Journal of Modern Optics
Vol. 54, Nos. 2–3, 20 January–15 February 2007, 163–189

Photon counting arrays for astrophysics

F. ZAPPA†, S. TISA*†, S. COVA†, P. MACCAGNANI‡, R. SALETTI§, R. RONCELLA§, F. BARONTI§, D. BONACCINI CALIA¶, A. SILBER¶, G. BONANNO‖ and M. BELLUSO‖

†Politecnico di Milano, Milano, Italy
‡IMM-CNR (Microelectronics and Microsystems Institute of the Italian National Research Council), Bologna, Italy
§University of Pisa, Pisa, Italy
¶ESO (European Southern Observatory), Garching, Germany
‖INAF – Astrophysics Observatory of Catania, Italy

(*Received 9 February 2006; revised 4 April 2006; in final form 7 April 2006*)

A compact system for counting and time-tagging single photons is presented, based on a monolithic array sensor of 60 pixels able to detect single photons, namely the single-photon avalanche diode array (SPADA). First, the working principle and performance of the single-photon detector pixel is detailed, with particular attention paid to monolithic array integration. Then the electronics needed to quench each pixel after avalanche ignition, namely the active-quenching circuit (AQC) is discussed, since the features of this quenching electronics dramatically affect the operating conditions of the detector, hence its actual performance. The discussion then focuses on integration of the SPADA system into Astrophysics applications such as adaptive optics, fast-transient imaging and atmospheric layer sensing. The whole electronics necessary to control SPADA operating conditions and temperature is also described, together with the complete opto-mechanics used to focus the telescope pupil onto the detector. Finally, experimental results are reported.

1. Single-photon detectors

In many fields and in particular in astrophysical observations, a chronic problem is the photon-starving condition, which becomes severe when images are to be obtained in short acquisition times (from micro to milliseconds), as happens in hot areas of astrophysics: optical counterparts of high-energy gamma-ray bursts, study and interpretation of Supernovae bursts, adaptive optics (AO) [1] for real-time correction of atmospheric turbulence by means of pulsed lasers. CCDs are inherently unable to provide accurate measurements of such fast low-intensity transients at high frame rates, whereas monolithic solid-state arrays of photon counters would be suitable. While high-sensitivity CCDs are being developed by industry [2], and some

*Corresponding author. Email: tisa@elet.polimi.it

Journal of Modern Optics
ISSN 0950–0340 print/ISSN 1362–3044 online © 2007 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/09500340600742320

applications have resorted to the use of a set of discrete SPADs [3], the monolithic array of single-photon detectors presented in this work is state-of-the-art in both the scientific community and industry.

To respond to single photons, suitable detectors must provide output signals that are sufficiently high to be individually processed by electronic circuits. Therefore, only detectors with an internal mechanism that provides a high multiplication of charge carriers are suitable, namely vacuum tube photomultipliers (PMTs), solid-state avalanche photodiodes (APDs) and emerging electron-multiplying CCDs (EM-CCDs) [2]. In PMTs, the photocathodes available for the visible spectral range provide fairly good quantum efficiency and low noise, whereas cathodes for the red and near-infrared range have lower quantum efficiency and must be cooled to reduce the dark-count rate. PMTs are bulky, and so not suitable for assembly in large arrays, fragile, sensitive to electromagnetic disturbances and mechanical vibrations, require high supply voltages (2–3 kV) and are costly devices, particularly the high-performance models. EM-CCDs exploit an internal multiplication process to achieve sub-electron readout noise, thus being able to detect single photons [2]. Their quantum efficiency is very high, and they are inherently suited to imaging applications. However, due to their nature, they cannot provide frame rates higher than a few kiloframes per second, and cannot be used in extreme time-resolved measurements.

Semiconductor APDs have the typical advantages of solid state devices (small size, low bias voltage, low power consumption, ruggedness and reliability, suited to building integrated systems, etc.). In APDs operating in linear mode, the internal gain is not sufficient or barely sufficient to detect single photons. Instead, single photons can be detected efficiently by avalanche diodes operating in Geiger mode, known as single-photon avalanche diodes (SPADs). Silicon SPADs have been investigated extensively [4] and are nowadays well known and widely employed; considerable progress has been made in design and fabrication techniques, and devices with good characteristics are commercially available [5–8]. In this paper we present a compact two-dimensional imaging system for counting and time-tagging single photons, based on a monolithic array of 60 pixels, namely the single-photon avalanche diode array (SPADA).

1.1 *SPAD working principle*

Essentially, SPADs are p–n junctions biased at a voltage V_A above the breakdown voltage V_B. At this bias, the electric field is so high (higher than 3×10^5 V/cm) that a single charge carrier injected in the depletion layer can trigger a self-sustaining avalanche. The current rises swiftly (nanosecond or subnanosecond rise-time) to a macroscopic steady level, in the milliampere range. If the primary carrier is photogenerated, the leading edge of the avalanche pulse marks (with picosecond jitter) the arrival time of the detected photon. The current continues to flow until the avalanche is quenched by lowering the bias voltage to or below V_B: the lower electric field is no longer able to accelerate the carriers to impact-ionize with lattice atoms. Then the bias voltage must be restored, in order to be able to detect

another photon. The circuit that performs such operations is usually referred to as a quenching circuit [9].

In order to work as a photodetector, a diode must be able to remain biased above breakdown for a sufficient time, let's say longer than a few milliseconds. This means that the generation–recombination phenomenon, which would fire the avalanche, must be kept very low. Since thermally generated carriers can fire an avalanche, it is possible to observe output current pulses also when a SPAD is kept in the dark: such an average counting rate is called the dark-counting rate and is one of the key parameters in defining detector noise [5].

2. Avalanche quenching circuits

Once the SPAD is biased above the breakdown voltage and is triggered, current keeps flowing until the avalanche process is quenched by lowering the bias voltage to V_B or below. After a dead-time, the operating voltage must be restored in order to enable the SPAD to detect another photon. This operation requires suitable electronics, which has the following tasks:

(1) it senses the leading edge of the avalanche current;
(2) it generates a standard output pulse, synchronous with the onset of current;
(3) it quenches the avalanche by lowering the bias below the breakdown voltage;
(4) it restores the photodiode voltage to the operative level.

This circuit is usually referred to as a quenching circuit. The most commonly used circuit in studies on Geiger-mode avalanche photodiodes is the passive-quenching circuit: the avalanche current quenches itself simply by developing a voltage drop across a high-impedance load ($R_L > 100\,k\Omega$). Such a circuit is very simple and can easily be employed, but sets severe limitations on the maximum admissible photon counting rate and on detector performance in general [9]. In fact, it was the introduction of the active-quenching circuit (AQC) concept by S. Cova [10, 11] that opened the way to practical application of SPADs. Many AQC types have since been reported, with circuit structure and mounting that evolved from standard NIM cards [11, 12] to small SMT boards suitable for compact detector modules [6, 13, 14].

2.1 *Fully-integrated iAQC*

In recent years, work towards monolithic integration of AQCs was started at Politecnico di Milano. First, the core structure of the AQC was integrated, excluding only the high-voltage quenching driver [15]. In order to reduce the overall stray capacitance, improve timing performance and strongly reduce power dissipation, we designed and fabricated the first fully-integrated active-quenching circuit (iAQC) [16]. The circuit was the first reported in the literature and is under European and US patent [17]. For this present work we developed an improved version of the iAQC, whose microphotograph is shown in the left side of figure 1.

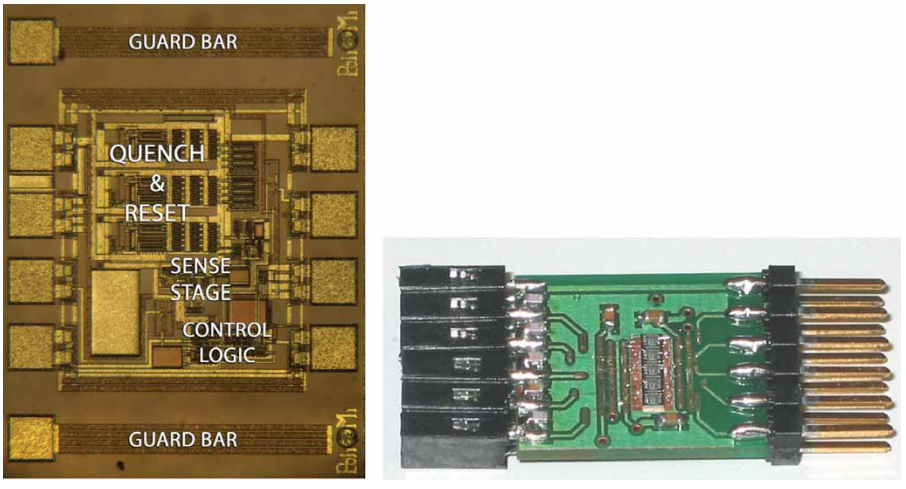

Figure 1. Micrographs of the iAQC (left) and compact board containing five iAQC channels. (The colour version of this figure is included in the online version of the journal.)

The overall dimensions are about 1×1 mm. In order to enable the integration of multichannel apparata, like the SPADA system discussed in this paper, five monolithic iAQC were built onto a single chip.

In the following we report the results of tests made on the iAQC driving thin SPADs with breakdown voltages ranging from 17 V to 35 V. The applied overvoltage was in the range from 5 V to 10 V. The very short dead-time is only 30 ns, thus leading to a maximum saturated counting rate of 30 Mcounts/s. Figure 2 shows the same measurement with a longer dead-time of about 400 ns, obtained by simply adjusting the voltage at a control pin. The maximum hold-off time achievable is 500 ns. Figure 3 shows the cathode waveforms and the corresponding output pulses obtained by illuminating the detector with steady weak light: the detector is triggered by subsequent single photons with a 200 ns dead-time. Figure 3 proves that the iAQC can be retriggered immediately after the cathode has been reset to the operating level. Figure 4 shows gated-mode operation: when the GATE level is high, the detector is free to operate and the circuit output delivers a standard TTL pulse; when the GATE is low, the SPAD is kept quenched and no avalanche can be triggered.

3. Astrophysics applications

The SPADA system was mainly aimed at three astrophysical applications that require ultra-sensitive arrays with a small number of detectors for the visible range: adaptive optics (AO), fast transient imaging (FTI) and layer sensing (LS). Adaptive optics was the driving force, and it determined many parameters of the system.

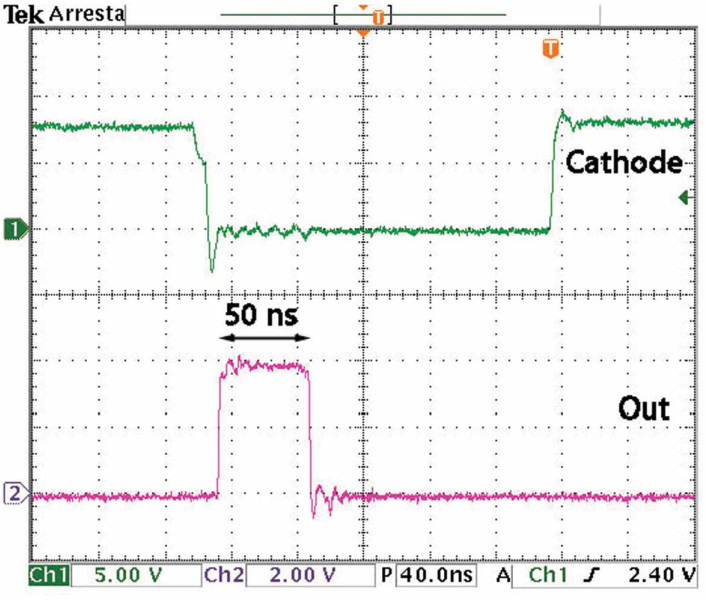

Figure 2. SPAD cathode waveform and iAQC output pulse, with 200 ns hold-off time. (The colour version of this figure is included in the online version of the journal.)

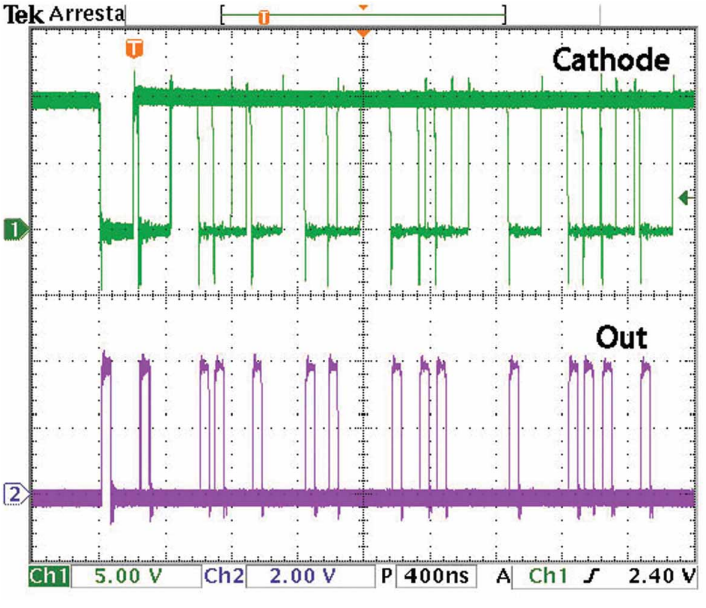

Figure 3. Example of free running operation of the SPAD. (The colour version of this figure is included in the online version of the journal.)

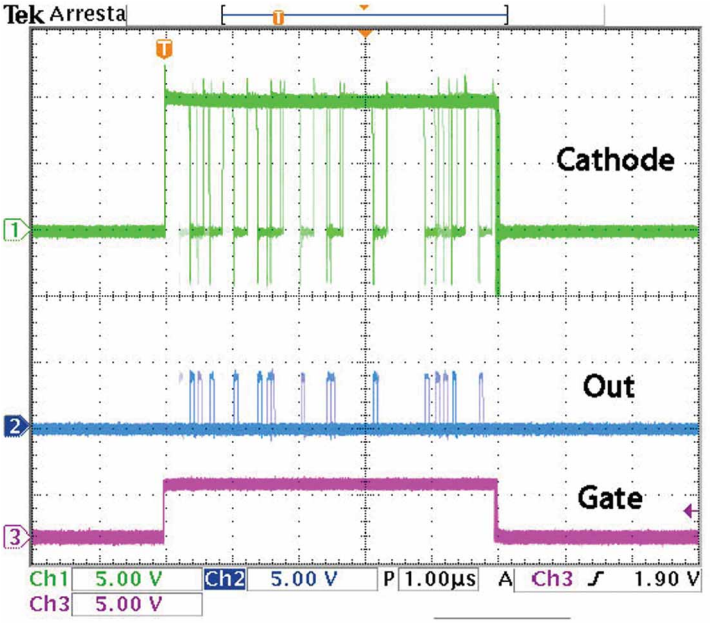

Figure 4. Example of gated-mode operation of the SPAD. The iAQC holds off the SPAD whenever the Gate signal is low. (The colour version of this figure is included in the online version of the journal.)

The following paragraphs provide a brief description of the applications and the requirements used to implement the data processing board of the SPADA system.

3.1 *Adaptive optics*

One of the toughest problems affecting ground-based telescopes is the presence of the atmosphere, which distorts the spherical wavefront, creating phase errors in the image-forming ray paths. Even at the best sites, ground-based telescopes observing at visible wavelengths cannot achieve an angular resolution in the visible better than telescopes of 10 to 20 cm diameter, because of atmospheric turbulence alone. The cause is random spatial and temporal wavefront perturbations induced by turbulence in various layers of the atmosphere; one of the principal reasons for flying the Hubble Space Telescope was to avoid this image smearing. In addition, image quality is affected by permanent manufacturing errors and by long timescale wavefront aberrations introduced by mechanical, thermal, and optical effects in the telescope, such as defocusing, decentring, or mirror deformations generated by their supporting devices.

Adaptive optics is the answer to this problem: a deformable mirror is inserted in the light path of the telescope, and its control signal is based on measurement of the incoming wavefront, performed by a suitable high-sensitivity detector.

Because of the high bandwidth and the small field to which correction can generally be applied, adaptive optics employs a small deformable mirror with a diameter of 8 to 20 cm located behind the focus of the telescope at or near an image of the pupil. The choice of the number of actuators, usually piezoelectric, is a trade off between degree of correction, use of faint reference sources and available light budget. A large number of actuators requires a similarly large number of subapertures in the wavefront sensor.

Two main methods are used to measure the degraded wavefront, the Shack–Hartmann device [18], which measures the slope of the wavefront from the positions of the images of the reference star given by each subpupil, and curvature sensing [18, 19], where the measured intensity in strongly defocused images directly gives the local curvatures of the wavefront. Correction in the Shack–Hartmann device is made with individual piezoelectric actuators. Correction in a curvature sensing system is accomplished with a bimorph adaptive mirror, made of two bonded piezoelectric plates. With both methods, wavefront sensing is done on a reference star (Natural Guide Star, NGS), or even on the observed object itself if it is bright enough and has sufficiently sharp light gradients. The measurement can be performed in the visible for observation in the infrared, or in the infrared itself. The control system is generally a dedicated computer that uses the measurements from the wavefront sensor to calculate the commands to be sent to the actuators of the deformable mirror. The calculation must be fast (within 0.5 to 1 ms), otherwise the state of the atmosphere may have changed, thus rendering the wavefront correction inaccurate.

The SPADA system was designed to operate as a curvature wavefront sensor (CWFS). As we will see, this choice and the need to make it a potential replacement for the MACAO [20] apparatus already in use at ESO's Very Large Telescope (VLT) determined many of its characteristics, especially the geometry of the monolithic detector. It has also a strong advantage over CCDs if a pulsed laser system is used thanks to its gating function and parallel readout, which allow faster loop cycles. Figure 5 depicts the simplified operating principle of a CWFS. By means of a moving membrane, the focusing plane of the incoming photon flux over the detector is changed with sinusoidal oscillations. The number of detected photons when such a plane is both before (counting A) and behind (counting B) the detector surface is collected. The measurement must be repeated over an integer number of oscillation periods. Eventually, for each pixel the CWFS provides either the accumulated A and B counts or the curvature signal:

$$C = \frac{A - B}{A + B}. \tag{1}$$

The value C is proportional to the local curvature of the wavefront, associated with the subaperture encompassed by the pixel in use. The servo drives the system at each loop step to have a flat wavefront, hence to have $C = 0$ for all pixels. The acquisition must be repeated in free running. The CWFS have to accomplish all these tasks: detect the incoming photon, count their occurrences in the two integration windows A and B, provide the curvature signal C for all the pixels, control the membrane that oscillates with a frequency ranging from 1.5 kHz to 3 kHz. On the other hand,

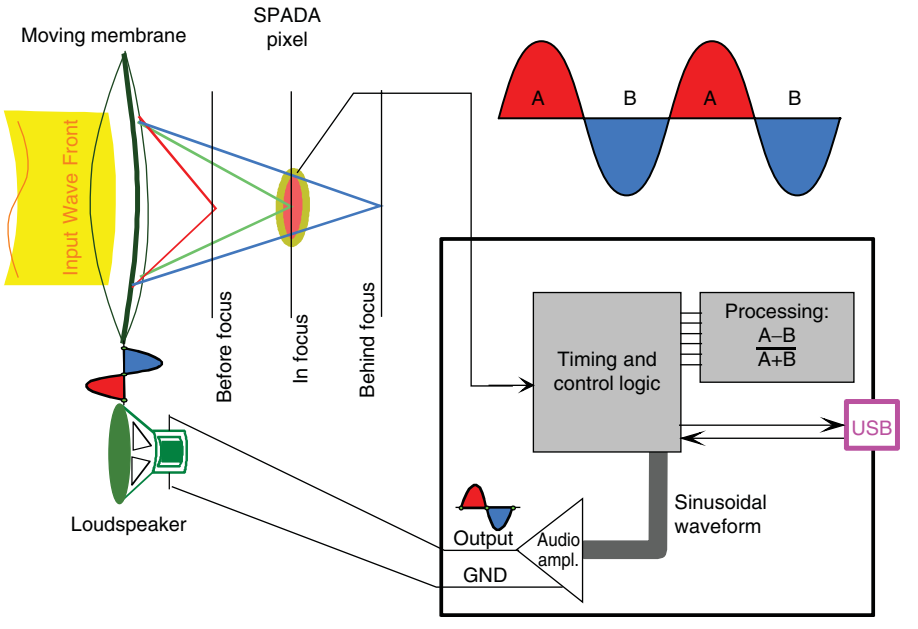

Figure 5. Schematic principle of operation of a curvature wavefront sensor (CWFS). (The colour version of this figure is included in the online version of the journal.)

the calculation of the signal to be applied to the deformable mirror to account for the atmospheric aberration is carried out by a specialized real-time computer.

The timing control logic must alternatively enable the counting of pulses corresponding to A or B semiperiod. The collections must continue for integer multiples of the sinusoid period; such multiples must range between 1 and 256. Therefore, the overall collection period can last a minimum of about 1/3 kHz=333 μs to a maximum of 256/1.5 kHz=171 ms. At the end of the collection interval, the processing electronics must upload the 60 items of curvature data (C values for all pixels). Hence, the maximum transfer rate requested is 60 channels × 16 bit/333 μs = 3 Mbps. In addition to the curvature signal, it must be possible to upload also the individual A and B counts every collection period. In conclusion a maximum 3·3 Mbps ≈ 10 Mbps is required.

3.2 *Fast transient imaging*

In fast transient imaging (FTI) applications, such as the study of transients of optical counterparts in high-energy gamma-ray bursts or supernovae bursts, it is important to acquire the image of the 60 pixel detector for each integration time window, adjustable from 10 μs to 100 ms. A key parameter in this application is the maximum attainable frame rate, since short time windows allows finer time-tagging of the incoming photon flux, with quite enough resolution for many

astronomical applications. The image consists of the number of photons detected by the corresponding pixel and counted by the board during the integration time. Therefore, every T_W time period, the 60 collected counts must be uploaded to a remote computer via a suitable high bandwidth link. The same link must be used to download settings and parameters to the board. A remote software displays the images and performs the required post-processing.

The integration time window must be programmable between 10 μs and 100 ms with 12 bits resolution. Each counter must be 16 bits wide and with saturation. In this way, at the theoretical maximum counting rate (20 Mcps) allowed by the iAQCs (at minimum dead-time of 50 ns), the counter will saturate for windows longer than 3.3 ms. In such a situation, the user can decide to reduce the integration window, thus improving also the time-tagging performance of the acquisition. Instead, with T_W shorter than 1 ms the counter will almost never exceed 14 bits, while at $T_W = 10\,\mu s$ 8 bits will be sufficient. Therefore also the uploading to the remote computer can be adapted to an effective number of bits given by:

$$n_{eff} = \log_2(T_W \cdot 20\,\text{Mcps}). \tag{2}$$

Timing and control logic should latch the content of each counter into the corresponding latch. In this way, no photon will be missed. The content of all 60 latches could be shifted out in serial order. At the shortest T_W interval, this means a maximum bit rate of:

$$F_{max} = \frac{60\,\text{pixels} \cdot 16\,\text{bit}}{10\,\mu s} = 96\,\text{Mbps}. \tag{3}$$

This is the maximum transfer rate that it has to be sustained during data uploading. As a consequence, a high bandwidth link is needed to accomplish the task. However, the effective transfer rate could be reduced if an adaptive algorithm was adopted, as previously suggested.

3.3 *Layer sensing*

Natural Guide Star (NGS) adaptive optics suffers from the fact that it is not always possible to find a suitable NGS close enough to the portion of sky under investigation. Normally only 1% of the sky has an available NGS for current AO systems. The most promising way to overcome the lack of sufficiently bright natural reference stars is the use of artificial reference stars, also referred to as laser guide stars (LGS). These are patches of light created by the back-scattering of pulsed laser light by sodium atoms in the high mesosphere, or in the case of Rayleigh LGS, by molecules and particles located in the low stratosphere. Such an artificial reference star can be created as close to the astronomical target as desired, and a wavefront sensor measuring the LGS wavefront is used to correct the atmospheric aberrations on the target object.

Nevertheless, there are still a number of physical limitations with an LGS. A problem, focus anisoplanatism, also called the cone effect, became evident very early on. Because the artificial star is created at a relatively low altitude,

back-scattered light collected by the telescope forms a conical beam, which does not cross exactly the same turbulence-layer areas as the light coming from the distant astronomical source. This leads to a phase estimation error. For even larger telescopes, such as the 100 m OWL under investigation by ESO, a solution to this problem is mandatory.

Attempts to overcome the cone effect are under study at different institutes [21], which use atmosphere's layers sensing and pulsed laser beams. The sensor timing is extremely important, in a sort of two-dimensional LIDAR technique. The key features are that a parallel (or nearly parallel) laser beam is projected from the full primary aperture and that sensing takes place on the upward path. The methods therefore rely on an observable modulation of the scattered intensity by turbulence-induced phase distortions during upward propagation of the laser beam. The SPADA system can be used to detect the small number of photons that are back-scattered by the different layers.

Concerning the functions required of the SPADA system, this application is a mixture of the two previous ones. The system should wait for a synchronization pulse at 7 kHz repetition rate, that signals the firing of the laser pulse; the detector is then gated-off for the subsequent 2.5 μs in order to mask the fluorescence of the optics, due to the powerful laser pulse; then the SPAD is gated on and the board records 10 samples lasting 13 μs each for each SPADA pixel, corresponding to 10 atmosphere layers, during the upward path of the laser beam. The acquisition is repeated a user-definable number of times. Eventually, the cumulative acquisition of the 10 sets of 60 pixels each is uploaded to the remote computer at the end of the acquisition, when the measurement will be restarted. At the same time, the board generates a sinusoidal signal, with the same specifications as listed for the AO application. This signal drives a pupil plane membrane mirror, of the type used in curvature systems such as MACAO systems. The sinusoidal motion of the membrane mirror creates for one half of the period a progressive conjugation of the sensor with different atmospheric heights above the telescope primary. It does so by acting as a gentle negative lens which collimates the layer conjugate. By adjusting the membrane mirror period and phase parameters, it is possible to obtain perfect tracking of the pulsed laser beam conjugate while propagating upward. Only sensors like SPADA can grab the images of the layers with the timing required.

The maximum communication bit rate for this application is almost equivalent to that required for FTI, setting at about 74 Mbps.

4. System integration

Figure 6 shows the architecture of the whole SPADA system. It is composed of the SPADA silicon chip mounted on a micro-machined holder, the matched optomechanics for focusing and alignment, a detection electronic board based on the iAQCs, a metallic housing, a custom data processing electronic board, and the control software. The complete system during bench testing is shown in figure 7. The data pocessing board acquires counts from the SPADA

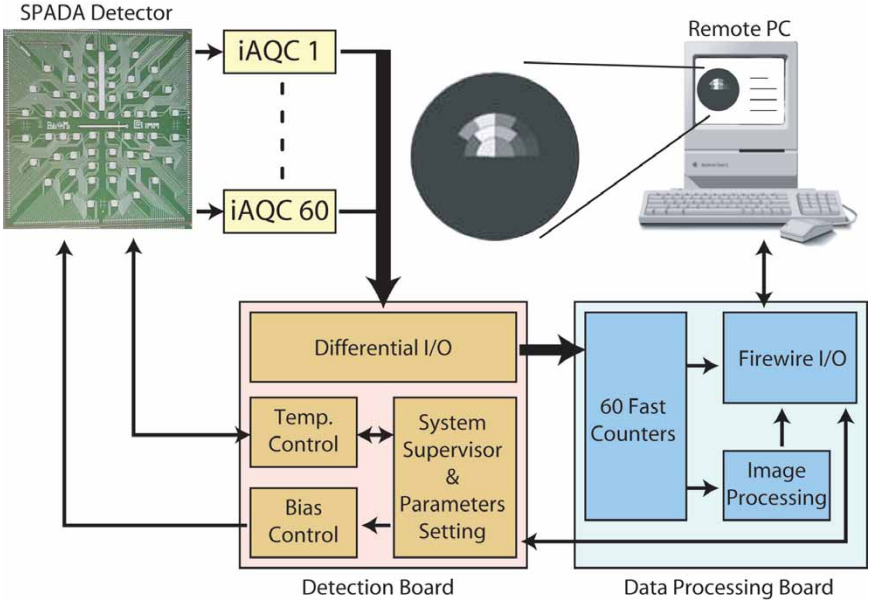

Figure 6. Architecture of the SPADA system. (The colour version of this figure is included in the online version of the journal.)

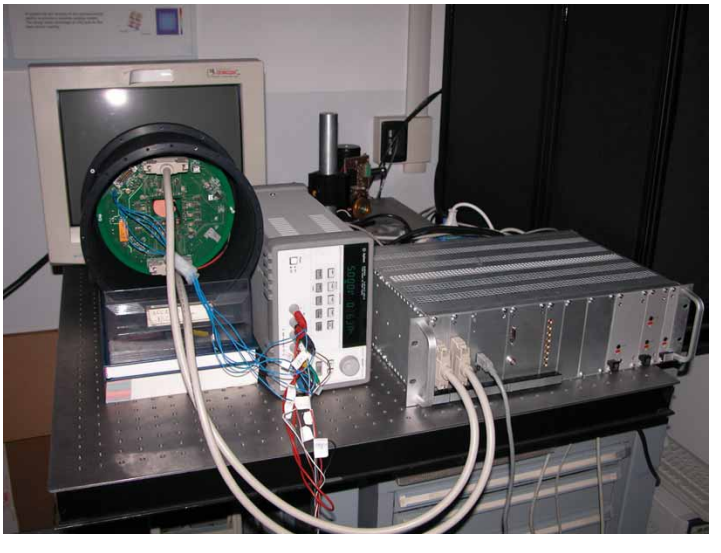

Figure 7. Rack containing the complete data processing electronics and power supplies (on the right) and the detection board (on the left). (The colour version of this figure is included in the online version of the journal.)

sensor (via the detection board), generates the sync signals and sinusoid for AO and LS applications, and handles hardware and software gates and interlocks. The board counts photon signals at each of the 60 inputs, generates also the required user-programmable time windows for the different applications and uploads processed data to a remote computer through a firewire connection that is able to sustain the required maximum data rate. Given the tasks to be performed, the board needs some dedicated logic for pixel counting, together with a processor for data processing and data transfer. The choice was to build a dedicated data acquisition system based on an Orsys board mod. C6713, which is equipped with a TMS320C6713 DSP and a Virtex-II FPGA [22].

4.1 *SPADA detector*

The SPADA sensor is made of 60 SPAD elements arranged in a circular concentric geometry, as shown in figure 8. This disposition is the same as that used for the MACAO system, and is optimum for the 60 subaperture adaptive optics apparatus designed by ESO. The inner circle has a diameter of 2.53 mm, whereas the outer has a diameter of 14.49 mm. This implied a square shape for the SPADA detector, with dimensions of 18×18 mm, in order to accommodate the bonding pads. The 60 pixels are at the focus loci of small spherelets, placed at the field focus of a 60 subaperture segmented lenslet array, about 16 mm in diameter, that ensures that the entire telescope's pupil is projected onto the 60 pixels.

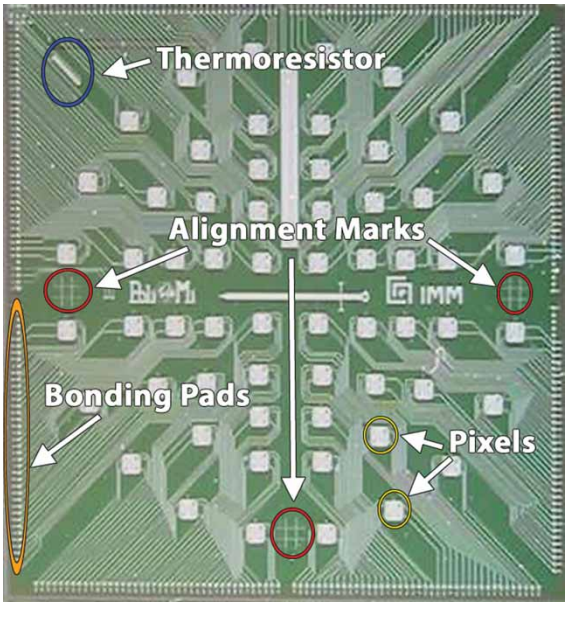

Figure 8. Photograph of the SPADA chip. Dimensions 18 mm × 18 mm. (The colour version of this figure is included in the online version of the journal.)

In order to maximize the chance of having a system with better attainable performance during the prototyping phase, the SPADA chip contains four sets of 60 pixels, with diameters of 20 μm (active area of about 300 μm^2), 35 μm (962 μm^2), 50 μm (2,000 μm^2), and 75 μm (4,400 μm^2), as shown in figure 9. The input optics is aligned and focused onto one set of SPADs (i.e. one single pixel for each element) and only those 60 pixels are bonded to the holder and reach the detection board. The set to use was selected after characterization of SPADA performance and yield, and for a good compromise between diameter and dark-counting rate the choice was to use the 50 μm diameter set. The alignment is performed with the help of three alignment marks, such as the one visible on the right side of Figure 9.

The SPADA chip also includes an integrated thermoresistor that is used by the temperature controller of the detection board to provide closed-loop regulation of the operating temperature of the detector. This assures constant detection efficiency over the whole measurement, and permits one to obtain a lower dark-counting rate if required by the application. It should be noticed that detector cooling is not necessary to guarantee correct behaviour, but it is a simple, viable method to further enhance it. This is not true for CCDs, which usually require strong cooling to attain reasonable noise level, or for thick SPADs that require cooling to avoid faults, due to the much higher power dissipation. Figure 8 shows a photograph of the SPADA sensor with all the aforementioned features highlighted.

4.2 *Detection electronics*

The detection electronics was designed to work under two operating conditions: together with the developed data processing board (as already shown in figure 6) or connected and one to one compatible with the MACAO system developed by ESO. A schematic block diagram of the detection electronics is shown in figure 10.

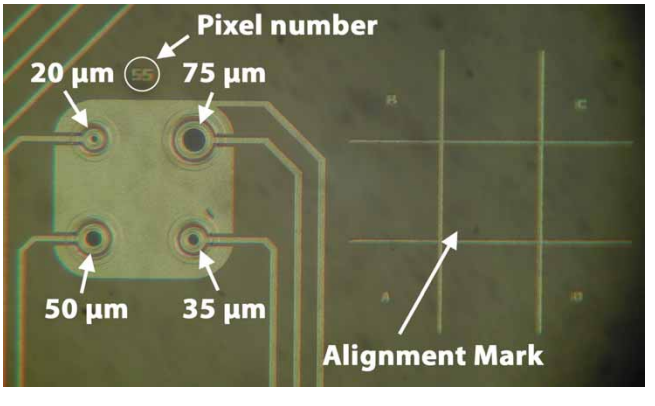

Figure 9. A pixel of the array, showing the four different diameter SPADs. Each detector has its own cathode, whereas anodes are shared. On the right side is shown one of the three marks that permit alignment of the optics with the chosen set of detectors. (The colour version of this figure is included in the online version of the journal.)

The board is built around 60 integrated active quenching circuits (iAQC), one for each pixel of the SPADA sensor, and outputs 60 differential signals through two SCSI-like connectors (68 pins each). As we have already seen, the iAQCs are fabricated as integrated chips containing five iAQCs each. For easier replacement in the case of a faulty iAQC, each chip is assembled on a compact printed circuit board as shown in figure 1, on the right, with dimensions 1.5×2.5 cm. The connectors on the left side of the board are mated to those from the SPADA, thus resulting in the shortest possible path between SPADs and iAQCs. The connectors on the right side mate with connectors on the detection board, providing power supplies and signal from and to the iAQCs. A total of 12 iAQC boards are needed to drive the 60 SPADA pixels, and they are arranged along a square, three for each side of the SPADA holder, as shown in figure 11.

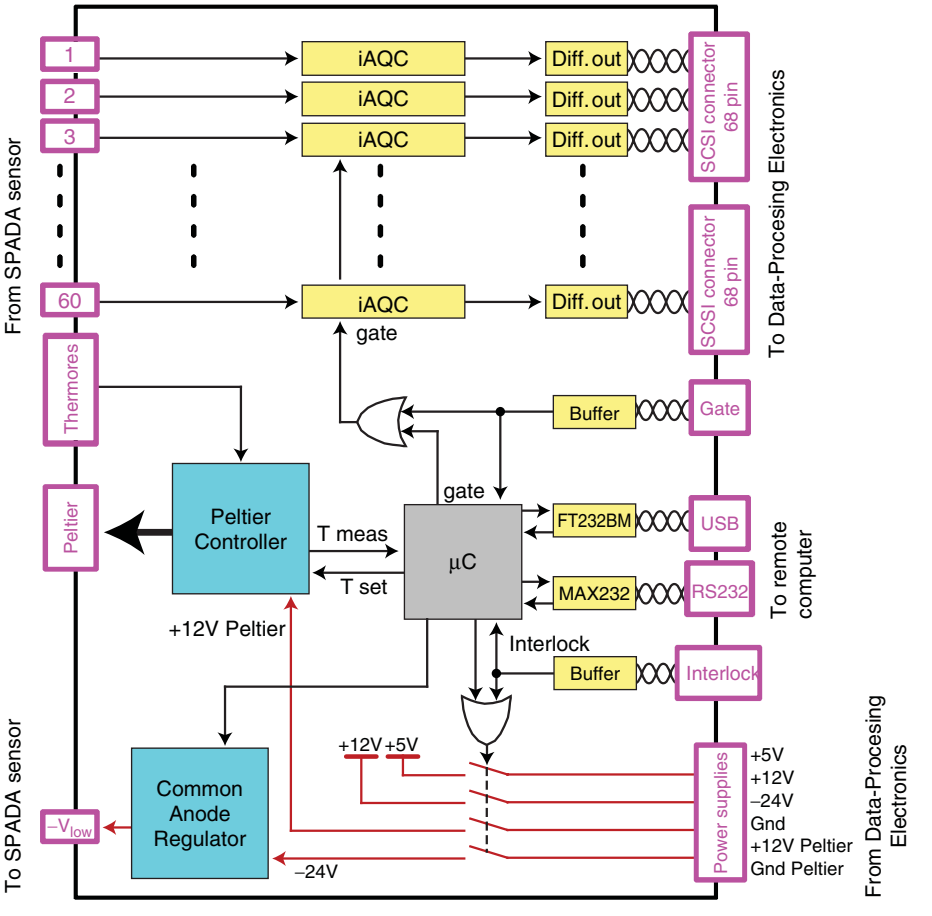

Figure 10. Schematic block diagram of the detection board. (The colour version of this figure is included in the online version of the journal.)

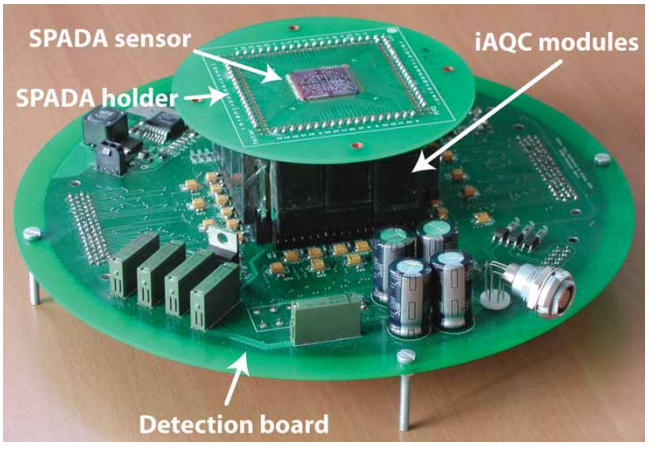

Figure 11. Complete detection board assembled with the SPADA detector. (The colour version of this figure is included in the online version of the journal.)

All iAQCs have a common Gate input, for enabling/disabling the SPADA sensor (for instance for the layer sensing application). The hold-off duration of all iAQCs is adjustable from about 10 ns to 500 ns, through a common setting controlled by the microcontroller. The detection board includes a temperature controller that drives the Peltier cooler based on the reading of the SPADA integrated thermoresistor. The temperatures of both side of the Peltier stage are also monitored for diagnostics and to prevent hazard conditions. The Peltier is powered by a dedicated +12 V supply, due to its high power demand. An 8 bit microcontroller manages all settings, diagnostics, and communication from the detection electronics to either the data processing board or the MACAO equipment. Commands and readings (temperature, overvoltage, hold-off duration) are transferred via RS-232 or USB interface.

In order to provide access for the cold finger, the detection board has a 4×4 cm hole in the centre, with the iAQC stamps mounted vertically between the detector and the board. Figure 12 shows the board with the different subsections highlighted. The board has a diameter of 20 cm and it is built on a six-layer-PCB, with two plane layers for ground and power supplies.

4.2.1 iAQC interfacing and settings. For compatibility reasons with the existing MACAO system, the 60 TTL outputs from the iAQCs are fed to a set of quad-differential RS-485 buffers, model DS26C31, mounted in close proximity to the iAQC socket (see figure 12), thus improving the noise immunity of signals. The 120 lines are then connected to two high-density HD68 SCSI connectors.

The common Gate of all iAQCs enabling the switch-off of all 60 pixels can be controlled either by the microcontroller or externally through a SMC connector and an optocoupler PC357NT, thanks to a wired-OR interconnection: whenever one of the two signals is high, the detector is switched off. The microcontroller can do this

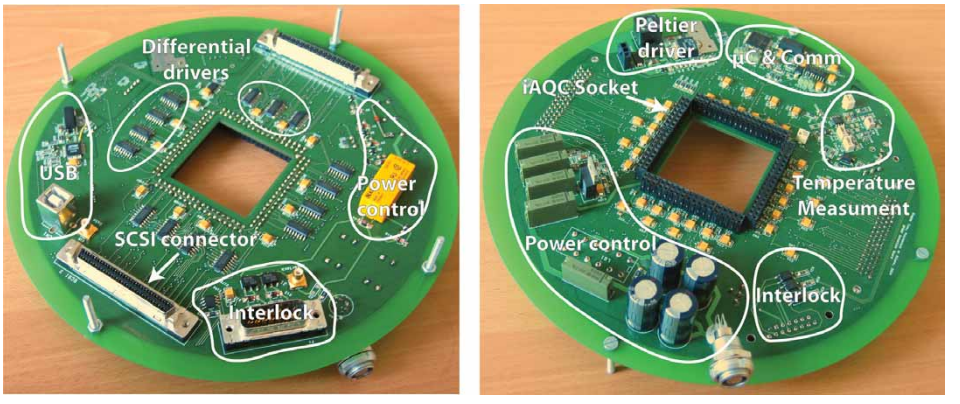

Figure 12. Photographs of the two sides of the detection board, with the different subsections highlighted. (The colour version of this figure is included in the online version of the journal.)

on an user request provided through the serial connection, or as the result of a failure in the board.

The common hold-off setting is carried out through the use of a digital potentiometer AD5220, with 128 taps. It is controlled by the microcontroller through a three-wire connection that permits one to increment or decrement the current value of the DigPot. The desired setting can be saved in the microcontroller's EEPROM so that the hold-off is automatically adjusted at power-on of the system.

The negative supply, common to all SPADs anodes, is adjusted using a linear voltage regulator LM337. This regulator was preferred to a switching regulator in order to have no ripple on the supply and to avoid any disturbance to the critical iAQC section caused by the commutation of the power MOSFETS. In addition, the low current requirement of this supply (less than 120 mA), results in moderate power dissipation of the regulator even with the largest voltage drop applied. The stage achieves a -22.8 V to -13.8 V range for the SPADA anode. Since the positive iAQC supply is fixed at $+12$ V, this corresponds to an overvoltage of up to 8 V for a SPAD with V_{BD}=26.8 V at room temperature, such as those in the SPADA detector.

4.2.2 Temperature control loop. SPAD performance and in particular the dark-counting rate can be greatly enhanced by moderate cooling (about -10 to -20°C). However, since the breakdown voltage is temperature dependent, it is crucial that the temperature remains constant regardless of any external stimulus: a change in breakdown at a fixed applied bias would result in a change in the overvoltage and thus a change in the detection efficiency. For this reason the detection board includes a temperature control loop.

The temperature controller is composed of an integrated thermoresistor, a signal reading and amplification stage, a software PID controller that runs in the PIC microcontroller, a DC–DC Peltier driver and a Peltier stage. The energy dissipated within a SPAD with 25 V breakdown and a 5 V overvoltage is

$W = V{\cdot}I{\cdot}t_{ON} = 30\,V \times 15\,mA \times 5\,ns = 2.25\,nJ$ per avalanche pulse of about 5 ns duration. At the maximum counting rate of 30 Mcps this corresponds to a power dissipation of 68 mW for each SPAD and 4 W for the whole array operating at saturation. This is the heat that the Peltier must be able to remove, to which all thermal losses should be added. The chosen Peltier stage, model CP1.0-71-05 from Melcor, meets these requirements, being a single-stage TEC with a maximum heat removal capability of 18 W, a maximum ΔT of 75°C and dimensions 23×23 mm. The system design included the possibility to use a cold finger with a flowing coolant to easily remove both the heat produced by the SPADA and that produced by the Peltier itself.

4.2.3 Component assembly. As already outlined, the SPADA system is composed of two assemblies, one containing the data processing board and the power supplies for the system, and the other hosting all the detection electronics and the fore-optics. Figure 13 shows a cross-section of the housing. The upper part hosts the fore-optics and ends with the coupling flange that permits connection to the telescope. The central part protects the SPADA detector mounted on the rectified holder and the detection electronics. Finally, the back plate holds all connectors needed and the attachment for the cold-finger coolant. The entire housing is vacuum proof, since during operation it will be filled with dry N_2 to avoid any condensation when the detector is operated below 0°C. The choice to fill the housing with dry N_2 and not only the detector package was made on account of the difficulties in assuring a tight seal and also in consideration of the small distance between the detector and the optics.

The fore-optics is derived from that used in the MACAO system [20]. It is composed of a custom 60 subaperture lenslet array (made by Heptagon) and 60 ball lenses. The 60 subaperture lenslet array is made of two back-to-back keystones shape apertures in order to realize the 45 mm focal length needed with less sag on each lenslet. The 60 ball lenses are placed in the focal plane of the lenslet array, and they are used to focus the light into the SPADs' active areas, from which they are at a distance of only 60 μm. In order to assure perfect alignment of these optical elements with the detector, they are mounted on precisely micro-machined holders, whose relative positions have been chosen after having measured the effective placement of the SPADA on its holder. Figure 14 shows the plate that hosts the 60 ball lenses, together with the clamp that keeps the lenses in position. It should be noted that, as is visible in figure 13, all the optics are fixed onto the rectified plate to which the SPADA is glued (light grey shapes). This ensures that any manipulation of the external case (dark grey shapes) or of the detection board will not affect the optics alignment.

The complete detection electronics, assembled and ready for insertion into the enclosures, is shown in figure 15. The exterior wall of the enclosures can be removed without having to dismount any part of the electronics, thus allowing easy troubleshooting. The overall dimensions of the detection head are about 30 cm in height and 20 cm in diameter.

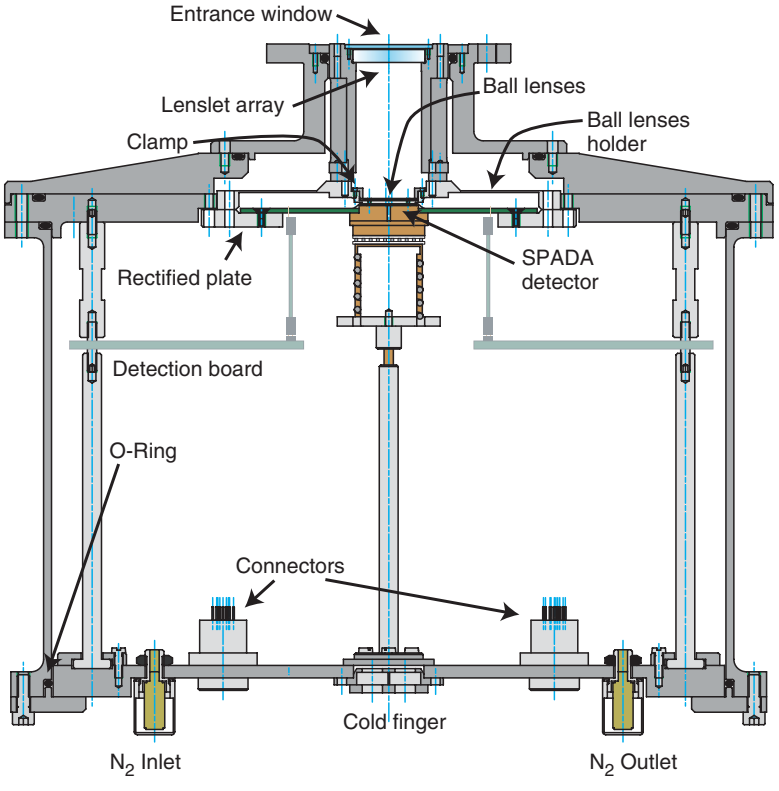

Figure 13. Cross-section of the complete SPADA detection head assembly. During operation it is filled with dry nitrogen to avoid condensation. Sealing is guaranteed by the use of O-rings and vacuum-tight connectors. (The colour version of this figure is included in the online version of the journal.)

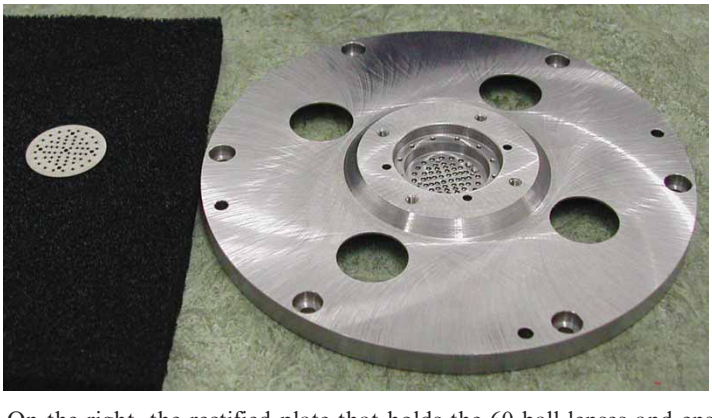

Figure 14. On the right, the rectified plate that holds the 60 ball lenses and ensures perfect alignment of the SPADA detector with the fore-optics. On the left, the clamp that keeps the ball lenses in position, just 60 μm above the detector. (The colour version of this figure is included in the online version of the journal.)

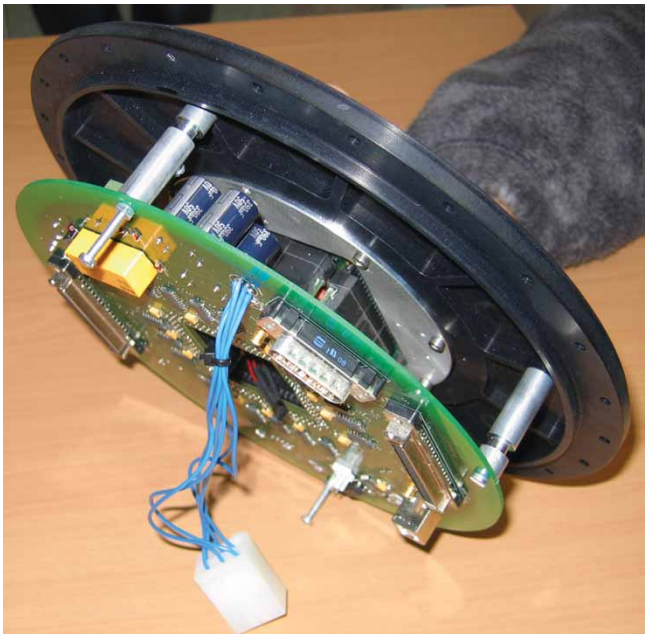

Figure 15. Complete assembled detection head, composed of the detection board, the 60 iAQCs and the SPADA detector mounted in its holder. (The colour version of this figure is included in the online version of the journal.)

5. Experimental results

The performance of the SPADA detectors was investigated through many tests on several different silicon wafers. In particular, breakdown uniformity over the same chip, dark counts, photon detection efficiency and optical crosstalk were measured. These characteristics resulted to be reasonably uniform over different SPADA. All the results reported are for the 50 μm pixels, which proved to be a good compromise between diameter of the active area and noise.

The breakdown voltages for a typical SPADA chip at room temperature show a total spread of less than ±0.5 V. As we will see in a while, this spreading is truly irrelevant for dark counting rate, though it has a minor effect on the detection efficiency uniformity. Figure 16 shows the measured dark counts for a typical SPADA detector at three overvoltages and at room temperature. For easier reading, pixels were ordered for increasing dark counting rate at 5 V overvoltage. It can be noted that all pixels are below 40 kcps at 5 V overvoltage, with a third of them being below 1000 cps. Dark counts are expected to decrease by about 15 times at a temperature of −10°C, on the basis of dark count measurements at low temperature performed on a single SPAD from the same wafer. As is visible in figure 17 (left and centre), it was not possible to find any correlation between dark counts and breakdown voltages. Photon detection efficiency (PDE) was measured in the wavelength range from 400 nm to 1000 nm, using a calibrated monochromator

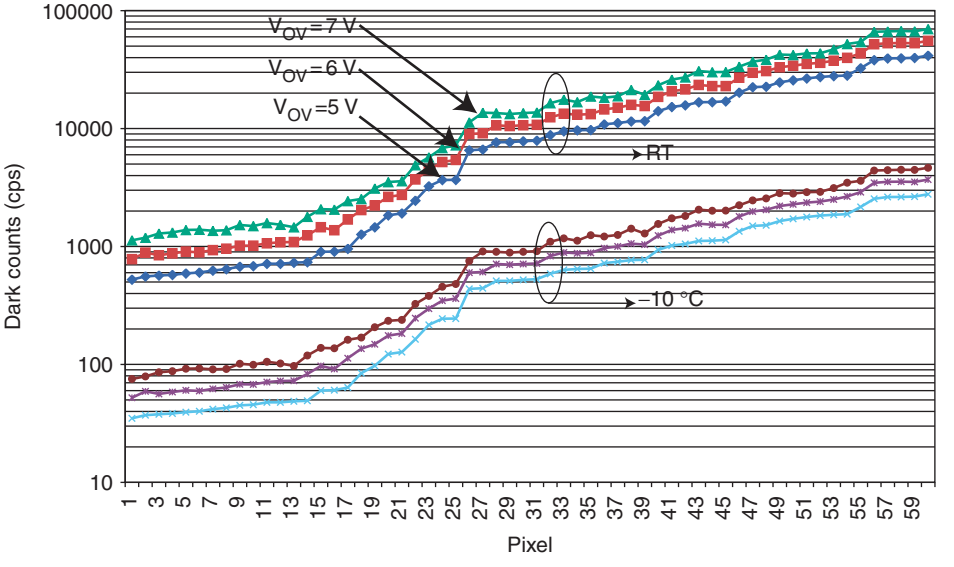

Figure 16. Dark counting rate at room temperature (upper set) for all the 50 μm pixels of a typical SPADA detector at three different overvoltages. The pixels are ordered for increasing dark counts. Expected values at −10°C (lower set), based on known temperature dependence of single SPAD from the same wafer is also reported. (The colour version of this figure is included in the online version of the journal.)

and an integrating sphere to achieve uniform optical power over the entire SPADA detector. There is a broad peak ranging from 400 nm to 700 nm, higher than 30% even at 5 V overvoltage, and also a not negligible few per cent in the near-infrared range. An even better sensitivity can be obtained through the use of an antireflection coating, as shown in figure 18. Figure 17 (left and right) shows the breakdown voltages of a typical SPADA chip, and the photon detection efficiency for the same chip when biased 5 V above the average breakdown. Pixels with lower breakdown, and thus higher effective overvoltage, have higher PDE, and vice versa. This indicates that spreading in the PDE is mainly due to the non-uniform biasing of the detectors in the SPADA system, which uses a common anode, and that the process variations are truly negligible in terms of PDE uniformity over a chip.

Silicon p–n junctions emit photons when operated in the avalanche regime [23]. In a monolithic array of detectors, photons emitted from a SPAD can trigger an avalanche in another detector, thus causing correlations among array elements. For this reason, we measured the optical crosstalk with high accuracy using the coincidence detector shown in figure 19. This is implemented using an external FPGA, thus allowing the simultaneous monitoring of all 60 channels of the SPADA system. In order to measure the crosstalk between two detectors ($SPAD_A$ and $SPAD_B$), the output pulses of their associate iAQCs are first reshaped, extending the corresponding pulse widths to ΔT_A and ΔT_B, and then they are sent to an AND gate. When both $SPAD_A$ and $SPAD_B$ detect an event within a time interval $\Delta T = \Delta T_A + \Delta T_B$, the AND gate produces a coincidence pulse.

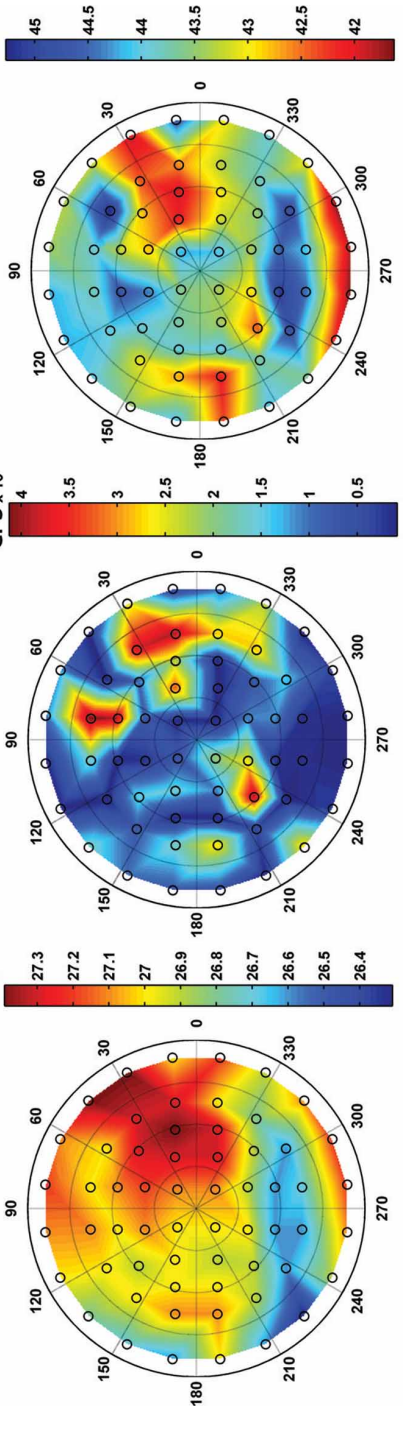

Figure 17. Breakdown voltage uniformity of a typical SPADA chip (left). Dark counts at 5 V overvoltage (centre), showing no correlation with breakdown voltage. Photon detection efficiency (right) at 5 V overvoltage, showing slight correlation with breakdown voltage. (The colour version of this figure is included in the online version of the journal.)

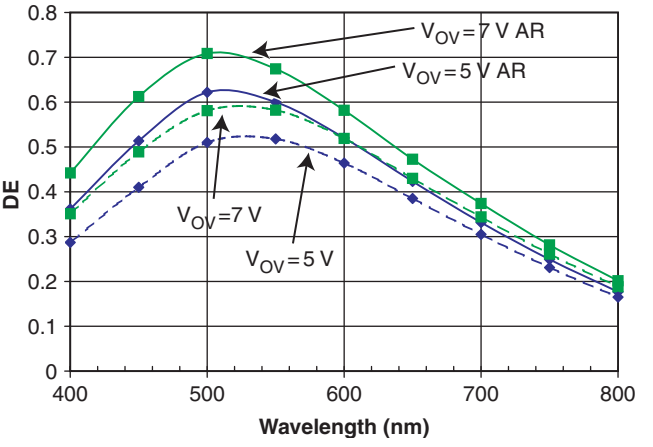

Figure 18. Photon detection efficiency (PDE) for a 50 μm pixel of a typical SPADA detector at different overvoltages, with an optimum anti-reflection coating, made using 78 nm of SiO_2 and 50 nm of Si_3N_4. Dashed lines represent the PDE with no specific anti-reflection coating. (The colour version of this figure is included in the online version of the journal.)

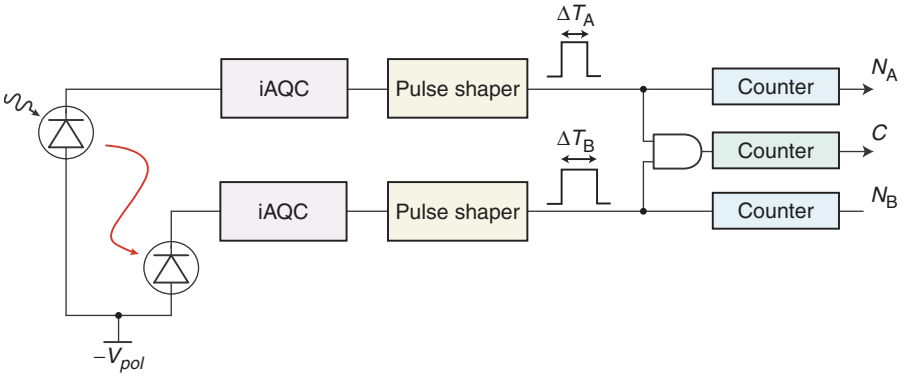

Figure 19. Block diagram of the coincidence detector implemented using an FPGA. (The colour version of this figure is included in the online version of the journal.)

Provided that ΔT is larger than the duration of the avalanche current pulse, all optical crosstalk events generate a coincidence pulse. Three counters accumulate the pulses generated by $SPAD_A$ and $SPAD_B$ and the coincidence pulses for a given measurement time T. The total counts N_A (N_B) includes both the actual counts generated by $SPAD_A$ ($SPAD_B$), and the counts caused by crosstalk from $SPAD_B$ ($SPAD_A$).

The total coincidence counts C includes all the optical crosstalk events, but also uncorrelated dark counts occurring within the time interval ΔT. The average number of these uncorrelated events in the measurement time T is given by

$$C_{uc} = \frac{N_A \cdot N_B \cdot \Delta T}{T}. \quad (4)$$

Therefore, the actual number of crosstalk events in the same time T is given by:

$$C_C = C - C_{uc} = C - \frac{N_A \cdot N_B \cdot \Delta T}{T}. \tag{5}$$

The optical crosstalk probability P_C between $SPAD_A$ and $SPAD_B$ can therefore be calculated as follows:

$$P_C = \frac{C_C}{N_A + N_B - C_C} \cong \frac{C - N_A N_B \Delta T/T}{N_A + N_B}. \tag{6}$$

The accuracy of the optical crosstalk measurement can be evaluated by assuming that the distribution of the coincidence counts is Poissonian. In this case, the standard deviation σ of the coincidence counts would be equal to $\sqrt{C}$. The accuracy in the estimation of the optical crosstalk events C_c is therefore given by

$$\frac{\Delta C_C}{C_C} = \frac{\pm 3\sigma}{C_C} = \frac{\pm 3\sqrt{C}}{C - N_A N_B \Delta T/T}. \tag{7}$$

By performing long measurements ($T \sim 20$ min) we were able to attain accuracy better than 5%.

Figure 20 shows the optical crosstalk probability as a function of distance between 50 μm elements for a SPADA detector, measured at an overvoltage of 5 V. A worst-case crosstalk probability of 4×10^{-4} was measured between two SPADs spaced at 1.5 mm. Strategies for further reducing the crosstalk probability are currently under investigation.

Photon detection efficiency can be increased by operating the SPADA at a higher overvoltage. One may thus think that the higher the overvoltage, the better the

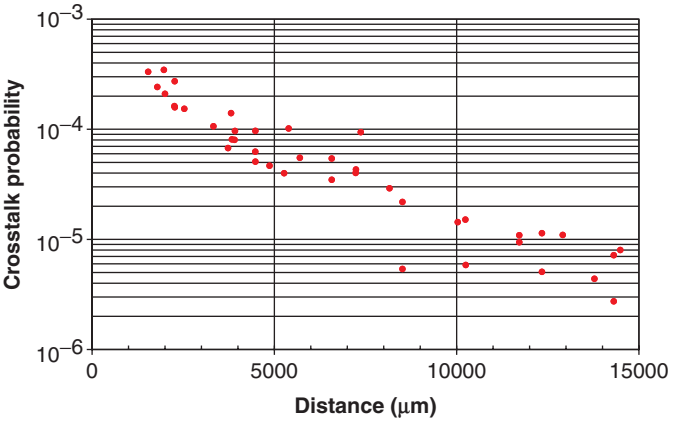

Figure 20. Measured crosstalk probability for a SPADA detector as a function of the distance between 50 μm pixels. (The colour version of this figure is included in the online version of the journal.)

performance of the system. However, this is not true, since the dark counts and hence the intrinsic noise of the detector also increases with the overvoltage, thus impairing the SNR. Moreover, the noise due to the statistical nature of the incident radiation must be accounted for, and the relative weight of the two noise source (detector and radiation) on the SNR depends on the duration of the measurement, following the equation

$$SNR = \frac{\eta \cdot n_s T_M}{\sqrt{\eta \cdot n_s T_M + n_D T_M}}. \tag{8}$$

In order to determine the optimum overvoltage value, a diagram such as the one reported in figure 21 is useful. It represents the minimum signal photon rate needed to give SNR=1 as a function of the integration time, and thus it indicates the maximum sensitivity of the detector for a given operating condition. Figure 21 shows curves at three different overvoltages for a typical SPADA detector at room temperature and at 550 nm, both for the best pixel (high PDE, low dark count) and for the worst pixel (low PDE, high dark count). The sensitivity limit imposed by the radiation statistic is also reported, since this cannot be exceeded by any detector. For each pixel, an optimistic curve corresponding to PDE=100% and to a noise corresponding to 7 V overvoltage is reported, in order to provide an upper limit to

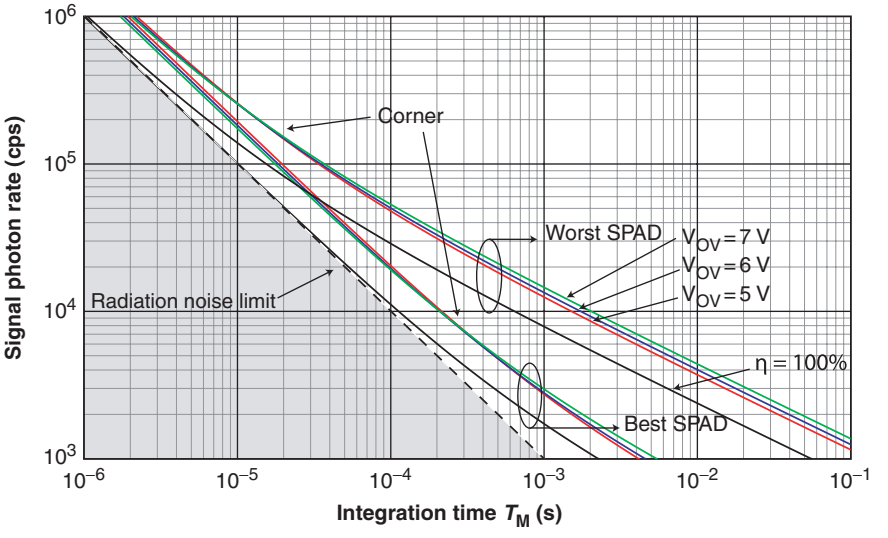

Figure 21. Minimum signal photon rate needed to attain a SNR = 1, as a function of the integration time, for the best and the worst 50 μm pixels of a typical SPADA chip, at room temperature and at 550 nm. For each pixel, the curves for three different overvoltages are reported, together with the one that would correspond to optimistic behaviour with PDE = 100% and the same dark counts as 7 V overvoltage. As a reference, the limit imposed by the radiation intrinsic noise is reported. (The colour version of this figure is included in the online version of the journal.)

the increase in sensitivity that one could ideally achieve. It is possible make the following points:

(a) The curves have two regions, one dominated by detector noise (right-hand side) and proportional to $T_M^{-1/2}$, the other by radiation noise (left-hand side) proportional to T_M^{-1}.
(b) The corner between the two regions is inversely proportional to the dark count, so it moves towards longer integration times as the dark count decreases.
(c) Increasing the overvoltage is effective in increasing sensitivity only if we are working in the left side (noise radiation limited). Otherwise the corresponding increase in dark count impairs the advantage of higher PDE.

From these considerations, it is possible to conclude that working at high overvoltage is useful only at short integration times, when the radiation noise overwhelms detector noise. However, short integration times require strong signal to achieve a good SNR. If the source is too weak to allow short measurement times, the best approach to increase sensitivity, rather than raising the overvoltage, is to cool down the detector to reduce the dark count (thus shifting to the right the region dominated by radiation noise) and if possible applying an anti-reflection coating to boost the PDE to the maximum.

Finally, in order to demonstrate the fast imaging capability of the system, we coupled the detector to a standard camera objective to focus a real-world image on the SPADA. Then, we positioned the detector in front of a CRT monitor, completely dark except for some lines at low luminosity. Figure 22 shows the counts

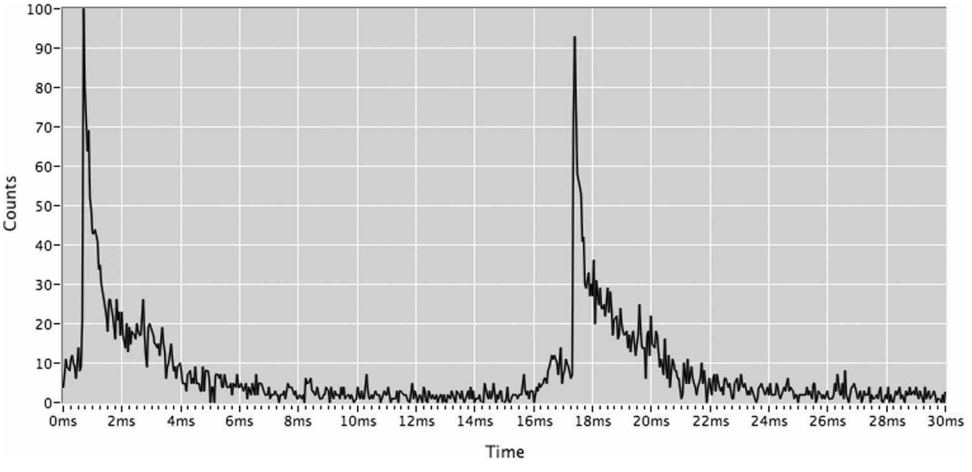

Figure 22. Example of fast transient acquisition. The detector was placed in front of a dark CRT, with only one horizontal line highlighted. The graph reports the counts collected by a pixel of the detector during 50 μs integration windows. The refresh of the CRT at ~60 Hz and the exponential decay of the phosphorous are clearly visible. (The colour version of this figure is included in the online version of the journal.)

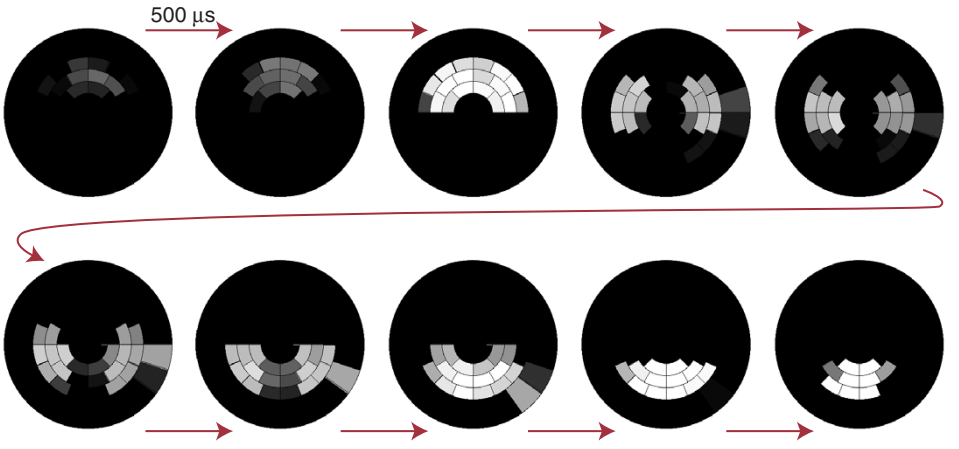

Figure 23. Example of fast transient imaging. The SPADA detector was placed in front of a dark CRT, with only one circular line highlighted. The figure shows ten frames of the SPADA image, acquired with a 50 µs integration window and frames separated by 0.5 ms from each other. The motion of the electron beam is clearly visible as it paints the circle on the CRT with a refresh frequency of ~60 Hz. (The colour version of this figure is included in the online version of the journal.)

accumulated in 50 µs integration windows by a pixel of the detector, when only a horizontal line is highlighted on the monitor. The fast frame capability of the system catches without any problem not only the refresh rate of the monitor (16.6 ms, corresponding to 60 Hz), but also the exponential decay of the phosphorous, showing a time constant of a few milliseconds. Figure 23 shows ten frames separated by 500 µs of the entire detector, when only a circular ring centred on the SPADA is highlighted on the monitor. It is possible to see how, from one frame to the other, the electron beam moves along the screen to paint the circle.

5. Conclusions

We have discussed the design and fabrication of an imaging system based on monolithic arrays of single-photon avalanche diodes to be employed in Astrophysics applications with fast frame rates. We have detailed the design and realization of both the detection and the data processing boards, and we have discussed the integration of the whole system into real astrophysical and adaptive optics applications. The SPADA imager shows state-of-the-art performance, achieving single-photon sensitivity in the visible range, together with low noise, high frame rate (up to 20 kframes/s) and parallel readout. Moreover, the diodes are extremely robust and work at low voltages, making them much more suitable for field and space instrument applications than existing photon counting systems. On-field performance within adaptive optics control loops is currently under investigation. For future high-sensitivity imaging detectors, SPAD arrays with a larger number of elements and subapertures, with square and linear layouts, are under development.

Acknowledgments

This work was supported by the Italian Ministry of University and Research (MIUR-PRIN project No. 2002021224). The authors wish to thank Mr Sergio Masci for the excellent work done in the assembly of the detection head.

References

[1] R.K. Tyson, *Principles of Adaptive Optics* (Academic Press, San Diego, 1991).
[2] M. Downing, N. Hubin, M. Kasper, *et al*., in Proceedings of SDW2005, Taormina, 2005.
[3] D. Bonaccini, F. Rigaut, A. Glindemann, *et al*., Proc. SPIE **3353** 224 (1998).
[4] S. Cova, M. Ghioni, A. Lotito, *et al*., J. Mod. Opt. **51** 1267 (2004).
[5] M. Ghioni and G. Ripamonti, Rev. Sci. Instrum. **62** 163 (1991).
[6] Photon counting detector module, datasheet, Micro Photon Devices, Bolzano, Italy. Available at http://www.micro-photon-devices.com.
[7] H. Dautet, P. Deschamps, B. Dion, *et al*., Appl. Opt. **32** 3894 (1993).
[8] SPCM-AQ single-photon counting module data sheet, Perkin Elmer Optoelectronics. Available at: http://opto.perkinelmer.com.
[9] S. Cova, M. Ghioni, A. Lacaita, *et al*., Appl. Opt. **35** 1956 (1996).
[10] P. Antognetti, S. Cova and A. Longoni, in Proceedings 2nd Ispra Nuclear Electronics Symposium, Stresa May 20–23, 1975, pp. 453–456, Euratom Publication EUR 537e (1975).
[11] S. Cova, A. Longoni and A. Andreoni, Rev. Sci. Instrum. **52** 408 (1981).
[12] R.G. Brown, R. Jones, J.G. Rarity and K.D. Ridley, Appl. Opt. **26** 2383 (1987).
[13] M. Ghioni, S. Cova, F. Zappa and C. Samori, Rev. Sci. Instrum. **67** 3440 (1996).
[14] H. Dautet, P. Deschampes, B. Dion, *et al*., Appl. Opt. **32** 3894 (1993).
[15] F. Zappa, M. Ghioni, S. Cova, *et al*., *IEEE Trans. Instrum. Meas*. **496** 1167 (2000).
[16] F. Zappa, A. C. Giudice, M. Ghioni and S. Cova, in Proceedings of ESSCIRC 2002, Firenze, Italy.
[17] F. Zappa, S. Cova and M. Ghioni, US Patent Appl. No. 09/797,974, filed 5 March 2001; European Patent Appl. No. 01200852.2-2217, filed 6 March 2001.
[18] F. Rigaut, B.L. Ellerbroek and M.J. Northcott, Appl. Opt. **36** 13 (1997).
[19] F. Roddier, Appl. Opt. **23** 1223 (1988).
[20] R. Arsenault, J. Alonso, H. Bonnet, *et al*., *Proc. SPIE* **4839** 174 (2003).
[21] D. Bonaccini Calia, R.M. Myers, *et al*., *Proc. SPIE* **5490** 1315 (2004).
[22] http://www.orsys.de/327c6713.htm.
[23] A.L. Lacaita, F. Zappa, S. Bigliardi and M. Manfredi, *IEEE Trans. Electron Dev*. **40** 577 (1993).