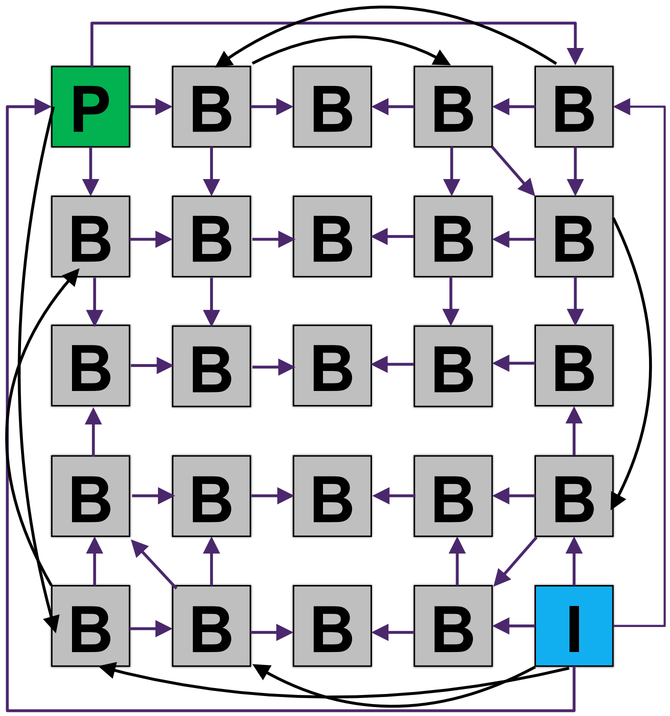
38.18% BD-Rate of LF-MVC (5 x 5 Views) ✓

Does not scale well ×

LIGHT FIELD – Existing Compression Methods

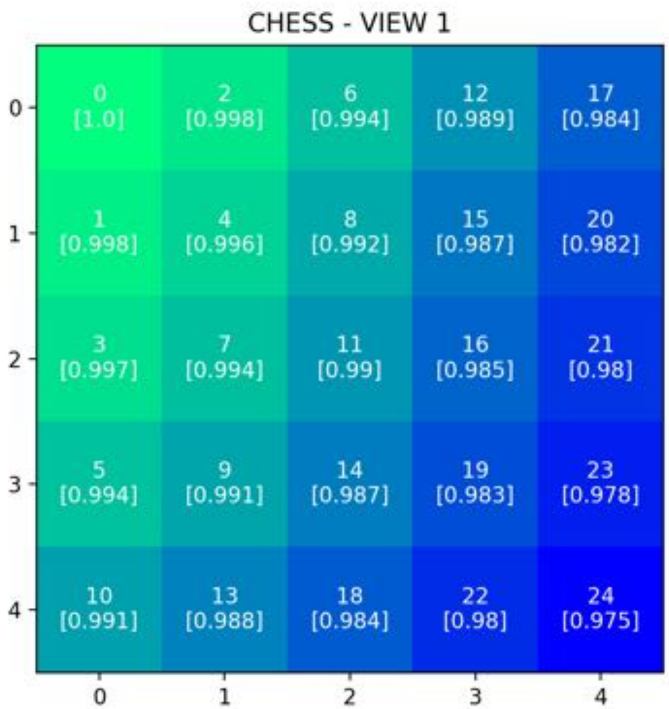
Full Scheme

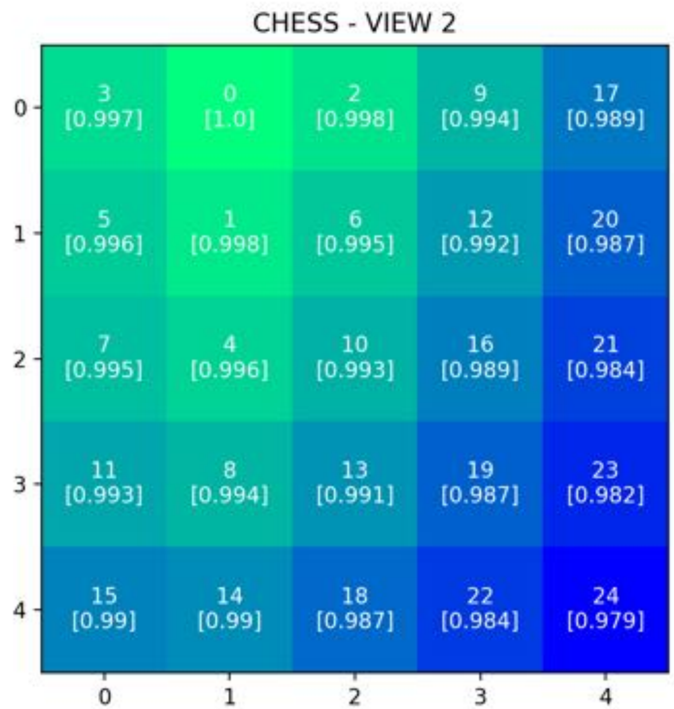
(Avramelos et al. 2019)

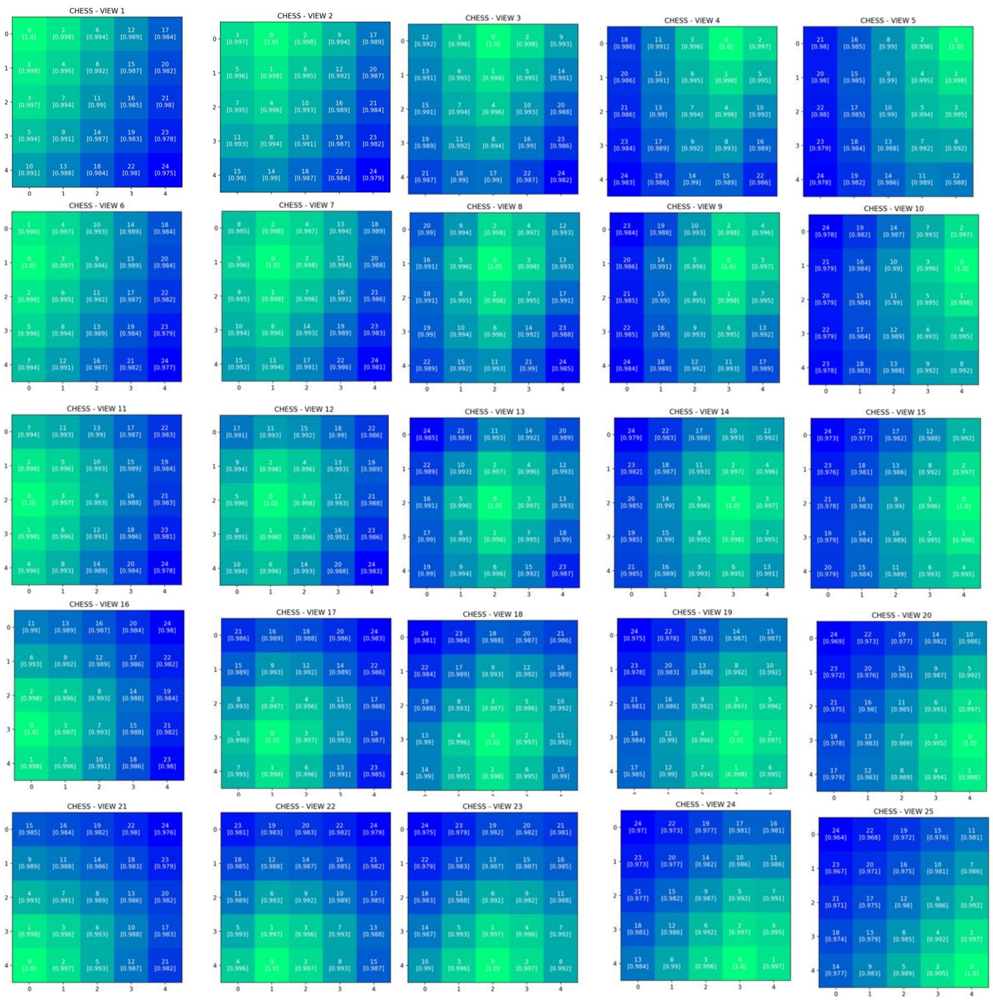


- Maximizes B frames so compression efficient ✓
- Again I frame at a corner ×
- Diagonal references ×
- Predicting frames from far ×
- Performs 24% better than LF-MVC and 15% worse than Khoury's method
- Does not scale well ×

SSIM heatmap for Chess Pieces (5x5)





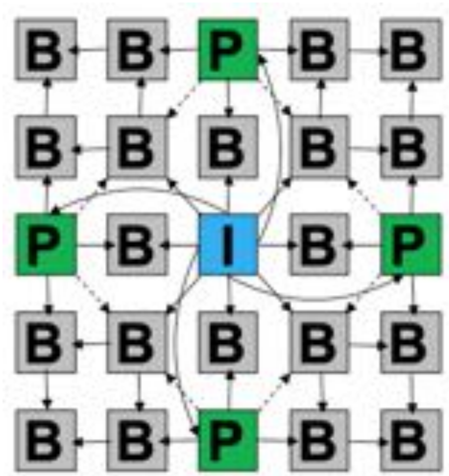


.982	**.985**	**.985**	**.984**	**.976**
.983	.985	.986	.985	.981
.984	.99	.991	.989	.985
.984	.987	.99	.986	.984
.98	.981	.984	.982	.984

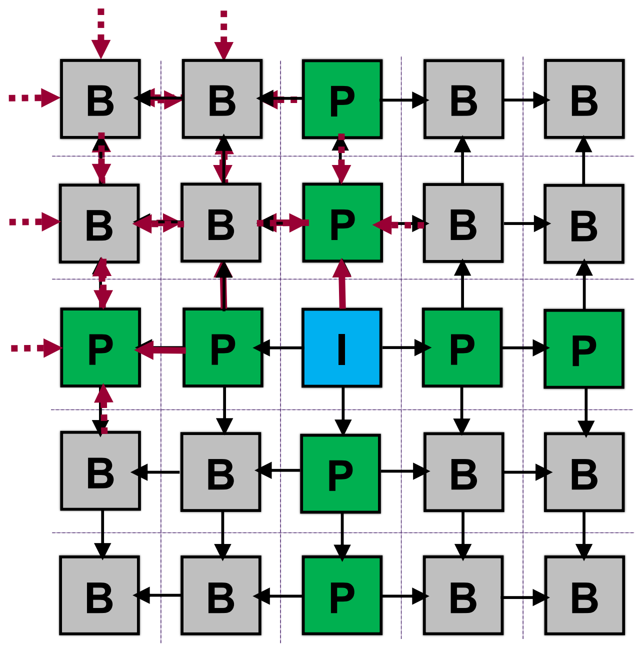
Avg of all views

Empirical results over 9 test sequences

SSIM based Compression Method



Universal Pseudo-sequence (UPS) SSIM based Compression Method



Unlike most of the PSB prediction structures and coding orders discussed above, the proposed Universal Pseudo-sequence (UPS) based structure takes full advantage of both horizontal and vertical correlation among the views

SSIM based Compression Method

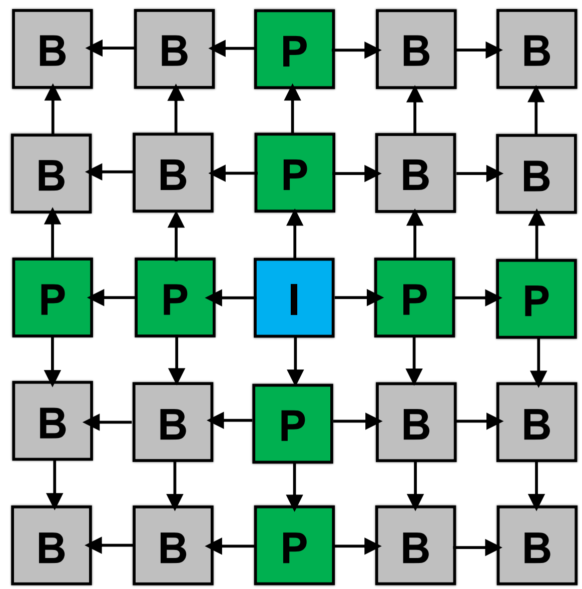
Has highest similarity with its neighbors

SSIM of all the views

Highest one is I-frame

SSIM calculation makes smallest possible P frames

To maintain our structure, next level of P frames are predicted



5x5 views

B frames with vertical and horizontal references

Leverage bidirectional correlation

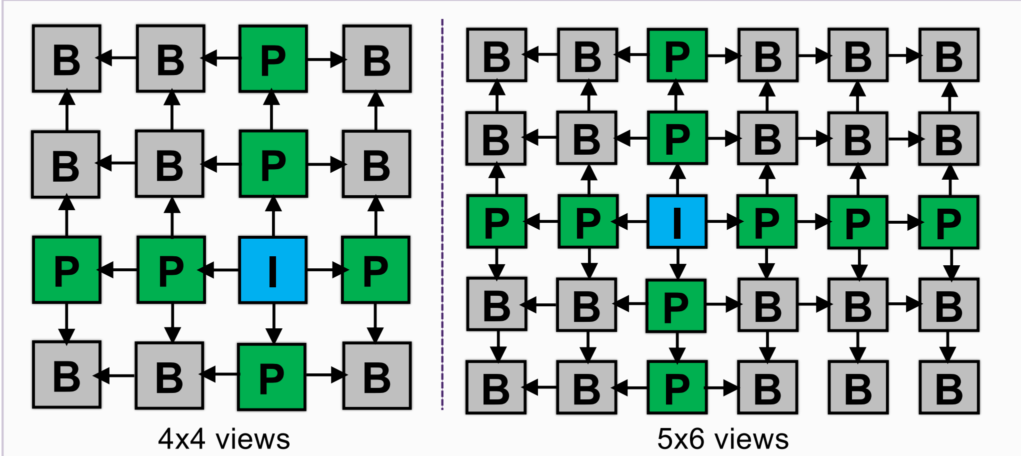
Scalable to number of views

Scalable to view arrangements

SSIM based Compression Method

Can be extended to any number of views

Feasible for most camera array structures

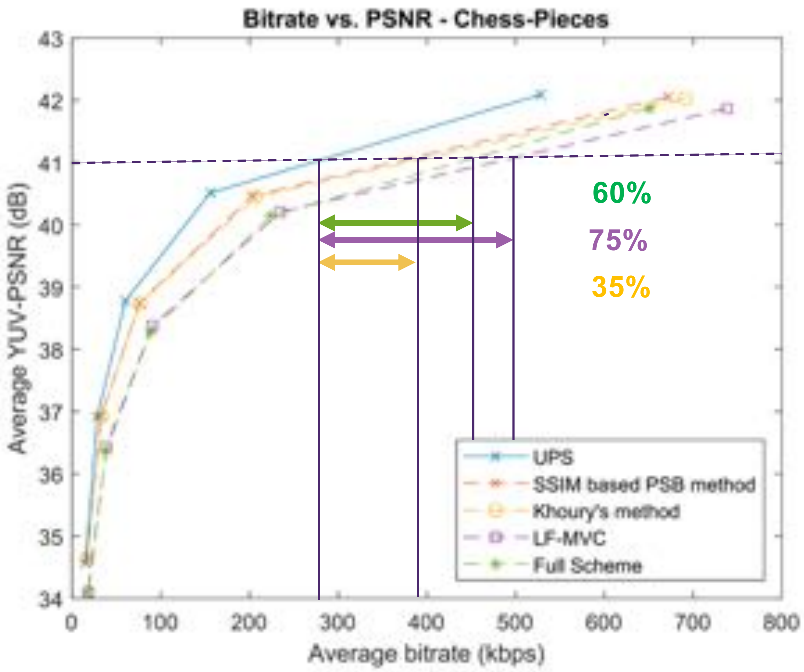


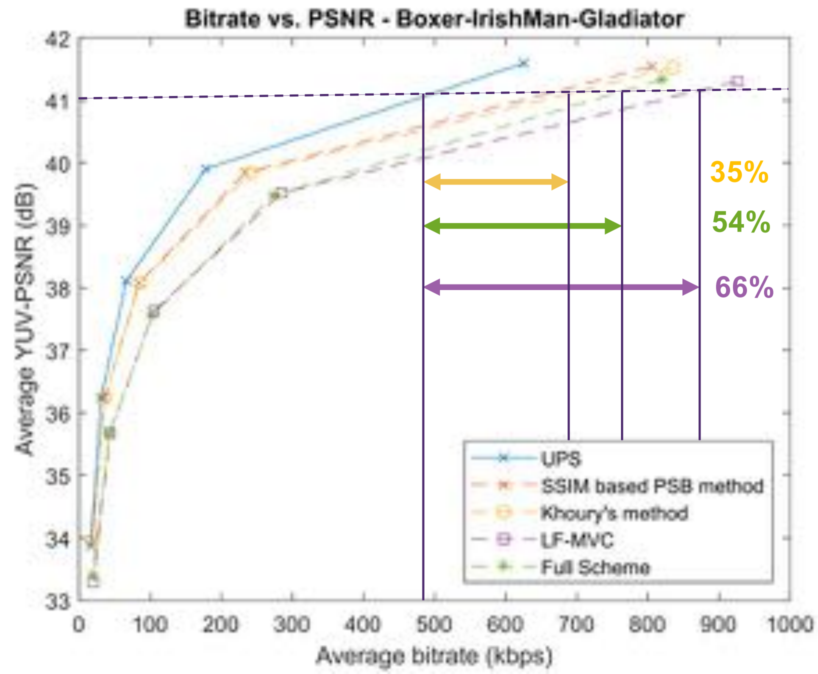
Circular

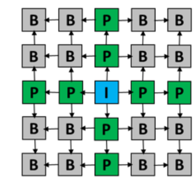
Hexagonal

Rectangular

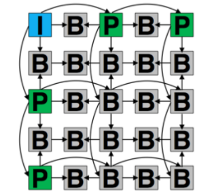
Performance – Microlens content (5x5)



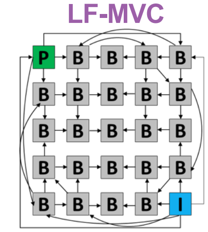




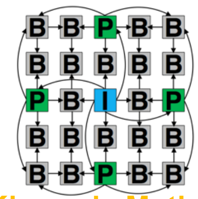
UPS method



LF-MVC



Full Scheme

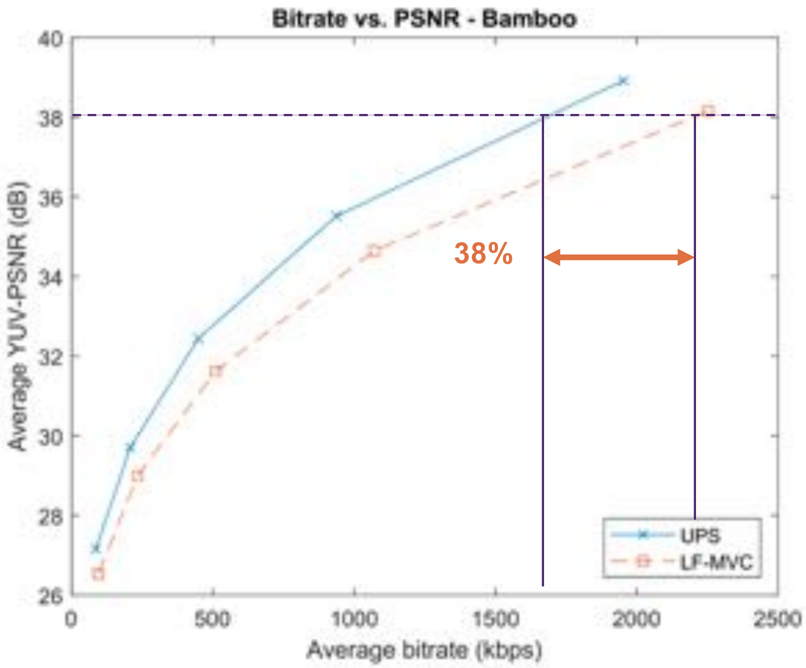


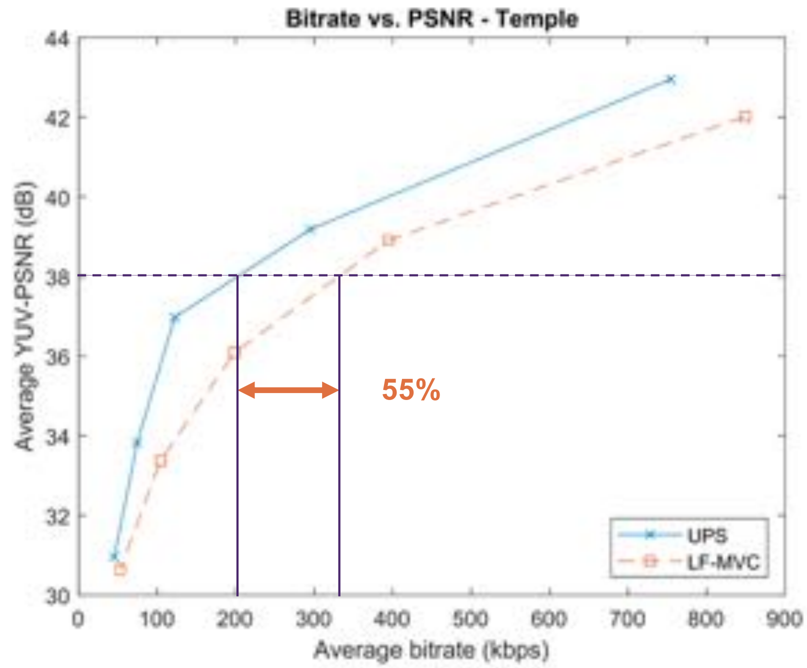
Khoury's Method

Raytrix R8

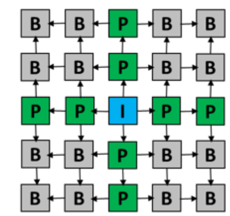
Resolution	**1920x1080**
FPS	**30; 300 frames**
Baseline	**4mm**

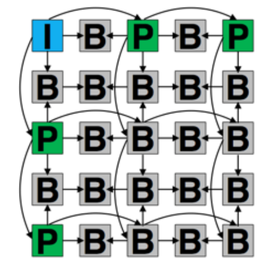
Performance – Animated content (3x3)



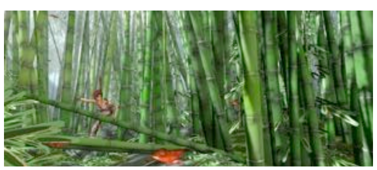


UPS method



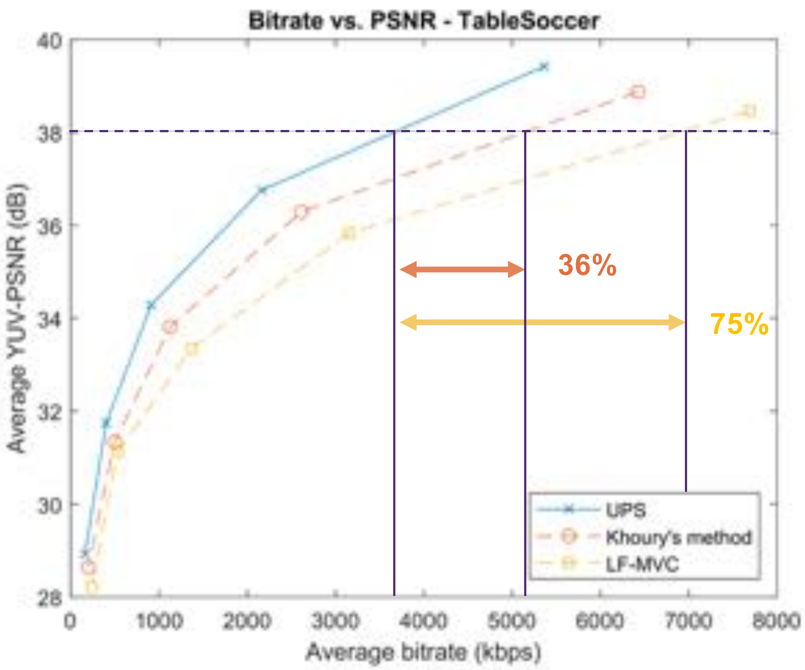


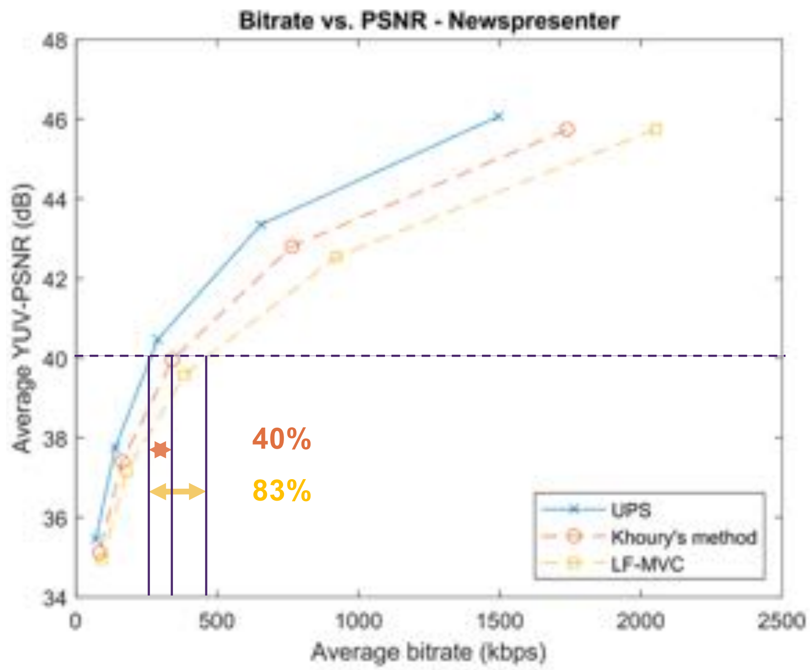
LF-MVC

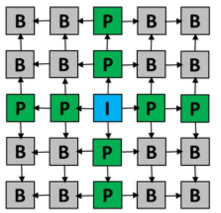


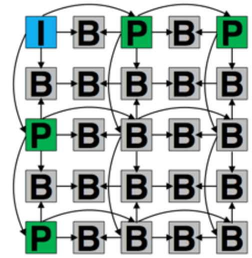
Resolution	1024x436
FPS	20; 100 frames

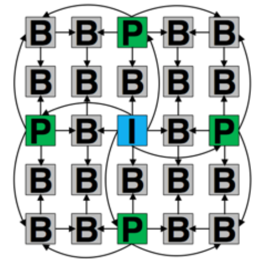
Performance – Camera Array content (3x5) and (4x4)

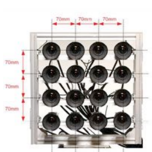


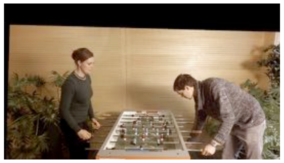










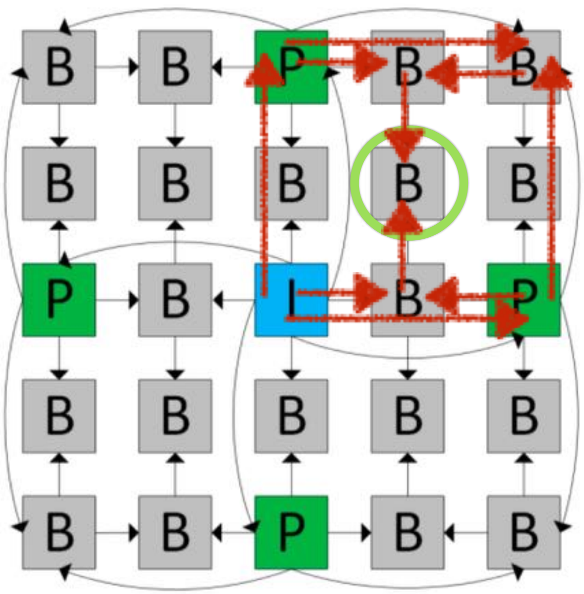


Resolution	1920x1080
FPS	30; 300 frames
Baseline	6 to 10 cm

Light Field Video Compression for Random-access efficiency

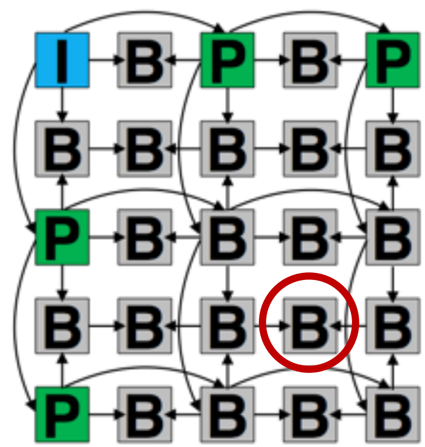
Light Field – Random Access matters

Khoury's Method



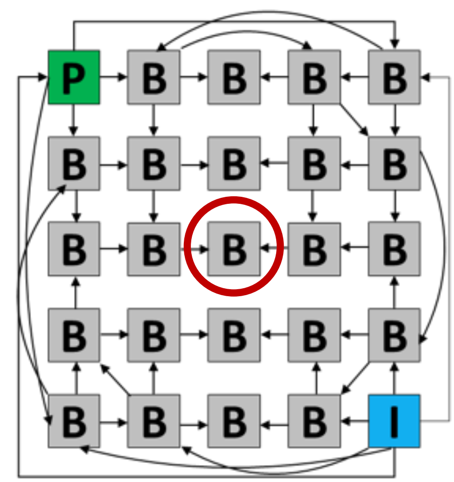
Worst Frame to Decode: 6

LF-MVC



Worst Frame to Decode: 11

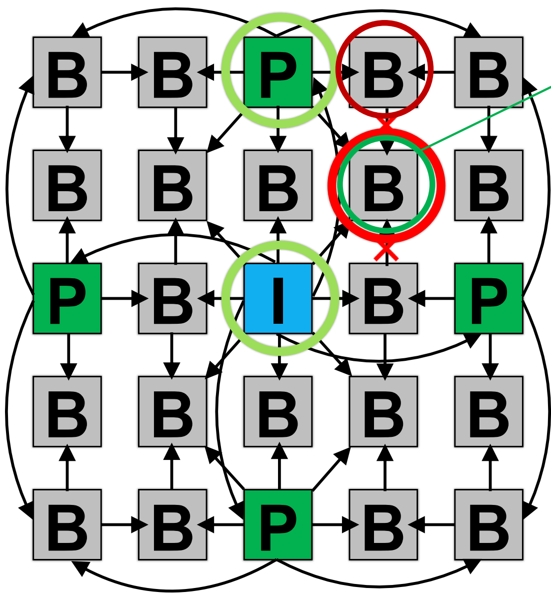
Full Scheme



Worst Frame to Decode: 18

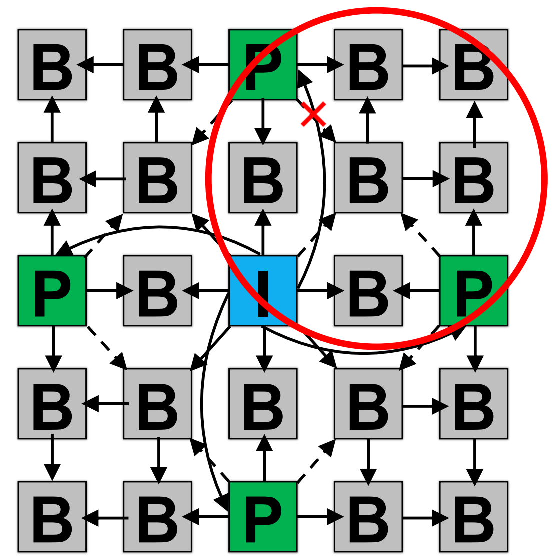
Proposed Methods

Reduced worst case random-access from 6 to 4



Random-access reduces to **2** from **6**

Diagonal Reference Based Prediction structure



SSIM Assisted Diagonal Reference Based Prediction structure

Highest SSIM

Increases compression efficiency by **15%**

Average random-access increases from **2.56 to 2.72**

Random Access performance

Method	Random-access Complexity	
	Average	Worst-case
LF-MVC	4.92	11
Khoury's	3.2	6
Full scheme	6.08	17
Diagonal reference	**2.56**	**4**
SSIM assisted RAE	**2.72**	**6**
UPS	**3.84**	**8**

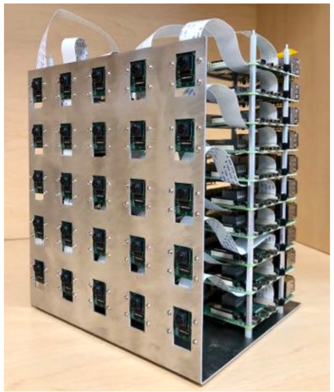
Best Random-access

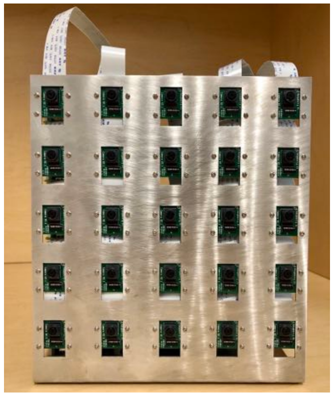
Best Random-access to compression trade-off

N. Mehajabin, S. R. Luo, H. Wei Yu, J. Khoury, J. Kaur, and M. T. Pourazad, "An Efficient Random Access Light Field Video Compression Utilizing Diagonal Inter-View Prediction," in 2019 IEEE International Conference on Image Processing (ICIP), 2019; N. Mehajabin, M. Pourazad, and S. Member, "SSIM Assisted Pseudo-sequence-based Prediction Structure for Light Field Video Compression," in IEEE International Conference on Consumer Electronics, 2020

Refocusing

Performance – Animated content

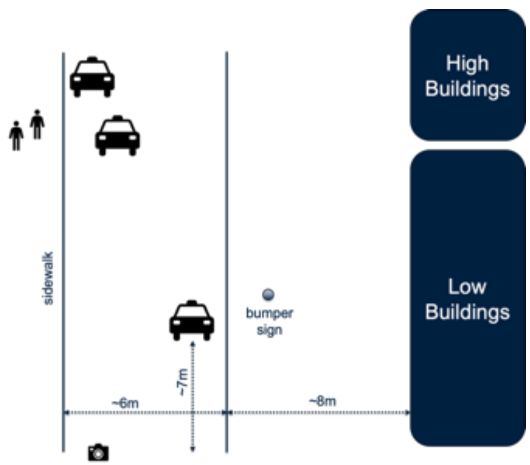




The dense camera array consists of 25 camera modules fixed on a stainless steel stand

Our Light Field Camera array – Outdoors content

- The video content is captured on the University campus, with the scene including walking people, bicycles, vehicles and buildings as background.

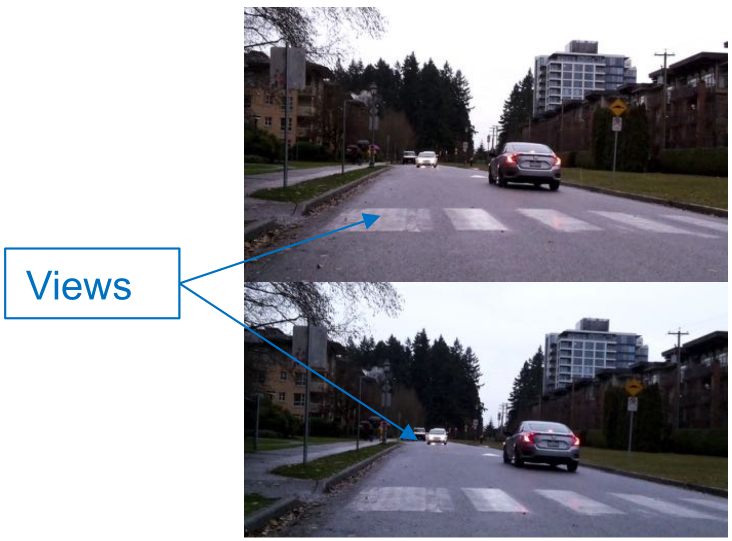


Sketch of the captured scene

Snapshot of the captured scene

Our Light Field Camera – Outdoors content

The frames from each camera are not perfectly aligned.



Our Light Field Camera array – Outdoors content

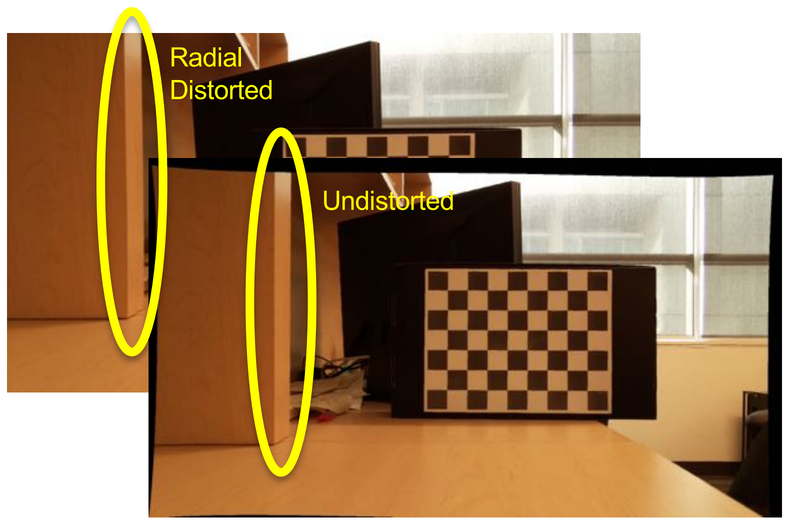
Color differences exist in the frames captured by different camera.

Camera1

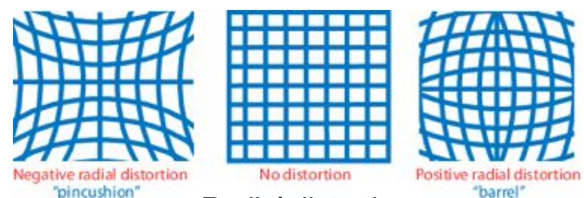
Camera2

Our Light Field Camera array – Outdoors content

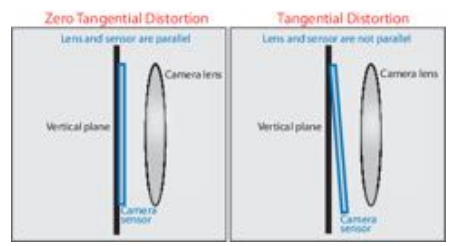
Different types of geometry distortion for single camera and the camera array.



Single camera distortion



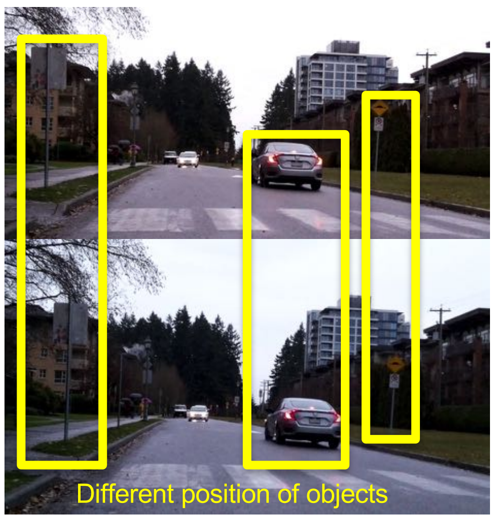
Radial distortion



Tangential distortion

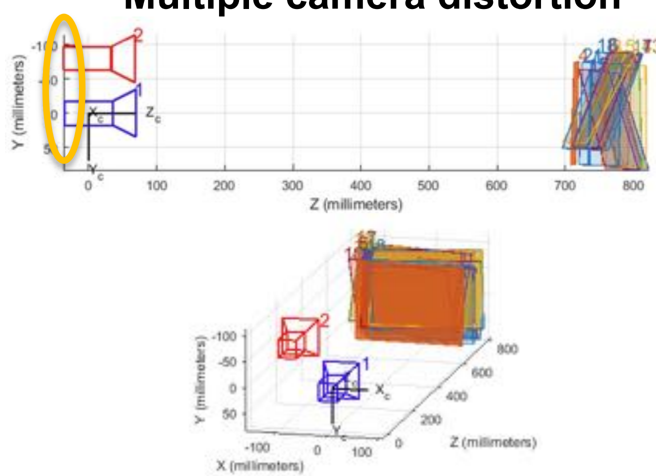
Problems - Geometry Distortion (2)

Different types of geometry distortion for single camera and the camera array.



This type of distortion is described by **extrinsic** parameters.

Multiple camera distortion



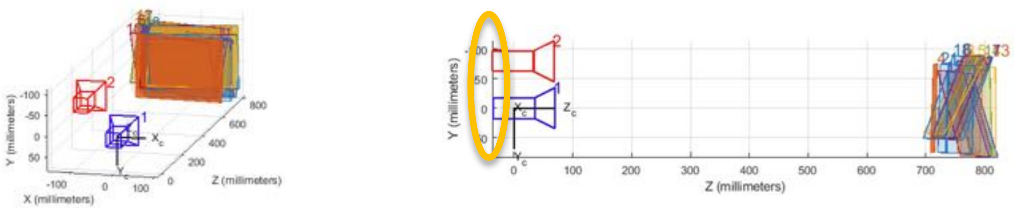
Rotation and translation between each two camera pairs.

Our Light Field Camera array – Outdoors content

Different types of geometry distortion for single camera and the camera array.

Images captured by 5 cameras in the **SAME** row. Lens distortion is already fixed.

Multiple camera distortion



Rotation and translation between cameras

This type of distortion is described by **extrinsic** parameters.

Performance – Visualization

Visual comparison before and after calibration and correction

(a) Uncorrected, view 19

(b) Uncorrected, view 20

(e) Corrected, view 19

(f) Corrected, view 20

(c) Uncorrected, view 24

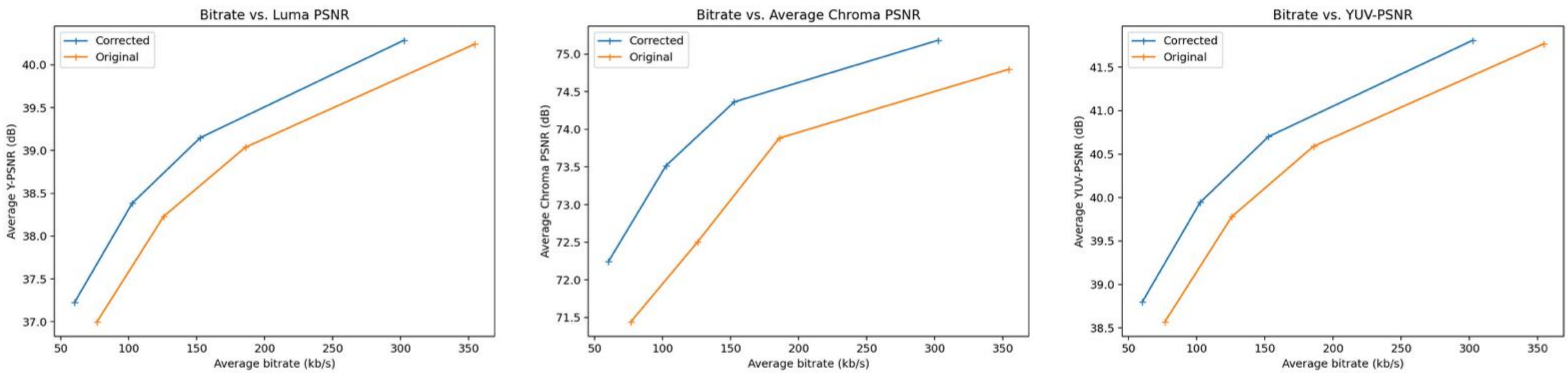
(d) Uncorrected, view 25

(g) Corrected, view 24

(h) Corrected, view 25

Performance – Quantitative Comparison

Quantitative comparison before and after calibration/correction



(a) Luma PSNR (b) Chroma PSNR (c) YUV-PSNR

The bitrate vs. PSNR at four different QP levels (25, 28, 30, and 33)

On average, for the same bitrate the objective visual quality (PSNR) increases by **0.519dB**, while for the same quality the bitrate savings are **30.275%**.

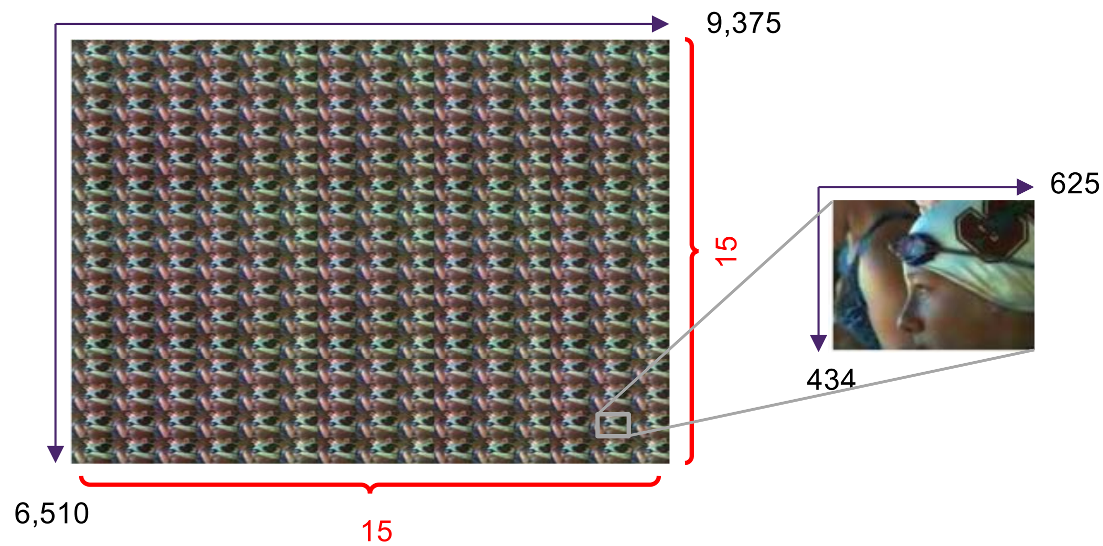
U N I V E R S I T Y   O F   B R I T I S H   C O L U M B I A

Light Field View Synthesis

Plenoptic Camera - challenges

+ High *Angular* Resolution

- Low *Spatial* Resolution

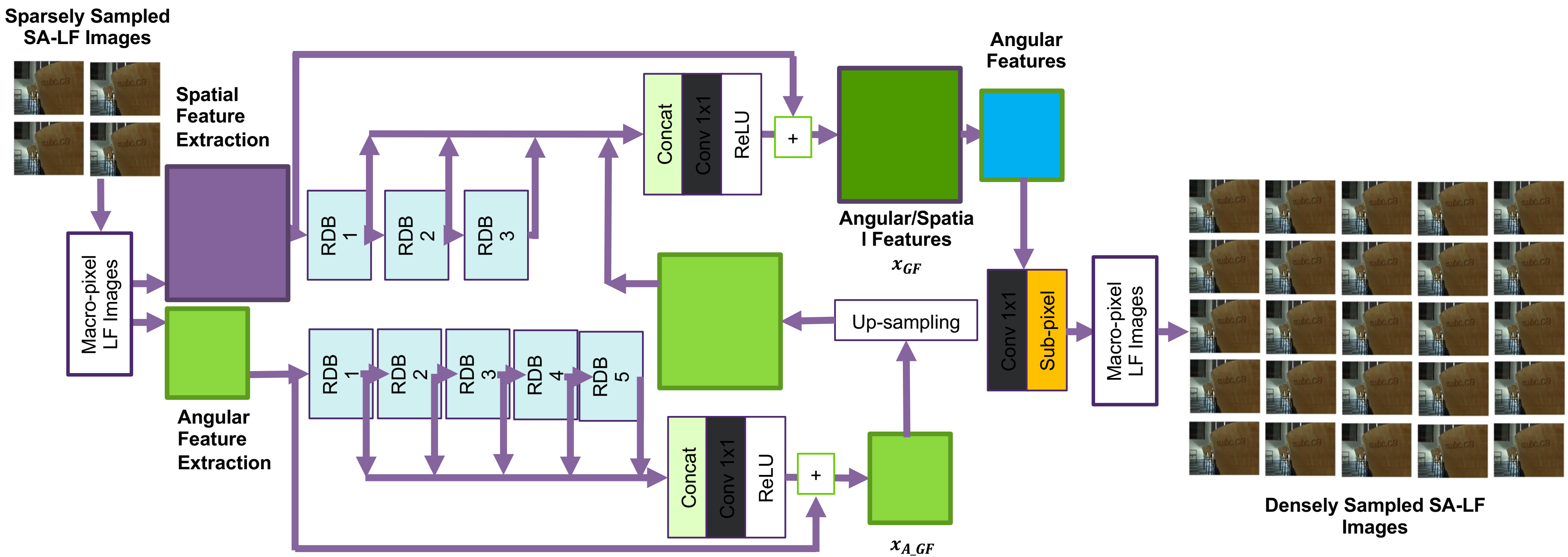


Huge amount of data to transmit

Transmission End

Receiver End

Network for View Synthesis



LF Data Representation

3) Epipolar Image Plane (EPI)

Improving Up-sampled Quality

Low-resolution Sup-aperture Images

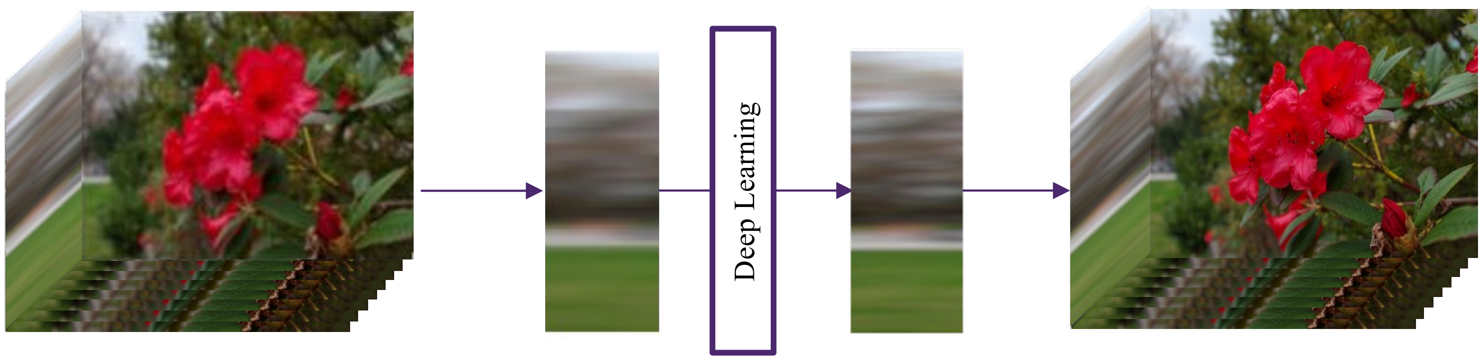
Low-resolution Up-sampled Sup-aperture Images

EPI representation of up-sampled LR images showing discontinuity in disparity

EPI representation of HR images showing continuity in disparity

High-resolution Up-sampled Sup-aperture Images

Improving Up-sampled Quality



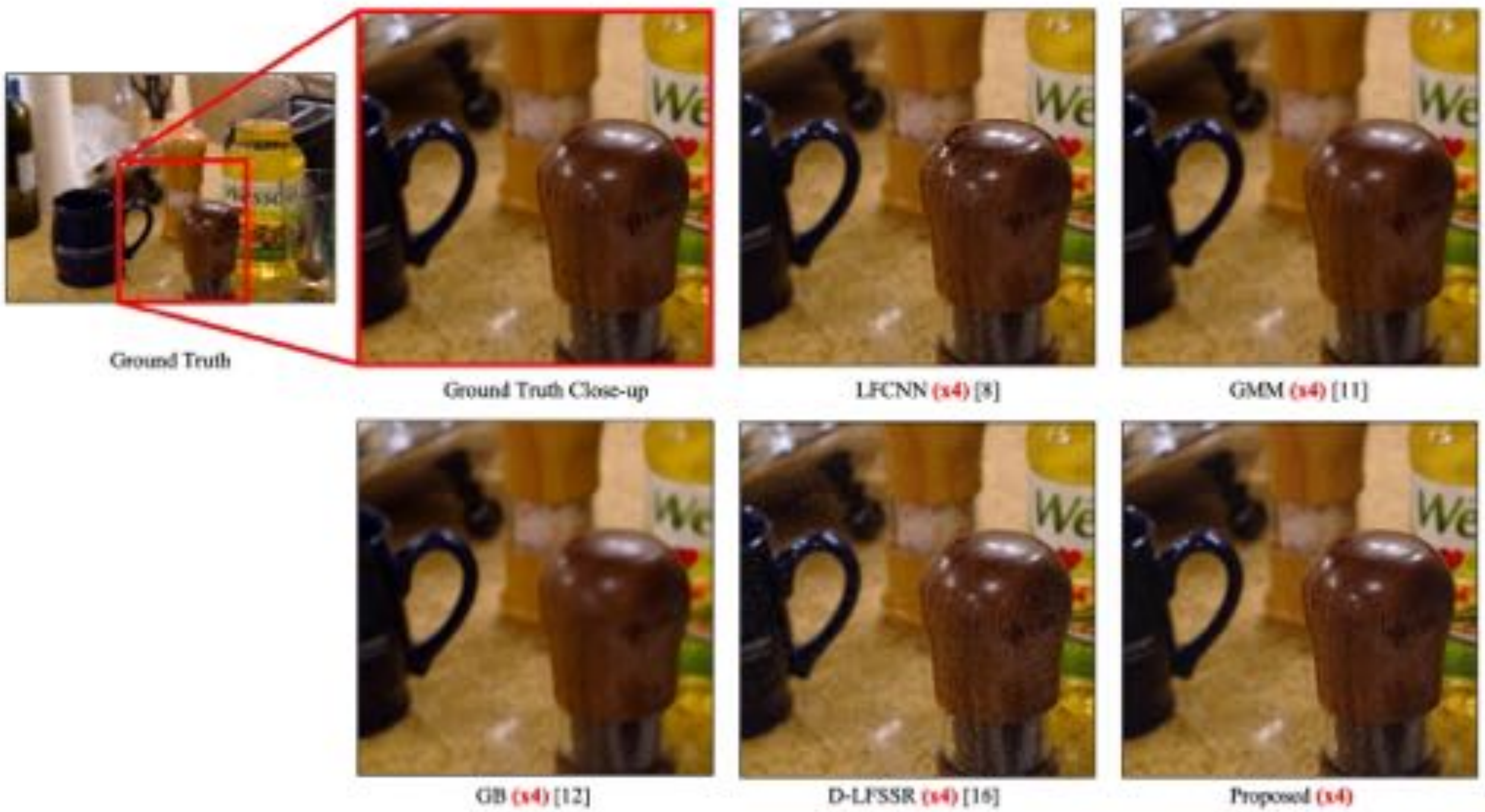
Low-resolution Up-sampled Sup-aperture Images

EPI representation of up-sampled LR images showing discontinuity in disparity

EPI representation of HR images showing continuity in disparity

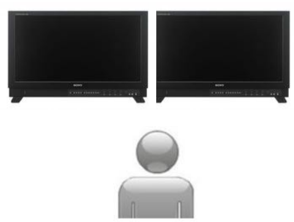
High-resolution Up-sampled Sup-aperture Images

Results

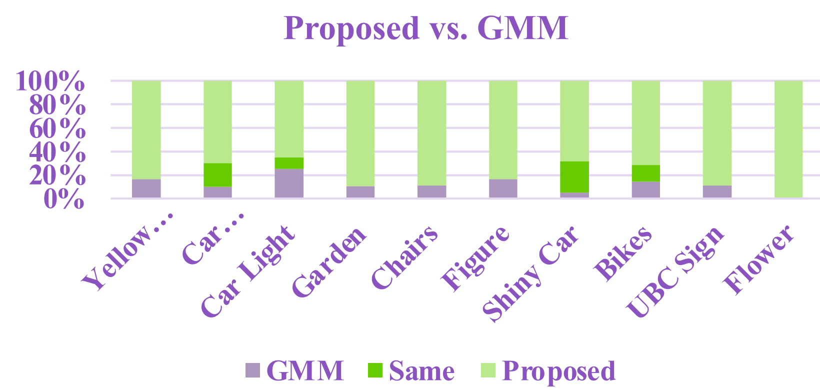

Ground Truth
Ground Truth Close-up
LFCNN (x4) [8]
GMM (x4) [11]
GB (x4) [12]
D-LFSSR (x4) [16]
Proposed (x4)

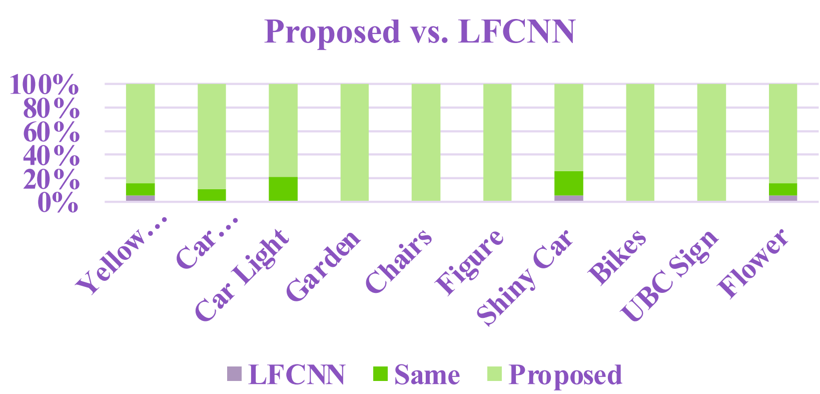

Ground Truth
EPI
LFSSR
ResLF
LF-ATO
Distg-Block
Proposed

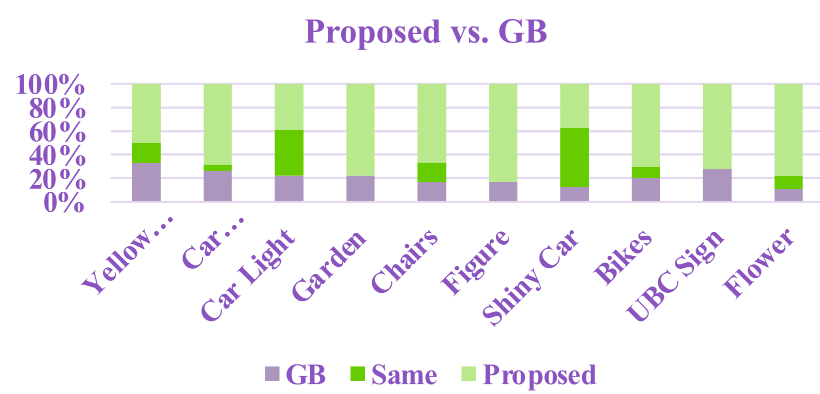
Subjective Results



Side by side evaluation
based on Recommendation BT.500-13 DSIS
Number of subjects: 18
Number of sequences: 10
Subjects asked which method they prefer

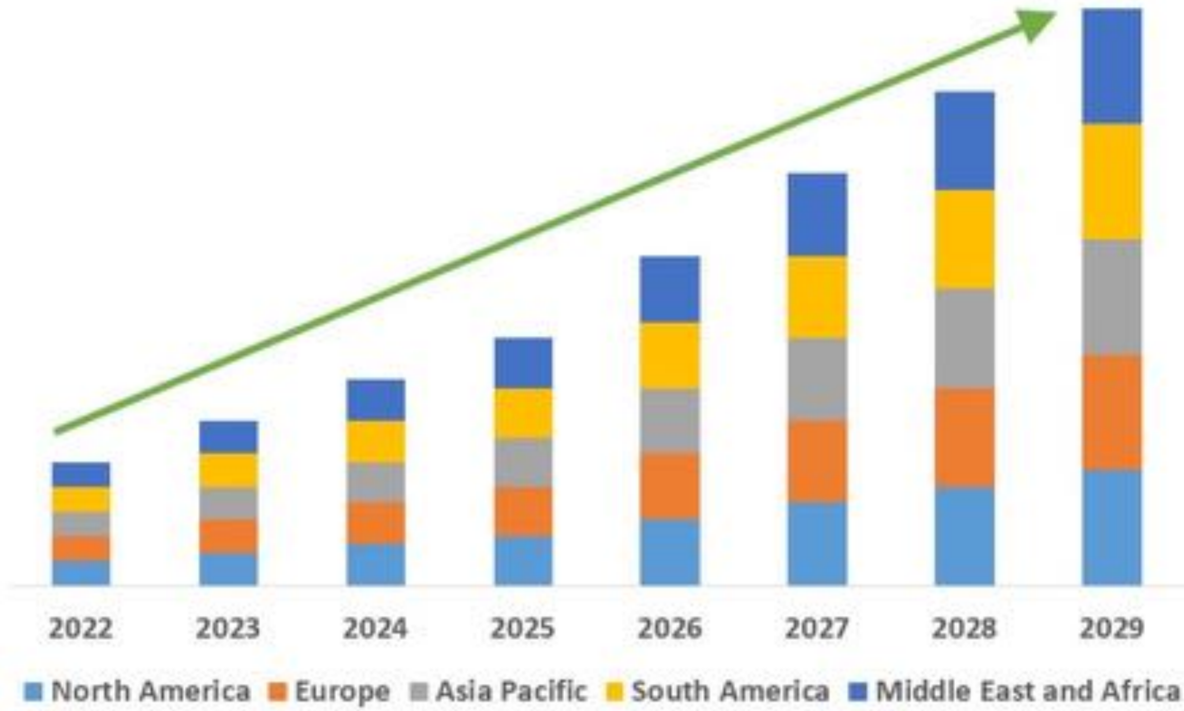






Industry & Market

Market



2022
2023
2024
2025
2026
2027
2028
2029
North America
Europe
Asia Pacific
South America
Middle East and Africa

Industry

magic leap

Leia Inc.

DIMENCO

FYR

LIGHT FIELD LAB®

FoVI3D

IARIA

Magic Leap

Multi-Planar

magic leap

- Multiplanar AR HMD (Head Mounted Display)
- Dual plane per eye
- LCOS SLM
- Waveguide optics


Leia
Lenticular
Leia Inc.

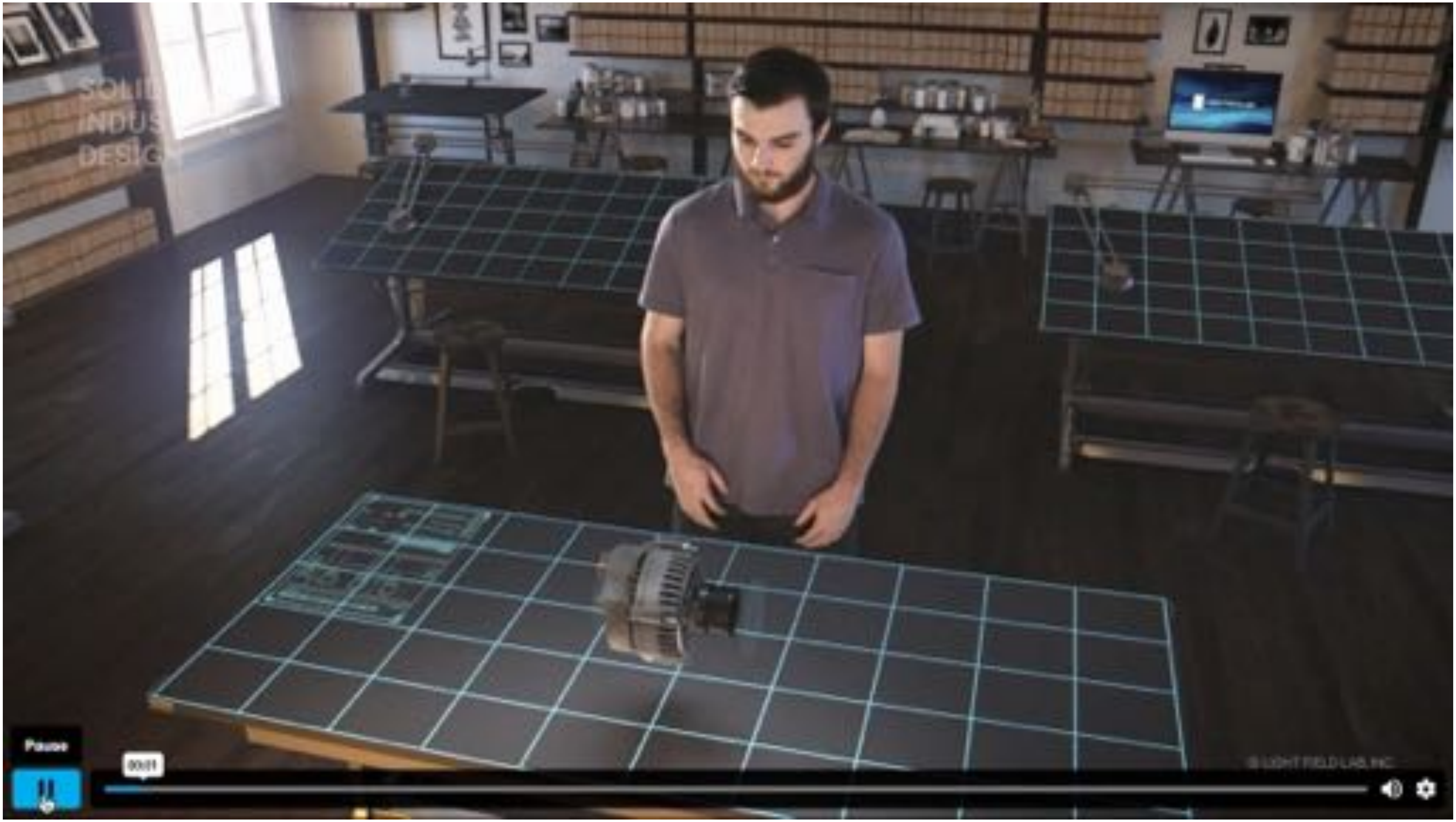
Leia

Lenticular

Leia Inc.

RED MEDIA MACHINE HYDROGEN

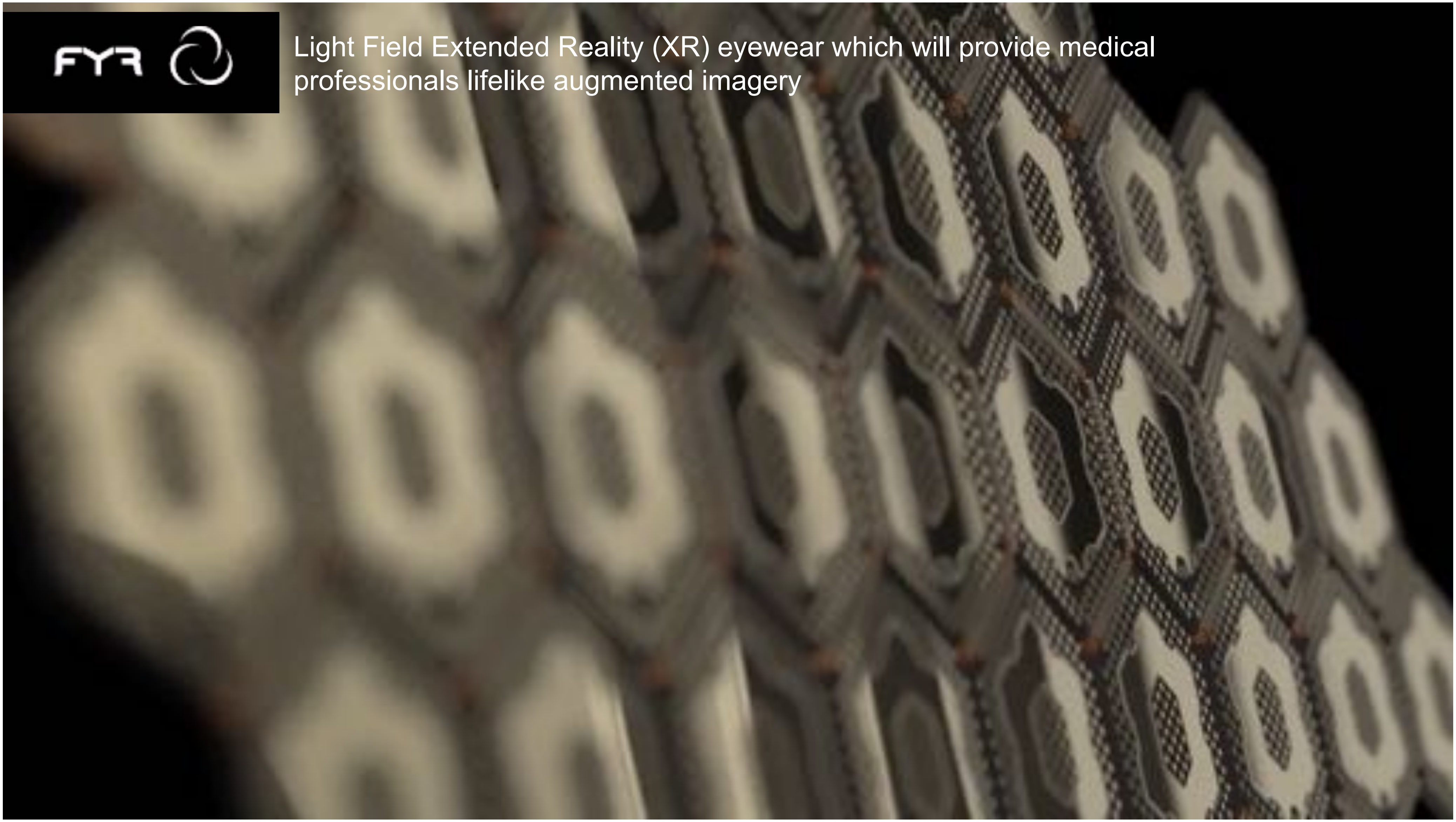
LIGHT FIELD LAB.
Escape the Screen
We're creating a world where content escapes the screen and merges with reality.
TIME
BEST
INVENTIONS
2022


Pause


DIMENCO


SONY

FoVI3D


FYR
Light Field Extended Reality (XR) eyewear which will provide medical professionals lifelike augmented imagery

CONCLUSION


VISUAL ACQUISITION WILL HAVE MOST DISRUPTIVE IMPACT ON MANY TRENDS & MARKETS
AI

Thank you!

Contact:

panos@ece.ubc.ca
http://dml.ece.ubc.ca