Photometric Redshifts

Masayuki Tanaka
(NAOJ)

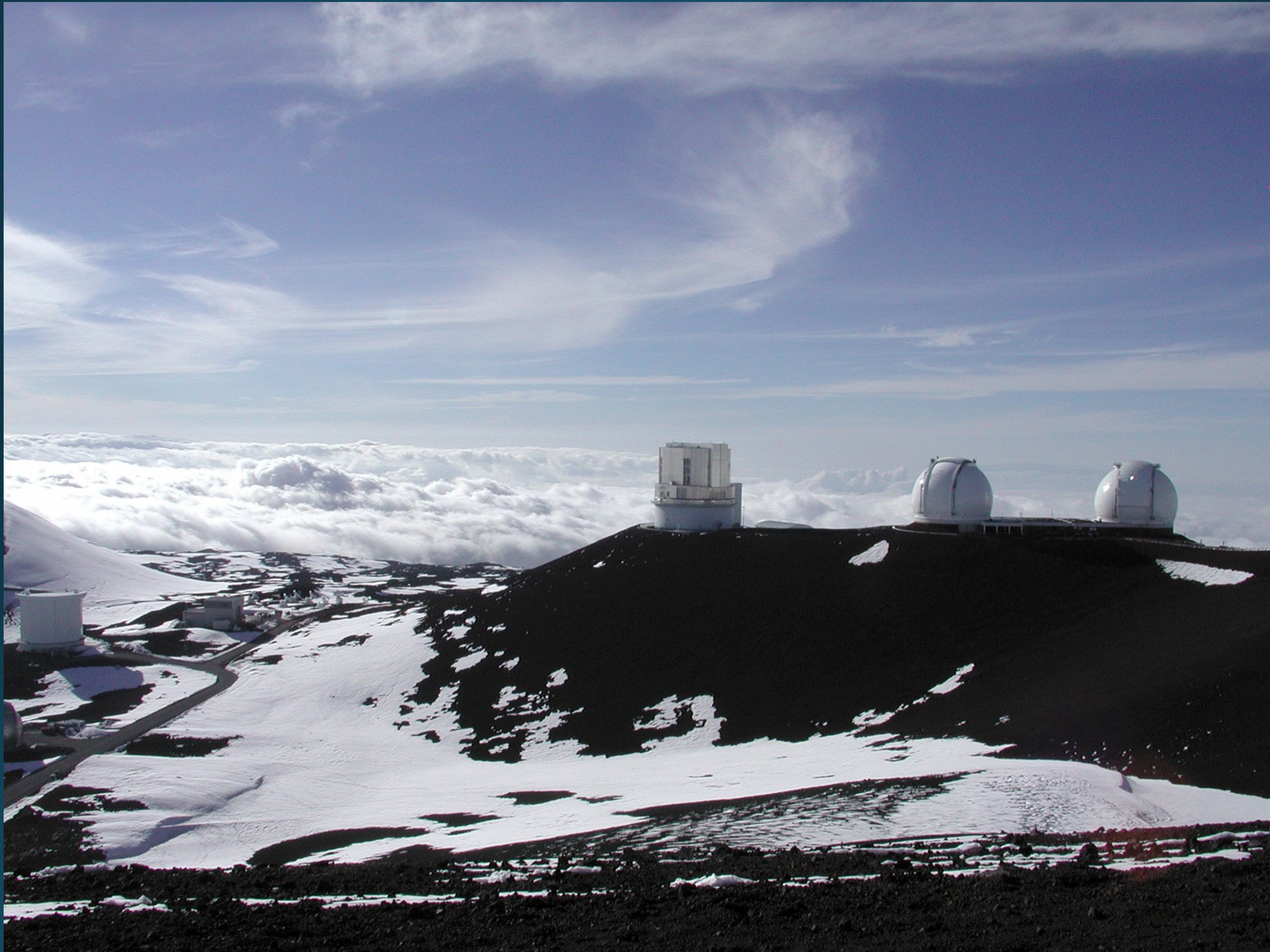
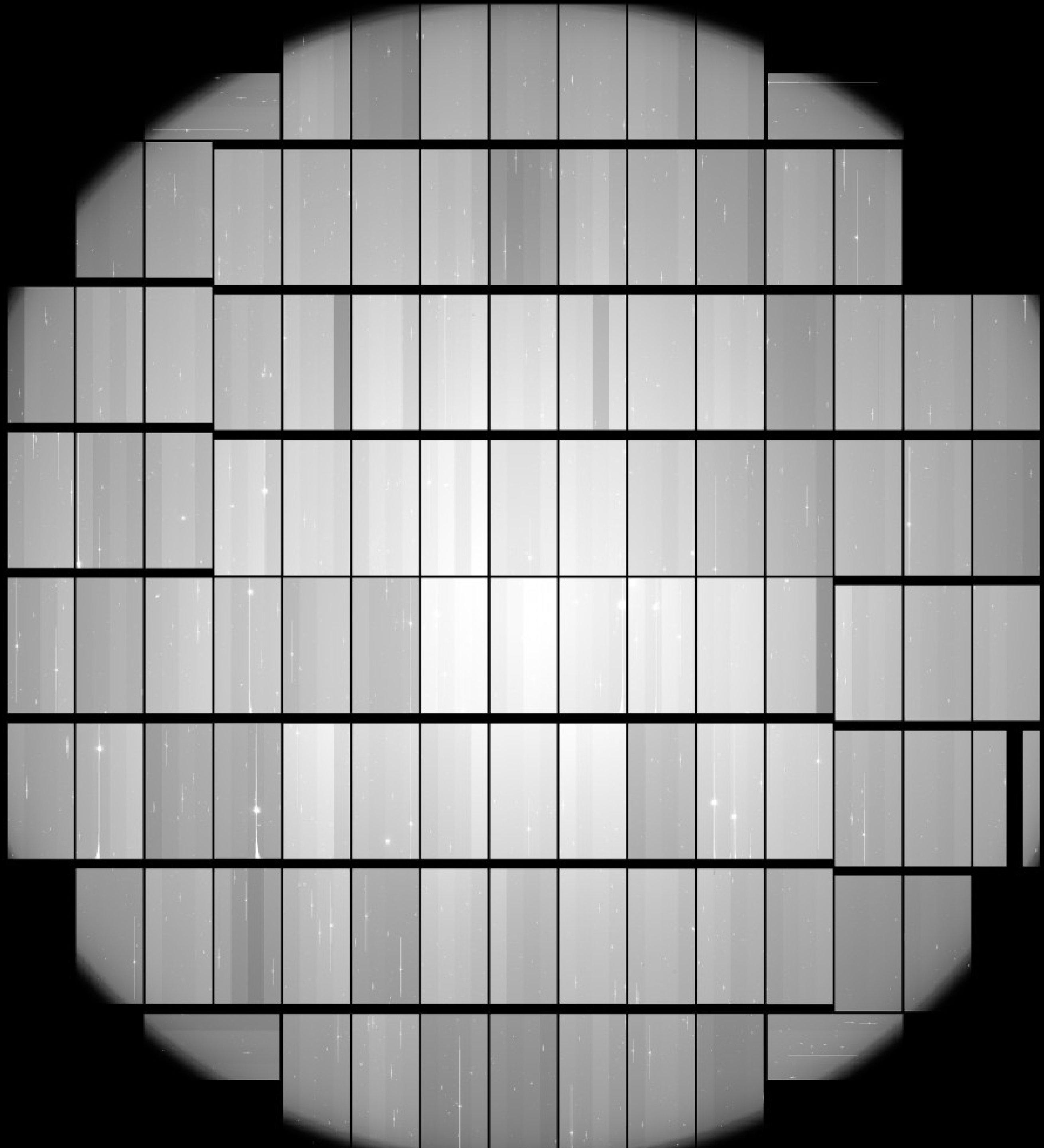
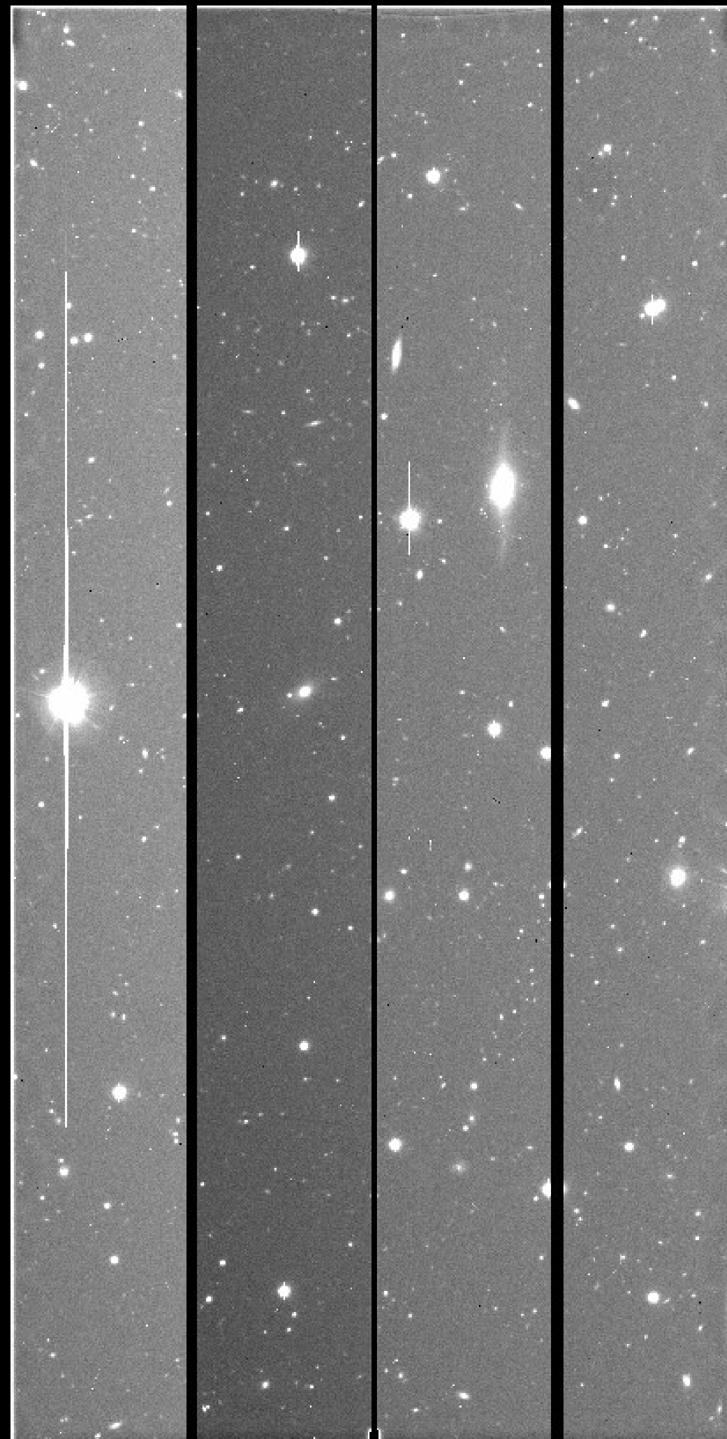


Photo by Y. Utsumi

Photo by Y. Utsumi

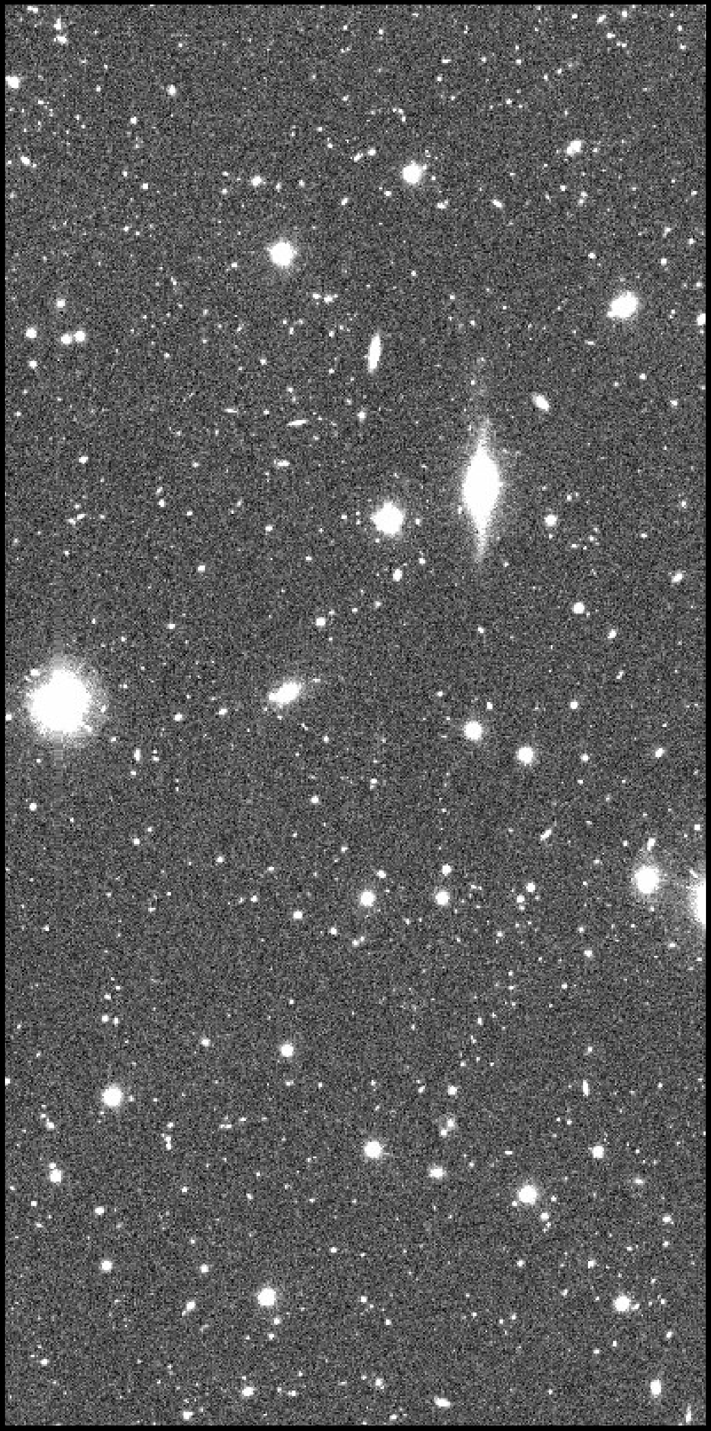


Raw



Raw

Processed



Raw

↓

Processed

↓

Stacked

Observables:

– flux (in a given filter)
– size / shape
– projected position
– time variability

From observables to physical quantities

Are we happy with these direct observables?

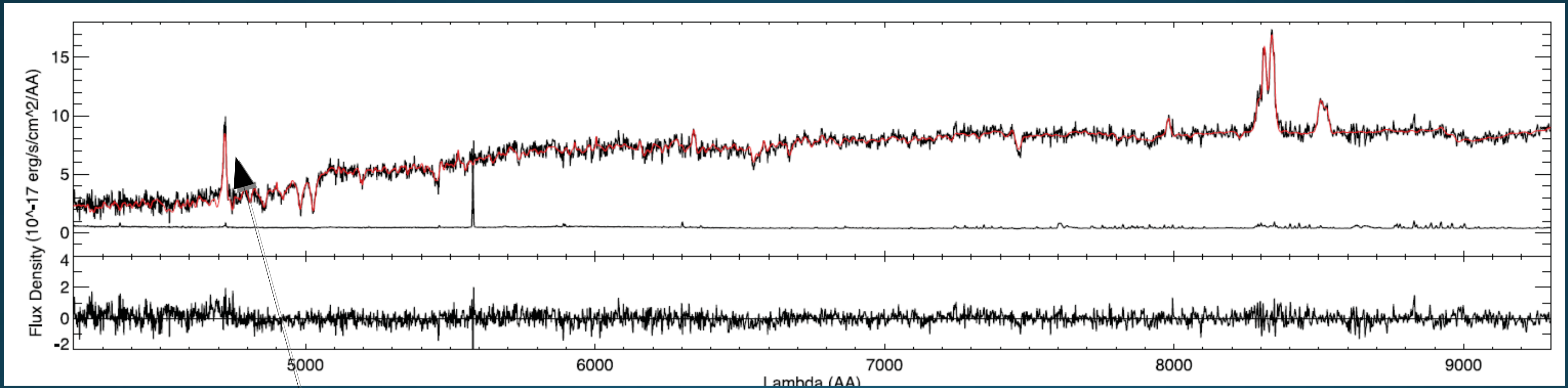
NO!

We need luminosities instead of apparent fluxes!
We need physical sizes instead of apparent sizes!

We need distance information!

How to measure distances?

There are several distance measures, but for objects at cosmological distances, we do spectroscopy:



This is [OII] at rest-frame wavelength of 3727 Angstrom.
It is observed at 4700 Angstrom.

$z = 4700 / 3727 - 1 = 0.26$

But, spectroscopy is very expensive

In addition to redshifts, spectra contain a lot of information about the galaxies. However, spectroscopy is **very expensive** in terms of telescope time.

	HSC	PSF
i=22.5 objects	10sec	1 hour
Objects / FoV	10^4-5	2400

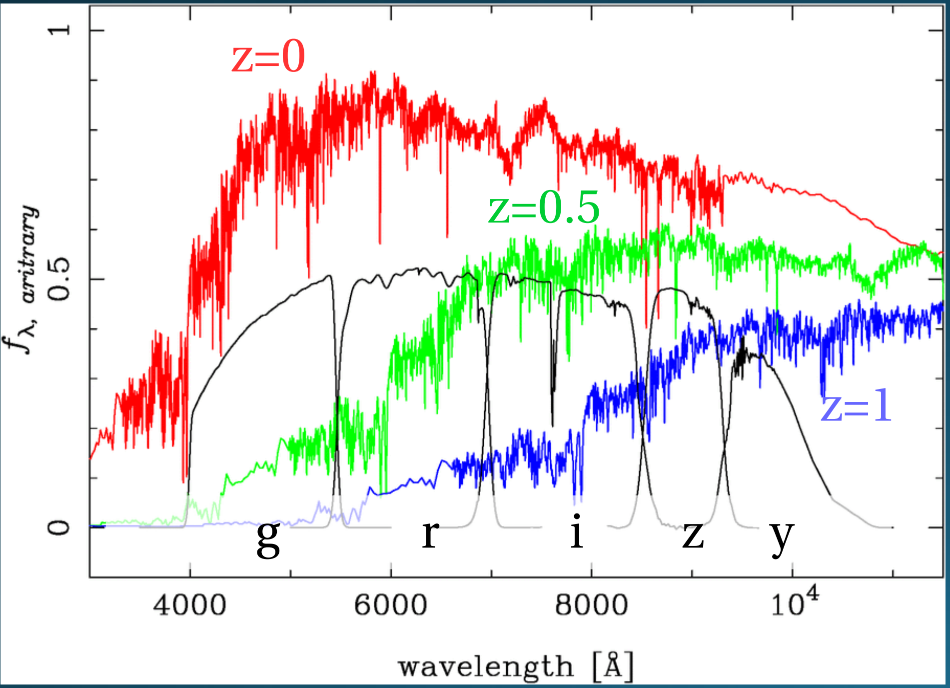
In addition, many of the objects that HSC detects are fainter than the spectroscopic sensitivity limits.

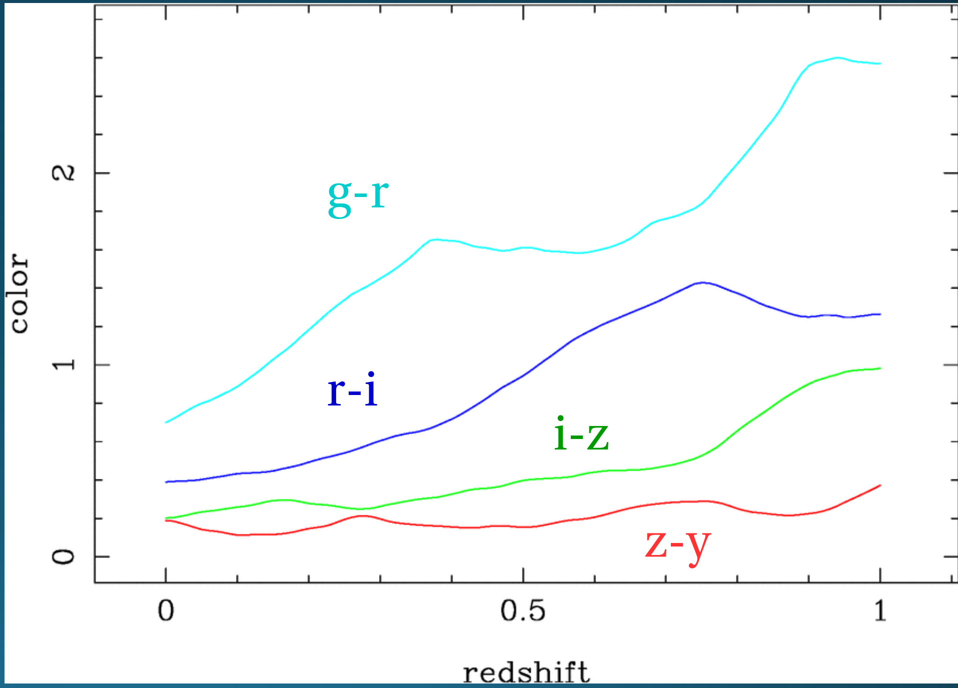
Can we use photometric information to infer redshifts?

Yes! That is photometric redshift.

The idea

We probe different rest-frame wavelengths of objects at different redshifts. We can use colors to infer redshifts.





Photometric redshift

In short, photometric redshift is a technique to make mapping between observables and redshift.

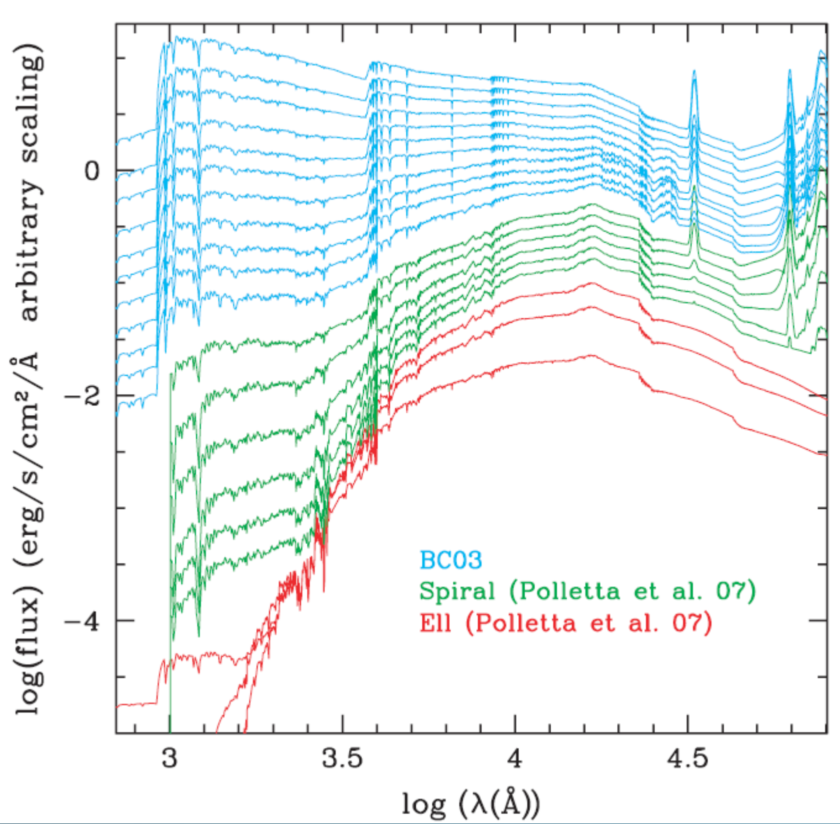
1 – template fitting:
We use spectral templates of galaxies. We put them at various redshifts, compute colors of these redshifted templates, and compare them with observed colors of galaxies.

2 – numerical fitting:
We assume some function (e.g., polynomials) to make the mapping using spectroscopic redshifts:
$z = a * m1 + b * m2 + c * m1 * m2 + \ldots$

3 – machine learning:
Generalized form of #2. We use spectroscopic objects and let a machine learn and make the mapping by itself.

4 – clustering redshifts:
We use spatial information. Potentially a very powerful technique.

1 – Template fitting



Ilbert et al. 2009

Spectral templates can be either from observations or stellar population synthesis models.

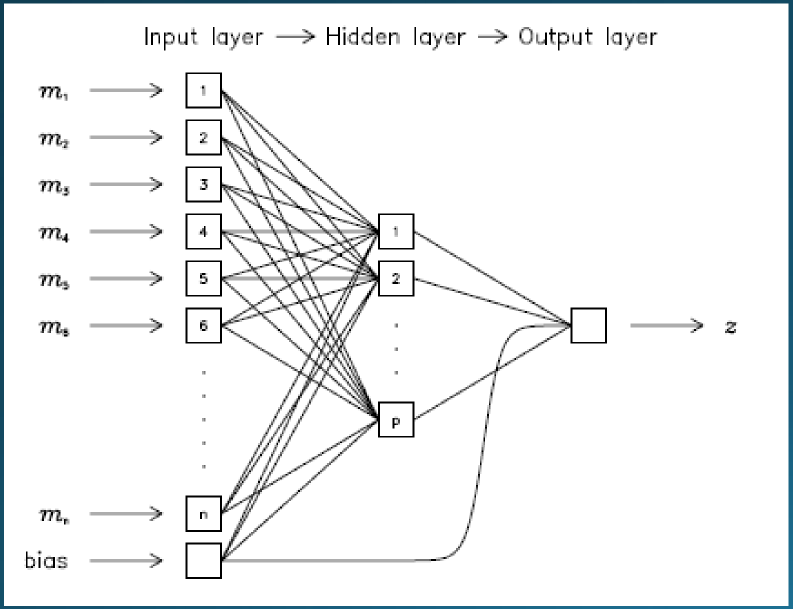
Pros: we 'expect' to go fainter than the spectroscopic limits provided that our understanding of galaxy spectra is reasonable.

Cons: templates may not include all types of galaxies.

3 – Machine-Learning

We feed a 'training' sample of spectroscopic objects to a machine-learning code and let the machine learn by itself.

Pros: If trained well, it works better than template fitting methods.
Cons: The training spectroscopic sample has to represent an input catalog to which you apply your code.



$$u_j = \sum_i w_{ij} g_i(u_i),$$

$$g_j(u_j) = 1/[1+\exp(-u_j)]$$

Collister and Lahav 2004

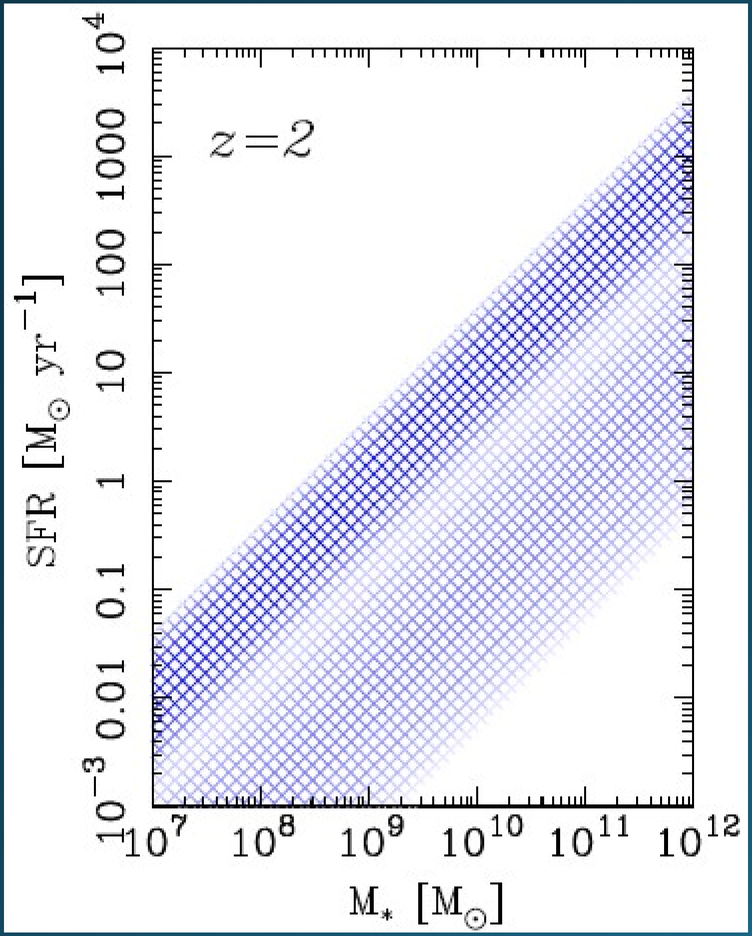
There are other techniques applied to photo-z: random forest, deep learning, etc.

Template fitting vs machine-learning

	Template-fitting	machine-learning
Accuracy	OK	Good
Spec-z demand	Low-Medium	High
Faint mags?	Yes	No

It depends strongly on science cases, but in order to fully exploit the imaging data from HSC, I think we should be using **both** techniques and combine them to make the 'best' photo-z estimates.

Improved template fitting



One of the physical priors used

Tanaka 2015, ApJ, 801, 20

$$P(z, G|m) \propto \int d\alpha P(m|z, G, \alpha) P(z, G, \alpha).$$

$$P(z, G, \alpha) = P(z) P(\mathrm{SFR}|M_*, z) P(\tau_V|\mathrm{SFR}, z) \times P(\mathrm{age}|M_*, z).$$

- Template fitting
- Galaxy templates with BC03, QSO and stellar templates
- Template error function to reduce systematic templates mismatches
- Bayesian priors
- Output:
  - redshift
  - stellar mass
  - star formation rates
  - dust attenuation
  - rest-frame mags

Machine-Learning

Some of us have been in touch with a machine-learning expert and are making a preliminary analiys (this project will be announced to HSC later). Some of the main goals are

1 – to ask which machine-learning code gives the best performance
2 – to use photometric information that has been less exploited so far such as shapes and sizes.

We just started a collaborative work and plan to make progress in the coming years.

Template fitting Machine-Learning Clustering?

'Master' photo-z catalog

Summary

- Information from imaging data are limited and we need redshifts to do science.
- But, spectroscopy is so hard and we need an easier way – photometric redshift.
- There are two major techniques for photo-z: template fitting and machine-learning.
- There are pros and cons in these techniques and we should be using both.
- We eventually combine multiple photo-z estimates into master photo-z estimates.

Photo by Y. Utsumi