1-bit heuristic adaptive quantizer (HAQ) for on chip image compression in CMOS image sensors

Michael Barrow[a], Amine Bermak[a], Shoushun Chen[b]
[a]ECE Department, Hong Kong University of Science and Technology [b]EEE Department, Nanyang Technological University
Email: mjbarrow@alumni.ust.hk, eebermak@ust.hk, eechenss@ntu.edu.sg

***Abstract*—This paper presents an algorithm for implementing a single bit adaptive quantizer based on fast boundary adaptation rule (FBAR). The peak signal to noise ratio (PSNR) gain and performance gain of the algorithm over prior designs is found to be larger than that displayed by prior art compared with a reference FBAR implementation. A maximum increase of 1.44*db* was seen. In addition, the new design facilitates an improved bits per pixel ratio (bpp) when integrated with the QTD compressor utilized in previous prototypes. The presented algorithm is hardware friendly and designed for low power implementation, with simulation results also showing an improvement of relative energy cost over previous work. Experimental evidence for image sizes ranging from 64x64 pixels to 512x512 pixels and the heuristic adaptive quantizer (HAQ) algorithm are detailed in this paper.**

I. Introduction

Image sensors are increasingly prevalent in today's electronic system designs with the most widespread application of CMOS image sensors is in portable devices [1]. In 2010 camera phones alone accounted for 62% of total CMOS image sensor sales [2]. With the ever increasing image processing demands of new mobile handsets, and the myriad of emerging novel applications, extensive processing is required in today's image sensor application.

Compression facilitates low energy manipulation of a captured image on the sensor and also reduces the energy required to transfer image data to different parts of a hypothetical device utilizing the sensor. This allows for lower power sensors and more extensive processing on chip.

Compression is therefore an instrumental technique used to meet the identified demands of low power and extensive processing in emerging CMOS image sensor designs.

Today's system design space features numerous well known compression algorithms and standards such as JPEG, SPHIT etc. which utilize the compression techniques of a discrete cosine transform (DCT)[3] or discrete wavelet transform (DWT)[4]. However these compression methods require extensive memory and processing which lies beyond the power and area budget for the typical portable application.

Recently there have been several proposed CMOS image sensor designs featuring on chip compression [5], [6], [7] which feature low power consumption and on sensor image compression.

This work contributes a hardware friendly "heuristic adaptive quantizer" (HAQ) architecture for sensors in order they better target emerging CMOS image sensor applications.

II. FBAR theoretical basis

The architecture described in [5] provides quantization research contributions based on predictive and adaptive manipulation of quantizer thresholds. Quantization introduces distortion which must be minimized in an optimum fashion. This design and later evolutions [6], [7] feature sub-optimal implementations of a boundary adaption rule or BAR quantizer.

There are several methods of minimising distortion in an N-point scalar quantizer. Distortion is commonly expressed as a product of rth power of possible threshold differences between an input signal and the probability that this delta could happen. According to the literature the most common method used to minimize distortion is the "Holder norm and its rth power"[8]

$$\Delta(x, Q(x)) = D_r \equiv \sum_{i=1}^{N} \int_{x_i-1}^{x_i} |x - y_i|^r \, p(x) dx \qquad [8] \quad (1)$$

The most commonly used powers are; r=1 (mean absolute error) r=2 (mean square error).

For the powers of r=1 or r=2 one may use the Lloyd I algorithm [8] however the reviewed sensors use the FBAR rule defined by (2) as the basis for quantizer threshold adaptation because it has comparable accuracy but far faster convergence then Lloyd I.

$$\Delta x_j = \eta \left(\sum_{k=j+1}^{N} \delta_k^r \frac{\mathbb{1}_{R_k}}{N-j} - \sum_{k=1}^{j} \delta_k^r \frac{\mathbb{1}_{R_k}}{j} \right) j = 1, \ldots, N-1 \qquad [8] \quad (2)$$

In order to reduce complexity (number of computations) and required storage, prior art implements FBAR with $r = 0$ when adapting threshold levels. This has the advantage of reducing the required computations for threshold adaption to $O(\eta/(N(N-1)))$ [8] where N is the number of Quantization regions. In the reviewed literature only one bit quantizers are implemented, so only two input comparisons are required to adapt quantizer thresholds. Such an implementation, henceforth referred to as "FBAR_0" is expressed as follows:

$$\Delta x = \eta \left(\mathbb{1}_{R_2} - \mathbb{1}_{R_1} \right) \qquad [7] \quad (3)$$

The cost of this simplification is that when $r = 0$ it is not possible to minimise distortion, so SNR is reduced. This is because input variance is not used to compand the quantization levels around commonly occurring signal levels

within the pixel array (or η is highly unlikely to be optimum). Because the variance terms are removed in $FBAR_0$, prior art has implemented a heuristic rule to re-introduce a measure of it [5]. This rule will henceforth be referred to as the "Λ rule".

III. Description of $FBAR_0$ Heuristic Rules

Prior art features evolutionary improvement in performance, characterized by two distinct image sensor architectures [5], [7]. The first, henceforth referred to as prototype one ([5]) focused on low power and sacrificed SNR for simplified architecture. The second generation henceforth referred to as prototype two ([7]) reduced power by removing a large amount of registers.

In consideration of the implemented $FBAR_0$ versus the general form (see eq.2), this work contributes an improved algorithm and improved architecture focussed on improving PSNR with minimal power cost. Given the interrelation, dependence and hardware overlap of the sensor system blocks described in prior art [5], [6], [7], this work focused on the single bit quantizer block. This benefits in being both most isolated in terms of its functionality and the main source of noise. In consideration of spatial continuity in images, prior art implements a heuristic for fast convergence of very large pixel intensity gradients (increasing η), but not one for very small gradients. We see the case for both large and small gradients by considering the example image in fig 1:

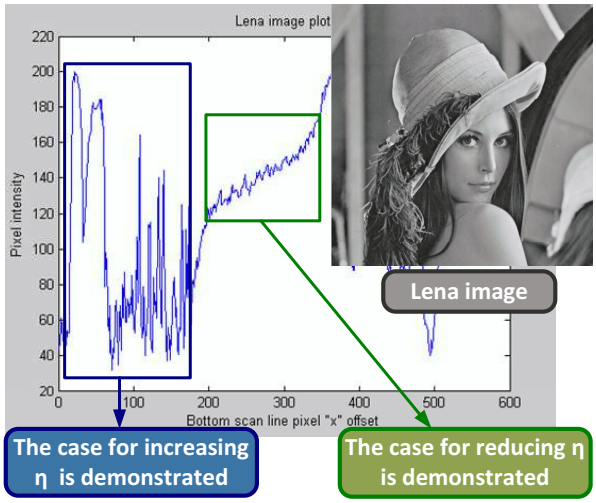

Fig. 1. a scan line of the "Lena" image [9] demonstrating the hypothesis

Sharp contrast between the shoulder and hair of the Lena image corresponds to a large gradient. This could benefit from the heuristic in Prior art (see [7]). However the minimal contrast in the region of Lena's back conversely corresponds to minimal intensity gradient. Previous work [6] employs the hilbert curve when scanning an image pixel array, so Lena from fig 1 is presented to the prior art quantizer in a spatially contiguous manner, compounding the case for a heuristic rule targeting such regions.

To summarize, the hypothesis explored in this contribution is that PSNR could be improved in the general case if:

- The original rule to increase η when the local gradient is large is preserved
- A new rule is used to reduce η when the local gradient is small is introduced
- Both rules are applied in a complimentary way

The theoretical foundation of this new rule is identical to that of the Λ rule in the cited literature, in that introducing a measure of local variance to $FBAR_0$ can help speed convergence, however it operates on the opposite extreme of image intensity variation.

IV. Proposed Implementation

A new η augmentation heuristic rule, henceforth referred to as the "γ rule" was identified for a one bit quantizer. In addition a modified $FBAR_0$ implementation was derived that includes both γ and Λ rules. The result is a newly contributed algorithm of

$$FBAR_0 = f(\eta, \Lambda, \gamma, N)$$

where N must be 2 (for a one bit quantizer). An implementation, known as the "HAQ algorithm" used to characterize the γ heuristic is presented below in fig2.

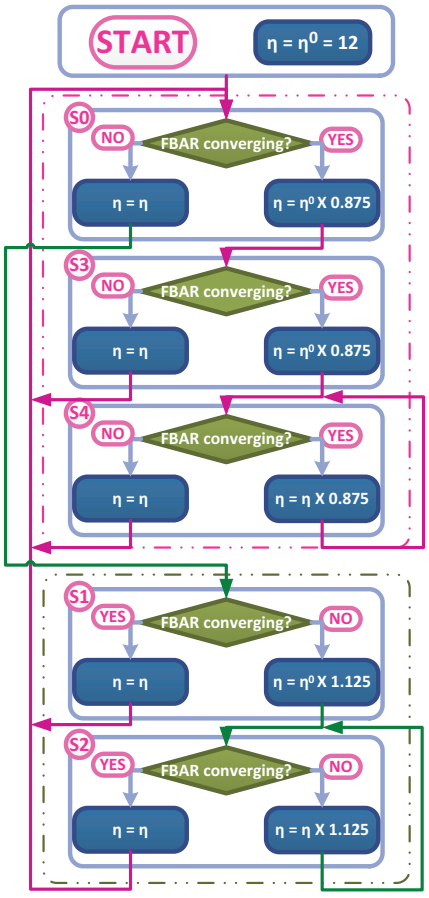

Fig. 2. Flow chart of the HAQ algorithm

Convergence of $FBAR_0$ is measured over two samples and is the previous quantizer output bit xored with the current output bit

TABLE I

SIMULATED PSNR (DB) AND BIT-PER-PIXEL (BPP) FOR 10 TEST IMAGES. ALL IMAGES WERE SCANNED USING A HILBERT CURVE AND SOURCED FROM A SCIENTIFIC REPOSITORY [9]. MEAN PSNR AND BPP WERE OBTAINED BY AVERAGING DATA FROM SWEEPING η_0 FROM 1 TO 50. "R"($psnr/bpp$) IS THE QUALITY TO COMPRESSION RATIO AND MEASURE OF OVERALL PERFORMANCE.

Operation modes	size of test images											
	64x64			128x128			256x256			512x512		
	PSNR	**BPP**	**R**	**PSNR**	**BPP**	**R**	**PSNR**	**BPP**	**R**	**PSNR**	**BPP**	**R**
$FBAR_0$	19.89	0.83	23.96	21.34	0.81	23.35	22.07	0.80	27.59	18.74	0.76	24.66
$FBAR_0$+Λ	19.74	0.79	24.99	21.23	0.74	28.69	21.94	0.73	30.06	18.50	0.67	27.61
HAQ	20.67	0.76	27.20	22.47	0.71	31.65	23.34	0.70	33.5	22.23	0.61	36.44

TABLE II

POST PLACE/ROUTE SIMULATED POWER AND DELAY FIGURES OF THE HAQ QUANTIZATION ARCHITECTURE AND TWO REFERENCE ARCHITECTURES IMPLEMENTED IN TSMC 0.18μm TECHNOLOGY. ENERGY AND LATENCY COST ARE CALCULATED RELATIVE TO THE PREVIOUS HEURISTIC IMPLEMENTATION FOR EACH DESIGN. MEASUREMENTS WERE OBTAINED WITH IDENTICAL HDL SYNTHESIS AND PLACE/ROUTE TOOL SETTINGS

Architecture	Heuristic	Power(μW)	Period(nS)	Energy(nJ/τ)	Energy Cost(nJ/τ)%	Latency Cost%
$FBAR_0$	1	7.281	6.370	0.0464	0%	0%
$FBAR_0$+Λ	2	7.687	9.920	0.0763	64%	56%
HAQ	3	8.309	12.270	0.1019	34%	23%

A. Algorithm key contribution

It is of critical importance to highlight the "Hold off" in this algorithm. The γ rule reduces η in a less aggressive manner then the Λ rule increases it, as shown by the inclusion of step **S3**. Although the Λ rule is preserved in its original form from prior art, it is applied less aggressively as shown by the large number of steps that leave η unaltered. These steps embody the aforementioned "Hold off". Either rule has a second chance at being applied before η is radically altered by the other. This enforces a complementary behaviour between the two and prevents rapid oscillation between applying both of them on consecutive input samples. Experimental implementations without hold off only exhibited a marginal PSNR improvement.

Fig 2A demonstrates the application of the Λ rule in the contributed algorithm. It is seen that after the quantization threshold has risen once the rule is in the "hold off state" **S3**. After a second consecutive rise, the heuristic will be applied **S4** this continues for consecutive samples that the quantizer threshold is below the input signal. The heuristic is also applicable to consecutively lower input signals.

Fig 2B demonstrates the application of the γ rule in the contributed algorithm. If $FBAR_0$ begins to converge, i.e. the quantizer threshold must be increased and reduced in consecutive cycles, the γ rule limits threshold change to $\eta_0 \times 0.875$ This is seen in **S1**. The rule is then in the "hold off state". If $FBAR_0$ (3) continues to converge, the heuristic will be applied to all consecutive oscillating samples and the algorithm stays in **S2**.

B. Hardware architectural implementation

The HAQ algorithm was implemented as a finite state machine (FSM). The optimized "γ" and "Λ" can be applied in a mutually exclusive fashion by using the same "η" adjusted right three bits ($\eta \gg = 3$) and either added (for "Λ") or subtracted (for "γ"). This simplicity results in a low power and latency increase relative to prior art. (see table II).

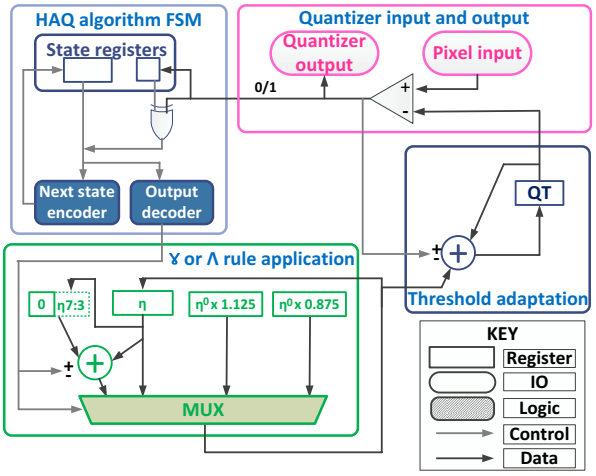

Fig. 3. "HAQ architecture". "QT" is the quantization threshold register. All data busses are eight bit except the comparator output bus, which is one bit. The "State registers" are segmented into a 3-bit block required for storing the HAQ algorithms state, and a 1-bit block to indicate $FBAR_0$ convergence.

Fig 3 demonstrates the HAQ algorithm implemented for low hardware overhead with respect to prior art [7]. The critical path from the heuristic rule to the quantizer threshold or "QT" register is only extended by the additional time needed for data to propagate through the widened MUX. On the control path, additional delay is incurred by the FSM output decoder logic or next state logic. In terms of additional register count, three additional one bit registers are required in the algorithm FSM.

V. EXPERIMENTAL RESULTS

Comparative experiments with prior art using test images were carried out to asses the fitness of HAQ architecture as a replacement. Output data was fed into a QTD compressor as described in prior art ([5]) and the design was also synthesised with TSMC 0.18μm technology (See fig 5). A variety of image sizes were tested to determine how well the design scales.

A. PSNR improvement

It was found that the HAQ architecture offered an additional gain in PSNR of up to $+1.144db$ over "FBAR$_0$ architecture" and up to $+0.764db$ over "FBAR$_0$+Λ architecture". It was observed that the average increase in PSNR over FBAR$_0$+Λ architecture is larger than the increase seen in when the Λ rule is introduced to augment standard FBAR$_0$ (see table I).

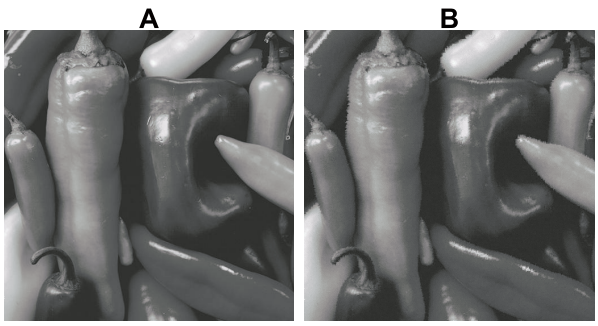

Fig. 4. "Peppers2" image [9] and quantized "Peppers2" image.

Fig 4A shows an original image. Fig 4B shows an image quantized with the HAQ algorithm. The quantized image, has a PSNR of $33.03db$ and was encoded at $0.77bpp$ using QTD compression. η_0 was 8.

In addition to PSNR improvement, compression ratio was found to improve by $0.04bpp$ on average compared with Λ rule prior art. Table I shows the HAQ algorithms performance improvement with respect to prior art (see [6]). "R" is the combined merit of PSNR and bpp ($PSNR/bpp$) and is the performance measure. This data demonstrated that the HAQ algorithm provides a gain in performance much larger then that of prior art compared with standard FBAR$_0$ in all resolutions considered in prior art. The data also demonstrated the trend of improved performance at higher resolution observed in [6], however HAQ performance scales better with an average four times improvement at the highest resolution. Average performance results are significant in demonstrating an improvement, regardless of input image and "η_0" (see fig 2) which may not be changed during field use of an image sensor.

B. Relative power and latency cost

It was found through hardware synthesis that the gain in PSNR and bpp from adding the "γ" rule incurred a low power and latency cost, relative to prior art implementing only the Λ rule. In summary of table II, HAQ architecture incurs a 23% latency cost respective to 56% for previous architecture. Energy per cycle cost also exhibits an improvement of a 34% increase compared with 64%.

VI. Conclusion

The hypothesis that FBAR$_0$ could be improved with a heuristic to adapt the convergence based on detecting small signal variance was shown to be true in a single bit quantizer. An FBAR$_0$ heuristic is shown to improve the performance of an existing single bit adaptive η quantizer. The HAQ algorithm

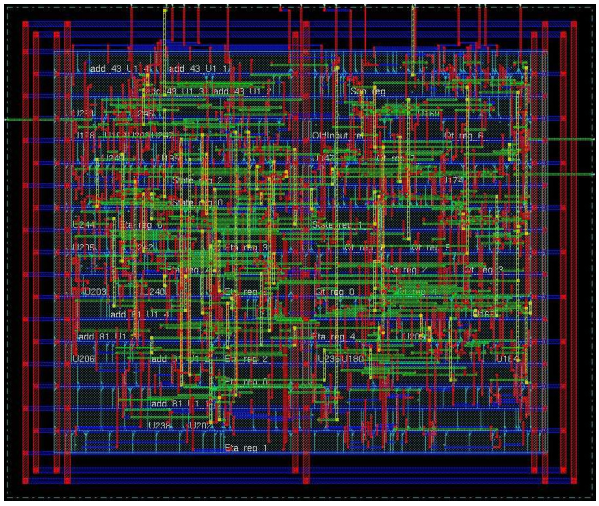

Fig. 5. Layout of HAQ architecture used in characterizing hardware performance. The architecture was synthesised with TSMC $0.18\mu m$ technology

is contributed as an improvement over prior art heuristic quantizers with PSNR demonstrated as having up to a $1.144db$ increase. Both PSNR and Performance were found improve by a larger margin than prior designs compared to a standard FBAR$_0$. Improvements make good use of the hardware design space as the added energy cost is around half that of prior art. In summary, designs able to leverage HAQ will better target today's low power CMOS image sensor applications.

Acknowledgment

The support of Wang Yang is acknowledged

References

[1] El Gamal, A., "Trends in CMOS image sensor technology and design," *Electron Devices Meeting, 2002. IEDM '02. Digest. International*, pp.805-808, December 2002.

[2] Lineback, R., IC Insights Inc, Scottsdale, AZ, USA., "*O-S-D Report 2011*"', 2011. [Online], Available: http://www.icinsights.com/data/articles/documents/256.pdf"

[3] Kawahito et al., "CMOS Image Sensor with Analog 2-D DCT-Based Compression Circuits," *IEEE Journal of Solid-State Circuits*, Vol. 32, No.12, pp.2029-2039, December 1997.

[4] Vishwanath, M. and Owens, R.M. and Irwin, M.J., "VLSI architectures for the discrete wavelet transform," *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, pp.305-316, May 1995.

[5] Shoushun, C., Bermak, A. Yan, W. Martinez, D., "Adaptive-Quantization Digital Image Sensor for Low-Power Image Compression", *Circuits and Systems I: Regular Papers, IEEE Transactions on*, Vol. 54, Number 1, pp. 13-25, Jan 2007

[6] Shoushun Chen and Bermak, A. and Yan Wang., "A CMOS Image Sensor With On-Chip Image Compression Based on Predictive Boundary Adaptation and Memoryless QTD Algorithm", *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, Vol. 19, Number 4, pp. 538-547, April 2011.

[7] Chen Shoushun and Bermak, A. and Wang Yan and Martinez, D., "A CMOS Image Sensor with combined adaptive-quantization and QTD-based on-chip compression processor", *Custom Integrated Circuits Conference, 2006. CICC '06. IEEE*, pp. 329-332, Sept 2006.

[8] Martinez, D. Van Hulle, M. M., "Generalized Boundary Adaptation Rule for Minimizing rth Power Law Distortion in High Resolution Quantization", *Neural Netw.*, Vol. 8, Number 6, pp. 891-900, 1995

[9] Signal and Image Processing Institute, University of Southern California, Los Angeles, CA, "*The USC-SIPI image database*", 2011. [Online], Available: http://sipi.usc.edu/database/inex.html