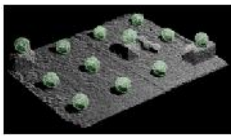
CARATTERIZZAZIONE DI SENSORI ATTIVI E PASSIVI LOW-COST

Fabio Menna, Fabio Remondino,
Silvio Del Pizzo, Erica Nocerino

3D Optical **M**etrology (**3DOM**)
Bruno Kessler Foundation (**FBK**), Trento, Italy
Email: fmenna@fbk.eu
Web: http://3dom.fbk.eu




Commercial low-cost sensors





- Active
  - **Kinect**, David 3D Scanner, NextEngine, etc. (triangulation-based)
  - D-Imager3D, Swissranger, CamCube, etc. (TOF-based)
  - Direct delivering of 3D data
  - Often coupled with open-source software (SDK) for sensor control, data acquisition, gesture recognition, surface modeling, etc.

- Passive (image-based)
  - **Fujifilm real 3D W series**, Bumblebee, etc.
  - Provide for 2D images or directly 3D data using some real-time processing



Commercial low-cost sensors

Microsoft Xbox Kinect

- *10 million sold in 5 months since November 2010*
- *Cost < 150 €*

Fujifilm REAL 3D W1

- *Cost < 250 €*

Commercial low-cost sensors

Microsoft Xbox Kinect

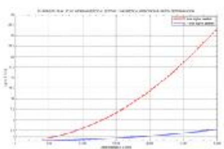
- *Active triangulation based plus speckle pattern decorrelation*

  Nominal baseline 74 mm

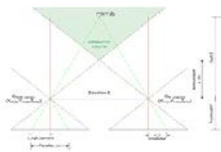
Fujifilm REAL 3D W1

- *Passive triangulation based*

  Nominal baseline 77 mm



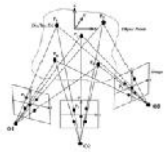
- *Standardized calibration procedures*
- *Theoretical accuracy analysis*
- *Characterization*
- *Performance evaluation*

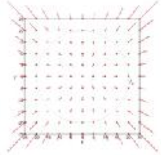


Sensor characterization

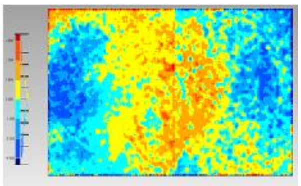
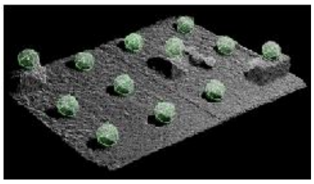
- Investigate the performances of the instrument
- Determine its intrinsic parameters
- Evaluate the accuracy potential

- Calibration of **imaging sensors** → photogrammetric bundle adjustment





- Calibration of **active sensors** → flatness measurement error, best fitting, etc.

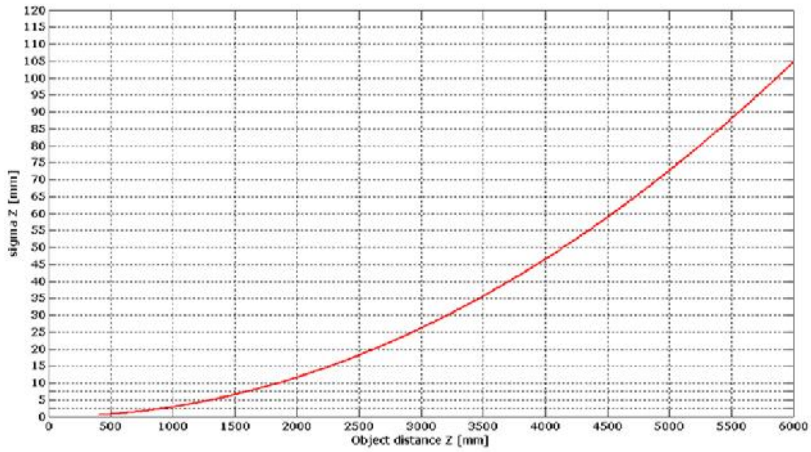

Low-cost active sensors

- Active - Microsoft Kinect (ca 100 Eur)
  - 10 milioni di unita' vendute in 5 mesi dal Novembre 2010
  - Nuvole di punti 3D fino a 30 fps alla risoluzione VGA (ca 300k points)
  - 2 CMOS sensors + proiettore di luce strutturata
  - Principio di misurazione basato sulla combinazione di triangolazione e la decorrelazione di speckle pattern
  - Distanze operative suggerite: 1.2-3.6 m
  - 3 brevetti alla base (Prime Sense LTD):
    - "Three-dimensional sensing using speckle patterns" - US Patent (2009)
    - "Range mapping using speckle decorrelation" - US Patent (2008)
    - "Method and systems for object reconstruction" - International Patent (2007)

	RGB camera	**IR camera**
Sensor	Aptina MT9M112	Aptina MT9M001
Sensor type	CMOS	CMOS
Sensor size (active imager)	3.58 mm x 2.87mm	6.66 mm x 5.32 mm
Pixel size	2.8 µm	5.2 µm
Raw image format	1280x1024 px	1280x1024 px
Output image resolution	640x480 px	640x480 px
Nominal focal length	2.9 mm	6 mm
FOV H	63 degrees	57 degrees
FOV V	50 degrees	45 degrees

Triangulation measurement principle

- **Close-range sensors** -> triangulation measurement principle

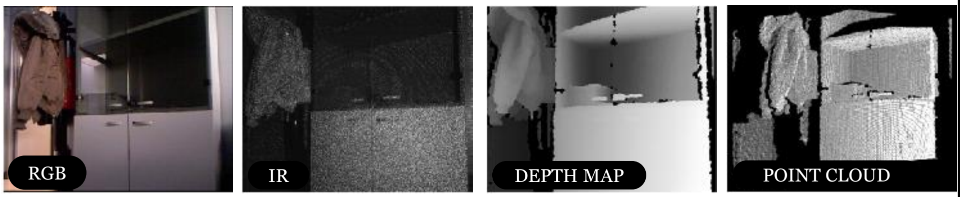


$$\sigma_Z = \frac{Z^2}{f \cdot b} \cdot \sigma_{px} \approx \frac{Z^2}{f \cdot b} \cdot \varepsilon \cdot \mu$$

- μ is the pixel size of the imaging device;
- ε is the sub-pixel resolution

KINECT characterization

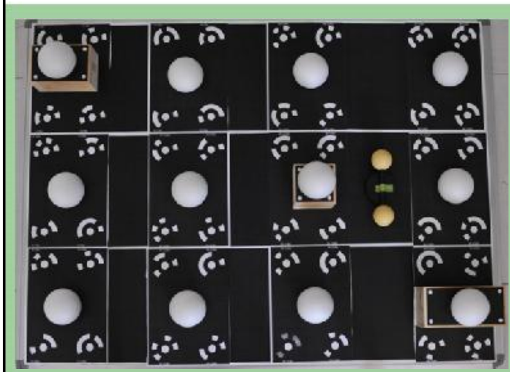
- Kinect delivers primarily disparity maps (640 x 480 px) at 11 bit
- Metric coordinates (X, Y, Z) derived with different approaches / equations (OpenNI, Libfreenet, etc.)
- Multi-sensor device:
  - Calibration of the passive imaging sensors (IR and RGB)
  - Calibration of the depth map
- Accuracy evaluation with respect to higher accurate sensors



[Menna, F., Remondino, F., Battisti, R., Nocerino, E., 2011: *Geometric investigation of a gaming active device*. Proc. of Videometrics, Range Imaging and Applications XI, SPIE Optical Metrology, Vol. 80850G-1-15]

KINECT characterization

- Calibration of the passive imaging sensors



CALIBRATION FRAME:

(1000 mm x 800 mm x 200 mm)

- **12 SPHERES** (90 mm diameter)
  - 9 spheres on a main plane
  - 3 spheres on wooden blocks
- **60 CIRCULAR CODED TARGETS**
- **1 SCALE BAR**

TESFIELD MEASURED WITH A 24 MPX NIKON D3X CAMERA:

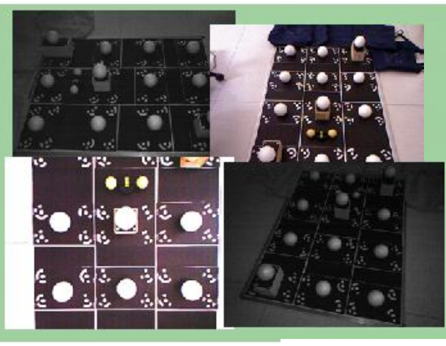
overall theoretical precision of the computed object coordinates resulted in

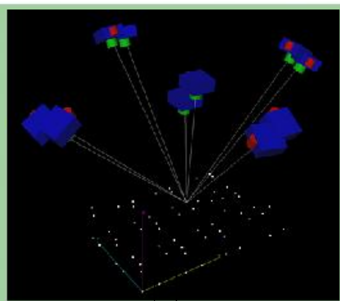
σxy = 0.015 mm, σz = 0.023 mm for circular targets

σxy = 0.04 mm, σz = 0.06 mm for the spheres

KINECT characterization

- Calibration of the passive imaging sensors



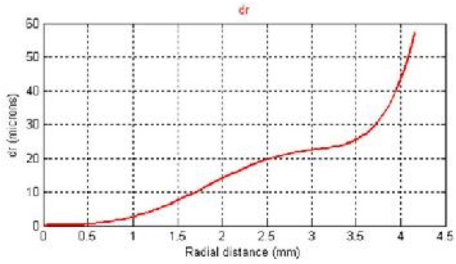


P1 and P2 not statistically significant for the RGB sensor

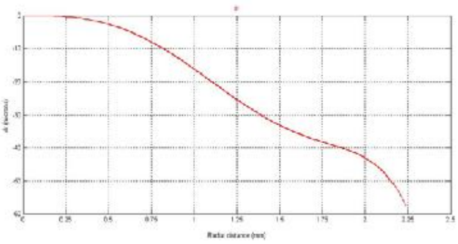
	RGB camera		IR camera	
	Value	Std	Value	Std
f	2.9114 mm	0.003 mm	6.0792 mm	0.007 mm
x_0	0.0346 mm	0.0007 mm	0.0488 mm	0.005 mm
y_0	-0.0315 mm	0.0008 mm	0.0480 mm	0.005 mm
k1	-2.310e-002	5.3e-004	3.253e-003	1.8e-004
k2	7.720e-003	2.7e-004	-3.720e-004	2.3e-005
k3	-8.246e-004	4.6e-005	1.347e-005	9.7e-007
P1	-	-	-2.708e-004	4.4e-005
P2	-	-	-1.999e-004	4.2e-005

KINECT characterization

- Calibration of the passive imaging sensors – Radial distortion profiles



IR CAMERA



RGB CAMERA

- IR camera is used with down-sampled image resolution (640x480) → 10μm pixel size
- Radial distortion at the border is less than 5 pixel → **internal correction?**

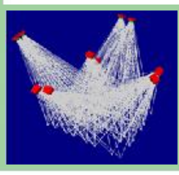
KINECT characterization

- Calibration of the passive imaging sensors – Relative orientation

ASYMMETRICAL RELATIVE ORIENTATION

RGB camera to IR camera

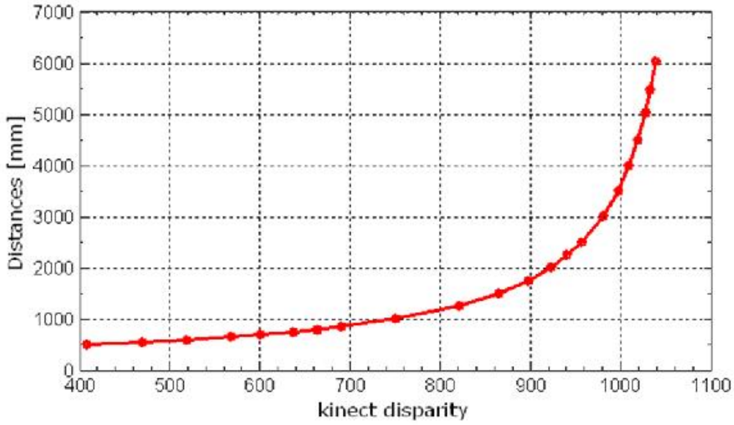
	IR camera	RGB camera
X [mm]	0	25.3869
Y [mm]	0	1.4971
Z [mm]	0	-4.3238
ω [degrees]	0	0.022308
φ [degrees]	0	0.217726
κ [degrees]	0	-0.274297



- Roto-translation of the cameras in order to fix the perspective center of the **IR camera as origin** and its Euler angles to null
- Relative orientation useful to correctly map the color information onto the point cloud

KINECT characterization

❑ Calibration of the depth map

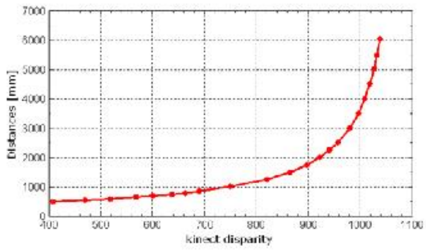


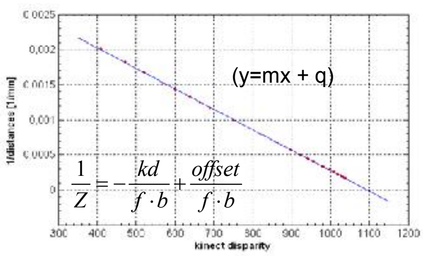
22 Kinect's depth maps of a plane orthogonal to the optical axes of the sensor are compared with the distances measured using a Leica Disto A6 (accuracy ±1.5 mm)

KINECT characterization

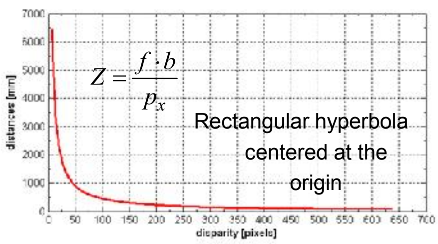
❑ Calibration of the depth map

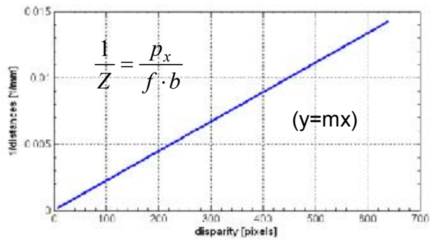
KINECT DEVICE





STANDARD STEREO TRIANGULATION SYSTEM

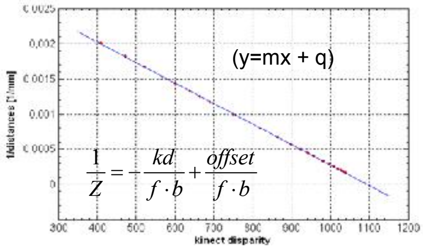


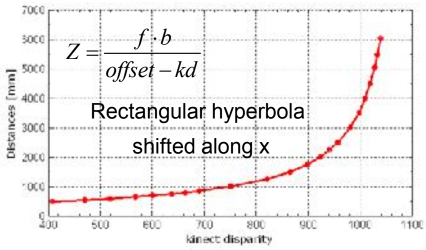


KINECT characterization

- Calibration of the depth map

KINECT DEVICE





$$\begin{cases} m = -\frac{1}{f \cdot b} \\ q = \frac{offset}{f \cdot b} \end{cases} \Rightarrow offset = -\frac{q}{m}$$

- m = -2.911482·e-6;
- q= 0.003187
- offset= 1094.7

- b=-1/(8·f·m) =73.448 mm

$$Z = \frac{1}{m \cdot kd + q}$$

$$X = (x - x_0 + \Delta x_{AP}) \cdot \frac{Z}{f}$$

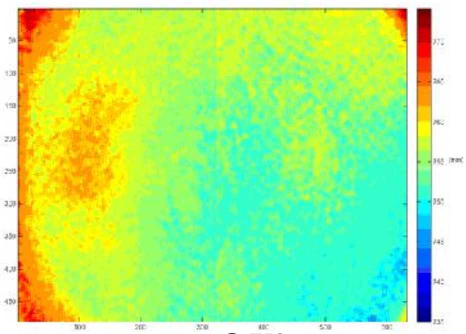
$$Y = (y - y_0 + \Delta y_{AP}) \cdot \frac{Z}{f}$$

Metric 3D coordinates of a point P (X,Y,Z) in the Kinect point cloud

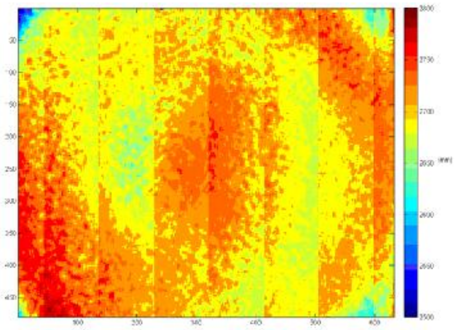
KINECT geometric performance tests

PRECISION OF THE KINECT IN DEPTH MEASUREMENTS:

Comparison between a **flat wall** and the **point cloud delivered by Kinect**



range map @ 750 mm



range map @ 2750 mm

Error > 40 mm @750 mm versus a theoretical precision of σ_z=1.5 mm expected

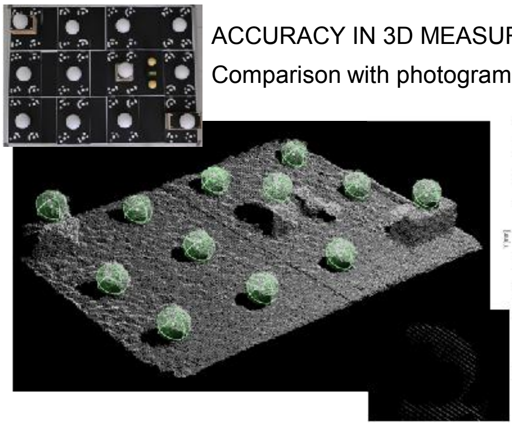
Error > 300 mm @2750 mm versus a theoretical precision σ_z=20 mm expected

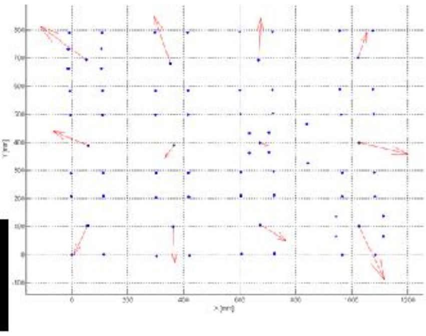
LARGE DEPTH ERRORS AT THE BORDERS

KINECT geometric performance tests

ACCURACY IN 3D MEASUREMENTS:

Comparison with photogrammetry on the 3D calibration frame



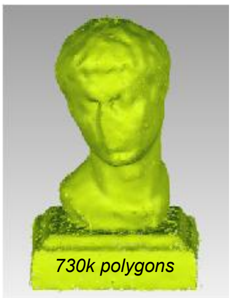


- ❑ Calibration frame recorded from 10 different positions at an average distance of 1.5 m
- ❑ 3D point clouds delivered by the Kinect compared with the photogrammetric 3D coordinates:
  - **uncorrected**: $\sigma_{XY} \approx 4$ mm, $\sigma_Z \approx 2$ mm and a systematic scale factor of 1.015
  - **corrected** (with the calibration parameters): $\sigma_{XY} \approx 2.5$ mm, $\sigma_Z \approx 2$ mm

KINECT geometric performance tests

ACCURACY IN 3D MODELLING:

3D comparison on a small statue (white matte plastic, ca 35 cm height and 20 cm wide)



KINECT at 50cm (theoretical $\sigma_z \approx 0.7$ mm)

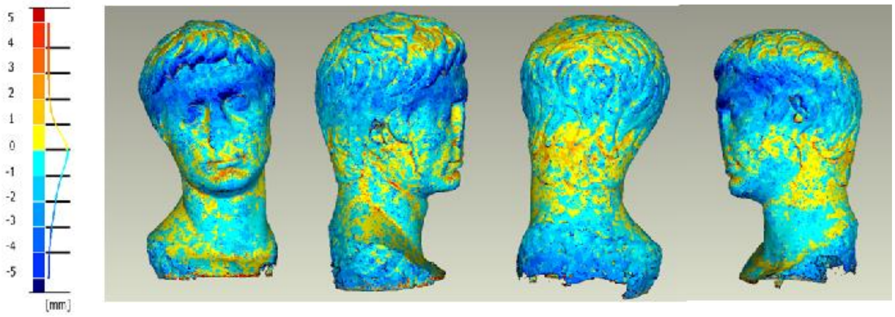
NEXTENGINE (accuracy $\sigma_z \approx 0.3$ mm)

- 19 Kinect range maps, aligned into a unique surface model (std = 1 mm)
- NextEngine 3D laser scanner used as reference
- Comparison after ICP alignment of the 2 datasets (std = 0.8 mm)

KINECT geometric performance tests

ACCURACY IN 3D MODELLING:

3D comparison on a small statue



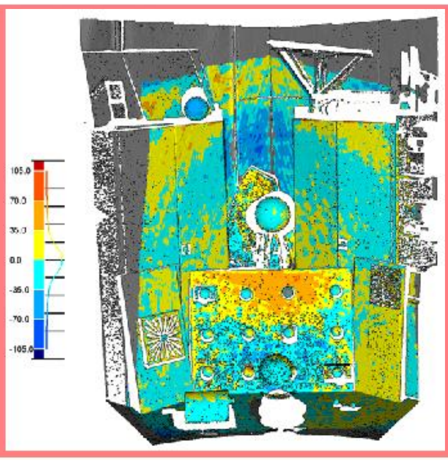
- Differences follows a Gaussian distribution slightly shifted toward the negative values probably due to a scale effect
- **Band artifacts** visible on the 3D point clouds

KINECT geometric performance tests

ACCURACY IN 3D MODELLING: 3DOM LAB CORNER

Scene volume 2m x 2m x 3m

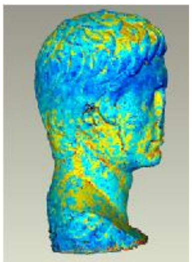
Theoretical precisions variable in the range [7,35] mm

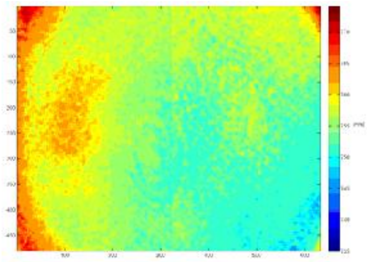


- Scene recorded with a TOF Laser Scanner Leica Scanstation 2
- The 3D point cloud of the kinect aligned on the scanstation point cloud (std=20 mm)
- Histogram of the Euclidean distances between the two point clouds follows a Gaussian distribution within the expected theoretical accuracy (min/max ±3σ)

KINECT considerations

- ❑ Active sensors with metric performances in the cm range up to 2-3 m (if corrected)
- ❑ Flatness problems at the borders of the acquired area
- ❑ Z-values (depth) improvement with Look-up-Table approach
- ❑ Mobile mapping possibilities with SLAM approach for real-time range mapping
- ❑ Strange striping effects when comparing the Kinect clouds with some reference data
- ❑ Best active low-cost sensor available on the market (data quality vs price vs performances)





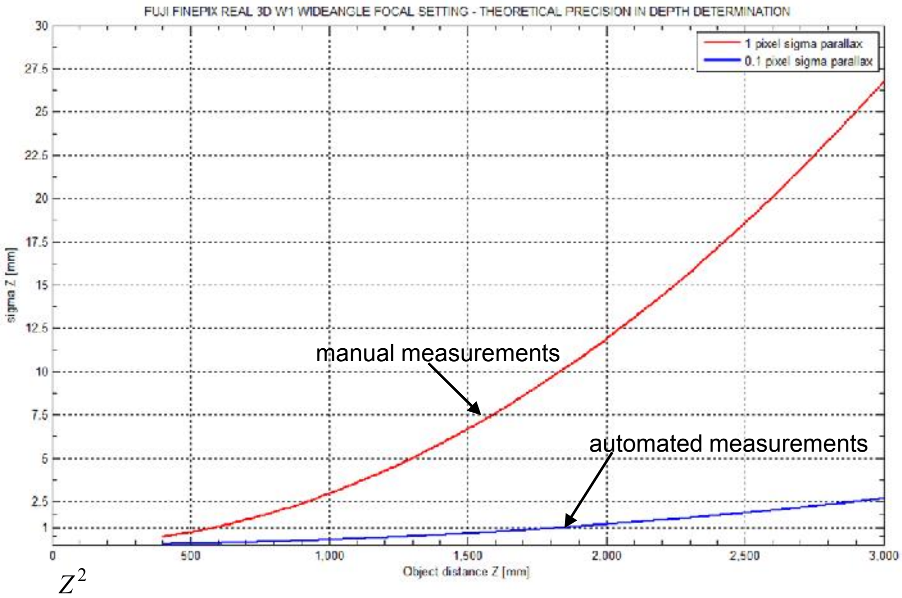
Low-cost passive sensors

- ❑ Passive (image-based) – Fujifilm Real 3D W1 (ca 500 Eur)
  - ❑ Twin-lens CCD system
  - ❑ 10 Mpixel / camera
  - ❑ 1/2.3" CCD
  - ❑ ca 7 cm baseline
  - ❑ 3D LCD monitor
  - ❑ 1 shot, 2 images → 3D data (Agisoft Stereo)

FUJIFILM characterization

- ❑ Twin-lens CCD system (1/2.3” CCD)
- ❑ 10 Mpixel / camera (3648x2736 pixel)
- ❑ ca 7 cm **baseline**
- ❑ Provide for a **stereo pairs** which can be directly converted into 3D “metric” data
- ❑ Calibration of both imaging sensors to improve the **metric performances**
- ❑ Uncertainty of the pixel size dimension:
  - ❑ 1/2.3” CCD → pixel size = 1.5 micron
  - ❑ Exif tag: 1.7 micron
- ❑ Calibration at **4 different focal length** settings:
  - ❑ Widest (minimal focal length)
  - ❑ Widest + 1 step
  - ❑ Middle (widest + 4 steps)
  - ❑ Tele (widest + 9 steps, i.e. max focal length)

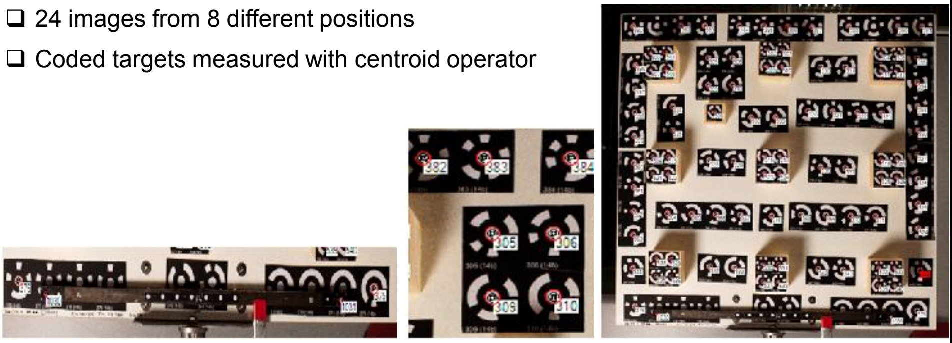
FUJIFILM characterization



$$\sigma_Z = \frac{Z^2}{f \cdot b} \cdot \sigma_{px}$$

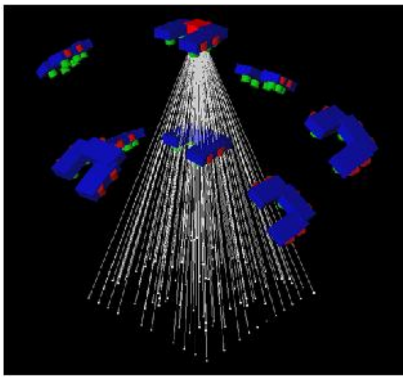
FUJIFILM characterization

- Testfield:
  - rigid square panel
  - 800 x 800 x 200 mm
  - ca 110 circular coded targets
  - 1 scale bar
  - 3D coordinates of the targets measured with a Nikon D3x (theoretical precision sigma_xyz = 0.022 mm)
- 24 images from 8 different positions
- Coded targets measured with centroid operator



FUJIFILM characterization

- Bundle adjustment results for minimal focal length

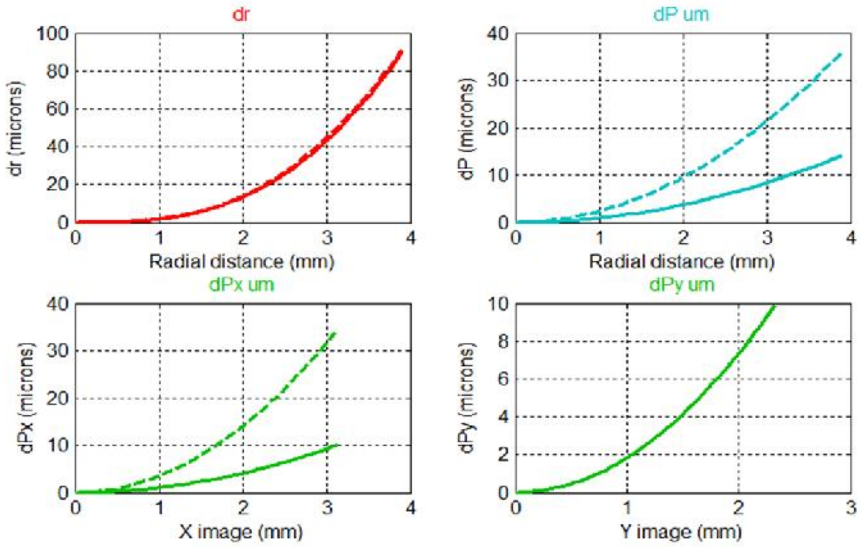


	LEFT camera		RIGHT camera	
	Value	Std	Value	Std
f	7.3469 mm	0.0005 mm	7.3655 mm	0.0005 mm
x_0	-0.0121 mm	0.0006mm	-0.0787 mm	0.0006 mm
y_0	-0.0974 mm	0.0006 mm	-0.0580 mm	0.0005 mm
k1	1.707e-003	9.5e-006	1.805e-003	8.4e-006
k2	-1.282e-005	5.6e-007	-1.768e-005	5.1e-007
k3	-	-	-	-
P1	1.683e-004	3.8e-006	1.058e-003	3.2e-006
P2	2.867e-004	3.4e-006	-1.735e-004	2.9e-006
A	1.2466e-002	3.658e-005	1.2596e-002	8.677e-005
S	-	-	-	-

- sigma0 = 0.21 px
- K3 and Skew factor (non-orthogonality) not significant
- Affinity scale factor in X necessary to reach convergence in the bundle adjustment
- RMSE X,Y,Z = 0.048 mm, 0.042 mm, 0.062 mm (reference is Nikon D3X)

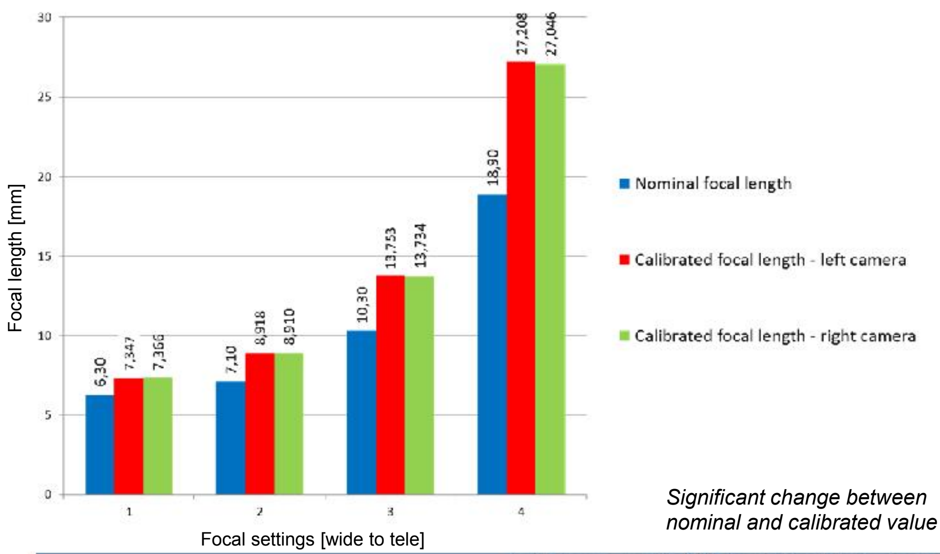
FUJIFILM characterization

- Bundle adjustment results for <u>minimal focal length</u> – distortion curves (left & right)



FUJIFILM characterization

- Bundle adjustment results for <u>different focal settings</u> – nominal vs calibrated

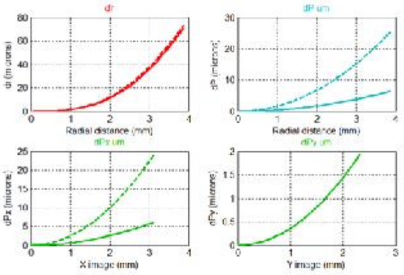


Significant change between nominal and calibrated value

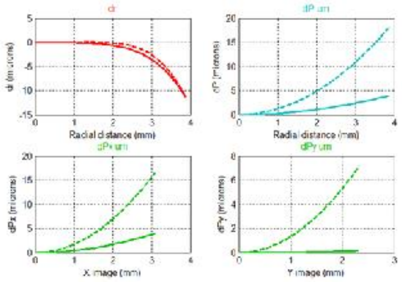
Fabio Menna - Low Cost 3D: Sensori Algoritmi e Applicazioni

FUJIFILM characterization

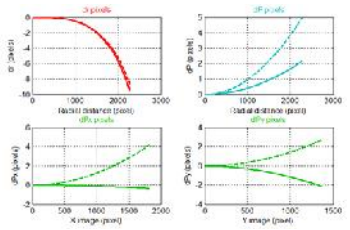
- Bundle adjustment results for different focal settings – distortion curves



Focal length = 8.9 mm



Focal length = 13 mm

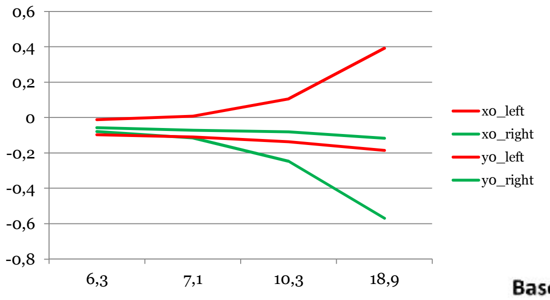


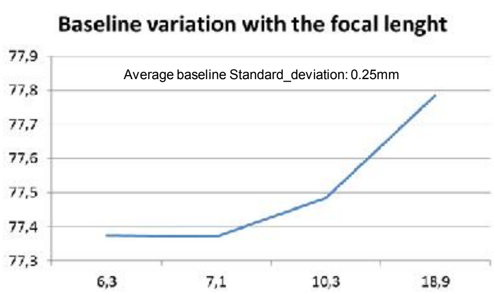
Focal length = 27.2 mm

- Similar behavior for radial curves of left and right camera
- Different behavior for the tangential distortions
- At minimal zoom radial distortion is up to 80 microns

FUJIFILM characterization

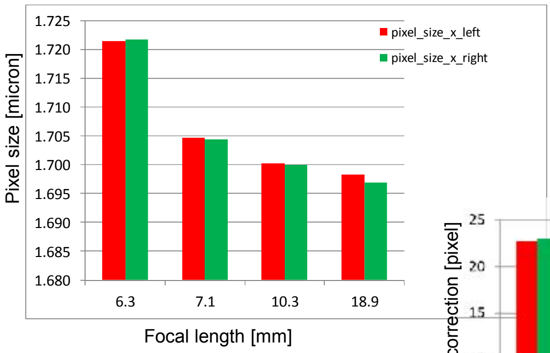
- Bundle adjustment results for different focal settings – principal point and baseline behavior

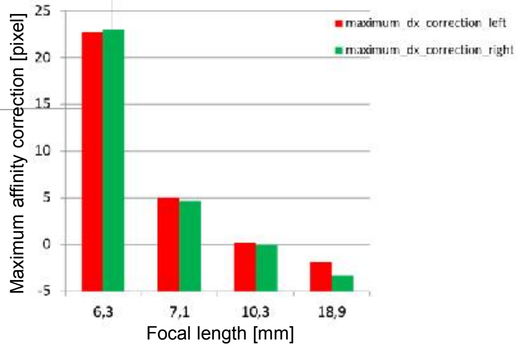




FUJIFILM characterization

❑ Bundle adjustment results for <u>different focal settings</u> – Affinity factor





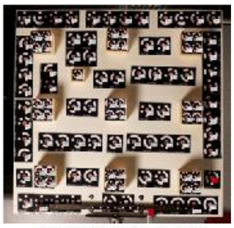
- Nominal pixel size: 1.7 micron
- Very high residuals without affinity factor

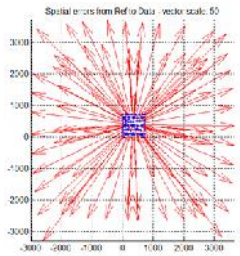
FUJIFILM characterization

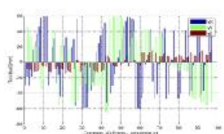
ACCURACY IN 3D MEASUREMENTS:

Comparison with photogrammetry on the 3D calibration frame

- A stereo pair taken from a frontal position (1.5m) is used to forward triangulate the 3D points of the testfield.







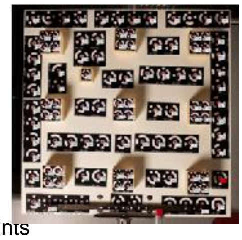
- COMPARISONS:
- IO and EO from a self calibration using only the single pair scale from MPO tag
- IO from EXIF, relative orientation, scaling from MPO tag
- IO and mean EO from the bundle adjustment with self cal
- IO from EXIF, relative orientation, scaling on the object
- IO from bundle adj, EO from relative orientation, scale from mean baseline
- IO and mean EO from the bundle adjustment with self cal (3 months before)

FUJIFILM characterization

ACCURACY IN 3D MEASUREMENTS:

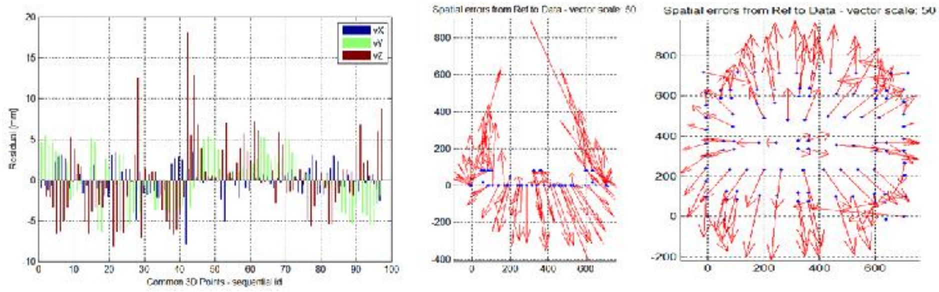
Comparison with photogrammetry on the 3D calibration frame

IO and EO from a self calibration using only the single pair scale from MPO tag



Automatic orientation and calibration in agisoft sterescan, IO and EO imported in photomodeler to mark and triangulated the coded target points

RMSE X,Y,Z = 1.84 3.33 4.69 mm



FUJIFILM characterization

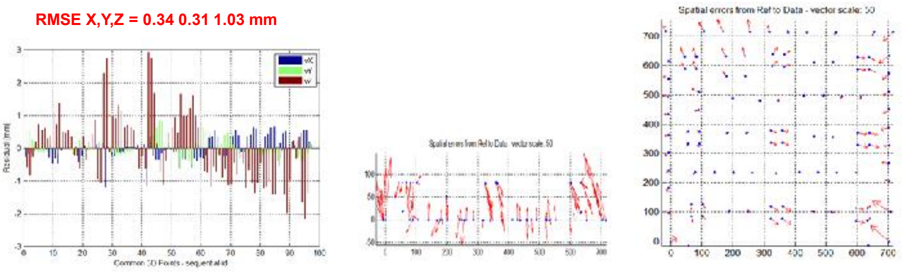
ACCURACY IN 3D MEASUREMENTS:

Comparison with photogrammetry on the 3D calibration frame

IO from bundle adj, EO from relative orientation, scale from mean baseline

Automatic calibration in agisoft lens, IO imported in agisoft photoscan for automatic orientation, IO and EO imported in photomodeler to mark and triangulated the coded target points (scale on the object)

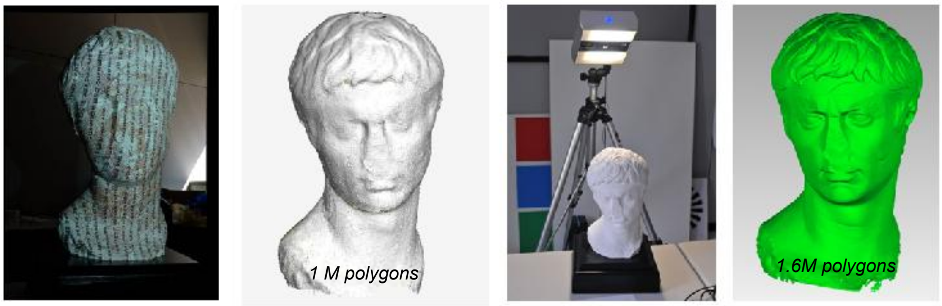
RMSE X,Y,Z = 0.34 0.31 1.03 mm



FUJIFILM characterization

PERFORMANCES IN 3D MODELLING:

3D comparison on a small statue (white matte plastic, ca 35 cm height and 20 cm wide)

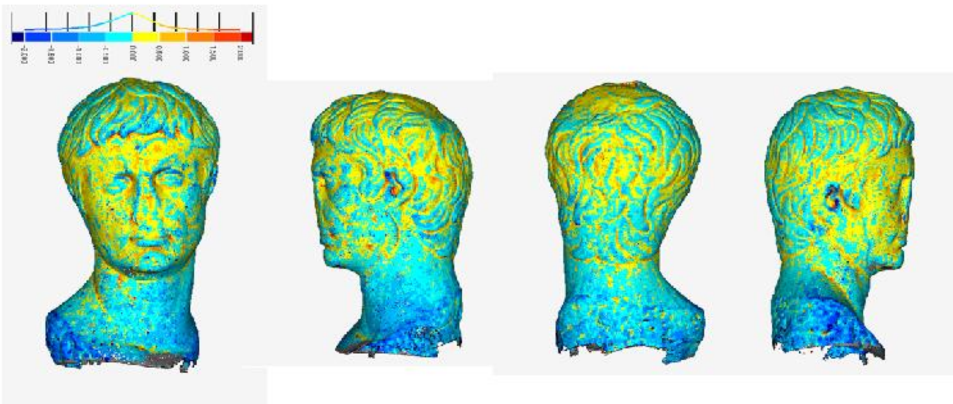


Fujifilm at ca 50cm (pattern projection) *NEXTENGINE* (accuracy $\sigma_z \approx 0.3$ mm)

- 31 stereo-pairs oriented and matched (2.5 Mil. points)
- Fuji point clouds registration / alignment
- NextEngine 3D laser scanner used as **reference**
- Comparison after ICP alignment of the 2 datasets (std = 0.35 mm)

FUJIFILM characterization

PERFORMANCES IN 3D MODELLING:



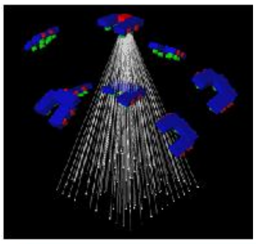
- Geometric differences follows a Gaussian (std = 0.35 mm)

FUJIFILM considerations

- Nominal focal length differs a lot with the increase of the focal length from the calibrated value
- Baseline quite constant
- Significant affinity factor to allow the bundle convergence (up to 20 px at minimal zoom)
- Similar behavior for the radial distortion curve (up to 80 microns at the borders)
- For 3D reconstructions with a single shot is OK but we know its limitations now
- As the camera is supposed to be used with a single shot (ca 1.5 m):

  RMSE X,Y,Z = 0.3 mm, 0.3 mm, 1 mm have been achieved

Slightly better numbers than the Kinect



Conclusions

- **Different sensors and packages** available to allow 3D recording and reconstruction in a low-cost mode
- Low-cost concept is attracting many **non-experts** to the 3D market with
  - positive aspects: enlarge the use of 3D & continuous development/improvement
  - negative aspects: misuse of 3D, idea that everyone can get 3D models, neglect of theoretical fundaments, etc.
- More investigations are needed to deliver definitive conclusions
- Best practices required
- More sensors and software are appearing on the market but …
  don't use them as black boxes ☺