A bottom-up approach to spectroscopic data reduction

Carlo Izzo, Yves Jung, Pascal Ballester

Major requirements for automatic data reduction

Robustness

- managing unexpected situations
- supplying information about what went wrong
- actual monitoring of the instrument

Flexibility

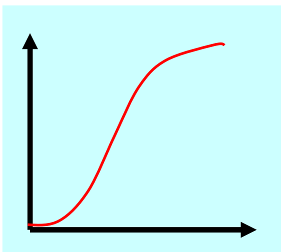
- using general algorithms
- applying instrument-independent DRS strategies
- low maintenance cost

Reference objects are used for data calibration:

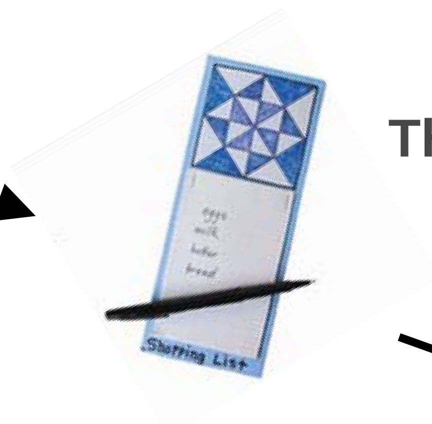
- Stars
  - astrometric calibration
  - photometric calibration
- Spectral lines (arc lamp, sky)
  - wavelength calibration
- Spectral edges (flat)
  - spectral curvature
- Slit positions (pinhole mask)
  - mask-to-CCD transformation
- ...

How to identify reference objects?

The top-down approach:



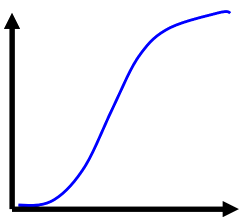
The model



The catalog

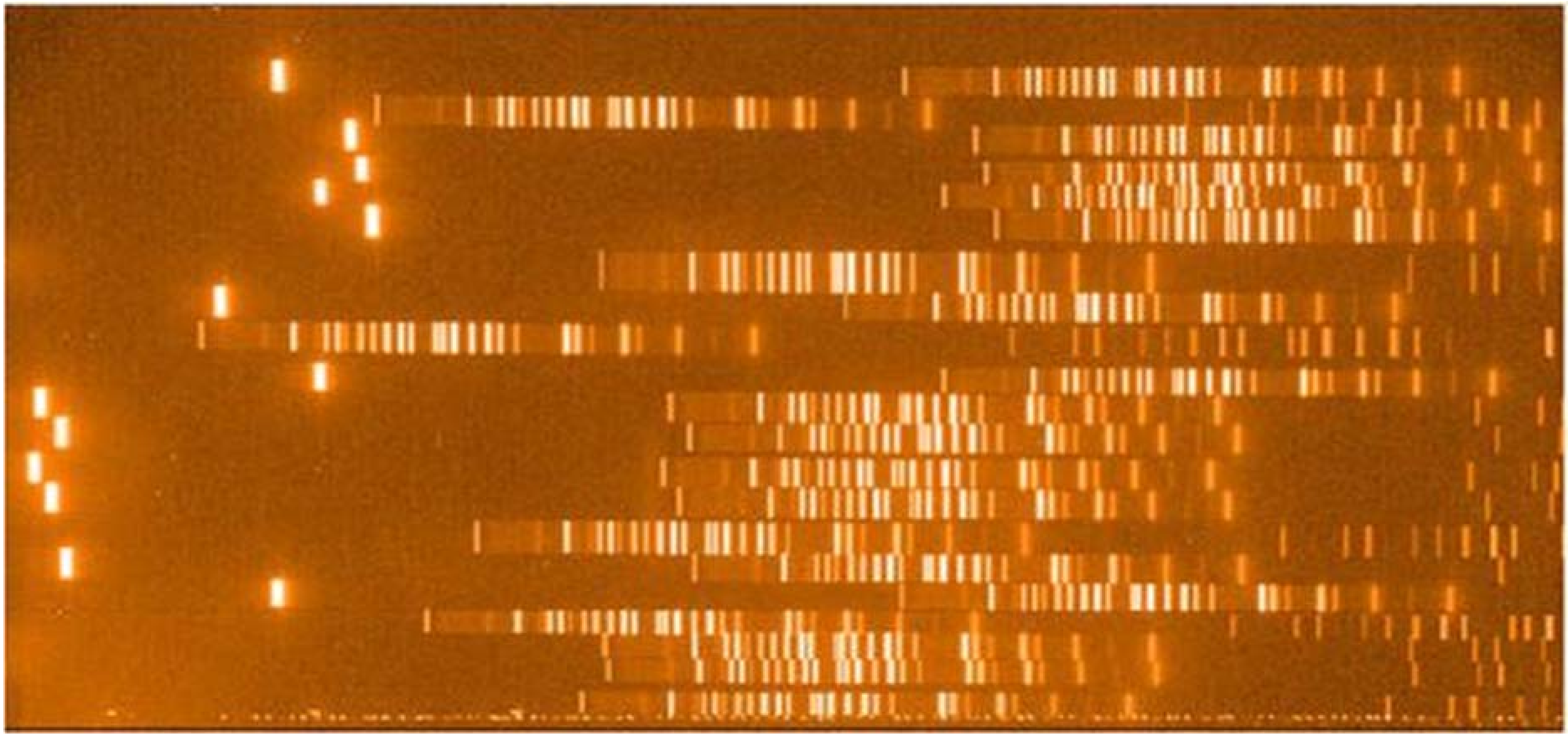
The objects

calibration

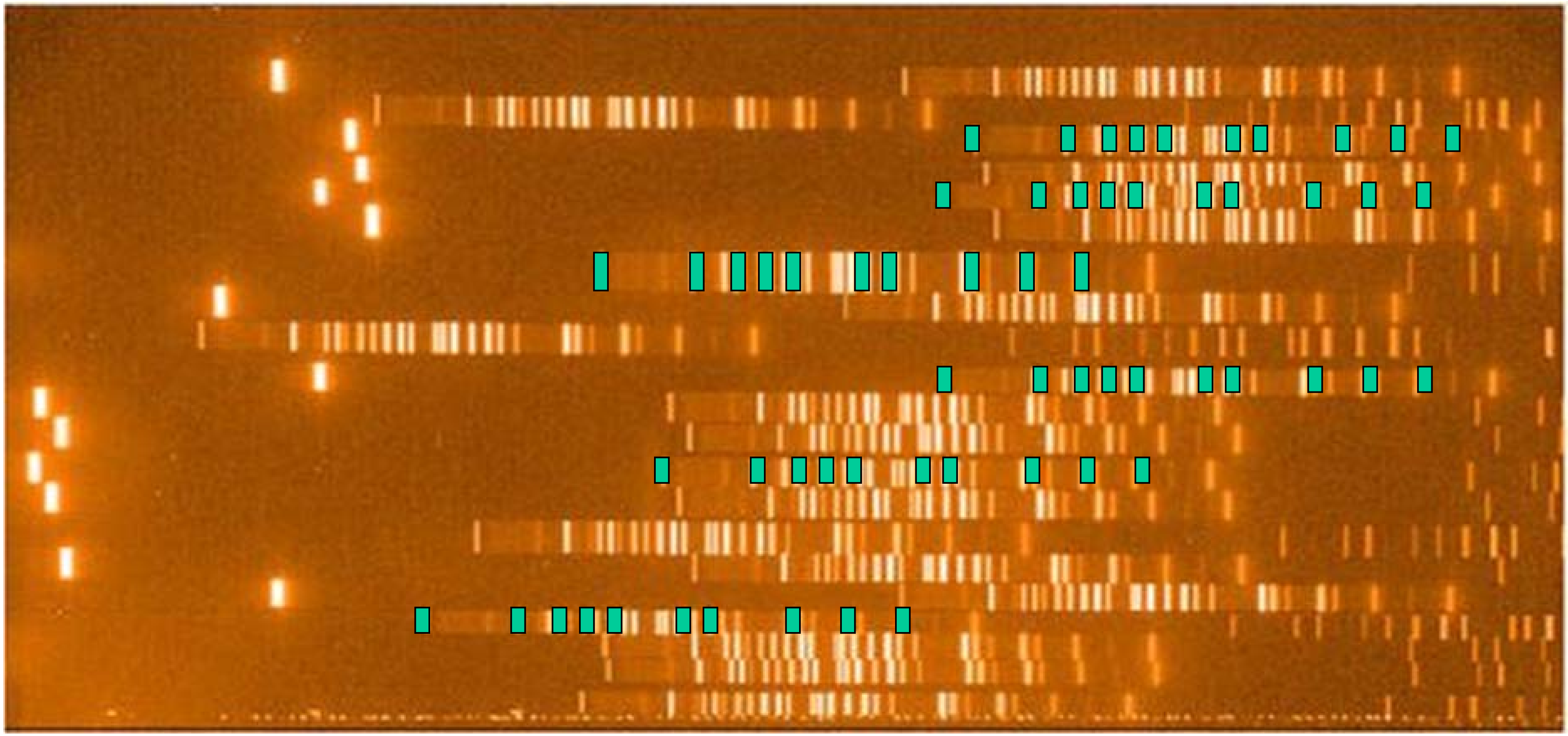


The improved model

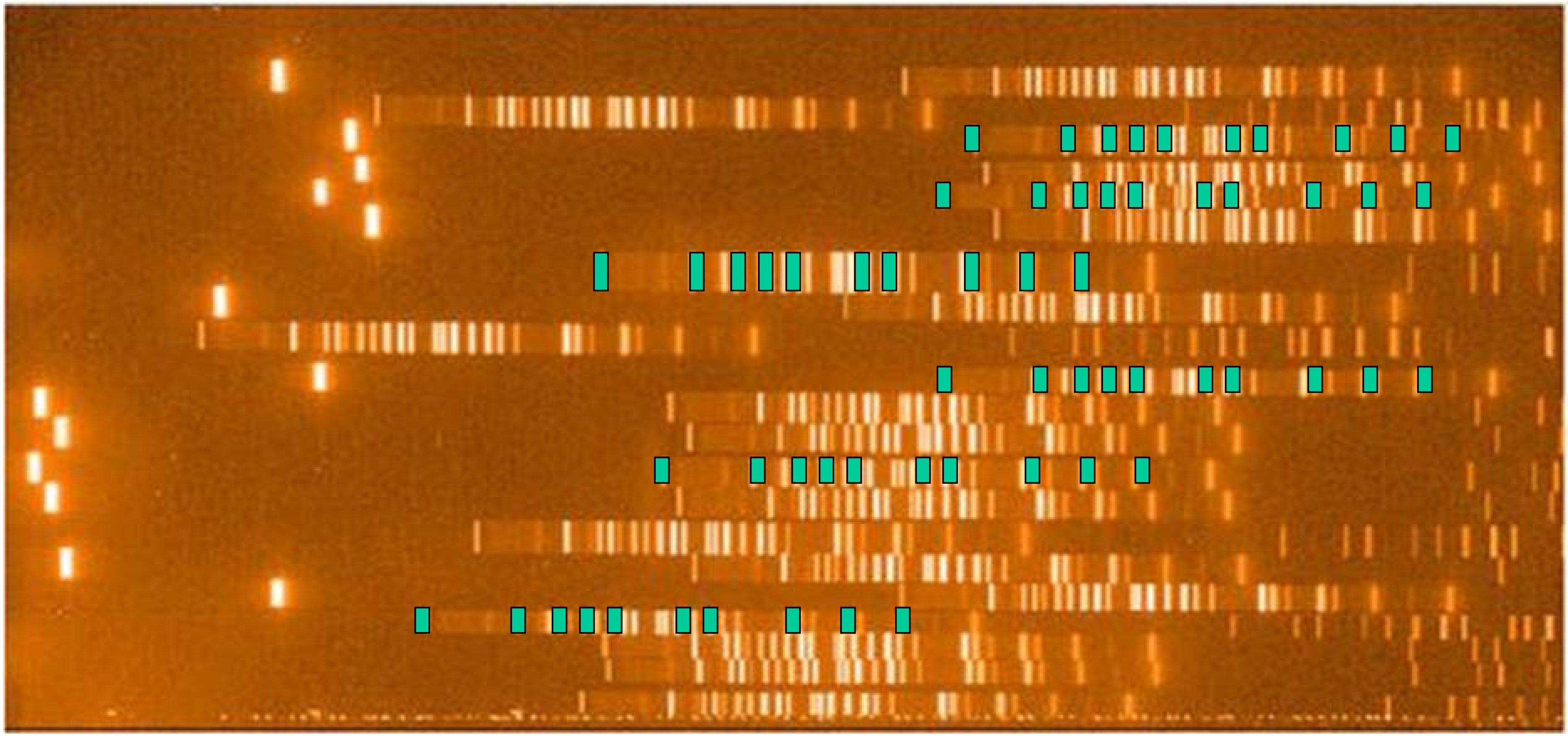
A typical MOS arc lamp exposure



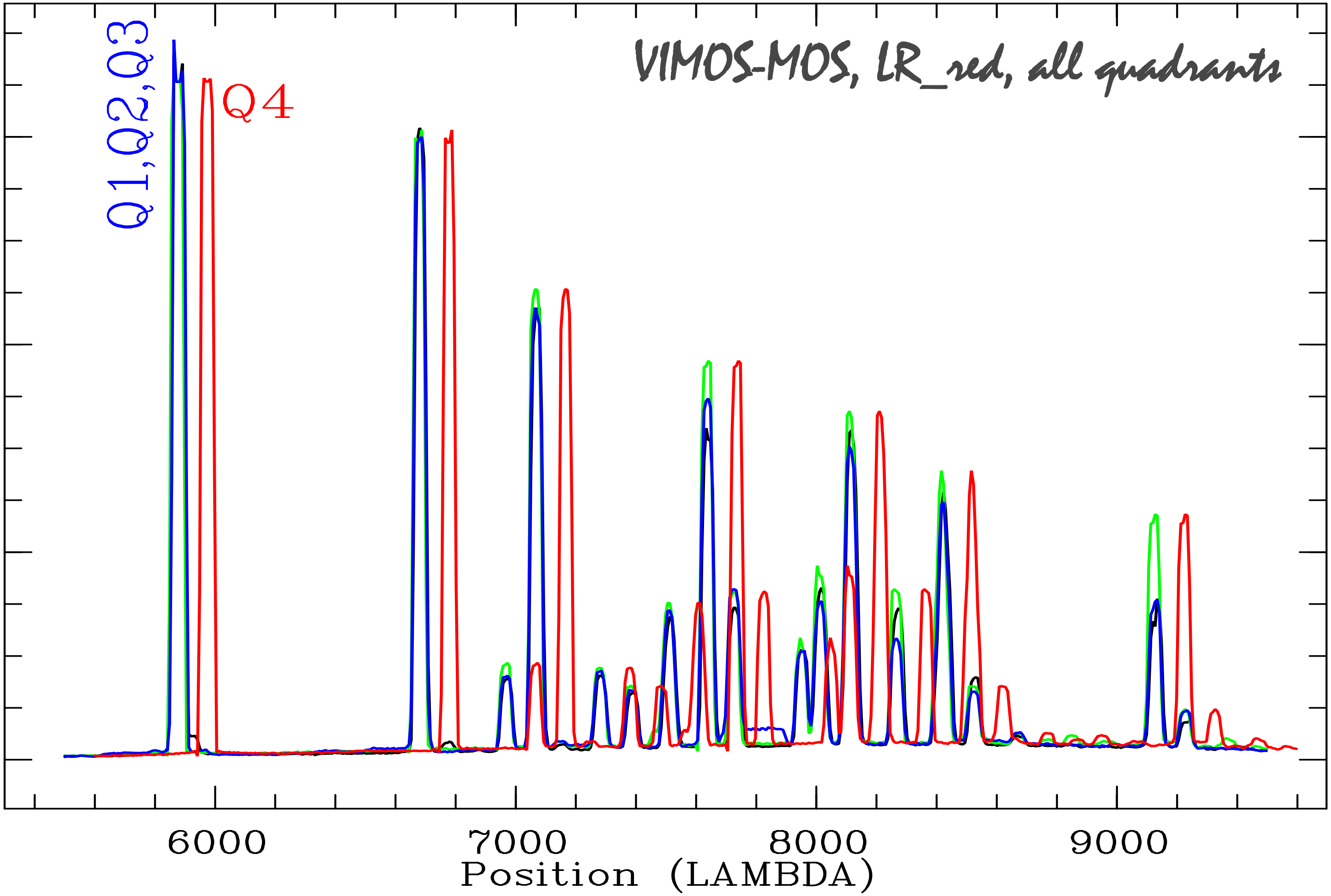
Using first-guess model to find reference lines…



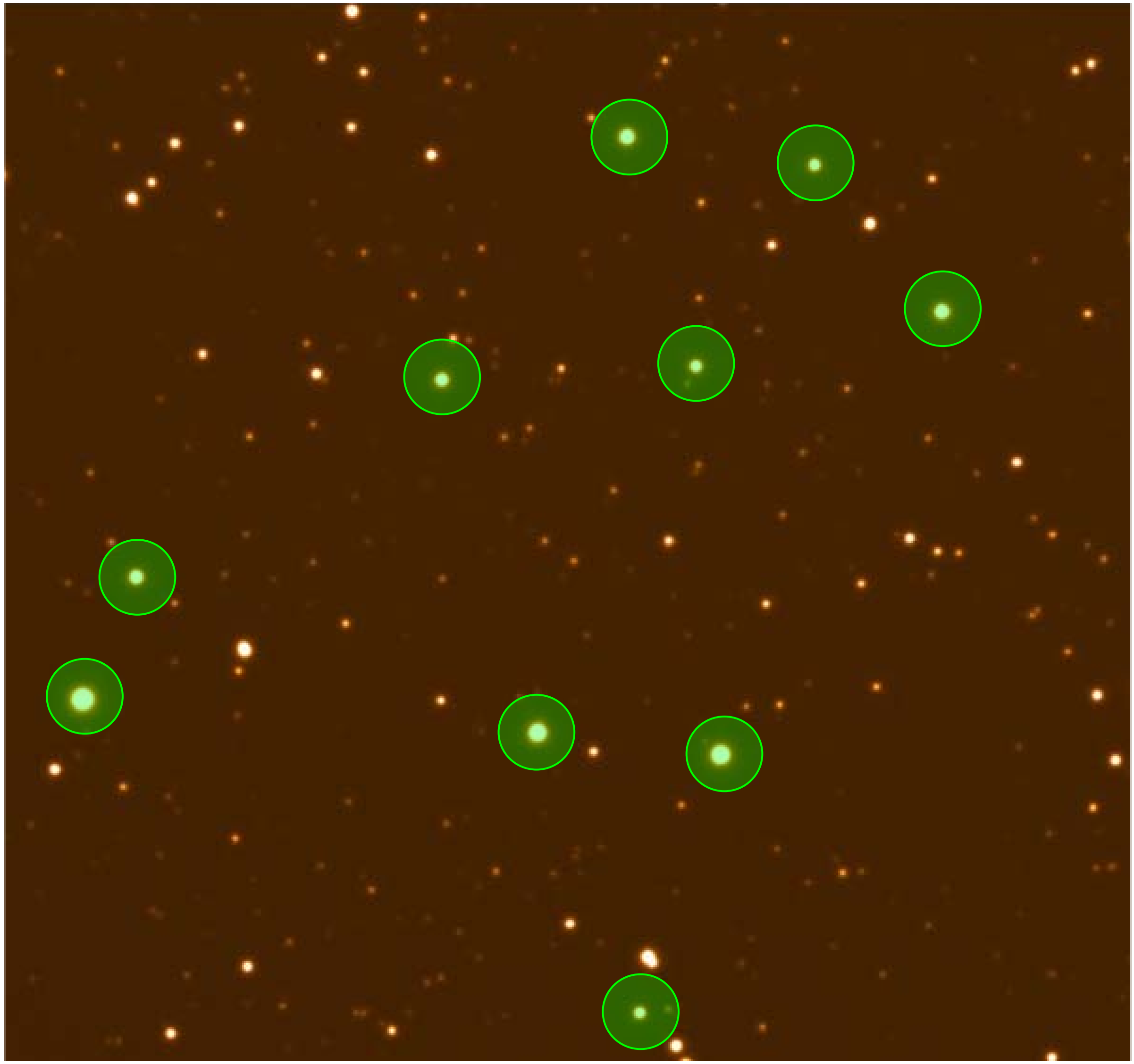
Earthquake!



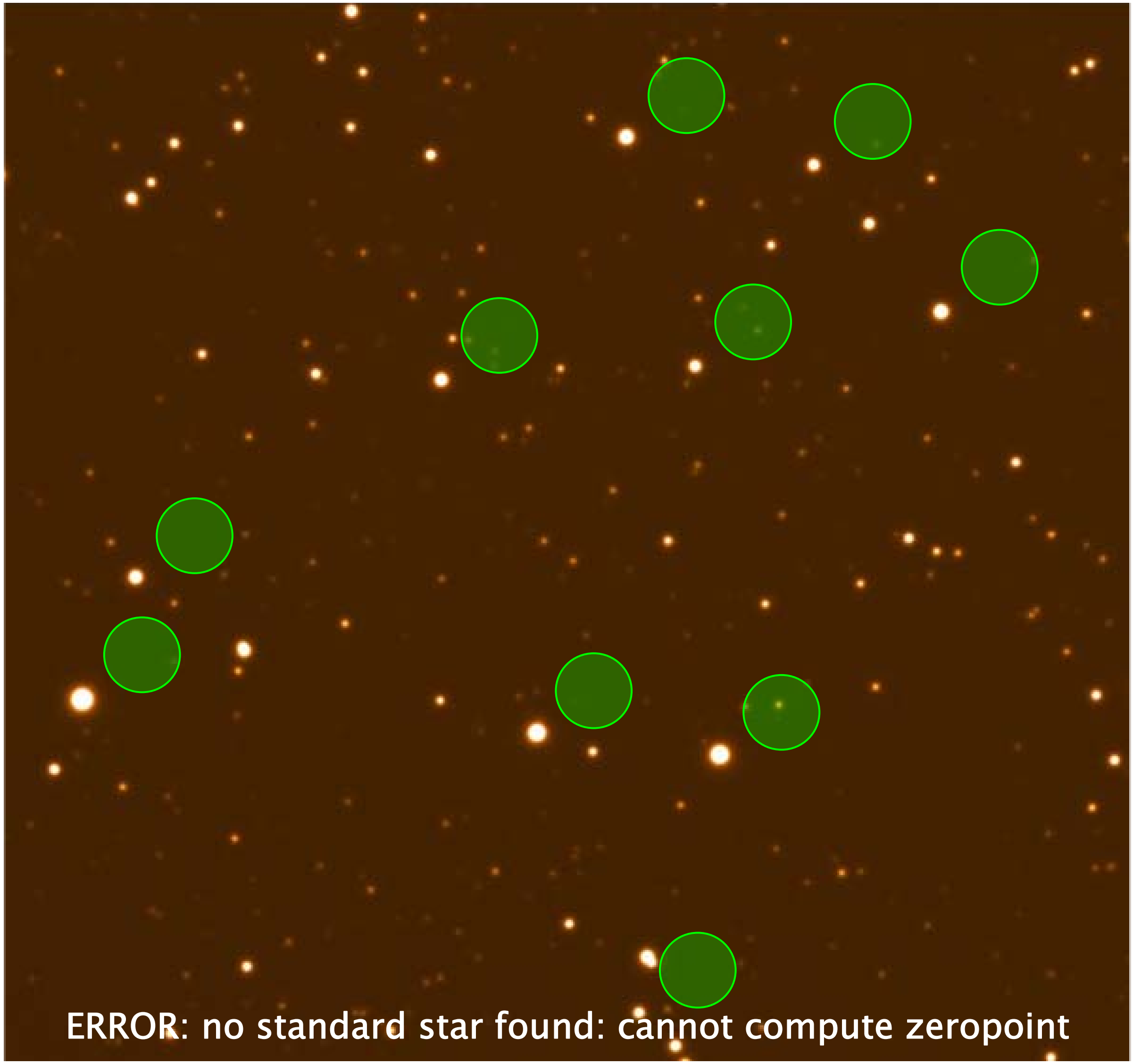
False matches confirm expectations…



Finding standard stars



Finding standard stars



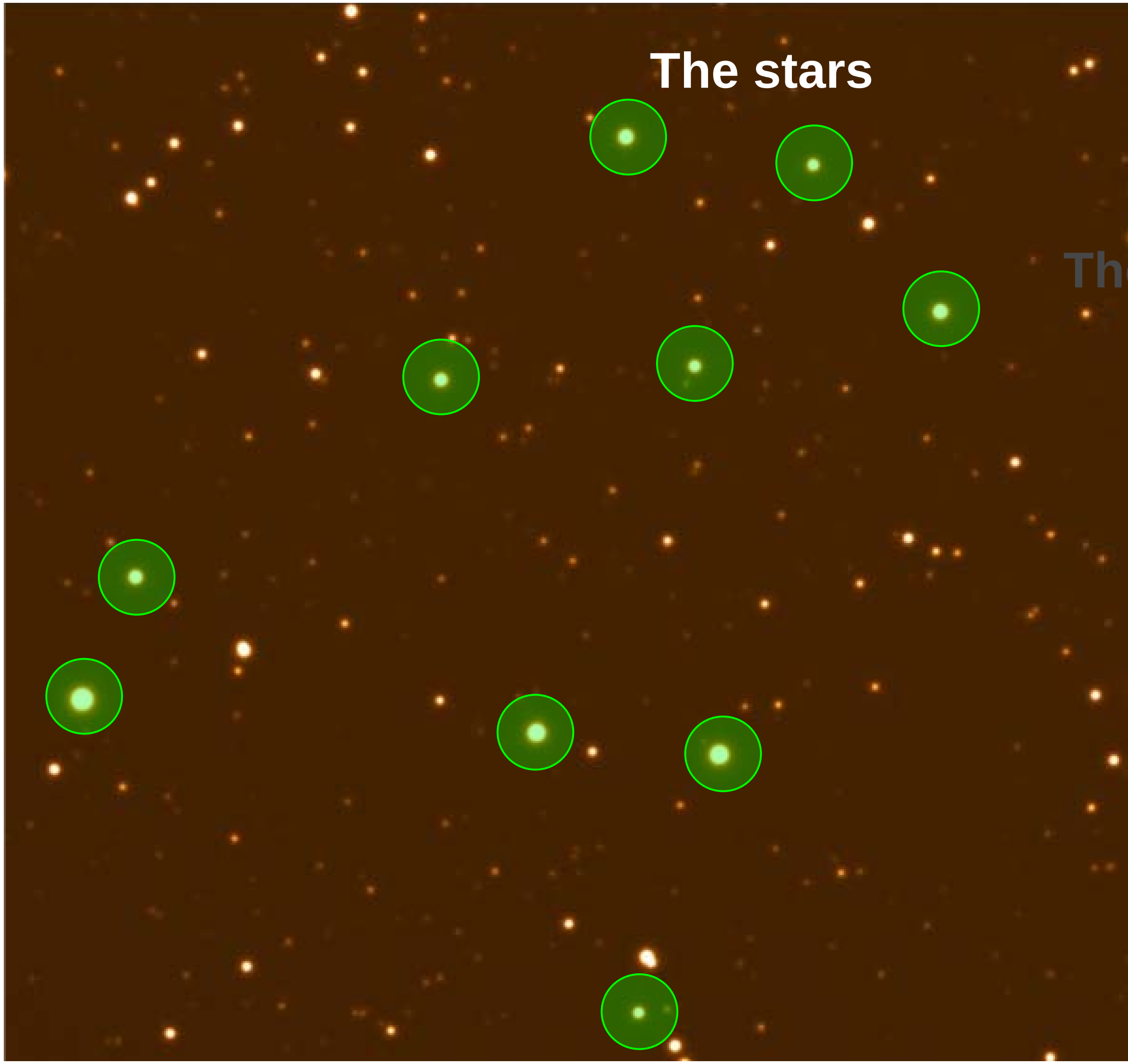
Advantages of the top-down approach

- The safest approach for stable instruments
- The only possible approach if very few reference objects are available

Disadvantages of the top-down approach

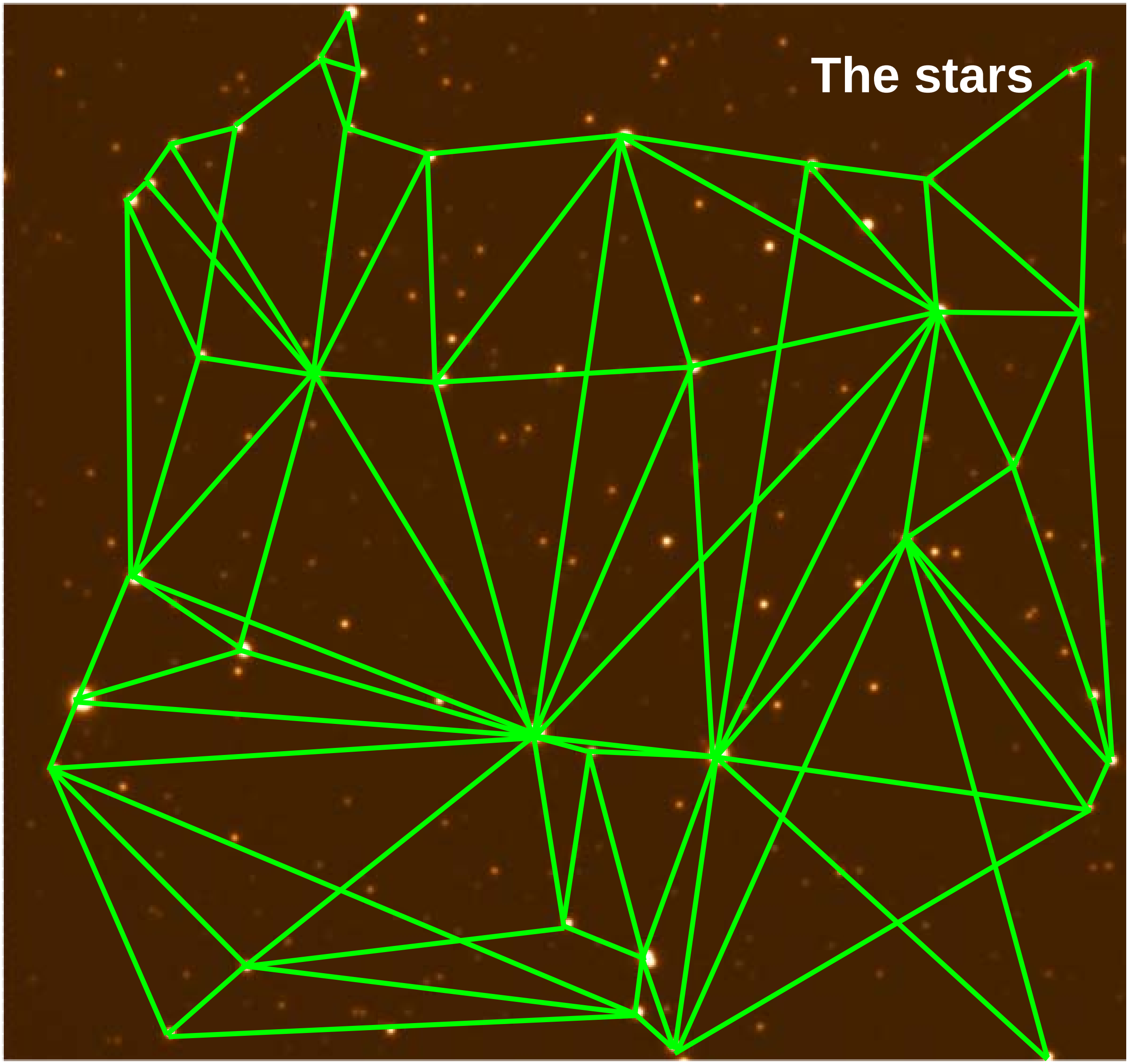
- Not robust: it fails with unstable instruments
- Not flexible: it requires frequent reconfiguration
- Biased toward expectations
- Misled by contaminations

Finding standard stars

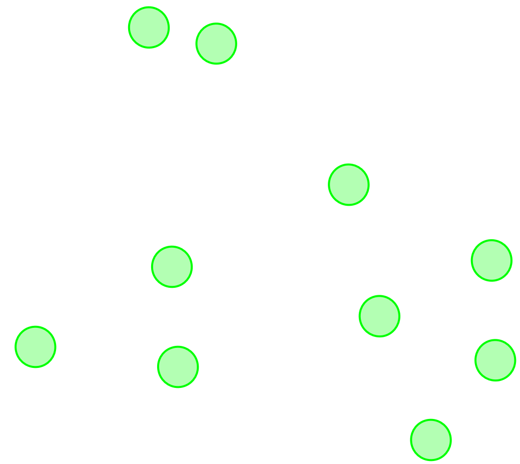


The catalog

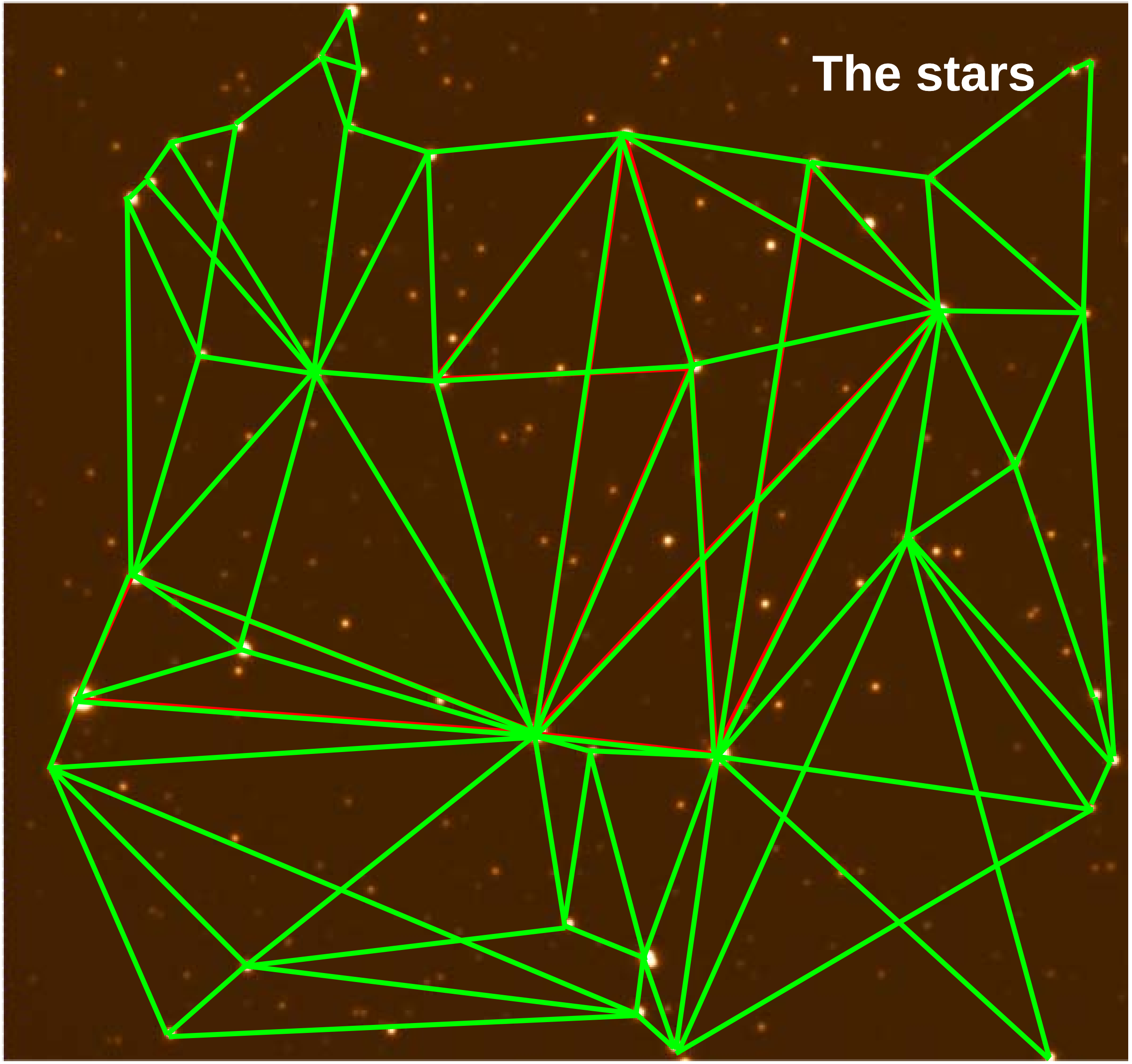
Finding standard stars

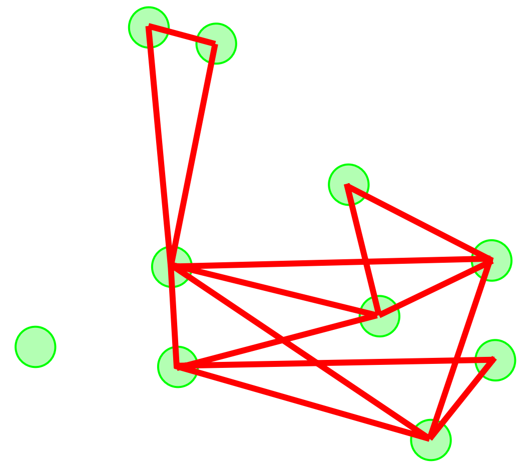


The catalog
(*i.e.*, the pattern)



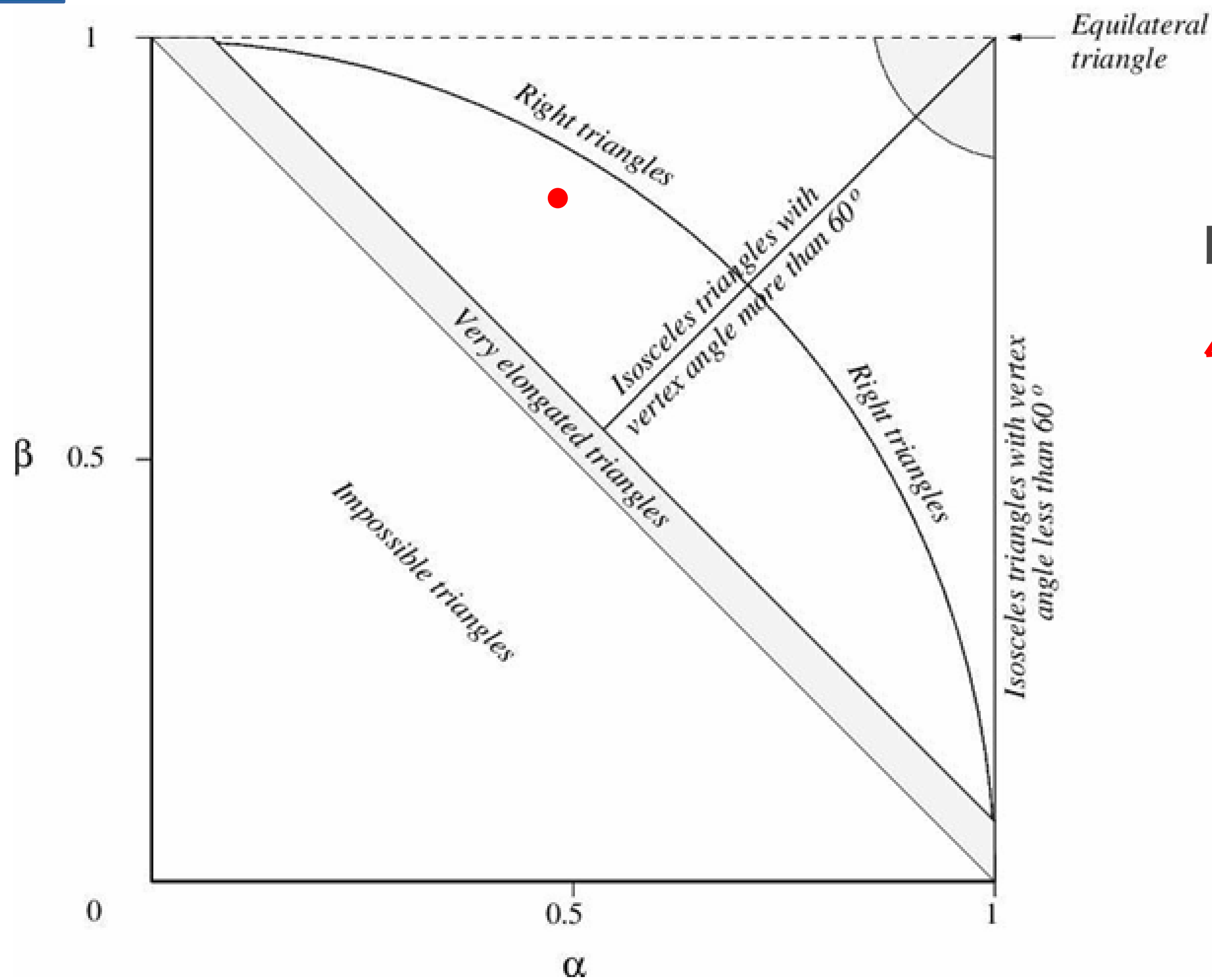
Finding standard stars

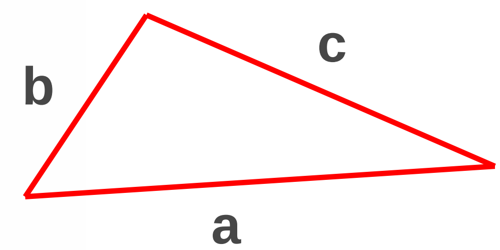




Point pattern-matching (2D)

G. S. Cox *et al*. (1991) "A New Method of Rotation, Scale and Translation Invariant Point Pattern Matching Applied To the Target Acquisition and Guiding of an Automatic Telescope"

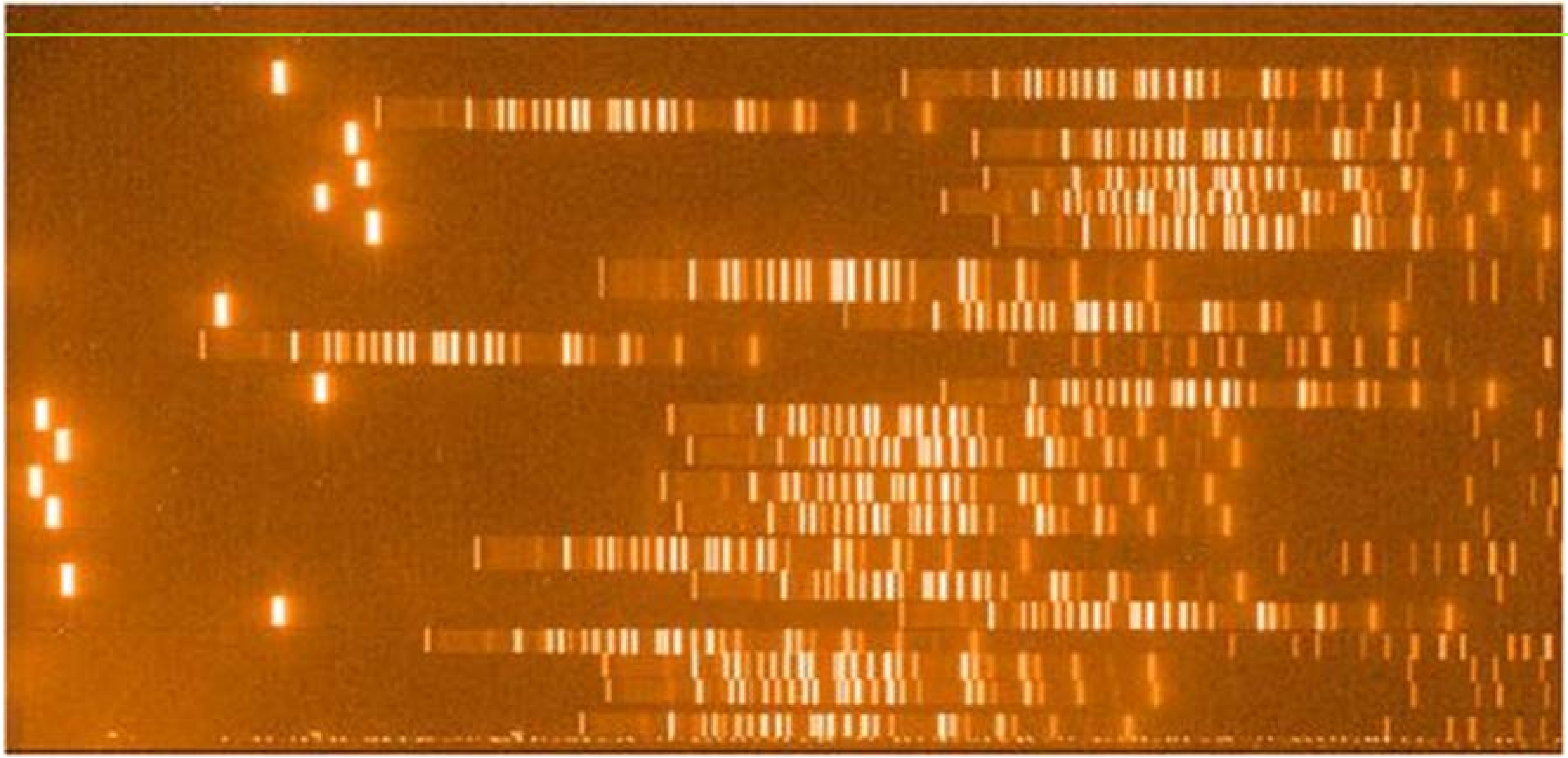




$$\alpha = \frac{b}{a}$$

$$\beta = \frac{c}{a}$$

Looking for peaks



Looking for patterns

The pattern: wavelengths

- ...
- 5400.562
- 5460.742
- 5764.419
- 5769.598
- 5790.656
- 5852.488
- 5875.620
- 5881.900
- ...

The data: pixel positions

- ...
- 1220.64
- 1253.23
- 1299.44
- 1304.07
- 1339.30
- 1400.33
- 1450.28
- 1457.32
- 1471.00
- 1496.21
- ...

Looking for patterns

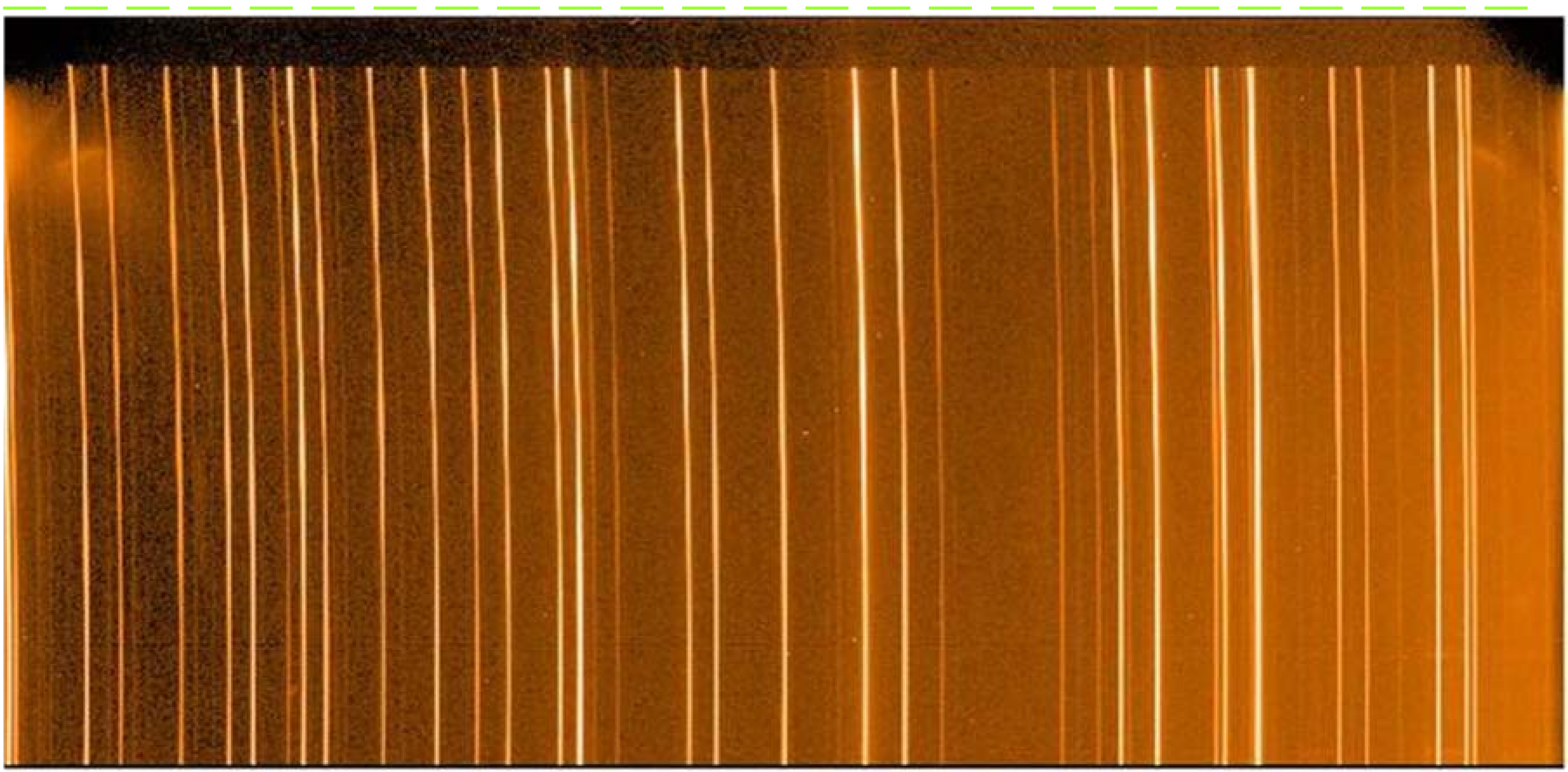
The pattern: wavelengths

- ...
- 5400.562
- 5460.742
- 5764.419
- 5769.598
- 5790.656
- 5852.488
- 5875.620
- 5881.900
- ...

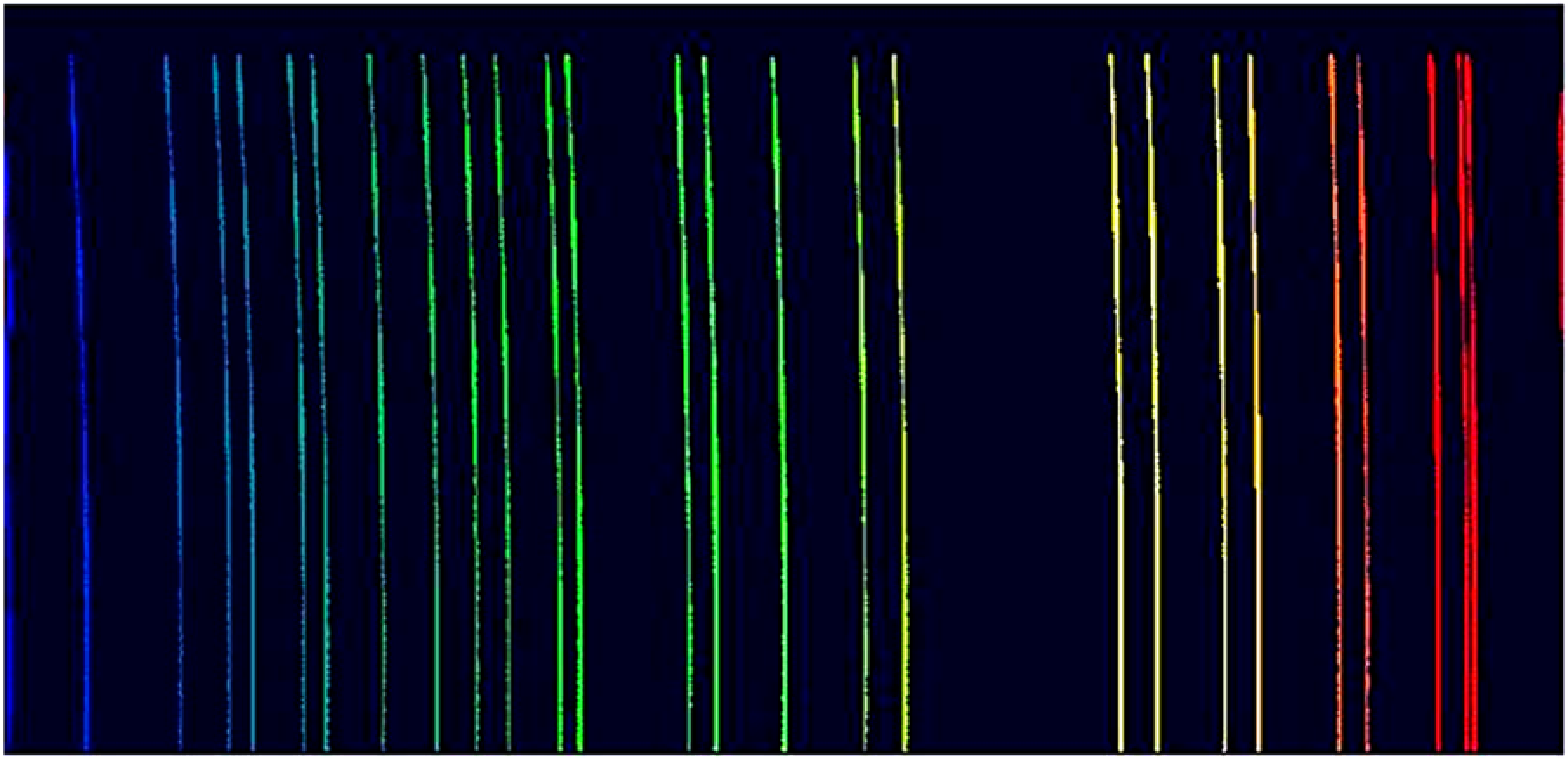
The data: pixel positions

- ...
- 1220.64
- 1253.23
- 1290.44
- 1304.07
- 1339.30
- 1400.33
- 1450.28
- 1457.32
- 1471.00
- 1496.21
- ...

A simple case: calibrating a single spectrum

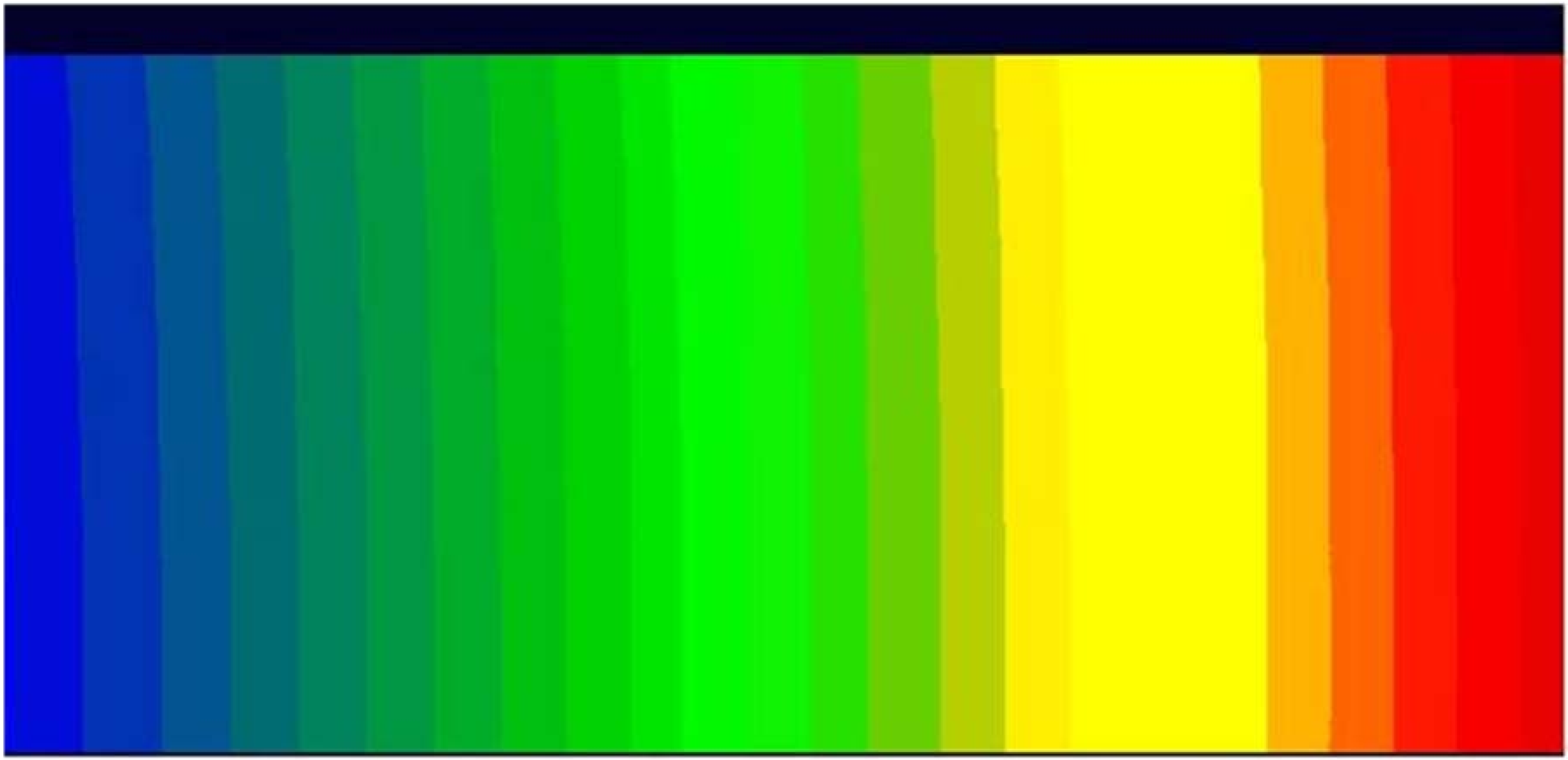


Identifying arc lamp lines



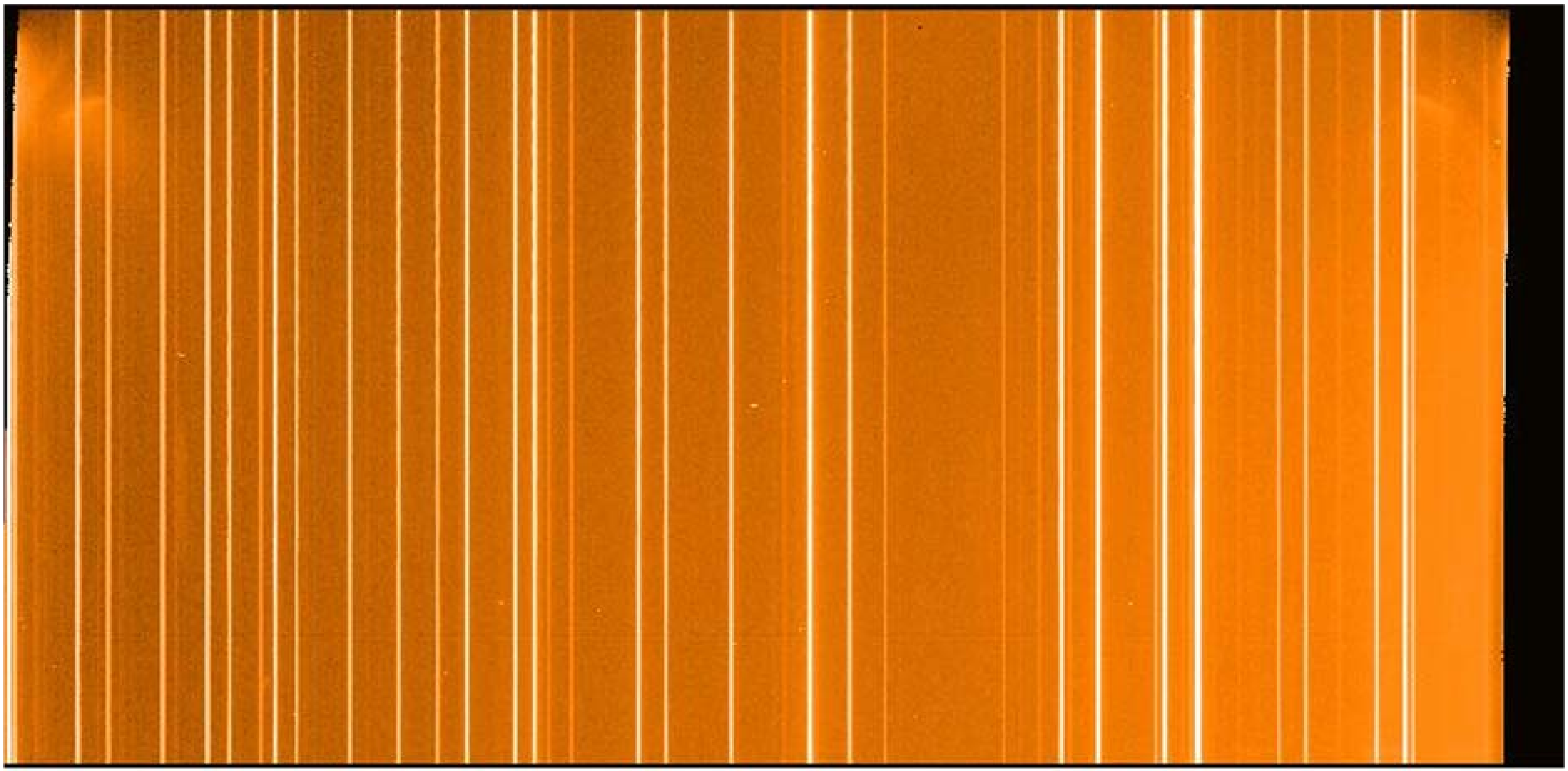
Wavelength map

Mean accuracy: 0.05 pixel

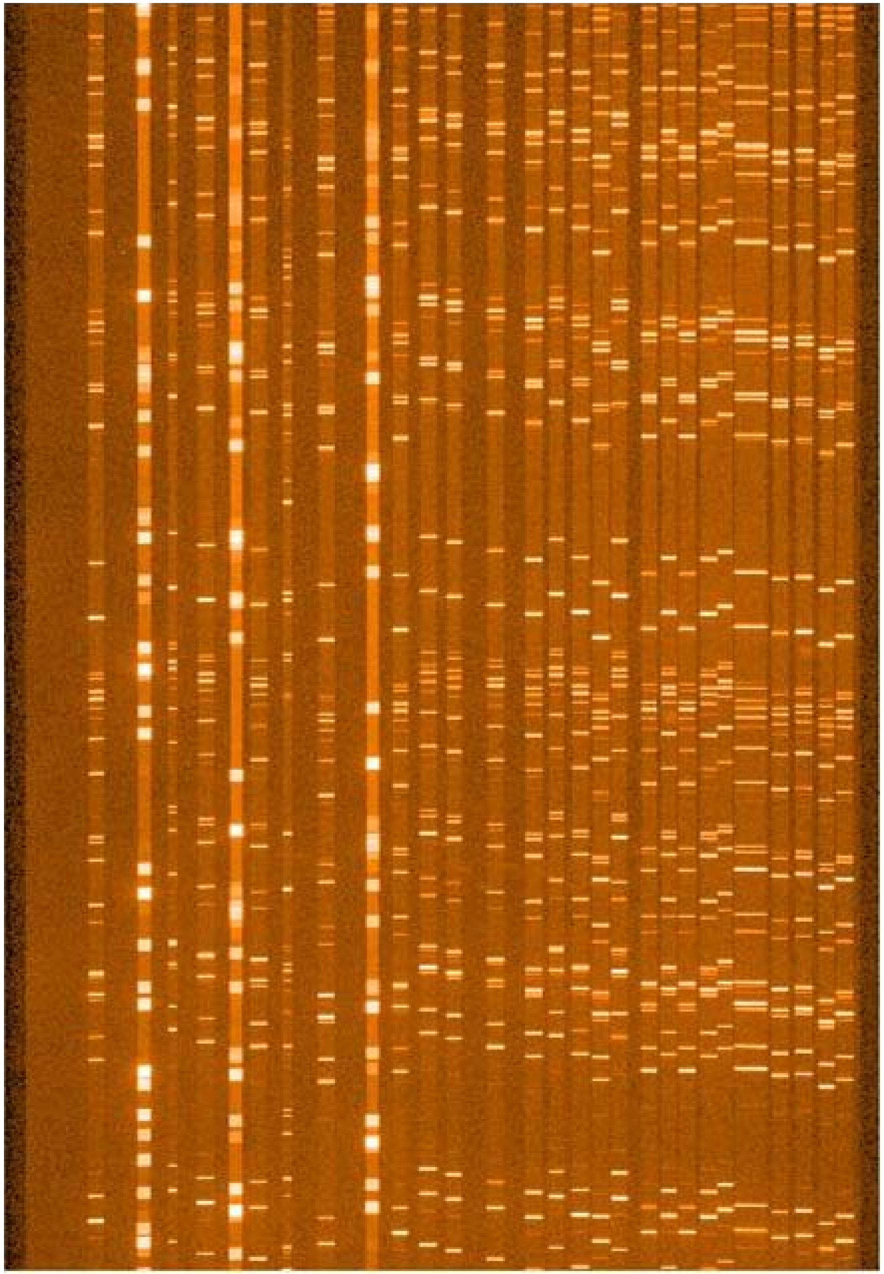


Resampled spectrum

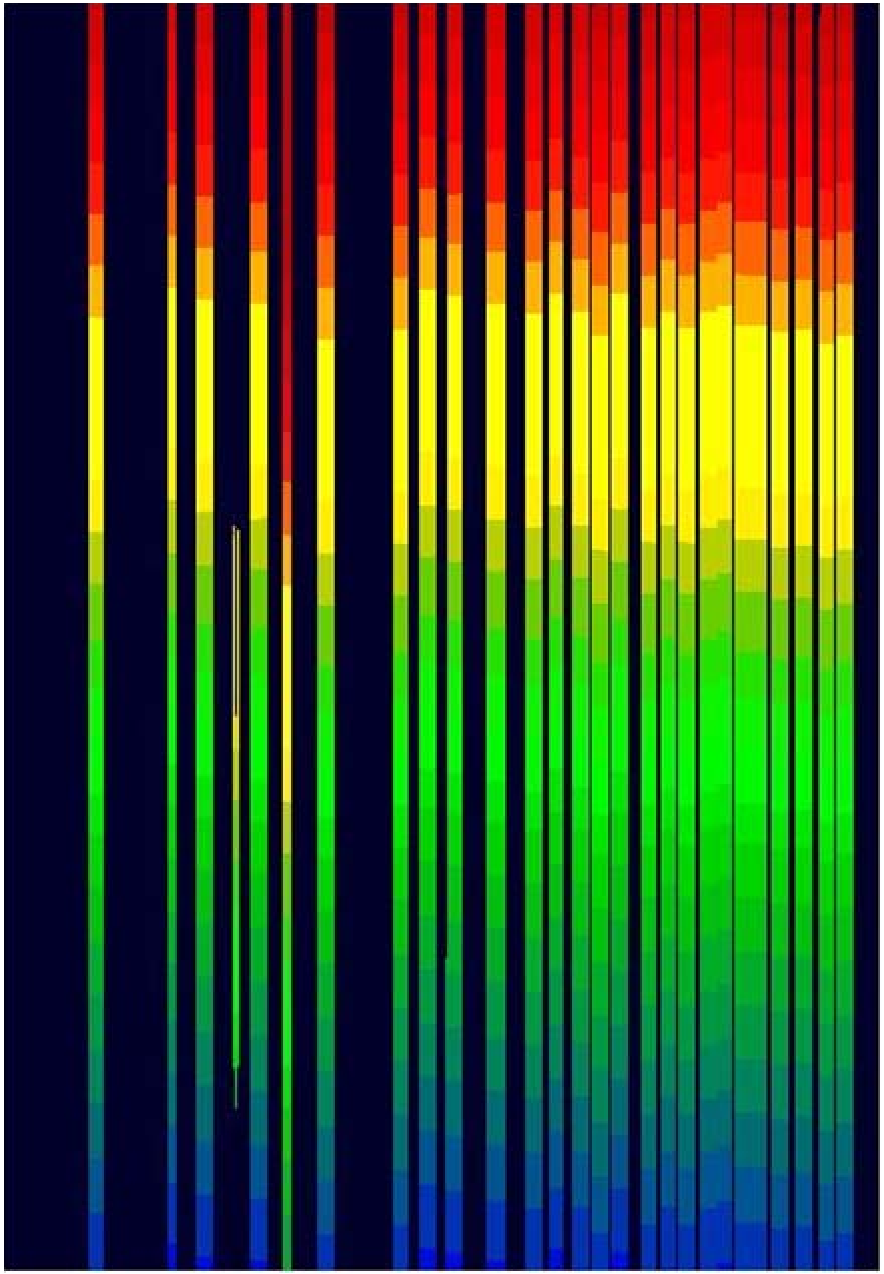
Mean accuracy: 0.05 pixel



A case with many spectra (VIMOS GRIS_HRred)

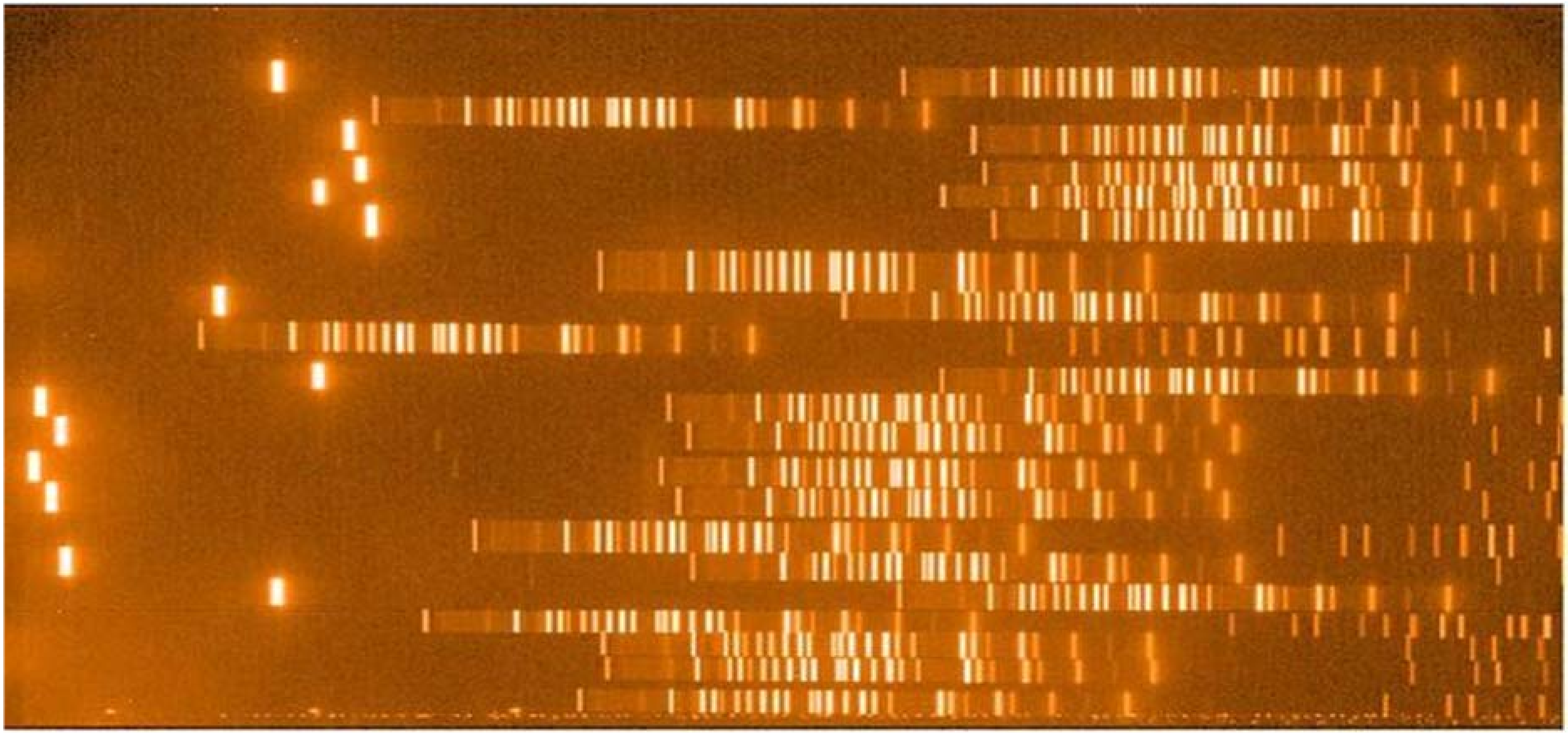


Wavelength map



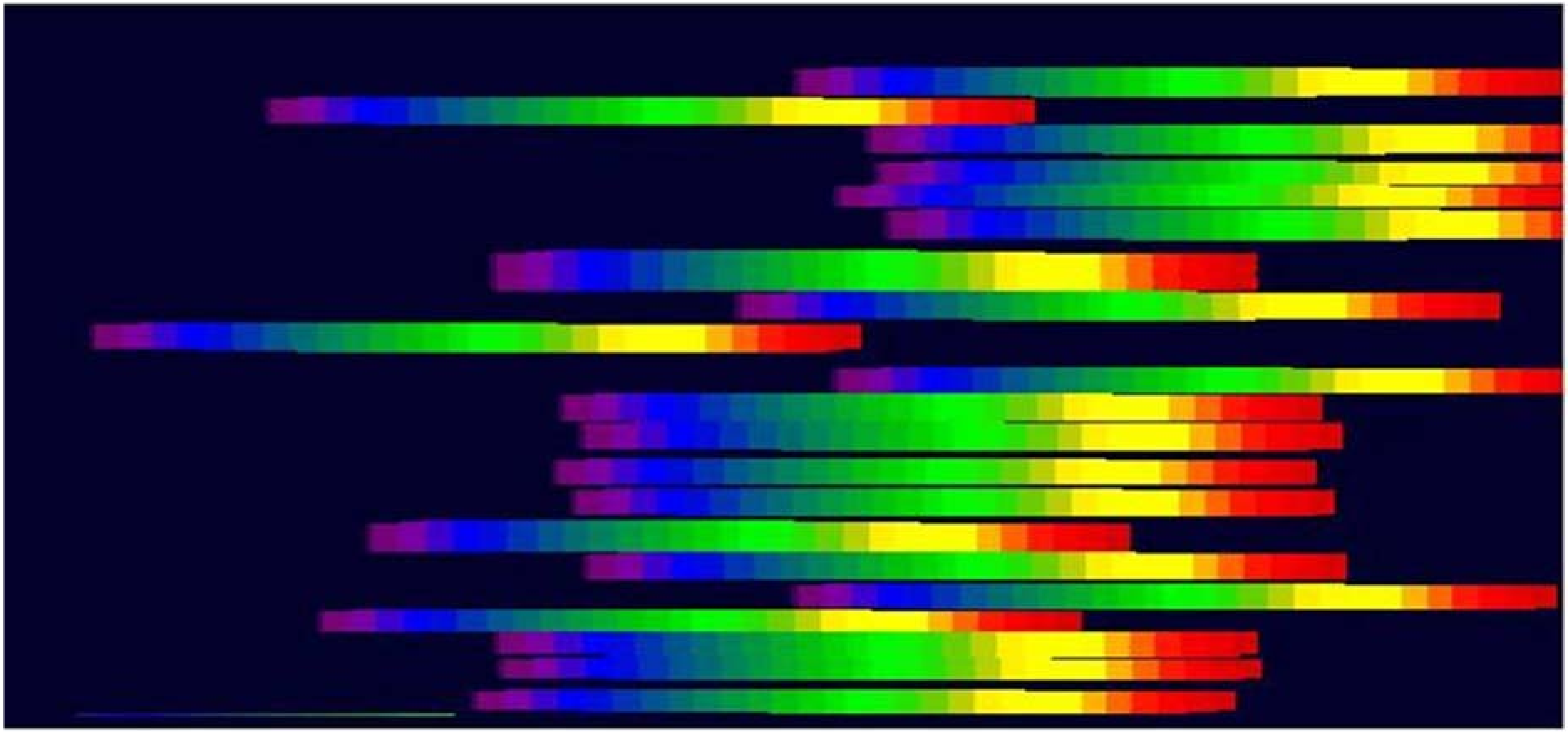
Mean accuracy: 0.07 pixel

Another example (FORS2-MXU GRIS_150I)



Wavelength map

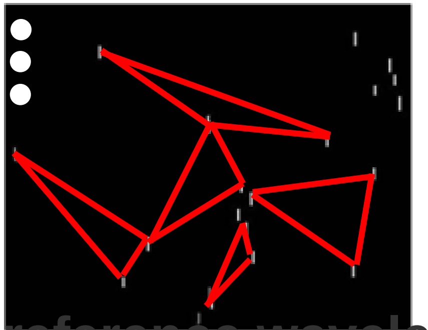
Mean accuracy: 0.05 pixel



Identifying the spectra

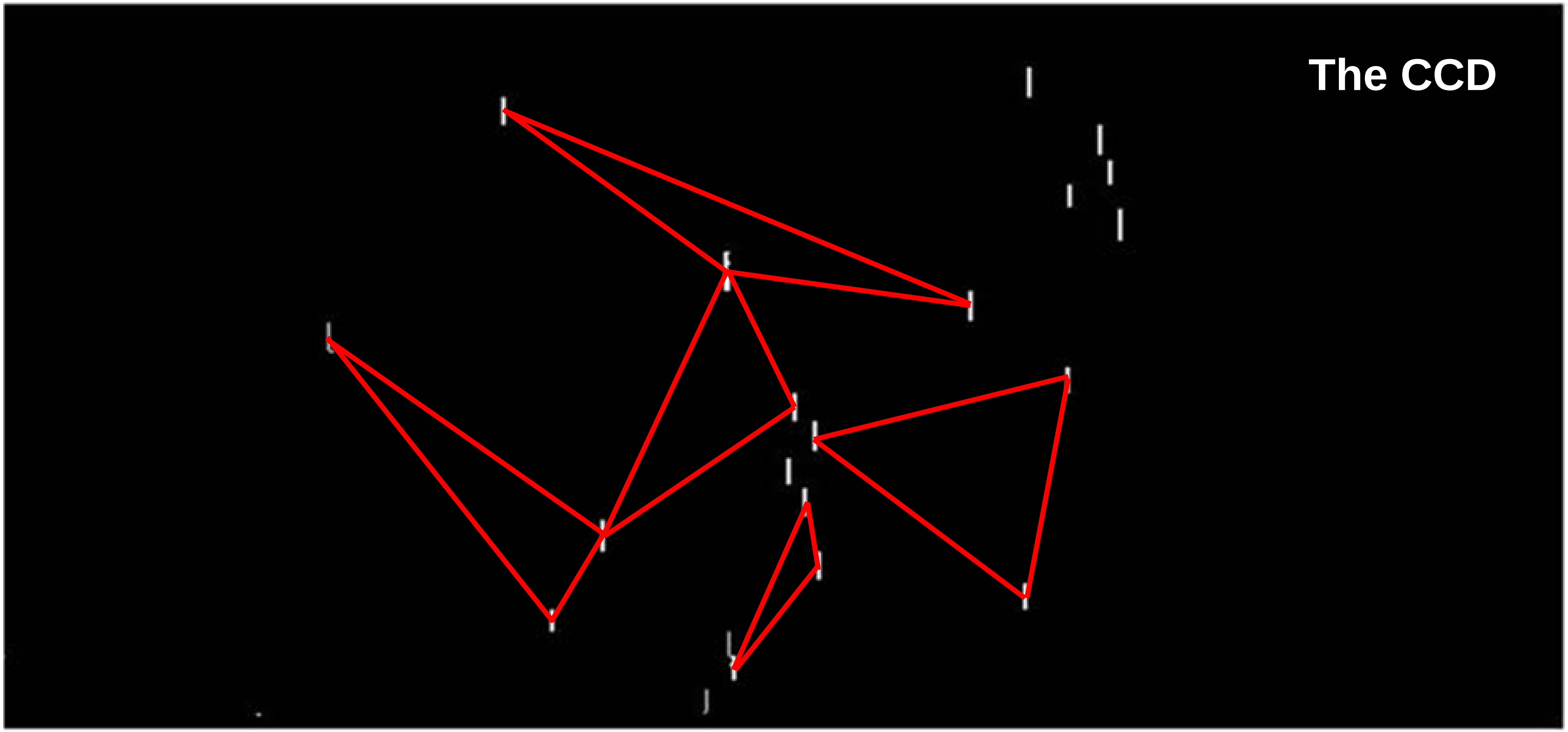
Select the reference wavelength: in this example, λ = 7000.00 A

Identifying the slits

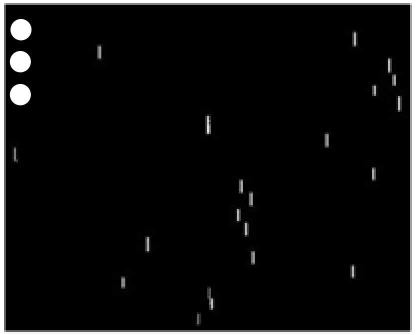


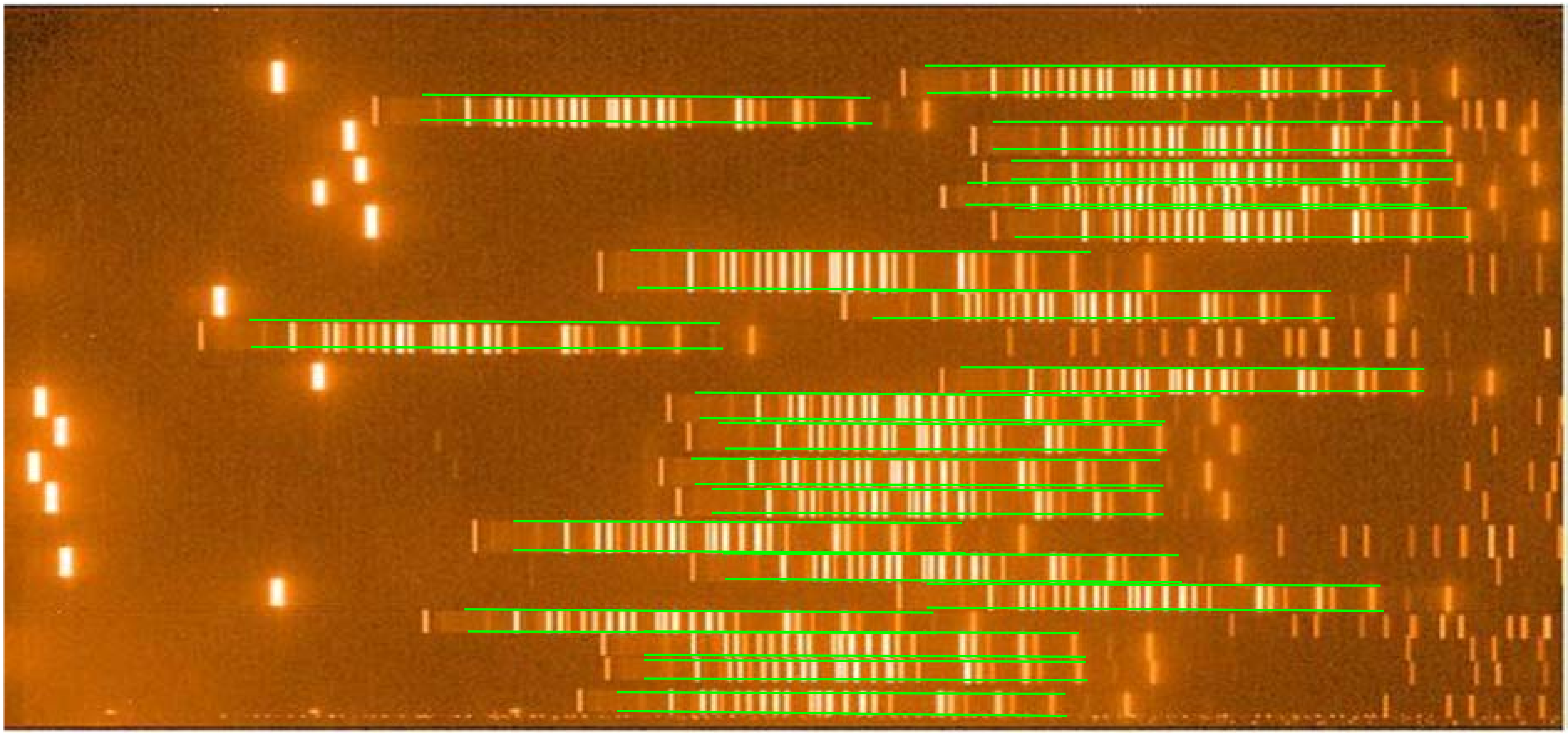
The mask

Select the reference wavelength: in this example, λ = 7000.00 A



Measuring the spectral curvature



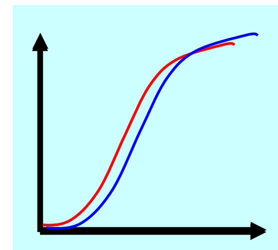


Rectified image

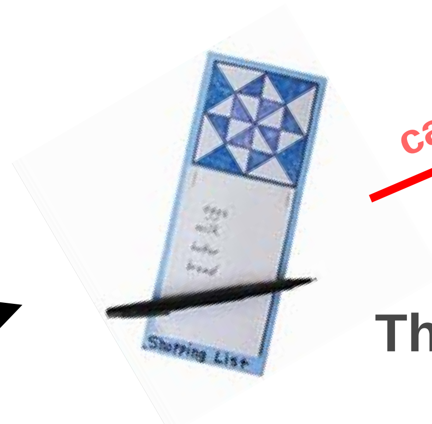
How to identify reference objects?

The bottom-up approach:

The generated model



calibration



The pattern

The candidate objects

Robustness

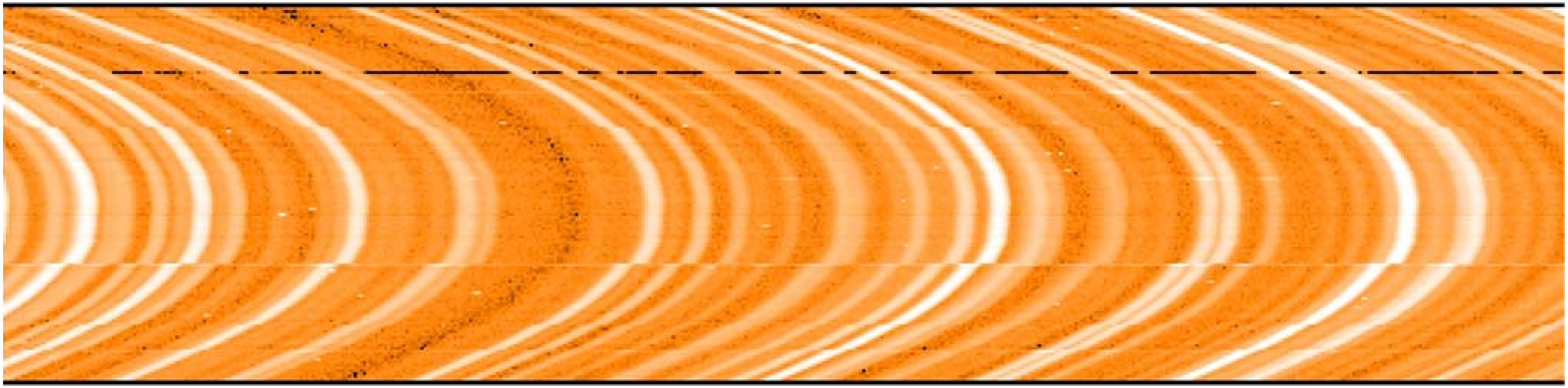
- **This approach can cope with unexpected position and/or number of spectra**
- **Reference lines are more safely identified for being part of a pattern, rather than for being close to some expected position**

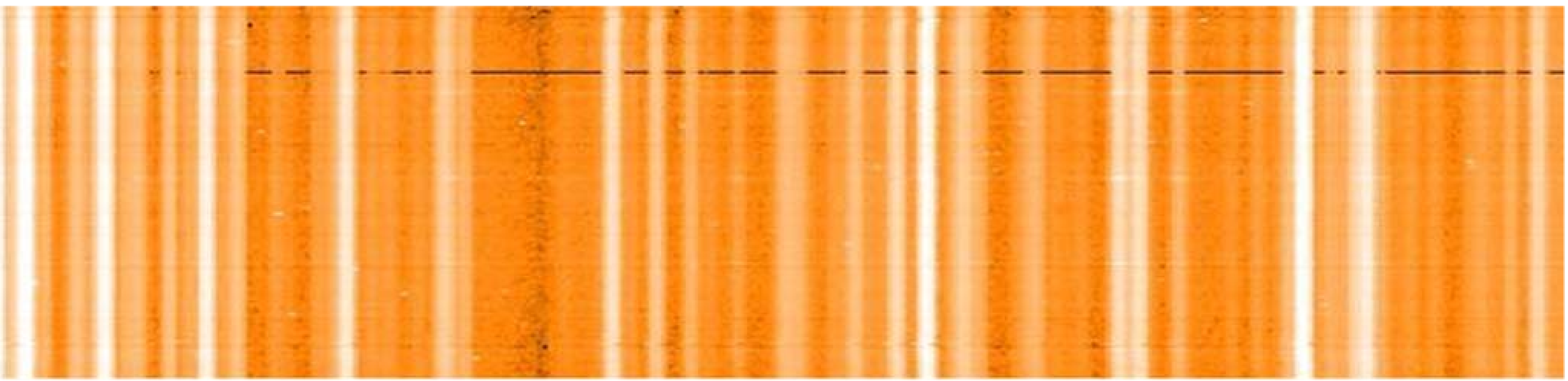
Flexibility

- This approach is (MOS) instrument-independent
- Low maintenance cost, even at instrument upgrades (new chip, new grism, new lamps…)
- This approach may be applied to extracted products from any kind of spectroscopic data (not just MOS, but also IFU, echelle, etc.).

GIRAFFE Medusa1_H525.8nm

Mean accuracy: 0.10 pixel



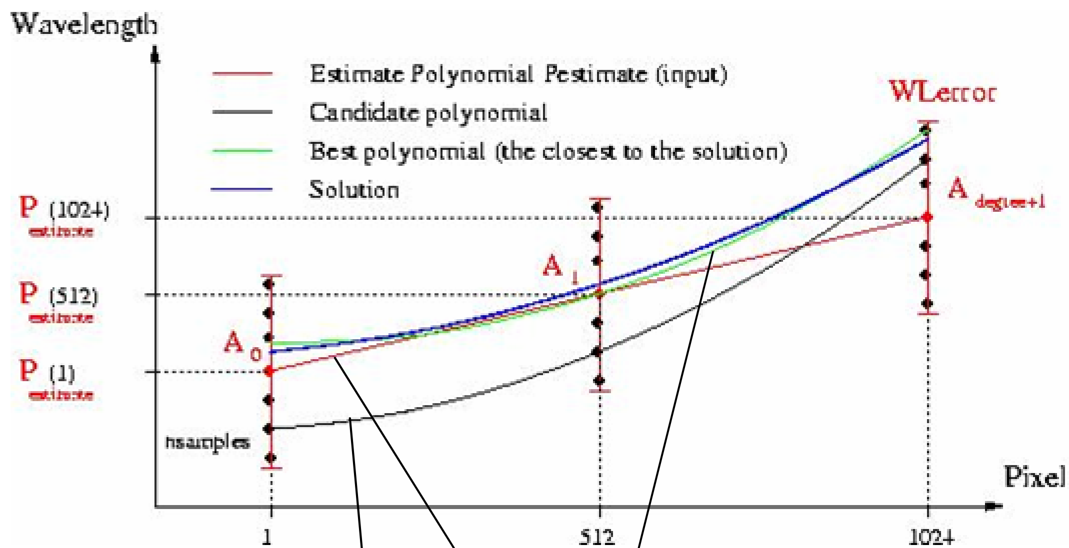


SDD

Disadvantages of the bottom-up approach

- **As the top-down approach depends on a model, the bottom-up approach depends on the data!**
- **This approach is a black box: as for any program based on a bottom-up strategy (*e.g.*, using trained neural networks), it is often difficult to find the reason of a failure.**

The CRIRES data reduction challenges

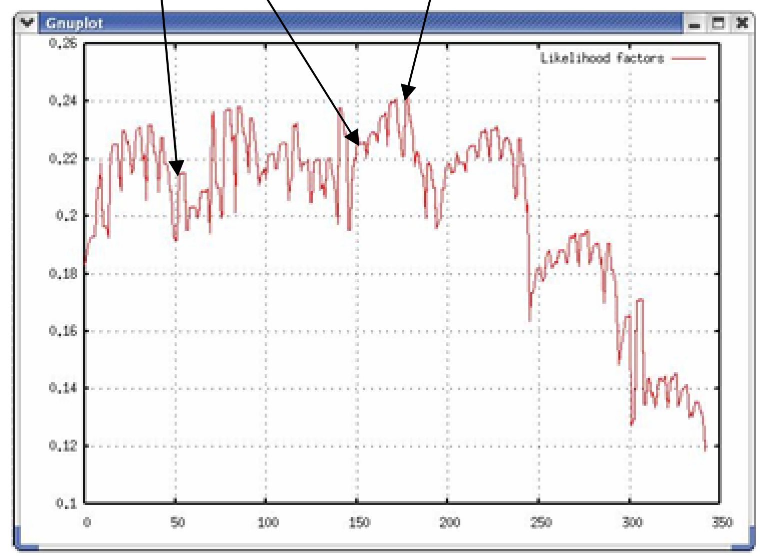


Inputs:

- The Estimate Polynomial
- The Wavelength Error of the estimate (WLerror)
- The degree of the searched polynomial (degree)
- The number of samples (nsamples)
- The lines catalogue (OH, Gas cell, Lamps, Hitran)
- The signal to calibrate (in pixels)

Algorithm:

- Consider degree+1 positions Ai regularly spaced, and nsamples points spread within WLerror around these positions
- For each possible sequence of points (nsamples^(degree+1) possibilities), the interpolation polynomial is created and considered as candidate
- The candidate polynomial is used to convert the signal to calibrate from pixels to wavelengths. This signal is compared to the signal generated from the catalogue. A likelihood coefficient in computed
- The best likelihood parameter gives the best candidate, i.e. the polynomial that is the closest to the solution
- A second pass (or more) is used to refine the solution with the first pass solution used as estimate, with a smaller WLerror and a higher degree.



Thanks to:

Lars Lundin (ESO/SDD)

Kieran O'Brien (ESO/PSO)

Emmanuel Jehin (ESO/PSO)

Ralf Palsa (ESO/SDD)

Marguerite Pierre (CEA, Saclay, France)

Stefano Cristiani (INAF, Trieste, Italy)

Christophe Adami (LAM, Marseille, France)

Harald Kuntschner (ESO/ST-ECF)

Martino Romaniello (ESO/DFO)

Vanessa Doublier (ESO)

Sandro Villanova (DdA, Padova, Italy)

Stefano Bagnulo (ESO/PSO)

Burkhard Wolff (ESO/DFO)

Emanuela Pompei (ESO/LSO)

Ivo Saviane (ESO/LSO)

Looking for patterns

The pattern: wavelengths

- ...
- 5400.562
- 5460.742
- 5764.419
- 5769.598
- 5790.656
- 5852.488
- 5875.620
- 5881.900
- ...

The data: pixel positions

- ...
- 1220.64
- 1253.23
- 1299.44
- 1304.07
- 1339.30
- 1400.33
- 1450.28
- 1457.32
- 1471.00
- 1496.21
- ...

Looking for peaks