

Human Activity Recognition SoC for AR/VR with Integrated Neural Sensing, AI Classifier and Chained Infrared Communication for Multi-chip Collaboration

Yijie Wei, Xi Chen, Jie Gu

Northwestern University, Evanston, IL, USA

Abstract

This paper presents a distributed multi-chip human activity recognition system for Virtual Reality (VR) and Augmented Reality (AR) applications. A comprehensive solution is delivered including AI core for classification, analog sensing for neural activity detection and infrared data communication for multi-chip collaboration. A 65nm test chip is fabricated and distributed across the body area to demonstrate the low power, low latency, and camera-free features of the target applications.

Introduction

VR and AR applications have recently experienced significant growth driven by gaming, workplace assistance, and social networking. VR/AR offers a new level of virtual immersion to users by seamlessly blending the real and digital worlds. However, as shown in Fig. 1, current VR/AR systems primarily rely on conventional techniques such as joysticks, IMU gloves along with external cameras for motion tracking. These methods suffer from low resolution for sophisticated gestures from users and use of cameras which often have limited view-of-sight and face challenges in a non-stationary environment. Instead, neural activities from human's physiological signals e.g., electromyogram (EMG), provide high-resolution and high-fidelity information to user's body movement [1]. Combining local neural activities with global positioning of human limbs provides a promising solution for camera-free high-resolution human activity sensing. In this work, we present a distributed multi-chip solution that simultaneously captures 4D information, i.e., gestures, limb position and continuous temporal movement of body activities of the user. The contributions of this work include: (1) a fully-integrated SoC chip is delivered including neural sensing of EMG signals, distance measurements of human limbs, and a body area communication for multi-chip collaboration demonstrated by a 65nm test chip; (2) Reconfigurable AI accelerator for both NN and LSTM is integrated for low power real-time body activity recognition; (3) A special infrared (IR) daisy-chained communication is developed for low-cost multi-chip collaborative computing. To the best of our knowledge, this is the first distributed multi-chip solution for camera-free activity sensing, well poised for AR/VR applications.

System Implementation

The proposed system is shown in Fig. 1. Up to four SoC chips (nodes) can be deployed on users' limb and collaborate on the activity classification. Each body node detects the user's local gestures by sensing and classifying the EMG signals from specific muscle groups while also measuring the relative distance between neighboring nodes. The distributed information is collected through a chained communication scheme passing into the last node for final recognition of a variety of body activities such as punching, shooting, and arching, detecting both temporal and spatial movement of multiple human limbs. An infrared (IR) data communication and distance measurement circuit are utilized through a daisy-chain protocol for multi-chip collaboration.

The signal flow for acquiring and classifying local body gestures/movements is shown in Fig. 2. Fully integrated 6-channel low-noise amplifiers with a tunable gain from 35dB -

55dB are deployed to sense EMG signals. The collected EMG signals are processed through mixed-signal circuitry for digitalization and extraction of time-domain features, such as mean, variance, zero crossing, and histograms, from each input channel. A 3-layer fully connected neural network (FCNN) is employed for local gesture recognition based on the extracted EMG features. For a body movement with a temporal series of gestures, e.g. punching, a long short-term memory (LSTM) network is deployed for classification. The LSTM considers all gesture results from FCNN and distance measurements among multiple nodes to derive the final classification results.

Fig. 3 shows the top level SoC block diagram. After the 6-channel analog front-end (AFE) and feature extraction, time-domain features are sent to an AI core which contains both a classifier and a local feature storage memory for recording data for offline training. The compute engine in AI core contains a reconfigurable neural network accelerator which comprises 80 MACs, weight SRAM banks, and other associated caches. As shown in Fig. 3, the NN accelerator computes both FCNN and LSTM along with activation functions, e.g., Tanh, Sigmoid, and dedicated process sequence control. For inter-chip communication, the analog data path includes IR data transceivers, distance sensing with power detector and ADC, and clock-data-recovery (CDR) circuits.

Fig. 4 illustrates the daisy-chain data communication for sequentially sharing data among chips and distance sensing. Each chip is programmed with a chip ID representing the communication order. After gesture classification at local site, each chip node takes turns to transmit its local result to the next node who receives the data and performs distance measurement to its preceding node. After receiving the data, the receiving node appends its gesture label and distance results to the bit stream and transmits it to the next node. The final node engages its LSTM network to process the gesture labels and distance measurements from all the nodes to determine the user's body activity in both temporal and spatial dimensions. A verification procedure is used to enhance error tolerance through a verification pattern and sender ID. Data will only be accepted when the receiver matches the sender ID and verification pattern. The daisy-chain communication enables low-power and low-latency data exchange since only neighbors with short distance are communicating with only gesture label and distance shared. The communication circuitry only consumes 20 μ W and takes less than 4 ms to finish within four interconnected nodes, suitable for low-cost real-time application. To account for clock mismatch, a CDR circuit utilizing a PWM modulated data signal and a delay lock loop (DLL) is used, as shown in Fig. 5. To minimize interference, 300 kHz was selected for data transmission and 3 MHz was used for distance measurement separated in time. The data signal is processed through a CDR circuit in the first 250 μ s, followed by the distance signal which is processed through a high pass filter, power detector, and converted to 8-bit data by an ADC. As the IR LED has a 120° illumination angle limitation, the LEDs beams are positioned to cover the receiver at the full range of arm rotation movement to maintain stable communication as shown in Fig. 5. A distance power detector

is used based on the distance-power relationship of IR signals.

Measurement Results

A 65nm test chip has been fabricated. Fig. 6 shows the measurement results of IR inter-node communication waveform, including initial signals for receiver's DLL to lock, verification pattern, data signal of first three nodes, and a 3 MHz distance measurement signal, completed within 500 μ s for fast communication and distance sensing. The average power of the SoC is 135 μ W, dominated by digital AI core which only consumes 88 μ W at 2 MHz. The AFE consumes 18 μ W while the communication transceivers and CDR consumes 20 μ W on average. The measured waveforms of EMG signals from LNA, and the distance readouts from ADC are also shown in Fig. 6. The distance based on IR signal is accurately captured within a range of 70 cm, making it suitable for localized body area sensing. Fig. 7 shows the demonstration of four example activities with activity matrix, PCB board and die photo. Three sensing nodes were placed on the subject's forearms and upper arm, with the final LSTM node located on the left forearm. FCNN/LSTM models were off-line trained. The multi-dimensional activity classification clusters separated by local gesture, movement and distance show the device achieved an 85% accuracy on activity classification tasks including hand waving, shooting, arching, and punching. Table 1 compares the proposed design with prior arts showing similar or better energy efficiency [3-6]. While existing works only focus on single-chip local bio-recording or standalone neural network processing for bio-signals, this work, for the first time, delivers a multi-chip comprehensive solution covering sensing, multi-chip communication and AI classification for camera-free human activity tracking targeting VR/AR applications.

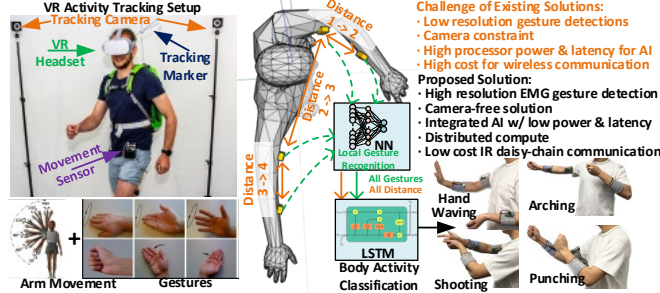


Fig. 1. Human body activity tracking for VR/AR application with challenges of existing solution [2] and proposed multi-chip system.

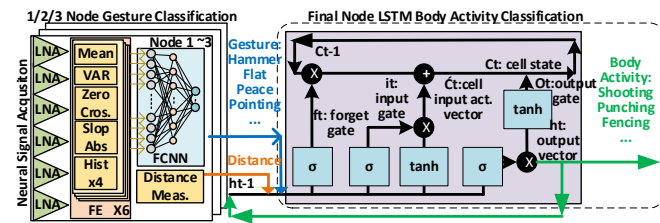


Fig. 2. Local gesture classification through FCNN and LSTM-based body activity classification flow.

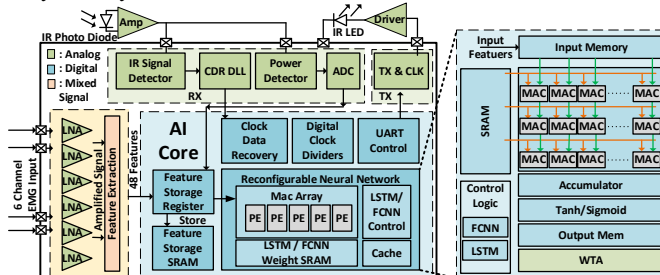


Fig. 3. Top-level SoC block diagram and reconfigurable AI core.

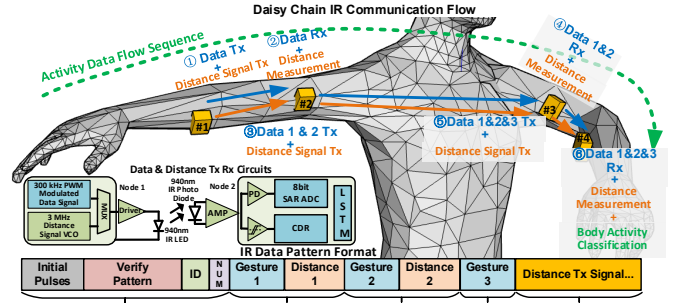


Fig. 4. Multi-chip communication and distance measurement scheme.

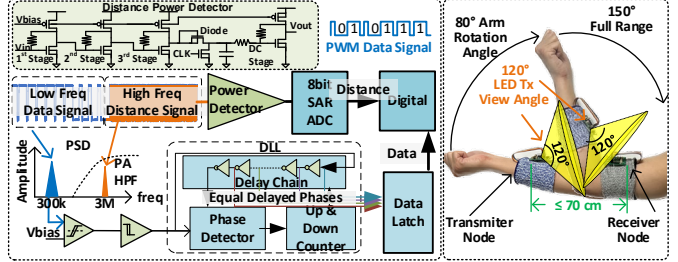


Fig. 5. Data and distance processing flow (left), diagram of IR signal transmission view angle with arm rotation and receiver (right).

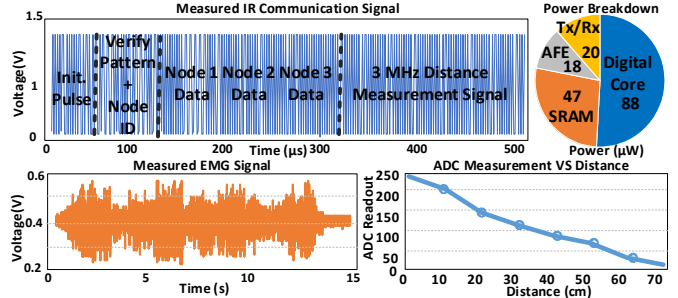


Fig. 6. Measurement results: Inter-node communication data signal (top-left), chip power breakdown (top-right), amplified EMG signals (bottom-left) and ADC readout vs distance (bottom-right).

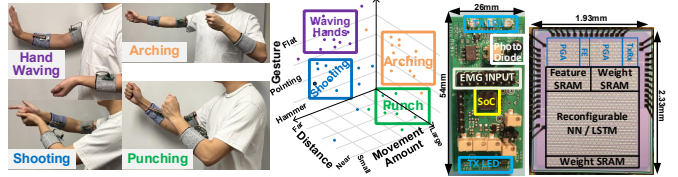


Fig. 7. Experimental setup of 3 nodes with different activities (left) and classification matrix (middle), demo PCB, and die photo (right).

Table 1 Comparison Table with Prior Arts

	CICC'20[3]	ISSCC'21[4]	JSSC'15[5]	JSSC'14[6]	This Work
Technology	180	65	180	65	65
Area(mm ²)	16	1.74	2.025	2	4.5
Power(μ W)	N/A	86.7	24	218	135
Task	EEG classification	Bio-signal processing	EMG sensing & control	Wireless body temp sensing	Gesture & body activity classification
Classifier	DNN	NN	N/A	N/A	LSTM/NN
Frequency	N/A	2.5 MHz	N/A	N/A	2 MHz
SRAM	16kB	73kB	N/A	N/A	40kB
Energy/class	10.13uJ	5.25uJ	N/A	N/A	2.91uJ(FCNN) 7.20 uJ(LSTM)
# of ch	2	N/A	1	1	6
Topology	LNA+ADC	N/A	LNA+ADC	ADC	LNA+FE
Gain(dB)	50-64	N/A	43-57	N/A	35-55
Total	3.26	N/A	7.5	3.6	18
Power(μ W)	N/A	N/A	53.6	215	20
Data Rate	N/A	N/A	1Mb/s	250kb/s	300kb/s
Method	N/A	N/A	RF OOK	RF OOK	PWM

Acknowledgements This work was supported by the National Science Foundation under CNS-1816870 and CCF-2208573.

References

- [1] H. Chen et al. SMC, 2017. [2] J. Auda et al. CHI, 2019. [3] A. R. Aslam et al. CICC, 2020. [4] J. Liu et al. ISSCC, 2021. [5] H. Bhamra et al. JSSC, 2015. [6] L. Xia et al., JSSC, 2014.