NOVEL DETECTION METHODS FOR BIOPHOTONICS SENSING

BY

PIOTR MACIEJ WARGOCKI

A THESIS SUBMITTED TO MACQUARIE UNIVERSITY

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS AND ASTRONOMY

Nov 2018

MACQUARIE University
SYDNEY·AUSTRALIA

Statement of Originality

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

_____ Date: _____

Piotr M. Wargocki

Abstract

The appearance or increase of the fluorescent signal in biological media (whether it is natural autofluorescence or fluorescence of artificially introduced fluorophores) can implicate reaction to specific conditions or diseases. In the first part of this thesis, we explored the possibilities of using smartphone as a point-of-care readout device that could replace specialised laboratory equipment like a spectrophotometer or colorimeter at a fraction of the cost. We present two methods that, using the minimum of additions to the phone itself, are sensitive enough to be used in medical diagnostics of clinically relevant conditions including arthritis, cystic fibrosis and acute pancreatitis. In the second part, we developed the highest resolution wide field Fluorescence Lifetime Imaging (FLIM) system based on a Single Photon Avalanche Diode sensor array. The device was capable of detection of sub-nanosecond lifetimes over an area of about 20cm^2, and was able to measure lifetimes of fluorescent signals lower than one photon per pixel per excitation pulse. We proved that the system can reliably distinguish between biologically relevant concentrations of free and bound forms of Nicotinamide Adenine Dinucleotide (NADH). An increment of bound NADH in cells, which is connected to increased cell metabolism, is a proven cancer marker for most cancer types. We characterised the system performance and provided two methods of data analysis appropriate for measurements in a biological tissues or surface tumours. We project that our system is suitable for development into a real-time fluorescence lifetime camera, able to operate at up to 20 Hz with presented performance level.

Acknowledgement

What a journey that was!

I would like to thank my supervisors: Ewa Goldys for initiating it and David Spence for ongoing help to finish it. It was a pleasure to work with both of you!

Gentlemen, my wife – Zofia. Thank you for letting me doing this and supporting me the whole time. I love you.

To my parents – no words can express the amount of things I would like to thank you for.

Thank you Ondra for all your help in the lab, you saved my emotional stability many times!

Much of my research would have been impossible without my collaborators – thank you for all the opportunities, and never-ending support.

Big thanks to all my friends and colleagues from Macquarie University – without you this time would me much less fun!

Finally, I would like to thank Centre for Nanoscale BioPhotonics for creating such an amazing environment to do science! I was proud to wear a CNBP badge attached to a CNBP t-shirt.

Contents

1. Theoretical introduction

1.1. Fluorescence – Physical principles

Fluorescence is the result of a three-stage process that occurs in fluorophores: Excitation, excited-state lifetime and fluorescence emission. Excitation happens when photon of energy hv_{ex}, from an external source such as laser or light-emitting diode (LED), hits the fluorophore. The fluorophore absorbs this photon achieving an excited electronic singlet state ${S_1}'$

$$S_0 + hv_{ex} \rightarrow S_1', \qquad (1)$$

Immediately after excitation, the fluorophore undergoes conformational changes, which typically last nanoseconds and has two important consequences. One part of the energy is dissipated, yielding a relaxed singlet excited state (S_1), from which fluorescence originates. A second part is emitted as non-fluorescent energy, such as heat (vibration), collisional quenching or fluorescence resonance energy transfer (FRET). The ratio between emitted and absorbed number of photons by a fluorophore is called the quantum yield and will be discussed later. The subsequent and final stage is fluorescence emission. A photon of energy hv_{em} is emitted, returning the fluorophore to its ground state S_0.

$$S_1 \rightarrow S_0 + hv_{em} + heat \qquad (2)$$

These processes are illustrated on a Jablonski diagram (Figure 1). Horizontal lines represent molecule energy levels; arrowed lines are energy transfers. For more detailed description of phenomenon of fluorescence, please refer to (1).

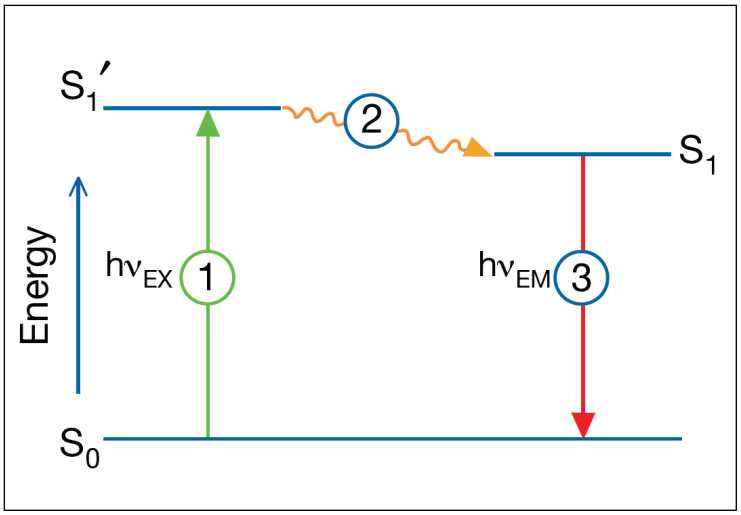


Figure 1. Jablonski diagram (2).

Because of energy dissipation described in equation 2, the energy of emitted photon is lower, therefore its wavelength is longer than the excitation photon hv_{ex}. This difference between maxima of wavelengths of excitation and emission spectrum of the fluorophore is called the Stokes shift (Figure 2). It owes its name to Sir G.G. Stokes who described it in 1852 (3). The Stokes shift is the fundamental basis of fluorescence techniques because it separates emission photons from excitation photons, effecting easier detection.

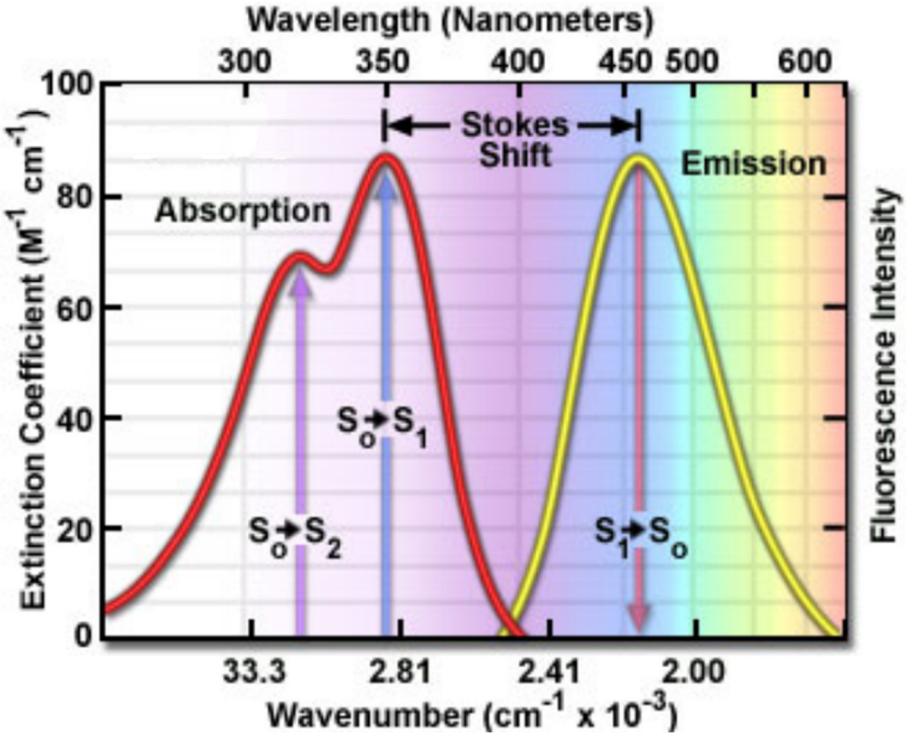


Figure 2. Stokes shift, presented using the example of quinine (1).

1.2. Characteristics of fluorescence emission

1.2.1. Excitation - Emission spectra

Most of the fluorophores used in science have emission spectra in the visible wavelength range or neighbouring ultraviolet or infrared spectra. Excitation is synonymous to absorption and therefore it can be measured using absorption techniques. Excitation and emission spectra typically have a simple peaked distribution (Figure 2). The emission intensity is proportional to the magnitude of the fluorescence excitation spectrum at the excitation wavelength; however the emission spectrum is independent of excitation intensity (Figure 3). This second phenomenon was explained by Michael Kasha in 1950 and is known as Kasha's rule (upon excitation into higher electronic and vibrational levels, the excess energy is quickly dissipated, leaving the fluorophore in the lowest, common to all molecules, vibrational level of S_1) (4). In 1926, Vavilov reported that that quantum yields are generally independent of the excitation wavelength (5). For many fluorophores, the emission spectrum looks like a mirrored copy of the excitation spectrum. This phenomenon can be explained looking at horizontal lines at Figure 3. In general, the emission intensity is proportional to excitation intensity, at the excitation wavelength. This is known as mirror image rule (1). Quinine, whose spectrum is presented in Figure 2, is one of the few exceptions from this rule.

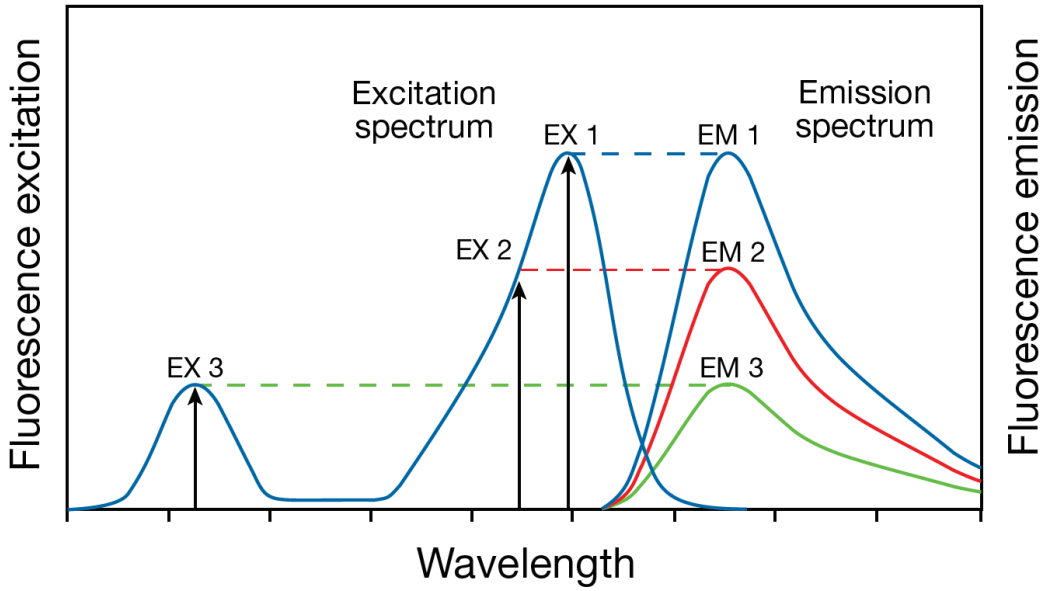


Figure 3. Excitation of a fluorophore at three different wavelengths (EX 1, EX 2, EX 3) does not change the emission profile but does produce variations in fluorescence emission intensity (EM 1, EM 2, EM 3) that correspond to the amplitude of the excitation spectrum (2).

1.2.2. Quantum yield

The quantum yield is a measure of the emission efficiency of the fluorophore. It is the ratio of emitted and absorbed number of photons. The maximum possible efficiency of 1, which means the brightest emission, can be decreased by non-radiative decay to S_0. The fraction of fluorophores that decay through emission, the quantum yield Q, is given by:

$$Q = \frac{\Gamma}{\Gamma + k_{nr}} \tag{3}$$

in which Γ is the radiative decay rate, and k_{nr} is the decay rate due to ever-present non-radiative processes, leading to a quantum yield Q always smaller than unity. Maximum quantum yield is achieved when non-radiative decay is substantially smaller than radiative decay $k_{nr} \ll \Gamma$.

1.2.3. Fluorescence lifetime

Fluorescence lifetime quantifies the time a fluorophore stays in an excited state before emitting photons. Referring to the Jablonski diagram (Figure 1), it is the average time before processes undergoing between excited singlet state $S_1{}'$ and relaxed state S_1 finish. Generally, fluorescence lifetime has an order of magnitude of $10^{-9}s$ (nanoseconds). Lifetime is given by:

$$\tau = \frac{1}{\Gamma + k_{nr}} \quad (4)$$

Because fluorescence emission is a random process, the lifetime is an average value of the time spent in the excited state. For a single exponential decay 63% of the molecules decay faster than $t = \tau$ and 37% decay at $t > \tau$ (1). Single exponential decay of fluorescence intensity as a function $\mathrm{I(t)}$, of time can be represented as an exponential function:

$$I(t) = I_0 e^{-t/\tau} \quad (5)$$

where $I(t)$ is a fluorescence intensity at time t, I_0 is initial fluorescence intensity and τ is fluorescence lifetime.

1.2.4. Fluorescence Quenching

Quenching is a process which results in a decrease in intensity of fluorescence emission of the fluorophore. There are many reasons and reactions that can result in a quenching effect. The most common are:

- Chemical quenching - the presence of iodide (I^-)(6), molecular oxygen (7), or other chemicals. These substances that can occur in the solvent, can create chemical bonds influencing excited state reactions, and altering emission processes. Temperature and pH might affect these reactions as well.
- Energy transfer – degradation of the excited state due to transfer of some portion of the energy to the nearby fluorescent molecule, known as Förster Resonance Energy Transfer – FRET (8)
- Collisional quenching - Some molecules returning to the ground state, encounter molecules of the quencher and are physically stopped from fluorescing. This phenomenon is described by the Stern-Volmer dependence:

$$\frac{F_0}{F} = 1 + K[Q] = 1 + k_q \tau_0 [Q] \quad (6)$$

In this expression F_0 is the original intensity of the fluorophore, F is the intensity after quenching, K is the Stern-Volmer quenching constant, k_q is the biomolecular quenching constant, τ_0 is the unquenched lifetime, and Q is the quencher concentration. The Stern-Volmer quenching constant K indicates the sensitivity of the fluorophore to a quencher.

- Static quenching – the part of the quenching that occurs before excitation of the molecules, i.e., in the ground state. Chemical reactions between fluorophore and quencher, happening before excitation, can change the excitation wavelength or even change the fluorophore into a non-fluorescent molecule, stopping the fluorophore from emission in certain conditions.
- Concentration quenching - in order to enhance the luminescence efficiency, it seems that the fluorophore concentration should be as high as possible. However, it is found (9) that the luminescence efficiency decreases if a specific value known as critical concentration is exceeded. With the increase of dopant concentration, the ion centres are sufficiently close together to transfer the excitation energy from one to another, causing an imperfection which may act as an energy sink.

Quenching phenomenon can be used to measure physical and chemical reactions. It is the basis of Förster resonance energy transfer (FRET) (8) which can be used to determine if two fluorophores are within a certain distance of each other.

1.2.5. Fluorescence Photobleaching

Photostability is one of the most important properties of a sample, describing how long the fluorophore is able to fluoresce. Photobleaching, an irreversible destruction of the fluorophore, terminates this ability. Almost all fluorophores photobleach under illumination, especially in fluorescence microscopy where the light intensities are high (Figure 4). Photobleaching originates from the triplet excited state, which is created from the singlet state via an excited-state process called intersystem crossing (10). Attempts to recover triplet states as well as charge-separated states through electron-transfer reactions (11), or chemical extension of the photostability (12) have been undertaken; however these are often specific to certain fluorophores and hard to achieve in in-vivo measurements.

The most effective method to minimise effects of photobleaching is to maximise sensitivity of the detection device, so the power of excitation can be reduced. Finding more photostable alternatives is a constant goal.

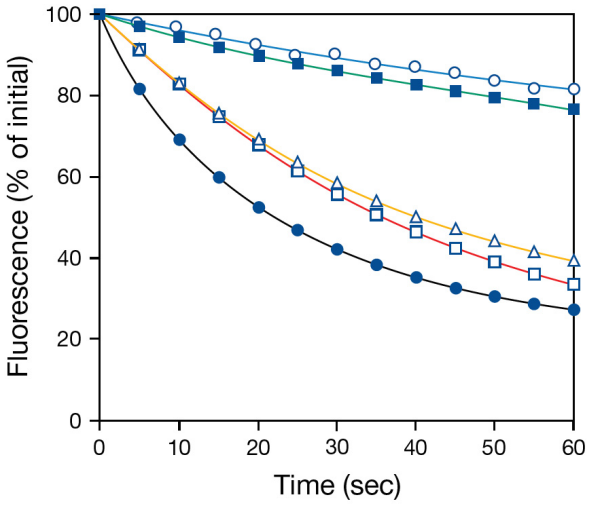


Figure 4. Comparison of photostability of green-fluorescent antibody conjugates, from the most photostable, respectively: Alexa Fluor® 488, Oregon Green® 514, BODIPY® FL, Oregon Green® 488, fluorescein (2).

1.3. Fluorophores

Currently we are aware of thousands of fluorophores. They can be broadly divided into two main classes—intrinsic and extrinsic. Intrinsic are the ones that occur naturally in systems we are studying, like aromatic amino acids, NADH or flavins. Extrinsic fluorophores are those added to the sample to make it fluorescent. This procedure is called labelling.

Currently, one of the most commonly used extrinsic fluorophores is fluorescein, or its amine reactive isothiocyanate derivative - fluorescein isothiocyanate (FITC). In its original form it is often used as a laser dye gain medium. As FITC, it is popular for labelling of antibodies. A wide selection of fluorescein-labelled antibodies is commercially available, and used in fluorescence microscopy and in immunoassays. In water, fluorescein has an absorption maximum at 494 nm and emission maximum of 512 nm.

Another commonly used group of fluorophores are rhodamines, for example Rhodamine B and Rhodamine 6G. Factors such as the large number of commercially available kits for staining or tracing, solubility in water, ethanol and methanol, fairly high quantum yield and wide spectra that are different between derivatives, make rhodamines extensively used in biotechnology.

Many other fluorophores are widely commercially available. They differ in excitation and emission spectra, lifetimes, photostability, solubility in different solvents, signal strength (quantum yield), reactivity etc. Some of them, like Alexa Fluor 488, are designed to be an alternative to another fluorophores (in this case – fluorescein), providing better performance in some areas like photostability or quantum yield. Other popular fluorophores groups include: Alexa fluor, Cy, Oregon, Coumarin, BODIPY, and Hoechst.

1.4. Instrumentation for fluorescence measurements

In this part we will discuss techniques and instruments used for different types of fluorescence measurements. We will also have a closer look at their structure and will describe their principles of operation. This is an important step since in the next chapter we will be presenting our own systems to read some of these signals.

1.4.1. Spectrophotometer/Spectrofluorometer

The most specific of the fluorophore attributes is its spectrum. The Excitation/Emission spectrum graph, like the one presented in Figure 3, informs us how to bring fluorescence to life and observe its consequences. The excitation spectrum, which typically is on the left side of the graph, with shorter wavelengths and higher energy, equivalent to absorption, can be measured using a spectrophotometer. An excitation spectrum is measured by scanning across all the absorption wavelengths, recording emission at a fixed wavelength. The emission spectrum, distanced from the excitation spectrum by the Stokes shift to the right, with longer wavelengths and lower energy, can be measured using a spectrofluorometer. An emission spectrum is a wavelength distribution measured with a specified, single excitation wavelength. Nowadays, these two systems are often combined in one machine that allows us to measure one or the other.

A typical spectrophotometer (Figure 5) is constructed of a light source, excitation and emission monochromators for wavelength selection, a sample chamber designed to optimise energy transfers, a readout system, typically preceded with additional filter and polarizers - for data collection and finally – a digital operating system that will control all the elements, but also gather and present results in a typical graph

representation of signal as a function of wavelength (1). We will take some time to have a closer look at each of these elements.

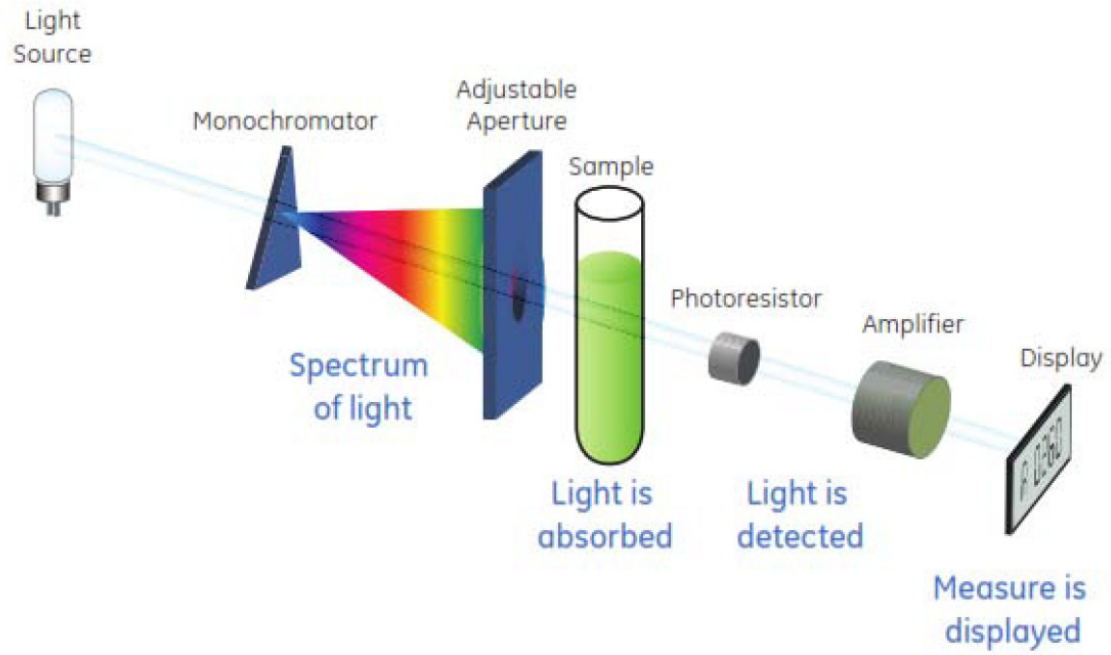


Figure 5. Schematic diagram of spectrofluorometer (13).

1.4.1.1. Light source

The most common light source used in steady-state spectrophotometers are continuous xenon arc lamps. They provide high and relatively even power (Figure 6a) in the range between 250 nm - 700 nm. Because of this high power of the light, high pressure (about 10 atmospheres) inside the lamp and high input power to it, they need to be kept in a housing, separating the lamp from the rest of the system, absorbing generated heat, and providing additional safety levels. Readouts from systems using Xe lamps are generally corrected for the imperfect spectral distribution across wavelengths (1). Pulsed Xe lamps suffer from even greater spectral variation (Figure 6a), however, they produce less heat, higher power in the UV range and consume less power (1).

The newest members of the group of sources, and the ones gaining rising popularity are light emitting diodes (LED) (14). LEDs are small and consume little power. They emit relatively low power of light, but also produce almost no heat, therefore can be placed close to the sample. Different LEDs emit different light spectra, but because of their size, it is easy to produce an array of LEDs that will cover the whole required

band (Figure 6b). There are ongoing efforts to produce white LEDs with the whole emission spectrum (15), with first successes currently used mainly in flashlights.

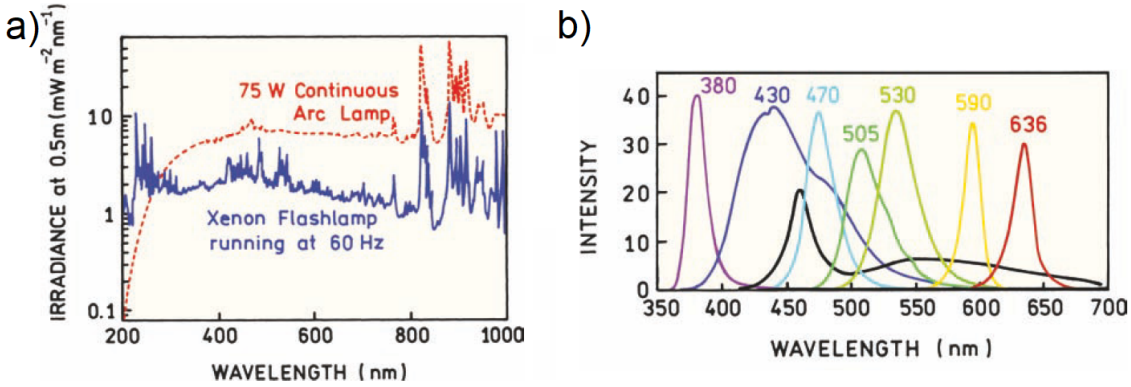


Figure 6. Light sources used for excitation in spectrofluorometers. a) comparison of emission of continuous and pulsed Xenon lamp; b) spectra of different LEDs, with a white LED in black (1).

1.4.1.2. Monochromators and filters

The monochromator in the spectrofluorometric system is the part that further selects a specific wavelength, from the excitation or emission light. This is done by dispersion of polychromatic or white light. Dispersion can be accomplished using a prism or diffraction grating, where typically the second one is chosen in spectrofluorometry. A monochromator for fluorescence spectroscopy should have high dispersion efficiency, low stray light levels (leaks of light with unwanted wavelength) and high light transmission to be able to detect low light signals across the whole spectrum. Resolution of the monochromator is a less important factor.

In addition to dispersion gratings, filters are used for correcting for imperfections of the grating, or even replacing it completely for some specific uses. Nowadays, among the huge variety of commercially available filters, almost any desired specification can be found.

1.4.1.3. Sample chamber

A typical sample chamber in a spectrofluorometer is constructed to hit the centre of the cuvette with excitation light and collect the emission light at an angle of 90°. This setup minimises leaking of the excitation light and decreases the inner filtering effects due to high optical densities or to sample turbidity.

Cuvettes are typically cubic in shape and made of a material that that is transparent to radiation in the spectral region of interest. The commonly used materials for different wavelength regions are: quartz or fused silica for UV, silicate glass for above 350 nm to 2000 nm, plastic for the visible region, or polished NaCl or AgCl for wavelengths longer than 2 µm.

1.4.1.4. Photomultiplier tubes

Photomultiplier tubes (PMTs) are able to detect the fluorescence signals at a wide wavelength range (200 nm to 900 nm) with high sensitivity. This reason, amongst others, is why they have become the most popular detector for modern fluorometers. The PMT is capable of detecting an individual photon, generating a burst of millions of electrons that can be detected as a pulse. PMTs can operate in two modes: photon counting or as a current source. In the first mode each arriving photon results in a discrete count; in the current source mode, the number of photons is integrated, generating a proportionate current. PMTs can be distinguished by several properties such as type of dynode chain, spectral or temporal response, and size.

1.4.1.5. Data acquisition and analysis/corrections

The final and crucial element to the spectrofluorometric device is its operating system. Typically, a computer application, connected to the firmware inside the device, serves a couple of important purposes throughout the whole operation process. At the beginning, it controls all the elements, synchronising their work, and setting them in correct positions. At the end, it collects data gathered by the readout system and processes it into a data output of chosen type. Once the data are collected and processed, one more important step is performed – data correction.

Data correction is a process where raw data from the readout instrument are multiplied by a function (or series of functions) to cancel all known readout errors, which might be introduced by any of the elements, or by known phenomena occurring during the readout process (Figure 7). Errors of the elements can be recognized and described in two ways: either measuring the difference between ideal expected output and real output generated by the element, or by temporarily replacing one element in the system with another with known, close to ideal character, and comparing outputs of the whole system. An example of such an element error and its

correction is an excitation lamp with imperfection in power distribution (Figure 6a). An example of a phenomenon that might occur during spectrophotometric readout is the Raman scattering.

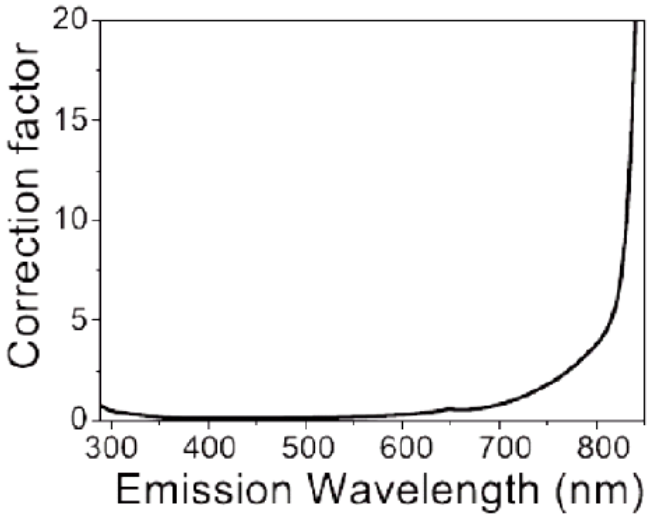


Figure 7. Example combined correction function for Fluorolog Tau 3 system (figure from the manual of horiba Fluorolog Tau 3 manual).

1.4.2. Fluorescence Lifetime Imaging Microscopy (FLIM)

One of the most important characteristics of a fluorophore, along with its spectra, is its fluorescence lifetime. By detecting different lifetimes, it is possible to distinguish spectrally-indistinguishable fluorophores. Fluorescence lifetime is independent of dye concentration, photobleaching, light scattering and excitation light intensity. At the same time fluorescence lifetime *is* dependent on things we might like to measure, such as ion concentration, oxygen concentration, molecular binding and interactions. Fluorescence lifetime imaging microscopy, as an imaging technique, determines the average fluorescence exponential decay time for many pixels and produces a map of fluorescence lifetime data. Therefore, FLIM allows us to perform accurate measurements of ion concentration, FRET analysis, or molecular changes in some reagents. FLIM is a very common technique to detect autofluorescence signals, also from living cells. Typical applications of FLIM in biology range from cells to complex tissues utilising both exogenous fluorescent labels like green proteins and endogenous autofluorescent molecules.

To acquire a fluorescence lifetime image, fluorescence lifetime detection has to be executed for many spatial locations in the sample. For a single-pixel detector typically a scanning stage moves the sample or the beam to build up an image, often using a modified microscope. For detector based on many pixels, one may be able to acquire lifetimes simultaneously from many points of the sample.

There are two main methods of lifetime detection: frequency-domain and time-domain. Systems based on frequency-domain measurements have the sample excited with (typically) sine-wave modulated light intensity. The modulation frequency is often around 100 MHz, so the period is comparable to the decay time τ. The excited sample emits with a forced response at the frequency of modulation, but fluorophores lifetime delays the emission in time relative to the excitation, shown as the shift to the right in Figure 8a. The induced phase shift (Δϕ), is used to calculate the decay time.

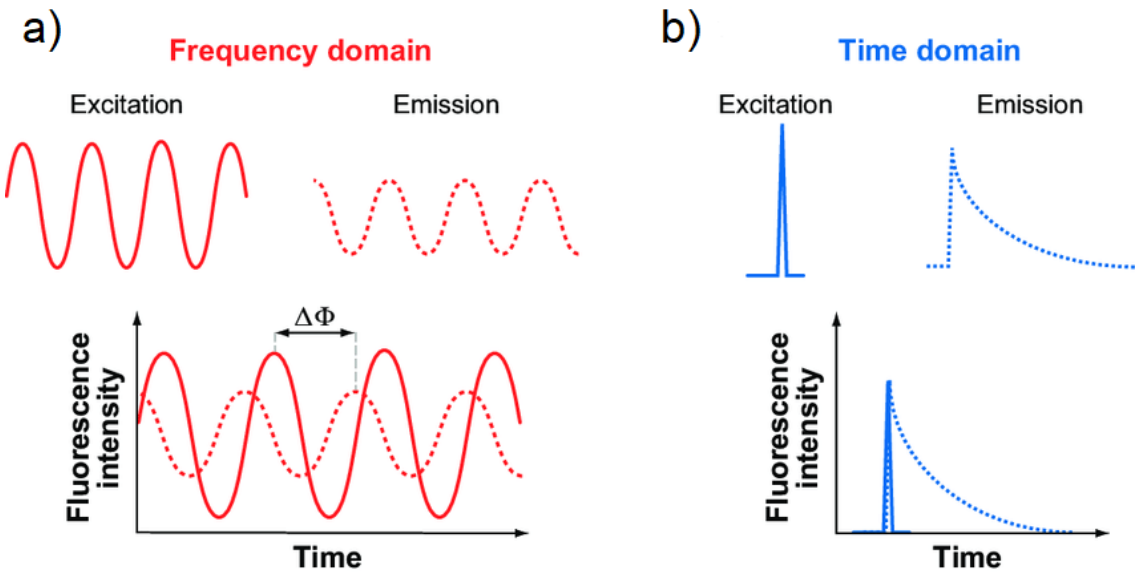


Figure 8. Principle of operation of FLIM system based on a) Frequency-domain, b) time-domain technique (16)

In the alternative time-resolved method, the sample is excited with a pulsed laser with pulse duration as short as possible, and preferably much shorter than the lifetime. For bright fluorescence, the time-dependent fluorescence intensity can be measured following the excitation pulse, and the decay time τ is calculated from the plot of intensity as a function of time after the pulse (Figure 8b). For weaker fluorescence, the most commonly used time-domain FLIM technique uses time-correlated single photon counting (TCSPC). TCSPC is based on repeatedly measuring the time between

excitation pulse of a laser and the arrival of a fluorophore-emitted photon at the detector. Both phenomena are precisely timed to calculate the delay, and the statistics of this delay reveal the lifetime.

The other popular method of time-domain FLIM is a time-gated measurement. In this approach, a pulsed laser excites the sample whose emitted fluorescence decreases in an exponential manner. After each excitation, a readout detection gate of a specified length opens at a specified time after the laser pulse. The measurement is repeated many times with the gate shifted to different delays (Figure 9). Each gate shift generates one data point for each sensor, generating a record of the decaying signal intensity for the whole array. The exponential function fitted to these points allows for the fluorophore lifetime calculation.

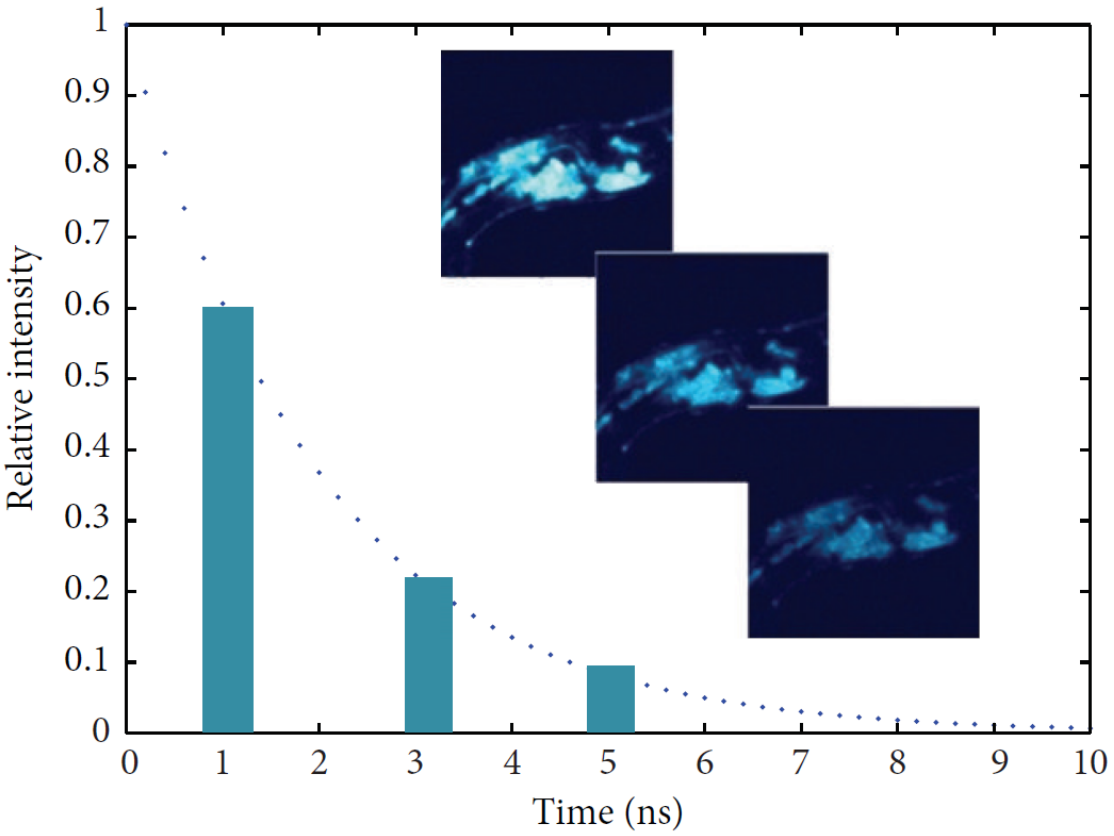


Figure 9. Scheme of operation of time gated fluorescence measurement. A number of measurements with gates at different delays generate data to which an exponential function can be fit, to calculate the lifetime Each image presents sample with decreasing intensity presented on consecutive bars. Dotted line shows the decay curve of the emission. (17)

There are many other popular measurement methods based on fluorescence like Fluorescence Resonance Energy Transfer (FRET) (18) microscopy or Fluorescence

Correlation Spectroscopy (FCS) (19). There are also different approaches to well established methods, like multiphoton excitation (20), or spectroscopy based on photobleaching (21). All of them, however, mostly share the same operation principles and elements discussed here.

1.5. Colours

Different types of emission measured in following work, will be indirectly represented as colours. Colours, as we can see them are electromagnetic waves of a length between about 400 nm and 700 nm. All the colours visible to human eye are within this spectrum. Even though dividing and naming colours is a cultural and contractual process, most people agree on common list that identifies six main bands: violet, blue, green, yellow, orange and red. Sources of light can be monochromatic, a polychromatic band, or a more complex spectrum. Mixtures of different wavelengths create different variations of colours. A typical monochromatic light source is a laser. The best known source of polychromatic light, which consist of all the colours with its resulting colour is close to white, is the sun.

The colour of an object interpreted by a human is a very complicated function of its illumination, structure and material properties, the environment, and the human eye and brain. The reflected/emitted light it is received at the retina by three types of cone cells with sensitivities as shown on Figure 10, and simplified to three components. The short-wavelength cone cells read mostly the blue range of colours; the medium-wavelength cone cells are receiving and interpreting the green-yellow spectrum; the long-wavelength cone cells, largely overlapping medium ones, also add the red part of the visible spectrum to the reception range.

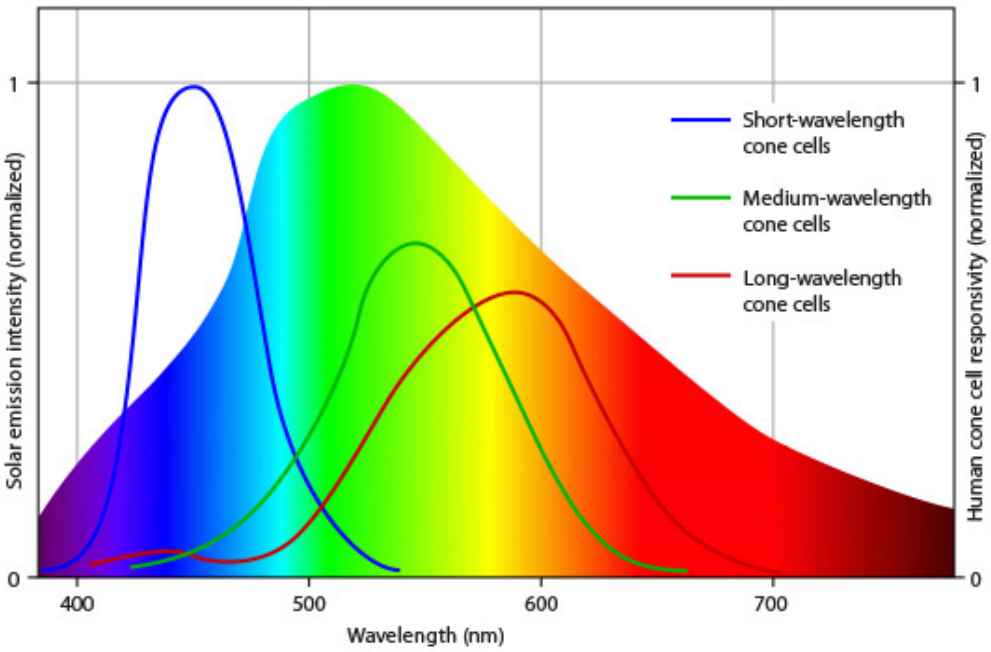


Figure 10. Spectral ranges of human cone cells responsible for interpreting colours, compared to the normalised spectrum of sunlight, shown across the visible spectrum of electromagnetic radiation.

It is estimated that a normal person can distinguish about 10 million colours (22), but this is a personal characteristic. Colour-blind people, who have reduced structure of cones, can not only see significantly fewer colours but also interpret them differently, while people with some mutations can distinguish even up to 100 million different colours.

People have long tried to reproduce colours using available knowledge and available dyes. Originally, the three colours that were chosen to be 'primary' colours, from which all other colours can be mixed, were red, yellow and blue. In art, these colours are still very often used as primary. Subsequently, in the 19th century, the yellow colour from the trinity was exchanged with green, later anchored with connection to human cone cells. At about the same time, chemical companies, trying to develop the most efficient process of printing colour photographs, established cyan, magenta and yellow as their primaries. Different colour systems so began to be used to represent the whole spectrum.

Digitalisation of colours is a whole new area of expertise. Photographers are using specific compressed and uncompressed file formats, for easier, or more accurate and editable version of their images. Colour television has had many image standards – first analogue (PAL, NTSC), then digital like DVBT. Digital displays in computers or

smartphones with different production technologies are using different colour systems. Printers, operated by digital systems, needed a proper way of communication with them. Colours needed standardisation.

Many different digital colour representations were introduced, representing different approaches and capabilities, though mostly based on the same principle of using functions with three values to represent the largest colour gamut. The representation that has gained the most popularity was presented in 1996 by Microsoft and HP: the sRGB (standard Red Green Blue) colour space. A standardised version of previously introduced RGB, sRGB was subsequently acknowledged by the International Electrotechnical Commission as IEC 61966-2-1:1999 (23). Widely used in most digital display systems, it started to also be used in capturing systems, including scientific systems. Capturing colours, because of previously described external factors that influence targets, is more complicated than displaying them. Furthermore, for many reasons, not all systems use sRGB as their base colour space. Standardisation in capturing, presenting and recalculating colours between different colour spaces became a necessity. To make possible comparison of images or colours captured under different lighting, standard illuminants (24) were introduced. A standard illuminant is a theoretical source of visible light with a specific profile (its spectral power distribution). They will be described in detail in the following chapters.

1.6. Colorimetry

A colorimeter is a device that in a non-direct way measures a colour of a liquid sample, naturally or artificially stained. It is based on the Beer-Lambert law that relates the attenuation of the light of specific wavelength dependant on the medium that it is going through. The general Beer-Lambert law determines absorbance (A) as:

$$A = a(\lambda) \times b \times c, \tag{7}$$

where $a(\lambda)$ is a wavelength-dependent absorptivity coefficient, b is the path length and c is the analyte concentration. Some colorimeters use the transmittance coefficient instead of the absorbance. Both, according to the Beer-Lambert law are directly related to the sample concentration and its colour. In a colorimeter, a beam of light with a specific wavelength controlled by filters, is passed through a sample via a

series of lenses, which navigate the coloured light to the measuring device. This analyses the colour compared to an existing standard.

Comparing a colorimeter to a spectrofluorometer, it is fair to say that the former is a limited version of the latter. While spectrofluorometry measures the transmittance and reflectance as a function of wavelength, the colorimetry measures the absorbance of specific colours. Colorimeters also operate only in the visible spectrum, while spectrofluorometers are typically extended to UV and IR as well. An advantage of a colorimeter is that, since the device is less complex, it is also significantly cheaper.

1.7. Smartphones in science

The first part of this thesis will investigate the utilisation of smartphones as an alternative method of colorimetric and spectrofluorometric measurement.

Since about 2007, when first mass production smartphones were released and gained popularity, the general idea of what the phone should be has drastically changed. Not only have they become small computers in our pockets, they were equipped with technology that regular computers did not have – microphones and cameras. In time, the number of sensors and their quality continued to grow. Smartphones gained accelerometers, gyroscopes, GPS modules for accurate positioning, light and proximity sensors, additional microphones to perform noise cancellation, and more. The quality of the photo cameras also got better each year, with the release of almost every new flagship model. Optics in the camera lenses became closer to the professional ones, resolution increased drastically, and the quality of the photosensitive matrices and their capabilities of low light level imaging increased. Over more than 10 years of development, not only the hardware, but also the software operating all these sensors became significantly better, more reliable and more user friendly. This opportunity could not have been missed by science.

Figure 11. A dual-mode mobile phone microscope using the on-board camera flash and ambient light (25).

Smartphones have found their way into almost all branches of science that could have any use of what they offer. The ability to localise the phone and instantly upload data to the cloud, including all the potentially important additional information about it like location or attributes of measurement, expands possibilities greatly: beginning with social sciences where analysis of human behaviour and customs can be collected (like social media analysis or even just questionnaire distribution), through using the computational power of the device by scientists (e.g., SETI@HOME, BOINC) and citizen science where regular people help to collect data for scientists, finishing with intentional use of smartphone sensors as readout device by scientists. This last group uses almost all available sensors, however, without any doubts, is dominated by photonics uses of the smartphones camera. With the add-on of proper peripherals and using the smartphone as a readout sensor, almost all of the previously mentioned fluorescence and colorimetric techniques can be implemented, and at significantly lower cost, often miniaturised and made portable. The moment we are at is still just the beginning of this process, but even now, systems based on smartphones can provide comparable results to professional solutions. Environmental and biological sciences are already taking advantage of Point-of-Care Testing (POCT) smartphone solutions, using smartphones as portable microscopes (Figure 11) or spectroscopes. Using images to quantify or qualify shapes or colours allows us to measure microorganisms, proteins, bacteria and molecules, to assess composition of targets or detect medical conditions. The list of potential applications is almost as long as the combined list of all previously described systems and is still growing.

In our work we will present proofs of concept of such devices, presenting setups to use smartphone-based systems to work as a spectrofluorometer or as a colorimeter.

1.8. Single Photon Avalanche Diode (SPAD)

The second part of this thesis will use an array of single photon avalanche diodes for time-gated FLIM, which we briefly introduce here.

For single photon detection, PMTs were for a long time the preferred device used in spectrofluorometers and other light sensitive measurement devices. These, however, have historically been bulky and very power hungry.

Solid-state diodes, when operating in the avalanche mode are known as avalanche photodiodes (APDs). When APDs operate in a Geiger mode (i.e., well above the breakdown voltage), they are called Single Photon Avalanche Diodes (SPADs). As semiconductor processes improved, in 1960s Cova and others (26) began experimenting with SPADs.

Until recently, manufacturing of SPAD detectors was a very specialised and fragile process, which limited their availability and use. SPADs manufactured in the standard CMOS process is a recent innovation (27) and a breakthrough that enabled their wide adoption. Figure 12 shows an example of a single pixel of an array of SPAD detectors, fabricated in CMOS technology.

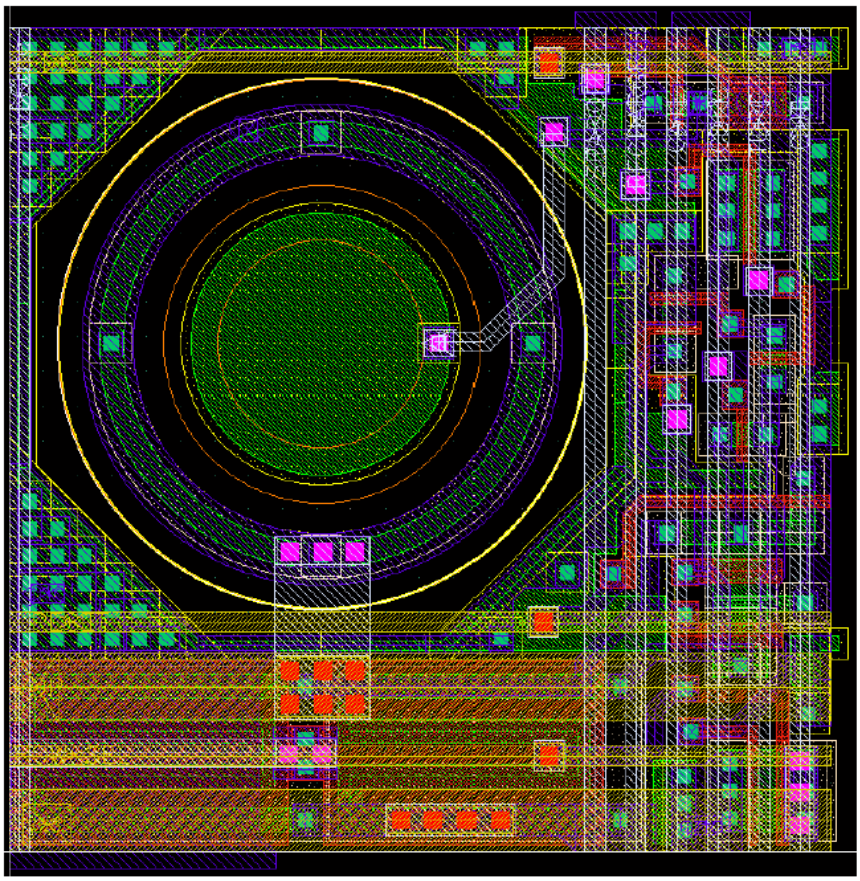


Figure 12. Typical layout of a CMOS fabricated, single pixel, SPAD detector.

SPADs can be used in most situations as a replacement for PMTs as single photon counters. Nowadays SPADs are used in quantum cryptography, spectroscopy, LIDAR applications, DNA analysis, particle measurements, fluorescence microscopy and obviously single photon counting. Most of the methods in this non-exhaustive list have something in common – they are imaging techniques. Imaging using just one sensor requires scanning to build up an image, which is time and resource consuming. Recently, CMOS technology has advanced to create relatively cheap SPAD arrays, constantly growing in size. Beginning with as few as a couple of pixels, nowadays SPAD arrays can achieve over 65 000 pixels of single photon sensitive detectors. Very high speed of operation makes them desirable in techniques like Positron Emission Tomography (PET), or FRET. Measurements of photon arrival time makes depth scanning possible. Finally, when used in time gated mode and synchronised with excitation source, SPAD arrays can be used as a FLIM system.

In our studies we have used one of the largest SPAD arrays available to date, to construct a fluorescence lifetime imaging system. The technology of the SPAD itself as well as more specific details about our setup will be described in following chapters.

This chapter presented a general background to support the following practical parts. Fluorescence and light emission is a connecting factor for of all of the research to be presented. The second chapter will focus on a use of a smartphone to quantify light emission of various medically relevant assays; the third chapter will present a complex setup based on the SPAD sensor, used to measure fluorescence lifetime with an aim to enable cancer detection.

2. Smartphone projects

2.1. Smartphones in science

The Pew Research Center reported that in 2017 89% of American adults, at least occasionally, uses the internet or owns a smartphone and 69% of them are users of social networking sites (28). The popularity of smartphones is a relatively new phenomenon and has a major influence on many aspects of people's lives (29-33). Smartphones popularised citizen science where amateurs help the scientists in their research collecting and/or processing data as part of a scientific enquiry (34-39)

Finally, smartphones can be used directly as a measurement device by the scientist, providing data in place of professional readout systems. Of course usage of consumer electronics (especially of imaging devices) in science was explored before the era of smartphones. In 2007, a paper on using DSLR cameras as a spectrophotometric device was published (40). A series of tests including quantum efficiency, dynamic range or linearity was performed, proving that it worked despite some limitations. Smartphones were next logical step as a cheaper, more accessible and more flexible device.

The smartphones were used to measure occupational noise (41), heartbeat pattern recognition (42) or physiological audio signals such as breathing sounds (43)

By comparing capabilities of modern smartphones to scientific readout systems, it is easy to realise that they include most of the main elements that the typical imaging system is constructed of. Most importantly, as well as providing sensors, smartphones can also serve as many peripherals, like a computational device, a control system and a data acquisition device. Depending on the use, the purpose or the measured target, specific add-ons can complete the setup.

2.2. Smartphone imaging in biophotonics

Imaging techniques are playing a major role in biomedical sciences and biotechnology. These communities are also paying attention to the transition from professional laboratories to point-of-care testing (POCT).

The smartphone was proposed for use in biophotonics for the first time in 2008 (44). The smartphone was mentioned there as one of the options for acquiring an image of a paper-based urine test, as an alternative to a portable scanner, a desktop scanner and a digital camera. All these devices, connected to a computer and the internet, were able to send the image to the specialist that would assess it. No local analysis was done on the computer or the smartphone. Not surprisingly, the smartphone was the least accurate method tested in this publication, even though the advantages of availability of the device was realised.

One of the pioneers of developing actual smartphone readout devices in biomedical imaging is the UCLA based group under the leadership of Professor Aydogan Ozcan. In 2010 they published an article entitled “Lensfree microscopy on a cellphone” (45). Using an attachment to the phone, they used an additional LED to illuminate the sample, which scattered the light and created a hologram (a shadow cast on the sensor) of each object on the detector array of the cell phone. Using digital image reconstruction, they reported imaging of single particles as small as a couple of micrometres. It was the very initial proof of concept, enabling just sensing of the presence of a particle in a low noise environment.

In the same year the same group presented detection of a fluorescent particle using a very similar setup (46). Instead of measuring scattered light, fluorescent emission was measured (Figure 13). The sample was excited in the same way as it was illuminated before, with an LED from the side. An additional filter was used to achieve a dark field environment and an additional lens was added to the attachment, improving the collection efficiency of the fluorescent signal.

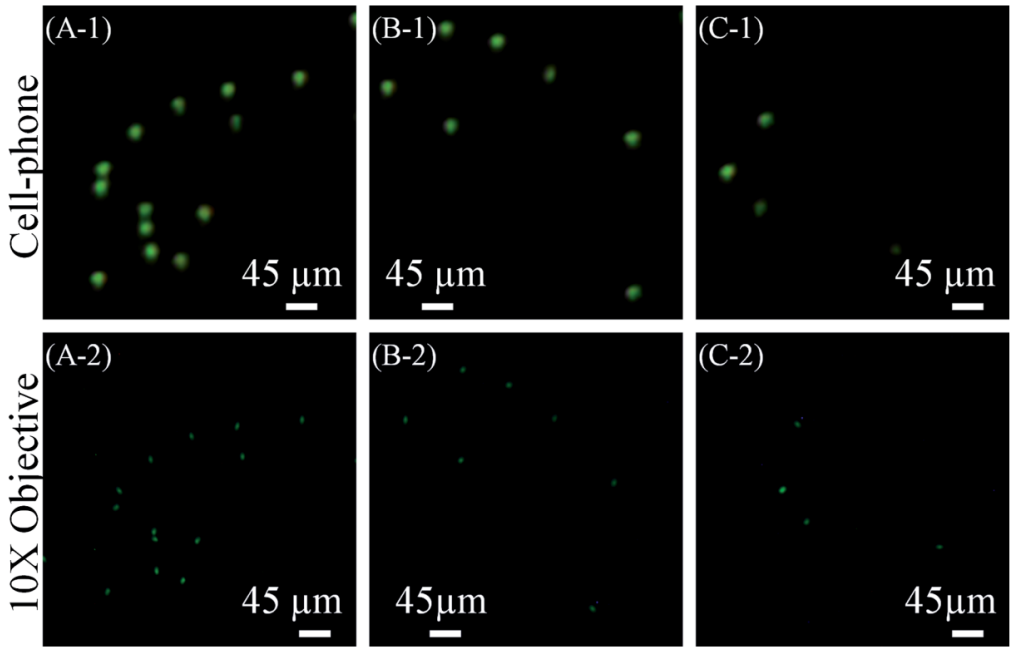


Figure 13. (Top) Giardia Lamblia cysts that are imaged using the fluorescent cell-phone microscope. (Bottom) Microscope objective (10, NA ¼ 0.25) images of the same samples are also provided for comparison purposes. (46)

In 2011 Ozcan's group published an evolution of the same setup – optofluidic fluorescent imaging cytometry (47). Instead of the steady sample slide, a microfluidic channel with a sample in motion was installed. The video analysis of the particles could estimate with fair precision the density of white blood cells in human blood samples.

In the same year, another group from the University of California joined the race and presented similar concepts of mobile microscopy and spectrography (48). The difference in their approach was use of typical optical elements (i.e., lenses, transmission grating) to achieve the desired functionality. In the case of microscopy, they used a 1 mm diameter ball lens to achieve about 350-times magnification (Figure 14). To use the phone as a spectrometer, they first attached a transmission grating and then another slit with a width of approximately 1 mm at the distal end of the tube. Using this setup, they reported a 300 nm bandwidth with a limiting spectral resolution close to 5 nm.

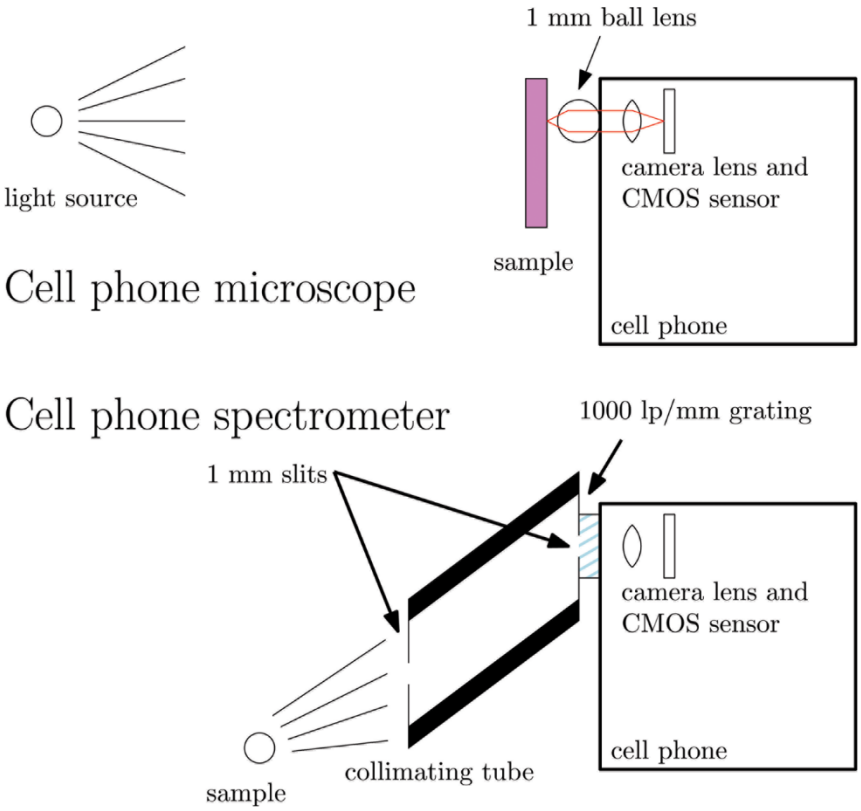


Figure 14. Top panel shows the cell phone microscope achieved by adding a ball lens to the cell phone camera system. Lower panel shows the cell phone spectrometer, constructed by adding a grating and collimating tube to the cell phone camera.(48)

In the following years Ozcan's group continued presenting new solutions for detecting different particles and medical conditions like Escherichia coli using quantum dots (49), a food allergen testing platform that images and automatically analyses colorimetric assays performed in test tubes (50) or albumin testing in urine using its auto fluorescence (51).

Over the next few years, the number of applications of smartphone readout systems multiplied in fields such as clinical chemistry (measuring levels of glucose in whole blood (52), serum (53), and urine (54) samples) , immunoassays (avian influenza (55), sexually transmitted diseases (56, 57)), nucleic acid detection (HSV-2 (56), single molecule of HCV (58)) and biosensing (detection of L-lactate in oral fluid and tears (59), sensitive measurements of pH (60), detection of cholesterol (61) or steroid hormone (62)) .

2.3. The smartphone camera

The imaging sensors in mobile phones have developed significantly in recent years. However, the general construction and technology behind each upgraded model stays mostly unchanged. While different camera modules are used in various smartphones providing different specification and capabilities, their principles of operation are almost identical. In this part we will present the typical smartphone camera sensor and its most important characteristics.

A digital photographic camera is constructed of two general parts: the sensor and the lens system. In the miniaturised smartphone environment these elements come together on one small chip. In smartphones, CMOS technology (complementary metal-oxide-semiconductor) won over the competitor CCD (charged-coupled device) sensor manufacturing technology because of CMOS's lower power consumption. Each CMOS pixel is a photodetector that captures photons, amplifies the signal and converts it to a digital form, so recording light intensity. To capture pictures in colour, a colour filter array is used. The most widely used CFA pattern in the image acquisition industry - a Bayer pattern is emulating the human eye (Figure 15). It employs a repeating 2 x 2 pattern consisting of one blue, one red and two green (RGBG) filters similar to two cones collecting green spectrum data. Each colour filter physically covers the area of one photosensitive pixel. The process of merging these 3 interlaced channels in order to obtain a full-colour image is called demosaicing.

Demosaicing is a process of colour interpolation and can be performed using many different algorithms like AMaZE (Aliasing Minimization and Zipper Elimination), LMMSE (linear minimum mean square-error estimation), or EAHD (Adaptive Homogeneity-Directed) to name just a few. The base of most of these algorithms is a convolution of the original image with an appropriate kernel. Two kernels are required: one (Fg) for estimating the missing green values and one (Fc) for estimating missing red/blue values. These kernels for the linear interpolation are:

$$Fg = \frac{1}{4} \times \begin{bmatrix} 0 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 0 \end{bmatrix}, Fc = \frac{1}{4} \times \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad (8)$$

and for the different type, cubic interpolation are:

$$Fg = \frac{1}{256} \times \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -9 & 0 & -9 & 0 & 0 \\ 0 & -9 & 0 & 81 & 0 & -9 & 0 \\ 1 & 0 & 81 & 256 & 81 & 0 & 1 \\ 0 & -9 & 0 & 81 & 0 & -9 & 1 \\ 0 & 0 & -9 & 0 & -9 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad 9)$$

$$Fc = \frac{1}{256} \times \begin{bmatrix} 1 & 0 & -9 & -16 & -9 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -9 & 0 & 81 & 144 & 81 & 0 & -9 \\ -16 & 0 & 144 & 256 & 144 & 0 & -16 \\ -9 & 0 & 81 & 144 & 81 & 0 & -9 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -9 & -16 & -9 & 0 & 1 \end{bmatrix} \quad (10)$$

Afterwards, many methods use additional post demosaicing algorithms like Local Colour Ratio Based Post Processing. The main goal of this algorithm is to correct unnatural changes in hue by smoothing the colour ratio channels. For the moment, let $\mathrm{R}(\mathrm{p},\mathrm{q}), \mathrm{G}(\mathrm{p},\mathrm{q})$, and $\mathrm{B}(\mathrm{p},\mathrm{q})$, be the functions representing the red, green, and blue colour channels, respectively, where (p,q) is the pixel location. The first part of the algorithm adjusts the green colour channel based on the red and blue colour channels. The example is the equation (one of six possible combinations) used to adjust the $\mathrm{G}(\mathrm{p},\mathrm{q})$ value according to surrounding B pixels at locations in ζ:

$$\zeta = \{(p+1,q),(p,q+1),(p-1,q),(p,q-1)\} \quad (11)$$

$$G(p,q) = -\beta + [G(p,q)+\beta] \times \underset{(i,j)\in\zeta}{mean} \frac{B(i,j)+\beta}{G(i,j)+\beta} \quad (12)$$

where β is a nonnegative constant, as defined by the normalized colour ratio model (63). Corresponding equations are applied to all combinations of channels.

Despite all the efforts in developing better demosaicing algorithms, some typical problems can still be seen. False colour artefacts manifest along edges, where the interpolation algorithm stepping across the edge fails. As a result, we can observe slight discolorations around the edges, or false colour changes on and around small objects. The other common problem is a zippering artefact, which is another name for

edge blurring that occurs in an on/off pattern along an edge. It results in blurring the edges through adding gradating coloured borders.

Smartphone manufacturers do not reveal which demosaicing algorithms were used in their products. Due to the fact that most of them suffer from the same issues, it will be desirable to avoid edges when performing colour assessment.

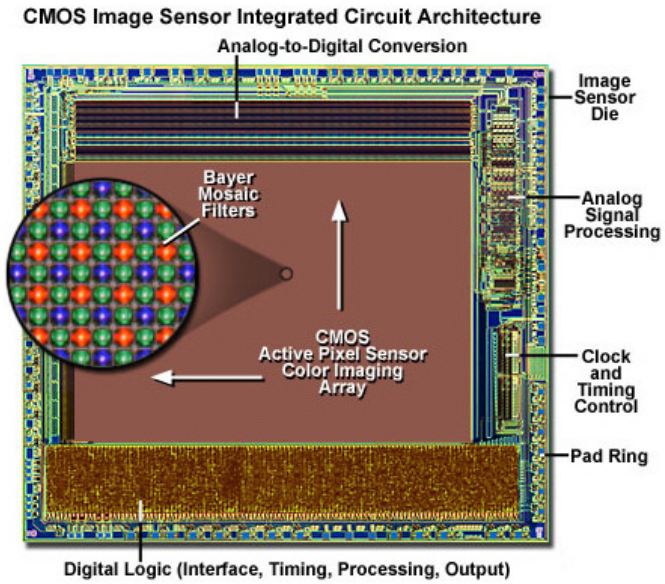


Figure 15. Structure of a CMOS manufactured photosensitive imaging sensor with the close-up on the Bayer Mosaic filter (64).

The most recognised characteristic that describes the camera is its number of megapixels, which is the number of millions of photodetectors in the array. Early smartphones were equipped with sensors of a size of about 0.2 Mpx. This number rose quickly to a typical size between 12 and 16 Mpx, with bigger matrices presented, but not popular. It is important to understand that the number of megapixels is not the most important factor determining image quality. Nevertheless, we will be using native/maximal image resolution settings to maximise the number of pixels to analyse.

Probably the most important physical characteristic of the sensor from a scientific point of view is the pixel size. The larger the pixel size is, the more photons it will be able to collect, which translates to better signal collection in low light level situations. Currently, we can observe a trend of increasing pixel size, with a typical size of about 1.6 microns compared to about 1 micron from couple of years ago.

The focal length is the exact distance where the lens focuses light rays, hence where the sensor needs to be. The focal length determines the field of view and the magnification of the imager. With the technological progress over the years, lens design has evolved from a simple double Gauss or Cooke triplet to many moulded plastic aspheric lens elements made with varying dispersion and refractive indices (Figure 16). Smartphone cameras with the focal length between 24-30 mm (of 35 mm-equivalent focal length) fall into the wide-angle lens bracket. Both field of view and magnification are very important factors in terms of mobile microscopy. A wider initial angle typically means a larger field of view in digital zoom conditions.

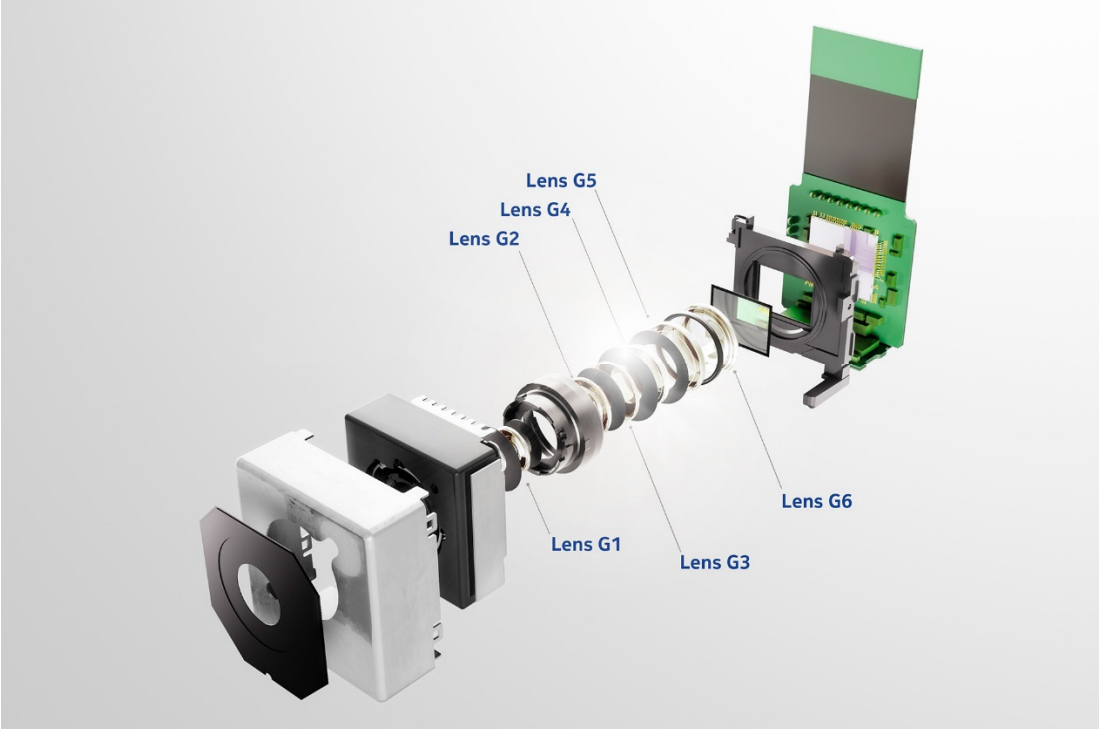


Figure 16. Modern, complex set of lenses in a smartphone's camera (64).

Last but not the least physical feature of the phone camera is its aperture or f-number. The f-number is a ratio between the focal length (f) and the size of the aperture (D), and tells us how much light can pass through to the sensor.

$$N = \frac{f}{D} \tag{13}$$

An f-number N=2, commonly named f/2, denotes that the focal length is twice as large as the aperture. A higher f-number has a wider depth of field, typically sharper images, less chromatic aberration and weaker bokeh (blur effect in planes that are out

of focus). Smartphone manufacturers tend to keep the f-number about f/2, which allows more light to hit the detector for better low light level imaging, while keeping previously mentioned parameters at satisfactory levels.

2.4. Image Acquisition

The process between capturing photons by an imaging array to the final image file as we can see it on the smartphone or the computer is complicated and realised of many steps. Each step can introduce its own errors or intentional changes, resulting in a different image. This makes comparing different images a complex procedure. Here we will try to analyse subsequent steps and their influence on the image. Unfortunately, manufacturers do not provide detailed information about each step and do not grant any control over many of them.

First, we need to realise that the smartphone camera is not designed for scientific use. We can only use options and functions that are provided by the manufacturer and the software company. We do not have free access to all the variables and theoretical functionality that could be possible, as for instance DSLR cameras are providing. Secondly, these cameras are designed to take the most appealing photographs, not the most accurate images. This means that, even though natural tone of colours is one of the factors being evaluated by the DXO mark (65) (the largest company comparing photographic cameras and smartphone cameras), many manufacturers will choose to saturate their colours, or add softening filters to the images, to make them more attractive rather than accurate.

The factors that have influence on the image can be divided into four main categories: hardware, software, user influence and environment influence.

Different smartphones are equipped with different sensors, from different manufacturers. Currently, the biggest provider is Sony with its IMX series, which can be found in almost every new flagship model of the largest smartphone manufacturers (for the list of sensors and smartphones that used them please refer to Wikipedia (66)). Sensors can be sorted by their price/quality but also date of release, providing different characteristics. Even phones with exactly the same sensor can at the end produce significantly different images: for example the IMX362 sensor used in

the HTC U11 phone scored 90 points in the DXO test, while a Google Pixel 2 phone with same sensor scored 98 points (65).

A second significant hardware stage is the Image Signal Processor (ISP). Its primary task is to perform demosaicing, which is interpolating colours based on the readout from Bayer pattern. This step on its own can be performed with different nuances as discussed above. In addition, this stage more and more often is responsible for autofocusing, exposure and white balance, as well as correcting for lens imperfections like vignetting or colour shading imparted by the imperfect lens system.

Here, the hardware stage fluently mixes with the software stage. After the initial signal processing corrections, a second round of edits are done on the regular processor of the phone, which can be influenced by the software or the user. Parameters like colour saturation, brightness, shutter speed, ISO setting or more complicated image processing algorithms like High Dynamic Range (HDR) imaging, are controlled and applied. Using all these variables, the automatic algorithm tries to capture a picture with the best quality. Subsequently all the aesthetic changes to the image are applied, like additional denoising or sharpening filters, changes to the colour saturation, etc. Some of these parameters can be user-controlled in manual mode, so the automatic software actions are now mixed with influence from the user.

Before the button is pressed and the (digital) shutter 'opens', most of the stock camera applications on the smartphone allow manual setting of some of the parameters of the image acquisition and processing. Depending on the application, automatic or manual setting may be desirable. Some of these options are:

ISO

In traditional analogue photography, ISO was a parameter describing film sensitivity. The same variable was transferred to digital photography, changing to sensitivity/increasing amplification of the sensor. Digital sensors have their sensitivity determined by the exposure required to produce a given characteristic of the output signal. Out of five possible methods presented by ISO, only two are now common: zero-bias photovoltaic mode resulting in a photoelectron build-up on the output, or reverse-biased photoconductive mode where absorbed photons release a

photoelectron that directly contributes to the current flowing through the diode. The current readout is sent to the analog-to-digital converter, which results in higher exposure received and a brighter pixel. ISO speed ratings are expressed in terms of the luminous exposure H:

$$H = \frac{qLt}{N^2} \quad (14)$$

where L is the luminance of the scene, t is the exposure time, N is the aperture f-number and

$$q = \frac{\pi}{4} T v(\theta) \cos^4 \theta \quad (15)$$

is a factor dependent on transmittance of the lens T, the vignetting factor $v(\theta)$ and the angle relative to the axis of the lens θ. ISO is often measured using the exposure required to saturate the photosites - the depletion region within the photodiodes on the sensor. According to the standard ISO 12232, the equation to define saturation-based speed is

$$S_{sat} = \frac{78 lx \cdot s}{H_{sat}} \quad (16)$$

where $H_{sat} = L_{sat} \times t$. The number 78 is chosen to deal with specular reflections that would appear brighter than a 100% reflecting white surface. The value of S_{sat} (rounded to the closest standard ISO number) is an ISO determinant. (67).

The higher the ISO number, the more sensitive a sensor is, e.g., ISO800 is 8 times more sensitive than ISO100. One must bear in mind that while a higher ISO value amplifies low signals, it also amplifies the noise.

Shutter speed

In smartphone cameras there are no physical shutters. The shutter is initiated by simply turning the sensor on and off. For the 'faster shutter', the sensor is active for a shorter period of time and collects less light. If this setting is user-controllable on the phone, it is very desirable to select long exposure times for static low light level imaging so that as much signal as possible is collected.

White balance

White balance is an extremely important algorithm that helps to achieve natural colours in photography regardless of external illumination. Human sight has such a compensation, and even with significantly different illumination we can recognise colours properly and perceive them normalised to some extent. A linear and straightforward digital sensor initially sees colours as they are and needs white balance to imitate human ability to adjust to the illumination 'colour temperature'.

The colour temperature describes illumination in terms of the spectrum of light which is radiated from a blackbody with that surface temperature. Figure 17 shows that a temperature of about 5000K has an approximately even distribution, which results in colour close to white. Temperature closer to 3000K has significantly more of the red wavelengths, whereas a higher temperature of about 9000K illuminates more in blue.

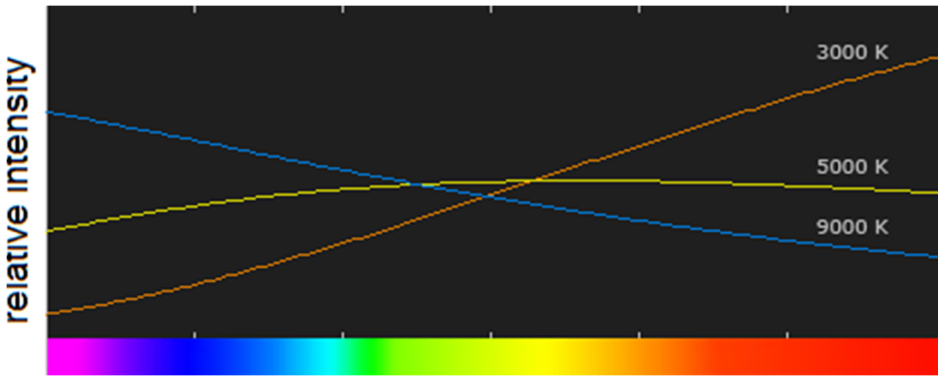


Figure 17. Normalised spectra of intensities of different colour temperatures, in kelvin .(68)

Table 1.Typical names for white balance compensation options

Colour Temperature	Light Source
1000-2000 K	Candlelight
2500-3500 K	Tungsten Bulb (household variety)
3000-4000 K	Sunrise/Sunset (clear sky)
4000-5000 K	Fluorescent Lamps
5000-5500 K	Electronic Flash
5000-6500 K	Daylight with Clear Sky (sun overhead)
6500-8000 K	Moderately Overcast Sky
9000-10000 K	Shade or Heavily Overcast Sky

White balance is a compensation of the temperature of the scene illumination to minimize unwanted colour cast on the image. Automatic white balance is a procedure performed by the camera by finding a neutral colour region (grey or ideally white) and

‘guessing’ scene temperature to be able to normalise all the colour channels properly. Typically, the user can define one of the preloaded scene illumination schemes manually. They are represented either by the temperature value or by a typical light source for common lighting conditions, as presented in Table 1.

White balance is the camera’s most powerful reaction to the last group of factors influencing the image: environmental factors. The temperature and intensity of the illumination are the most important of these factors in both photography and scientific use. To cope with this issue, one more concept - the standard illuminant – is needed. We will discuss it shortly.

All elements presented so far are crucial for proper imaging and need to be in perfect balance to take a proper and accurate photograph. The image obtained this way then needs to be saved.

2.5. Digital image file and its standards

2.5.1. Compression

The standard file format for non-lossless photography is JPEG. JPEG stands for Joint Photographic Experts Group, the name of the committee that created this lossy standard of compression. The compression is based on the Discrete Cosine Transform (DCT), justified using a perceptual model based loosely on the human psychovisual system, discarding high-frequency information, i.e., sharp transitions in intensity and colour hue. This standard gained popularity for two reasons: the degree of compression can be adjusted, allowing a selectable trade-off between storage size and image quality; and it still achieves very good image quality even with lossy compression of about 10 times in size. For more technical information about JPEG compression, please refer to the specification (69).

2.5.2. The colour space

A colour space is the defined range of colours in a given analogue or digital representation. A colour model can be arbitrary, for example with specified colours with names or numbers. Alternatively, a colour model can be structured

mathematically, describing triplets of numbers with mathematically specified dependencies to wavelengths or a different colour model, with a gamut positioned on a CIE 1931 xy chromaticity diagram (like in Figure 19), which is called mapping. The reference standard for all colour spaces was set in 1931 by the International Commission on Illumination (**CIE** for its French name) (70), who were the first to define quantitative relationship between distributions of wavelengths in the electromagnetic visible spectrum and colours perceived by the human eye. The CIE XYZ 1931 colour space is considered a full colour space, while most others can represent only a defined subset of the full colour space.

The CIE XYZ tristimulus values are based on the cone responses of the human eye. Each variable does not represent real colour (they cannot be generated in any light spectrum). Instead, Y represents luminescence and the XZ plane will contain all possible chromacities at that luminescence. XYZ values are given by functions:

$$X = \int_{\lambda} L_{e,\Omega,\lambda}(\lambda)\bar{x}(\lambda)\, d\lambda \quad (17)$$

$$Y = \int_{\lambda} L_{e,\Omega,\lambda}(\lambda)\bar{y}(\lambda)\, d\lambda \quad (18)$$

$$Z = \int_{\lambda} L_{e,\Omega,\lambda}(\lambda)\bar{z}(\lambda)\, d\lambda \quad (19)$$

Where $L_{e,\Omega,\lambda}$ is a spectral radiance, $\bar{x}(\lambda), \bar{y}(\lambda)$ and $\bar{z}(\lambda)$ are the colour matching functions shown in Figure 18, λ is the wavelength of the equivalent monochromatic light (measured in nanometres), and the standard limits of the integral are $\lambda \in [280,780]$.

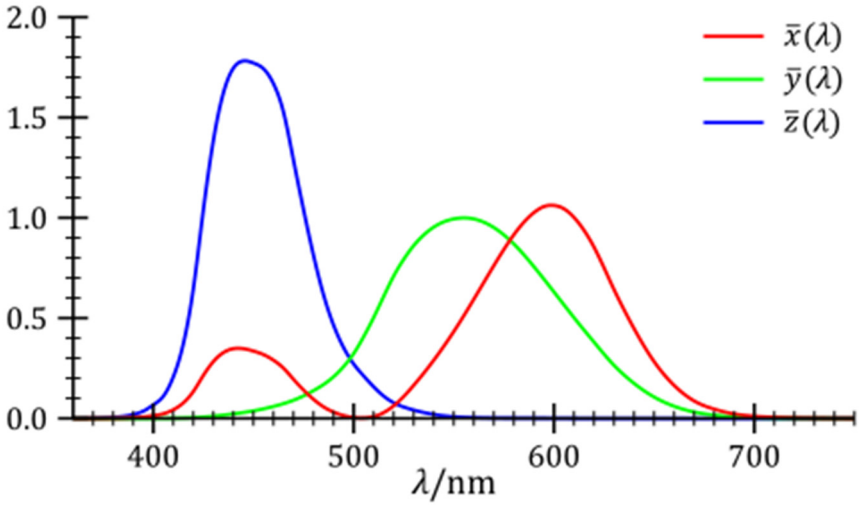


Figure 18. The CIE standard observer colour matching functions (43).

Since both the human eye and CIE representation have 3 variables, the visual representation of this colour space should be three-dimensional. To conveniently display such a 3D space, it is standard practice to present 2D graphs as chromaticity of a colour specified by the two derived parameters x and y, two of the three normalized values being functions of all three tristimulus values X, Y, and Z, for given luminescence Y.

$$x = \frac{X}{X+Y+Z} \tag{20}$$

$$y = \frac{Y}{X+Y+Z} \tag{21}$$

$$z = \frac{Z}{X+Y+Z} = 1 - x - y \tag{22}$$

This procedure is widely used to compare colour spaces in practice, and such a graph is presented in Figure 19.

Another important colour space, also defined by CIE is CIELAB (or CIE L*a*b) introduced in 1976. CIELAB was designed to be perceptually uniform with respect to human colour vision. It other words, it means that the function of the change of the colour is linear in respect to the human perception, not with number of photons. CIELAB colours are defined relative to the white point, so values do not define absolute colours unless the white point is also specified. White point can be defined as

one of the standard illuminants that we will describe shortly. The lightness value, L* is normalised according to the white point and varies between 0 and 100 (100 being the brightest). The a* axis presents the change between green and red, ranging from negative values for green and positive values for red. The b* axis presents the change between blue and yellow, ranging from negative values for blue and positive values for yellow. Since CIELAB was derived from the CIEXYZ colour space, which is a basic colour space, presenting transformation equations between them is the most typical and the most commonly used. The CIELAB–CIEXYZ conversions goes as follows:

$$L^* = 116f\left(\frac{Y}{Y_n}\right) - 16 \tag{23}$$

$$a^* = 500\left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right) \tag{24}$$

$$b = 200\left(f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right) \tag{25}$$

where

$$f(t) = \begin{cases} \sqrt[3]{t} & if\ t > \delta^3 \\ \frac{t}{3\delta^2} + \frac{4}{29} & otherwise \end{cases} \tag{26}$$

$$\delta = \frac{6}{29} \tag{27}$$

Here, X_n, Y_n and Z_n are the CIE XYZ values of the normalised reference white point. For standard illuminant D65 and normalisation to $Y_n = 100$, these values are:

$$X_n = 95.047$$

$$Y_n = 100$$

$$Z_n = 108.883$$

Smartphones typically save images in sRGB colour space. sRGB is a standard of RGB proposed by HP and Microsoft cooperatively in 1996, subsequently standardized by the IEC as IEC 61966-2-1:1999 [44]. The sRGB colour space defines a gamut of colours that can be represented by 24 bit (8 bit i.e., 256 levels per colour channel) chromaticities of the red, green, and blue primaries (Figure 19). Even though a smartphone camera records in RGB, for the JPEG compression algorithm the image is converted to the $Y'C_BC_R$ colour space. $Y'C_BC_R$ is a family of colour spaces developed

specifically for compression, eliminating RGB system redundancy and allowing for decreased resolution of single components. Slightly different spaces are used for different purposes like analogue, digital or high definition television signal, or for the most accurate retention of luminance information. There is also a separate $Y'C_BC_R$ colour space for JPEG format, tailored for the use of full 8 bit variables with its own conversion equations expressed to six decimal digits of precision from RGB to $Y'C_BC_R$:

$$Y' = 0 + (0.299 \times {R'}_d) + (0.587 \times {G'}_d) + (0.114 \times {B'}_d) \quad (28)$$

$$C_B = 128 - (0.168736 \times {R'}_d) - (0.331264 \times {G'}_d) + (0.5 \times {B'}_{d)} \quad (29)$$

$$C_R = 128 + (0.5 \times {R'}_d) - (0.418688 \times {G'}_d) - (0.081312 \times {B'}_{d)} \quad (30)$$

and from $Y'C_BC_R$ to RGB:

$${R'}_d = Y' + 1.402 \times (C_R - 128) \quad (31)$$

$${G'}_d = Y' - 0.344136 \times (C_B - 128) - 0.714136 \times (C_R - 128) \quad (32)$$

$$C_R = Y' + 1.772 \times (C_B - 128) \quad (33)$$

where apostrophe denotes gamma corrected values.

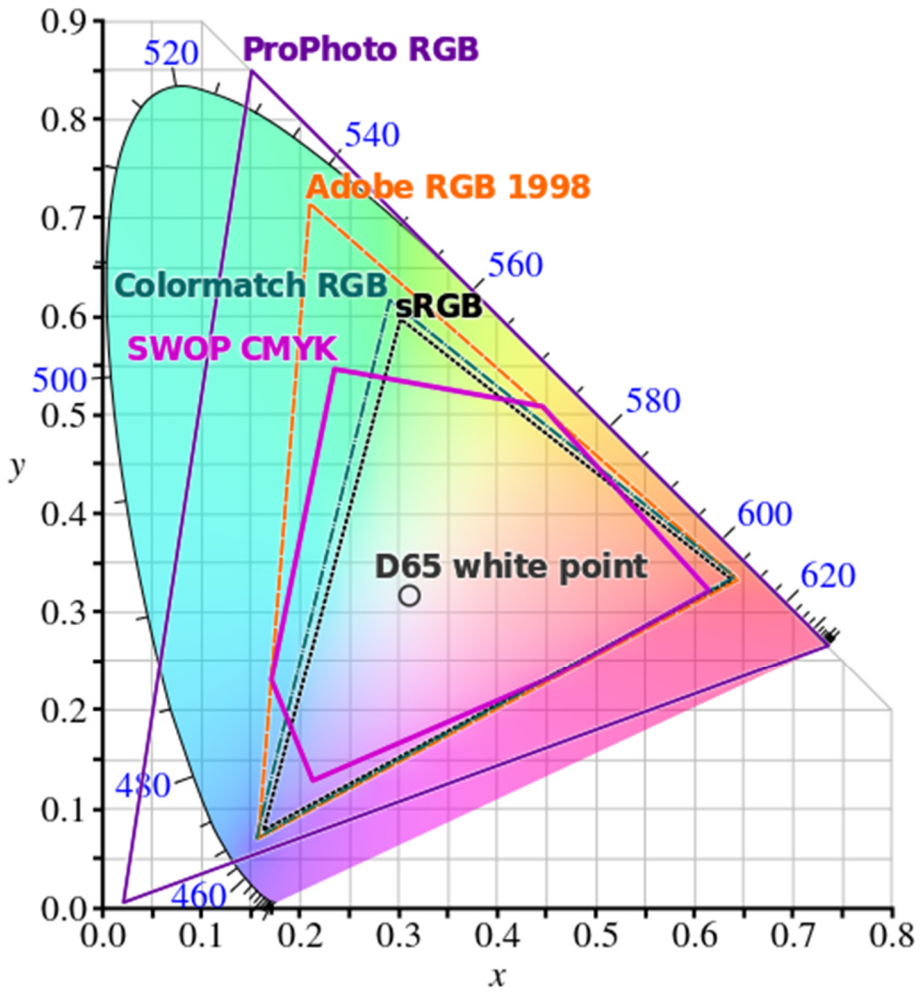


Figure 19. Comparison of some RGB and CMYK colour gamuts on a CIE 1931 xy chromaticity diagram. Blue numbers represent corresponding wavelengths in nm of the colours at the edges of the graph. This diagram displays the maximally saturated bright colours that can be produced by a computer monitor or television set. Triangles represent the range of colours that each colour representation is capable of presenting (71).

2.5.3. Gamma correction

A gamma correction is a nonlinear scaling used to adjust the linear colour representations to a nonlinear one that better reflects the response of the human eye to the digital screen. In case of sRGB, the gamma correction function was designed for CRT displays.

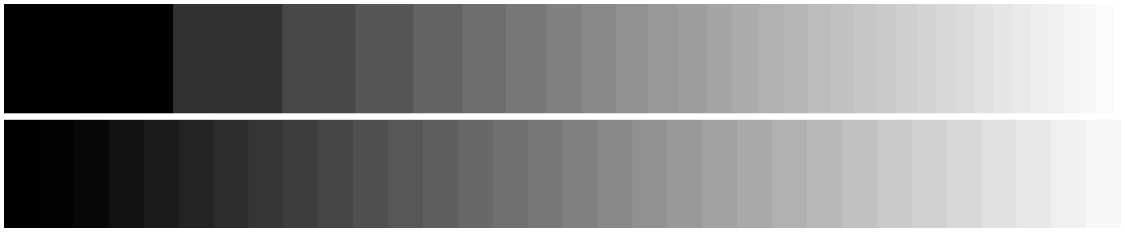


Figure 20. Gradients in linear colour grading (top) and sRGB gamma corrected colour space (bottom).

Gamma correction is a function that enables use of more or fewer bits in extreme areas. The human eye can distinguish more tones in very dark conditions than in very bright. Once the gamma correction function gradually increases this tone resolution with a darker image (Figure 20). The general gamma correction function is

$$V_{out} = V_{in}{}^{\gamma} . \tag{34}$$

The curve on a log–log plot is a straight line, with slope equal to γ. The sRGB colour space though, has an untypical gamma correction, becoming linear after a specific intensity. This is represented by the following formula:

$$I_{srgb} = \begin{cases} 12.92 I_{linear}, & I_{linear} \leq 0.0031308 \\ (1+a) I_{linear}^{\frac{1}{2.4}} - a, & I_{linear} > 0.0031308 \end{cases} \tag{35}$$

where I represents RGB colour channel values and $a = 0.055$. This formula is presented in Figure 21. The grey line presents the function for $\gamma = 2.2$, which is a common approximation of equation 34. In this study we have used the full formula.

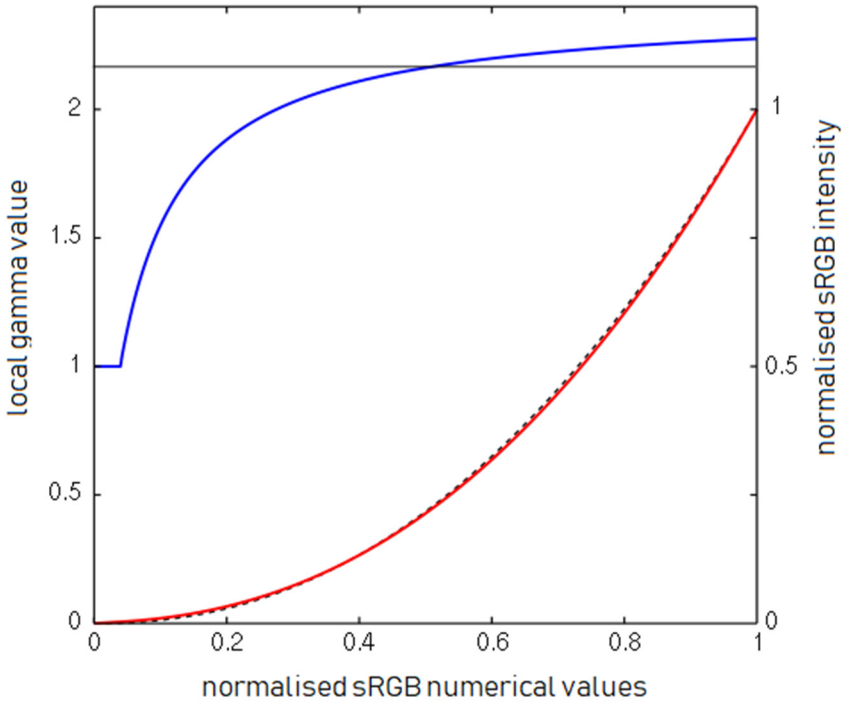


Figure 21. Plot of the sRGB intensities versus sRGB numerical values (red), and this function's slope in log-log space (blue), which is the effective gamma at each point. Below a compressed value of 0.04045 or a linear intensity of 0.00313, the curve is linear so the gamma is 1. Behind the red curve is a dashed black curve showing an exact gamma = 2.2 power law (23)

2.5.4. Standard illuminants

JPEG compression requires conversion between RGB and $Y'C_BC_R$ colour spaces. Since different colour spaces represent different gamut and intensity resolutions, any conversion between them needs a common standardised reference point. In fact, even comparison of different images taken under different illumination conditions within the same colour space, requires a common reference point. This is why in 1931 the International Commission on Illumination introduced standard illuminants (70). An illuminant is a mathematical representation of a theoretical light source, used for calculating tristimulus values from a spectrophotometric measurement. The relative power distribution of a real light source might also be used for calculation, but is difficult to standardise. The spectral power distribution varies between light types. Daylight (illuminant D65), for example emits a relatively small amount of the UV light and peaks around 460 nm, with slowly decreasing power for longer wavelengths. To

determine an object's colour using a standardised method, calculations must be based on one of the illuminants.

CIE has introduced a number of standard illuminants, for example: Illuminant A that is intended to represent typical, domestic, tungsten-filament lighting; Illuminant C represents average daylight with a correlated colour temperature of 6774 K; Illuminant E is an equal-energy radiator; Illuminant series F models various types of fluorescent lighting. Some illuminants are presented in Figure 22.

CIE standard illuminant D65 has now replaced illuminant C as a more accurate representation of an average noon sky daylight. Its colour temperature is 6504K. It is widely used for a general evaluation of colour, but more importantly it provides a visual correlation with spectrophotometric readouts (for this colour temperature the human eye perceives neutral colours), and introduces a standard for global industry. It is also chosen to define the white point in sRGB colour representation. We will use D65 illuminant to compare colours in all subsequent experiments, even if taken under different, definable illumination, in order to keep a constant reference point.

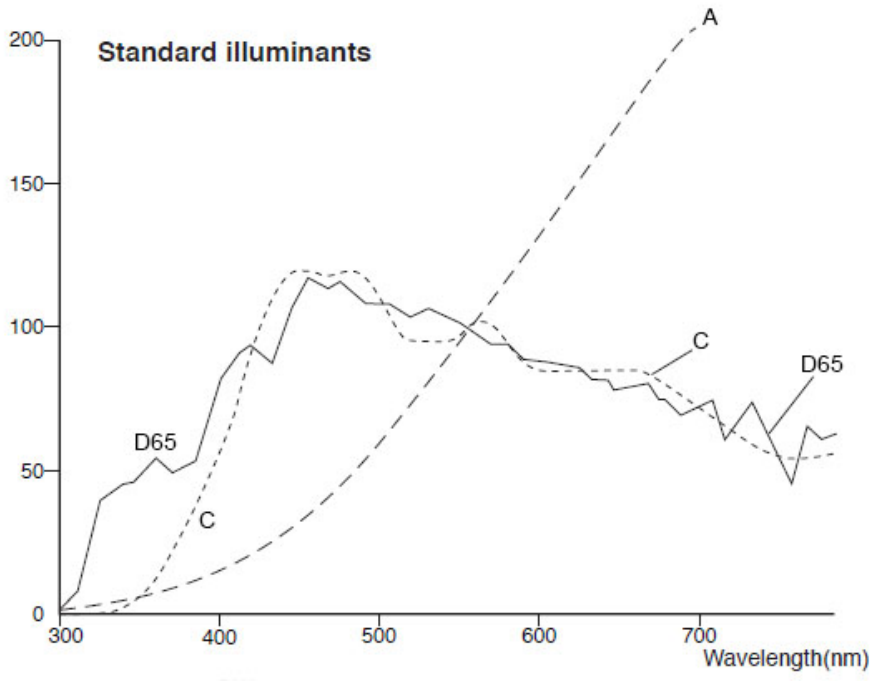


Figure 22. Relative spectral power distributions (SPDs) of CIE illuminants A, C, and D65 from 300 nm to 780 nm (72)

2.6. Quantification of spectrofluorometric and colorimetric signals

Performing a spectrophotometric readout using a smartphone is in principle not very different from a regular scientific readout. Quantification itself comes down to a proper processing, analysis and interpretation of the output image. We will use the phone to investigate assays that require either readout of intensity or readout of colour.

In the case of fluorometry, the original red, green and blue colour channels for each pixel of the image, each with a value between 0 and 255, in non-direct way represent counts of photons of specific wavelengths (or spectra). Proper preparation of the whole readout setup, then proper understanding of phenomena occurring in the phone, and finally correct image processing and analysis of calculated signals, should produce arbitrary values that will allow for quantification of emitted light; from there an application of Beer-Lambert's law (explained in section 1.4) allows the designation of the concentration of the target assay.

In the case of colorimetry, instead of measuring transmittance/absorbance of the target at a given wavelength, the visible colour under given illumination conditions is the determinant of assay concentration. By performing measurement using standard illumination, and then additionally characterising that illumination, we can normalise colours to allow reliable concentration detection

For both of these types of measurements, we can prepare working curves for each assay type, mapping the properties of a prepared set of assays of known concentrations, in the same way as would be done for a standard spectrophotometer or colorimeter.

2.7. Medically Relevant Assays with a Simple Smartphone and Tablet Based Fluorescence Detection System - the concept

Fluorescent assays are widely used across the life sciences. Fluorescent assay readout can be carried out by a variety of devices. Most commonly, a microplate reader is used where a number of wells (frequently 96) filled with the assay solution can be examined at once (73, 74). In most cases, assay readout requires professional research laboratory equipment. There is a need to develop easy to use, portable and affordable systems for assay readout suitable for point-of-care applications.

In this work, we decided to develop a cell phone-based portable bioassay platform with a goal of using a minimal number of additional components, and using commonly available components where possible. Our final system uses a screen from a commercial tablet as an excitation light source, linear polarisers that separate the excitation light from the fluorescence signal being read out, a smartphone as a camera that takes still pictures of an assay well, and a black box protecting the assay against ambient light. Specialised software for the smartphone is not required and setting up the standard commercial tablet and smart phone is straightforward. The assay readout is simply carried out by taking an image of the assay well. After image acquisition, the pictures are post processed to quantify the result. In order to assess the system's performance and sensitivity, two kinds of assays for the detection of trypsin and collagenase were performed using this smart phone-based device. Both used fluorescein isothiocyanate (FITC) dye as a reporter, which has an excitation peak at 495 nm and emission peak at 519 nm (Figure 25).

2.8. The setup

The device used to perform the measurements was constructed as shown in Figure 23. The well slide with assay samples was placed on the surface of the tablet screen with a polariser on top. The excitation source was a bright single square displayed on the tablet screen. The tablet screen used emitted linearly polarised light and the polariser was aligned perpendicularly in relation to the tablet screen polarization. The tablet,

assay well, and polariser were placed inside of a box with an opening for the camera; this box shielded the detection system from ambient light. The phone was placed on the top of the box, aligned to look through a hole to capture images.

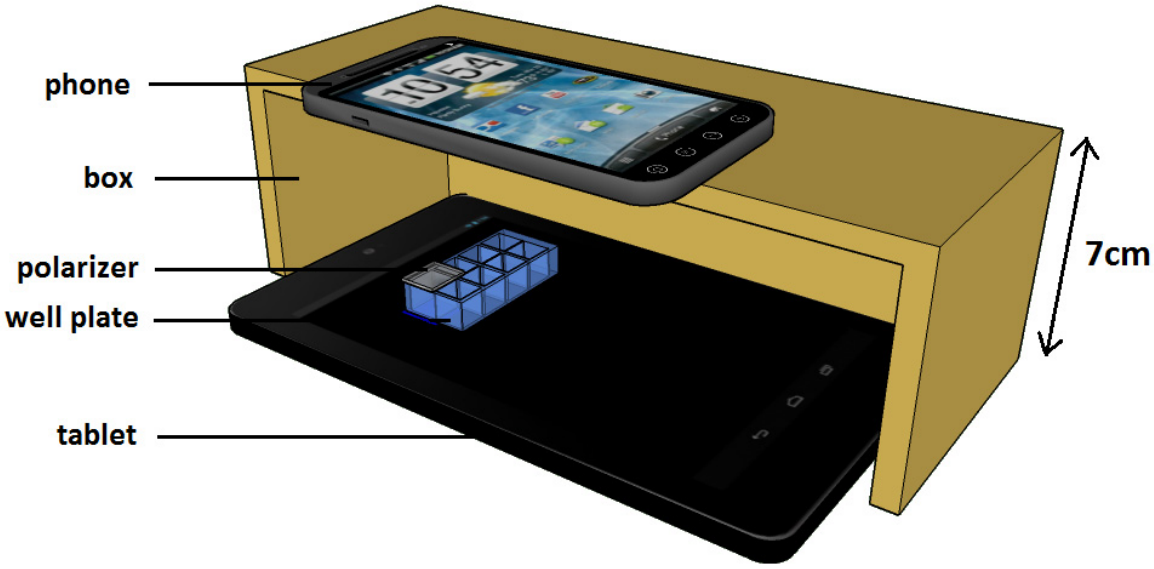


Figure 23. Schematic diagram of the device, with the box shown cut away.

We will now discuss each of the elements separately, presenting them and their role. Then we present the results of experiments to determine optimal conditions for the final setup and then proof-of-concept measurements of real assays.

2.8.1. Samples

Commercial assays to detect collagenase and trypsin typically use a sandwich immunoassay approach with two antibodies to analyse the amount of collagenase or trypsin in samples. Radio-immunoassays have also been used to detect serum trypsin level (75), however, fluorescent reporters, where a fluorophore selectively binds to a specific region or functional group on the target molecule, are much safer and more convenient to use especially in point-of-care situations. In this study we used commercial assays for trypsin and collagenase with a very common fluorescein isothiocyanate (FITC) dye as a reporter. This dye has excellent fluorescence quantum yield, good water solubility, and it reacts with amino groups of most proteins. It is

commonly applied in a variety of assays, including for FITC-casein (76), fibrinogen (77), Ca2+-ATPase (78), NO (79), C-reactive protein (80), and many other biomolecules. Figure 24 presents its excitation/emission spectra compared to the typical screen RGB colour emission.

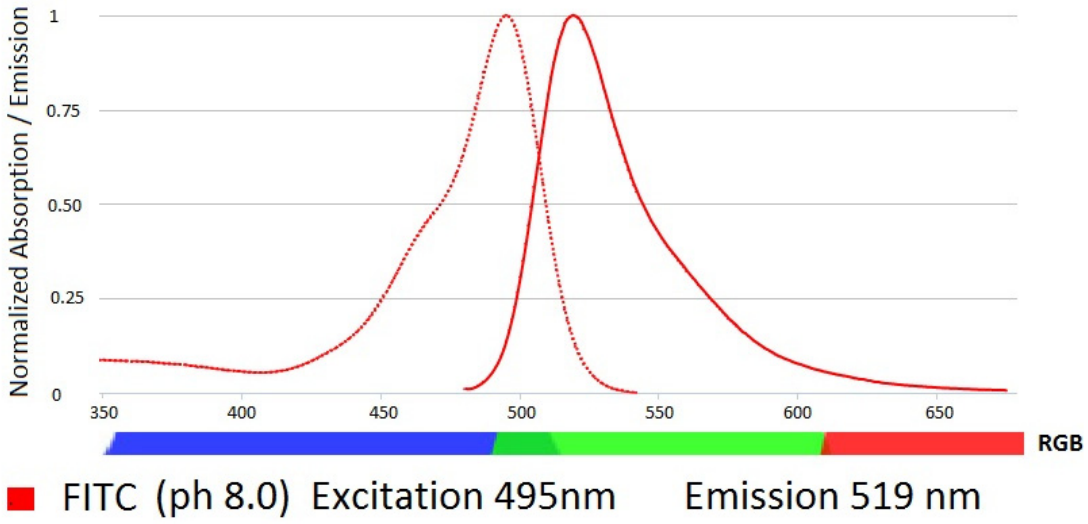


Figure 24. Excitation and emission spectra of fluorescein in comparison to a typical RGB colour emission range for LCD screens. Spectral data downloaded from Fluorophores.org.

The samples for the analysis were prepared as follows. We used the Pierce fluorescent protease assay kit (catalogue #23266, Thermo Scientific, Inc. Hudson, NH, USA) and type 1 collagenase assay kit (B-Bridge International, Inc., Cupertino, CA, USA, catalogue #AK07) and collagenase from Clostridium histolyticum (Sigma-Aldrich, St. Louis, MO, USA, Sigma Prod. No. C0130). For the trypsin assay, trypsin solution with different concentrations was mixed with FTC-Casein solution, followed by incubation at room temperature for 60 min. For the collagenase assay, collagenase enzyme crude from Clostridium histolyticum (Sigma Aldrich) was prepared at different concentrations (60, 40, 20, 10, 5, 2.5, 1.25) µg/mL in TESCA buffer (50 mM TES, 0.36 mM calcium chloride, pH 7.4 at 37°C). The type 1 collagenase assay kit (B-Bridge international, Inc. USA) was used to detect collagenase concentration in prepared solutions. In order to prepare the substrate solution for the enzymatic reaction, 1 mL of fluorescent-labelled collagen and buffer A (supplied in collagenase kit) were mixed and kept on ice until needed. One hundred microlitres of substrate solution and collagenase enzyme solutions with different concentrations were added into a

microtube and mixed thoroughly. Samples were then incubated at 35°C for 2 h. Six hundred microlitres of cooled buffer B (supplied with the collagenase kit) was added to each tube and kept in ice for 15 min. All tubes were centrifuged at 10,000 rpm for 10 min. The fluorescence was measured from the supernatant (the liquid lying above a solid residue after centrifuging). In order to perform measurements using the smartphone system, solutions of trypsin and collagenase assays were added into 8 well chamber slides (BD Biosciences). Each well was a cube with a 9 mm size (Figure 25). Two hundred and fifty microlitres of solution was placed in each well and this yielded 1.4 mm of the sample height.

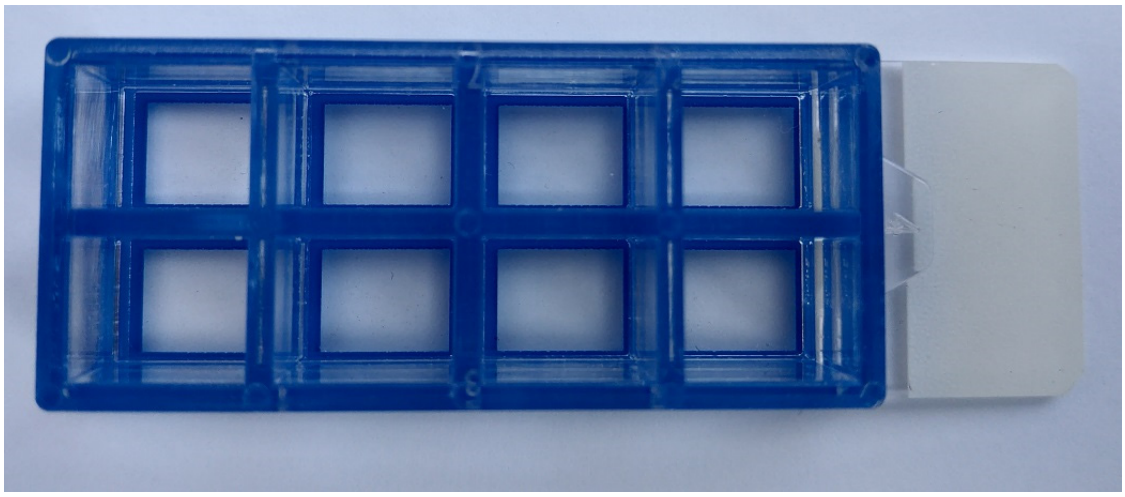


Figure 25. Well slide used for measurement

2.8.2. The tablet

The tablet used in our device was an Asus Nexus 7 (2013) with full HD (1200 × 1920) resolution and an IPS-LCD screen. Its brightness at its maximum white output is $583\ \mathrm{cd/m^2}$ (81). The tablet was set to display a single square on an otherwise black background. The square was slightly larger in size than the single well being read out, with a side 13 mm long (this is 165 px given the tablet specification of 323 PPI). This square was blue, with an RGB value of (0;0;255). Since LCD displays work based on emission through three different coloured filters, they emit light in three primary colours in the RGB system, where each of the colours has its own emission spectrum. A full range of colours is achieved using mixture of the three emitters with different intensities. When displaying one of the primary colours, we are expecting only one

emission peak. We measured the emission of the previously mentioned, blue-only square with the spectrophotometer. This emission is presented in Figure 26.

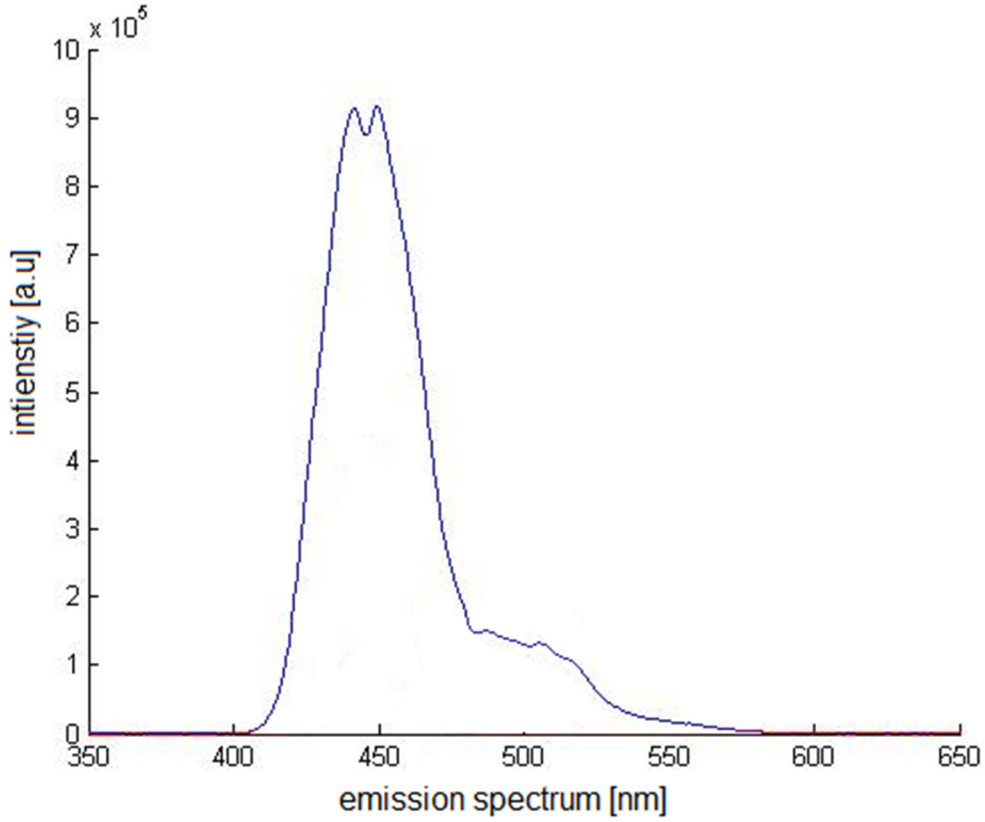


Figure 26. Emission of the blue only square presented on the screen of the tablet. The dip on top of the peak is due to the sensor saturation.

The emission spectrum of the blue pixel of the tablet matches the excitation spectrum of FITC very well covering the spectrum between 400 nm and 520 nm. The spectrum is almost identical in terms of wavelengths, with slightly shifted peak value (Figure 24 and Figure 26). High brightness of the tablet screen (100%) was used in the further studies, to induce the most FITC emission.

2.8.3. The Polariser

A polariser is an essential part of our device. The polariser used for this work was a Moxtek PFU04C wire grid polarizer, with size of 12.5 mm × 12.5 mm. This wire grid polariser is characterised by high contrast, large acceptance angle, and broadband performance. Due to this construction it has excellent durability and long lifetime, and its crossed transmittance is low (0% to 0.5%). The polarizer was placed on the top of

the well with the sample, to prevent blue excitation light reaching the camera and interfering with the readout. Since the tablet screen was already polarized, only one polarizer was used in this instance. In order to verify the effectiveness of this polariser in blocking the excitation light, the following experiment was conducted: The blue square on the tablet screen at its full intensity of 255 was covered with the polarizer at 90 degrees relative to the direction of screen polarisation (Figure 27). An image was taken using the phone and the signals in the green and blue readout channels were measured. In this case, we registered the mean green and blue pixel values of respectively 7 and 9, which was about double the background noise value of 4. Without the polariser the signal was just saturated, at 255. The difference of 3/255 levels bleeding into the green readout channel was smaller than 2% of the full signal. The difference of 5/255 into the blue channel readout was also slightly smaller than 2%. The polariser was indeed effective to block excitation light. We note that the final background noise measured with neutral fluid in the well in place was even lower, as discussed later in the text.

Figure 27. The polarizer, crossed with the screen polarization, covering the blue square displayed on the tablet that is an excitation source for FTIC. Surrounding colour is a black on 100% screen brightness. A small blue dot is visible where the polarizer was damaged.

2.8.4. The box

The box, even though obviously the simplest part of the setup, still served a couple of important purposes. Most importantly, the box was protecting the phone camera against ambient light. Secondly, the box supported the phone directly above the sample and set the distance between the phone and the sample (placed on the

tablet). The height, allowing the phone to properly focus on the sample, was set to 7 cm.

2.8.5. The phone

In this work, a HTC EVO 3D (X515) cell phone was used to capture assay images with its built-in camera. The camera settings need to be controlled by the user to a sufficient degree, and to ensure repeatable readout pixel intensities the settings for the ISO value and the white balance were chosen manually. To determine the best white balance option, an experiment was conducted where a green square (0;255;0 in the RGB colour model) was displayed on the tablet screen, covered with the box and images were taken using the phone with different settings.

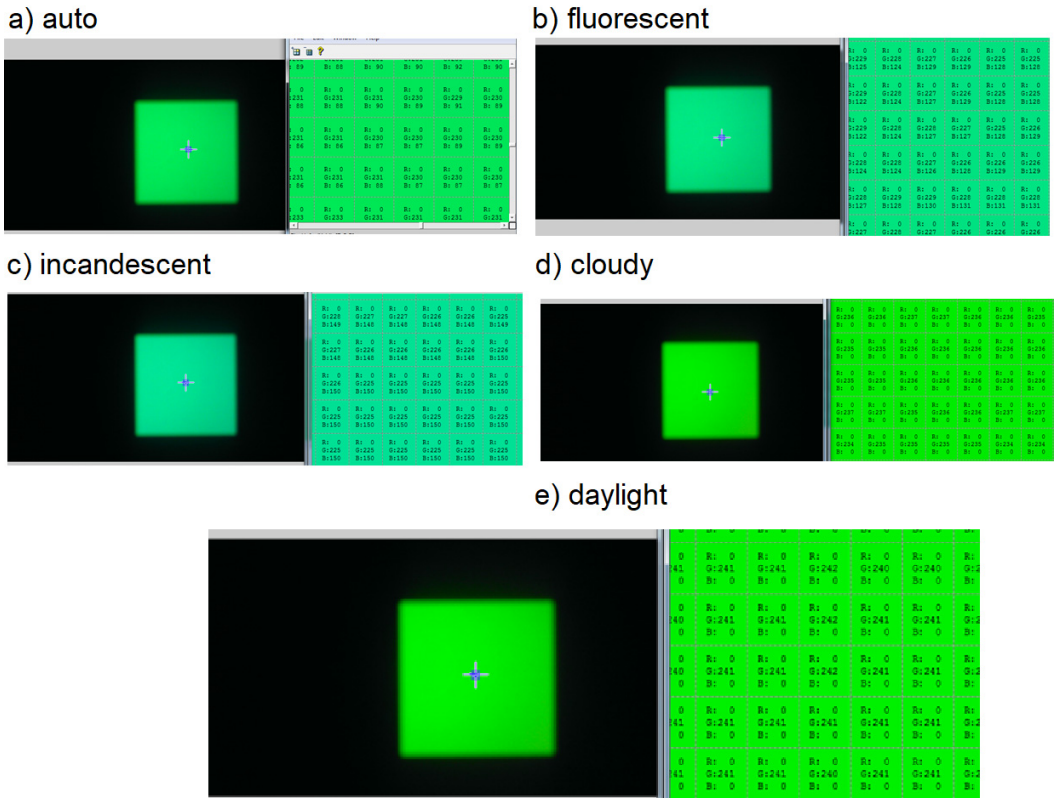


Figure 28. RGB values for images taken with phone with different white balance options, of green square displayed on tablet screen.

We looked for the settings to provide the highest green pixel value and lowest red and blue pixel values: Since we know that pixel readout in the sensor is based on physical RGB filters, the setting that gives the simplest output is most likely the one that is the

least processed by photo enhancing algorithms, and therefore the least distorted to provide an accurate readout. These conditions were best met by the 'daylight' option. Photos for different named white balance settings are shown in Figure 28, with Figure 29 showing a cross section through the 'daylight' image in Figure 28e, showing the three colour channels separately.

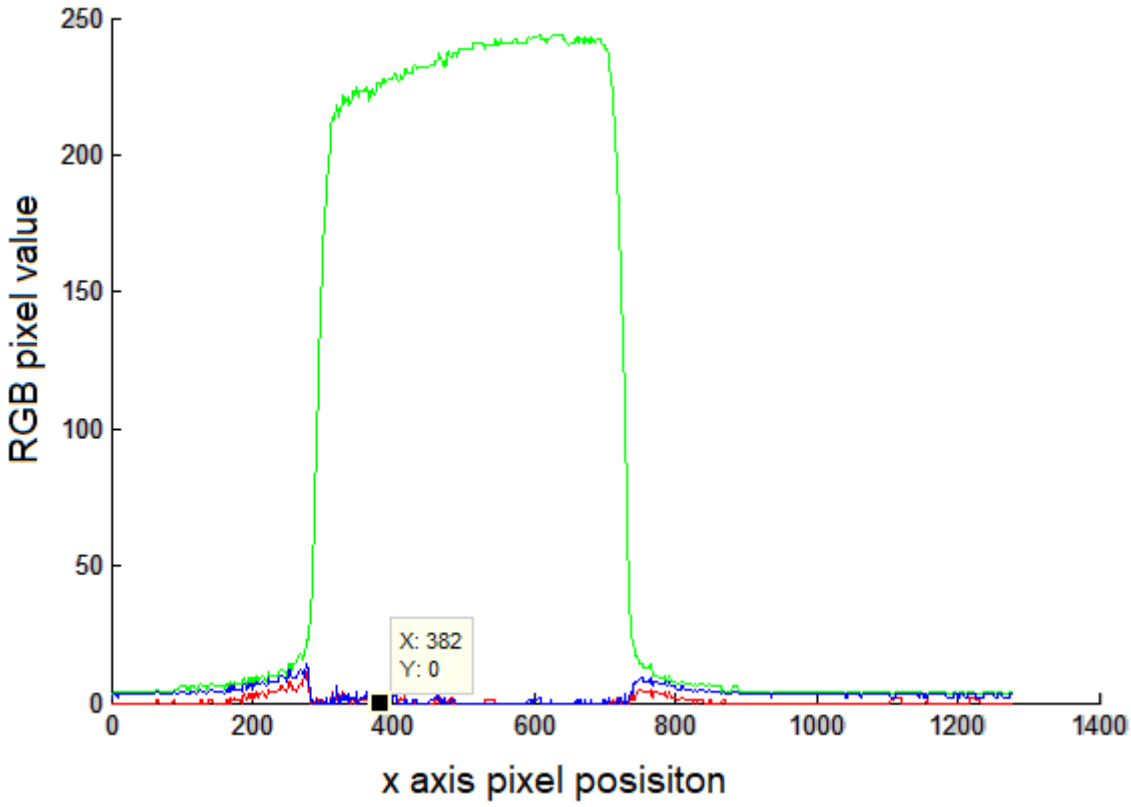


Figure 29. Cross-section through the green square displayed on the tablet screen for the optimal 'daylight' white balance option, showing the three colour channels separately.

Figure 29 shows that indeed the green channel readout is close to maximum 255 counts, while both other channels have values of zero or close to zero. The slight increase of all the channels close to the border is a result of slight bleeding of neighbouring pixels and imperfection in focus of the image. Non-zero values of the black areas displayed around the green square are a consequence of the noise of the camera and the fact that in contrast to AMOLED displays, LCD screens cannot completely turn off the black pixels, and thus some light is bleeding through them. It appears that the bleeding is stopped when the strong green light is emitted through them on purpose.

The settings for both ISO and image resolution were chosen to be the maximum available. An ISO value of 800 provided images with the highest amount of signal, without introducing significant noise that would interrupt measurements. The image resolution of 5 Mpx was native to the camera, and choosing other resolution would be just downscaling of the saved image rather than increasing photon detection capabilities.

All the options that could be manually set are presented in Table 2.

Table 2. Manual camera settings for image collection of the used HTC Evo3D smartphone.

Phone camera setting	option
ISO	800
White balance	daylight
Image resolution	5 Mpx
Flash	off
Automatic correction	off
Automatic sharpening	off

2.9. Initial system testing and calibrating and rejected ideas

Once the parameters were determined, testing of initial ideas and basic characteristics began. Here we will describe the initial concept verification, as well as some ideas that finally were replaced with other solutions.

2.9.1. Camera distance and focus

The standard configuration for a smartphone camera is focus at infinity. This design assures proper focusing on objects at different distances (with the help of active focusing), above a certain minimum. This 'minimum focus distance' in our case is about 7 cm.

Experiments using smaller distances were conducted while searching for the simplest setup. With a smaller distance, the sample will be out of focus, but given that we don't need to spatially resolve within a sample but rather just tell one sample from its neighbour, this is not necessarily a problem. First, the smallest possible distance was tested, using the smartphone directly lying on the tested object (which at the time

was a green square displayed on the tablet screen), with two images shown in Figure 30. The difference between the two displayed intensities of the square was clearly recognisable, and most of the internal surface of the shape had uniform texture that could be used as an area of analysis. However, this idea was abandoned for other practical reasons; it was too easy to break the polarizer or violate the sample with this contact method.

The next attempt was taken from the distance of about 2 cm, where the whole square of the final size was visible on the picture, shown in Figure 28. This idea was replaced with a proper focus distance, since once the phone needed to be separately supported away from the sample, there was no particular advantage in simplicity using 2 cm rather than the in-focus distance of 7 cm.

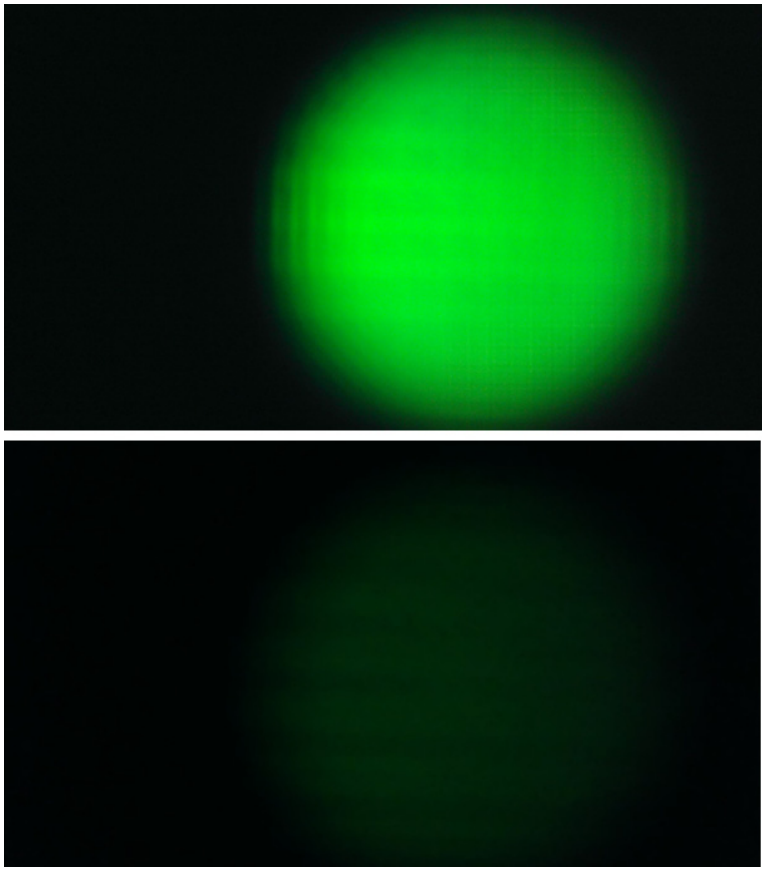


Figure 30. Images of the green square taken with the phone in contact with the screen displaying the square. Upper image: 100% brightness, lower: 10% brightness of the screen.

2.9.2. Quantification method testing

Before starting measurements on actual biological samples, some tests of the system behaviour and readout algorithm were conducted. The first most important question

was whether the system will be able to reliably read out a simulated signal intensity change and if this can be within biomedical standards of signal to noise ratio (SNR).

The following experiment was conducted. A green square of the size of the well was displayed at the tablet screen, to simulate the fluorescence from an assay. The intensity of illumination of the screen (brightness) was gradated from 100% to 0% (there was still a discernible colour displayed at a setting of 0%) in steps of 10, and the green channel response of the smartphone camera was recorded. The square was covered with a different polariser (with higher transmittance but more resistant to damage) to decrease signal intensity to test the system, but also to test the SNR determinant of the readout. The signal strength here was about twice stronger than the medical samples presented later, and so in the relevant range. The results are presented in Figure 31.

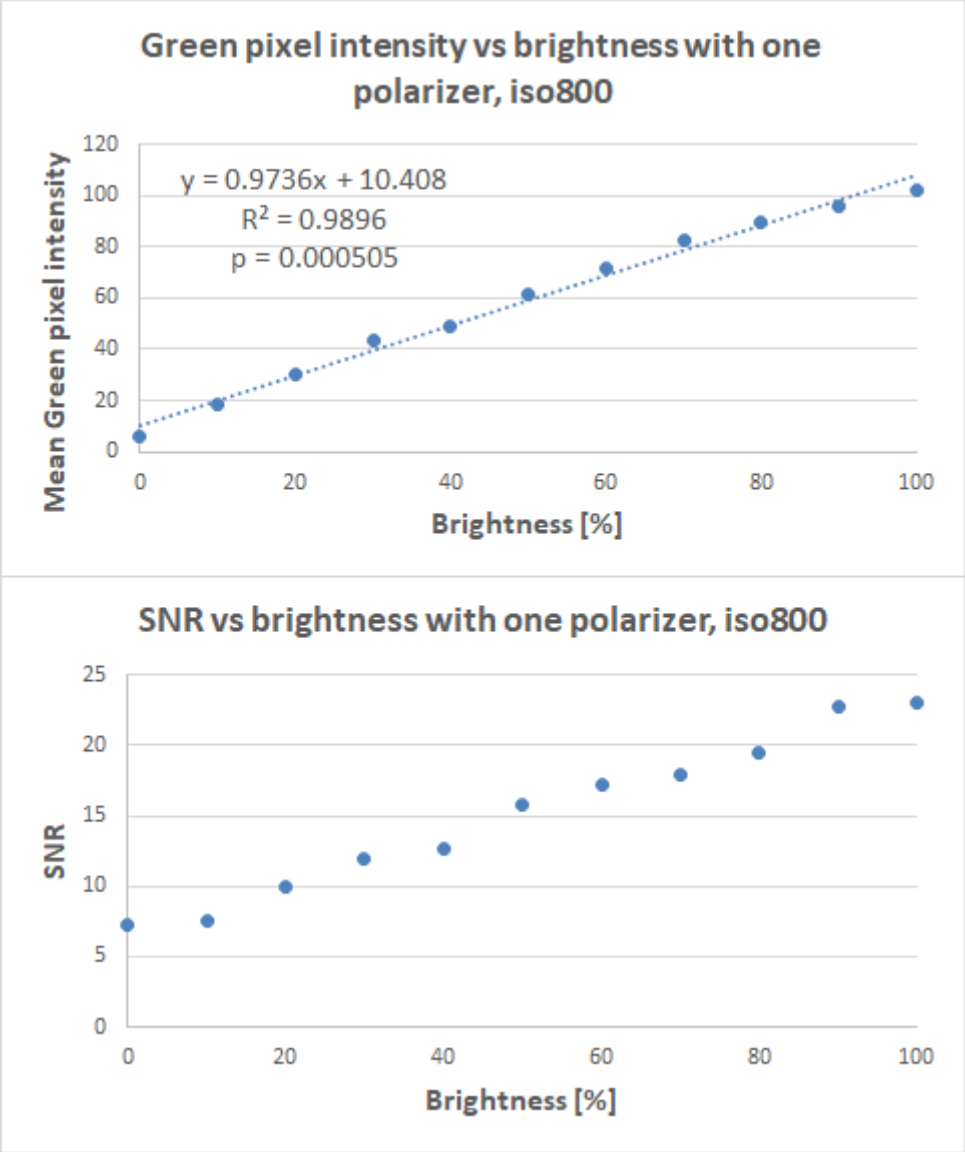


Figure 31. Test setup for green pixel intensity changes with respective SNR calculation.

As shown in Figure 31, the camera was able to detect changes in pixel intensity with almost linear response. The signal to noise ratio was defined here as the mean readout signal value divided by the standard deviation of measured pixels, and was testing low light level readout distortions introduced by the sensor. The SNR decreased with decreasing pixel intensity with a lowest value of 7. These results were very satisfactory, and led to further investigation.

In the early versions of the algorithm, the signal level was acquired from the cross section as presented in Figure 32. This method allowed us to measure the noise level on the black part of the screen, and average over part of the signal area of the image.

This method was replaced with analysis of a square area, with the greater number of pixels in the averaging assuring statistically more accurate measurement. We also used an independent background measurement, measured using a blank sample (well filled with solvent) as a better estimation (one including the bottom of the well and neutral in colour fluid) of the background in the area of analysis.

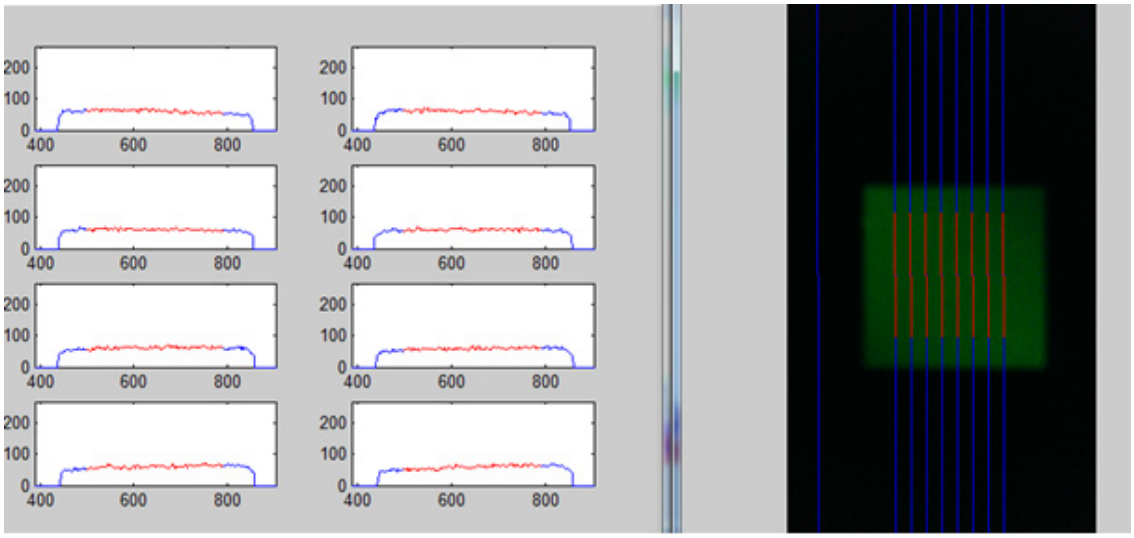


Figure 32. Image presenting cross sections going through the sample. Parts marked in red were analysed as the signal part. Noise from this part was used to calculate SNR. An additional cross section next to the sample was tested to observe noise behaviour.

The last feature that was considered was automatic sample area detection. If the excitation square was set to be larger than the polarizer size, it created a blue border around the well (Figure 33). This border was attempted to be automatically detected and treated as the cropping square for the Area of Analysis (AOA) determination. The method however, proved to be imperfect, especially in cases of imperfect sample distribution and polarizer positioning. This idea was later replaced with a manual AOA selection for more control.

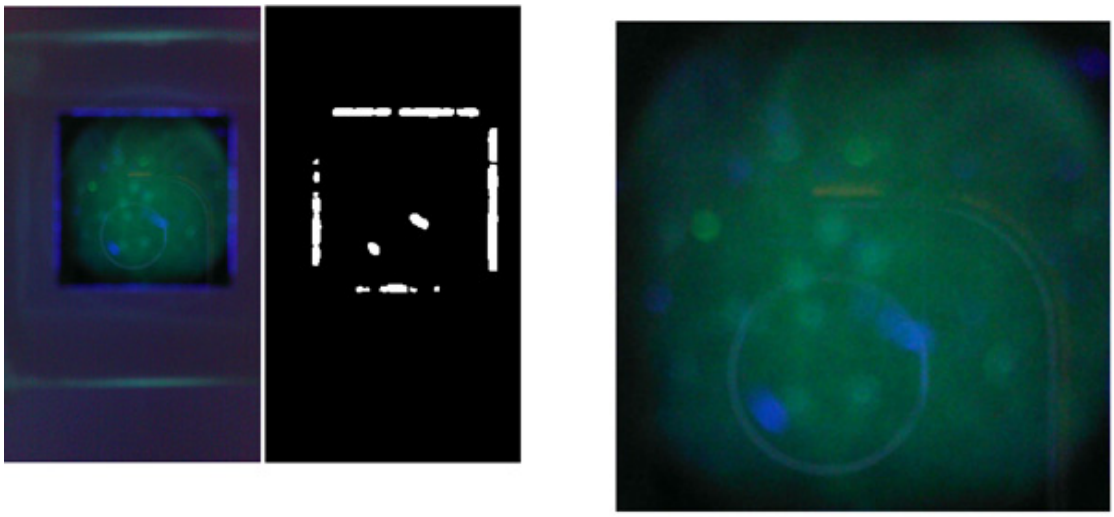


Figure 33. Algorithm of automatic sample detection based on the blue border around the sample. On the left, original sample image, in the middle mask created with detection of the blue elements, on the right: image cropped with the mask.

2.10. Taking a measurement and calculating the result

In Figure 34 we present the block diagram of the protocol for taking measurements. We previously described sample preparation steps. The data collection part including the characteristics of each element was also already presented.

In a single well-slide (Figure 25) all needed concentrations plus an additional neutral sample for the background signal can be prepared. Each sample was imaged two times, assembling the whole setup from the scratch to give an estimate of sensitivity to user - variables such as improper sample position on the screen, imperfect crossing of polarisers. Subsequently, images were transferred to the Matlab script for data processing. Each image was sampled three times with a reselected AOA window. The final result is an average value of three measurements as described.

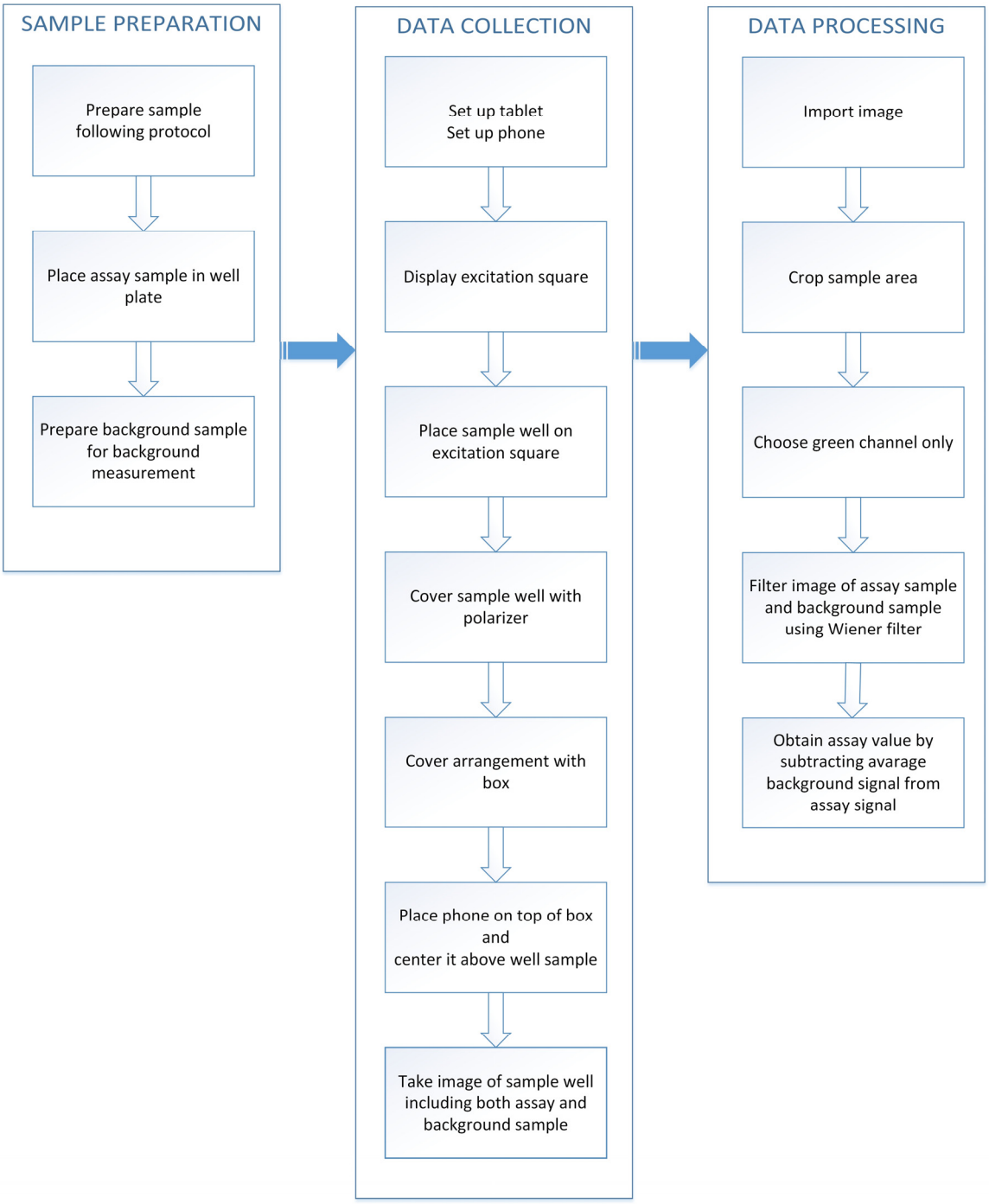


Figure 34. The assay workflow.

Accurate quantification of the assay signal, especially at low analyte levels, was found to require some additional image processing, which was performed in the following way. The area of analysis (AOA) was determined from the original image. The AOA used here was a 200 px × 220 px rectangle, cropped from the signal part of the image (Figure 35a). Subsequently, the resulting image was split into 3 colour channels – red,

green and blue. Since the signal emission for fluorescein is green, neither the blue nor red channels contain useful response for our assays. The green channel was selected for further analysis, and Figure 35b shows the intensity of green channel as a grayscale image. In the next step, adaptive filtering (82) was used to improve image quality, reduce noise and smoothing any artefacts appearing for instance because of the dust on the polarizer. (Figure 35c).

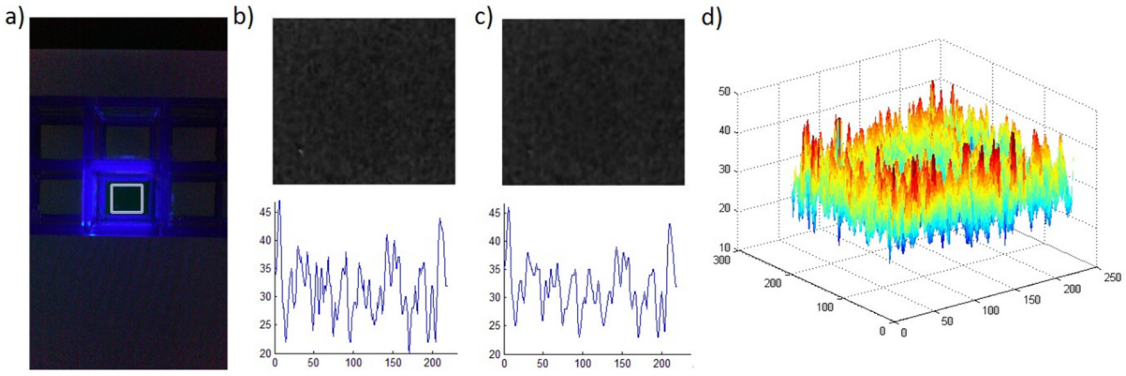


Figure 35. Image processing steps. (a) Initial image with the AOA rectangle highlighted; (b) Green channel image (in grayscale) and a representative cross-section of screen intensity showing noise amplitude; (c) Image filtered with adaptive Wiener filter and a representative cross-section of intensity showing noise amplitude; (d) Pixel intensity map of the AOA after adaptive filtering.

The adaptive filtering used was a Wiener filter, which estimates the local mean and variance around each pixel, as follows:

$$\mu = \frac{1}{NM}\sum_{n_1,n_2\in\eta} a(n_1,n_2) \qquad (36)$$

and

$$\sigma^2 = \frac{1}{NM}\sum_{n_1,n_2\in\eta} a^2(n_1,n_2) - \mu^2 \qquad (37)$$

where η is the NxM local neighbourhood of each pixel in the image a. We used $N = M = 9$. The Wiener filter then creates pixelwise estimates:

$$b(n_1,n_2) = \mu + \frac{\sigma^2 - {\gamma_n}^2}{\sigma^2}(a(n_1,n_2) - \mu) \qquad (38)$$

where ${\gamma_n}^2$ is the noise variance. The noise variance is the average of all the local estimated variances. Figure 35d shows a typical AOA pixel intensity after Wiener filtering. The results were considered reliable when approximately uniform intensity

distribution across the whole AOA was observed without noticeable outlying peaks. The filtering decreased STD by almost 20%.

The final algorithm step was added to cap individual pixel values at a level three times the mean signal value, to prevent effects such from hot pixels. This level was chosen after studying the occurrence of such hot pixels during the experiment.

The assay signals were produced in the following way from the green pixel values. The average pixel intensity and standard deviation values were calculated from the assay AOA after adaptive filtering. We then established the background by imaging the assay well filled with buffer only without any fluorescent substances, again using an appropriate AOA and adaptive filtering. The assay signal is the average pixel intensity of the assay AOA minus the average pixel intensity of the background AOA. The SNR was calculated for each sample as an average filtered pixel intensity of six measurements (three of each image) with subtracted background noise and divided by the standard deviation of mean values of the images for given sample. The low detection limit was considered to be achieved when the signal to noise ratio had the value of three.

2.11. Results and discussion

Several different sets of samples were carefully prepared and examined using the device described here, and also with a specialised fluorometer to compare our results to the standard method. The results for trypsin and collagenase are presented in Figure 36. Figure 36a shows the results of trypsin assay measured on our smartphone system, while Figure 36b shows the same trypsin assay samples examined by standard fluorometry. Another sample set with lower concentrations of trypsin were examined in order to determine the low detection limit; these results are presented in Figure 36c. (Error bars on data points appearing in both this figure and Figure 36a are the same, but they appear larger than in Figure 36a due to a different scale on the vertical axis.) Figure 36d shows the results for the same low trypsin density sample characterised by standard fluorometry.

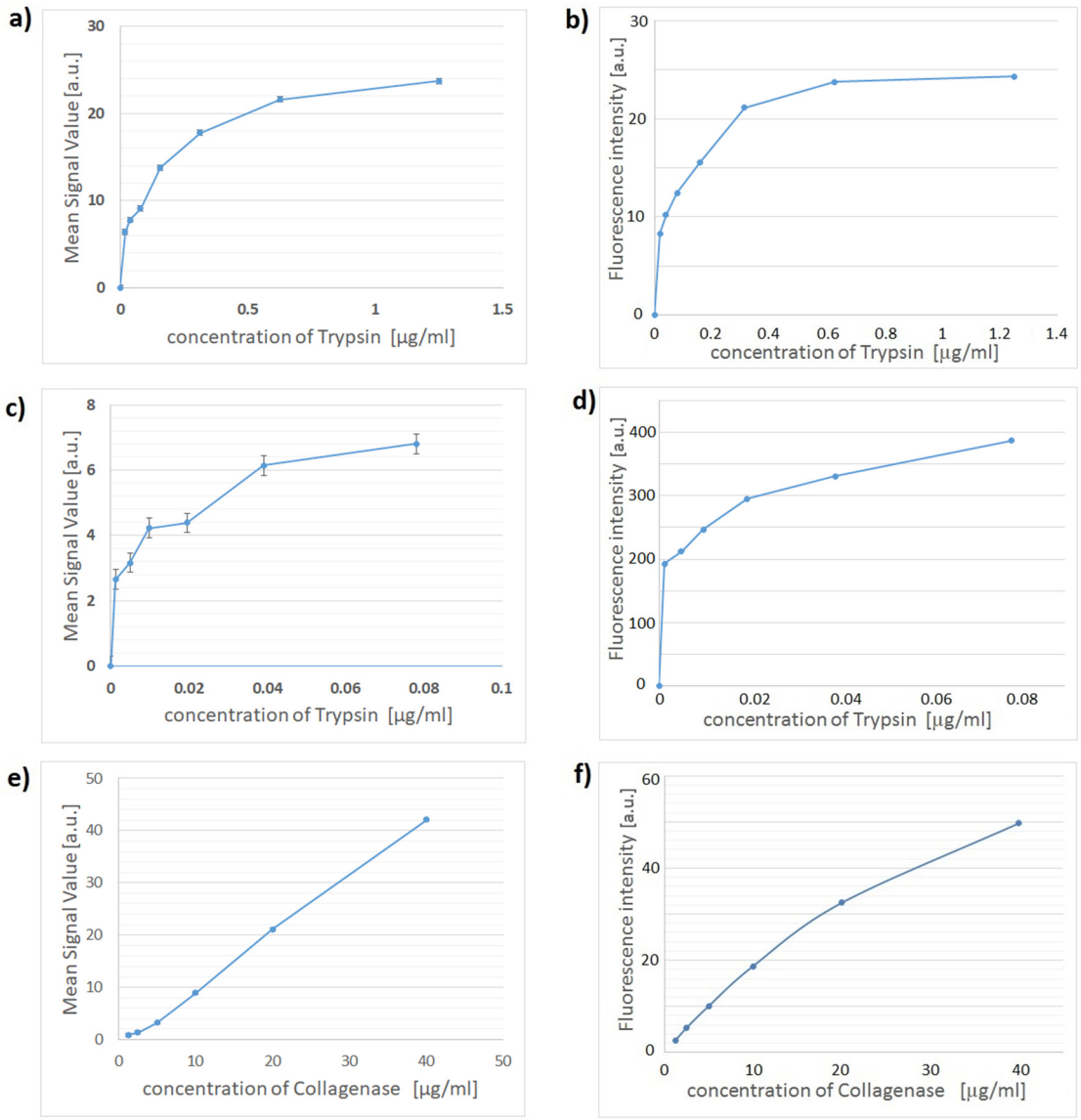


Figure 36. Assay signal as a function of analyte concentration for trypsin and collagenase assays. Except panel c, error bars are too small to be visible (a) Trypsin assay with smartphone device; (b) Trypsin assay with Cary Eclipse readout with photomultiplier detector voltage option set to 'Low'. Uncertainty of each data point is 0.1; (c) Trypsin assay with smartphone device at low concentrations; (d) Trypsin assay at low concentrations with Cary Eclipse readout with photomultiplier detector voltage option set to 'Medium'. Uncertainty of each data point is 2.7; (e) Collagenase assay with smartphone device; (f) Collagenase assay with Cary Eclipse readout with photomultiplier detector voltage option set to 'Low'. Uncertainty of each data point is 1.1.

Figure 36c, presents the lowest concentration measured with the smartphone - 1.22 ng/ml. The measured background noise of 4 was subtracted from the signal, setting the lowest measured concentration to the value of 2.65 a.u. The standard deviation between measurements, presented as error bars was 0.3 a.u, which sets the value for noise to 0.6 (doubled STD). The Signal to Noise Ratio for this sample was 4.4, which is close to the limit of 3. The total quantity of detected trypsin was 500 pg

within a 250 µL sample. This value is comparable to the low detection limit of 400 pg suggested in the manufacturer's manual for the assay, which is based on the use of a microplate reader. Thus, the presented device has been able to measure low concentrations of trypsin in combination with fluorescein as a fluorescent reporter.

The concentration resolution for this measurement based on the measurement errors was 0.14 ng/ml, which makes the sensitivity of the device smaller than the sensitivity of the Cary Eclipse, which in general is a high-sensitivity spectrophotometer. This resolution however, is more than sufficient for the medical use.

The examination of trypsin level in blood is a standard procedure for newborns, because increased trypsin level may indicate cystic fibrosis (83, 84). The typical level of trypsin for a healthy child below one year of age is below 200 ng/mL (83), while a trypsin level in a range of 200–1000 ng/mL is an indication to carry out additional tests. The detection range of interest for cystic fibrosis application is, therefore, between 200 ng/mL and 1800 ng/mL (83). With our low detection limit of 1.22 ng/mL the smartphone device is over 100 times more sensitive than what is required for the detection of cystic fibrosis.

We now turn to the colleganenase results in Figure 36e. By using our device, we were able to achieve the low detection limit for collagenase of 1.25 µg/mL with a SNR = 3 and standard deviation 0.23 µg/mL. This limit of 1.25 µg/mL corresponds to 0.31 µg in total in the sample. The clinically relevant concentration range of collagenase in synovial fluid is between 1.6 - 11.7 µg/mL in patients with rheumatoid arthritis and patients with different grades of joint inflammation (85, 86). The sensitivity based the error for this assay is 0.31 µg/mL. We see then that our low-detection limit concentration of 1.25 µg/mL is just below the range of clinically relevant values. No signal or our minimum signal suggests lack of the medical condition; anything higher can be accurately quantified, allowing for instance for the differentiation between inflammation class (85).

We emphasise that the device presented here is largely technology platform-independent. Alternative cell phone models with a different camera can also be used in this device, provided the camera can be prevented from making automatic

adjustments of the relevant settings. For example, Apple iPhones and Android devices (e.g., Samsung Galaxy series) are suitable replacements. Other types of mobile devices can also be used as the light source in our setup, for example Apple iPads, Samsung Galaxy Tab family, or even various models of cell phones. The key parameters determining their suitability are screen brightness, colour projection (gamut) and polarization of the screen. Many commercial devices are compared in relation to these parameters on specialized Internet websites (81). In alternative devices with different screen resolution or screen size, the size of the excitation square should be adjusted. The screens in some mobile devices (such as iPads) produce light polarized at a 45-degree angle, so the polarizer needs to be set properly for the excitation light to be fully extinguished. If the screen is not polarized, or insufficiently polarized, then two polarizers should be used.

We also emphasise that the system described here will tolerate various modifications (e.g., excitation light wavelength, readout channel) without affecting core functionality, and it can be can adapted to become a universally applicable fluorescence detection device. Due to this flexibility, the device will also be able to be used with alternative fluorescent assays and other fluorophores.

With increasing availability and wide adoption of smartphone technology, including in the developing world, come new opportunities for their application in biomedical diagnostics, with multiple authors reporting being able to use smartphone-based systems for medically relevant assays (57, 87-90). However, in many cases, this requires fairly complex technologies in addition to a phone. One of the limitations is the sample preparation procedure required for specific assays, which usually is managed by a specialised add-on “dongle” (57, 88) whose complexity depends on the assay reporting scheme. For example, in (57), a silver enhancement step, following binding of gold-labelled antibodies, required complex fluidics. The complexity of fluidics grows with multiple requirements, especially if an analyte amplification step is required (88). The readout itself is typically carried out by a combination of standard optical components, such as excitation lasers, filters, and lenses (87), or optical filters and optical gratings (89) with the phone camera only performing the function of a detection device. In this context, the presented system for fluorescent assays sits at

the lower end of the scale for complexity and cost. It uses the RGB channel split and a polariser instead of optical filters and a screen as the light source. Hence, in the present system, all required system engineering is contained in its software, which is freely available at no cost.

2.12. The computer application

An important outcome from this project is a standalone graphic-user interface driven application for our system with all the required functionality to carry out assays. This can be downloaded from the website http://cnbp.org.au/smartphone_biosensing. A screen-grab from the application is shown in Figure 37. The application carries out all the steps mentioned in the text, including calibration and background measurements, and sample images used in this study are also supplied. With this software, any phone or tablet can be tried to measure any fluorescein isothiocyanate (FITC) assay, so other researchers can very easily employ and/or expand upon our work here.

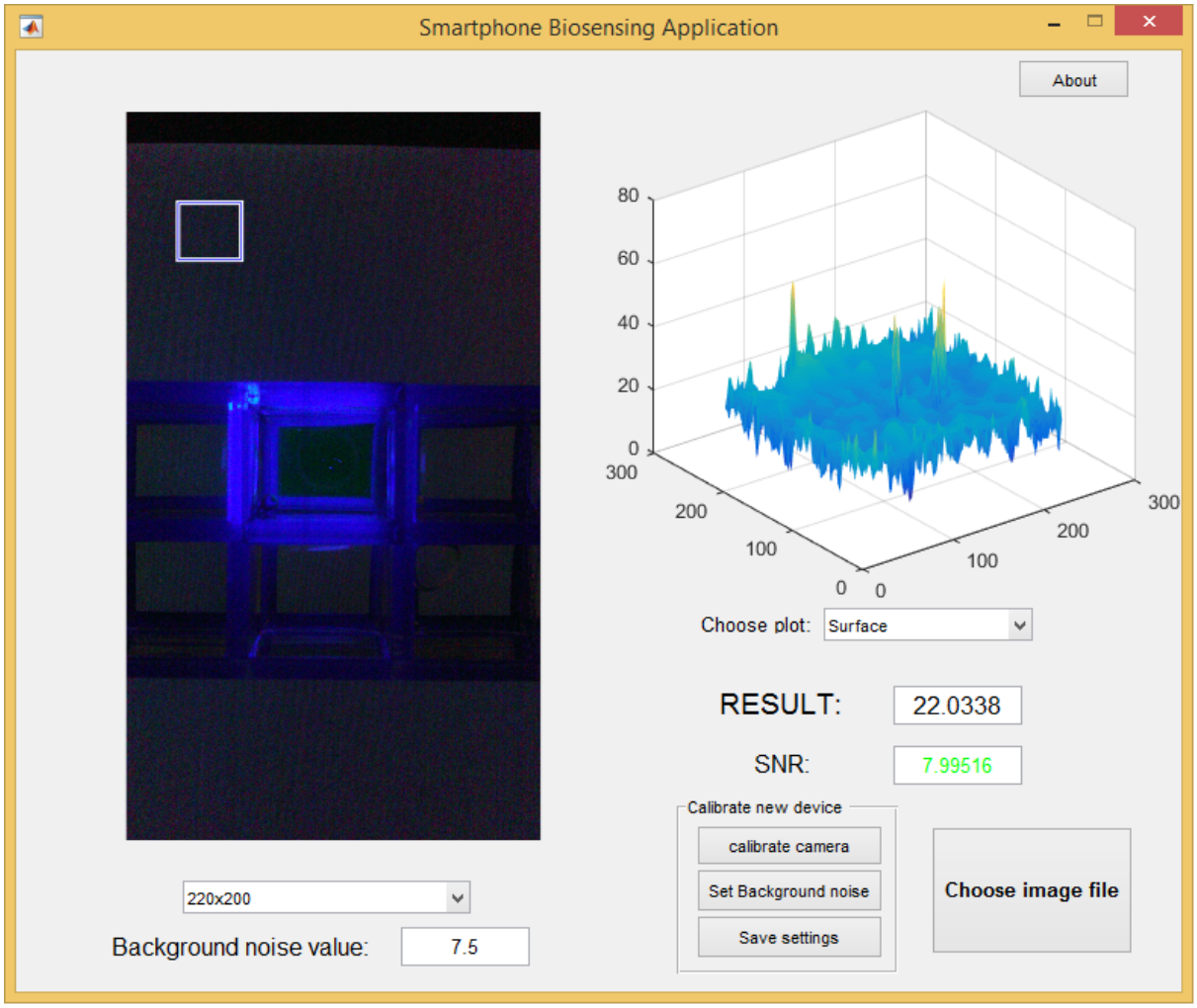


Figure 37. Main window of the application, presenting the image sampling area, showing the 3D graph to check the usability of the image, with the final result presented and SNR coloured in green to highlight the readout correctness.

2.13. Assessing respiratory inflammation using smartphone imaging: a proof of concept study - the concept

In the previous section we demonstrated that even without specialised attachments, smartphones can be used to measure fluorescent signals such as those from chemical assays in solution. In this section we take this idea further and apply the smartphone technology to the area of colorimetric assays for disease diagnostics and monitoring. Instead of using a specific light source and isolating the assay from ambient light as we did in the previous research, we decided here to use ambient lighting, and to utilise smartphone automatic image adjustments and additional colour normalisation to achieve consistent results. Instead of avoiding external illumination by adding the complexity of an in-system light source, we actively use external illumination and measure and compensate for it. Use of specific standards enabled comparison of assay samples between different photographs, captured under varying illumination conditions.

Traditionally, colorimetric assays are quantified using direct measurement of transmittance (or absorbance) at a given wavelength. We instead assessed assay concentrations based on the normalised colour of the assay sample. We used a white sheet of paper as a colour reference point and a system of software tests to indicate to the user whether images have been taken under suitable conditions to assure reliable results.

Here we focus on assays for the cytokine IL-8 and the protease enzyme neutrophil elastase (NE), which both serve as reliable markers of inflammation in patients with Cystic Fybrosis. Experiments were conducted on clinical samples in cooperation with Telethon Kids Institute in Perth. Two popular smartphones were used as examples.

2.14. The Setup

The idea behind this study was to test how well a phone can perform with real medical samples of a colorimetric assay, without needing to build a physical setup nor use a professional, expensive colour calibration device like Macbeth ColorChecker. A white sheet of paper was used as a background for the photographs, serving as a reference

point for colour correction. Samples were placed in standard 96-well plates. Then, images were taken with the phone, ensuring that samples to be measured were close to the middle of the image and that the white sheet of paper was visible somewhere and not obscured by the plate. An example of the readout procedure is presented in Figure 38.

All measurements of the medical samples were taken opportunistically during regular tests routinely performed as part of the existing AREST CF early surveillance program. Time sensitive samples were imaged between other measurements. These data collection conditions did not allow us to take large numbers of images nor for example vary lighting parameters, and some data points had image issues limiting the validity of the results. These readout imperfections will be explained in the text.

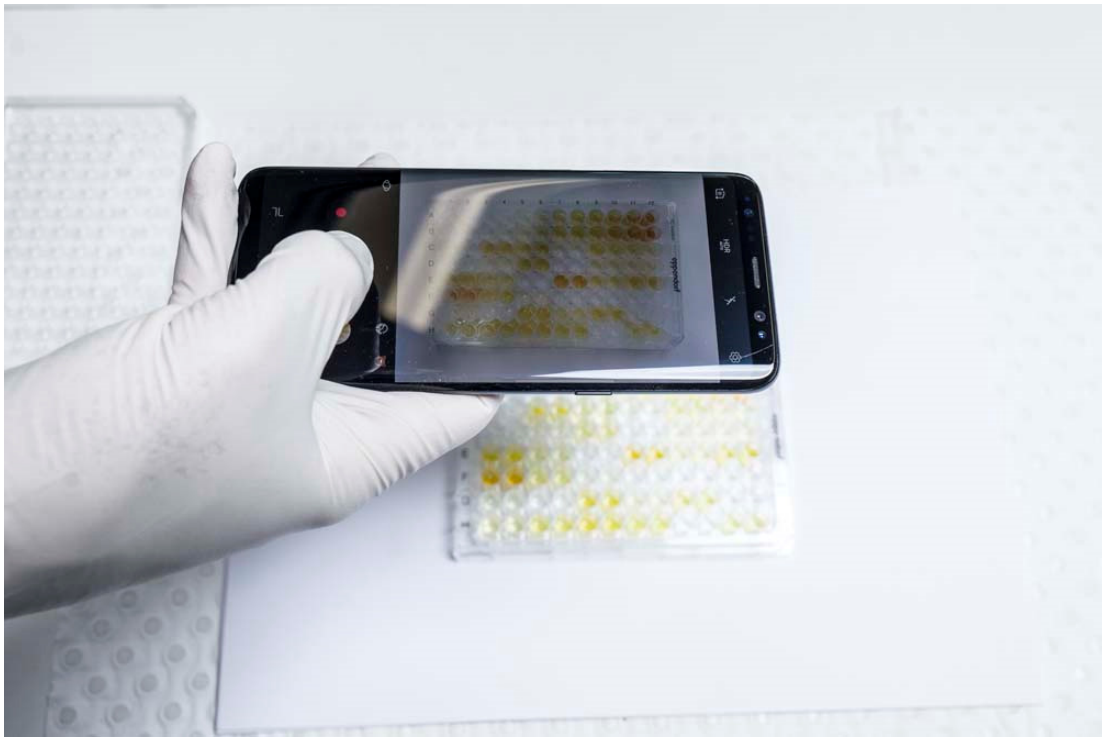


Figure 38. Taking an image of the IL-8 sample. The sample to be measured is about in the middle of the image, lighting is approximately regular, and the white sheet of paper is visible around the sides of the well plate.

2.14.1.Samples

Children with CF attending Princess Margaret Hospital for Children (Subiaco, Western Australia) and Royal Children's Hospital (Parkville, Victoria) were recruited to the Australian Respiratory Early Surveillance Team for Cystic Fibrosis (AREST CF) early

surveillance program after signed informed consent for the use of samples in research studies was obtained from a parent or legal guardians. The AREST CF early surveillance program has been approved by the research ethics committees at each of the participating institutions (Melbourne HREC25054; Perth 1762/EP). In the main experiment bronchoalveolar lavage fluid (BALF) samples were tested from 32 children with CF, prepared by instilling and aspirating normal saline into the right middle lobe under general anaesthesia, during their clinically directed annual surveillance visit (91, 92). The second and third aspirates were pooled and centrifuged at 330 g for 5 minutes at room temperature. The resulting supernatant was aliquoted and immediately stored at -80 °C.

Commercial assays for both IL-8 and NE were performed, and quantified in the standard way using a measurement of the absorbance at a given wavelength. The absorbance readings were performed on assay plates by a Multiskan FC (Thermo Fisher Scientific, Waltham, MA, USA) microplate spectrophotometer.

IL-8 levels were measured using a commercial assay kit (BD OptEIA #555244, BD Biosciences, San Diego, CA, USA), performed in accordance with manufacturer's protocol, including the addition of 50 µL 2 N H_2SO_4 to stop the reaction before measuring absorbance at 450 nm. The levels of NE were determined as previously described (93). Briefly, 50 µL of sample BALF was applied in duplicate to a 96-well plate. Added to this was 5 µL of a solution containing 1 mM n-methoxysuccinyl-ala-ala-pro-val p-nitroanilide (Sigma Aldrich Australia, Castle Hill, NSW, Australia) in 0.2 M Tris pH 8.0. After 40 minutes incubation at 37°C, the absorbance at 405 nm was measured. Standard curves were prepared using the same approach, by using solutions with known concentrations. Both IL8 and NE assays were performed in Costar half-area 96 well plates (Corning Inc., New York, NY, USA). A new standard curve sample set was prepared for each new batch of biological samples for both assays.

To compare to this standard approach, we used the smartphone to assess changes in the colour intensity as an alternate indicator of the biomarker concentration. In both assays, the colour changes from very faint to more vivid yellow, with greater range for the IL-8 assay. The IL-8 assay is stated by the manufacturer to be limited to a

maximum detectible concentration of 0.2 ng/mL, which is insufficient for medically relevant levels for Cystic Fibrosis evaluation. In our study, we prepared reference concentrations much higher than this stated maximum, and measured them with our system and with the standard absorption measurement method. We found that the smartphone colour method could give reliable results far beyond the manufacturer's recommendations using the standard method.

2.14.2.The white sheet of paper

The white sheet of paper served two major purposes in this study. First, since the well slide with samples is transparent, it provides a uniform background for the measurement. Secondly, it provides a colour reference point for the image normalisation to allow comparison of colour intensity between different images. By using a piece of paper as a white reference, working curves prepared in one set of lighting condition can still be reliably used to interpret images of real samples taken under different lighting conditions, by comparing the white reference in each of the photos.

Ideally grey colours (including white) are characterised by all RGB channel values to be equal (e.g., R=200, G=200, B=200). Not all the papers have the same colour though. The colour of the paper can be described with three parameters: whiteness, brightness and shade. The CIE measure of whiteness is a measurement of the light reflected across the visible spectrum. For a perfect reflecting non-fluorescent white material, the CIE L*a*b Lightness (L) variable has a value of 100. Many papers though, have additions of Optical Brightening Agents (OBA) which are designed to absorb and refluoresce the light from the non-visible range (mainly ultra-violet) back in the visible spectrum. These papers might visually appear brighter, 'more white' for a human eye, while in reality they will be reflecting/emitting more in the blue spectrum, deviating from the reflection balance of white/grey colours. Brightness is a parameter that describes the ratio between the amount of illumination and emission light of the paper in the visible spectrum. By TAPPI standard (94), equal amount of light would have a value of a 100, but because of OBAs, similarly to the whiteness, this indicator tends to be slightly higher at about 110-120. Paper shade is again a measure of reflectivity of the paper, but this time is not represented as a number but rather by

the actual spectrum. The spectrum of a white/grey paper sheet will be about equal at all wavelengths as presented on Figure 39a.

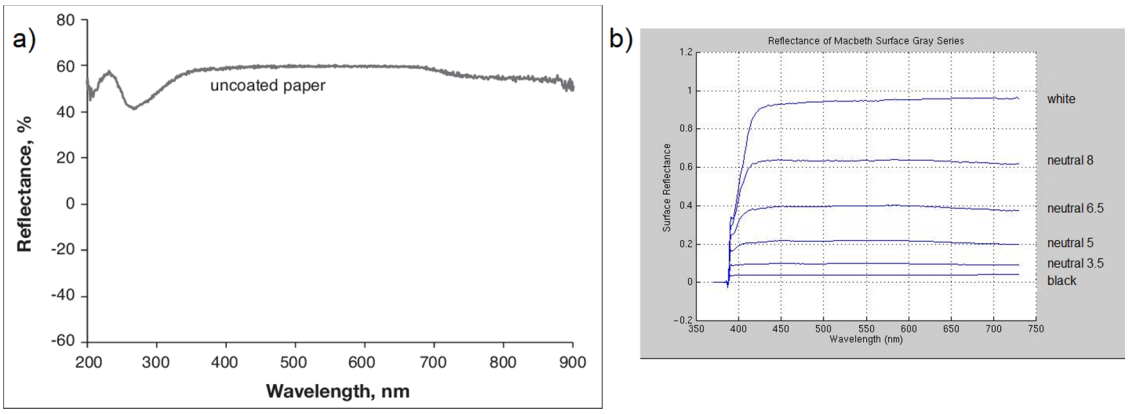


Figure 39. Comparison of reflectance of a white sheet of paper (95) and grey series patches of the original macbeth colour checker.

The ColorChecker chart, manufactured by Gretag Macbeth is commonly used as a colour reference target for photographic and video production work. We measured the reflectance of a grey series of this chart. Because of the similarity of the reflectance of the neutral (in colour) sheet of paper, and Neutral-8 field in the original Macbeth colour checker (Figure 39b), tristimulus colour value of this field was used as a reference point. The colour of this field under Standard illuminant D65 in sRGB colour space is R=200,G=202,B=202 (96). We used this point as potentially the closest to the actual colour of the paper; however, any reasonable arbitrary reference colour point could be used.

For our method, the brand of the sheet of paper under the wells should be chosen that does not indicate 'extra whiteness', which in reality implies a blue tint. If stated by the manufacturer, brightness and whiteness of 100 is desired. Other desirable phrases that vendors might use include 'true white' or 'neutral white'. If other type of a white paper must be used, it would be necessary to create new working curves for assays, for accurate concentration recognition.

2.14.3.The phone

In this study, we used two phones from the most popular manufacturers in Australia, the US and the UK (97): an iPhone 6S (Apple, Cupertino, CA, USA) and a Galaxy Note 4

(Samsung Group, Seoul, South Korea). Since we are using a single reference point for the colour normalisation procedure, we are able to compensate for the white balance only, not the complete colour response of the sensor. During our initial studies, we discovered that the iPhone 6s saturates colours more intensely than the Galaxy Note 4, which resulted in a different shape of the working curve prepared using known concentrations of our assays. To avoid additional readout errors related to an incorrect working curve function, we decided to use separate working curves for these two phones.

2.15. Choosing a colour space

The camera sensor originally captures light using RGBG (RGB with doubled green sensor) as described in section 2.3. While undergoing internal image processing (which is not disclosed by manufacturers in detail but we can suspectit to be based on standard procedures), the signal is transferred to different colour spaces like $Y'C_BC_R$ (section 2.5.4), or probably CIE L*a*b for the white balancing (section 2.3). The final image file, however, is again saved in sRGB. Using the colour space in which the camera physically operates suggests the smallest distortions and data losses. sRGB though, can be recalculated to different colour spaces, whose values can represent different phenomena, be differently scaled, or differently processed, which might provide a more accurate indication of the changes in the assays that we are trying to quantify. We have considered some of the popular colour spaces to assess whether they would provide a better indicator for assay concentration calculation.

2.15.1.Working curve experiments and the golden standard

To assess different colour models, we prepared an IL-8 working curve sample set as instructed by the manufacturer, including additional concentrations higher than the maximum stated by the manufacturer (Figure 40). We distinguish these samples into two subsets: in part 1 – wells from 1 to 7, we cover the suggested concentration range and the colour changes from bright to vivid yellow. In part 2 – wells from 8 to 12, we exceed suggested concentration range and the colour changes from vivid to dark yellow. When discussing the results for different colour spaces we will refer to part 1 and part 2 of the working curve sample set. The difference in the characteristics

of the change of the colour for these two parts is very clearly distinguishable in most of the colour spaces.

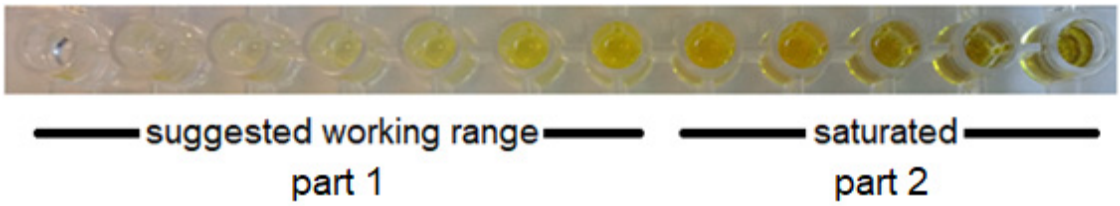
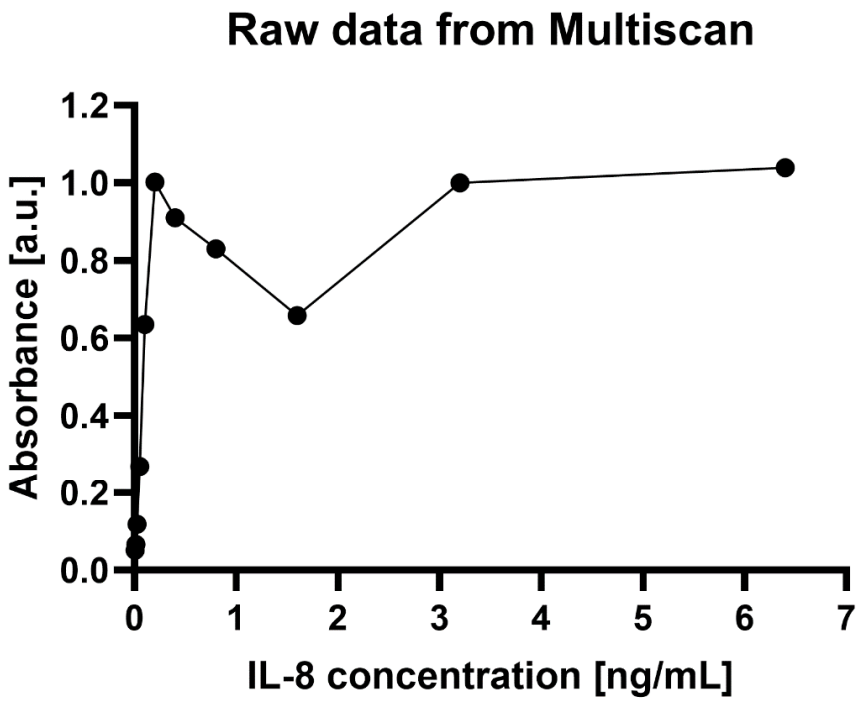


Figure 40. Working curve sample set for IL-8 assay. Wells from 1 to 7 are representing manufacturers working range for this assay, wells from 8 to 12 are additional higher concentrations.

We measured this sample set using the manufacturer's stated absorbance measurement, using an Enspire 2300 (PerkinElmer, Waltham, MA, USA). Figure 41 presents this golden standard working curve acquired with the Enspire 2300 in absorbance measurement mode. Despite obvious changes in colour characteristics, the absorbance measurement does not provide differentiation for samples after number 6, whose concentration is 0.2 ng/mL.

Subsequently, we measured all the samples using the Galaxy Note 4 (Samsung Group, Seoul, South Korea) smartphone, following our protocol that will be described in the next section. Each sample was characterised by 3 numbers, one for each RGB channel, representing one colour in the sRGB colour space. Then, we performed recalculation of the sRGB colours of each different sample into different colour spaces, acquiring new tristimulus numbers appropriate to that colour space. We performed all the calculations with respect to D65 standard illuminant. We examined each channel of all resulting colour spaces to assess its usefulness in serving the purpose of a working curve for this assay.

All the following graphs in this section will have the same X axis presenting concentration of IL-8 assay in ng/mL. The Y axis will always present the value of the specific channel of discussed colour space and will be presented in its own units. This information will not be on the graphs to preserve space. The channel will be specified in the title. The range of channels will always be specified in the text.



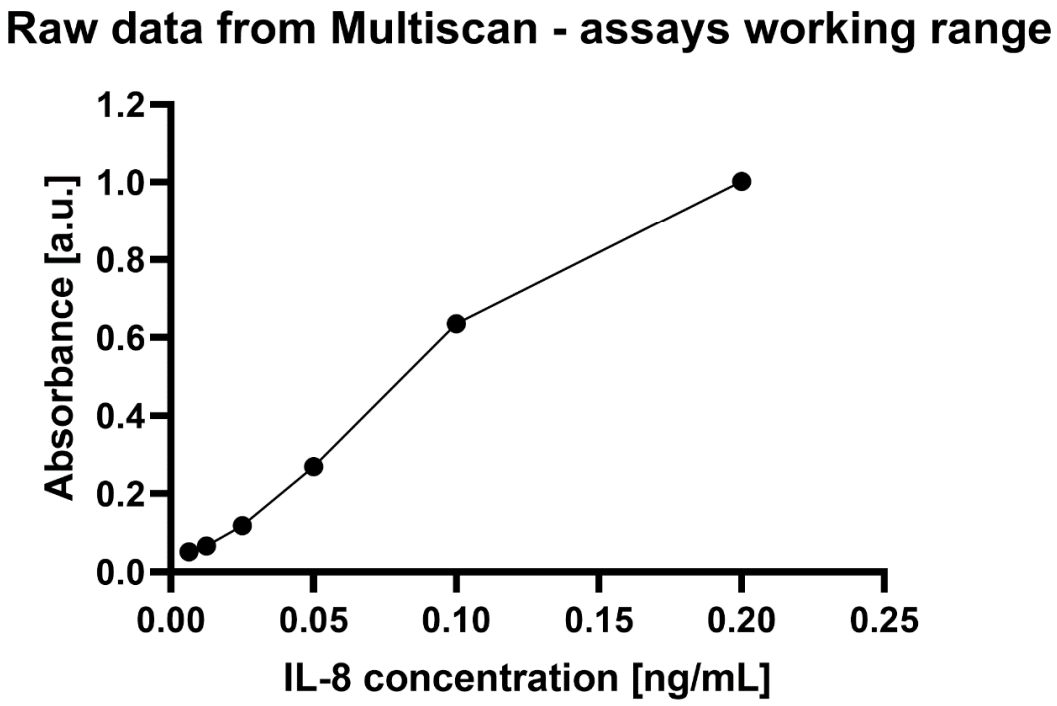


Figure 41. Working curves acquired using Enspire 2300 absorbance mode at wavelength of 405 nm. Upper panel presents absorbance of all assay samples (part 1 and part 2). Lower panel is presents only the working range suggested by the manufacture (part 1 samples only).

2.15.2.sRGB colour space

sRGB is the original colour space of the image saved by the smartphone. We already described it in more detail in sections 2.3 - 2.5 as a native colour space in which smartphones operate. For the working curve consideration, the colour of the assay does not directly fit any of the RGB primary colours - yellow colour in RGB format is a combination of the red and green primaries. RGB is an additive system, which means that adding all the colours will result in white. This implies that change between faint yellow and vivid yellow as seen in our assay will be encoded mostly in the third colour channel – blue, starting with high values and decreasing with 'whiteness'. All the channels have integer values between 0 and 255.

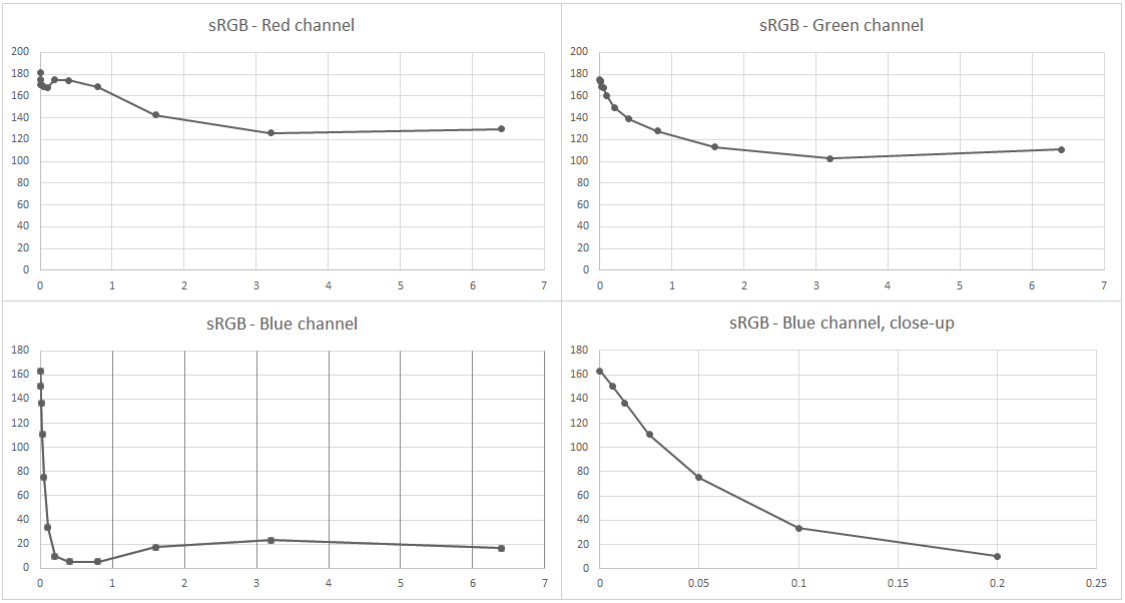


Figure 42. Values for each channel of sRGB colour space of consecutive samples.

On Figure 42 we can observe expected behaviour of all the colour channels. For concentrations below of 0.2 we can observe fairly stable values of red and green channels, with rapidly changing blue channel (close-up on Figure 42 – bottom right panel). For higher concentrations, values of the blue channel are not usefully linked to concentration any more. The blue channel has a variability of about 170/255 levels which is about 65% of the full range. We can observe regions of useful change in green and red channels. The green channel has a greater (60/255) variability than the red channel (40/255) but both of them are less sensitive than the blue channel, and the usefulness of both channels is limited to concentrations of under 3.2 ng/mL.

2.15.3.HSV colour space

HSV (hue, saturation, value) is a transformation of the sRGB colour space, and their components are derived from sRGB components. In this model based on cylindrical coordinates as shown in Figure 43, colours of each hue are arranged azimuthally around a central axis of grey which ranges from black at the bottom to white at the top, with colour saturation increasing with radial distance. HSV was an attempt to accommodate more traditional and intuitive colour mixing models (98).

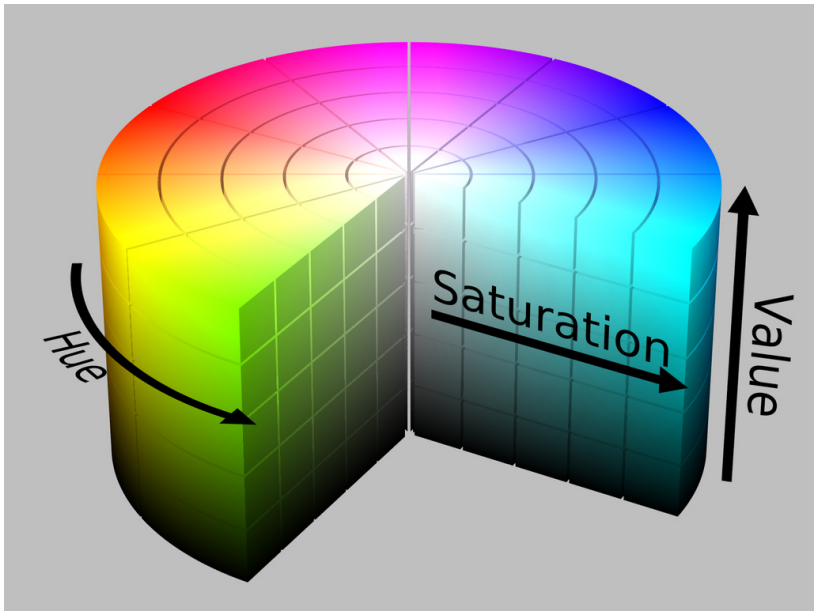


Figure 43. The HSV cylinder (99)

The computer science applications that favour the HSV colour model include object detection, object recognition (face recognition), text, content-based image retrieval, and analysis of medical images. Most of these uses however, are based on the grayscale thresholding for which it is easier to use the HSV saturation component than to analyse all three RGB channels. In colour recognition HSV is less efficient because of the simplicity of its design (for low computational consumption). HSV often causes confusion between parameters where a systematic change of one value does not always lead to expected perceptual outcome. Calculation of the HSV values from sRGB colour model goes as follows:

$$H := \begin{cases} 0, & if\ MAX = MIN \leftrightarrow R = G = B \\ 60° \times \left(0 + \frac{G-B}{MAX-MIN}\right), & if\ MAX = R \\ 60° \times \left(2 + \frac{B-R}{MAX-MIN}\right), & if\ MAX = G \\ 60° \times \left(4 + \frac{R-G}{MAX-MIN}\right), & if\ MAX = B \end{cases} \quad (39)$$

If $H < 0°\ then\ H := H + 360°$

$$S_{HSV} := \begin{cases} 0, & if\ MAX = 0 \leftrightarrow R = G = B = 0 \\ \frac{MAX-MIN}{MAX} & if\ MAX \neq 0 \leftrightarrow R = G = B \neq 0 \end{cases} \quad (40)$$

$$V := MAX \quad (41)$$

The H component has values between 0° and 360°, but are presented here normalised to between 0 and 1. The S and V values range from 0 to 1. In our case, determining which component will be the best for the assay calculation is not obvious since white and black are not at opposite ends of any component. Since the brightest colours are at Saturation = 0, we are expecting the first part of the assay sequence to be encoded here. Black colour is Value = 0, which means the second part of the assay sequence should be decreasing in this component. All HSV channels for IL-8 assay are presented on Figure 44.

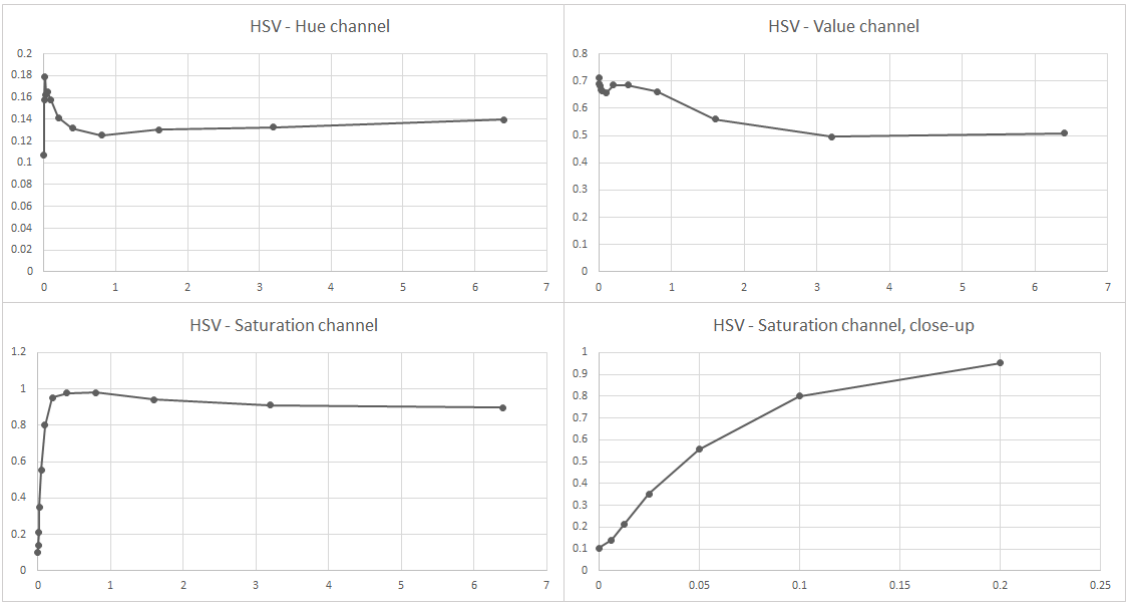


Figure 44.Values for each channel of HSV colour space of consecutive samples.

As expected, we can observe that the Hue channel is not descriptive at all, sightly oscillating around the value for yellow colour. The Saturation channel presents high variability in the first part of the assay, covering over 80% of the range of this channel. The Value channel responses to changes in the second part, but both ends of the graph are very flat (have a small spectral resolution) and the whole graph covers only 18% of its range.

2.15.4.CIE XYZ colour space

We described the CIE XYZ colour space in detail in section 2.5.2. Calculation of XYZ values from sRGB channels requires two steps. Initially we need to linearize sRGB using inverse gamma correction function:

$$I_{linear} = \begin{cases} \frac{I_{sRGB}}{12.92}, & I_{sRGB} \leq 0.04045 \\ \left(\frac{I_{sRGB}+a}{1+a}\right)^{2.4}, & I_{sRGB} > 0.04045 \end{cases} \quad (42)$$

where $a = 0.055$ and I represents one of three RGB channels, followed by a matrix multiplication of the linear values to get XYZ:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} R_{linear} \\ G_{linear} \\ B_{linear} \end{bmatrix} \quad (43)$$

Since the CIE XYZ colour space is based on the cone responses of the human eye, which is roughly similar to RGB in principle, we expect similar behaviour of these channels. X and Y channels combined will produce a yellow colour, while the remaining Z channel should manage its brightness. The unit of the tristimulus values X, Y , and Z defined such that $\mathrm{Y} = 1$ is the brightest white that a colour display supports. The corresponding white point values for X and Z can then be calculated using the standard illuminants. All the channels have a range between 0 and number close to 1 (depending on the normalisation). Outcomes are shown in Figure 45.

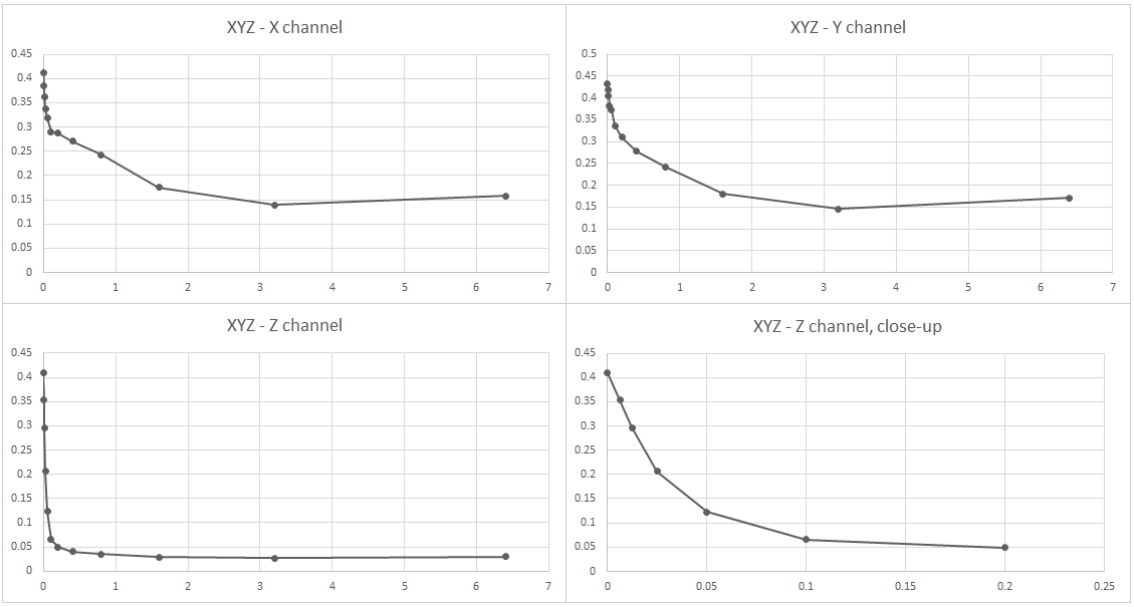


Figure 45. Values for each channel of XYZ colour space of consecutive samples.

Despite similarities to the sRGB colour space, the XYZ is different in many aspects, including greater gamut and different positioning of three variables in the general colour space. Y is typically named Luminescence, Z corresponds to blue stimulation and X is a mix of cone response curves chosen to be nonnegative. Channels X and Y present variability across the whole assay, but covering about 30% of the range and presenting rather low resolution. The Z channel behaves similarly to the blue channel in sRGB, changing rapidly in the first part, but completely flatten for concentrations higher than 0.2. The Z channel utilises about 35% of the range of the colour channel for concentrations between 0 and 0.2.

2.15.5.CIELAB colour space

We initially described the CIE L*a*b colour space in section 2.5.2. In short – it was developed to match human vision, rather than present accurate linear light intensity. The advantage in the case of our assay is that one of the channels represent lightness, which should be an appropriate descriptor of the gradation of the colour in the assay.

Transformation from sRGB is not straightforward and consists of a couple of steps. First we convert sRGB to linear RGB (equation 42) and derive XYZ values from linear RGB values as in section 2.15.4. This is followed by the XYZ to LAB transformation presented by equations 23-27.

The Lightness channel has a range of 1 to 100, while both *a and *b components are in a range between −128 to 127. LAB channels for our test assay are presented in Figure 46.

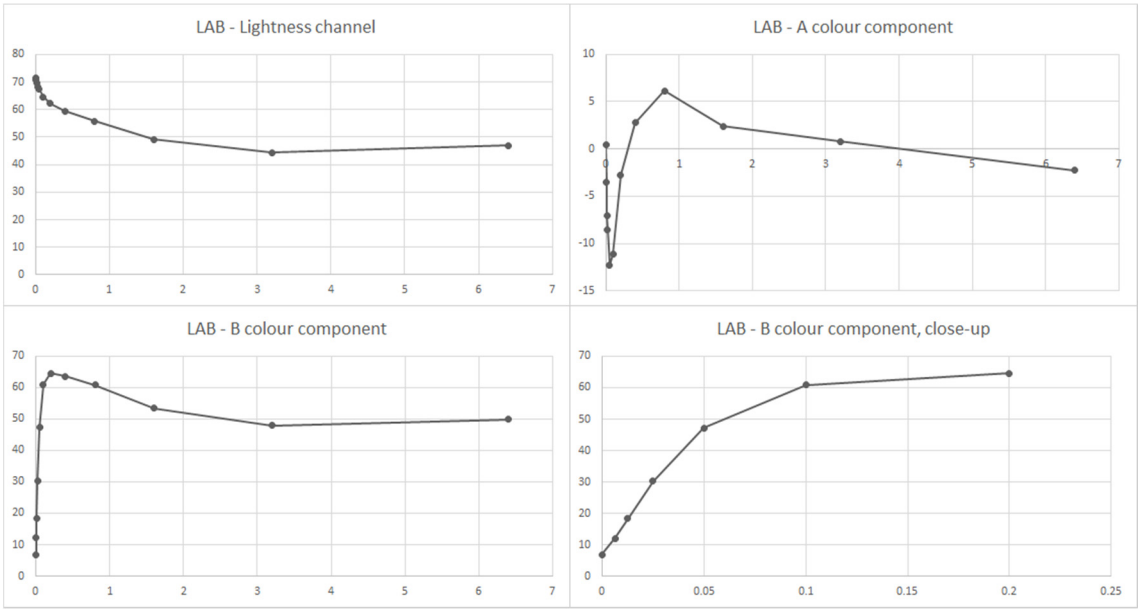


Figure 46. Values for each channel of LAB colour space of consecutive samples.

The Lightness channel, as expected presents, differences in subsequent samples across the whole range of samples, however its resolution is rather small covering about 30% of the range of the channel, drastically flatting in the second part. Component *b, which presents transformation between blue and yellow colour presents a steep slope in the first part, with a range of almost 60/255 covering about 24% of the range of the channel. Component a* presented a negligible and irregular variability.

2.15.6.YCbCr colour space

YCbCr is another colour space that is derived from RGB colour space, described previously in section 2.5.2. This space was mainly used for analogue TV signal transmission and is not widely applied in computer applications. This is another colour space that has a specific component describing luminance –Y, and so from its definition, this channel might be expected to have increased resolution compared to two other channels. Effectively the Y channel has a range between 16-235, while the

Cb and Cr channels have a range between 16-240 (the remaining bits are used for metadata information). Results for the YCbCr colour space are presented in Figure 47.

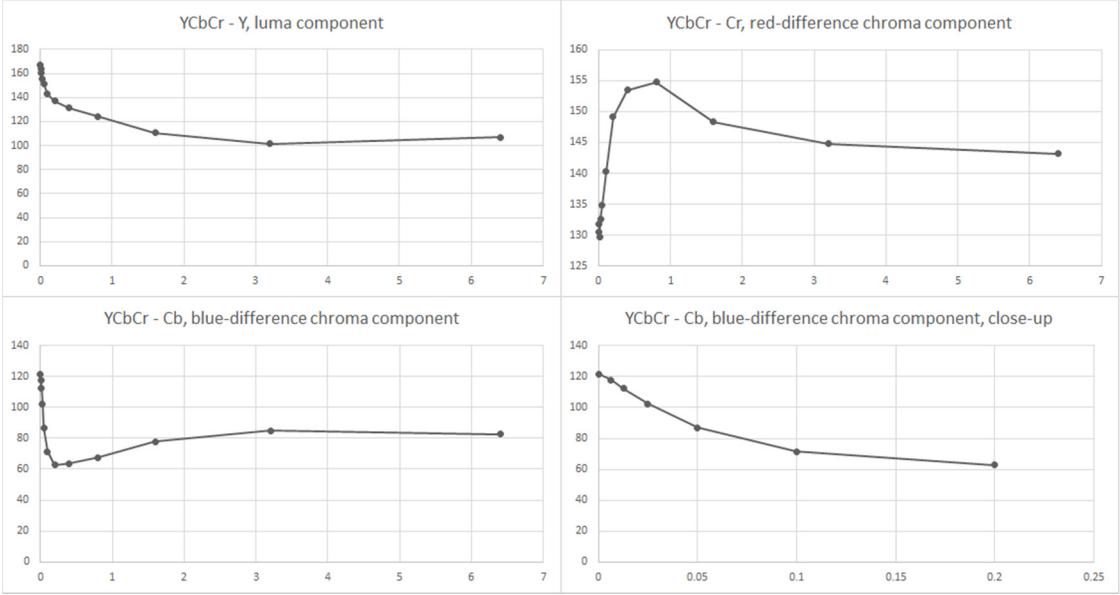


Figure 47. Values for each channel of YCbCr colour space of consecutive samples.

The Luminescence component similarly to other colour spaces presents difference in the first part where colour changes from bright to vivid yellow. The resolution decreases drastically for the second part but would still allow detection. The useful part of the signal covers about 30% of the range of the channel. The Cb channel also shows good variability of the signal for the first part, with even steeper slope, but the range of the signal in this channel covers only about 27%.

2.15.7.Summary and choice

The ideal colour space for our assay would be one that would have an approximately linear response in one whole channel, covering all the changes in colour with steep slope and high dynamic range to assure high readout resolution. In reality, based on all the readouts, we can realise that the assay can be clearly separated into two parts, first where yellow colour changes between bright and vivid, and second where vivid yellow turns to dark yellow.

Colour spaces offering greater gamut (e.g., CIE XYZ) in reality did not provide it since colours were calculated from smaller sRGB colour space. Since ultimately all the colour

spaces offered 8 bits per channel, increased gamut decreased the resolution of useful colours in it.

Some colour spaces like XYZ or LAB offered channels that were responsive to the whole spectrum of changes of the sample. The resolution of these changes was small though, potentially not providing good enough distinction between similar concentrations.

Most of the representations offered a channel that present changes in colour in the first part well. Colour spaces offering a channel of light intensity (HSV, LAB) were expected to do well, although spaces with different channel division typically also offered a channel that behave in the similar manner such as Blue in sRGB and Z in XYZ.

If the colour change of the assay covers significant range of the channel, it has a direct influence on the steepness of the slope of the function that describes the change, therefore on the effective resolution of the readout. The Saturation channel from the HSV colour representation offered the best coverage of over 80% for the first part of the assay. The second best colour space in the first part of the assay was the blue channel from the original sRGB colour space, offering coverage of over 65%.

The colour space that presented the best resolution for the second part (the last 6 samples) was the green channel from sRGB colour space with a dynamic range of about 50/255, which is about 20% of the range. The second best was the V channel from HSV colour space with a range of about 18%

The choice of the colour space was settled between sRGB and HSV. HSV offered slightly better dynamic range for the first part, but slightly worse for the second. sRGB had an advantage of being the native colour space of the camera meaning it did not introduce any additional errors or limitations caused by the conversion to smaller gamut or smaller bit resolution colour space. Also, sRGB for both assay parts provided graphs with a more regular shape, which allowed for better function fit, decreasing concentration interpolation errors.

Eventually, we decided to use the original sRGB colour space for concentration calculation, using the obvious choice of the blue channel for low concentrations, but

using a ‘greyscale value’ rather than the green channel for higher concentrations. To explain this, consider the simulation below showing similar results to our assay changes. Figure 48 on the X axis presents colour bar similar to the ones that can be found in graphical programs, generated by the RGB functions presented on the plot. At the bottom our sample set is there for comparison.

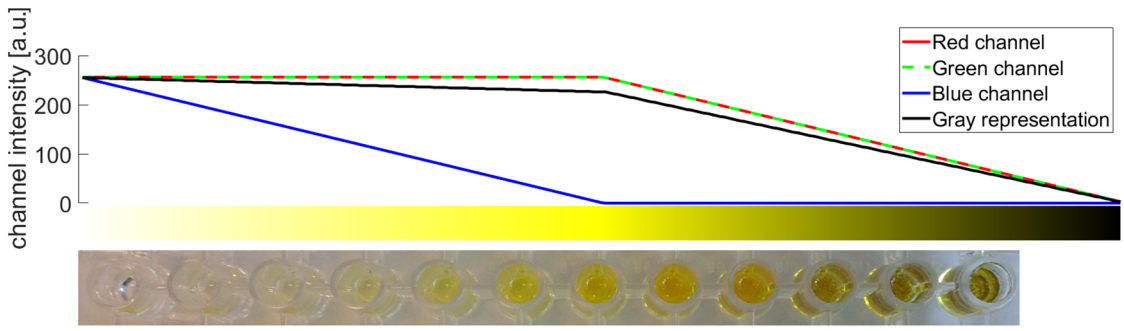


Figure 48. Behaviour of all three RGB channels, (plus additionally grey representation) for the colour change from faint to dark yellow, compared to IL-8 assay colour gradation.

As expected the first part of the graph is a change in the blue channel, while the second part is covered by equal change of both green and red channels. The graph in addition to the RGB colour channels, presents one more function – greyscale representation.

Greyscale colour representation is not a separate colour space. It contains only one channel of chromacity (Y), typically with 255 levels for 8 bit variable and shade changing from white to black, with no colours. From sRGB it is calculated using a simple linear equation (100):

$$Y = 0.2126\,R + 0.7152\,G + 0.0722\,B \qquad (44)$$

We see that the greyscale value is most strongly related to the green channel. Because the value of the grey representation is also dependant on the blue channel, in the case of our assay it has slightly smaller values than red and green channels. Because of averaging between all channels, it is more resistant to small illumination temperature changes than the green channel alone.

2.16. Process for measurement and software for calculating results

2.16.1.Taking an image

We describe the process of taking images, applied to clinical samples as well as the previously presented working curve images. Separate assay plates with IL-8 and NE were imaged on a white sheet of paper, using an iPhone 6S (Apple, Cupertino, CA, USA) and a Galaxy Note 4 (Samsung Group, Seoul, South Korea). Sets of images of assay plates were taken at a height of 20 cm, at approximately a 90-degree angle, such that at least 9 wells on the microplate were imaged and the unobstructed parts of the white sheet of paper were visible. Both smartphones had the cameras set to default settings, including automatic exposure, automatic white balance, and autofocus. The images were taken under office fluorescent lights (Figure 49). Our normalisation procedure accommodates the most common illumination types like sunlight, cloudy day and standard types of bulbs.

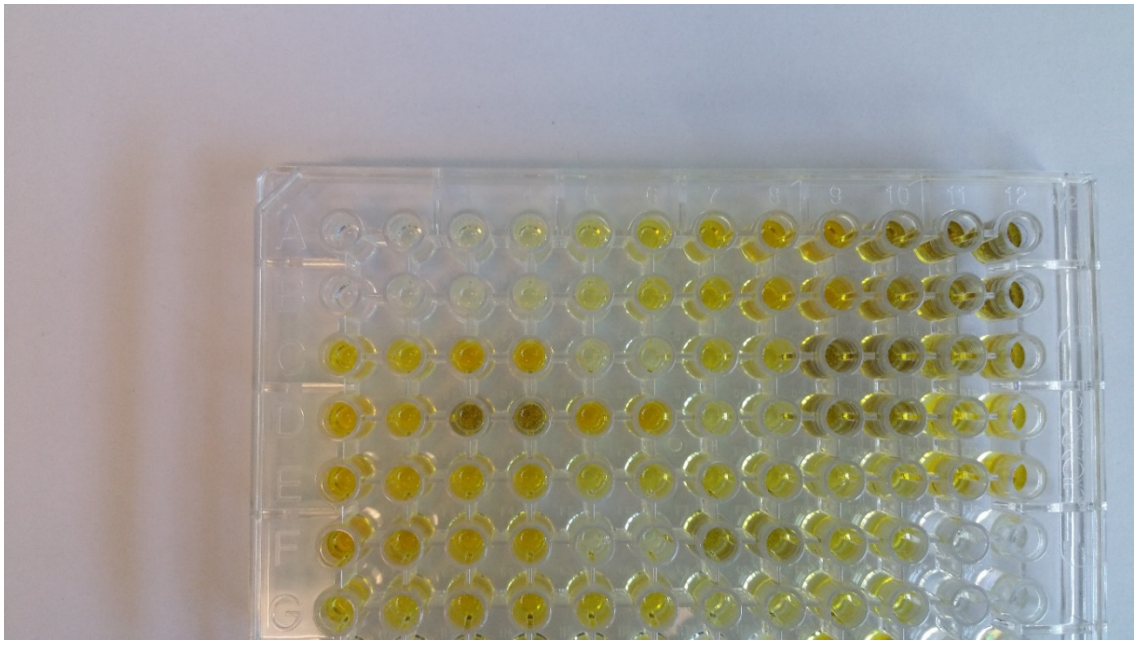


Figure 49. Sample image of the IL-8 assay photographed for the measurement.

2.16.2. Image normalisation

To enable absolute colour comparisons, standard illumination conditions have been introduced, referred to as the D65 standard illuminant (24). D65 is a reference point for the sRGB colour space and all further calculations will be processed with respect to it.

The colours of an image captured with a smartphone can vary significantly not only depending on the illumination conditions, but also on the device used. To quantify the colour readout as objectively as possible, we performed several operations to relate the captured images to a common benchmark.

The real image is modified by a smartphone and its camera application according to the phone default settings. In particular, the smartphone software applies an automatic white balance which estimates the illumination temperature and adjusts the image colours to represent them in the "most natural" way (68). The image brightness is adjusted by the automatic exposure time chosen by the smartphone.

In order to relate the smartphone-modified image to a common benchmark, we used normalisation procedures based on the blank sheet of white paper. This blank sheet is assumed to yield standard RGB values $(\mathrm{R} = 200, \mathrm{G} = 202, \mathrm{B} = 202)$ under illumination of the D65 standard illuminant (24). We applied three different normalisation procedures in each specific case as follows.

2.16.2.1. sRGB normalisation

Using these standard values, we recalculate all image sRGB values based on the information taken from a section of our image showing the blank white paper. To this aim, the image data in all three sRGB channels are multiplied by suitable correction factors to enforce the sRGB colour values of the paper to be equal to $\mathrm{Iref}\,(\mathrm{R} = 200, \mathrm{G} = 202, \mathrm{B} = 202,\ G = 202)$ where last G stands for additional channel for the greyscale image calculated using equation 44. This operation provides two important corrections: (a) it compensates for differences in the lighting temperature and/or colour of light and for the smartphone image processing algorithms; and (b) it adjusts image brightness of all images to a common level so that further analysis can be performed.

Our normalisation algorithm involves the following steps:

- The image is linearized using equation 42

- The user is asked to select a region of the image where the blank sheet of paper is visible. If the illumination varies across the image the user must make this selection where paper is similarly illuminated as the sample, typically just above or below the sample (Figure 50a).

- This selected and cropped region is decomposed into separate RGB colour channels.

- Additionally, this region is converted into a greyscale image using formula equation 44, which creates an additional greyscale channel.

- The mean pixel value $Ipaper_{RGBG}$ over the selected region is calculated for each RGB and greyscale channels.

- The correction factor is calculated for each channel separately as follows

$$Cf_{RGBG} = \frac{Ipaper_{RGBG}}{Iref_{RGBG}} \quad (45)$$

- These correction factors are then applied to the whole image.

$$Icorr_{RGBG} = I_{RGBG} \times Cf_{RGBG} \quad (46)$$

- Corrected image $Icorr_{RGBG}$ is delinearized using equation 35.

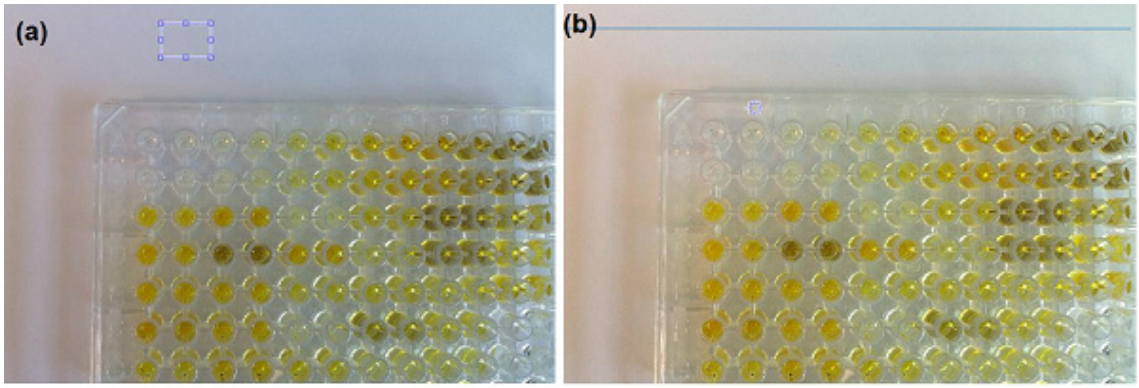


Figure 50. (a) Selecting an appropriate region of the blank sheet of paper to perform colour normalisation of the image. (b) Colour corrected image, with lighting uniformity check line visible (procedure described in section 4.4.3.2.)

2.16.2.2. Bradford normalisation with selected reference point

The second procedure in principle is very similar to procedure 1, but it uses different calculations for the normalisation. Instead of working in the original sRGB representation, the values of the sRGB are transformed to the XYZ colour space including image linearization as presented in section 2.15.4. Then the Bradford method is applied to scale (normalize) colours. The Bradford method is a chromatic adaptation algorithm that can be implemented as a linear transformation of a source colour X_s, Y_s, Z_s into a destination colour X_d, Y_d, Z_d, by a linear transformation $[M]$

$$\begin{bmatrix} X_d \\ Y_d \\ Z_d \end{bmatrix} = [M] \begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix} \tag{47}$$

which are dependent on the reference neutral $(Iref)$ and measured neutral $(Ipaper)$. The Bradford method defines the cone response with following matrices

$$[M] = \begin{bmatrix} 0.9869929 & -0.1470543 & 0.1599627 \\ 0.4323053 & 0.5183603 & 0.0492912 \\ -0.0085287 & 0.0400428 & 0.9684867 \end{bmatrix} \times \begin{bmatrix} \frac{\rho_D}{\rho_S} & 0 & 0 \\ 0 & \frac{\gamma_D}{\gamma_S} & 0 \\ 0 & 0 & \frac{\beta_D}{\beta_S} \end{bmatrix} \times \tag{48}$$

$$\times \begin{bmatrix} 0.8951000 & 0.2664000 & -0.1614000 \\ -0.7502000 & 1.7135000 & 0.0367000 \\ 0.0389000 & -0.0685000 & 1.0296000 \end{bmatrix} \tag{49}$$

After normalisation, the image is transformed back to the linear RGB colour space

$$\begin{bmatrix} R_{linear} \\ G_{linear} \\ B_{linear} \end{bmatrix} = \begin{bmatrix} 3.2404542 & -1.5371385 & -0.4985314 \\ -0.9692660 & 1.8760108 & 0.0415560 \\ 0.0556434 & -0.2040259 & 1.0572252 \end{bmatrix} \tag{50}$$

followed by the gamma correction to delinearize the image (equation 35)

The advantage of this method over the sRGB linearization lies in the cone response which considers changes in all the related RGB channels for each given (even monochromatic) illumination. In other words, a pure red colour illuminated on the sensor based on the cone response (Figure 18) could have influence on the green channel as well. In reality, we are observing images collected by the sensor, not the

human eye. We do not know what demosaicing procedure was used in the smartphone but we do know that red, green, and blue channels were collected separately. Additionally, the blue channel that was used for most of the concentration recognition has the smallest common part with other colours. Any increase in accuracy of image normalisation for our use is expected to be small, with great increase of computational power. This method was applied to iPhone 6 elastase measurements since the colour change for this assay had the smallest range and we tried to minimise all possible errors to maximise readout accuracy.

2.16.2.3. Bradford normalisation with thresholded reference point

This procedure is almost identical to the previous procedure with the difference with respect to the treatment of the reference grey. Instead of selecting a section of the white sheet of paper, the algorithm uses the mean values (excluding the 10% of highest and lowest pixel values), of the whole image to estimate the greyscale level of the image. This method assumes that most of the image shows the paper with just a few small samples. This approach decreases the number of steps performed by the user. It was applied to Samsung galaxy 4 elastase measurements for the same reasons of accuracy and because of favourable conditions of the taken image.

Both Bradford methods were added to the system at later point, after the IL-8 assay was tested. We did not repeat IL-8 measurements since the gain would be small and it was not required by the assay.

2.16.3.Rejecting poor quality images

Additionally, we perform an image check, to avoid incorrect readouts if the image taken for analysis has detectable flaws. We reject the images in two check steps.

2.16.3.1. Incorrect automatic white balance

When the user selects area where the blank sheet of paper is visible, the algorithm acquires RGB values for this part of the image ($Ipaper$). If any channel value is different from the other two channels by 25 or more, it implies a significantly problem with the phone's internal automatic white balance procedure, which disqualifies the image from being reliable for the measurement. We chose this number as 10% of the range of 255 levels. The user is then requested to take another picture, considering

visual correctness of colours visible on the screen of the smartphone, or changing illumination conditions to one of the standards that camera compensates for correctly (Table 1).

2.16.3.2. Significantly variable illumination

We also implemented algorithm checking consistency of the illumination across the horizontal axis of the image. We check a horizontal cross section through the whole image passing through the central point of the square that the user selected as the white sheet of paper. This cross section is shown as a line on Figure 50b. The cross section is filtered with a gauss filter (101) for smoothing. We used an overlapping sampling window to acquire a mean pixel value from each 20 subsequent pixels, moving the window by 10 pixels in each step along the cross section. This procedure was applied from both ends of the image simultaneously, moving towards the centre. The values returned by the window for each colour channel separately were analysed. If the value was greater than 150 (towards white colour), the standard deviation of 20 pixels within the window is smaller than 3 (uniform surface), and the pixel values differences between channels are not greater than 25 levels (grey colour), the sample is marked as positive and the window is moved. If 5 consecutive samples meet these conditions, the area is considered to be a sheet of paper. Based on these checks, the X coordinate of the beginning and the end of what is detected as the sheet of paper is recorded. The function returns colour values for all channels at the beginning and the end of the detected paper and draws the line as shown on Figure 50b for visual confirmation. If any of the channels has a difference between the beginning and the end of the sheet of the paper greater than 15, the image is considered faulty because of inconsistent illumination. The value of 15 was chosen based on the potential impact on the readout in suggested conditions. If the sheet of paper covers at least 80% of the image size and change of the illumination is fairly linear, it means that the colour changes spatially about 1 level per size of the sample well. Assuming an attempt of a proper normalisation where the section of the paper was selected with the same illumination conditions (typically above or below the sample), it allows potential error of about 1/256 level because of the illumination inconsistency.

2.16.4.Sample analysis

After the image is accepted and normalised, we perform sample readout. The measurement starts with a selection of a rectangular Area Of Analysis (AOA) by the user (Figure 51a). The AOA is a section with a single sample image cut from the image captured with a smartphone (Figure 51b). Depending on the size of the image and distance between sample and camera, different sizes of the cropping rectangle can be chosen. We refer to the cropped image as CI. $P_{(x_1,y_1)}$ denotes the upper left corner of the cropping square. The resulting image is separated into 3 sRGB channel images $CI_R + CI_G + CI_B$ (Figure 51e), and recalculated to form an additional single channel of greyscale representation CI_{GR} (equation 44). All presented algorithms were applied to the greyscale channel as well. CI_{ch} represents one of the mentioned four possible channels.

This step creates four square matrices of a given size. In the next step, adaptive filtering is used to improve image quality and to reduce noise in each channel. To this aim we apply a two dimensional Wiener filter (equations 36).

The Wiener filter then creates pixelwise filter kernels:

$$b(x,y) = \mu + \frac{\sigma^2 - v^2}{\sigma^2}(CI_{ch}(x,y) - \mu), \quad (51)$$

where μ is the local mean, σ^2 is the local variance, and v^2 is the noise variance - the average of all the local estimated variances.

This filter is applied to each channel separately (Figure 51f and Figure 51h)

$$CI_{ch}^F = CI_{ch} * b(x,y). \quad (52)$$

Here, CI^F is the filtered image, $*$ denotes the convolution operation and b is the filter kernel.

Subsequently, a round binary mask is calculated (Figure 51c). The radius of the circle r inscribed in the cropped square is slightly smaller than the half of the square side s so the unwanted artefacts at the border of the image are eliminated

$$r < 0.98 \times \frac{s}{2}. \quad (53)$$

The mask is applied to each channel separately (Figure 51g):

$$CI_{ch}^{M} = CI_{ch}^{F} \circ M \tag{54}$$

where M is a round binary mask with ones for pixels fulfilling the condition (equation 51) and zeros elsewhere and $\circ$ is an element-wise multiplication.

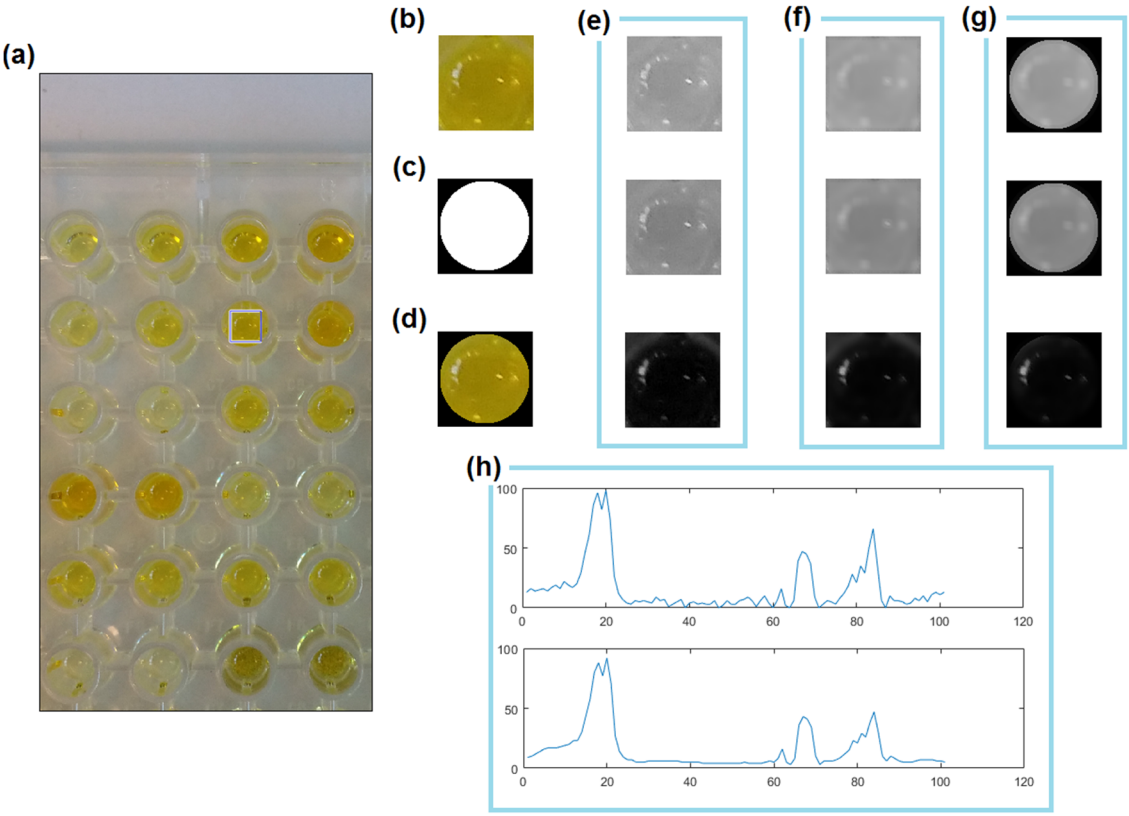


Figure 51. Algorithm steps for probing. a) Selection of the AOA; b) Cropped image; c) Calculating a mask for the cropped image; d) Simulated effect of a mask applied to each channel; e) Separation of channels; f) Filtering each channel with the Wiener filter; g) Applying a mask to each channel; h) Cross section through unfiltered and filtered image, (green channel).

Because the fluid surface in the well typically creates a meniscus, it behaves like a wide-angle lens, creating visible reflections of any strong source of light around the well. These reflections are visible as a colour change, introducing readout errors. To overcome this problem, an additional anti-glare system was designed and applied. The anti-glare system detects significant colour changes using Otsu's threshold selection method (102). Once identified, these areas are added to the mask and excluded from the calculation area (Figure 52).

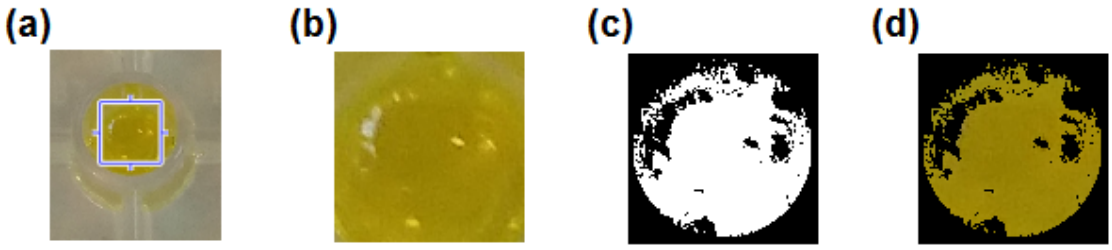


Figure 52. Resolving the problem of reflections. (a) Selection of the AOA, (b) magnified AOA with visible light reflections, (c) Round mask with anti-glare system included, (d) Final mask applied to the AOA, excluding unwanted regions from the readout.

The mean values of these prepared matrices are calculated

$$P_{ch}^{I} = \frac{1}{X}\sum_{x=1}^{X}\frac{1}{Y}\sum_{y=1}^{Y}P_{x,y}\,, ch = CI_R, CI_G, CI_B, CI_{GR} \quad (55)$$

where P^I denotes the mean pixel intensity, x and y are pixel indices and $P_{x,y}$ are the nonzero pixels of the calculated colour channel matrix. The result of a single AOA is a set of four numbers, one for each sRGB channels and the associated greyscale representation value.

2.16.5.User experience summary

Here we summarise the procedure from the user perspective. The user carries out the assay of the tested sample and takes a smartphone image. To this aim, the sample well is placed on white sheet of paper, under fairly uniform, natural sunlight or artificial interior illumination. In order to obtain the assay readout, the user uploads the image to our program. First, the user selects an area of the image with the white sheet of paper visible. It is always a good idea to select an area with illumination the most similar to that of the sample to measure; typically, directly above or under it. The image is normalised and displayed in a recalculated colour version. If the image does not meet the quality criteria, it is rejected, with an appropriate message displayed. Then, the user drops the floating window to identify the AOA on the sample image to select it. The calculations are then performed automatically. The normalisation and image processing procedures are executed. The values for the blue channel of sRGB representation and greyscale representation are acquired. The algorithm checks what phone was used to take an image and checks which assay type the user chose. Based on these variables the program proceeds with the correct working curve. In the case

of the IL-8 assay, both blue and greyscale channels are investigated. When the greyscale value is lower than 133 (for normalised image, meaning that the analyte concentration is elevated above the regular range), the greyscale working curve is chosen. The threshold value is chosen to be identical for both: Samsung and iPhone. Finally, after the calculations are completed, the analyte concentration, potentially with initial descriptive interpretation of the result, is displayed to the user.

2.16.6.Tests of colour normalisation

The first experiment conducted was testing the smartphone readout under different illumination conditions with and without our image normalisation procedure. The first scenario presents the situation where images were taken according to our instructions under fluorescent illumination in the office (D50 standard illuminant) and then under sunlight illumination outdoors (D65 standard illuminant), but the normalisation procedure was skipped. Only the smartphone's white balancing and automatic adjustments took place. Secondly the same images were normalised using our sRGB method and measured afterwards. We used consecutively increasing concentrations of IL-8 for this measurement. The results are presented in Figure 53. Please note that the IL-8 assay used for these test was prepared following different instructions than in all other presented experiments with medical samples, and therefore concentrations had different signal levels. However, the shape and behaviour of the channels for the working curves are similar, and so the results are relevant to our general measurements.

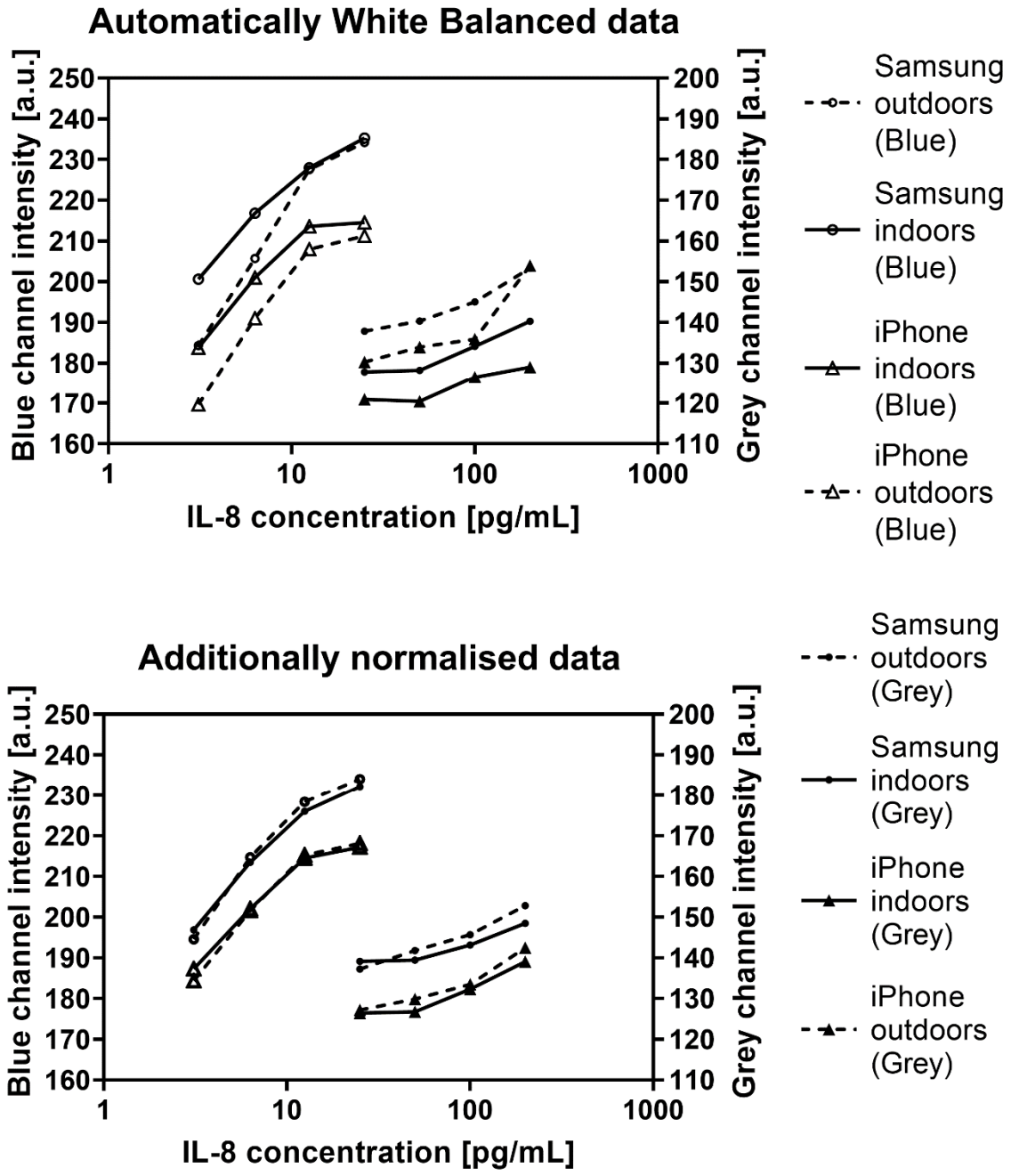


Figure 53. Consecutive IL-8 concentrations measured with two smartphones. Solid lines represent measurement performed in the office, dotted line outdoors. Empty markers are for the blue working curve, solid markers are for the grey working curve. Top graph shows raw smartphone data while the bottom shows data after our normalisation procedure.

Presented in Figure 53, measured points were taken from different images, which means that phone could perform slightly different adjustments to each one of them, so allowing the greatest change of us observing an inconsistent readout. The first set of curves (empty markers) is for the blue channel readout, while the second (solid markers) is for the grey channel - as used later to prepare working curves. The top

panel shows that the iPhone kept its automatic adjustments more regular, consistently lowering (but decreasing the difference with the increasing concentration) the blue channel readout compared to the measurements taken under artificial light. The Samsung phone initially lowered the readout values for the first four samples to match it later for both illuminations. The amount of the yellow colour (higher sample concentration) appearing in the image appears to have been changing Samsung's adjustments. The mean standard deviation between data points for the iPhone measurements for original images was 4.20 for blue data points and 7.15 for grey (5.68 average for both), while for the Samsung phone was 5.11 for blue and 5.77 for grey channel data points (5.43 average for both).

The bottom panel presents the same data points after image normalisation using our sRGB method. In the case of iPhone outdoors measurements, readouts were increased by a mean value of 9.9 (blue), 7.3 (grey), 8.6 (both), while the indoors measurements on average were increased by 2.0 (blue), 2.5 (grey) and 2.28 (both). This suggests that automatic adjustments of the iPhone were better under the artificial lighting – probably because it was more standardized than unknown conditions of the sunlight.

For Samsung measurements, the mean difference between raw and normalised data for measurements taken outside was 5.1 (blue), 2.3 (grey) and 3.7 (both), while inside the numbers were 3.3 (blue), 5.1 (grey) and 4.2 on average. The internal adjustment of the phone for images taken indoors were more accurate for the blue channel; however, the grey channel difference was higher (one of the other channels was more skewed). The mean difference for external measurements was smaller than the iPhones for the blue channel, but higher for the grey, hence for both.

The mean standard deviation between data points between additionally normalised images taken outdoors and indoors decreased to 0.70 (blue), 1.05 (grey) and 0.88 (averaged) in the case of the iPhone and to 1.57 (blue), 1.39 (grey) and 1.48 (averaged) in the case of Samsung phone.

The experiment was conducted in clearly different but fairly standard illumination types. Neither artificial light nor sunlight can be assumed fully known, because of the

many differences in bulbs and reflections, or altitude and weather conditions, but they should be close to estimated expected conditions. Both phones applied their automatic white balances and other adjustments and as shown in Figure 53 top panel, they both did a reasonable job. However, a standard deviation of 5.68 with maximum difference between same samples of 24 levels (for grey point) for iPhone, and 5.43 standard deviation with maximum difference between same samples of 22 (for blue channel) would introduce a significant error in the concentration readout. After our additional normalisation procedure, mean standard deviation decreased over 6 times in the case of iPhone to the value of 0.88 and over 3 times for Samsung measurements to the value of 1.48. The maximum difference between the same samples were respectively 3.38 (grey) and 4.32 (grey), decreasing respectively over 7 times for the iPhone and almost 4 times for the Samsung phone. The mean difference between same sample points for the original data and treated with our normalisation algorithm is 1.7 for the iPhone and 2.4 for the Samsung phone.

The blue channel normalisation went better than the grey channel. Since the grey channel is a combination of all the available channels, one of the remaining green or red channels must have had greater variation. In conclusion, the normalisation procedure in tested conditions did not provide ideal results, but decreased potential errors 4-7 times, bringing colour readouts very close to each other.

2.17. Analysis of clinical data

We now move on to present more detailed testing and benchmarking using clinical samples and procedures described in section 2.14.1 and steps presented in the section 2.16. Separate working curves, as described were generated for each smartphone and the assays tested. We found that the blue channel of the sRGB colour space provided the best resolution and detection range for most the IL-8 standard curve (Figure 54 and Figure 54c) and for the entire NE standard curve (Figure 54e, Figure 54f). Therefore, the blue channel was used in all algorithms across most of the examined concentration ranges. However, we found that at concentrations of IL-8 higher than 0.1 ng/mL, the results of the smartphone assay were represented more accurately by the greyscale channel (Figure 54b, Figure 54d). Therefore, the algorithm for IL-8 incorporates the results obtained from the blue channel at lower concentrations (up

to 0.1 ng/mL) and from the greyscale channel elsewhere (up to 3.2 ng/mL). The working curves have swapped axes due to better function fits for data arranged this way.

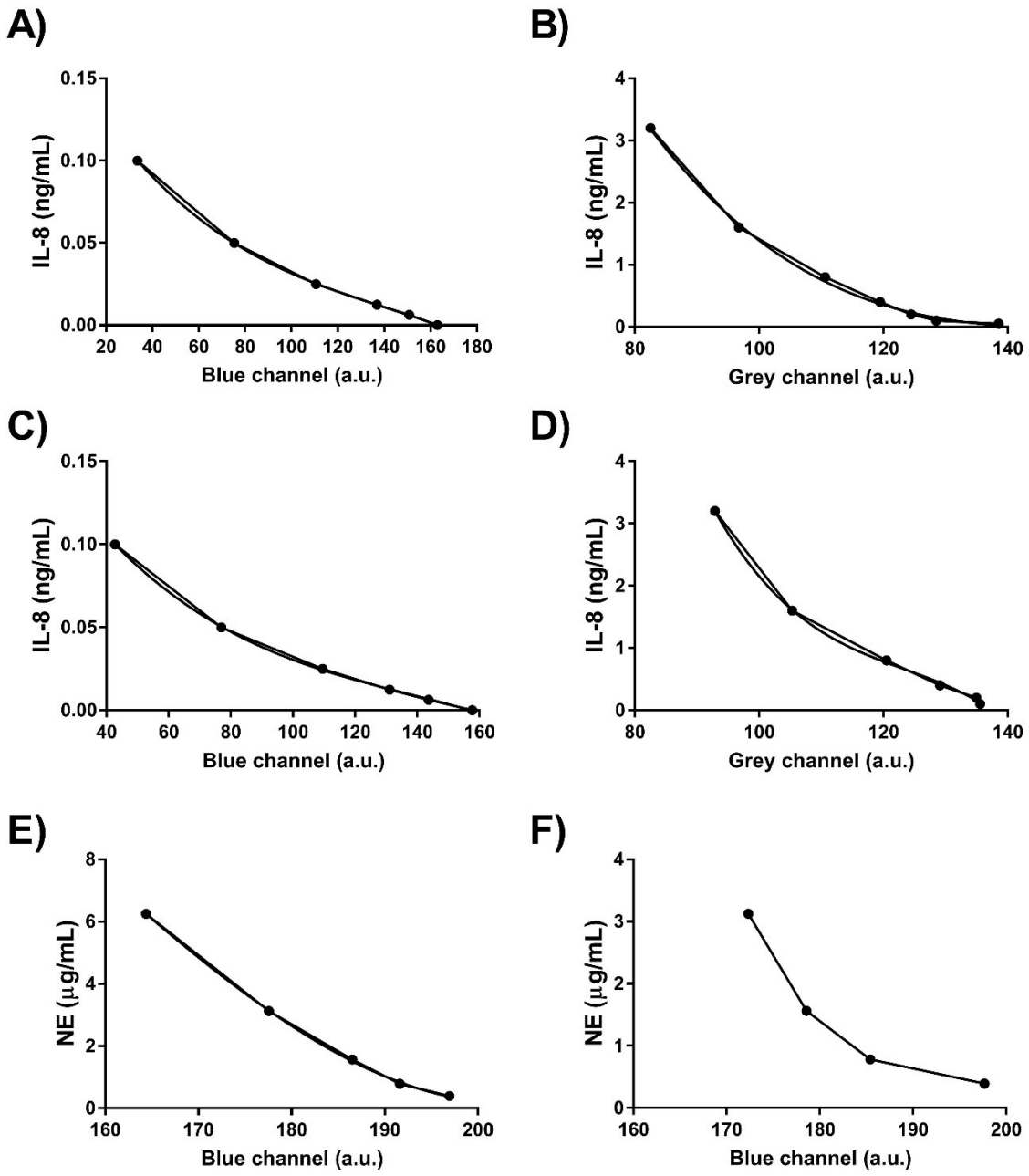


Figure 54. Standard curves generated by the algorithm for (A) Galaxy Note 4, Il-8 assay, blue channel for low concentrations; (B) Galaxy Note 4, Il-8 assay, greyscale channel for high concentrations; (C) iPhone 6S, Il-8 assay, blue channel for low concentrations; (D) iPhone 6S, Il-8 assay, greyscale channel for high concentrations; (E) Galaxy Note 4, NE assay, blue channel; (F) iPhone 6S, NE assay, blue channel

We then used the described smartphone algorithms to calculate the concentration of standard solutions based upon the image analysis of those wells in the 96 well plate. The corresponding standard working curves were also produced for the IL-8 and NE assays using the standard approach of microplate spectrophotometry. These are compared in Figure 55. Close agreement between the standard curves of the smartphone algorithms and the spectrophotometer (Figure 55) was found within the working ranges of the spectrophotometric assays (0 – 0.2 ng/mL for IL-8 and 0 - 6.25 μg/mL for NE). We emphasise that the working range of the IL-8 assay read by the spectrophotometer is lower than typical values in clinical BALF samples (mean 0.23 – 1.11 ng/mL dependent upon infection, (103)). However, our smartphone assay does not have this limitation and it is able to reliably measure IL-8 up to 3.2 ng/mL. (Figure 54b and Figure 54d) For the NE assay, the spectrophotometric readout gave a working range of 0 – 6.2 μg/mL and both smartphone types had a similar working range (0.4 – 6.2 μg/mL).

The smartphone working curves were taken in best-case uniform illumination conditions. The comparison of the working curves (Figure 55) should be treated as the accuracy capability of the smartphone sensor in close to ideal circumstances. The following examination of medical samples was executed in the opposite worst-case situation of non-uniform illumination and with images containing many samples. Non-uniform illumination decreases the accuracy of white paper correction (samples under a slightly different illumination because of a significant distance from the reference point). Many samples on the same image decreases the accuracy of smartphone automatic correction (images contain larger regions of colour influence).

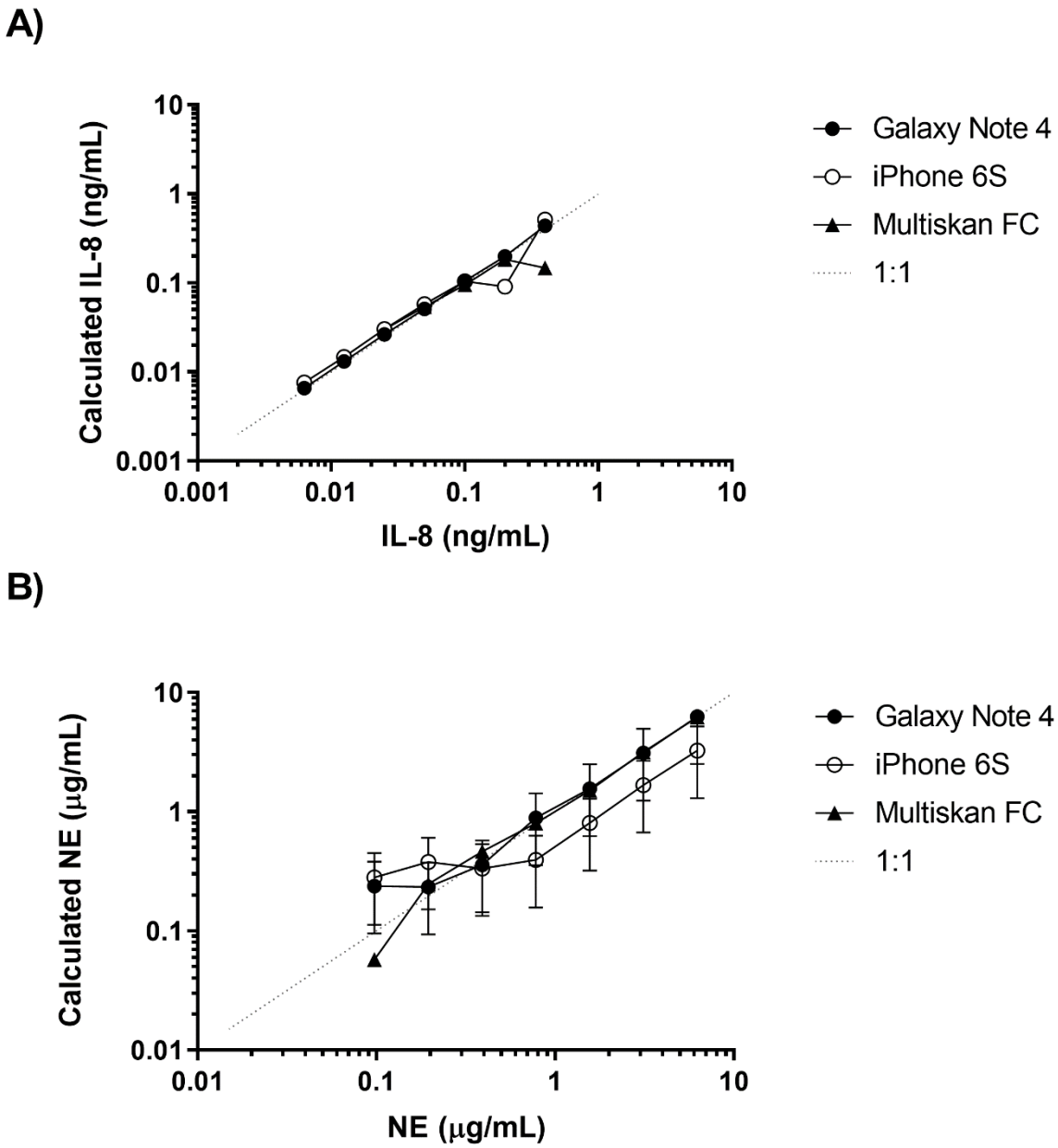


Figure 55. Concentrations of the (A) IL-8 and (B) neutrophil elastase (NE) standards were calculated (y-axis) from the working curves of each device using raw image analysis date (iPhone 6S; Galaxy Note 4) or absorbance values (Multiskan FC) and plotted against the actual concentration of the solution (x-axis). Error bars for the IL-8 plot are small enough to be covered with data points. The last data point for the IL-8 concentration of 0.4 falls out of the Multiscan range. The iPhone working curve for IL-8 changes at the 0.1 ng/mL concentration to grey channel, where the fit has decreased accuracy, hence calculation imperfection.

We finally studied the 32 BALf samples from clinical trials using the IL-8 and NE assays, measured with both the smartphone and standard spectrophotometry. The concentrations of IL-8 and NE calculated by the smartphone algorithms were compared to the gold standard spectrophotometry analysis using Bland-Altman plots

(Figure 56). The Bland-Altman plot is a graphical method of comparing two measurement methods, plotting differences between corresponding points against their average value. This figure shows good agreement between both readout methods in the working range of the spectrophotometric assay for IL-8 (0-0.2 ng/mL). The corner figure presents the same graph with an addition of higher concentration samples measured by the phone, but saturated at the constant value by the spectrophotometer (hence increasing the difference).

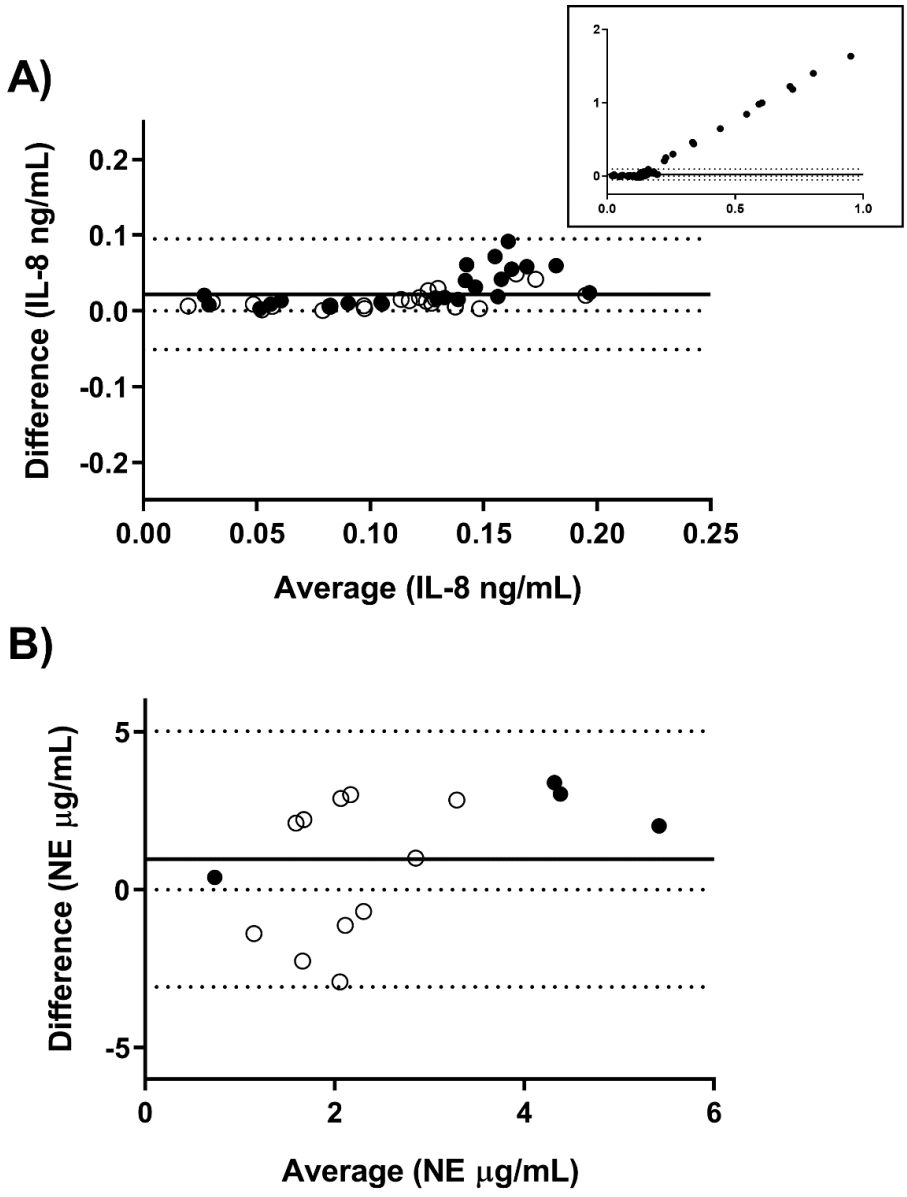


Figure 56. Calculated concentrations of IL-8 and neutrophil elastase (NE) using smartphone image analysis were compared to Multiskan FC spectrophotometer measurements by Bland-Altman. A) IL-8 concentrations as measured by Samsung Galaxy Note 4 (closed circles) an and iPhone 6S (open circles); B) NE concentrations as measured by Samsung Galaxy Note 4 (closed circles) and iPhone 6S (open circles). The thick line represents the mean difference, and dashed lines represent limits of agreement.

Small and regular differences can be observed in a range between 0-0.14 ng/mL for the Il-8 assay. For higher concentrations, readout imperfection, especially for the Samsung phone can be observed to overestimate sample concentration. This results from the change of the working curve at this signal intensity to grey, for which dynamic range is significantly smaller than the blue ones (deviation from the pixel value causes larger change in the concentration readout), and the function approximation slightly deviates from the designated points (Figure 54b). The working curves could be adjusted for the higher accuracy in this region (probably compromising other regions) if this would be preferred. The average difference between the spectrophotometer and the phone readout was 16%. The additional corner figure presents a number of samples that were measured with the phone but could not be measured with the spectrophotometer. The highest measured concentration of IL-8 in clinical sample was 1.77 ng/mL.

Concentration determination resolution can be derived using the working curve and the difference between readouts. The maximum difference from the first experiment of 4.32 counts translates into a difference of 0.0032 ng/mL. The mean error value from the Bland-Altman is 0.018 ng/mL, which can be translated into a difference of 26 counts. This difference will be explained in the discussion. Both are more than sufficient for the clinical use.

For the NE assay, once samples were in the detection range of the device, the samples were correctly recognized as positive (indicating the medical condition), but with low accuracy. This lower accuracy for the NE assay is due to several factors. The overall colour intensity of the NE assay and its variation with analyte concentration is smaller than in IL-8 case, and more difficult to read. The whole range of tested concentrations have a blue channel variation of 30 levels, which is 4 times less than IL-8 for low concentrations and about 6 times in total (including grey working curve). Since our normalisation procedures are basically the same, it is safe to assume similar errors. Average difference between the spectrophotometer and the phone readout was 60%, which is presented on Figure 56b and as large error bars on Figure 55b. This could be addressed by changing the assay chemistry. Secondly, the microplate spectrophotometer used in the laboratory to read assays carries out the absorbance

measurements. The change in diffuse reflectance which is measured by the smartphone reflects varying absorbance, but not in a proportional way and in the high absorbance range the reflectance changes tend to be small. Finally, the most accurate smartphone readout requires that the sample well is located in the middle of the image. In this study, nine sample wells were analysed from a single image which caused additional readout errors for the off-centred sample wells. Due to working curve and interpolation imperfection in the case of the iPhone measurement, these introduced larger errors than the Samsung phone.

2.17.1.Discussion

Here we demonstrate that it is possible to measure clinically relevant levels of CF biomarkers using commercially available assay kits and a smartphone with custom-made software. We tested two common smartphone brands, whose differences in sensor response were managed by using separate working curves for each brand. Higher accuracy of the smartphone readout, especially for the NE assay, is anticipated under standardised illumination that excludes external sources of light. Overall, the accuracy of the smartphone readout is sufficient to distinguish concentrations of IL-8 typically seen in BALF from CF patients (103). For example, elevated IL-8 concentrations above 550 pg/mL may indicate ongoing infection and/or inflammation (103). In the case of neutrophil elastase, free activity is not present in the normal lung and is predominantly undetectable in the first years of life with CF (104). Studies identifying NE activity as a predictor of structural and functional lung disease (105) have utilised spectrophotometric measurements with a lower limit of sensitivity than the smartphone approach (~0.1 µg/mL versus 0.4 µg/mL respectively). Therefore, a measurement within the range of the smartphone standard curve (ie > 0.4 µg/mL) would still serve as a useful indicator of inflammation and potential structural disease progression.

This initial study to assess the potential of smartphones to read laboratory assays does have several limitations. We did not assess a large set of samples, but instead opportunistically imaged assays that were being routinely performed as part of the existing AREST CF early surveillance program. This means we cannot interpret the

ability of the smartphone measurement to replicate prior established associations between these biomarkers and disease.

We performed this assessment under conditions that were on the edge of the rejection algorithms. Because of the large number of samples on each image, the automatic adjustment of the phone was often normalising to the yellow cast of samples, which was not indicative of illumination colour and was causing automatic white balance to fail. Images were taken in changing illumination conditions, which can be observed on Figure 50a (yellow tint from the sun on the left and blue tint from the bulb on the right). Finally, because of the large slide of wells, not every image had the required white sheet of paper with the same illumination condition as the measured sample. If samples were measured in stable illumination condition and in small batches, this would increase accuracy greatly. Initial experiments where samples were measured under different illumination conditions, proved that the normalisation method works well even for significant illumination changes if images allow for a proper execution of the procedure. Results presented on Figure 56a should be considered as close to the worst case scenario. This figure presents only correlation between working range for both devices. We again state that spectrophotometer is not able to read concentrations high enough for clinical BALF range detection (mean 0.23 – 1.11 ng/mL dependent upon infection, (103)), while the smartphone can (Figure 54).

The simplicity of the proposed method causes many potential errors (like previously mentioned irregular illumination, improper photograph, reflections in the sample etc.) to occur and sum up, even if they are recognised and taken care of. These errors, even combined – if significantly smaller than the signal range, will ensure proper readout with sufficient confidence, like in the case of IL-8 assay. If the signal range is small, like in the NE assay, the same errors will make accurate concentration determination impossible.

There remain additional obstacles to the application of CF biomarkers in routine home monitoring. Firstly, lung sampling is limited in the first years of life to invasive methods like bronchoalveolar lavage and regular non-invasive monitoring of lung disease only becomes possible once induced or expectorated sputum becomes tolerable. Exhaled

breath condensate is an alternative modality for non-invasive sampling, although biomarker concentrations can be orders of magnitude lower than BAL or the source airway fluid (106). This poses the additional issue of biomarkers being below the limits of current assay detection, even using expensive laboratory equipment such as spectrophotometers or plate readers to obtain the assay readout. Secondly, assays to detect biomarkers are often technically challenging or require specialist materials to achieve accurate results. However, many of these factors can be overcome with a dedicated approach. For example, the development of mobile testing kits such as the pin prick blood glucose test for diabetes (107), has demonstrated that it is possible to adapt clinical tests to make them available for at-home diagnostics.

Beyond the specific setting explored in this study, our findings further add to the burgeoning field of innovating point-of-care testing for use in low resource areas (108). It has been recognised for several years that improving diagnostics in these areas offers significant potential in improving global health outcomes (109). Many other respiratory diseases also feature NE activity as a key pathological mechanism and of these, chronic obstructive pulmonary disease (COPD) poses an enormous burden (110, 111). Therefore, the refinement of our methodology into a non-invasive biomarker assay of neutrophil inflammation severity that can be performed anywhere but is accurate and prognostically relevant, may be of significant value in complementing COPD care in low resource areas.

2.18. Chapter summary

In this chapter we presented two different readout systems based on smartphones as the main (imaging) sensor. Each system was designed to measure a different type of assay, and compared with measurements from the 'gold standard' device. In both cases the aim was to minimise the number of physical additions to the system, keeping the measurement as simple as possible for the user. Both systems were tested on medical assays, and the colorimetric system was also tested on actual clinical samples. They both achieved very good results compared to gold standard devices, with the colorimetric system additionally extending useful range of the assay by 18 times. Both systems acheived these results at very low cost, assuming the user already has a smartphone. A flow-chart for the fluorescent system was presented in Figure 34, and a flow-chart for the colorimetric system is presented in

Figure 57 below.

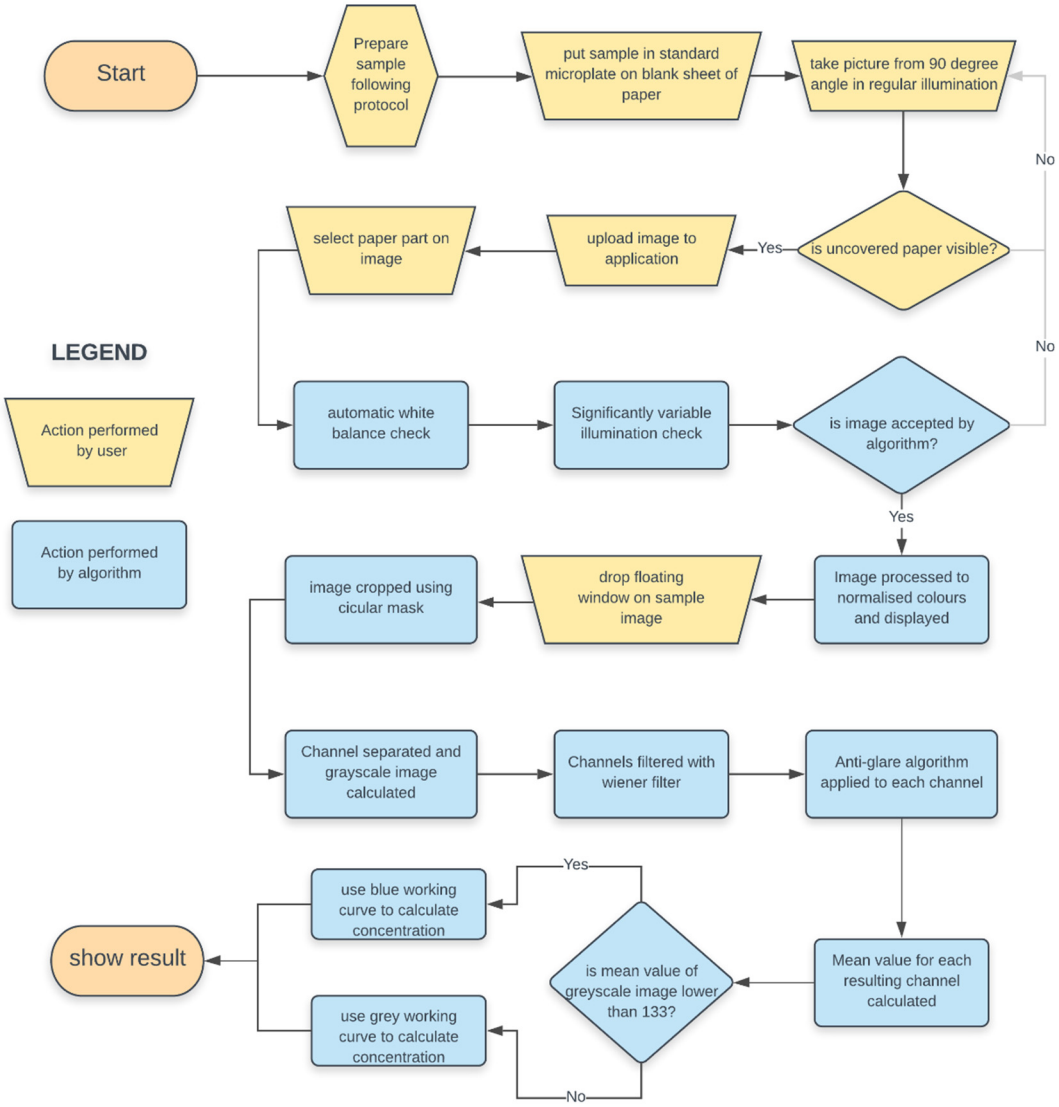


Figure 57. Flow-chart for the colorimetric system.

3. Fluorescence lifetime imaging using a single photon avalanche diode array – Cancer detection based on NADH state

3.1. The use of fluorescence lifetime (imaging) in biology

Fluorescence lifetime imaging (FLIM) is a technique that provides a spatial map of measured lifetimes, typically presented in the form of a colour coded image. FLIM can be applied to measure any process that influences the excited state fluorescence lifetime. It is often used for in vitro, live cell or live animal (in vivo) measurements. In contrast to other fluorescence-based microscopy approaches, measurements of fluorescence decay times depend on relative intensity values, therefore are largely unaffected by many factors that limit steady-state intensity measurements (112): the lifetime of the fluorophore does not depend on excitation power, angle, scattering, obscuring fluids or (within limits) the concentration of the fluorophore.

The most popular use of FLIM technology in biology is FRET-FLIM measurements. In FRET, some part of the energy is transferred in a nonradiative way from a donor fluorophore to an acceptor chromophore. This process causes a reduction of the donor's fluorescence lifetime, that can be measured in comparison to a control group. FRET is only observed at distances from 0-9 nm between two reagents. This distance is about equal to the range of protein size, hence occurs only upon physical interactions between donor and acceptor molecules. By binding proteins, DNA or lipids to acceptor and donor chromophores one can study molecular interactions in single cells.

Since FRET bases its signal on pairs of interacting fluorophores, anything that can be bound with suitable pairs can be imaged using FRET-FLIM, and the wide selection of such pairs broadens the possible applications. Some of the most common pairs are Fluorescein-Rhodamine (113) calcein-sulforhodamine B (114) , and Cy3-Cy5 (115)

One of the phenomenon that can be studied using FRET-FLIM approach is the Protein-Protein interactions scheme. Epidermal-growth factor receptor (EGFR or ErB1) is a

protein for which abnormal expression can signal Alzheimers disease or a variety of tumours. A series of phenomena in living cells were observed using FRET-FLIM, like oligomerization (116) and phosphorylation (117) of EGFR, or propagation (118) and signalling (119). FRET-FLIM has been also proven to be useful in monitoring interactions between key proteins involved in Rab-prenylation disease (120).

Another branch of biological application of FRET-FLIM is DNA analysis. Properly stained, individual nuclei reveal non-homogeneous FRET patterns, allowing them to be visualised (121). Another study showed that SYTOX-dyes (a high-affinity nucleic acid stain) and fluorescently-tagged proteins can be monitored to observe DNA-protein interactions in single cells (122).

It is possible to perform quantitative FRET-FLIM analysis. FLIM can provide a quantitative estimation of the FRET signal arising upon the formation of a protein complex, and can also be utilised to measure intermolecular distances. Assuming two lifetime states: a non-interacting unquenched donor and an interacting quenched donor, deconvolving FLIM data allowed observation of the phosphorylation state of ErbB1 receptors tagged with green fluorescent protein in breast cancer cells (123) and phosphorylation of EGFR (124).

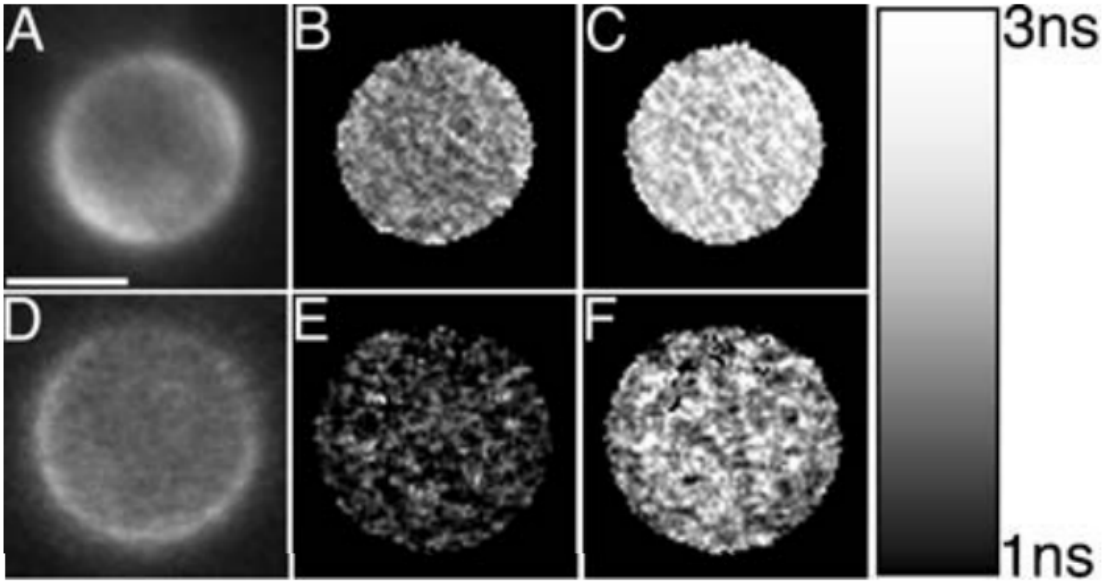


Figure 58. FLIM-FRET analysis of Cowpea protoplasts co-expression for various fluorescent proteins, expressing different lifetimes. (125)

Using fluorescent dyes that react with specific ions allows measurement of local concentrations of such ions with FLIM (Figure 58). For instance, calcium can be quantified by measuring increased fluorescence lifetimes of calcium-green or Quin-2 (126) or Calcium-Crimson upon calcium-binding (127). Dyes with a pH-dependent lifetime like BCECF enable identification of microdomains of two distinct pH values within the uppermost layer of the epidermis (128), or carboxyfluorescein to image deep microbial biofilms (129). The usefulness of the fluorescent probe RTDP for the quantitative imaging of oxygen in single cells was investigated utilizing FLIM. The results indicate that the fluorescence quenching by oxygen is a dynamic quenching process but simple calibration techniques allow for the oxygen concentration quantification (130).

The ethidium blue fluorophore is an example of a dye which when bound to DNA expresses a different lifetime than when bound to the membrane. This difference was observed using FLIM to discriminate between these two (131). The same phenomenon was used to study the chromatin-conformation in living cells using green fluorescent nucleic acid stains like SYTO13 (132) or YOYO-1 (121).

As a final example area, FLIM techniques were implemented in various skin and tissue cancer detections, targeting different autofluorescent molecules. In 1993 scientists found that tumours (ovarian carcinoma) marked with photosensitisers reveal strongly extended fluorescence lifetimes (133). In 1997 a different group used a CW laser with a wavelength of 514 nm to excite flavins in the human bladder, using frequency-domain FLIM (section 1.8) to record areas with differences in lifetimes, indicating tumorous tissues (134). Scientists from London in 2013 showed that malignant and benign breast cancer tissues can be recognized by the difference in autofluorescence lifetime (135). In February of the next year, a different group from University College London published a paper where they distinguished lifetimes between two spectrally-identical molecules: NAD (which is a key determinant of cellular energy metabolism) and its phosphorylated form NADP (which plays a central role in biosynthetic pathways and antioxidant defence). The results suggest that enzyme-bound NADPH has a significantly longer fluorescence lifetime than enzyme-bound NADH within the cellular environment. The proportion of enzyme-bound NADPH and NADH present in

live tissue (on cellular level) determined the lifetime of their combined fluorescence decay (136). One more British group from Imperial College London presented a non-microscopic system (with a field of view larger than a centimetre) to detect cancer regions in tissues, fitting a single-exponential curve to the externally gated (using gate signal intensifier) signal, recorded with a CCD camera (137).

3.2. Single Photon Avalanche Diode (SPAD)

In our work, we will measure fluorescence lifetimes using a Single Photon Avalanche Diode (SPAD) sensor array. SPAD sensors (section 1.8) have gained attention because they are small, easy to integrate in a large array, and are insensitive to magnetic fields, making them suitable for medicine and space applications (138). From the early 2000s, it was possible to fabricate SPADs in standard CMOS technology. Researchers started to develop compact and cost-effective SPAD image sensors in different CMOS technology nodes, typically optimising one parameter sacrificing efficiency of others. Dozens of chips in different CMOS production processes like 800 nm (139), 350 nm (140), 180 nm (141), 130 nm (142), and 90 nm (143) were presented. A fair comparison between them is not a simple task because of lack of unified tests and rules. There are, however, some important standard properties that typically are measured and presented.

- Photon detection efficiency (PDE) is defined as the ratio of the number of detected photons to the number of photons incident on the photoactive area. This ratio depends on absorption probability and on triggering efficiency (144).
- Noise is a complex matter and consists of the signal's own Poisson statistics as well as false counts. False counts that are uncorrelated to real counts are called the Dark Count Rate (DCR), while correlated false counts come from optical and electrical crosstalk and afterpulsing (an improperly induced count caused by the previous readout) (144).
- Dead-time is a delay enforced after each measurement in order to reduce afterpulsing, at the cost of lowered count rate.

- Timing jitter (photon-timing precision) is the statistical spread of the output pulse compared to the true photon arrival time, and is quoted by the Full-Width at Half Maximum (FWHM) of the distribution histogram (145).
- Fill factor is a ratio of active, photon sensitive area of the sensor compared to the whole area.

SPADs optimised for best performance in any subset of these standard properties might be desired for different applications.

Another important factor when comparing different SPADs is the number of single pixels in the array. The great advantage of the CMOS process over regular custom made SPADs is its scalability, enabling manufacturing of chips with a greater number of pixels. Different sizes and shapes of arrays have been published, starting with the small ones consisting of 8 pixels (146) or 32 pixels (4x8 array) (27), the most commonly seen 32x32 pixels (147, 148), through to more irregular shapes like 4x112 (149) or 1024x8 (150), finishing with the largest arrays like 128x128 pixels, which sums up to over 16,000 individual pixels (151).

SPAD sensors are commonly used in variety of applications where photon counting or photon arrival time is required. In 2010 small SPAD array was used to present an approach to high-throughput Fluorescence Correlation Spectroscopy (FCS) which enabled a one order of magnitude improvement in acquisition time compared to multispot confocal microscopy (146). SPADs have been proven to work in Positron Emission Tomography (PET), providing required low-noise performance, precise timing and very fine control over threshold levels (152, 153). The previously mentioned sensor (150) was demonstrated to read Raman signals as well as serving as a Laser-Induced Breakdown Spectroscopy (LIBS) system. SPADs were also reported to serve as three-dimensional imaging laser radars featuring single-photon sensitivity (154) (Figure 59), or phase-resolved imagers for extracting either three-dimensional depth-resolved images or lifetime maps, by measuring the phase shift between a modulated excitation light and the reflected photons (155). Finally, SPADs can be used to detect fluorescence lifetime, as we discuss in the next section.

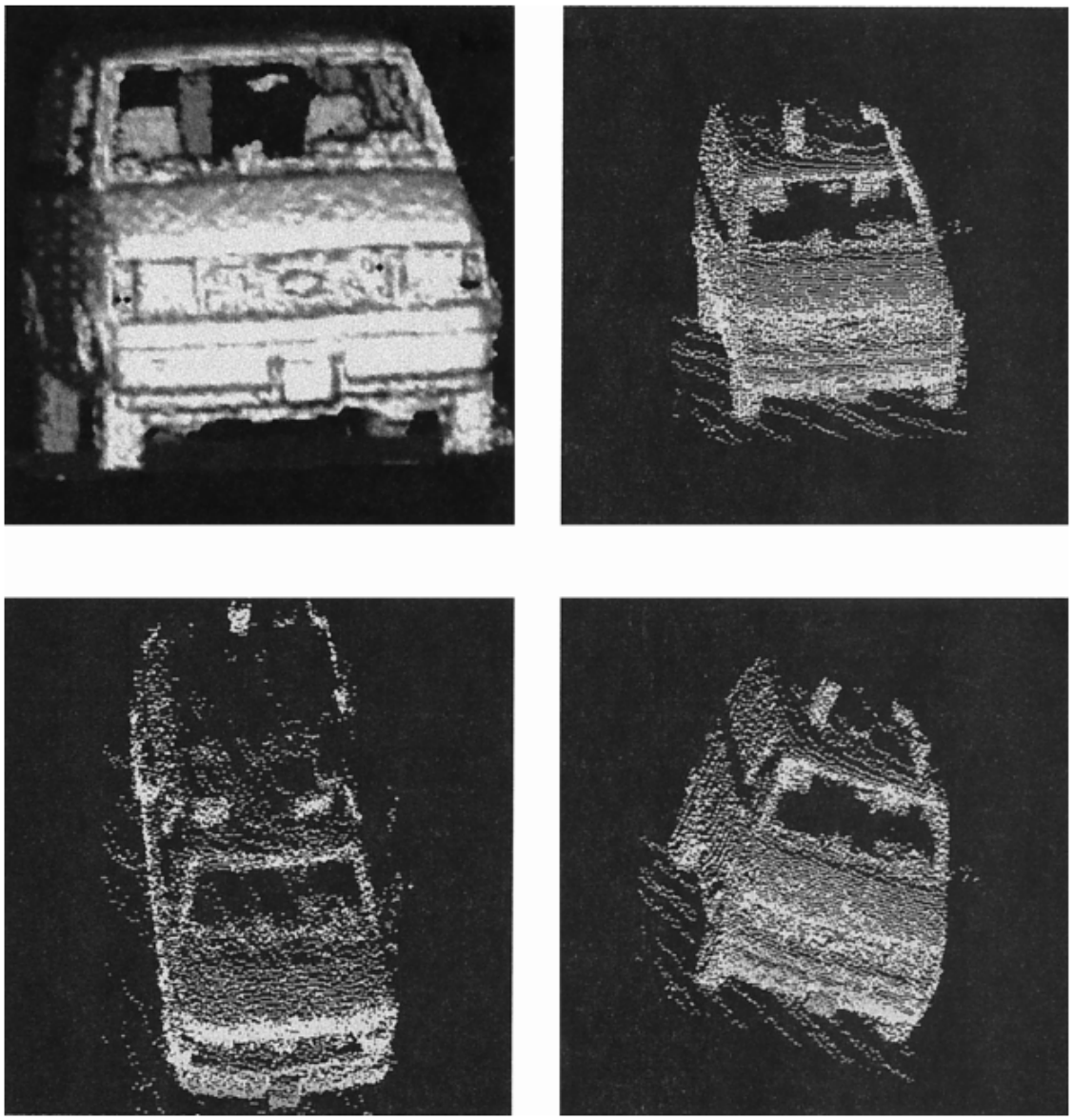


Figure 59. Image of the Chevy van computed from range-coincidence processing acquired using the SPAD sensor.(154). An example of using SPAD as a 3d imaging system.

3.3. FLIM on SPAD

Fluorescence lifetime imaging performed on SPAD sensors has been presented before, though mostly as a proof of concept. Because of differences between sensors, there is a significant variety of the designed scheme of operation, suitable applications, and level of performance. In 2009 a group from Italy presented a 64 pixel SPAD, with strictly limited number of 4 non-uniform gates in predefined positions, presenting a measurement of Alexa 488 fluorophore with lifetime of 4.1 ns (156). In 2010 a group from the UK presented more extensive lifetime determination research, measuring a

number of popular fluorophores in a range of 1.6 to 16 ns of decay time. The effective pixel resolution of 32x16 did not allow for wide field imaging (157). The same group proposed a different lifetime calculation method called the 'center-of-mass method" which was a hardware implemented method suitable for a single-exponential decay time (158). Finally, in 2014 a group from Switzerland presented the largest SPAD array to date, with resolution of 512x128 pixels, reporting measurement of four popular fluorophores with single-exponential decay lifetimes (159). We used this sensor to build our FLIM SPAD system

3.4. The concept

A typical FLIM system using a single detector employs a laser scanning microscope and time-correlated single photon counter (TCSPC). While this is a powerful technique for cell biology, its small field of view (typically less than 1 mm) and rather long scanning time makes it impractical for imaging the larger areas of tissues. This issue, combined with the relatively high cost, makes this approach impractical for many (especially biological) applications.

We decided to build a FLIM system capable of detecting lifetimes in the (sub)nanosecond regime, using the largest image sensor based on single-photon avalanche diodes (SPADs) published to date (with a resolution of 512 x 128 pixels) (159). By taking a series of gated measurements and changing the gate delay with respect to the excitation pulse, fluorescence lifetimes can be retrieved for each individual pixel. Over 65 thousand independent SPAD pixels creates a detailed spatial map of fluorescence lifetimes in the field of view of the sensor. Our final system has a large field of view of over 20 cm^2 which allows for imaging of significant areas from a distance of about 20 cm. Detection of single exponential lifetimes allowed for lifetime-based fluorophore determination, while a bi-exponential fit made determination of mixed fluorophores possible. In the case of bi-exponential measurement it was possible to estimate a quantitative ratio of two fluorophores. We used our system to measure autofluorescent signals of reduced nicotinamide adenine dinucleotide (NADH) in two forms – free and bound to L-Malate Dehydrogenase (L-MDH), where the second form has prolonged lifetime. These two forms naturally exist in the human

body and are related to the cell metabolism. Increased metabolism, leading to increased level of bound form of NADH is a cancer determinant (160), and so our quantitative measurement could be used as a real-time, in-vivo measurement technique to visualise cancer tissues.

In the next sections, we present the basic design of our system, the detailed principles of the SPAD array operation and its full characterisation, and our investigations and validation of the best ways to retrieve single- and bi-exponential fluorophore lifetimes.

3.5. The Setup

All our measurements were based on a similar illumination and timing system, but using different samples for different experiments. A simplified scheme of the first version of the setup is presented in Figure 60. A modelocked 10 ps pulsed 532 nm or 355 nm laser illuminated the sample as well as a photodiode that provides a timing signal to synchronize the SPAD readout with the laser pulses. The sample emission signal is filtered to block the excitation wavelength and collected by simple optics attached to the FPGA motherboard board that houses and controls the SPAD array. Now we will discuss all the elements in more details.

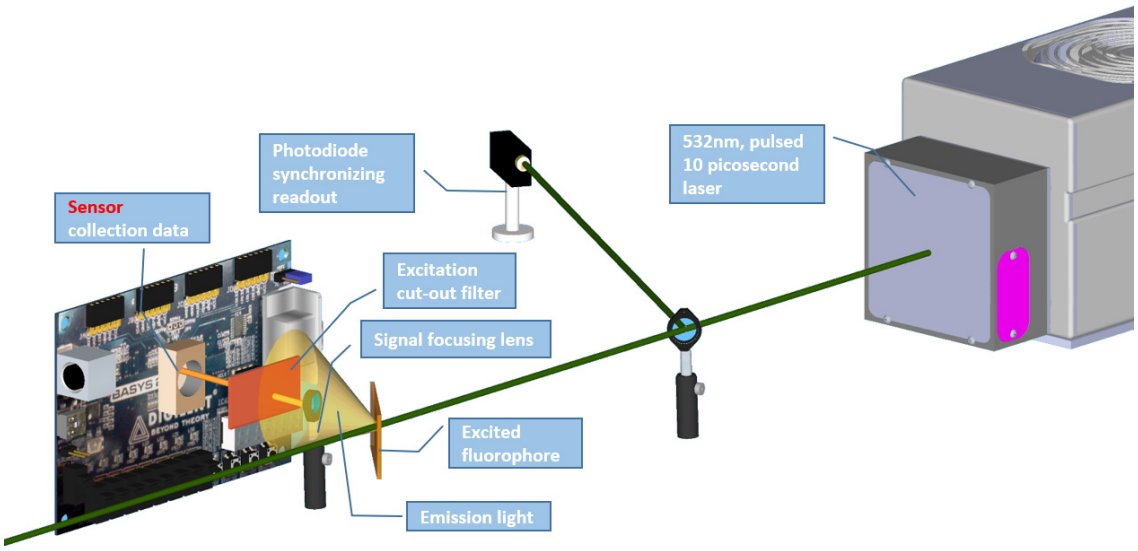


Figure 60. Simplified image of the initial setup for measurements of a fluorescent plate with the SPAD-FLIM system.

3.5.1. The SPAD sensor

In our studies we used SwissSPAD sensor (159). It is fabricated in a standard high-voltage 0.35 µm CMOS process. Each pixel has a side of 24 µm with a round sensitive area with a diameter equal to 6µm. The nominal fill-factor of 5 % was increased by microlenses to up to 50 % (depending on the light distribution) of the pixel area for the collimated light. The Photon Detection Probability with microlenses reaches 46 %. it's target wavelengths are between 400 – 850 nm. It has 1-bit memory and is driven by a global shutter. The circuit of a single pixel is presented in Figure 61 (top)

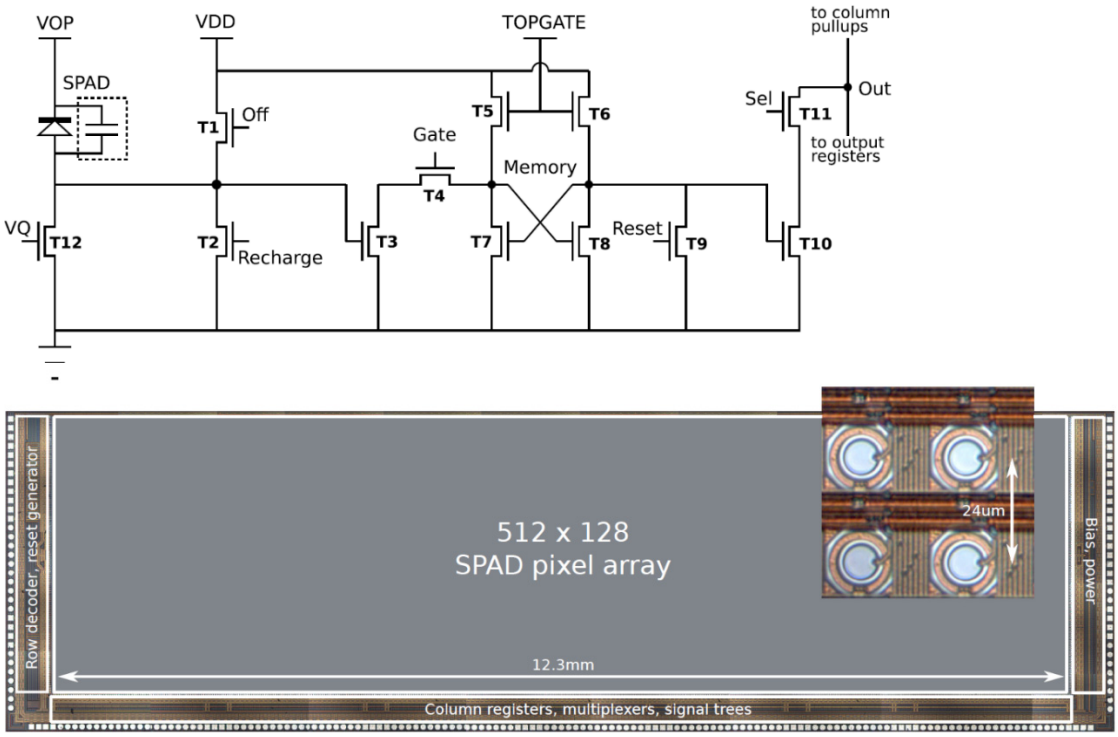


Figure 61. Top: Transistor level schematic of the pixel circuit. The SPAD is shown together with its junction capacitance. T12 can be used for passive quenching and separates the SPAD from ground. Transistors T1 and T2 control the SPAD bias and are used to switch the SPAD on and off. T4 controls the access to the NMOS-latch formed by T7 and T8, loaded by T5 and T6. T9 is used to reset the storage latch, previously set by T3. Finally, T10 is used to transfer the memory value to the output line through the row select transistor T11. Bottom: SwissSPAD die micrograph with the SPAD-array in the centre and logic on three sides.

Transistors T1, T2 and T4 forms a gating circuit. Each cycle contains three phases – off, recharge and gating phase. A cycle starts with gating turned off, followed by recharge phase that restores SPAD to Geiger operating conditions and finally gating phase, where the pixel can be triggered by an incoming photon (gating will be described in more details in following sections). T4 prevents registration of false events from

turning the SPAD on and off. Transistors T5-T8 form a 1-bit memory block that store whether or not the pixel was triggered during the gating phase, and transistors T9-T11 are responsible for data readout.

An array of 512 x 128 pixels is formed on a chip as shown in Figure 61 (bottom). The chip size is 13.5x3.5 mm^2 and it is placed on a FPGA board that runs it. To achieve high timing precision such that all pixels are simultaneously active, matched signal trees are used. The SwissSPAD can generate a maximum data rate of 10.2Gbps over the 128 output lines. Because of the USB2 limitation of 480 Mbps, grayscale images are first stored in large SO-DIMM buffer DRAMs and subsequently transferred to the computer. The most important parameters of the chip are presented in Table 3. For more detailed information regarding the sensor, please refer to (159).

Table 3. Basic parameters and performance figures of the SwissSPAD sensor

Chip size:	13.5x3.5 mm2
Technology:	AMS HV 0.35 µm 4M
Resolution:	512x128
Pixel pitch:	24 µm
Fill Factor:	~40% (with microlenses)
PDE:	~15% (with microlenses)
Dead time:	100 ns
DCR (mean for pixel):	206 Hz
Target wavelength:	400-850 nm

The SwissSPAD chip is based on a FPGA board that provides computing power with fast clocks to run the chip and additional peripherals like internal memory or I/O ports. However, this integrated part is not the full system, but just a receiver. For proper operation, external elements like power supplies, a laser, a photodiode or filters are still required. The sensor is not a commercial product, but rather a prototype provided to us by the AQUA research group from École polytechnique fédérale de Lausanne. It is not fully characterised nor repeatable, and hence before attempting reliable readouts, optimal operating conditions needed to be established. Figure 62 presents an early version of the setup connected to adjustable power supplies to control the behaviour of the SPAD, computer power supply to power the FPGA board, and a fast photodiode to synchronise the readout with laser pulses.

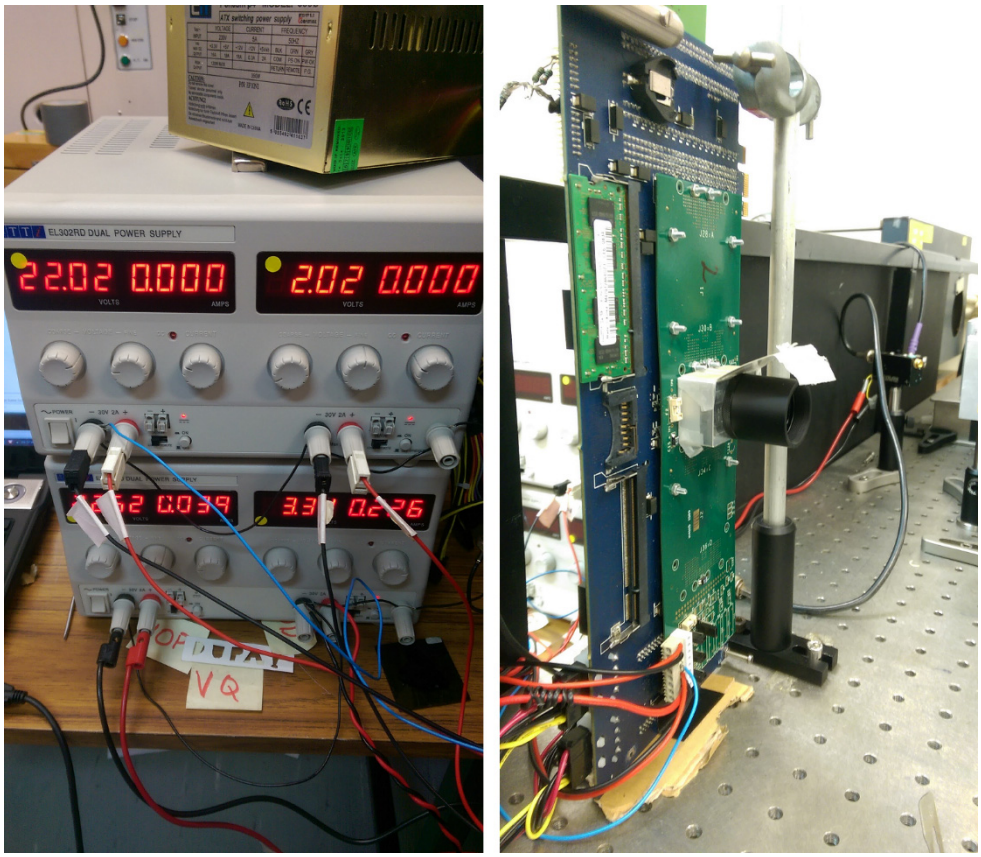


Figure 62. The SwissSPAD sensor based on the FPGA motherboard, with cables connected to adjustable power supplies and a computer power supply. Early version of the optical lens is attached. Fast photodiode synchronising the readout pattern with laser pulses visible in the background.

3.5.2. The laser and the photodiode

The laser used in our setup was a pulsed, 10 picosecond laser with 80 MHz repetition rate and maximum power of 2 W (Spectra-Physics VNGD200-80-HM532). The laser excites the fluorophores, with the short pulse duration providing a sharp start-point for their exponential decay. A beam splitter was used to deliver some of the laser beam to a photodiode, which was terminated with 50 Ω and was adjusted to produce a response with the maximum value of 3.3 V. This signal was used to trigger the SPAD cycles, synchronising them to the excitation time.

Details of the setup were changing depending on the experiment and will be described in the appropriate following sections. In general, the laser synchronised the system readout pattern and illuminated samples; emitted fluorescence was filtered for the excitation wavelength and collected by optical elements that gathered it on the sensor.

For final experiments, we generated a wavelength of 355 nm from the same laser, and used this to excite NADH samples. To generate 355 nm, we used the fact that the laser generates accessible output beams at both 532 nm and 1064 nm. The two pulses were spatially and temporally overlapped using translation stages and a combining mirror that was highly-reflecting (HR) at 1064 nm and highly transmissive (HT) at 532 nm. The pulses were then focused into a Beta Barium Borate (BBO) crystal, cut for phase matching of the sum-frequency generation process for the 355 nm wavelength. After the BBO the beams are collimated and spatially separated using a Pellin Broca prism (Figure 63).

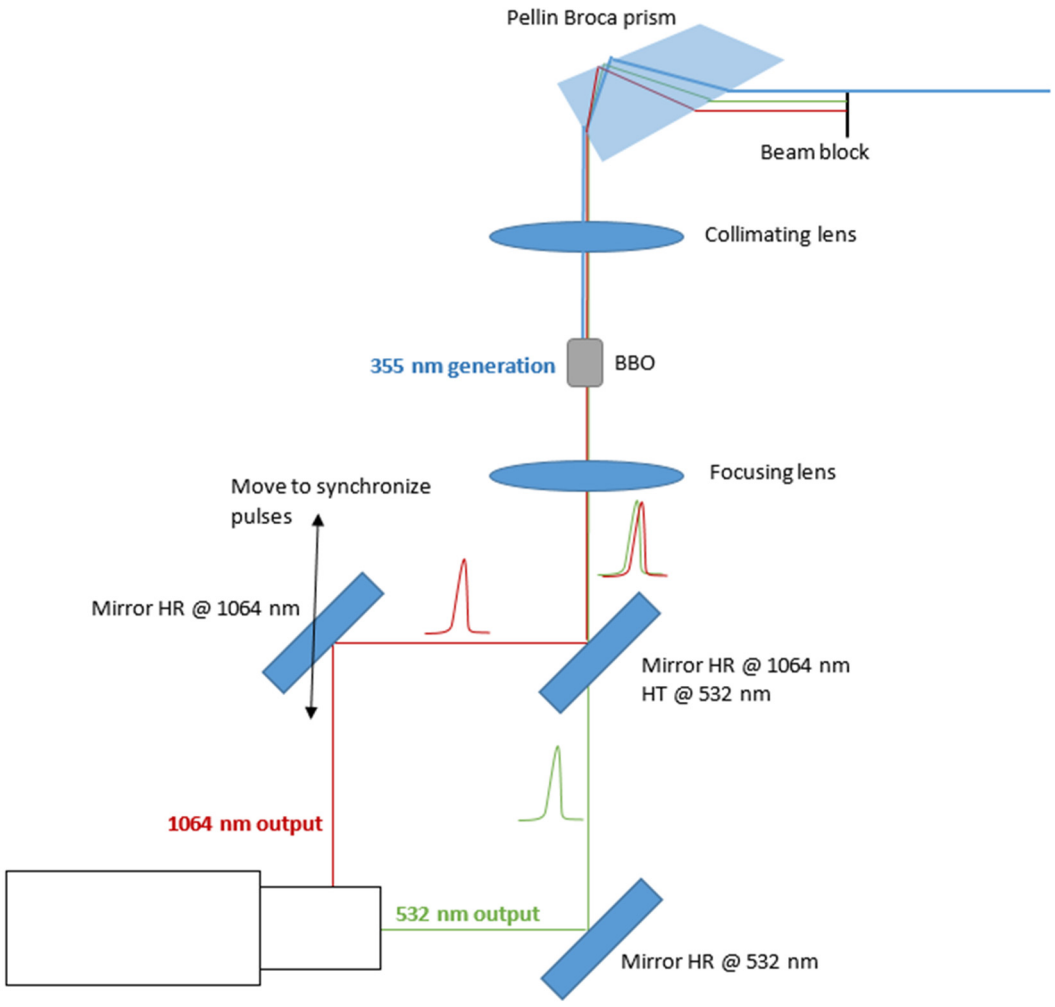


Figure 63. Scheme of the laser setup for generating the 355 nm wavelength.

3.6. The principle of operation of the gating mode of the SPAD sensor

The SwissSPAD sensor can operate in two basic modes: real time filming and time-gated measurement. We used the time gated mode to develop our FLIM system only. Time-gated mode of operation is strictly synchronised with laser pulses.

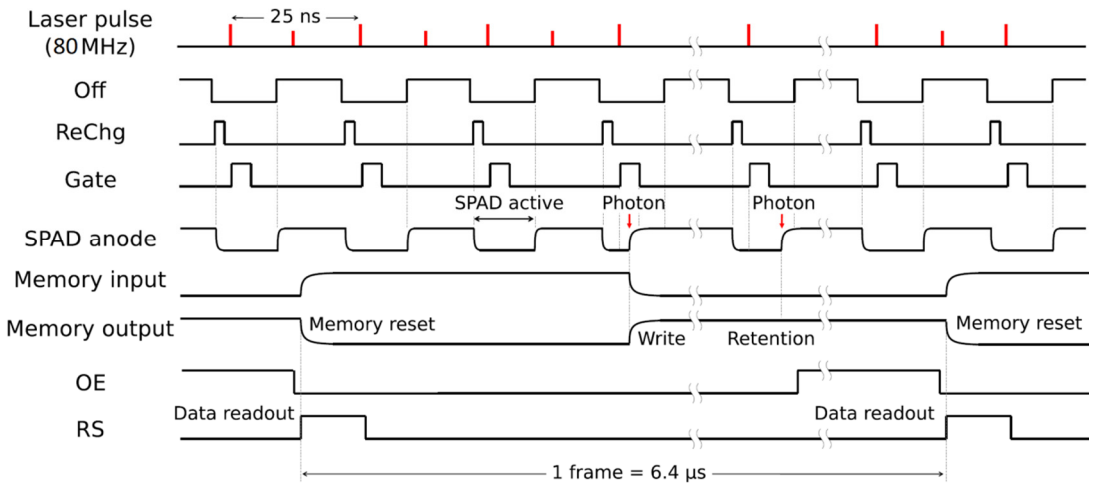


Figure 64. Timing diagram for pulsed illumination imaging. The gating statuses (Off, ReChg, GATE) are derived from the reference clock supplied by the illumination system (here a picosecond laser with 40 MHz repetition rate). Output enable (OE) and reset (RS) signals are used to control the chip readout.

Figure 64 presents the timing of basic internal timing components in respect to the laser pulses. Each gating cycle is constructed of three stages: off state, recharge state and gating time. The timing of the gate with respect to the laser trigger could be adjusted to allow sampling of the fluorescence as a function of time after the laser excitation. The laser we used in our setup generated pulses at a rate of 80 MHz which corresponds to 12.5 ns between pulses. This time was shorter than the minimum time for the SPAD's full readout cycle, and so we decided to trigger the SPAD at half this frequency. By triggering off every second laser pulse, we treated the laser as 40 MHz, with 25 ns between consecutive SPAD cycles. This procedure resulted in there being two laser pulses within each cycle, and so depending on where the gate was positioned with respect to the trigger, either one could be detected. For generating the pulses for the gating sequence, the FPGA internal fast clock with a frequency of 200 MHz was used. It allowed us to adjust the timing of the gating signals at multiples of 2.5 ns, with an additional fine shift (a single gating sequence step) with a resolution

of a 256th clock period or around 20 ps. For a full period of the cycle (25 ns) we can thus set 1280 different temporal positions (frames) of the gating sequence with a 20 ps resolution.

There are a number of settings that influence the sensitivity and readout capabilities of the system, and we will now present the most important ones. All the data will be presented for a single pixel; remember that all these procedures are performed simultaneously on all pixels in the array.

3.6.1. The single readout

During one laser cycle of 25 ns one single readout is performed. The sensor detects photons only for during 'gate' state of each cycle which lasts for about 30% of the time of the cycle (depending on the setting). Each pixel of the system has a 1-bit memory, which means that during one cycle it can only record the detection or lack of detection of a photon. If two photons hit the sensor during one cycle, this still counts only as a single detection and our signal count becomes inaccurate: we want to keep the probability of photon detection low enough to avoid this situation. Because the active area of each pixel is about 40% (increased by microlenses), the time of the 'gate' state is also a fraction of the whole cycle and biological signal levels are relatively low - photon detection probability for each pixel is low, of order 1%.

3.6.2. Repetitions

Because the probability of a photon detection during one cycle is significantly smaller than 1%, we performed a number of detection repetitions to determine a single 1-bit detection. The repetition parameter states how many cycles of the same measurement will be performed to try to detect a single photon. One single detection during all the repetitions will set the detector to positive state (detection). Again, if more than one repetition sees a photon this will result in a lost count. The repetition variable has the biggest influence on the acquired signal amplitude as well as the signal to noise ratio. We used this value to keep the readout signal level in the proper range (proper range will be described later).

3.6.3. Max counts per frame value

To achieve higher dynamic range than just 1-bit, the procedure described in previous sections is repeated a number of times, specified by the Max Counts Per Frame (MCPF) parameter. The final pixel count is the sum of detections, and so the MCPF value defines how many potential levels can each pixel achieve - the dynamic range of the readout. If all the parameters are kept at correct levels, the outcome will give a statistical measure of the single-cycle detection probability, representing the strength of the measured signal.

On graphs in Figure 65, MCPF was set to 1000. Traces achieved maximum intensities of about 70% the range. For later measurements (unless stated otherwise) we used the value of 256 levels (8-bit) as a compromise between sufficient accuracy and the speed of operation. Increasing MCPF has a linear effect on increasing the collection time. For the MCPF value of 256 (256 repeated 1-bit measurements) and repetition value of 200 (200 attempts to see a photon per 1-bit measurement), a total number 51200 measurements over 51200 detection cycles are performed to achieve a single value for one gate timing position. This procedure, performed for all the pixels in the array, produces an 8-bit greyscale image of the size of the array.

3.6.4. Shifting the gating sequence

The measurement procedure is performed for a specific single gating cycle position. The recorded signal depends on how the light on the sensor changes with time, where the gate is positioned and how long the gate is ‘open for’. To determine how the signal depends on the time after the laser pulse, we perform gate shifts.

For the full time-gated readout, the gating is moved in turn to all positions across the cycle by steps of 20 ps each, repeating the measurement procedure and recording the values for each position. The previously calculated number of single measurements, for one gate position equal to 51200, is performed 1280 times to map the whole 25 ns period. The total number of single measurements for this example equals 65,536,000. This produces 1280 greyscale images of signal intensity for different time delays after the laser pulse. For a single pixel plotted against shift, the plot as we can see in Figure

65 can be obtained. To understand the shape of this plot we must consider the process of convolution.

3.6.5. Convolution of the laser pulse and the gate of the device

Mathematically speaking, convolution is an integral of the overlap of two functions as one of them, time-reversed, is shifted across the other one. The resulting third function expresses how the shape of one function influences the shape of the second one. This can be written as:

$$[f * g](t) = \int_0^t f(\tau)g(t-\tau)d\tau, \qquad (56)$$

where f and g are functions and τ is a time difference.

For engineers, convolution is typically used to describe the output of a linear, time-invariant (LTI) system, such as ours. Shifting the gate across the cycle as described in previous section is the execution of convolution theory. For us, the convolution tells us how the signal on our SPAD as a function of gate delay depends on the temporal shape of the light signal and temporal shape of the gate. If a system's gate were much shorter than the measured signal, the operation would be basically a sampling of temporal shape of the light signal, taking a continuous time signal into discrete values representing an approximation of the signal in the digital environment. Unfortunately, in the case of our system the gate (10 ns) is longer than the fluorescence decay signals, which requires use of convolution to interpret the result.

This dependency, however, works both ways. If the second function is much shorter than the first one – the resulting third function will be basically sampled first function. This means if we observe a signal lasting three orders of magnitude less than the time of the gate, like a laser pulse of our picosecond laser, the resulting function will present a very good approximation of the temporal shape of the gate of the system. This shape can be called the Impulse Response Function (IRF) of the system, or just the shape of the gate. Examples of such gate shapes acquired in this manner are presented in Figure 65. Knowing the gate shape allowed us to understand the measured fluorescence signals as the convolution with the now-known gate function.

3.6.6. The gating sequence

The electronic gating sequence is a combination of the states presented in Figure 64 for one cycle. A FPGA's fast internal 200 MHz clock allows to place the edges of the gating signals at multiples of 2.5 ns. With the 25 ns cycle time this gives us 10 possible state changes locations for each (recall that fine timing adjustments are also available to shift the whole timing pattern with a resolution of 1/256 of this clock - 20 ps). Certain off and recharge times are necessary for a correct electrical charge handling to ensure a proper readout. We can choose the length of the gate. The gating time is a trade-off between photon collection time and sampling resolution; the shorter the gate, the better resolution. The gate length, however, also influences the internal readout behaviour and reliability, limiting how short the gate can usefully be set. We tested over 20 electronic gating sequences, observing the IRF of the system, and chose the one with the most reliable outcome such that the gate shape was repeatable, high amplitude, and with steep edges. Figure 65a presents a few gate shapes resulting from different sequences for comparison. Figure 65b shows the shape of the selected optimum gate on a separate graph, and Figure 65c shows the associated optimal gating sequence. It is worth realising that while the gate should be 10 ns according to the state signals, in reality it is shorter by about 20 %, lasting for about 8 ns in our case. Recall that due to laser frequency being halved for the purpose of sensor synchronisation, the gate is mapped out twice for the full range of gate shifts because of two laser pulses falling into each cycle.

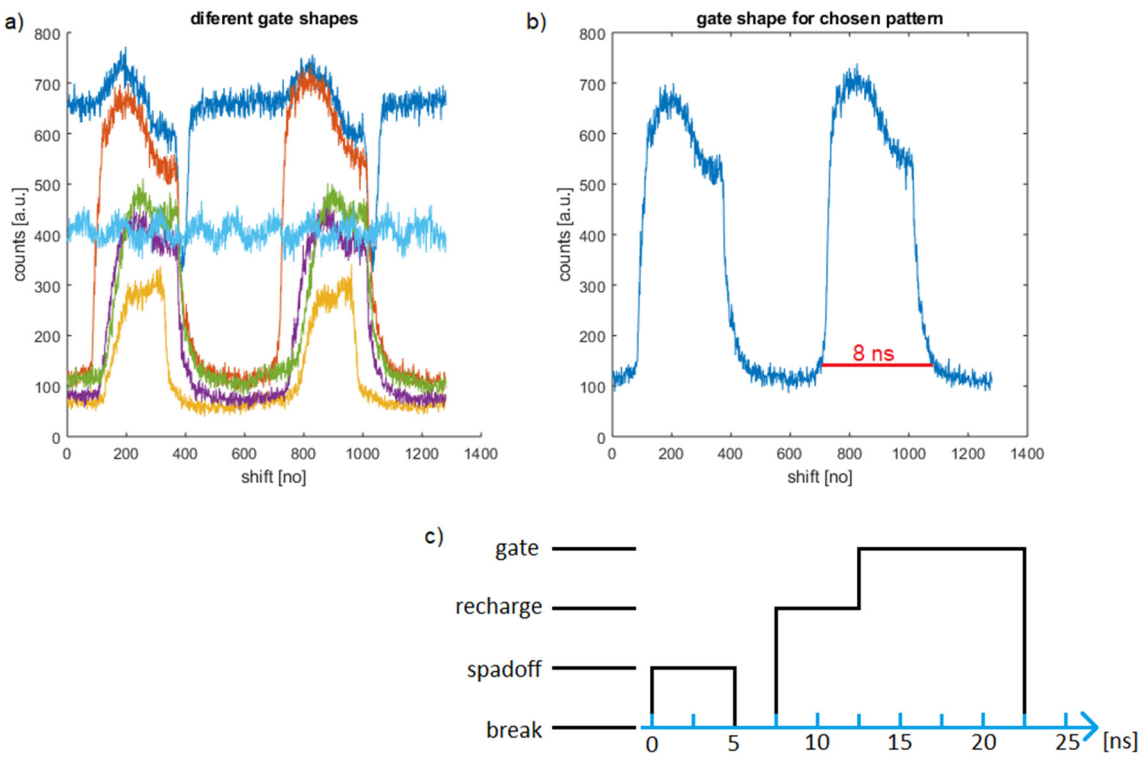


Figure 65. a) Gate shapes acquired using gating sequences. b) Gate shape for the chosen sequence. c) Representation of the chosen gating sequence.

3.6.7. Correction factor and signal amplitude limit

With an increasing number of photons hitting the sensor, the probability of sensor saturation also increases. The situation where signal strength changes in time might lead to the improper shape of the function, if part of it were saturated (keeping constant saturated value, where it should be changing). The phenomenon is even more deceiving in the case of a 1-bit sensor where the total number of summed counts might not hit the maximum value (not all 1-bit measurements are 1), but at least some single 1-bit measurements that are constructing it are 'locally saturated', meaning two photons arrived but only could be counted. This can lead to a change of the shape of the trace, without a warning signal of clearly saturated readout. We observed that gates were changing in shape with increased signal intensity, explained by this local saturation. The gating sequence chosen in the previous section is optimal, however the shape of the chosen gate presented in Figure 65 is incorrect due to sensor saturation.

Dainty and Shaw presented a model (161) proving the proportion of cells in a large array that receive n photons is governed by

$$P(n) = \frac{r^n e^{-r}}{n!}, \tag{57}$$

where n is a number of received photons, r is the average number of photons per pixel.

In our case of single photon detector, we record a 1 if it receives one *or more* photons, which can be presented as:

$$P(\geq 1) = \frac{count}{MCPF} = 1 - P(0). \tag{58}$$

where $count$ represents number of real counts out of $MCPF$ possible levels. Since $P(0)$ is given by:

$$P(0) = e^{-r}, \tag{59}$$

we then find:

$$\frac{count}{MCPF} = 1 - e^{-r}, \tag{60}$$

For $r \ll 1$, $count/MCPF$ is equal to r, and our measured intensity is linear with intensity. For higher r, our count becomes saturated. Inverting the equation above, we can find the relation between r and our count:

$$r = -\ln\left(1 - \frac{count}{MCPF}\right). \tag{61}$$

For Max Counts Per Frame equal to $MCPF = 256$, we can derive the equation for the correction of saturation problem, for our sensor in the typical scenario:

$$correctedcount = -256 \times \ln\left(1 - \frac{count}{256}\right). \tag{62}$$

This formula will be applied to all the measurements performed by our system. Because it changes the readout value in a nonlinear manner, it slightly changes the function shape. Figure 66 shows this behaviour on the example of the gate shape. The shape of the gate and the acquired trace is crucial for system accuracy. We will discuss this issue in the following sections.

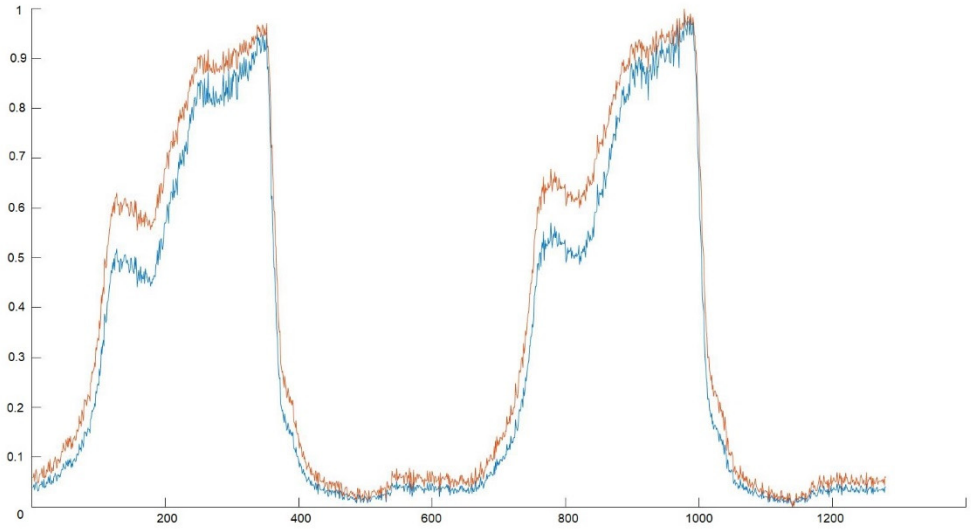


Figure 66. Normalised gate shapes, with original trace in red and corrected trace in blue.

Another important factor shown by this equation is the range of the signal values for which the response of the sensor is fairly linear to the amount of collected photons. Plotting equation 62 (Figure 67) shows where the collection accuracy is about linear and where saturation becomes a problem. While we can in principle correct for saturation, we report later that our SPAD detector does not in reality saturate quite in this way (so revealing some underlying issues with the detector operation). For this reason, we actually attempt to keep signals within the linear range. Signals with a value of 30 (out of 256 levels) suffer from the detection error of about 6% (equation 62). This value increases to about 11% for intensity of 50 and over 27% for the signal value of 100. We find about 10% error to be the highest acceptable error for the readout deviation. To keep optimal shapes of the trace without the correction, the range between 0 and 50 counts for the MCPF = 256 should be kept. Analogously, for the MCPF = 1000 as in the example from Figure 65, the maximum reliable readout value (giving about 10% error) would be 200. Signals with an amplitude of over 600 counts (like presented in Figure 65) are saturated, and therefore of a different, incorrect shape.

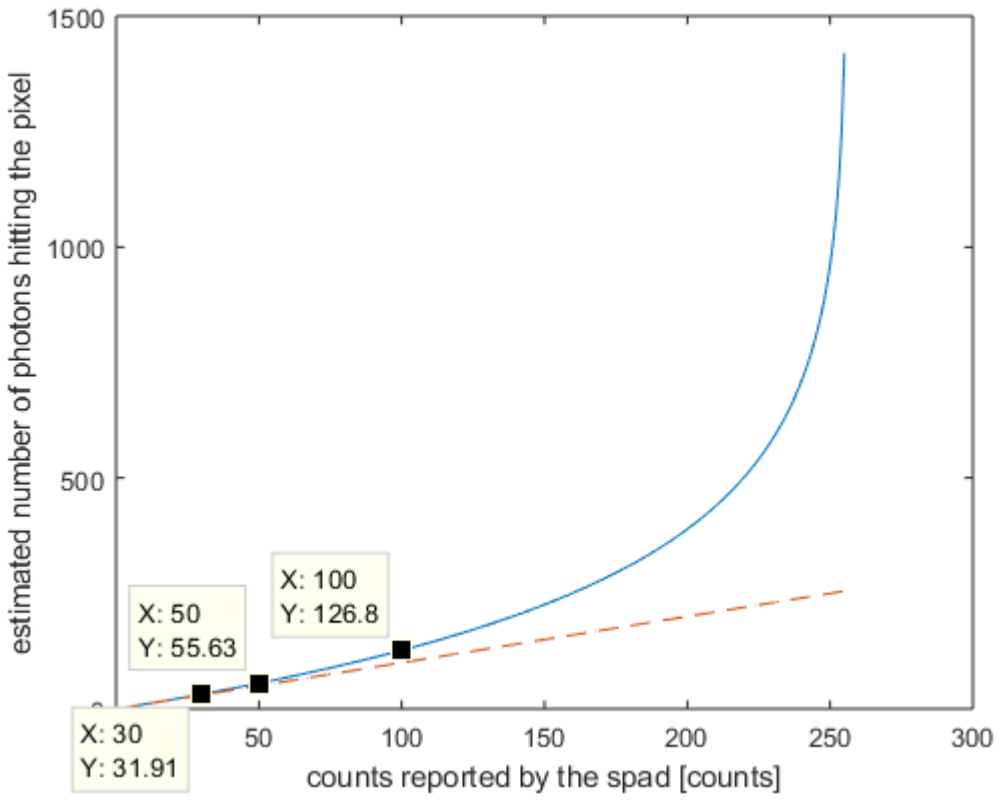


Figure 67. Dainty and Shaw's model in a form presented in Equation 62, plotted to present the photon collection response of the system. Red dotted line presents ideal linear 1:1 response.

3.7. The lifetime retrieval algorithm

We have now presented the basics of operation of our system including the mechanism of using a gate moved with respect to the synchronised laser induced to record a temporal record related to the fluorescence decay. Once the gate is properly designed, retrieved and characterised, we can proceed to use it to retrieve the nanosecond lifetime of the fluorophore. In this section we will present the four main steps required to carry this out, all of which were performed using MATLAB. Some additional procedures used in later sections will be omitted here and then described in those relevant sections.

3.7.1. Sensor noise and filtering procedure

The swissSPAD sensor we used has over 65 thousand individual pixels. Because the manufacturing process is still relatively new and imperfect, some of the pixels are slightly more or less sensitive than the others. Hot pixels are the ones that always

show some value, even when most other pixels are close to zero. We implemented a denoising procedure to compensate for these pixels.

First, we carried out a measurement in conditions where over 90% of the pixels are not reading any signal. Analysing a single frame (an image with specified MCPF) of the readout, we spatially identified pixels which returned a value of 4 or more. These pixels were recognised as hot pixels. The top panel of Figure 68 presents these hot pixels for our array.

Subsequently, the image was filtered with a 2-dimensional median filter with a kernel of 3x3 (82). To avoid blur and other artefacts of filtering, we replaced only the noisy pixels with equivalent pixels from the filtered image. The result is presented in the bottom panel of Figure 68.

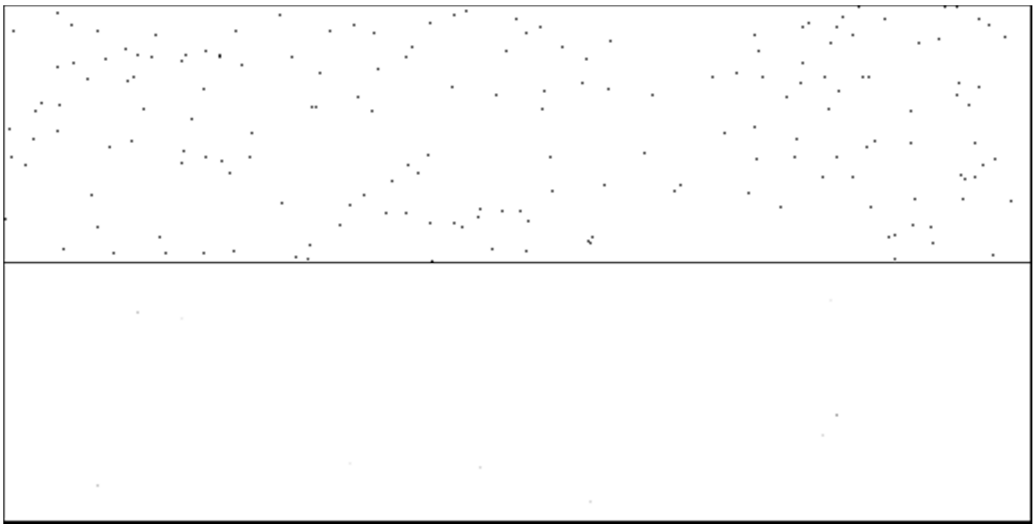


Figure 68. Image taken in dark conditions to localise hot pixels of the system. Both images are presented in inverted colours. Top panel presents image with visible noisy pixels, bottom panel is the same image after denoising procedure.

The number of pixels that will meet the criteria for a hot pixel will increase with increased number of repetitions of 1-bit readouts because of false counts. For the typical repetition number of 200 we found that 309 pixels met the criteria to be identified as hot pixels. This is less than 0.5% of all pixels, and so has negligible impact on the imaging capability of the sensor.

3.7.2. Pixel averaging

The denoised image is passed to the function that prepares it for the fitting procedure that retrieves the fluorescence lifetime. Depending on the goal of the measurement, this can be executed in several ways.

The basic method is passing the time series for each pixel value to process the whole Field Of View, retrieving the fluorescence lifetime for each pixel separately. Because we operate in the single photon regime, these traces will be characterised with the strongest noise, but also with the best spatial resolution. Effectively, each pixel represents about 0.2 mm for our imaging conditions.

The second method allows the user to choose the area over which the signal will be averaged and the fitting procedure will be executed on a time series of average values. This method is designed for the trace observation, testing the fitting procedure and acquiring the gate trace. The gate function of the system used for the fitting procedure should be noise free and as accurate as possible. One of the ways of optimising this trace is to acquire this gate from a large amount of averaged pixels. Gates presented in Figure 65 are acquired from a square of 5x5 pixels, while less noisy gates from Figure 66 are acquired from the average signal value over a square of 25x25 pixels.

The third method is a combination of the previous two and computes lifetimes for the whole image but averaged in blocks of 5x5 pixels. This procedure significantly increases the SNR of the readout allowing us to retrieve accurate lifetimes from significantly smaller signals, and still keeping the spatial resolution below 1 mm for our imaging conditions.

3.7.3. Correction factor

At this stage, the correction factor presented in equation 62 is applied to each pixel value of all the images. The influence of the correction factor on the shape of the trace will be discussed in more detail in section 3.10.2.

3.7.4. Fitting procedure

The fitting procedure is the core of the algorithm, and also the most complex part. Recall that the measured time signal is a convolution of the gate function and the

decay function describing the fluorescence: we must model the fluorescence and then use its convolution with the measured gate function to fit the measured trace. In section 3.6.6 we showed how we acquired the gate function, which is stored in the algorithm (Figure 69a). As presented in section 1.2.3, single exponential decay of fluorescence intensity as a function of time can be represented as an exponential function:

$$f_d(t) = A \cdot e^{-\frac{t}{\tau}}, \tag{63}$$

where A is an amplitude, t is time and τ is the decay time of the fluorophore – fluorescence lifetime. Figure 69b shows an example of equation 63, with τ = 6.8 ns We use an x-axis of the frame number, with frames spaced in time by 19.53 ps: this is 25 ns (time per frame) divided by 1280 (frames per cycle) . Because of the nature of the exponential function, the signal decreases quickly at the beginning and then remains with low, but until some point – significant values. To properly model the behaviour of the fluorophore measured by our system, we summed values of 20 copies of the decay curve each shifted in time by 12.5 ns (representing a train of laser pulses at 80 MHz) to mimic overlapping of consecutive fluorescent decay signals. This is presented in Figure 69c and as can be seen, the trace stabilises its amplitude from the 4th peak onwards. Subsequently, we perform a convolution of this fluorescence signal and the gate trace to imitate time-gated measurement of the system. The resulting function is presented in Figure 69d. The early and late peaks are not valid because of the boundary conditions for the convolution. We chose one full cycle of 1280 consecutive values from the middle section of this plot, where the shape is stabilised. This is presented in Figure 69e.

We fit this model curve to the measured data using least-squares curve-fitting. The model has 3 variables to optimise – amplitude, y offset and the most importantly τ – the lifetime. Equation 63 updated with the y offset parameter (oy) is as follows:

$$f_d(t) = A \cdot \left(e^{-\frac{t}{\tau}} + oy\right). \tag{64}$$

If we replace all the variables to be optimised with the vector x, we can write the minimisation function down as:

$$min_x \left|\left|F(x, model) - signal\right|\right|_2^2 = min_x \sum_i (F(x, model_i) - signal_i)^2. \qquad (65)$$

Figure 69f shows an example trace acquired with the system and the function from the panel e, fitted to it in red. Optimised parameter τ of the fitted function is the designated lifetime of the measured fluorophore in units of numbers of frames (the value of the x axis).

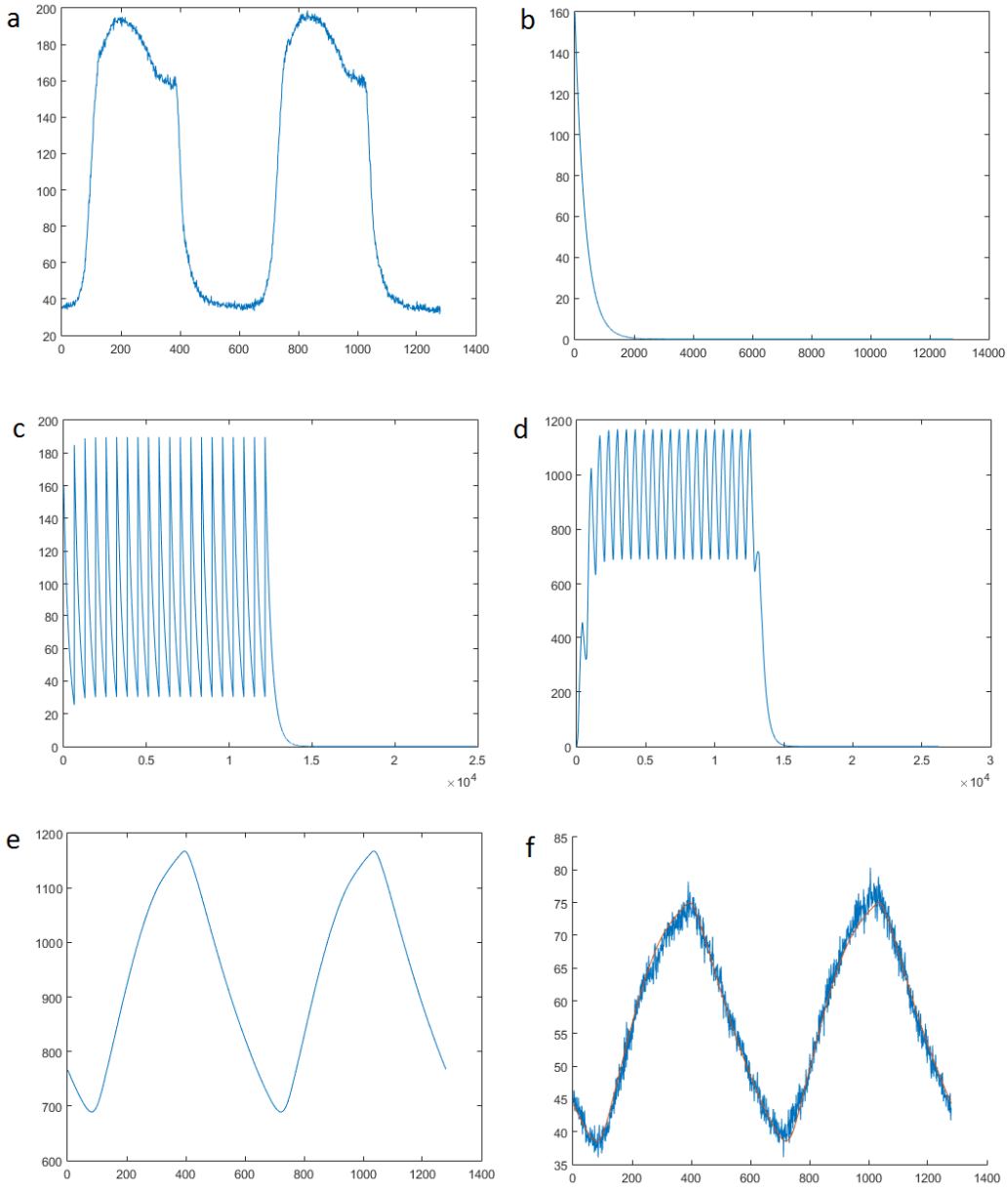


Figure 69. a) Gate function measured by the device, b) modelled 6.8 ns decay curve with amplitude similar to the gate, c) model repeated 20 times, summing all the signals, d) convolution of the gate and repeated model, e) proper cut of one modelled measure period, f) model fitted to the real trace.

To present the result in nanosecond we need to convert it as follows:

$$\tau[ns] = \frac{\tau[samples]}{1280[samples]} 25[ns] \,. \qquad (66)$$

The optimisation function requires initial values of parameters to be provided, like we did in the step presented on panel b of Figure 69. Choosing values close to those expected increases fitting efficiency but does not change the outcome (for unlimited fitting iterations).

3.8. Lifetime detection limit

With the lifetime calculation procedure prepared, we were able to perform a simulation of the trace behaviour for different modelled exponential lifetimes. Doubled laser frequency causing collection of two pulses during one readout cycle effectively decreases maximum lifetime possible to determine. The second pulse (shape on the graph) results in the gate function collecting an additional decay function which affects the shape of the resulting convolution trace. Averaging over additional (disturbing) signal results in more flat trace, hence harder to distinguish. The time between two pulses is 12.5 ns and the gate length is about 8 ns. This means that the decay time of 4.5 ns will be scanned by the gate function completely, before entering the second pulse. Collection of the second pulse while still averaging over the previous one (for lifetimes longer than 4.5 ns) does not yet indicate uselessness of the trace, but the determination capability will start dropping. We prepared two sets of simulated outcome traces (convoluted with a collected gate): for very short lifetimes between 0.1 – 1 ns with a step of 0.1 ns to observe low detection limit capabilities, and for lifetimes between 1 – 8 ns to observe the lifetime where we expect to observe changes. These traces are presented in Figure 70a and Figure 70c. Figure 70b and Figure 70c presents the same trace with simulated noise.

We can observe that the retrieval resolution is the highest for the shortest lifetimes, theoretically going as low as 0.1 ns. For lifetimes longer than 5 ns, traces get flatter and closer to longer lifetime traces. We estimate that 7 ns is the high lifetime limit of accurate detection.

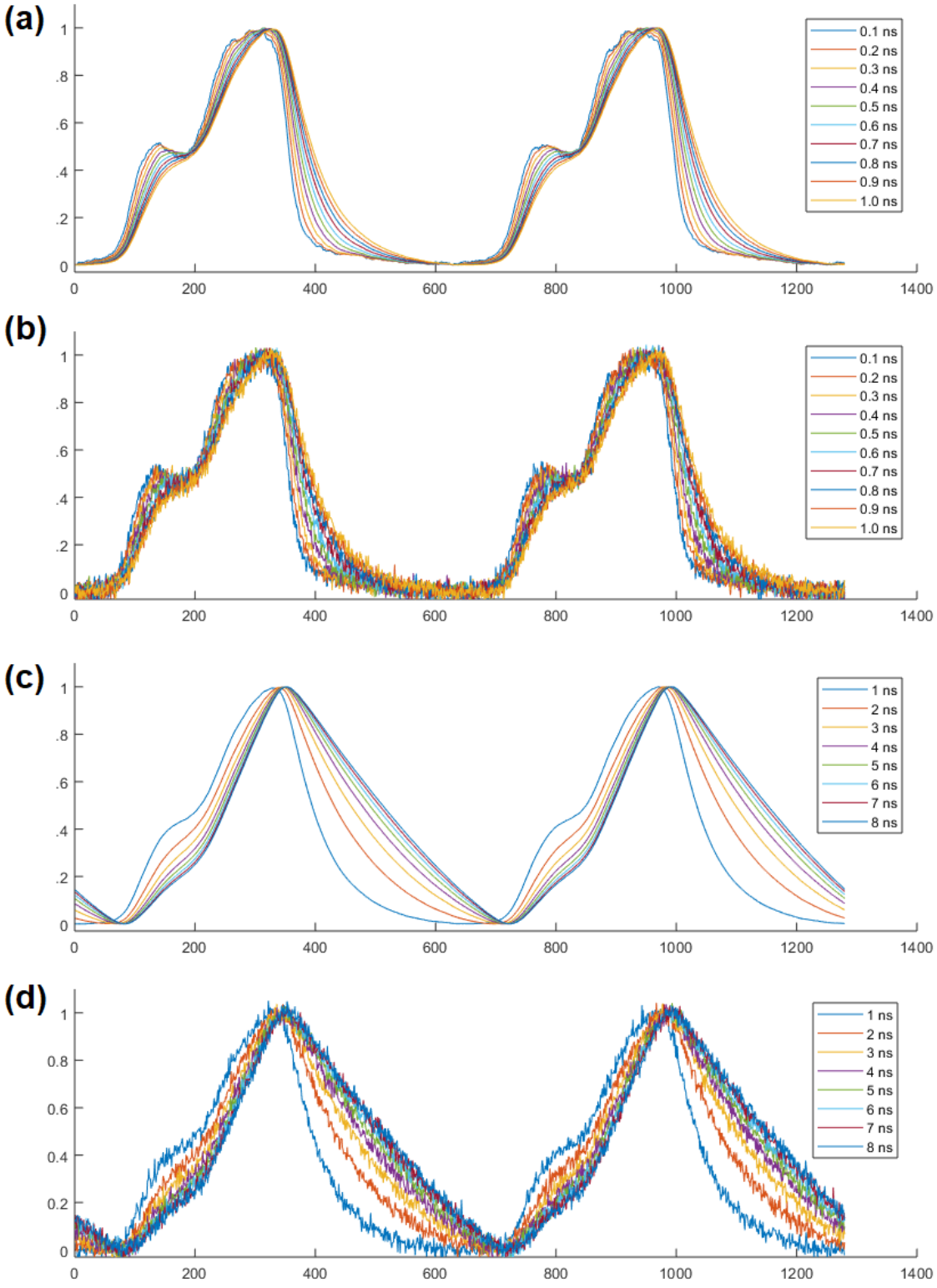


Figure 70. Simulated lifetimes convoluted with the IRF to estimate lifetime retrieval limits. a) traces between 0-1 ns, b) same traces with simulated noise, c) lifetimes between 1-8 ns, d) same traces with simulated noise

3.9. The first fluorescence lifetime retrieval

At a very early stage we performed a first fluorescence lifetime measurement of a Perspex plate doped with a fluorescent dye (chosen as a convenient solid target). The measurement was in fact that already presented in Figure 69e. We quantified the fluorescence lifetime independently using a fast photodiode and a LECROY WM8620A (Teledyne LeCroy, New York) oscilloscope, retrieving a reference lifetime of 6.79 ns.

Our setup and lifetime retrieving algorithm was not finalised back then. The setup was constructed exactly as shown in Figure 60. A collimated 532 nm beam excited the fluorescent perspex. The perspex had an angle of 45° in respect to the excitation beam and the sensor, to reflect most of the excitation light away from the sensor. The fluorescence emission of the plate was gathered by the standard 5 times objective from a distance of about 2 cm from the sample. We used a high pass filter with cut-off wavelength of 550 nm to eliminate the remaining excitation beam from the readout. The signal captured by our system is presented in Figure 71.

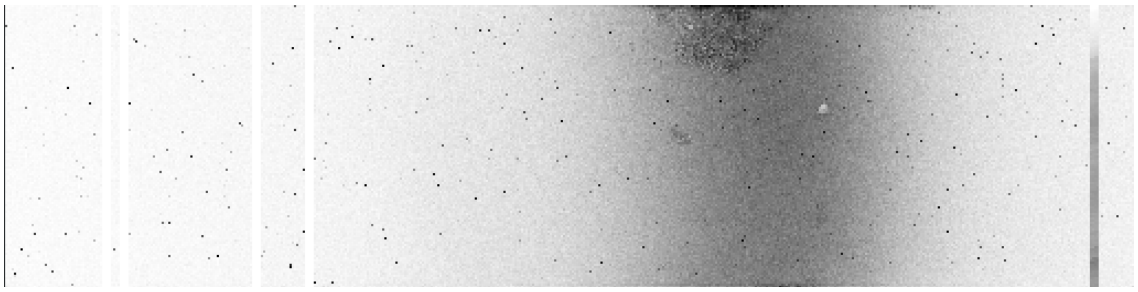


Figure 71. Fluorescence signal captured by the SPAD system. Vertical line visible on the right is caused by the pin connection problem, these pixels are not used in the measurement. Image presented in inverted colours.

The analysis did not include the denoising function at this early time – an area without any hot pixels was selected for the measurement. The correction factor was also not implemented yet. We used the somewhat saturated gate function presented on Figure 69a. Because the fluorescence signal of the perspex was strong, the readout was also somewhat saturated (y offset of almost 40 – Figure 69f). The long lifetime of the fluorophore helped to smooth the resulting trace (throughout the convolution). Figure 69f presents a good fit with the signal averaged over the 5x5 pixel square.

Lifetimes retrieved from individual pixels with deviations from the mean value are presented in Table 4.

Table 4. Lifetimes retrieved from a 5 by 5 array of individual pixels on left. On the right, % deviation from the mean value of corresponding pixels.

[ns]				
6.29	6.88	6.72	6.42	7.43
6.57	6.28	7.49	7.82	5.79
6.21	6.43	6.56	7.39	6.45
6.79	6.57	6.81	6.64	7.64
6.51	7.02	7.04	6.76	6.48

[%]				
−7.39	1.37	−1.01	−5.52	9.35
−3.17	−7.55	10.27	15.20	−14.67
−8.51	−5.28	−3.36	8.87	−4.95
−0.03	−3.28	0.26	−2.19	12.50
−4.17	3.35	3.67	−0.50	−4.51

The lifetime measured from the averaged pixels and the fit presented in Figure 69Figure 58f was 6.73 ns. The mean value of lifetimes presented in Table 4 is 6.76 ns. The mean error for single pixel measurement is smaller than 6% with maximum error of 15%.

As an initial proof of principle, this was a pleasing result, with retrieved lifetimes consistent with our reference measurement and with relatively low noise despite somewhat saturated traces, and yet-to-be optimised analysis.

3.10. Gates behaviour for different signal levels

Ideally the shape of the gate of the system should be an even top hat shape. Perfectly steep edges would provide a good replica of the decay curve of the fluorophore during the process of convolution, and the even top of the function would ensure simple scaling of the shape for different amplitudes of the measured signal, therefore simple fitting. In our analysis method, we also are assuming that the gate *shape* does not change as the gate delay is changed; in our setup with two laser pulses within the range of gate delay this means we should see two identical gates on the IRF function. Figure 72 shows two ideal functions: the gate with the exponential decay on top panel, and the product of the convolution of these functions. For decay functions with 'longer lifetime', the shape will start to look more like that shown in Figure 69f.

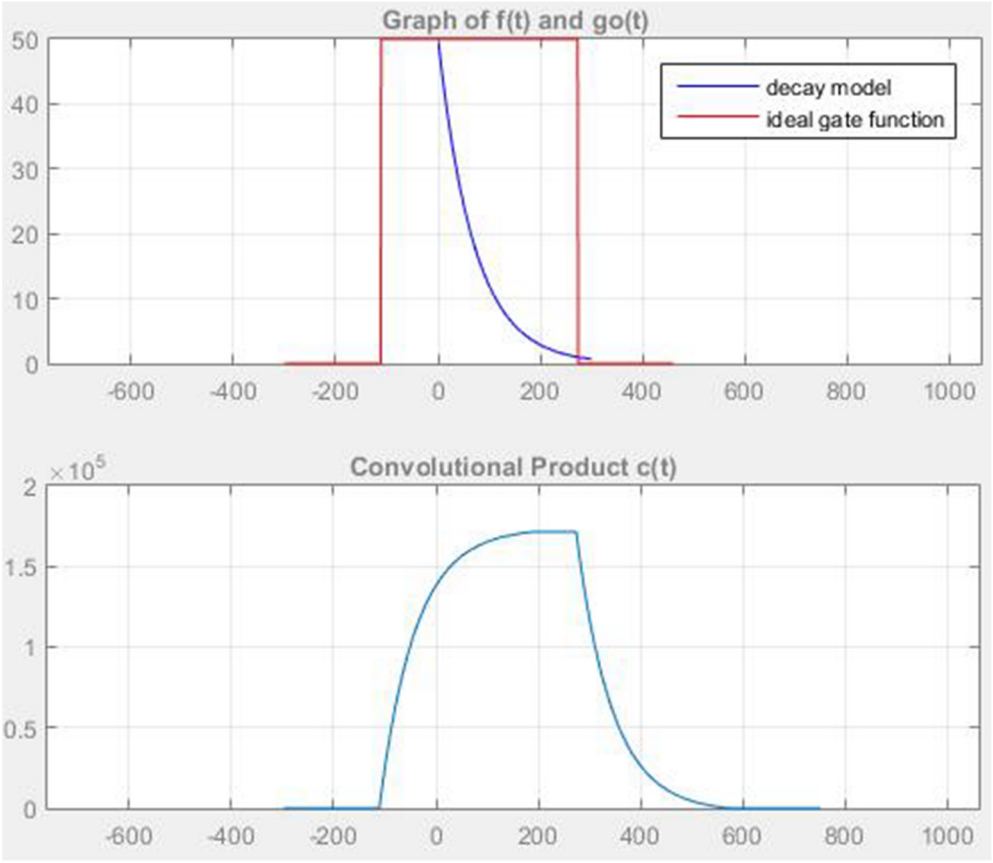


Figure 72. Top: Exponential decay curve model in blue and ideal top hat gate shape in red. Bottom: Product of convolution of two functions presented on the top.

Unfortunately, the system does not provide an ideal gate shape, and worse, there are signs that the gate is not unchanged with gate delay. We carried out a simple experiment where for constant readout conditions we changed the laser signal intensity hitting the sensor through a cube polarizer and half-wave plate. We took 10 measurements with increasing signal to observe the gate shape behaviour. Figure 73 presents these gate traces on one graph. A couple of conclusions can be drawn:

- Both edges are steep but not ideal. The leading edge (which for the process of convolution is reversed into a trailing edge) is steeper which is desired. The edges are similar for all the intensities.
- The shape of the top of the function changes with the intensity. Initially lower at the beginning and higher at the end, and then the reverse for higher intensities.
- The lowest achievable value rises with the overall function intensity.

- The two gate shape measurements for the same intensity (due to the first and second laser pulses within the delay range) are different, both in shape and amplitude.

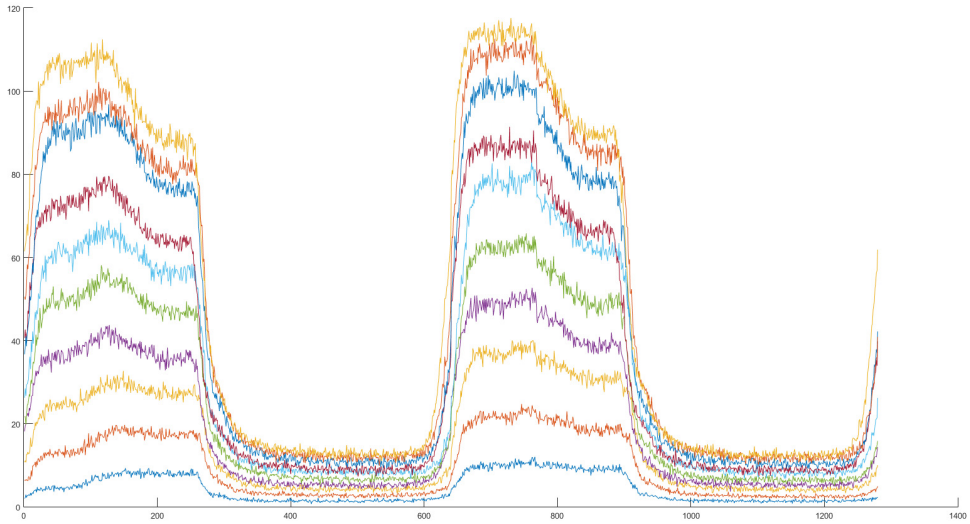


Figure 73. Gate shape traces with different intensities collected for identical readout conditions.

We tested and implemented some additional procedures and restrictions to try to eliminate or compensate for these differences to ensure the most reliable fitting.

3.10.1. The 'Hold off' procedure and the system response to trace shifting

The irregular shape of two gate shapes for a single measurement suggested one of two potential problems: a different system response for different gate shifts, or problems with electrical stability of pixels during the cycle (improper shapes for different readout sequences if insufficient time spent in particular states - Figure 65). First, we performed a test where we introduced a 'hold off' to the readout. The hold off is a variable setting of a number of initial unused cycles for each group of repetitions. It addresses the same electrical phenomenon as the dead time characteristics, allowing for an additional time between measurements. For example, if we set 100 repetitions (100 attempts to see a photon into the 1-bit sensor), hold off = 1 means that the first cycle will not set the gating status, leaving 99 active cycles and

giving a little extra time to the sensor to stabilise after recording the previous readout. We performed a measurement of 4 different signal intensities for hold off equal to 1,2, and 5. Each panel of Figure 74 presents different intensity with traces for all the hold off values on it for comparison. The first and the most important result of this experiment is that the shape of two gate measurements for a single readout is significantly closer to each other comparing to gates from Figure 73. This gives us confidence that the gate shape is now much more stable as a function of gate delay. The second finding is that the shape for different hold off values (different colours on each panel) is almost identical for all gates, decreasing slightly in overall amplitude because of the smaller amount of repetitions. The conclusion from this finding is that the first cycle for each group of repetitions must be significantly different for different gate delays, and this somehow is related to the error seen in Figure 73. From now on, all the measurements were performed with the hold off set to 1.

The hold off procedure improved trace stability substantially, however, did not eliminate the problem completely. The shape of two gate measurements is very similar between each other, yet the amplitude of the first local maximum (the 'shoulder') still seems to be slightly higher for the second measurements. Also, the difference between the first local maximum and the top of each gate is changing with the overall amplitude of the signal (between panels).

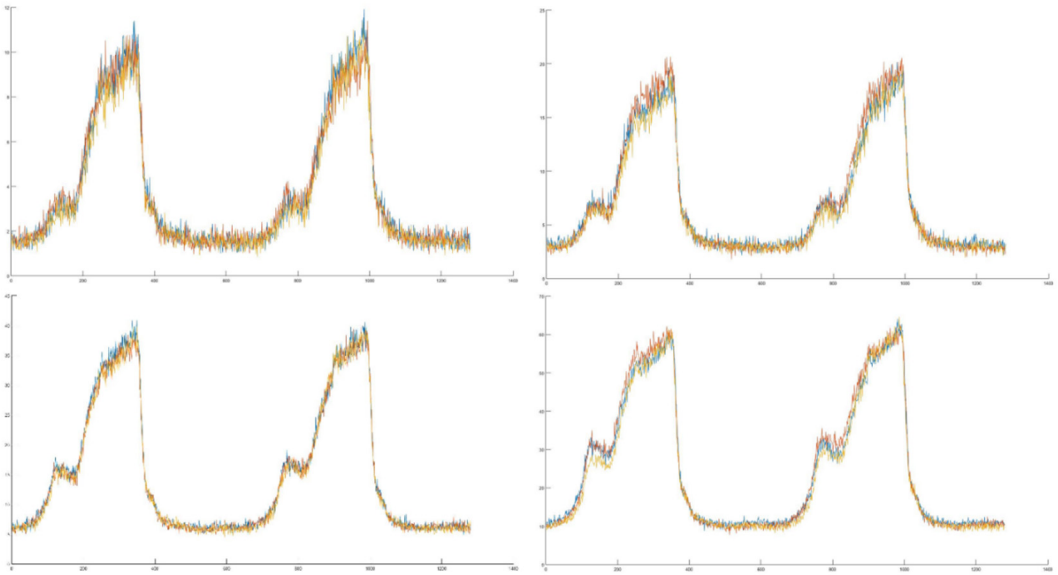


Figure 74. Four examples of gate traces with different hold off values. Blue colour is hold off = 1, red - 2 and yellow is hold off = 5.

To further investigate the change in gate shape with gate delay, we took number of identical measurements of the gate signal but starting the gating sequence at different shifts from the laser pulse timing signal. We took a total number of 10 measurements, shifting each starting position by 128 samples (this would 'swap' the position of the gate measurements after 5 shifts - a total shift of 640 so half of the whole cycle - and return to original position after 10 shifts). Figure 75 shows consecutively shifted gating traces, which were numerically aligned back to overlay each other for shape comparison in panel c. As seen on this panel, all 10 gates are basically identical with a small variance of noise. We conclude then that the number of shifts (delay) from the beginning of the readout does not seem to have any strong influence on the gate shape with our optimised settings.

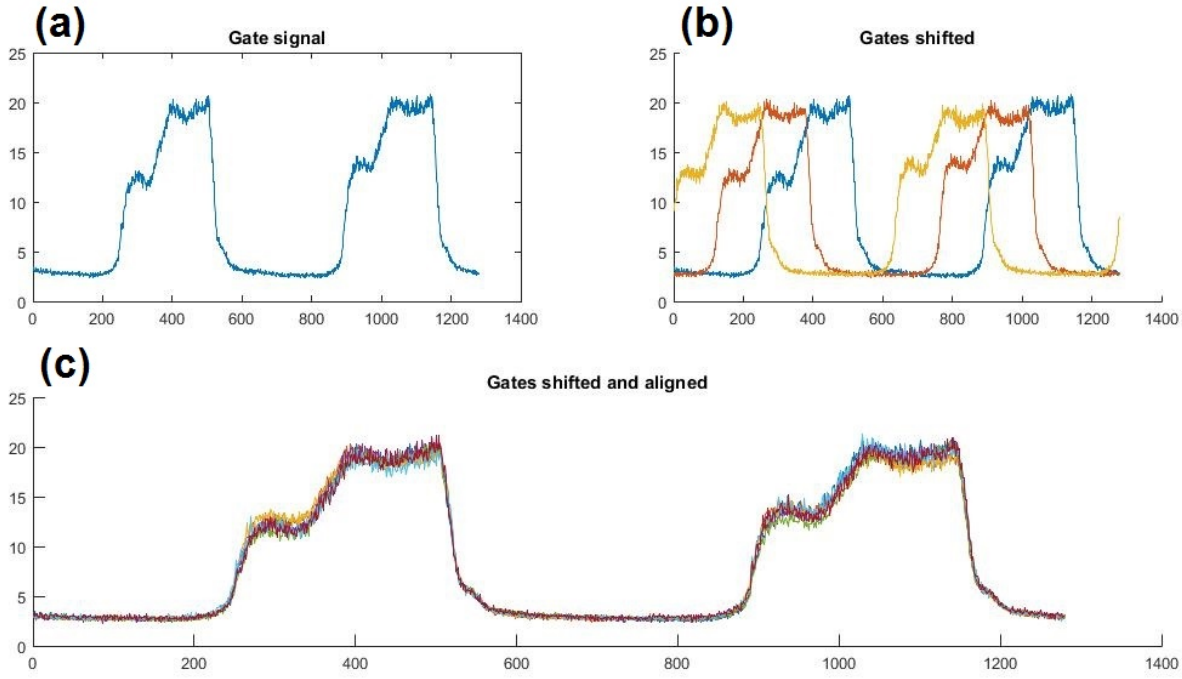


Figure 75. (a) A single, regular gate measurement; (b) Three gate traces measured in identical conditions, each starting the measurement 128 samples further than the previous measurement; (c) 10 consecutively shifted traces, numerically aligned to overlap each other.

Finally, we compared two gates from the same measurement in detail. Now using the hold off procedure, their shapes became very similar, yet not identical. To compare them, we cut the trace in half and plotted both gate shapes from a single measurement on top of each other for comparison. We found that there is a very

specific point where the first gate measurement suddenly drops and in the same (respective) point, the second gate measurement goes up in the signal intensity. This phenomenon was observed in every measurement (Figure 76 a,b,c,d) and is probably the reason why actual fluorescence measurements have different intensities for two gates (Figure 76e). This appears to be a bug in the SPAD design – at these delay points the sensitivity of the pixel abruptly changes. We were unable to find a cure in software, and so it may be a hardware problem that we cannot address.

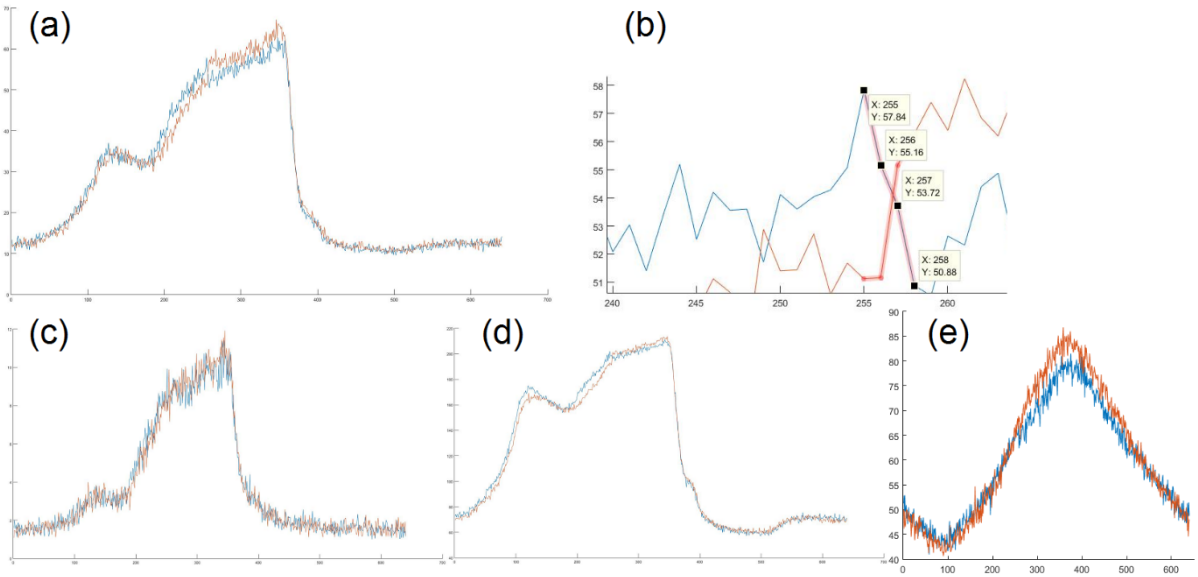


Figure 76. Two gate shapes from the same measurement on top of each other. Panels a,c,d show gates for different laser powers. Panel b is a close-up at the exact point of the phenomenon occurrence. Panel e presents measurement of the fluorescence signal influenced by the phenomenon.

3.10.2.Influence of the correction factor on the gate shape

We described the correction factor in section 3.6.7. The influence of the correction on the signal is not linear; therefore the shape of the trace is influenced. As we presented in Figure 66, the correction increased the amplitude difference between two local maxima of a single gate shape. It was expected that correcting for saturation would entirely explain differences in the gate shape for different signal intensities. To examine this, we carried out another experiment, recording gate traces for different laser intensities, this time with the hold off set to one. Figure 77 presents these gates (a), and leads us through consecutive analysis steps: the traces are smoothed for clarity of view (b), normalised to compare to their different shapes (d), corrected for

saturation (c) and again normalised (e). Comparing panel d and e, we can observe that differences between local maxima of the first peak of each gate are decreased by the saturation correction, yet still not eliminated. It appears that there is a real change in the sensitivity on the sensor in different lighting conditions, with the shoulder recording an excess number of additional counts as intensity is increased. It seems that the shoulder part of the trace reports more counts than we expect with increasing signal. Most likely it is an electrical issue of the chip (each pixel circuit), potentially an increased vulnerability to afterpulsing due to rising number of photons hitting the sensor.

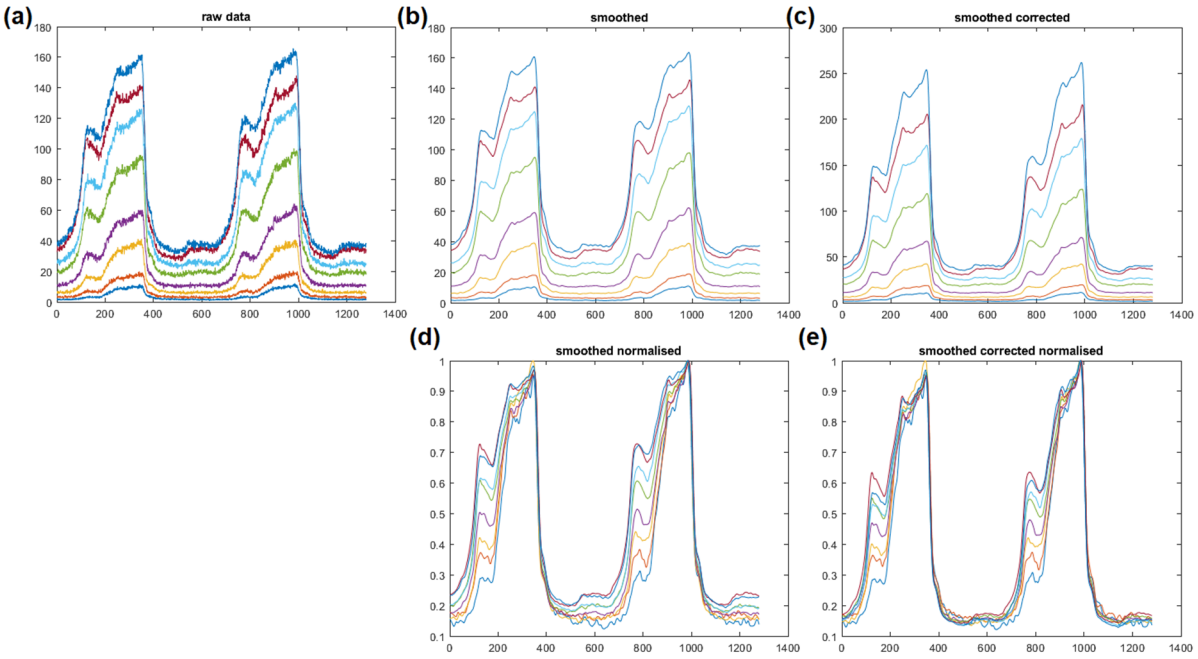


Figure 77. Differences in the gate function shapes for different intensities a) raw data, b) denoised, c) denoised and with correction applied, d) denoised and normalised, e) denoised, normalised after correction, applied.

3.10.3.Gating behaviour conclusions

The shape of the gate of the SwissSPAD is far from the ideal shape presented in Figure 72. The real gate shape suffers from a number of issues, only some of which we managed to eliminate. The significant difference in the shape of two visible gates from the same measurement was resolved by using the hold off procedure (Figure 73 and Figure 74). Applying the saturation correction to all measured traces decreased, but didn't eliminate, the problem of the changing amplitude ratio between two local

maxima of the gate shape (Figure 77d,e). The difference between the first and the second visible gate shape of a single measurement also still exists, showing that the gate shape does change somewhat as the gate delay is changed.

The biggest complication arising from presented issues is the difficulty of a proper function fit for different signal intensities. Because the shape of the gate changes with the amplitude in a way that cannot be linearly scaled, accurate fit of a fluorescence signal significantly different in amplitude than the gate used to prepare the model becomes inaccurate. To overcome this problem, we measured the gate shapes 5 different times with amplitudes increasing by 10 counts within the range of 0 – 50 counts. The fitting algorithm analyses the amplitude of the lifetime signal to be fitted and chooses a proper gate file to prepare the model to perform the fitting procedure. These gate shapes are presented in Figure 78.

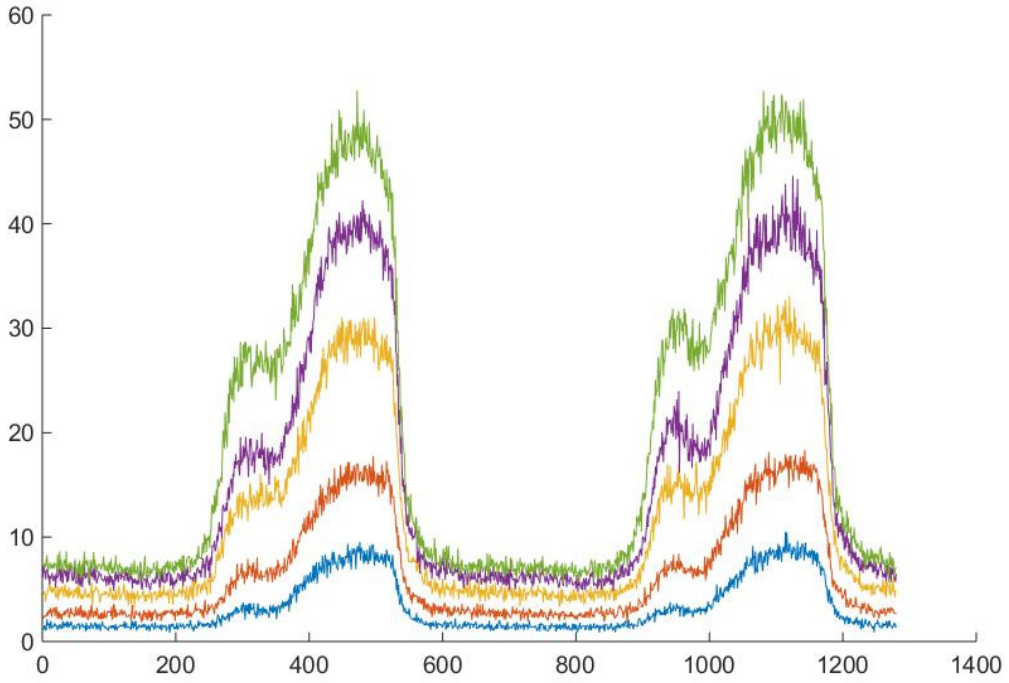


Figure 78. The set of the gate traces with different amplitudes to prepare models for the lifetime fitting procedure.

The second issue that we were not able to eliminate completely was the difference between gate shapes at different delays within one measurement. We proved that the issue occurs after a number of shifts performed rather than corresponding to the

relative position from the laser pulse (Figure 75). To provide a good fit to the signal collected as presented so far, it is important to always use the corresponding part of the model. Using the second visible gate shape to fit the first peak of fluorescence trace would not provide a proper measurement. The timing of the acquisition of the gate should always match the timing of the acquisition of the fluorescence trace, and the first gate used to model the first fluorescence peak.

Addressing all these phenomena, the real shape of the gate for the proper range of intensity and the resulting model of the signal of the fluorescence lifetime of the fluorophore is presented in Figure 79.

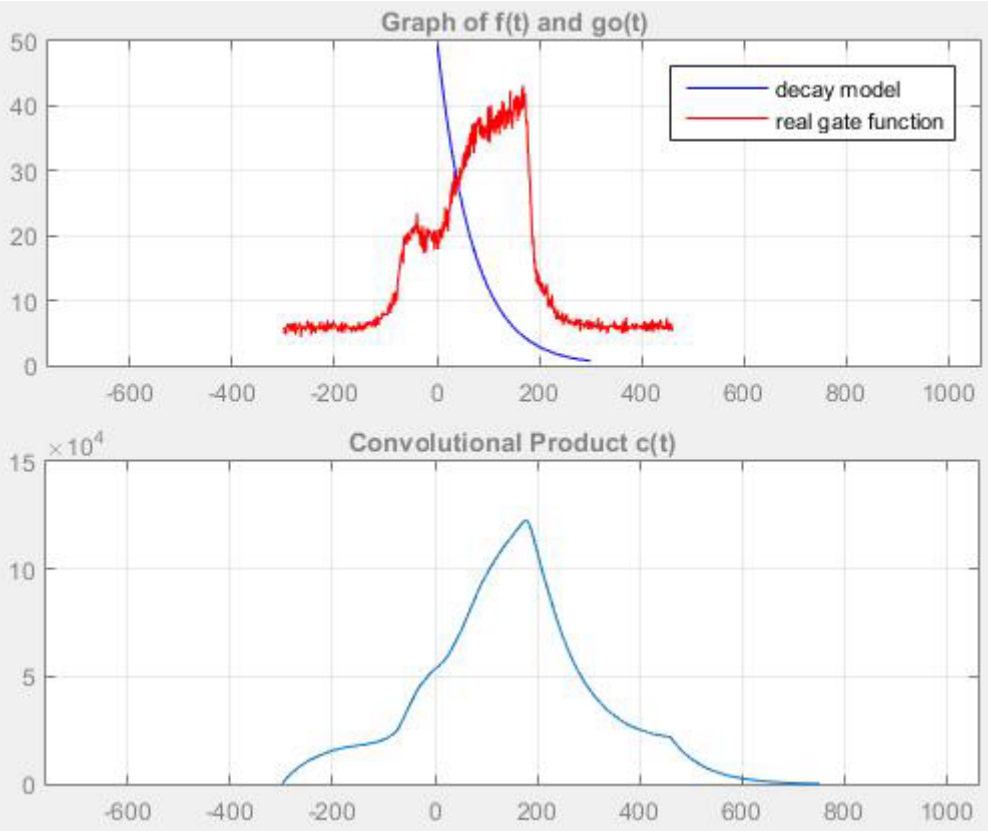


Figure 79. The real shape of the gate acquired from the system and resulting model calculated by convoluting the gate with an exemplary lifetime (presented for a half of the typical whole trace showing two peaks).

3.11. The imaging of fluid fluorophores

After proving the principle of operation on the doped perspex on just a few pixels on the array, and understanding capabilities and limitations of the gating procedure of the SPAD, we moved forward towards imaging. The focusing lens in front of the SPAD

array was replaced with a proper, specifically assembled camera objective that allowed us to take sharp images from the distance of about 18cm. We expanded the collimated laser beam to a circle with a diameter of 9cm, covering the whole field of view of the camera. The power of the expanded beam was below 5 mW/cm^2. We used a high pass filter with a cut-off wavelength of 560nm as an emission filter (Figure 80b).

We used two fluorophores in separate solutions, and imaged them side by side. The fluorophores that we used in the experiments were Fluorescein and Rhodamine B. 1 mL of each was placed in the rectangular plastic, 8 wells cuvette (Figure 80a). Rhodamine B had a concentration of 0.3 mM and the Fluorescein solution was 1 mM. Excitation-emission spectra for both dyes are presented in Figure 80 panel c and d respectively. While Rhodamine B has a high absorption at 532 nm, the Fluorescein is at the very edge of the spectrum with hardly any absorption (about 3%). In addition to very low excitation, a significant part of the Fluorescein emission is blocked by the emission filter. We expect the fluorescein signal to be very low. The Fluorescein has a fluorescence lifetime of 4 ns (162) and the Rhodamine B has a fluorescence lifetime of 1.68 ns (162).

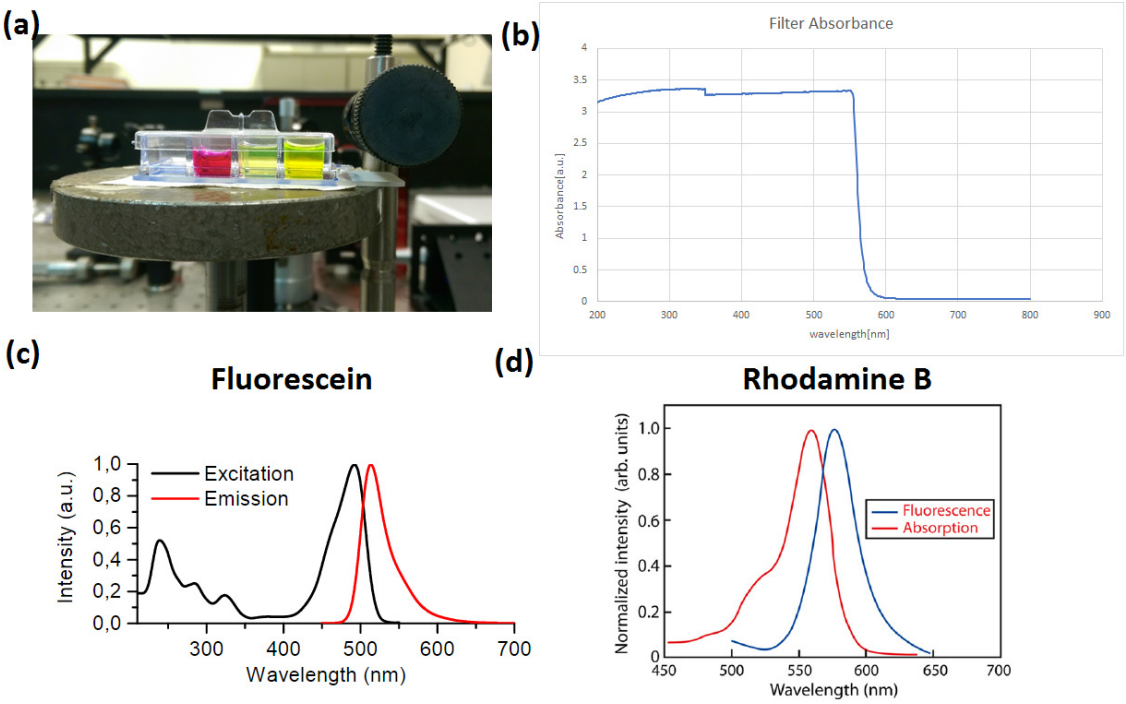


Figure 80. a) Solution of Rhodamine B and two solutions of Fluorescein in 8 well plastic cuvette. The image was taken with a regular camera from about the perspective of the sensor b) Absorption spectrum for the filter, with cut-off wavelength at 560nm. c) Excitation emission spectrum for Fluorescein. d) Excitation emission spectrum for Rhodamine B

All the measurements were taken using MCPF = 256 and repetition of 200. Because of the low signal levels, we used averaging of 5 by 5 pixels for lifetime calculation, which increased the SNR. The fitting procedure was performed only for one trace peak (slightly extended half of the full length of the measurement) to test the accuracy of this solution, allowing for faster measurement (acquiring about half of the original data). Images were masked to present lifetimes only for relevant parts of the image where fluorophores were visible. The image presented in panel a of Figure 80 was taken with a regular camera from the location of the sensor and presents a similar image to the one that the sensor recorded. The actual frame of the same fluorophores collected by the SPAD is presented in Figure 81a. The emission from the middle fluorescein sample (with concentration of 0.1 mM) was too weak for the current measurement conditions (very weak excitation) and was excluded during the masking procedure. Colour coded results showing lifetimes retrieved as a function of position for the measurement are presented in Figure 81b.

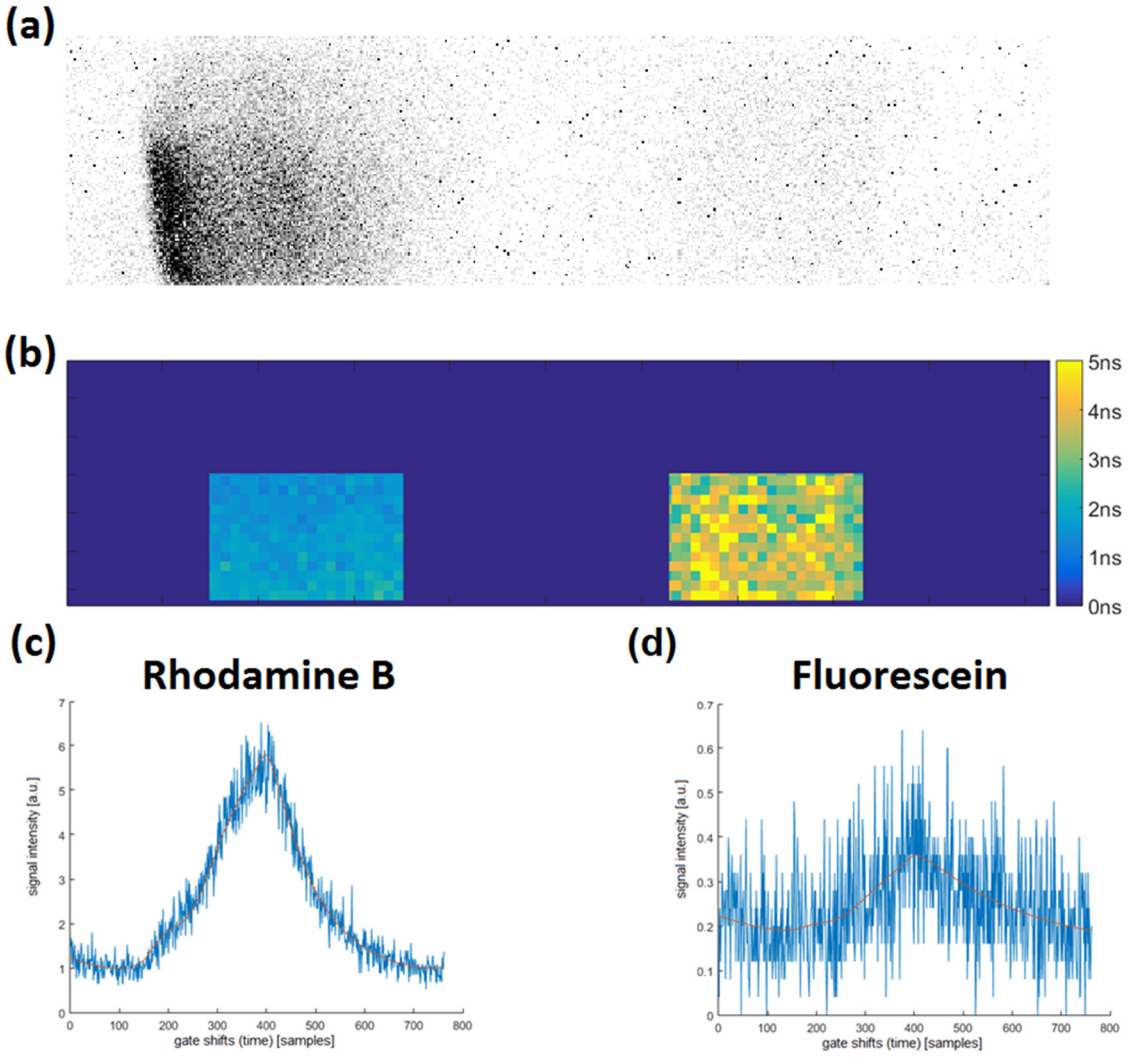


Figure 81. Lifetime measurement of fluid Rhodamine B and Fluorescein solutions.a) one frame as seen by the sensor, presented as 8-bit grayscale image with between readout values 0-10 counts (inverted colours) b) Spatial image of two wells with fluorophores, colour coded based on calculated lifetimes. c) Example trace with fit function for Rhodamine B. d) Example trace with fit function for Fluorescein

The Rhodamine B part of the image shows a stable result with a mean lifetime across the image of 1.69 ns with standard deviation of each measurement of 0.26 ns and a Standard Error of Mean of 0.016 ns. SEM is calculated as a standard deviation divided by the square root of number of samples, and assumes that our measurements have a random error around the true value. The amplitude of the signal is about 5 counts (Figure 81c) which is low. Because of the pixel averaging procedure, the noise is averaged, therefore decreased, which helps the accuracy of the readout. Using the averaged pixels (visible in Figure 81b) makes this signal level completely sufficient to accurately designate the lifetime for the area. An example fit for the trace for one averaged 5x5 area is presented in Figure 81c.

On the right-hand side of Figure 81b, the measurement of Fluorescein can be seen. The mean lifetime across the area of 3.8 ns is less accurate than for Rhodamine B. The standard deviation of 1.0 ns and standard error of means of 0.0625 ns are significantly larger as well. The explanation of these differences is the signal amplitude. Figure 81d shows that the amplitude of the signal is about 0.5 counts, with the noise level almost equal to the signal level. The signal amplitude below 1 reflects the pixel averaging, meaning that only some of the averaged pixels received a single photon during this measurement. Even for such an extremely low signal, we are able to retrieve lifetime with a satisfying accuracy. Differentiation between two fluorophores is clear and indisputable. All the results are gathered in Table 5. Both measurements agree with published values for those fluorophores.

Table 5. Lifetime results with statistical information of fluid Rhodamine B and Fluorescein solutions.

Fluorophore	Rhodamine B	Fluorescein
Literature lifetime [ns]	1.68	4
Mean Measured lifetime [ns]	1.69	3.8
Signal amplitude [counts]	5	0.5
Standard Deviation [ns]	0.26	1
Standard Error of Means [ns]	0.016	0.0625

Figure 82 presents the distribution of lifetime readouts for the Rhodamine B and Fluorescein samples. The image was analysed using averaging of 5x5 pixels, which decreased the noise and made very low signal useful. The effective pixel size remained below 1 mm. More or less averaging can be used, setting the trade-off between readout accuracy and spatial resolution. Both histograms have a distribution close to normal with Rhodamine distribution narrower (more accurate) because of the higher signal level. Based on the histograms, we conclude that the system is capable of fluorescence lifetime imaging and can distinguish between lifetimes of presented fluorophores, even for low and very low signal levels.

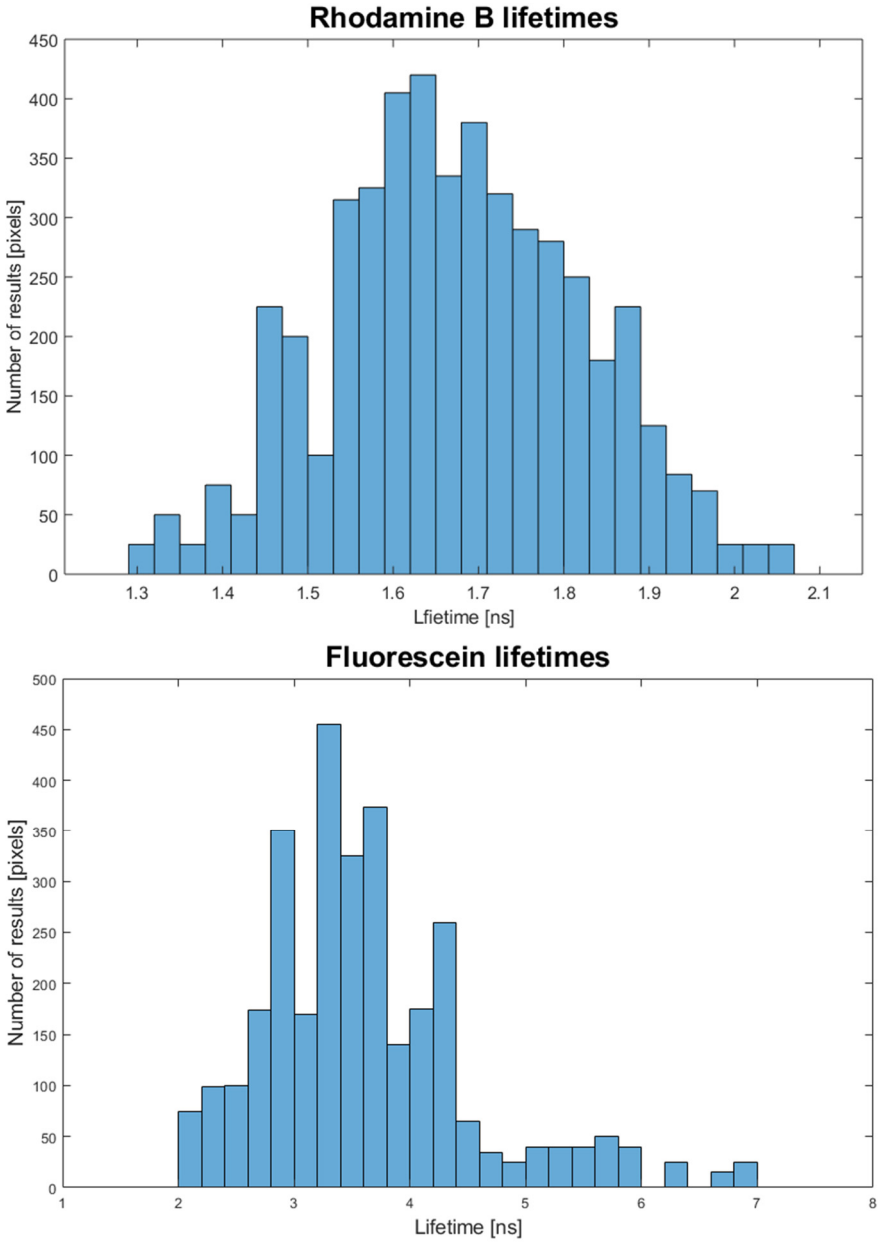


Figure 82. Histogram presenting distribution of measured lifetimes from the area of the Rhodamine B (top panel) and Fluorescein (bottom panel) samples presented in Figure 81b

3.12. Imaging of free and protein-bound NADH

3.12.1. Reduced nicotinamide adenine dinucleotide

Reduced nicotinamide adenine dinucleotide (NADH) is a coenzyme naturally existing in human cells. It is involved in, and used to study redox status of tissues. Disruption in the redox equilibrium may cause many pathological conditions including cancer. When redox homeostasis breaks, metabolism increases which leads to increase of anoxic reactions in the system. One of the reactions is binding of the NADH to proteins, changing its form, making the NADH a cancer marker. Significantly increased amount

of the bound NADH compared to the free form, suggests tumorous changes (163, 164). NADH is auto fluorescent, highly excitable at 355 nm, and emits light in a range of about 400 nm to 600 nm (Figure 83). The issue is that both forms, free and bound NADH have basically identical spectral responses, thus cannot be distinguished from each other on this basis (165). The difference between forms is in the fluorescence lifetime. Free NADH has a lifetime of about 0.4 ns (166). The lifetime of the bound NADH, depending on the binding element might vary, but is always longer. In our case, bound to malate dehydrogenase, the lifetime is equal to 1.0 ns.

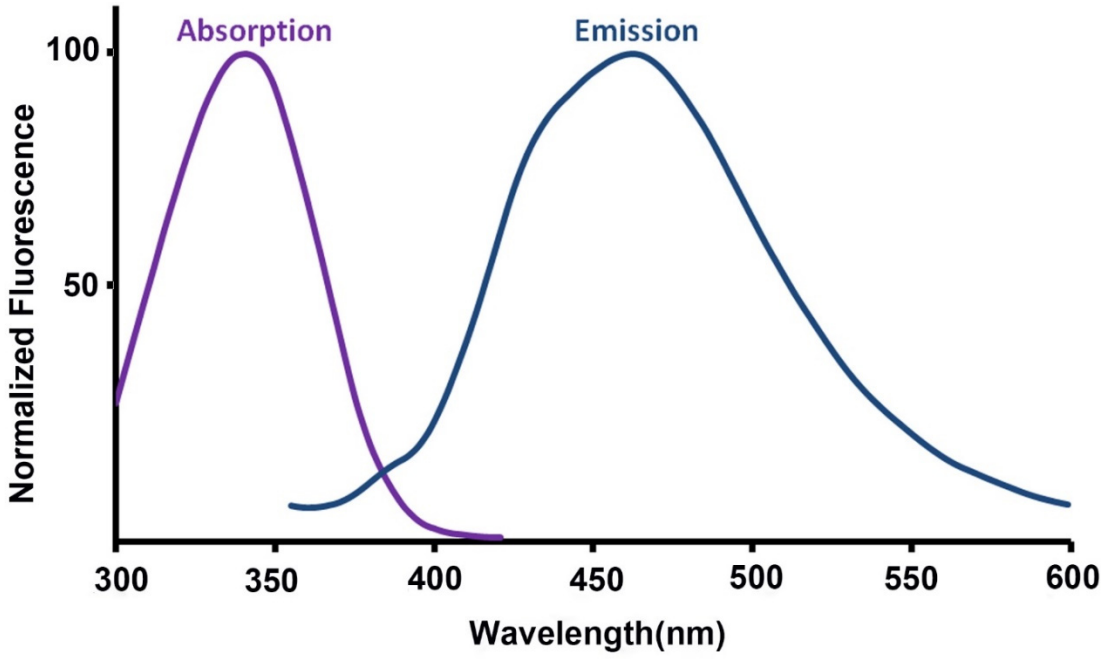


Figure 83 Excitation and emission spectra of NADH (adapted from (165))

We prepared our samples according to Lakowicz's instructions from (166). Free NADH was simply NADH (Sigma-Aldrich, St. Louis, MO, USA, Sigma, SKU: 10107735001) diluted in water in a ratio to obtain the desired concentration. Bound NADH, was bound to L-MDH (Sigma-Aldrich, St. Louis, MO, USA, Sigma, SKU: 10127914001) in a molar ratio 2:1 to NADH in a 100 mM MOPS (Sigma-Aldrich, St. Louis, MO, USA, Sigma, SKU: M1254) buffer (pH 7.0). In a test tube L-MDH was mixed with the MOPS powder. Then, a proper solution of NADH in water was added to complete the volume. The product was then mixed in a shaker to dissolve all chemicals.

3.12.2.The noise measurement, the SNR and the dynamic range

We measured the noise level choosing a part of the trace between two gates, where its (mean) value is the most stable (Figure 84). The noise was measured from the section of 100 samples, from 6 different areas/pixels of the image (600 samples altogether). Each signal was measured at the same areas/pixels, from the same trace section. The noise was calculated as doubled standard deviation of described samples.

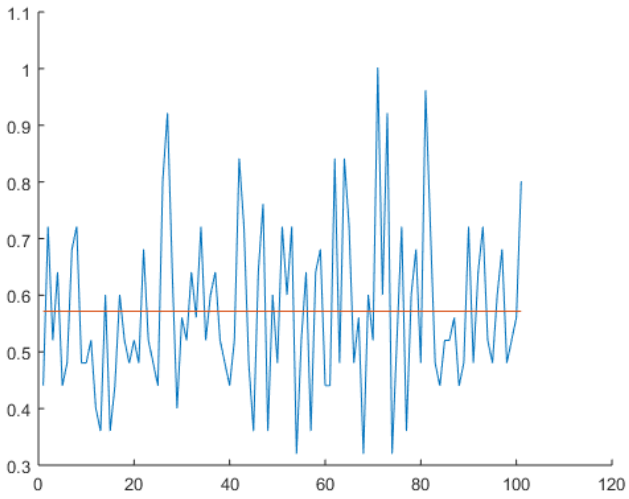


Figure 84. Part of the trace from which the noise measurement was taken. Red line presents the mean value.

The Signal to Noise Ratio was calculated as follows:

$$SNR = 10\log_{10}\frac{\sigma^2_{signal}}{\sigma^2_{noise}}\ dB \qquad (67)$$

where σ^2_{signal} is the signal power and σ^2_{noise} is the noise power calculated as described. The SNR is presented in decibels dB.

The dynamic range of each measurement for a single pixel equal to MCPF = 256. Pixel averaging (over 25 pixels) allows the signal to obtain an additional range of values between natural numbers, and averages out the noise to some degree. The noise averaging makes lifetime determination from signal amplitudes lower than 1 possible.

3.12.3. Changes in the FLIM setup

We used the same laser as an excitation and synchronisation source as in previous measurements, but used the wavelength of 355 nm as described in section 3.5.2. The beam was expanded using two lenses to the shape similar to the FOV of the camera. The generated 355 nm beam at the sample was 5 mW/cm^2. Samples were examined in 3 mL quartz cuvettes, typically containing about 1 mL of solution. The filter to block excitation light was a long-pass filter with a cut-off wavelength of 360 nm.

3.12.4. Single exponential measurements

3.12.4.1. Preparations and settings

Because NADH fluorescence is weaker than the previously tested fluorophores, we increased the sensitivity of the system. While keeping the same number of possible counts of MCPF=256, we increased the repetitions value to 2000 – so allowing 2000 attempts to measure a photon into the 1-bit buffer, then repeating this 256 times. This repetition value was chosen to keep all the measurements in the range of useful amplitudes.

We measured 4 different concentrations of free NADH: 40 µM, 20 µM, 10 µM and 5 µM. Subsequently we measured bound NADH of a concentration of 10 µM for a lifetime comparison. Free NADH samples were measured about 24 hours after preparation, giving lower signal intensities due to photobleaching over time, especially for lower concentrations that photobleach faster. Bound NADH was measured directly after preparation. The previously-measured ratio of signal intensity between bound and free NADH is 2:1 (166). In our case, the difference was up to 4:1, suggesting about 50% photo bleaching of the signal.

A difference compared to the measurements in the previous section is the fitting model. We decreased the part of the trace that is used for the fitting procedure further, to attempt to limit the effect of gate shape anomalies. Now, instead of using the whole peak which lasted over 700 samples, we chose only 200 samples of the decaying slope of the time trace. Using this part, which is a relatively regular and visually representative part of the trace, made the fitting slightly less vulnerable to small changes of the gate shape.

3.12.5. The results

An example of a frame taken by the sensor is presented in Figure 85. This is 1 out of 1280 images taken for the whole measurement. For the fitting procedure we are using only 200 images. This one is taken at the peak signal value for visibility. Because of the increased number of repetitions, the noise is also bigger. Some additional, more sensitive pixels are visible, whose value was still too small to be compensated by the algorithm.

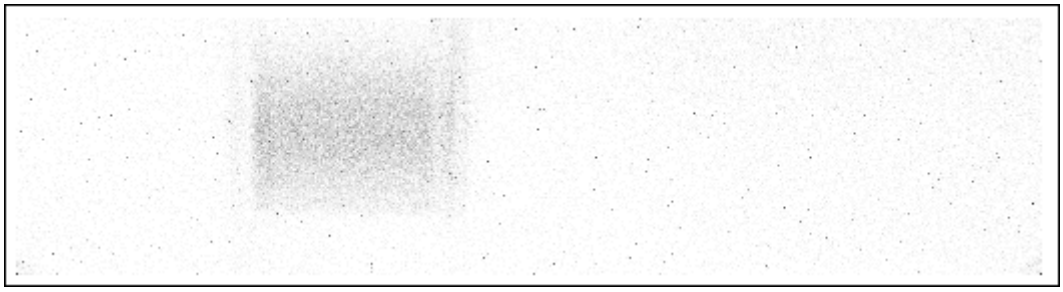


Figure 85. One frame of the cuvette with 40 µM solution of free NADH. Image presented in inverted colours.

Examples of the fit functions for a single 5x5 averaged pixel signal for all measured free NADH solutions are presented in Figure 86.

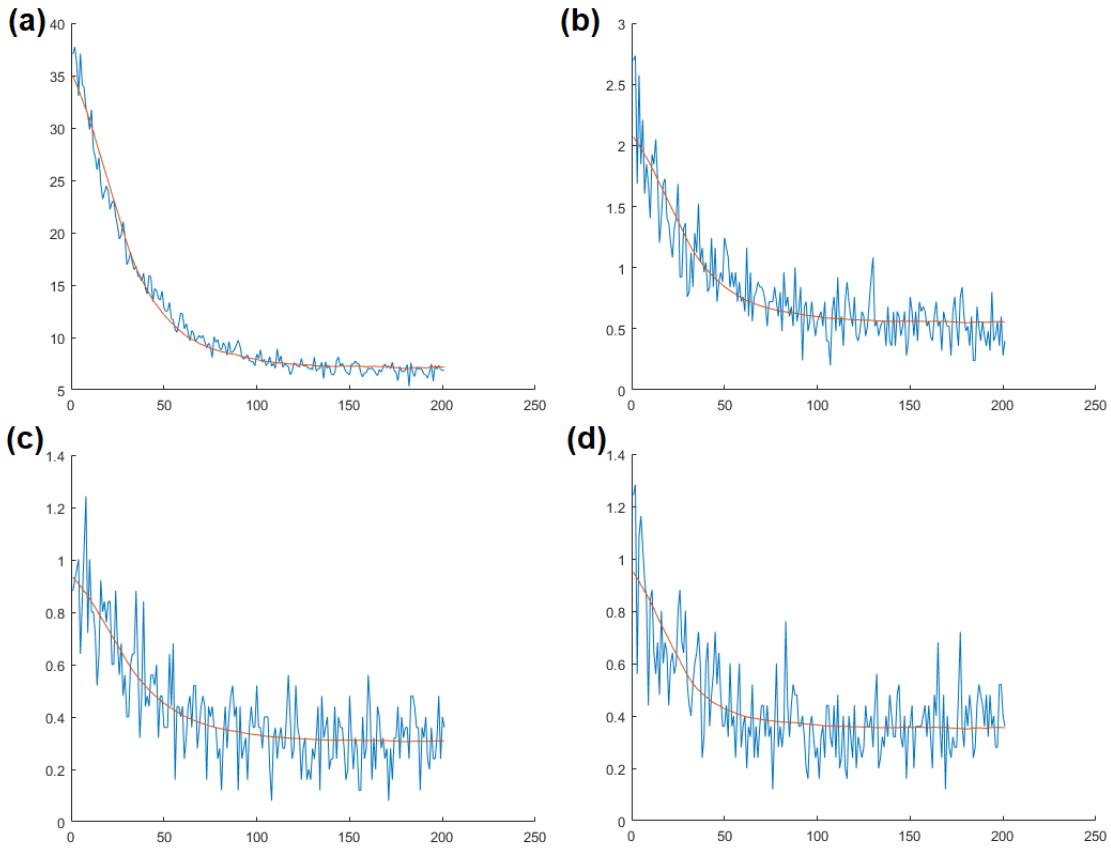


Figure 86 Example of the fit functions for a) 40 µM free NADH, b) 20 µM free NADH, c) 10 µM free NADH, d) 5 µM free NADH. Graphs also present trace and noise levels.

The results of the measurement of 40 µM free NADH are presented in Figure 87. A high signal intensity of about 35 counts and relatively low noise of 0.82 counts makes the measurement reliable and with relatively small spread. The mean lifetime value retrieved in the measured AOA is 0.37 ns with standard deviation of 0.026 ns.

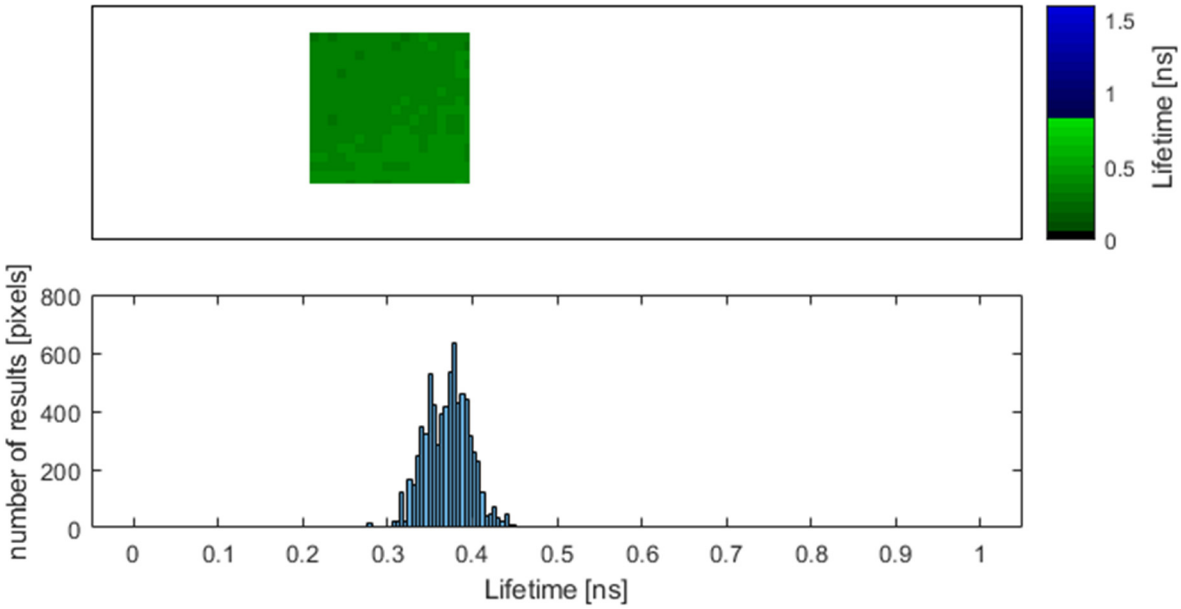


Figure 87. Results for 40 µM free NADH. Top panel presents masked, colour coded image of lifetimes (inverted colours); Bottom panel present distributing of measured lifetimes.

The signal level dropped drastically for 20 µM solution of free NADH, due to additional photobleaching of the weaker concentration. An amplitude of about 2 counts (Figure 86b) is still sufficient to determine lifetime with noise of 0.24 count. The colour coded image and the distribution of measurements are presented in Figure 88. The mean measured lifetime equals 0.39 ns and the standard deviation of the measured pixels is 0.092 ns.

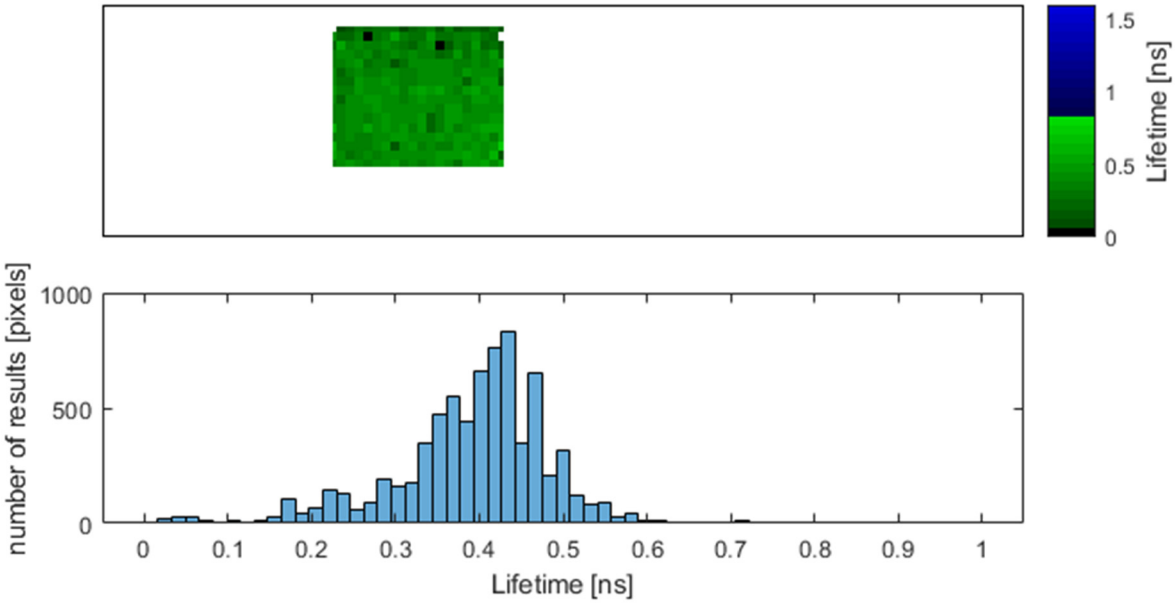


Figure 88.Results for 20 µM free NADH. Top panel presents masked, colour coded image of lifetimes (inverted colours); Bottom panel present distributing of measured lifetimes.

The 10 µM free NADH solution has an even lower signal intensity, reaching an amplitude of about 1 count on average and noise of 0.2 count. The trace with its fit is presented in Figure 86c. For a signal this low, the mean measured lifetime moves away from the reference value to 0.32 ns, with a standard deviation of 0.209 ns. Both colour coded image and the histogram though, show mostly good distribution of the measurements, which are disturbed by results close to zero where the signal was insufficient (smaller than presented in Figure 86).

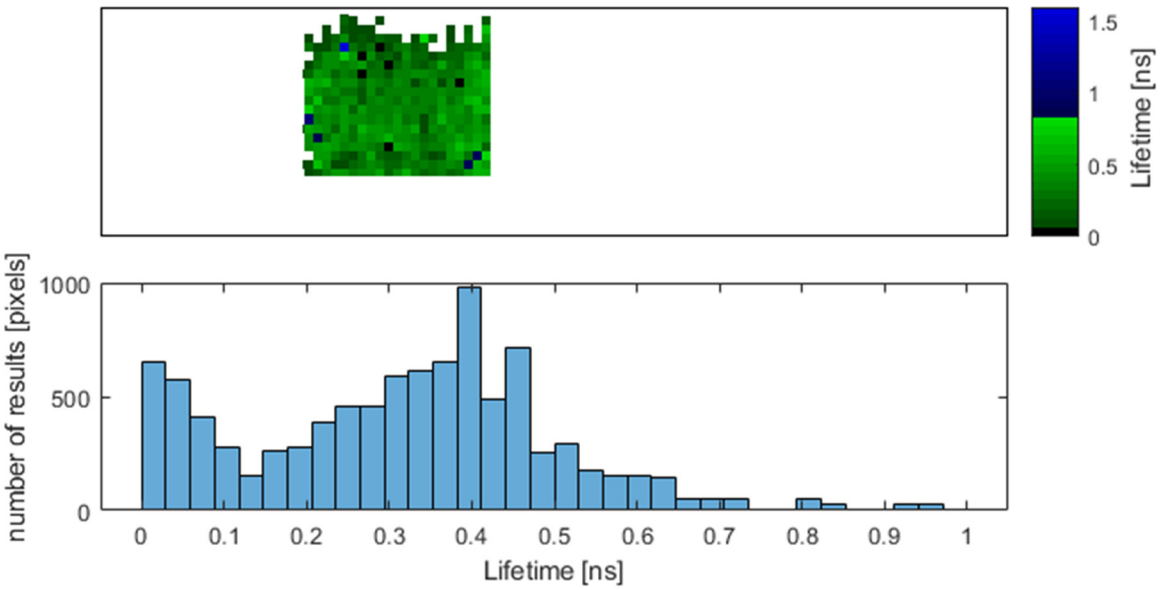


Figure 89. Results for 10 µM free NADH. Top panel presents masked, colour coded image of lifetimes (inverted colours); Bottom panel present distributing of measured lifetimes.

Finally, the lowest measured free NADH concentration was 5 µM. The signal decreased again to less than 1 count with a noise level of 0.19 count (Figure 86d). Spatial results for the measurement of this concentration are presented in Figure 90. The mean lifetime obtained from the whole are is 0.27 ns, with a standard deviation of 0.151 ns.

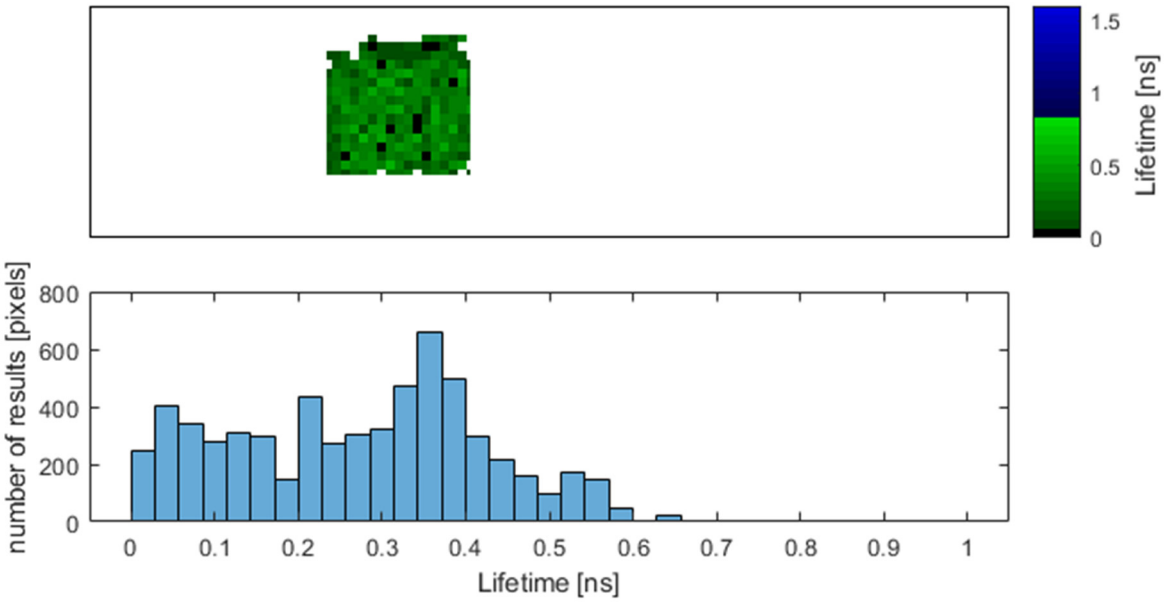


Figure 90. Results for 5 µM free NADH. Top panel presents masked, colour coded image of lifetimes (inverted colours); Bottom panel present distributing of measured lifetimes.

Subsequently, we measured a 10 µM concentration of bound NADH. The signal amplitude of about 8 counts was high enough to provide high readout accuracy with a noise value of 0.32 counts. The mean lifetime value of all measured pixels was 0.88 ns with a standard deviation of 0.061 ns.

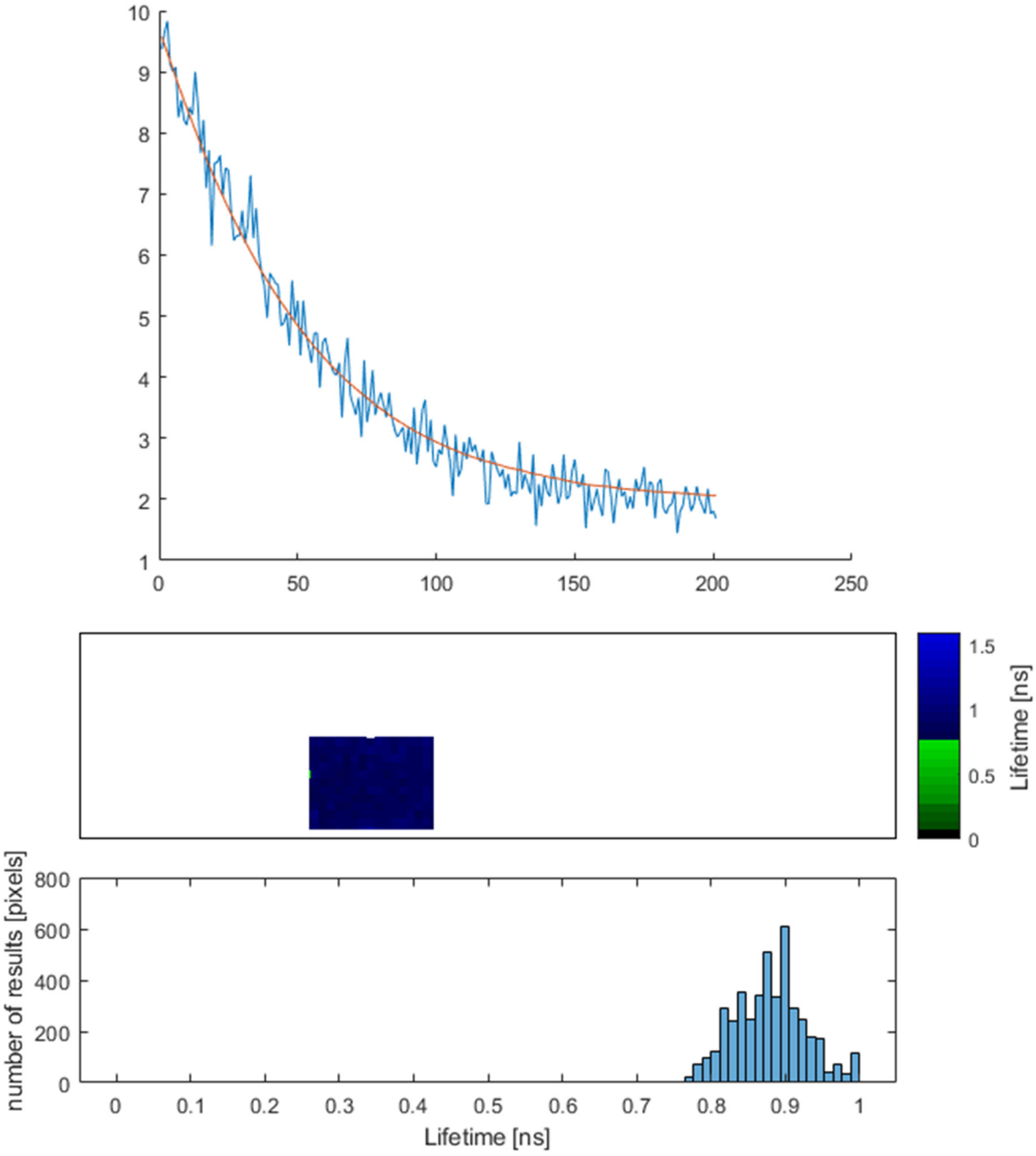


Figure 91. Example of the fit functions and results for 10 µM bound NADH. Top panel presents masked, colour coded image of lifetimes (inverted colours); Bottom panel present distributing of measured lifetimes

We also took measurements of selected samples with a single pixel trace fit (without averaging) for comparison. Examples of the traces, amplitude and fits are shown in Figure 92.

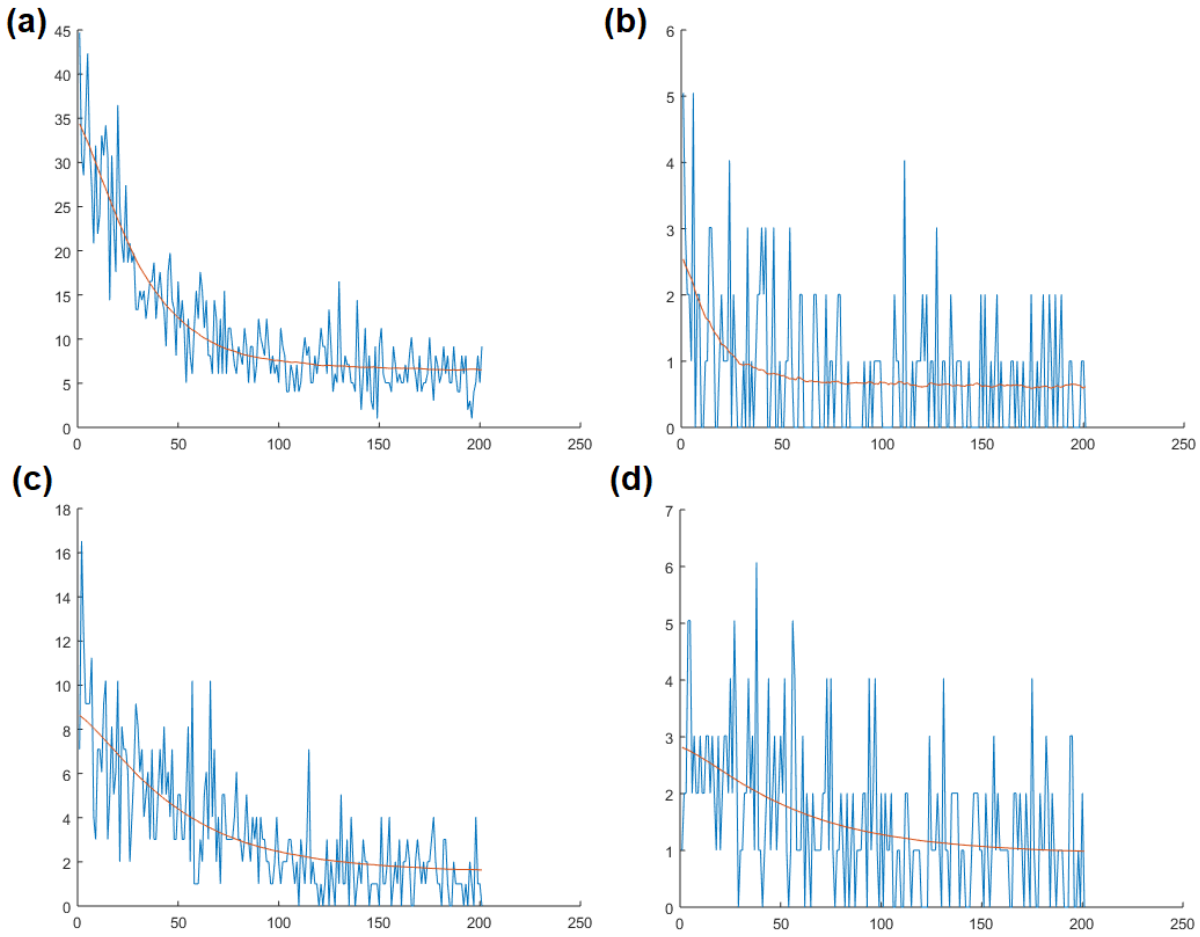


Figure 92. Example of the fit functions for single pixel measurements for a) 40 µM free NADH, b) 20 µM free NADH, c) 10 µM bound NADH, d) 10 µM bound NADH, for a pixel with smaller signal amplitude. Graphs also present signal and noise levels.

Figure 93 presents the same readout as Figure 87, but calculated for single pixels instead of groups of 5x5 pixels. The mean lifetime value of all presented pixels was 0.37 ns with standard deviation of 0.099 ns. The signal noise was 4.17 counts.

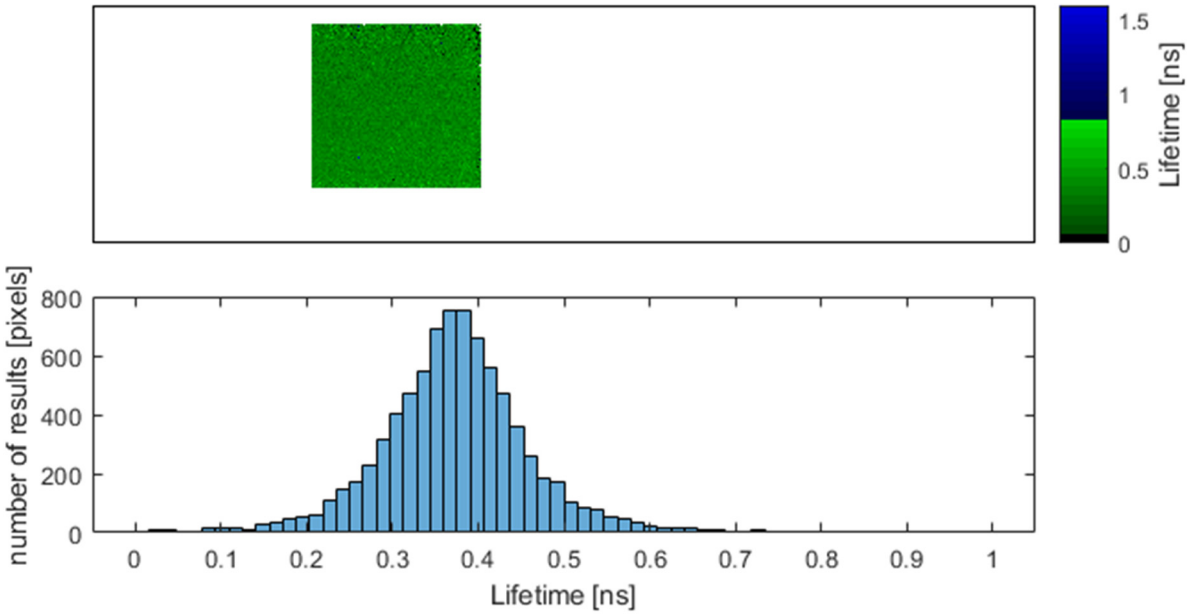


Figure 93. Results for 40 µM free NADH, single pixel measurements. Top panel presents masked, colour coded image of lifetimes (inverted colours); Bottom panel present distributing of measured lifetimes.

Figure 94, similar to the previous figure, present results of the single pixel measurements of the same readout as presented in Figure 88. The mean lifetime value equals 0.45 ns and the STD increased to a very large 0.384 ns. The noise of the signal was 1.65 counts.

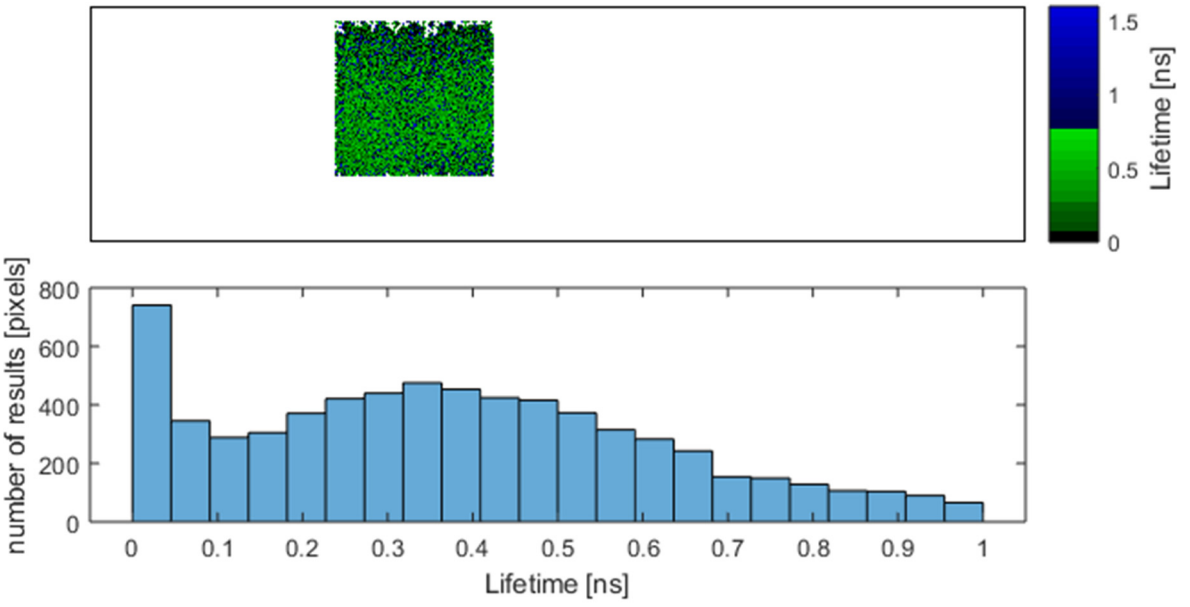


Figure 94. Results for 20 µM free NADH, single pixel measurements. Top panel presents masked, colour coded image of lifetimes (inverted colours); Bottom panel present distributing of measured lifetimes.

The final single pixel measurement was performed for the 10 µM bound NADH sample (Figure 95). The mean obtained lifetime is 0.89 ns with a standard deviation of 0.209 ns and a noise of 1.65 counts.

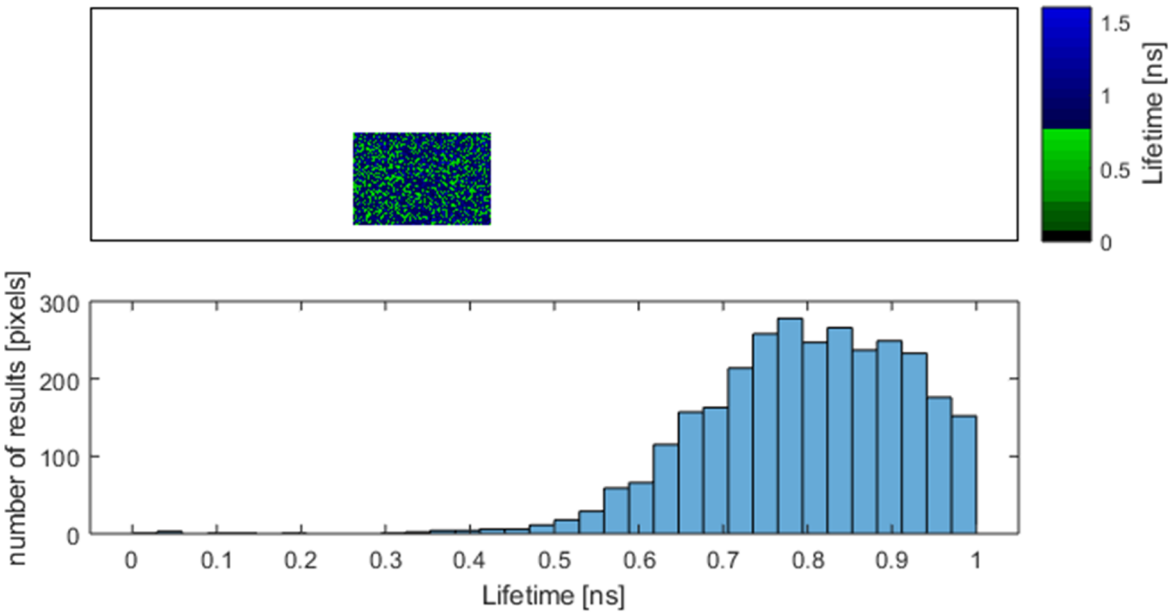


Figure 95. Results for 10 µM bound NADH, single pixel measurements. Top panel presents masked, colour coded image of lifetimes (inverted colours); Bottom panel present distributing of measured lifetimes.

3.12.5.1. Signal, noise, and accuracy analysis

All the measurements present are expected to have Gaussian distribution of acquired lifetimes, with the weakest data having additional values close to zero indicating a failure of the fitting. For the signal amplitude of the highest measure concertation (about 35 counts) we were able to measure the lifetime with the mean accuracy error of about 7%. The SNR was high, equal to 27.7 dB. A large signal level drop was observed for the twice-weaker solution of 20 µM. The amplitude of the signal of about 2 counts, and the noise of 0.31, still allowed for the reliable readout with SNR = 16.3 dB. The STD of 0.092 ns for the mean signal value of 0.39 provides a mean accuracy error of about 20%. It increased 3 times, compared to the signal level decreasing over 17 times and the SNR decreasing almost 2 times. Subsequent measurements of 10 µM and 5 µM free NADH solutions, have similar signal and noise levels, therefore the SNR. The mean calculated lifetime decreased in both cases, analysing histograms, because of the pixels with insufficient signal returning lifetimes

of 0 or close to 0. The general distribution remained similar with the addition of a peak around mentioned 0. The spread of given results increased, but more towards 0 than higher than expected lifetime. In both cases maximal retrieved lifetimes were about 0.6 ns.

The 10 µM bound NADH with signal amplitude of over 7 clearly showed higher lifetime than the free form. The mean value of 0.88 ns had a standard deviation of 0.061ns. The STD increased about 2 times for the amplitude changing 5 times, compared to the results of 40 µM free NADH. The SNR though was closer to the value for the 20 µM solution with the value of 16.8 dB.

The measured lifetime of the bound NADH is most likely lowered compared to the literature value of 1.0 ns for two reasons. An imperfect gate model was used for the fitting procedure (small changes in the system gate shape for different measurements), which consequently lowered all the measurements (of free NADH as well) and possible imperfection of the sample. If not all the NADH molecules were properly bound, the free ones in the mixture still emitted with the lifetime of 0.4 ns, which shortens the overall lifetime of the mixture, measured with the single exponential fit. We later investigate bi-exponential fitting to retrieve concentrations of each of bound and free NADH in a mixture.

Observing histogram spreads of measurements for free (Figure 89) and bound (Figure 91) 10 µM NADH's, the two distributions are almost entirely distinguishable, allowing for very reliable discrimination between these two forms.

The comparison of analysis of 25-pixel averages to single pixels shows, as expected, the benefit of averaging to increase the SNR. The obvious advantage of single pixel measurements is its higher spatial resolution, utilizing the maximum spatial capabilities of the sensor. In our setup, a single pixel has a scale length 0.2 mm in the image plane, whereas the 25-pixel averages have a scale length of 1 mm. Averaging will always present a trade-off between resolution and lifetime-retrieval accuracy, and best settings will depend on the signal strength and the resolution required by the application. Figure 92 presents traces acquired from single pixels, with the fit, allowing us to see the signal amplitudes and visualise the noise. The noise is the most obvious

change compared to averaged pixels and it increased almost 4 times in the case of 40 µM solution and almost 6 times for 20 µM. The histograms, because of the increased number of analysed traces, present a smoother distribution but with a wider spread. The STD increased almost 5 times for 40 µM and almost 4 times for 20 µM. The measurement for 40 µM solution was still accurate and compared all the measured factors, similar to the measurement of averaged 20 µM free NADH solution. This suggests that signal about 17 times smaller can be measured when averaging is used.

The measurement of 20 µM free NADH with single pixel fitting (Figure 94) did not work. On the histogram a similar trend of retrieving lifetimes close to zero was observed as for low concentration averaged measurements. The SNR below 1 dB (signal slightly higher than noise) is definitely insufficient for determination between free and bound NADH. Determination of lifetimes exceeding that difference of about 0.4 ns would still be possible though.

Figure 92c and Figure 92d present traces from two different pixels for the same measurement. The difference in amplitude is almost triple. This highlights another important role of pixel averaging: single pixels can have slightly different sensitivity and can collect different amount of light, which might lead to significantly different lifetime retrieval conditions. The averaging compensates for this issue.

All the results are gathered and presented in the Table 6. All the readouts were gathered from areas of about 5000 pixels.

Table 6. Results of measurements of free and bound NADH with single exponential fit. ^ denotes measurements for a single pixel.

NADH Solution	Mean τ [ns]	STD [ns]	Standard Error of Means	Signal amplitude [counts]	noise [counts]	SNR [dB]
free 40 µM	0.37	0.026	0	35	1.45	27.7
free 20 µM	0.39	0.092	0.001	2	0.31	16.3
free 10 µM	0.32	0.209	0.002	1	0.26	11.7
free 5 µM	0.27	0.151	0.002	0.7	0.25	9.1
bound 10 µM	0.88	0.061	0.001	7.3	1.06	16.8
free 40 µM^	0.37	0.099	0.001	35	5.77	15.7
free 20 µM ^	0.45	0.384	0.004	2	1.82	0.8
bound 10 µM^	0.89	0.209	0.002	7.3	2.32	10.0

The difference between reference lifetimes of free and bound NADH is 0.6 ns. From our measurements this difference is decreased to 0.5 ns because of the lower lifetime retrieved for bound NADH. This means the middle point between 0.4 and 0.9 is 0.65 ns. This can be chosen as a threshold for classification of the state of the NADH. Ideally, neither of the lifetime distributions would reach it and the separation would be full. Because of the high number of measurements (pixels) we can assume that small overlap of the very ends of the spreads (less than 5 per cent of the measurements) can happen.

In Figure 96a we plotted a standard deviation of the measured lifetime distributions as the estimation of the spread of readouts (Interquartile range was not representative because of the changing characteristics of the deviation for lower concentrations) against Signal to Noise Ratio of the same signal for averaged pixels (blue) and single pixel (orange) measurements. Visually, from histograms we can tell that the spread reaches the threshold highest lifetime value of about 0.65 ns for the STD = 0.2 ns (Figure 89) (and exceeds it by far for the STD = 0.38 ns (Figure 94)). From Figure 96a we can read that this point for the averaged pixels measurement will be met for the SNR of about 0 dB which means the signal is equal to the noise. From Figure 96b we can read that this might occur for the signal as low as about 0.1 counts.

The same point for the single pixel measurement will be reached for the SNR of about 10 dB, which is about the case of the measurements Figure 95 (for which the spread was slightly larger). From Figure 96b we can read that this SNR should be reached for the signal with the amplitude of about 10, which will be the smallest acceptable signal intensity for this case. (The measurement presented in Figure 95 achieved this SNR for a smaller amplitude, but the spread is higher than expected).

From Figure 96a we can also learn that the spread of single pixel readouts increases over three times faster with a decreasing SNR than for averaged values.

Comparing the noise from measurements that were performed for both single and averaged pixels, we can see that averaging decreased it about 5 times. The averaging allowed for measurements of signals at least 10 times smaller (based on measured samples), up to 65 times smaller (based on the estimations).

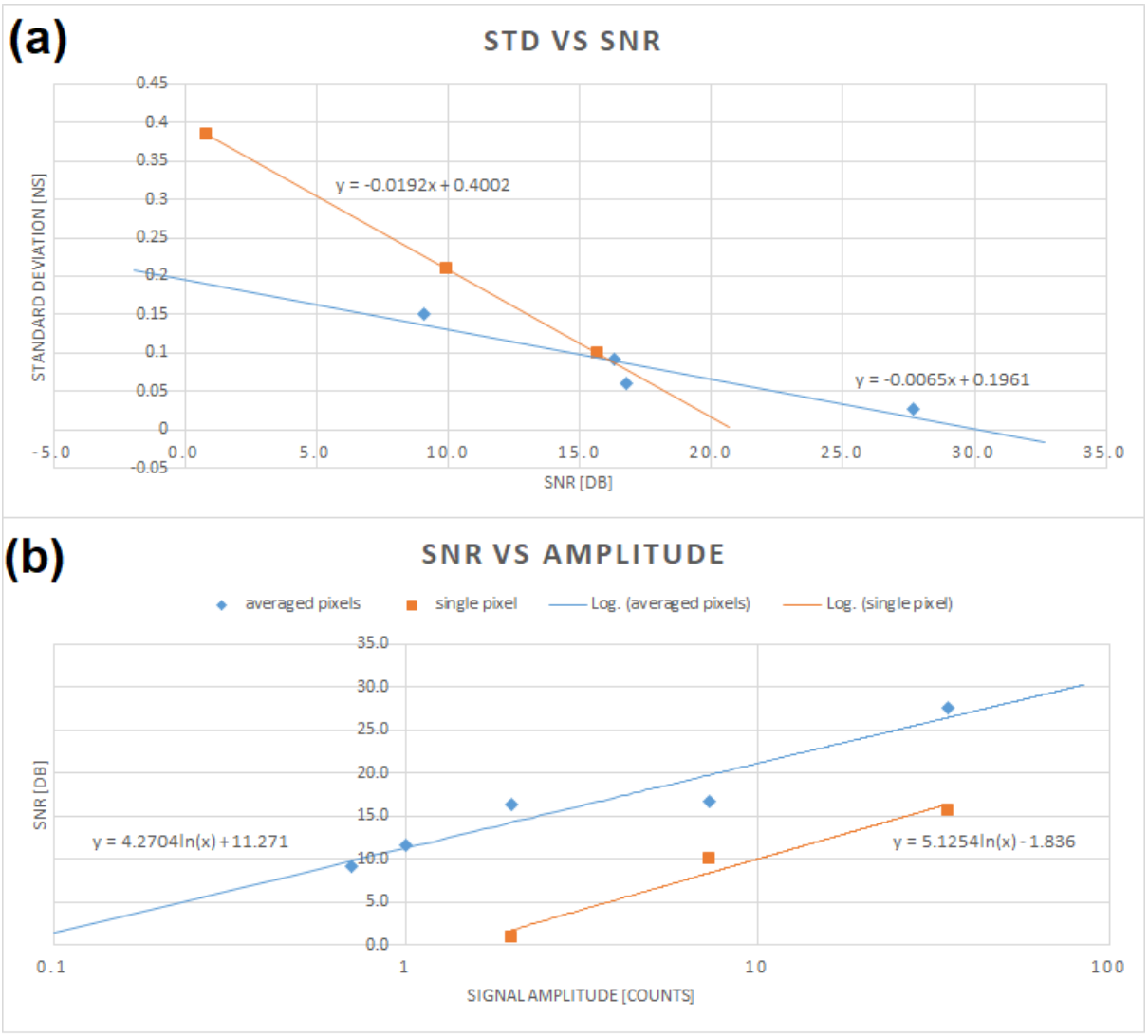


Figure 96. (a) standard deviation of measurements plotted against its Signal to Noise Ratio, (b) Signal to Noise Ratio plotted against the signals amplitude. Blue lines are for averaged pixels and orange lines are for single pixel measurements.

3.12.5.2. NADH lifetime differentiation from a single image

To confirm our findings, we carried out additional measurements with two cuvettes next to each other, with free and bound NADH imaged at once. The following figures present images as seen by the sensor on top, colour coded lifetime images in the middle and combined histograms of both lifetimes at the bottom. Figure 97 and Figure 98 present the same 30 µM bound and free NADH samples, measured with repetition of 2000 (standard) and 1000 (lowered for accuracy test) respectively. Figure 99 shows 10 µM bound and free NADH measured with standard parameters.

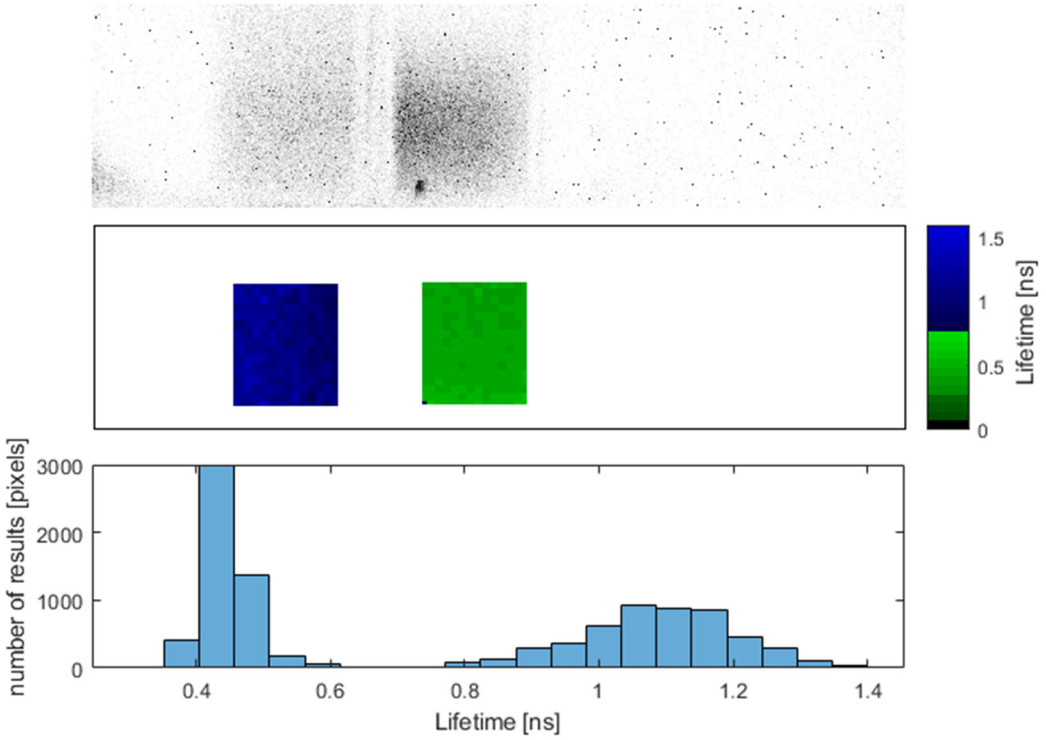


Figure 97. Bound and free NADH, both with concentrations of 30 µM. Top panel shows the frame from the system (inverted colours). Middle panel presents colour coded lifetime map. Bottom panel shows common histogram of all the presented lifetimes.

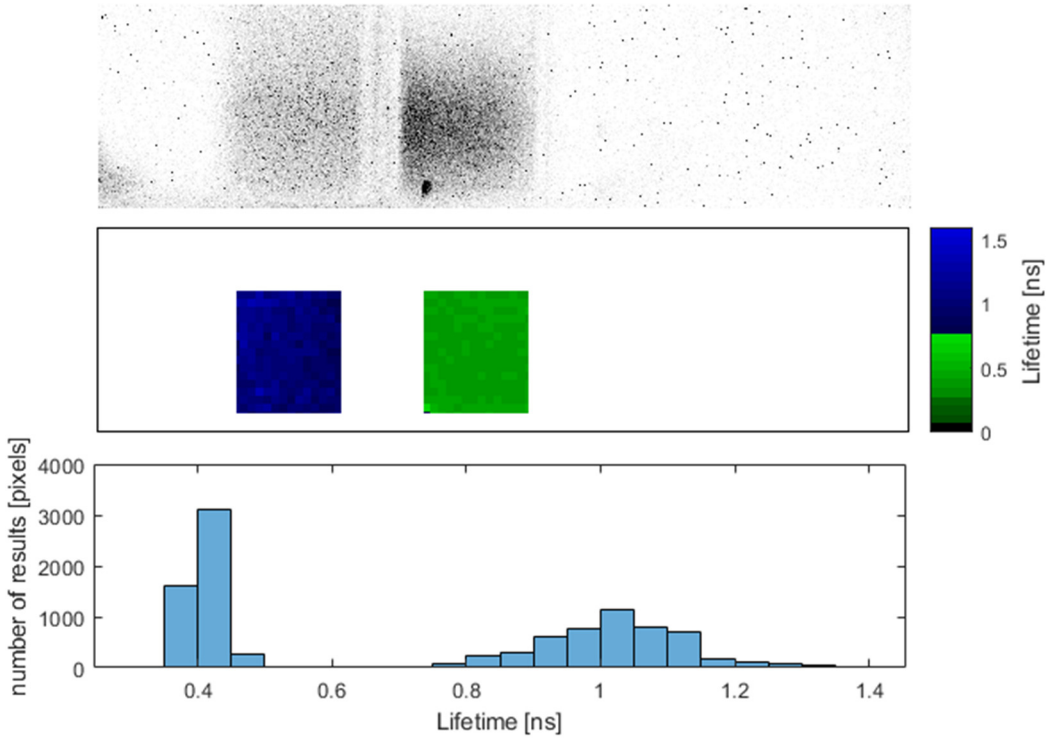


Figure 98 Bound and free NADH, both with concentrations of 30 µM measured with lowered repetition number of 1000. Top panel shows the frame from the system (inverted colours). Middle panel presents colour coded lifetime map. Bottom panel shows common histogram of all the presented lifetimes.

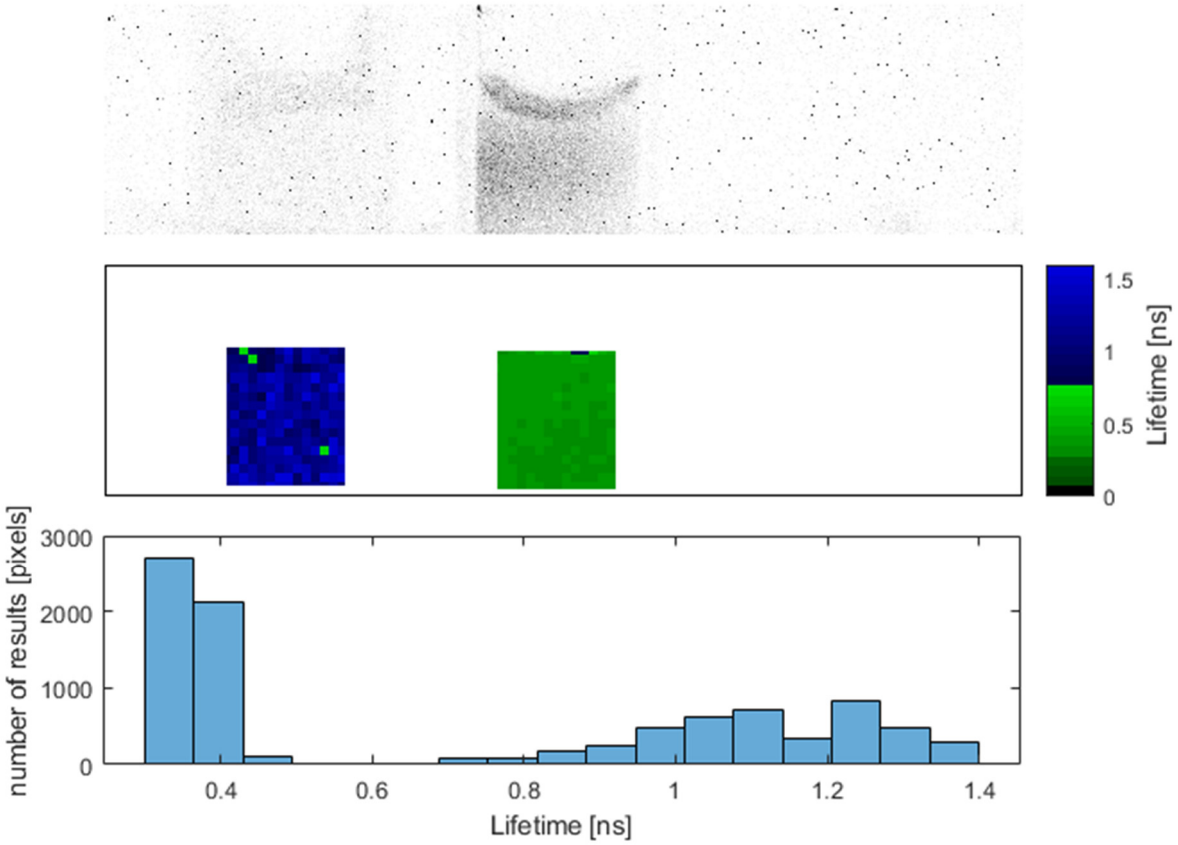


Figure 99. Bound and free NADH, both with concentrations of 10 µM. Top panel shows the frame from the system (inverted colours). Middle panel presents colour coded lifetime map. Bottom panel shows common histogram of all the presented lifetimes.

All the selected areas are of the same size: 65x75 pixels which equals to 4875 pixels. Designated results from these measurements are presented in Table 7.

Table 7. Results for a single exponential fit of free and bound NADH measured at the same time.

NADH Solution	**Mean** τ [ns]	**STD [ns]**	**Signal amplitude [counts]**	**Standard Error of the Means**
bound 30 µM	1.03	0.105	8	0.001
free 30 µM	0.41	0.034	32	0
bound 30µM low	1.09	0.113	4	0.002
free 30 µM low	0.45	0.038	18	0.001
bound 10 µM	1.2	0.348	1.5	0.005
free 10 µM	0.37	0.047	30	0.001

Because of the problem with chemicals, signals of the bound NADHs were strongly supressed. Instead of being up to twice larger, they were significantly smaller in all cases. However, according to the expectations this did not prevent SPAD from correct readouts. All the designated parameters drop in accuracy with decreasing amplitude

of the signal, but all of them are sufficient for full separation of measured lifetimes. All designated lifetimes are within the error range of proper literature values.

3.12.5.3. Single exponential lifetime for mixed NADHs

The final experiment using a single exponential fit was performed for different mixtures of free and bound NADH in one cuvette. Initially in a cuvette was 480 µL of pure 30 µM bound NADH. After the measurement, 160 µL of the same concentration of free NADH was added and the measurement was repeated. The same procedure was repeated one more time before the sample photo bleached. Lifetime measurements of these three consecutive measurements are shown in Figure 100.

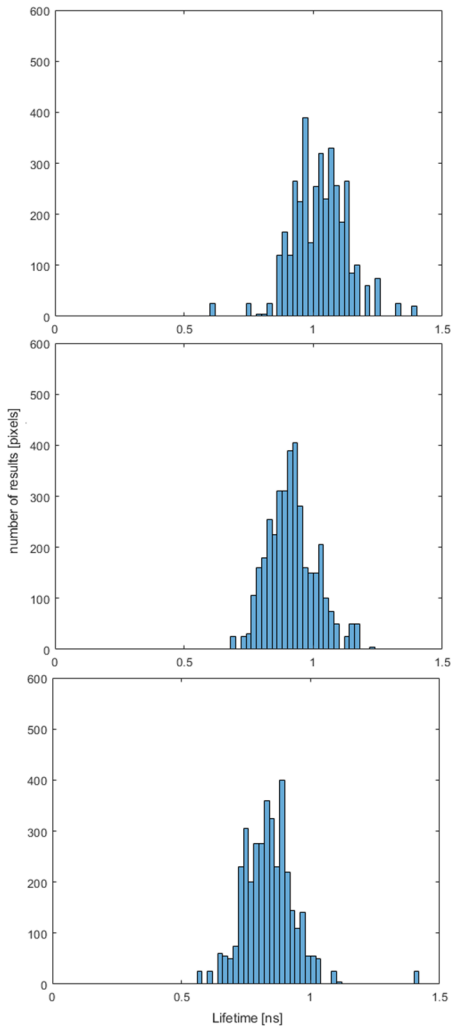


Figure 100. Bound NADH consecutively diluted with free NADH, measured with single exponential fit.

The mean measured lifetimes of these samples were respectively 1.02 ns, 0.92 ns, and 0.84 ns. All the standard deviations were very close to 0.1 ns. We can observe that the

measured lifetime slowly moves in the direction of smaller lifetimes with addition of the NADH with smaller lifetime. The last sample was a mixture of 3:2 of bound to free NADH. Additionally, bound NADH is normally prepared with double the amount of binding protein to assure 100% of bounds in the mixture; when free NADH added, additional binding is happening, decreasing the amount of free and increasing amount of bound NADH in the mixture. As expected, single exponential approximation of the mixture of two lifetimes, results with a value in between them, indicating in a quantitative sense the ratio of fluorophores. We can observe a decreasing mean lifetime with increasing amount of free NADH (with shorter lifetime) in the solution.

3.12.6.Biexponential measurements

The final step of this study was an implementation of a biexponential decay function to fit measured NADH with a new model. Free and bound NADH both have single exponential decays with lifetimes as previously stated. Using a biexponential model that looks for two signal component with different lifetimes should, for a mixture of free and bound NADH, return both lifetimes and indicate the relative strengths of their signals. Practically, bound NADH solutions almost always had some unbound NADH as well which could be detected through signal unmixing with the biexponential fit. Fitting the biexponential model to pure free or pure bound NADH samples should return two equal (or similar) mean lifetimes equal to the lifetime of a single exponential fit, with standard deviations similar to each other.

3.12.6.1. The biexponential model

The model is an updated formula from equation 63, which now presents as follows:

$$f_d(t) = A_1 e^{-\frac{t}{\tau_1}} + A_2 e^{-\frac{t}{\tau_2}} + C \qquad (68)$$

The new, second exponential functions adds two new parameters to the fitting function: amplitude and decay time of the second exponential signal component, A_2 and τ_2 respectively (parameters A and τ from the single exponential model was replaced with A_1 and τ_1). The Y offset C is common for both parts and does not add any new parameters. While τ_1 and τ_2 should provide two lifetimes of fluorophores in the mixture, their amplitudes A_1 and A_2 should provide the relative intensity of each

signal, which can be used to calculate the relative amount of each fluorophore in the mixture. The rest of the fitting procedure remained unchanged

3.12.6.2. Results

We took a measurement with a free NADH and the mixture of 2:1 free and bound NADH in cuvettes next to each other and analysed them with a biexponential fitting function. We introduce an additional element to the histogram – weighted counts. Since every biexponential result returns two pairs of lifetimes and their relative amplitudes, the counts can be weighted with their amplitudes – this means that if one of the two lifetimes has a much smaller amplitude, its weighting on the histogram is correspondingly reduced. Ideally, we expect two identical lifetimes returned for the single exponential free NADH sample, and two lifetime components of the mixture with amplitudes designating relative amount of one to the second one.

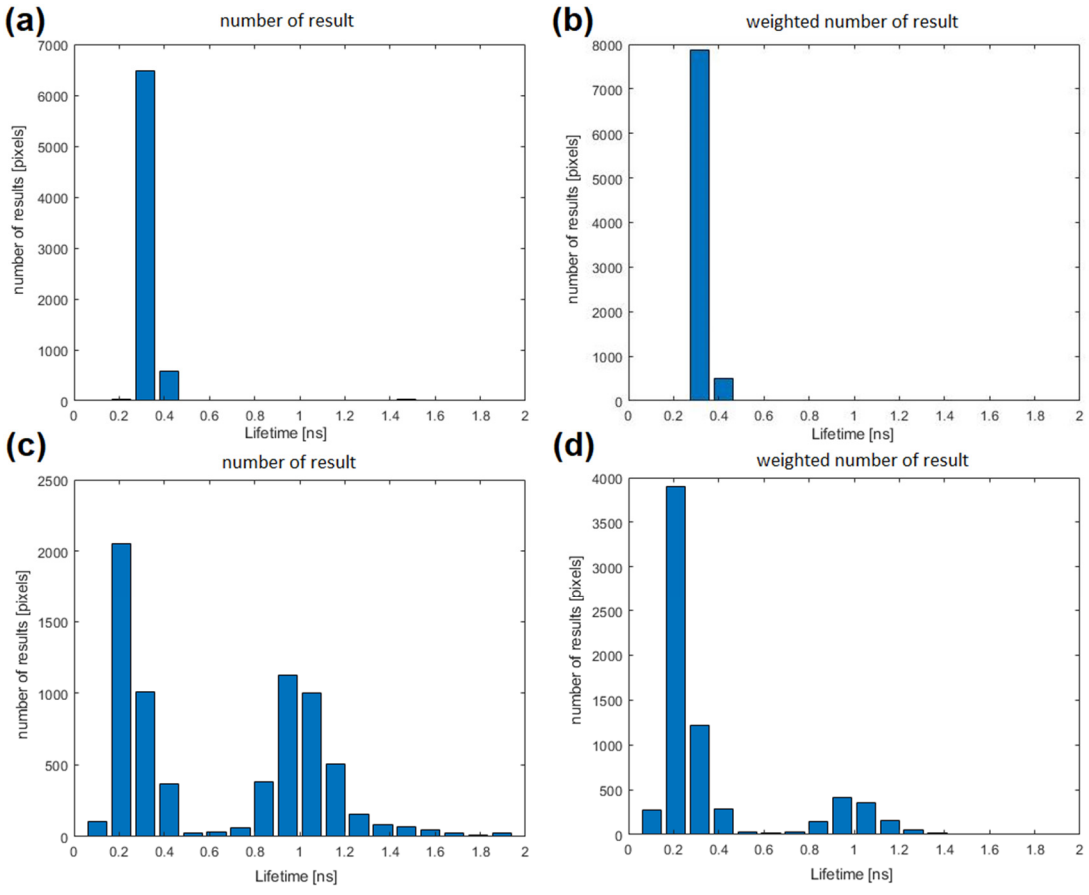


Figure 101. Lifetimes retrieved with biexponetial fit from: a) free NADH, regular histogram. b) the same free NADH, weighted histogram. c) NADH mixture, regular histogram. d) the same NADH mixture, weighted histogram.

Figure 101 presents results of biexponential fit with all the designated lifetimes on a single histogram. Figure 101a and Figure 101b show a measurement of the free NADH. Designated lifetimes were 0.34 ns with STD = 0.020 ns, and 0.36 ns with STD = 0.322 ns. Both lifetimes are very close to each other, but standard deviations are very different. The STD of the second one is probably massively increased by one mistaken lifetime visible at 1.5 ns. All the lifetimes are presented in one histogram and are very close to each other (falling mainly into two bins). Weighted histogram presented in panel b, decreases values of all the small bins, stating that these values had smaller amplitude – in this case these were probably mistaken.

Figure 101c, presents two lifetimes designated from the mixture. Clear separation between their lifetimes can be observed. Mean designated lifetime of free NADH was 0.26 ns with STD = 0.076 ns, and mean designated lifetime of bound NADH was 1.07 ns with STD = 0.315ns. The weighted histogram presented on panel d, shows decreased amplitude of the lifetimes of the bound NADH almost three times, meaning their designated amplitude in the pairs of lifetimes were relatively smaller to free NADH. This behaviour was not exactly as expected since this mixture should result in about 1:1 of free to bound molecules due to additional binding in an excessive amount of proteins from the bound NADH. The bound NADH STD (designated by the algorithm) also seems to be confirming that it had smaller relative amplitude. Lifetimes deviating more from the mean became smaller (meaning their amplitudes were smaller than others) as expected.

Measurement of the biexponential signal is significantly harder than single exponential decay. Higher signal amplitude (dynamic range) and SNR are required for the correct lifetime determination. We estimate this amplitude to about 30 counts with corresponding SNR for our algorithm to distinguish two lifetimes from the decay curve. Below this number, the algorithm was typically returning two similar lifetimes between expected ones. Biexponential fit is also less robust to small gate shape changes which happen to our system, introducing more errors. This is the most probable reason of lowered free NADH mean lifetimes. The algorithm worked mostly as expected, except significant decrease of the bound NADH readouts. With this experiment we confirm that the system is capable of the biexponential signal

unmixing, although some more analysis of optimal parameters and system behaviour is desired.

3.13. Measurement time

During this study we focussed on trying to optimize readout performance for best results, often using a large number of measurements, which resulted in long collection and calculation times. These times though, can be significantly shortened. In this section, we consider the time that might be needed to acquire an accurate FLIM image.

There are three parameters directly influencing the SPAD readout time:

- Repetitions – number of attempts to try to catch a photon for a single 1-bit readout
- Max Counts Per Frame – number of times that Repetitions will be executed to achieve MCPF possible readout levels
- Shifts – number of different gate delay shifts to be acquired (data points on X axis on graphs with signals – 1280 for most of presented).

All three parameters are multiplied by each other to achieve the total number of single pixel detection attempts. The minimum possible time would be the length of the cycle (25 ns) multiplied by this total number of attempt; however, the real time taken is in fact far longer.

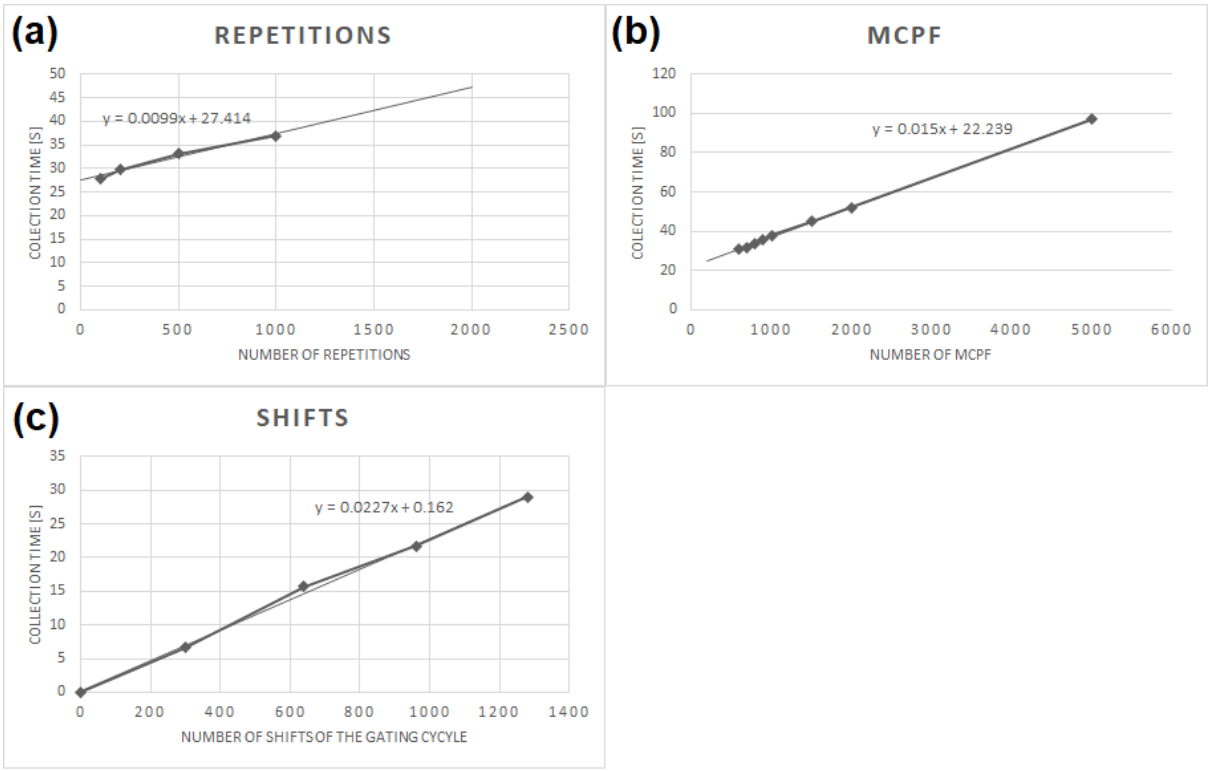


Figure 102. Collection times of the sensor in functions of different parameters.

Figure 102a shows how the actual collection time depends on repetitions for these same other settings of MCPF = 256 (our typical number) and 1280 shifts (full scan). Knowing that one laser cycle lasts 25 ns, for the number of repetitions = 200 the total number of required cycles lasts 1.6384 s; Figure 102a shows that it actually takes about 30 seconds. When the number of repetitions is increased to 2000, so the number of performed cycles is multiplied by 10, the required number of cycles lasts 16.384 s, so is about 15 seconds longer; Figure 102a shows it actually takes about 47 seconds for the system to finish, so is actually 17 seconds longer. Clearly additional repetitions take close to the minimum possible time, so do take place efficiently on consecutive cycles. The y-intercept of the Figure 102a shows that there is a 27 s constant operation time for the system for these settings, on top of which the actual collection time is added.

When analysing Figure 102b, we can see similar dependency as a function of Max Frames Per Second, suggesting a constant operational time about 22 seconds (for repetitions = 150 and shifts = 1280).

Figure 102c which presents collection time as a function of number of shifts (for repetitions = 150 and MCPF = 500) shows that the additional operation time is close to linearly proportional to shifts, and so must represent the time taken to read and save each image in sequence. The theoretical minimum collection time for each shift is 0.38 ms, but the gradient shows that each shift actually takes 23 ms. The great majority of the time then is required for the system to perform all the starting and finishing procedures, but predominantly for sending the file to the computer.

All the measurements carried out with the system, independently of the MCPF or repetition number, are saved as 16 bit values in a file with an additional header. This file has a size of 167,772,272 bytes for 1280 shifts, and is proportionately smaller for fewer shifts. Because of the USB 2.0 transfer speed limitation, this file is saved in the internal RAM memory of the system first, then transferred to the computer via USB 2.0 protocol.

When 8-bit images as we use (MCPF=256) are collected, the amount of data generated by the sensor is half of the file size actually used, and is just under the USB 2.0 transfer capability. This data could be sent directly to the computer for analysis (or analysed on chip) if the SPAD software had the option of an 8-bit file. Further speed up can be achieved by reducing the number of shifts. In the NADH experiments (section 3.12), we used only 200 shifts. Currently, collecting this trace takes 5.1 s, and with 8-bit files could be 2.6 s.

The collection time could in principle be shortened even further. Having a continuous finely spaced set of data for different gate delays is favourable but not essential for the fitting procedure. For a single exponential fit and ideal data, three points are required, because of three degrees of freedom (amplitude, lifetime, and offset). Sampling the signal with three groups of 10 consecutive data points (to average the noise) as shown in Figure 103, would again drastically decrease the collection time. This would allow for 2-3 Hz collection frequency for 2000 repetitions (low level signals) and over 20 Hz for 200 repetitions (higher intensity signals) for a single exponential fit. This means video-rate acquisition of image data.

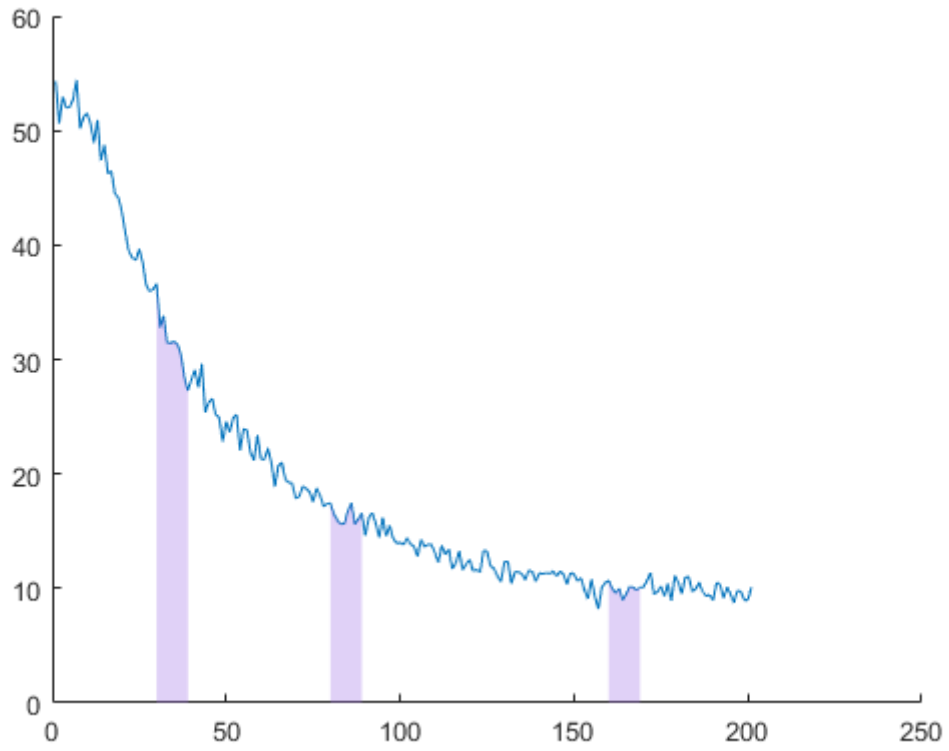


Figure 103. Representation of three clumps of data from the time trace that could be collected only for a fitting procedure.

The second part of the time needed to present results is lifetime calculation time. Again, we have not attempted to make this fast, and we used a time consuming method of least square optimisation. Typically, about 300 steps are needed to find the best fit in Matlab performed on the CPU (or GPU, the difference was very small) takes about 0.8 s for each point in the image. To calculate the whole image of averaged pixels, about 30 minutes is required. However, this number, can be decreased drastically as well.

From the computational point of view, if proper GPU computing were used (utilising all available GPUs which Matlab cannot do), these calculations could be completely parallelised, increasing slightly calculation time for a single pixel, but executing it for all of them at once (or divided into smaller groups, depending on the number of available GPUs).

The least square fitting procedure could also be significantly optimised, by setting a minimum step size (e.g., limiting lifetime step to 0.01 ns), or using neighbouring pixel results as an entry point for the next fit. Normalising all the traces and gates to 1 would allow to decrease number of variables to optimise to just τ (lifetime) for a single exponential fit and four degrees of freedom (two amplitudes and two lifetimes) for the biexponential procedure.

Finally, an optimised lookup table could be used to fit the trace with only a limited set of predefined decay lifetimes to find which of those are present and in what ratios. For example, using just two functions for theoretical free and bound NADH decay times would be extremely fast for analysing the biexponential data from section 3.12.6.

This approach has been studied, written to be performed on a FPGA system and a single SPAD pixel (167), and was subsequently expanded by the same group to a 32x32 array (168) and proven to achieve speed of over 30 Hz for a single exponential fit and over 100 Hz for lifetimes shorter than 1 ns (as in our case).

So, we are confident that possible improvements in both acquisition and analysis times can achieve sub-second images, which promises possibility of real-time readout with the system.

3.14. Discussion and conclusions

Using our SwissSPAD, a 512x128 pixel sensor we have built a wide-field imaging system, collecting weak fluorescent signals from a distance of about 20 cm with a FOV of 20 cm^2. We estimated the detectable lifetime range based on the trace modelling with the systems IRF to be in a range of 0.2 – 7 ns. This covers over 85% of fluorescent dyes, probes and labels that are frequently used for biological applications and in biomedical research, listed by the Fluorescence Foundation (169). We report that the detection accuracy decreases for longer lifetimes, because of the interference of subsequent laser pulses.

Initially, we reproduced the experiment of our colleagues (159) using only a few pixels to detect fluorescence lifetime of a very high intensity signal, without imaging. We

confirmed the proof of principle even for the lifetime region of relatively lower resolution.

Before we moved to the second stage of the lifetime determination, we carefully analysed the system's impulse response function. We carried out a number of experiments characterising the Impulse Response Function behaviour for various conditions. As a result, we decreased inconsistencies in the readout sensitivity depending on a gate position, through applying extra time for the sensor to recharge after recording each 1-bit information. We applied a count correction factor, which corrected for nonlinear saturation of the sensor. Some sensor's disadvantages, like a changing gate shape with the signal amplitude and with the gate position, were deeply analysed and solutions to these problems were applied in the updated fitting procedure.

For the second part we built a proper imaging system utilising the whole array of the SPAD and measured lifetimes of fluorophores with signals lower by 7 or more times. Lifetime retrieval of Rhodamine B over the area of about $2cm^2$ returned was very accurate with the mean $\tau = 1.69$ ns (1.68 ns literature value), with a STD of values of individual locations of 0.26 ns, and a Standard Error of Means of 0.016 calculated from the whole area. The signal had an amplitude of only 2% of the full range, or 10% of the designated useful range. Measured from the same image and equivalent area, the lifetime of Fluorescein was $\tau = 3.8$ ns (literature value 4 ns) with a STD of values of individual locations of 1 ns, and a Standard Error of Means of 0.0625 ns calculated from the whole area. Lower accuracy was caused by an extremely low signal level equivalent to 0.5 counts in amplitude, which is 1% of the useful range. This provided evidence of capability of lifetime retrieval from extremely low recorded signals with acceptable accuracy.

Lifetime retrieval from extremely low signals was possible due to applied signal averaging over 5x5 pixels. Decreased spatial resolution (effectively 102x25 px) was still larger than most available SPADs for single pixel measurements and still provided an effective pixel size smaller than 1 mm, allowing for imaging of small objects.

Measurements of two forms of NADH, free and bound with lifetimes respectively 0.4 ns and 1.0 ns with different concentrations of free NADH were carried out next. Excellent precision of the measurement was achieved for 40 μM free NADH with standard deviation of only 0.026 ns. This deviation was achieved for a SNR = 28 dB. For comparison, the same measurement analysed in a single pixel manner achieved SNR = 15.7 dB and STD = 0.099 ns. Our analysis shows that the SNR rises logarithmically with increasing signal amplitude, very quickly reaching a high ratio. Growth tempo is almost identical for both averaged and single pixels with a difference of 12 dB between respective amplitudes. Standard deviation grows almost three times faster with decreasing SNR for single pixel measurements.

The lowest measured free NADH concentration was 5 μM with a signal amplitude of 0.7 and STD = 0.151 ns. The mean lifetime was determined as 0.27 ns. We emphasise that this value was lowered by pixels with insufficient signal (lower than 0.7 counts) returning lifetimes close to 0. The histogram shows that the majority of pixels were measured accurately. We also emphasise that measured concentration yielded a lowered signal due to photo bleaching over the time of 24 hours. The signal strength of different NADH concentrations should decrease linearly (166), and thus a signal of 5 μM should be about 5 times higher, suggesting given results to be appropriate to about 1 μM free NADH concentration. Achieved results allow for the full separation of free and bound NADH readouts for the lowest measured concentrations for single exponential fit and averaged pixels. For the single pixel measurement, an amplitude of about 10 times higher is required to maintain the same accuracy, which translates to about 10 times higher NADH concentrations.

The following experiment with imaging free and bound NADH at the same time confirmed the complete lifetime separation is possible across the range of presented concentrations and signal amplitudes. It also pointed out a problem with bound NADH samples, which caused lower signals, and hence larger lifetime deviations. For the literature behaviour, bound NADH with twice the signal of free form would characterise with even smaller spread than free NADH, presenting even greater separation.

The final single exponential experiment presented the capability of this fit to average lifetimes of mixed fluorophores. We prepared three mixtures, adding small amounts of free NADH to the bound one. Because of the process of NADH binding, we expect the added free NADH to bind some more, so increasing the amount of free NADH in the mixture slowly. Figure 100 presents the expected shift of the mean lifetime value towards lower lifetimes with increasing free NADH amount in the mixture.

Subsequent experiments with biexponential fitting presented mostly expected behaviour. Single exponential free NADH sample was designated with two similar and close to literature lifetimes. The NADH mixture was unmixed with two almost separate lifetime groups. A slightly shifted (compared to the literature) short lifetime is most likely a result of the imperfect fitting trace. Relative amplitude based weighting lifted strong readouts closest to the mean values and supressed weak ones further from it. An unexpected finding was that the weighted histogram decreased readout of bound NADH while increasing free NADH. From the ratio of weighted counts close to 1:1, it changed to over 5:1 in favour of free NADH. This behaviour would be in agreement with the quenched bound NADH proposal, assuming that NADH in the mixture was bound but emitting a very faint signal. This would cancel the free NADH signal but would introduce only a small amount of the bound NADH, until some point. Once the critical mass of free NADH in the solution is exceeded, new molecules of free NADH would not be binding to free molecules of proteins and would rapidly elevate free the NADH signal in comparison to the bound form. We estimate that a proper biexponential trace unmixing requires a signal of about 30 counts to work properly. Adding appropriate limitations to the fitting function might increase accuracy and decrease required amplitude.

Analysing the speed of our system, we suggest that the current rather long time of data collection and lifetime calculation in our studies can be shortened multiple times by limiting of a number of gate positons and using published fast methods of lifetime calculation. A speed between 1-20 Hz seems possible keeping the presented sensitivity of the system.

We emphasise that the SwissSPAD is a prototype system, and had remaining imperfections and unintended performance characteristics. The system's IRF was far

from perfect and the shape of the gate changed sometimes for unspecified reasons. Small changes of electric current, imperfections on electrical and signal connectors, change of ambient temperature and even pixel to pixel variation could affect the shape of the gate, making the fitting procedure far more complex than ideal. Dealing with these issues, either electrically or in any part of the firmware/software/analysing script would be very beneficial for the systems reliability and repeatability.

Direct comparison between performance of different SPAD-FLIM systems is not easy, as everyone carries out different experiments. Even for the same measured fluorophores, factors like concentration, volume, excitation laser power, optics or distance will introduce variation. Typical reported STDs are between 0.2 – 0.4 ns (156, 157, 170) for fluorophores with nanosecond lifetimes, which is at the upper range for our measurements. These results were always acquired using many times smaller SPAD arrays, collecting signals attached to the microscope. The largest field of view of analysed papers was reported as 400 x 200 μm (157) which is about 2500 times smaller than presented by us. In the same paper the scientist reported that the STD of 0.2 ns was achieved for SNR = 26 dB, where the same STD in our case was achieved for single pixel measurement and SNR of 10-12 dB.

We now consider the application of our system to analysing biological samples. While we presented that our system is capable of determination of lifetimes of two mixed fluorophores with the biexponential fitting procedure, a biological environment typically is far more complex, presenting more than two lifetimes at once. As proven many times (137, 171), single exponential estimations of complex autofluorescent biological signals can be a good estimator cancerous regions for skin, breast, brain and other cancers.

In 2009 it was reported that the NADH concentration in a breast cancer tissue equals 168 μM, while for the healthy tissue it is 99 μM (172). Our robust discrimination of free and bound NADH for just 10 μM solutions suggest that our system in the current condition has the sensitivity for surface cancer detection. More limited FLIM systems that ours have been used previously to investigate this idea. In 2014 scientists published a paper with extensive analysis of NADH and NADPH using FLIM system (136). Imaging of single cells under the microscope allowed them to perform detailed

structural analysis of lifetimes, including biexponential measurements for some specific parts of cells. Their commercial TCSPC scanning FLIM system provided measurements with accuracy varying between 0.1 – 1 ns. In 2010 in a paper about brain tumours, scientist presented analysis of lifetimes for cancer and non-cancer regions of the brain tissue (including NADH analysis) for a field of view of 4 mm^2 (171). The system that is the closest to our approach was presented in 2010 (137). A complex system built around a CCD camera, presented a time-gated FLIM, providing a cm sized field of view for observing autofluorescent signals of cancer tissues under the excitation of 355 nm laser. The system provided an acquisition time of about 20 s. Lifetimes, however, were determined with significantly lower precision (standard deviation between 0.4 – 1.1 ns). Also, the degree of complexity of the system was significantly higher, compared to our integrated SPAD sensor.

The 355 nm laser illumination at the NADH sample is below 5 mW/cm^2. 355 nm is classified as a UVA radiation, which is a part of the emission spectrum of the Sun. In the peak hour of the day the sun emits 4.7 mW/cm^2 of the UVA radiation (173) which is comparable with our illumination power.

The flow-chart of the system is presented in Figure 104 below.

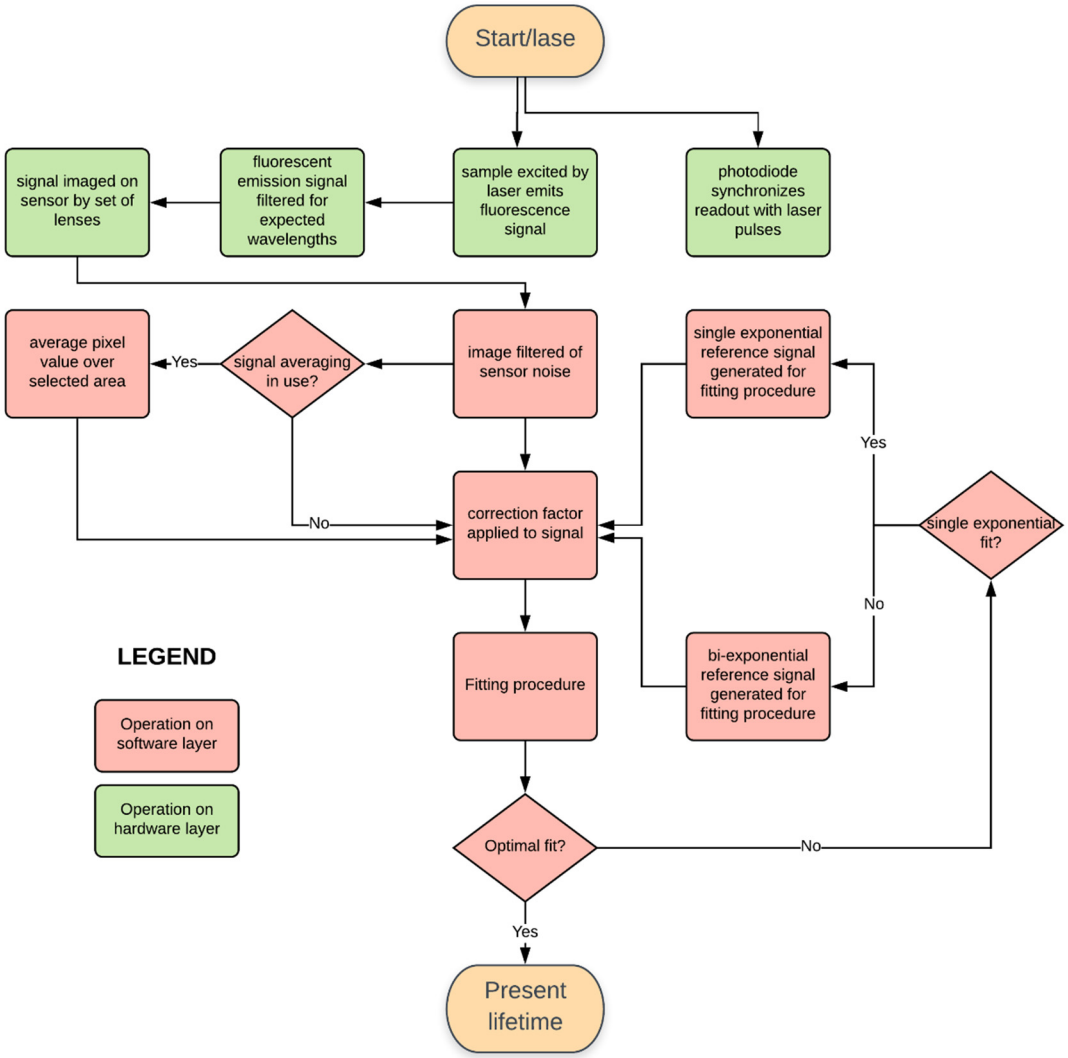


Figure 104. Flow-chart of the SPAD system.

4. Conclusions and Outlook

We have presented research in two areas, connected by the use of fluorophores in biomedical studies, but different in almost everything else. The smartphone part focused on the use of everyday electronics to provide medically relevant readout devices for home use, while the SPAD part utilised a cutting edge prototype SPAD array to construct an advanced imaging system for a professional use. Both approaches came with unique sets of advantages and disadvantages, making both projects so distinctively engaging.

Employing smartphones as light detection devices came to the world of science naturally since sensors not so different from the ones in phones are often used in professional readout devices. Indisputable advantages of smartphones are popularity and high quality of embedded cameras. We decided to take our own approach to the problem, which did not follow the main trend of 3d-printed attachments with incorporated optical elements inside. Instead, we put emphasis on simplicity and utilising commonly available elements.

In the first smartphone experiment, based on the fluorescence detection, we proved that all the parts of the typical spectrophotometric readout system, can be substituted with simple elements, most of which the typical modern household is equipped with. We used a tablet screen as an excitation source (replacing specifically designed lasers, diodes or lamps), tablets screen polarisation along with an additional polariser instead of set of filters, a simple cardboard box as an element providing a structure and shielding the detection system from ambient light, and finally a smartphone replacing a specialised sensor.

We showed that our system can compete with a spectrophotometer, being many times more expensive, for reading out the presented commercial assays. While less accurate, the smartphone still distinguished between very low concentrations of trypsin and collagenase assays. The low detection limit for the trypsin was 100 times smaller than required for the assessment of cystic fibrosis, and exceeded the required range by over 20 times with resolution allowing for over 2600 readout levels within

the range, which was more than sufficient to provide clinical information. The collagenase assay, which is an indicator of (inter alia) rheumatoid arthritis and joint inflammation, had significantly lower signal intensity but also covered relevant ranges for these medical conditions. The working curves produced by the smartphone and the spectrophotometer were almost identical in shape and the response of our system was very similar to that gold standard.

The second smartphone project pushed the concept of setup simplicity to its extreme, necessarily increasing the complexity of analysis. The method of acquiring data was simply to take a photo with the smartphone in reasonably uniform and typical illumination conditions (checked if within acceptable limits), with the sample placed on the white sheet of paper which served as a background and a colour reference point for the system. We studied an assay for IL-8; elevated IL-8 concentrations above 550 pg/mL may (103) indicate ongoing infection and/or inflammation which is a common ailment for cystic fibrosis infant patients. The only available commercial assay for IL-8 detection is designed to be read out with an absorbance measurement, and has a suggested upper detection limit of just 200 pg/mL. Our decision to perform a colour analysis instead of an absorbance analysis increased the assay working range 18 times. Our system was capable of distinguishing concentrations up to 3600 pg/mL, so enabling the use of this assay to detect the condition. Clinical samples measured with the system saw IL-8 concentrations of up 1770 pg/mL.

We believe that emphasising simplicity is the most appropriate approach for smartphone based systems, because of the possibilities that it brings. We envision that the next step after Point-Of-Care Testing (in presence of the medically trained person) scheme, will be patients' self-testing for first indications of common diseases, or for regular monitoring of very specific conditions (like cystic fibrosis). This step calls for the most available, easy to use yet still accurate solutions, accessible for everyone.

Simple quantitative chemical and medical testing using smartphones could bring analysis to places it is currently not available. POCT is only the beginning of the 'on the spot' testing applications. Simple extensions of the idea include for example veterinarians being able to test animals in rural areas, or small facilities (e.g., food production, breweries or distilleries) testing their product's quality in the place where

it is produced. The very low cost of the solution could encourage testing in places where it is not currently considered. Ultimately, reliable rapid colour analysis using a smartphone in any situation could bring a new era of consumer tests (e.g., colour based freshness of meat or ripeness of fruits). Our work is a step toward accounting for the many variables in such an uncontrolled test environment to ensure reliable results.

Constant rapid development of smartphones brings many opportunities but also has drawbacks. The multiplicity of smartphones on the market, with new ones arriving each month, brings a range of variations in their sensors' performance. Methods that are applicable to most phones may not work (or will drastically decrease in performance) for others. The impossibility of testing all the models will always limit the market to the certain extent. On a positive note, cameras in smartphones are still improving at a very rapid pace, allowing for more accurate scientific use each year. Relatively new to high class smartphones is raw (uncompressed linear data) image acquisition, which will be a great help in scientific image analysis.

The next reasonable step towards more accurate smartphone colour determination will be in-depth analysis of the variables hidden in metadata of each file saved by the smartphone. Precise information about the image resolution or used colour space, but also an exposure time, F-number or focal length can be found. Others of these variables are records of the image processing done by the phone, potentially helping to understand and undo any changes, to assist absolute colour comparison. The help of the sensor and smartphone manufacturers in understanding how the image is processed would be invaluable, though it Is very unlikely that they will share their secrets to the public or even scientists.

The SPAD-FLIM project was divided into 3 main parts. In the first part, we built a first version of the system, applied a simple version of the lifetime retrieval algorithm and performed a proof of concept measurement using a fluorophore whose lifetime was confirmed by a reference measurement. The second part resulted from the fact that the sensor was a cutting-edge, new prototype device which was not fully tested and characterised. We spent a significant amount of time for IRF analysis, detection and characterisation of unwanted behaviour, and developing and applying corrections. For

the third and final part, we constructed a fully operational wide-field FLIM system, with the field of view of about 20 cm^2, imaged from the distance of about 20 cm. Our target fluorophores were NADH in a free and bound forms, which can indicate cancer cells of breast, brain or skin tumours in the human body. We tested the accuracy of the lifetime determination for a range of medically-relevant concentrations of fluorophores. Signal noise and lifetime standard deviation were characterised for two modes of data analysis – single pixel and averaged pixels analysis. Both of these modes were similar to, or outperformed, other published SPAD-FLIM systems. In terms of a field of view, our system improved by three orders of magnitude (158). The resolution, even when effectively decreased by pixel averaging was still at least two times higher than other largest SPAD (170) (10 times without averaging).

Comparing our performance with literature data, we suggest that our system is capable of a surface cancer detection based on NADH lifetime analysis in a real time regime (1 - 20 Hz). Comparing performance of our system with a published CMOS FLIM system with similar field of view (137), ours presents an integrated 'on-chip' solution against complex system built of many separate parts. The reported accuracy was 2 – 4 times better in favour of our system.

The next step towards cancer screening with our system will be determining capabilities and limits while imaging live and fixed, artificially grown cancer cells, and imaging of matching healthy cells. Subsequently, tissues samples with both healthy and cancer regions, can be studied to accurately represent the complex target environment of human body. Finally, optimisation of the collection time and time of the lifetime determination algorithm (as proposed) will open new doors for the system application in biomedical sciences.

Potential of real time lifetime analysis allows us to consider applying such a system in an actual surface cancer detection. Scanning the whole body for melanoma changes would be possible with presented field of view, providing digital images with colour coded cancerous tissues. The system could be used during brain or breast cancers operations, where fixed at the position of the surgical extraction, it could highlight cancer tissues to be removed on the screen. Ultimately, Augmented Reality should be considered. The field of view of the camera could be specifically marked for the AR to

recognise this part. Subsequently, lifetime information retrieved by the SPAD could be sent to the AR headset and applied as a digital layer in the previously marked area, where the lifetime determination is executed. This way, using the simplest AR technique, the medical doctor could see a highlight of cancer tissues on top of the regular vison. This solution would save the need of looking at the digital screen instead of the patient. A view of this future is given in Figure 105.

Figure 105. A visualisation of the medical doctor during an operation, wearing augmented reality goggles connected to the outcome of the SPAD sensor to highlight cancer tissues on top of the regular vision. Photograph: Medical Realities/PA

Contribution

In the first smartphone project Ayad G. Anwer contributed to collagenase assays.

In the same project Wei Deng contributed to trypsin assays.

In the second smartphone project Luke W. Garratt prepared IL-8 and NE clinical samples and acquired images with smartphones for analysis.

In the SPAD-FLIM project Ayad G. Anwer and Saabah B Mahbub contributed to NADH preparation.

References

1. Lakowicz JR. Principles of Fluorescence Spectroscopy, (1999). Kluwer Academic/Plenum Publishers, New York; 2004.
2. Haugland RP. The handbook: a guide to fluorescent probes and labeling technologies: Molecular probes; 2005.
3. Stokes GG. On the change of refrangibility of light. Philosophical transactions of the Royal Society of London. 1852;142:463-562.
4. Kasha M. Characterization of electronic transitions in complex molecules. Discussions of the Faraday society. 1950;9:14-9.
5. Eftink MR, Ghiron CA. Fluorescence quenching studies with proteins. Analytical biochemistry. 1981;114(2):199-227.
6. Lehrer S. Solute perturbation of protein fluorescence. Quenching of the tryptophyl fluorescence of model compounds and of lysozyme by iodide ion. Biochemistry. 1971;10(17):3254-63.
7. Lakowicz JR, Weber G. Quenching of fluorescence by oxygen. Probe for structural fluctuations in macromolecules. Biochemistry. 1973;12(21):4161-70.
8. Förster T. Zwischenmolekulare energiewanderung und fluoreszenz. Annalen der physik. 1948;437(1-2):55-75.
9. Dexter D, Schulman JH. Theory of concentration quenching in inorganic phosphors. The Journal of Chemical Physics. 1954;22(6):1063-70.
10. Lamola AA, Hammond GS. Mechanisms of photochemical reactions in solution. XXXIII. Intersystem crossing efficiencies. The Journal of Chemical Physics. 1965;43(6):2129-35.
11. Vogelsang J, Kasper R, Steinhauer C, Person B, Heilemann M, Sauer M, et al. A Reducing and Oxidizing System Minimizes Photobleaching and Blinking of Fluorescent Dyes. Angewandte Chemie International Edition. 2008;47(29):5465-9.
12. Giloh H, Sedat J. Fluorescence microscopy: reduced photobleaching of rhodamine and fluorescein protein conjugates by n-propyl gallate. Science. 1982;217(4566):1252-5.
13. Vista S. Use of Humic acid in agriculture2015.
14. Hart SJ, JiJi RD. Light emitting diode excitation emission matrix fluorescence spectroscopy. Analyst. 2002;127(12):1693-9.
15. Landgraf S. Use of ultrabright LEDs for the determination of static and time-resolved florescence information of liquid and solid crude oil samples. Journal of biochemical and biophysical methods. 2004;61(1-2):125-34.
16. Wang Y, Shyy JY-J, Chien S. Fluorescence proteins, live-cell imaging, and mechanobiology: seeing is believing. Annu Rev Biomed Eng. 2008;10:1-38.
17. Liu C, Wang X, Zhou Y, Liu Y. Timing and Operating Mode Design for Time-Gated fluorescence lifetime imaging microscopy. The Scientific World Journal. 2013;2013.
18. Jares-Erijman EA, Jovin TM. FRET imaging. Nature biotechnology. 2003;21(11):1387.
19. Thompson NL. Fluorescence correlation spectroscopy. Topics in fluorescence spectroscopy: Springer; 2002. p. 337-78.
20. Xu C, Zipfel W, Shear JB, Williams RM, Webb WW. Multiphoton fluorescence excitation: new spectral windows for biological nonlinear microscopy. Proceedings of the National Academy of Sciences. 1996;93(20):10763-8.
21. Meyvis TK, De Smedt SC, Van Oostveldt P, Demeester J. Fluorescence recovery after photobleaching: a versatile tool for mobility and interaction measurements in pharmaceutical research. Pharmaceutical research. 1999;16(8):1153-62.
22. Judd DB. Color in business, science and industry. 1952.
23. Commission IE. IEC 61966-2-1, 1999. Multimedia systems and equipment–Colour measurements and management–Part. 1999:2-1.
24. Standardization IOf. ISO 11664-2:2007 (CIE S 014-2/E:2006)

Colorimetry -- Part 2: CIE standard illuminants. 2007.
25. Orth A, Wilson E, Thompson J, Gibson B. A dual-mode mobile phone microscope using the onboard camera flash and ambient light. Scientific reports. 2018;8(1):3298.
26. Cova S, Ghioni M, Lacaita A, Samori C, Zappa F. Avalanche photodiodes and quenching circuits for single-photon detection. Appl Opt. 1996;35(12):1956-76.
27. Rochas A, Gosch M, Serov A, Besse P, Popovic R, Lasser T, et al. First fully integrated 2-D array of single-photon detectors in standard CMOS technology. IEEE Photonics Technology Letters. 2003;15(7):963-5.
28. Poushter J, Bishop C, Chwe H. Social Media Use Continues to Rise in Developing Countries but Plateaus Across Developed Ones. Pew Research Center. 2018;22.
29. Lanaj K, Johnson RE, Barnes CM. Beginning the workday yet already depleted? Consequences of late-night smartphone use and sleep. Organizational Behavior and Human Decision Processes. 2014;124(1):11-23.
30. Kwon M, Lee J-Y, Won W-Y, Park J-W, Min J-A, Hahn C, et al. Development and validation of a smartphone addiction scale (SAS). PloS one. 2013;8(2):e56936.
31. Oulasvirta A, Rattenbury T, Ma L, Raita E. Habits make smartphone use more pervasive. Personal and Ubiquitous Computing. 2012;16(1):105-14.
32. Demirci K, Akgönül M, Akpinar A. Relationship of smartphone use severity with sleep quality, depression, and anxiety in university students. Journal of behavioral addictions. 2015;4(2):85-92.
33. Wang D, Xiang Z, Fesenmaier DR. Smartphone Use in Everyday Life and Travel. Journal of Travel Research. 2016;55(1):52-63.
34. Anderson DP, editor Boinc: A system for public-resource computing and storage. Grid Computing, 2004 Proceedings Fifth IEEE/ACM International Workshop on; 2004: IEEE.
35. Sullivan BL, Aycrigg JL, Barry JH, Bonney RE, Bruns N, Cooper CB, et al. The eBird enterprise: an integrated approach to development and application of citizen science. Biological Conservation. 2014;169:31-40.
36. Kong Q, Allen RM, Schreier L, Kwon Y-W. MyShake: A smartphone seismic network for earthquake early warning and beyond. Science Advances. 2016;2(2).
37. Silvertown J. A new dawn for citizen science. Trends in ecology & evolution. 2009;24(9):467-71.
38. Whiteson D, Mulhearn M, Shimmin C, Cranmer K, Brodie K, Burns D. Searching for ultra-high energy cosmic rays with smartphones. Astroparticle Physics. 2016;79:1-9.
39. Anderson DP, Cobb J, Korpela E, Lebofsky M, Werthimer D. SETI@ home: an experiment in public-resource computing. Communications of the ACM. 2002;45(11):56-61.
40. Hoot JE, editor Photometry with DSLR cameras. Society for Astronomical Sciences Annual Symposium; 2007.
41. Kardous CA, Shaw PB. Evaluation of smartphone sound measurement applications. The Journal of the Acoustical Society of America. 2014;135(4):EL186-EL92.
42. Yen T-H, Chang C-Y, Yu S-N, editors. A portable real-time ECG recognition system based on smartphone. Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE; 2013: IEEE.
43. Comtois G, Salisbury JI, Sun Y, editors. A smartphone-based platform for analyzing physiological audio signals. Bioengineering Conference (NEBEC), 2012 38th Annual Northeast; 2012: IEEE.
44. Martinez AW, Phillips ST, Carrilho E, Thomas III SW, Sindi H, Whitesides GM. Simple telemedicine for developing regions: camera phones and paper-based microfluidic devices for real-time, off-site diagnosis. Analytical chemistry. 2008;80(10):3699-707.
45. Tseng D, Mudanyali O, Oztoprak C, Isikman SO, Sencan I, Yaglidere O, et al. Lensfree microscopy on a cellphone. Lab on a Chip. 2010;10(14):1787-92.

46. Zhu H, Yaglidere O, Su T-W, Tseng D, Ozcan A. Cost-effective and compact wide-field fluorescent imaging on a cell-phone. Lab on a Chip. 2011;11(2):315-22.
47. Zhu H, Mavandadi S, Coskun AF, Yaglidere O, Ozcan A. Optofluidic Fluorescent Imaging Cytometry on a Cell Phone. Analytical Chemistry. 2011;83(17):6641-7.
48. Smith ZJ, Chu K, Espenson AR, Rahimzadeh M, Gryshuk A, Molinaro M, et al. Cell-phone-based platform for biomedical device development and education applications. PloS one. 2011;6(3):e17150.
49. Zhu H, Sikora U, Ozcan A. Quantum dot enabled detection of Escherichia coli using a cell-phone. Analyst. 2012;137(11):2541-4.
50. Coskun AF, Wong J, Khodadadi D, Nagi R, Tey A, Ozcan A. A personalized food allergen testing platform on a cellphone. Lab on a Chip. 2013;13(4):636-40.
51. Coskun AF, Nagi R, Sadeghi K, Phillips S, Ozcan A. Albumin testing in urine using a smart-phone. Lab on a Chip. 2013;13(21):4231-8.
52. Comina G, Suska A, Filippini D. Autonomous chemical sensing interface for universal cell phone readout. Angewandte Chemie International Edition. 2015;54(30):8708-12.
53. Chun HJ, Park YM, Han YD, Jang YH, Yoon HC. based glucose biosensing system utilizing a smartphone as a signal reader. BioChip Journal. 2014;8(3):218-26.
54. Wu Y, Boonloed A, Sleszynski N, Koesdjojo M, Armstrong C, Bracha S, et al. Clinical chemistry measurements with commercially available test slides on a smartphone platform: Colorimetric determination of glucose and urea. Clinica Chimica Acta. 2015;448:133-8.
55. Yeo S-J, Choi K, Cuc BT, Hong NN, Bao DT, Ngoc NM, et al. Smartphone-based fluorescent diagnostic system for highly pathogenic H5N1 viruses. Theranostics. 2016;6(2):231.
56. Chin CD, Cheung YK, Laksanasopin T, Modena MM, Chin SY, Sridhara AA, et al. Mobile device for disease diagnosis and data tracking in resource-limited settings. Clinical chemistry. 2012:clinchem. 2012.199596.
57. Laksanasopin T, Guo TW, Nayak S, Sridhara AA, Xie S, Olowookere OO, et al. A smartphone dongle for diagnosis of infectious diseases at the point of care. Science translational medicine. 2015;7(273):273re1-re1.
58. Rodriguez-Manzano J, Karymov MA, Begolo S, Selck DA, Zhukov DV, Jue E, et al. Reading out single-molecule digital RNA and DNA isothermal amplification in nanoliter volumes with unmodified camera phones. ACS nano. 2016;10(3):3102-13.
59. Calabria D, Caliceti C, Zangheri M, Mirasoli M, Simoni P, Roda A. Smartphone–based enzymatic biosensor for oral fluid L-lactate detection in one minute using confined multilayer paper reflectometry. Biosensors and Bioelectronics. 2017;94:124-30.
60. Lipowicz M, Garcia A. Handheld device adapted to smartphone cameras for the measurement of sodium ion concentrations at saliva-relevant levels via fluorescence. Bioengineering. 2015;2(2):122-38.
61. Oncescu V, Mancuso M, Erickson D. Cholesterol testing on a smartphone. Lab on a Chip. 2014;14(4):759-63.
62. Ogirala T, Eapen A, Salvante KG, Rapaport T, Nepomnaschy PA, Parameswaran AM. Smartphone-based colorimetric ELISA implementation for determination of women's reproductive steroid hormone profiles. Medical & biological engineering & computing. 2017;55(10):1735-41.
63. Lukac R, Plataniotis KN, editors. A normalized model for color-ratio based demosaicking schemes. Image Processing, 2004 ICIP'04 2004 International Conference on; 2004: IEEE.
64. Schiesser T. Know Your Smartphone: A Guide to Camera Hardware: TechSpot, Inc.; 2014 [Available from: https://www.techspot.com/guides/850-smartphone-camera-hardware/.

65. Cardinal D. How DxOMark Mobile scores smartphone cameras: DxOMark Image Labs; 2018 [Available from: https://www.dxomark.com/dxomark-mobile-scores-smartphone-cameras/.
66. Encyclopedia WTF. Exmor: Wikimedia Foundation Inc.; 2018 [Available from: https://en.wikipedia.org/wiki/Exmor.
67. Standardization IOf. Photography — Digital still cameras — Determination of exposure index, ISO speed ratings, standard output sensitivity, and recommended exposure index. ISO 12232:2006; 2008.
68. McHugh S. Understanding White Balance: Cambridge in Colour; [cited 2018 26/06]. Available from: http://www.cambridgeincolour.com/tutorials/white-balance.htm.
69. Wallace GK. The JPEG still picture compression standard. IEEE transactions on consumer electronics. 1992;38(1):xviii-xxxiv.
70. Smith T, Guild J. The CIE colorimetric standards and their use. Transactions of the optical society. 1931;33(3):73.
71. Encyclopedia WTF. sRGB: Wikimedia Foundation Inc.; 2018 [Available from: https://en.wikipedia.org/wiki/SRGB.
72. Konica Minolta Sensing Americas I. Understanding Standard Illuminants in Color Measurement: Konica Minolta Sensing Americas, Inc.; 2018 [Available from: https://sensing.konicaminolta.us/blog/understanding-standard-illuminants-in-color-measurement/.
73. Lorenzen A, Kennedy SW. A fluorescence-based protein assay for use with a microplate reader. Analytical biochemistry. 1993;214(1):346-8.
74. Wang H, Joseph JA. Quantifying cellular oxidative stress by dichlorofluorescein assay using microplate reader1. Free Radical Biology and Medicine. 1999;27(5-6):612-6.
75. Lequin RM. Enzyme immunoassay (EIA)/enzyme-linked immunosorbent assay (ELISA). Clinical chemistry. 2005;51(12):2415-8.
76. Spencer RD, Toledo FB, Williams BT, Yoss NL. Design, construction, and two applications for an automated flow—cell polarization fluorometer with digital read out: Enzyme-inhibitor (antitrypsin) assay and antigen—antibody (insulin—insulin antiserum) assay. Clinical chemistry. 1973;19(8):838-44.
77. Pappenhagen AR, Koppel J, Olwin JH. Use of fluorescein-labeled fibrin for the determination of fibrinolytic activity. The Journal of laboratory and clinical medicine. 1962;59(6):1039-46.
78. Pick U, Karlish SJ. Indications for an oligomeric structure and for conformational changes in sarcoplasmic reticulum Ca2+-ATPase labelled selectively with fluorescein. Biochimica et Biophysica Acta (BBA)-Protein Structure. 1980;626(1):255-61.
79. Suzuki N, Kojima H, Urano Y, Kikuchi K, Hirata Y, Nagano T. Orthogonality of calcium concentration and ability of 4, 5-diaminofluorescein to detect NO. Journal of Biological Chemistry. 2002;277(1):47-9.
80. Devaraj S, Du Clos TW, Jialal I. Binding and internalization of C-reactive protein by Fcgamma receptors on human aortic endothelial cells mediates biological effects. Arteriosclerosis, thrombosis, and vascular biology. 2005;25(7):1359-63.
81. Klug B. Nexus 7 (2013) - Mini Review: Purch; 2013 [Available from: https://www.anandtech.com/show/7176/nexus-7-2013-mini-review.
82. Lim JS. Two-dimensional signal and image processing. Englewood Cliffs, NJ, Prentice Hall, 1990, 710 p. 1990.
83. Crossle J, Elliot R, Smith P. Dried-blood spot screening for cystic fibrosis in the newborn. The Lancet. 1979;313(8114):472-4.
84. Heeley A, Heeley M, King D, Kuzemko J, Walsh M. Screening for cystic fibrosis by dried blood spot trypsin assay. Archives of Disease in Childhood. 1982;57(1):18.

85. Maeda S, Sawai T, Uzuki M, Takahashi Y, Omoto H, Seki M, et al. Determination of interstitial collagenase (MMP-1) in patients with rheumatoid arthritis. Annals of the rheumatic diseases. 1995;54(12):970.
86. Clark I, Powell L, Ramsey S, Hazleman B, Cawston T. The measurement of collagenase, tissue inhibitor of metalloproteinases (timp), and collagenase—timp complex in synovial fluids from patients with osteoarthritis and rheumatoid arthritis. Arthritis & Rheumatism: Official Journal of the American College of Rheumatology. 1993;36(3):372-9.
87. Ming K, Kim J, Biondi MJ, Syed A, Chen K, Lam A, et al. Integrated quantum dot barcode smartphone optical device for wireless multiplexed diagnosis of infected patients. Acs Nano. 2015;9(3):3060-74.
88. Edwards AD, Reis NM, Slater NK, Mackley MR. A simple device for multiplex ELISA made from melt-extruded plastic microcapillary film. Lab on a Chip. 2011;11(24):4267-73.
89. Yu H, Tan Y, Cunningham BT. Smartphone fluorescence spectroscopy. Analytical chemistry. 2014;86(17):8805-13.
90. Vashist SK, van Oordt T, Schneider EM, Zengerle R, von Stetten F, Luong JH. A smartphone-based colorimetric reader for bioanalytical applications using the screen-based bottom illumination provided by gadgets. Biosensors and Bioelectronics. 2015;67:248-55.
91. Sly PD, Brennan S, Gangell C, de Klerk N, Murray C, Mott L, et al. Lung Disease at Diagnosis in Infants with Cystic Fibrosis Detected by Newborn Screening. Am J Resp Crit Care. 2009;180(2):146-52.
92. Stick SM, Brennan S, Murray C, Douglas T, von Ungern-Sternberg BS, Garratt LW, et al. Bronchiectasis in infants and preschool children diagnosed with cystic fibrosis after newborn screening. J Pediatr. 2009;155(5):623-8 e1.
93. Berry LJ, Sheil B, Garratt L, Sly PD, Arest CF. Stability of interleukin 8 and neutrophil elastase in bronchoalveolar lavage fluid following long-term storage. J Cyst Fibros. 2010;9(5):346-50.
94. TAPPI. Diffuse brightness of paper, paperboard and pulp (d/0) (ultraviolet level D65), Test Method T 579 om-18. TAPPI; 2012.
95. Qian X, Chen J, An X. Polypyrrole-coated conductive paper prepared by vapour-phase deposition method. Appita Journal: Journal of the Technical Association of the Australian and New Zealand Pulp and Paper Industry. 2010;63(2):102.
96. Pascale D. RGB coordinates of the Macbeth ColorChecker. The BabelColor Company. 2006:1-16.
97. POUSHTER J. Smartphones are common in advanced economies, but digital divides remain. Fact Tank [Internet]. 2017. Available from: http://www.pewresearch.org/fact-tank/2017/04/21/smartphones-are-common-in-advanced-economies-but-digital-divides-remain/.
98. Smith AR. Color gamut transform pairs. SIGGRAPH Comput Graph. 1978;12(3):12-9.
99. SharkD. HSV color solid cylinder. In: cylinder.png Hcs, editor. Wikimedia commons: Wikimedia commons. p. The HSV color model mapped to a cylinder. POV-Ray source is available from the POV-Ray Object Collection.
100. BT RI-R. Studio encoding parameters of digital television for standard 4: 3 and wide-screen 16: 9 aspect ratios. 1995.
101. Tomasi C, Manduchi R, editors. Bilateral filtering for gray and color images. Computer Vision, 1998 Sixth International Conference on; 1998: IEEE.
102. Otsu N. A threshold selection method from gray-level histograms. IEEE transactions on systems, man, and cybernetics. 1979;9(1):62-6.
103. Gangell C, Gard S, Douglas T, Park J, de Klerk N, Keil T, et al. Inflammatory responses to individual microorganisms in the lungs of children with cystic fibrosis. Clinical infectious diseases. 2011;53(5):425-32.

104. Garratt LW, Sutanto EN, Ling K-M, Looi K, Iosifidis T, Martinovich KM, et al. Alpha-1 antitrypsin mitigates the inhibition of airway epithelial cell repair by neutrophil elastase. American journal of respiratory cell and molecular biology. 2016;54(3):341-9.
105. Mayer-Hamblett N, Aitken ML, Accurso FJ, Kronmal RA, Konstan MW, Burns JL, et al. Association between pulmonary function and sputum biomarkers in cystic fibrosis. Am J Resp Crit Care. 2007;175(8):822-8.
106. Effros RM, Biller J, Foss B, Hoagland K, Dunning MB, Castillo D, et al. A simple method for estimating respiratory solute dilution in exhaled breath condensates. Am J Resp Crit Care. 2003;168(12):1500-5.
107. Newman JD, Turner AP. Home blood glucose biosensors: a commercial perspective. Biosensors and bioelectronics. 2005;20(12):2435-53.
108. Derda R, Gitaka J, Klapperich CM, Mace CR, Kumar AA, Lieberman M, et al. Enabling the development and deployment of next generation point-of-care diagnostics. PLoS neglected tropical diseases. 2015;9(5):e0003676.
109. Jani IV. How point-of-care testing could drive innovation in global health. The New England journal of medicine. 2013;368(24):2319.
110. Carter RI, Ungurs MJ, Mumford RA, Stockley RA. Aα-Val360: a marker of neutrophil elastase and COPD disease activity. European Respiratory Journal. 2013;41(1):31-8.
111. López-Campos JL, Tan W, Soriano JB. Global burden of COPD. Respirology. 2016;21(1):14-23.
112. Chen HM, Chiang CP, You C, Hsiao TC, Wang CY. Time-resolved autofluorescence spectroscopy for classifying normal and premalignant oral tissues. Lasers in Surgery and Medicine: The Official Journal of the American Society for Laser Medicine and Surgery. 2005;37(1):37-45.
113. Bacskai BJ, Skoch J, Hickey GA, Allen R, Hyman BT. Fluorescence resonance energy transfer determinations using multiphoton fluorescence lifetime imaging microscopy to characterize amyloid-beta plaques. Journal of biomedical optics. 2003;8(3):368-76.
114. Oida T, Sako Y, Kusumi A. Fluorescence lifetime imaging microscopy (flimscopy). Methodology development and application to studies of endosome fusion in single cells. Biophysical journal. 1993;64(3):676-85.
115. Bastiaens P, Jovin TM. Microspectroscopic imaging tracks the intracellular processing of a signal transduction protein: fluorescent-labeled protein kinase C beta I. Proceedings of the National Academy of Sciences. 1996;93(16):8407-12.
116. Gadella T, Jovin TM. Oligomerization of epidermal growth factor receptors on A431 cells studied by time-resolved fluorescence imaging microscopy. A stereochemical model for tyrosine kinase receptor activation. The Journal of cell biology. 1995;129(6):1543-58.
117. Wouters FS, Bastiaens PI. Fluorescence lifetime imaging of receptor tyrosine kinase activity in cells. Current Biology. 1999;9(19):1127-S1.
118. Verveer PJ, Wouters FS, Reynolds AR, Bastiaens PI. Quantitative imaging of lateral ErbB1 receptor signal propagation in the plasma membrane. Science. 2000;290(5496):1567-70.
119. Lidke D, Nagy P, Barisas B, Heintzmann R, Post JN, Lidke K, et al. Imaging molecular interactions in cells by dynamic and static fluorescence anisotropy (rFLIM and emFRET). Portland Press Limited; 2003.
120. Larijani B, Hume AN, Tarafder AK, Seabra MC. Multiple factors contribute to inefficient prenylation of Rab27a in Rab prenylation diseases. Journal of Biological Chemistry. 2003.
121. Murata S-i, Herman P, Lakowicz JR. Texture analysis of fluorescence lifetime images of AT-and GC-rich regions in nuclei. Journal of Histochemistry & Cytochemistry. 2001;49(11):1443-51.

122. Cremazy FG, Manders EM, Bastiaens PI, Kramer G, Hager GL, Van Munster EB, et al. Imaging in situ protein–DNA interactions in the cell nucleus using FRET–FLIM. Experimental cell research. 2005;309(2):390-6.
123. Verveer PJ, Squire A, Bastiaens PI. Improved spatial discrimination of protein reaction states in cells by global analysis and deconvolution of fluorescence lifetime imaging microscopy data. Journal of microscopy. 2001;202(3):451-6.
124. Clayton A, Hanley Q, Verveer P. Graphical representation and multicomponent analysis of single-frequency fluorescence lifetime imaging microscopy data. Journal of microscopy. 2004;213(1):1-5.
125. Vermeer J, Van Munster E, Vischer N, Gadella Jr T. Probing plasma membrane microdomains in cowpea protoplasts using lipidated GFP-fusion proteins and multimode FRET microscopy. Journal of microscopy. 2004;214(2):190-200.
126. Lakowicz J, Szmacinski H, Nowaczyk K, Lederer W, Kirby M, Johnson M. Fluorescence lifetime imaging of intracellular calcium in COS cells using Quin-2. Cell Calcium. 1994;15(1):7-27.
127. Periasamy A, Wodnicki P, Wang XF, Kwon S, Gordon GW, Herman B. Time-resolved fluorescence lifetime imaging microscopy using a picosecond pulsed tunable dye laser system. Review of scientific instruments. 1996;67(10):3722-31.
128. Hanson KM, Behne MJ, Barry NP, Mauro TM, Gratton E, Clegg RM. Two-photon fluorescence lifetime imaging of the skin stratum corneum pH gradient. Biophysical journal. 2002;83(3):1682-90.
129. Vroom JM, De Grauw KJ, Gerritsen HC, Bradshaw DJ, Marsh PD, Watson GK, et al. Depth penetration and detection of pH gradients in biofilms by two-photon excitation microscopy. Applied and environmental microbiology. 1999;65(8):3502-11.
130. Gerritsen HC, Sanders R, Draaijer A, Ince C, Levine Y. Fluorescence lifetime imaging of oxygen in living cells. Journal of fluorescence. 1997;7(1):11-5.
131. Dong C, So P, French T, Gratton E. Fluorescence lifetime imaging by asynchronous pump-probe microscopy. Biophysical journal. 1995;69(6):2234-42.
132. van Zandvoort MA, de Grauw CJ, Gerritsen HC, Broers JL, oude Egbrink MG, Ramaekers FC, et al. Discrimination of DNA and RNA in cells by a vital fluorescent probe: lifetime imaging of SYTO13 in healthy and apoptotic cells. Cytometry. 2002;47(4):226-35.
133. Kohl M, Neukammer J, Sukowski U, Rinneberg H, Wöhrle D, Sinn H-J, et al. Delayed observation of laser-induced fluorescence for imaging of tumors. Applied Physics B. 1993;56(3):131-8.
134. Wagnieres G, Mizeret J, Studzinski A, Van den Bergh H. Frequency-domain fluorescence lifetime imaging for endoscopic clinical cancer photodetection: apparatus design and preliminary results. Journal of Fluorescence. 1997;7(1):75-83.
135. Tadrous PJ, Siegel J, French PM, Shousha S, Lalani EN, Stamp GW. Fluorescence lifetime imaging of unstained tissues: early results in human breast cancer. The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland. 2003;199(3):309-17.
136. Blacker TS, Mann ZF, Gale JE, Ziegler M, Bain AJ, Szabadkai G, et al. Separating NADH and NADPH fluorescence in live cells and tissues using FLIM. Nature communications. 2014;5:3936.
137. McGinty J, Galletly NP, Dunsby C, Munro I, Elson DS, Requejo-Isidro J, et al. Wide-field fluorescence lifetime imaging of cancer. Biomedical optics express. 2010;1(2):627-40.
138. Prochazka I, Hamal K, Sopko B. Recent achievements in single photon detectors and their applications. Journal of Modern Optics. 2004;51(9-10):1289-313.
139. Zappa F, Tisa S, Gulinatti A, Gallivanoni A, Cova S, editors. Monolithic CMOS detector module for photon counting and picosecond timing. Solid-State Device Research Conference, 2004 ESSDERC 2004 Proceeding of the 34th European; 2004: IEEE.

140. Villa F, Bronzi D, Zou Y, Scarcella C, Boso G, Tisa S, et al. CMOS SPADs with up to 500 μm diameter and 55% detection efficiency at 420 nm. Journal of Modern Optics. 2014;61(2):102-15.
141. Mandai S, Fishburn MW, Maruyama Y, Charbon E. A wide spectral range single-photon avalanche diode fabricated in an advanced 180 nm CMOS technology. Optics express. 2012;20(6):5849-57.
142. Palubiak D, El-Desouki MM, Marinov O, Deen MJ, Fang Q. High-speed, single-photon avalanche-photodiode imager for biomedical applications. IEEE Sensors Journal. 2011;11(10):2401-12.
143. Webster EA, Richardson JA, Grant LA, Renshaw D, Henderson RK. A single-photon avalanche diode in 90-nm CMOS imaging technology with 44% photon detection efficiency at 690 nm. IEEE Electron Device Lett. 2012;33(5):694-6.
144. Zappa F, Tisa S, Tosi A, Cova S. Principles and features of single-photon avalanche diode arrays. Sensors and Actuators A: Physical. 2007;140(1):103-12.
145. Cova S, Lacaita A, Ghioni M, Ripamonti G, Louis T. 20-ps timing resolution with single-photon avalanche diodes. Review of scientific instruments. 1989;60(6):1104-10.
146. Colyer RA, Scalia G, Rech I, Gulinatti A, Ghioni M, Cova S, et al. High-throughput FCS using an LCOS spatial light modulator and an 8× 1 SPAD array. Biomedical optics express. 2010;1(5):1408-31.
147. Guerrieri F, Tisa S, Tosi A, Zappa F. Two-dimensional SPAD imaging camera for photon counting. IEEE Photonics Journal. 2010;2(5):759-74.
148. Niclass C, Rochas A, Besse P-A, Charbon E. Design and characterization of a CMOS 3-D image sensor based on single photon avalanche diodes. IEEE Journal of Solid-State Circuits. 2005;40(9):1847-54.
149. Niclass C, Sergio M, Charbon E, editors. A single photon avalanche diode array fabricated in 0.35-μm CMOS and based on an event-driven readout for TCSPC experiments. Advanced Photon Counting Techniques; 2006: International Society for Optics and Photonics.
150. Maruyama Y, Blacksberg J, Charbon E. A 1024$\,\times\, $8, 700-ps Time-Gated SPAD Line Sensor for Planetary Surface Exploration With Laser Raman Spectroscopy and LIBS. IEEE Journal of Solid-State Circuits. 2014;49(1):179-89.
151. Maruyama Y, Charbon E, editors. An all-digital, time-gated 128X128 spad array for on-chip, filter-less fluorescence detection. Solid-State Sensors, Actuators and Microsystems Conference (TRANSDUCERS), 2011 16th International; 2011: IEEE.
152. Veerappan C, Bruschini C, Charbon E, editors. Sensor network architecture for a fully digital and scalable SPAD based PET system. Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2012 IEEE; 2012: IEEE.
153. Tétrault M-A, Lamy ED, Boisvert A, Thibaudeau C, Kanoun M, Dubois F, et al. Real-time discrete SPAD array readout architecture for time of flight PET. IEEE Transactions on Nuclear Science. 2015;62(3):1077-82.
154. Albota MA, Heinrichs RM, Kocher DG, Fouche DG, Player BE, O'Brien ME, et al. Three-dimensional imaging laser radar with a photon-counting avalanche photodiode array and microchip laser. Applied Optics. 2002;41(36):7671-8.
155. Bronzi D, Villa F, Tisa S, Tosi A, Zappa F, Durini D, et al. 100 000 frames/s 64× 32 single-photon detector array for 2-D imaging and 3-D ranging. IEEE journal of selected topics in quantum electronics. 2014;20(6):354-63.
156. Pancheri L, Stoppa D, editors. A SPAD-based pixel linear array for high-speed time-gated fluorescence lifetime imaging. ESSCIRC, 2009 ESSCIRC'09 Proceedings of; 2009: IEEE.
157. Li D-U, Arlt J, Richardson J, Walker R, Buts A, Stoppa D, et al. Real-time fluorescence lifetime imaging system with a 32× 32 0.13 μm CMOS low dark-count single-photon avalanche diode array. Optics express. 2010;18(10):10257-69.

158. Li D-U, Rae BR, Andrews R, Arlt J, Henderson RK. Hardware implementation algorithm and error analysis of high-speed fluorescence lifetime sensing systems using center-of-mass method. Journal of biomedical optics. 2010;15(1):017006.
159. Burri S, Maruyama Y, Michalet X, Regazzoni F, Bruschini C, Charbon E. Architecture and applications of a high resolution gated SPAD image sensor. Optics express. 2014;22(14):17573-89.
160. Tata DB, Foresti M, Cordero J, Tomashefsky P, Alfano M, Alfano R. Fluorescence polarization spectroscopy and time-resolved fluorescence kinetics of native cancerous and normal rat kidney tissues. Biophysical journal. 1986;50(3):463.
161. Dainty J, Shaw R. Image Science: Principles. Analysis and Evaluation of Photographic-Type Imaging Processes (London: Academic). 1974.
162. Magde D, Rojas GE, Seybold PG. Solvent dependence of the fluorescence lifetimes of xanthene dyes. Photochemistry and Photobiology. 1999;70(5):737-44.
163. Salmon JM, Kohen E, VIALLET P, Hirschberg JG, Wouters AW, Kohen C, et al. Microspectrofluorometric approach to the study of free/bound NAD (P) H ratio as metabolic indicator in various cell types. Photochemistry and photobiology. 1982;36(5):585-93.
164. Butte PV, Qiyin F, Javier JA, Yong WH, Pikul BK, Black KL, et al. Intraoperative delineation of primary brain tumors using time-resolved fluorescence spectroscopy. Journal of biomedical optics. 2010;15(2):027008.
165. Patterson GH, Knobel SM, Arkhammar P, Thastrup O, Piston DW. Separation of the glucose-stimulated cytoplasmic and mitochondrial NAD (P) H responses in pancreatic islet β cells. Proceedings of the National Academy of Sciences. 2000;97(10):5203-7.
166. Lakowicz JR, Szmacinski H, Nowaczyk K, Johnson ML. Fluorescence lifetime imaging of free and protein-bound NADH. Proceedings of the National Academy of Sciences. 1992;89(4):1271-5.
167. Li D-U, Bonnist E, Renshaw D, Henderson R. On-chip, time-correlated, fluorescence lifetime extraction algorithms and error analysis. JOSA A. 2008;25(5):1190-8.
168. Li D-U, Walker R, Richardson J, Rae B, Buts A, Renshaw D, et al., editors. FPGA implementation of a video-rate fluorescence lifetime imaging system with a 32× 32 CMOS single-photon avalanche diode array. Circuits and Systems, 2009 ISCAS 2009 IEEE International Symposium on; 2009: IEEE.
169. Foundation TF. Lifetime Data of Selected Fluorophores. ISS, Inc.; 2018.
170. Li DD-U, Ameer-Beg S, Arlt J, Tyndall D, Walker R, Matthews DR, et al. Time-domain fluorescence lifetime imaging techniques suitable for solid-state imaging sensor arrays. Sensors. 2012;12(5):5650-69.
171. Sun YH, Hatami N, Yee M, Phipps JE, Elson DS, Gorin F, et al. Fluorescence lifetime imaging microscopy for brain tumor image-guided surgery. Journal of biomedical optics. 2010;15(5):056022.
172. Yu Q, Heikal AA. Two-photon autofluorescence dynamics imaging reveals sensitivity of intracellular NADH concentration and conformation to cell physiology at the single-cell level. Journal of Photochemistry and Photobiology B: Biology. 2009;95(1):46-57.
173. Balasaraswathy P, Kumar U, Srinivas C, Nair S. UVA and UVB in sunlight, optimal utilization of UV rays in sunlight for phototherapy. Indian Journal of Dermatology, Venereology, and Leprology. 2002;68(4):198.