Advanced 3D Technologies and Architectures for 3D Smart Image Sensors

Pascal Vivet[1], Gilles Sicard[1], Laurent Millet[1], Stephane Chevobbe[2], Karim Ben Chehida[2], Luis Angel Cubero MonteAlegre[1], Maxence Bouvier[1], Alexandre Valentian[1], Maria Lepecq[2], Thomas Dombek[2], Olivier Bichler[2], Sebastien Thuriès[1], Didier Lattard[1], Cheramy Séverine[1], Perrine Batude[1], Fabien Clermidy[1]
CEA, LETI, Minatec Campus, 38054 Grenoble, France ; CEA, LIST, 91191 Gif-sur-Yvette, France
{firstname.lastname@cea.fr}

***Abstract*—Image Sensors will get more and more pervasive into their environment. In the context of Automotive and IoT, low cost image sensors, with high quality pixels, will embed more and more smart functions, such as the regular low level image processing but also object recognition, movement detection, light detection, etc. 3D technology is a key enabler technology to integrate into a single device the pixel layer and associated acquisition layer, but also the smart computing features and the required amount of memory to process all the acquired data. More computing and memory within the 3D Smart Image Sensors will bring new features and reduce the overall system power consumption. Advanced 3D technology with ultra-fine pitch vertical interconnect density will pave the way towards new architectures for 3D Smart Image Sensors, allowing local vertical communication between pixels, and the associated computing and memory structures. The presentation will give an overview of recent 3D technologies solutions, such as Hybrid Bonding technology and the Monolithic 3D CoolCube™ technology, with respective 3D interconnect pitch in the order of 1μm and 100nm. Recent 3D Image Sensors will be presented, showing the capability of 3D technology to implement fine grain pixel acquisition and processing with ultra-high speed image acquisition and tile-based processing. As further perspectives, multi-layer 3D image sensor based on events and spiking will reduce power consumption with new detection and learning processing capabilities.**

Keywords*—Smart Imager, Vision system, 3D technology, Neural-network, spiking computing.

I. Introduction

Image processing demand is constantly growing in many fields of applications like security [1], automotive [2], manufacturing or automation. Embedding such a processing on silicon ("camera on chip") is also a constant demand in order to minimize cost, power, footprint, and maximize speed.

A vision chip, which embeds an image sensor and processing elements including Artificial-Intelligence (AI) features onto a single chip, can answer these limitations. It can provide efficient computation acceleration on the first processing step, combined to a data flow reduction. Being able to integrate all these features into a single chip is facing various challenges: 3D system partitioning, technological constraints, pixel / compute partitioning, data representation to offer low power AI features, design integration, etc. In order to achieve a low power device, the main challenge is to build an imager-computing cube, breaking the rules of current imager architectures.

Regarding the pixel⇔compute partitioning within the imager, several implementations are possible, as depicted in Fig. 1. Embedding a standalone processing unit like in Fig. 1-A allows obtaining a higher bandwidth and a lower power consumption. However, it does not benefit from parallel capabilities. A column processing array implementation (Fig. 1-B) brings parallelism and preserves high pixel fill factor and fine pitch. However, the computation of co-located pixels are limited due to the column data stream and it is not suitable for region of interest operations. Embedding a processing element (PE) in the pixel array is another solution. Fig. 1-C assigns one PE to a cluster of pixels. Such macro-pixel structure (MPX) brings higher parallelism and scalability enabling large framerate capability. The major drawback is the silicon area penalty and mixed signal constraints. Fig. 1-D illustrates one PE per pixel for a maximal parallelism, high speed computing, but poor fill factor and high pixel pitch. It limits the computing capabilities.

3D stacking integration should tackle fig 1. C & D drawbacks by folding pixels and processing elements. Moreover, 3D technology uses one silicon wafer per layer of the 3D circuit, enabling each to be processed in the CMOS technology most suited for its application: imaging technology for the imager, advanced node for the digital processing, etc.

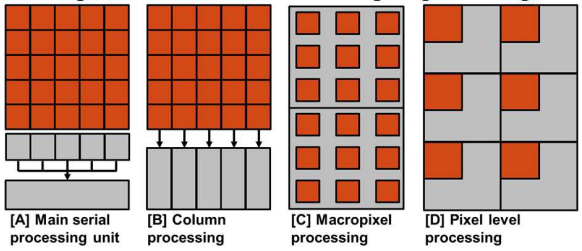


Fig. 1. Examples of pixels (red) and processing elements (Gray) partitioning in several vision chips

In this overview paper, we firstly present the possibilities offered today by advanced 3D technologies, then we provide a detailed review of the evolution of 3D integrated images, including smart fine grain imagers. An overview of neural-network architectures, both covering conventional and spiking neurons computing mode is presented, and the link with the potential pixel architecture. As a perspective, a smart 4-layer fine grain imager architecture is proposed, presenting its main principle, the potential application scopes and the associated architecture and design challenges.

II. 3D Advanced Technology for Smart Imagers

3D technology is offering a wide spectrum of integration schemes [3], as presented in Fig. 2. The standard 3D integration scheme using μ-bumps and Through-Silicon-Vias(TSVs) with 3D-interconnect pitches in the 40 μm range are nowadays a

mature technology, and are used for system integration in many "High Performance Computing" products, with, for instance, Hybrid Bonding Memory (HBM) DRAM cube, or in 2.5D passive interposers for GPU architectures. Nevertheless, the proposed pitch of μ-bumps and TSVs is not very aggressive, and allows only chip-level or core-level integration schemes. For imager applications, more aggressive 3D technologies are required. Two complementary technologies are proposed and envisaged [4]: CuCu Hybrid Bonding and Monolithic 3D.

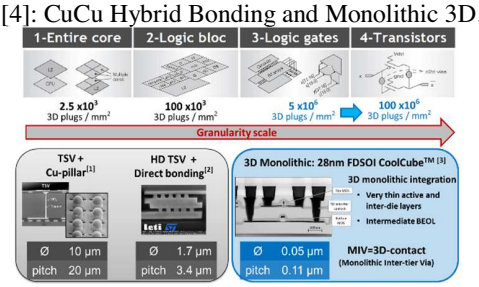


Fig. 2. 3D technology overview and evolution, wrt. interconnect pitch

A. CuCu Hybrid bonding

For CMOS imagers, 3D integration has been a strong enabler to improve form factor and filling factor, to increase image quality and to integrate more logic within the imager. By using wafer level stacking, the imagers have evolved from a single layer device, i.e. a Front Side Imager (FSI); to a dual layer device, i.e. a Back Side Imager (BSI), only integrating the CMOS imager assembled on a bare wafer, with light from the back side for improved image quality; to finally a real dual layer device, a 3D stacked Back Side Imager, integrating the CMOS imager on top of its logic device layer.

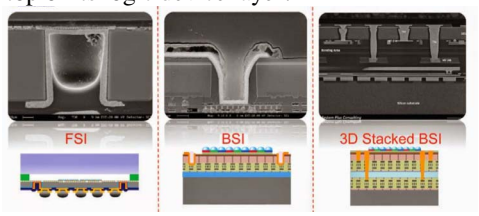


Fig. 3. 3D technology evolution of 3D imagers [5]

In opposition to Fig. 3-C where large TSVs were used to connect the 2 layer Wafer-to-Wafer structure, the 3D technology is now based on a CuCu Hybrid Bonding Technology, such as presented in Fig. 4, using Cu pads and offering more aggressive technology pitches.

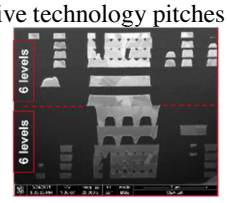


Fig. 4. CuCu Hybrid Bonding interface (Wafer-to-wafer bonded interface)

CuCu hybrid bonding pitches in the range of 10 μm are nowadays common in current CMOS imager products [6], while more aggressive pitches in the range of 1.44 μm were recently achieved and demonstrated [7]. The RC-delay cost is reduced and can be compared to a regular BEOL-via. Nevertheless, using such 3D interconnect pitches, the current CMOS imagers are still using a folded architecture (see section III.A), with sequential column access to the pixels, and not a fully parallel 3D access way (as presented in section III.C).

B. Monolithic 3D

Monolithic 3D (M3D) integration consists in stacking active device layers on top of each other in a sequential manner, (Fig. 5) [8, 9]. It differs from 3D Packaging where the tiers are fabricated in parallel followed by a stacking or bonding step. The sequential flow offers unique 3D connectivity opportunities: as the top active patterning is defined by lithography, the alignment accuracy and feature size of stacked tiers and inter-tier interconnections are only limited by stepper resolution and not by bonding alignment accuracy like 3D packaging. However it comes at the cost of thermal budget constraints for top layer processing, thus the CoolCube™ denomination [9]. Using such M3D technology, 100nm 3D interconnect pitch is achieved.

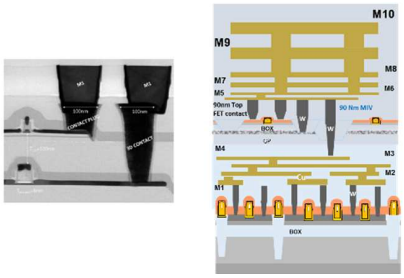


Fig. 5. Monolithic 3D : Cross section of silicium devices and BeoL

C. Technology perspectives for Smart Imagers

M3D technology allows not only to integrate dense logic & memory layers [33], but also to integrate heterogeneous technologies, like MEMS/NEMS for tight coupling of sensing and computing. In the context of CMOS imagers, an optimized 3D pixel may be envisaged, that could be coupled to its associated distributed vision processing. Finally, in order to integrate a multi-layer imager structure as the one proposed Section V, a combination of both 3D technologies, CuCu Hybrid bonding and Monolithic 3D, could be proposed. To achieve a 4 layer imager, the integration by CuCu bonding of two 2-layers M3D structures could be envisioned.

III. Advanced 3D Imagers

A. State-of-the-art Overview

3D stacking technologies applied to image sensors enable new considerations regarding floor-planning and data communication between the two tiers. Since the hard limitation is the respect of the regular pixel array, three main options are depicted in Figure 6.

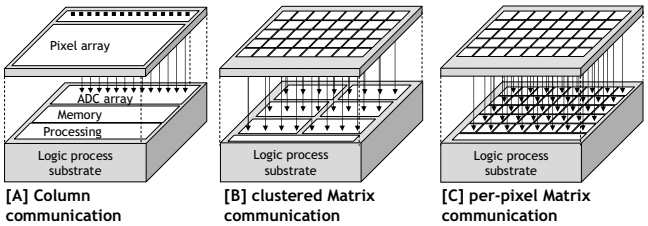


Fig. 6. Communication options between the two layers in 3D-stacked image sensors.

First, the column communication (Fig. 6-A) is a conservative approach where ADCs and pre-processing circuits are folded under the image sensor. This topology provides high design flexibility and benefits from high data bandwidth at column level and dedicated technological nodes. Application examples cover various fields such as image color motion tracking in vision chips for image analysis with low latency [10], tunable frame rate for high speed videos thanks to column readout subsampling [11], event driven image sensor for low power applications [12], and slow motion videos recording capability enabled by embedded DRAM in a 3 layer stacked image sensor [13]

Those works demonstrate the benefits of 3D stacking for a new range of embedded features while reducing the total chip area. However they do not provide fine grain control of pixel regions in the focal plane. To address this point the clustered matrix topology (Fig. 6-B) can be used, where a cluster of pixel from the top tier is assigned to a dedicated readout circuit on the bottom layer. This topology is suitable for region control as detailed in work [14], demonstrating captures with tunable resolutions driven at cluster granularity. Such implementation also brings a massively parallel readout of the pixel array, which leads to an even higher framerate than previously, as presented in [15]. Note that since the readout structures must share the same area as the pixel array, additional peripheral circuits will lead to a total area larger than the focal plane.

With recent improvement regarding 3D via pitch reduction, especially with direct bonding, stacked connectivity is now achievable at pixel level (Fig. 6-C). The limitation of this option comes from the area consumption of the feature embedded in the bottom layer, since we want to dedicate the top tier for the photodiode for maximal sensitivity. Hence this solution is currently suitable for large pixels applications, like SPADs. Two examples of 3D stacked SPAD image sensors showing direct [16] and indirect [17] TOF measures exhibits pixel sizes of 19.8μm and 7.83μm respectively. Some other works investigated the fabrication of stacked photodiodes to enhance the total pixel sensitivity. The works presented in [18] exploit this solution for multispectral image sensor featuring visible RGB plus additional Infra-Red capability.

B. A Fast Burst 2-layer Imager

In more specific fields, burst image sensors are designed to reach multi-million frames per seconds for scientific applications. At such speeds the sensor is unable to stream out the data and must store the whole video in embedded memories. CCD sensors are commonly used, with the ability to record up to one hundred frames in in-pixel capacitor array memories.

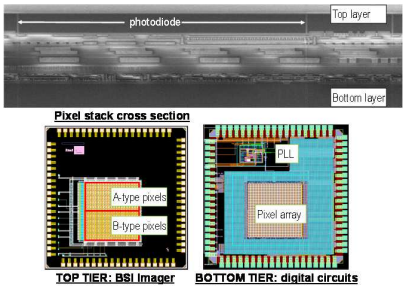


Fig. 7. Burst 2 layer layout views, with pixel cross section detail [20]

However emerging CMOS designs starts to compete with them by providing more flexibility and lower power consumption at cost of lower pixel fill factor. 3D stacking suppresses this limitation by allowing a dedicated pixel partitioning with large photodiode with high fill factor on the top tier, and memories on the bottom tier allowing in-pixel compact analog [19], and even now digital [20] storage (see Fig. 7).

C. A Smart 2-layer Retina

As discussed, 3D stacking implementation is suitable to embed processing structures close to the sensor. With the increasing need of image analysis regarding automotive or emerging augmented reality applications, such integration may be suitable to perform computing tasks ahead of time with low latency feedback. Up to now, very few works take fully advantage of 3D stacking parallelism for image analysis applications. A novel prototype of such vision chip is proposed in [21-22][21]. It combines a back side illuminated image sensor featuring fully parallel in-focal-plane readout circuit with a matrix of 3072 programmable processing elements on the bottom tier, demonstrating a 5500fps image sensor with low latency image analysis.

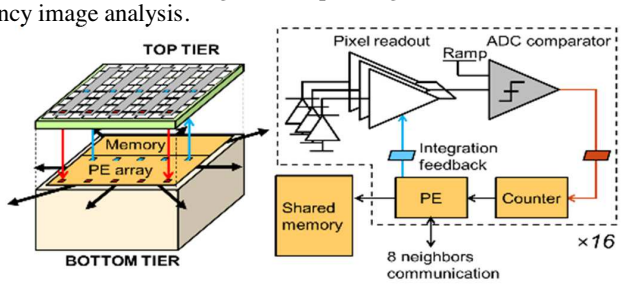


Fig. 8. Bloc diagram of the scalable Macropixel 3D structure

The main characteristic of the design is the use of scalable clustered structure called macropixel (MPX) as detailed in Fig. 8, removing data stream bottleneck between the image sensor and the processing cores. A single MPX structure is a combination of a 16×16 pixel array (or 64×64 pixels in high resolution mode), associated to 16 SIMD 8bit processor cores. The acquisition configuration (pixel dynamics, integration time) and the programming of the chip can be set at MPX granularity level. This flexibility allows ether a fully SIMD execution mode or a heterogeneous mode by running simultaneously several programs on different areas (Fig. 9).

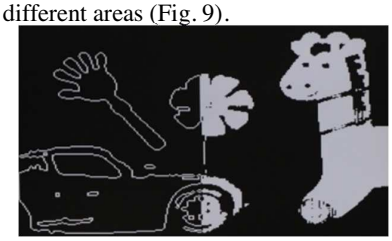
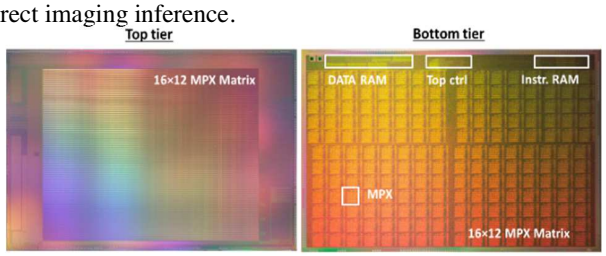


Fig. 9. vision chip in heterogeneous execution mode : edge detection computing on the left, and threshold on the right

The circuit has been implemented in a 130nm technology, and 3D stacked using CuCu Hybrid Bonding using a 7μm pitch. The chip cross section and die micrographs are given in Figure 10.The vision chip prototype has demonstrated local pixel neighborhoods based computation capabilities for applications such as motion detection and edge detection. The efficiency of

the current architecture can be improved by increasing its embedded memory capacity thanks to a 3rd 3D layer integration. Some image processing functions need a global interpretation level that may require extra layers such as an efficient general purpose CPUs or dedicated accelerators like neuro-engines for direct imaging inference.



Item	Data
Technology	130nm 1P7M/130nm 1P7M
3D bonding	Cu-Cu direct bonding
Chip / Sensor area	187 / 113 mm²
Resolution	1024 x 768
Pixel size	12µm x 12µm
Pixel fill factor	75%
Frame Rate	340fps @ 0.78 Mpixels 11bits 5500fps @ 0.05 Mpixels 9bits
PE Nb @ Frequency	3072 @ 80 MHz
Power consumption	70mW (top) + 720mW (bottom)
Performances	61Gops @8b img

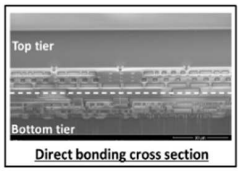


Fig. 10. Vision chip micrographs: top and bottom tiers, 3D cross section and main prototype characteristics.

IV. Neural Networks for Imagers

A. Introduction

There is currently a growing interest in Neural Networks (NNs) due to their impressive ability at autonomously extracting statistically relevant patterns in large amounts of data. More specifically, convolutional NNs (CNNs) can reach above human performance in image classification tasks, and are well-suited to 3D integration below a pixel matrix thanks to their local information processing [23]. Near-sensor NNs evaluation may enable to reduce the amount of communicated data to further processing units, or avoid the requirement of a cloud connection to treat relevant data, which solves some privacy and safety complications. The networks would be trained off-chip and only the inference would occur within the smart imager.

B. Efficient implementation of Neural Networks

Due to the high complexity of Deep Neural Networks (DNNs), their integration into embedded systems remains challenging. Their structure and the amount of parameters obtained as results of the learning phase make the inference phase compute - and memory - intensive. In addition, even if their computation requires simple energy-efficient operators, their number is tremendous and requires a complex interconnection and memory hierarchy architecture incompatible with embedded implementation. Efficient processing of modern neural networks for smart imager can be performed by programmable accelerator such as PNeuro [24]. It is a scalable energy-efficient programmable accelerator designed for the inference phase of DNN processing chains. Simple programmable processing elements architecture in SIMD clusters perform all the operations needed by DNN (convolutions, pooling, non-linear functions). Beyond the integration close to the imager, 3D integration technology can improve memory density that is a key challenge for hardware implementation of neuronal networks.

To fully exploit the PNeuro accelerator, N2D2 framework [25] has been developed. It integrates a complete methodology that conciliates both the formal and spiking neural models. It offers the basic functions of classical deep learning frameworks to explore formal DNN, and integrate code generation modules for specific accelerator such as PNeuro and for many CoTS targets using standalone CUDA, CuDNN, OpenMP, OpenCL, OpenCL for HSL and C for HLS programming models. This is an adequate tool to study conventional and spiking Neural Network architectures within a dedicated imager structure.

C. Spiking Neural Network

Spiking Neural Networks (SNNs) differ from the previous presented conventional networks in the sense that they process trains of spikes instead of static numeric value [26]. They are well suited to exploit event-driven information, for instance coming from dynamic vision sensors [27], which present several interesting properties. These sensors allow high dynamics and very small reaction time, with an overall low power for the reached frequency available. SNNs are appealing for hardware designers because neuron evaluation requires adders only, and inter-neuron communication is efficient, with the only data transmitted being the address of the emitting neuron [28]. There is thus a growing belief that the combination of event-driven sensors with local SNN processing is realizable and efficient so that it can offer reduced power budgets, well adapted to smart imager and vision system. Belhadj et al. [29] have proposed and simulated at the layout level a 2-tier 3D integrated circuit using micro-bump 3D interconnects dedicated to the evaluation of a 2-layer SNN, on the MNIST benchmark [30]. They show that the 3D architecture (see Fig. 11) enables energy and speed gains of more than 60x with respect to a GTX480 GPU evaluation when each layer is evaluated on a single tier. They also compare their circuit to a similar 2D design, in which case they reach gains in time, energy and area of at least 1.5x. This example shows the benefit of using 3D technology to efficiently implement a tightly coupled neural network, in that case a spiking neural network.

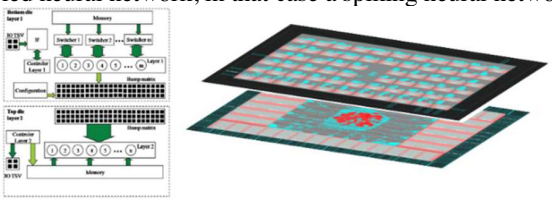


Fig. 11. 2 layer spiking neural network implemented into a 2 layer 3D-circuit

D. Towards Spiking Pixels

In order to build an efficient Smart Imager, it is required to co-design the pixel and the associated neuro-processing, with interesting opportunities to be studied. Instead of using classical pixels, pixel intensity can be codified in pulses like in the case of a "Time to first spike" scheme [31], in which the time that the pixel's reading voltage drops from an initial state below a threshold value encodes the light intensity.

Spikes can also indicate asynchronous events, using the Address Event Representation (AER), which can then be processed asynchronously. Then, useful information can efficiently be extracted, as proposed by [32]. They presented a smart sensor which implemented AER for object tracking applications : a complete spiking system can be envisaged, starting from a spiking pixel to spiking neuro- computing.

V. Towards a 3D multi-layer Smart Image Sensor

A. Application Targets

Smart imagers taking advantages of advanced 3D integration can solve system bottlenecks in various application domains going form very low power IoT solutions to high complexity image based embedded applications.

- In IoT domain, tightly integrated features like event based wake-up modes, variable image resolution with adapted processing, adaptive sensor parameters or even energy harvesting pave the way for fully autonomous smart sensor systems.
- On high-end applications like sensing systems for autonomous vehicles, 3D integrated imagers enable complex computing for scene perception at high pixel rate with very competitive power budget.
- Another key differentiator of such sensors, thanks to the 3D integration, is the low achievable latency and high compactness for applications in the AR/VR domain.

For all these domains, smart imagers brings also two key advantages: privacy and energy reduction. Privacy is a common concern in all the application domains of connected systems. Integrated smart imagers have advantageous properties to this respect, since no raw data is going outside the integrated imager device. This enables easier protection or anonymization of the interpreted output data. The image is kept local, while only extracted features would be exchanged with the cloud, etc. Smart imagers will reduce the system energy. By locally extracting information within the imager, there is no need to store the raw or compressed images in distant memories/servers: the interpreted information is computed and stored locally, reducing the energy and bandwidth needs of communication means (radio, system infrastructure, etc).

B. Architecture proposal

In the context of these targeted applications, based on current 3D technologies, by re-using the concept of fine grain 3D imagers and integrating neuro-based computing, it is envisioned to propose and design a multi-layer smart image sensor, offering in-situ energy efficient detection and learning processing capabilities. To address the major limitations of previous architectures, a dedicated memory layer is added, and thanks to an optimized neuro-based processing layer and fine grain communication scheme in between, we target to achieve one further step in power efficiency and low latency.

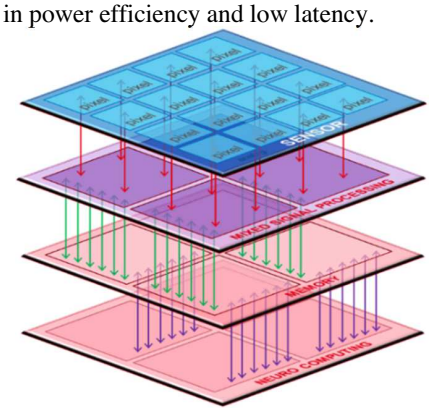


Fig. 12. 3D multi-layer smart image sensor architecture overview.

The overview of this complex integrated vision system is given in Fig. 12 and is presented below in more details:

- The first layer would be composed of the photosensitive array. To achieve high image quality, BSI 3D integration are required so that pixels can be access from the back side. Depending on application target (see below), the pixel could be either an existing fully qualified industrial pixel, or a dedicated pixel (spiking pixel as an example for direct integration with the spiking neuro-layer). This pixel-only layer allows a low cost dedicated imaging technology.
- The second layer would be dedicated to all required near pixel processing, using mixed mode circuits. The structure must be clearly partitioning at pixel level or cluster of pixels, for logic re-use, multi-scale resolution, and wake-up mechanisms based on region of interest. This layer should be better implemented using a mature low cost CMOS technology (65nm for instance). Fine grain 3D integration between the first two layers allow fine pixel connectivity.
- The third layer would be composed of mostly a memory array, to store pixels, spiking events, Neural Network weights, and some intermediate frames. The memory layer must be partitioned in line with the computing tiles presented below.
- The fourth layer would be dedicated to digital computing for smart neuro-applications for image recognition, vision system, etc. Two neural network computing systems may co-exist and be integrated jointly, depending on the targeted application classes, using Convolutional Neural Networks or using Spiking Neural Network.

These two last layers would be implemented using advanced technology nodes (28nm or even 10nm CMOS for efficient digital computing and memory density). Dense 3D interconnect pitch between memory & compute layers will allow fast and energy efficient memory access from compute engines, maintaining SIMD-like parallelism, and reducing the memory wall effect.

C. Challenges and trade-offs

The proposed 4-layer smart imager architecture requires a strong co-design between all the proposed layers, while proposing disruptive mechanisms, addressing challenging technology integration aspects, in order to be integrated in a reduced form factor, cost efficient, and energy efficient smart object. The design tradeoffs will obviously depends on the application domain target, and the main identified challenges can be classified as follows:

- In terms of overall system sizing, the imager resolution and pixels pitch are critical characteristics as they will fix the size of all layers. This size has to be balanced with the needed memory size and needed digital complexity. The global partitioning of the object must be studied with the full system as a target, by balancing pixel size, memory size, and computing capabilities.
- The parallelism level of processing is challenging regarding the power consumption and the latency. This choice will impact the pixel cluster size and will define the complexity of the first level of processing on the second tier. Moreover, using an industrial pixel could strongly limit the processing complexity on the second layer.
- The data encoding of pixel information is a major concern. While spiking pixels attached to energy efficient spiking

neural networks accelerator will offer an event driven imager for environment monitoring, it may not be an option for autonomous driving. Combining partly the conventional and spiking information is an open research topic that could bridge different application sub-classes.

- Regarding the 3D logical and physical design, many challenges arise, such as fine grain 3D logic-on-memory physical design [33], Design-for-Test, power delivery, and thermal aspects.

VI. Conclusion

In the current context of imager and vision systems, this paper presents an overview of the recent advances in 3D technologies available in CEA Leti, proposing two complementary technologies: CuCu Hybrid bonding and Monolithic 3D. A detailed state of the art of 3D-stacked imagers is given showing the current integration trends, which are still mostly using column based architectures. A smart imager using fine grain pixel 3D connectivity and parallel SIMD computing is presented, validating the approach. By using an additional memory layer and integrating neuro-based accelerators, a 4-layer imager is proposed. Application perspectives are discussed, while the architecture and design challenges are identified. This paves the way toward a multi-layer 3D image sensor based on frames and spikes to reduce system energy, integrating neural network processing capabilities.

Acknowledgment

This work was partly funded thanks to the French national program "Programme d'Investissements d'Avenir, IRT Nanoelec" ANR-10-AIRT-05, and by the European H2020 project 3D-MUSE (780548).

References

[1] A. Suponenkovs, A. Sisojevs, G. Mosns, J. Kampars, K. Pinka, J. Grabis, A. Locmelis, and R. Taranovs, "Application of image recognition and machine learning technologies for payment data processing review and challenges", AIEEE 2017, pp. 1–6.

[2] M. Takemura, T. Shima, and S. Muramatsu, "Embedded image recognition systems for advanced safety vehicles," in 2015 Symposium on VLSI Circuits (VLSI Circuits), pp. C146–C147.

[3] V. Pavlidis, I. Savidis, E. Friedman, "Three-Dimensional Integrated Circuit Design", 2nd Edition, Morgan Kaufmann Publishers, ISBN: 9780124104846.

[4] Chapter 7, "3D Integration in VLSI Circuits: Implementation Technologies and Applications", Edited by Katsuyuki Sakuma. ISBN-13: 978-1138710399. ISBN-10: 1138710393.

[5] Yole Development report 2015

[6] S. Lhostis et al., "Reliable 300 mm Wafer Level Hybrid Bonding for 3D Stacked CMOS Image Sensors", ECTC'2016.

[7] J. Jourdon et al., "Hybrid bonding for 3D stacked image sensors: impact of pitch shrinkage on interconnect robustness", IEDM 2018.

[8] P. Batude et al. "Advances in 3D CMOS Sequential Integration", IEDM 2009.

[9] P. Batude et al., "3D Sequential Integration: Application-driven technological achievements and guidelines", IEDM 2017.

[10] Yamazaki T., Katayama H., Uehara S. , et al., « A 1ms High-Speed Vision Chip with 3D-Stacked 140GOPS Column-Parallel PEs for Spatio-Temporal Image Processing", ISSCC 2017, pp82-83

[11] Chou P., Chang C., Mhala M. M., Liu C. C., et al., "A 1.1μm pitch 13.5Mpixel 3D-Stacked CMOS Image Sensor Featuring 230fps Full-High-Definition and 514fps High-Definition Videos by Reading 2 or 3 Rows Simultaneously Using a Column-Switching Matrix", ISSCC 2018, pp. 88-90

[12] Kumagai O., Niwa A., Hanzawa K., Kato H., Futami S., et al., "A 1/4-Inch 3.9Mpixel Low-Power Event-Driven Back-Illuminated Stacked CMOS Image Sensor", ISSCC 2018, pp. 86-88

[13] Haruta T., Nakajima T., Hashizume J. , et al., "A 1/2.3inch 20Mpixel 3-Layer Stacked CMOS Image Sensor with DRAM", ISSCC 2017.

[14] Takahashi T., Kaji Y., Tsukuda Y., et. "A Stacked CMOS Image Sensor With Array-Parallel ADC Architecture", JSSC 2018, pp. 1061-1070

[15] Arai T., Yasue T., Kitamura K. , et al. "A 1.1μm 33Mpixel 240fps 3D-Stacked CMOS Image Sensor with 3-Stage Cyclic-Based Analog-to-Digital Converters", ISSCC 2016, pp 126-128

[16] Ximenes A. R., Padmanabhan P., Lee M., Yamashita Y., Yaung D. N., Charbon E, "A 256x256 45/65nm 3D-Stacked SPAD-Based Direct TOF Image Sensor for LiDAR Applications with Optical Polar Modulation for up to 18.6dB Interference Suppression", ISSCC 2018, pp96-98

[17] Abbas T. A., Dutton N. A. W., Almer O., Pellegrini S., Henrion Y., and Henderson R. K., "Backside Illuminated SPAD Image Sensor with 7.83μm Pitch in 3D-Stacked CMOS Technology", IEDM 2016, pp 8.1.1-8.1.4

[18] Y. Takemoto, K. Kobayashi, M. Tsukimura, N. Takazawa, H. Kato, S. Suzuki, J. Aoki, T. Kondo, H. Saito, Y. Gomi, S. Matsuda, and Y. Tadaki, "Multi-storied photodiode CMOS image sensor for multiband imaging with 3D technology", IEDM 2015

[19] Suzuki, Manabu; Suzuki, Masashi; Kuroda, Rihito; Sugawa, Shigetoshi, "A Preliminary Chip Evaluation toward Over 50Mfps Burst Global Shutter Stacked CMOS Image Sensor", Electronic Imaging, Image Sensors and Imaging Systems, 2018, pp. 398-1-398-4.

[20] Laurent Millet ; Margaux Vigier ; Gilles Sicard ; Wilfried Uhring ; Nils Margotat ; Fabrice Guellec; Sebastien Martin, "A 5 Million Frames Per Second 3D Stacked Image Sensor With In-Pixel Digital Storage", ESSCIRC, 2018, pp62-65.

[21] L. Millet et al., «A 5500FPS 85GOPS/W 3D Stacked BSI Vision Chip Based on Parallel in-Focal-Plane Acquisition and Processing». 2018 IEEE Symposium on VLSI Circuits, Honolulu, HI, 2018.

[22] L. Millet et al. « A 5500FPS 85GOPS/W 3D Stacked BSI Vision Chip Based on Parallel in-Focal-Plane Acquisition and Processing», IEEE Journal of Solid State Circuits 2019.

[23] M. F. Amir, J. H. Ko, T. Na, Duckhwan Kim, et S. Mukhopadhyay, « 3-D Stacked Image Sensor With Deep Neural Network Computation », IEEE Sensors Journal, vol. 18, no 10, p. 4187–4199, 2018.

[24] A.Carbon, J.-M. Philipe, O. Bichler, et al., « PNeuro : a scalable energy-efficient programmable hardware accelerator for neural networks » DATE 2018, pp.1039-1044

[25] https://github.com/CEA-LIST/N2D2/

[26] W. Maass, « Networks of spiking neurons: The third generation of neural network models », Neural Networks, vol. 10, no 9, p. 1659–1671, 1997.

[27] C. Brandli, R. Berner, M. Yang, Shih-Chii Liu, et T. Delbruck, « A 240 × 180 130 dB 3 μs Latency Global Shutter Spatiotemporal Vision Sensor », IEEE Journal of Solid-State Circuits, vol. 49, no 10, p. 2333–2341, 2014.

[28] G. Indiveri et S.-C. Liu, « Memory and Information Processing in Neuromorphic Systems », Proceedings of the IEEE, vol. 103, no 8, p. 1379–1397, 2015.

[29] B. Belhadj, A. Valentian, P. Vivet, M. Duranton, L. He, et O. Temam, « The improbable but highly appropriate marriage of 3D stacking and neuromorphic accelerators », ESWEEK, CASES, 2014.

[30] Y. Lecun, L. Bottou, Y. Bengio, et P. Haffner, « Gradient-based learning applied to document recognition », Proceedings of the IEEE, vol. 86, no 11, p. 2278–2324, 1998.

[31] X. Guo, X. Qi and J. G. Harris, "A Time-to-First-Spike CMOS Image Sensor" in IEEE Sensors Journal, vol. 7, no. 8, pp. 1165-1175, 2007.

[32] F. Gómez-Rodríguez, L. Miró-Amarante, F. Diaz-del-Rio, A. Linares-Barranco and G. J. Robotics, "Real time multiple objects tracking based on a bio-inspired processing cascade architecture," ISCAS 2010, pp. 1399-1402.

[33] P. Vivet, S. Thuriès, O. Billoint, S. Choisnet, D. Lattard, E. Beigné, P. Batude, "Monolithic 3D : an alternative to advanced CMOS scaling, technology perspectives and associated design methodology challenges" ICECS'2018, Bordeaux, France, Dec 2018