

# **BUILDING INFORMATION MODELING IN NEAR-REAL-TIME CON RANGE CAMERA: ANALISI DELLE POTENZIALITÀ DELLO STRUCTURE SENSOR**

## **BUILDING INFORMATION MODELING IN NEAR-REAL-TIME BY RANGE CAMERAS: ANALYSIS OF THE STRUCTURE SENSOR POTENTIALITIES**

Roberta Ravanelli <sup>a</sup>

<sup>a</sup> Università di Roma “La Sapienza”, Area di Geodesia e Geomatica (DICEA), via Eudossiana 18, 00184 Roma  
roberta.ravanelli@uniroma1.it

**PAROLE CHIAVE:** Building Information Modeling, Range camera, Structure Sensor, tempo reale

**KEY WORDS:** Building Information Modeling, Range camera, Structure Sensor, real-time

### **RIASSUNTO**

Il presente lavoro si propone di investigare le potenzialità di una range camera a basso costo, lo Structure Sensor di Occipital, nella realizzazione di un Building Information Modeling (BIM) in tempo (quasi) reale, in particolare per quanto concerne la valutazione dell'accuratezza nel rilievo di ambienti interni.

Le range camera sono sensori ottici attivi, a basso costo e di facile utilizzo, in grado di misurare distanze di un elevato numero di punti con un alto frame rate (30-60 Hz) e possono essere quindi utilizzate come scanner tridimensionali per generare nuvole dense di punti praticamente in tempo reale. Lo Structure Sensor è la prima range camera appositamente progettata per interagire con dispositivi mobili.

L'applicazione Room Capture, fornita gratuitamente dalla stessa Occipital all'interno dello Structure SDK, è stata pertanto utilizzata per ricostruire in tempo reale i modelli tridimensionali di tre ambienti interni, caratterizzati da differenti forme e dimensioni. Per ogni stanza sono stati acquisiti più modelli tridimensionali e per ogni modello è stata realizzata una sezione orizzontale in prossimità del soffitto, in modo da estrarre le planimetrie al netto di oggetti presenti a livello del pavimento e misurare alcune grandezze caratteristiche (lati, perimetro e superficie) degli ambienti. Tali grandezze sono state confrontate con quelle ottenute attraverso un tradizionale rilievo architettonico, realizzato tramite una cordella metrica.

I risultati ottenuti dimostrano che è effettivamente possibile utilizzare lo Structure Sensor per la realizzazione di modelli tridimensionali e di planimetrie di ambienti interni in tempo (quasi) reale con accuratezza media pari a 5 cm, risultando pertanto idonee per una rappresentazione alla scala 1:200.

### **ABSTRACT**

This work investigates the potentialities of a low-cost structured light range camera, the Occipital Structure Sensor, for the construction of a Building Information Modeling (BIM) in (near) real-time, with the specific aim of evaluating its accuracy in surveying indoor environments.

Range cameras are active imaging sensors, low-cost and easy to use, able to natively measure the distances of several points at high frame rate (30 - 60 Hz) and therefore they can be used as 3D scanners to easily collect dense point clouds practically in real-time. The Structure Sensor is precisely the first range camera specifically designed for mobile devices.

The Room Capture application, provided within the Structure SDK by Occipital itself, has been thus used to acquire in real-time the 3D models of three indoor environments, characterized by different shapes and dimensions. For every room, several 3D models have been captured and for each model a horizontal section has been cut in proximity of the ceiling, in order to extract the corresponding 2D map net of floor-level objects and to uniquely measure the principal quantities (sides, perimeter and area) of the scanned environments. These quantities have been compared with those obtained through a traditional architectural survey, performed by means of a measuring tape.

The achieved results show that it is effectively possible to use the Structure Sensor to reconstruct 3D models of indoor environments and to generate the corresponding planimetric maps in (near) real-time. In fact, the reconstructed planimetric layouts present a mean accuracy of 5 cm, suitable for 2D maps at scale 1:200.

### **1. INTRODUZIONE**

Le range camera sono sensori ottici attivi, a basso costo e di facile utilizzo, in grado di misurare distanze di un elevato numero di punti con un alto frame rate (30-60 Hz). Ad ogni acquisizione, esse producono una mappa di profondità della scena ripresa, un'immagine in cui ad ogni pixel è associata la propria distanza rispetto a un riferimento specifico, solitamente associato al sensore stesso (generalmente il piano immagine). A partire dalla mappa di profondità, le range camera generano anche una nuvola densa di punti dell'ambiente scansionato,

ossia un insieme di coordinate tridimensionali di un numero elevato di punti in un sistema di riferimento legato allo strumento stesso.

Pertanto, in accordo con la definizione fornita in (Boehler e Marbs, 2002), le range camera possono essere considerate, a tutti gli effetti, veri e propri scanner tridimensionali. Inoltre, algoritmi di *Simultaneous Localization And Mapping (SLAM)*, come per esempio *KinectFusion* (Izadi et al., 2011), (Newcombe et al., 2011)) consentono di sfruttare al meglio i dati tridimensionali e l'elevato frame rate che tali sensori offrono, stimando continuamente la posizione dello strumento

durante il processo di scansione e fondendo le mappe di profondità catturate da differenti punti di vista non appena queste vengono acquisite. In questo modo, si ottiene facilmente un modello tridimensionale completo, praticamente in tempo reale. Per di più, la tecnologia sulla quale le range camera si basano è in continua evoluzione e presto tali strumenti saranno integrati all'interno di dispositivi comuni come smartphone e tablet, consentendo quindi il loro utilizzo anche assieme ad altre tipologie di sensori (si vedano, ad esempio, (Benedetti et al. 2016) e (Ravanelli et al., 2016)).

Le caratteristiche appena descritte rendono quindi le range camera un utile strumento nell'ambito del rilievo architettonico di ambienti interni. Le nuvole di punti da esse prodotte offrono infatti un'elevata potenzialità di descrizione geometrica dell'ambiente ricostruito, proprietà che assume un'importanza sempre maggiore anche in virtù della recente diffusione del Building Information Modelling (BIM). Al giorno d'oggi occorrono invero prodotti del rilievo facilmente interpretabili e soprattutto da cui poter agevolmente ricavare rappresentazioni di qualsiasi natura e genere. In questo senso, l'impiego delle range camera per la realizzazione di modelli tridimensionali e di planimetrie da essi derivati, grazie alla facilità d'uso e rapidità di ricostruzione tridimensionale di questi strumenti, può offrire significativi vantaggi rispetto ai metodi tradizionali, in particolar modo quando sono coinvolti utenti non esperti nell'utilizzo di tecniche geomatiche o quando si rende necessario un rapido intervento.

Tuttavia, prima di impiegare sistematicamente le range camera in questo ambito, è necessario valutare la qualità metrica della ricostruzione della geometria tridimensionale da esse fornita. Il presente lavoro si inserisce precisamente in questo contesto, prefiggendosi di investigare le potenzialità di una promettente range camera a basso costo, lo Structure Sensor di Occipital (Figura 1), per la realizzazione di planimetrie in tempo quasi reale di ambienti interni, incentrando l'analisi sulla valutazione dell'accuratezza e della precisione dello stesso.

Lanciato su Kickstarter nel settembre 2013, dove ha raccolto 1.3 milioni di dollari in soli 45 giorni, lo Structure Sensor è la prima range camera appositamente progettata per interagire con dispositivi mobili. Caratterizzato da dimensioni compatte e una batteria interna, lo Structure Sensor può essere connesso a un tablet o a uno smartphone iOS, rendendo la tecnologia delle range camera facilmente accessibile a un pubblico sempre più vasto.

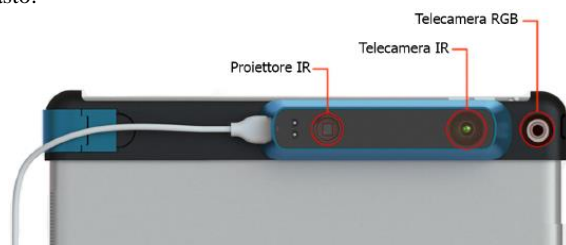


Figura 1. Componenti dello Structure Sensor: il proiettore infrarosso e la camera infrarossa; la camera a colori non è parte del sensore, ma è quella fornita dal dispositivo (smartphone o tablet) al quale è connesso.

In particolare, lo Structure Sensor è una range camera a luce strutturata e stima le distanze basandosi sul principio della triangolazione attiva. Il sensore è pertanto costituito da un proiettore laser infrarosso e una camera infrarossa sensibile alla stessa banda. Il primo proietta un pattern infrarosso sulla superficie degli oggetti che si intendono rilevare, laddove la seconda registra come l'ambiente deforma il pattern originale,

ottenendo di conseguenza la geometria tridimensionale (forma e dimensioni in unità metriche) della scena ripresa. Differentemente da altre range camera attualmente disponibili sul mercato, lo Structure Sensor non possiede una camera a colori, per cui la texture della scena viene acquisita dalla camera a colori del tablet/smartphone al quale è connesso.

## 2. ACQUISIZIONE DEI MODELLI

Come accennato in precedenza, lo scopo del presente lavoro è la valutazione dell'accuratezza dello Structure Sensor nella realizzazione di planimetrie di ambienti interni. L'applicazione *Room Capture*, sviluppata dalla stessa Occipital, è stata quindi utilizzata per ricostruire i modelli tridimensionali di tre ambienti, caratterizzati da differenti forme e dimensioni (in ogni caso con lati massimi contenuti entro 8-10 metri), di seguito denominate *Aula Grande*, *Aula Tesisti* e *Aula Piccola*, situate presso l'Area di Geodesia e Geomatica dell'Università di Roma "La Sapienza". L'applicazione *Room Capture*, gratuita e di facile utilizzo, consente infatti di realizzare modelli tridimensionali di ambienti semplicemente inquadrandone le varie zone con lo Structure Sensor. L'applicazione è parte integrante dello Structure Software Development Kit (SDK) fornito da Occipital: il codice sorgente è disponibile in forma di esempio e può essere modificato dagli sviluppatori secondo le particolari esigenze dello specifico caso d'uso.

In particolare, la generazione del modello 3D si basa sul tracking, il processo per cui il sensore stima il proprio movimento in relazione all'oggetto/ambiente da scannerizzare. Sono infatti necessarie più acquisizioni da differenti punti di vista per modellare la geometria dell'intera superficie di un oggetto e/o ambiente. L'operatore deve pertanto muoversi lentamente lungo la stanza da rilevare, scansionandone tutte le pareti. Nella maggior parte dei casi il tempo necessario per la scansione è di alcuni minuti, tempo che può aumentare a seconda della forma e complessità dell'ambiente. Inoltre è possibile controllare immediatamente la qualità della ricostruzione tramite la funzionalità di realtà aumentata grazie alla quale il modello appare in tempo reale sullo schermo del dispositivo mobile. In questo modo l'operatore può migliorare istantaneamente i risultati della scansione, acquisendo nuovamente le aree problematiche e riempiendo le eventuali lacune. Inoltre, l'operatore può impostare le dimensioni del volume di scansione, che dovrebbe essere allargato o diminuito in funzione delle dimensioni della stanza da rilevare.

Prima di procedere con la scansione, lo Structure Sensor è stato calibrato utilizzando l'applicazione *Calibrator* di Occipital al fine di consentire una corretta sovrapposizione dei dati geometrici catturati dal sensore con i dati di colore acquisiti dalla camera RGB del tablet. È importante sottolineare che la calibrazione è essenziale anche per la buona riuscita della scansione, in quanto può influenzare il corretto funzionamento del tracking, il quale infatti può sfruttare sia la geometria che il colore della scena.

Per il presente lavoro, sono stati acquisiti più modelli tridimensionali di ognuno dei tre ambienti, precisamente 7 per l'*Aula Piccola* e 4 per l'*Aula Tesisti* e l'*Aula Grande*. Prima di tutto si è reso necessario identificare una procedura generale di scansione al fine di ottenere modelli di buona qualità degli ambienti. Infatti, l'allineamento incrementale di frame successivi può generare un progressivo accumulo di errori nel processo di tracking, causando una non corretta stima della traiettoria percorsa dal sensore durante il processo di scansione. In tal caso, quando alla fine della scansione vengono catturate nuovamente aree già precedentemente acquisite, queste possono

essere collocate in una errata posizione globale, e la traiettoria stimata non si chiude. Ciò può comportare una rilevante deformazione del modello 3D finale, la cui geometria in questo caso non corrisponde alla reale forma dell'ambiente rilevato (si veda ad esempio la Figura 2).



Figura 2. Modello della stanza *Aula Piccola*, con tracking errato (sinistra) e corretto (destra).

È pertanto indispensabile limitare i movimenti non necessari del sensore che possono disturbare il tracking. I modelli 3D sono stati quindi acquisiti con l'operatore seduto su una sedia girevole collocata in posizione centrale rispetto all'ambiente da rilevare, mantenendo il sensore il più possibile vicino all'asse di rotazione della sedia stessa. In questo modo i movimenti di traslazione sono stati minimizzati e la traiettoria percorsa dal sensore è risultata per lo più circolare, come la stessa Occipital consiglia. Infine, le dimensioni del volume di scansione sono state impostate in maniera tale da consentire la massima copertura della parte superiore delle pareti con una minima inclinazione del sensore, al fine di ricostruire esclusivamente i soffitti delle stanze, ossia le parti sgombre da mobili e ostacoli vari generalmente presenti a livello del pavimento.

### 3. REALIZZAZIONE DELLE PLANIMETRIE E RISULTATI

I tre ambienti sono stati rilevati tramite il metodo di scansione circolare descritto nel precedente paragrafo. I modelli 3D risultanti sono stati quindi elaborati per estrarre delle planimetrie che consentissero di misurare univocamente le grandezze caratteristiche (lati, perimetro e superficie) delle stanze. Tutto ciò allo scopo di comparare tali grandezze con quelle di riferimento, ottenute attraverso un tradizionale rilievo architettonico realizzato, in questo caso, utilizzando una comune cordella metrica. È importante sottolineare che il rilievo tradizionale è più complesso e meno rapido da eseguire rispetto alla scansione con lo Structure Sensor.

Per ogni stanza sono stati acquisiti più modelli tridimensionali. In particolare, il software *Slic3r* (Hodgson et al., 2017), gratuito e open source, è stato utilizzato per tagliare ogni modello 3D lungo una sezione orizzontale parallela al piano del soffitto e molto vicina ad esso (Fig. 3). Le Figure 6, 7 e 8 mostrano un esempio delle planimetrie ottenute per ogni ambiente.

Per valutare l'accuratezza di ogni modello ricostruito, sono stati dapprima calcolati gli scarti tra le lunghezze dei lati misurate sulle planimetrie ricavate dai modelli generati dallo Structure Sensor e quelle ottenute tramite il rilievo tradizionale. Successivamente, gli scarti così ottenuti sono stati mediati su tutti i lati per ottenere l'errore sistematico del sensore nella ricostruzione del singolo modello; allo stesso modo è stata quindi calcolata la deviazione standard degli scarti. In questa maniera, è stato possibile calcolare l'RMSE per ogni modello, per caratterizzarne l'accuratezza.

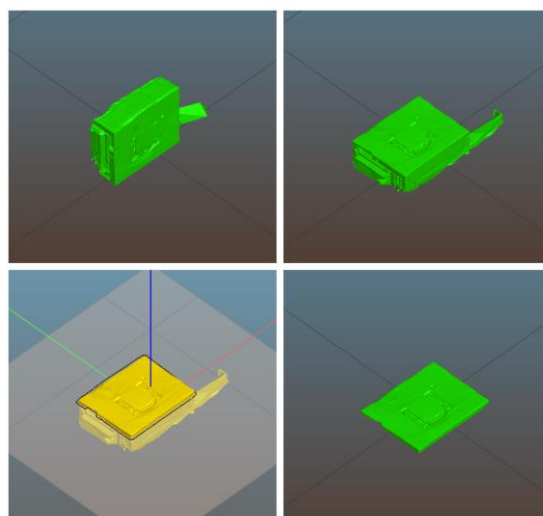


Figura 3. Sezionamento dei modelli tramite il software *Slic3r*.

La stessa procedura è stata eseguita per tutte le aule prese in esame e i risultati sono riportati rispettivamente in Figura 4 e Tabella 1.

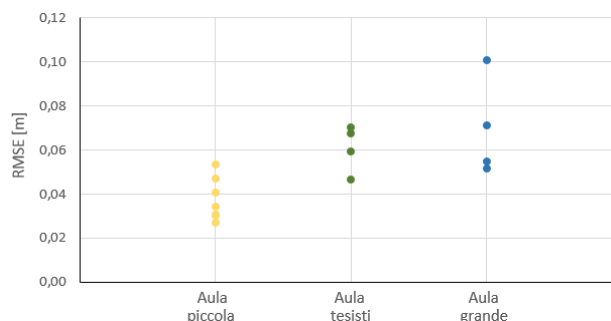


Figura 4. RMSE dei lati per le planimetrie ottenute tramite lo Structure Sensor.

	RMSE [m]		
	<i>Aula Piccola</i>	<i>Aula Tesisti</i>	<i>Aula Grande</i>
Modello 1	0.03	0.06	0.05
Modello 2	0.05	0.05	0.07
Modello 3	0.03	0.07	0.05
Modello 4	0.05	0.07	0.10
Modello 5	0.04	-	-
Modello 6	0.03	-	-
Modello 7	0.03	-	-
Media	0.04	0.06	0.06

Tabella 1. RMSE dei lati per le planimetrie ottenute tramite lo Structure Sensor.

In particolare, si può notare come l'RMSE assuma complessivamente valori compresi tra 3 e 10 cm, piuttosto contenuti se confrontati con le dimensioni dei lati degli ambienti. Il livello di accuratezza ottenibile è pertanto quantificabile mediamente intorno a 5 cm. I risultati evidenziano inoltre un lieve trend dell'RMSE in funzione delle dimensioni degli ambienti (Figura 4), non inatteso a causa del già sottolineato accumulo degli errori nel corso della scansione. Inoltre, per caratterizzare ulteriormente l'accuratezza del sensore, sono state confrontate anche le aree ed i perimetri delle aule. In Tabella 2 e Tabella 3 sono riportati i risultati ottenuti, che confermano ulteriormente le ottime prestazioni dello Structure Sensor.

	Aree [m <sup>2</sup> ]		
	Aula Piccola	Aula Tesisti	Aula Grande
Riferimento	13.26	23.36	26.96
Modello 1	13.20	23.19	26.84
Modello 2	13.27	23.55	26.83
Modello 3	13.41	23.58	26.49
Modello 4	13.28	23.77	26.40
Modello 5	13.32	-	-
Modello 6	13.16	-	-
Modello 7	13.24	-	-
Media scarti	-0.01	-0.16	0.32
Dev. Std. Scarti	0.08	0.21	0.20
RMSE	0.08	0.27	0.38
RMSE (%)	0.57%	1.14%	1.40%

Tabella 2. Statistiche sulle aree.

	Perimetri [m]		
	Aula Piccola	Aula Tesisti	Aula Grande
Riferimento	14.65	19.37	21.01
Modello 1	14.61	19.24	20.99
Modello 2	14.74	19.45	21.13
Modello 3	14.68	19.62	21.04
Modello 4	14.61	19.57	20.95
Modello 5	14.59	-	-
Modello 6	14.56	-	-
Modello 7	14.63	-	-
Media scarti	0.01	-0.11	-0.02
Dev. Std. Scarti	0.06	0.15	0.07
RMSE	0.06	0.18	0.07
RMSE (%)	0.39%	0.93%	0.35%

Tabella 3. Statistiche sui perimetri.

Per rimuovere almeno in parte gli effetti sistematici noti per questo tipo di sensori e migliorarne di conseguenza l'accuratezza, è comunque necessario procedere ad una calibrazione. Da prove ripetute effettuate in laboratorio si è quindi determinata sperimentalmente una funzione di calibrazione per sensori del tipo a tempo di volo e a luce strutturata (Ravanelli, 2017). In particolare, le tre più famose e/o promettenti range camera a basso costo, vale a dire la Kinect v1 e la Kinect v2 di Microsoft e lo Structure Sensor, sono state caratterizzate da un punto di vista geomatico per valutare la qualità metrica dei dati di profondità da esse acquisiti. Le indagini svolte hanno dimostrato che tali sensori presentano una precisione e un'accuratezza che sono funzione sia del principio operativo adottato dal singolo dispositivo che della profondità, ossia della distanza (non inclinata) sensore - oggetto. Sono stati pertanto identificati due diversi modelli per la precisione e l'accuratezza rispetto alla profondità: lineari per il sensore a tempo di volo (Kinect v2) e parabolici per quelli a luce strutturata (Kinect v1 e Structure Sensor). Infine, l'efficacia del modello di calibrazione è stata dimostrata essere globalmente conforme al modello di precisione identificato per ognuno dei tre sensori. In Figura 5 si riporta l'andamento complessivo dell'errore di misura dello Structure Sensor all'aumentare della distanza di acquisizione, assieme al modello di calibrazione ipotizzato.

L'applicazione di tale modello di calibrazione comporterebbe un miglioramento in termini di accuratezza dell'ordine di qualche centimetro nella ricostruzione delle planimetrie, tanto maggiore quanto più elevate sono le dimensioni dell'ambiente da rilevare (aumento della distanza media sensore - parete). Ad esempio, effettuando una scansione ad una distanza media dalle pareti di circa 2.5 m, come è avvenuto per il rilievo dell'Aula

Tesisti, si potrebbe ottenere una diminuzione dell'errore di poco superiore ai 2 cm.

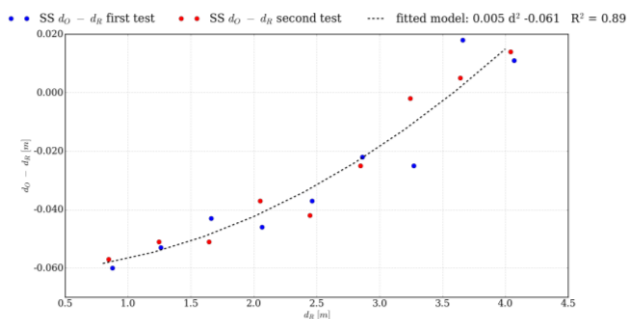


Figura 5. Andamento dell'errore di misura dello Structure Sensor all'aumentare della distanza di acquisizione (due prove di calibrazione indipendenti).

#### 4. CONCLUSIONI

I risultati ottenuti dimostrano che è effettivamente possibile utilizzare lo Structure Sensor per la realizzazione di planimetrie di ambienti interni in tempo (quasi) reale. Le planimetrie ricostruite presentano infatti un'accuratezza media pari a 5 cm e comunque quantificabile sempre al meglio di 10 cm, risultando pertanto idonee per una rappresentazione alla scala 1:200. Quest'ultima è una caratteristica molto importante, dal momento che il sensore è a basso costo e di facile utilizzo e può quindi essere facilmente impiegato da utenti non esperti per la raccolta di dati relativi ad ambienti indoor nel più ampio contesto della generazione di informazioni geografiche volontarie (VGI).

Infine, come sviluppi futuri, sarebbe interessante inserire il modello di calibrazione del sensore all'interno del software di scansione, oltre ad ampliare il numero dei test di scansione tridimensionale.

#### RINGRAZIAMENTI

L'autrice è sinceramente grata al Prof. Mattia Crespi e all'Ing. Andrea Nascetti per il sostegno, gli incoraggiamenti ed i consigli ricevuti durante la stesura di questo lavoro.

#### BIBLIOGRAFIA

- Benedetti, E., Ravanelli, R., Moroni, M., Nascetti, A. e Crespi, M., 2016. Exploiting Performance of Different Low-Cost sensors for Small Amplitude Oscillatory Motion Monitoring: Preliminary Comparisons in View of Possible Integration. *Journal of Sensors*.
- Boehler, W. e Marbs, A., 2002. 3D scanning instruments. *Proceedings of the CIPA WG 6*, pp. 9-18.
- Hodgson, G., Ranellucci, A. e Moe, J., 2017. Slic3r manual, <http://manual.slic3r.org/>.
- Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S., Freeman, D., Davison, A. e Fitzgibbon, A., 2011. KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera. *Proceedings of the 24<sup>th</sup> annual ACM symposium on User interface software and technology*.
- Newcombe, R., Izadi, S., Hilliges, O., Molyneaux, D., Kim, D., Davison, A., Kohli, P., Shotton, J., Hodges, S. e Fitzgibbon,

A., 2011. KinectFusion: Real-Time Dense Surface Mapping and Tracking. *Proceedings of the 10<sup>th</sup> IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*.

Ravanelli, R., Nascetti, A. e Crespi, M., 2016. Kinect V2 And RGB Stereo Cameras Integration For Depth Map Enhancement. *The International Archives of the Photogrammetry, Remote*

*Sensing and Spatial Information Sciences XLI-B5*, pp. 699–702.

Ravanelli, R., 2017. 3D modeling by low-cost range cameras: methods and potentialities. *Tesi di dottorato*.

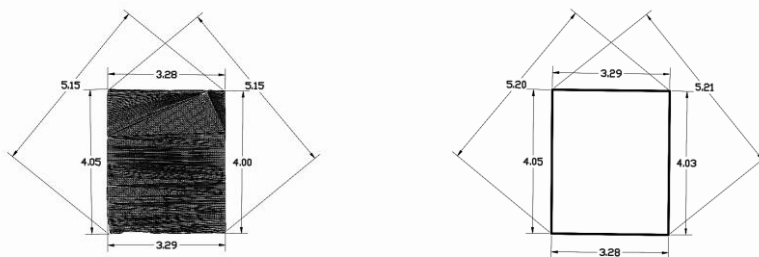


Figura 6. Planimetrie della stanza *Aula piccola*: a sinistra quella ottenuta con lo Structure Sensor, a destra quella ottenuta con il rilievo tradizionale.

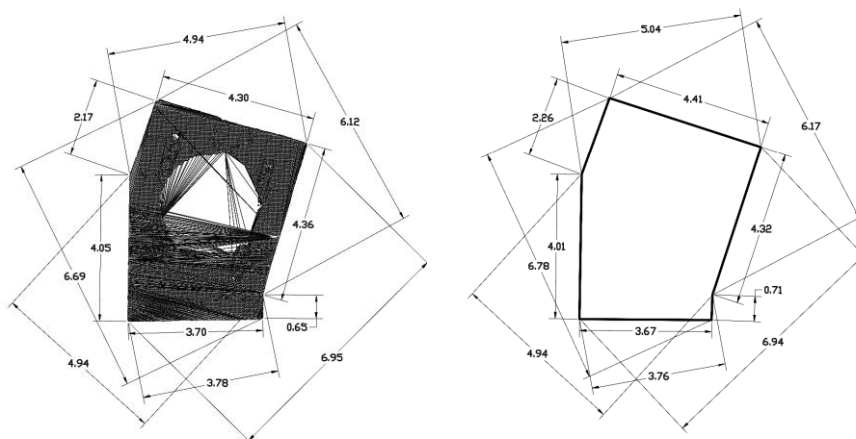


Figura 7. Planimetrie della stanza *Aula Tesisti*: a sinistra quella ottenuta con lo Structure Sensor, a destra quella ottenuta con il rilievo tradizionale.

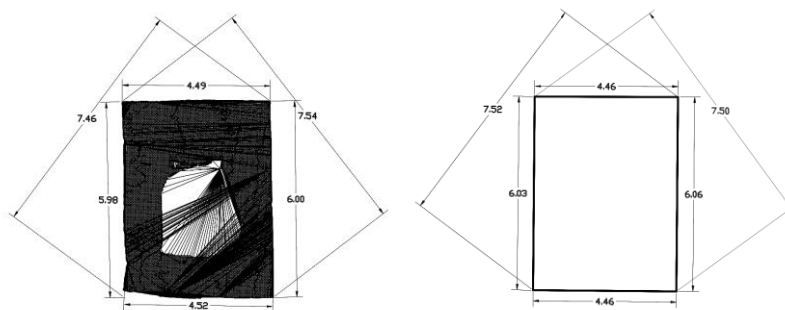


Figura 8. Planimetrie della stanza *Aula grande*: a sinistra quella ottenuta con lo Structure Sensor, a destra quella ottenuta con il rilievo tradizionale.