# RAM-less Streaming Architecture for Efficient Spherical Panorama Systems

2017 GRO Project Theme: Mobile/Wearable Technology  
2017 GRO Project Sub-Theme: Ultra-low Power 360-degree Video Camera

PI: Robert LiKamWa  
Arizona State University   
School of Arts, Media & Engineering | School of Electrical, Computer & Energy Engineering  
PO Box 875802, Tempe, AZ 85287 | likamwa@asu.edu | 1-480-965-2686

# Project Summary

Recent advances in image sensor technology and lens designs have led to the emergence of portable spherical panorama systems, such as the Samsung Gear 360, capturing information to process into 360ºx180º images and videos. However, ***real-time, energy-efficient*** generation of ***high-resolution*** spherical panoramas remains a substantial challenge, as standard computational architectures are incapable of efficiently processing large amounts of data. Because energy consumption generates heat, creating imaging artifacts, e.g., lens warping, spherical panorama systems are constrained by a tight energy budget. This has led commercial implementations to offloading-based designs, in which stitching is done on the smartphone, and not in real-time. These implementations are incapable of scaling to large resolutions, due to limited and energy-expensive network bandwidth. Our proposed research will create designs that efficiently scale to ultra-high resolutions into the future based around streaming spherical panorama architectures that do not rely on DRAM, memory storage, or other expensive I/O during processing. We estimate that this will create a sub-watt architecture to generate and transmit 4K 360º video under 1 watt on the portable spherical capture device.

We propose two key research objectives: (i) Establish an efficient **RAM-less streaming architecture** to direct fisheye sensor data into equirectangular output; and (ii) Study the effects of **early in-sensor compression** to reduce the transmission of data across sensor interfaces. These research objectives, which we fully describe in the project description, not only form the basis of an efficient ultra-low-power architecture for 4K 360º video, but also present opportunities scale to larger input resolutions, not limited by memory bandwidth or network bandwidth. Over the course of the year-long project, we will implement our designs on a low-power FPGA to validate our assumptions and test the limits of scalability.

Success in the year-long initial stage of this project will translate to further research opportunities along a three-year trajectory. In the second year, we plan to investigate how real-time video analytics can guide regional compression of the 360º video, satisfying network bandwidth while maintaining perceptual quality and visual saliency. In the third year, we plan to investigate the implementation of scalable designs in chip development, exploring architectural limitations in the bottlenecks of image sensor bandwidth and networking bandwidth. Over the three-year project, the overarching goal will lead to solutions for efficient system architectures for high-resolution 360º in real-time, high quality, and high spatiotemporal resolution.

*Keywords: image processing; high-resolution imaging; compression; energy-efficiency; streaming architecture*

# Description of Project

Project Duration: 09/01/2017 – 08/31/2018

We propose to develop a specialized streaming hardware architecture without random-access memory to reduce the energy consumption and improve the performance of portable spherical panoramic capture devices. The goal of the project is to enable system-level energy efficiency, enabling high spatiotemporal capture resolutions, despite constrained energy and heat requirements. Our work leans on key observations that DRAM and storage consume proportionally more power than computational operations; DRAM access consumes 60-85% of the energy for specialized image processing architectures [1]. Furthermore, continuous I/O to DRAM resources limits the scalability to high-resolution for imaging architectures. The research project will build insights towards effective RAM-less streaming architectures and prototype FPGA-based implementations and evaluations of a specialized accelerator for spherical panoramic capture.

Thus, over the course of the proposed year-long project, we target the following research objectives:

* Establish an efficient streaming architecture to direct fisheye sensor data into equirectangular output
  + Eliminate dependence on DRAM and storage for all panoramic processing stages
  + Design efficient parallelism for feature detection, correspondence, projection mapping and blending
* Study the effects of early in-sensor compression to reduce the transmission of data across sensor interfaces
  + Develop projection algorithms to map compressed input into compressed output without decompression
  + Investigate region-based compression ratios to reduce data transfer while guaranteeing output fidelity

Such objectives will pave the way for energy-efficient and high-performance spherical capture by: (i) reducing the data movement within the capture device, (ii) reducing the required network bandwidth to the smartphone, and (iii) reducing the computational burden on the smartphone.

*Samsung-GRO-Figures/1.pdf  
Fig. 1: Conventional system architectures vs proposed RAM-less streaming architecture*

## Significance of Research

* *

*Fig. 2: Left: Pair of spherical fisheye images. Right: Equirectangular-projected output panorama*

Successful completion of the project’s research objectives will significantly transform the capabilities of future panoramic capture devices. Specifically, compared to existing system architectures in industry and in academic literature, the proposed research will be the first to perform panoramic stitching operations without the use of DRAM, memory cards, or other I/O resources for processing. The unprecedented energy efficiency will create new opportunities for on-device spherical fisheye processing at ultra-low power and high spatiotemporal resolution. Our conservative estimates declare ***4K 360º video streaming under*** ***1 watt on the portable spherical capture device****,* which compares favorably to commercial Gear 360 solutions, at 2x higher streaming resolution with improved quality and 2x lower power consumption, and performing the panorama generation entirely on the capture device and streaming to the mobile device.

## Team Experience

Leveraging experience on low power computer vision system architectures at the application level [2], operating systems level [3], and sensor architecture level [4], the team is poised to pursue novel architecture for streaming panoramic capture. In 2015, PI LiKamWa interned with the System-on-Chip team of the Samsung Mobile Processor Innovation Lab over a 4-month period, building mobile hardware architecture for computational photography.

# Background

**Mapping fisheye capture data to equirectangular projection:** Fisheye lenses allow image sensors to capture images within an ultra-wide hemispheric field of view. With two fisheye-lensed sensors that capture complementary fields of view each of over 180º, the pair of captured images can be processed to achieve over a spherical 360ºx180º area, as shown in Figure 2. The equirectangular projection format is a common format for 360ºx180º images, allowing remapping to other projections for convenient viewing. To create equirectangular images, the paired fisheye capture data undergoes multiple stages:

1. ***Projection Mapping:*** The equirectangular image is populated by sourcing image pixels from the fisheye images along a (spherical coordinate → polar coordinate) projection map. As projected pixel coordinates typically fall between integer pixel coordinates, the algorithm typically either pulls a nearest-neighbor pixel or a bilinear combination of a neighborhood of pixels.
2. ***Correspondence:*** As the two fisheye cameras do not precisely occupy the same point in space, objects at the edges of fisheye images appear in different positions in the images, dependent on their distances from the camera. This phenomenon is called the parallax effect. To ensure that objects appear properly, a correspondence algorithm identifies matching visual features across image pairs, warping the projection to reduce object seams in the image.
3. ***Blending:*** Even after projection and correspondence suggest image overlay coordinates, intensity variations from misalignments still occur between the two projected images at the stitching boundary. The blending stage combines the images through a weighted sum of pixel values to generate a seamless 360º image with a smooth transition.
4. ***Compression:*** To reduce the bandwidth at the capture, networking, or storage interface, images can be compressed into representations that use smaller file sizes. Lossy compression schemes, e.g., JPEG/MPEG, allow dramatic reductions in file size by discarding information that is considered to be perceptibly irrelevant.

**Conventional offload-based system architecture:** Energy consumption creates a hard limit, or we will heat sensors and lenses, which degrades the quality of image capture. A conventional memory-based computational architecture on the capture device would consume prohibitively high energy consumption per frame, bottlenecked by memory operations on the many pixels of images with high spatiotemporal resolution. This has motivated commercial system implementations to leverage networked devices, e.g., smartphones for the Gear 360, to perform the processing stages, while the portable capture device is only responsible for capturing and transmitting lightly-compressed images. The offload-based system architecture is illustrated in Figure 1a. Unfortunately, ***the scalability of offload-based solutions is limited by the network interface***. Networking for hiqh-quality panoramas is already prohibitively energy-expensive (500 mJ per 15MP frame), and slow (2 frames per second), requiring a networking power consumption of over 1 watt to achieve required bandwidth. An additional burden to the network interface and other system resources is that ***input fisheye images must be captured and transferred with high quality.*** Resolution and compression artifacts in the input image will greatly deteriorate the perceived quality of the output image. Thus, image compression on the capture device is limited to low compression ratios, as fisheye capture data, further straining energy-efficiency. Finally, ***smartphone SoC architectures are ill-equipped for panoramic stitching***, due to limited parallelism and a heavy dependence on memory loads and stores. This causes an overhead of roughly 4 seconds to generate a spherical panorama on a Galaxy S7.

On a specialized device, such as the Gear 360, it is possible to incorporate application-specific hardware to enhance the device’s features, e.g., Figure 1b. However, such hardware must be designed carefully; conventional architectures that rely on DRAM and storage consume substantial energy consumption, heating up the device, and draining the battery. Our proposed RAM-less streaming architecture, illustrated in Figure 1c, aims to perform the processing stages on the spherical capture device itself at sub-watt power efficiency. Generating the equirectangular image before transmission allows larger compression ratios for the transmitted output image while preserving perceived quality. Finally, the computational load on the networked smartphone is also reduced and high quality images are readily available for real-time streaming, i.e., without processing latency.

## Related Work

Although previous architectures have been incapable of meeting energy-efficiency requirements for high-resolution spherical panorama capture, much work has investigated hardware-accelerated image processing. Feature-detection and feature matching implementations [5] [6] [7] have leveraged streaming-based FPGA-based hardware accelerators to perform computation intensive tasks with high performance. We can use insights from these works, e.g., line buffer structures, to accelerate design exploration. However, our proposed research pursues challenges and opportunities to efficiently integrate these solutions into a low-power panorama capture design.

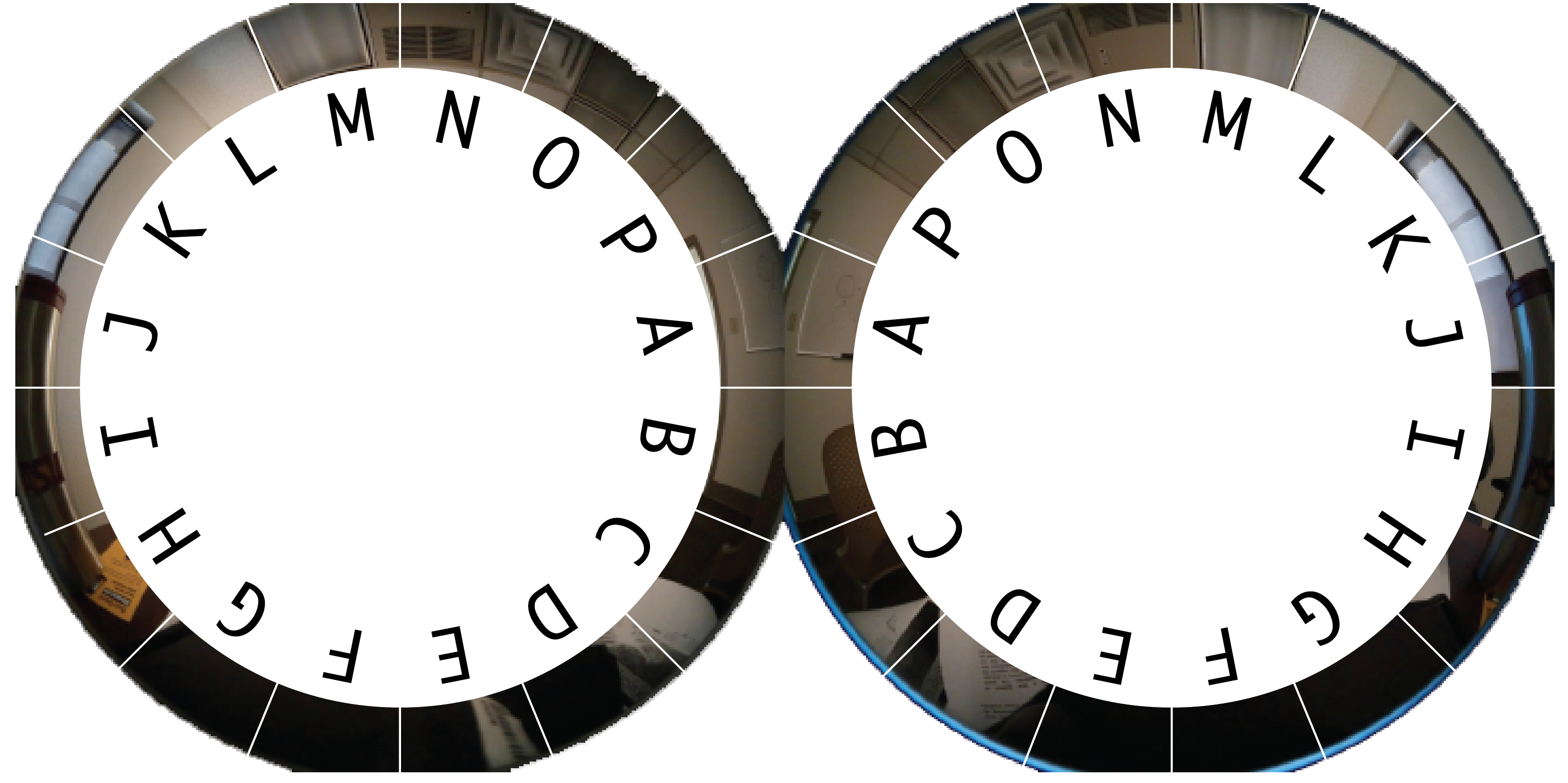
In doing so, we must avoid the scalability pitfalls of standard DRAM-based implementations. For example, [6] proposes a real-time streaming architecture for panorama stitching, but their design’s performance and efficiency is bandwidth-limited as they heavily rely on DRAM for image buffering. (Notably, feature-based correspondence alignment is absent from their work.) The DRAM bottleneck on energy-efficiency has been laid out in previous works, including Darkroom [1], which finds that DRAM consumes 60-85% of total power in image processing architectures. Based on this insight, we propose a RAM-less streaming architecture to satisfy energy-efficiency and performance requirements.

# Research Plan and Technical Approach

Towards a streaming panoramic capture architecture, we propose to execute a two-thrust research plan to:  
(i) design ***streaming hardware*** to efficiently perform projection mapping ***without random-access memory***; and   
(ii) design systems and algorithms for early ***in-sensor compression***with ***region-based compression quality***.

By reducing I/O-dependence and dataflow bandwidth, these thrusts will form the basis of a hardware accelerator towards ultra-low power panoramic capture with high spatiotemporal resolution.

## Thrust 1: Streaming Architecture for Spherical Panoramic Projection Mapping



*Fig. 3: Feature correspondence is regionally paired across fisheye images, e.g., a correspondence pair appears in J. both images. Consequently, each correspondence hardware block needs to work only on small regions of pixels.*

Towards an ultra-low power system architecture, a chief goal is to eliminate dependence on DRAM and file access while achieving the tasks necessary to generate high-quality equirectangular images from fisheye sensor data. Our key insight is that the processing stages are highly parallelizable, with high spatial locality to both input and output data. We can leverage this insight to design streaming image correspondence operations, which can be used to guide streaming projection and blend operations.

### Streaming feature detection and correspondence

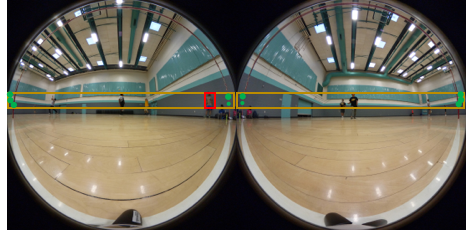
We propose an architecture for feature detection and correspondence that makes effective use of spatial locality towards line buffers and divide-and-conquer, allowing an independence from random-access memory. We expect that features detection and correspondence operations can be derived from standard corner-based algorithms, but propose to study novel hardware designs for area-efficiency and energy-efficiency.

As shown in Figure 3, we observe that relevant feature comparison and matching across pairs of frames occur near the same frame rows, isolated to overlapping regions of interest. As feature detection and description is a spatially-local operation, we only have to store pixel data from a limited number of rows. Through our proposed research, we will study the tradeoff between feature correspondence quality and number of pixel rows to store. This will allow a memory-less architecture, while minimizing line buffer hardware requirements.

The hardware resources and energy consumption of feature detection and matching hardware units has a quadratic dependence (*N2*) on the size of the input region (*N*) – we assume that features are evenly distributed across input regions, and the hardware must account for all combinations of feature pairs across the two images. Fortunately, while depth-disparity causes features to appear in slightly different pixel locations, we can declare paired regions in each fisheye frame where features are likely to match. We can leverage this observation to reduce hardware area by reducing the size of the input region to (*N/k*). This can allow us a variety of design optimizations, such as building *k* parallel feature detection/matching units with (*N/k*) input region size (*k*(*N/k*)2 = *N2/k* < *N*2). Through our proposed research, we will study these and other divide-and-conquer decisions, as well as studying the appropriate size of input region to minimize hardware area for feature detection.

While we plan to execute the feature detection and matching directly on fisheye-captured pixels, we may require a projection to a rectangular format for sufficient feature matching. While this would be a vital implementation detail for correctness of operation, such projection is not central to our work. Instead, our research contributions target the design of area-efficient and energy-efficient feature detection and correspondence hardware by reducing input dependencies.

Equirectangular Output:  
**

Input Fisheye Frames:  


*Fig. 4: Generating an equirectangular output patch (top) requires streaming input dependencies (bottom), including mapped input region (red square) and nearby feature correspondences (green dots). Potential dependencies (yellow region) must be temporarily held in line buffers.*

### Streaming projection and blend operations

After obtaining the image correspondences, we will need to efficiently guide and combine input fisheye image pixels to their outputs in the spherical panorama format. To form a spherical panorama, pixels are interpolated from an input fisheye images from indices sourced from a projection map. Spherical panorama formats typically use a equirectangular projection, which has equally distributed resolution along spherical angular axes, and can be re-projected to various viewpoints. The streaming hardware will mildly adjust the projection map based on image correspondence information. After projecting both input images to common output plane, the blending operations will use a weighted sum to combine pixel values. The blend weights can also be scaled to correct for optical intensity variation artifacts of fisheye captures. Thus, after identifying nearby correspondences, the system will have sufficient information to project and blend fisheye input pixels into their output destinations in an equirectangular output. As shown in Figure 4, we expect that a blended equirectangular output pixel depends on source frame pixels and correspondence derived from pixels.

Our proposed research will study how much correspondence information is required to reliably generate equirectangular input. Reducing dependency on correspondence information will allow smaller pixel buffer sizes, and thus, smaller hardware area. However, insufficient correspondence information can result in inaccurate output images with seaming artifacts. This will lead to a key question: ***How many buffered pixels are required for projection mapping to be reliably applied?***The answer to this question has substantial implication on area-efficiency, performance, and energy-efficiency of the hardware unit. We will study the data dependencies and its relationship to pixel location in the equirectangular output. If the projection and blend operations require correspondence information derived from spatially distant input pixels, it may be possible to leverage correspondence information from the previous frame pairs. We will study captured fisheye frame sequences to explore this possibility.

Thus, Thrust 1 seeks to explore research contributions related to energy-efficient and area-efficient streaming hardware to generate equirectangular data from captured fisheye frame pairs. As we eliminate dependence on random-access memory, we will pursue research challenges and opportunities related to streaming data dependencies. To explore and validate the research contributions of Thrust 1, we will use an FPGA board to design the streaming hardware for feature detection, image correspondence, and projection and blending operations. This will allow us to characterize relationships between many factors, including resolution, quality, energy consumption, and hardware resources.

## Thrust 2: System architecture for early image compression

To further reduce the energy consumption of the system architecture proposed in Thrust 1, we target ***the image sensor physical interface*** as a bottleneck to energy efficiency. As temperature requirements force a substantial distance between image sensors and processing units, sensor data transactions are notoriously energy-expensive. At full input resolution of 15 MP at 30 frames per second, using a typical Low-Voltage Differential Signaling (LVDS) physical interface consumes 2W to satisfy the >3 Gbps bandwidth, e.g., [8]. This is prohibitively high, and the Gear 360 typically is constrained to capture video at one half of this available resolution, while still consuming multiple watts of average power consumption. We explore the use of in-sensor compression to assist in an energy reduction, as compared in Table 1. Existing hardware solutions for JPEG and MPEG compression are plentiful and sufficient, dropping data bitrates by substantial compression ratios (e.g., 8:1, 23:1, 46:1, etc.) and some image sensors integrate real-time JPEG encoders into their package [9]. As shown in Table 1, this can have dramatic savings on interface power consumption. Using this as our basis, Thrust 2 proposes system-oriented research for placement and utilization of compression hardware close to the sensor. Thrust 2 aims to approach research objectives of (i) processing without decompression and (ii) region-based compression quality.

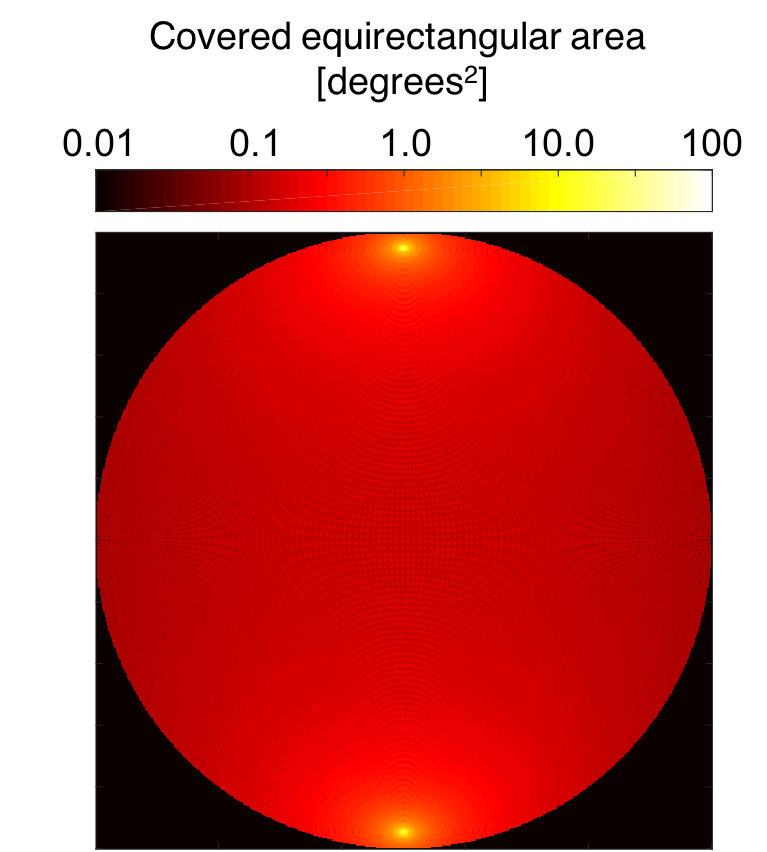
Table : Compression vs. Filesize, bandwidth, and Interface power consumption

|  |  |  |  |
| --- | --- | --- | --- |
|  | File Size | Bandwidth at (30 fps) | LVDS Interface  Power Consumption |
| Raw Bayer 15 MP | 15 MB | 3.6 Gbps | 2 W |
| High quality JPEG (CR=8:1) | 1.9 MB | 456 Mbps | 450 mW |
| Med. Quality JPEG (CR=23:1) | 652 KB | 157 Mbps | 350 mW |
| Low quality JPEG (CR=46:1) | 326 KB | 78.2 Mbps | 100 mW |

### Processing Without Decompression

While compression allows for reduced data rates over energy-expensive interfaces, the need for decompression hardware prior to the processing unit reduces the energy-efficiency gains and adds to the hardware complexity. Thus, we seek to design our streaming processing hardware with complete independence from decompression operations. This will require adaptation of the feature detection, image correspondence, and projection and blending operations to operate on compressed blocks.

JPEG encoders buffer and compress 8x8 blocks of pixels into 8x8 blocks of discrete cosine transform coefficients that are quantized and run-length encoded into compression vectors, which are read out sequentially. Previous works have explored algorithmic operations on the compressed discrete cosine transform representation of the data [10]. It is possible to run several operations on fully-encoded vectors, including multiplying vectors by constant values, adding constant value to a vector, or performing pixel-wise multiplications across blocks. Our research challenge is to leverage these and other compressed arithmetic techniques to avoid the overhead of decompression and the dependence on random-access memory. Towards area-efficiency and energy-efficiency, we also propose to study algorithmic dependencies between the number of buffered JPEG vectors and the quality of output.



*Fig. 5: Covered equirectangular area by input region. Certain input regions, e.g., north and south poles, map to large areas in the equirectangular projection. These regions have greater sensitivity to quality, e.g., compression ratio.*

### Region-based compression quality

In generating an equirectangular output image, the projection stretches some small input pixel regions to cover large output regions, as illustrated in Figure 5. Thus, some regions of the spherical capture frames are more quality-sensitive than others. Intuitively, we can vary the compression quality of the JPEG blocks depending on the location of the block in the input frame. Increasing compression ratios where possible will allow a reduction in data rate across the sensor interface, and thus a further improvement in energy efficiency. Thus, through our proposed research, we plan to study the dependence of equirectangular output quality on input frame compression quality on a region-basis, using MATLAB simulations to evaluate quality of output using standard image processing metrics, e.g., structural similarity. We will then adapt our streaming hardware architecture to leverage region-based compression quality while ensuring reasonable quality output.

As with Thrust 1, we will evaluate Thrust 2 through a hardware design and implementation on an FPGA to evaluate the implications of compression ratios against various input frame resolutions, output image resolutions, and image stitching quality metrics. Thrust 2’s research objectives will result in efficient fisheye frame transmission and processing through reducing the data transmission from the sensor to the processing unit.

# Estimated Power and Performance

Our proposed research thrusts aim to reduce data traffic across the image sensor interface, the memory interfaces, and the network interfaces. Here we issue a comparative estimate to current system architectures.

The current system architecture of the Samsung Gear 360 (2017) entails two 8.4 MP camera modules, a MicroSD card slot, WiFi and Bluetooth chipset, and microprocessor. From a user perspective, the Gear 360 captures spherical fisheye images, storing compressed 15 MP .jpg images and 4096 x 2048 .mp4 videos on the SD card. These files can be transmitted to the user’s smartphone over WiFi for processing into equirectangular formats. Alternatively, image data can be streamed from the Gear 360 to the smartphone in real-time. It is evident that in both cases, the smartphone performs all of the processing to unwarp and stitch the input fisheye images. The Gear 360 consumes 4 W of power consumption when recording 2560x1280@30fps.

Meanwhile, a conventional RAM-based specialized architecture for spherical panoramic projection would be unable to perform 4K 360º processing on the capture device without overloading power requirements. Based on scaling an existing architecture in the literature [5] to 4K 360º systems, its resource requirements and memory bandwidth needs require a high-end FPGA architecture, e.g., Virtex7, consuming 11 W of power consumption, of which 6.6 W is due to memory access.

As shown in Table 2, our proposed RAM-less streaming architecture requires substantially fewer resources, due to its hardware reuse for streaming operations. As described in Thrust 1, each stage of processing, e.g., feature detection, operates on N rows of pixel data, allowing substantial division of hardware resources. However, our architecture will also require hardware duplication to achieve streaming bandwidth. We conservatively estimate that altogether, our proposed architecture will reduce LUT count by 16x and fully eliminate DRAM dependence, at the expense of line buffers to store rows of data. Our proposed architecture will fit on an ultra-low power FPGA, such as the IGLOO2, which satisfies the core frequency requirements of our future designs.

Table : REsource requirements for panoramic architectures

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Number of LUTs | Registers | Line Buffers | DRAM |
| Conventional RAM-based architecture | 20k | 15k | - | 250 Mb |
| Proposed RAM-less Streaming Architecture | 1.25k | 15k | 4.7 Mb | - |

Our estimates of running our proposed architecture on an IGLOO2 FPGA suggest a computational power consumption of 100 mW, a sensor capture power consumption of 350 mW, and a networking power consumption of 400 mW, leading to a sub-watt total power consumption for streaming panoramic images in real time over a network. This amounts to less than half of the power consumption of the current Gear 360 system, while removing dependence on the mobile phone. Studying the actual resource usage is a research challenge that will be approached through a full implementation of our designs on an FPGA. Furthermore, we recognize that streaming bandwidth is a standing issue with low-power architecture designs, but hope to address it outside of the scope of this proposal.

# Milestones

Through our year-long research plan, we propose to execute a combination of simulations and hardware design tasks towards an efficient architecture for spherical panorama systems with no random-access memory dependence.

### Thrust 1

* Month 1: Simulate algorithms for streaming feature detection and correspondence (MATLAB)
* Month 2: Simulate algorithms for streaming projection mapping (MATLAB)
* Month 3: Design architecture for streaming feature detection and correspondence (HDL)
* Month 4: Design architecture for streaming projection mapping (HDL)
* Month 5: Implement architecture for streaming input from sensors to FPGA (HDL)
* Month 6: Evaluate Power, Performance, Quality

### Thrust 2

* Month 7: Simulate algorithms for streaming projection mapping on compressed data (MATLAB)
* Month 8: Simulate algorithms for streaming feature detection and correspondences on compressed data (MATLAB)
* Month 9: Design architecture for streaming projection mapping on compressed data (HDL)
* Month 10: Design architecture for streaming projection mapping on compressed data (HDL)
* Month 11: Investigate and integrate region-based sensitivity for compression ratios (MATLAB, HDL)
* Month 12: Evaluate Power, Performance, Quality

# Expected Outcomes and Results

Successful completion of the work carries broad implications, both commercial and academic. By reducing the dependence on random access memory transactions, memory card transactions, and network transactions, the proposed architectural contributions will break down barriers of energy consumption and I/O latency towards scalability to high resolutions and high frame rates. Through an FPGA-based design, we conservatively estimate an achievement of ***4K video streaming under 1 watt*** ***on the portable spherical capture device***, and an ***eliminated computational burden on the smartphone***. A further order of magnitude of power savings can come from ASIC implementation. This tangible outcome will revolutionize the adoption of portable spherical capture devices as they can be used longer with sustained battery life and no overheating concerns of both the smartphone or the capture device. Furthermore, as the portable spherical capture device can create its own panoramas, users can immediate access previously captured photos and videos without waiting for conversion processing time.

We plan to disseminate our research results through academic journals and conferences. Portable spherical capture is a largely untapped opportunity in the architecture community, and new results and intuitions will broadly influence the development of other sensor capture systems. In particular, the research contributions of the proposed energy-efficient streaming architecture will foster further research and development towards reducing I/O transactions across system interfaces and reducing memory transactions within a streaming architecture.

# BUDGET

We request $99,927 to fund two full-time M.S./Ph.D. students under research assistantships (including salary and tuition remission) over the course of 12-months. Budgeted costs include Facilities and Administrative overhead. Supporting equipment, e.g., FPGA, DSP boards, sensors, will be purchased from the PI’s existing research funds.

|  |  |  |
| --- | --- | --- |
| Cost Categories | Period 1  9/1/2017 - 8/31/2018 | Cumulative |
| Total Salary, Wages and ERE: | $39,496 | $39,496 |
| Other Direct Costs: | $33,692 | $33,692 |
| 8. Tuition Remission | $33,692 | $33,692 |
| Direct Costs: | $73,188 | $73,188 |
| Indirect Costs (67.7%): | $26,739 | $26,739 |
| Total Direct and Indirect Costs: | $99,927 | $99,927 |

# REFERENCES

|  |  |
| --- | --- |
| [1] | J. Hegarty, J. Brunhaver, Z. DeVito, J. Ragan-Kelley, N. Cohen, S. Bell, A. Vasilyev, M. Horowitz and P. Hanrahan, "Darkroom: compiling high-level image processing code into hardware pipelines," *ACM Transactions on Graphics,* vol. 33, no. 4, 2014. |
| [2] | R. LiKamWa and L. Zhong, "Starfish: Efficient concurrency support for computer vision applications," in *Proc. ACM Annual International Conference on Mobile Systems, Applications, and Services (MobiSys)*, 2015. |
| [3] | R. LiKamWa, B. Priyantha, M. Philipose, L. Zhong and P. Bahl, "Energy characterization and optimization of image sensing toward continuous mobile vision," in *Proc. ACM Annual International Conference on Mobile Systems, Applications, and Services (MobiSys)*, 2013. |
| [4] | R. LiKamWa, Y. Hou, J. Gao, M. Polansky and L. Zhong, "RedEye: analog ConvNet image sensor architecture for continuous mobile vision," in *Proc. ACM/IEEE International Symposium on Computer Architecture (ISCA)*, 2016. |
| [5] | J. Wang, S. Zhong, L. Yan and Z. Cao, "An embedded system-on-chip architecture for real-time visual detection and matching," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 24, no. 3, pp. 525-538, 2014. |
| [6] | Y. Xu, Q. Zhou, L. Gong, M. Zhu, X. Ding and R. K. F. Teng, "High-Speed Simultaneous Image Distortion Correction Transformations for a Multicamera Cylindrical Panorama Real-time Video System Using FPGA," in *IEEE Transactions on Circuits and Systems for Video Technology*, 2014. |
| [7] | L. Yao, H. Feng, Y. Zhu, Z. Jiang, D. Zhao and W. Feng, "An Architecture of Optimised SIFT Feature Detection for an FPGA Implementation of an Image Matcher," in *Proc. International Conference on Field-Programmable Technology*, 2009. |
| [8] | T. Instruments, "DS25BR120 Datasheet: 3.125 Gbps LVDS Buffer with Transmit Pre-Emphasis," 2013. [Online]. Available: http://www.ti.com/lit/ds/symlink/ds25br120.pdf. |
| [9] | OnSemi, "MT9D131 Datasheet: 1/3.2-Inch System‐On‐A‐Chip (SOC) CMOS Digital Image Sensor," Nov 2016. [Online]. Available: https://www.onsemi.com/pub/Collateral/MT9D131-D.PDF. |
| [10] | B. C. Smith and L. A. Rowe, "Algorithms for maniplating compressed images," *IEEE Computer Graphics and Applications,* vol. 13, no. 5, pp. 34-42, 1993. |
| [11] | W. Zhu, L. Liu, G. Y. S. Jiang and S. Wei, "A 135-frames/s 1080p 87.5-mW Binary-Descriptor-Based Image Feature Extraction Accelerator.," in *IEEE Transactions on Circuits and Systems for Video Technology*, 2016. |
| [12] | F.-C. Huang, S.-Y. Huang, J.-W. Ker and Y.-C. Chen, "High-Performance SIFT Hardware Accelerator for Real-Time Image Feature Extraction," in *IEEE Transactions on Circuits and Systems for Video Technology*, 2012. |

# APPENDIX A: FACILITIES AND EQUIPMENT

The Mobile Systems Research Studio, directed by PI Robert LiKamWa, focuses on software and hardware systems for sensing and actuation on mobile devices. Built on operating systems, computer architecture, and machine learning research, we create low-level system designs to balance the efficiency and performance of sensing, processing, and acting on data. The Mobile Systems Research Studio operates across two departments at Arizona State University: the School of Arts, Media and Engineering (AME) and the School of Electrical, Computer and Energy Engineering (ECEE).

Through the generous support of the National Science Foundation, Microsemi Corporation, and NVIDIA Corporation, the research studio has grown to support 2 PhD-track students and 6 undergraduate students. Our equipment includes a set of cameras, development kits, mobile devices, and computing workstations.

Cameras



*Panoramic capture of the   
Mobile Systems Research Studio laboratory space*

* Gear 360 (2016)
* Gear 360 (2017)
* FLIR Imaging camera

Development Kits

* Microsemi FPGA Advanced Development Kit
* Jetson TK1
* Jetson TX2

Mobile Devices



*Motion Analysis Lab/Intelligent Stage   
immersive motion capture and visualization space*

* Samsung Galaxy S8
* Microsoft HoloLens
* Google Nexus 5X
* (20x) NVIDIA Shield Tablet K1

Computing Workstations

* (2x) Dell Precision Tower 5810
* Dell XPS 15”
* MacBook Pro with Touch Bar

In addition to its 10-person research laboratory space, The Mobile Systems Research Studio has access to the Motion Analysis Lab/Intelligent Stage[[1]](#footnote-1), a research lab and performance space dedicated to motion analysis and interactive, multimodal feedback development. The lab has two sections, each with two independent high end and standard motion-capture systems. This environment could be used to test and visualize effective scenes for panoramic capture. The facility has controllable projection systems, lighting, and a pressure-sensitive floor that are integrated to the motion-capture system. This initiative was awarded an NSF infrastructure grant in 2005, which allowed for its continued expansion.

The School or Arts, Media and Engineering also has an electronics workshop for student use for prototyping, testing board-level electronics. This includes oscilloscopes and soldering stations.

# Appendix B: Curriculum Vitae of Robert LiKamWa, PI

# 

# Robert LiKamWa

Assistant Professor  
Arizona State University  
School of Electrical, Computer and Energy Engineering  
School of Arts, Media and Engineering·

<http://roblkw.com>

[likamwa@asu.edu](mailto:likamwa@asu.edu)

Research Interests

Low power sensing on mobile devices.  
Mobile computing; Operating systems; Energy management; Context-awareness; Vision systems;

Education

Ph.D., under Dr. Lin Zhong, Electrical & Computer Engineering, Rice University — August 2016  
 Thesis: Continuous Mobile Vision: Rethinking the Vision Sensing Pipeline for Energy Efficiency  
M.S., under Dr. Lin Zhong, Electrical & Computer Engineering, Rice University — May 2012  
B.S.E.E. (cum laude), Electrical & Computer Engineering, Cum Laude, Rice University — May 2010  
Minor, Computational and Applied Math, Rice University — May 2010

Honors/Awards

* Best Presenter Award at Ph.D. Forum @ MobiSys (6/2014)
* Best Paper Award at MobiSys (6/2013)
* Best Paper Award at PhoneSense @ SenSys (11/2011)

Selected Conference & Workshop Papers

“RedEye: Analog ConvNet Image Sensor Architecture for Continuous Mobile Vision”   
 Robert LiKamWa, Yunhui Hou, Julian Gao, Mia Polansky, Lin Zhong   
 ISCA ’16: Proc. of the 43rd int’l symposium on computer architecture

“Starfish: Efficient concurrency support for computer vision applications”   
 Robert LiKamWa, Lin Zhong   
 MobiSys '15: Proc. of the 13th annual int’l conf. on mobile systems, applications, and services

“Energy proportional image sensors for continuous mobile vision” ***(Best Paper Award)***  
Robert LiKamWa, Bodhi Priyantha, Matthai Philipose, Lin Zhong, Paramvir Bahl  
 MobiSys '13: Proc. of the 11th annual int’l conf. on mobile systems, applications, and services

“MoodScope: Building a mood sensor from smartphone usage patterns”  
 Robert LiKamWa, Yunxin Liu, Nicholas D. Lane, Lin Zhong   
 MobiSys '13: Proc. of the 11th annual int’l conf. on mobile systems, applications, and services

“Reflex: using low-power processors in smartphones without knowing them”  
 Felix Xiaozhu Lin, Zhen Wang, Robert LiKamWa, Lin Zhong   
 ASPLOS ‘12: Proc. of the 17th int’l conf. on arch. support for programming languages and op. systems

Other Papers & Posters

Poster: “Temperature-driven task migration to balance energy efficiency and thermal noise of sensor processing workloads” ***(Best Poster Award)***  
 Venkatesh Kodukula, Robert LiKamWa  
 ACM HotMobile ’17: Workshop on Mobile Computing Systems and Applications

“Draining our Glass: An energy and heat characterization of Google Glass”  
 Robert LiKamWa, Zhen Wang, Aaron Carroll, Felix Xiaozhu Lin, Lin Zhong  
 APSys '14: Proc. of 5th Asia-Pacific workshop on systems

Invited Paper: “Efficient image processing for continuous mobile vision”  
 Robert LiKamWa, Yunhui Hou, Peter Washington, Lin Zhong  
 SID Display Week, Imaging Technologies and Applications

“Efficient image processing for continuous mobile vision” ***(Best Presentation Award)***  
Robert LiKamWa  
 MobiSys PhD Forum '14: Proc. of the MobiSys 2014 PhD Forum Workshop

Poster: “Retrofitting computer vision libraries for concurrent support on mobile devices”  
 Robert LiKamWa, Lin Zhong  
 MobiCom '14: Proc. of the 20th annual int’l conf. on mobile computing and networking

“Styrofoam: A tightly packed coding scheme for camera-based visible light communication”  
 David Ramirez, Robert LiKamWa, Jason Holloway  
 VLCS @ MobiCom '14: Proc. of the 1st ACM workshop. on Visible Light Communication Systems

“Can your smartphone infer your mood?” ***(Best Paper Award)***  
Robert LiKamWa, Yunxin Liu, Nicholas D. Lane, Lin Zhong  
 PhoneSense '11: Proc. of the second int’l workshop on sensing applications on mobile phones

Demo: “SUAVE: Sensor-based User-Aware Viewing Enhancement for mobile device displays”  
 Robert LiKamWa, Lin Zhong  
 UIST '11: Adjunct Proc. of the 24th annual ACM symp. on user interface software and technology

Ph.D. Dissertation: Rethinking the Sensing Pipeline for Continuous Mobile Vision

Vision Library Support for Efficient Concurrency

Rice University with Lin Zhong

The future of continuous mobile vision envisions multiple computer vision-based applications concurrently running without user engagement. However, running a multitude of applications calling resource-hungry vision algorithms will strain a wearable device with limited resources. We design Starfish, a split-process software system that enables applications to share redundant computations, reducing the overhead of computer vision function calls. Starfish splits the vision library from an application into a separate process and securely and efficiently caches library function calls, improving the performance and efficiency of concurrent vision tasks.

*Full Paper at MobiSys 2015*

Image Sensor Control for Energy-Proportional Capture

Microsoft Research Redmond with Matthai Philipose, Bodhi Priyantha, Victor Bahl  
Rice University with Lin Zhong

A major hurdle to frequently performing mobile computer vision tasks is the high power consumption of image sensing. We experimentally and analytically characterize the energy consumption of CMOS image sensors. With this knowledge, we create system-level driver-controlled power optimizations for capturing images and video relevant to machine vision tasks.

*Best Paper at MobiSys 2013*

Vision Sensor Hardware Architecture for Deep Analog Processing

Rice University with Yunhui Hou, Yuan Gao (undergrad), Mia Polansky (undergrad), Dr. Lin Zhong

Sensor readout is a key bottleneck to vision energy efficiency. We design a novel sensor that performs early analog processing to reduce the sensor readout. We discover that vision workloads map well to analog hardware, due to robustness to noise and repetitiveness of execution. We then provide an architecture that allows deep analog vision processing with fixed complexity. Through cyclic reuse of modules in a column-based topology, we limit the design complexity of the chip while provisioning for iterative execution before readout. As opposed to prior vision sensors and neuromorphic processors discussed in Related Work, deep processing allows us to target readout as the fundamental boundary to sensor and system efficiency.

*Full Paper at ISCA 2016*

Transparent Compiler/Runtime for Heterogeneous CPU Architecture (co-author)

Rice University with Felix Xiaozhu Lin (lead), Zhen Wang & Lin Zhong

Reflex is our suite of compiler and runtime support tools for efficient smartphone sensing. A distributed heterogeneous architecture promises energy-efficiency to sensing applications, but programming software on such an architecture is difficult. Reflex not only manages deployment and execution of code for heterogeneous resources, but also creates a software shared memory among distributed code.

*Full Paper at ASPLOS 2012*

Energy and Heat Characterization of Wearable Systems

Rice University with Zhen Wang, Aaron Carroll, Felix Xiaozhu Lin, Lin Zhong

The Google Glass small form factor hampers its potential: (1) battery size, and therefore lifetime, is limited by a need for the device to be lightweight, and (2) high-power processing leads to significant heat, which should be limited due to the compact form factor and proximity to the user's skin. We study the power and thermal characteristics of Glass as an exemplar optical head-mounted device. We share insights and implications to limit power draw with the goal of increasing the safety and utility of head-mounted devices.

*Workshop Paper at APSys 2014*

*Master’s Thesis: Building a Mood Sensor from Smartphone Usage Patterns*

Mood Inference over Smartphone Usage Patterns

Microsoft Research Asia with Yunxin Liu , Nic Lane   
Rice University with Lin Zhong

Our MoodSense system infers a user's mood based on information already available in today’s smartphones, including website visitations, app usage, and communication via SMS, e-mail, and phone calls. The service enhances context-awareness by providing clues about mobile users’ mental states.

*Best Workshop Paper at PhoneSense 2011 (co-located with SenSys), Full Paper at MobiSys 2013*

Internships

* Spring 2015: Samsung Mobile Processor Innovation Lab (Richardson, TX)
  + Internship under Manish Goel in System-on-Chip Team
* Summer 2013: Microsoft Research (Redmond, WA)
  + Internship under Matthai Philipose in Mobility & Networking Research
* Summer 2012: Microsoft Research (Redmond, WA)
  + Internship under Victor Bahl in Mobility & Networking Research
* Summer 2011: Microsoft Research Asia (Beijing, China)
  + Internship under Yunxin Liu in Wireless & Networking Research

Patent Applications

Energy-proportional image sensor (Microsoft)

*Robert LiKamWa, Nissanka A. Bodhi Priyantha, Matthai Philipose, Lin Zhong, Paramvir Bahl*  
2013/2/19; US; 13/770,031

Energy saving mechanisms of image sensor circuitry (e.g., in a camera). Image quality data, such as provided by an application, is processed to make energy consumption of image sensor circuitry more proportional to output image quality by controlling power saving mechanisms of the image sensor circuitry.

Wireless electronic pegboard setup for quantification of dexterity (Shriner’s Hospital; Rice University)

*Steven E. Irby, Dillon P. Eng, Rachel Jackson, Allison C. Scully, Jessica Scully, Robert LiKamWa,*   
*Marcia K. O'Malley, Z. Maria Oden, Gloria R. Gogola, Avery L. Cate*  
2012/4/13; US; 13/446,610

An electronic pegboard setup for assessing patient dexterity. In certain embodiments, the pegboard setup may be wireless and may employ pegs equipped with sensors that that allow tracking of the motion of the peg in three-dimensions and over time, providing quantitative motion path data to assess patient dexterity.

Hand muscle measurement device

*Shuai Xu, Gloria R. Gogola, Graham Sattler, Sridhar Madala, Robert LiKamWa*  
2010/10/8; US; 13/500,607

The present disclosure relates to an integrated system for measuring hand strength and dexterity. Specifically, the integrated system allows for measuring of hand muscle strength through the pinch-grip test, of intrinsic hand muscle strength, and testing of performance in various dexterity tests.

Grants

National Science Foundation, CNS-1657602 CRII: CSR: System Support for Reactive Sensor Operation for Efficiency and Performance, $174,950, 2017-2019, PI.

Teaching

* Fall 2017, EEE 598: Mobile Systems Architecture
* Spring 2017, AME 112: Computational Thinking
* Fall 2016, EEE 598: Mobile Systems Architecture

Professional Service

* 2017: MobiCom ’17 Technical Program Committee
* 2017: MobiSys ’17 External Technical Program Committee
* 2017: MobiSys ’17 Web Chair
* 2017: HotMobile ’17 Technical Program Committee
* 2016: Visible Light Communication Systems (VLCS) ’16 Workshop Panel Chair
* 2016: MobiCom ’16 Social Chair
* 2015: MobiSys ’16 Poster/Demo/Video Regional Chair (North America)
* 2015: MobiCASE '15 Technical Program Committee
* 2015: MobiSys '15 Publicity Chair
* 2015: IPSN '15 Shadow Technical Program Committee
* 2014: MobiCASE '14 Publicity Chair

1. https://artsmediaengineering.asu.edu/about/facilities/intelligence-stage-istage [↑](#footnote-ref-1)