A 16 × 256 SPAD Line Detector With a 50-ps, 3-bit, 256-Channel Time-to-Digital Converter for Raman Spectroscopy

Ilkka Nissinen, *Member, IEEE*, Jan Nissinen, Pekka Keränen, David Stoppa, *Senior Member, IEEE*, and Juha Kostamovaara, *Senior Member, IEEE*

Abstract—A 16 × 256 element single-photon avalanche diode array with a 256-channel, 3-bit on-chip time-to-digital converter (TDC) has been developed for fluorescence-suppressed Raman spectroscopy. The circuit is fabricated in 0.35 μm high-voltage CMOS technology and it allows a measurement rate of 400 kframe/s. In order to be able to separate the Raman and fluorescence photons even in the presence of the unavoidable timing skew of the timing signals of the TDC, the time-of-arrival of every detected photon is recorded with high time resolution at each spectral point with respect to the emitted short and intensive laser pulse (~150 ps). The dynamic range of the TDC is set so that no Raman photon is lost due to the timing skew, and thus the complete time history of the detected photons is available at each spectral point. The resolution of the TDC was designed to be adjustable from 50 ps to 100 ps. The error caused by the timing skew and the residual variation in the resolution of the TDC along the spectral points is mitigated utilizing a calibration measurement from reference sample with known smooth fluorescence spectrum. As a proof of concept, the Raman spectrum of sesame seed oil, having a high fluorescence-to-Raman ratio and a short fluorescence lifetime of 1.9 ns, was successfully recorded.

Index Terms—SPAD detector, Raman spectroscopy, time gating.

I. INTRODUCTION

MEASURING the chemical environment and molecular composition of a medium is important in many applications, e.g. in the agricultural, food, oil and pharmaceuticals industries [1]–[4]. An advanced technique for such measurements is Raman spectroscopy, which can be used to measure samples of various types, e.g. solids, liquids and gases, and allows flexible sample interfacing and preparation [5]. It is also a practical way of identifying aqueous solutions, because the Raman spectrum of water itself is very low [6].

In a typical Raman measurement setup a non-destructive, narrow-band CW (continuous wave) laser beam excites the
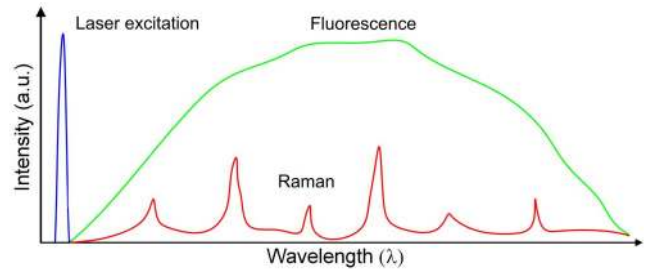
Fig. 1. Raman and fluorescence spectra.

sample and the Raman spectrometer reads out the combination of stray light, fluorescence and vibrational spectrum emitted by the sample. A small number of incident photons (typical scattering probability $\sim 10^{-8}$) interact with the vibrational states of the sample molecules and are then shifted from their original excitation wavelength [7]. This shift can be recorded with a 2-D CCD (charge-coupled device) array under a spectrograph or grating, resulting in a Raman spectrum if the fluorescence background and stray light are not the dominating sources, as illustrated in Fig. 1 (red curve).

One major reason, why conventional Raman spectroscopy is not widely used in some otherwise potential applications is the high fluorescence background, as shown in Fig. 1 (green curve) [8]. This high-level background caused by a substance absorbing the laser light and reemitting it (not a scattering process) masks the weaker Raman signal partly or even totally, and thus the Raman signal-to-noise ratio is typically degraded dramatically. This results in a long measurement time (typically several minutes or tens of minutes). In addition, the fluorescence background does not indicate the molecular information in the form of a Raman signal, and thus it cannot be used to distinguish molecular structures. Raman scattering and fluorescence emission can thus be regarded as two competing phenomena and reduction of the fluorescence background in Raman spectroscopy should be considered an important step towards extending the use of Raman spectroscopy to new fields of application [8].

Some techniques have been presented for reducing the effect of the fluorescence background in traditional Raman spectroscopy, e.g. by selecting of a longer laser wavelength for excitation (whereupon the fluorescence will decrease) together with signal averaging [8]. These techniques can solve the
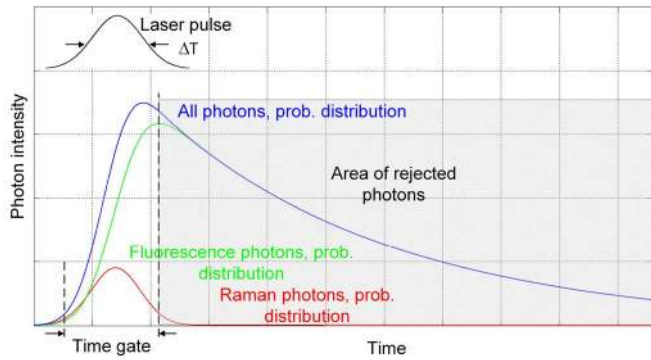
Fig. 2. Principle of fluorescence suppression by means a pulsed laser and time gating.

fluorescence problem in some materials but do not provide a generic solution to the fluorescence background problem because the Raman signal decreases in proportion to the fourth power of the wavelength, so that the signal-to-noise ratio tends to remain low, especially with high fluorescence samples [9]. In addition, while signal averaging can improve the SNR, the recording time also increases above the permitted limits in practical measurement cases.

Luckily, the time responses of these two phenomena, Raman scattering and fluorescence emission are different. Raman photons scatter from the sample instantly, whereas the fluorescence photons are typically emitted with a time constant of a few nanoseconds or more. This opens up a possibility for reducing the level of fluorescence by using a short laser pulse instead of a CW laser to illuminate the sample and then collecting the scattered photons only during the laser pulse [9], as depicted in Fig. 2. This time-gating technique results in a suppression of the fluorescence which is roughly proportional to the ratio of the time-gate length to the fluorescence lifetime of a sample [9]. To achieve a meaningful suppression (a factor of 10 or more) with samples having fluorescence lifetimes in the nanosecond range, the time-gate width, and thus also the laser pulse width, should be in the sub-ns range. The time-gating can be realized by means of a Kerr gate [10] or an ultra-fast gated intensified CCD [11], [12], for example. These measurement systems are used mainly in laboratories, however, but not in the field, since they are not affordable or portable.

As is well-known, CMOS technologies can be used to fabricate photodetectors and also single-photon avalanche detectors (SPAD). As mentioned above, the probability of Raman scattering is very low, which makes SPADs suitable for use as detectors in Raman spectroscopy, as demonstrated in [13] and [14]. A SPAD can be triggered by a single photon and it gives a digital output signal with a jitter of ~100 ps (FWHM), so that the signal intensity can be derived simply by counting the number of photons at each spectral point, provided that the photon detection probability is markedly less than one. Thus single-photon detection-based Raman spectroscopy works in principle in the same way as any other time-correlated single-photon measurement [15]. In addition, CMOS technologies allow the time gating and reading electronics to be integrated into the same die as the

SPAD array, enabling the recording of the whole Raman spectrum [16]–[20]. As a result, the size of the whole Raman setup can be reduced markedly, paving the way to off-laboratory "low-fluorescence" Raman spectroscopy for use in potential applications within the fields of agriculture, food, pharmaceuticals and the oil industry, and also in crime investigation, for example.

There are some earlier studies which have used relatively large CMOS SPAD arrays and ns-scale time-gating in Raman spectroscopy. The time resolution of a TDC and measurement rate of 430 ps and 19 000 lines/s, respectively, were achieved with a 2×256 SPAD array in [19]. In [17], a detector array 8×1024 SPADs was demonstrated with a time gate width and measurement rate of 700 ps and 1052 spectra/s, respectively. Also, successful measurement of the Raman spectra were demonstrated with a sample having some fluorescence background level. A time gate of 3.5 ns was used with a single SPAD in [18], but the time position of the gate with respect to the Raman signal could be adjusted so that an effective time gate of 250 ps could be achieved.

Time gate widths of a few hundred picoseconds or even a couple of nanoseconds are adequate in e.g. many mineralogical applications where the typical fluorescence lifetimes are longer than 10 ns [17]. However, to make SPAD-based Raman spectroscopy more practicable for most of the potential in situ applications with highly fluorescent samples having nanosecond-scale fluorescence lifetimes, a shorter time gate width in the deep sub-ns range (~100…200ps) is needed to achieve a high enough fluorescence rejection ratio. The laser pulse length should obviously be matched with the time gate width, and a relatively large SPAD array and high spectrum recording rate would also be needed.

A time gate width of approximately 100 ps was achieved with a 2×128 SPAD array with three adjustable time gates to collect both the Raman and fluorescence photons [21], while in [22] a 4×128 SPAD array with a 512-channel TDC was used to measure the time-of-arrival of photons with a resolution and spectrum rate of 78 ps and 380 000 spectra/s, respectively. The latter one has however a measurement range of 300 ps only with the resolution of 78 ps. These receiver circuits show quite reasonable time gating performance, but the number of spectral pixels in an array is only moderate. Moreover, the results presented in [21] and [22] showed that the timing skew of the time gates throughout the array (~140 ps) was quite large compared to the time gate width [21] or to the time measurement range of the on-chip TDC [22]. The timing skew obviously introduces distortion to the measured Raman spectrum, since the spectral points are not sampled simultaneously in time. This is illustrated in Fig. 3 where the time position of the laser pulse (and thus the Raman peak) does not change as a function of the wavelength, but the time gate width and position vary at different wavelength points due to the skew in the timing signals (for clarity, only one time gate is shown in Fig. 3). In other words, the relative positions of the Raman peaks vary inside the time gate, causing a timing-skew-dependent variation in the recorded Raman signal (number of counts within the bin). As a result, the recorded Raman spectrum
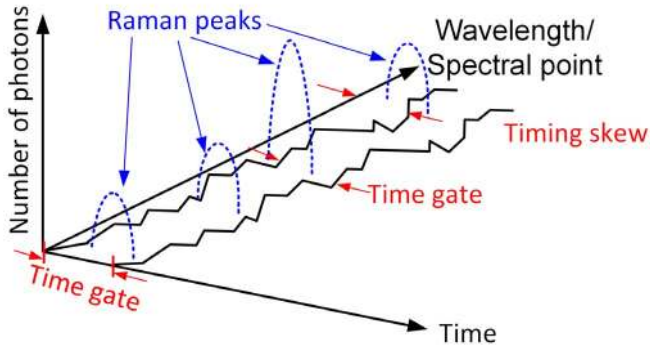
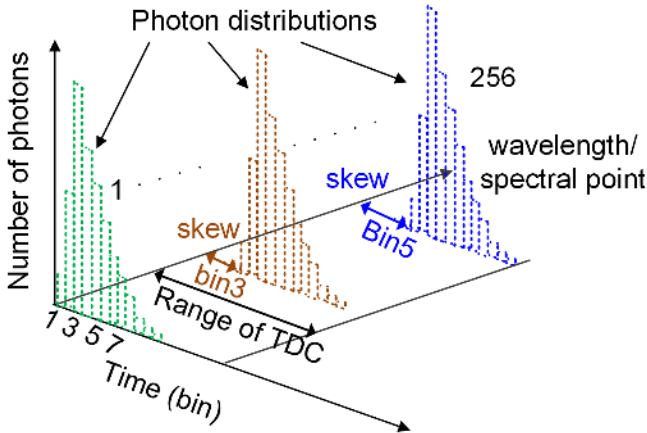Fig. 3.   Principle of the timing skew as a function of wavelength.

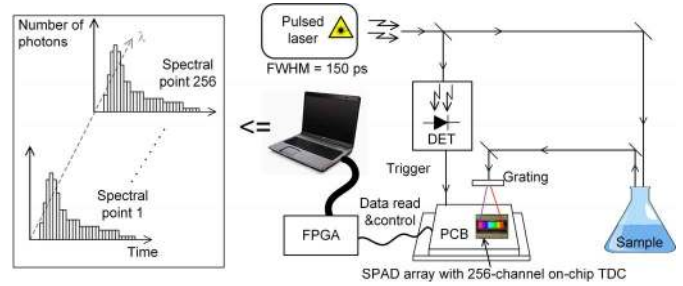Fig. 4.   Principle of recording the times-of-arrival of the photons at each spectral point.

Fig. 5.   Block diagram of a time-gated Raman spectrometer consisting of a pulsed laser (Full Width Half Maximum of 150 ps), a trigger detector (DET), a grating, a printed circuit board (PCB) with a SPAD array IC, a Field-Programmable Gate Array (FPGA) circuit and a laptop.

is distorted due to the timing skew. In addition to distorting the recorded spectrum, the timing skew deteriorates also the signal-to-noise ratio of the measurement, especially if the sample has a high fluorescence background with a short, nanosecond-scale fluorescence lifetime [23].

The sampling error introduced by the timing skew is further increased if the number of spectral points, and thus the number of SPADs in a line array, is increased. As a result, if sub-ns time gating is aimed at with 256 spectral points, for example, a straightforward time gating approach based on a single or double (for fluorescence background compensation) sampling windows may not be practicable any more. Another approach would be to record the time-of-arrival of each detected photon at every spectral point with high-enough time resolution, so that the complete time history of the detected photons (N = f(t)) would be available for each spectral point. In this kind of a system, the measurement range of the time-to-digital conversion has to be set so that the meaningful part of the distribution can be fully recorded even in the presence of the timing skew, see Fig. 4. When all the Raman photons are collected in each spectral point, in principle no information is lost due to the timing skew and the time alignment of the spectra at different wavelength locations can be achieved from the measured data by post-processing.

One straight-forward technique for post-processing is to use a calibration measurement with a sample providing a known

smooth (as a function of wavelength) fluorescence background. Since the timing skew effects (variations in the absolute time position and widths of the individual bins of the various TDC channels) as well as the variation in the photon detection probability of the individual SPADs in the row would cause systematic variation to the recorded spectrum, the measured result from the "smooth" spectrum can be used to partly compensate for these effects. As a first-order approximation, the timing along the wavelength spectrum could be aligned by the bins gathering the maximum number of counts (As shown in Fig. 4 the bin collecting the maximum number of counts is changing along the wavelength). Then, based on the known continuity (along wavelength scale) of the reference spectrum, the weighting factors for each spectral points could be derived. The amount of improvement achieved in the quality of the spectrum depends of course on the resolution and precision of the TDC used.

In this paper, a custom-designed 16 × 256 single-photon avalanche diode (SPAD) array which includes also a 256-channel, 3-bit on-chip time-to-digital converter (TDC) on the same silicon die is presented. The design of the chip follows the above design principles and the effectiveness of the suggested techniques are demonstrated using samples (olive oil and sesame seed oil) which are known to be difficult to measure with conventional continuous wave (CW) Raman spectroscopy due to the high fluorescence background and relatively short fluorescence time constant. In addition, different oil samples with various fluorescence level enable an efficient way to verify the effectiveness of the fluorescence suppression of a specific Raman setup. The circuit was manufactured in a 0.35 μm high-voltage CMOS technology.

In the following, the basic principle of the time-gated Raman spectroscopy is described in Section II. Section III describes the architecture and electronics of the chip developed. The measurement results are given in Section IV, with specific emphasis on the effect of the inhomogeneity of the timing on the performance of the measurement system. The conclusions are given in Section V.

II. TIME-GATED RAMAN SPECTROSCOPY USING A CMOS SPAD LINE ARRAY

A block diagram of the time-gated Raman spectroscope is shown in Fig. 5. The device consists of a 532 nm pulsed laser

(a custom made Q-switched solid state laser) as an illuminator, with a short optical pulse width of ∼150 ps (FWHM) and a pulse energy of 0.2 μJ, some optics, a diffraction grating and a SPAD-based line sensor with a 256-channel, 3-bit TDC. The pulse rate of the laser is approximately 50 kHz. To derive the Raman spectrum, the sample to be measured is illuminated by the short optical pulse from the laser. Simultaneously a synchronization signal (Trigger in Fig. 5) is generated optically by a fast optical detector (DET in Fig. 5) to bias the SPADs to the Geiger mode so that they are able to detect individual photons. In addition, a delayed start signal is sent to the TDC so that the times of arrival of the photons can be measured by the channels of the TDC at each spectral point. An FPGA circuit controls the operation and reads the measured time intervals from the detector chip. The Raman photons scattered from the sample have different wavelengths and thus diverge upon encountering the diffraction grating, so that the position of each SPAD in the line sensor corresponds to a specific range of wavelengths. A Matlab program is then used to derive histograms of the times-of-arrival of the photons as a function of the spectral point, as shown in Fig. 5.

III. THE SPAD-BASED LINE SENSOR WITH A 256-CHANNEL 3-BIT TDC

A block diagram of the proposed line sensor with 16 × 256 SPADs and a 256-channel on-chip TDC is shown in Fig.6. The SPAD array, which consists of 16 columns and 256 rows (spectral points), is presented in the middle of the figure. 256 rows cover the wavenumber range of approximately 1600 1/cm which is adequate e.g. for oil samples [24]. Each row consists of 16 SPADs which all collect photons to one point in the Raman spectrum. The number of SPADs in each row has been matched to the area of photon distribution of the used grating. The columns are placed in groups of two so that the front-end electronics, a hot pixel elimination circuit (HPE in Fig. 6) and an OR gate, can be placed after every other column, as shown in the zoomed area in Fig.6. The front-end electronics include active quenching and loading transistors and a sensing element, an inverter. Each pixel in the row is connected through an OR gate tree to the input of a register, as shown in Fig. 6. The hot pixel elimination circuit can be used to disable SPADs with an abnormally high dark count rate. Hot pixel selection can be programmed serially before the measurement.

The loading and quenching pulses for the SPADs (Load and Quench in Fig. 6) are generated from four off-chip rising edges, the delays in which relative to each others can be adjusted independently by programmable off-chip delay units with a resolution of 250 ps. The Load and Quench pulses are used to bias the SPADs to the Geiger mode or quench them, respectively. In order to achieve simultaneous loading and quenching signals for all the SPADs, a tree-like wiring and buffering system from the pulse generator is used.

The 256-channel time-to-digital converter (TDC), shown in Fig. 6, is placed next to the SPAD array. Two identical 7-phase delay lines are used to deliver the phases of one common delay line (replica DLL in Fig. 6) all the way through
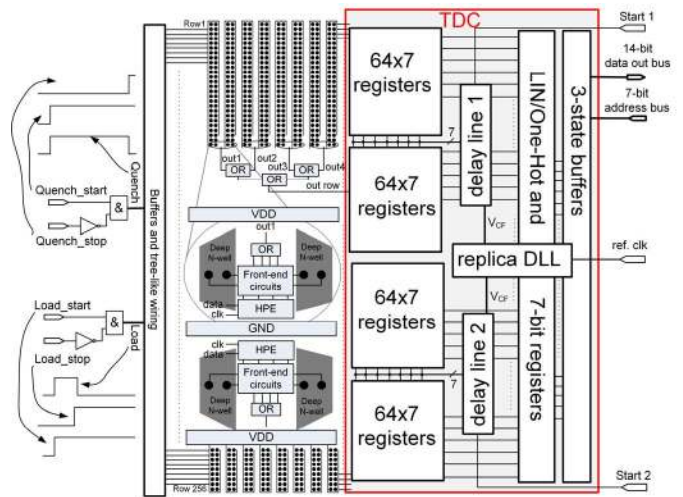

Fig. 6. Block diagram of a time-gated 16 × 256 SPAD line sensor with a 256-channel TDC consisting of a replica delay locked loop (DLL), two delay lines, a linear-to-one hot coder (LIN/One-Hot), registers and three-state buffers.

the chip, and therefore both delay lines have to deliver their phases to only half of the registers, as shown in Fig. 6. This is supposed to reduce the timing skew of the phases of the delay lines caused by mismatch between the wiring and the buffers. Each spectral point in the array has its own channel in the TDC. The TDC operates by storing the state of the phases of the delay line in a 7-bit register for each spectral point as a response to the detected photon. Temperature and supply voltage variations in the TDC are compensated for by using a replica delay-locked loop which is delay-locked to an off-chip reference clock. The resulting control voltages, VCF, are used to stabilize the actual delay lines 1 and 2. The resolution, and thus also the dynamic range, of the TDC can be adjusted by changing the frequency of the reference clock within the range 50–100 MHz, giving a resolution of approximately 50–100 ps and a dynamic range of 350–700 ps. Since the Raman photons appear during the short laser pulse (FWHM = 150 ps), the above measurement range of the TDC of less than 1 ns was assumed to be sufficient with regard to the expected timing skew.

The thermometer codes of the TDCs are converted to one-hot codes and saved in the corresponding cell of the 7-bit register. Reading of the data is controlled by an FPGA circuit which sets the 7-bit addresses of the registers and reads the data from the SPAD IC to the FPGA circuit using a 14-bit parallel Data out bus (shown in Fig. 6). This bus makes it possible to read the results of two 7-bit registers (results from two rows) per FPGA clock cycle (50 MHz), and thus the reading of the results of every row takes 128 clock cycles (256 rows) resulting in a reading time of approximately 2.5 μs. For this reason the maximum laser pulse rate can be approximately 400 kHz and the maximum spectrum frame rate 400 kframes/s.

A. The SPAD Detector and Front-End

A schematic diagram of a single row (16 parallel SPADs = one spectral point) of a SPAD array with the corresponding
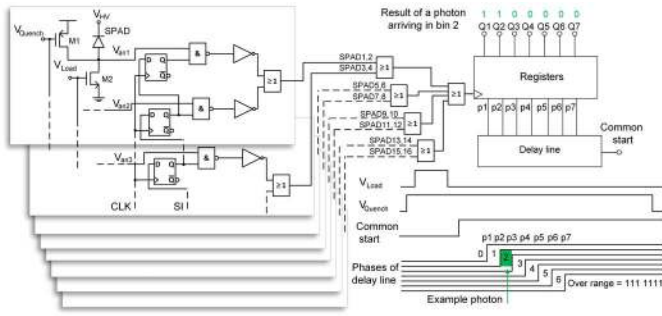
Fig. 7. Schematic and timing diagrams of a single SPAD row.

Fig. 8.   a) Block, b) schematic and c) timing diagrams of the TDC delay line.

electronics is shown in Fig. 7. The SPAD structure is based on a p+/DNWELL diode with a PWELL guard ring as described in [25] and [26]. Each pixel has its own quenching and loading transistors. There are eight pairs of SPADs in an each row, arranged in such a way that each pair is connected through an OR gate to the next OR gate of an adjacent pair of SPADs and so on. This equalizes the timing skew between the SPADs in a single spectral point. The possible timing skew between the parallel pixels at a single spectral point would cause additional error into the time domain measurement, because the SPADs in a row are randomly triggered by photons entering the area of a single spectral point (the light intensity is almost uniformly distributed over the whole area of a row by an output slit in the spectrograph). In addition, a serially programmable 1-bit memory cell (D-flip flop in Fig.7) is placed next to each pixel to disable the outputs of hot pixels.

A timing diagram of the operation of the array is also shown in Fig. 7. Before the measurement cycle all the SPADs are kept quenched (the anode node V_{an1} is set at 3.3 V) by means of the PMOS transistor M1, the gate signal of which (V_{Quench}) is kept at the ground. Then the SPADs cannot detect photons because the voltage above them is less than the breakdown voltage. In the next phase, a synchronization signal from the laser starts the measurement cycle and the gate signal of the PMOS transistor (V_{Quench}) goes high, leaving the anode of the SPAD (V_{an1}) to float for a short moment until it is discharged to the ground by the signal (V_{Load}) at the gate of the NMOS transistor M2. The voltage over the SPADs is now larger than the breakdown voltage and they are ready to detect individual photons. At the same time, the time-to-digital conversion is started by the start signal (Common start in Fig. 7) which generates seven phases in the delay line (p1 – p7), the states of which are stored in a register by an arriving photon, as shown in Fig. 7. Note also that the SPAD can be triggered by a dark count between the rising edge of the loading pulse and the first rising edge of the delay line, resulting in a recorded state of 000 0000, because the voltage over the SPAD is immediately above the breakdown voltage after the rising edge of the loading pulse. Thus this zero result (bin 0) is not used in the time-of-arrival measurements because of the higher noise level and longer effective photon capturing time window. The laser pulse, and thus the Raman response, is adjusted to occur after the first rising edge (p1) by the common delay control of the phases of the TDC (the electrical delay in the
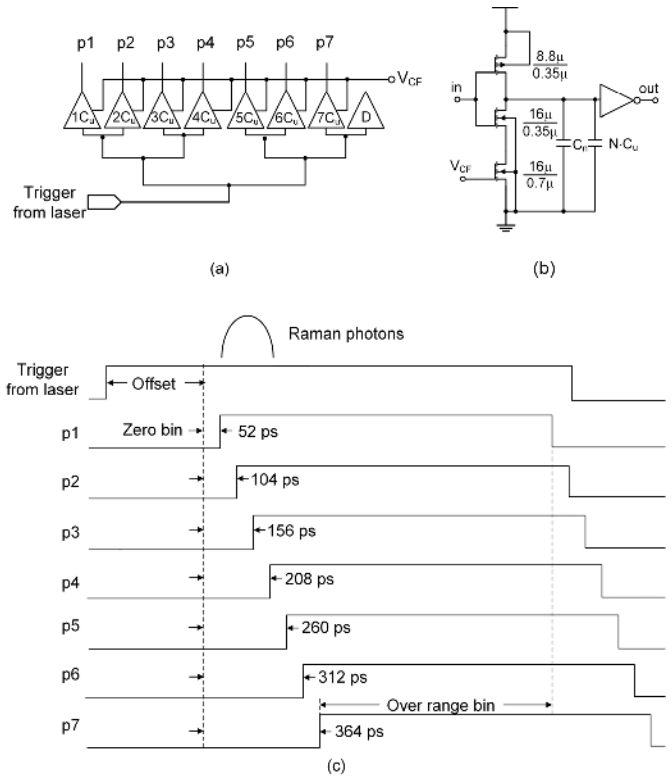
triggering signal from the laser has to be adjusted by an off-chip adjustable delay unit to compensate for the optical delay in the photons passing from the laser to the sample and back to the SPAD detector.).

The SPADs are quenched by the quenching signal V_{Quench} and kept quenched until the next synchronizing signal from the laser, and thus the after-pulsing probability can be kept low even at the highest possible pulse rate of 400 kHz.

B. Time-to-Digital Converter

The time-to-digital converter is based on parallel-scaled delay lines [27] (delay line 1 and delay line 2 in Fig. 6), a block diagram of which is shown in Fig. 8 a). The delay lines are stabilized using the control voltage V_{CF} from a replica delay line, the operation of which is explained below. Both delay lines consist of seven parallel delay elements, a schematic diagram of which is shown in Fig. 8 b). The delay elements have a different number of unit capacitances as the load, as shown in Fig. 8 a) and b). The time difference between adjacent phases is now caused by the difference of one unit capacitance between adjacent elements, resulting in a nominal 52 ps delay difference between all the phases (p1 – p7), as shown in Fig. 8 c). Note that the total propagation delay of the buffers is the sum of the offset delay (caused by the nominal capacitance C_n shown in Fig. 8 b)) and the delay caused by the number of unit capacitances as an additional load. These seven phases are triggered by a trigger signal from the laser starting the measurement of the time-of-arrival of a photon, and the states of the phases are stored in
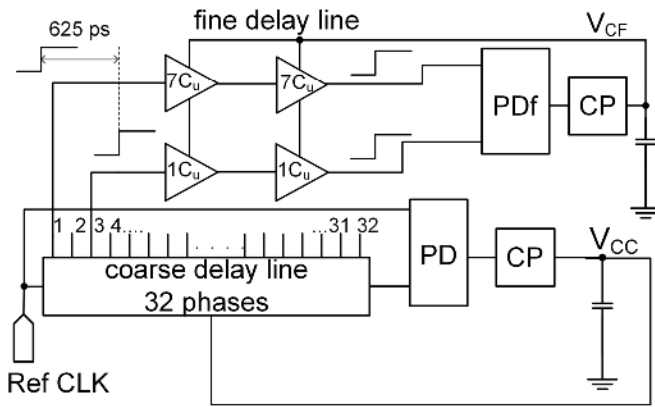
Fig. 9. Block diagram of a replica delay-locked loop consisting of a fine delay and a coarse delay lines, phase detectors (PDf and PD), charge pumps (CP) and a reference signal (Ref CLK).

Fig. 10. Photograph of a printed circuit board with the detector chip.

7×256 registers by the detection of a signal from the SPADs of each spectral point, as shown in Fig. 6. When the delay in the optical path is compensated for with an external electrical delay, as explained above, the arrival of the Raman photons overlap with the TDC phases, as shown in Fig. 8 c).

The delay lines 1 and 2 are stabilized by means of a replica delay-locked loop (replica DLL in Fig. 6), a block diagram of which is shown in Fig. 9. A 32-phase coarse delay line is locked nominally to a 100 MHz reference clock signal (Ref CLK in Fig. 9), resulting in a delay difference of 312.5 ps between adjacent phases in the delay line [28]. The coarse DLL is used only to stabilize the delays in the fine delay line based on parallel-scaled delay elements (shown in Fig. 8 b)). The delay of a parallel-scaled delay element with a load of 12 unit capacitances $(7 + 7 - 1 - 1)$ is set to be equal to the nominal delay difference between the 1st and 3rd phases of the DLL ($2*312.5 = 625ps$) and thus the LSB of the TDC (based on a single unit capacitance) corresponds to 52ps at the reference frequency of 100 MHz. Note that the control voltage of the fine delay line (V_{CF}) is also used to stabilize the actual delay lines of the TDC.

As was mentioned above, the delays in the delay lines, and thus the resolution of the TDC, can be changed by altering the frequency of the reference clock. The nominal LSB of 52 ps is achieved at a 100 MHz clock frequency, but the LSB can be degraded to 104 ps by using a 50 MHz clock if a longer measurement range is needed. The most accurate sampling of the Raman photons can be achieved with the highest resolution, but a lower resolution and thus a longer measurement range, will enable a larger timing skew to exist throughout the spectral range.

IV. MEASUREMENT RESULTS

The time-gated 16×256 SPAD line detector with a 256-channel, 3-bit TDC and having adjustable resolution between 52 ps and 104 ps was fabricated in a 0.35 μm High Voltage CMOS technology. A photograph of the printed circuit board with the detector chip is shown in Fig. 10. The size of the chip is 9 mm $\times$ 3 mm. The breakdown voltage of the
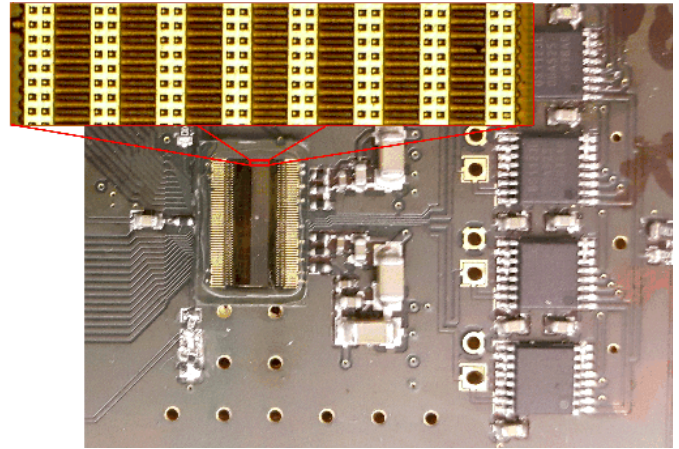
SPAD was measured to be approximately 19 V and the fill factors of the line detector along the spectral axis and the whole array are 26% and 12%, respectively.

The dark count rates (DCR) of the SPADs were measured in the time-gated operation mode (during normal Raman measurement, but without any sample) and hot pixels were eliminated so that the best possible performance in the Raman measurement could be achieved. Hot pixels were eliminated so that the dark count probability (DCP) of an each spectral row was less than 10^{-3}. The resolution of the TDC was set to 100 ps, resulting in an effective dynamic range of approximately 600 ps (52 ps was not used here because of the reduced time range for the Raman measurement, as explained below). After elimination of the hot pixels 12 out of the 16 spatial SPADs available were selected to keep detector area identical in every spectral row, resulting in a total receiving aperture of approximately 3900 μm^2 (12 $\times$ 18 μm $\times$ 18 μm). The average dark count probability (DCP) of the 256 spectral rows within the whole effective dynamic range of the TDC ($\sim$600 ps) was measured to be approximately $3 \cdot 10^{-4}$ implying that the SPADs in each spectral row will be triggered by noise approximately 30 times when 100 000 laser pulses are shot at the sample In addition, the dark count noise levels of all the spectral rows were stored, so that the average noise level could be subtracted from the Raman spectrum.

A. Timing Skew Measurement

The timing skew of the phases of the actual delay lines throughout the line detector was measured by means of a pulsed laser (FWHM 60 ps), which was used to illuminate the whole area of the line detector, and the time range of the TDC (7 phases) was swept over the laser pulse in a synchronized measurement, as is shown in Fig. 11 a) and b). The trigger signal from the laser produces two other signals (Load and Start in Fig 11 a) and b)) which both are delayed with adjustable coaxial cable delay units so that the delays in the optical and electrical paths are matched. The SPADs are loaded into the Geiger mode by V_{Load} and all the phases of the delay line (DL phases p1 – p7 in Fig 11 a) and b)) arrive before the laser pulse, as is shown in Fig. 11 b). All the SPADs
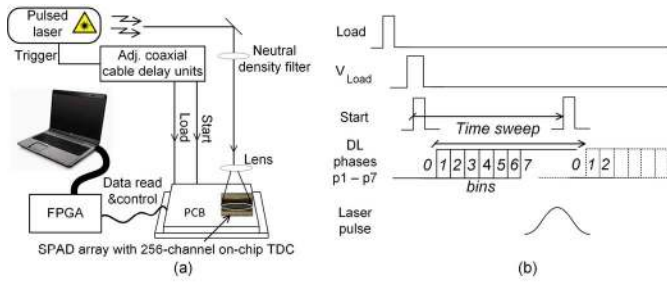
Fig. 11. a) Block and b) timing diagrams of a setup measuring the timing skew SPAD array with a TDC.
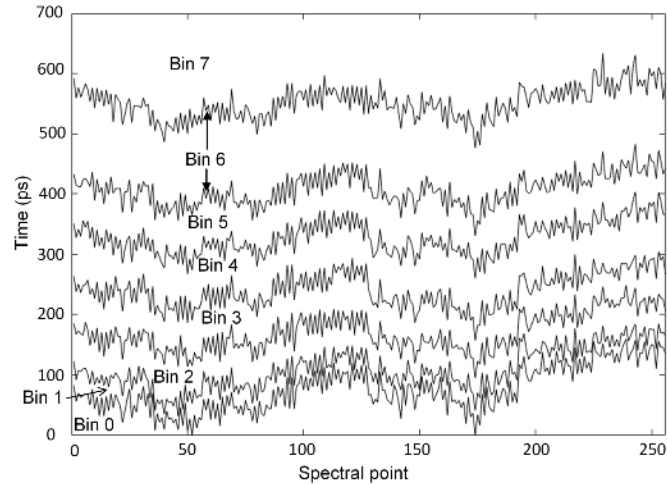

Fig. 13. Timing skew of the Raman spectrum as a function of the spectral point.

Fig. 12. Timing skews of the phases of a TDC.

Fig. 14. Impulse response functions of the whole Raman system at three different spectral point, 82, 118 and 148.

triggered by the laser pulse are then collected into bin 7 of the TDC (the over range bin). After that, only the start signal and thus the phases of the delay line (the loading pulse was kept in its original position), were swept over the whole laser pulse to record the time position of the laser pulse envelope as a function of the spectral point. The photon detection probability was adjusted with a neutral density filter to be less than 2% in order to limit pile-up distortion. The timing skews of the phases of the actual delay lines as a function of the spectral point are shown in Fig. 12. The maximum skew was ±75 ps. All the Raman photons can be collected within the dynamic range of the TDC, but the highest resolution of ∼50ps cannot be used without losing Raman photons and introducing distortion into the spectrum. As explained above, the timing skew causes an additional error source degrading the signal-to-noise ratio of the Raman spectrum, because the Raman signal is sampled differently at different spectral points. This sampling error is constant between measurements, however, and thus it can be partly compensated for by using reference sample measurement. The key requirement for this is that the TDC range should cover the whole time range of the Raman photons, including the skew effects, since only then no information is lost during measurement.

In addition, the impulse response function (IRF) of the whole Raman device was assessed by using a real Raman setup (shown in Fig. 5) with a fluorescence-free sample (a mixture of a toluene and an acetonitrile). An adjustable coaxial delay
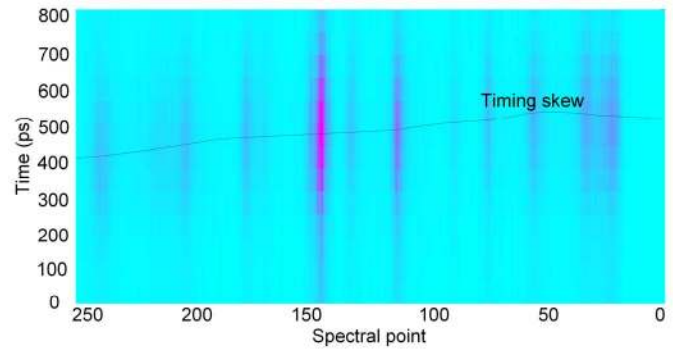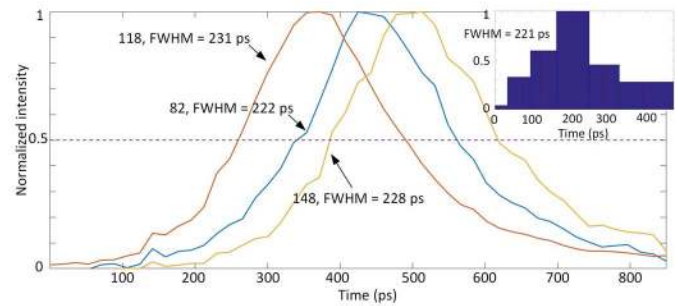
unit was used to sweep the phases of the delay line over the Raman scattered photons with steps of approximately 17 ps, as shown in Fig. 11 (Raman scattered photons occur during the laser pulse). Instead of using the whole dynamic range of the TDC to collect photons, only the fifth bin of the TDC was used during each step, to achieve linear sweeping over the whole dynamic range. The fluorescent-free sample had ten Raman peaks, which can be seen in Fig. 13 as different brightnesses of the purple colour at different spectral points (the brightness shows the intensity of the Raman signal). The IRFs of the whole Raman system are shown in Fig. 14 derived at three different spectral points 82, 116 and 148. The full width of half maximum values of IRF were calculated to be approximately 225 ps as shown in Fig. 14. The variation of the time position of peak amplitudes is showing the timing skew of the SPAD array. Additionally, timing skew line as a function of spectral point has been drawn to illustrate the variation of the maximum amplitudes of Raman peaks in time domain in Fig. 13. The FWHM of the IRF was also derived from the data of the on-chip TDC (LSB adjusted to be 50 ps) measuring the time-of-arrivals of Raman photons during normal operation. The histogram of the time-of-arrivals of Raman photons at the single Raman peak is shown in small figure in Fig. 14 with a FWHM value of 220 ps derived from the histogram.

B. Raman Measurements

Two samples, olive oil and sesame seed oil, having a high fluorescence background and nanosecond-scale fluorescence

TABLE I
COMPARISON OF TIME-GATED RAMAN SPAD SENSORS

	This work	[21] Nissinen et al.	[18] Deen et al.	[17] Maruyama et al.	[19] Kufsak et al.	Unit
Technology	0.35μm HVCMOS	0.35μm HVCMOS	0.13μm CMOS	0.35μm HVCMOS	0.13μm CMOS	
No. of pixels	16 x 256	2 x(4)x 128	1	8 x 1024	2 x 256	
Fill factor along spectrum	26	23	N/A	4.9 & 44.3	43.7	%
Pitch along spectrum	35	32	N/A	24	~24	μm
Min. Gate width	52[a]	100	250	700	430[a]	ps
Max. spectrum rate of sensor	390 (50)[b]	50[b]	N/A	N/A	19	kspectra/s
FWHM of IRF	225	N/A	165	1100	620 (>1000)[c]	ps

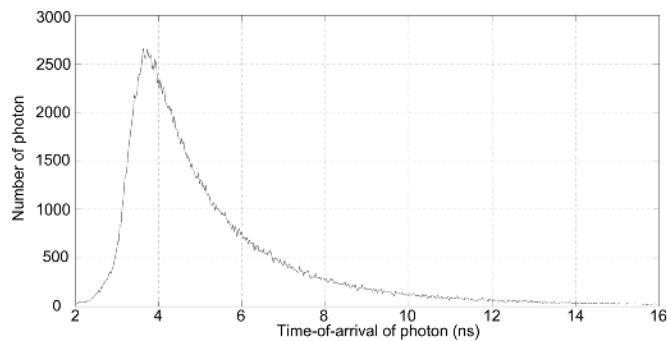[a]Actually the LSB of the on-chip TDC [b]The laser pulse rate is limiting [c]Often this value

Fig. 15. Scattered photon distribution as a function of time, measured with a pulsed Raman system and an external TDC (LSB = 24 ps) at the Raman peak. The result is a convolution of the instrument response function of the system and the Raman and fluorescence scattered photons.

Fig. 16. Raman spectra of olive oil a) without any correction, b) with residual fluorescence correction and c) with reference sample correction.

lifetimes, were used to test the functionality of the time-gated Raman spectrometer chip developed here. The fluorescence level of the sesame seed oil was measured to be one decade higher than that of the olive oil, and the time constants of the fluorescence lifetimes of the sesame seed and olive oils were 1.9 ns and 2.5 ns, respectively. The time distribution of the scattered photons from a sesame seed oil sample measured with a Raman system having an external TDC (resolution 24 ps) at a single spectral point is shown in Fig. 15. This distribution thus shows both the Raman and fluorescence photons (fluorescence lifetime ∼2 ns). In the Raman measurements the resolution of the TDC was adjusted to be 104 ps in order to collect all the Raman photons excited by the laser pulse with a width of 150 ps (FWHM) within the usable bins of the TDC while also taking into consideration the timing skew through the whole spectral
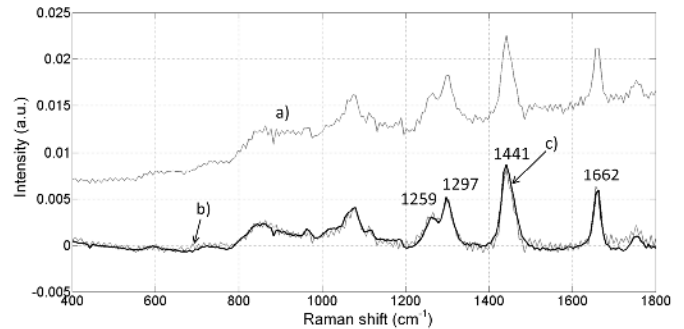
range. The pre-measured DCR level of each spectral point was subtracted from the Raman results presented below.

The Raman spectra of an olive oil sample derived by summing all the photons collected in bins 1 to 6 (raw spectrum) are shown in Fig. 16 a), and after a residual fluorescence correction in Fig. 16 b). This procedure corrects the baseline of the spectrum but does not improve the signal-to-noise ratio [14]. As shown in Fig. 16 b), the residual fluorescence level can be subtracted from the result but the spectrum contains still a sawtooth-like error caused by the mismatch of the stop channels of the TDC. This kind of clear variation between adjacent spectral point (adjacent channels of the TDC) can be explained by the mismatch of the layout of adjacent spectral points, because the power supplies of front-end electronics were routed between adjacent spectral point so that a supply voltage and ground were following each other as shown in Fig. 6. Note also the remarkable similarity of this variation to that shown in Fig. 12.
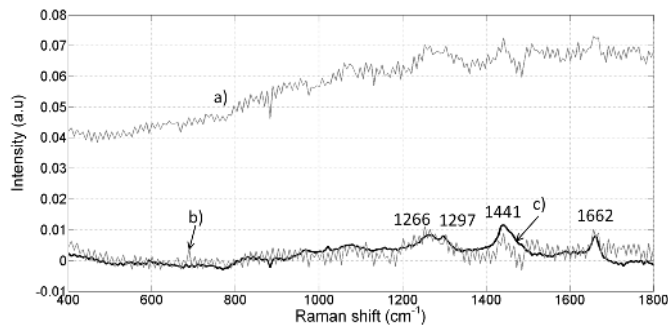
Fig. 17. Raman spectra of sesame seed oil a) without any correction, b) with residual fluorescence correction and c) with reference sample correction.

The Raman spectrum after correction by means of a reference sample measurement is shown in Fig. 16 c). This correction is based on a premeasured spectrum from a strongly fluorescent sample that reveals the inhomogeneities between the spectral points (since pure smooth fluorescence would imply no variation between adjacent spectral points). This reference measurement largely removes the sawtooth-like error from the spectrum, as is evident from Fig. 16 c).

The effectiveness of this reference sample correction can be seen more clearly in a more challenging Raman measurement using sesame seed oil as a sample. The recorded Raman spectra measured with the same procedure as in Fig. 16 are shown in Fig. 17 a), b) and c). With this sample the Raman spectrum can barely be distinguished from the raw data or even from the residual fluorescence-corrected spectrum, but it is, however, distinguishable after applying the reference sample correction, as is shown in Fig. 17 c). The Raman peaks of the olive and sesame seed oils can be found in their proper spectral regions, as can be seen by referring to [24].

The comparison of state-of-the-art SPAD Raman sensors (where Raman spectra have been measured) in Table I shows that the advantage of the design presented here over previous ones lies in its deep sub-ns time resolution in detecting photons and the low jitter of the whole system, which enables efficient fluorescence suppression even with strongly fluorescent samples having ns-scale fluorescence time constants, as was shown above. The FWHM value of the IRF of the whole Raman system is indicating that other jitter sources are in the same order as the FWHM of the laser pulse (150 ps). The best FWHM of IRF is achieved in [18], but the Raman setup used there is based on a single SPAD element and thus the value is not directly comparable to values achieved by using SPAD arrays.

V. DISCUSSION AND CONCLUSIONS

A CMOS receiver chip including 16 × 256 SPAD detectors and an on-chip 256-channel TDC having a time resolution of 52…104 ps was designed for single photon detection-based Raman spectroscopy. The SPAD line detector with a TDC array allows the measurement of the time-of-arrival of the detected photons simultaneously over the whole spectral range. The 3-bit TDC range is selected so that the measurement system can tolerate the timing skew of the timing signals over the whole spectral range. It is also shown that with reference measurement from a known sample with "smooth" fluorescence spectrum the effect of circuit non-idealities can in principle be mitigated. As a proof of this concept, the Raman spectrum of a highly fluorescent sesame seed oil sample (fluorescence lifetime of 1.9 ns and fluorescence-to-Raman signal ratio of 333) was successfully measured using a laser pulse with a width of ∼150 ps (FWHM) and the IRF (FWHM) of the whole setup of 225 ps. To the best of the knowledge of the authors, this is the most challenging sample whose Raman spectrum has been successfully recorded using a SPAD-based Raman sensors. The timing skew of the circuit was however slightly higher than expected and thus the highest resolution of ∼50ps couldn't be used to derive the time domain histograms of Raman photons at every spectral point efficiently. This, of course, limits the available performance improvement.

The conclusion arising from this work is that efficient fluorescence suppression, which is the key issue in Raman spectroscopy, can be achieved by using a high-precision TDC (<50ps) for the photon lifetime measurements and by setting the TDC range to exceed the expected timing skew of the chip, and preferably also ensuring that some portion of the fluorescence spectrum is also sampled. This enables one to effectively remove the fluorescence background (based on the fluorescence part of the spectrum) and to correct the spectral point-to-point inhomogeneity effects by applying post processing to the accurate time domain histogram data. The present design is by no means a perfect Raman receiver in this respect, but hopefully it paves the way to overcoming the effect of circuit non-idealities in SPAD-based time-gated Raman spectroscopy.

REFERENCES

[1] D. Tuschel, "Raman spectroscopy of oil shale," *Spectroscopy*, vol. 28, no. 3, pp. 20–27, Mar. 2013.
[2] T. Rojalin *et al.*, "Fluorescence-suppressed time-resolved Raman spectroscopy of pharmaceuticals using complementary metal-oxide semiconductor (CMOS) single-photon avalanche diode (SPAD) detector," *Anal. Bioanal. Chem.*, vol. 408, no. 3, pp. 761–774, Jan. 2016.
[3] Y.-S. Li and J. S. Church, "Raman spectroscopy in the analysis of food and pharmaceutical nanomaterials," *J. Food Drug Anal.*, vol. 22, no. 1, pp. 29–48, Mar. 2014.
[4] M. E. Darvin, I. Gersonde, M. Meinke, W. Sterry, and J. Lademann, "Non-invasive *in vivo* determination of the carotenoids beta-carotene and lycopene concentrations in the human skin using the Raman spectroscopic method," *J. Phys. D, Appl. Phys.*, vol. 38, no. 15, pp. 2696–2700, Aug. 2005.
[5] R. Hopkins, S. H. Pelfrey, and N. C. Shand, "Short-wave infrared excited spatially offset Raman spectroscopy (SORS) for through-barrier detection," *Analyst*, vol. 137, no. 19, pp. 4408–4410, Aug. 2012.
[6] O. S. Khalil, "Spectroscopic and clinical aspects of noninvasive glucose measurements," *Clin. Chem.*, vol. 45, no. 2, pp. 165–177, Feb. 1999.
[7] E. B. Hanlon *et al.*, "Prospects for *in vivo* Raman spectroscopy," *Phys. Med. Biol.*, vol. 45, no. 2, pp. R1–R59, Feb. 2000.
[8] M. L'Heureux, "Analysis of the state of the art: Raman spectroscopy," *Spectroscopy*, vol. 30, no. 6, pp. 44–49, Jun. 2015.

[9] R. P. Van Duyne, D. L. Jeanmaire, and D. F. Shriver, "Mode-locked laser Raman spectroscopy. New technique for the rejection of interfering background luminescence signals," *Anal. Chem.*, vol. 46, no. 2, pp. 213–222, Feb. 1974.

[10] P. Matousek *et al.*, "Fluorescence suppression in resonance Raman spectroscopy using a high-performance picosecond Kerr gate," *J. Raman Spectrosc.*, vol. 32, no. 12, pp. 983–988, Dec. 2001.

[11] D. V. Martyshkin, R. C. Ahuja, A. Kudriavtsev, and S. B. Mirov, "Effective suppression of fluorescence light in Raman measurements using ultrafast time gated charge coupled device camera," *Rev. Sci. Instrum.*, vol. 75, no. 3, pp. 630–635, Mar. 2004.

[12] E. V. Efremov, J. B. Buijs, C. Gooijer, and F. Ariese, "Fluorescence rejection in resonance Raman spectroscopy using a picosecond-gated intensified charge-coupled device camera," *Appl. Spectrosc.*, vol. 61, no. 6, pp. 571–578, Jun. 2007.

[13] I. Nissinen *et al.*, "A sub-ns time-gated CMOS single photon avalanche diode detector for Raman spectroscopy," in *Proc. IEEE ESSDERC*, Helsinki, Finland, Sep. 2011, pp. 375–378.

[14] J. Kostamovaara, J. Tenhunen, M. Kögler, I. Nissinen, J. Nissinen, and P. Keränen, "Fluorescence suppression in Raman spectroscopy using a time-gated CMOS SPAD," *Opt. Exp.*, vol. 21, no. 25, pp. 31632–31645, Dec. 2013.

[15] L. Pancheri and D. Stoppa, "A SPAD-based pixel linear array for high-speed time-gated fluorescence lifetime imaging," in *Proc. IEEE ESSCIRC*, Athens, Greece, Sep. 2009, pp. 428–431.

[16] I. Nissinen, A.-K. Länsman, J. Nissinen, J. Holma, and J. Kostamovaara, "2×(4×)128 time-gated CMOS single photon avalanche diode line detector with 100 ps resolution for Raman spectroscopy," in *Proc. IEEE ESSCIRC*, Bucharest, Romania, Sep. 2013, pp. 291–294.

[17] Y. Maruyama, J. Blacksberg, and E. Charbon, "A 1024×8, 700-ps time-gated SPAD line sensor for planetary surface exploration with laser Raman spectroscopy and LIBS," *IEEE J. Solid-State Circuits*, vol. 49, no. 1, pp. 179–189, Jan. 2014.

[18] Z. Li and M. J. Deen, "Towards a portable Raman spectrometer using a concave grating and a time-gated CMOS SPAD," *Opt. Exp.*, vol. 22, no. 15, pp. 18736–18747, Jul. 2014.

[19] A. Kufcsak *et al.*, "Time-resolved spectroscopy at 19,000 lines per second using a CMOS SPAD line array enables advanced biophotonics applications," *Opt. Exp.*, vol. 25, no. 10, pp. 11103–11123, May 2017.

[20] A. T. Erdogan, R. Walker, N. Finlayson, N. Krstajic, G. O. S. Williams, and R. K. Henderson, "A 16.5 Giga events/s 1024×8 SPAD line sensor with per-pixel zoomable 50ps-6.4ns/bin histogramming TDC," in *Proc. Symp. VLSI Circuits*, Kyoto, Japan, Jun. 2017, pp. 292–293.

[21] I. Nissinen, J. Nissinen, P. Keränen, A.-K. Länsman, J. Holma, and J. Kostamovaara, "A 2×(4)×128 Multitime-gated SPAD line detector for pulsed Raman spectroscopy," *IEEE Sensors J.*, vol. 15, no. 3, pp. 1358–1365, Mar. 2015.

[22] I. Nissinen, J. Nissinen, J. Holma, and J. Kostamovaara, "A 4×128 SPAD array with a 78-ps 512-channel TDC for time-gated pulsed Raman spectroscopy," *Anal. Integr. Circuits Signal Process.*, vol. 84, no. 3, pp. 353–362, Sep. 2015.

[23] I. Nissinen, J. Nissinen, P. Keränen, and J. Kostamovaara, "On the effects of the time gate position and width on the signal-to-noise ratio for detection of Raman spectrum in a time-gated CMOS single-photon avalanche diode based sensor," *Sens. Actuators B, Chem.*, vol. 241, pp. 1145–1152, Mar. 2017.

[24] Process Instruments Inc., *Raman Spectra of Cooking Oils*. Accessed: Mar. 14, 2017. [Online]. Available: https://vw-process-instruments-inc.storage.googleapis.com/uploads/2016/04/28/files/PI_Raman_Cooking_Oils.pdf

[25] A. Rochas *et al.*, "Single photon detector fabricated in a complementary metal–oxide–semiconductor high-voltage technology," *Rev. Sci. Instrum.*, vol. 74, no. 7, pp. 3263–3270, Jul. 2003.

[26] D. Mosconi, D. Stoppa, L. Pancheri, L. Gonzo, and A. Simoni, "CMOS single-photon avalanche diode array for time-resolved fluorescence detection," in *Proc. IEEE ESSCIRC*, Montreux, Switzerland, Sep. 2006, pp. 564–567.

[27] A. Mäntyniemi, T. Rahkonen, and J. Kostamovaara, "An integrated 9-channel time digitizer with 30 ps resolution," in *Proc. ISSCC*, San Francisco, CA, USA, Feb. 2002, pp. 266–465.

[28] T. E. Rahkonen and J. T. Kostamovaara, "The use of stabilized CMOS delay lines for the digitization of short time intervals," *IEEE J. Solid-State Circuits*, vol. 28, no. 8, pp. 887–894, Aug. 1993.

Ilkka Nissinen (M'09) was born in Oulu, Finland, in 1976. He received the M.Sc. (Eng.) and Dr.Tech. degrees in electrical engineering from the University of Oulu, Finland, in 2002 and 2011, respectively. He has been a Post-Doctoral Researcher with the Circuits and Systems Research Unit, University of Oulu, since 2011. His research interests include the design of time interval measurement architectures for the integrated sensors of pulsed time-of-flight laser technologies.

Jan Nissinen was born in Oulu, Finland, in 1976. He received the M.Sc. (Eng.) and Dr.Tech. degrees in electrical engineering from the University of Oulu, Finland, in 2002 and 2011, respectively. He has been a Post-Doctoral Researcher with the Circuits and Systems Research Unit, University of Oulu, since 2011. His research interests include the design of analog and mixed-signal integrated circuits for pulsed time-of-flight laser rangefinders and for pulsed Raman spectroscopy.

Pekka Keränen was born in Tampere, Finland, in 1985. He received the M.Sc. (Tech.) and Dr.Tech. degrees in electrical engineering from the University of Oulu, Finland, in 2010 and 2016, respectively. He is currently a Post-Doctoral Researcher with the Circuits and Systems Research Unit, University of Oulu. His main research interest is the design of time interval measurement circuits.

David Stoppa (M'97–SM'12) received the Laurea degree in electronics engineering from the Politecnico di Milano, Italy, in 1998, and the Ph.D. degree in microelectronics from the University of Trento, Italy, in 2002. From 2002 to 2012, he has been teaching courses on analogue electronics and microelectronics with the Telecommunications Engineering Faculty, University of Trento. He was a Group Leader of the Smart Optical Sensors and Interfaces Group from 2010 to 2013. From 2014 to 2017, he was the Head of the Integrated Radiation and Image Sensors Research Unit, Fondazione Bruno Kessler, where he has been a Research Scientist since 2002. In 2017, he joined AMS, where he is currently in charge of the research and development of next generation range-sensors. He has authored or co-authored over 120 papers in international journals and presentations at international conferences, and holds several patents in the field of image sensors. His research interests are mainly in the field of CMOS integrated circuits design, image sensors, and biosensors. He was a recipient of the 2006 European Solid-State Circuits Conference Best Paper Award. Since 2011, he has been serving as a Program Committee Member of the International Solid-State Circuits Conference and the SPIE Videometrics, Range Imaging, and Applications Conference, and was a Technical Committee Member of the International Image Sensors Workshop, in 2009, 2013, 2015, and 2017, respectively. He was a Guest Editor for the IEEE JOURNAL OF SOLID-STATE CIRCUITS special issues on ISSCC'14 in 2015. He has been serving as an Associate Editor since 2017.

Juha Kostamovaara received the Ph.D. degree in electrical engineering from the University of Oulu, Finland, in 1987. He currently holds a full professor position in electronics with the University of Oulu. His main research interests include the development of pulsed time-of-flight devices, circuits, and systems for electronic and optoelectronic measurements.