



Visualità digitale applicata a metodologie di rilievo integrato. Sinergie collaborative tra sperimentazione e tecnologia

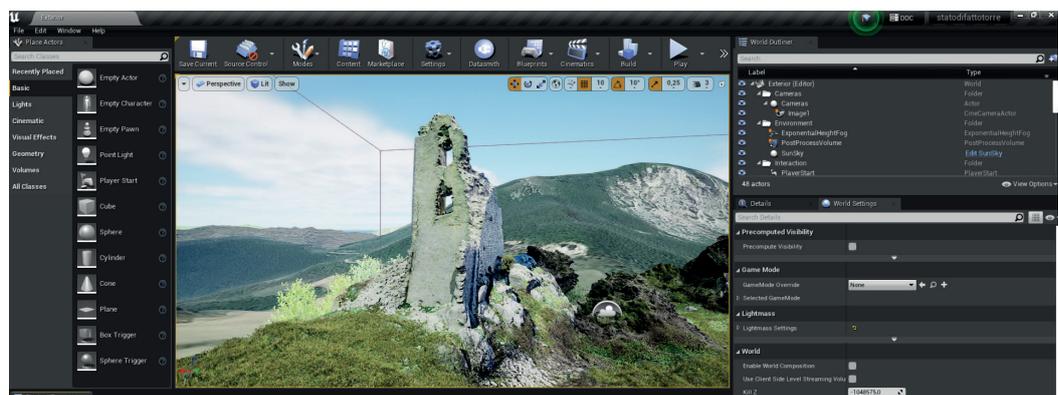
Maurizio Perticarini
Alessandro Basso

Abstract

La progressiva adesione alla *Computer Vision* ha valorizzato l'uso dell'immagine come mezzo principale di comunicazione, integrando automatici processi interpretativi che travalicano i confini strutturali legati al luogo e al tempo, per poter accedere direttamente a processi di condivisione di linguaggio che consentono a tutti di acquisire nuovi strumenti percettivi che offrono, di fatto, una migliore adattabilità alle modalità interpretative di partecipazione culturale, relative a concetti dalla eccezionale versatilità. Tecnologia e *Computer Vision*, quindi, imbastiscono una relazione strutturale tra Rappresentazione e sperimentazione, tra arte e scienza, alimentando una mutua collaborazione tra le varie aree disciplinari. Il paper in questione vuole indagare quindi tali questioni, ponendo a confronto nello specifico del rilievo digitale tre casi studio dalle intrinseche caratteristiche, in relazione a *workflow* di fotomodellazione, ottimizzazione degli asset 3d ed operazioni di migrazione virtuale in spazi interattivi. Il contributo mira ad analizzare diversi sistemi applicativi per la fotogrammetria 3d a basso costo attraverso una pipeline precisa, chiarendo come tre diverse metodologie di acquisizione spaziale e restituzione digitale possano ricercare i medesimi risultati percettivamente fotorealistici, impiegabili nel cyberspazio virtuale, documentando come uno spazio riconfigurato possa combinare molteplici fasi di indagine fino ad avvicinarsi a una visione grafica di sintesi oggettiva. La riconfigurazione digitale si rivela così, con tutti i suoi limiti e paradossi, un metalinguaggio potente e versatile in continua evoluzione.

Parole chiave

Sensori Lidar, rilievo low-cost, image based, Computer Vision, real time rendering



Screenshot del rendering in tempo reale tratto da Unreal Engine.

Introduzione

La tecnologia, come strumento per facilitare la progettazione di idee, alla quale l'umanità nel corso dei secoli si è rivolta scommettendo risorse, lavoro e ricerca, diviene, alla luce di una comune riflessione sul valore del dialogo culturale, un vero e proprio linguaggio attraverso il quale è possibile non solo trasmettere informazioni, delineando le prevalenti sinergie teorico-applicative che identificano l'ambito della Rappresentazione architettonica e non, ma soprattutto strutturare interscambi partecipativi di natura funzionale che mettono in relazione ambiti all'apparenza estremamente diversificati. L'adesione alla Computer Vision, sempre più preponderante negli ultimi tempi, ha valorizzato l'uso dell'immagine, che è sempre stata il mezzo principale di comunicazione il cui percorso evolutivo è certamente connaturato con quello dell'uomo, integrando automatici processi di interpretazione che travalicano i confini strutturali legati al luogo e al tempo, per poter accedere direttamente a processi di condivisione che consentono a tutti di utilizzare nuovi strumenti percettivi capaci di offrire, di fatto, una migliore adattabilità alle modalità interpretative di condivisione culturale, relative a concetti dalla eccezionale versatilità. Tecnologia e Computer Vision, quindi, imbastiscono una solida relazione strutturale tra la Rappresentazione e la sperimentazione, tra l'arte e la scienza, alimentando, durante questa Terza Rivoluzione Digitale, malgrado un periodo storico che continua ad essere difficile sotto molti punti di vista, la collaborazione tra le varie aree disciplinari che utilizzano il disegno come solida risorsa al fine di ottenere risultati concreti. Restringendo il campo sul tema del rilievo digitale, il contributo oggetto della ricerca vuole indagare, mediante confronti e approfondimenti sperimentali, quanto sia efficace la modalità di progettazione e di rappresentazione che coinvolge i nuovi strumenti grafici per restituire, ricercando un univoco risultato interpretativo, una percezione dei luoghi digitalizzati sovrapponibile a quella reale e comprensibile da tutti in maniera fluida e naturale. Sebbene l'obiettivo sia quello di generare elaborati che vogliono ricalcare lo stato di fatto spaziale, non tutte le modalità di rilievo digitale offrono una immediata possibilità di lettura, alcune necessitano di lavoro di ottimizzazione per essere comprese ed interpretate da chi non possiede un solido background culturale, altre modalità poi, per loro natura grafica, sono decodificabili seguendo esclusivamente specifici schemi interpretativo-culturali, molto lontani dalla percezione cognitivo-visuale che segue nello specifico esclusivamente indizi Fisiologici ed indizi Psicologici. Con il passare del tempo si diffonde però la tendenza comune di aderire ad una vera e propria clonazione in scala che possa essere interpretata visivamente in maniera diretta e che quindi possa essere il più possibile coerente, a livello percettivo di profondità e acquisizione cromatica, allo spazio reale in oggetto. Ad avvantaggiare questa tendenza c'è l'evoluzione tecnologica degli hardware ma anche la componente umana dell'immedesimazione, che travalica l'effe-



Fig. 01. a) Fotografia dell'edicola e dell'altare in Sanata Maria ad Cryptas b) Fotografie equirettangolari della sala principale della Fornace c) Fotografia della torre prima del sisma del 2016. Immagine degli autori.

tiva capacità sensoriale e che funziona molto bene in sinergia con le attuali metodologie di rilievo digitale. Tale capacità di interpretare la profondità di spazi digitali, clonati o inesistenti, è il punto di forza che permette di rendere la dimensione virtuale "reale", ed è a questo surrogato di realtà basata sul fotorealismo, al di là della possibilità di acquisire misure e dati, al quale puntano tutti gli strumenti next-generation di clonazione virtuale, di rilievo digitalizzato dello spazio. Si mettono a confronto quindi tre casi studio - condotti durante gli ultimi due anni - che utilizzano tre diverse metodologie di acquisizione spaziale low-cost [1] e restituzione digitale con l'obiettivo di ottenere i medesimi risultati percettivamente fotorealistici, impiegabili in migrazioni su piattaforma interattiva, identificandone risultati e criticità sulla base di workflow diversificati.

Tre casi a confronto

La ricerca pone l'attenzione su tre casi studio che hanno caratteristiche diverse, sia formali che spaziali: un oggetto architettonico inserito in uno spazio chiuso e poco illuminato (edicola votiva nella chiesa di Santa Maria ad Cryptas) (Fig. 01a), un ambiente interno di forma complessa (Fornace del Canova) (Fig. 01b) e un edificio caratterizzato da un'altezza elevata (Torre di Forca di Penne) (Fig. 01c). Il rilievo di questi tre manufatti storici si basa su tre diverse tecnologie per il rilievo, riuscendo a conferire un discreto fotorealismo agli oggetti acquisiti malgrado la velocità di acquisizione ed elaborazione dei dati. Tre approcci differenti che hanno come requisito comune l'impiego di tecnologie low-cost alla portata di tutti, semplici da usare e capaci di ottenere oggetti validi per essere visualizzati in uno spazio interattivo. Il primo caso descritto è quello del rilievo eseguito per mezzo di sistema *portable Lidar* dell'edicola posta a sinistra dell'altare centrale della chiesa di Santa Maria ad Cryptas di Fossa; il secondo caso si riferisce al rilievo eseguito mediante fotogrammetria basata su immagini sferiche dell'interno della Fornace del Canova a Roma, un tempo impiegata dall'artista per cuocere i suoi modelli in creta, e oggi sede di un'associazione culturale, Canova 22; il terzo caso descrive il rilievo eseguito per mezzo di fotogrammetria aerea (Drone Dji Mini 2) della Torre di Forca di Penne, torre medievale che sorge in posizione strategica nel valico omonimo in provincia dell'Aquila. I tre casi sono stati messi a confronto e studiati sia dal punto di vista dell'attendibilità con il reale, che dal punto di vista della velocità di esecuzione del rilievo. Ciascuna delle tecnologie presenta dei punti di forza e dei punti deboli, a seconda di fattori intrinseci di rilevamento e tipo di oggetto da acquisire.

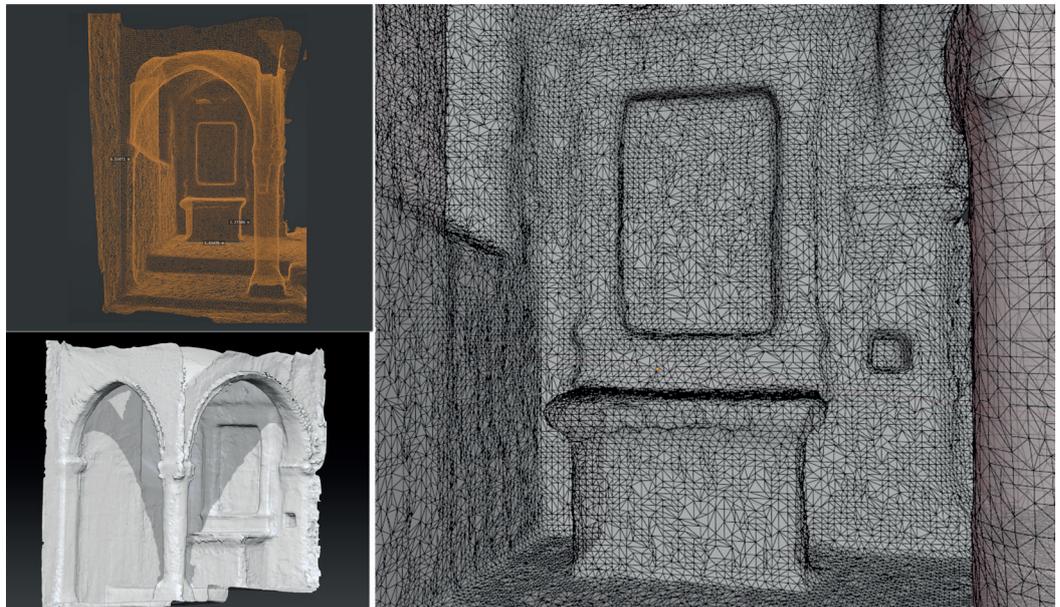


Fig. 02 a) Dense Cloud e mesh realizzata con Polycam b) Wireframe della mesh non ottimizzata. Immagine degli autori.

La chiesa di Santa Maria ad Cryptas e l'uso del Lidar per il rilievo di un oggetto architettonico: La chiesa di Santa Maria ad Cryptas prende il nome da alcune grotte vicine. Si trova a Fossa (AQ) all'interno della Valle Subequana, ai piedi del Monte d'Ocre. Fu edificata nella seconda metà del 1200 dai monaci cistercensi. Il piccolo altare al di sotto di un'edicola tardo seicentesca, oggetto dello studio, risale all'epoca di costruzione del complesso architettonico e non ha subito i rimaneggiamenti che hanno trasformato gran parte del resto della chiesa nel periodo barocco. Nel rilievo di altare e edicola si è sperimentato l'utilizzo dell'applicazione LIDAR per iPhone [2] che sfrutta il sensore 3d per la realizzazione di nuvole di punti e mesh. Ad oggi, molti produttori di smartphone hanno incluso sensori di profondità ma il Lidar di Apple, sviluppato in collaborazione con Lumentum, risulta essere una evoluzione delle precedenti tecnologie integrate, diverso dai più canonici sistemi laser scanner che hanno un sensore ingombrante che ruota su sé stesso, differenziandosi dai precedenti sistemi per una stabile scansione continua. Ha una portata di 5 metri e una precisione che va da 3 mm a 9 mm. L'edicola e il piccolo altare sono stati scansionati utilizzando l'applicazione *Polycam* che consente una scansione che va da 13 mm a 80 mm di dimensione Voxel e una semplificazione dallo 0% al 99% (Fig. 02a). È stata utilizzata la predisposizione di semplificazione "Oggetto", che ha ridotto le tempistiche del processo di scansione in real time, durato 5 minuti, e l'elaborazione della mesh, composta da 54,9k vertici, in circa 3 minuti (Fig. 02b). È stato quindi possibile risalire perfettamente alla forma originaria in quanto la nuvola di punti copriva un'area piuttosto circoscritta attraverso un numero consistente di punti a disposizione. L'asset digitale ottenuto, senza ulteriori passaggi oltre a quelli canonici della creazione dei dettagli tramite mappe di profondità (*Normal-Bump-Displacement*), risulta essere pronto per l'eventuale inserimento su piattaforme di rendering interattivo RT come Unity o Unreal Engine (Fig. 03). In sintesi, tale sistema accelera il flusso di lavoro consentendo di ottenere un modello mediamente ben strutturato impiegabile subito in dinamiche di visualizzazione digitale.

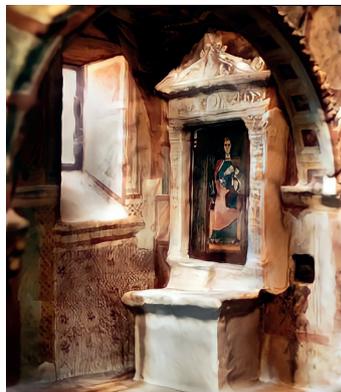
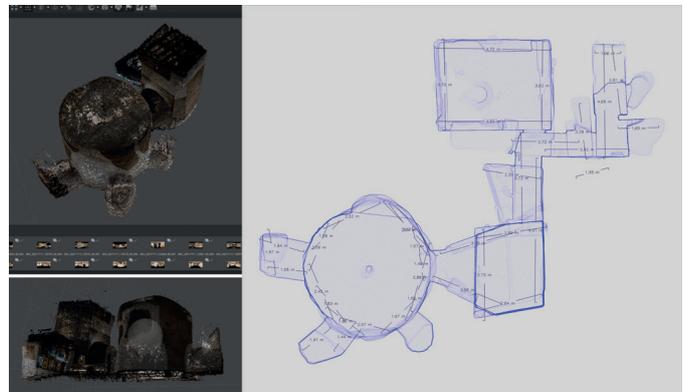


Fig. 03 RTRendering in Twinmotion del modello non ottimizzato, esportato da Polycam con texture e UVMap. Immagine degli autori.

Fig. 04 Viste prospettive della Dense Cloud in Metashape e pianta quotata della Fornace. Immagine degli autori.



La Fornace del Canova ed il rilievo integrato 360°: Il secondo caso di studio propone il rilievo dell'Ex Fornace Canova, nel cuore di Roma già dai primi anni del 1600. La fornace ha una forma circolare con tre nicchie laterali e un'aula centrale sveltante verso l'alto con un piccolo lucernario aperto alla sommità. La struttura irregolare a pianta centrale, seppur ben conservata, presenta alcune caratteristiche critiche per un rilievo fotogrammetrico 3D, come la mancanza di illuminazione sufficiente e locali concatenati piuttosto ristretti a quote variabili che convergono nella grande sala fornace. La ricerca, quindi, propone un rilievo effettuato mediante strumentazioni low cost e tempi molto brevi, con applicazione di una metodologia speditiva esclusivamente supportata da tecniche di fotomodellazione 3d terrestre, senza ricorrere ad alcun utilizzo di laser scanner (TLS), strumenti LIDAR, né mini-droni da interno, sfruttando nello specifico esclusivamente dati fotografici integrati di diverso tipo, panoramiche HDR 360°, foto grandangolari *fisheye* e fotografie formato *frame* a lunghezza focale di circa 50 mm. In particolar modo per gli scatti *frame* e grandangolari è stata utilizzata una Sony

Reflex DSC-HX400V Cybershot con funzionalità HDR mentre per gli scatti equirettangolari una Insta360 One x2, una fotocamera sferica, dotata di uno stabilizzatore di movimento e un giroscopio a sei vie, composta da due obiettivi *fish-eye* con un formato di 6080 x 3040 (2:1), con un'apertura di F2. 0, ISO da 100 a 3200. Prima delle effettive fasi di automodellazione i dati fotografici necessari ai processi image based sono stati sottoposti ad un accurato procedimento di *machine learning AI* al fine di correggere il bilanciamento colore, contrasto/luminosità e distorsione prospettica, tutte operazioni che snelliscono i calcoli per algoritmi SFM. L'integrazione di scatti derivanti da fotocamera digitale ad obiettivo variabile con fotogrammi prodotti da camera HDR 360° presuppone un il bilanciamento automatico dell'esposizione luminosa esterno-interno e, grazie alla visione panoramica, una più semplice gestione degli elementi per la ricostruzione 3d, ottimizzando le fasi di cloud points editing, meshing e texturing, al fine di generare, con il giusto compromesso qualità/tempo di sviluppo, un modello virtuale di alta qualità malgrado le condizioni di scarsa illuminazione e la mancanza di utilizzo di droni (Fig. 04). Il modello generato è così subito utilizzabile su piattaforme di esplorazione virtuale RTR come Unreal Engine 4/5, per tour virtuali interattivo esplorativi, o impiego in archivi digitali tridimensionali online o off-line (Fig. 05).



Fig. 05 Rendering in Unreal Engine 4 del modello poligonale con texture. Immagine degli autori.

L'edificio oggetto di studio presenta alcune peculiarità: i numerosi farette per l'illuminazione interna e la piccola apertura posta in alto nell'ambiente principale generano una condizione di controluce difficilmente bilanciabile con tecniche di correzione dell'esposizione e post-produzione. La similarità degli elementi dell'aula centrale a forma irregolare e le nicchie ad arco incassate nelle pareti circolari, come l'altezza dei solai arrotondati, sono elementi che per essere correttamente rilevati richiederebbero un numero considerevole di scatti da terra oppure l'utilizzo di un piccolo drone da interno. Inoltre, dovrebbe essere strutturato un piano di rilievo complesso per superare la difficoltà nel mettere in relazione i punti/pixel omologhi delle immagini digitali in allineamento in base alla loro qualità colorimetrica. Per ovviare a queste problematiche si è proposto un nuovo approccio di rilievo, aderendo ad una metodologia alternativa per ottenere in modo rapido e efficace, senza l'utilizzo eccessivo di chunks o marker in fase elaborativa, un modello in scala dello spazio, misurabile e fedele all'originale in modo da contenerne tutti i suoi segni di riconoscibilità: gli esclusivi dati 360 all'interno del software Metashape permettono nell'immediato di strutturare un modello complesso completo, rispettando livelli di scala e rapporti di proporzione tra gli elementi, mentre i dati

frame, che vengono nell'immediato associati dal software alla nuvola di punti generata dalle foto sferiche, conferiscono qualità visiva di dettaglio agli elementi rilevati. L'integrazione delle immagini 360 è quindi finalizzata esclusivamente alla funzione di *targeting* oggettuale e completamento strutturale del modello, riducendo molto le lacune e migliorando l'allineamento generale di tutti i dati fotografici utilizzati.

La Torre di Forca di penne e l'uso del drone: La Torre di Forca di Penne si trova in Abruzzo, in posizione baricentrica rispetto a due aree geografiche distinte, nel mezzo del valico che, assieme a quello di Popoli, costituiva l'unico passaggio della dorsale appenninica abruzzese. A pianta quadrata, con un'altezza in origine di circa 20 metri, situata su un piano roccioso, riversa ad oggi in condizioni critiche dal punto di vista strutturale, a causa dei due terremoti del 2009 e del 2016.

Nell'intorno della struttura sono visibili tracce di nervature che testimoniano la presenza in passato di architetture annesse alla torre. Le tre finestre (oggi solo due, la terza crollata durante l'ultimo terremoto) della parete rimasta in piedi, lasciano intendere che la torre fosse costituita da quattro piani fuori terra. L'attento sopralluogo dell'area di interesse del rilievo ha portato a considerare efficace il rilievo per mezzo di sistema UAV. Il luogo dove si trovano i resti della torre, infatti, ha una conformazione idonea per il rilievo fotogrammetrico eseguito per mezzo di drone, in quanto il manufatto si erge su uno sperone che guarda verso l'ampia vallata circostante, aperta su tutti i lati e priva di ostacoli. L'ampio spazio circostante facilita le manovre aeree del drone e consente di operare con una buona profondità di campo per la corretta sovrapposizione dei fotogrammi. Unico impedimento è stato il forte vento che soffiava da nord nel momento del rilievo. Il drone utilizzato per l'acquisizione delle fotografie è il dji Mini 2, un drone di fascia economica, provvisto di funzioni GPS, che si distingue però per l'ottima resistenza alla velocità del vento.

La fotocamera integrata ha un sensore 1/2,3" CMOS con 12MP, un obiettivo con FOV di 83° con formato di 35mm equivalente 24mm, un'apertura di f/2.8 e una distanza di ripresa di 1 m a ∞ . Le fotografie scattate per il rilievo sono di 4000x2250 pixels ad una risoluzione di 72 dpi, il tempo di esposizione è 1/500 sec, ISO 100, la lunghezza focale è di 4 mm. Il piano di volo è stato quindi strutturato in due fasi: la prima in cui si è scelto di eseguire degli scatti concentrici seguendo un tracciato elicoidale che andasse ad inquadrare tutti i lati del manufatto fino ad arrivare alla sommità; la seconda fase ha previsto dei cambi di quota che hanno interessato le quattro facce della torre e i quattro spigoli fra di esse, partendo dalla sommità della torre fino ad arrivare a circa 3 metri dalla base. Il GSD corrisponde a 8 mm ed è stato calcolato in base ai parametri della fotocamera e le dimensioni in pixel delle immagini [3].

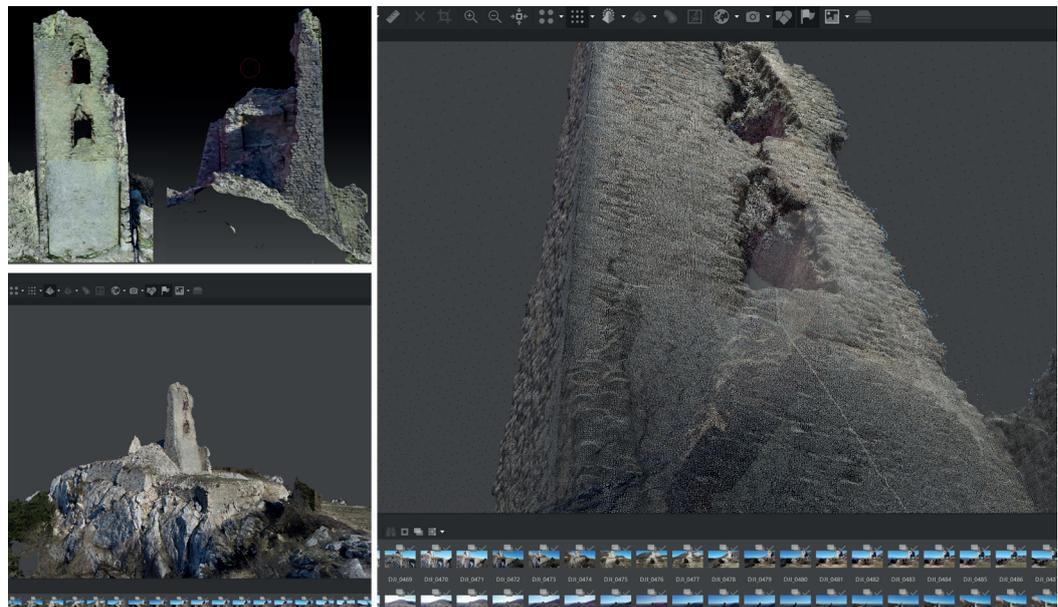


Fig. 06 a) Viste frontale, sinistra e prospettica della mesh con texture in Metashape, b) Dettaglio della dense cloud in Metashape. Immagine degli autori.

In questo modo si sono ottenuti degli scatti sequenziali delle superfici verticali della struttura che hanno incrementato il livello di dettaglio nei punti di maggiore interesse. Il risultato è un rilievo a più livelli di dettaglio del manufatto e di una porzione di terreno circostante. Il software utilizzato per l'elaborazione dei dati è Metashape, con il quale sono state gestite le quattro canoniche fasi per l'elaborazione del modello tridimensionale basate sull'algoritmo *Structure from Motion* (Fig. 6a). Il processo di manipolazione *della dense cloud* ha previsto una pulitura dai punti in eccesso e non utili alla costruzione del modello 3d (Fig. 6b) arrivando così ad una più coerente ed asciutta costruzione del modello mesh texturizzato nella fasi finali dell'elaborazione. Si nota come la mesh prodotta dal processo automatico sia costituita da una fitta maglia irregolare di triangoli che poco si presta all'esportazione su programmi come Unreal Engine o Twinmotion, ne consegue quindi una fase finale di retopology in quad format per mezzo della plug-in Zremesher di Zbrush e la creazione di asset modulari in grado di essere facilmente gestiti nelle successive fasi di composizione su piattaforma Unreal attraverso operazioni di *instance Cloning* (Fig. 7).



Fig. 07 Rendering fotorealistico in Unreal Engine 4. Immagine degli autori.

Conclusioni

Un veloce confronto sulle tre diverse metodologie applicative mette in luce alcuni aspetti comuni relativi alla facilità di acquisizione dei dati fotografici unitamente a una discreta precisione nei risultati - seppur non paragonabile a quella ottenibile mediante procedure di laser scanning. Il punto focale della ricerca è costituito da un equilibrato compromesso velocità-qualità di acquisizione spaziale in due dei tre casi studio: il Lidar malgrado qualche imprecisione, ha fornito in tempo reale un modello utilizzabile direttamente su piattaforme virtuali senza ricorrere ad ulteriori operazioni correttive, mentre il sistema integrato *Frame-Equirectangular Data*, relativo al secondo caso, ha semplificato di molto le fasi di editing costruttivo di uno spazio complesso *indoor* senza richiedere un massiccio utilizzo di *Chunks*. Il terzo caso, invece, ha fornito risultati esteticamente migliori, malgrado una ottimizzazione poligonale degli *asset* più complessa, dovuta soprattutto all'ampia scala del rilievo. In tutti i casi però si riscontra una buona corrispondenza metrico-proporzionale con i manufatti reali, con un margine di errore che varia circa da 0,5-1,5 cm, per quanto riguarda i primi due casi analizzati, e 3-5cm per la macro-scala del terzo caso studio. Considerando il rilievo image based e non laser scanner, i risultati sono stati piuttosto soddisfacenti. In conclusione, considerando l'obiettivo di ottenere modelli per una migliore navigabilità grafica su piattaforme virtuali,

l'utilizzo dei nuovi software di rilievo permettono di generare un modello 3d completo, texturizzato e misurabile, a scala variabile di approfondimento, pienamente compatibile con le dinamiche interattive, risultando una risorsa che potrebbe costituire una valida alternativa ai sistemi di condivisione delle informazioni finalizzate alla valorizzazione del *Cultural Heritage* e degli archivi 3d online e off-line.

Note

[1] Nell'ultimo decennio gli strumenti di rilievo digitale hanno avuto un cospicuo abbassamento dei costi ed hanno incrementato la loro facilità di utilizzo e la qualità dei risultati. Basti pensare alle moderne fotocamere presenti sui dispositivi mobili, sui droni o alla tecnologia Lidar dei nuovi iPhone, per avere contezza dell'enorme incremento qualitativo che questi dispositivi hanno avuto negli anni, pur essendo strumenti rivolti ad un'utenza commerciale e non professionale.

[2] Il Sistema Lidar presente nei dispositivi Apple di ultima generazione - nel caso descritto è stato utilizzato un iPhone 12 Pro Max - è il risultato di una ricerca iniziata un decennio fa. È un sistema basato su algoritmi V-SLAM (Visual Simultaneous Localization and Mapping): il dispositivo considera le proprie coordinate spaziali rilevando allo stesso tempo l'ambiente circostante attraverso più sensori. Il modello di elaborazione dei dati si basa sul calcolo in tempo reale delle nuvole di punti e delle mesh per mezzo della triangolazione e non attraverso operazioni in post produzione come nella fotogrammetria 3d.

[3] $GSD = d : D = f : H = 8 \text{ mm}$ dove d è la dimensione del pixel stabilita in base al rapporto tra la dimensione del sensore e il numero di pixel del fotogramma; dimensione del sensore $1/2,3''$, lunghezza del sensore in millimetri $6,4$, lunghezza in pixel dello scatto: 4000 , rapporto: $0,0016$; D è la distanza dei punti centrali tra due pixel (8mm); f è la lunghezza focale (4mm) e H è la distanza dalla superficie verticale della torre (20000).

Riferimenti Bibliografici

Basso, A. (2020). *Ambienti virtuali per nuove forme di comunicazione*. ROMA:Aracne Editrice.

Basso, A. (2019). The survey of the interiors of the Santo Stefano Rotondo complex through 3D photomodelling from integrated data with 360° HDR panoramas. In *The art of drawing/the drawing of art*. UID 2019 International congress.

Barazzetti, L. Previtali, M. Roncoroni, F. (2017). 3D modelling with the Samsung Gear 360. In: 2017 TC II and CIPA-3D Virtual Reconstruction and Visualization of Complex Architectures. International Society for Photogrammetry and Remote Sensing, 2017. p. 85-90.

Brusaporci, S. (2016) *Le murature nell'architettura del versante meridionale del Gran Sasso*. Roma: Gangemi Editore, pp. 99-100.

Brusaporci, S. (2011). Architetture cistercensi nell'Abruzzo aquilano: Misure, geometrie, proporzioni. Cistercian Architecture in the L'Aquila region of the Abruzzi: Measurements, geometries, proportions, in *Disegnare idee immagini*: 43, 2, 2011 Roma: Gangemi.

Cai, Z. Jungong H. Li L. Ling S. (2017). RGB-D Datasets Using Microsoft Kinect or Similar Sensors: A Survey. In *Multimedia Tools and Applications* 76(3): 4313–55

Cruz, L. Djalma L. Luiz V. (2012). *Kinect and RGBD Images: Challenges and Applications*. In Proceedings: 25th SIBGRAPI - Conference on Graphics, Patterns and Images Tutorials, SIBGRAPI-T 2012 (August): 36–49.

Di Vincenzo, B. (2012) Archeologia nel cratere sismico, in *Quaderni di Archeologia d'Abruzzo. Notiziario della Soprintendenza per i Beni Archeologici dell'Abruzzo*, 4, pp. 239-244: 240.

Ghisetti, A. (2017). *L'Aterno Pescara: Mistero di un fiume*. L'Aquila: Textus Edizioni.

Luigini, A, et al. (2020). Heritage education through serious games. A web-based proposal for primary schools to cope with distance learning. In *VITRUVIO*, vol. Vol 5, No 2 (2020), p. 72-85. < <https://polipapers.upv.es/index.php/vitruvio/article/view/14665>> (consultato il 12 febbraio 2022).

Mataloni, G. (2021). *Applicazioni geometriche nella modellazione 3d della chiesa di Santa Maria ad Cryptas*. In Santa Maria ad Cryptas. Storia, arte, restauri, pp. 155-159, Napoli: Editori Paparo.

Matsuda, N., Oliver, C., Mohit G. (2015). *MC3D: Motion Contrast 3D Scanning*. In 2015 IEEE International Conference on Computational Photography, ICCP 2015 - Proceedings.

Pérez Ramos, G., Robleda, P. (2016) *Only image based for the 3d metric survey of gothic structures by using frame cameras and panoramic cameras*. In *isprs-archives-XLI-B5-363-2016*. < <https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLI-B5/363/2016/> > (consultato il 4 febbraio 2022).

Perticarini M, Marzocchella V, Mataloni G (2021). *A Cycle Path for the safeguard of Cultural Heritage: Augmented reality and New LiDAR Technologies*. In Arena, A, et al. (a cura di): *CONNETTERE CONNECTING un disegno per annodare e tessere drawing for weaving relationships*. DISÉGNNO - OPEN ACCESS, p. 2571-2579, MILANO: FrancoAngeli.

Autori

Maurizio Perticarini, Dipartimento di Architettura e Disegno Industriale, Università degli Studi Luigi Vanvitelli II Università di Napoli, maurizio.perticarini@unicampania.it
Alessandro Basso, Università degli Studi di Camerino, alessandro.basso@unicam.it

Per citare questo capitolo: Perticarini Maurizio, Basso Alessandro (2022). *Visualità digitale applicata a metodologie di rilievo integrato. Sinergie collaborative tra sperimentazione e tecnologia/Digital Visualization applied to integrated survey methodologies. Collaborative synergies between experimentation and technology*. In Battini C., Bistagnino E. (a cura di). *Dialoghi. Visioni e visualità. Testimoniare Comunicare Sperimentare. Atti del 43° Convegno Internazionale dei Docenti delle Discipline della Rappresentazione/Dialogues. Visions and visuality. Witnessing Communicating Experimenting. Proceedings of the 43rd International Conference of Representation Disciplines Teachers*. Milano: FrancoAngeli, pp. 2734-2750.



Digital Visualization applied to integrated survey methodologies. Collaborative synergies between experimentation and technology

Maurizio Perticarini
Alessandro Basso

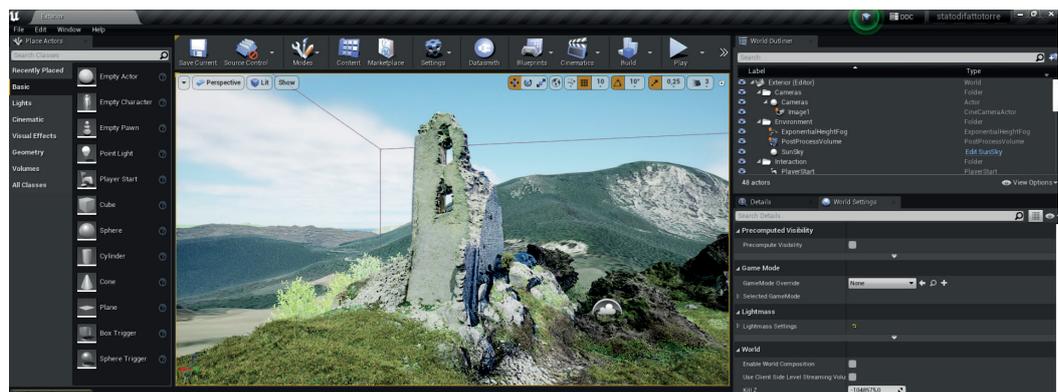
Abstract

The progressive adherence to Computer Vision has enhanced the use of the image as the main means of communication, integrating automatic interpretative processes that go beyond the structural boundaries related to place and time, in order to have direct access to processes of language sharing that allow everyone to acquire new perceptual tools that offer, in fact, a better adaptability to the interpretive modes of cultural participation, related to concepts of exceptional versatility. Technology and Computer Vision, therefore, establish a structural relationship between representation and experimentation, between art and science, fueling a mutual collaboration between the various disciplinary areas. The paper in question wants to investigate these issues, comparing the specific digital survey with three case studies with intrinsic characteristics, in relation to photomodeling workflows, optimization of 3D assets and operations of virtual migration in interactive spaces. The contribution aims to analyze different application systems for low-cost 3D photogrammetry through a precise pipeline, clarifying how three different methodologies of spatial acquisition and digital restitution can seek the same perceptually photorealistic results, which can be used in virtual cyberspace, documenting how a reconfigured space can combine multiple phases of investigation to approach a graphic vision of objective synthesis. Digital reconfiguration is thus revealed, with all its limits and paradoxes, as a powerful and versatile metalanguage in continuous evolution.

Keywords

Lidar sensor, low-cost survey, image based, Computer Vision, real time rendering

Screenshot of the real time rendering in Unreal Engine.



Introduction

Technology, as a tool to facilitate the design of ideas, to which mankind over the centuries has turned betting resources, work and research, becomes, in the light of a common reflection on the value of cultural dialogue, a real language through which it is possible not only to transmit information, outlining the prevailing theoretical-applicative synergies that identify the field of Architectural Representation and other areas, but above all to structure participatory interchanges of a functional nature that relate areas apparently extremely diverse. The adherence to Computer Vision, increasingly prevalent in recent times, has enhanced the use of the image, which has always been the main means of communication whose evolutionary path is certainly connected with that of man, integrating automatic interpretative processes that transcend the structural boundaries related to place and time, in order to have direct access to sharing processes that allow everyone to use new perceptual tools capable of offering a better adaptability to the interpretive methods of cultural sharing, related to concepts of exceptional versatility. Technology and Computer Vision, set up a solid structural relationship between Representation and experimentation, between art and science, feeding, during this Third Digital Revolution (despite a historical period that continues to be difficult in many respects) the collaboration between the various disciplinary areas that use drawing as a solid resource to achieve concrete results. Restricting the field to the theme of digital survey, the contribution that is the object of this research aims to investigate, through comparisons and experimental in-depth studies, how effective are the design and representation methods that involve the new graphic tools in order to return - in a fluid manner and seeking a univocal interpretative result - a perception superimposed on the real one and comprehensible to all of the digitised places. Although the aim is to generate graphic works that trace the actual state of affairs, not all digital survey methods offer an immediate possibility of interpretation. Some need to be optimised in order to be understood and interpreted by those who do not have a solid cultural background, while other methods, due to their graphic nature, can only be decoded by following specific cultural-interpretative schemes, which are very different from visual-cognitive perception that only follows physiological and psychological indications. With the passing of time, however, there is a common tendency to move towards a true cloning proportional to reality that can be interpreted visually in a direct manner and that can therefore be as coherent as possible (in terms of depth perception and chromatic acquisition) with the real space. This trend is helped by the technological evolution of the hardware, but also by human identification that goes beyond the five senses and works very well in synergy with current digital survey methods. This ability to interpret the depth of digital spaces, whether cloned or non-existent, is the strong point that makes it possible to make the virtual dimen-



Fig. 01 a) Photograph of the aedicule and altar in Santa Maria ad Cryptas
b) Equirectangular photographs of the main room of the Furnace
c) Photograph of the tower before the 2016 earthquake. Image by authors.

sion "real", and it is this surrogate of photorealistic reality that all the next-generation tools for virtual cloning and digitized space surveying (in addition to the possibility of acquiring measurements and data) are aiming at. The research therefore compares three case studies - conducted over the last two years - that use three different low-cost spatial acquisition [1] and digital restitution methodologies with the aim of obtaining the same perceptually photorealistic results that can be used in migrations on an interactive platform, identifying results and criticalities on the basis of diversified workflows.

Three case studies compared.

The research focuses on three case studies that have different characteristics, both formal and spatial: an architectural object inserted in a closed and dimly lit space (votive aedicule in the church of Santa Maria ad Cryptas) (Fig. 1a), an interior environment with a complex shape (Fornace del Canova) (Fig. 01b) and a building characterised by a high height (Torre di Forca di Penne) (Fig. 01c). The survey of these three historical artefacts is based on three different survey technologies, managing to give a discrete photorealism to the acquired objects despite the speed of data acquisition and processing. Three different approaches that have in common the use of low-cost technologies that are accessible to everyone, simple to use and capable of obtaining objects that are valid to be displayed in an interactive space. The first case described is the survey carried out by means of a portable Lidar system of the aedicule to the left of the central altar of the church of Santa Maria ad Cryptas in Fossa; the second case deals with the survey carried out by means of photogrammetry based on spherical images of the interior of the Fornace del Canova in Rome, once used by the artist to bake his clay models, and today Canova 22, seat of a cultural association, ; the third case describes the survey carried out by means of aerial photogrammetry (Drone Djl Mini 2) of the Torre di Forca di Penne, a medieval tower that stands in a strategic position in the homonymous pass in the province of L'Aquila. The three cases were compared and studied from the point of view of both reliability with reality and speed of execution of the survey. Each of the technologies has its own strengths and weaknesses, depending on the intrinsic factors of the survey and the type of object to be acquired.

The Church of Santa Maria ad Cryptas and the use of Lidar for the survey of an architectural object: The church of Santa Maria ad Cryptas takes its name from some nearby caves. It is located in Fossa (AQ) in the Valle Subequana, at the foot of Monte d'Ocre.

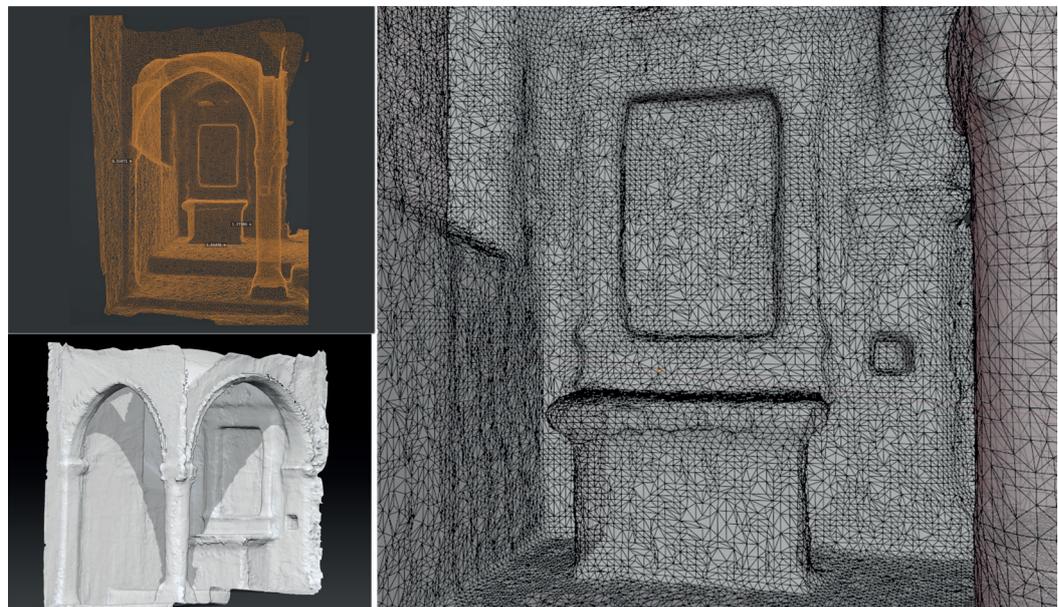


Fig. 02 a) Dense Cloud made with Polycam and mesh made with Polycam b) Wireframe of the mesh Image by authors.

It was built in the second half of the 1200s by Cistercian monks. The small altar below a late 17th-century aedicule, the subject of the study, dates back to the time of construction of the architectural complex and has not undergone the changes that transformed much of the rest of the church in the Baroque period. In the survey of the altar and aedicule, the use of the LIDAR application for iPhone [2] was experimented with, which uses the 3D sensor to create point clouds and meshes. To date, many smartphone manufacturers have included depth sensors, but Apple's Lidar, developed in collaboration with Lumentum, is an evolution of the previous integrated technologies, different from the more canonical laser scanner systems that have a bulky sensor that rotates on itself and different from previous systems for a stable continuous scan. It has a range of 5 meters and an accuracy of 3 mm to 9 mm. The aedicule and the small altar were scanned using the 'Polycam' application which allows scanning from 13 mm to 80 mm Voxel size and a simplification from 0% to 99% (Fig. 02a). The 'Object' simplification arrangement was used, which reduced the time of the realtime scanning process to 5 minutes and took 3 minutes to process the mesh consisting of 54.9k vertices (Fig. 02b). It was therefore possible to perfectly trace the original shape of the environment, as the point cloud covers a rather circumscribed area through a consistent number of available points. The digital asset obtained, without further steps other than the canonical ones of the creation of details through depth maps (Normal-Bump-Displacement), is ready for the eventual insertion on RT interactive rendering platforms such as Unity or Unreal Engine (Fig. 03). In short, the system accelerates the workflow, allowing to obtain an average well-structured model that can be used immediately in digital visualization contexts.



Fig. 03 RTRendering in Twinmotion of no-optimized model, exported from Polycam with texture and UVMap. Image by authors.

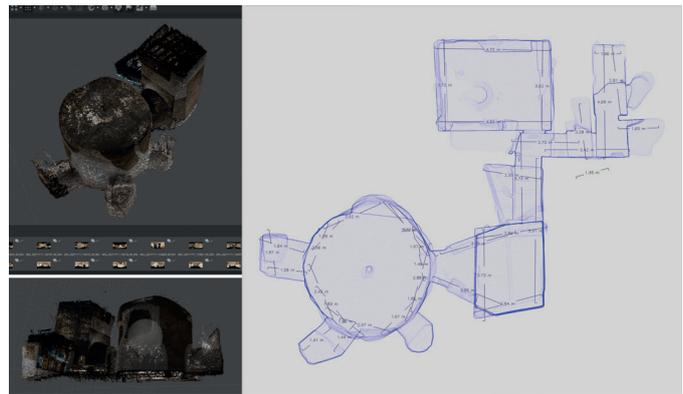


Fig. 04 Perspective views of the dense Cloud in Metashape, and dimensional plan of the Furnace. Image by authors.

La Fornace del Canova ed il rilievo integrato 360°: The second case study proposes the survey of the Ex Furnace Canova, in the heart of Rome since the early 1600s. The furnace has a circular shape with three side niches and a central hall rising upwards with a small open skylight at the top. The irregular structure with a central plan, although well preserved, presents some critical characteristics for a 3D photogrammetric survey, such as the lack of sufficient lighting and rather narrow linked rooms at varying heights that converge in the large furnace hall. The research, therefore, proposes a survey carried out using low-cost instruments and very short timescales, with the application of a rapid methodology supported exclusively by terrestrial 3D photomodelling techniques, without resorting to any use of laser scanners (TLS), LIDAR instruments, or indoor mini drones, but using only integrated photographic data of various types: 360° HDR panoramas, wide-angle fisheye photos and frame format photographs with a focal length of about 50 mm. In particular, for the frame and wide-angle shots, a Sony Reflex DSC-HX400V Cybershot with HDR functionality was used, while for the equirectangular shots an Insta360 One x2, a spherical camera, equipped with a motion stabilizer and a six-way gyroscope, composed of two fisheye lenses with a format of 6080 x 3040 (2:1), with an aperture of F2.0, ISO from 100 to 3200. Before the automatic modelling stages, the photographic data required for the image-based processes were subjected to a

thorough AI machine learning process in order to correct the color balance, contrast/brightness and perspective distortion, all of which streamline the calculations for SFM algorithms. The integration of variable lens digital camera shots with shots produced by 360° HDR camera requires an automatic external-internal light exposure balance and (thanks to the panoramic view) an easier management of the elements for 3D reconstruction, optimizing the cloud points editing, meshing and texturing phases, in order to generate a high-quality virtual model with the right quality/development time compromise, despite the low light conditions and not using drones (Fig. 04). The generated model is immediately usable on RTR virtual exploration platforms such as Unreal Engine 4/5, for interactive explorative virtual tours, or use in online or offline three-dimensional digital archives (Fig. 05).



Fig. 05 Rendering in Unreal Engine 4. Image by authors.

The building studied has some peculiarities: the numerous spotlights for internal lighting and the small opening at the top of the main room generate a backlight condition that is difficult to balance with exposure correction and post-production techniques. The similarity of the irregular-shaped elements of the central hall and the arched cavities embedded in the walls, as well as the height of the rounded ceilings, are elements that would require a considerable number of shots from the ground or the use of a small indoor drone to be detected correctly. In addition, a complex survey plan would have to be structured to overcome the difficulties of the photomodelling software in coordinating the numerous images and relating homologous points according to their colorimetric quality. To overcome these problems, a new survey approach was proposed, using an alternative methodology to quickly obtain, without the excessive use of chunks or targets, a scaled model of the space, measurable and faithful to the original: the exclusive 360 data within the Metashape software allow the immediate structuring of a complete complex model, respecting scale levels and proportion ratios between the elements, while the Frame data, which are immediately associated by the software to the point cloud generated by the spherical photos, give visual quality of detail to the elements detected at a lower height and therefore closer to a virtual observer. The integration of the 360 images is therefore aimed exclusively at the function of object targeting and structural completion of the model, greatly reducing the gaps and improving the general alignment of all the photographic data employed.

The Torre di Forca di Penne and the use of the drone: The Torre di Forca di Penne is located in Abruzzo, in a central position with respect to two distinct geographical areas, in the

middle of the pass which, together with that of Popoli (PE), constituted the only passage of the Abruzzo Apennine Mountain ridge. With a square plan, originally about 20 meters high, located on a rocky plateau, today it is in critical conditions from a structural point of view, due to the two earthquakes of 2009 and 2016. Around the structure are visible traces of ribs that testify to the presence in the past of architecture annexed to the tower. The three windows (today only two, the third collapsed after the last earthquake) of the still standing wall suggest that the tower consisted of four floors above ground. The careful inspection of the area of interest led to consider the survey by means of the UAV system as the most effective. The place where the remains of the tower are located, in fact, has a suitable conformation for the photogrammetric survey from drone, because the artifact stands on a spur overlooking the wide surrounding valley, open on all sides and free of obstacles. The large surrounding space facilitates the aerial maneuvers of the drone and allows to operate with a good depth of field for the correct overlapping of the frames. The only impediment was the strong wind that blew from the north at the time of the survey. Even the glare of sunlight caused some discomfort, as there was clear sky during the whole time of the survey. The drone used for capturing the photographs is the DJI Mini 2, a low-cost drone with integrated GPS that stands out for its excellent resistance to wind speed. The integrated camera has a 1 / 2.3 "CMOS sensor with 12MP, an 83 ° FOV lens with a 35mm equivalent 24mm format, an aperture of f / 2.8 and a shooting distance of 1m at ∞ . The photographs taken for the survey are 4000x2250 pixels at a resolution of 72 dpi, the exposure time is 1/500 sec, ISO 100, the focal length is 4mm. The flight plan was therefore structured in two phases: the first in which it was decided to perform concentric shots following a helical path that would frame all sides of the building up to the top; the second phase involved vertical changes in height that affected the four faces of the tower and the four edges between them, starting from the top of the tower up to about 3 meters from the base. The GSD corresponds to 8 mm and was calculated based on the camera parameters and the pixel dimensions of the images[3]. In this way, sequential shots of the vertical surfaces of the structure were obtained, which increased the level of detail in the points of greatest interest.

The result is a multi-level detailed survey of the building and a portion of the surrounding land. The software used for data processing is Metashape, with which the four canonical phases for processing the three-dimensional model based on the Structure from Motion algorithm were managed (Fig. 06a). The process of manipulating the dense cloud has provided for a cleaning of excess points that are not useful for the construction of the 3D model (Fig. 06b) (the survey involved very large open spaces), arriving at a more consistent construction

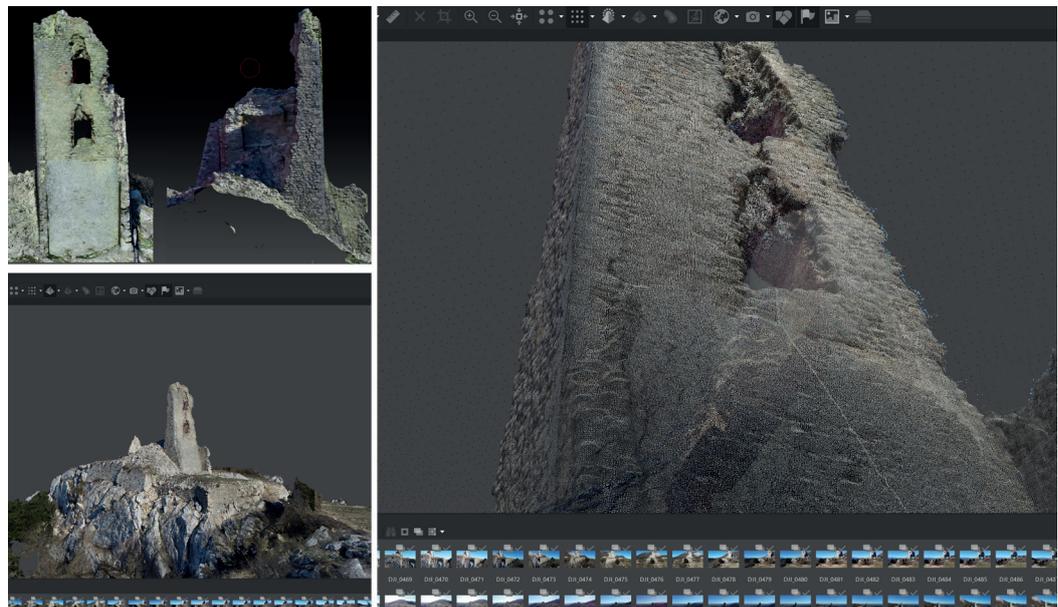


Fig. 06 a) Front, left elevations and perspective views of the Mesh textured in Metashape, b) Detail of the dense cloud in Metashape. Image by authors.

of the textured mesh in the final stages of processing. The mesh produced by the automatic process is made up of a dense irregular mesh of triangles that is not very suitable for exporting to programs such as Unreal Engine or Twinmotion, therefore a final phase of retopology in quad format follows by means of Zbrush and the creation of modular assets that can be easily managed in subsequent operations (Fig. 07).

Conclusions

A quick comparison on the three different application methodologies highlights some common aspects relating to the ease of acquisition of photographic data together with a fair precision in the results - although not comparable to that obtainable through laser scanning procedures. The focal point of the research is constituted by a balanced compromise between speed and quality of spatial acquisition in two of the three case studies: the Lidar, despite some inaccuracies, provided in real time a model that can be used directly on virtual platforms without resorting to further corrective operations; regarding the integrated Frame-Equirectangular Data system, relating to the second case, it has greatly simplified the constructive editing phases of a complex indoor space without requiring a massive use of Chunks. The third case, on the other hand, provided aesthetically better results, despite a more complex polygonal optimization of the assets, mainly due to the large scale of the survey. In all cases, however, there is a good metric-proportional correspondence with the real artifacts, with a margin of error that varies from about 0.5-1.5 cm, as regards the first two cases analyzed, and 3-5cm for the macro-scale of the third case study. Considering the image based and non laser scanner survey, the results were quite satisfactory. In conclusion, considering the objective of obtaining models for better graphic navigability on virtual platforms, the use of the new survey software allows the generation of a complete, textured and measurable 3D model, with a variable scale of study, fully compatible with the dynamics. interactive, resulting in a resource that could constitute a valid alternative to information sharing systems aimed at enhancing the Cultural Heritage and online and off-line 3D archives.



Fig. 07 Photorealistic Rendering in Unreal Engine. Image by authors.

Notes

[1] In the last decade digital surveying tools have become cheaper and have increased their ease of use and the quality of results. It is enough to think of the modern cameras present on mobile devices, drones or Lidar technology of the new iPhone, to get an idea of the enormous increase in quality that these devices have had over the years, although they are tools aimed at a commercial user and not professional.

[2] The Lidar System present in the latest generation of Apple devices - in the case described, an iPhone 12 Pro Max was used - is the result of research begun a decade ago. It is a system based on V-SLAM (Visual Simultaneous Localization and Mapping) algorithms: the device considers its own spatial coordinates while detecting the surrounding environment through multiple sensors. The data processing model is based on the real-time calculation of point clouds and meshes by means of triangulation and not through postproduction operations as in 3d photogrammetry.

[3] $GSD = d/D = f/H = 8 \text{ mm}$ where d is the size of the pixel established on the basis of the ratio between the size of the sensor and the number of pixels in the frame; sensor size $1/2.3''$, sensor length in millimeters 6.4, length in pixels of the shot: 4000, ratio: 0.0016; D is the distance of the central points between two pixels (8mm); f is the focal length (4mm) and H is the distance from the vertical surface of the tower (20000)

References

- Basso, A. (2020). *Ambienti virtuali per nuove forme di comunicazione*. ROMA:Aracne Editrice.
- Basso, A. (2019). The survey of the interiors of the Santo Stefano Rotondo complex through 3D photomodelling from integrated data with 360° HDR panoramas. In *The art of drawing/the drawing of art*. UID 2019 International congress.
- Barazzetti, L. Previtali, M. Roncoroni, F. (2017). 3D modelling with the Samsung Gear 360. In: 2017 TC II and CIPA-3D Virtual Reconstruction and Visualization of Complex Architectures. International Society for Photogrammetry and Remote Sensing, 2017. p. 85-90.
- Brusaporci, S. (2016) *Le murature nell'architettura del versante meridionale del Gran Sasso*. Roma: Gangemi Editore, pp. 99-100.
- Brusaporci, S. (2011). Architetture cistercensi nell'Abruzzo aquilano: Misure, geometrie, proporzioni. Cistercian Architecture in the L'Aquila region of the Abruzzi: Measurements, geometries, proportions, in *Disegnare idee immagini*: 43, 2, 2011 Roma: Gangemi.
- Cai, Z. Jungong H. Li L. Ling S. (2017). RGB-D Datasets Using Microsoft Kinect or Similar Sensors: A Survey. In *Multimedia Tools and Applications* 76(3): 4313–55
- Cruz, L. Djalma L. Luiz V. (2012). *Kinect and RGBD Images: Challenges and Applications*. In Proceedings: 25th SIBGRAPI - Conference on Graphics, Patterns and Images Tutorials, SIBGRAPI-T 2012 (August): 36–49.
- Di Vincenzo, B. (2012) Archeologia nel cratere sismico, in *Quaderni di Archeologia d'Abruzzo. Notiziario della Soprintendenza per i Beni Archeologici dell'Abruzzo*, 4, pp. 239-244: 240.
- Ghisetti, A. (2017). *L'Aterno Pescara: Mistero di un fiume*. L'Aquila: Textus Edizioni.
- Luigini, A, et al. (2020). Heritage education through serious games. A web-based proposal for primary schools to cope with distance learning. In *VITRUVIO*, vol. Vol 5, No 2 (2020), p. 72-85. < <https://polipapers.upv.es/index.php/vitruvio/article/view/14665>> (consultato il 12 febbraio 2022).
- Mataloni, G. (2021). *Applicazioni geometriche nella modellazione 3d della chiesa di Santa Maria ad Cryptas*. In Santa Maria ad Cryptas. Storia, arte, restauri, pp. 155-159, Napoli: Editori Paparò.
- Matsuda, N., Oliver, C., Mohit G. (2015). *MC3D: Motion Contrast 3D Scanning*. In 2015 IEEE International Conference on Computational Photography, ICCP 2015 - Proceedings.
- Pérez Ramos, G., Robleda, P. (2016) *Only image based for the 3d metric survey of gothic structures by using frame cameras and panoramic cameras*. In *isprs-archives-XLI-B5-363-2016*. < <https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLI-B5/363/2016/>> (consultato il 4 febbraio 2022).
- Perticarini M, Marzocchella V, Mataloni G (2021). *A Cycle Path for the safeguard of Cultural Heritage: Augmented reality and New LiDAR Technologies*. In Arena, A. et al. (a cura di): *CONNETTERE CONNETTING un disegno per annodare e tessere drawing for weaving relationships*. DISÉGNNO - OPEN ACCESS, p. 2571-2579, MILANO: FrancoAngeli.

Authors

Maurizio Perticarini, Dipartimento di Architettura e Disegno Industriale, Università degli Studi Luigi Vanvitelli II Università di Napoli, maurizio.perticarini@unicampania.it
Alessandro Basso, Università degli Studi di Camerino, alessandro.basso@unicam.it

To cite this chapter: Perticarini Maurizio, Basso Alessandro (2022). *Visualità digitale applicata a metodologie di rilievo integrato. Sinergie collaborative tra sperimentazione e tecnologia/Digital Visualization applied to integrated survey methodologies. Collaborative synergies between experimentation and technology*. In Battini C., Bistagnino E. (a cura di). *Dialoghi. Visioni e visualità. Testimoniare Comunicare Sperimentare. Atti del 43° Convegno Internazionale dei Docenti delle Discipline della Rappresentazione/Dialogues. Visions and visuality. Witnessing Communicating Experimenting. Proceedings of the 43rd International Conference of Representation Disciplines Teachers*. Milano: FrancoAngeli, pp. 2734-2750.