100 years of photon-counting: the quest for the perfect eye

Sébastien Morel\textsuperscript{1} and Swapan K. Saha\textsuperscript{2}
\textsuperscript{1}: European Southern Observatory, avenida Alonso de Córdova 3107, Casilla 19001, Vitacura - Santiago, Chile
\textsuperscript{2}: Indian Institute of Astrophysics, Koramangala, Bangalore, India
(e-mail: \textsuperscript{1} smorel@eso.org; \textsuperscript{2} sks@iiap.res.in)

1. Introduction

The problem of the nature of light has been addressed for a long time. In the 17th century, René Descartes, after having described the geometrical laws of optics, suggested that the light would consist of moving particles. Later, Isaac Newton agreed with this idea. However, it was dismissed at the same time by Christian Huygens, and later at the beginning of the 19th century by Augustin Fresnel, who established the laws of wave optics. The corpuscular hypothesis was rehabilitated thanks to Max Planck’s works: in order to determine the black-body radiation law for short wavelengths, Planck suggested in 1901 that radiative transfers would occur through “packets” of energy named “quanta”. In 1905, Albert Einstein [1] explained the photo-electric effect: when exposed to light, certain materials emit electrons (called “photo-electrons”). The photo-electric effect was first observed by Heinrich Hertz in 1887. From the idea of Planck, Einstein described light as consisting of elementary particles called \textit{lichtquanten}, then “photons”.

The existence of photons means that, for a given collecting area, there exists a physical limit on the minimum light intensity for any observed phenomenon. The perfect detector would then be one that is able to detect individual each photons (or “photo-event”) in an image plan, thus giving the maximum possible signal-to-noise ratio. This ability is called “photon-counting”. We will see that such a detector can almost be built, by approaching the ideal “quantum efficiency” (ratio of the number of counted photons over the actual number of incoming photons) QE = 1, and the ideal “false detection rate” (or electron noise) FD = 0.

The main applications of such detectors is in astronomy. However, progress in high-energy (gamma) particles have been crucial for the development of photon-counting in the visible spectrum. We therefore report in this review the history of photon counting systems, mostly in the field of astronomy.

2. Photon-counting sensors and photon-counting cameras

In this article, we will refer as “photon sensor”, a device that produces a sole signal if a photon of the visible spectrum has been detected within its field of view, regardless of its angle of incidence. Hence, such a device is able to count how many photons it has received, but cannot give any information about the angle of incidence for each photon. The classical charge-coupled device (CCD), a light sensitive electronic chip made up of p-type silicon, is not a photon-counting detector, even if one photon produces one electron with QE ≈ 0.9, and having analog-to-digital converter set where one step (ADU) corresponds to one electron. This is because their readout noise is too important to determine whether a photon (or no photon at all) has been received in an image element when the light level is low (less than 1 photon per pixel and per frame), i.e. FD ≫ 0.

One way to characterize a photon-counting device (sensor or camera) is its “pulse height distribution” (PHD). It is the statistics of the intensity of the output signal (usually expressed in
electrons) triggered by a photo-event. The PHD is represented by plotting a curve number-of-counts vs. output signal intensity. Usually, a PHD curve shows a peak which can be characterized by its normalized peak-to-valley (PV) ratio, and its normalized full-width at half-maximum (NFWHM). Ideally, PV should tend to ∞ and NFWHM to zero. Of course, the intensity corresponding to the peak has to be much larger than the maximum level of readout noise of the sensor or camera (anode, CCD chip, ...) whose output signal is the photon information carrier and that terminates the chain in the photon-counting device. In these conditions, FD → 0.

![Diagram](image)

**Figure 1:** Example of a usual PHD. The dashed curve is the noise statistics of the detector readout. PV = P/V and NFWHM = W/S. The usual PHD is a Dirac’s peak.

Note that it is difficult to write a “linear” history of photon-counting imaging. The photon-counting cameras for low-energy photons, as we know them today, are the result of parallel progress in different fields of research (e.g., gamma-ray imaging, night vision, photometry) which did not directly attempt at low-energy photon counting imaging, therefore it separately brought the elements (micro-channel plates, image intensifiers, sensitive position anodes, etc...) that we find today in photon-counting cameras. Basically two problems had to be overcome, before the integration would be achieved, i.e.,

1. detecting photons individually,
2. measuring the photon positions in an (x, y) focal plane.

Two parallel ways were therefore possible:

1. From an image intensifying device, increase the gain to reach the quantum limit (i.e., to solve problem 1).
2. From a single element photon sensor (photomultiplier), increase the image resolution by miniaturizing and multiplying the basis element (i.e., to solve problem 2).

Actually, modern photon-counting cameras (whatever the technology they use) inherit from both approaches that were used alternatively (for example, from photomultipliers to MCPs, or from low gain MCP-equipped imaging devices to photon-counting cameras).

Besides the sensitivity and the image resolution, another aspect that had to be improved was the temporal resolution. High angular resolution (HAR) optical astronomy (speckle [2], long-baseline
interferometry [3]) requires to know the date of occurrence of each photo-event within less than 20 ms. It is notable that the first photon-counting cameras, which though able to reach the temporal resolution of a TV camera (50 Hz typically), were rather used as still cameras, employed for classical long exposure time astronomical applications.

3. History of photon-counting sensors

3.1. Gas detectors

In 1916, Elster and Geitel [4] imagined a device to detect photo-electrons, based on the design of the α particles counter by Rutherford and Geiger [5]. It consisted of a gas-filled bulb, containing a cathode and an anode. An electron inside the bulb could cause, by ionization of a gas, a discharge (measured by an electrometer), and therefore be detected, because of the high voltage between both electrodes. The Geiger-Müller device [6] for counting radioactive particles, based on the previous system, was adapted in 1932 by Locher [7] to count visible photons. He built several counters consisting of a gas bulb in which the cathode had a three-quarters-of-cylinder shape (coated with photoelectric material on the inner side to form a “photocathode”), and the anode was a simple wire aligned on the cylinder axis (Fig. 2). Several alkali materials, associated with hydrogen were used as photocathodes. Potassium was the more sensitive, but also the more noisy (the measured dark-count was 122/min., whereas it was 5.8/min for cesium, and 4.7/min for sodium). The best gas was helium because of its high minimum ionization potential. For all these counters, the maximum rate, limited by the reading electronics was less than 300 counts/s. Because of their weak precision, gas detectors as stellar photometry sensors were abandoned in the 1940s.

3.2. Secondary emission at work: the photomultiplier tubes

Early research on devices, that are able to emit a large number of electrons for each electron received, attempted to detect the faint luminous flux produced by photoelectric effect. This led to the construction of “photomultiplier tubes” (PMTs), yielding a measurable electric current from a
faint light source. PMTs are still manufactured today. They use the property of certain materials emitting several electrons when hit by an electron. This phenomenon is called “secondary emission”. By cascading several elements made of such materials called “dynodes”, one can obtain average high electron multiplication factors. An electron from the photocathode, ejected by a photon, creates a “snowballing” electron cloud through the dynode path. These electrons end up hitting an anode emitting several electrons when hit by an electron. This phenomenon is called “secondary emission”. whose output current is large enough to be measured. Guiding electrons in the tube is done by a strong (many kV) electrostatic field. The history of the PMT started with the discovery of the secondary emission. The first device to use secondary emission was the “Dynatron”, a negative resistor system used for oscillators, designed by Hull [8] in 1918. One of the first electron multipliers based on secondary emission was imagined by J. Slepian in 1923. Slepian’s design inspired V. Zworykin (the pioneer of television tubes), who built in 1936 a 12-dynode electron multiplier [9]. These dynodes were made of a silver + zirconium + cesium mixture. The gain of his PMT was claimed to be $10^{12}$ (which seems a little exaggerated).

![Photocathode](image.png)

Figure 3: (a) MS-10 photomultiplier, (b) continuous dynode photomultiplier, and (c) micro-channel.

In 1940, the Société Française Radioélectrique company introduced a PMT named “MS-10” [10]. It featured 10 dynodes made of Ag-Cs$_2$-O-Cs providing a gain ranging from 4,000 to 12,000. This tube used a magnetic field (10 to 20 mT) in order to apply to electrons a Lorentz’s force making them “bounce” from one dynode to the next (Fig. 3-a). Applications of this tube have been used in mechanical-scanning of television (using the Nipkow’s disk). Though the MS-10 system was quickly discarded, it was a step towards miniaturization of the PMTs, thanks to the idea of using a magnetic field.
In 1960, Heroux and Hinterreger [11] simplified the design of the magnetic PMT by using only one dynode which consisted of a coating on a long plate. The electrons bounced and were multiplied on this plate during their travel from the cathode to the anode (Fig. 3-b). G Goodrich and W Wiley, from the Bendix company, built in 1961 a similar device [12], a few millimeters thick, providing a $10^7$ electron gain. The miniaturization process was going on well. We note that the continuous dynode multiplier idea was already patented by Farnsworth in 1930 [13].

In 1962, Goodrich and Wiley achieved a major breakthrough by inventing the “micro-channel” [14], a simple glass tube whose inner side was coated by a secondary emission semi-conductor (Fig. 3-c). Goodrich and Wiley noticed that the electron gain of a micro-channel does not depend on its diameter, but on the ratio $a = \text{length/diameter}$, in a proportional way. Their prototypes had a diameter of 0.56 mm with $a$ ranging from 40 to 100. With such dimensions, parallel assembling of micro-channels in arrays for image intensification became feasible. However, the gain of micro-channels is limited by the positive charges left by the secondary electron cascade, which goes against the electric field that is applied at the ends of the micro-channel [15]. The gain can even decrease if $a$ increases [16]. The maximum electron gain of a micro-channel is a few $10^4$.

3.3. Bundles of photon-counting sensors

The Digicon [17] was one of the first alternatives to Lallemand tubes (see Sect. 4.2) for astrophysics, and provided photon-counting ability. Since it measured photon positions in one dimension, it was not a real imaging device. Spectrometry was thus its main application. Instead of bombarding a photographic plate, accelerated electrons in the Digicon collides with an array of photodiodes (38 for the first Digicon prototype). For each diode, the binnarized signal resulting from the collision incremented a 16-bit register. A Digicon, with a larger number of diodes (viz. an improved resolution) has been used in the Hubble Space Telescope. We can also notice that Herrmann and Kunze [18] presented in 1969 a photon-counting spectrometer working in the UV, and featuring an array of 40 miniaturized PMTs.

From the Digicon concept, Cuby et al. [19] investigated the “electron-bombarded CCD” concept at the end of the 1980s. It consists of a CCD array placed in a vacuum tube with a photocathode. Electrons, that are ejected from the photocathode by photons, are accelerated by a 25-kV voltage to make them hit the CCD pixels. Each accelerated electron liberates a charge of around 7500 electrons in the CCD. The CCD they used was a Thomson TH31513 linear array. The characteristics of the PHD were $\text{PV} = 0.33$ and $\text{NFWHM} = 0.22$. Since these devices were not used for HAR astronomy, they were quickly replaced by high-performance CCDs.

4. Towards more sensitive imaging devices

4.1. History of image recording in astronomy

Hand-drawing from eye observations had been used in astronomy since Galileo. Photography was used in astronomy as early as 1850, when W. Bond and J. Whipple took a Daguerreotype of Vega. Silver bromide dry emulsions were used first around 1880. However, one problem of photographic plates was the “dark background” effect: if a very faint object was observed, and whatever the exposure time was, the object was drown in a background from the plate brighter than the object. To overcome this problem, scientists imagined how to make the observed objects “more luminous” before the plate. Also, another problem for astronomy was to measure (and not eye-guess) the flux in each point of the plate (whether it represented a field image or a spectrum). The microphotometer was imagined by E Pickering in 1910 for that purpose [20]. Photographic plates were scanned
by the microphotometer which locally measured, using a photoelectric cell, the intensity from the illuminated plate [21]. Later, television cameras, directly mounted on telescopes, were imagined to replace the photographic plate + microphotometer systems [22].

4.2. The Lallemand tube

On the 6th July 1936, at the Science Academy of Paris, André Lallemand introduced a new imaging device he had invented [23]. It consisted of a 35 cm long glass tube, with a potassium photocathode at one end, and a 8 cm diameter zinc sulfide monitoring screen at the other end (which could be replaced with a photographic plate for recording the image). Electron focusing was performed by an electrostatic lens made by an inner silver coating on the tube, and by a magnetic lens consisting of a 10 cm diameter coil fed with a 0.5 A current. The accelerating voltage inside the tube was 6 kV, giving intensified images by the collision of accelerated photo-electrons onto the screen or the plate. Lallemand already thought about the possibility of building second generation (MCP-equipped) image intensifiers, because he suggested to replacing the plate with an electron multiplying device and mentioning Zworykin's works on this subject. However, Lallemand tubes, which used the original photographic plate system (to bring the image intensity above the dark background level of the photographic plate), has been regarded as “remarkable” and widely used in astronomy since 1936.

![Diagram of a Lallemand tube](image)

Figure 4: A Lallemand tube used for astronomy. The ring magnet 1 is used to move the hammer and break the glass bulb containing the photocathode. The ring magnet 2 is used to bring the photocathode behind the input window. The ring magnet 3 is used to change the photographic plate.

The first Lallemand tube to be commissioned on a telescope was described in 1951 [24]. Similar tubes (Fig. 4) have been employed until the beginning of the 1970s for faint object imaging. However, the sensitivity of Lallemand tubes did not reach the quantum limit because of the dark threshold of the photo plates. Lallemand tubes were made of glass and therefore very fragile and not convenient for operation on telescopes.

4.3. Image intensifiers

The early Lallemand system, using a monitoring screen (usually named “phosphor”), onto which the energy of each accelerated electron from the photocathode was converted into a burst of photons (spot), was actually imagined in 1934 by Holst et al. [25] from the Phillips Company (The Nether-
lands). They used a simpler system by using “proximity” focusing without electronic or magnetic lenses: the photocathode and the phosphor were merely spaced by a few millimeters. This gave a poor spatial resolution. However, many image converters for infrared imaging were built during World War 2, following their design [26]. Works of Lallemand and Holst et al. led therefore to the industrial production of “first generation” (‘Gen-I”) image intensifiers, from the 1950s. To overcome the problem of photon gain limitation of these intensifiers, cascades of Gen I intensifiers were used for high sensitivity cameras. Note that these devices did not allow photon-counting even if the phosphor was placed before a TV camera to quantify the signal: the PHD of a Gen-I intensifier has a negative exponential shape with no peak. Moreover, for a Gen-I intensifier, the output energy that is generated by any photo-event is statistically weaker than the electronic noise of a TV camera. A decisive step was brought by the application of the micro-channel to the image intensifier.

Research on how to build arrays of micro-channels (see Sect. 3.2), known today as “micro-channel plates” (MCPs), were carried out in the 1960s. In 1969, teams of Manley [27] from the Mullard Research Laboratory (UK), and of Eschard [28] from LEPA (France) introduced operational MCPs, ready to be mounted in image intensifiers that will be called “second generation” (Gen-II). Progress in miniaturization had gone on: the channels of Manley’s MCP had a 40-micron diameter. Compared to Gen-I image intensifiers, Gen-II offer a larger gain, but a smaller quantum efficiency, due to the fact that some electrons, ejected from the photocathode by a photon, do not enter any micro-channel. The right combination consists in a cascade of a Gen-I intensifier to have a good QE for photo-event detection, then a Gen-II that intensifies the spot at the exit of the Gen-II.

There are also now “third generation” (Gen-III) image intensifiers. They are actually Gen-II (MCP-equipped) image intensifiers, with an AsGa photocathode that offers a larger QE (0.3) than multi-alkali photocathodes. AsGa photocathode for Gen-III image intensifiers were introduced by Rouaux et al. in 1985 [29].

One recent alternative to the MCP is the microsphere plate [30]. It consists of a cluster of glass beads whose diameter is about 50 μm each. These beads have a secondary emission property. Electrons are therefore multiplied when they cross a microsphere plate. Compared to MCPs, microsphere plates require a less drastic vacuum (10^{-2} Pa), a reduction of ion return (see Sect. 6.1) and a faster response time (about 100 ps). The drawback is a poor spectral resolution (2.5 lp/mm). Hence, they can be used for PMTs, but not for image intensifiers.

4.4 Intensified TV cameras: the first visible photon-counting cameras

Let’s go back to the early 1970s. Any TV camera (generally using either Isocon or SEC technology) that was equipped with a stack of cascaded Gen I + Gen-II image intensifiers was able to produce TV frames at the quantum limit, hence in photon-counting mode. However, because of the sparse nature of photon-counting mode images, it was necessary, in order to optimize the data processing, to convert each frame into a set of (x, y) coordinates of the photo-events. Nevertheless, photon-counting cameras without photon coordinate extraction were mostly used as still cameras by integration of the TV frames. Such system were employed in the 1970s. For example, Lowrance et al. [31] designed an intensified TV camera system for photon-counting, in which the output signal was digitalized and added to an image memory with 16 bits per pixel.

The first photon counting camera using the coordinates of the photo-events for the visible spectrum was designed by Boksenberg and Burgess [32] in 1972. It was similar to the systems like the one built later by Lowrance, except that each time a photo-event was detected in a frame (when the TV signal was above a threshold), its coordinates (given by line and column counters) were used to address a digital number in a RAM and to increase it. The advantage was a reduction of the number of additions to be performed in real time.
5. The photon position encoding problem: the real challenge of designing a photon-counting camera

5.1. Introduction of the problem

As we see, the Boksenberg’s system was also used as a still camera, though it was able to give the photon coordinates. The interest of using the image information in the form of coordinate sets came from the HAR techniques for astronomy, which require to take images with exposure times shorter than the atmospheric turbulence coherence time $\tau_0$. Usually, $\tau_0 < 20$ ms. At this exposure time, the SNR in each frame is so low that image intensification is usually required. In the case of speckle imaging, a problem in the early 1970s was the data processing. Computers were not powerful enough for real-time processing and video recorders were expensive. One of the first cameras for speckle imaging, built by Gezari et al. [33] in 1972, was an intensified film movie camera. For photon-counting cameras used in HAR, one challenge is to convert as fast as possible the position in the image plane of an incoming photon into a set of coordinates $(x, y)$ that are digital signals. A lot of ingenious systems have been invented for this purpose. We are going to review some of them.

![Figure 5: Three systems for measuring the photon position by Iredale et al. [35].](image)

It is important to notice that this problem was addressed by H. Anger in 1952, for medical gamma imaging [34]. Because of the large area of Na I(Tl) scintillators that convert each gamma photon into a burst of visible photons, it is possible to mount a PMT array downstream of a scintillator (Fig. 9-a). The secondary photons spread on the photocathodes of the PMTs. The combination of the analog signals given by the PMTs gives the $(x, y)$ coordinates of the gamma photon.

The problem of photon position encoding in the image intensifiers was addressed by Iredale et al., [35] for one dimension case ($x$ coordinate to be estimated). They presented in 1969 the results of three possible optical setups (Fig. 5). In the first setup, the spot on the intensifier output (corresponding to a photo-event) was re-imaged as a line along $y$ (by mean of a cylindrical lens) onto a binary code mask. This mask was located at the entrance of a stack of fiber optics, each fiber having...
a rectangular cross-section. The output of each fiber fed a PMT. Therefore, the PMT outputs gave the binary value of the photo-event coordinate. In the second setup, the spot was re-imaged onto a neutral density whose attenuation was varying along $x$. A PMT behind the neutral density gave a signal with an amplitude proportional to the photo-event $x$-position. In the third setup, the spot was re-imaged with a certain defocus on a stack of fibers connected to PMTs, like in the first setup. By combination of the analog signals at the PMT outputs, $x$ was found out (like for the Anger gamma camera). This latest setup was considered as the most accurate, and Iredale proposed to extend it for 2-D imaging.

5.2. Cameras using image intensifiers

![Classical design of an intensified-CCD photon-counting camera. The represented Gen-II intensifier features proximity focusing (which reduces the image distortion).](image)

The classical design of the image intensifier (Fig. 6) consists of a stack of Gen-I + Gen-II image intensifiers, whose phosphor output image illuminates a CCD frame array detector. A real-time electronics is used to extract the photon coordinates from the frames. One example of such a camera used for HAR astronomy is the “CP40”, built by Blazit [36] in 1986. It actually featured a mosaic of four $384 \times 288$-pixel CCD chips (Thomson TH7861) fed by fiber optic conduits and taper connected to the fiber optics output window of the second image intensifier. The diameter of the entrance photocathode was 40 mm (hence the name of the camera). Coordinates were extracted in real time from the CCD frames by a dedicated electronics and sent to a computer system, either for speckle imaging or for dispersed fringe imaging for long baseline interferometry. The rate of the CCDs was 50 frames per second (fps). The maximum count rate for artifact-free images was about 25,000 ph/s. Above this value, holes in the autocorrelation of the images occurred due to the centering electronics.

A lighter version of this camera called “CP20” [37], featuring only one CCD chip and 20-mm photocathode intensifier was built from the CP40 design. The same French team has recently built the “CPng” (new generation) camera [38]. It features a Gen-III (AsGa photocathode) image intensifier, coupled to a Gen-II image intensifier, and a 262 fps CCD camera (Dalsa) with a $532 \times 516$-pixel resolution. The processing electronics consists of a real-time computer which extracts the photo-event positions. The software can extract these positions at sub-pixel resolution and offers a
The Precision Analogue Photon Address (PAPA) camera [39] is another camera using image intensifiers. Its originality is to use a relatively small number of PMTs to encode the photo-event position. The image of the spot, corresponding to a photo-event, at the phosphor output of an image intensifier is collimated by a primary lens, and then replicated as imagelets by a system of secondary lenses (Fig. 7). Behind these lenses are located binary masks and then a PMT. Each mask gives a Gray-code bit, either of the $x$ or the $y$ photo-event coordinate. The re-imaged spot may either be blocked or not by a mask, the PMT thus giving a signal that is binarized to yield a value, either 0 or 1. This value corresponds to a Gray-code bit of $x$ or $y$. One of the secondary lens + PMT system has no mask and is used to detect the presence of a photo-event by sampling the outputs of the other PMTs. For a $2^N \times 2^N$-pixel resolution, $2^N + 1$ secondary lens + mask + PMT sets are required. Though having a count rate close to $10^6$ ph/s, the PAPA camera suffered from many artifacts in the images due to the complexity of the tuning of its optics [40].

The DELTA camera [41] was a project attempting to use the sparse nature of the frames at the output of an intensifier phosphor (if the frame time is short enough). Spots at the output of the phosphor (corresponding to photo-events) can be projected onto three axes. The coordinates of these projections allow to retrieve the $(x, y)$ coordinates, even if there are more than one photo-event in a frame. Concretely, the image of the phosphor is re-imaged, by means of a cylindrical optics (c.f., Iredale et al. [35] experiments) onto three 1024-pixel linear CCD arrays (Thomson TH7809A) noticeable by a very short frame time (2.6 µs). One problem came from the coupling efficiency between the phosphor and the linear CCDs. These CCDs had to be coupled to compact Gen-II image intensifiers. Because of the complexity of the system and the lack of support (especially when Thomson decided to stop the production of the TH7809A), the DELTA camera project had to be abandoned.

The “Diamicon” (DIffraction-Adressed Mask Iconographer) is an attempt to modernize the PAPA camera, by using a simpler optics [42]. The concept features only one lens, coupled to with a
Figure 8: The Diamicon camera.

special diffraction filter (Fig. 8). This optics re-imaged each spot from the image intensifier phosphor as a cross pattern. A binary mask is located in front of fibers feeding an array of PMTs. According to the position of the cross-pattern on the mask, some PMTs are illuminated, yielding a signal that codes the \((x, y)\) photo-event coordinates. For an array of \(N \times N\) PMTs, the resolution of a Diamicon camera would be \(N(2^{N-1} - 1) \times N(2^{N-1} - 1)\) pixels. Hence, if \(N = 8\), the resolution would be \(1016 \times 1016\) pixels. Since for an equivalent resolution, more PMTs are required for a Diamicon than for a PAPA camera, PMTs could be replaced by an array of avalanche photodiodes (see Sect. 6.3).

5.3 Special-anode cameras

We describe here a variety of photon-counting camera that consists of a tube with a photocathode, one or several MCPs, and a special anode to determine the photo-event \((x, y)\) coordinates.

In 1966, Anger [43] invented the “wedge-and-strip” camera concept. In this system, the target for the cloud of electrons coming out from the MCP is a four-electrode \((A, B, C\) and \(D)\) anode with special shapes. The local areas of \(A\) and \(B\) vary with \(y\) and the local areas of \(C\) and \(D\) vary with \(x\) (Fig. 9-b). The intensity on each electrode is proportional to the number of electrons received by the electrode, and therefore to the local area. The difference between \(A\) and \(B\) yields \(y\), while the intensity difference between \(C\) and \(D\) yields \(x\). Cameras based on this concept have been developed at University of California Berkeley [44] and at University of London [45]. The problems of the wedge-and-strip technique comes from the limitations on the anode capacitance, which restricts the maximum count rate 40 000 ph/s (for the Univ. London’s camera) and from the defocussing (required to spread the cloud of electrons onto the anode) which is sensitive to the ambient magnetic field.

Another solution for photon position encoding is the resistive anode, used in the “Ranicon” camera, invented and tested by Lampton and Paresce [46]. It consists of a single resistive plate located at the exit of an MCP. By measuring the voltages at four edges of the plate anode, the position of the photo-event can be estimated. In 1980, a prototype of Ranicon obtained a \(150 \times 150\)-pixel resolution [47]. To increase the resolution, it was necessary to increase the electron gain by
using a stack of MCPs. With three MCPs instead of two, the resolution of a Ranicon camera can be multiplied by two [48]. Using 5 MCPs, 1024 × 1024-pixel Ranicon cameras were built [49].

While Ranicon cameras were being developed, Kellog et al. [50] worked on another concept of camera called “Photicon”. The first prototype was commissioned in 1977. The Photicon used as an anode, two orthogonal networks of 17 wire electrodes each (Fig. 10). By interpolating the intensities at the electrode exit, a 128 × 128-pixel resolution could be obtained. As performance is not so good compared to the Ranicon system, the Photicon concept was discarded.

Nevertheless, the MAMA camera [51] (Multi-Anode Microchannel Array) concept was based on the Photicon, but later evaluated by using a different wiring of the electrodes to digitally encode the photon position [52]. The idea is to slightly defocus the electron cloud at the exit of the MCP, so it falls into two wire electrodes for each coordinate (x or y). One electrode is used to encode the coordinate divided by an integer N, and the other encodes the coordinate modulo N. Fig. 11 shows an example with N = 3. Hence, the number of electrodes to encode X possible coordinate values is N + (X mod N). To reach X = 1024, if N = 32, then 64 wire electrodes are needed (128 for a 2-D imager). The MAMA camera presented in 1985 had a 256×1024-pixel resolution [53]. If the analog signal at the output of the electrodes is processed, the resolution can reach 4096 × 4096 pixels.

Another system for position encoding is the delay-line anode [54]. One of the first prototypes of this system was designed by Sobottka and Williams [55] in 1988. It consists of a ceramic plate...
(152 mm × 152 mm), on which are wined two orthogonal pairs of coils (Fig. 12-a). Each pair is used to encode a coordinate $x$ or $y$. Within a pair, a coil is exposed to the electrons from a MCP, while the other is isolated and is used as a reference (Fig. 12-c). The difference of current at an end of the coil pair is used to trigger a ramp generator, and at the other end to stop this generator. The voltage at the output of the generator, when it is stopped, depends on the delay between the pulses received at both ends, and therefore to the position of the electron cloud on the exposed coil. This system (Fig. 12-b) allows a high count rate ($10^6$ photons/s). The problem of the system is the size of the anode target, larger than any MCP and requiring a distortion-free electronic lens. From a similar concept, Friedman et al. built in 1996 a camera using two serpentine delay-lines [56] (Fig. 13-a). The FWHM resolution was 32 μm (the size of the anode was 80 mm × 80 mm), and the count rate was 110 000 ph/s.

An hybrid system wedge-and-strip + delay-line has been experimented by Lampton et al. [57]. In this case, $x$ was determined by a zig-zag delay-line electrode, while $y$ was determined by measuring the charge partition on two electrodes (Fig. 13-b). Therefore, flaws of the wedge-and-strip system appeared. To correct them in order to have a better $y$ resolution, a low-capacitance double delay-line was imagined by Raffanti et al. [58]. They also designed in 1994 a more accurate system for measuring the time delay, reaching a 4-ps RMS resolution [59].

One of the latest types of special anode is the “vernier” anode [60], which is an hybrid of the wedge-and-strip technique and the MAMA anode: a set of three superimposed strip electrode anodes, with different orientations, gives 3 analog values $A$, $B$ and $C$, from which can be computed $x \mod N$, $y \mod N$, $x/4$ and $y/4$. Despite of a 10-μm announced resolution, the count rate of a vernier anode camera is limited, like other special-anode cameras, by the anode remanence.

To conclude on the special anode cameras, we will notice that it is possible to improve them using the induced-charge technique [61]. In this, a resistive plate inserted between the MCP exit and the anode is used to induce electrons in the anode. The electrons are then received from the MCP. This method has been tested with delay-line and wedge-and-strip anodes [62], and reduces the readout noise and the image distortion.

### 5.4. Position-sensitive photomultipliers

The position-sensitive photomultipliers technology uses dynodes, like classical PMTs. A position-sensitive photomultiplier tube (PSPMT) consists of an array of dynode chains, packed into a vacuum tube. Like in the Photicon, currents measured on electrodes at the output of the last dynodes are interpolated to find out the position of the photo-event. A photon-counting camera, based on PSPMT,
Figure 12: (a) Sobottka & Williams’ delay-line anode, (b) camera using this anode, and (c) Readout electronics of this camera.

installed at the exit of an image intensifier, has been built by Sinclair and Kasevich [63] in 1997. The count rate is important (500,000 photons/s), but the resolution poor (360 μm FWHM with a 16-mm image field). Another camera based on the same concept has been built in 2000 by Pruksch and Fleischmann [64]. The resolution of their camera only reaches 128 × 128 pixels and presents some distortion in the image.

6. Solid state technologies: the future of photon-counting

6.1. The need for replacing tubes

All the devices that we have described above are based on the photoelectric effect, i.e., the emission in the vacuum of an electron from a photocathode by a photon. These devices have several problems:

1. The quantum efficiency is rather low (around 0.1 for multi-alkali photocathodes, 0.2 for AsGa photocathodes). It also tends to decrease with time due to the interaction of residual gas molecules with the photocathode [65].

2. They present false counts due to thermoionic emission (random electron emission from the photocathode due to the temperature). Electron emission may also be due to chemical interaction of residual gas molecules with the photocathode [66].

3. Residual gas molecules in the tube may be ionized by an electron. In this case, the positive ion
will hit the photocathode and liberate several electrons. This phenomenon called “ion return” causes artifacts in the image that are noticed by bright spots (brighter than those caused by a photo-event).

4. Their constructions require a high vacuum, implying a fragility of the devices. Also very high voltage for power supply are required for operating an image intensifier, causing problems of electrical insulation.

Therefore, research in solid state components has been initiated in order to find alternative to vacuum tubes in photon detection.

Research in solid state electronic imaging detector started at the end of the 1960s. Boyle & Smith introduced the concept of charge coupled semiconductor in 1970 [67]. In 1973, Fairchild Inc. introduced the first commercial charge coupled detector (CCD). It had a $100 \times 100$ pixel resolution. CCDs replaced TV camera tubes for the photon-counting cameras using image intensifiers, as we have seen in Sect. 5.2. However, the possibility to have a fully solid state photon-counting camera is quite recent.

6.2 Low light level charge coupled device (L3CCD)

The low light level charge coupled devices (L3CCD), recently developed [68] using both front- and back-illuminated CCD, do not employ any external image intensifiers, but uses on-chip gain technology to multiply photon-generated charge which can allow a signal to be detected above the noisy readout. It consists of a normal two-dimensional CCD, either in full frame or in frame-transfer format and is provided with Peltier cooling system that is comparable with liquid nitrogen cooled cryostats. The image store and readout register are of conventional design operating typically at 10 volts, but there is an extended section of gain register (Fig. 14), between the normal serial register and the final detection mode which operates at much higher amplitude (typically at 40-50 volts). This large voltage creates an avalanche multiplication which thereby increases the number of electrons in the charge packets, thus producing gain. Such a gain, $G$, is achieved by generating secondary electrons via impact ionization, which may be more than 1000×. The probability of generating a secondary electron is dependent on the voltage levels of the serial clock and the temperature of the CCD. Such a probability is low; typically it ranges from 0.01 to 0.016. Albeit the total gain is very
Figure 14: Structure of a typical L3CCD sensor.

high due to a large number of pixels in the multiplication register. The gain, $G$, is expressed as,

$$G = (1 + p)^N,$$

in which $N$ is the number of pixels in the multiplication register and $p$ the probability of generating a secondary electron.

Following the extended register, the L3CCD has an output amplifier like a conventional CCD. The L3CCD generates an output event with a high signal-to-noise (S/N) ratio from a single electron detected by the CCD imaging array. Since there is no spatial or temporal dispersion due to the amplification, it allows much higher mean photon arrival rates. The readout noise of the amplifier is a voltage noise within the output transistor. The gain factor may be adjusted by varying the amplitude of the high voltage clock, thus allowing a trade-off between full well capacity of the system and readout noise. With a multiplication gain, $G$, the total detected noise, $S_N$, referenced to the image area is given by,

$$S_N = \sqrt{SF^2 + S_D F^2 + N_r^2 / G^2},$$

in which $S$ is the total number of photons arriving at each pixel, $S_D$ the total dark signal including spurious charge, $F$ the excess noise factor (typically the value of $F$ is $\sqrt{2}$), and $N_r$ the detector read noise.

Another important factor influencing electron multiplication gain is that of cooling the CCD; the more cooler the CCD, a primary electron generates a secondary electron in the silicon, which gives rise to higher on-chip multiplication gain (Fig. 15). However the spurious charge that generates due to the movement of electrons through the multiplication register’s pixels (the sharp inflection in the clock waveform produce a secondary electron in absence of the primary electron), occurs in numbers. Typically a single spurious electron is generated for the transfer of every 10 pixels, resulting a value of 0.1 e/pixel/frame. The exposure has no effect on such a spurious charge.
Incoming photons follow Poisson statistics and have an inherent noise known as photon shot-noise. The on-chip gain L3CCD is a stochastic process which adds to this shot-noise that is derived from the fundamental quantum nature of light and constitutes the theoretical noise limitation of any low-light level imaging application. The statistics of photon arrival rates mean that the uncertainty in the signal in a pixel. A detected signal of \( S \) photons has a photon shot-noise of \( \sqrt{S} \) photons. The uncertainty in the signal increases to \( \sqrt{S^2} \), which is equivalent to the half of the detective quantum efficiency of the sensor. This reduction is eliminated by using a true photon-counting system in which this event is recognized as a single photon. All the output signals above a threshold are, generally, counted as photon events provided the incoming photon flux is of a sufficiently low intensity that no more than one electron is generated in any pixel during the integration period. The dark noise is zero, and the gain is set at a suitable level with respect to the amplifier read noise. By digital signal analysis of an L3CCD output current, it can be possible to drastically reduce the readout noise. Such a method, an alternate to the classical correlated double-sampling has been presented by Gach et al. [69] in 2003.

Saha and Chinnappan [70] reported an electron multiplying CCD (EMCCD) camera system having 576 \times 288 pixels of size 20X30 \( \mu \text{m} \) in the image area, which was procured from Andor Technology. This was used for their speckle interferometric programs at the 2.34 meter Vainu Bappu Telescope (VBT), Vainu Bappu Observatory (VBO), Kavalur, India. This camera is a front-illuminated one with 45\% efficiency and is provided with Peltier cooling system that operates to -60\° C with air-cooling and with further additional water circulation, it reaches to -75\° C. The performance of this cooling system is comparable with liquid nitrogen cooled cryostats. This EMCCD has the provision to change gain from 1 to 1000 by software. The noise at 1 MHz read rate is \( \sim 2 \) e rms. Each pixel data is digitized to 16 bit resolution; the data can be archived to a Pentium PC.
6.3. Avalanche photodiodes

Avalanche photodiodes [71] (APDs) are the most common solid state photon-counting sensors. They are based on the ionization of a high-voltage pn+ junction, triggered by a single photo-electron (Fig. 16-b). To be able to count photons, the APD must be used in “Geiger mode”: the voltage has to be more than around 200 V to allow a chain reaction (avalanche) of electron liberation. The problem is then to stop the avalanche. In one design (“passive quenching” [72]), a serial-mounted resistor decreases the voltage of the APD when a current from the APD crosses this resistor (Fig. 16-c). The problem of passive quenching comes from the capacitance of the system that limits the bandwidth. However, 400-ps time resolution for photon-counting can be obtained with a passive-quenching APD with a special wiring [73]. The second technique to stop the avalanche is called “active quenching” [74]; in this case, the power supply of APD is controlled by a system measuring the output current.

Making integrated arrays of APDs working in photon-counting mode is a challenge because of the photon-emission phenomenon that is caused by the electron avalanche. These photons may trigger avalanche in the pixels in their vicinity [75]. In 1987, Trakalo et al. [76] built an integrated linear APD array with 32 pixels, but working in sub-Geiger mode (gain = 60). However, problems related to APD arrays in Geiger mode seem to be overcome. A 8 × 8-pixel APD array working in Geiger mode has been presented by Vasile et al. [77] in 1998.

6.4. Superconducting tunnel junctions

This kind of photon-counting sensors [78], born from research in X-ray detectors [79] is based on a stack (Fig. 16-a) of different materials (Nb/Al/Al2O3/Al/Nb). It has the property to get a charge proportional to the energy of an incoming photon. They have therefore an interest for spectroscopic applications. First prototypes of superconducting tunnel junction (STJ) detectors [80] had QE = 0.5 and a count rate of the order of 2500 photons/s. The photon-counting performances of the STJs have been improved by using niobium instead tantalum. In this case [81], PV → ∞ and NFWHM = 0.05 (for λ = 250 nm). The spectral resolution is 8 nm at λ = 200 nm, and 80 nm at λ = 1000 nm). The main problem of the STJs is the very low temperature that they required (370 mK). Moreover, making STJ array detectors for imaging is a challenge. A 6 × 6-pixel STJ array has nevertheless been made and used in astronomy [82].
7. Conclusions

The photon has been “invented” by Albert Einstein almost 100 years ago, in 1905. The discovery of the quantum nature of light, beyond the explanation of the photo-electric effect, has revolutionized how ultra-sensitive light detectors can be imagined. The existence of a quantum limit in light detection has led to a quest, on through the 20th century (and still going on), for the “perfect detector”. Thanks to the progress in solid-state technology, future photon-counting detectors will asymptotically match the performance of the perfect detector (characterized by $\text{QE} = 1$, $\text{FD} = 0$) to yield a comparable quantity of information.

The history of photon-counting, that we have reported, can be seen as a tribute to the genius of many scientists, whether they are (or were) chemists, able to find the materials to build the components, or opto-electronicians, having imagined a fantastic variety of systems in order to find the position of a photo-event in an image plane. Various works in high-energy photon detection (X-rays, gamma particles,...) have been extremely useful for visible photon detection.

Until a few decades ago, astronomers used photographic technique to record images or spectra of celestial objects. Owing to low quantum efficiency of the photographic emulsion, usage of the modern cooled charge coupled device (CCD) camera system became necessary in the fields of astronomy. It is being used as an imaging device in other scientific fields like biomedical science and in commercial applications like digital cameras as well.

The short-exposure speckle imaging as well as the adaptive optics system that corrects the perturbations introduced by the atmosphere in real time, require a sensor. In spite of the high sophistication of modern photon counting cameras that are engineered to address the challenges of ultra-low light level imaging applications, they are unable to target the photon-starved astronomical objects. The sampling rate of these sensors scales with the turbulence in the atmosphere up to kHz and is limited by the number of photons received in a short-exposure. The performance relies on the characteristics of such sensors, e.g., (i) the spectral bandwidth, (ii) the quantum efficiency, (iii) the detector noise that includes dark current, read-out and amplifier noise, (iv) the time lag due to the read-out of the detector, and (v) the array size and the spatial resolution.

In the field of optical interferometry with diluted apertures, the time resolution of these cameras should reach 1 msec. Development in this field in future may have far reaching impact on interferometric imaging that requires detection of very faint signals and reproduction of interferometric visibilities to high precision [83].

References

9 Zworykin V K, L’Onde Électrique, 15 (1936), 265.