Physical and Cost Comparison of Smartphone Laser Autofocus Solutions

P.L. Bonanno

Yole Group, 55 rue de Leinster, 44240 La Chapelle sur Erdre, France
+33 2 40 18 09 16, peter.bonanno@yolegroup.com

***Abstract*— As interest in smartphone 3D imaging wanes, camera designers have not forgotten the importance of active autofocus to high-end smartphone photography. Accordingly, most manufacturers have abandoned 3D camera modules and opt instead for less costly laser autofocus-capable rangers such as the VL53L5. In contrast, Apple has stayed with the LiDAR module, which in addition to its 3D imaging functions is also fully capable of laser autofocus. While all these solutions rely on SPAD-based dToF, they have vastly different requirements and manufacturing costs. In this work, we perform teardown and detailed physical analysis of three smartphone laser autofocus solutions, the iPhone 15 Pro LiDAR module by Sony, VL53L5 from STMicroelectronics and TMF8821 from ams AG. We determine the physical and technical differences between the true 3D imaging camera and multi-zone rangers and then perform a reverse costing analysis to estimate the mass production costs of each of the three components.**

Keywords— Laser Autofocus, SPAD, dToF, Reverse Costing

I. INTRODUCTION

The years 2016 to 2019 saw an explosion in smartphone 3D imaging with over 20 smartphones confirmed by our teardown team to contain such technologies. During this period, a variety of players implemented a variety of solutions, including structured light, indirect time-of-flight and SPAD-based direct time-of-flight (dToF) systems. These capabilities presented novel possibilities such as improved augmented reality, object recognition, measuring lengths and surfaces, and creating 3D maps of rooms and objects. Also among these uses was to aid focusing in smartphone photography by providing distance information about the scene. This is often called "laser autofocus," and is especially useful in low-light situations where techniques based on visible light struggle.

Around 2021, things shifted in the Android camp. Smartphone makers decided that the expense and volume of these modules were not justified, and they fell out of use almost overnight. Laser autofocus, however, was seen as critical for premium smartphones, where the photography experience is paramount. This need came to be filled by small "depth ranging" modules, especially the FlightSense series from STMicroelectronics. Our teardowns show that the current ultra-premium smartphone landscape is dominated by multi-zone rangers such as the FlightSense VL53L5, which uses SPAD pixels grouped into functional zones to produce a low-resolution distance map of a scene, enabling simultaneous tracking of multiple subjects and scene depths. Ams AG has also introduced their multi-zone ranger TMF882X series, which has seen some success with Chinese smartphones.

This shift did not take place in the iPhone series. Apple continues to believe in full 3D imaging for smartphones and has kept both the front-facing Face ID system and, in their more premium models, the rear-facing dToF LiDAR (light detection and ranging) module. Recently, in the iPhone 15 Pro, the module has been rebuilt around a fully Sony-furnished system (all active dies are provided by Sony).

In this work, we present a physical and cost comparison of three laser autofocus solutions by Sony (iPhone 15 LiDAR), STMicroelectronics (VL53L5) and ams AG (TMF8821). Physical characteristics are based on analyses such as teardown, cross-sectional analysis and circuit delayering. We then use a simplified version of our reverse costing analysis to estimate and compare the costs of these solutions.

II. METHODS

A. Physical Analysis

Physical Analysis is performed to obtain information such as die dimensions, SPAD dimensions, process identification, optical layer chemistry and thicknesses, etc. Most of the analysis used for this article is not presented here but is available as part of our reverse-costing reports. We prepare cross sections and delayering samples using our confidential polishing, contrasting and etching processes. We then image these samples using optical and scanning electron microscopy (SEM) and characterize their chemistry using energy-dispersive X-ray (EDX) spectroscopy.

B. Costing Analysis

Reverse costing starts by using physical analysis from our laboratory to identify the technologies and determine the manufacturing processes used to produce each part of a component. These findings, when combined with other information such as supply chain analysis, consumable costs and insider information obtained by our market intelligence team, allows us to simulate the front-end, back-end and assembly manufacturing steps that occur during mass production. We use the best available information, often from industry insiders, but where we miss reliable intel, our experience allows us to substitute reasonable assumptions.

For the simplified version for cost comparison of the modules in this work, we use calculated wafer costs, potential good dies per wafer (PGDPW) and yields rounded to the most significant figure. This allows us to present and discuss a useful cost comparison and gives the reader a certain transparency of our assumptions and how they are incorporated into the cost model.

The simplified cost model for a fabbed die is

$$diecost = (wafercost / (PGDPW \cdot yield)) \quad (1),$$

where the *wafercost* is the cost to produce, test, and dice a wafer, *PGDPW* is the number of dies yielded per wafer if 100% of dies pass testing, and *yield* is the pass-rate of the test.

To create finished modules, the fabbed dies need to be assembled onto a printed circuit board (PCB) substrate along

with the housing and optics and undergo a final test. Some functionality and specs cannot be tested until this point, and so a significant number of components may be disposed of in this step. In practice, some costly parts of the BOM can potentially be salvaged, but our simplified model will assume that the module is totally lost on failing the final test.

The simplified cost model for an assembled and tested module is

$$(diecosts + BOMcost + assemblycost) / yield \quad (2),$$

Where *diecosts* is the cost of all active dies calculated by (1), *BOMcost* is the cost of all other components, *assemblycost* is the cost to assemble everything into a module and apply the final test, and *yield* is the pass rate of this test.

In practice, the profit margins paid to foundry partners, die suppliers and assemblers can more than double the effective cost of a die or component, which is why supply chain analysis is critical to a proper reverse costing. Our simplified model will include this parameter for wafer costs but ignore it otherwise.

III. RESULTS & DISCUSSION

A great deal of physical analysis is required to perform reverse costing, but here we will present a brief overview to give an idea of the important module components and SPAD processes. Selected findings are collected in Table 1 at the end of the section. All SPAD structures included in these modules have been presented elsewhere and the interested reader is invited to refer to the references list to learn more.

A. iPhone LiDAR

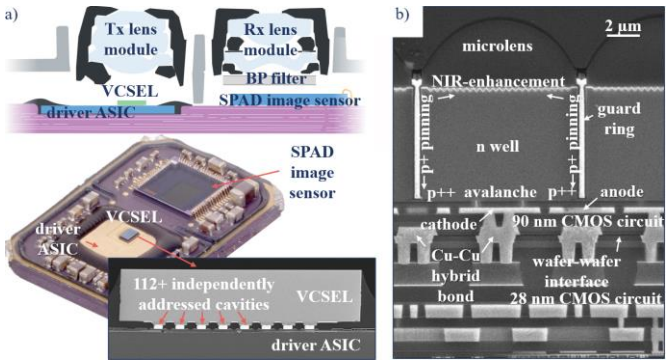

Fig. 1. The iPhone 15 Pro LiDAR module by Sony. a) Overview of the module. The schematic (top) shows the placement of the optics over the three active dies. Below is a view of the module with the optics removed to reveal the active dies on the PCB. A cross-sectional SEM image of the VCSEL bonding shows that each mesa is gold bumped to the driver ASIC. b) Cross-sectional SEM image through the pixel array of the SPAD image sensor. The functional regions and circuit connections are annotated.

Fig. 1 presents a physical overview of the iPhone 15 Pro LiDAR module. The 8x14 dot pattern is created by the VCSEL's 112 independently addressed cavities. The VCSEL is bottom emitting, meaning that emission is through the backside opposite the anodes. This allows it to be flip-chip gold bumped to the driver ASIC so that each anode can be independently addressed without wirebonds interfering with the emitted light. All the active dies are from Sony.

A cross sectional SEM image of the SPAD pixel is shown in Fig. 1b. The SPAD and circuit structure is well described by Sony's 2018 patent [1]. NIR-enhancement of the pixel is achieved by a light-trapping inverted pyramid array structure. [2] Cu-Cu hybrid bonding with 5 µm pitch connects the SPAD structure in the 90 nm CMOS sensing circuit to the operating transistors and readout in the 28 nm CMOS circuit below. The later circuit also performs histogramming.

To estimate the cost of the LiDAR module, we use the following parameters, which have been rounded to the nearest most significant figure: SPAD Image Sensor (*wafercost* = $5,000, *PGDPW* = 9,000, *yield* = 80%), VCSEL + driver ASIC (*wafercost* = $6,000, *PGDPW* = 7,000, *yield* = 70%), *BOMcost* = $3, *assemblycost* = $1 and *yield* = 90%. The VCSEL and driver ASIC are costed together, because the VCSEL testing is realized only after bumping to the ASIC. Plugging these values into equations (1) and (2) gives a module cost of $6.60.

B. VL53L5

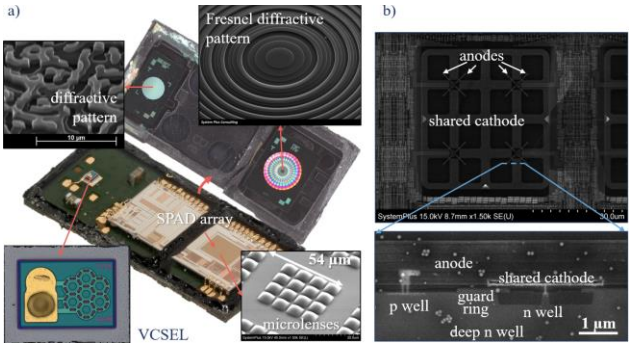

Fig. 2. The VL53L5 multizone ranging module from STMicroelectronics. a) Overview of the module. The module housing containing the flat optics is separated from the PCB and closeups of the diffractive patterns are shown. There is also a zoom on the VCSEL die and SEM imaging of the microlenses over the SPAD pixels. b) SEM-imaging of the SPAD pixel layout and structure. The functional regions and circuit connections are annotatedated.

Fig. 2 presents a physical overview of the FlightSense VL53L5 multi-zone ranging module from STMicroelectronics. The module has two active dies: the VCSEL and a SPAD ASIC that in addition to SPAD detection, also drives the VCSEL and performs histogramming. The optics are quite developed. VCSEL emission passes through a diffractive pattern that shapes the five beams into a smooth rectangular intensity distribution to efficiently illuminate the scene in the field of view. The scattered light returning to the module passes first through a flat optic with a Fresnel-like pattern that focuses the light onto the SPAD pixel array. The glass die bearing the pattern also has the bandpass filter. Each pixel then has its own microlens to maximize fill factor.

SPAD pixels are grouped into 4x4 macro-pixels, which share a cathode, but are independently addressed by their anodes. The anode track lengths are all identical and are independently addressed by quenching and pulse-shaping circuitry. While the diodes are independent, "multi-zone ranging" is achieved by grouping the diodes logically into "zones" of 4 or more macro-pixels to customize the trade-off between dynamic range and resolution. The SPAD structure and CMOS 40 nm technology has been presented. [3-5]

To estimate the cost of the VL53L5, we use the following parameters, which have been rounded to the nearest most significant figure: SPAD ASIC (*wafercost* = $2,000, *PGDPW* = 10,000, *yield* = 90%), VCSEL (*wafercost* = $2,000, *PGDPW* = 200,000, *yield* = 90%), *BOMcost* = $0.1, *assemblycost* = $0.2 and *yield* = 90%. Plugging these values into equations (1) and (2) gives a module cost of $0.60

C. TMF8821

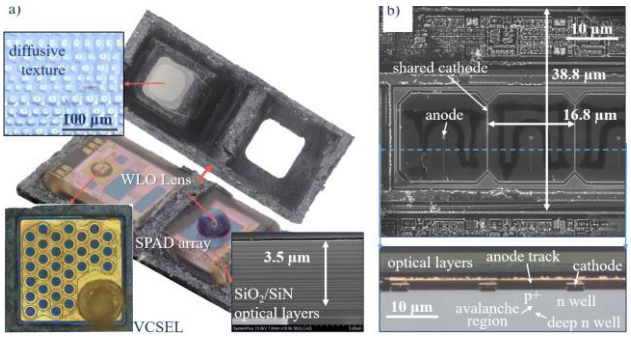

Fig. 3. The TMF8821 multizone ranging module from ams AG. a) Overview of the module. The module housing contains a microlens array to act as diffuser for the VCSEL emission. All other optics are on chip, including deposited optical layers that act as bandpass filter and a WLO lens on the SPAD array region. b) SEM-imaging of the SPAD pixel layout and structure. The top image shows the planar structure with the metal layers removed, and the bottom image shows the cross-section with the metal layers and optical bandpass layers intact. The functional regions and circuit connections are annotated.

ams's solution is the TMF882X multi-zone ranger series. The TMF8821 is presented in Fig. 3. The module leans on wafer-level optics (WLO) technology acquired from Heptagon and VCSEL technology acquired from Princeton Optronics. VCSEL emission is diffused by a WLO microlens array, and light scattered back from the scene is focused onto the SPAD array by a single on-chip WLO lens. Interferometric optical layers deposited directly onto the SPAD array act as a bandpass filter. The SPAD pixels are achieved using the high-voltage 55 nm CMOS DMOS process that has been presented previously. [6]

To estimate the cost of the TMF8821, we use the following parameters, which have been rounded to the nearest most significant figure: SPAD ASIC (*wafercost* = \$3,000, *PGDPW* = 10,000, *yield* = 90%), VCSEL (*wafercost* = \$3,000, *PGDPW* = 200,000, *yield* = 90%), *BOMcost* = \$0.02, *assemblycost* = \$0.1 and *yield* = 90%. Plugging these values into equations (1) and (2) gives a module cost of \$0.50.

D. Comparison

Key findings are summarized in Table 1. The LIDAR module is over 10x as costly to produce as the mult-zone ranging modules and also takes up a lot more space in the smartphone. Producing the dot pattern relies on a large and complex (low-yield) VCSEL and collecting it relies on a large, high-resolution detecting area. Accordingly, the VCSEL in the LiDAR has over 10x the total cavity area as that in the VL53L5, and the SPAD pixel array has around 5x the total detecting area. To accommodate these large areas, the optics are correspondingly larger and more costly.

With the TMF8821, ams AG seeks to offers a less clostly option for multi-zone ranging. They leverage their in-house WLO and VCSEL production capacity to make a module with very simple optics, that while less costly are surely less efficient than those in the VL53L5, especially at focusing light into the SPAD pixels. To further reduce cost, they opt for a detecting area of around half the size and make up for it by using a very large and powerful VCSEL with almost 5x more total cavity area than that in the VL53L5. Comparing the performance of the two ranging modules is beyond the scope of this work, but one can imagine that the ams module consumes more power.

TABLE 1. KEY FINDINGS OF PHYSICAL & COSTING ANALYSIS

	iPhone LiDAR	***VL53L5***	***TMF8821***
estimated cost	***\$6.60***	***\$0.60***	***\$0.50***
module vol. (LxWxH=V)	10.3x7.6x3.0 = 235 mm^3	6.4x3.0x1.5 = 29 mm^3	4.6x2.0x1.4 = 13 mm^3
VCSEL cavities	112 (indep.)	5	33
VCSEL total cavity area[a]	12,650 μm^2	1,150 μm^2	5,400 μm^2
transmission optic	refractive lens module	diffractive optic	multi-lens diffusor
emission shape	8x14 dot pattern	square	diffuse
baseline	3.0 mm	4.0 mm	2.4 mm
reception optics	refractive lenses, bandpass filter plate, on-chip µlenses, NIR enhanced	Fresnel-like flat lens, on-chip µlenses	WLO refractive lens, on-chip interferometric layers
SPAD process	90nm on 28nm CMOS with in-pixel DBI [1]	40nm CMOS [3-5]	55nm high-voltage CMOS & DMOS [6]
pixel pitch	10.1x10.1 μm^2	54x54 $\mu m^{2,b}$	38.8x16.8 μm^2
detecting area per pixel	68 μm^2 (µlens area)	1100 μm^2 (µlens area)[b]	132 μm^2
SPAD array resolution	10,672 (116 x 92)	140 (14 x 10)[b]	612 (34 x 18)
total detecting area	0.73 mm^2	0.15 mm^2	0.08 mm^2

[a.] Cavity estimated by aperture in the anode

[b.] Per 4x4 macro-pixel

IV. CONCLUSION

We conducted physical and costing analysis of three laser autofocus solutions, from Sony, STMicroelectronics and ams AG to understand and compare the technologies, design choices and mass production costs. As expected, the LiDAR module by Sony is vastly more costly to manufacture than the two multi-zone ranging modules, more than 10x as much in fact. This helps to explain the abandoning of 3D imaging by Android smartphones. Among the two rangers, the VL53L5 is the ultra-premium industry standard, and the TMF882X series leverages ams's unique strengths to challenge it on cost. We can't speak to performance comparison.

REFERENCES

[1] Sony Semiconductor Solutions Corp, "Sensor chip and electronic apparatus." JP2016231585A, 2018.

[2] S. Yokogawa, et al. "IR sensitivity enhancement of CMOS Image Sensor with diffractive light trapping pixels." Sci Rep 7, 3832 (2017). https://doi.org/10.1038/s41598-017-04200-y

[3] S. Pellegrini et al., "Industrialised SPAD in 40 nm technology," IEEE International Electron Devices Meeting (IEDM) 2017, doi: 10.1109/IEDM.2017.8268404.

[4] S. Pellegrini. "Industrialized SPADs in deep-submicron CMOS technology." ISSW 2018.

[5] Bruce Rae and Pascal Mellot, "A SPAD-based, Direct Time of Flight, 64 Zone, 15fps, Parallel Ranging Device Based on 40nm CMOS SPAD Technology." ISSW 2018.

[6] R. Kappel. "Multizone, Multiobject D-TOF System in 55nm." ISSW2018.