Digital imaging systems

Linus Falk

January 23, 2023

Contents

1 Course concepts

The aim of the course is to give a unified perspective on the variety of digital imaging technologies that have been developed the last decades. Different aspects of the imaging technologies will be discussed such as:

- What they are imaging
- How
- With what quality
- For which applications

After this course you should be able to

- Describe the **physics** and **techniques** behind modern imaging techniques
- Describe the basic principles for sample preparation in relation to the imaging technique.
- Reason and analyze around possibilities and limitations in resolution with regards to:
  - density
  - space
  - time
  - spectrum
- Describe how the techniques affect the image and subsequent interpretation and analysis
- Reason about suitability of different imaging techniques in combination with image processing and machine learning for different applications.

The subject of imaging is interdisciplinary and cover a lot of subjects:

- Physics
  - Optics, wave propagation
  - Solid state sensing principles
- Electronics
  - Circuit designs
  - Sensor technology
  - Signal processing
- Mathematics
  - Geometry
  - Fourier analysis
- Computer science

No one is an expert on all imaging technologies and the course therefore consists of lectures of several guest lecturers. They will try to keep it in the common structure.

2 What is an image

An image is a multidimensional sample of the reality that consists of:

$$D = F(x, y, z, w, t) \tag{1}$$

- Densitometry **D**, signal intensity
- The spatial dimensions **x,y,z**
- Spectral dimension **w** – wavelength
- time **t** - the temporal dimension

We don't have any 5D sensors so we need to **multiplexing**. Since the subject is digital images, each dimension and the function value must be **quantized** into a limit range of discrete values.

The densiometric aspect **D** or intensity, what physical property is being imaged? is it a reflection, transmission of light, a density distribution of a molecule, a surface topology or an elastic property of an object? How well can we describe this property? and what physical effect do we use to measure: photo resistance, inducted charge or photon counting?

We create images from signals that can be of different types: electromagnetic waves (light, thermal, x-rays), pressure variations (sound) or contact forces (Braille).

We can use more than only the visible part of the electromagnetic spectrum, we can use all of it with different techniques.

Emission, excited emission, transmission or reflection

Imaged physical properties can be categories into : **Emission** (Astronomy, Autoradiography), **Exited emissions** (Flourescense), **transmission** (light microscopy, film scanning, classical x-rays) and **Reflections** (Normal photography, Document scanning, satellite sensors)

Consider where/what is the light source. In case of **emission** we have a well defined spectral properties, coming directly from the source. **Transmitted light** on the other hand is exponentially absorbed with a logarithmic intensity that is directly proportional to the absorbing matter. In the case of **reflection** is the surface orientation as well as the material properties and the direction and the spectral characteristics of the illumination that determines the signal. There is a need to differentiate between diffuse and specular reflection.

The illumination can be controlled in the case of a **Active** sensor system. This can be done either all at once for the whole scene or by scanning a pixel or a line at a time. An important example of a **Active** sensor system is: **LIDAR:** "laser imaging, detection and ranging".

Densitometric aspects: Resolution

It is important to consider what the densitometric resolution is and what is the signal to noise rate SNR. What pixel depth can we get or in other words how many greylevel do we get and are all of them meaningful. And what is the actual property that is measured?

- Material density
- Density
- Energy
- Photon count
- Topographic elevation

The contrast resolution is another aspect, the image should have a correct exposure time such that we use the most of the dynamic range of the sensor, not blowing up the brighter parts or under expose the image not showing any details in the darker parts of the image. Using most of the dynamic range of the sensor should result in white noise in the least significant bit. The optimal use of bits is therefore where we have about 1 bit of noise. If the conditions for imaging allows it, it can be possible to increase signal to noise ratio by taking the average of multiple exposure according to:

$$\sqrt{\text{number of exposures}} \tag{2}$$

We also need to have in mind the spatial consistency and ask: Will all positions in the image give the same density value for the same signal, there could be random or systematic variations where systematic variations can be caused be for example defects or imperfections in the instrument/sensor that can be corrected for by calibration or other methods. One tool for correcting imperfections are: shading correction.

Imperfections is not only present considering spatial consistency where different part of the sensor register differently, we can have non linearity behavior of the registered sensor that must be corrected for with calibration. We need to know if the grey value is linearly or logarithmically related to the physical property we are interested in. In the field of normal photography is the intensity registered by our sensor linearly related to the reflected light and in transmission imaging is the light absorption logarithmically related to the amount of material the light passes through. If calibration is needed then it will be of importance to investigate of stable it is over time.

The spatial dimension (x,y,z)

Questions regarding the spatial dimensions to have in mind: How are the spatial dimensions **x,y,x** mapped into the image? Is the image a slice, a projection a depth map or something else? Are there any distortions that make the image not geometrically correct? What is the spatial resolution. Is it possible to get more than 2 dimensions with the available technology?

The spatial dimension can be mapped as **projection** that gives a 2D image of reflections from visible surfaces (in 3D) for the sensor or a transmission through the object. A **distance** image give explicit information as seen from a single point (2 1/2 D). With a **slice** we select a slice from a volume. **Tomographic**

reconstruction is a method to compute information about the internal density structure using measurements of numerous line integrals.

The spatial resolution is limited in different ways in analogue and digital images. In Analogue there are constraints of the aperture of the lens and the wavelength of the light while in digital images the limiting factor is the sampling according to the sampling theorem. Under sampling can give problems such as aliasing and when it is caused be poor sampling the result is often worse then when it is because of limited resolution. It is therefore common to have low pass/blurring filters to prevent sharper images than can be digitized. There have been recent inventions that describes the ways of going beyond the resolution limits. example??

Distance images is a wat of representing 3d in 2D by making measurements of the distance to the surface of the object to the sensor for all points in the image. This can be done using either **passive** sensors with: Parallax camera or stereo images. With an **active** it can be done with the measurement of time of flight (lidar, radar, ultrasound, laser) or with triangulation, structured light.

Creating 3D images or image volumes is a rapid growing area in digital imaging mostly driven by medicine. In contrast to the 2D images these can not be viewed directly and it's necessary to use special visualization efforts to interpret the result. These types of imaging system generates very large data sets. Imaging volumes can also be done be psychically slicing (and destroying the sample in the process) and the result would be a 2D image of this thin slice. Another category of volume imaging is **Tomograpy** that includes: x-ray(CT scan, magnetic resonance (MRI), Emission (SPECT and PET), electron microscopic (EMT) and Optical coherence (OCT what is this?)). Confucal microscopy, ultrasound and holography are also volume imaging techniques.

Reconstructed images like tomographic reconstruction is a type of inverse problem with multiple dimensions that estimates a system from a finite number of projections. Examples of where this technique is used are:

- Transmitted X-rays, Computer Tomography (**CT**)
- Radioactive decay, Emission Tomography
  - PET
  - SPECT
- MRI, emitted excited radio frequency
- SAR (synthetic aperture radar)
  - CARABAS - long wave - incoherent

The spectral dimension w

In all imaging we need to limit the spectral range we are imaging by choosing a range in a wavelength interval. Different spectral ranges often give different image contrast. The signal in each of the pixels is the result of a convolution between:

- The spectral distribution of the illumination

- The spectral absorption/reflection properties of the object.
- The spectral sensitivity function(s) of the sensor.

The visual perception by the human eye is the result of the application of the three different spectral sensitivity ranges of the cones. Most imaging system are designed to be optimized for human color reproduction fidelity. There are many different illumination and reflection functions that can give the same color experience.

For capturing a color image we need three spectral samples which can be captured in different ways where the **Bayer filter** is the most common. This pattern leads to loss of resolution but is restored by interpolation of some sort. The alternative to having the **Bayer filter** is to have three sensor chips each color, splitting the image with a prism. This solution is expensive and requires also high mechanical precision. A third option is to stack the three sensing layers for the different color channels which have the benefit of not losing resolution/not need of interpolation.

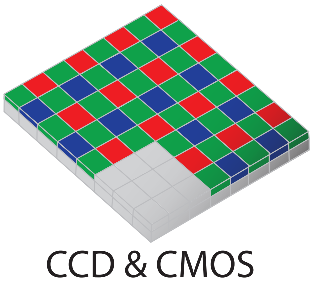

Figure 1: Bayer filter

What limit us to only use three spectral channels in image analysis is conventional thinking and that there are many cost effective camera in this range. When we register more than one spectral channel we need to use multiplexing, switching between channels either spatially or temporally. Ultimately it should be the application that decides how wide the spectral range should be and the number of channels. It is possible to achieve spectral imaging in a method similar to the Bayer filter, there are cameras with 9 (3×3) visible light channels and 16 (4×4) infrared channels that together can give us 25 channels simultaneously. As the number of channels increase to several hundred we call it **imaging spectroscopy** or hyperspectral imaging. This type of imaging creates a large amount of data and needs effective transmission and compression.

A solution to capture a large number of spectral channels is to have a rotating filter wheel in front of a wide band, single channel sensor/sensing chip. This requires the scene to be stationary since there will be several images registered.

Temporal aspects (t)

Each image will register something for defined time interval and if there is only one single temporal slice we have a still image, while if there are multiple exposures we get a film/movie. For each captured image or exposure we need to define the exposure time, the following questions should be considered:

- Does it give motion blur, how fast is the scene/object moving?
- Can it be varied freely?
- How will the quality of the image be affected by the exposure time? high ISO sensitivity more noise for example.

During the exposure is each pixel, line or image exposed for a certain time, so we need to consider the motion in the scene and of the camera relative to it and the light intensity can by limited. There are however ways to freeze a moving object in the image by using a flash or follow the object in motion with the camera. There are even special solutions with sensors that have electronic object following.

Using sequences of images can be useful when measuring a motion. It is then crucial with timing:

- Repetition time
- Exposure time
- Data transfer and storage time
- Influenced by resolution in all five dimensions.

Sequences can also be used to detect changes in scenes by subtracting the previous stored reference image. For a sequence to be perceived as continuous by the human visual system it is necessary to have a repetition frequency of at least 25 Hz.

Multiplexing is needed since the intensity must be measured for all pixels to get a complete image matrix and we have multiple spectral channels (color RGB) to consider in the integration process. There are no effective sensors that can capture all dimension at once and we need to multiplex the light collection. The amount of parallelism in this light collection is an aspect imaging system and is strongly influenced by how the light is handled (economically) by the system.

Area integration

Area integration in imaging technology that uses a 2D image sensor to register the light for the whole image in parallel is by far the most common. It started of with image tubes and was later replaced with **CCD** (Charge Coupled Device) that are now being replaced with **CMOS** (Complementary Metal Oxide Semiconductor) that could be replaced by new technologies like QIS (photon counting devices) in the future.

Area integrations gives the best light collection efficiency since it collects "all the light". It also have a rigid geometry and gives stable and predictable imaging geometry that don't suffers from a lot of distortions. There is no need for mechanical motions and they can be mass-produced and therefore inexpensive. The drawbacks with area integration is that it need an even illumination of the whole image surface and there may be varying sensor sensitivities that need to be corrected for. It is also difficult to achieve high fill factor since there are other things that competes for space on the 2D area, multispectral scanning can therefore also be hard to achieve. The image size will be limited by the sensor size.

CMOS and CCD matrices/sensors are available in many different version with a variety of megapixel resolutions for special applications. In some applications they can be cooled to very low temperatures that allows for long integration times when the light levels are very low. In contrast they can also be used to capture fast events like laser flashes.

Linewise integration

This kind of sensors are often find in scanners and can be moved to capture an image of the object, or the object can be moved over the sensor. It is also common in remote sensing like satellites that move along the earth surface.

The technical advantages with line integrations are much better pointwise integration but worse economy compered to area. In motion the orthogonal line will be "frozen" which can be useful in some applications. It allows the use of the other dimension in a 2D sensor for the wavelength, for RGB and hyper spectral scanners ? read more. It possible to create whats called "intelligent sensors" by having a processor for each pixel. The disadvantages are similar to area integration, it needs an even illumination along the scan line, it needs corrections of sensor sensitivity along the line of pixels. The 1D fill factor is important but not often a problem meaning?. There is also a risk for x-y inhomogeneity because of the widely different methods of scanning.

Point-wise integration

With this methods we register light from one pixel at a time and it requires motion to create the image. Here we most distinguish between what is moving:

- The illumination
- The sensor
- The object

The method is mainly used in stationary conditions such as microscopy or scanning film or paper/document. Examples of applications are Drum scanner which produces very high quality scans of document or film. Other applications:

- Flying spot scanner
- Microscopy
  - Fluorescens
  - Confocal
  - Multi-photon
  - Moving stage

Some of the advantages are that it gives maximal possibilities for optimization of the measurement of each pixel and that i can have optimized optical path and sensor. No differences between the sensor properties of the different pixel sensors, i.e. only one pixel. It also have the advantage of not having a limit to the image size. The disadvantages are that it uses the incoming light poorly and need some sort of complex mechanical system for scanning which also results in it being very slow compared to the other methods.

Multiplexing for volume imaging

A few words about multiplexing and volume imaging; It is possible to register single voxels, a line, a plane at a time, the whole volume in parallel. Collecting data in the Fourier domain or through other transforms can be done. Today most tomographic systems collect data from the whole volume or from multiple planes simultaneously which is fast and therefore saves time and signal economy.

Wavefront imaging

Is new form of imaging that have high information density. The Wavefront sensing measure the amplitude and phase of the incoming optical field simultaneously. It is still under development but show promising results.

Stored intermediate analogue image

This category is mainly of historic interest and a few example are listed here:

- Photographic film
- Polaroids
- Magnetic tape
- Semiconductor materials (image plates)

3 Photography

The focus of this course are ways of creating images for scientific or medical applications but its worth noting that a majority of all digital images created are digital photography. Digital photography has grown tremendously the last 20 years and analog photography is now limited for some small applications or enthusiast.

Description of digital photography

Digital photography is an optical imaging technology and the aspects of it are the same as for analogue cameras: Lens quality, aperture, depth of focus and focal length.

The lenses come in large variety of sized and qualities and it should be chosen with the application in mind. Some basic guidelines are that: bigger is better and glass is better than plastic. Compact lenses has become better and there is a trend towards more compact lenses such as liquid based lenses that can change focal length and focus very rapidly.

The depth of focus depends on the lens focal length and aperture. A smaller aperture will result in a larger depth of focus but less light is reaching the sensor and therefore need longer exposure time.

Choosing between a fixed focal length or a zoom lens is depending on the application. For general hobby photography is a zoom lens often to prefer since it doesn't require the user to change lenses when the scene or application changes. Zoom lenses are in general not as good as a fixed lens but instead offer the flexibility. Some geometric distortions to the resulting image is caused by the lens and takes the form of **Pincushion** or **Barrel** distortion which can be corrected for in image processing.

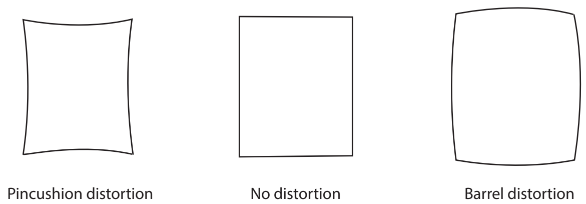

Figure 2: Example of caption

There are two types of sensors domination in the area of digital photography: **CCD** and **CMOS**. CCD sensors are matrices of photosites each comprised by a photodiode which converts light into a charge and region that can hold the charge. The charges are shifted/moved out of the sensor as a bucket brigade and converted to a digital signal at the end of the circuit. In CMOS sensors similar to the CCD are there photodiodes that converts the light into electrons/ a charge but there is also a reset and select transistor with an amplifier section. This mean that the amplification is done at each pixel.

In the **full frame** CCD sensor are the charge holding region integrated with the light sensing region. When the light is collected is happen over the entire imager but it needs to be shut off after that so the charge can be moved through

CCD	**CMOS**
Power consumption 2-5W	Power consumption 20-50mW
More light sensitive approx 1 lux	Less light sensitive approx 5 lux
Less digital noise	Uses same silicon design as other electronics
Better light sensitivity, up to 85% quantum efficiency	Any subset of pixels can be read out
More color depth, more dynamic range	Younger technology
Requires special chips with higher voltage, higher cost	CMOS pushing CCD:s out of the market

Table 1: Comparison CCD vs CMOS

a horizontal charge transfer register a charge voltage conversion, amplifier and analog to digital (A/D) converter. This results in a almost 100% fill factor of the sensor but there is need for an external shuttering. In the **interline** type of CCD sensor are the charge holding region shielded from the light meaning there is no need for a external shutter to keep light out. This type of sensor architecture results in lower fill factor but can be compensated with micro-lenses what is that?. A third type of CCD sensors are **frame transfer** sensors which have like the **full frame** photosites that cover almost all of the sensor/sensing area. The difference is that the charges are shifted quickly to a equally sized charge holding region that are shielded from the light. Thanks to the high speed is no shutter need and the readout can be done while a the light for a new image is captured.

The **CMOS** sensor is built with the same technique as processors and memory arrays. This architecture allow for readout of the entire array or only parts of it with a simple X-Y address. The fill factor is decreased compared to the full frame CCD since there is more electronic in each pixel. This can be compensated by either Micro-lenses or using thin sensors that can be exposed from the back side. CMOS sensors can ha

For a long time have megapixels been a selling point for sensor, the pitch has been: more is better. But this is not always true. A small sensor size with a high number of pixels can lead to decrease in sensitivity and worse signal/noise.

Comparing priorities between consumer and scientific cameras we can see that consumer cameras generally focus on having good looking pictures while scientific camera priorities correct, quantitative pictures. The increasing use of digital photography has pushed on the development of cameras in the consumer segment, this has also resulted in much value for scientific needs also. Although the trend has now shifted towards smaller sensors since most of them are fitted to mobile phones. This creates a divergence between consumer cameras and scientific cameras that are not faced with the same constraints as mobile phone cameras. As today there is still a wide range of digital cameras for scientific applications. The sensors range from a couple of euros to 10 000 euros in price, but a typical camera can be bought for around 1000 euros.

A paradigm shift might be taking place with **Photon counting** sensors being developed. These sensors count each photon separately and is as sensitive as it is possible to be. This allow for new capabilities such as trade off in sensitivity and resolution that can be dependent on the scene. It will allow for motion blur compensation for multiple targets and high apparent SNT for a low photo flux.

Operational steps

We start of with the lens that most often includes a IR blocking filter and an optical anti-aliasing filter. The focal length is often adjustable (zoom lens) and the focus is controlled by the focus motor. Exposure and focus measuring is done by pressing the shutter button halfway. The lens focuses the light from the scene onto the sensor. The analog signal that is produced by the sensor is then converted to a digital signal (A/D converter). Then the shutter button is fully pressed down the the image is captured and stored on in DRAM. To long exposure times can lead to motion blur in the image which can be solved by active image stabilization, either by moving the lens or the sensor. Exposure control can be used to exposing different parts of the image differently to optimize quality of the image.

The final high-resolution image is processed by a digital image processor in the camera. The first step is to **de-mosaicing** since most cameras use the Bayer color filter array, here is interpolation used to fill in the missing color values for each pixel. An algorithm decides if the "missing" color values are in a smooth area or along an edge to determine the value for each missing color value. This process results in a full-color image but not a perfect one. To improve the result there is need for white balancing to compensate for spectral variations in the illumination of the scene. Both daylight from the sun and indoor lightning provides white light is the daylight more "high energy" in the blue portions while indoor lightning is more "high energy" in the red portion of the spectrum. An algorithm is used to analyze the scene and adjust the red and blue signal strength to match the green signal strength in white and neutral parts of the image. Continuing with color correction that is needed since the sensor has a different sensitivity than our eyes and the colors can for the human eye be perceived as unsaturated without correction. The color correction compensate for this and transform the output image to the output color space (often sRGB) that is ready to be displayed on a monitor.

Densitometric

Geometrical

Spectral

Temporal

History

4 SEM - Scanning Electron Microscopy

Description of SEM

The electron microscopy is an instrument that allow us to create images/visualize organic and inorganic structures with impressive magnification.It is an invaluable instrument in the engineering and development of new materials where nano-meter sized imaging is very important. The SEM can also be used in chemical analysis and accessing the crystalline structures of materials.

In comparison to light microscopy is the depth of field much larger in scanning electron microscopy thanks to the narrow electron beam. The resulting images get a 3D appearance due to this large field of depth that is very useful when examining the surface structures.

Operational steps

The idea behind the technique is to generate a beam of energetic electrons with emission from an electron source. This is typically done by heating up a metal in vacuum and accelerate the electrons with a electric field. The electrons are stopped from leaving the atoms of the metal be an energy barrier between the emission tip and the surrounding vacuum. By adding an electric field in the vacuum the energy barrier is reduced to a "slope", but the electrons still need to perform the "work" W to pass this barrier. Luckily there is an quantum effect called tunneling allowing electrons to tunnel through this barrier out into to vacuum.

Picture of tunneling

The energy of the electrons in this beam is measured in 1eV which is equivalent to what a electron in an electric field generated by 1 V would have. This energy is denoted E_0 and is often around 0.1 to 30keV. The electron beam is after acceleration modified/reshaped by lenses and apertures to reduce the diameter of the beam and to scan this beam into a raster of x-y coordinates which it sequentially is placed in. This coordinates are discrete but closely spaced. The lenses are no ordinary glass lenses since the electrons would pass through them or lose to much energy (wouldn't focus them either). Instead are they modified with electrostatic lenses and electromagnetic coils into desired shape and properties.

At each of the raster coordinates in the scan pattern are two types of outgoing electrons from the specimen created: **back-scattered electrons (BSE's)** and **secondary electrons (SE's)**. The BSE's are electrons that emerge from the specimen with a lot of the initial energy after interacting with the atoms electric fields of the atoms in the specimen, scattering and deflection. The SE's are electrons that emerge from the specimen surface after the electron beam have ejected them from the atoms in the sample. These electron escapes with very low energy compered to the typically high energy electron beam. They are in the range of 0-50eV with the majority below 5eV.

Electron-sample interaction, the interaction volume.

The secondary electrons are often measured with a Everhart-Thornley detector that is sensitive for both SE's and BSE's while the BSE's are measured with a dedicated BSE detector that is not sensitive for the SE's. The signal for each of the detectors are measured at each of the coordinates in the x-y raster.

The location and the intensity is recorded and correspond to a gray value in the x-y coordinates. Other sensors can also be used to capture for example X-ray signals that also can be emitted from the specimen due do the electron beam.

Densitometric

In order to create contrast between the different features of the sample there is need for a signal need to distinguish them. The BSE's and SE's give us different information about the specimen.

The secondary electron energy

Geometrical

Spectral

Temporal

History