
Copyright © 2014 The Authors

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

Content must not be changed in any way or reproduced in any format or medium without the formal permission of the copyright holder(s)

When referring to this work, full bibliographic details must be given

http://eprints.gla.ac.uk/96924

Deposited on: 03 October 2014
On Evolution of CMOS Image Sensors

Luiz Carlos Piuva Gouveia  
School of Engineering  
University of Glasgow  
Glasgow, U.K. G12 8LT  
Email: Luiz.Gouveia@glasgow.ac.uk

Bhaskar Choubey  
Department of Engineering Science  
University of Oxford  
Oxford, U.K. OX1 3PJ  
Email: Bhaskar.Choubey@eng.ox.ac.uk

Abstract—CMOS Image Sensors have become the principal technology in majority of digital cameras. They started replacing the film and Charge Coupled Devices in the last decade with the promise of lower cost, lower power requirement, higher integration and the potential of focal plane processing. However, the principal factor behind their success has been the ability to utilise the shrinkage in CMOS technology to make smaller pixels, and thereby have more resolution without increasing the cost. With the market of image sensors exploding courtesy their integration with communication and computation devices, technology developers improved the CMOS processes to have better optical performance. Nevertheless, the promises of focal plane processing as well as on-chip integration have not been fulfilled. The market is still being pushed by the desire of having higher number of pixels and better image quality, however, differentiation is being difficult for any image sensor manufacturer. In the paper, we will explore potential disruptive growth directions for CMOS Image sensors and ways to achieve the same.

Index Terms—Image Sensors, CMOS, CCD, Wide dynamic range, 3D integration, hyperspectral imaging

I. INTRODUCTION

Electronic imaging has been flourishing for the last decade, having practically suppressed traditional imaging techniques of film-based cameras. In fact the evolution of digital imaging has been so fast that even a mature imaging technology like Charge Coupled Devices (CCD) has already been surpassed by CMOS based image sensors. The overall image sensor market as a whole has presented one of the largest compound annual growth rate of all electronic application markets [1].

In this paper, we review the evolution of CMOS image sensors and explore potential differentiators in image sensors of future. Section II reviews the evolution of CMOS image sensors till date. The principal differentiator till date has been that of the number of pixels and Section III presents approaches to continue this scaling of pixel count. However, with these reaching optical limits as well as straining the storage and bandwidth capacity, it is high time to explore other features of image sensors. Three such potential disruptors are presented in next sections in the form of dynamic range (Section V), speed (Section IV), spectral response (Section VI) and computation (Section VII).

II. EVOLUTION

The earlier digital cameras were almost exclusively made from Charge Coupled Devices as the manufacturing process of these devices is optimised for imaging. However, microelectronic manufacturing is expensive and hence the cost of these image sensors was typically high. This meant the produced digital cameras were expensive. In addition, these also suffered from high power consumption and little or no functional integration.

Most integrated circuits including computing and communication chips are however made in CMOS processes. Attempts have been made to manufacture imaging circuits in CMOS processes from early days of digital imaging. However, these sensors suffered from higher noise levels, which limited their ability to produce good quality images. Nevertheless, the feature size of CMOS processes has consistently reduced over the years following the empirical Moore’s law. This means that these processes offer the potential of making smaller pixels and thereby more pixels per chip. In addition, processes with lower feature size of transistors also led to lower power consumption of a typical transistor made in these processes. More importantly, however, with a larger number of integrated circuits being manufactured in these processes, the average cost of individual chips made in these processes is significantly lower than that of CCDs.

Therefore the low manufacturing cost, lower power requirement and potential integration with other functionality drove the development of techniques to improve the image quality of image sensors build in CMOS process. This was achieved by development of CCD like buried photodiode in the CMOS process, which enabled Correlated Double Sampling (CDS) for reduction of temporal noise in the CMOS Image Sensor (CIS). Figure 1 shows a typical active pixel sensor circuit in a CMOS process [2]. In addition to the buried photodiode, this pixel has a reset transistor (M1), a buffer source follower (M2) and a select switch (M3). This pixel led to a surge in low cost digital imaging and its embedding in devices like toys, mobile phones as well as computers.

Another method to improve an image quality is to increase the spatial resolution in the imager. The effective resolution of an image sensor is calculated from test charts and Modulation Transfer Function (MTF) test results. However, it is a common practice to define it in terms of number of available pixels in the sensor. The number of pixels in an image sensor has constantly increased over the years from few thousand pixels to several millions of pixels per chip [3], [4]. In addition to the sheer resolution of images, high resolution imagers also...
allow for additional functionality like pixel binning to improve performance in low light and windowing or digital zoom, thereby enabling cameras with limited optical components. The later feature has been the driving force behind mobile phone cameras.

III. RESOLUTION

The leading market differentiator in digital imaging till date has been the number of pixels in a camera. Despite some saturation in the market, the pixel count is still one of fundamental sales pitches for any image sensor. In order to further increase CIS resolution without excessively compromising pixels’ charge recording capacity (or Full Well Capacity - FWC), pixel sharing is becoming a increasingly popular design. Pixel sharing aims to share parts of the pixels that can be used by other pixels in the vicinity, increasing the detector area and, therefore, the pixel’s charge capturing capacity. This is particularly suitable for use with 4-T APS pixels, where the reset, the source follower and the row selection transistor can be shared while the TX gate is used to isolate the photo-diode of each pixel. Existing sensors present 2.5T, 1.75T and 1.5T pixel sharing configurations, where 2, 4 and 6 adjacent pixels share the common pixel transistors.

Photo-diode peripheral utilization method can also be used to increase the FWC [5]. For technologies with low feature size, the capacitance density of photo-diode periphery is greater than of its area. Therefore, by making small openings on the photo-diode area, one can increase the lateral area increases further despite reducing the surface area. Apart from pixel topologies and layout techniques, more fabrication specific methods are also being applied. The most notable is the Back Side Illumination (BSI) sensors instead of the more traditional Front Side Illumination (FSI) pixel design [6]. The metallic routing required by the pixel’s electronic components interferes with the light path in FSI imagers. Figure 2 shows a typical metal stack on top of an image sensors as well as the non-photosensitive part of the pixel. Inverting the light incidence as shown in figure 3 provides an opportunity to increase the fill-factor as well as allow for better connectivity between pixels. Such BSI pixels, however, require thinning of the substrate and application of anti-reflective layers.

Other techniques with the potential of improving the pixel’s optical properties may involve the development of light guides (or light pipes[7]), reduction of the metal layer stack on FSI sensors, and photo-diode stacking. The former technique utilises the light wavelength penetration depth [8]. By stacking 3 photo-diodes, each individual pixel can capture information related to three primary colours and thus avoid using three of more pixels needed for the Bayer colour pattern. Alternative detectors like quantum dots have also been proposed as potential solutions [9].

As the pixel pitch shrinks bellow the length of the vertical light path, cross-talk between the pixels becomes more prominent. Source of cross-talk include the light diffraction from metal layers (for FSI designs) and the light diffusion on the substrate. Light-guides (for FSI), microlenses and Front-side Deep-Trench Isolation (F-DTI) [10] can improve the optical properties. Front-side Deep-Trench Isolation not only electrically isolated pixels by creating a SiO$_2$ wall but also acts as light reflector due to the difference of Si and SiO$_2$ refraction indexes.

Typically, increasing the spatial resolution comes at the cost of reduction in the pixel size to keep the imager size constant (to fit established formats). However, some applications, particularly biomedical and scientific imaging, require large image sensor formats. One application is that of x-ray imaging, where lens technology does not allow high optical gains. Large format image sensors are also required when the image quality loss due to pixel reduction and lens design is prohibitive (for example in astronomy or aerial reconnaissance imagers) or to enable photographic features like shallow depth-of-view. Although the definition of large format imagers varies according to the application sector, one can consider any image sensor larger than 4x5 inches as a large format sensor.

Such large format image sensors have two challenges. One is that of the readout due to their excess area, which leads to increased parasitic capacitance [11]. Secondly, as the die of the image sensor is limited in size by the wafer size itself,
increasing the format of such image sensors requires multi-chip stitching [12].

IV. HIGH SPEED

Although the image quality is the main characteristic of still image cameras, the speed of image capture is another crucial parameter, particularly for video applications. ITU high-definition recommendation 4k UHD and 8k UHD (Rec. 2020[13]) requires both high resolution – 8.3 and 33.2 million pixels respectively – and high speed capture – up to 120 fps. Beyond high-definition video, high speed image sensors are also required for machine vision, 3D vision and scientific applications for instance. For such applications the speed spans from hundreds to millions of frames per second.

The speed of a CIS is typically increased by speeding-up the signal chain on a Multiple-Inputs, Single-Output (MISO) system. Successive improvements have been based on the analogue-to-digital conversion (ADC) placement on the signal chain: from initial analogue output systems (with digitalization performed off-chip) to most common column-based ADCs to in-pixel ADC CIS in increasing digital parallelism. Multiple outputs can also be used to increase the total output throughput as well as different ADC architectures, from slower slope ADC to faster SAR-ADCs [14] and sigma-delta ADCs [15].

For very fast acquisition CIS, internal data storage is required to implement temporary frame storage. This storage – usually located inside the pixel – allows for burst mode operation where a small number of very fast sequences of images are captured [16]. Tweaking the fabrication process can also help the design of very fast CIS. For example, in a time-of-flight 3D vision CIS, the photo-diode can be implemented with a specific vertical diffusion profile to allow for fast charge transfers from the photo-diodes [17].

V. DYNAMIC RANGE

Conventional CMOS Image Sensors provide a limited dynamic range (typically about 3 decades of intensity). However, this is often insufficient for natural scenes containing both dark and bright regions. Such scenes can be captured by the human retina as it can dynamically sense up to 120 dB through adaptation. Several approaches have been proposed to address this limitation. Most common technique uses multiple image captures using different exposure times with conventional CIS [18], [19]. The final image is formed by post-capture computational algorithms. Multiple capture however leads to limited frame rate and blurred images. Methods to avoid those limitations involve the use of specific pixel designs, including logarithm pixels [20], [21], multi-mode operations [22], capacitance adjustment, frequency-based and time-based pixel operation and selective integration time [23], [24].

The low light performance of CMOS image sensors is limited by the leakage current in the diode available in the process. Many applications in automotive, security and surveillance, medical and scientific fields, however require capturing images with very low light intensity. One technique to improve the performance is to cool the CIS using cryogenic devices. This can also significantly reduce the noise levels and therefore, increase the low light sensitivity. However, it is not practical in many applications. Leakages in CMOS diodes owe their origins to the thermal generation-recombination followed by drift and parasitic leakage currents due to defects near the isolation regions. For the latter case, diffusion implants or a polysilicon layer surrounding the photo-diode[25]. The increase in sensor resolution tends to increase the noise problem, particularly the low frequency noise (1/f) and Random Telegraph Signal noise from the source follower transistor becomes prominent at these levels [26]. High-gain column-level amplifiers and Correlated Multiple Sampling (CMS); however, can be utilised to reduce the effects of these noise sources[27].

These techniques are useful to reduce the noise and improve the low light sensitivity, but not to the point of being able to measure individual photons. For this purpose, special imager have been developed such as electron multiplier CCD [28] and single photon avalanche detectors (SPADs)[29]. SPAD pixels works in Geiger mode, where each incoming photon trigger an avalanche effect on the near breakdown-biased photo-diode that generates a current impulse. This impulse then updates a digital counter and the final value of this counter over a time windows determines the amount of photon collected. SPADs development is still in experimental phase and faces challenges of increasing fill-factor and reducing the dark counting rate, jitter, low resolution and operational voltage.

VI. SPECTRUM SENSITIVITY

A. Colour sensitivity

In most current CIS applications, the extraction of colour is fundamental, not only for consumer application but also in a variety of control and detection cases such as spectroscopy. Colour detection, separation and fidelity is a required feature for such applications. The traditional way to implement colour detection is to apply three different organic colour dye filters onto the same pixel design. The transmission coefficient ranges from about 70-90% on the band pass of the filter and about 30-50% on overall visible spectrum [30]. These filters are applied on adjacent pixels so each individual pixel turns into a colour channel. When grouped into three primary colours set, it impacts the resolution, the overall light absorption (QE) and colour crosstalk.

A traditional colour unit uses a set of 2x2 pixels with a Bayer pattern where one pixel is the blue channel, another is the red and the last two make the green channel (RGGB), as shown in figure 4. This implies a quarter of the original monochromatic imager and less than half its QE. To keep the same resolution and QE, diode stacking can be used with the current trade-off in crosstalk [31]. Placing the colour filter as close as possible of the photo-diode, designing coloured micro-lenses and pixel isolation helps to reduce the (optical) cross talk on Bayer patterns. To further increase the quantum efficiency, other colour pattern rather can be used such as incorporating white (clear) light channels (RGBW/RGBC) or the complete removal of green filters (RCCB).
colour splitters are another alternative to maintain the QE is using colour splitters rather than colour filters[30]. On every second pixel a splitter is placed resulting in white minus read (W-R) for this pixel and white plus red (W+R) for adjacent pixel. The pattern is obtained with a similar pattern with blue splitter and further post-processing. More experimental approaches include implementation of selective colour absorption using nanowires [32] and plasmonic colour filters [33].

**B. Beyond visible spectrum**

Most image sensor applications correspond to the human retinal response to a scene and, therefore, most CIS have been employed to light detection on the visible spectrum (typically from 400nm to 700nm wavelength band). This has also been made possible by the spectral characteristics of silicon, as shown in Figure 5. Nevertheless, ability to detect beyond the visible spectrum will be a defining feature of next generation image sensors.

CIS are now directly sensitive to X-rays over a good energy band without the need for a scintillator. This is due to the reduction in thickness of the polysilicon layer in typical CMOS process. As stated earlier, the application of CIS as radiography detector requires the use of large format sensors due the difficult to design lenses for X-rays.

![Bayer Pattern of filters for colour imaging](image1.png)

**Fig. 4. Bayer Pattern of filters for colour imaging**

Integration of infrared (IR) detectors, will enable a range of new applications for CMOS image sensors. IR detectors are used to trace sources of heat. Typical applications include military and security (night vision), health (temperature checking) and industrial and engineering (gas and heat leakage). As silicon is sensitive to near infrared region, it can be used as an IR detectors. On the other hand, this may affect the visible performance of cameras and hence typical digital cameras often use an infrared cut filter to avoid IR detection. For detection of mid- and far-IR spectrum, one requires integration of other materials with silicon. For mid-IR, materials include intrinsic semiconductors such as MTC and InSb, extrinsic varieties as Si:As and Si:In, black silicon and micro-bolometers (using either VOx or amorphous silicon).

Direct integration of these materials with CMOS process is difficult, if not impossible due to cost. A potentially simple solution is to use 3D integration of CMOS based visible detectors with IR detectors using techniques like flip-chip bonding. For Far-IR (also known as Terahertz) imagers, options include integrated antenna-based [34] and meta-material/bolometer-based approaches [35].

**VII. COMPUTATIONAL IMAGING**

Despite the advances in CMOS imaging, one of the promises of CMOS Image Sensor was the ability to undertake image processing at the acquisition stage itself as one can design circuits to do so in CMOS technology. However, the pixels of present day camera can at best provide a buffer in term of signal processing. This is so as the majority area of a pixel has to be devoted to a photodiode to capture the incident light, leaving very little for transistors circuits, as shown in the typical layout of figure 1.

A number of focal plane array circuits have been proposed; however, these circuits have a large number of transistors per pixel which in turn reduces the optically sensitive area and hence the fill factor of the pixel. As a result these pixels with very poor optical response. In addition, the reduction in the cost of individual cameras means that systems are being deployed that contain many cameras. Unfortunately, it is increasingly common for the amount of image and video data that is captured by a system to be so large that it is impossible to gather or store all the data and/or extract all the useful information contained in the data that is gathered or stored.

Increasing the number of pixels in a camera has also put excessive stress on the analogue to digital converter used to convert the analogue output of pixels to a digital form suitable for transmission and storage. To some extent, this has been mitigated by placing one or even two converters in each column. However, with multi-million pixel cameras, even these approaches fail to obtain high speed images. This is further problematic when using large focal plane stitched imagers for biomedical application like the x-ray imaging.

To improve the performance of these pixels, our group is currently working in a new imaging stack, wherein we will utilise back-side illumination of the sensor chip using thinned wafer, as shown in Figure 3. These back-side-illuminated (BSI) pixels are more efficient than conventional pixels, and so despite their additional cost, they are slowly becoming a mainstream technology. Further reductions in the amount of data generated by the image sensors will be obtained using
flip-chip technology to bump-bond a back-side illuminated image sensor containing an array of pixels to a second chip containing other circuits, as shown in figure 6. The first approach is to bond a small array of pixels per bump bond. The simplest design is to connect the image sensors to arrays of Analogue to Digital Converters. The bump-bonding technology means that there will be one ADC for each small array of pixels. This will immediately increase the speed of conversion, resulting in an ability to gather images at higher frame rates, particularly for multi-million pixel image sensors. Furthermore, this will provide more space for design of ADC compared to a conventional column parallel design, providing us with the ability to design better (faster, lower power, lower noise and higher resolution) and even intelligent converters. A more interesting design, however, will be to design image processing circuits onto the second chip.

Fig. 6. Three dimensional stacking of image sensor and processing circuits

Such designs will also provide us an opportunity to explore comparative pixels as proposed above by our group [36], [37]. These pixels utilise a comparator inside each pixel, as shown in figure 7. An externally applied reference voltage is used to stop the recording of integrating pixel voltage. By suitably designing the pixel reference, one can obtain any monotonically increasing transduction function from this pixel. Of particular interest to us is the tone mapped response, wherein a wide dynamic range image can not only be captured, but directly displayed as well [38].

Another approach to reduce the stress on readout and bandwidth with increasing pixel count is to compress the image. Traditional compression algorithms have been implemented on-chip such as Discrete Cosine Transform (DCT) [39] and predictive coding [40] or more specific, feature extraction-based [41]. However those traditional compression algorithms are a post acquisition computation. In other words, the image is firstly captured in full resolution and compressed afterwards, penalizing both energy and time consumption.

Compressed sensing (CS) has a different approach for compression. The compression of the image is performed on-the-fly, i.e., at the same time it capture the image [42], sampling only the important k components of the image. No previous assumptions on the image characteristics are required other than its sparcity. Compressed sensing outputs are composed of a series of M “measurements” (y_{Mx1}) produced by some transformation (measurement matrix \Phi_{MxN}) onto the image with N pixels (x_{Nx1}) with k < M << N. It turns out that the best measurement matrix \Phi has random coefficients. In CS the computational overhead is transferred to the decoding part of the process.

As the compression and image acquisition are inseparable, traditional CIS cannot be used and specific CIS designs have to be used. One of the first CS oriented sensors are single-pixel sensors associated with an array of Digital Micromirror Devices (DMD)[43]. Implementations using more conventional multi-pixel arrays have been implemented using coded aperture cameras, switched-capacitor integrators[44] and column \Sigma\Delta ADCs [45]. Such sensing has been applied to a wide range of applications but more specifically where images are known to be sparse as in medical (MRI and CT), security (radar) and astronomy.

VIII. CONCLUSION

In this paper, we have reviewed the evolution of CMOS image sensors with a view on potential future impact of the technology. CMOS Image Sensors utilise the reduced cost of chip manufacturing in typical CMOS processes. They also utilise the reducing feature size of these processes as per the Moore’s law. This has enabled them to continuously increase the pixel count. However, this alone is no more a good differentiator for image sensors. Therefore, different functionality is required from next generation of image sensors. In this paper, we have reviewed the potential approaches to wider dynamic range as well as focal plane processing.

REFERENCES


