Imperial College London

Facility for Imaging by Light Microscopy

FILM

Observing Life As It Happens

MICROSCOPY DAY 2011:

Understanding and handling image data

Martin Spitaler

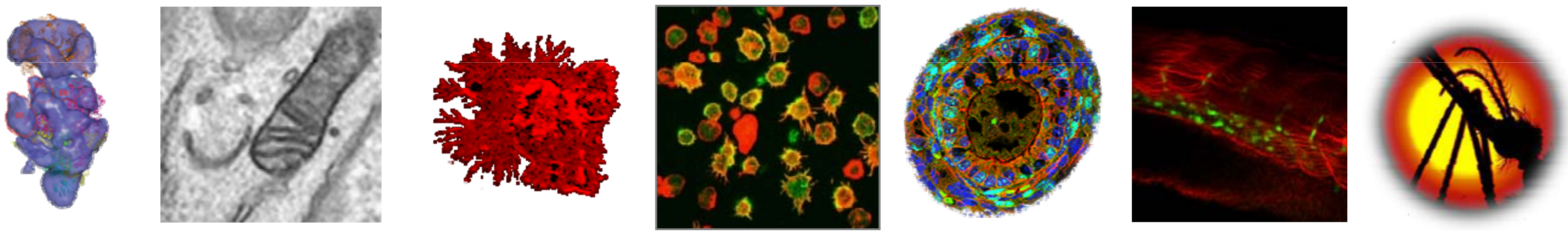
Understanding Images (Martin Spitaler)

- What's in an image file:
  - Pixel data
  - metadata

- Getting the data into the file: Image acquisition
- Image file formats
- Using images: Image visualisation and presentation

Handling Images (Chris Tomlinson & Mark Woodbridge)

- Omero image database
- Xperimenter experiment annotation system



Pixel (or binary) data with information about the sample

- One or more frames of pixels (XY, XZ)
- Each frame typically consists of a two-dimensional array of pixel values
- Pixel values can be:
  - light intensity
  - array of intensities (PALM, STORM)
  - array of fluorescence lifetimes
  - in the future: correlated data, e.g. exposure times (CMOS), mass spectra, …
- Frames are stacked in one or more specific orders:
  - Channel (colour, lifetime, …)
  - Z stack
  - Time
  - XY position in a plate

What's in an image file: Meta data

Meta data make sense of the pixel information

- Image type (TIFF, LSM, CXD, ...)
- pixel dimensions (size, time point, focus position)
- Hardware settings:
  - objective lens
  - excitation light source:
    - type (laser, lamp)
    - intensity
    - excitation and dichroic filters
    - ...
  - emission settings:
    - emission filters
    - detector gain and offset
    - pinhole size
    - sampling speed / exposure time
    - ...

What's NOT in an image file: Experimental data

Meta data that make sense of the pixel information but are NOT in saved in the file:

- sample (cell line, organism, ...)
- coverslip / slide / dish, glass vs. plastic, coating, ...
- sample preparation (live, fixation method, ...)
- labelling (fluorescent protein, antibodies, chemical dyes, ...)
- labelling protocols (blocking, incubation times, concentrations, transfections, ...)
- other sample preparation steps:
  - live:
    - CO_2 / buffer
    - temperature
    - medium / buffer
  - fixed:
    - mounting medium
    - antifade

Pixel data types

- **dimensions:**
  - XY
  - Z
  - T
  - intensity channels (fluorescence, Flim, polarisation, ...)
  - bit depth / dynamic range
    - 8-bit = 256 intensity values
    - 12-bit = 4096 intensity values
    - 16-bit = 65536 intensity values

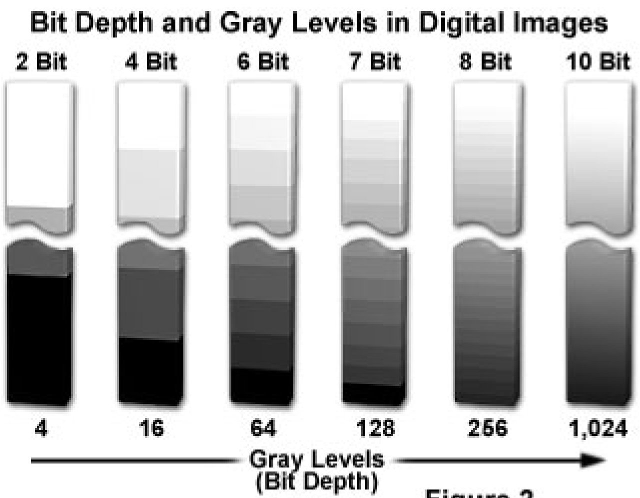


Figure 2

source: microscopyprimer online

File formats: overview

File formats used for microscopic image:

- **Proprietary formats** (Zeiss *.LSM, Leica *.LEI and *.LIF, Volocity, SimplePCI *.CXD)
- **General image formats**
  - **TIFF** (tagged image file format); de-facto standard, best for scientific images
  - **OME-TIFF** (hybrid format containg pixel data in TIFF format and metadata in XML format, contained in the TIFF header)
  - **GIF** (graphics interchange format); lossless for 8-bit single channel, otherwise *very* bad quality
  - **JPG** (Joint Graphics Expert Group); small size, but extremely lossy compression
  - **PNG** (Portable Network Graphic); intermediate size, lossless but slow compression

File formats: TIFF

TIFF (tagged image file format; copyright owned by Adobe (originally Aldus):

Advantage:

- raw data format
- basis for original proprietary formats
- uncompressed or lossless compression (LZW, ZIP and others) (but partially restricted by Adobe copyright)
- can save metadata (like in OME-TIFF)
- can contain multiple channels, stacks, time points, ...

Disadvantage:

- large size

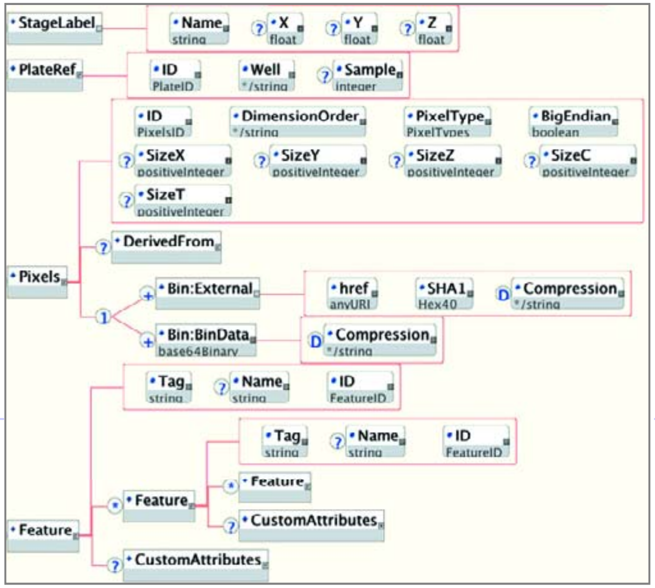
OME-TIFF (Open Microscopy Environment TIFF):

Advantage:

- based on widely supported TIFF format
- open standard
- XML-based metadata format understandable for human and computer interpretation
- universal format, supported by most software

Disadvantage:

- only directly written by few microscopes → needs file conversion



GIF (graphics interchange format):

Advantage:

- lossless compression for dimensions
- ubiquitously accessible
- relatively small
- lossless for single channel 8-bit images (e.g. only GFP), see below

Disadvantage:

- **indexed** format, i.e. only max. 256 values for colour and intensity, very lossy if more than 1 channel

PNG (portable network graphics):

Advantage:

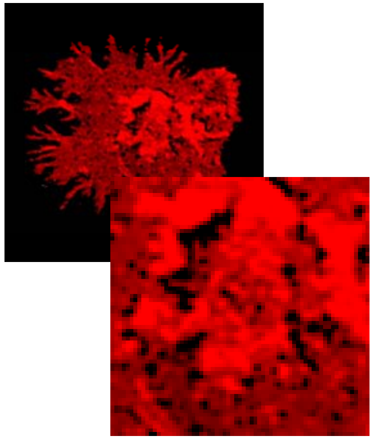
- lossless compression
- ubiquitously accessible
- relatively small

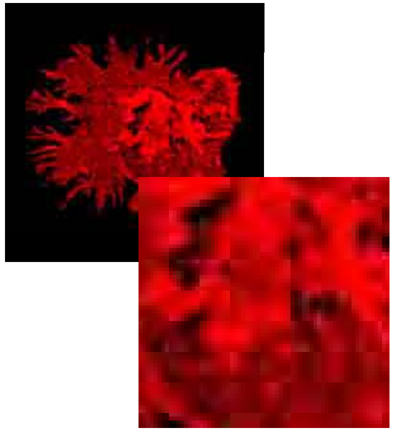
Disadvantage:

- slightly slow compression

File formats: JPG

JPG (joint graphics expert group):





Advantage:

- very small
- ubiquitously accessible

Disadvantage:

- open standard, copyright-free
- entirely based on minimising impact on human perception (except lossless version)
- lossy multistep compression:
    - conversion RGB → YCbCr (Y = brightness; CbCr chrominance)
- downsampling (reducing resolution)
- Block splitting (deviding image into 8x8 blocks, creates artifacts if xy resolution is not multiple of 8)
- Discrete cosine transformation (transforms blocks to a linear combination of these 64 squares)

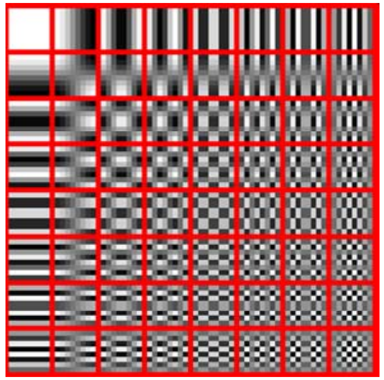


Image visualisation: lookup tables (LUTs)

- **Colours: intensity maps** (Look-up tables)

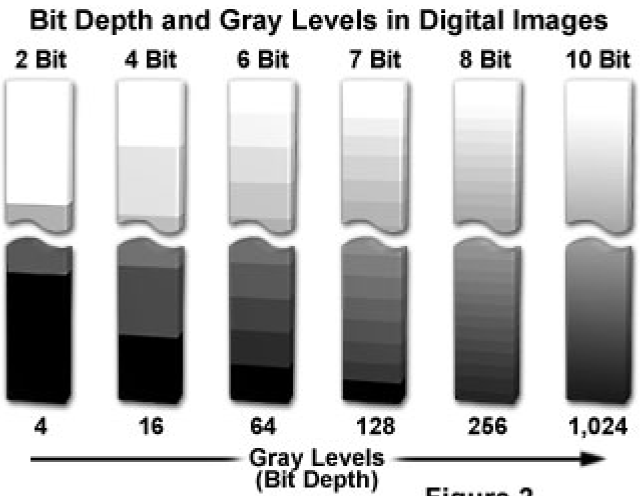
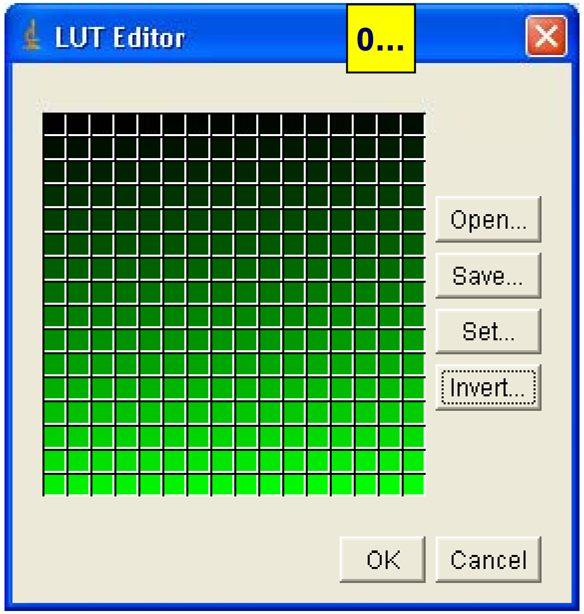


Figure 2



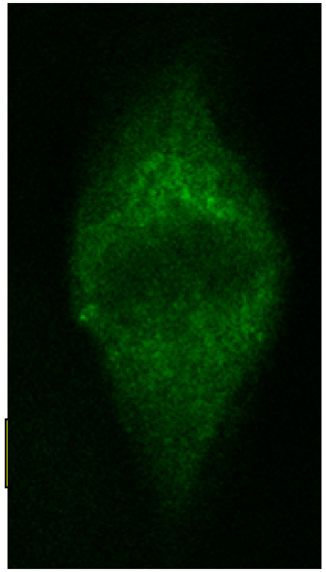
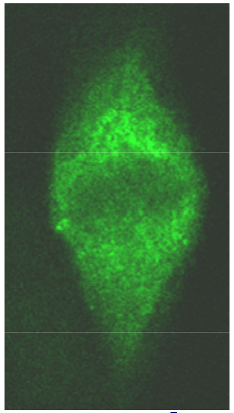


Image visualisation: ideal



Sample

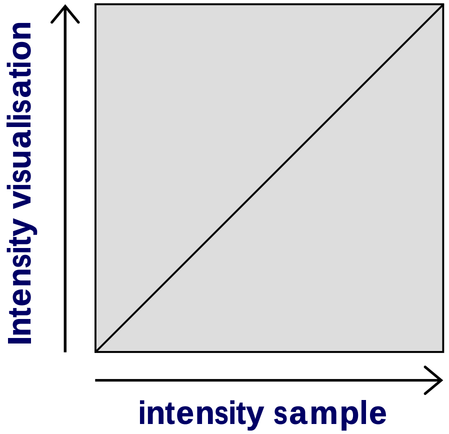
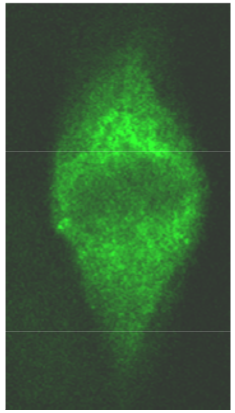
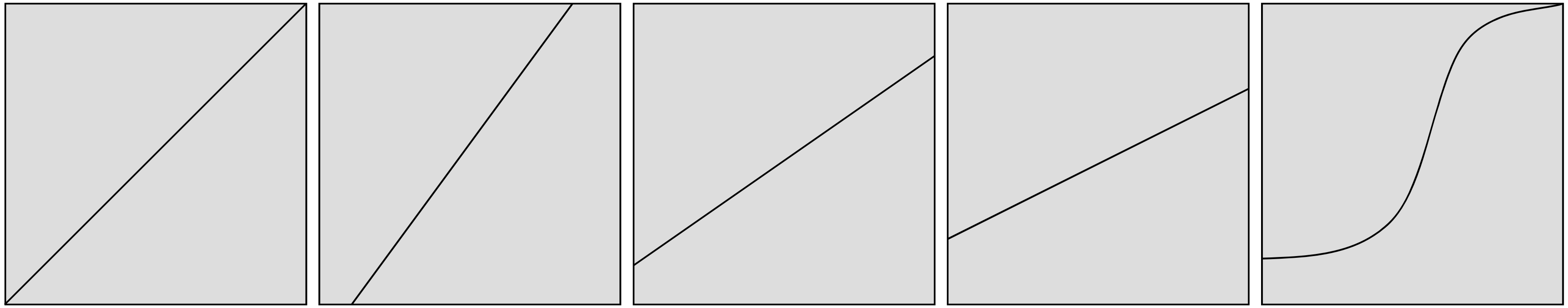


image acquisition



visualisation, presentation, publication

Image visualisation: reality



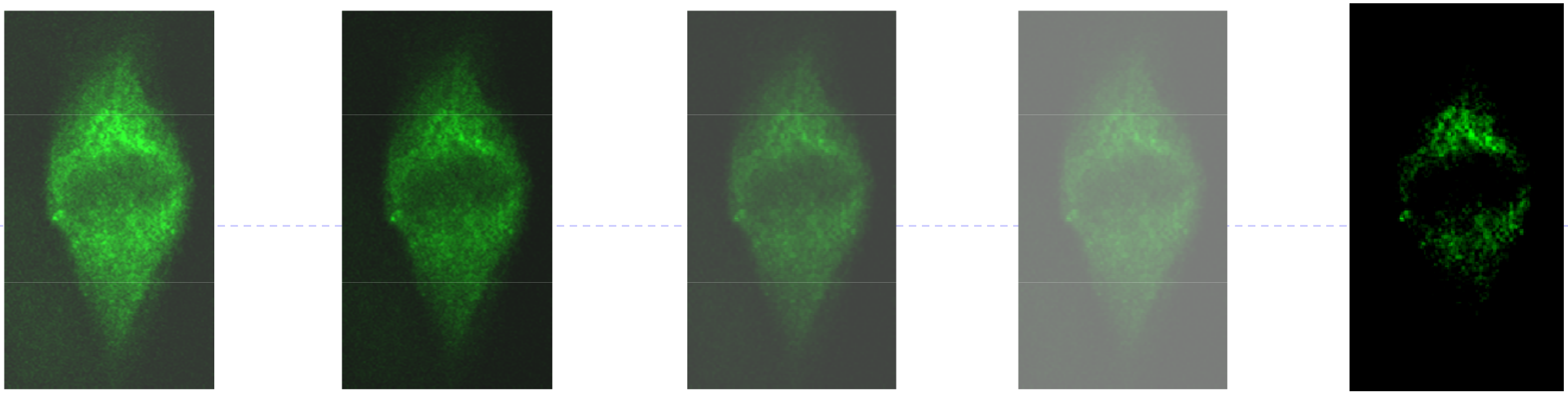
sample

Excitation
(intensity, wavelength, transmission of optics, quantum efficiency, bleaching)

sensor
(eye,PMT, CCD) sensitivity, exposure time, saturation, noise,...)

file
(bit depth, compression)

presentation
(projector, printer, ...) dynamic range, colour balance, resolution



Look-Up Tables (LUTs)

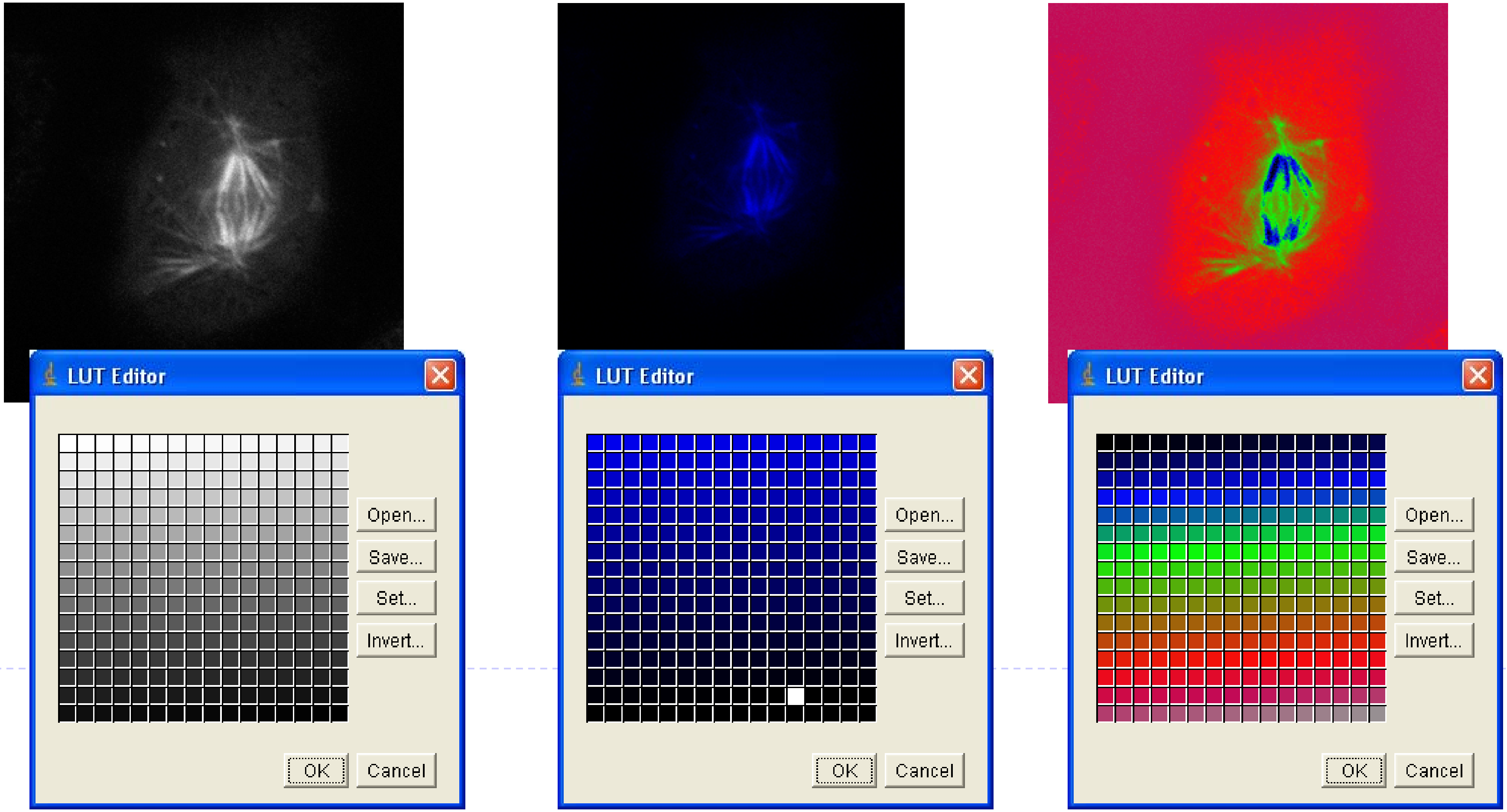


Image visualisation: LUTs

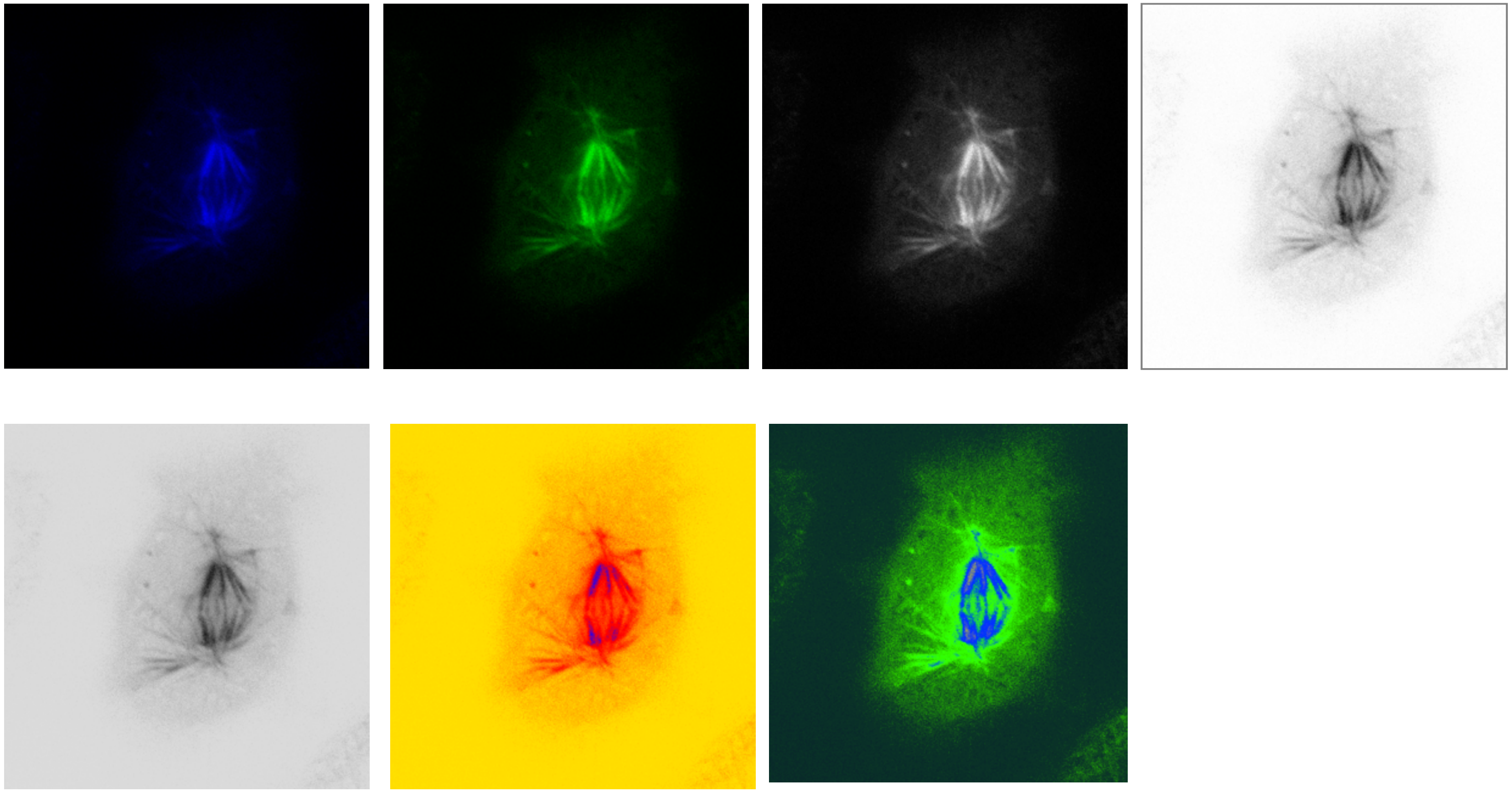
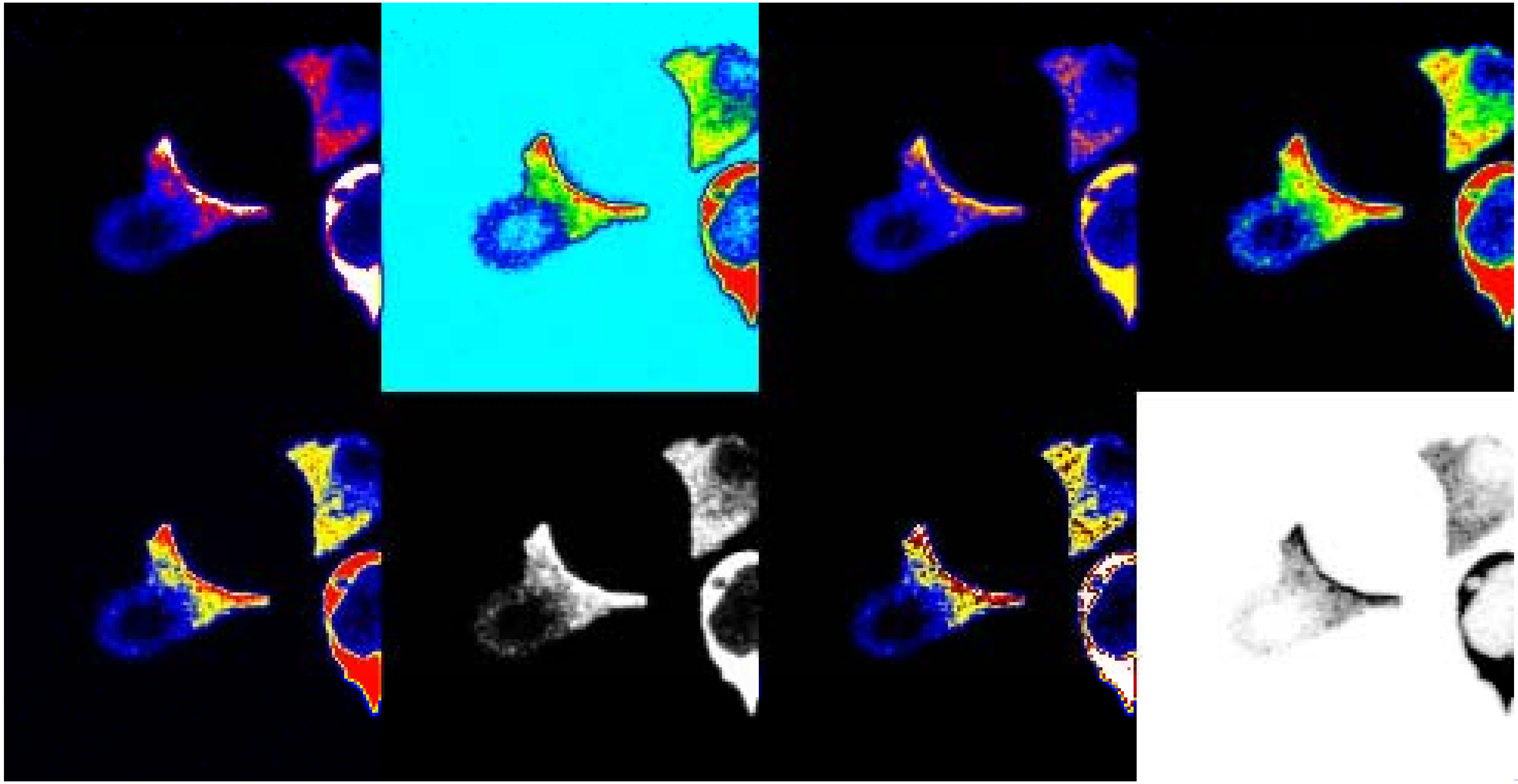
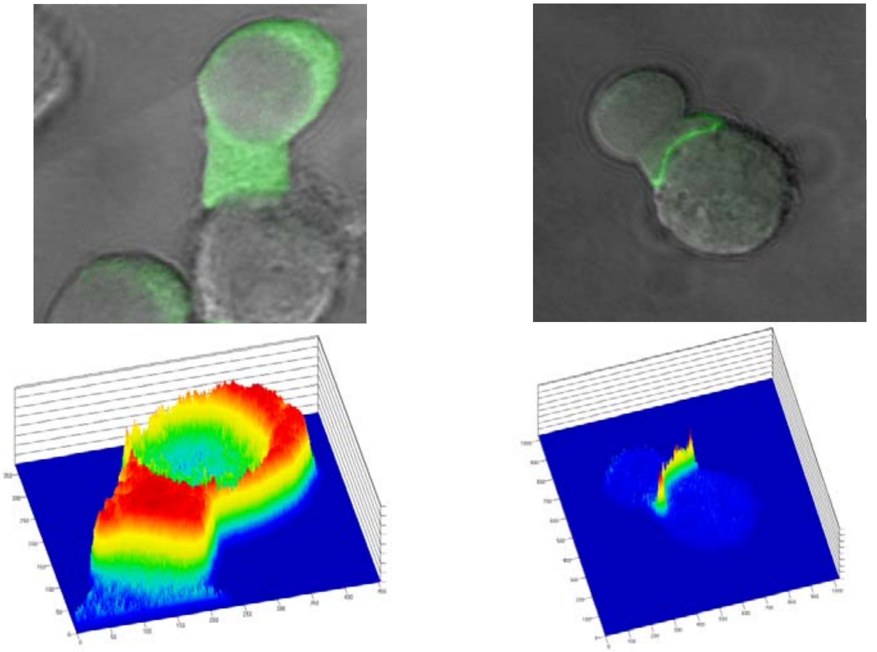


Image visualisation: LUTs





The problem...

single confocal image	1.5 MB
single widefield image	6 MB
XYZ stack	250 MB
XY movie	1 GB
96-well plate screen (XY)	6 GB
XYZ movie	30 GB
PALM / STORM single image	10-30 GB
multi-position XYZ movie	50 GB - 2 TB

the solution!