Impact of TSV area on the Dynamic Range and Frame Rate Performance of 3D-Integrated Image Sensors

Adi Xhakoni[1], David San Segundo Bello[2], Georges Gielen[1]
[1]K.U.Leuven, Dept. Elektrotechniek ESAT- MICAS
Kasteelpark Arenberg 10, B-3001 LEUVEN, Belgium
[2]IMEC, Kapeldreef 75 - B-3001 Leuven, Belgium
Tel. : +32 16 328618, e-mail: adi.xhakoni@esat.kuleuven.be

Abstract—**This paper introduces a 3D-integrated image sensor with high dynamic range, high frame rate and high resolution capabilities. A robust algorithm for dynamic range extension with low sensitivity to circuit non-idealities and based on multiple exposures is presented. The impact of the TSV diameter over the dynamic range and frame rate performance is studied allowing the choice of the best 3D technology for the required performance.**

Keywords-component; 3D integration, CMOS image sensor, high frame rate, high dynamic range

I. INTRODUCTION

Imagers capable of high frame rate and high dynamic range (DR) are required in more and more applications such as robotics, automotive, security etc. High frame rates simplify the image processing algorithms and can trade time resolution with spatial resolution [1] while high dynamic range allows high image fidelity in different light conditions due to the increased capability of the sensor to distinguish between bright and dark areas in the same image.

High frame rate is difficult to achieve for sensors with high resolution. If column processing is used, elements such as the ADCs have to process a higher number of pixels. Also, the settling time of the output of the pixels is increased due to the increased column capacitance which is proportional to the number of pixels sharing the column. This leads to either an increase in power consumption or a decrease in frame rate.

The dynamic range can suffer from the increase in sensor resolution since multiple readouts and captures are needed to synthesize a high dynamic range frame. Although many dynamic range extension techniques have been developed, most of them show low frame rates [2] or limited low light performance [3].

In this paper we introduce a 3D-integrated CMOS image sensor which allows both high frame rate and high dynamic range. Further, we explore the impact of the size and number of the through silicon vias (TSV) in the performance of the system. Our goal is to guide the designer in the choice of the best interconnect technology which fulfills the required specifications of the imager system under design.

This paper is organized as follows. Section II describes the 3D-integrated image sensor system. Section III describes the dynamic range extension algorithm implemented in this work and exploits its low sensitivity to circuit non idealities. Section IV discusses the pixel grouping. Section V provides simulation results showing the impact of the TSV size in the dynamic range and frame rate. Finally, conclusions are presented in section VI.

II. 3D-INTEGRATED IMAGE SENSOR

A. System implementation

The 3D-integrated image sensor system proposed in this paper is represented in fig. 1. It consists of three layers or tiers. The upper layer (tier 0), consists of the photo-detector elements, which are pinned photodiodes in our case [4]. This first layer is connected face-to-face to tier 1 via bump bonds. This avoids the need for TSVs at tier 0 and requires backside illumination process. Tier 1 contains sample and hold (S&H) circuits for global shutter implementation [5], digital logic and comparators for dynamic range extension and analog/digital buffers for interfacing to tier 2. Through-silicon vias (TSVs) and bumps connect back-to-face tier 1 to the bottom layer (tier 2). Tier 2 includes analog to digital converters (ADCs) and memory banks to allow digital frame storage, as well as digital buffers for interfacing with digital processing hardware.

The main benefit of this implementation is the possibility to choose the best technology for each layer according to the functionality. Tier 0 is optimized for sensing (e.g. CMOS Image Sensor technology), tier 1 is optimized for analog performance while tier 2 is optimized for digital performance. Unlike monolithic imagers, where most of the processing is typically performed at column (or even full pixel array) level, the 3D-stacking allows an arbitrary (within certain limits) grouping of pixels for processing. It follows that in analog domain the readout speed does not depend on the resolution of the sensor. The bottleneck in high frame rate performance is therefore moved to the digital domain where fast drivers and DSP systems are required to process the increased amount of data.

B. Tier 1 implementation

The dynamic range extension and the high frame rate are dependent mainly on the implementation of tier 1, which

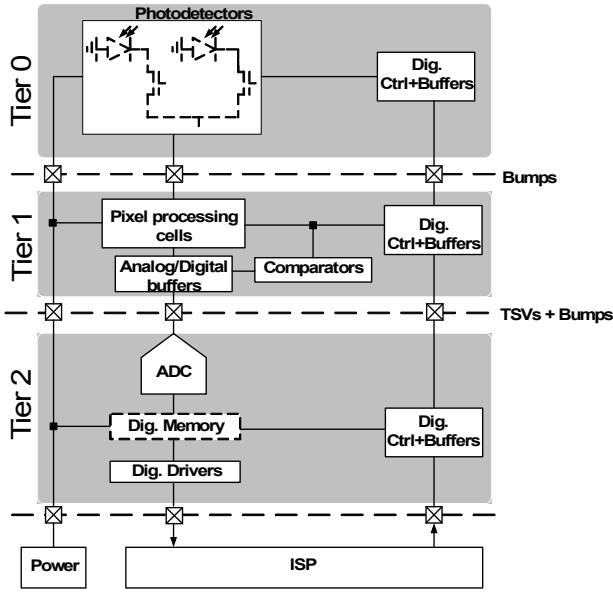contains the TSVs. This tier will be the main subject of this analysis.

Fig. 1. Proposed 3D-integrated imager architecture

Each pixel at tier 0 has its own corresponding pixel processing cell at tier 1 which contains logic gates and an SRAM cell to store a "veto" bit for the dynamic range extension algorithm and two sample and hold (S&H) capacitors: one is used to store the pixel signal value while the other stores the pixel reset noise value in order to implement a correlated double sampling [4]. The differential signal at the two S&H capacitors is buffered via source followers to the ADC in tier 2 through the TSVs. Each comparator used in the DR extension algorithm requires one TSV to send digital data to tier 2. At the cost of a higher complexity, the TSVs between source followers and ADC could be shared with the comparators, allowing smaller pixel grouping. Since all the groups of pixels work in parallel, one only multiplexer is sufficient to perform the TSV sharing operation.
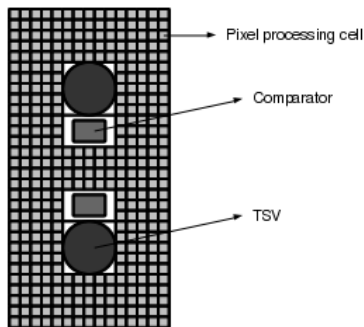

Fig. 2. TSV placement within the group of pixels at tier 1

III. DYNAMIC RANGE EXTENSION

A. Algorithm

A common method to increase the dynamic range of image sensors is the "multiple exposures method" [6]. The whole sensor integrates for different exposure times regardless of the incoming light intensity. The images with the different exposure times are usually stored in a digital memory and the final frame is reconstructed by choosing the capture with the closest value to saturation.

The total dynamic range extension value is shown in (1), where $T_{int}(i)$ is the integration time of capture i, $T_{max}=\max(T_{int}(i))$ and $T_{min}=\min(T_{int}(i))$ for i=1...n. Switching between adjacent exposures means lower number of photons collected. Therefore, an SNR dip is seen at the switching points (2).

$$DR_{ext} = 20\log_{10}\frac{T_{max}}{T_{min}} \tag{1}$$

$$SNR_{dip} = 10\log_{10}\frac{T_{int}(i)}{T_{int}(i+1)} \tag{2}$$

Although two captures are sufficient to extend the dynamic range, the SNR dip would compromise the picture quality at mid-light if the dynamic range extension is large. More captures per frame are therefore required for low SNR dips. This requires a large digital memory and multiple analog-to-digital conversions reducing the frame rate and the maximum integration time for low light detection, and also increases the power consumption per synthesized frame.

In this paper we propose an on-chip implementation of the multiple exposure algorithm which solves most of the drawbacks of the conventional method and that is suitable for high frame rate.

In order to simplify the processing and reduce the SNR dips, the total integration time is divided into sub-integration times $T_{int}(i)$ where each $T_{int}(i)$ is chosen according to:

$$T_{int}(i) = T_{max}\left(\frac{1}{2}\right)^{n-1} \tag{3}$$

with T_{max} representing the maximum integration time of the frame and n the number of captures needed for a certain DR enhancement. The number of bits encoding the DR extension is equal to n-1.

The voltage of the first capture with integration time T_{max} is compared with a reference voltage which is usually close to the lowest allowable output of the pixel processing cell. The reason is that the pixel output voltage decreases at the increase of light intensity. If the output of the pixel processing cell is lower than the reference voltage, the best integration time for the pixel is the longest one and a veto signal is sent by the comparator to the SRAM memory. This will forbid any further captures from being stored in the pixel memory until the next frame starts and the SRAM is reset. If the pixel processing output is higher than the reference, the next capture with half integration time is stored in the pixel S&H. The procedure is repeated until the shortest integration time. At each comparison, the comparator sends its output (one bit) to a digital memory located in tier 2 implementing a floating point dynamic range extension[3].

As seen in (1), the DR extension of this method depends on the ratio of T_{max} and T_{min}. T_{max} is chosen according to the frame

rate and low light level detection required by the application while T_{min} depends on the processing speed which will be explained in paragraph V.

B. Low sensitivity to circuit non idealities

The maximum SNR of image sensors is limited by the photon shot noise which is equal to $\sqrt{m}$, with "m" representing the number of electrons generated in the well capacity of the photodiode. A typical numbers of full well capacity is about 2fF and the maximum number of electrons is in the order of 10ke- to 11ke-[7]. It means that the typical maximum SNR is in the order of 40dB. In case the reference voltage of the comparator used in the dynamic range extension algorithm is increased, the entire fill up of the well capacitance of the photodiode will be omitted decreasing the maximum obtainable SNR. Since the number of electrons generated by the photodiode is lowered, the photon shot noise is decreased too. For instance, an increase of the reference voltage by half of the comparator input range corresponds to a drop of the maximum SNR by only 3dB. A SNR above 32dB is considered as excellent quality according to ISO12232. The low light detection is not compromised and at the highest detectable light, the SNR follows the ideal one since no comparison is done at the last capture.

IV. PIXEL GROUPING

At tier 1, a small grouping would be advantageous for both dynamic range extension and maximum frame rate since it would allow higher parallelism. Ideally, an arbitrary grouping of pixels could be chosen. However, factors such as the processing elements for each group (the comparator in our case), TSV pitch and area, and ADC area will limit the minimum size of the group. To ease the bump-bonding and signal routing complexity, the area of the group of pixels at tier 1 must be equal to the area of the same group of pixels at tier 0. It follows that a pixel pitch of e.g. 10µm at tier 0 imposes a pixel processing cell pitch below 10µm at tier 1, since part of the tier 1 area is used to accommodate the comparators and the TSVs.

To minimize the area lost by the TSVs, two of them are used per group of pixels, as shown in fig. 2. Both are multiplexed, allowing the transmission of the differential analog signal of the sample and hold capacitors towards the ADCs and successively the transmission of the output of the comparators.

The minimum grouping of pixels is chosen according to:

$$N_{pixels} \geq \frac{2A_{TSV} + 2A_{CMP}}{(1-ff)A_{PIXEL}} \tag{4}$$

where A_{TSV}, A_{CMP} represent the TSV area and the comparator area respectively, and A_{PIXEL} represents the area of the single pixel at tier 0. ff is the ratio between the area of the pixel processing cell at tier 1 and the pixel at tier 0 and represents the fill factor of the pixel processing elements over the entire area of the group of pixels. We use a fill factor of 0.9 as a reasonably conservative estimate and an estimated area for the comparator of 5 µm by 5 µm (technology dependent). The comparator small area is due to the relaxed offset requirements as explained in section III. The area of the group dedicated to the TSVs and the comparators is therefore 10% of the total area of the group of pixels at tier 1. By keeping a constant pixel

processing cell pitch, the design freedom is in the size of the sample and hold capacitors. A high fill factor of the pixel processing cell corresponds to an increase of the S&H capacitors value, leading to a decrease of the thermal noise therefore an increase of the inherent dynamic range of the readout system of the sensor.

Following the same reasoning, the area of the ADC at tier 2 should be equal or smaller than the area of the pixel grouping at tier 1. Therefore, the second equation used to determine the minimum number of pixels per group is thus:

$$N_{pixels} \geq \frac{A_{ADC}}{A_{PIXEL}} \tag{5}$$

where A_{ADC} is the ADC area. The ADC used in this simulation has an area of 78µm·39µm which is the same as the one shown in [7].

V. TSV SIZE IMPACT

A key point in determining the dynamic range and frame rate performance of the sensor is the minimum processing time (T_P) which sets the minimum integration time:

$$T_P = t_{DF} + N_{pixels} \frac{t_{PC} + t_C + t_M}{2} \tag{6}$$

where t_{DF} is the time required by the pixel to transfer the photo-generated charges from the photodiode to the floating diffusion node [4], t_{PC} is the time needed by the comparator to access the S&H value, t_C is the time required by the comparator to perform the comparison and t_M is the time required by the comparator driver to write the pixel SRAM memory. The time needed to transfer the voltage of the floating diffusion to the sample and hold capacitor is not taken into account since it is negligible compared to t_{DF}. t_{PC} is usually dominant since the source follower connected to the sample and hold capacitor that has to drive the capacitance of the bus connecting it to the comparator is typically small. t_{PC} is calculated as follows:

$$t_{PC} = c \cdot 2\sqrt{N_{pixels}} \tag{7}$$

where c is a time coefficient proportional to the pixel select switch capacitance and is technology dependent. t_{PC} benefits from the increase of the reference voltage of the comparator as explained in paragraph III. In this analysis we increase the reference voltage to half of the input range of the comparator. In this case, ideally, a 50% of sample and hold output settling would be enough for a correct integration time choice. We take a value of 80% of settling. This way, also circuit non idealities such as the comparator offset do not influence the choice of the integration time. The output settling reduction allows a drastic decrease of the T_P. T_P is also N_{pixels} dependent which is TSV and ADC area dependent as well as shown in (4) and (5). Note that during the analog to digital conversion, the settling error of the S&H output must be lower than the accuracy of the analog to digital converter. The analog to digital conversion of the pixels of frame "i" is done during the first capture of frame "i+1" with integration time T_{max}. The maximum frame rate limitation is as shown in (8) and (9):

$$FR \leq \frac{1}{\sum T_{int}(i)} \tag{8}$$

$$FR \leq \frac{1}{N_{pixels} \cdot T_{AD}} \tag{9}$$

$\sum T_{int}(i)$ depends on T_P (TSV and ADC area dependent) and on the required DR extension. The values of the time parameters discussed above are extracted from circuit simulations and the analysis is done with the following assumptions.

- Two TSVs per group of pixels
- One comparator per TSV
- 10μm pixel pitch at tier 0
- The TSVs and the comparators occupy 10% of the group of pixels area → 9μm pixel processing cell pitch at tier 1
- S&H capacitor of pixel processing cell sized to provide 72dB signal to thermal noise ratio

In case of careful pixel design at tier 0, the inherent dynamic range is expected to be thermal noise limited at 72dB.

At 8 bit DR extension, correspondent to a total 120dB dynamic range, a maximum frame rate of 1928 frames/s is reached (fig. 3). Up to 15μm TSV diameter, an ADC speed of 0.2MS/s is sufficient for the application since the frame rate is limited by T_{MAX} as shown in (5). At larger TSV diameters, the ADC speed has a more pronounced effect on the frame rates. The 0.2MS/s ADC limits the frame rate above 15μm TSV diameter while the 1MS/s ADC limits the frame rate above 40μm. At 1MS/s ADC conversion speed, a TSV diameter of 40μm would be a good choice given the better yield and hence lower cost compared to smaller size TSVs.
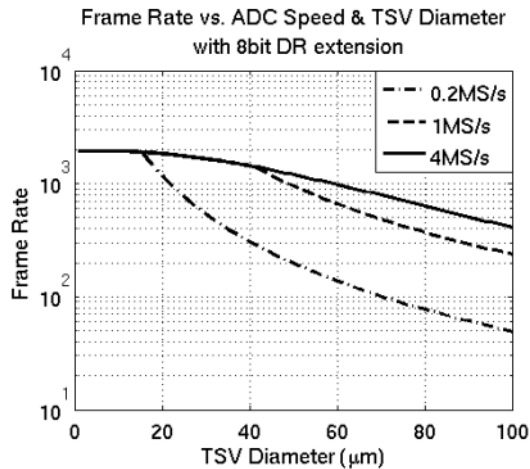

Fig. 3. Maximum frame rate at different ADC speed and TSV diameter when 8bit dynamic range extension implemented

The system could be set to work at increased frame rate by deactivating the dynamic range extension algorithm. The frame rate is, at any TSV size, ADC speed dependent. A maximum speed of 26460frames/s can be reached with a 4MS/s ADC with a TSV diameter up to 12μm (fig. 4). This is due to the minimum grouping size which is ADC area dependent at TSV diameters below 12μm (see (5)).

Compared to state of the art image sensors 2D technology implementations[2],[8], the 3D-integrated system proposed in this work is expected to increase both dynamic range and frame rate.

VI. CONCLUSIONS

A 3D-integrated image sensor system with high dynamic range and high frame rate capabilities has been presented. The concept of pixel grouping allows the maintaining of an ideally constant frame rate and dynamic range at variable sensor resolution. Reducing the TSV size decreases the minimum grouping of pixels allowing an increased parallelism. The paper shows a method to quickly estimate the impact of the TSV diameter in the frame rate and dynamic range performance of a 3D-integrated image sensor.
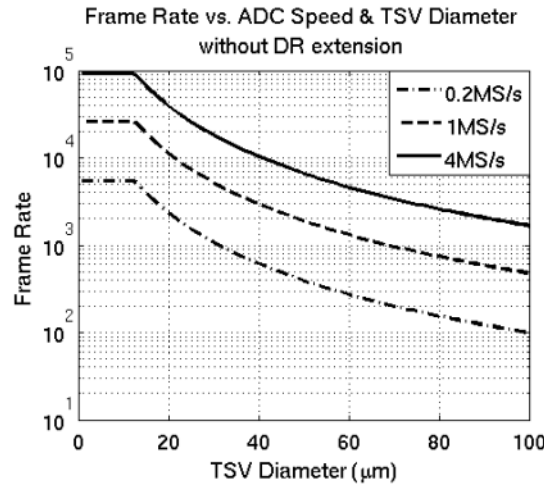

Fig. 4. Impact of TSV diameter and ADC speed in frame rate performance when no dynamic range extension is used

REFERENCES

[1] Masatoshi Ishikawa, "New Application Areas Made Possible by High Speed Vision," in *IISW*, Hokkaido, 2011.

[2] Xinyang Wang, Bram Wolfs, Guy Meynants, Jan Bogaerts, "An 89dB Dynamic Range CMOS Image Sensor with Dual Transfer Gate Pixel," *IISW*, Hokkaido pp. 248-251, 2011.

[3] Yang, D.X.D.; El Gamal, A.; Fowler, B.; Hui Tian; , "A 640×512 CMOS image sensor with ultra wide dynamic range floating-point pixel-level ADC," *ISSCC* , 1999 pp. 308 - 309.

[4] Albert Theuwissen, "CMOS Image Sensors: State-Of-The-Art and Future Perspectives," in *ESSCIRC*, 2007, pp. 21-26.

[5] J. Bogaerts, K. Ruythooren, A. Gvozdenovic, K. Van Esbroeck, B. Ceulemans, W. Ogiers, G. Arsinte, X. Wang, G. Meynants, "High speed 36Gbps 12Mpixel global pipelined shutter CMOS image sensor with CDS," in *IISW*, Hokkaido, 2011.

[6] Spivak, A.; Belenky, A.; Fish, A.; Yadid-Pecht, O., "Wide-Dynamic-Range CMOS Image Sensors—Comparative Performance Analysis," *Electron Devices, IEEE Transactions on*, pp. 2446 - 2461, 2009.

[7] Y. Chae ,J. Cheon, S. Lim, M. Kwon, K. Yoo, W. Jung, D. Lee, S. Ham, G. Han, "A 2.1 M Pixels, 120 Frames/s CMOS Image Sensor with Column-Parallel Delta Sigma ADC Architecture," *Journal of Solid State Circuits*, March 2011.

[8] K. Yasutomi, S. Itoh, and S. Kawahito, "A 2.7e- temporal noise 99.7% shutter efficiency 92dB dynamic range CMOS image sensor with dual global shutter pixels," in *ISSCC*, 2010, pp. 398 - 399.