Analisi comparative tra i modelli 3D fotogrammetrici creati dall'elaborazione diretta dei dati del sensore e dalle immagini RGB dopo il trattamento *raw-conveter*

Alessio Cardaci, Pietro Azzola, Michele Bianchessi, Ruggero Folli e Simone Rapelli

¹ Dipartimento di Ingegneria e Scienze Applicate, Università degli studi di Bergamo alessio.cardaci@unibg.it; pietro.azzola@unibg.it, m.bianchessi@studenti.unibg.it; r.folli@studenti.unibg.it; <u>s.rapelli@studenti.unibg.it</u>

Abstract.

Il caso studio del *Torresino da Polvere* di via Beltrami, in città alta a Bergamo, ha fornito il pretesto per uno studio comparativo tra i modelli 3D generati dall'elaborazione diretta dei dati del sensore fotografico e quelli ottenuti dal processamento di immagini dopo i tradizionali trattamenti *raw-converter*. Interessanti considerazioni hanno evidenziato delle singolari peculiarità di alcuni software nella gestione del dato grezzo, sia per un'elevata accuratezza delle restituzioni, sia per un'inaspettata rapidità di calcolo.

Parole chiave: 3D architectural survey, Torresini da polvere, Bergamo.

1 Introduzione

Il formato comunemente indicato come *raw* contiene le informazioni legate all'intensità luminosa misurata dal sensore fotometrico, un 'negativo digitale" privo di modifiche ed alterazioni. I dati *raw* sono delle sequenze di *bit* salvati all'interno di *files* proprietari secondo codici specifici stabiliti dai diversi costruttori e, tra loro, incompatibili gli uni con gli altri. I produttori di camere digitali, per motivi meramente commerciali, sono stati poco inclini nell'adottare uno standard universale come è stato dimostrato dal limitato 'successo' del format *dng* perché boicottato dai grandi marchi quali, ad esempio, Canon®, Nikon® e Sony®.

Una fotografia, anche della stessa scena e nelle stesse condizioni di illuminazione, sarà quindi differente se acquisita con fotocamere diverse, sia perché dipendente dalle diverse sensibilità spettrali dei sensori, sia dalla differente 'conservazione' dell'informazione (Nguyen 2014). In più il dato *raw* non rappresenta un'immagine strutturata e visibile grazie a pixel colorati ma, bensì, una lista di valori registrati dai



Fig. 1. Il 'Torresino da Polvere' di via Beltrami a Bergamo; a sinistra lo scorcio del modello 3D (la proiezione del modello fotogrammetrico sulla nuvola di punti 3D *laser scanner*), a destra immagini fotografiche dell'architettura dopo gli interventi di restauro degli anni '80

rivelatori fotosensibili per mezzo di una matrice a filtri - sensibile alle frequenze del rosso, del verde e del blu - disposti secondo il noto e consueto *Array di Bayer* (Faser 2004).

La creazione di una figurazione realistica è il frutto di una conversione operata attraverso algoritmi *raw-converter* che trasformano i valori di luminosità di ogni fotorecettore (eseguendo in modo sequenziale un processamento di *mapping linear values, white balancing, demosaicing, colour space conversion* e gamma correction) in una tabella ordinata di dati RGB (Rojtberg 2017).

La principale funzione dei software *raw-converter* è di 'stimare' la quantità di luce catturata dalla camera e renderla visibile attraverso l'informazione colorimetrica; un processo fortemente condizionato sia dall'elettronica della camera, sia dalle scelte dell'operatore (McHugh 2018).

La trasformazione del dato grezzo in codice *dng*, *tif* o *jpg* è quindi un'azione che comporta sia una alterazione dell'informazione (causata dall'interpretazione del valore) che una perdita di parte dei contenuti (a seguito di compressione e/o della riduzione dalla profondità di bit, generalmente da 14 ad 8, nonché di operazioni di riduzione del rumore e aumento del contrasto e della nitidezza).

Un processo che in letteratura è valutato come un valore aggiunto in quanto permette di migliorare la qualità delle fotografie, sia consentendo la correzione degli errori effettuati in fase di cattura (regolazione delle tonalità di colore, aumento dei contrasti, filtraggio del rumore di fondo, ...), sia per la possibilità di fusione di scatti multipli con elaborazioni *merging* e/o HDR (Verhoeven 2010).

La ricerca proposta, prendendo a pretesto il caso studio del 'Torresino da Polvere' di via Beltrami in città alta a Bergamo, ha voluto condurre uno studio comparativo tra i modelli 3D fotogrammetrici estratti dall'elaborazione diretta dei dati del sensore fotografico con quelli, invece, ottenuti dal processamento di immagini RGB dopo il tradizionale trattamento *raw-converter*.

2 Il rilievo 3D laser scanner e la rete topografica.

Il rilievo 3D del 'Torresino da Polvere' di via Beltrami, una piccola architettura dalla geometria regolare e legata alle Mura Venete della città, ha costituito quindi la cornice di studio necessaria per la sperimentazione (fig. 1). La forte presenza di vegetazione sulla copertura e il parziale interramento di due dei suoi lati, non ha consentito la restituzione del modello integrale dell'edificio ma solamente di una parte di esso; tratti murari comunque più che soddisfacenti per delle riflessioni puntuali ed attendibili.

Le misurazioni sono state eseguite con più strumentazioni e diverse metodologie di acquisizione, prassi ormai consolidata nelle nuove procedure di rilievo integrato (Cardaci & Versaci 2018).

E stata approntata una rete topografica le cui coordinate dei vertici, materializzati con prismi posti su tripodi fissi, sono state determinate con t*otal station* e operando un centramento forzato, metodo che permettendo l'intercambiabilità tra segnale e strumento ha consentito una indeterminatezza inferiore al millimetro.

La risoluzione delle equazioni iperdeterminate della rete (sia per eccedenza delle misurazioni delle distanze che quelle angolari) è stata effettuata attraverso software di compensazione ai minimi quadrati; nel caso specifico è stato impiegato il noto prodotto commerciale MicroSurvey® *Star*Net* perché in grado di coniugare la semplicità di utilizzo e le essenziali funzioni ad una grande affidabilità e un elevato rigore. Dai vertici della rete, attraverso intersezioni in avanti e rilevazioni distanziometriche, sono state acquisite le coordinate dei centri di target cartacei posizionati sulle pareti dell'edificio; essi, in numero elevato e ben studiati nella geometria e nella distribuzione spaziale, hanno costituito i *Ground Control Points* (GCP) del sistema di georeferenziazione dei modelli *3D laser scanner* e fotogrammetrici.

Il rilievo topografico è stato integrato con un rilievo a sensori attivi al fine di avere un numero molto più elevato di punti a coordinate note; la nuvola di punti ottenuta ha costituito, infatti, il riferimento metrico di riferimento per il confronto dei vari modelli. Al fine di una ottimizzazione dei tempi sono state registrate le sole informazioni di riflettanza e non il dato colorimetrico; il fine della sperimentazione era infatti la verifica dell'accuratezza dei modelli che rendevano superfluo la conoscenza del colore.

La creazione di un'unica nuvola globale di progetto è stata effettuata con un iniziale allineamento delle singole scansioni sui target prima presentati, e quindi un avvicinamento delle varie porzioni per mezzo di algoritmi 'a riconoscimento di forma'. L'uso di riferimenti esterni come sfere o scacchiere ha consentito una registrazione più rigida delle scansioni nonché una verifica delle imprecisioni sulla base delle differenze dei target misurati con il *laser scanner* e la posizione dei vertici della rete topografica (Liscio 2016; Cox 2015). La registrazione delle scansioni è stata operata attraverso il software proprietario dello strumento, il Faro® *Scene*, e la verifica delle accuratezze è stata effettuata di per mezzo dell'applicativo della ATS® *Quality Manager* in grado effettuare il calcolo delle tensioni, della compensazione e restituire le statistiche dei *Ground Control Points* (GCP). È stata così accertata l'ottima sovrapposizione delle singole nuvole con un'approssimazione media contenuto all'interno della precisione strumentale (± 1.5 mm).



Fig 2. Il modello 3D laser scanner utilizzato per il confronto dei modelli fotogrammetrici.

La nuvola globale, fusione delle singole scansioni dopo esser state sottoposte ad una riduzione del rumore e ad un campionamento dei punti per eliminare quelli sovraposti (perché molto prossimi tra loro all'interno sfera ideale di un millimetro di raggio), ha quindi restituito una nuvola densa che è stata impiegata per il confronto delle nuvole fotogrammetriche (fig. 2).

Lo scostamento tra una geometria 'certa' e affidabile (quella del *3D laser scanner*) e il modello virtuale ricreato sulla base dei diversi trattamenti fotogrammetrici, ha permesso di valutare, dunque, il modello metricamente più accurato.

3 Il processamento dei dati raw e delle immagini RGB

La generazione dei modelli 3D fotogrammetrici è stata ottenuta dal processamento dello stesso set di cattura fotografica. La sperimentazione si è basata sulla realizzazione di cinque artefatti virtuali: due generati dall'elaborazione diretta dei dati del sensore salvati sulla scheda di memoria della camera fotografica (*RAW Image Sensor* e *JPG Image Sensor*), tre creati dalle immagini RGB dopo la trasformazione dei dati grezzi (*JPG Raw-Converter, TIF Raw-Converter* e *DNG Raw-Converter*). La conversione è stata attuata per mezzo del software Adobe@ *Camera Raw* applicando a tutte le fotografie del set gli stessi parametri di trasformazione e, quindi, salvandoli in formati differenti: *dng, tif* e *jpg*.

Le informazioni grafiche del *RAW Image Sensor e delle immagini RGB JPG Raw-Converter*, *TIF Raw-Converter* e *DNG Raw-Converter* sono state catturate/conservate la stessa risoluzione di circa 21 Mega-Pixel (5616x3744), mentre quelle *JPG Image Sensor* leggermente inferiore, pari a 5.2 Mega-Pixel (2784x1856). Questo a causa delle impostazioni della camera fotografica utilizzata per l'esperienza, una reflex Canon@

EOS 5D mark II con ottica fissa Canon@ 24mm/1.4, settata in modo da ottenere dei file di dimensione contenute.

Lo sviluppo digitale delle informazione *raw* del sensore è prassi consigliata - a volte obblgata perchè il dato grezzo non è importabile in molti appliocativi commerciali - in ambito fotografico sia perché in grado di correggere eventuali errori durante la cattura (errata esposizione, bilanciamento del bianco, ...), sia perché permette di ottenere immagini RGB 'personali' più realistiche grazie alla manipolazione della temperatura cromatica, della regolazione delle luci e delle ombre, della riduzione del rumore e del miglioramento della nitidezza (Carpiceci 2012).



Name	Minimum Quality	Average Quality	Maximum Quality	Image Aligned
JPG Raw-Converter	60,00	2027,50	3995,00	76,26 %
TIF Raw-Converter	98,00	1550,00	3002,00	76,98 %
DNG Raw-Converter	48,00	1103,00	2158,00	74,10 %
JPG Image Sensor	402,00	2221,00	4040,00	97,84 %
RAW Image Sensor	309,00	2393,50	4478,00	100,00 %

Tab 1. L'indice di qualità delle immagini: tabelle e grafici comparativi tra i vari formati.



Fig 3. Parametri imposti per ogni formato nel processamento del set di immagini.

Le prime verifiche hanno subito evidenziato che il trattamento *raw-converter* sembra apportare solo un miglioramento della qualità visiva, percepibile dall'occhio del fotografo ma non dai software di ricostruzione 3D. Test più accurati e finalizzati alla valutazione della qualità hanno infatti evidenziato come nelle immagini processate sia sensibilmente aumentata la sfocatura e le aree con *texture* con colorazione uniforme, come dimostrato dalle tabelle e dai grafici riportati (Tab. 01).

Il valore dell'indice di qualità è stato restituito da specifici algoritmi di elaborazione di analisi multiscala con scomposizione in frequenza (Pirotti et al. 2006), soluzione più affidabile rispetto ai più semplici applicativi comunemente impiegati di rilevamento diretto dei bordi. Un metodo rigoroso finalizzato a conoscere le caratteristiche complessive di ogni singola immagine (Alfio et al. 2020); nella restituzione in tabella, sono stati riportati il *minimum value*, l'*average value* e il *maximum value* dell'intero set costituito da 140 immagini. I risultati migliori, come intuibile, sono stati quelli relativi alle catture dirette *raw* ma, ha sorpreso, il confronto tra i due gruppi in *jpg* con parametri molto simili tra loro, come a dimostrare che il trattamento *raw-converter* - per questo formato - sia indipendente dall'algoritmo di conversione impiegato.

Il peggioramento dell'indice di qualità è stato anche rimarcato dai primi passaggi con i software *3D Image Based Reconstruction*, in cui è stata evidente una riduzione di circa il 30% delle immagini adoperate per il posizionamento esterno delle camere. Il maggiore numero di fotografie non allineate, dall'osservazione attenta dei modelli, è stata provocata sia dal minor numero di *keypoint* riconosciuti dai software, sia dalla maggiore imprecisione nel posizionamento dei centri dei target; caratteristica che, come verrà di seguito sottolineato, sarà cagione di una minor accuratezza del modello. Infatti, lo scostamento tra le coordinate dei riferimenti attribuiti dal software (per la minore nitidezza e una maggiore sfocatura dei bordi) e quelle forzate dall'assegnazione diretta dei valori misurati con la rete topografica, è all'origine di un maggior errore di riproiezione RMS.

La sperimentazione si è basato sull'utilizzo dell'ultima versione del software 3DFlow® Zephyr per la creazione dei modelli; esso infatti è uno dei pochi applicativi commerciali in grado di poter importare i dati raw. I test hanno potuto sfruttare le possibilità offerte dal nuovo motore Windows Imaging Component (WIC) Raw Format in grado di supportare la maggior parte dei formati grezzi dei maggiori produttori di apparecchiature fotografiche. In particolare il software impiega le potenzialità del codec pack sviluppato da Microsoft® nell'ambito del progetto chiamato LIBRAW, una libreria software gratuita e open source (libraw.org) in grado di processare i file grezzi di molte case produttrici mantenendo i metadati, la profondità bit-data (sino a 32 bit per canale) e la high dynamic range del sensore.

La scelta di un unico applicativo penalizza la generalizzazione del problema in quanto i risultati possono essere condizionati dalla struttura del programma; consci di questa limitazione gli autori stanno conducendo le stesse prove con più software (al fine di poter avanzare delle ipotesi generali non dipendente dagli applicativi impiegati) ma questo, al momento, esula dalla trattazione del saggio perché i riscontri sono ancora acerbi e meritano di un ulteriore approfondimento. In linea di massima, comunque, sembrano confermare le ipotesi ad oggi avanzate.





Tab 2. Parametri imposti per ogni formato nel processamento del set di immagini.

La creazione dei modelli fotogrammetrici ha seguito il consueto workflow contraddistinto, innanzitutto, da una scelta univoca dei parametri per i diversi step di elaborazione per tutte le ricostruzioni. Sono stati impostati dei valori standard per le variabili che hanno regolato la creazione degli artefatti, evitando estremizzazioni e situazioni particolari che avrebbero potuto alterare la generalità dei risultati favorendo uno specifico formato rispetto ad un altro (fig. 3). La prima fase ha visto il riconoscimento, da parte del software, di punti particolari nelle fotografie e quindi il successivo abbinamento di essi tra i fotogrammi omologhi; sulla base di queste indicazioni, 3DF Zephyr ha continuato operando il posizionamento spaziale delle camere e la conseguente generazione della nuvola sparsa. I passi successivi hanno riguardato la costruzione della nuvola densa, la conversione del modello discontinuo a punti in quello continuo mesh e, infine, la riproiezione sulle facce dei singoli triangoli delle immagini RGB (dopo un accurato blending per garantire uniformità di esposizione e colore). La georeferenziazione del modello e la sua 'messa in scala' è stata effettuata sulla base delle coordinate spaziali dei Ground Control Points (GCP) che hanno anche assolto al ruolo di Quality Control Points (QCP); come nelle procedure che hanno riguardato i dati acquisiti dai sensori passivi, infatti, alcuni di essi sono stati impiegati per il confronto delle diversità tra le coordinate dei modelli e quelli della rete topografica (tab. 2).



Fig 4. Confronto tra i modelli: la nuvola di punti sparsa, quella densa e il modello mesh (da sinistra verso destra); ricostrzione dal set di immagini *JPG Raw-Converter*, *TIF Raw-Converter*, *DNG Raw-Converter*, *JPG Image Sensor e RAW Image Sensor* (dall'alto verso il basso).



Tab 3. Confronto delle accuratezze tra i modelli in finzione dell'indice RMS.

Al fine dell'ottenimento di un miglior risultato invece di procedere con il processamento dell'intero set di immagini all'interno di un unico ambiente – più gravoso in termini di tempo e risorse hardware – si è optato per delle lavorazioni parziali delle immagini per poi riunire i modelli. In particolare sono stati ricostruiti separatamente i due fronti mantenendo come elementi comuni di giunzione le aree prossime allo spigolo e la copertura. Questo artificio ha inoltre permesso di ottenere un maggior numero di punti (sia della nuvola sparsa che di quella densa) nonché delle *mesh* più ridotte e più 'interpretative' della reale geometria dell'edificio, grazie alla doppia elaborazione dei fotogrammi comuni (fig. 4).

Il confronto tra i modelli mette in risalto, con assoluta chiarezza, come le ricostruzioni migliori sono quelle ottenute dai dati grezzi, sia in termini di accuratezza dei modelli che nella percentuale di parti ricostruite. La ricostruzione della copertura, da scatti acquisiti in condizioni non ottimali con la camera inclinata e più distante, è garantita dalle informazioni del sensore (*raw* e *jpg*) e non possibile con le immagini RGB dopo i trattamenti *raw-converter*.

Importante sottolineare i risultati ottenuti dal formato *JPG Image Sensor* perché frutto di immagini con risoluzione molto più bassa e forte rapporto di compressione. Questo è infatti alla base dei minori tempi di elaborazione del set e del minor numero di punti e di triangoli costituenti nuvole e *mesh*. Un modello che seppur meno accurato (non deve indurre il errore la tabella 3 perchè comparata con set di fotografie di diversa risoluzione) ha comunque ricostruito nell'interezza il modello, comprensivo del tetto.

4 Conclusioni

La fotogrammetria *3D Image Based Reconstruction* si fonda sull'utilizzo delle foto digitali ma molti software non permettono l'elaborazione diretta del file grezzo ma solo di immagini RGB.

La prima considerazione è la convenienza dell'utilizzo del dato grezzo per la creazione di artefatti virtuali basati su catture fotografiche, perché in grado di coniugare un'elevata accuratezza metrica a tempi di processamento comunque contenuti. Il mancato sviluppo digitale degli scatti, seppur consentendo di ridurre 'tempo e lavoro', è vincolo per una maggiore attenzione in fase di cattura della scena; infatti non è possibile la correzione di eventuali errori 'di presa' perchè, ad oggi, non è possibile apportare modifiche ai *files raw* all'interno dei software.

La trasformazione esterna del dato grezzo (non operata dal processore della camera e/o con software proprietari) ha invece evidenziato una generale 'perdita' di qualità dell'informazione; gli stessi dati *jpg* grezzi salvati direttamente dalla camera consentono la creazione di modelli molto più accurati che non quelli ottenuti a seguito delle trasformazioni in *raw-converter* e salvati nei formati *dng*, *tif* e *jpg*.

Riferimenti bibliografici

- 1. AAVV.: LibRaw raw images decoder library: project Goals and Objectives. Available online at: http://www.libraw.org.
- Alfio, V. S., Costantino, D., & Pepe, M.: Influence of Image TIFF Format and JPEG Compression Level in the Accuracy of the 3D Model and Quality of the Orthophoto in UAV Photogrammetry. Journal of Imaging, 6(5), 30 (2020).
- 3. Cardaci A., Versaci A.: Rilievo e restauro: un binomio imprescindibile: approcci metodologici e applicazioni operative finalizzate alla conoscenza e alla conservazione del patrimonio culturale. Aracne, Roma (2018).
- 4. Carpiceci M.: Fotografia digitale e architettura; elaborazioni con le odierne attrezzature fotografiche e informatiche. Aracne, Roma (2012).
- Cox R.: Real-world comparisons between target-based and targetless point-cloud registration in Faro Scene, Trimble RealWorks and Autodesk Recap. (2015). Available online at: https://eprints.usq.edu.au/29195/
- Fraser B., Understanding Digital Raw Capture. Adobe Systems Incorporated, San Jose (2004).
 Online at:
- https://www.adobe.com/digitalimag/pdfs/understanding_digitalrawcapture.pdf
- Liscio E., Hayden A., Moody J.: A comparison of the terrestrial laser scanner & total station for scene documentation. In Journal of Association for Crime Scene Reconstruction, 20 (2016): 1-8.
- 8. McHugh S. T.: Understanding Photography: master your digital camera and capture that perfect photo. No Starch Press, San Francisco (2018).
- Nguyen, R., Prasad, D. K., Brown, M. S.: Raw-to-raw: Mapping between image sensor color responses. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2014) (pp. 3398-3405).
- Pirotti F., Vettore A., Guarnieri A.: Algoritmi di classificazione contextual con GRASS. In Geomatics Workbooks, n.6 (2016). Online at: http://www.geolab.polimi.it/volume-6/

- Rojtberg, P.: Processing RAW images in Python, (2017). Available online at: https://www.researchgate.net/profile/P_Rojtberg/publication/314239357_Processi ng_RAW_images_in_Python/links/58bd3abda6fdcc2d14e64497/Processing-RAW-images-in-Python.pdf
- 12. Verhoeven, G. J. J.: It's all about the format–unleashing the power of RAW aerial photography. In International Journal of Remote Sensing, 31(8), 2009-2042, (2010).