A 65nm 2.02mW 50Mbps Direct Analog to MJPEG Converter for Video Sensor Nodes using low-noise Switched Capacitor MAC-Quantizer with automatic calibration and Sparsity-aware ADC

Gaurav Kumar K[1], Gourab Barik[1], Baibhab Chatterjee[1,2], Sumon Bose[3], Shovan Maity[3], Shreyas Sen[1]

[1]Purdue University, [2]University of Florida, [3]Quasistatics Inc., USA

Recent expansions in camera deployments for applications like surveillance, healthcare, autonomous vehicles, robots necessitate low-power realizations even with large data volume. Standard digital video cameras generate ~Gbps data that are usually compressed to ~10's of Mbps, often in the camera (e.g. Omnivision 5640) using compression schemes, like lossy MJPEG before high energy/bit wireless communication. For instance (Fig. 1), a 2K RGB 8b video at 30fps will generate ~1.5Gbps that is reduced to ~75Mbps (MJPEG: per-frame compression) or to ~6 Mbps (H.264: both per-frame and inter-frame compression). While H.264 enables lower data rate, it suffers from complex decoding (incurs high latency), and poor error-tolerance, thus MJPEG is preferred for latency-sensitive error-prone communication of compressed video with real-time inference needs. However, such in-camera ADC with Digital compression suffers from 1) high data volume (~Gbps) generated, consuming high ADC energy; 2) high compute energy required by the in-sensor digital accelerator to compress the data, increasing the power consumption of the processing modules in these nodes to ~10's mW. This work draws on the recent progress of mixed-signal in-sensor computing to address the above challenges in these high-power and high data-volume nodes and presents a "**direct Analog to MJPEG Conversion (d-AJC)**" technique that (1) compresses the analog pixel values from imagers in the mixed-signal domain using Switched Capacitor (SC)-based multiply and accumulate (MAC) units, and Quantizer, before (2) digitizing them with an event-driven sparsity-aware ADC with run-length encoding, effectively compressing Gbps data in SC domain with 10's of Mbps ADC, leading to an order of magnitude system power benefit. With gaining prominence of low-power analog camera sensors (e.g., 25mW Omnivision 6946), along with the development of low-cost processing schemes, d-AJC, it is envisioned to solve issue of high-power imager nodes.

d-AJC employs SC-based mixed-signal circuits for compression before digitization, thereby processing the images at/near the source (imager), which is more power efficient than digital implementations, due to approximate analog computing like [1]-[2]. Fig. 1 (top-right) shows the underlying operations in JPEG compression. It employs 2D Discrete Cosine Transform (DCT) on an 8X8 block of image to transform it to frequency domain, which can also be expressed as a matrix multiplication (primarily, a series of MAC operations) of the input image block with the DCT matrix (A). The Quantization (Q) block takes in DCT output and Q matrix to perform element-wise division. The input Q-matrix determines the degree of compression of MJPEG (Chosen Q_{50} to trade-off compression and quality). The Zigzag (ZZ) Traversal block reorders and serializes the quantized 8X8 matrix such that the significant samples occupy the beginning while the near zero (insignificant) samples are present at the end. Finally, Run Length Encoding (RLE) compresses the image data into fewer bits. We utilize the principle of charge redistribution in SC circuits to realize low-power MAC and division units that form the basis of 2D DCT and Q block, respectively. The values of multiplicands and divisors are stored in ratio of capacitors. The automatic Q calibration, discussed later, takes care of any variation (mismatch/PVT) in these capacitances (used in both DCT and Q), as will be shown in Fig. 5. By pushing the MJPEG conversion at/near the sensor, the presented work achieves ~4X lower power consumption than its digital counterparts and >20X lower ADC conversion energy by enabling it for only significant samples obtained after processing, which only contribute to <5% of the total compressed analog samples. Unlike event driven cameras that can be <10mW but only produce data during spiking events, this work targets power reduction of general-purpose video nodes from ~10's mW to sub-10mW, while providing continuous MJPEG output.

The system level implementation of the d-AJC IC is shown in Fig. 2. The IC serially takes in analog pixel voltages as inputs (column-wise 8X8 image block), performs the MJPEG compression before feeding to the ADC. The 2D DCT architecture is realized in two stages, comprises of DCT unit cells, using charge/voltage mode SC MAC units without any active power-hungry integrators operating with complex clock phases. Each DCT unit cell comprises of eight accumulator slices (P0-P7), containing a set of eight capacitors (C0-C7), corresponding to each row of matrix A that samples the input in non-overlapping phases.

Die micrograph.
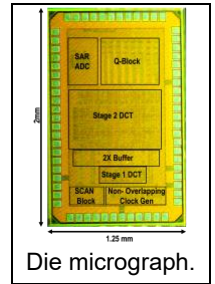
Once all the capacitors sample the input, they are accumulated at C_{Acc} to compute the desired MAC operation, which is then sampled by the next stage before resetting. An intentional attenuation (μ), introduced due to charge sharing, is utilized to avoid saturation, which is later partially compensated using a voltage buffer with gain between two stages. Buffers are carefully placed after two sampling capacitors to ensure the total accumulated input referred KT/C noise remain within limits. Stage 1 DCT takes nine (9) data cycles to produce the output for each input column, which is supplied to the stage 2 DCT that run on slower clock phases, generated by a digital synchronization block, to sample output of the stage 1 DCT. Thus, a total of 72 (=*9 cycles X 8 columns*) cycles are required to compute the 2D DCT of the image block.

The design of key subsystems is shown in Fig. 3, including the Q block, ZZ Traversal and control unit of the RLE+ADC. The Q block employs the principle of charge redistribution between capacitances to divide the output of 2D DCT with the coefficients of Q matrix. The division is performed in two phases to limit the range of the required Q capacitors (C_{Qij},) to reduce overall area of the chip to < 2.25mm[2]. However, the area can be further scaled down considerably for lower technology nodes with higher capacitor densities. Note that the intentional attenuation of the stage 2 DCT is already incorporated in the Q matrix during design time, avoiding the need for extra gain amplifiers for compensation. The output of the ZZ traversal block, implemented using 64:1 analog muxes and digital controller is fed to Differential Ended to Single Ended (DE to SE) amplifier with a gain of 3X to compensate the residual attenuation (from Stage 1 DCT). The Strong-Arm latch-based comparator with PMOS input stage identifies the significant samples at the output of the DE to SE amplifier to enable a 10-bit SAR ADC, while the RLE is activated for insignificant samples.

Power consumption of the d-AJC IC is shown in Fig. 4, for different input data rates and supplies. Our architecture consumes 2.02mW power for 5MSps analog image data (*equiv. to* 360p *RGB* at 12fps or 480p *RGB* at 6fps) for 0.95V core supply. d-AJC is scalable to operate for prevailing higher resolution and high frame rate (HD or 4K 30fps) videos with high-BW buffers, and high-speed ADC. The bias current of the buffer is limited by the quality of the reconstructed image and the input data rate. Fig. 4 also presents the shmoo plot for the ASIC for different supplies and data rates. Fig. 5 exhibits the performance benefits of the system, and the Q matrix calibration that significantly improve (>9dB) the peak signal to noise ratio (PSNR) of a reconstructed image. Different set of 8X8 analog image sets, obtained from the attributes of Q and DCT matrices are sent to the IC for Q calibration. Comparing the actual output from the chip (for the custom input image block) with the expected output provides a direct estimation of the actual capacitance realized in the chip, which is later compensated at the time of decoding to enhance the PSNR of the reconstructed image (Fig. 5), as MJPEG allows concatenation of Q'' with the compressed output for decoding. For 480p RGB video input at 6fps, the effective data rate handled by the d-AJC IC is > 50Mbps, limited by the throughput of ADC.

Fig. 6 shows the measurement setup and compares the test chip performance with state-of-the-art JPEG encoders and DCT compression engines, exhibiting the lowest total Pixel to JPEG (DCT+ADC) Power (2.02 mW, 4X improvement, Fig. 5) and also the lowest 2D-DCT core power (148μW) power by 3X compared to [4] that uses digital circuits at NTV.

References:

[1] M. Miscuglio et al., Nat. Commun. Phys., 2021.
[2] J. Liang et al.,JSSC, 2021.
[3] N. Reynders et al., ISSCC, 2014.
[4] Y. Pu, et al., ISSCC, 2009   [5] M. Pankaala et al., TCAS Video., 2006
[6] S. Kawahito et al., ISSCC, 1997
[7] T. Kuroda, ISSCC1997

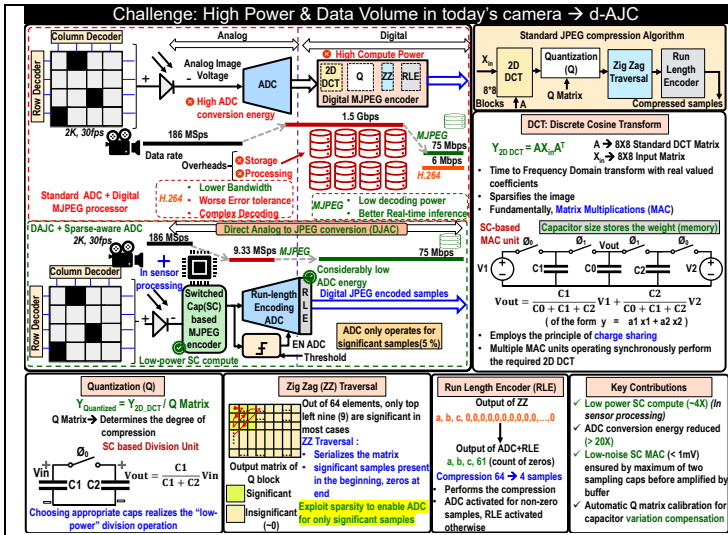## Challenge: High Power & Data Volume in today's camera → d-AJC

Fig 1: (left) Direct Analog to JPEG Conversion (d-AJC) with Sparse-Aware ADC as a low power alternative to standard high power VSNs; (right) Flow of the standard JPEG with SC-based MAC unit for DCT computation; (bottom) Charge sharing to implement division, Sparsity in output of ZZ traversal, functioning of the RLE with key contributions

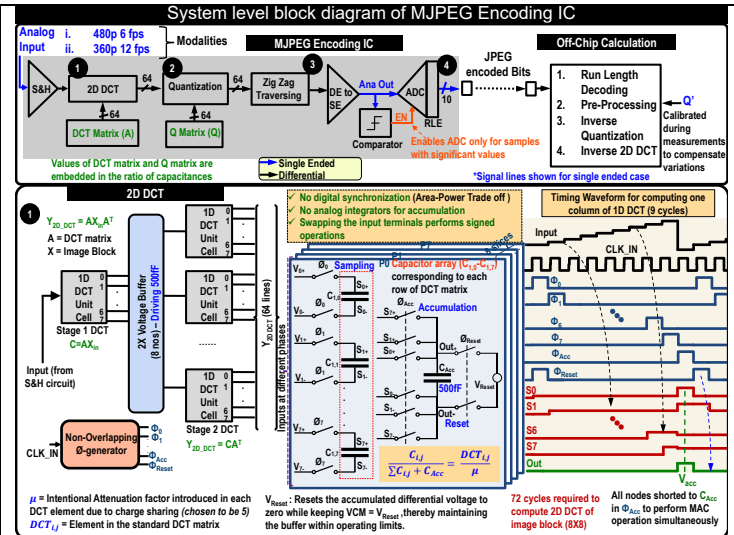## System level block diagram of MJPEG Encoding IC

Fig 2: (top) System level block diagram of MJPEG encoding IC with off chip calculations; (bottom) implementation of 2D DCT using SC based Charge mode MAC units and non-overlapping clock phases to perform sampling, accumulation and resetting of charge stored in the capacitors.
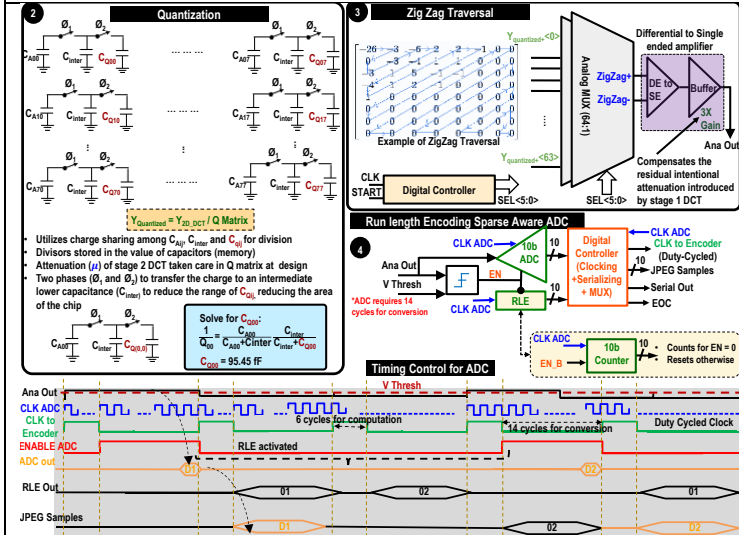

Fig. 3: (left) Circuit level implementation of the Quantization block using SC based divisions; (right) Implementation of Zig Zag traversal block with DE to SE amplifier; (bottom) Control circuit and Timing Diagram of Run length encoding ADC.
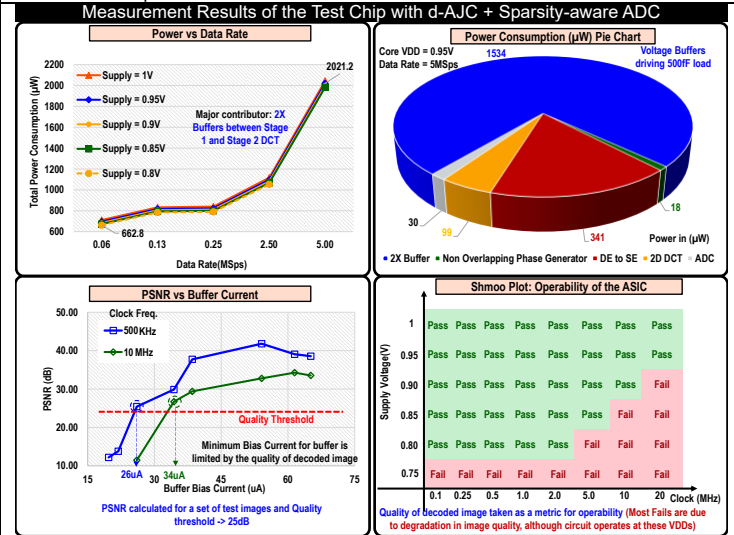
Measurement Results of the Test Chip with d-AJC + Sparsity-aware ADC

Fig. 4: Measurement Results of the test chip: (top left) Overall power consumption for different input data rates and supplies, achieving 2mW operation for 5MSps data; (top right) Power breakdown of constituent blocks; (bottom left) Quality-power tradeoff for 2X buffers; (bottom right) Shmoo plot showing operability of the ASIC

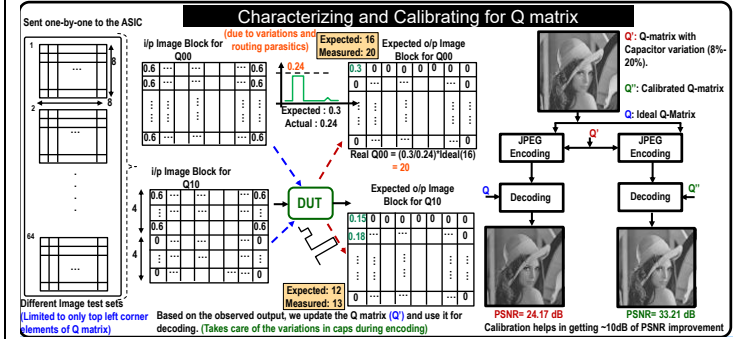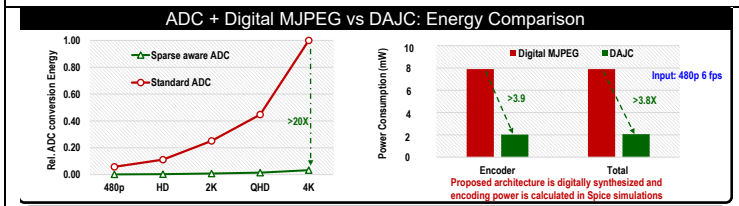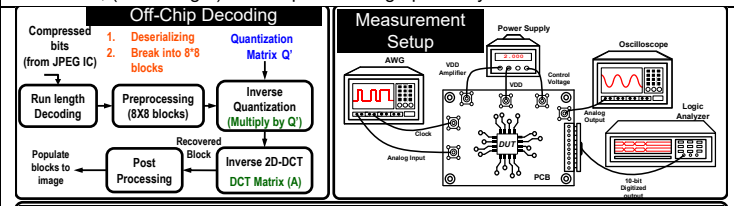## ADC + Digital MJPEG vs DAJC: Energy Comparison

Fig 5 (top) Performance benefit of d-AJC over Standard video sensing node for 480p 20fps data, achieving >20X lower ADC conversion energy and ~4X low-cost processing in SC domain; (bottom) Characterization and Calibration of Q matrix, showing ~10dB benefit in PSNR of reconstructed image

Off-Chip Decoding / Measurement Setup

Fig. 6 (top) Off chip decoding of JPEG encoded image with measurement setup; (bottom) Comparison with state-of-the-art JPEG encoders and compression engines, showing sub 2mW MJPEG encoder and >3X lower power 2D DCT compression over previous implementations

Comparison with the State-of-the-Art

		This Work	ISSCC'14 [#]	ISSCC'09	TCAS for Video'06	ISSCC'97	JSSC'96
Target Application		JPEG encoder	JPEG encoder	JPEG encoder	DCT compression	JPEG encoder	DCT compression
Implementation Style		Analog - Charge/Voltage	Digital	Digital	Analog - Current	Analog - Charge/Voltage	Digital
Clock Freq		5MSps, effective (50 Mbps)	5-275 MHz	2.5 MHz	No Clock	760 KHz	150 MHz
Power dissipation	Total Pixel to JPEG (DCT+ADC) Power	2.02 mW (with low-noise buffers + ADC) @ 480p 6fps	no ADC	no ADC	no ADC	Not Reported	no ADC
	2D DCT Power	148 uW (SC-DCT core power) @ 480p 6fps	30pJ/pixel @41 MHz 35pJ/pixel @ 100MHz 437 uW @ 480p 6fps	[3pJ(for 4 engines)+1.8pJ]* Npipeline (not reported)	2.5mW	10.8 mW- 2 core	10mW
	ADC Energy	0.05E	E	E	-	E	-
Power Supply		0.8-1	0.33	0.4(engine)/0.6 (Huff)	3	0.9	
Area		2.5 mm²	3 mm²	1.96 mm²	0.03mm² (core) total (not reported)	1.1 mm²	4 mm²
Technology		65nm CMOS	40nm	65nm	180nm	350nm	300nm

Without considering ADC conversion and voltage translators power (for NTV) that typically consume significant power at these operating frequencies
* For even comparison with DCT-only architectures, Intermediate Buffer Power omitted during calculation (includes DCT, Q, ZZ with DE to SE in the end)
* pJ/64 pixels not reported for ISSCC'09 (design specific)
*ADC Energy considered as "E" for all encoder architectures