Imaging Trends and the Future of
High Resolution Satellite Remote Sensing

Jonathan Dyer[a], Paul Boerner[b], Mark Robertson[a]

[a]Google, Inc. 1600 Amphitheatre Parkway, Mountain View, CA 94043
[b]Planet, 346 9th Street San Francisco, CA 94103

Abstract. Many of the technology and investment trends that ushered in the so-called "SmallSat Revolution" will continue or accelerate but there is a paucity of analysis available in the open literature regarding the system design implications of these trends. While these systems may not replace more traditional remote sensing systems in the near term, in many cases they are highly complementary such that interest and investment will continue to grow.

The authors identify the trends in imaging and micro-electronics driven by investments in such applications as smart phones and 4k TV and analyze the impact the associated technologies can have in new remote sensing spacecraft. The commercially-driven technologies present large opportunities for cost reduction and performance improvements but require new system architectures and approaches.

Keywords: Remote Sensing, Satellite, Imaging, Image resolution, CMOS Integrated Circuits, Systems Engineering..

1 Introduction

The last 15 years have seen a massive upheaval in digital imaging culminating in almost total obsolescence of film-based imaging. Initially driven by the pro-sumer DSLR and pocket camera market, the trend has continued and accelerated with the introduction and refinement of high quality cameras in to smart phones and the digitization of the TV and movie industries.

Satellite remote sensing systems underwent a similar revolution roughly 40 years earlier as film-based systems were phased out in favor of CCD-based electronic line scanners. While the specific demanding requirements of satellite remote sensing drove much of the early technology investment and development in digital electronic imaging, far more investment is now being applied in the commercial arena than in further refinement of now mature CCD-based line scan sensors.

We believe that the future of the satellite remote-sensing industry will be in leveraging the massive investment in commercial digital and computational imaging technologies rather than continued large-scale investment in standalone technologies for these systems and that this shift presents both opportunities and challenges to new systems.

2 Trends

In order to motivate the subsequent discussion, we identify and discuss the "SmallSat" philosophical trend and two technology trends uniquely enabling it in remote sensing systems.

2.1 Trend to Smaller Spacecraft

There is a large and growing interest in what can be accomplished with much smaller spacecraft. This trend is driven by

- Reduced appetite for the large and ballooning costs of traditional systems

- Recognition of the mission level advantages offered by constellations of satellites (homogeneous or with sensor disaggregation) enabled by lower cost systems

- Demonstration of high-performance remote sensing SmallSats developed at low cost by small, commercially funded teams

There is significant evidence indicating that mission cost is non-linearly correlated with spacecraft size and complexity. Indeed[?] has shown with high correlation that

$$C \sim m^{1.261},$$

implying cost C is super-linearly impacted by mass, m. Because for spacecraft systems $m \sim \chi^n$ where χ is a characteristic spacecraft dimension and n is typically between 2 and 3, we can also say

$$C \sim \chi^{3+} \tag{1}$$

Similar arguments can be derived from other cost scaling observations such as Meinel's Law[?] for telescopes which states

$$C \sim D_{ap}^{2.7}, \tag{2}$$

where D_{ap} is aperture diameter.

This polynomial relationship between size and cost indicates that there are significant savings available in shrinking system size down subject, of course to fundamental physical limits such as diffraction-limited resolution.

2.2 Sensor Technology

Historically, space-based remote sensing systems have utilized CCD-based Time Delayed Integration (TDI) line-scan sensors. Modern TDI sensors are a mature technology with several very desirable characteristics for high resolution remote sensing:

- Mechanism-free relative motion compensation through charge transfer

- High quantum efficiency and fill fractions

- Low noise

- Relatively high readout line-rates compatible with needs of high resolution systems

- Reasonable radiation tolerance

However the unique requirements associated with space-based sensors implies that the sensors are niche, low volume and typically very expensive to develop, integrate and use.

Meanwhile the large commercial investment in imaging sensors driven by the consumer photography/cinema, machine vision and smart phone industries has resulted in incredibly high performance image sensors available at high volume and low cost. The big trends that have developed due to this investment are:

- Movement from CCD to CMOS detector technology

- Drive to smaller pixels to enable compact cameras[?]

- Significant increase in detector read-out rates at low noise

CCD-based sensors still hold some performance advantages over CMOS sensors, but the capability gap has narrowed almost entirely. The manufacturing and read-out rate advantages offered by CMOS have led to its wholesale adoption in the consumer / commercial world.

When developing new remote sensing systems, the capabilities of modern commercial detectors promise the potential of smaller, less expensive systems with similar or better imaging performance to traditional TDI CCD-based systems. However to realize this promise, the imaging system designer must consider the unique capabilities and shortcomings of the technology.

2.3 Non-Imaging Commercial Electronic Trends

In addition to the rapid advances in commercial imaging technology, Moore's law has continued to drive massive increase in microelectronic integration density with the associated performance, cost and device size improvements.

Indeed in recent years much of the emphasis has been in large-scale functional integration into single die "system-on-chip" (SoC) devices that operate at very high performance/watt for the consumer devices industry. Such devices are uniquely enabling for small spacecraft in that, just as with consumer devices, they enable massive functionality in a small, low power package and limit the amount of functional engineering "glue" required, shortening design cycles. Stated more directly, the computational power most people have in their pockets today was unavailable in large, complex spacecraft ten years ago.

Similar arguments can be made for:

- High density solid-state storage (Flash memory)

- Dedicated signal- and image-processing hardware for compression or computational imaging

- Digital modulation and radio technologies

- Inertial sensors

Indeed, one can look at many of our personal consumer electronic devices and see the applicability to spacecraft in virtually every technology. And these technologies have investment rates orders-of-magnitude higher than dedicated aerospace technologies.

Historically there has been resistance to the utilization of commercial electronics technologies for space applications due to reliability concerns, both intrinsic and due to the space environment. Again, several developments make such resistance less defensible:

- Demonstration of high uptime and reliable operation of COTS-based avionics in LEO.[?]

- High demonstrated statistical reliability of commercially available components qualified to operate over thermal ranges much more severe than typical in-space environments, e.g. auto, industrial.

- Substantial improvement in CMOS latch-up susceptibility

- Pervasive integration of hardware single-event protection driven by sea-level single-event susceptibility of modern high integration devices

For small systems these trends are uniquely empowering and are more easily adoptable with lower cost (and hence less risk-averse) systems.

Furthermore the ability to incorporate high performance compute, storage and communications in space- and power-constrained systems invalidates some of the long-held constraints on other spacecraft systems such as the ability to perform near-real-time on-board image processing.

3 Remote Sensing Metrics

In order to compare different remote sensing systems, we need quantitative metrics. The metrics we will introduce are fundamentally related to the data produced by such systems and are in two categories - quality and volume. In a value-based assessment, quality and volume would form the numerator of a global value metric with cost in the denominator.

3.1 Image Quality

Image quality is a complex, multi-faceted subject and while there is no perfect model or metric, substantial work has gone into quantifying the image quality of high resolution remote sensing systems. Before jumping into integrated image quality models, we will define the geometric and radiometric components of image quality.
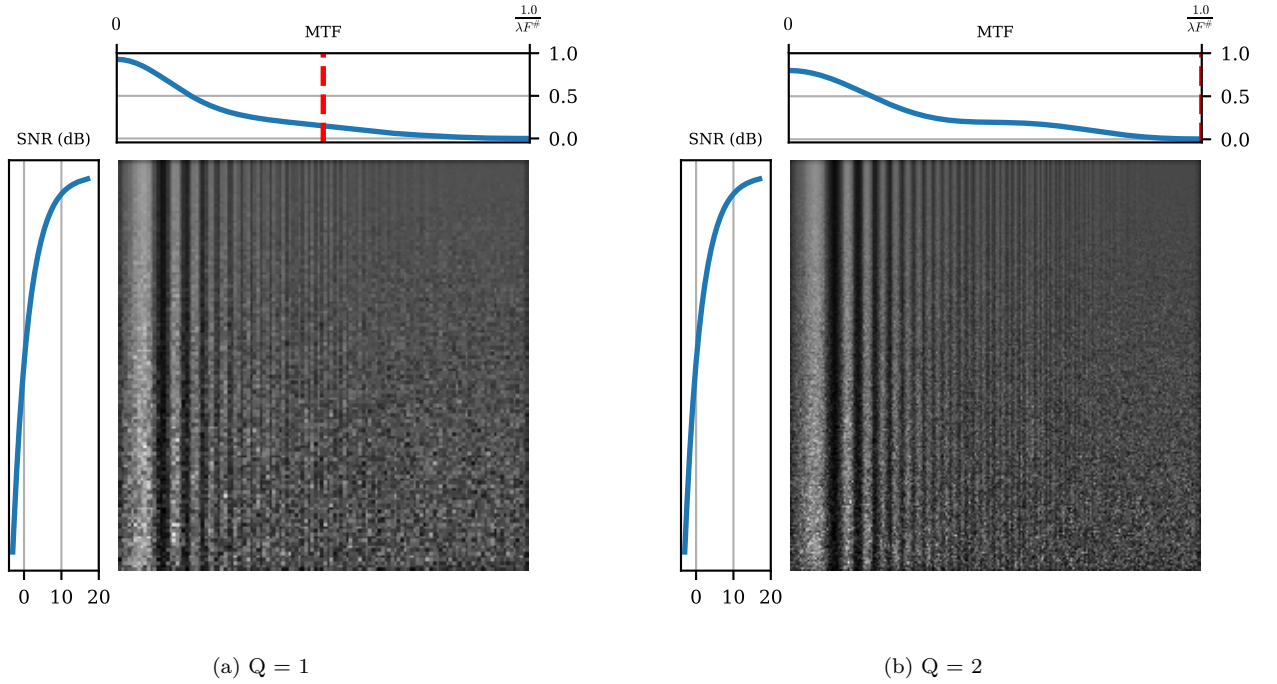
(a) Q = 1   (b) Q = 2

Fig 1: Visual resolvability vs SNR and MTF for two different values of Q. Note that dashed red line indicates Nyquist criteria.

3.1.1 Resolution, NIIRS and GIQE

Resolution is a key component of image quality but is a surprisingly difficult-to-define metric. Often Ground Sample Distance, GSD is used as a proxy for resolution under the assumption of near-critical sampling. Alternatively, Ground Resolution Distance, GRD, is used where GRD is a form of the diffraction-limited optical resolution such that given by the Rayleigh Criterion:

$$GRD = 1.22 \frac{\lambda R}{D_{ap}}, \tag{3}$$

where R is range to target (spacecraft altitude if Nadir-pointed), D_{ap} is the aperture diameter and λ is the weighted center-wavelength of the sampled spectral band.

Comparing image quality purely on GSD or GRD becomes problematic for systems with varying Q and/or SNR (see Section 8.1) because depending on the sampling or SNR regimes resolvability may first be constrained by sampling, diffraction or SNR. Figure ?? illustrates this graphically.

More holistic image quality formulations typically refer back to resolvability or the ability to distinguish objects or features of a given size in an image.

High resolution space-based systems are typically evaluated by quantitative "image interpretability" scales, the best known being the National Image Interpretability Rating Scale, or NIIRS.

NIIRS rating is described by specific scene intepretation standards, examples of which are shown in Table 1.[?]

NIIRS is a logarithmic scale traditionally evaluated by trained human image interpreters. However, much work has gone into developing semi-analytic regressions to predict NIIRS from basic image quality parameters such as sharpness, signal-to-noise ratio (SNR), etc. The Generalized Image Quality Equation (GIQE) is an example of such a function; we will use the 5th version, GIQE-5 which is defined in eq. (4).

$$NIIRS = 4.4 - 3.32 \log(GSD_m) + 3.32 \left[1 - e^{\frac{-5.308}{\Delta SNR}} \right] \log(RER_0)$$
$$- 0.402 \log(RER_0)^4 - 2.92/\Delta SNR - 0.069 N_{smear}, \tag{4}$$

4

NIIRS	GRD	Description
3	2.5m - 4.5m	Identify a road as divided or undivided Detect rows of automobiles in a parking lot
4	1.2m - 2.5m	Detect barriers/obstacles (barrels) on runways Distinguish between locomotives and railcars
5	0.75m - 1.2m	Identify individual lines painted on paved roads, parking lots Identify fallen utility poles
6	0.4m - 0.75m	Detect individuals, when not in a group Detect small road signs in an urban area

Table 1: Example NIIRS definitions

where GSD_m[1] is the ground sample distance in meters, ΔSNR is difference between SNR for a 15% and 8% target reflectance, RER_0 is the raw Relative Edge Response (without sharpening or MTFC) and N_{smear} is the number of pixels of linear smear.

GIQE-5 is fairly well validated again human interpreters[?] and nicely captures the trade-off between various imaging system parameters such as SNR and GSD. Note that GIQE-5 is logarithmic in resolution such that a halving of GSD (proportional to resolution when holding other parameters constant) results in a NIIRS improvement of 1.

Based on this definition, one can use GIQE-5 to define an effective ground resolvability distance GRD_{eff}

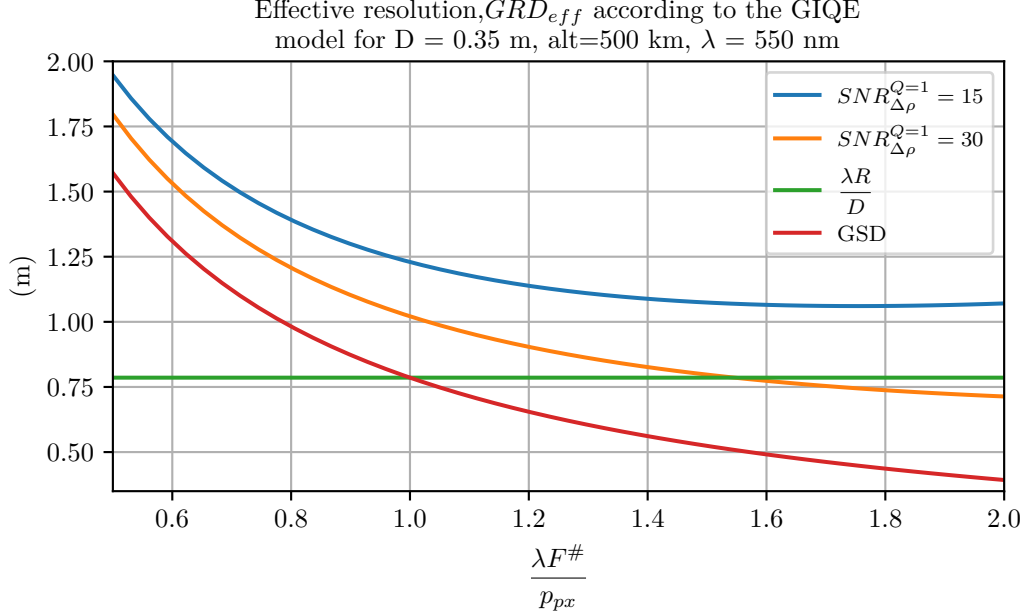$$GRD_{eff}(m) = 10^{\frac{4.4-NIIRS}{3.32}}. \tag{5}$$

Fig 2: Relationship of GRD and GSD as resolution metrics as a function of Q

Notice in Figure 2 that GRD asymptotes at $Q = 2$ due to diffraction but that GSD continues decreasing monotonically. This is because optical systems do not pass spatial frequency beyond the diffraction limit which is critically sampled at $Q = 2$.

---

[1]GSD units must be paid close attention to – many references list GIQE using units of inches for GSD. In this case, the first parameter in the equation is 9.7 rather than 4.4 because $4.4 - 3.32 \log\left(39.4\frac{in}{m}\right) \approx 9.7$

3.1.2 Dynamic Range (DR) and Signal-to-Noise Ratio (SNR)

Dynamic range and signal-to-noise ratio are closely related by subtly different measures of radiometric performance.

Dynamic range is defined as the ratio of maximum-to-minimum signal level that the system can capture or represent in an image

$$DR = \frac{s_{max}}{s_{min}}.$$

The minimum signal level is the noise-floor of the system, $s_{min} = \sigma$ where σ represents all non-signal-correlated noise sources in the system including read noise, dark current, quantization, etc.

The maximum signal is the maximum number of photo-electrons the system can capture for one output image pixel and is limited by the well depth of a pixel, $s_{max} = N_{e^-}^{FWC}$ for single-exposure systems. Note that in digitally oversampled systems (covered later), the digital oversampling factor directly expands dynamic range by increasing the effective well depth such that $s_{max} = N_s N_{e^-}^{FWC}$ where N_s is the oversampling ratio

$$N_s = \text{floor}\left(\frac{GSD \times LR}{V_{gnd}}\right) \tag{6}$$

where LR is the time-rate that lines (rows) are read from the detector in lines/second.

Oversampling also affects s_{min} and it can be shown that, for equally exposed samples

$$DR = \sqrt{N_s}\frac{N_{e^-}^{FWC}}{\sigma_{uncor}}. \tag{7}$$

Dynamic range is a very important quality parameter for remote sensing systems because the dynamic range of terrestrial natural scenes can be quite high.

Fig 3: The first two images have relatively low dynamic range. The first is under-exposed and the second over-. Note the details lost in shadows in the first and highlights (parking lot) in the second when compared with the third image which has about 90dB of DR

For systems that digitally oversample with unequal exposures, the dynamic range can be even larger as is shown in Section 5.6. Signal-to-noise ratio, SNR, like dynamic range depends critically on both the detector well depth and its noise characteristics. However signal-to-noise ratio is not purely a system-defined metric but depends also on the scene being imaged and as well as exposure parameters. SNR is defined as

$$SNR = \frac{s}{\sigma_{noise}}.$$

The noise, σ_{noise} consists of the same sources discussed for DR but also the shot-noise associated with the signal, s. Because arrival rate of incoming photons obey Poisson statistics, it can be shown that $\sigma_{shot} = \sqrt{s}$ such that

$$SNR = \frac{s}{\sqrt{\sum \sigma_i^2 + s}}.$$

In order to state an SNR, the signal, s, must be constrained by other system parameters and some exposure. Typically SNR is specified with respect to a specific target reflectance or top-of-atmosphere (TOA) radiance and a maximum or saturation reflectance or TOA radiance.

For a perfectly exposed scene with

$$\alpha = \frac{L_{typ}}{L_{sat}}, \tag{8}$$

SNR can be expressed as

$$SNR_\alpha = \frac{\alpha N_{e^-}^{FWC}}{\sqrt{\sum \sigma_i^2 + \alpha N_{e^-}^{FWC}}}, \tag{9}$$

where α is the ratio of target radiance (or reflectance) to the saturation value and $\sum \sigma_i^2$ is the RSS of noises other than shot noise.

Eq. (9), like eq. (7) for DR, can be expanded to include multiple digital exposures such that

$$SNR_\alpha = \sqrt{N_s} \frac{\alpha N_{e^-}^{FWC}}{\sqrt{\sum \sigma_i^2 + \alpha N_{e^-}^{FWC}}}. \tag{10}$$

For most modern detectors, uncorrelated noise sources are small compared with well depth so we can approximate eq. (10) as

$$SNR_\alpha \approx \sqrt{N_s \alpha N_{e^-}^{FWC}}. \tag{11}$$

The high dynamic range in natural overhead scenes presents a challenge in achieving high SNR. The SNR for a target at L_{typ} will scale inversely with L_{sat} forcing the designer to trade highlight dynamic range against SNR in much of the scene. Indeed DR and SNR requirements, taken together, become significant drivers of sensor performance requirements.

As mentioned earlier and discussed further in Section 5.6, digital oversampling allows the collection of so-called High Dynamic Range (HDR) images. HDR enables a system to achieve higher SNR over the full dynamic range of a scene.

3.2 Data Intensity

Data intensity, I_D, is a metric that impacts a wide range of spacecraft subsystems from on-board storage and data handling to downlink bandwidth. It is defined as the number of digital bits captured, processed or transmitted per image product output pixel

$$I_D \equiv \frac{N_{bit}}{N_{px}^{L1b}}, \tag{12}$$

where we assume an output product ("L1b") to consist of the fused, corrected but un-rectified product. For a typical imaging pipeline that includes compression we must choose exactly where in the system N_{bit} is defined. We will use the post-compression image data because it impacts the most system interfaces (storage, downlink, processing) and is generally directly proportional to the raw pixel data rate through the compression ratio, C

$$C \equiv \frac{N_{bpp}^{raw}}{N_{bpp}^{comp.}}.$$

Note also that any resampling operations (including image reconstruction techniques such as super-resolution) impact I_D because for non-unity resampling ratios the number of output pixels is affected. Thus we include a resampling factor

$$\eta_{rs} = \frac{GSD_{L1b}^2}{GSD_{L0}^2},$$

such that

$$I_D = \eta_{rs} N_s N_{bpp}^{comp.}. \tag{13}$$

In traditional remote sensing systems, I_D is largely driven by the balance of image quality requirements (driving smaller quantization steps and less compression) and data storage and downlink constraints (driving larger quantization steps and more compression).

For systems with well-matched analog and digital (quantization) dynamic ranges

$$N_{bpp}^{raw} = \log_2(DR^{px}) = \log_2\left(\frac{N_{e^-}^{FWC}}{\sigma_{uncorr.}}\right). \tag{14}$$

Dynamic range requirements are typically 60-75 dB ($\sim 1000:1$ to $\sim 5000:1$) driving N_{bpp}^{raw} of 10-12 bits. And industry experience has shown that modern image compression (such as JPEG2000) can achieve near visually loss-less performance at compression ratios of between 3:1 and 6:1. Thus it is reasonable to expect I_D between 2 bit/px and 4 bit/px for non-oversampling systems.

As noted previously, digital oversampling allows for system dynamic range performance superior to the inherent dynamic range of the sensor. However, the cost is in higher data intensity as shown in (15).

$$I_D = N_s N_{bpp}^{comp.} \tag{15}$$

Given a required dynamic range, DR_{req}, we can combine Equation (14), (7) to show that

$$I_D^{raw} = \left[\frac{DR_{req} \times \sigma_{uncorr.}}{N_{e^-}^{FWC}}\right]^2 \log_2(DR_{req}). \tag{16}$$

Equation (16) is important because it captures the inverse relationship of pixel full well capacity and data intensity in a digitally over-sampled system - a key result in defining pixel performance metric in Section 5.2.

3.3 Collection Capacity

For a spacecraft, system collection capacity is constrained by a delicate balance between image collection capacity and the ability to get the collected data to the ground. The latter is dictated entirely by data intensity, I_D (defined in Section 3.2), and communications bandwidth.

Image collection capacity in turn depends on a number of parameters including time over land, ACS agility, and area collection rate,

$$ACR = V_{gnd} W_{ct}, \tag{17}$$

where W_{ct} is the swath width or cross-track field of view projected to the ground and V_{gnd} is the apparent relative velocity of the ground in the imaging instrument's frame.

We will define a collection duty cycle metric that captures all of the impacts on collection capacity outside of the system ACR as

$$\beta = \frac{t_{collect}}{t_{per}}, \tag{18}$$

where t_{per} is the spacecraft orbital period and $t_{collect}$ is the collection time per orbit. Note that β is itself dependent on V_{gnd} in addition to a wide variety of targeting-related parameters such as time per orbit over land and in the sun, average target cross-track spacing (for targeted spacecraft), attitude control bandwidth, etc.

For continuous daytime collection over all land excluding Antartica, an absolute upper bound on β becomes ~ 0.12. For remote sensing systems constrained by slew as well as sun and time-over-land, typically
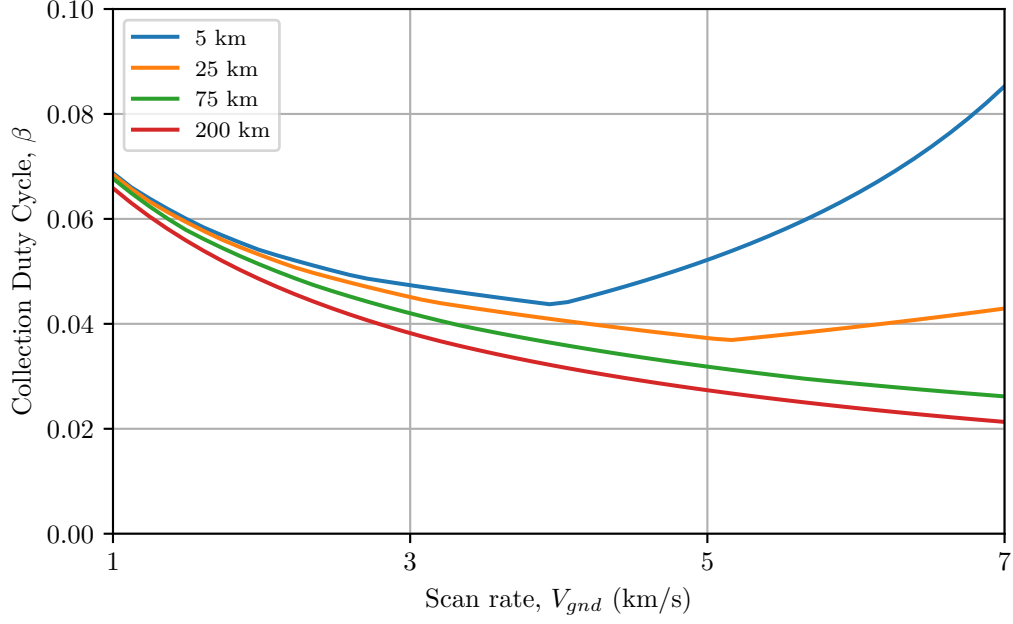
Fig 4: Collection Duty Cycle, β for a reasonably agile spacecraft. Note the different lines correspond to target cross-track spacing for a collect of 50km in-track length

$\beta < 0.1$. In order to generate intuition, Figure 4 captures how β varies as a function of cross-track target spacing and ground scan velocity for a typical, targeted high resolution spacecraft.

For small cross-track spacing, β largely depends on collection size and the relative ground-scan rate that the imaging system is capable of because imaging time dominates slew time. For larger cross-track spacing or smaller target size, the time to slew between targets becomes more important and spacecraft agility influences β more heavily.

Figure 4 is an upper bound in the sense that it assumes there will always be a desirable target to collect when the spacecraft has slewed cross-track which is likely not true for most real-world target collection decks.

We can similarly define a downlink duty cycle

$$\gamma = \frac{t_{d/l}}{t_{per}}, \tag{19}$$

where $t_{d/l}$ is time per orbit spent in downlink. γ is typically ~ 0.1 for a single ground station well matched in latitude to the spacecraft's orbit inclination and can be increased with additional, geographically diverse ground stations.

There are two primary capacity regimes that a spacecraft may operate in:

1. Downlink-constrained regime

2. Collection-constrained regime

For a specified downlink bandwidth $BW_{d/l}$, we can define a capacity constraint metric, k_{cap}

$$k_{cap} = \frac{\gamma BW_{d/l} GSD^2}{\beta I_D ACR}, \tag{20}$$

such that when $k_{cap} < 1$ the system is downlink constrained and when $k_{cap} \geq 1$, the system is collection constrained. The system is well balanced for k_{cap} near unity.

Equation (20) is also important because it directly exposes the tradeoffs inherent in collection capacity (ACR), image quality (GSD and I_D) and system data bandwidth ($BW_{d/l}$).

9

3.4 Etendue and Photon Efficiency

Optical Etendue or throughput, $A\Omega$, is a fundamental space-bandwidth capacity metric for optics. It is important because it captures the optical photon collection capacity of the optics,

$$P_{opt} = A\Omega \int_{\lambda_1}^{\lambda_2} L(\lambda)d\lambda = A\Omega L. \tag{21}$$

$A\Omega$ is also highly correlated with complexity and cost. A reasonable design objective is, given some optical etendue, extract the most imaging performance from the optics as possible. Photons can be lost spatially if the detector area does not cover the full field of view (at the focal plane) of the optic or if there is non-unity duty cycle in photon integration. Thus the effective usable photon power for imaging is

$$P_{img} = \eta_{ph}A\Omega L, \tag{22}$$

where

$$\eta_{ph} = \frac{A_{det}}{A_{fov}}\phi_{int}, \tag{23}$$

where A_{det} is detector area, A_{fov} is the area of the optical field of view at the focal plane, ϕ_{int} is the fraction of time that the detectors are actively integrating photons and L is incoming radiance.

3.4.1 Relationshp of η_{ph} to ACR and IQ

It can be shown that the total photo-electron generation rate, $\dot{N}_{e^-} \propto L\eta_{ph}A_{ap}\Omega$ and for systems operating in the shot-limited noise regime

$$\dot{N}_{e^-} = \frac{ACR \times SNR^2}{GSD^2}, \tag{24}$$

so that

$$\frac{ACR \times SNR^2}{GSD^2} \propto \eta_{ph}A_{ap}\Omega L. \tag{25}$$

$\eta_{ph}A\Omega$ is a fundamental imaging system capacity metric. Various system trades can be made between image quality and area collection rate but these trades are all constrained within the boundaries of this fundamental system capacity. In Section 5.4 we define a sensor metric related to η_{ph}.

4 Collection Modalities

Large orbital velocity is a large advantage of spacecraft for remote sensing over other platforms as it gives:

- Daily global access

- High collection rates

However it also presents large challenges in developing an imaging system that can collect sufficient light at the high relative scene velocity without incurring unreasonable blur.

Equation (25) shows that $\eta_{ph}A\Omega$ (and thus cost) is driven directly by ACR and therefore V_{gnd} (Equation (17)).

For systems whose integration times are limited by velocity (such as blur-limited or TDI systems with a fixed number of stages) $^?$ shows the even stronger dependence of achievable image quality and ground velocity

$$V_{gnd} \propto \frac{GSD^4}{SNR^2}.$$

This is a major challenge for high-resolution systems, and historically there have been many solutions to compensate for scene motion on spacecraft. These date back to the earliest film-based systems that used the film motion through the camera itself to compensate. Other mechanical solutions included gimbaled cameras, "back-scan" of the entire spacecraft bus and full-aperture scan mirrors, among others.

Indeed for systems with resolutions better than a few meters, some form of stabilization is essentially required to obtain sufficient SNR without significant bus back-scan to reduce relative ground velocity.

We will classify modern stabilization strategies into three buckets

1. Analog electronic charge transfer or time-delayed integration (TDI) stabilization

2. Digital oversampling stabilization

3. Analog optomechanical stabilization

Digital imaging spacecraft have traditionally used a "push-broom" architecture where a line-array sensor is pushed along the ground by the spacecraft's motion. Analog stabilization by charge transfer between rows synchronous with scene velocity allows for so-called Time Delayed Integration, TDI.

Development of large format CCD and CMOS framing detectors allows for "step-and-stare" architectures in which a scene snapshot is captured with a 2D framing sensor that is then moved further along the orbital path before another scene is captured. If the framing rate is high enough, the scene can be digitally oversampled such that a single point on the ground is captured with multiple exposures of the focal plane.
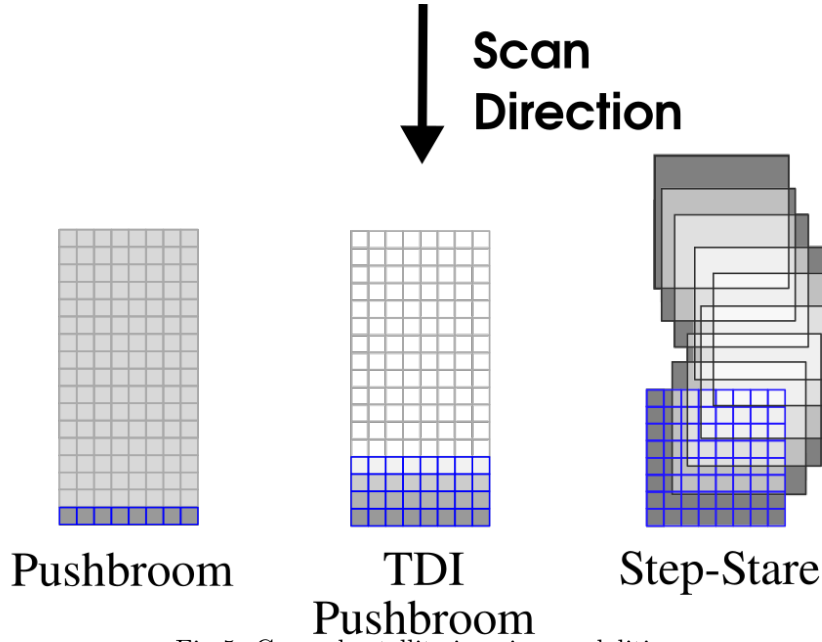

Fig 5: General satellite imaging modalities

Analog optomechanical stabilization relies on an optomechanical component with the ability to dynamically steer the payload line of sight repetitively so as to remove relative scene motion during integration. Fast scan mirrors (FSM) or excited dynamic motion of the entire instrument are two ways to achieve this,[?].[?]

The analog techniques are considered such because they occur before digital sampling of the detectors and are thus still limited by detector well capacity. Digital oversampling is not as constrained by FWC because each exposure resets the pixel and thus the FWC can be effectively expanded by as many redundant samples as can be obtained.

Both TDI pushbroom and step-and-stare allow for some form of stabilization or motion compensation which is required to obtain reasonable SNR in high resolution systems. In the case of Figure 5, analog charge transfer is depicted for TDI pushbroom, and digital oversampling (through overlapping frames) is depicted for step-and-stare. Pushbroom is shown mostly for historical reference as pure pushbroom collection is not feasible in systems with GSD much below 10 m.

To date, analog electronic or TDI stabilization has only been implemented successfully in CCD detectors. There is ongoing work to implement TDI in CMOS detectors but this technology is still immature – see,[?].[?]

Likewise digital oversampling is typically not practical using CCD sensors in high resolution systems due to the inherent limitations on read-out rate of CCD as compared with CMOS technology.

Finally it is important to point out that digital and analog stabilization can be combined in the same system. Such hybrid stabilization schemes could be in the form of analog opto-mechanical stabilization plus digital oversampling or a TDI CCD that is read out fast enough to allow digital oversampling. Indeed we

will argue later that such hybrid stabilization schemes uniquely leverage the capabilities of modern CMOS detectors, enable smaller systems and provide significant system-level flexibility.

4.1 Analog Stabilization with TDI

TDI systems are based on a unique property of CCD image sensors - their ability to move charge between nodes at very high charge transfer efficiency (CTI) and without incurring significant noise in the process. In fact TDI sensors essentially enabled the move away from film in the classified systems of the 1970's.

Typical TDI sensors will have anywhere from 4 to 256 "stages," each consisting of a row of pixels covering the swath width. The rows are oriented normal to the scan direction and charge is shifted row-to-row synchronously with the scene motion such that the effective integration time for a point on the ground is multiplied by the number of stages in the device. Figure 6 illustrates this mode for a 12-stage TDI detector in t-x space where x is in-track dimension.
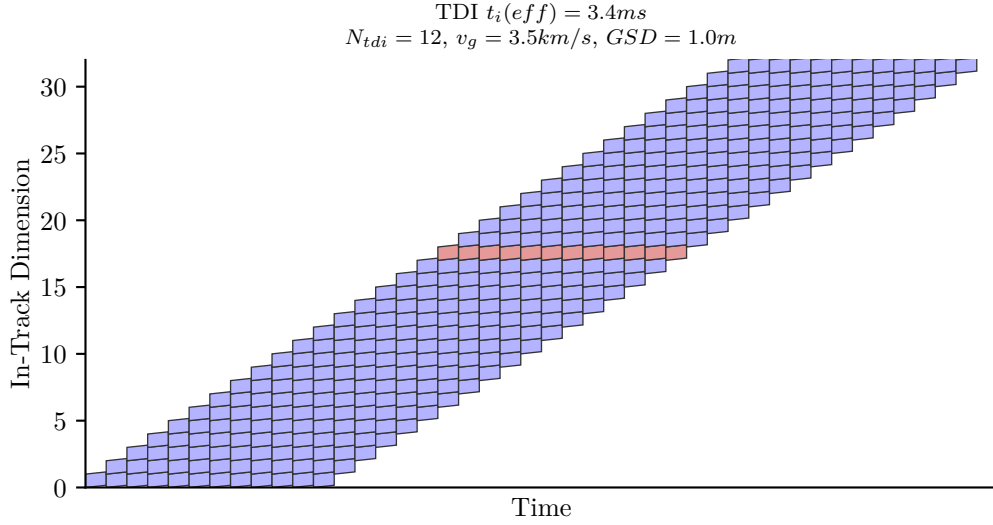

Fig 6: TDI t-x diagram. Red row indicates one in-track ground sample.

In general TDI system's photon collection efficacy is "well-depth limited" meaning that the duration of integration is typically limited by pixel well saturation before other factors such as number of TDI stages, dark current, or charge transfer efficiency. For an appropriately exposed scene, the well-limited signal level is:

$$s = \alpha N_{e^-}^{FWC}. \tag{26}$$

Noting that the exposure time for a single stage of a TDI system is extrinsically related to the linerate LR,

$$t_{int}^{line} = \frac{1}{LR}. \tag{27}$$

and noting that $t_{int} = N_{TDI} t_{int}^{line}$, we can combine with equation (51) to obtain an expression for optimal number of TDI stages:

$$N_{TDI} = \frac{N_{e^-}^{FWC} Q^2}{k_{pe} L_{sat}} \frac{LR}{QE} \propto \frac{N_{e^-}^{FWC} V_{gnd}}{GSD}. \tag{28}$$

While extremely powerful and well-proven over many missions, TDI detectors present challenges for low-cost, high resolution missions.

Historically missions have developed custom TDI detectors due to the unique performance and mission requirements of a spacecraft.[?] These include high line rates, large full-well capacity, low read-noise, high quantum-efficiency and radiation tolerance. The challenge with CCD-based detectors is data read-out rate which is often limiting for high resolution systems.

In Section 5.3 we argue that smaller pixels are desirable from a system-size perspective but this typically comes at the expense of pixel full-well capacity, especially in off-the-shelf detectors. While this limitation can be overcome through digital oversampling (Section 5.2), single TDI detectors inherently couple read-out with integration, precluding this.

Finally, line-of-sight stability is crucially important during integration in a TDI array[?] and the charge-transfer process itself incurs smear,[?] degrading image MTF. Digital oversampling partially mitigates both of these issues.

4.2 Digital Stabilization with Step-and-Stare

Step and stare relies on a high speed framing CMOS sensor to capture multiple, overlapping frames of the same point on the ground as the scene moves by. The redundant samples are then re-combined in post-processing to creat output data with a longer effective integration time – virtual TDI in a way.
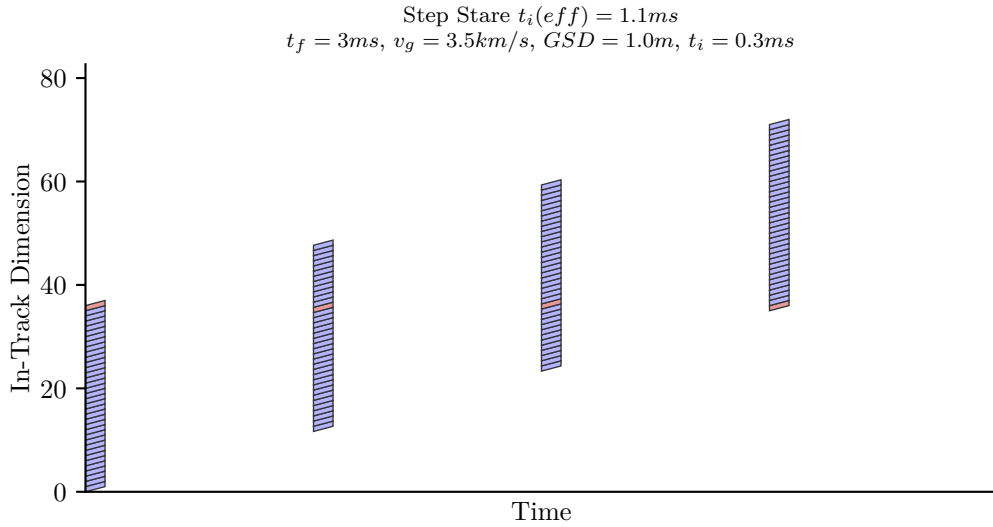

Fig 7: Step-and-stare t-x diagram. Red row indicates one in-track ground sample

Figure 7 illustrates this mode on a t-x diagram. Note that for visualization we illustrate a detector height of only 36 pixels and a very high framing rate whereas in reality most modern framing sensors are > 1000 pixels high and frames at about 1/6 of the rate shown. Figure 8 shows the same thing with realistic frame height and framerate.

It is worth noting in this diagram that the photon collection duty cycle, ϕ_{int} is small, especially compared with a traditional TDI system which is essentially 100%. This can be seen in the white gaps between integration windows. It is also worth noting the slanted integration periods indicative of motion blur during integration. Single frame integration time (and thus SNR) for step-and-stare is typically blur-limited (see 5.2). This is a problematic limitation and severe enough that systems GSD below about 3 m must reduce V_{gnd} with a bus "back-nod" in order to get sufficient signal without incurring substantial blur.

However digital oversampling through step-and-stare IS NOT limited by small pixel FWC, as the well is read out in each independent integration and then summed digitally. This means that, in principle, much higher dynamic range can be achieved with small pixels and at high SNR. To realize this requires analog stabilization to move into the FWC-constrained regime as will be discussed in Section 5.2.

4.3 Hybrid Stabilization

The final modality we will discuss is a hybrid of analog and digital oversampling. This combination eliminates the shortcomings each have alone and provides a great range of flexibility in both the design and operational space of imaging system.

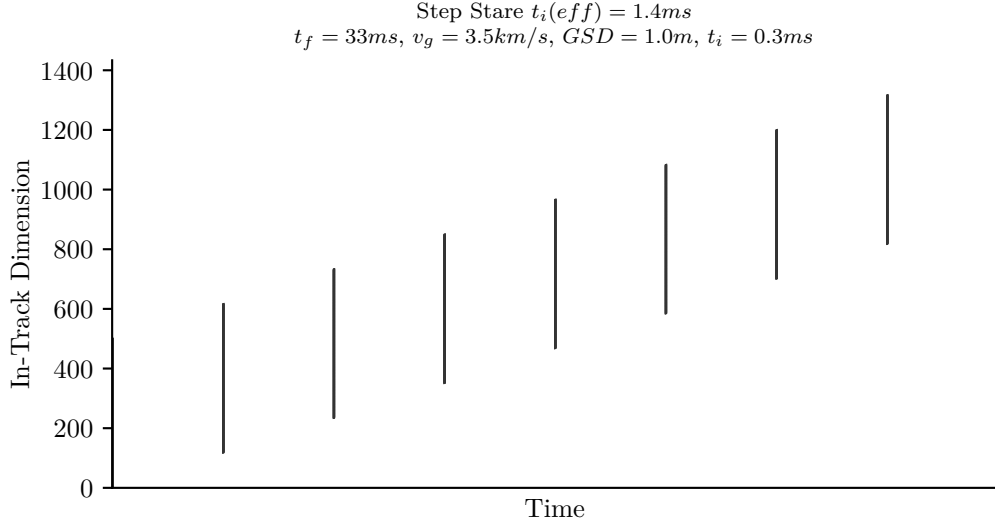More specifically, the issues hybrid stabilization addresses are:

Fig 8: Step-and-stare t-x diagram with realistic array height and framerate.

- Analog stabilization allows operation in FWC-constrained regime, thus extracting the most from optics $(A\Omega)$ and detectors (η_{ph})

- Digital oversampling enables HDR and "flatter" SNR over dynamic range as discussed in Section 5.6

- Digital oversampling allows for resolution-enhancement techniques such as superresolution

- Hybrid imaging systems inherently un-constrain other elements of the system in both development and operation allowing for image quality, area collection rate, data intensity, required ACS stability and more to be traded throughout the lifecycle of the system
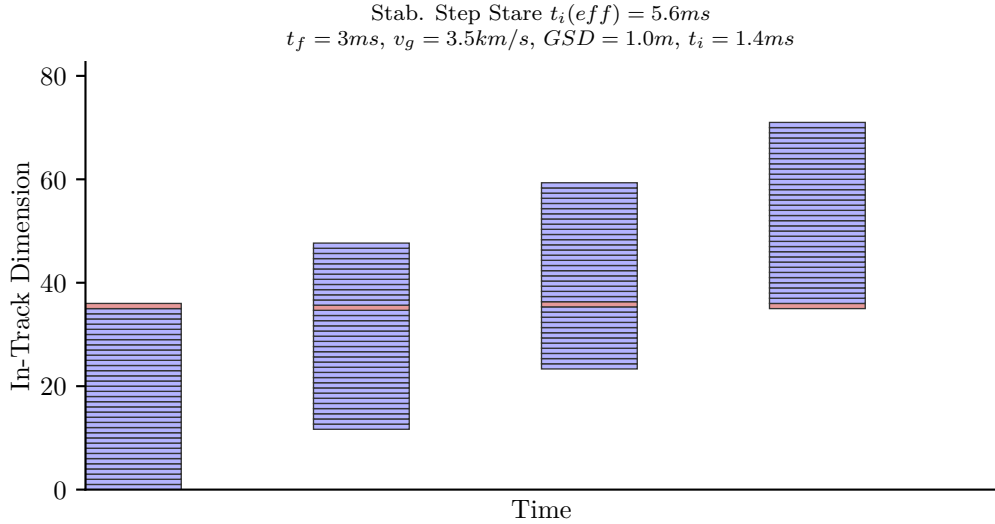

Fig 9: Stabilized Step-and-Stare t-x diagram. Red row indicates one in-track ground sample

Figure 9 illustrates such hybrid collection in t-x space. Note the tall collection columns characteristic of a step-and-stare system and the much larger photon collection duty cycle, ϕ_{int} characteristic of analog stabilization.

14

There are multiple system configurations that allow hybrid stabilization. The detector must be a staring format with enough rows and high enough framerate to allow for digital oversampling. While not out of the question in medium and low resolution systems, CCD's typically do not have the read-out rate to accomplish this in high resolution systems. For staring sensors the linerate, LR in equation (6) is given by

$$LR = (v_{pix}FPS)_{det}. \tag{29}$$

Analog stabilization in a staring sensor system can be accomplished in one of three ways

1. Optomechanical line-of-sight steering element within the optical chain (e.g. fast scanning mirror, detector translation mechanism)[?]

2. Entire payload or instrument-level actuation to achieve back-scan during integration[?]

3. Inter-row charge transfer in a staring sensor that supports it (TDI-mode)

The third option is attractive from a simplicity perspective but the authors are unaware of staring sensors that both support such a charge transfer mode and have sufficient digital read-out rates to enable digital oversampling in high resolution systems.

Alternatively, multiple TDI push-broom arrays with separate digital read-outs mosaic'd in-track could accomplish digital oversampling. The oversampling ratio in this case would simply be the number of independent detectors integrated.

The authors belief is that the most promising path to realization are high framerate CMOS detectors coupled with either option 1 or 2.

5 Performance Trade Space

The remote-sensing trade space is largely constrained by four mission needs:

1. Phenomenology

2. Image Quality

3. Collection Capacity

4. Cost

Phenomenology refers to the study of the phenomenon being sensed and goes beyond Image Quality to include things like the spectral content of the sensed data. We will largely discuss the trade space in terms of wide-bandwidth ($> 100nm$) collection in the near-visible spectrum ($\sim 400nm$ to $\sim 1000nm$) but will also point out where the trends in questions impact the ability to perform other phenomenology (e.g. SWIR or HSI). It is also worth pointing out that staring sensors enable video-mode collection, a phenomenology new to space-based remote sensing.

As discussed in Section 4 and 3.4 Image Quality (IQ) and Collection Capacity are tightly coupled at the system level. Equation (25) demonstrates this deep relationship for a set of other system parameters. In this section we develop this result as it pertains to more granular system parameters.

5.1 Q and Pixel Size

We will study the effect of pixel size through the lens of Q recognizing that[?]

$$Q = \frac{\lambda f}{D_{ap}p_{px}}.$$

Q is generally chosen as a trade between image quality, collection capacity and system size.

One thing demonstrated in a parametric study of the GIQE-5 is that image quality is often maximized at $Q > 1$ (Figure 10). This has also been validated in human-interpreter studies,[?,?] and is not unexpected considering that in a diffraction-limited system with $Q < 2$, there is contrast passed by the optical system
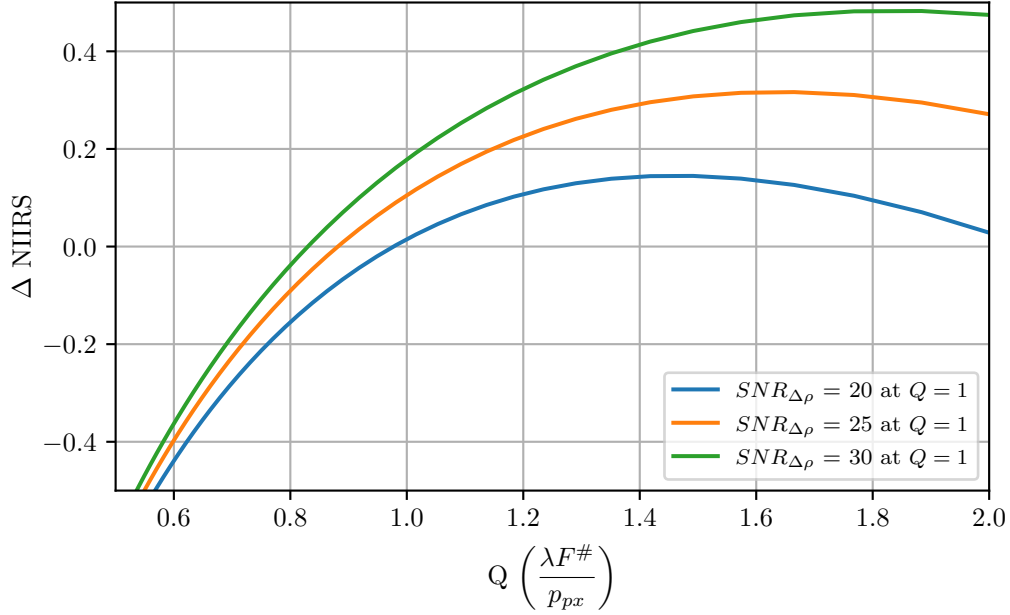
Fig 10: Parametric study of GIQE-5 shows how optimal image quality depends on Q. Note that this plot assumes $RER_0 = 0.3$ and $GSD = 1m$ for $Q = 1$ and that $SNR \sim 1/Q$. RER_0 is also scaled with Q based on an assumption that Q is varied by changing pixel size with fixed optics.
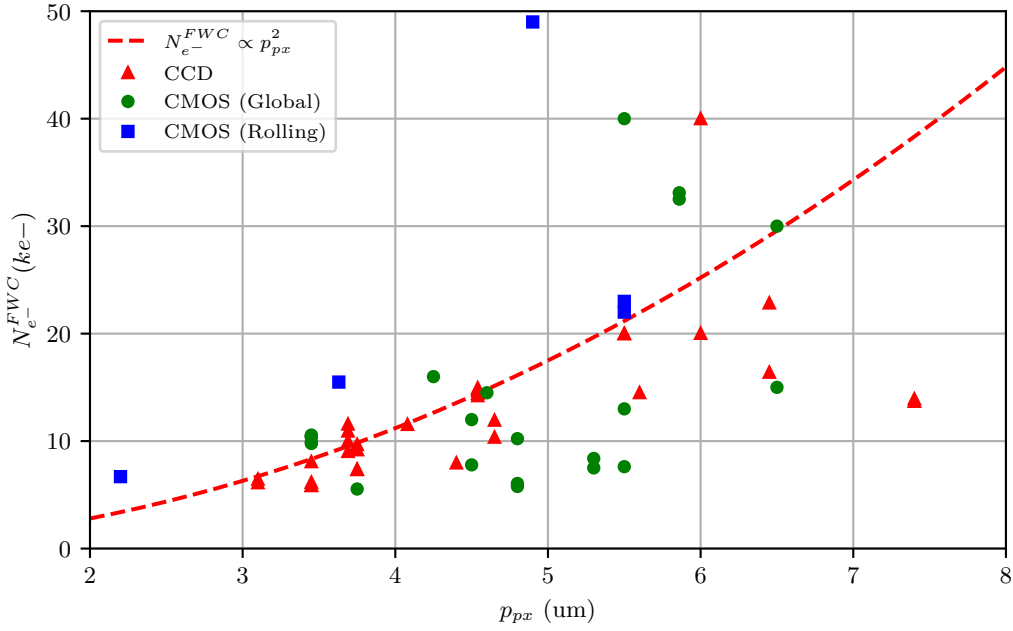
Fig 11: Relationship of $N_{e^-}^{FWC}$ to pixel size, p_{px}. Note that while the upper frontier generally follows Eq. (30), there are notable outliers especially in some of the new CMOS architectures

that is not sampled by the detector. At first blush it seems obvious that, given the cost and difficulty of building larger optics, systems would be designed at $Q > 1$.

However, there are two other system impacts of Q that argue against pushing it higher.

The first is that for CCD-based TDI systems operating in the FWC-limited regime, pixels with sufficient full well capacity (FWC) to meet dynamic range and SNR requirements have required relatively large pixels

compared with commercial state-of-the-art (see Figure 11).

Smaller pixels typically store less charge, both in the photodector and storage node. In fact to first order[?]

$$N_{e^-}^{FWC} \propto p^2, \tag{30}$$

and while there are process optimizations that can improve this, these typically are second order to simple area scaling. Figure 11 shows that Equation (30) describes the upper frontier reasonably well although several recent pixel architectures are significant outliers.

Q can be increased with longer focal lengths as well but this is often practically limited by packaging and size constraints at the spacecraft level.

The second challenge in pushing Q higher is that given other system constraints (e.g. SNR required, focal plane dimension)

$$ACR \propto \frac{1}{Q^2}. \tag{31}$$

Thus designers have tended to accept lower Q, and hence larger GSD and poorer image quality in exchange for manageable optical focal lengths, increased area collection capacity and acceptable SNR and DR.

Observation 1 ($Q > 1$). Image Quality is optimized for $Q > 1$ but because $ACR \sim 1/Q^2$ and systems were pixel-size limited, historical systems typically have been developed at $Q < 1$.

Happily, the trend towards capable commercial framing image detectors with smaller pixels and large read-out rates opens up the opportunity of driving towards higher Q's where image quality is optimized without significantly impacting system size.

5.2 Modality effect on η_{ph}

One way to conceptualize the effect of modality on η_{ph} is as the time-average (or integral over time) of the blue areas in Figures 6, 7, 8 and 9. Increasing in-track height of the blue bars with taller arrays helps improve this integral and reducing the gaps between them horizontally (in time) improves this. TDI has a time-duty cycle of 1 (no horizontal gaps) while step-and-stare typically have much taller arrays.

In order to quantify this it can be shown that

$$A_{fov} = \left(\frac{D_{ap}F^{\#}}{2} \right)^2 \Omega = \frac{f^2 \Omega}{4},$$

or

$$A_{fov} = \left(\frac{D_{ap}F^{\#}}{2} \right)^2 \pi \theta_{fov}^2 = \pi \frac{f^2 \theta_{fov}^2}{4}, \tag{32}$$

where θ_{fov} is the half-cone angle of the field of view in the small angle approximation.

From this, we can derive η_{ph} for systems in two regimes: blur constrained and full-well capacity constrained. The domains are separated by the following criteria:

$$N_{e^-}^{FWC} > \qquad \frac{k_{pe}L_{sat}(QE)N_{bl}GSD}{Q^2V_{gnd}} \rightarrow \qquad \text{Blur-constrained}$$

$$N_{e^-}^{FWC} \leq \qquad \frac{k_{pe}L_{sat}(QE)N_{bl}GSD}{Q^2V_{gnd}} \rightarrow \qquad \text{FWC-constrained}$$

Figure 12 is a regime map for a set of reasonable parameters.

For the full-well capacity-dominated regime

$$\eta_{ph}^{FWC} = \frac{4\lambda^2 h_{px} N_{e^-}^{FWC} LR}{\pi L_{sat} D_{ap}^2 \theta_{fov}^2 k_{pe} QE}, \tag{33}$$

which we can simplify to observe

$$\eta_{ph}^{FWC} \propto N_{e^-}^{FWC} LR. \tag{34}$$

17

η_{ph} regime for $k_{pe} = 2.5 \times 10^5$, $QE = 0.4$, $L_{sat} = 200$, $N_{bl} = 1.0$, $Q = 1.0$, $N_e^{FWC} = 20$ ke-
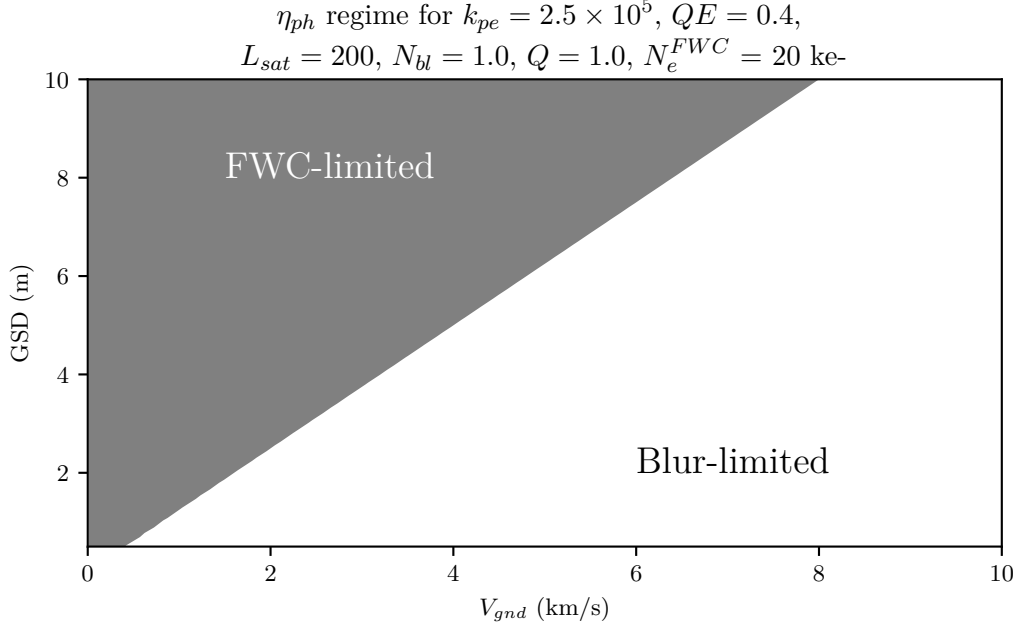
Fig 12: Representative regime map as a function of GSD and V_{gnd}

For high resolution systems without analog stabilization or TDI systems with a fixed number of stages, integration time is typically constrained by the high apparent motion in the focal plane rather than the limited FWC of the pixel (Figure 12). For digital step-and-stare systems without analog stabilization,

$$\eta_{ph}^{bl} = \frac{4\lambda^2 h_{px} N_{bl}(LR)(GSD)}{\pi D_{ap}^2 \theta_{fov}^2 Q^2 V_{gnd}}, \tag{35}$$

Or

$$\eta_{ph}^{bl} \propto \frac{N_{bl}(LR)(GSD^3)}{V_{gnd}}. \tag{36}$$

This scaling law is not good for high resolution systems, requiring high $\frac{LR}{V_{gnd}}$ (and commensurate large data intensity, I_D, $N_{bl} >> 1$ px or very low Q.

Likewise, for a TDI system with a fixed number of stages

$$\eta_{ph}^{TDI} = \frac{4\lambda^2 h_{px} N_{TDI}}{\pi D_{ap}^2 \theta_{fov}^2 Q^2}, \tag{37}$$

and we note that

$$\eta_{ph}^{TDI} \propto N_{TDI} GSD^2, \tag{38}$$

indicating that η_{ph} is highly sensitive to GSD when there is a non-FWC-based integration time constraint.

Observation 2 (Design for FWC-limited regime). Operation in the FWC-limited regime results in higher η_{ph} and thus should be a design goal.

In order to operate in the FWC-constrained regime for some resolution, effective scan speed, V_{gnd} must be reduced. While bus back-nod accomplishes this, it does so at the expense of ACR.

Observation 3 (The need for analog stabilization). Analog stabilization – defined as the dynamic cancellation of ground scan speed, V_{gnd} – is likely required to operate in the FWC-limited regime for high-resolution satellite remote sensing.

5.3 Focal Plane Dimensions

The required physical dimension of the imaging detector is a critical system-driver for a number of reasons. First, the larger required detector dimension implies either a physically large semiconductor or an array of smaller detectors.

Very few large monolithic detectors are produced commercially because typically applications requiring them are very specialized. They are also expensive due to yield issues associated with large semiconductor dies and often have low frame rates due to poor perimeter/area ratios.

Arrays of smaller off-the-shelf detectors at first seem an attractive alternative. However, such tiled arrays also suffer from several shortcomings. First, there is overhead in packaging and electrical connection such that it is difficult in practice to get high effective tiling fill fractions. This may lead to solutions requiring additional optical components such as mirrors to allow for reasonable physical layout of the detectors. The arrays must also be individually optically aligned to the optics which can be costly and time consuming. However if these problems are solved, tiled arrays are scalable and leverage COTS components.

Whether the solution is monolithic or a tiled array, it is clear that minimizing the physical dimension of the detector array is important in constraining system size and cost.

Expanding Eq. (32), it can be shown that, for a circular aperture, the illuminated focal plane radius is

$$r_{fp} = \frac{f\theta_{fov}}{2} = \frac{p_{px}R\theta}{2(GSD)}.$$ (39)

For a detector to span the full width of the illuminated area, it thus must span $2r_{fp}$.

Eq. (39) clearly illustrates how important pixel pitch, p_{px} is to minimizing the physical size of detector arrays and provides the second motivatation for the large push towards smaller pixels.

Observation 4 (Smaller Pixel). Smaller pixels enable driving system size down and Q up but require hybrid stabilization (digital + analog) to make practical due to inherently lower signal, S, obtained during exposure.

5.4 Detector Technology

Per Eq. (33), we note that

$$\eta_{ph}^{FWC} \propto h_{px}N_{e^-}^{FWC}LR.$$ (40)

High efficiency FWC-limited systems must either have larger full-well capacities or higher line-rates. We observed also that smaller pixels have positive system-level size (and thus cost) impacts and thus one desires focal planes with small pixels and high effective line-rates (where effective line-rates for staring arrays is defined in Eq. (29)).

We also showed in Eq. (16) that data pixel FWC is very important to dynamic range and data intensity. In light of these observations we define a sensor figure of merit

$$\psi_{px} = h_{px}N_{e^-}^{FWC}LR.$$ (41)

ψ_{px} is computed for a number of commercial-off-the-shelf sensors and plotted in Figure 13.

Observation 5 (CMOS detectors coupled with analog stabilization are enabling). The trend towards small pixels and high read-out rates enables CMOS detectors to extract more information from high performance, compact (and thus low cost) optics.

Of course, this additional information does come at the cost of increased data intensity as shown in Equation (42) and, empirically across all types of sensors, in Figure 14.

$$I_D \propto \frac{\psi_{px}}{\left(N_{e^-}^{FWC}\right)^3 ACR}$$ (42)

Systems that are downlink-constrained, therefore, will likely be unable to take advantage of the greater information throughput of newer sensors. Equation (42) indicates that in these cases, choosing a sensor with lower framerates and marginally larger pixel FWC is a good trade even if the overall ψ_{px} is substantially lower. Indeed if downlink bandwidth is the largest constraint, Equation (42) suggests that $N_{e^-}^{FWC}$ is probably the most important sensor metric.
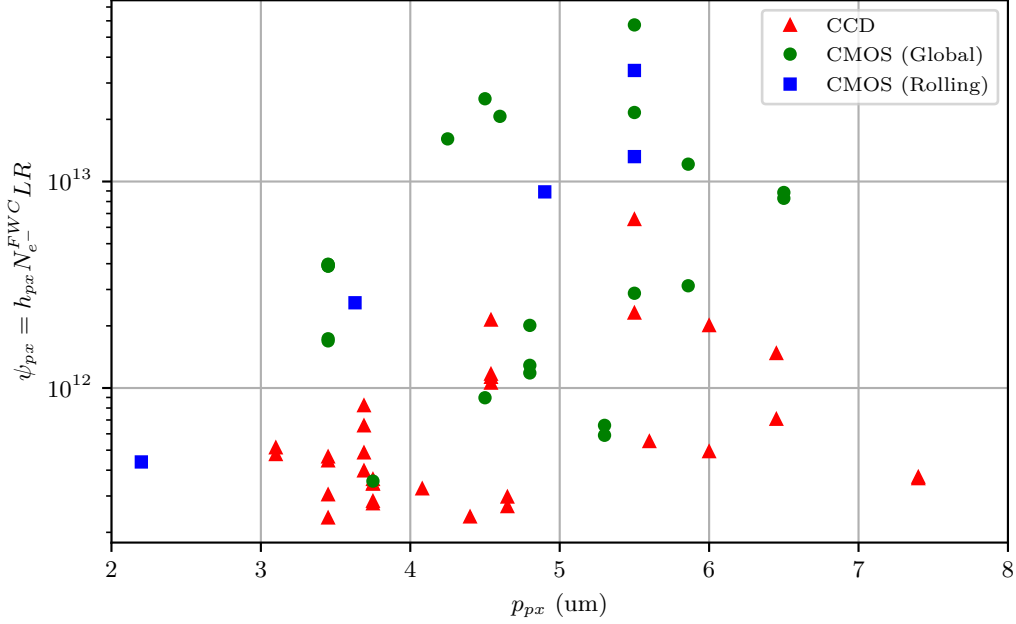
Fig 13: Detector figure of merit, ψ_{px}. Note log scale and general trend of higher ψ_{px} for CMOS sensors.

One way around this limitation is to shift more of the data reduction onto the spacecraft, for example by performing on-board stacking of digitally-oversampled data. As shown in Figure 15, stacking redundant samples prior to downlinking the data reduces the data intensity for a given quantity of information, with detectors that rely more heavily on framerate rather than well capacity realizing a more significant improvement.

The combination of high-frame-rate sensors and robust on-board processing offers the potential to dramatically increase the usable data that can be obtained from bandwidth-limited systems and enable full utilization of high ψ_{px} sensors.

5.5 Spectral Bands

There are a number of system parameters that can be tuned to take advantage of the additional information content available from high-framerate sensors. It can be used to increase area collection capacity (by increasing the ground scan speed of the boresight), improve dynamic range and signal-to-noise (see 5.6), or add diagnostic capability through additional spectral bands. A standard architecture for collecting multi-spectral data is to use a multi-stripe filter over the sensor, dividing it into a number of regions of approximately equal in-track extent. Each ground point is sampled as it moves through each of the spectral bands. This technique can be used for both TDI and push-framing sensors. For a system where SNR is FWC-constrained, the number of spectral bands that can be collected with a minimum signal-to-noise ratio SNR_{min} at L_{typ} is

$$N_{bands} = \frac{\text{floor}\left(\frac{N_{rows} \times GSD \times LR}{V_{gnd}}\right)}{\text{ceil}\left(\frac{SNR_{min}^2 \times \alpha}{FWC}\right)} \tag{43}$$

This is a fundamental characteristic of the system that depends on the selection of high-level system parameters (dynamic range, SNR and GSD) as well as details of the orbit and the sensor. Figure 16 shows the number of bands available in a 1m GSD system at 500 km. It is evident that only high-framerate sensors are capable of capturing true multispectral imagery at high resolution in this geometry. For a medium-resolution system, however, this constraint does not apply; as seen in Figure 17, a number of CCDs are capable of reading out quickly enough to capture up to 10 spectral bands.

20
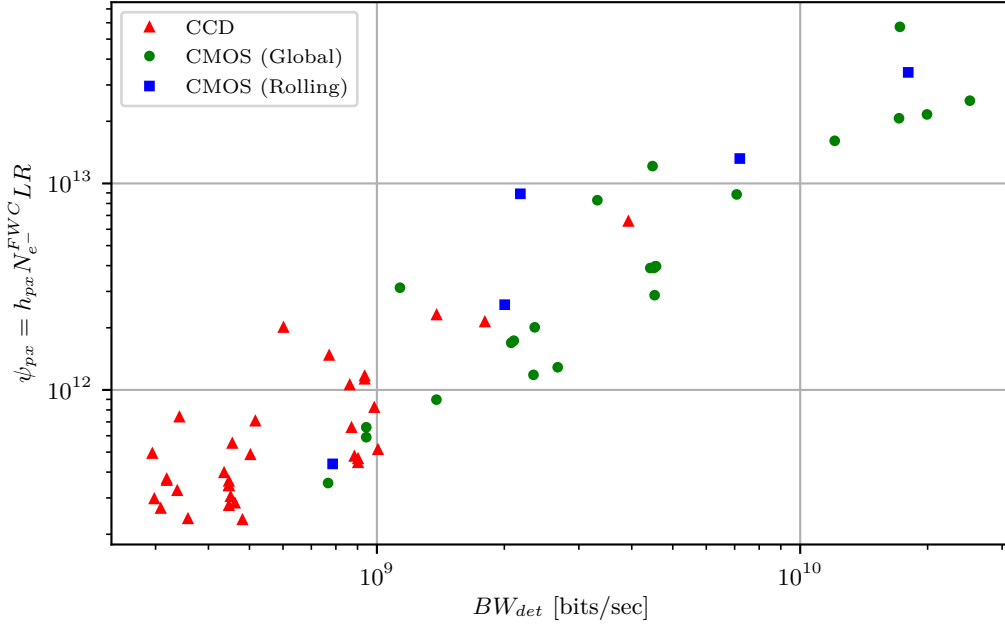
Fig 14: Detector figure of merit is closely correlated with Data Intensity (as expected from Equation (42)), with slight deviations from the trend due to subtle differences in sensor characteristics like aspect ratio and the ratio of well depth to bit depth. CCDs occupy the lower-left corner of the plot, suggesting they are applicable for bandwidth-starved systems that cannot do on-board data fusion.
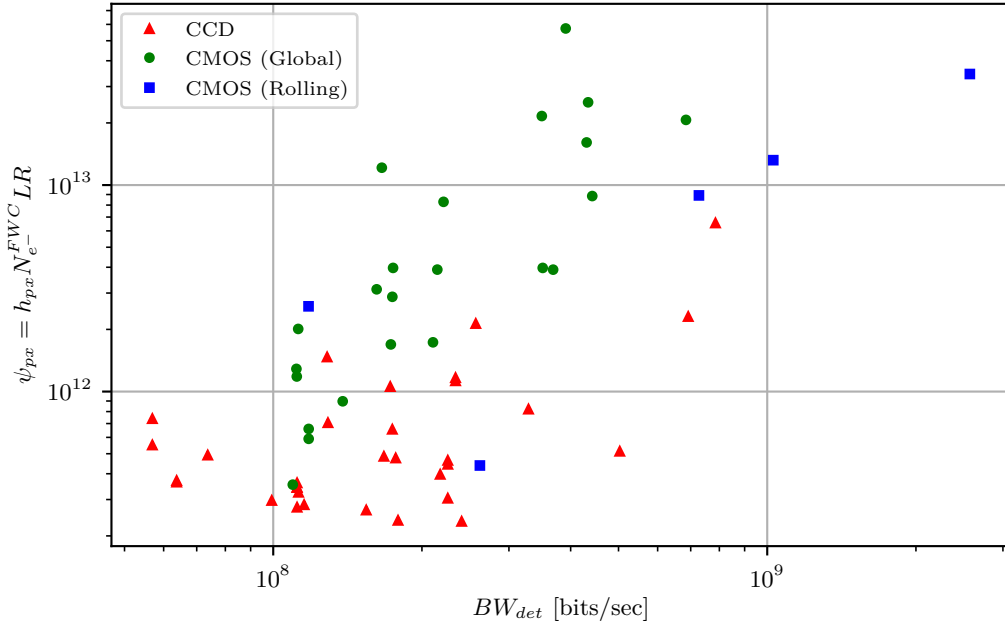
Fig 15: If digitally-oversampled images can be combined on-board, then the data intensity can be greatly reduced (with systems that rely heavily on digital oversampling benefitting the most from this improvement). Note that, in the above plot, some CMOS systems can provide substantially more useful information for the same amount of downlinked data as the best CCD systems.
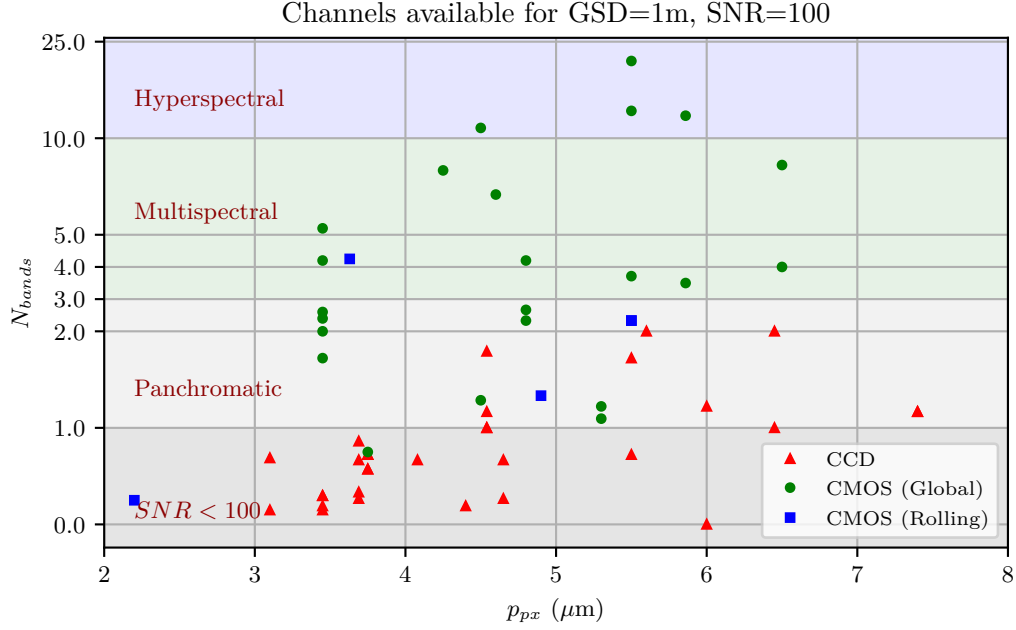
Fig 16: In this case, we assume a system with an orbital altitude of 500km, a GSD of 1m, and a FWC-constrained SNR. The focal plane is split into spectral bands of equal in-track extent, where the height of each spectral band is set such that the system is able to achieve an SNR of 100 at Ltyp, with alpha (Lsat/Ltyp) = 5 for each ground point. Note that most systems (CCDs and CMOS sensors) require some amount of digital oversampling to achieve the SNR requirement. Most CCDs struggle to sample each point in more than a single channel, while CMOS sensors offer the possibility of truly multispectral coverage.
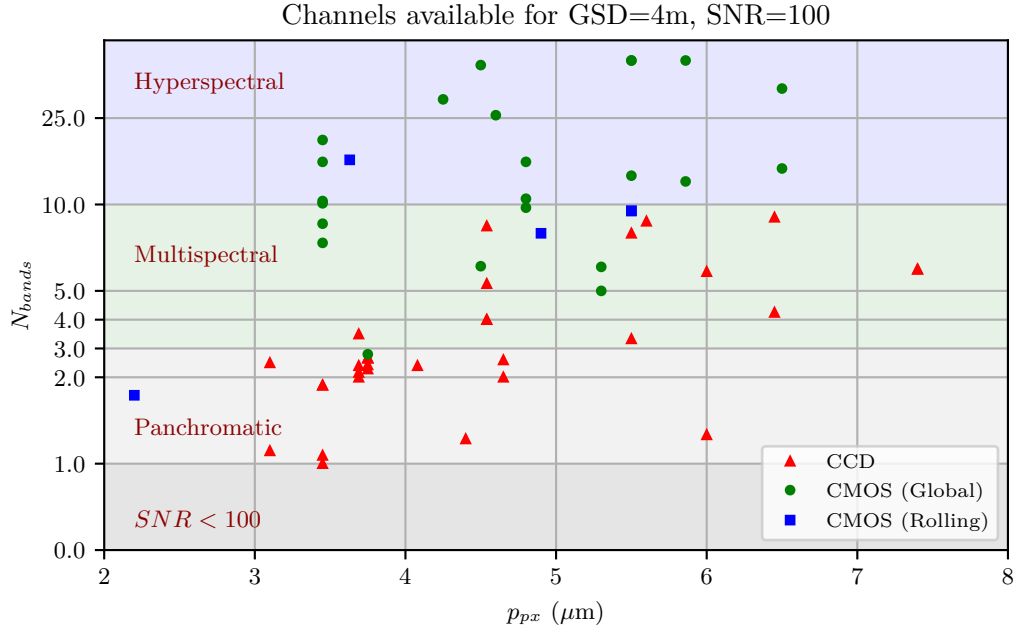
Fig 17: Note that, for a GSD of 4m, CCDs are a much more appropriate choice, as there are a number that are capable of multispectral imaging. While the plot suggests that the greater capacity of CMOS sensors may be leveraged by building hyperspectral imagers, in practice this approach requires bandpass filtering that may take the system out of the FWC-limited regime.

5.6 High Dynamic Range Imaging

As mentioned previously, digital oversampling allows improvements in SNR. Previous analysis assumed constant integration times for each of the redundant images, and it was shown that SNR scaled with the square root of the number of redundant samples. Here we demonstrate the advantages of digital oversampling using different integration times, and how it leads to high dynamic range imaging by shaping the SNR profile across a range of scene brightness.

In general photography, exposure bracketing – namely, taking multiple pictures of a scene with different exposures – has been around for a long time, and in the last few decades has evolved to the point today where we see typical consumer camera phones digitally combining images taken with different exposures in near real time.

When digitally oversampling with a constant integration time, every frame is capturing the same range of brightness values. The maximum brightness (without sensor saturation) is fixed, and SNR is simply being scaled according to the square root of the number of samples. Note that the SNR for very bright parts of the scene, which are already high due to the nature of photon shot noise, are scaled by the same amount as the dark parts of the scene, which have much lower SNR. Distribution of brightness is scene dependent, but it is often the case that a large fraction of scene content is relatively dark compared to the bright objects in the same scene. With constant integration time, to capture the dark content with high SNR requires either longer integration times at the cost of clipped scene highlights, or additional oversampling at the cost of data intensity (and perhaps collect capacity due to slower V_{gnd}).

Using different integration times resolves the above dilemma: Shorter integration times can capture the brightest areas in the scene, longer integration times can capture the darker areas in the scene, and digital combination of all exposures yields images that combine the best of the short and long integrations, and everything in between. Figure 18 demonstrates the concepts from the point of view of SNR. In the figure are the SNR profiles of three exposures: A "nominal" exposure (1x), an exposure with double the nominal integration time (2x), and an exposure with quadruple the nominal integration time (4x). The 1x exposure captures some range of brightness values, normalized to maximum value of 1.0, with an SNR dependence on brightness that assumes photon shot noise as the dominant noise source. The 2x integration captures only half the range of brightness values, but has 1.4 times the SNR over that range; similarly, the 4x integration only captures one quarter the full range of brightness, but has double the SNR over that range.

Combining the three exposures from Figure 18 yields an effective SNR profile as shown by the saw-tooth dashed line in the figure. This composite SNR assumes a straightforward maximum-likelihood approach to combining images taken with different integration times, and does not include any advanced processing such as super-resolution, deblurring, or noise reduction. Note that in this hypothetical case, some darker regions in a scene have higher SNR than some brighter regions of the scene, and is quite distinct from the SNR behavior of simple shot noise.

More than simply increasing the dynamic range of the output data product, allowing different integration times grants considerable flexibility in tuning image quality. Figure 19 illustrates the power of this approach. In a typical scene, the majority of the content has a luminosity much lower than the saturation value, and α is between 4-16. For a single-shot image, the signal-to-noise ratio for most of the pixels of interest is therefore 2-4 times lower than the FWC-limited SNR of the sensor. Even a relatively small amount of digital oversampling, using a sensor with a much smaller FWC, allows a system to achieve a higher SNR in the luminance range where most of the content lies.

Figure 20 shows the resulting distribution of SNR for all the pixels in the image. The simple HDR system produces a uniform, high SNR for the bulk of the pixels in the scene, while the single-shot image suffers from saturation of highlights and underexposure (low SNR) for much of the scene. Indeed, this is a best-case scenario for single-shot (or TDI) system, in that we have assumed that the integration time has been tuned perfectly to the desired saturation radiance; if there are errors in the exposure model, the SNR distribution will be shifted to produce either more severe saturation or more underexposure. The HDR scheme can be configured to optimize SNR over an extremely wide range of luminance values.

5.7 Comparison of Modalities

In summary, we have shown that the hybrid stabilization imaging modality has some strong advantages over pure analog or digital stabilization. Specifically hybrid stabilization
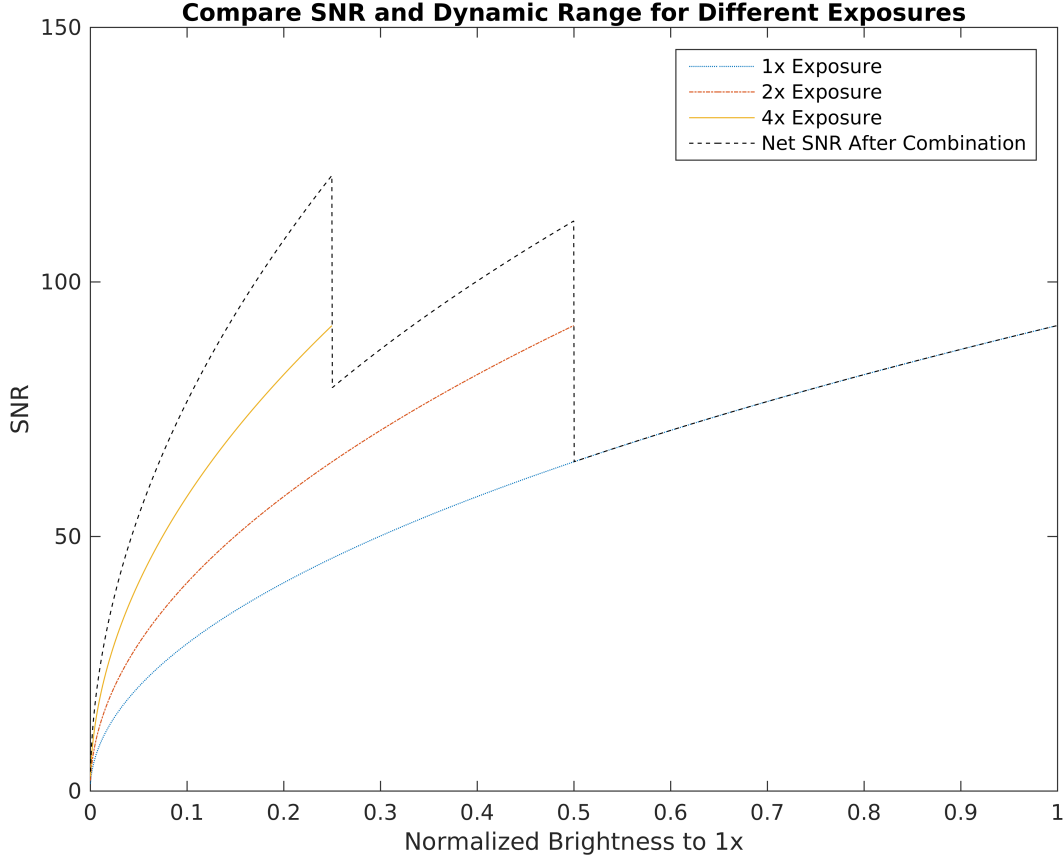
Fig 18: SNR and dynamic range for different integration times. The nominal 1x integration time captures a range of bright levels up to 1.0. Longer integration times capture smaller ranges of brightness, but with higher SNR. Combining the three image exposures gives the saw-tooth-shaped SNR profile.

1. Leverages the high ψ_{px} of newer sensors

2. Enables operation in the FWC-limited regime

3. Expands dynamic range and SNR beyond the native sensor capacity

4. Provides the designer with additional parameters that can be used to trade image quality, capacity and data intensity through the entire system lifecycle (even in operation)

Table 2 summarizes qualitatively the relative advantages and disadvantages of the three modalities.

	DR	SNR	η_{ph}	Stability Req'd	Data Intensity
Analog Electronic	-	+	-	- -	++
Digital Oversampling	+	+	-	++	- -
Hybrid	+	++	+	Tunable	Tunable

Table 2: Relative advantages and disadvantages of the modalities. DR is dynamic range, SNR is signal-to-noise ratio.
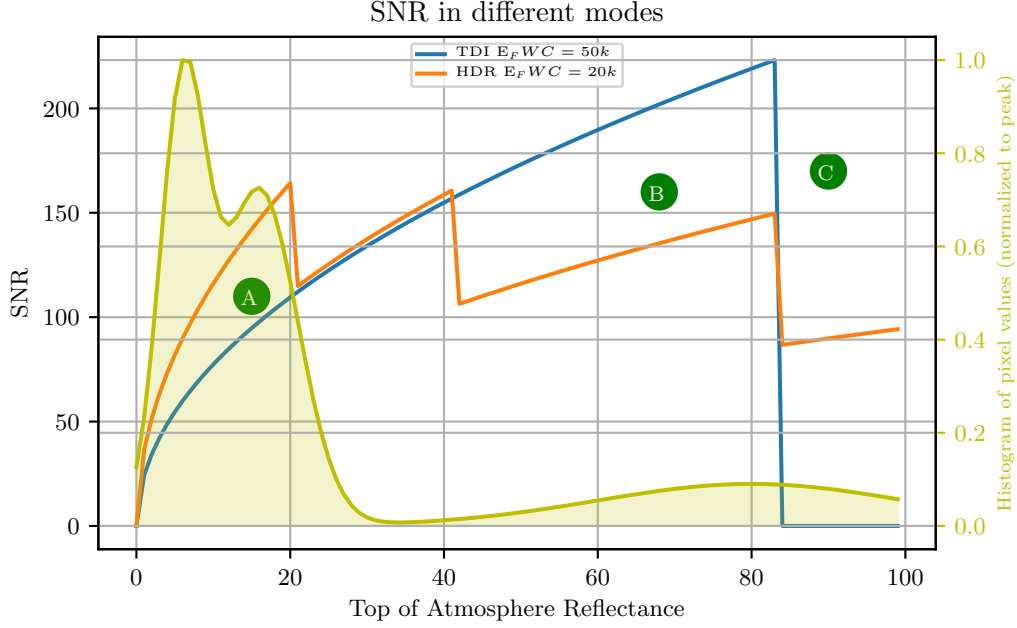
Fig 19: The SNR as a function of scene luminosity for a TDI-based system using a CCD with a FWC of 5e4 electrons, and for a digitally-oversampled system using 4 different HDR integration times on a CMOS sensor with a FWC of 2e4 electrons. Also shown is the luminosity histogram for a typical scene. In region A (relatively low luminosity), the stabilized HDR system is able to achieve a substantially higher SNR than the TDI system thanks to the ability to take long integrations. In region B (relatively high luminosity), the TDI system has higher SNR; however, there is minimal practical advantage, since there is relatively little scene content in this regime, and the extra SNR is above the "good enough" level that the HDR system can capture. In region C (highlights and specular reflections), the TDI system is saturated and reports no useful signal, while the HDR system is still able to discriminate contrast within the highlights.

6 Design Case Studies

In this section we take the earlier results and apply them to two systems: a very high resolution small satellite for point-collection and a high resolution large area collector for mapping. For both case studies, we use the following assumptions:

- Goal is to meet requirements while minimizing overall system size, which is driven directly by aperture diameter and focal plane size

- Imaging performance is defined by GIQE-based GRD_{eff} or NIIRS

- For simplicity, IQ is specified only for panchromatic but it is assumed the focal plane is split in in-track extent equally amongst 5 multi-spectral bands (including PAN) and all bands have identical GSD

- Dynamic range computed assuming equal exposures – the authors note that improved dynamic range is possible for digitally stabilized system with bracketed exposures as discussed in Section 5.6

Additional requirements and assumptions are captured in Table 3.

Using the SNR requirement at L_{typ} and saturation at L_{sat} from Table 3 as well as Equation (11) we can immediately determine that for any system meeting these requirements,

$$\left[N_s N_{e^-}^{FWC}\right]_{req} \geq \frac{SNR_{req}^2}{\alpha} = 9.8 \times 10^4 \, e^- \tag{44}$$

One design choice would be to use an analog-stabilized detector that does not allow digital oversampling (TDI push-broom or staring sensor with $LR < 2\frac{V_{gnd}}{GSD}$) and require $\mathbf{N_{e^-}^{FWC} \geq 9.8 \times 10^4 \, e^-}$.
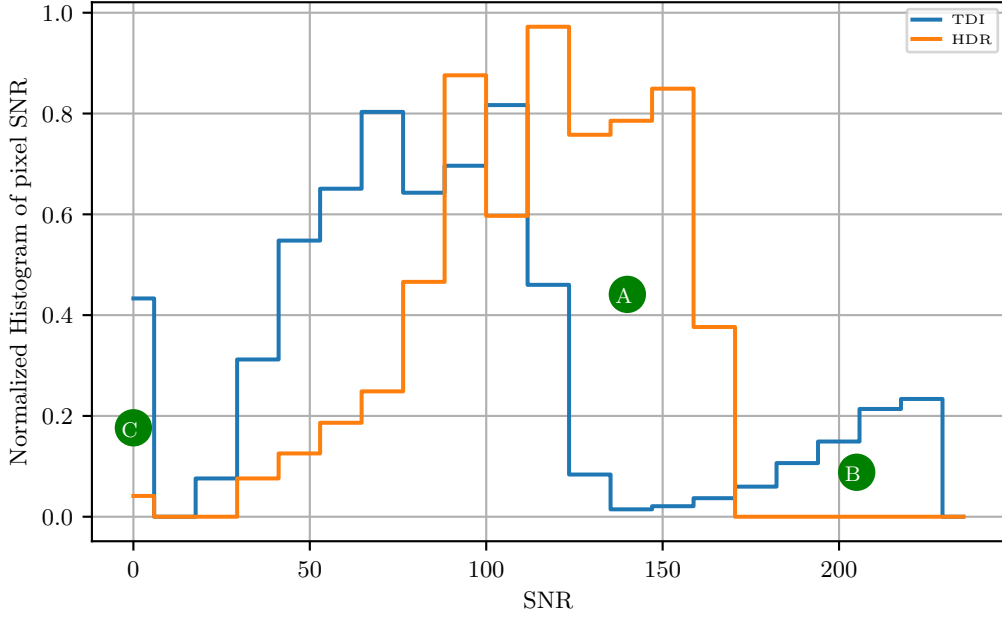
25

Fig 20: Histogram of SNR values assuming the scene luminosity histogram and luminosity-to-SNR mapping of the two systems in Figure 19. Note the HDR system observes the bulk of the scene in the "sweet spot" of high SNR, while the TDI system has a lower mean SNR along with a substantial "blind spot" for highlights.

Unfortunately, at the small pixel sizes there are no known available commercial sensors with sufficient well depth. Therefore a small-system based on COTS detectors is pushed to digital oversampling.

Additionally, in order to avoid operating in the blur-limited regime we will state without defense (see Figure 12) that all of these systems must have some form of analog stabilization (either TDI or opto-mechanical)

We can connect Equation (44) and our detector metric, ψ_{px} in Equation (41) to show that:

$$\psi_{px} \geq \frac{h_{px}N_\lambda V_{gnd}\left[N_s N_{e^-}^{FWC}\right]_{req}}{GSD \times N_{det}},\tag{45}$$

where we introduce N_λ to denote the number of unique spectral bands the focal plane is divided into in the in-track dimension and N_{det} to denote the number of unique detectors tiled in the focal plane.

6.1 50-cm Resolution Point Collection Small-Sat

This design point is for a minimum-size (low-cost) 50cm GRD_{eff} spacecraft with a relatively small cross-track swath width. High level requirements are captured in Table 4.

Note that the first thing we will look at is the required ψ_{px} for the focal plane. Given the GRD_{eff} (or equivalently, $NIIRS$), Q, obscuration and WFE requirements, we use GIQE5 to determine $D_{ap} = 0.48\,m$, $GSD = 42.5\,cm$ and $h_{px} \geq 17,698$.

Using Equation (45) we can determine $\psi_{px} \geq \mathbf{1.4 \times 10^{14}\,e^-/s}$ - a requirement higher than any single detector in our survey is capable of (Figure 13). The large number of cross-track pixels required is also a challenge given typical COTS sensor formats. Happily, we can split the FOV up among a tiled array of detectors to reduce the required ψ_{px} and h_{px} for any one detector.

The 4K TV standard means that assuming a single detector width of 4,000 pixels is a reasonable first assumption for tiling. To meet the h_{px} requirement, this would require four tiled detectors. And now the unit-detector ψ_{px} is reduced by a factor of four to $\mathbf{3.6 \times 10^{13}\,e^-/s}$. The only sensor in our survey that meets this requirement is the BAE LTN4625a with a $\psi_{px} = \mathbf{5.73 \times 10^{13}\,e^-/s}$.

26

Requirement	Value	Note
Altitude	500 km	
V_{gnd}	7.06 km / sec	
λ	450 nm to 800 nm	
λ_c	525 nm	
L	200 W/m^2-sr (sat.) 10 W/m^2-sr (typ.)	Over λ, see[?]
SNR	70	at L_{typ} for $\alpha = 0.05$
DR	80 dB	
$\sigma_{uncorr.}$	$< 10\,e^-$ RMS	
Optical Obscuration	0.3	By diameter
Optical WFE	0.15 λ RMS	At λ_c
N_λ	5	PAN+NRGB
N_{ct}	5	~ 20 kpix cross-track req.

Table 3: Case study requirements and assumptions.

As we noted before, analog stabilization is required to operate in the FWC-constrained regime making the overall system a hybrid stabilized system.

The full point design given this sensor choice produces the parameters captured in Table 5.

We re-iterate the following observations about this point design:

1. The minimum requirements for this system can only be met with small-pixel COTS sensors using hybrid stabilization and high-speed CMOS framing sensors

2. This system achieves a very high level of image quality (NIIRS 5.4, 42-cm GSD, high SNR) simultaneously with a high dynamic range of > 82 dB

3. The cost of this image quality perfmance is a relatively small swath width of 7.8 km and high data intensity of > 41 bit / pix uncompressed.

6.2 1-m Resolution Large-Area Collector

For this case study, the design point goal is a 1-m class GRD_{eff} with a large swath width for large area coverage. Image quality is still considered important and so none of the other requirements (SNR, N_λ, etc) are different than for the 50-cm system. The requirements for this system are captured in Table 6.

Following the analysis done for the 50cm case, we determine that the required ψ_{px} for this system is $\psi_{px} \geq \mathbf{9.9 \times 10^{13}\,e^-/s}$. Again this value exceeds any of the surveyed sensors so a tiled array is required.

We again assume a 4K-wide format sensor as a starting point and note that the cross-track pixel count requirement is $h_{px} \geq 23,966$ and to meet this requires at least six 4,000 pixel wide sensors.

Requirement	Value	Note
GRD_{eff}	0.5 m	at Nadir
NIIRS	5.4	From Eq. (5)
x_{ct}	7.5 km	
Q	1.3	Compromise between IQ, ψ_{px}, I_D
$N_s N_{e^-}^{FWC}$	98 ke^-	From SNR_α, α

Table 4: Specification for 50-cm small-sat point collector.

Parameter	Value	Note
GSD	42 cm	From GIQE optimization
D_{ap}	48 cm	From GIQE optimization
h_{px}	18432 (total) 4608 (per sensor)	From N_{ct}, GSD
x_{ct}	7.8 km	
ψ_{px}^{det}	$\geq 3.6 \times 10^{13}\, e^-/s$	
Focal Plane Width	96.9mm	
FPS^{req}	≥ 96	
Ns	3	
I_D^{raw}	41.3 bit / pix	Uncompressed
DR	82.8 dB	

Table 5: Dependent system parameters for optimized 50cm system

With six sensors, the unit-sensor ψ_{px} requirement is now reduced to $\psi_{px} \geq \mathbf{1.7 \times 10^{13}\, e^-/s}$. This performance is met by five of the surveyed sensors all of which are CMOS. The GPixel 3005 sensor has an electronic rolling shutter and is thus likely problematic.

Given the choice of sensors, we look at how other system parameters are affected by which sensor is selected in Table 7.

With the exception of the LTN4265a, the biggest downside of most of the sensors is the large data intensity, I_D, and relatively low DR – both driven by the small FWC.

Requirement	Value	Note
GRD_{eff}	1.0 m	at Nadir
NIIRS	4.4	From Eq. (5)
x_{ct}	$\geq$ 20 km	
Q	1.3	Compromise between IQ, ψ_{px}, I_D
$N_s N_{e^-}^{FWC}$	98 ke^-	From SNR_α, α

Table 6: Specification for 1-m large-area collector.

Sensor	N_{det}	FPS^{req}	x_{ct}	DR	I_D	$F^{\#}$
LTN 4625a	6	49	23 km	82.8 dB	41.3 bit / pix	13.6
Python 25k	5	75	21.2 km	68.2 dB	102.0 bit / pix	11.1
CMV12000	6	111	20.3 km	69.4 dB	92.2 bit / pix	13.6
CMV50000	4	49.5	26.3 km	72.8 dB	84.6 bit / pix	11.4

Table 7: Impact of sensor choice on other system parameters

It is also worth pointing out that although the CMV50000 nominally meets the ψ_{px} requirement when we assumed a six detector mosaic, to meet the cross track FOV requirement only requires four such detectors. And with only four detectors it does not meet the ψ_{px} requirement any longer due to its relatively low frame readout rate of 30 FPS.

Again, we chose the LTN4625a sensor and the resulting system parameters are compiled in Table 8.

6.3 30-cm Point Collector

For fun we will try one more case with the very agressive design goal of a 30-cm GRD_{eff} system. Iteration shows that some requirements must be relaxed to achieve this case with the sensors available to us and so we will relax the following requirements:

- Required SNR at $\alpha = 0.05$ reduced from 70 to 50

- Required dynamic range reduced from 80dB to 75dB

The other case-specific direct and derived requirements are captured in Table 9.

The model produces a feasible solution to this system design utilizing 5 tiled LTN4625a detectors with the parameters given in Table 10

This is an impressive result considering that this is a system with better GSD than commercial state-of-the-art (WorldView-3, WorldView-4) and that is likely significantly smaller and lower cost than those systems.

7 Conclusions

In summary, we have shown that

1. The trend in space-based remote sensing is to smaller sized system due to cost and scalability

Parameter	Value	Note
GSD	83 cm	From GIQE optimization
D_{ap}	25 cm	From GIQE optimization
h_{px}	27648 (total) 4608 (per sensor)	From N_{ct}, GSD
x_{ct}	23 km	
ψ_{px}^{det}	$\geq 1.7 \times 10^{13}\, e^-/s$	
Focal Plane Width	127.2 mm	
FPS^{req}	≥ 49	
Ns	3	
I_D^{raw}	41.3 bit / pix	Uncompressed
DR	82.8 dB	

Table 8: Dependent system parameters for optimized 1-m system

Requirement	Value	Note
GRD_{eff}	30 cm	at Nadir
NIIRS	6.1	From Eq. (5)
x_{ct}	≥ 5 km	
Q	1.3	Compromise between IQ, ψ_{px}, I_D
$N_s N_{e^-}^{FWC}$	50 ke^-	From SNR_α, α

Table 9: Specification for 30 cm point collector

2. Small pixel, high data rate commercial CMOS detectors enable high performance systems at reduced size & cost

3. To realize potential of modern area detectors, a combination of analog and digital stabilization (hybrid stabilization) is required

4. Hybrid stabilization with high speed CMOS sensors enables High Dynamic Range imaging in high resolution systems

Parameter	Value	Note
GSD	23 cm	From GIQE optimization
D_{ap}	89 cm	From GIQE optimization
h_{px}	23040 (total) 4608 (per sensor)	From N_{ct}, GSD
x_{ct}	5.3 km	
ψ_{px}^{det}	$\geq 3.4 \times 10^{13}\, e^-/s$	
Focal Plane Width	127.2 mm	
FPS^{req}	≥ 119	
Ns	1	
I_D^{raw}	13 bit / pix	Uncompressed
DR	78.1 dB	

Table 10: Dependent system parameters for optimized 30 cm system

5. Hybrid stabilization with high-speed CMOS sensors enable higher spectral dimensionality systems (multi- or hyper- spectral)

The potential of the trends identified here have yet to be fully realized and the authors are excited to see this potential realized.

8 Appendix A: Definitions

8.1 $\frac{\lambda F\#}{p}$ or Q

The system parameter Q is defined as[?]

$$Q = \frac{\lambda F\#}{p}. \tag{46}$$

and is essentially a measure pixel sampling in the spatial domain relative to the Nyquist criterion. Nyquist says that to completely reconstruct a band-limited system, the spatial sampling frequency.

$$\rho_s = 1/p \geq 2BW_{opt}$$

Because we always sample to DC, the optical bandwidth BW_{opt} is replaced by the diffraction cutoff frequency.

$$\rho_c = \frac{1}{\lambda F\#}$$

Thus Q is a measure of Nyquist sampling where $Q = 2$ is critically sampled. Most systems, however, operate at $Q < 2$ and hence incur some spatial aliasing. There is little reason to operate at $Q > 2$ as the optics do not pass spatial frequency information there so the additional sampling is essentially wasted.

Beyond a statement of the Nyquist criterion, Q is a very useful parameter by which to evaluate and compare systems. For more see.[?]

8.2 Photon Radiometry

In order to not have to carry full radiometry for every calculation, we will define a payload photo-electron gain,

$$k_{pe} \equiv \frac{\dot{N}_{e^-}^{std}}{L} \tag{47}$$

Where

$$L = \int_{\lambda_1}^{\lambda_2} L_\lambda(\lambda)d\lambda$$

is radiance at aperture in W/m^2-sr, QE is quantum efficiency, $\Delta\lambda$ is spectral bandwidth and $\dot{N}_{e^-}^{std}$ is photo-electron generation rate for the case of $QE = 1$, $Q = 1$.

Moving forward, we will use $k_{pe} = 2 \times 10^5$ e$^-$-sr-um-m^2/J[2]

Now we have a simple relationship useful in comparing systems radiometrically:

$$\dot{N}_{e^-} = \frac{k_{pe}L \times QE}{Q^2} \tag{48}$$

Typically remote sensing systems report signal to noise at a radiance a small fraction of the saturation radiance. We will define this ratio of typical to saturation radiance -

$$\alpha = \frac{L_{typ}}{L_{sat}} \tag{49}$$

For systems in which shot and read noise dominate

$$SNR = \frac{s}{\sigma_{rd} + \sqrt{s}} \tag{50}$$

where s is the signal magnitude in photoelectrons and σ_{rd} is read noise in RMS photoelectrons.

Finally, we come to a very useful relationship for a single exposure in eq. (51) -

$$s = \alpha k_{pe} L_{sat} \frac{t_{int}QE}{Q^2} \tag{51}$$

Note that in several of the modalities, digital over-sampling is utilized. In that case, the effective output signal level for a ground sample is

$$s^{eff} = N_s s$$

where N_s is the oversampling ratio.

9 Appendix B: Sensor Table

---

[2]Given $\Delta\lambda = 0.5$um, $L = 100$ W/m^2-sr-um, this equates to $\dot{N}_{e^-}^{std} = 1 \times 10^7$ e$^-$ / s or 5,000 photo-electrons accumulated in a 1ms integration time on a $Q = 1$, $QE = 0.5$ system

Model	Type	Shutter	Width	Height	FPS	Pixel Size (um)	QE (525 nm)	N_{rd}	$N_{e^-}^{FWC}$
Sony ICX424	CCD	Global	648	488	84	7.4	46	12.9	13932
Sony ICX424	CCD	Global	648	488	84	7.4	53	12.0	13701
Sony ICX414	CCD	Global	648	488	90	9.9	39	19.4	25949
Sony ICX618	CCD	Global	648	488	120	5.6	70	11.7	14508
Sony ICX693	CCD	Global	808	608	50	6.0	71	11.2	20024
Sony ICX204	CCD	Global	1032	776	31	4.7	42	12.1	11944
Sony ICX692	CCD	Global	1288	728	30	4.1	72	8.6	11551
Sony ICX445	CCD	Global	1288	964	30	3.8	62	10.3	9686
Sony ICX445	CCD	Global	1288	964	30	3.8	66	9.2	9196
Sharp RJ33J4CA3DE	CCD	Global	1288	964	30	3.8	70	5.4	7384
Sony ICX445	CCD	Global	1288	964	30	3.8	68	10.1	9231
Sony ICX445	CCD	Global	1288	964	31	3.8	61	7.6	7347
Sony ICX267	CCD	Global	1384	1032	18	4.7	53	11.5	10366
Sony ICX825	CCD	Global	1384	1032	45	6.5	73	8.3	22856
Sony ICX285	CCD	Global	1384	1036	30	6.5	56	11.9	16408
Sony ICX274	CCD	Global	1624	1224	15	4.4	59	8.3	7969
Sony ICX674	CCD	Global	1920	1440	26	4.5	67	9.4	14693
Sony ICX818	CCD	Global	1928	1448	13	3.7	76	10.5	10936
Sony ICX687	CCD	Global	1928	1448	15	3.7	79	9.7	11586
Sony ICX687	CCD	Global	1928	1448	26	3.7	68	10.2	9039
Sony ICX808	CCD	Global	2024	2024	18	3.1	75	9.3	6459
Sony ICX655	CCD	Global	2448	2048	8	3.5	60	9.4	5856
Sony ICX625	CCD	Global	2448	2048	15	3.5	58	8.7	6168
Sony ICX625	CCD	Global	2448	2048	15	3.5	57	8.2	5903
Sharp RJ32S4AA0DT	CCD	Global	2448	2048	7	3.5	57	5.5	8086
Sony ICX694	CCD	Global	2736	2192	13	4.5	73	10.5	14446
Sony ICX694	CCD	Global	2736	2192	25	4.5	74	10.9	14227
Sony ICX694	CCD	Global	2736	2192	13	4.5	72	10.9	14959
Sony ICX814	CCD	Global	3376	2704	9	3.7	75	9.4	9996
Sony ICX834	CCD	Global	4240	2824	7	3.1	78	10.9	6125
Kodak KAI-29050	CCD	Global	6576	4384	4	5.5	43	12.0	20000
Kodak KAI-47051	CCD	Global	8856	5280	7	5.5	43	10.0	20000
On Semi KAF-50100	CCD	Global	8176	6132	1	6.0	62	12.5	40000
Aptina AR0134	CMOS	Global	1280	960	52	3.8	77	6.6	5542
e2v EV76C560	CMOS	Global	1280	1024	60	5.3	59	25.1	8384
e2v EV76C560	CMOS	Global	1280	1024	60	5.3	61	25.3	7506
ON Semi VITA1300	CMOS	Global	1280	1024	150	4.8	61	26.3	10226
ON Semi PYTHON 1300	CMOS	Global	1280	1024	149	4.8	59	9.3	6057
ON Semi PYTHON 1300	CMOS	Global	1280	1024	170	4.8	59	9.2	5779
Sony IMX035	CMOS	Rolling	1328	1048	120	3.6	77	6.0	15491
e2v EV76C570	CMOS	Global	1600	1200	60	4.5	66	24.2	7788
Sony IMX174	CMOS	Global	1920	1200	162	5.9	76	6.8	32513
Sony IMX249	CMOS	Global	1920	1200	41	5.9	80	7.1	33105
Sony IMX252	CMOS	Global	2038	1536	121	3.5	76	2.3	10482
Sony IMX265	CMOS	Global	2048	1536	55	3.5	71	2.9	9777
CMOSIS CMV4000	CMOS	Global	2048	2048	90	5.5	53	16.8	7620
Sony IMX264	CMOS	Global	2448	2048	35	3.5	69	2.3	9869
Sony IMX250	CMOS	Global	2448	2048	75	3.5	76	2.4	10361
Aptina MT9P031	CMOS	Rolling	2592	1944	13	2.2	63	7.6	6693
Sony IMX255	CMOS	Global	4096	2160	43	3.5	71	2.4	10435
Sony IMX253	CMOS	Global	4096	3000	30	3.5	72	2.4	10563
BAE CIS2521	CMOS	Global	2560	2160	50	6.5	55	5.0	30000
BAE LTN4625a	CMOS	Global	4608	2592	120	5.5	55	5.0	40000
On Semi Python 25k	CMOS	Global	5120	5120	80	4.5	50	14.0	12000
CMOSIS CMV12000	CMOS	Global	4096	3072	132	5.5	45	12.5	13000
CMOSIS CMV20000	CMOS	Global	5120	3840	30	6.5	45	8.0	15000
CMOSIS CMV50000	CMOS	Global	7920	6004	30	4.6	45	8.8	14500
Sony A7R	CMOS	Rolling	7392	4920	5	4.9	50	4.3	49000
GPIXEL 3005	CMOS	Rolling	30000	5000	10	5.5	50	3.9	23000
GPIXEL 1205	CMOS	Rolling	12000	5000	10	5.5	50	2.9	22000
GSENSE 5130	CMOS	Global	5056	2968	67	4.2	50	5.0	16000

Table 11: All sensors considered in analyses