Tecnologie per la ricostruzione di modelli 3D da immagini

Marco Callieri
ISTI-CNR, Pisa, Italy

3 Marzo 2016

3D from Photos

Our not-so-secret dream: obtain a reliable and precise 3D from simple photos…

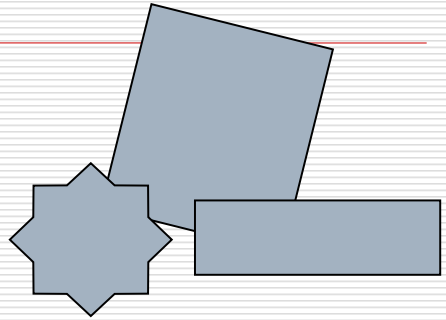
Why ?

Easier, less costly and more versatile than a 3D scanner

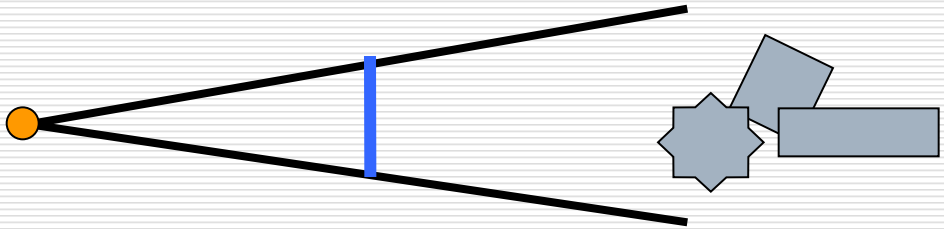
We can perceive 3D shapes from shapes, why should a computer not be able to do the same?…

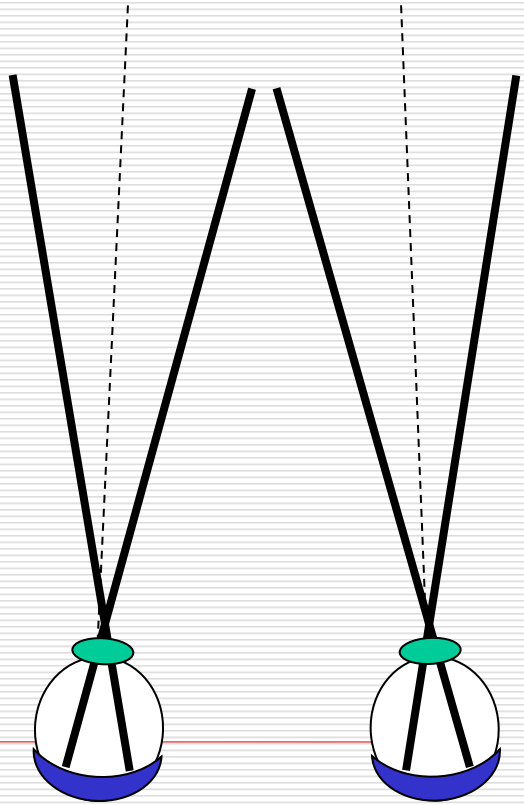
Perspective and Stereo

We do perceive the environment three-dimensionality thanks to our stereo vision AND the perspective projection that occurs in our eyes..



Using the geometric laws of this two phenomena we can build the 3D geometry of the scene



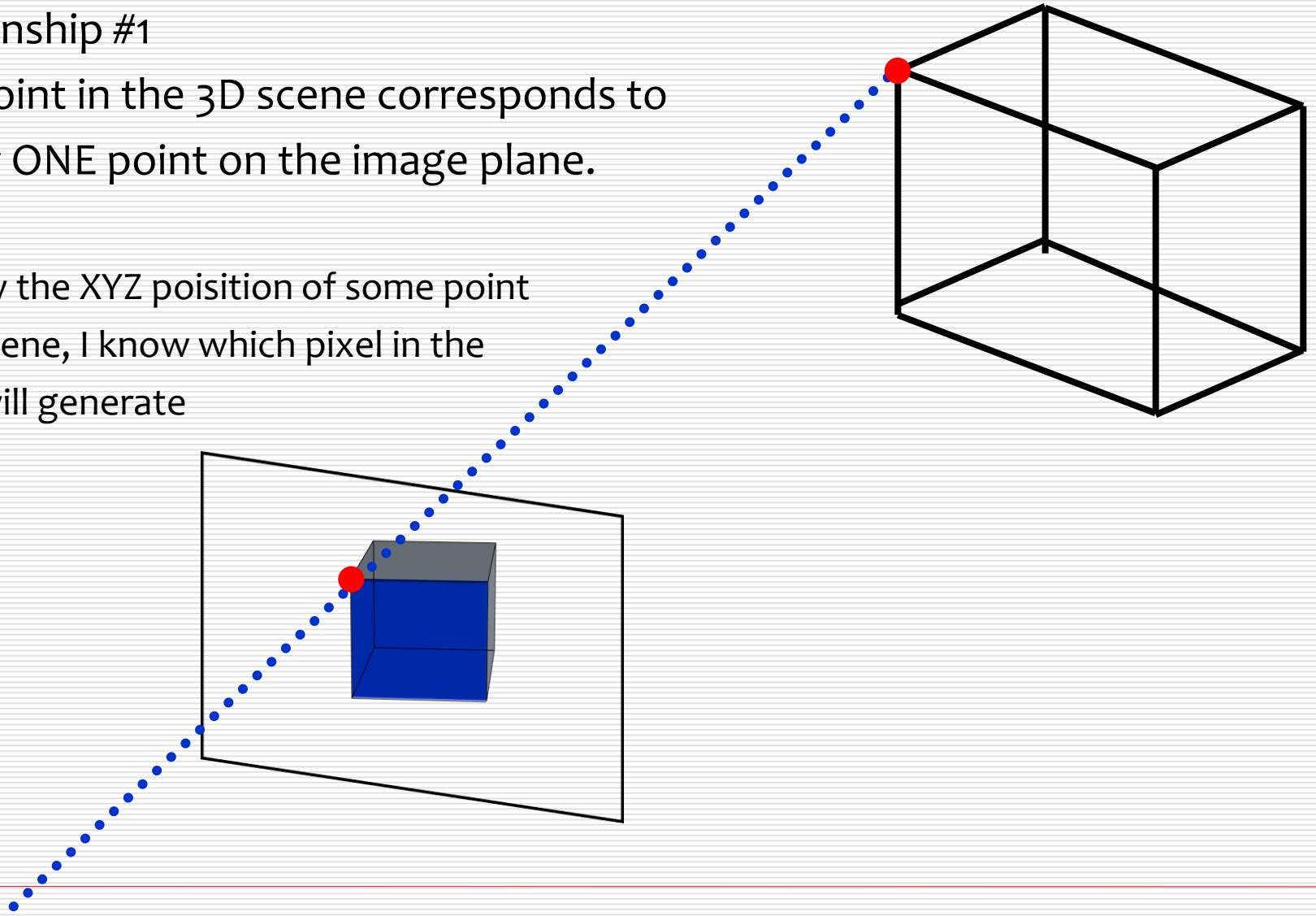


Perspective Projection

Relationship #1

Each point in the 3D scene corresponds to exactly ONE point on the image plane.

If I know the XYZ poisition of some point in the scene, I know which pixel in the photo will generate

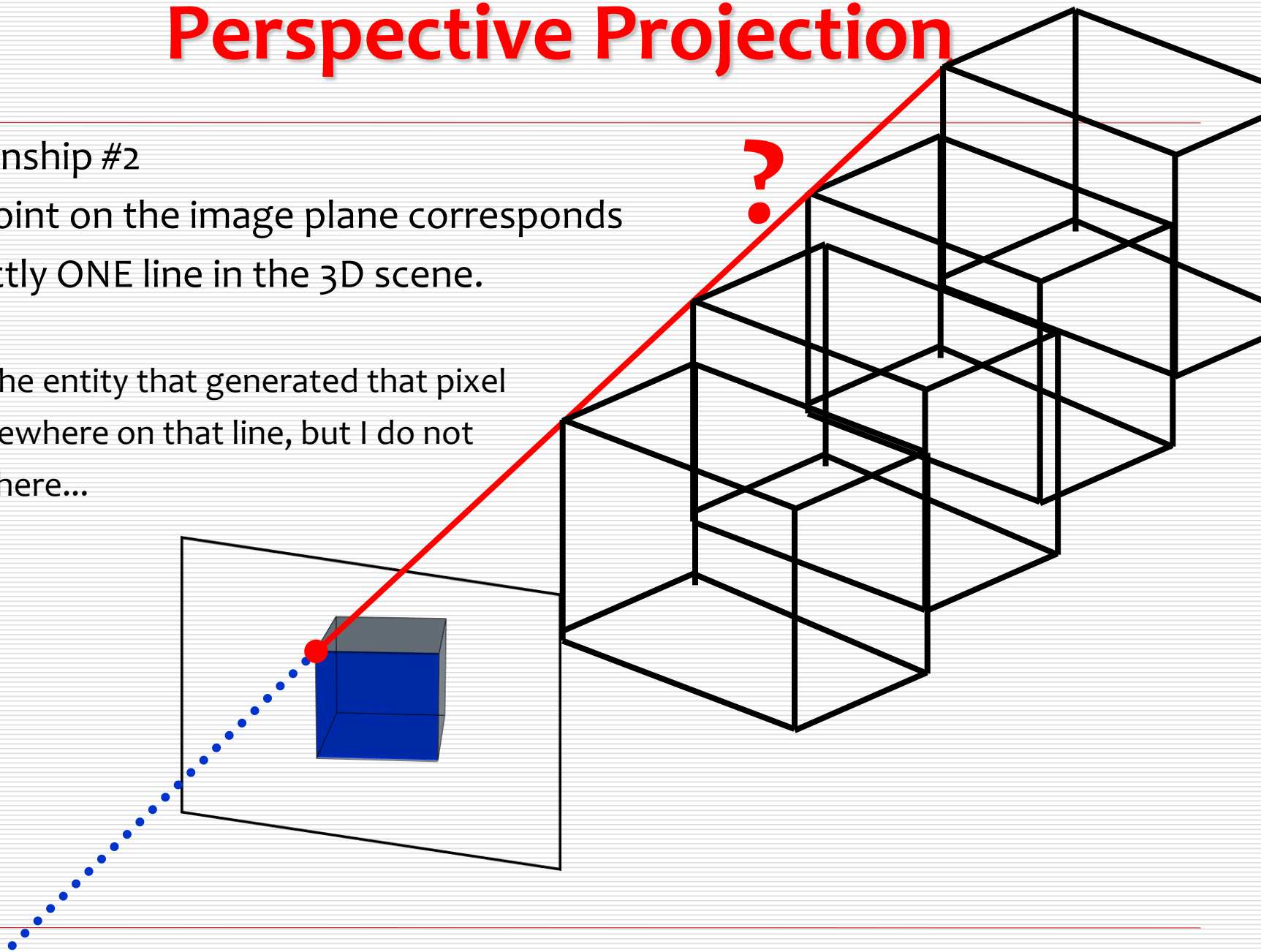


Perspective Projection

Relationship #2

Each point on the image plane corresponds to exactly ONE line in the 3D scene.

I know the entity that generated that pixel lies somewhere on that line, but I do not know where...

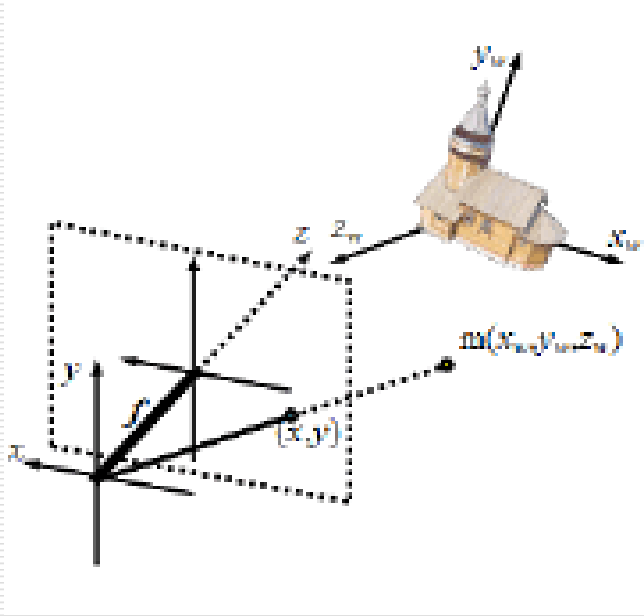


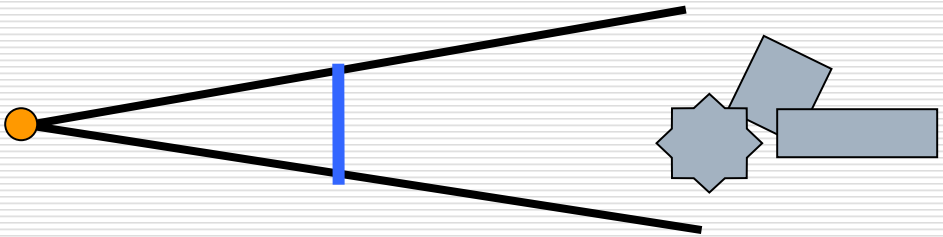
Perspective Camera

But to establish these correspondences, what do I need?

All starts here… a “camera” is defined by parameters:

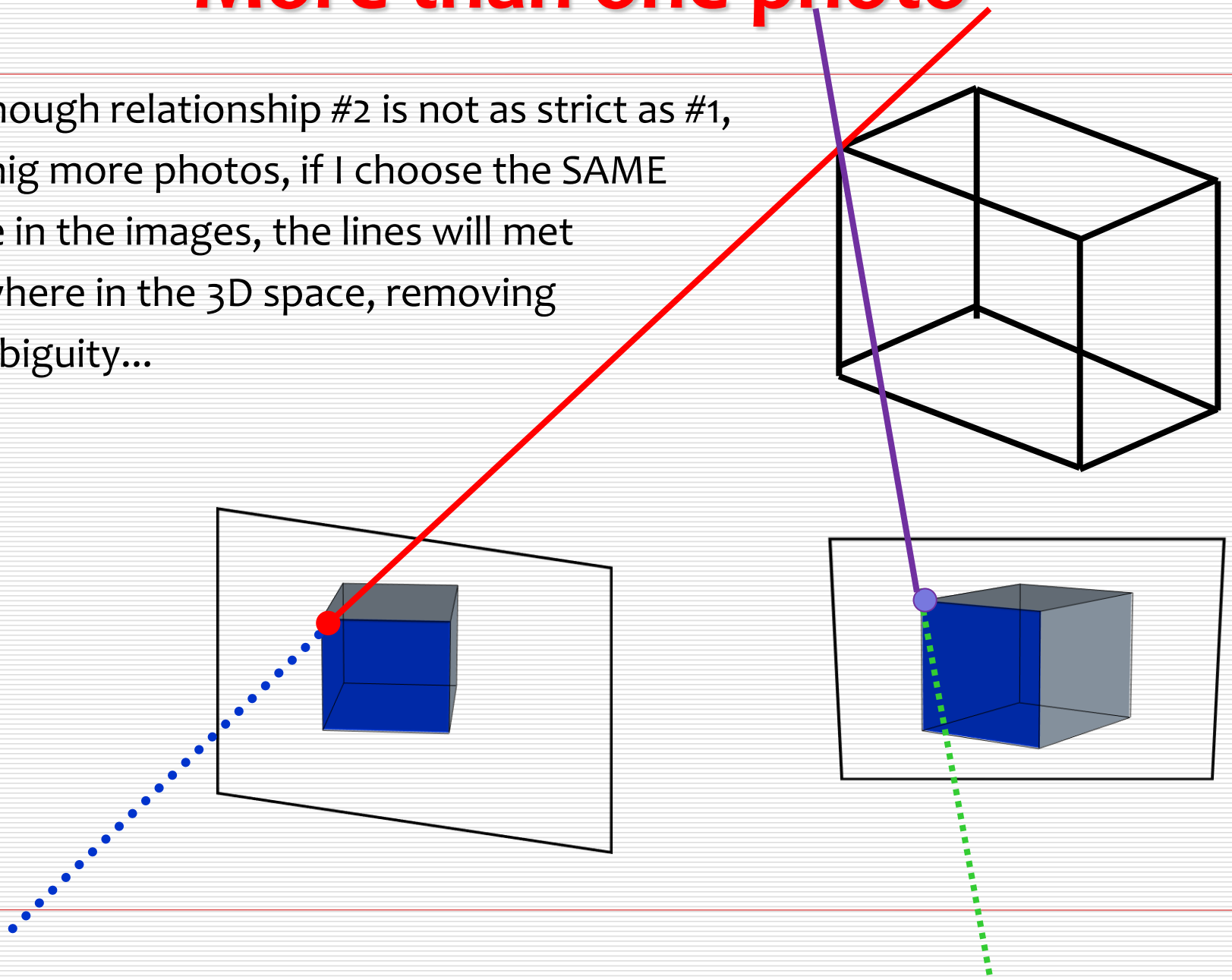
- Position
- Orientation
- Focal Length
- Viewport
- Distortion





More than one photo

Even though relationship #2 is not as strict as #1, combinig more photos, if I choose the SAME feature in the images, the lines will met somewhere in the 3D space, removing the ambiguity...



Ouroboros

But, in this way, It is a dog biting its tail!
In order to know the scene geometry, you do need to know something about the scene geometry…

We are lucky: by providing "semantic" information (non metric), it is possible to determine the camera Calibration and Orientation… the info needed is:

- (hand picked) image-to-image correspondences of points
- (hand picked) correspondences between image points and a known geometry

Camera Calibration

From some initial data (correspondences), determine all camera parameters

Completely mathematical process… however, not an easy one… there are lots of numerical problems.

Distortion

Very important component to obtain precise data

Generally, obtained by the use of some pattern.

Much easier mathematically if done separately from the extrinsics...

- It is possible to calibrate a camera and then use the intrinsics without calculating them every time
- On the market, pre-calibrated lenses and cameras (at a very high cost)

2 paths...

Two possible directions

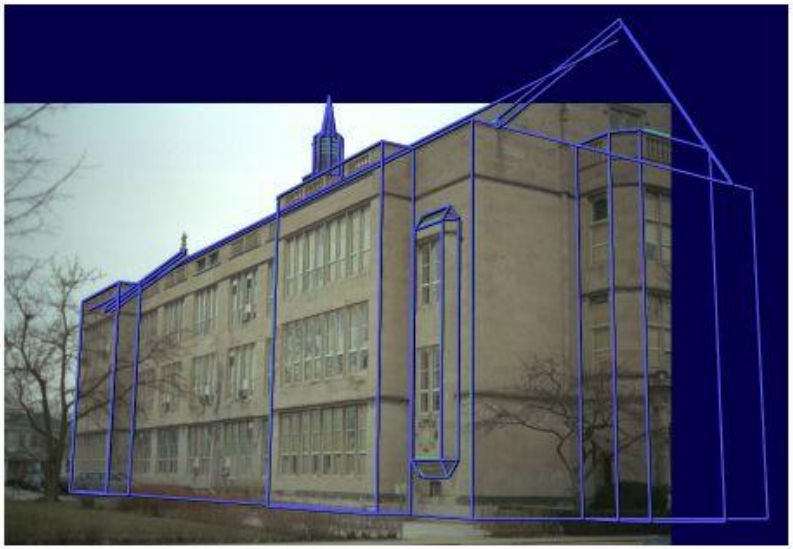
- Assisted Modeling
- Automatic Stereo Matching

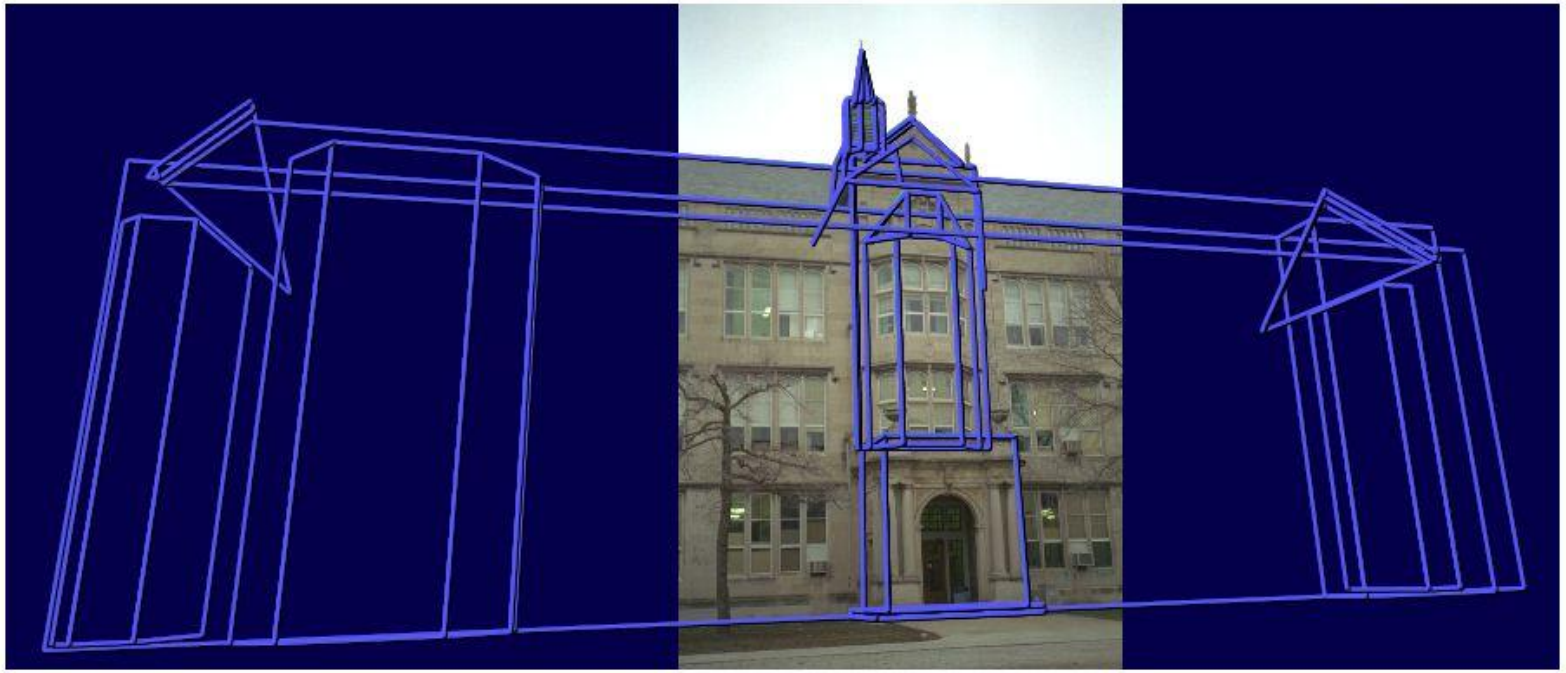
The underlying principles are the same... but they sprouted different kind of tools...

As usual, this distinction is becoming blurry...

...things ***do*** converge, after all

Assisted Modeling





Perspective & vanishing point

If in an image, the depicted object presents lines along two axis, we can easily detect vanishing point(s) and axis orientation…

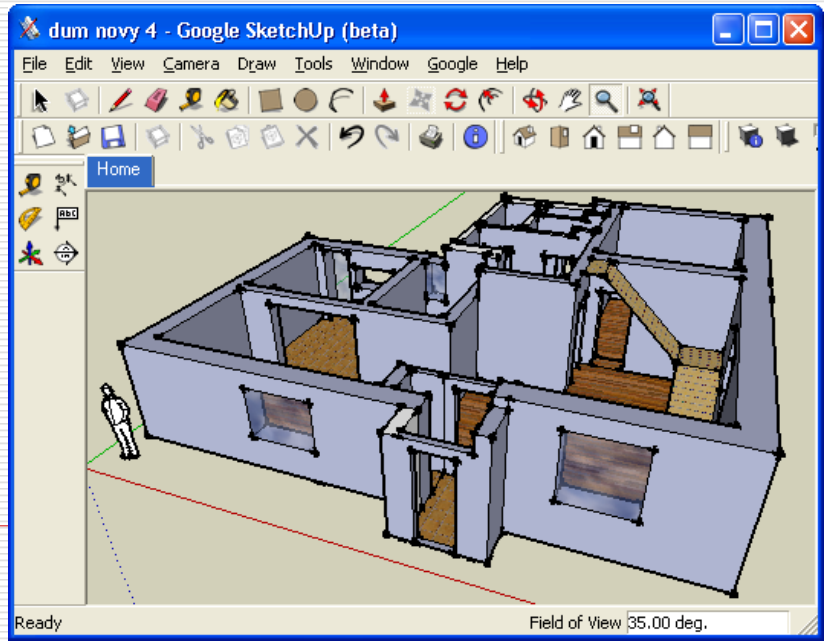
Not really much, but much better than nothing…

Sketchup

A very strange modeling tool.
Follows more the way a technical drawing is done on paper (reporting/referencing) than the usual 3D modeling metaphore.

Easier for people with a technical drawing/sketching background.
Easier for people with no experience in 3D modeling.

Focused towards modeling of buildings and mechanical entities...



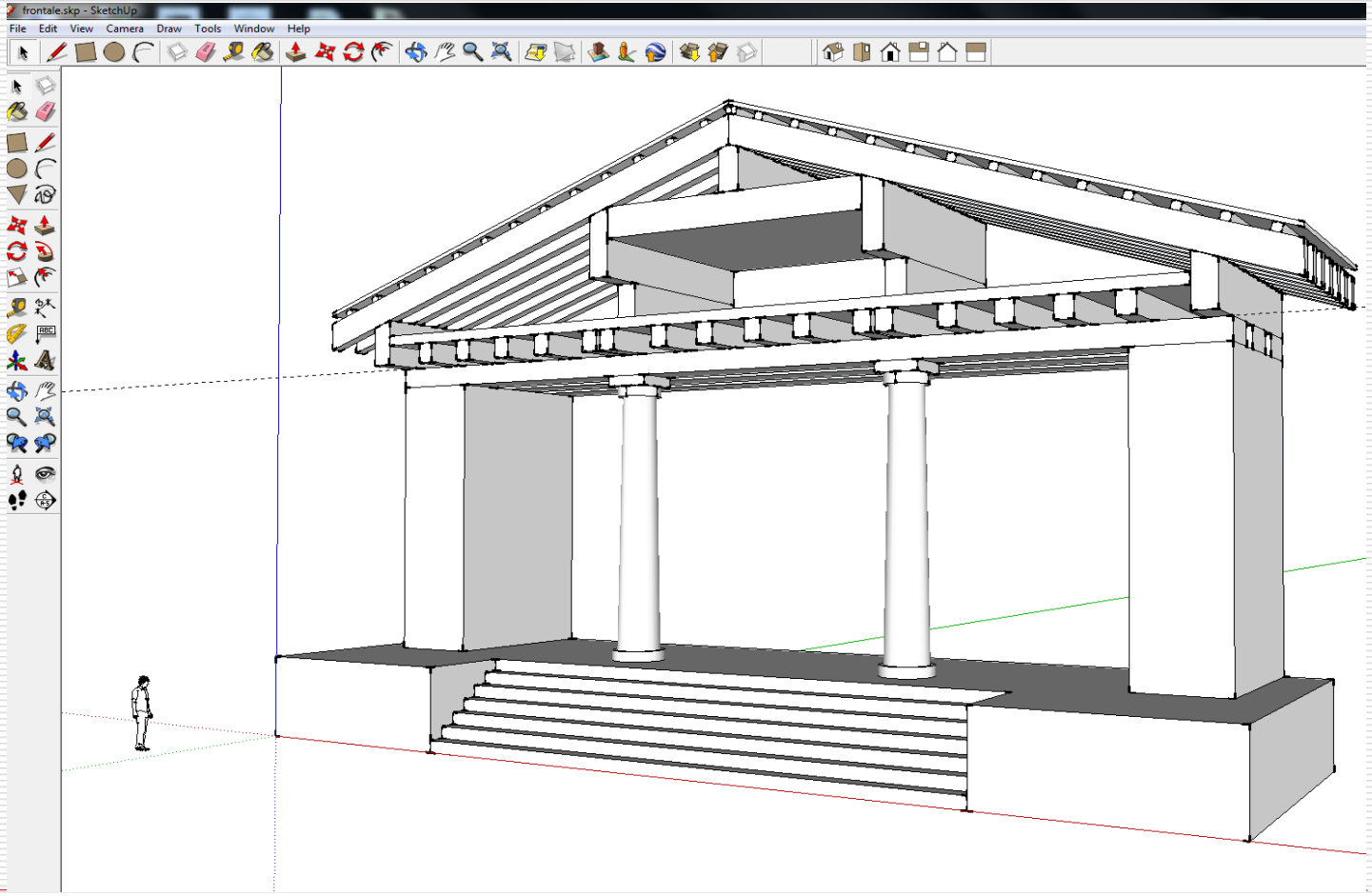
Acquired by Google some years ago, and then sold again...
distributed now as a semi-free tool

http://www.sketchup.com/

Sketchup – the Luni Temple

An ideal tool for very regular buildings... like this one

(regular does not equal *new*)



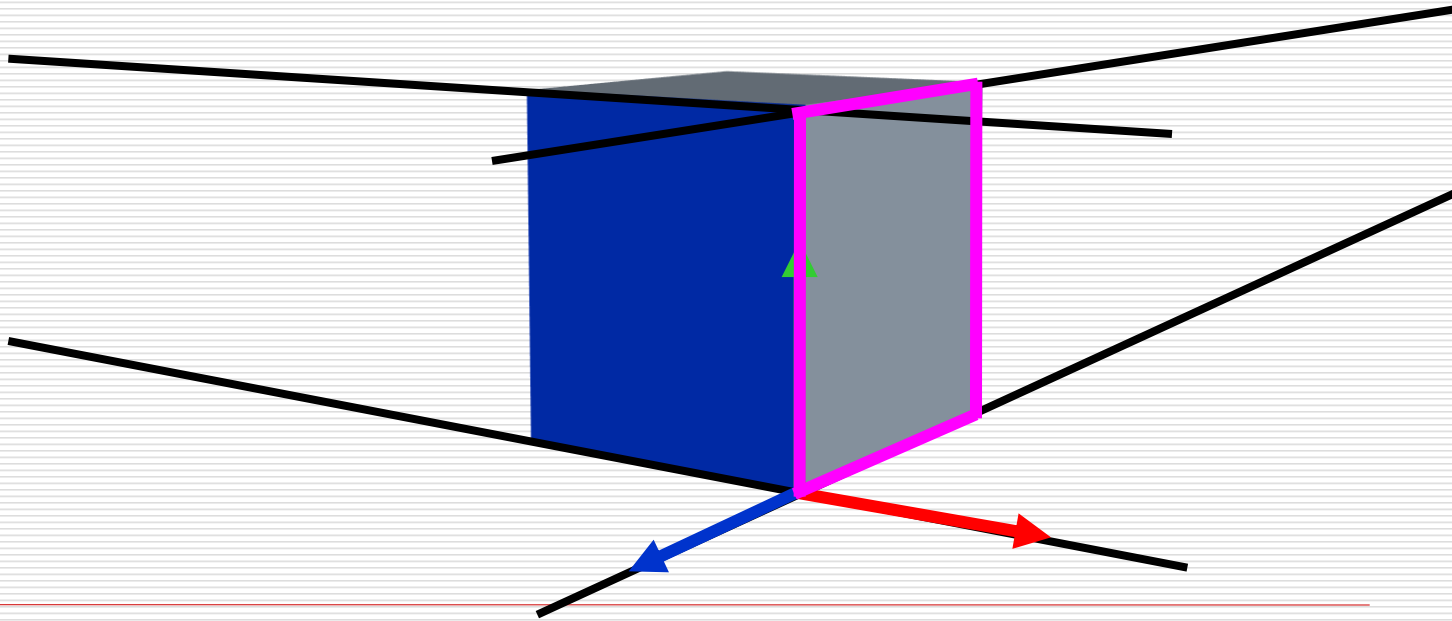
Sketchup Photo Match

Assited modeling from a SINGLE photo

- Calibration: axis and vanishing points markup
- Modeling: 3D drawing by axis/reference reporting

Partial calibration, with only a single photo, only the axis can be recovered.

SketchUp can be used to model by reporting/referencing



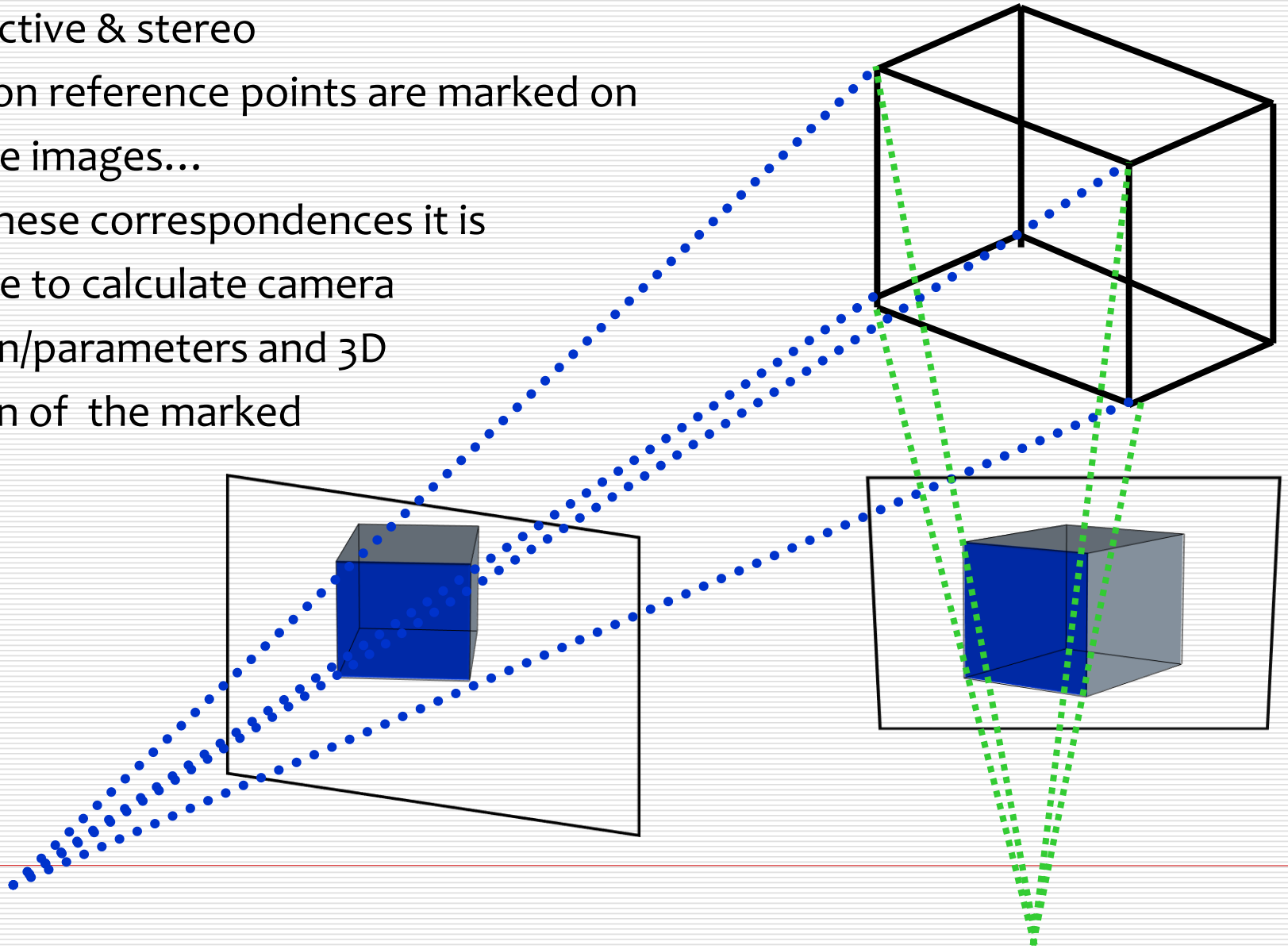
http://sketchup.google.com/

Photogrammetry

Perspective & stereo

Common reference points are marked on multiple images…

From these correspondences it is possible to calculate camera position/parameters and 3D location of the marked points



Photogrammetry

At first seems impossible, but with some care, the obtained precision is way below the projected size of a pixel.
For a building, it may be few millimeters (yes, it ***does*** depend on the size of the object)

Photogrammetry is the name of the principle; many different tools and approaches.

Often rely on calibrated cameras and markers...
Often used in conjunction with other measuring methods...

Photogrammetry

VI-9i (Polygon Editing Tool Ver.2.0)

< VI-9i Main standard accessories>

① Interchangeable Lenses (TELE, MIDDLE, WIDE)
② Field Calibration System
- Polygon editing software "Polygon Editing Tool ver.2.0"
- SCSI Cable

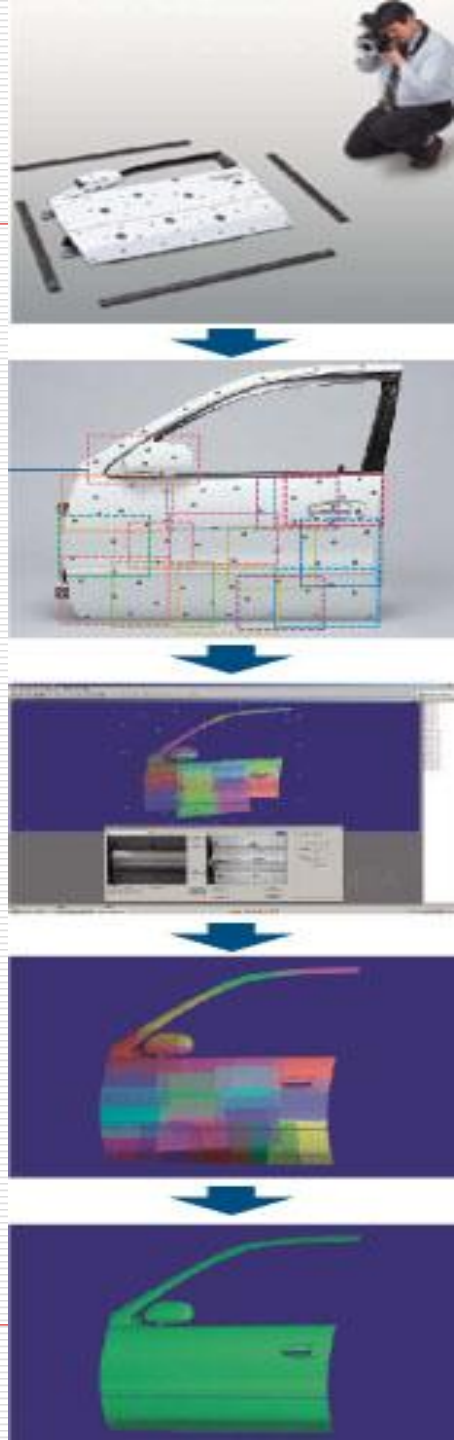
<Optional accessories>

③ Tripod Set
- Rotating Stage Set

※ Shape varies with sales region.

Photogrammetry System PSC-1 optional

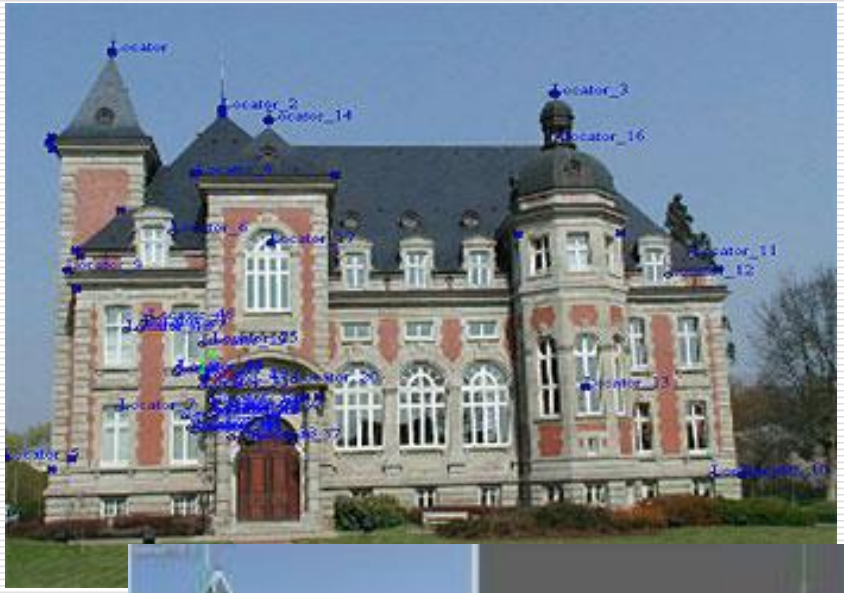
① Digital SLR Camera(14 megapixel)
② Ring Flash
③ Scale Bar Set
④ Camera Calibration System
⑤ Reference Marker Set
⑥ Coded Marker Set
- Compact Flash Card
- Photogrammetry software Photo Modeler KM



ImageModeler

Photogrammetry commercial tool

Points are marked on input images, camera are fully calibrated using these points, camera are calibrated and modeling can be done using the recovered 3D points

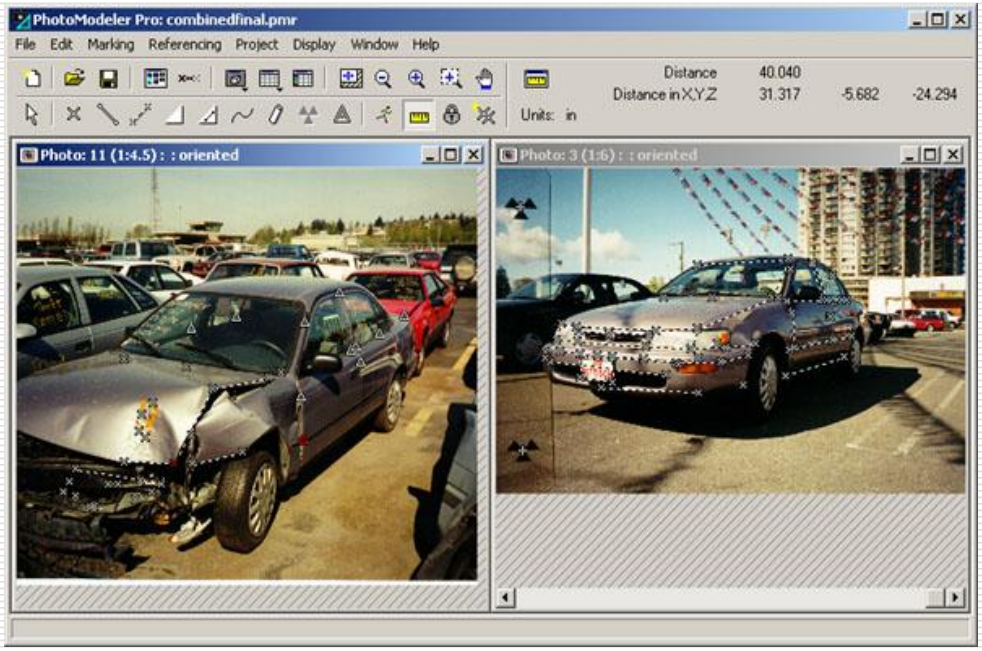


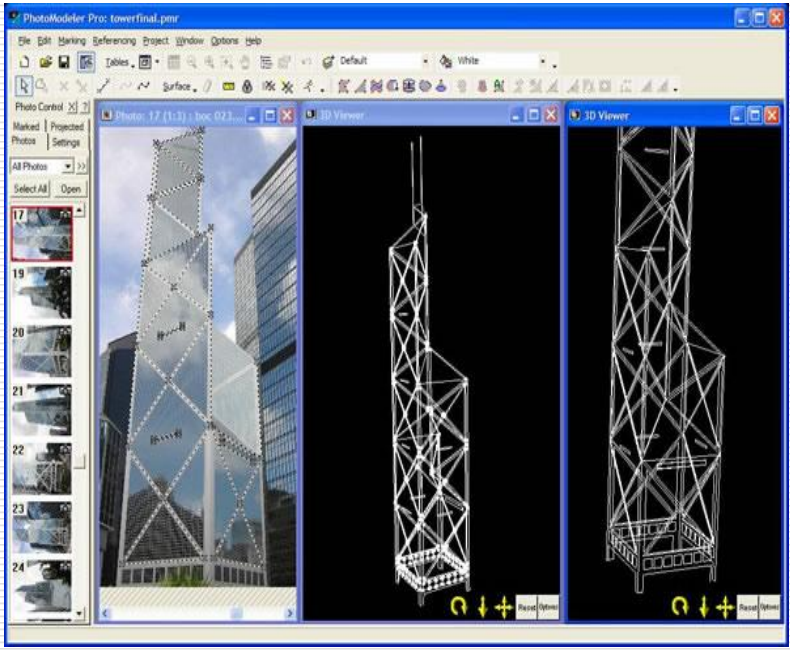
Acquired by Autodesk... now it costs three times the old price (with the same features)

PhotoModeler

Photogrammetry commercial tool

The tool for the professionals... Two steps: camera calibration (with markers and references) and camera pose estimation. Modeling and measuring with lots of different tools





Very, very, very complex to use...

XSIGNO

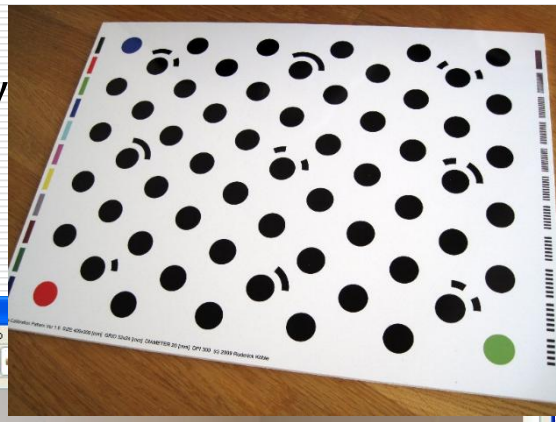
Shareware photogrammetry tool

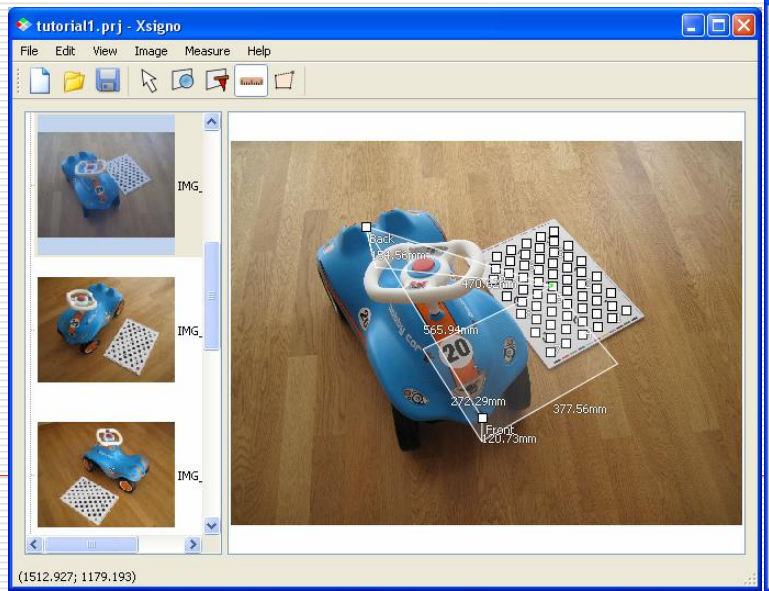
Oriented to measurement (and not modeling)

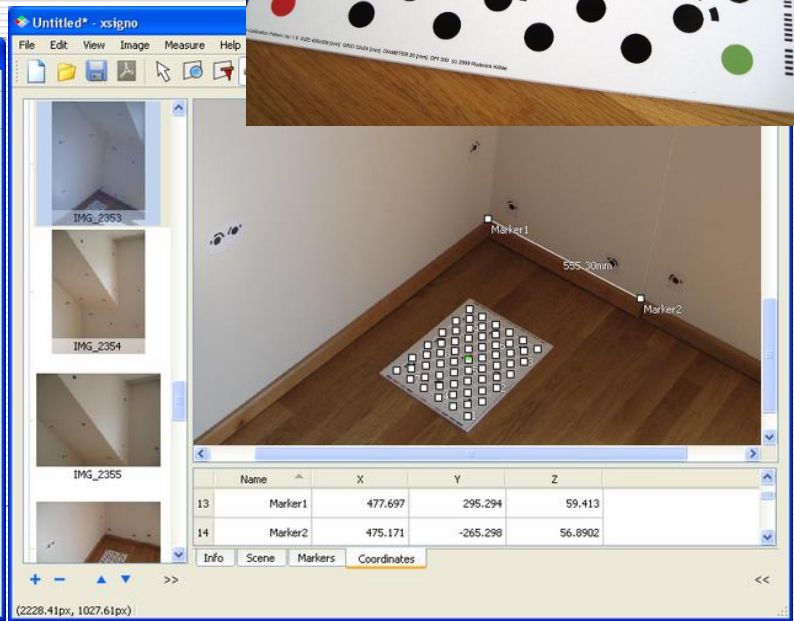
Intensive use of markers for camera calibration (intrinsics) and orientation (extrinsics)

Looks promising for a small software...

But the necessity to use markers may reduce its applicability



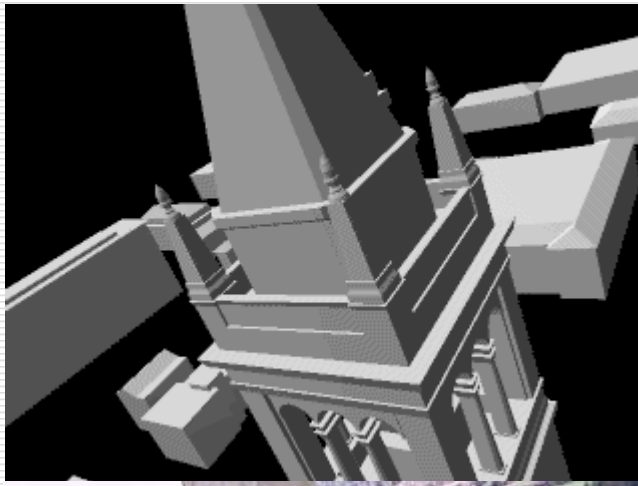




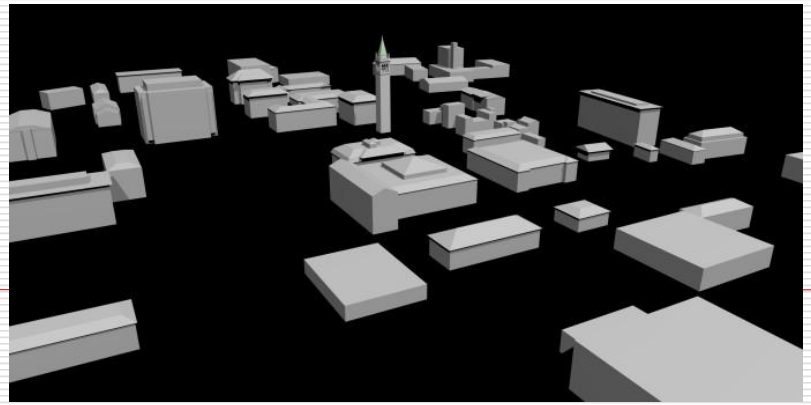
The “Campanile Movie”

A short movie from 1997, but with a large impact on the movie industry...

A photogrammetry-based model used for a video where real-world sequences are interleaved with digital renderings



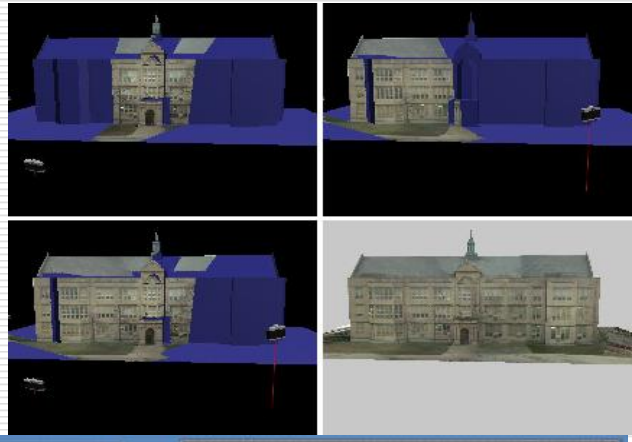
Seems quite naïve... but it was a demonstration of the maturity of image-based graphics...



Façade

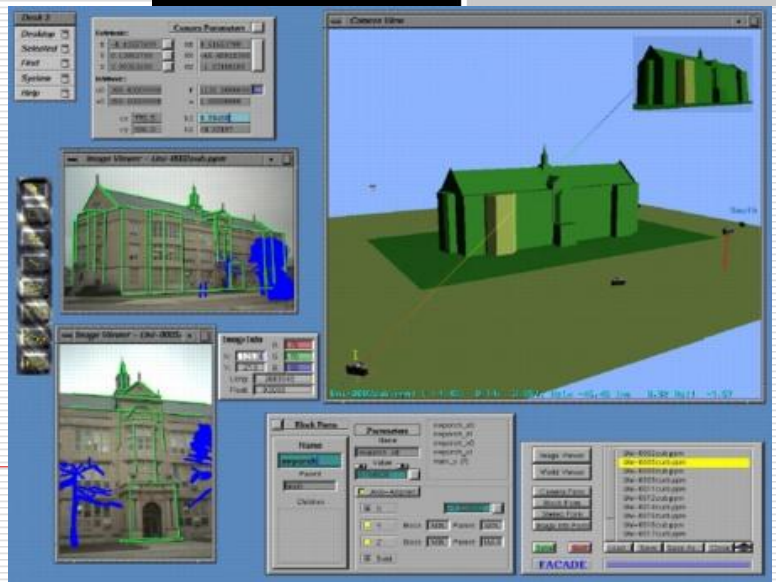
Component-based photogrammetry...

Correspondence points are used to place parametric “blocks” (cubes, pyramids, archs)



The basic principle is still photogrammetry, but it is focused on buildings (which represent ¾ of the market)

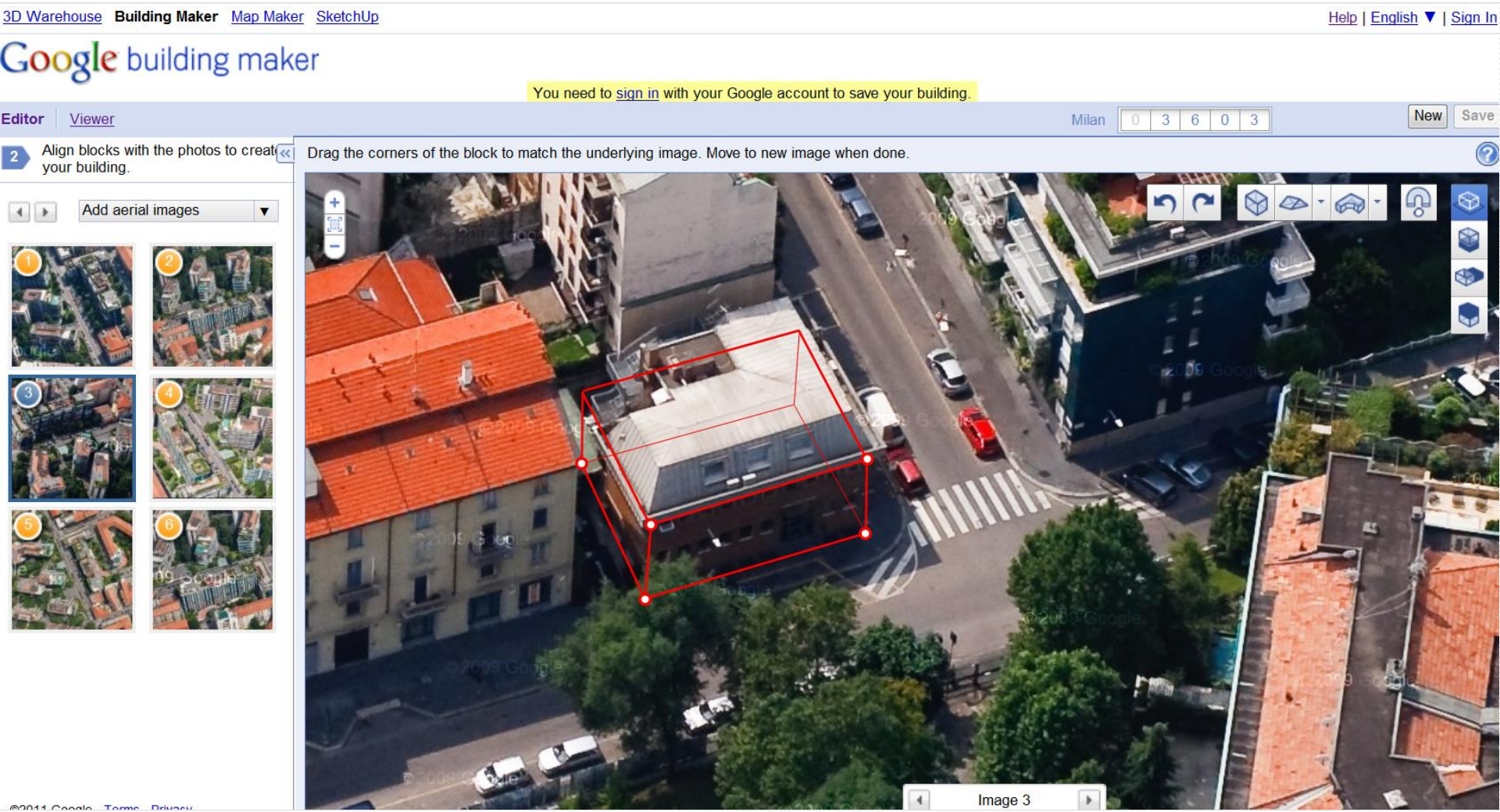
View-dependent texture mapping: color information is generated in realtime from the best photo resulting in a much more realistic animation



Google – Online modeling tool

Google Building Maker... a principle similar to Façade...

Parametric building blocks placed according to photos from different angles.



The sad news

NO free photogrammetry tools out there... sorry

IMAGEMODELER: acquired by Autodesk... Now component for other tools

PHOTOMODEL: very professional tool, good support, quite costly

XSIGNO: http://www.xsigno.com/

~~CANOMA: a very interesting tool acquired by Adobe and disappeared...~~

- try using demo/trial versions
- hope for the best...

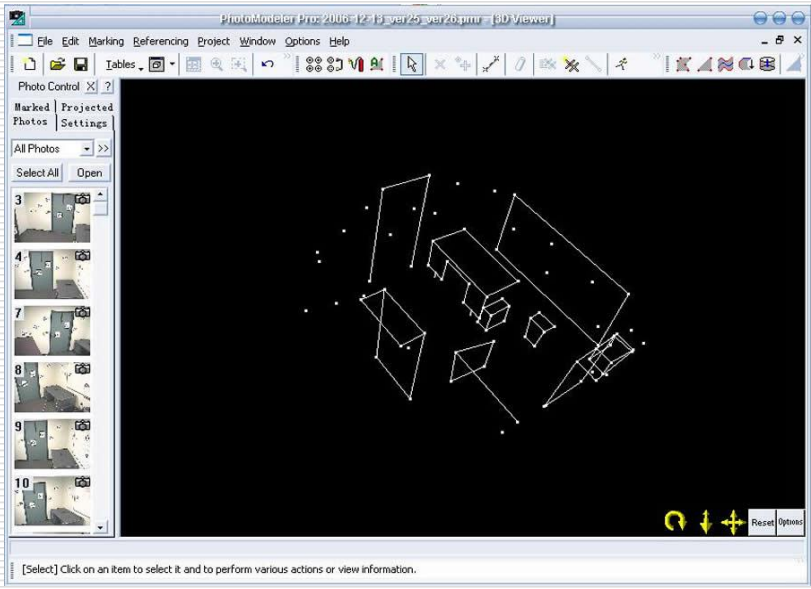
Recap

2 phases: CALIBRATION & GEOMETRY EXTRACTION

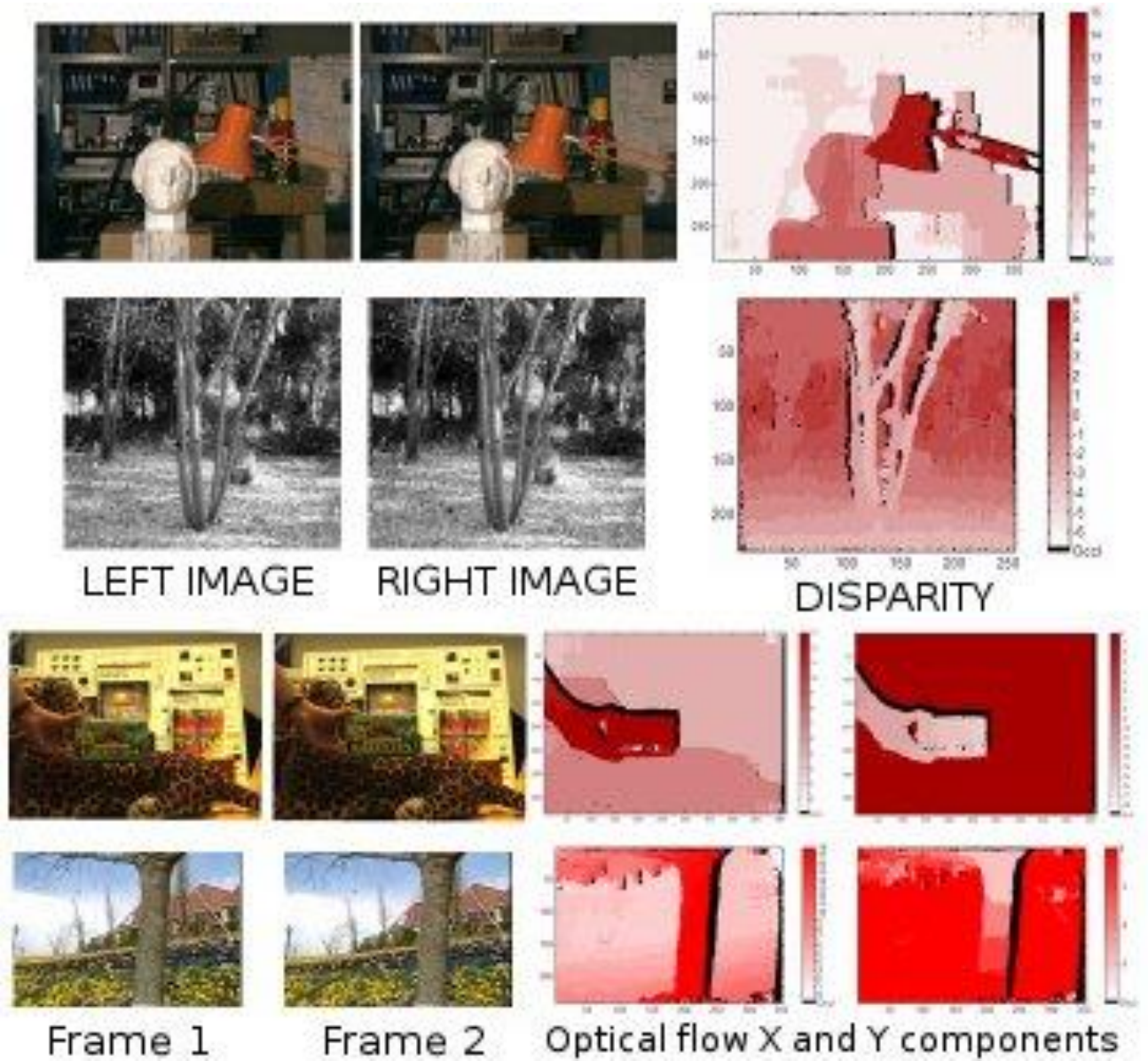
- both are based on correspondences
- first calibration step returns also some 3D points in the scene
- correspondences are made by user or recovered from markers in the scene

Results: a series of points in 3D

- very high precision
- only the marked points are recovered
- modeling is done on the recovered points



Automatic 3D extraction



Automatic

Idea: use some kind of automatic feature matching to perform a DENSE reconstruction...

Not just rely on user-picked points, but try to match the entire surface, **pixel by pixel**..

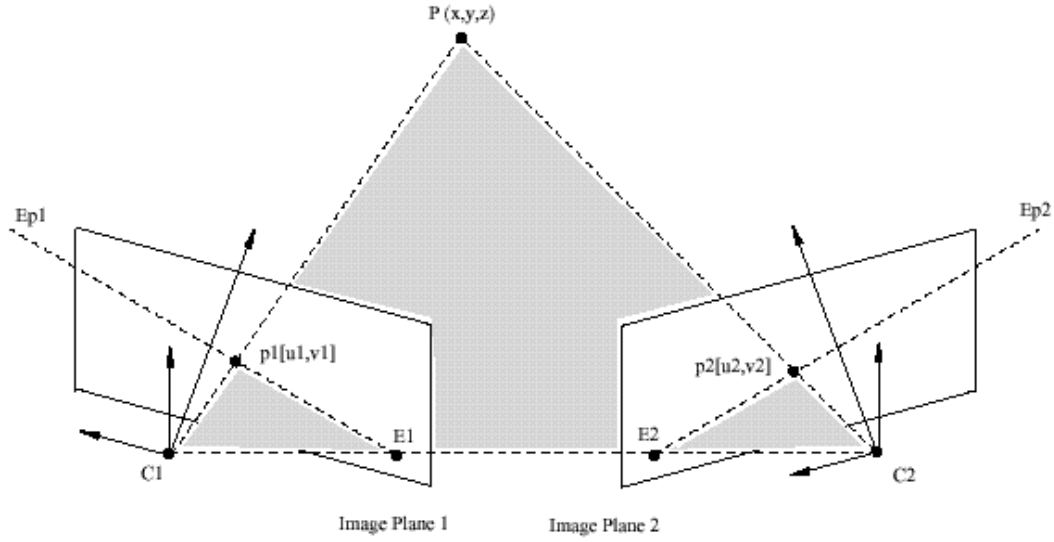
Basic situation:

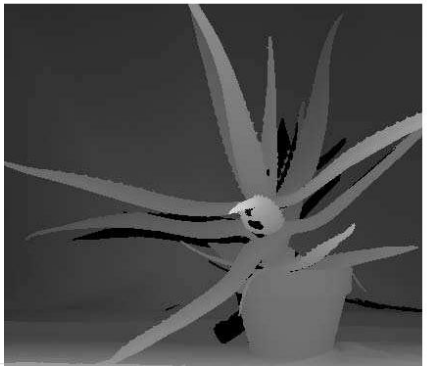
- camera calibration and position are known.
- automatic matching between image pixels for **DENSE** reconstruction

Automatic

Difficult ? Yes and no

- NO given registered cameras, a pixel P1 on image 1 will have its corresponding pixel P2 on image 2 laying on a known line. This is called EPIPOLAR line(thanks to the laws of perspective)
- YES changes in illumination, different exposition, flat-colored areas, noise and occlusion makes difficult to always pick the right candidate

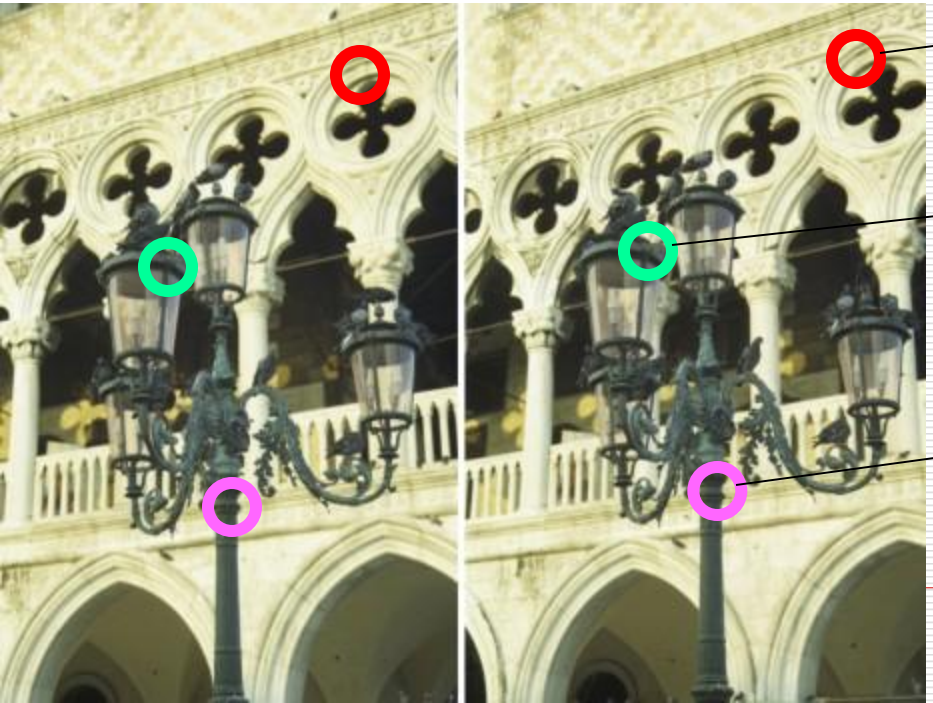




Shape from Stereo

Based on the same principle of human stereo vision: two sensors that perceive the world from slightly different position. From parallax it is possible to obtain a depth for each visible point

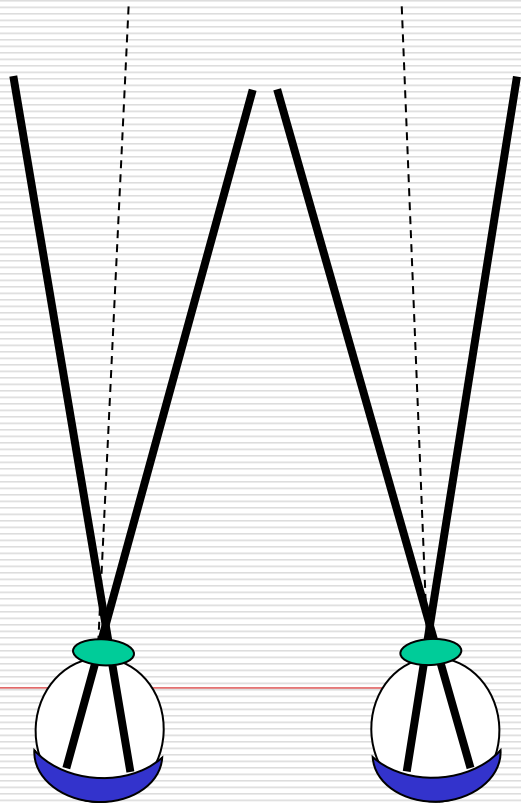
Our brain does this automatically… A machine can be programmed to do the same



Same position
=> background

High variance
=> close

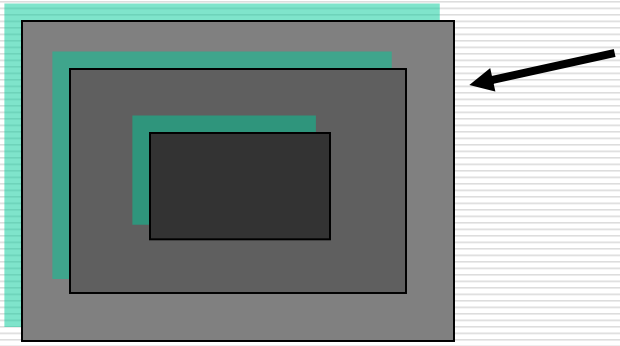
Mid variance
=> mid distance

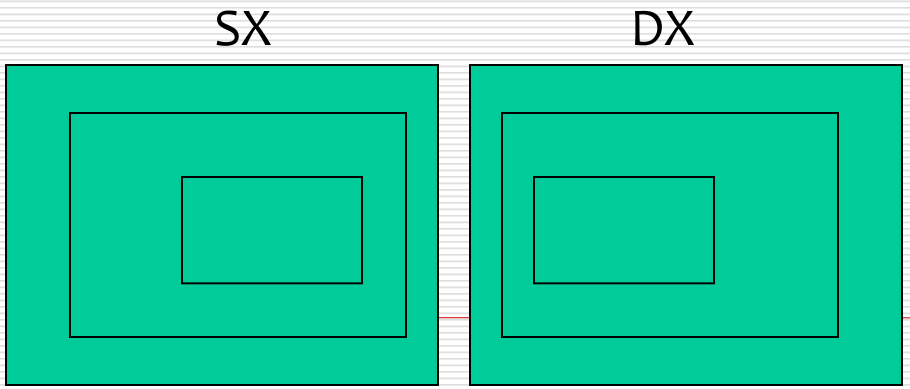


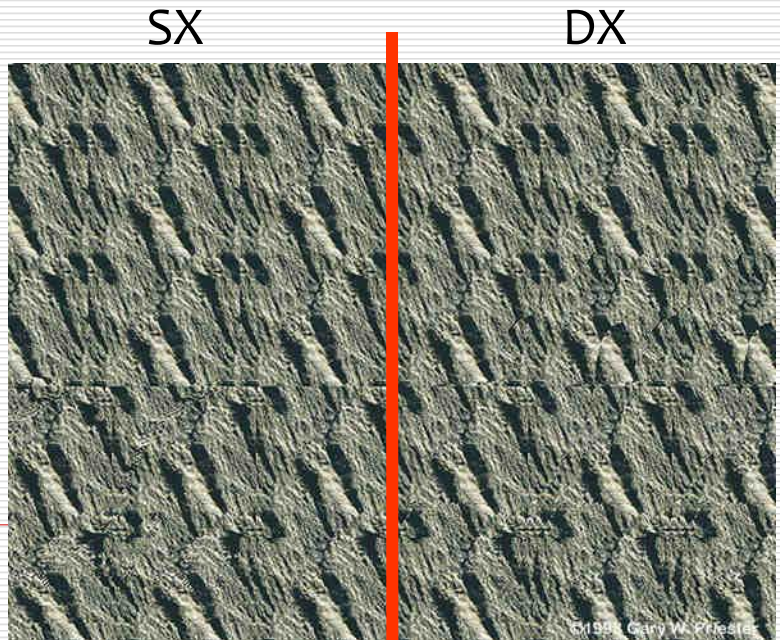
And i mean *automatically*

The brain does stereo match continuously... even when the input is random junk... this is the base of stereograms.

The patterns are constructed such that the matched points follows the disparity of a given depth map







Dedicated Devices

It is available on the market, some hw that is able to perform stereo matching (entirely or partially) and gives back a depth map. In practice it is formed by two (or more) synchronized cameras, plus a DSP.

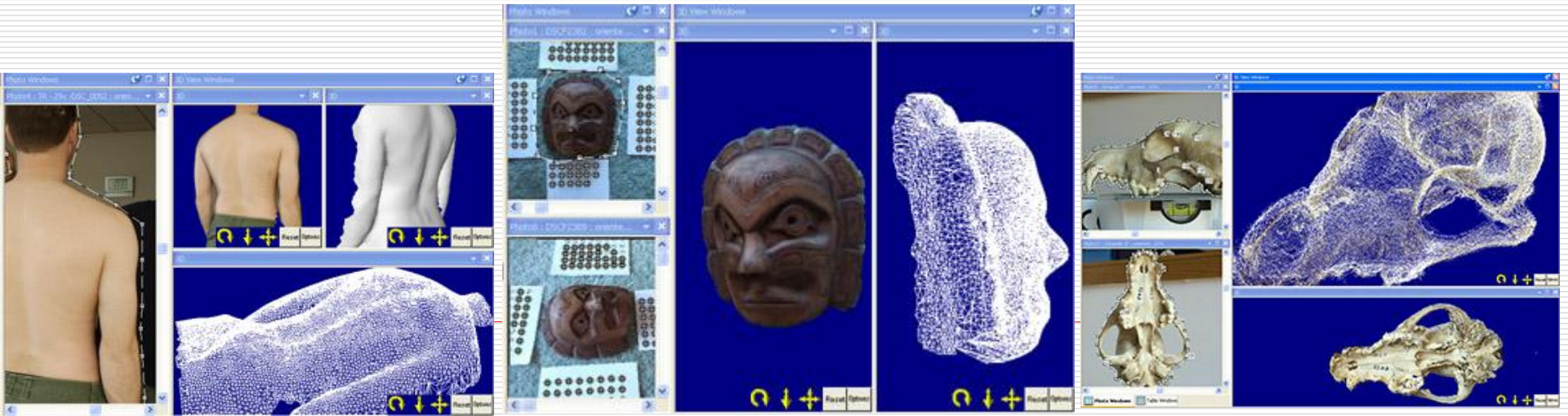
Normally used in robotics, because they are fast but not much precise… but hardware is getting better

PhotoModeler - dense

Photogrammetry commercial tool, 2nd version

The tool is the same as before... after doing the camera calibration, instead of using only user-picked points, the system does a dense-matching.

The result is quite similar to a range map... you need to do standard processing in order to obtain a 3D model...



MenciSoft

Commercial solution… italian product (quite unusual)
Three photos from a calibrated camera sliding over a very precise railing.
Easy to use, versatile (multi-scale), fast acquisition…

Good results, but very long processing time to obtain a final, complete model



MenciSoft




MenciSoft

- camera is calibrated (intrinsics)
- camera positions are known (extrinsics)

What remains is just DENSE matching...

The very regular camera placement helps the stereo matching

Result: fast, precise and reliable extraction of 3D data

The real pain is data processing :)

Fully automatic

A step further… also the camera position is unknown

1- automatic feature match between photos to obtain initial calibration
2- automatic dense matching to recover 3D data

The additional step is much more difficult than the dense step… finding GOOD features for image calibration is hard because, in this case, there is no epipolar geometry to help .

Modern image analysis techniques make it possible (local descriptors like SIFT)

Thanks for your attention…

Question Time