ILLIXR Consortium

Democratizing XR research, development, and benchmarking

Sarita Adve and Muhammad Huzaifa

University of Illinois at Urbana-Champaign

illixr@cs.illinois.edu

illixr.org

This work is supported in part by DARPA, NSF, and the Applications Driving Architecture (ADA) Research center (JUMP center co-sponsored by SRC & DARPA)

Team ILLIXR

ILLIXR students and developers

- Madhuparna Bhowmik
- Henry Che
- Rishi Desai
- Steven Gao
- Samuel Grayson
- Qinjun Jiang
- Muhammad Huzaifa
- Xutao Jiang
- Ying Jing
- Jae Lee
- Fang Lu
- Yihan Pang
- Joseph Ravichandran
- Giordano Salvador
- Bill Sherman
- Finn Sinclair
- Rahul Singh
- Boyuan Tian
- Lauren Wagner
- Henghzi Yuan
- Jeffrey Zhang

Consultations

- Ameen Akeel
- Wei Cui
- Aleksandra Faust
- Liang Gao
- Rod Hooker
- Matt Horsnell
- Amit Jindal
- Steve LaValle
- Steve Lovegrove
- David Luebke
- Andrew Maimone
- Vegard Oye
- Maurizio Paganini
- Martin Persson
- Archontis Politis
- Eric Shaffer
- Paris Smaragdis
- Chris Widdowson

Founding consortium members: Arm, Meta Reality Labs, Micron, NVIDIA

Founding sponsor: ADA research center, a DARPA/SRC JUMP center

Extended Reality (XR): The Next Interface

Virtual, Augmented, Mixed Reality

The next computing interface

Will transform science, medicine, education, …

Approximate	Current	Desired
Res (Mpixels)	**7**	**200**
Power (W)	**~7**	**0.1**
Weight (g)	**500**	**10**
…	**…**	**…**

BUT orders of magnitude gap

in power, performance, quality-of-experience

between current and desired systems

XR Systems: Challenges

Approximate	Current	Desired
Res (Mpixels)	**7**	**200**
Power (W)	**~7**	**0.1**
Weight (g)	**500**	**10**
…	**…**	**…**

Orders of magnitude gap

Power, performance, quality-of-experience (QoE)

Diverse expertise

graphics, vision, audio, video, optics, haptics, …

Cross-layer system co-design

hardware, compiler, OS, algorithm

Complex metrics

multiple, user-driven, end-to-end QoE metrics

Closed systems, few participants

No open reference systems or benchmarks

Large barrier to entry for open R&D

How can we democratize XR systems research, development, benchmarking?

ILLIXR: Illinois Extended Reality Testbed

ILLIXR: Open-source full system XR testbed

State-of-the-art XR components w/ modular runtime

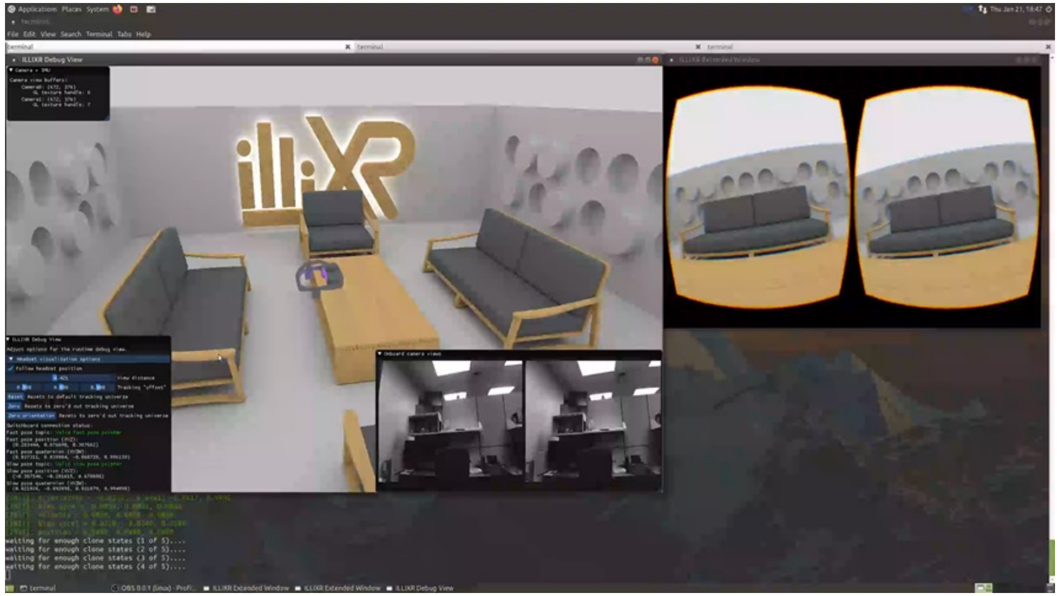
OpenXR compatible

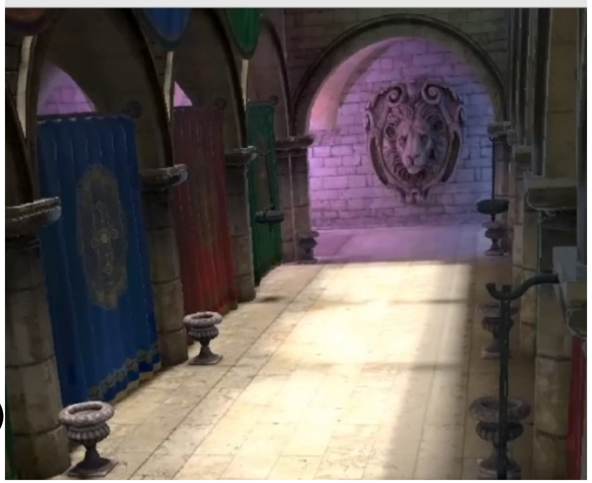
Extensive characterization and use for research

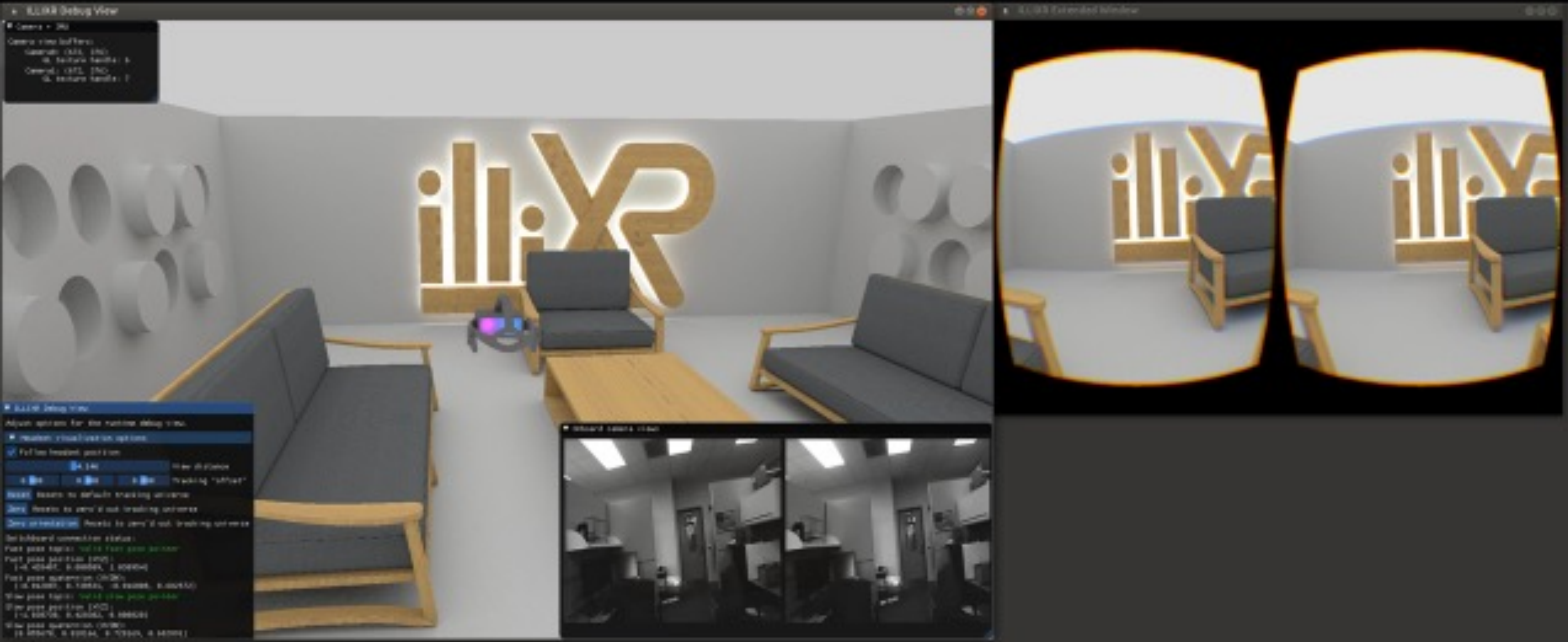
illixr.org

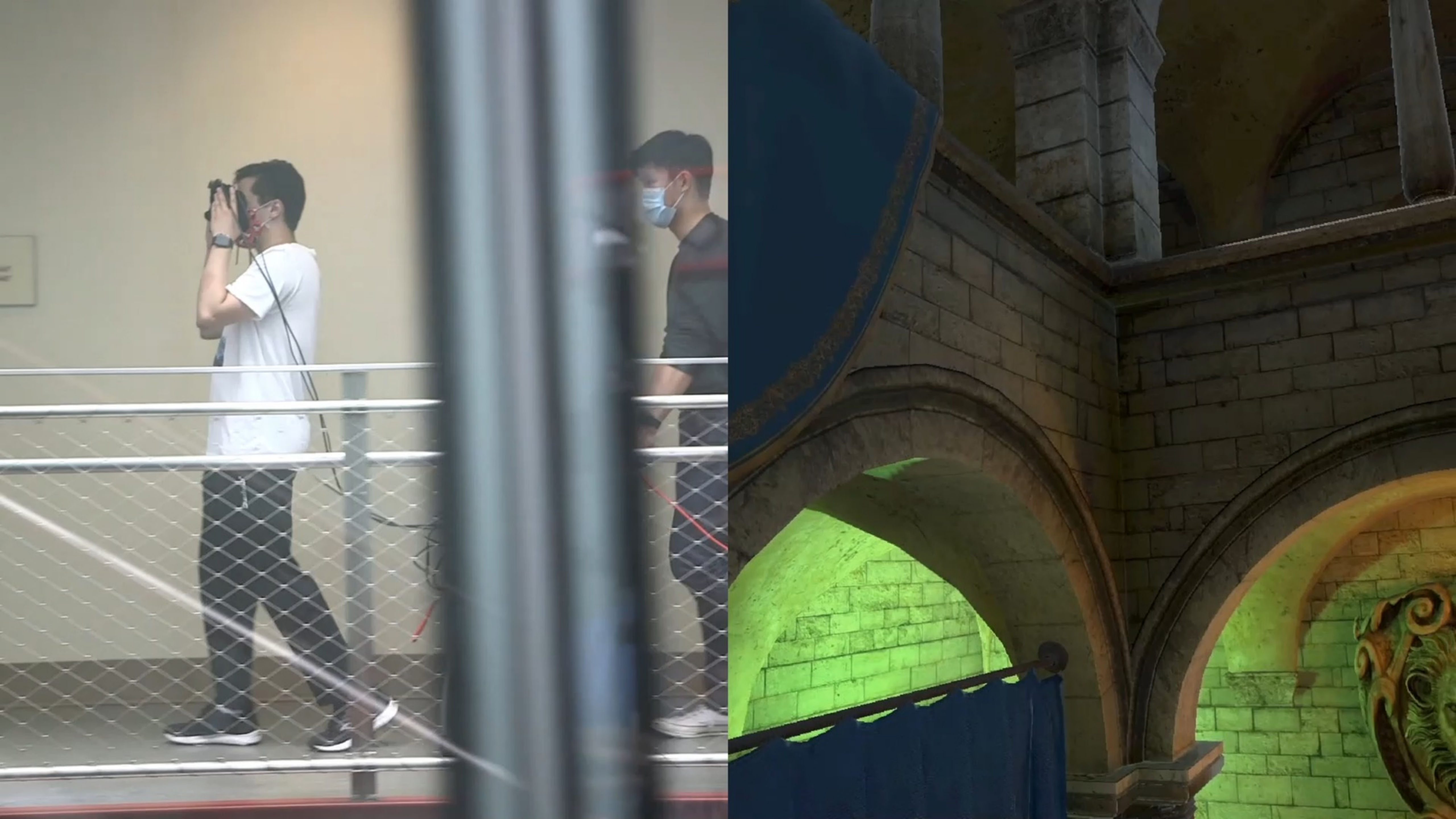
Huzaifa et al., IISWC'21 best paper,

IEEE Micro Top Picks'22 (top 12 papers from 2021 architecture conferences)






ILLIXR Debug View

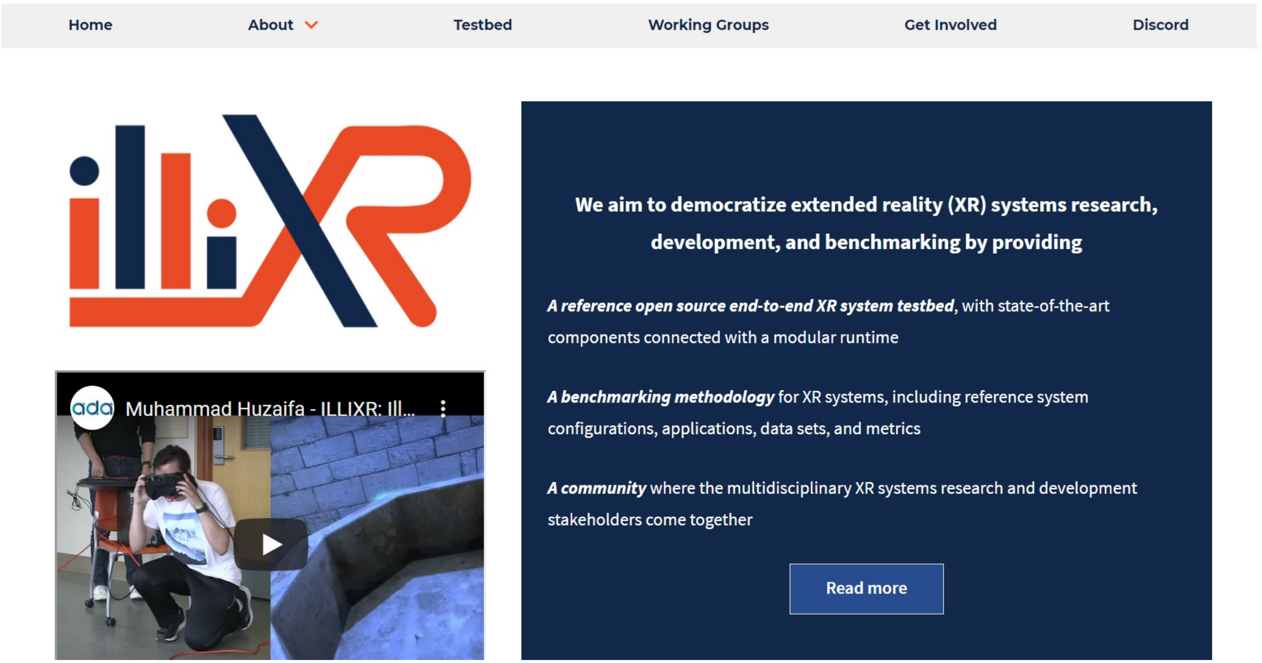


ILLIXR Consortium

ILLIXR Consortium w/ industry + academic partners

- Arm, Meta, Micron, North Star, NVIDIA, …

illixr.org



Goals

- Reference open-source testbed
  - Components and interfaces
  - Modular, extensible runtime
  - Telemetry
- Benchmarking methodology
  - Applications, data sets
  - System configurations
  - Metrics
- Build XR systems research and development community

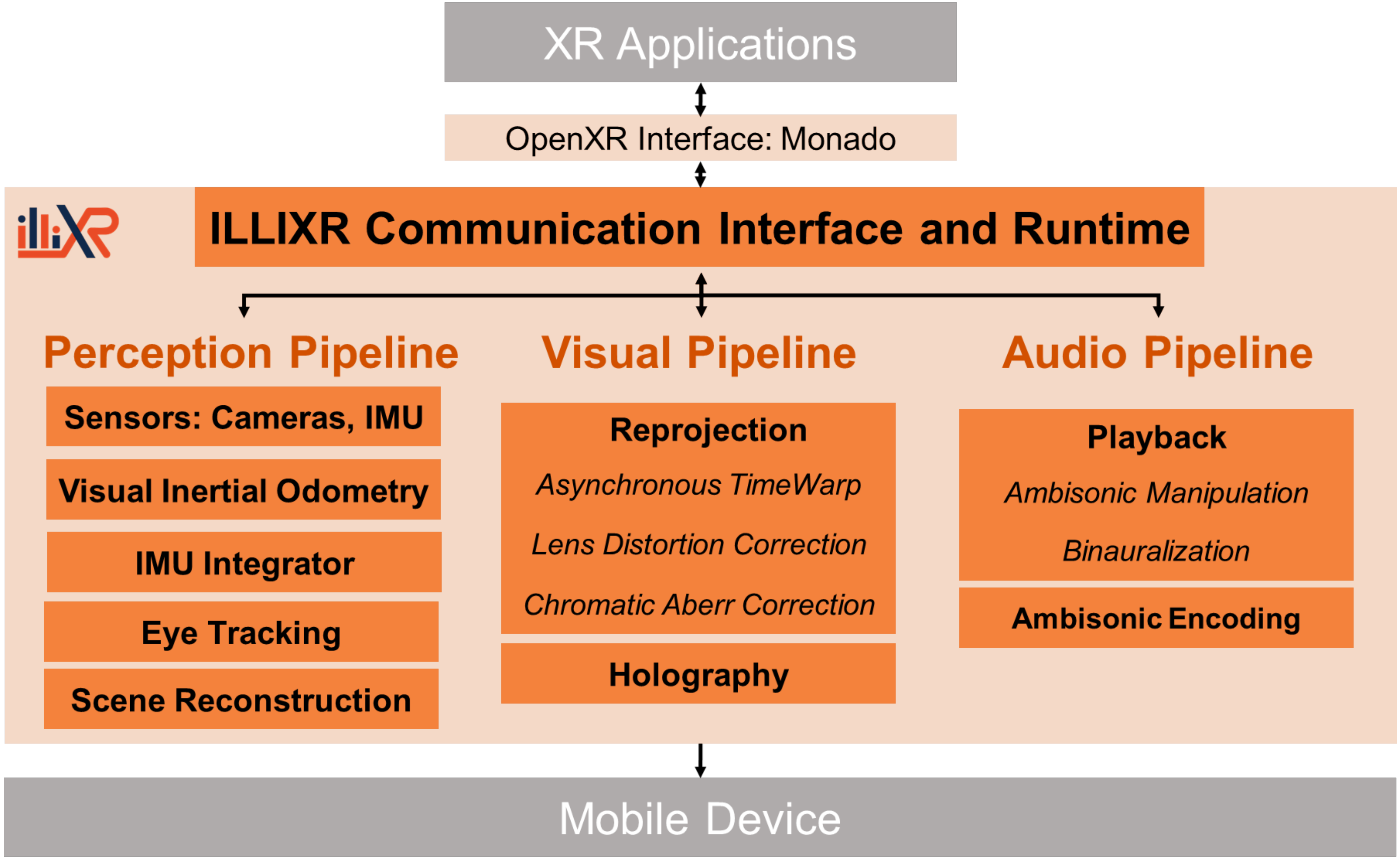
Now funded by NSF CISE community research infrastructure program

Join us: illixr@cs.illinois.edu, illixr.org, discord, weekly meetings

Outline

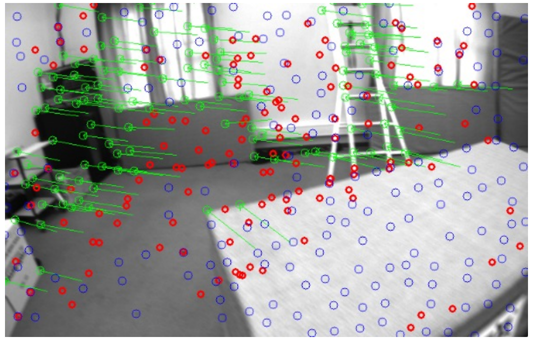
- Introduction
- ILLIXR Description
- Evaluation and Implications
- Ongoing Research with ILLIXR
- Ongoing Work on ILLIXR Testbed
- Consortium Organization and Goals

ILLIXR Overview

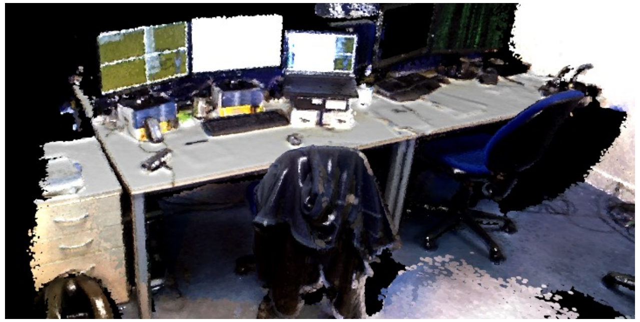


Perception Pipeline

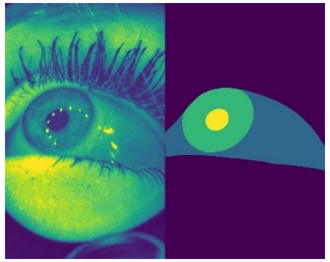
- Sensors: Camera, Inertial Measurement Unit (IMU)
- Visual Intertial Odometry (VIO)
  - Provides position and head orientation (pose)



- IMU Integrator
  - Provides high frequency pose estimates
- Pose Predictor
  - Extrapolates pose to future timestamp

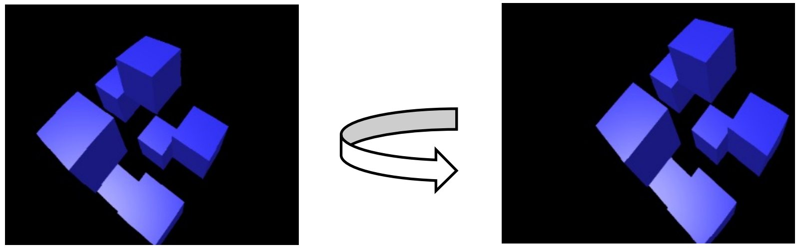


- Scene Reconstruction
  - Uses RGB-Depth camera to build dense 3D map of world
- Eye Tracking

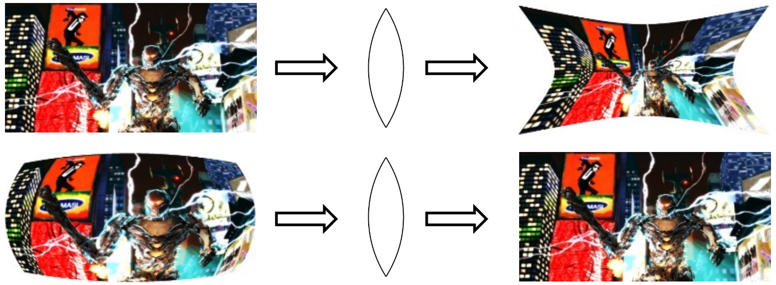


Visual Pipeline

- Asynchronous reprojection
  - Warp rendered frame to account for head movement during rendering
  - Uses latest pose estimate and prediction
  - Cuts motion-to-photon latency



- Lens distortion and chromatic aberration correction
  - Corrects for distortion due to curved lenses



- Computational holography
  - Vergence-accommodation conflict (VAC): eyes focused at fixed point, converge at different points
  - Computational displays w/ multiple focal planes can fix VAC: compute per-pixel phase shift

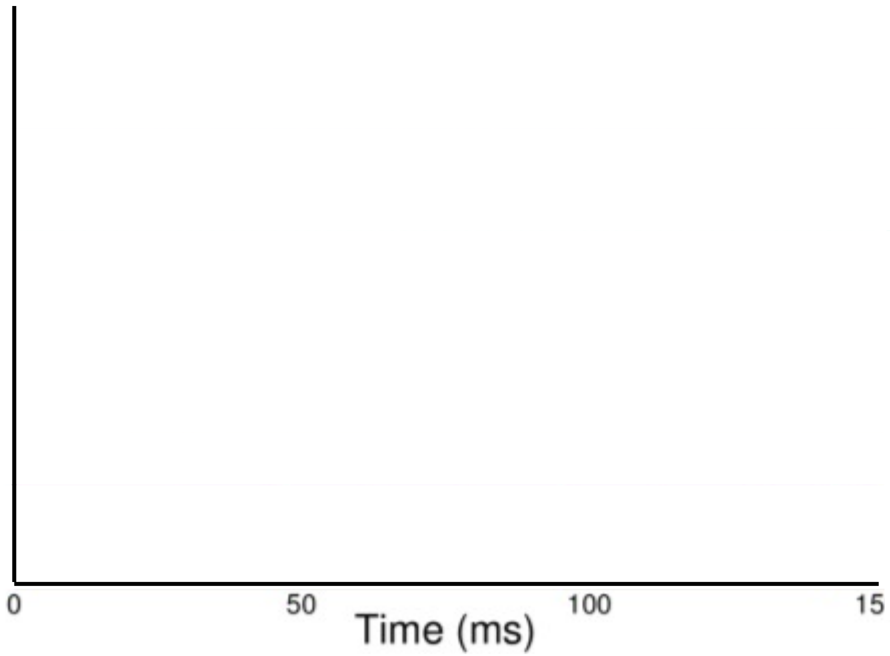
Audio Pipeline

- Audio encoding
  - Encodes multiple sound sources into Higher Order Ambisonics (HOA) soundfield
- Playback
  - Rotates and zooms HOA sound field for user’s latest pose
  - Performs binauralization to account for user’s ear, head, nose

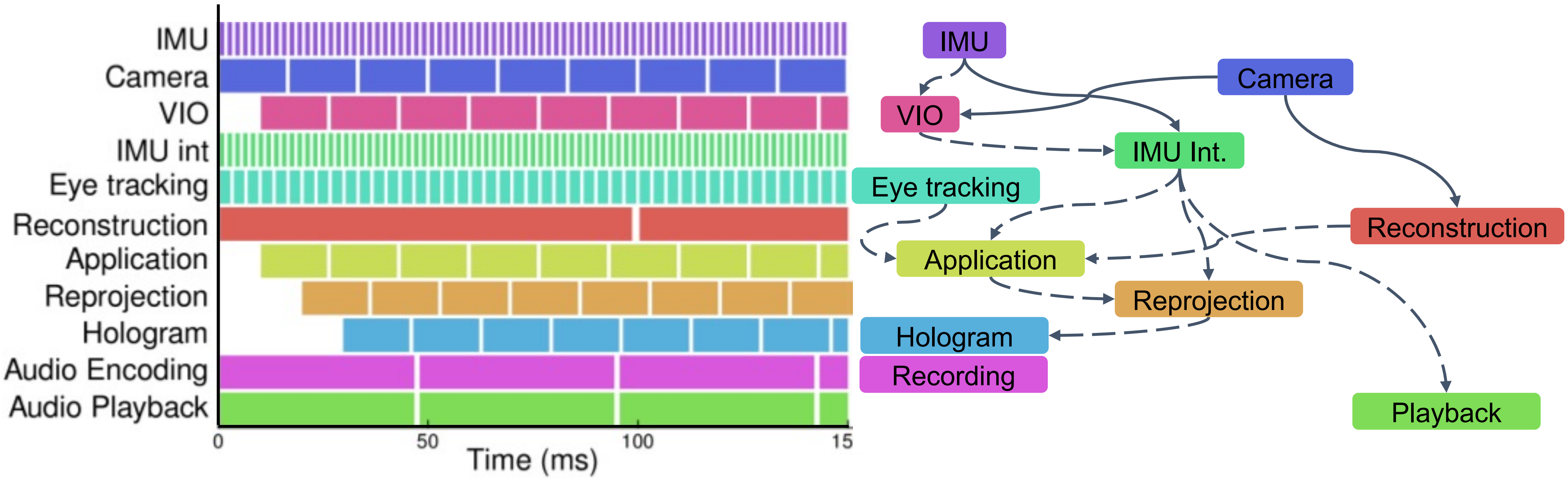
BUT XR is not just a collection of components

It is a SYSTEM

XR System Dataflow

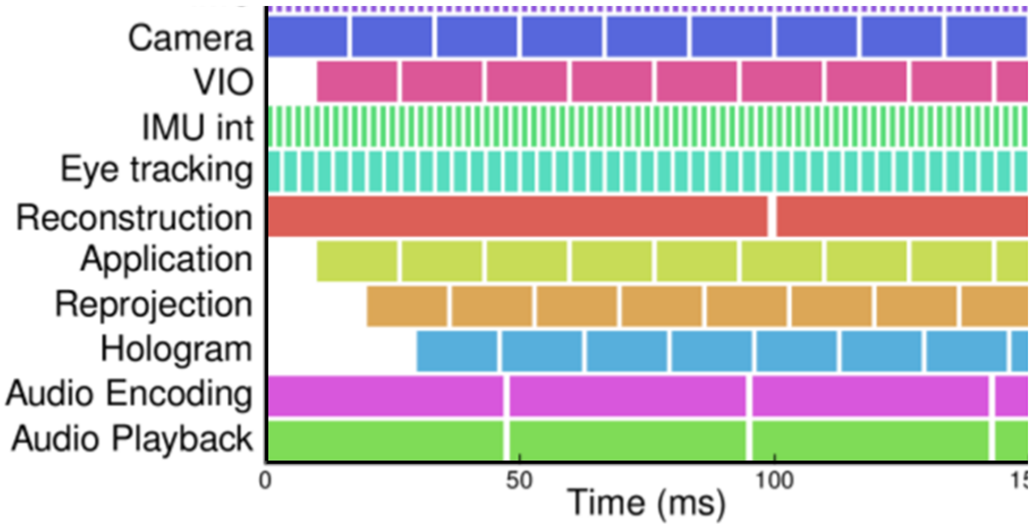


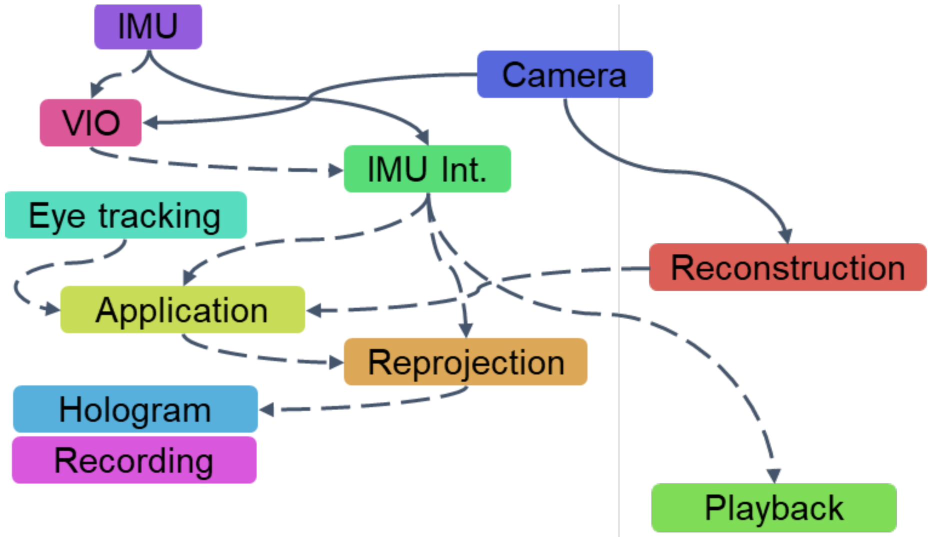
XR System Dataflow



Different components at different frequencies
Multiple interacting pipelines
Synchronous and asynchronous dependences
Multiple quality of experience metrics

ILLIXR Runtime





Modular, flexible architecture

ILLIXR components are plugins

Separately compiled, dynamically loaded

Easily swap/add new components, implementations

Efficient, flexible communication interface

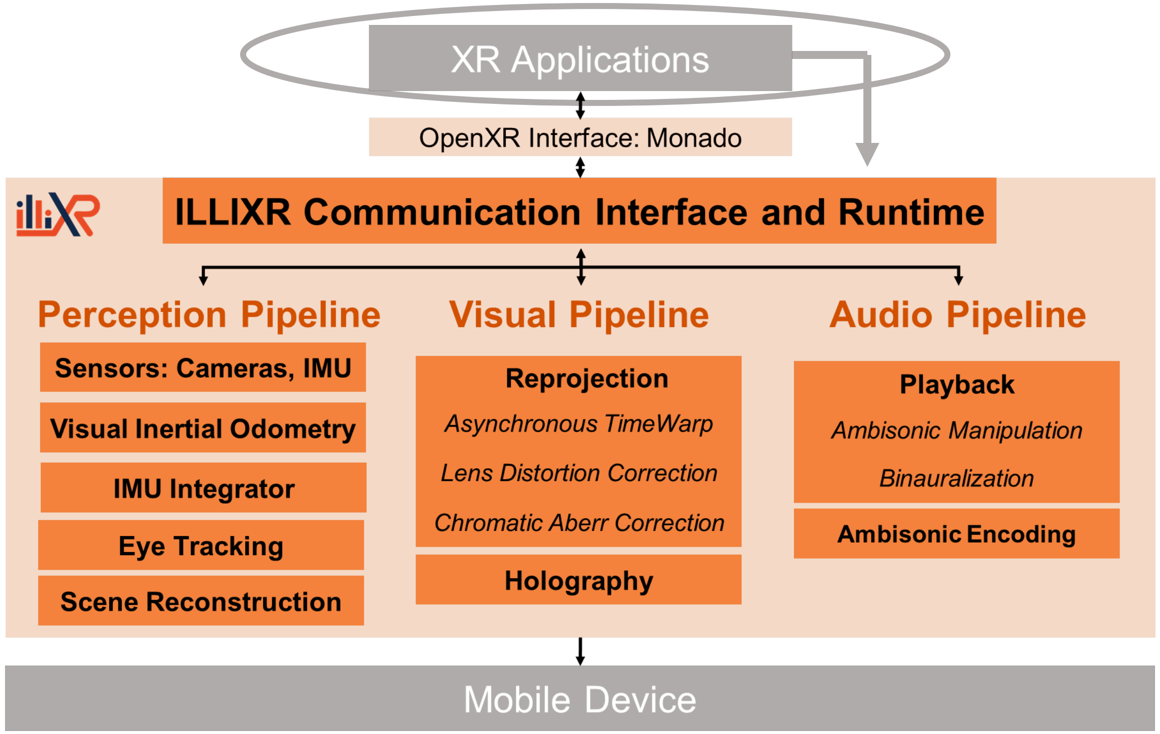
Component specifies event streams to publish, subscribe

Synchronous or asynchronous consumers

Copy-free, shared memory implementation

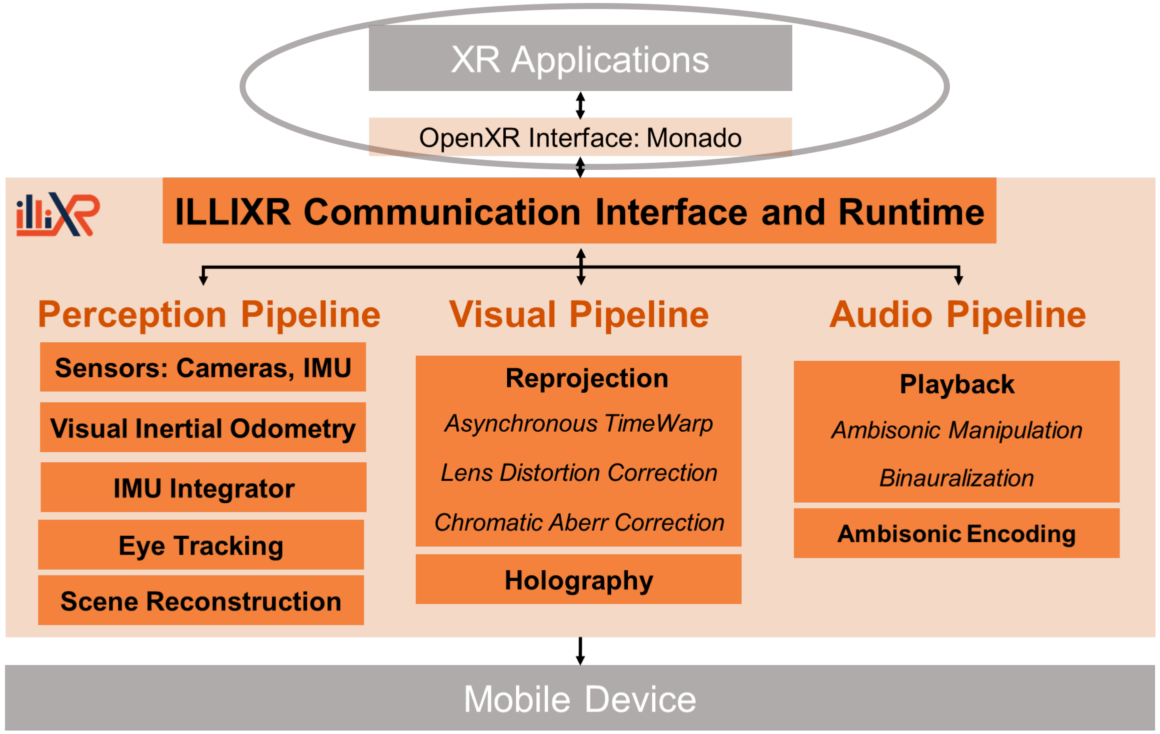
End-to-end system balances flexibility with efficiency

ILLIXR Applications



Can write XR applications directly to ILLIXR

ILLIXR Applications



Can write XR applications directly to ILLIXR

ILLIXR supports OpenXR applications

- Uses Monado implementation of OpenXR
- Today: Godot game engine with many apps
- Soon: Unity, Unreal, …

End-to-End Quality Metrics

- Motion-to-photon latency
  - Time from head motion to display (currently w/o display latency)

- Image quality: SSIM and FLIP

+ Extensive telemetry: Frame rates, missed frames, time distributions, power, …

ILLIXR Components and Systems Today

	Component	Algorithm	Implementation
Perception Pipeline	Camera	ZED SDK	C++
	Camera	Intel RealSense SDK	C++
	IMU	ZED SDK	C++
	IMU	Intel RealSense SDK	C++
	VIO	OpenVINS	C++
	VIO	Kimera-VIO	C++
	IMU Integrator	RK4	C++
	IMU Integrator	GTSAM	C++
	Eye Tracking	RITnet	Python, CUDA
	Scene Reconstruction	ElasticFusion	C++, CUDA, GLSL
	Scene Reconstruction	KinectFusion	C++, CUDA
Visual Pipeline	Reprojection	VP-matrix reproject w/ pose	C++, GLSL
	Lens Distortion	Mesh-based radial distortion	C++, GLSL
	Chromatic Aberration	Mesh-based radial distortion	C++, GLSL
	Adaptive Display	Weighted Gerchberg-Saxton	CUDA
Audio Pipeline	Audio Encoding	Ambisonic encoding	C++
	Audio Playback	Ambisonic manipulation, binauralization	C++

Systems

Desktop PC

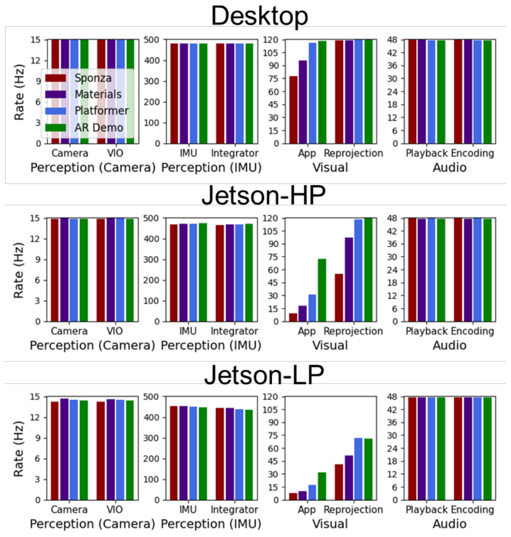
Jetson Xavier

Outline

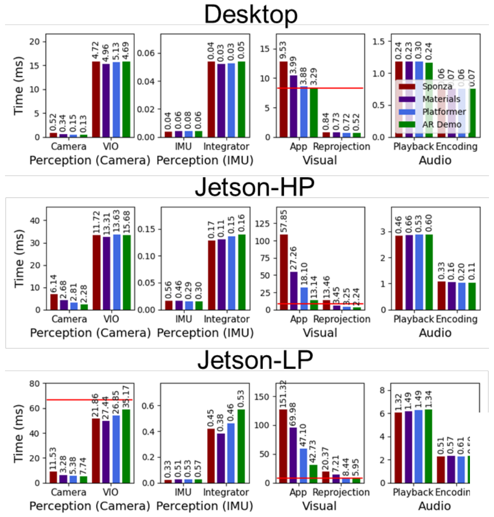
- Introduction
- ILLIXR Description
- Evaluation and Implications
- Ongoing Research with ILLIXR
- Ongoing Work on ILLIXR Testbed
- Consortium Organization and Goals

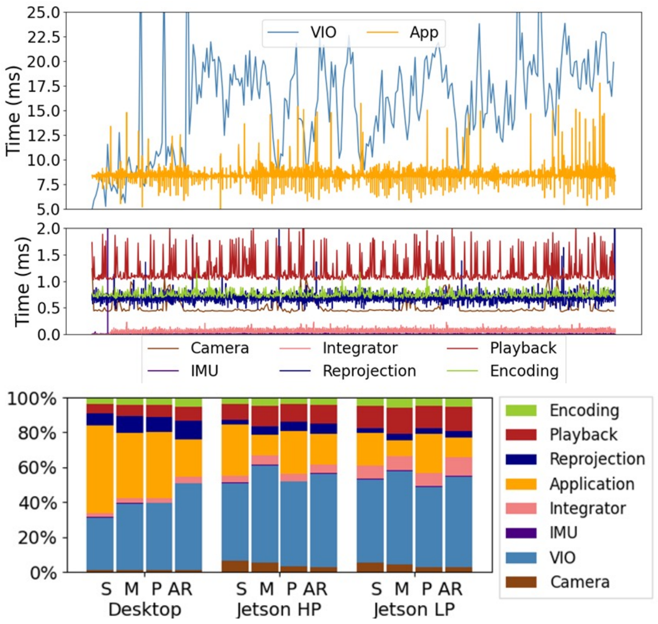
Results Summary

Frame Rate



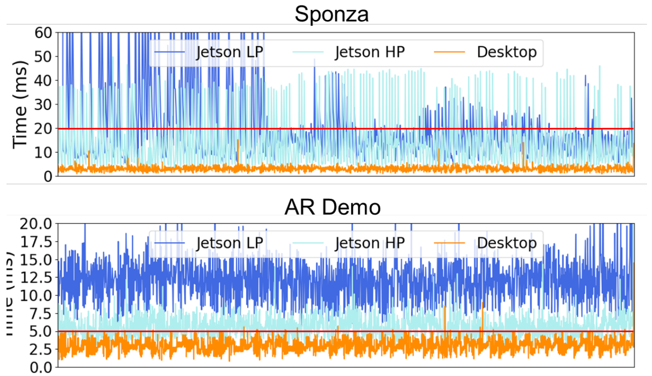
Execution Time & Distribution





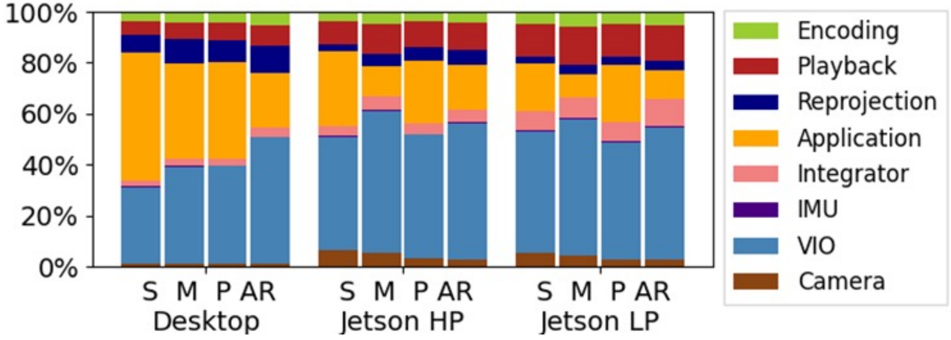
Quality of Experience

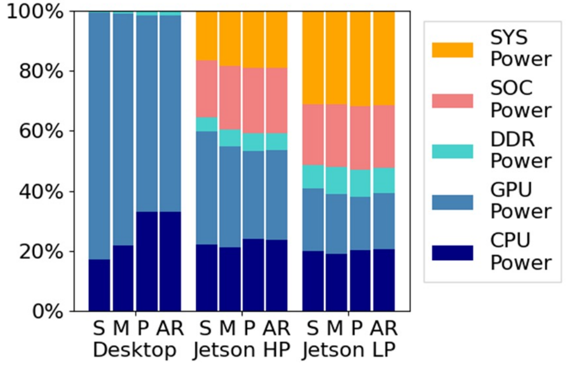
Application	Desktop	Jetson-HP	Jetson-LP
Sponza	3.1 ± 1.1	13.5 ± 10.7	19.3 ± 14.5
Materials	3.1 ± 1.0	7.7 ± 2.7	16.4 ± 4.9
Platformer	3.0 ± 0.9	6.0 ± 1.9	11.3 ± 4.7
AR Demo	3.0 ± 0.9	5.6 ± 1.4	12.0 ± 3.4



Platform	SSIM	1-FLIP
Desktop	0.83 ± 0.04	0.86 ± 0.05
Jetson-HP	0.80 ± 0.05	0.85 ± 0.05
Jetson-LP	0.68 ± 0.09	0.65 ± 0.17

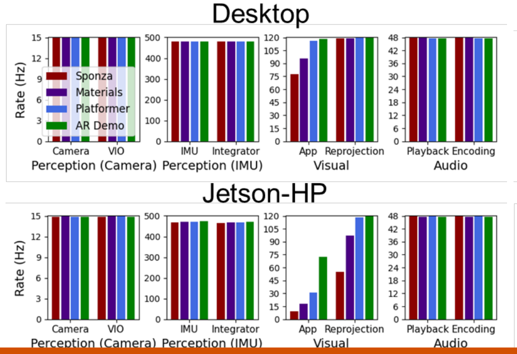
Power



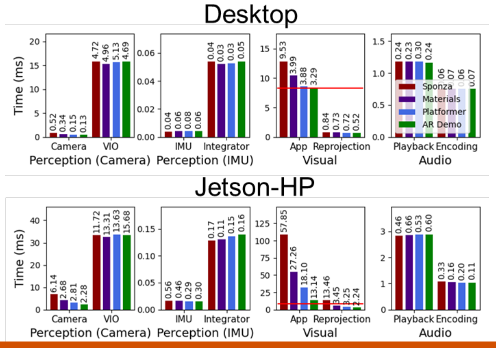


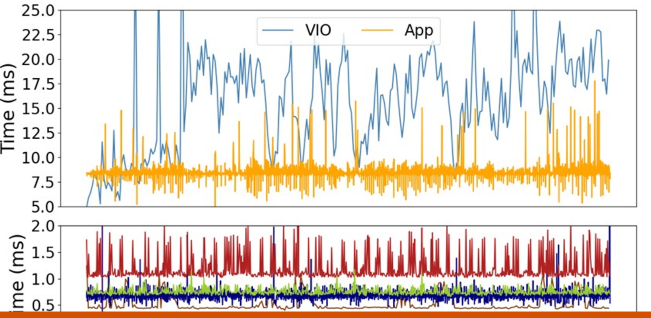
Results Summary

Frame Rate



Execution Time & Distribution

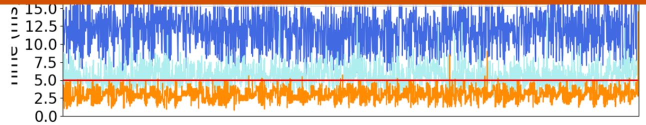




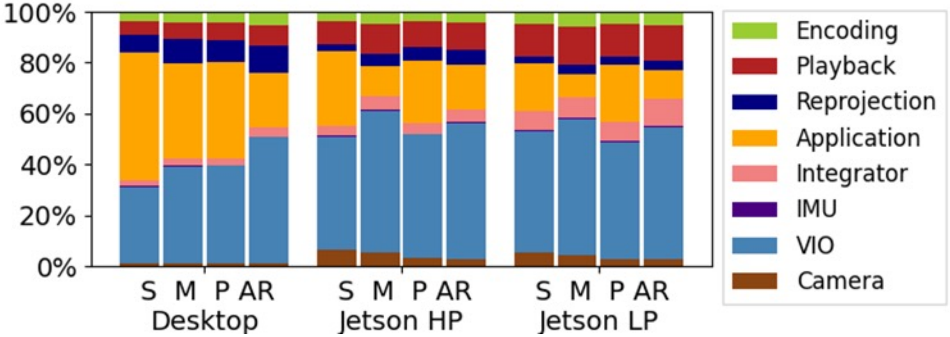
Quality of Experience

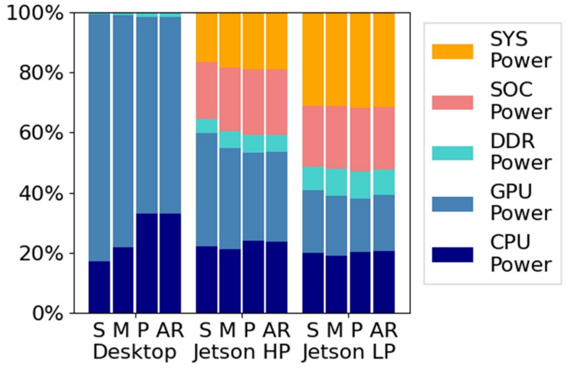
Application	Desktop	Jetson-HP	Jetson-LP
Sponza	3.1 ± 1.1	13.5 ± 10.7	19.3 ± 14.5
Materials	3.1 ± 1.0	7.7 ± 2.7	16.4 ± 4.9
Platformer	3.0 ± 0.9	6.0 ± 1.9	11.3 ± 4.7
AR Demo	3.0 ± 0.9	5.6 ± 1.4	12.0 ± 3.4

First published performance/power/QoE results for end-to-end XR system



Power





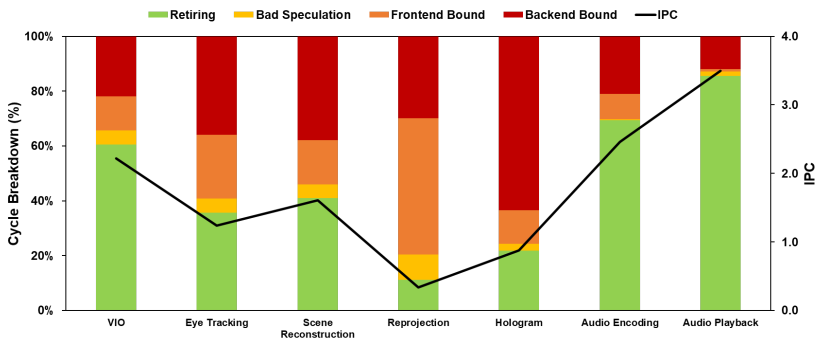
Platform	SSIM	1-FLIP
Desktop	0.83 ± 0.04	0.86 ± 0.05
Jetson-HP	0.80 ± 0.05	0.85 ± 0.05
Jetson-LP	0.68 ± 0.09	0.65 ± 0.17

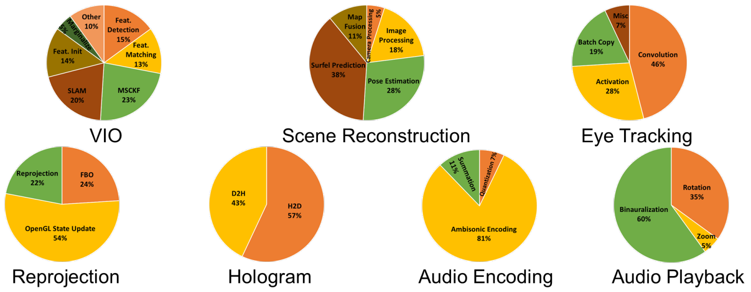
Summary of Implications for System Designers

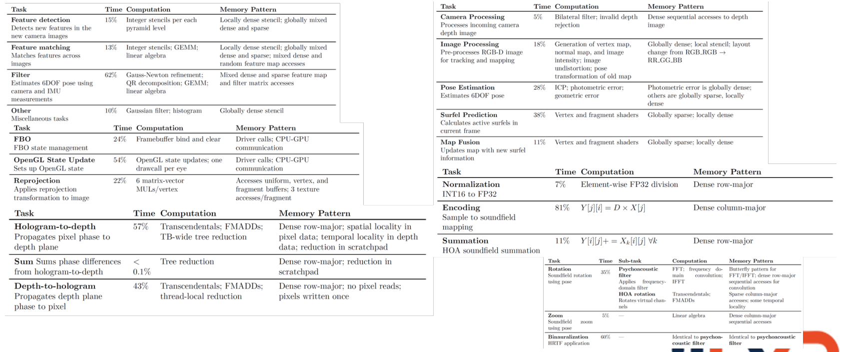
- Need to specialize hardware, software, *system*
- Must consider all application components in *system* together
- Must consider *system*-level hardware components; e.g., display and I/O
- Need to partition, allocate, and schedule *system* resources
- Must look at entire *system* to make QoE-driven tradeoffs

- Abundance of tasks and no single task dominates
  - ⇒ Need *automated* techniques to determine what to accelerate
- Impractical to build accelerator for every task
  - ⇒ Must build *shared* hardware
- Diversity of compute and memory primitives
  - ⇒ *Flexible* on-chip memory hierarchy
  - ⇒ *Flexible* accelerator communication interface
- Algorithms in flux
  - ⇒ Must design *programmable* hardware
- Different algorithms have different QoE vs. resource usage profiles
  - ⇒End-to-end QoE driven *approximate computing*

Standalone Components







Task	Time	Computation	Memory Pattern
Feature detection Detects new features in the new camera images	15%	Integer stencils per each pyramid level	Locally dense stencil; globally mixed dense and sparse
Feature matching Matches features across images	13%	Integer stencils; GEMM; linear algebra	Locally dense stencil; globally mixed dense and sparse; mixed dense and random feature map accesses
Filter Estimates 6DOF pose using camera and IMU measurements	62%	Gauss-Newton refinement; QR decomposition; GEMM; linear algebra	Mixed dense and sparse feature map and filter matrix accesses
Other Miscellaneous tasks	10%	Gaussian filter; histogram	Globally dense stencil

Task	Time	Computation	Memory Pattern
FBO FBO state management	24%	Framebuffer bind and clear	Driver calls; CPU-GPU communication
OpenGL State Update Sets up OpenGL state	54%	OpenGL state updates; one drawcall per eye	Driver calls; CPU-GPU communication
Reprojection Applies reprojection transformation to image	22%	6 matrix-vector MULs/vertex	Accesses uniform, vertex, and fragment buffers; 3 texture accesses/fragment

Task	Time	Computation	Memory Pattern
Hologram-to-depth Propagates pixel phase to depth plane	57%	Transcendentals; FMADDs; TB-wide tree reduction	Dense row-major; spatial locality in pixel data; temporal locality in depth data; reduction in scratchpad
Sum Sums phase differences from hologram-to-depth	< 0.1%	Tree reduction	Dense row-major; reduction in scratchpad
Depth-to-hologram Propagates depth plane phase to pixel	43%	Transcendentals; FMADDs; thread-local reduction	Dense row-major; no pixel reads; pixels written once

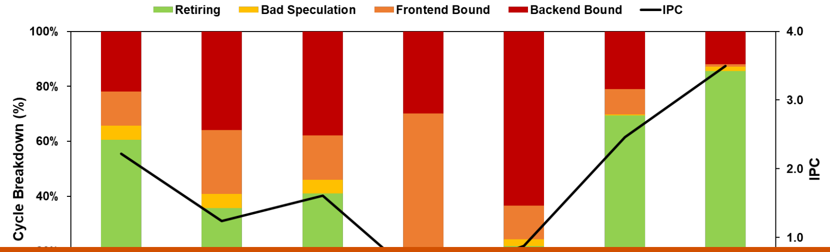
Task	Time	Computation	Memory Pattern
Camera Processing Processes incoming camera depth image	5%	Bilateral filter; invalid depth rejection	Dense sequential accesses to depth image
Image Processing Pre-processes RGB-D image for tracking and mapping	18%	Generation of vertex map, normal map, and image intensity; image undistortion; pose transformation of old map	Globally dense; local stencil; layout change from RGB_RGB → RR_GG_BB
Pose Estimation Estimates 6DOF pose	28%	ICP; photometric error; geometric error	Photometric error is globally dense; others are globally sparse, locally dense
Surfel Prediction Calculates active surfels in current frame	38%	Vertex and fragment shaders	Globally sparse; locally dense
Map Fusion Updates map with new surfel information	11%	Vertex and fragment shaders	Globally sparse; locally dense

Task	Time	Computation	Memory Pattern
Normalization INT16 to FP32	7%	Element-wise FP32 division	Dense row-major
Encoding Sample to soundfield mapping	81%	$Y[j][i] = D \times X[j]$	Dense column-major
Summation HOA soundfield summation	11%	$Y[i][j] += X_k[i][j] \ \forall k$	Dense row-major

Task	Time	Sub-task	Computation	Memory Pattern
Rotation Soundfield rotation using pose	35%	**Psychoacoustic filter** Applies frequency-domain filter	FFT; frequency domain convolution; IFFT	Butterfly pattern for FFT/IFFT; dense row-major sequential accesses for convolution
		HOA rotation Rotates virtual channels	Transcendentals; FMADDs	Sparse column-major accesses; some temporal locality
Zoom Soundfield zoom using pose	5%	—	Linear algebra	Dense column-major sequential accesses
Binauralization HRTF application	60%	—	Identical to **psychoacoustic filter**	Identical to **psychoacoustic filter**

Summary of Implications for System Designers

Standalone Components



- Need to specialize hardware, software, *system*
- Must consider all application components in *system* together
- Must consider *system*-level hardware components; e.g., display and I/O
- Need to partition, allocate, and schedule *system* resources
- Must look at entire *system* to make QoE-driven tradeoffs

ILLIXR = Rich playground for systems research

Enables QoE-driven, end-to-end hardware-software-algorithm codesigned systems research

- Diversity of compute and memory primitives
  - ⇒ *Flexible* on-chip memory hierarchy
  - ⇒ *Flexible* accelerator communication interface
- Algorithms in flux
  - ⇒ Must design *programmable* hardware
- Different algorithms have different QoE vs. resource usage profiles
  - ⇒ End-to-end QoE driven *approximate computing*

Reprojection Hologram Audio Encoding Audio Playback

Task	Time	Computation	Memory Pattern
Feature detection Detects new features in the new camera images	15%	Integer stencils per each pyramid level	Locally dense stencil; globally mixed dense and sparse
Feature matching Matches features across images	13%	Integer stencils; GEMM; linear algebra	Locally dense stencil; globally mixed dense and sparse; mixed dense and random feature map accesses
Filter Estimates 6DOF pose using camera and IMU measurements	62%	Gauss-Newton refinement; QR decomposition; GEMM; linear algebra	Mixed dense and sparse feature map and filter matrix accesses
Other Miscellaneous tasks	10%	Gaussian filter; histogram	Globally dense stencil

Task	Time	Computation	Memory Pattern
FBO FBO state management	24%	Framebuffer bind and clear	Driver calls; CPU-GPU communication
OpenGL State Update Sets up OpenGL state	54%	OpenGL state updates; one drawcall per eye	Driver calls; CPU-GPU communication
Reprojection Applies reprojection transformation to image	22%	6 matrix-vector MULs/vertex	Accesses uniform, vertex, and fragment buffers; 3 texture accesses/fragment

Task	Time	Computation	Memory Pattern
Hologram-to-depth Propagates pixel phase to depth plane	57%	Transcendentals; FMADDs; TB-wide tree reduction	Dense row-major; spatial locality in pixel data; temporal locality in depth data; reduction in scratchpad
Sum Sums phase differences from hologram-to-depth	< 0.1%	Tree reduction	Dense row-major; reduction in scratchpad
Depth-to-hologram Propagates depth plane phase to pixel	43%	Transcendentals; FMADDs; thread-local reduction	Dense row-major; no pixel reads; pixels written once

Task	Time	Computation	Memory Pattern
Camera Processing Processes incoming camera depth image	5%	Bilateral filter; invalid depth rejection	Dense sequential accesses to depth image
Image Processing Pre-processes RGB-D image for tracking and mapping	18%	Generation of vertex map, normal map, and image intensity; image undistortion; pose transformation of old map	Globally dense; local stencil; layout change from RGB_RGB → RR_GG_BB
Pose Estimation Estimates 6DOF pose	28%	ICP; photometric error; geometric error	Photometric error is globally dense; others are globally sparse, locally dense
Surfel Prediction Calculates active surfels in current frame	38%	Vertex and fragment shaders	Globally sparse; locally dense
Map Fusion Updates map with new surfel information	11%	Vertex and fragment shaders	Globally sparse; locally dense

Task	Time	Computation	Memory Pattern
Normalization INT16 to FP32	7%	Element-wise FP32 division	Dense row-major
Encoding Sample to soundfield mapping	81%	$Y[j][i] = D \times X[j]$	Dense column-major
Summation HOA soundfield summation	11%	$Y[i][j] += X_k[i][j]\ \forall k$	Dense row-major

Task	Time	Sub-task	Computation	Memory Pattern
Rotation Soundfield rotation using pose	35%	**Psychoacoustic filter** Applies frequency-domain filter	FFT; frequency domain convolution; IFFT	Butterfly pattern for FFT/IFFT; dense row-major sequential accesses for convolution
		HOA rotation Rotates virtual channels	Transcendentals; FMADDs	Sparse column-major accesses; some temporal locality
Zoom Soundfield zoom using pose	5%	—	Linear algebra	Dense column-major sequential accesses
Binauralization HRTF application	60%	—	Identical to **psychoacoustic filter**	Identical to **psychoacoustic filter**

Ongoing Research with ILLIXR

- Representing heterogenous parallelism in software
- Automated selection and generation of accelerator hardware and software
- Accelerator communication interfaces
- Automated approximation selection
- Cross-component co-design
- QoE-driven scheduling
- Computation offload
- QoE metrics
- XR algorithms
- Eye tracking + Holograms [Sivasubramanium et al., Micro'21]

Ongoing Work on Testbed

- New components: spatial reprojection, hand tracking, foveated rendering, …
- Off-loading, streaming, multiparty XR (ILLIXR + CONIX ARENA)
- Broaden hardware/software platforms supported: Dockerization, Android, …
- Create and curate datasets and applications
- Incorporate research results

Outline

- Introduction
- ILLIXR Description
- Evaluation and Implications
- Ongoing Research with ILLIXR
- Ongoing Work on ILLIXR Testbed
- Consortium Organization and Goals

Consortium Goals

- Reference open-source testbed
  - Components and interfaces
  - Modular, extensible runtime
  - Telemetry
- Benchmarking methodology
  - Applications, datasets
  - System configurations
  - Metrics
- Build XR systems research and development community

Weekly open meetings and working groups on various topics

illiXR

Executive Committee

Responsibilities

- Oversight of overall direction of the project
- Executive committee membership
- Approving working groups and their chairs
- Approving working group recommendations and outcomes
- Setting project policies, including governance and membership agreements
- Outreach for the project

Current members

- Sarita Adve, University of Illinois, Chair
- Muhammad Huzaifa, University of Illinois, Vice-Chair
- Ameen Akel, Micron
- Sean Eilert, Micron
- Art Goldberg, Arm
- Rod Hooker, Meta
- Matt Horsnell, Arm
- Vasileios Laganakos, Arm
- David Luebke, Nvidia
- Maurizio Paganini, Meta
- Anthony Rowe, CMU

Advisory Board

Current members

- Ameen Akel
- Sarita Adve
- Vikram Adve
- Valeria Bertacco
- David Brooks
- Rod Hooker
- Matt Horsnell
- Muhammad Huzaifa
- Steve LaValle
- David Luebke
- Sharad Malik
- Franziska Roesner
- Anthony Rowe
- Gu-Yeon Wei

Providing overall strategic vision and direction of the consortium

Working Groups

- Any representative of a consortium member organization may attend working group meetings
- Consortium members have voting privileges in the working groups (one vote per organization)
- Only consortium member representatives may chair a working group
- Participants from non-member organizations are welcome to join a working group, subject to lightweight approval and without voting or chair privileges

Planned Working Groups

- Metrics
- Interfaces
- Components
- Computation offload
- Distributed XR
- Configurations and benchmark reporting
- Applications
- Datasets
- Infrastructure
- Research related

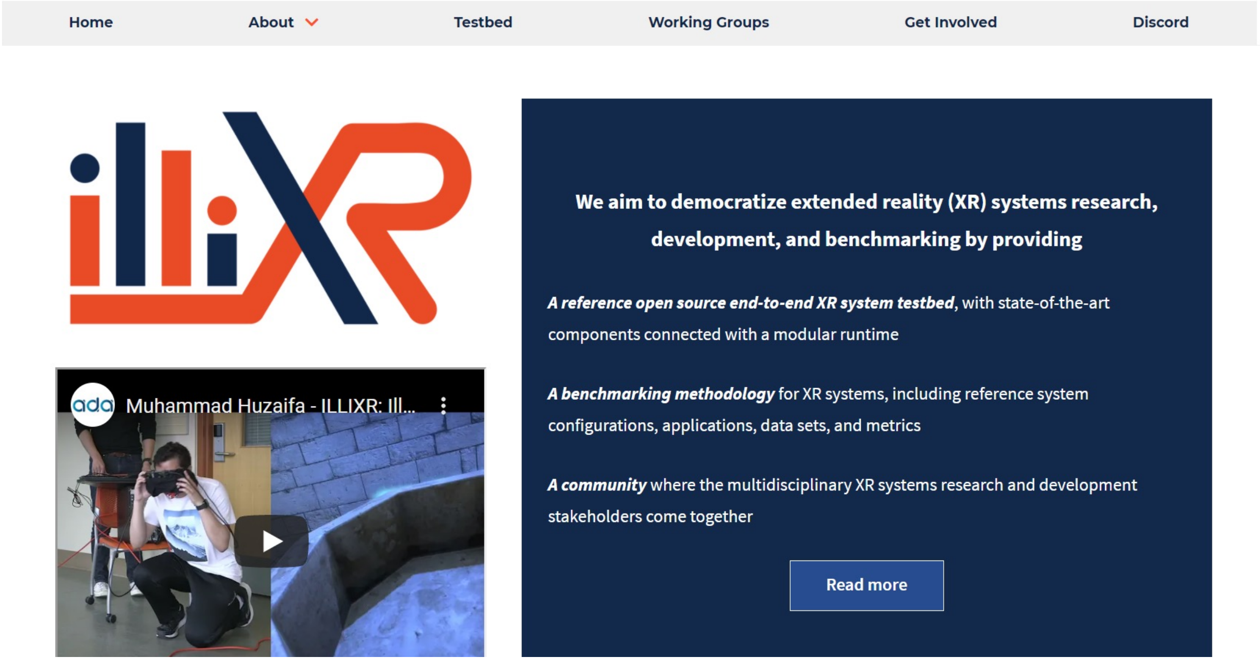
Next week

- Intro to ARENA (distributed XR), integrating w/ ILLIXR

Anthony Rowe, CMU

ILLIXR Consortium

illixr.org



Goals

- Reference open-source testbed
- Benchmarking methodology
- Build XR systems R&D community

Join us: illixr@cs.illinois.edu, illixr.org, discord, weekly meetings