Scalable Architecture for High-Resolution Video-Rate CMOS Imaging System on Chip

A. Joshi, D. Chiaverini, K. Jung, V. Douence, G. Wijeratne, G. M. Williams and M. Loose
Rockwell Scientific, 5212 Verdugo Way, Camarillo, CA 93012, U.S.A.
Phone: (805) 373-4533, Fax: (805) 373-4687, e-mail: ajoshi@rwsc.com

Abstract

A scalable architecture for high-resolution and high data rate CMOS imaging system on chip (iSoC) is designed. Resolutions such as 2 Mpixel, 12 Mpixel and 16 Mpixel have been implemented in 0.25μm CMOS processes. The architecture offers ultra-low noise performance (7-15e-) with video readout rates up to 320 Mpixel/sec and is suitable for scaling frame rate with window size by being able to use peak bandwidth at various resolutions. The image sensors are based on 4.8μm–10μm pixel size with soft reset and double sampling functionality to reduce pixel temporal and fixed pattern noise.

I. Introduction

Recent improvements in CMOS processing technology and use of novel design approaches are pushing state-of-the-art video sensors capable of >30 Hz frame rate into the mega-pixel resolution regime. HDTV imaging system on a chip having a resolution of 1920x1080 pixels running at 30 Hz progressive and 60 Hz interlaced frame rate is being commercialized [1]. Recently an HDTV sensor capable of 90Hz frame-rate with a video bandwidth of 225MHz has been reported [2]. Such high data rate imagers with large formats pose special challenges to the architecture design. Keeping the dynamic range high while maintaining low power dissipation and offering flexible readout features requires novel architecture design. Some of the earlier high speed imagers [3-4] achieve high data throughput by digitizing image data either using in-pixel ADCs or column parallel ADCs. Digitizing close to the source is power efficient for ultra-high speed imagers but leads to reduced functionality such as inability to scale the frame rate with window size. In addition, due to limited room for digitization circuitry such imagers' conversion time grows exponentially with bit-resolution and they suffer from poorer linearity. Thus the use of these imagers is limited to applications requiring lower dynamic range. In addition, the bandwidth of in-pixel and column parallel ADCs is fixed and the layout is pitch matched, which necessitates redesign of blocks for different applications.

These shortcomings have led to a third class of imagers being most widespread with either passive column buffers used to transfer charge onto capacitive trans-impedance video amplifier or active column buffers amplifying and buffering the voltage onto an ADC. The charge transfer amplifier's power grows super-linearly with imager resolution due to increase in bus capacitance at the amplifier input, which makes their voltage mode counterparts best suited for high resolutions.

This paper introduces a scalable architecture, based on voltage mode column buffers, developed to design high-resolution video-rate imaging system on a chip. Various resolution image sensors such as 2 Mpixel, 12 Mpixel and 16 Mpixel with video data rates up to 320 Mpixel/sec have been implemented.

II. Image Sensor Architecture

The sensors are based on a traditional 3-T pixel. Temporal noise primarily due to reset KT/C noise and the spatial fixed-pattern noise primarily due to source follower FET threshold variations limit the performance of a 3-T pixel. Soft reset is used to lower the reset noise below the KT/C limit and double sampling to reduce pixel fixed pattern noise. Hard reset precedes the soft reset to avoid image lag due to incomplete settling of the integration node.
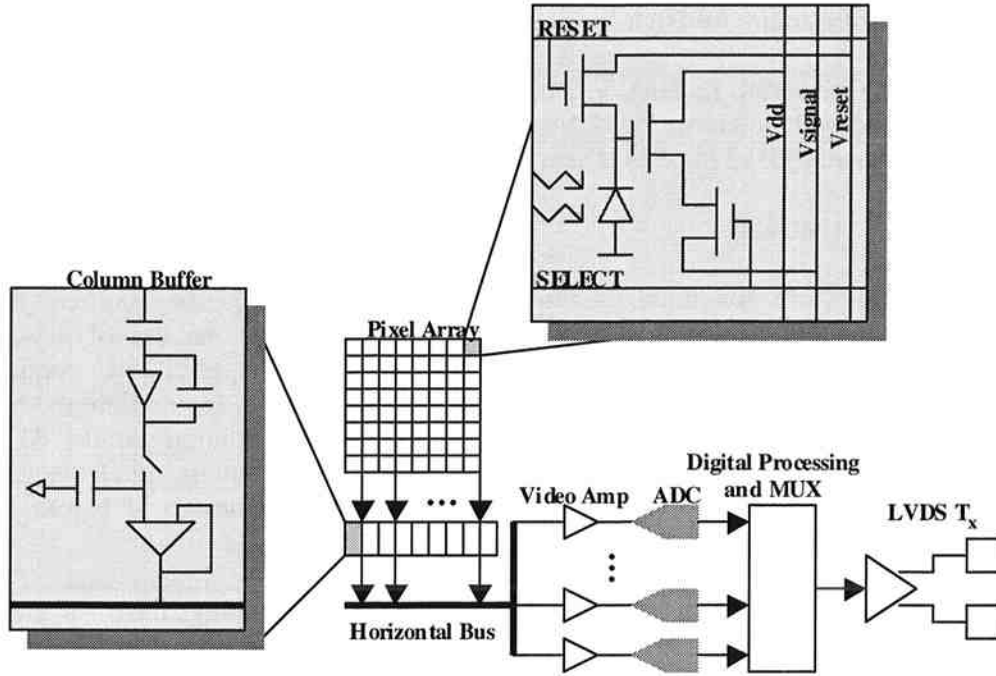
Figure-1 CMOS imaging system on a chip architecture for large format high resolution video

Column amplifiers are composed of a gain stage and a video amplifier (Figure-1). An amplifier's cycle time is limited by its slew time and its settling time, which can be given by:

$$t_{cycle} = t_{slew} + t_{settle} = \alpha \cdot V_{swing} \frac{C_L}{I_B} + \beta \cdot N \cdot \ln(2) \cdot \tau \cdot G \qquad (1)$$

where α and β are parameters depending on percentage of the swing the amplifier slews, N is the bit resolution of the circuit, τ is the settling time constant of the amplifier, C_L is the load capacitance, I_B is the bias current and G is the closed loop gain. Since τ is given by C_L/g_m where g_m is the amplifier transconductance and for input pair FETs in strong inversion g_m is given by $2 \cdot I_B/V_{dsat}$ (assuming amplifier gain bandwidth $g_m/C_L \ll f_T$ of the CMOS process allowing V_{dsat} to be constant and assuming slew time is less than settling time) equation-1 can be simplified as:

$$P_{amp} = V_{dd} \cdot \gamma \cdot I_B \propto \frac{C_L \cdot G}{t_{cycle}} \cdot \left[\begin{array}{c} \alpha' + \\ \beta' \cdot N \cdot \ln(2) \cdot V_{dsat} \end{array} \right] \qquad (2)$$

Using equation-2 the power dissipation of single stage column buffer amplifier can be compared with a 2-stage column buffer with gain stage running at line rate and a video

amplifier buffering the horizontal bus (equation-3).

$$\frac{P_{CB2}}{P_{CB1}} = \frac{1}{G} \cdot \left[1 + \frac{N_{CB}}{N_{LD}} \cdot \frac{f_{line}}{f_{pix}} \cdot \frac{2 \cdot C_{storage}}{C_{bus}} \right] \qquad (3)$$

Since all column buffer gain stages (N_{CB}) are on all the time while only 3-4 line driver video amps (N_{LD}) are on at a given time per ADC channel, N_{CB} x f_{line} ~ f_{pix} and N_{LD}=3-4 per output. For typical mega-pixel sensors with 10-12 bit dynamic range built in 0.25μm CMOS processes C_{bus} ~ 4-8x$C_{storage}$. Thus equation-3 can be simplified further to ~1/G. Due to the gain being decoupled from the video bandwidth the chosen implementation leads to lower power dissipation (for G>1). Each amplifier is auto-zeroed to eliminate vertical fixed pattern noise.

Pipeline ADCs are most power efficient at video rates. Such ADC's power can be divided into bias, digital and analog component power (equation-4).

$$P_{ADC}(f) = P_{bias} + P_{dig}(f) + P_{ana}(a + b \cdot f + c \cdot f^2) \qquad (4)$$

182

The dominant power component, the MDAC amplifier bias current's effect on conversion rate can be given by equation-5.

$$f_{conv} = \left[2 \cdot \left(t_{novr} + \frac{C_L}{I_B} \cdot \left[\frac{\left(N - \log_2\left(\frac{V_{swing}}{V_{settle}}\right)\right) \cdot \ln(2)}{2} \cdot \frac{1}{\sqrt{\frac{2 \cdot I_B}{K' \cdot \beta(I_B)}}} \right]^{-1} + \left(V_{swing} - V_{settle}\right) \right) \right]$$ (5)

ADC power scales linearly with conversion rate till limits of technology are reached. For imagers much faster than 100Mpix/sec having 12-bit dynamic range the benefit of using multiple ADCs is apparent from figure-2. This architecture uses 2 40MHz ADCs for 80Mpix/sec and 8 40MHz ADCs to achieve 320Mpix/sec data rates at < 4mW/M-pix-sec.
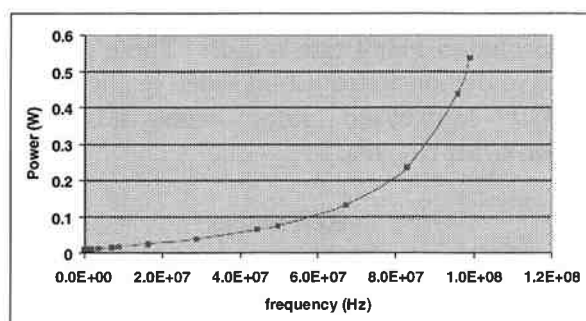

Figure-2 ADC power scales linearly with conversion rate till limits of technology (f_{clk} = 1%-10% of process f_t) are reached

The data in higher speed version of the chips is transmitted using fast LVDS ports. All the necessary sync signals are generated on-chip to enable easy integration into a camera. This system-on-chip approach results in a compact and low power camera, with good video quality at high speeds. The reference camera design (figure-3) employs simple FPGA interface for formatting and transmitting data over a CamLink interface as well as an analog output port for direct image display. The 12 Mpixel sensor offers 60p frame rate at 4 x HDTV (4x1280x720) resolution with highly programmable video timing making it suitable for various standards of high-end application such as high-resolution broadcast. In addition the sensor can run at several hundred to several thousand frames per

second at lower resolution making it suitable for high-speed imaging applications such as sporting event broadcasting and industrial imaging.
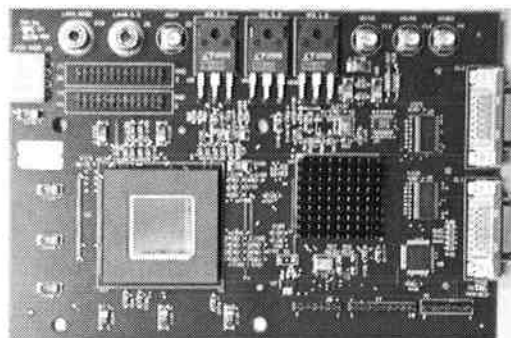

Figure-3 Reference camera board based on 16 Mpixel imaging system on a chip

III. Experimental Results

Contrast transfer function (CTF) of the fabricated sensors was measured using black and white square wave pattern. The data matches well with the model without microlenses, and microlensed parts have a further improved CTF (figure-4). No significant CTF difference in vertical or horizontal direction indicates that analog chain bandwidth is adequate.
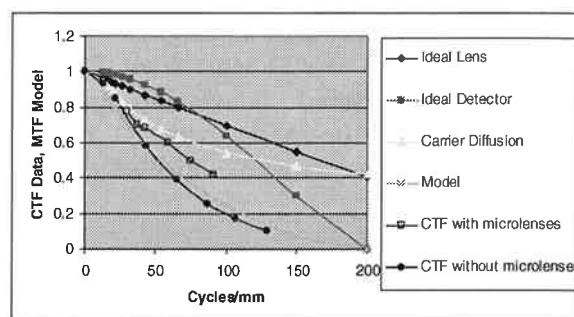

Figure-4 Contrast transfer function (CTF) data from 5μm pixel imagers

The measured QE fill factor product follows the model closely and with microlenses achieves a peak value of 70% (figure 5). The signal chain shows good linearity with 0-90% signal range having non-linearity less than ±1% (figure 6).
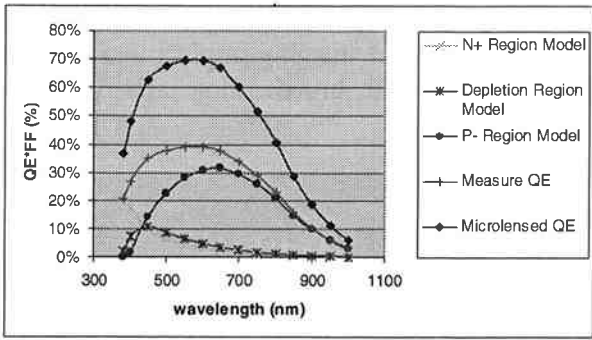
183

Figure-5 Model and measurements of QE x fill factor with and without microlenses from 5μm pixel imagers
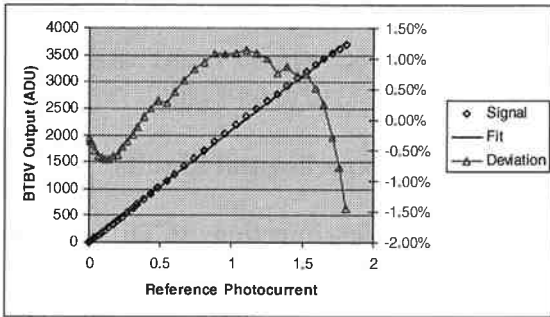

Figure-6 Measured transfer curve (left axis) and non-linearity (right axis) of the mega-pixel imager

The architecture offers ultra-low noise performance (7e- for high gain setting and 15e- for nominal gain setting) with video readout rates up to 320 Mpixel/sec. Figure-7 shows a measured conversion gain of approximately 10e-/ADU with ~1.5 ADU read noise (no signal) under nominal

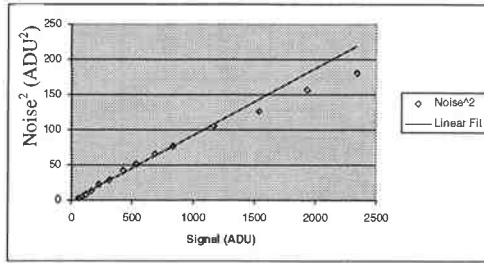conditions. Figure 8 shows some sample images taken by the designed sensors.

Figure-7 Measured noise and conversion gain

IV. Conclusions

An imaging architecture based on 3-T pixel, voltage mode column buffers and an array of ADCs has been developed to design high-resolution video rate sensors. These imagers have demonstrated 12-bit video quality at 80-320 Mpixel/sec rates with less than 4mW/Mpixel-sec.

References

[1] M. Loose et al., 2001 IEEE Workshop on CCDs and Advanced Image Sensors
[2] L. Kozlowski, et al., Proc. of 2005 IEEE ISSCC, San Francisco, U.S.A.
[3] A. Krymski et al., 2001 IEEE Workshop on CCDs and Advanced Image Sensors.
[4] S. Kleinfelder et. al., Proc. of 2001 IEEE ISSCC, San Francisco, U.S.A.

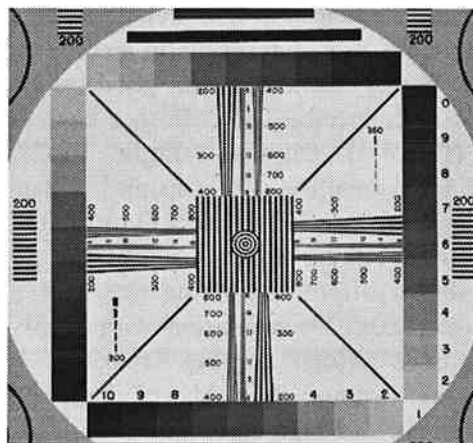Figure 8a. Standard test chart imaged under standard room lighting conditions

Figure 8b. HDTV test pattern imaged using an opto-aligner

184