



UNIVERSITÀ DELLA CALABRIA  
Dipartimento di Strutture

Antonio Lio

*Antonio Lio*

**FOTOGRAMMETRIA E FOTOMODELLAZIONE  
NEL RILIEVO DELL'ARCHITETTURA**

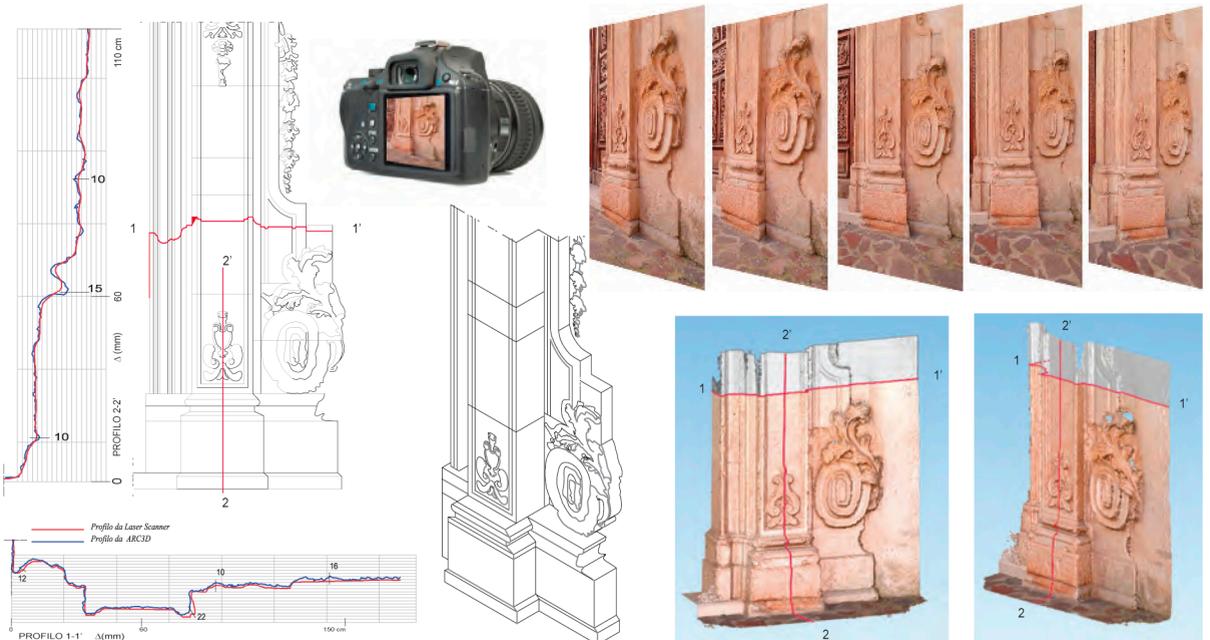
*Tesi di Dottorato in Ingegneria dei Materiali e delle Strutture*

*sezione: Metodiche dell'Architettura e della Costruzione*

*Scuola Pitagora in Scienze Ingegneristiche*

*Ciclo XXV (2009-2012)*

**S.S.D. ICAR 17-Disegno (settore concorsuale 08/E1)**



Coordinatore: *prof. Renato S. Olivito*

Tutor: *prof. Aldo De Sanctis*

*Renato S. Olivito*



UNIVERSITÀ DELLA CALABRIA  
Dipartimento di Strutture

Antonio Lio

**FOTOGRAMMETRIA E FOTOMODELLAZIONE NEL RILIEVO  
DELL'ARCHITETTURA**

Tesi di Dottorato in Ingegneria dei Materiali e delle Strutture  
**sezione: Metodiche dell'Architettura e della Costruzione**

Scuola Pitagora in Scienze Ingegneristiche

Ciclo XXV (2009-2012)

S.S.D. ICAR 17-Disegno (08/E1)

*Coordinatore: prof. Renato S. Olivito*

*Tutor: prof. Aldo De Sanctis*



<i>Premessa</i>	1
1 FOTOGRAFIE E ARCHITETTURA	
1.1 Origini della fotografia	3
1.2 La fotografia come “mezzo” per documentare/indagare un’architettura	9
2 IL MODELLO DIGITALE 3D	
2.1 Riflessioni sul modello 3D in architettura	15
2.2 Confronto tra sensori di misura attivi e passivi	23
3 LA FOTOMODELLAZIONE	
3.1 Introduzione	31
3.2 La ripresa fotografica	35
Qualità della fotografia: illuminazione, nitidezza, contrasto	35
3.3 Tecniche di ripresa fotografica	43
Ripresa ad assi convergenti	45
Ripresa ad assi paralleli	48
Ripresa panoramica, cilindrica e sferica	49
3.4 Acquisizione delle coordinate 3D da foto	51
La geometria interna della camera: i parametri intrinseci	52
Individuazione dei punti omologhi: accoppiamento 2D/2D	54
Distribuzione dei punti sull’immagine e nello spazio	55
3.5 I modi per calibrare ed orientare le camere	57
Calibrazione e orientamento progressivo di immagini convergenti	58
Calibrazione ed orientamento di immagini parallele su immagini convergenti	60
4 LA FOTOGRAMMETRIA	
4.1 Introduzione	63
4.2 <i>Close range photogrammetry</i> : fotogrammetria	64
Parametri di orientamento interno	66
Parametri di orientamento esterno	67

Operazione di resezione	68
<i>Matching</i>	69
4.3 Parametri di distorsione affine, radiale e tangenziale	69
Distorsione provocata dalle lenti dell'obiettivo	70
Problema della focalizzazione dei raggi luminosi	70
Aberrazioni	71
Aberrazioni cromatiche	71
Astigmatismo	73
Curvatura di campo	73
Distorsione Radiale	73
Distorsione tangenziale	74
Distorsione affine	75
4.4 Caratteristiche delle immagini digitali	76
Il sensore	77
Risoluzione radiometrica e geometrica	80
Formazione di immagini digitali a colori	81
4.5 <i>Structure and motion</i> . Ricostruzione 3D in automatico da fotografie	83
ARC3D	87
MeshLab	92
5 LE ESPERIENZE CONDOTTE	
5.1 Introduzione	97
5.2 Le esperienze condotte in laboratorio	99
5.3 Le esperienze condotte all'esterno	115
Il portale della chiesa di Santa Maria delle Vergini (Cosenza)	115
I mosaici della chiesa di Sant'Adriano (San Demetrio Corone, Cosenza)	130
<i>Conclusioni</i>	139
<i>Nota bibliografica</i>	145

## *Premessa*

Alla straordinaria ricchezza di complessi architettonici e urbani presenti nel nostro paese, spesso si contrappone purtroppo, una generalizzata scarsità di manutenzione continua e l'assenza di un'efficace politica volta alla conservazione e valorizzazione del patrimonio culturale; a questi aspetti, di ordine più generale, si aggiunge poi l'assenza di protocolli operativi condivisi e di direttive in grado di coordinare la raccolta di informazioni sullo stato di conservazione delle opere, che capillarmente mettano gli studiosi in grado di operare sui manufatti, in situazioni di "normalità", così come di rischio, o di emergenza.

Tale stato delle cose, imputabile evidentemente a diversi fattori, si è determinato, in larga parte, per l'inadeguatezza, o l'indisponibilità di un'opportuna banca-dati capace di fornire, in ogni momento, le informazioni necessarie, in relazione agli interventi da effettuare.

Anche per i complessi più famosi non esistono che documentazioni limitate, sulle caratteristiche più rilevanti (geometriche, strutturali, etc.) e in più di un'occasione il loro parziale danneggiamento ha, in buona sostanza, coinciso con guasti maggiori e costosi interventi di restauro.

A questo tipo di problemi, oggi, possono contribuire utilmente le nuove tecnologie di acquisizione e "manipolazione" dei dati (dalle tecniche di *Reality-based modeling*, all'uso di *laser scanner 3D*, alle tecniche di *Image-based modeling*), così come lo sviluppo dei nuovi *software* dedicati, che consentono di strutturare, unitariamente, informazioni di diversa natura (data-base), ed essere richiamati e selezionati, secondo le varie esigenze.

Il presente lavoro di ricerca ha affrontato alcuni di questi aspetti, studiando le nuove possibilità offerte dalle moderne tecniche di acquisizione e restituzione architettonica, per procedere alla definizione di protocolli operativi innovativi e di nuove metodologie per l'integrazione dei vari dati, con particolare riferimento al settore della documentazione, conservazione e divulgazione dei beni architettonici.

Per valutare criticamente l'impatto di queste tecniche sul rilevamento architettonico, e come si sono, "trasformate" le operazioni inerenti la disciplina del Rilievo Architettonico, si è eseguita una ricognizione ad ampio raggio delle esperienze fatte sinora, sia in campo nazionale che internazionale, al fine di conoscere lo stato dell'arte del settore e, più in generale, le possibilità delle nuove tecniche, viste come "fattori attivi" sia per la conoscenza, che per la conservazione e valorizzazione dei Beni Culturali. Si sono, di conseguenza, affrontate le problematiche indotte dall'utilizzo delle tecniche di acquisizione della forma (*image-based modeling e 3D photography*) e l'integrazione tra loro; in particolare, si sono studiate ed analizzate sia le tecniche ottiche di misura 3D attive (cioè laser scanner 3D), che le tecniche ottiche di misura 3D passive (fotomodellazione, fotogrammetrie e tecniche di acquisizione *multi-view stereo*): si è approfondita la conoscenza, studiandone vantaggi, limiti, principi di funzionamento, capacità operative e possibilità d'integrazione con le altre tecnologie, al fine di avere elementi certi per la scelta della tecnica più appropriata e per definire metodi e procedimenti adeguati agli scopi prefissati.

All'interno delle tecniche ottiche di misura 3D passive, si sono analizzate le differenti procedure di acquisizione tramite due casi-studio ed una serie di sperimentazioni di laboratorio: l'insieme delle sperimentazioni ci ha permesso di confrontare i risultati ottenuti con quelli del *laser scanner* e di confrontare e valutare la precisione ottenibile, in funzione dei diversi tipi di applicazione, degli scopi e della loro possibile integrazione.

*Parole chiave:* rilevamento architettonico, tecniche ottiche di misura 3D, modello digitale tridimensionale, fotomodellazione, fotogrammetria computer vision.

### 1.1 *Origini della fotografia*

Il francese Joseph Nicéphore Niépce interessato alla nuova arte incisoria, studia come eseguire incisioni automaticamente, senza il contributo dell'incisore che deve ricopiare l'originale; da tale studio ne scaturisce l'idea di utilizzare la camera oscura per riprodurre gli oggetti e nel 1826 riesce ad ottenere la prima immagine disegnata dalla luce, impiegando una lamina di rame e argento ricoperta da uno strato di bitume. Il materiale "fotosensibile" fu esposto alla luce, per diverse ore, sul fondo di una camera oscura. La lamina dopo l'esposizione fu immersa in un bagno di oli di lavanda, per dissolvere i frammenti di bitume, che non erano stati colpiti dalla luce. Ne venne fuori un'immagine in negativo, di ciò che la camera oscura aveva inquadrato, ossia il panorama visto dalla finestra della sua casa di campagna. L'inconveniente di tale risultato fu che l'immagine non si era fissata e si anneriva progressivamente al contatto con la luce.

Le sperimentazioni di Niépce, furono riprese e portate avanti dal figlio Isidore Niépce e da Louis Dagherre, che nel 1838, dopo diversi anni di sperimentazione, svilupparono un metodo più conveniente ed efficace per ottenere immagini con la luce. Il risultato prese il nome dello stesso inventore: "dagherrotipo".

Modificando la superficie di metallo della lastra – in parte riflettente a

specchio e in parte opaca –, per riprodurre i toni nelle zone scure ed avere immagini nitide, Louis Dagherre riuscì inoltre ad ottenere un'immagine latente (cosa che non riuscì a Niépce), riducendo i tempi di posa da 8 ore della prima foto, a 30 minuti, rendendo l'immagine visibile in seguito allo sviluppo effettuato con vapori di mercurio.



Fig. 1.1 – *Vista dalla finestra a Gras*. La prima foto di Joseph Nicephore Niépce.



Fig. 1.2 – *La prima foto realizzata da Louis Daguerre*. Si tratta di un'immagine del 1838 di un boulevard parigino e nonostante la via fosse trafficata, l'esposizione della lastra è stata di dieci minuti per cui il traffico cittadino è "scomparso" nella foto. Sono rimasti solo il tizio che si fa pulire le scarpe e il lustrascarpe.

Pochi anni dopo fu l'inglese William Henry Talbot a porre le basi della fotografia chimica, così come la intendiamo oggi, cioè quel procedimento che tramite un negativo permette di ottenere una o più stampe positive su carta:

Talbot inventa il procedimento negativo-positivo per la riproduzione delle fotografie dove sia il negativo, che la stampa positiva sono costituiti da una carta impregnata di cloruro d'argento (ioduro d'argento). Fondamentale è la scoperta che il sale d'argento, non alterato dall'azione della luce, può essere sciolto in diverse soluzioni.

Con la carta ai sali d'argento di Talbot, l'immagine s'impresiona in negativo, basta poi rifotografare il negativo di carta per invertire l'immagine, traducendola in positivo.

La diffusione e il sostegno scientifico della fotografia, sono da riconoscere ai francesi François Arago e Gay-Lussac, che nel 1839, in una relazione presentata alla Camera dei Deputati e All'accademia delle Scienze Francesi dichiarano: “... *le immagini fotografiche, essendo soggette nella loro formazione alle regole della geometria, permettono, con l'aiuto di un piccolo numero di dati, di risalire alle dimensioni delle parti le più elevate, le più inaccessibili degli edifici...*”.<sup>1</sup>



Fig. 1.3 – *Foro Romano*. Foto di A. R. Bisson, 1864.

---

<sup>1</sup>In C. CUNDARI, *Fotogrammetria architettonica*, Napoli Edizione Kappa, 1983, p. 159.



Fig. 1.4 – *L'Eretteo di Atene*. Foto di W.J. Stillman, 1868.

La nascita della fotografia d'architettura in Europa è rivolta soprattutto alla documentazione delle antiche rovine, ai resti dei monumenti delle civiltà passate e al loro coesistere con le città moderne. Di solito le immagini fotografiche dei resti di antichi monumenti, escludono la presenza umana, non per volontà del fotografo, ma per i tempi lunghi di esposizione, necessari per avere l'inquadratura voluta, e per la complessità e dimensioni delle attrezzature fotografiche.

In Italia i primi che si occupano di fotografia e comunicazione architettonica, sono i fratelli Alinari, con la loro azienda fondata a Firenze nel

1852. La nascita della fotografia d'architettura e la storia della loro azienda, sono legate da un percorso comune di evoluzione e crescita. I fratelli Alinari, con lo studio e la pratica, mettono a punto criteri per le riprese dell'architettura, che possono essere sintetizzati in:

- uso degli schemi della prospettiva rinascimentale nella composizione dell'inquadratura;
- isolamento dell'edificio dal contesto;
- il controllo dell'ortogonalità tra la linea d'orizzonte e le linee verticali, per evitare di veder "linee cadenti";
- utilizzazione di una luce generica, diffusa che non generi ombre profonde;
- eliminazione di ogni elemento distraente;
- esecuzione delle riprese da circa tre metri da terra.

Quest'ultima caratteristica, la ripresa da un punto più alto rispetto alla quota "zero", consentiva spesso di riuscire a mantenere la macchina "in bolla" e inquadrare tutto l'edificio. Nella produzione degli Alinari s'incontrano sovente riprese che "sacrificano" un pezzo di architettura, pur di non inclinare verso l'alto la fotocamera.

L'immagine che il fotografo vuole registrare sul supporto fotosensibile, è il "disegno" che la luce compie del mondo esterno. Mondo di cui il fotografo ha una conoscenza che è limitata alla sua capacità di osservare e di ricordare sensazioni analoghe, elaborate dalla cultura che possiede<sup>2</sup>. Il fotografo non fa altro che cercare di fissare con la fotografia, quello che in quel momento lo colpisce e gli provoca un'emozione. Roland Barthes osserva "*che una foto può essere l'oggetto di tre pratiche (o tre emozioni, o tre intenzioni)*"<sup>3</sup>, quella del fare "operator", quella del subire, "referente" e quella del guardare "spectator". Allo stesso tempo, l'immagine fotografica è costituita dalla cooperazione tra la realtà fisica, (l'oggetto della ripresa), e la capacità del fotografo di selezionare, comporre e organizzare il materiale in qualcosa di significativo. Alfred Stieglitz considera la fotografia come il risultato di una serie di operazioni tecniche che hanno lo scopo di comunicare "*l'equivalente di*

---

<sup>2</sup> R. L. GREGORY, *Occhio e cervello. La psicologia del vedere*, trad. it. Alberta Rebaglia, Milano, Raffaello Cortina, 1998, (ed. originale *Eye and Brain*, 1998).

<sup>3</sup> R. BARTHES, *La camera chiara. Nota sulla fotografia*, Torino Piccola Biblioteca Einaudi, 2003, p. 6.

*quello che il fotografo ha visto e sentito*"<sup>4</sup>, in cui il sentire è da intendere nel senso di emozioni.



Fig. 1.5 – *Campanile di Giotto a Firenze*. Foto ottenuta mediante il decentramento dell'obiettivo, inventato da Leopoldo Alinari. Fotografia dei fratelli Alinari, 1854.

---

<sup>4</sup> M. DOCCI, D. MAESTRI, *Manuale di rilevamento architettonico e urbano*, Roma-Bari Editori Laterza, 2009, p. 242.

## 1.2 La fotografia come “mezzo” per documentare/indagare un’architettura

Il riconoscimento di una forma, secondo la teoria percettiva dello “strutturalismo”, deriva da sensazioni di luce, ombre e toni, immagini e ricordi, che sono già stati associati in memoria. L’esperienza visiva non richiama solo la memoria visiva, ma anche quella tattile, uditiva, olfattiva e gustosa (ad esempio la vista di una mela rossa), da questo semplice esempio, si può dedurre che la percezione è programmata per avvenire per mezzo di processi associativi, rapidi e inconsci, che vanno oltre la memoria acquisita attraverso le precedenti esperienze.



Fig 1.6 - Sant'Ivo alla Sapienza, interno della cupola. Immagine volta a far comprendere la dinamica compositiva. Fotografia di Oscar Savio, 1967 (M. DOCCI, 2009).

La rappresentazione di una realtà architettonica, è un’operazione complessa, che richiede competenze critiche e tecniche differenti, e una conoscenza appropriata della stessa. La conoscenza di un oggetto architettonico è il frutto

di una vasta esperienza, che include osservazioni visive, sensazioni tattili, misure compiute in vario modo (quali quelle derivanti dalla comparazione con oggetti, quelle compiute muovendosi in relazione all'oggetto stesso), elaborazione di documenti storici e tante altre. La conoscenza si completa e prende forma, rappresentando l'oggetto architettonico, quando le percezioni sensoriali entrano in relazione con le elaborazioni concettuali di tipo storico-critico e tecnico, che sono già presenti nella nostra memoria. In tali circostanze, il potere della fotografia permette di immagazzinare l'immagine nella nostra memoria, in modo più stabile, più efficace e più forte delle altre impressioni sensoriali. Tale peculiarità rende la fotografia *“uno strumento insostituibile nel rilevamento architettonico e urbano vuoi come documentazione ai grafici di rilievo, vuoi come strumento ausiliario nelle operazioni di rilievo”*<sup>5</sup>; e ancora come strumento sul quale indagare l'architettura eseguendo rappresentazioni fotografiche che contengono elementi in grado di far comparire le proporzioni tra le parti, gli allineamenti, gli schemi compositivi ecc., l'impiego della fotografia come documentazione, acquista significato quando sarà maggiore il numero di fotografie che trasmettono informazioni visive, capaci di esprimere determinati significati, senza la preoccupazione di avere informazioni sovrabbondanti, in maniera analoga a quello che avviene in fase di creazione degli eidotipi, è preferibile avere delle informazioni sovrabbondanti, per evitare di dover riandare sul posto per acquisire i dati mancanti, in fase di restituzione.

Le fotografie affinché siano in grado di descrivere un'architettura, o essere usate come documentazione e appoggio per un rilievo, devono rispondere ad alcuni requisiti tecnici fondamentali, e le operazioni tecniche per le quali sono richieste una particolare attenzione sono: il controllo dell'inquadratura e la determinazione di una corretta esposizione. L'inquadratura, come dice la stessa parola, creare un quadro che rappresenta una posizione della realtà o della scena, quindi è necessario che in tale quadro vi siano contenute tutte le informazioni percepite dall'occhio e dalla mente, e che comprenda particolari aspetti, e metta in evidenza tratti caratteristici dell'opera architettonica indagata, come le regole o il dinamismo compositivo, le forme contenute dentro l'architettura, la struttura, lo stato di conservazione ecc., ogni elemento

---

<sup>5</sup> M. DOCCI, D. MAESTRI, *Manuale di rilevamento architettonico e urbano*, Roma-Bari Editori Laterza, 2009, p. 233.

d'interesse dovrebbe essere ripreso frontalmente, in prospettiva e di scorcio, affinché possano essere documentati anche gli elementi presenti nei sottoquadri, nascosti dagli elementi aggettanti. Le fotografie, dovrebbero essere prive d'inganni ottici, finzioni e ambiguità, che possono indurre l'osservatore in errori d'interpretazione, avere un elevato rigore geometrico, che è necessario per una giusta interpretazione dei rapporti dimensionali. L'esposizione deve consentire, di poter leggere la trama e ogni dettaglio esistente, sia nelle parti luminose che in quelle in ombra, e l'equilibrio tonale fra le diverse zone; sono queste a favorire la lettura delle diverse parti, che compongono l'architettura rappresentata. Quando si riprendono particolari architettonici molto complessi, come capitelli, cornicioni, bassorilievi ecc., le fotografie devono essere fatte con luce soffusa, per evitare che ombre troppo accentuate, vadano a ricoprire parte dell'oggetto. La condizione di luce favorevole per fare le riprese, è quella diurna, con il cielo leggermente velato e con il sole alto; se una simile condizione non è possibile, – come può accadere per le riprese di un interno, dove la condizione di luce naturale è debole –, ci si può aiutare con delle lampade, cercando di ripristinare al meglio la condizione di luce naturale, e facendo provenire la luce sempre dalla stessa direzione e in modo riflesso, non direttamente sull'oggetto.

L'impiego della fotografia, per il rilevamento, la documentazione e la rappresentazione dell'architettura, scaturisce dal principio che la fotografia, altro non è che una proiezione centrale, simile alla prospettiva<sup>6</sup>, se si trascurano alcune caratteristiche:

- Le aberrazioni indotte sul piano di quadro dall'obiettivo. L'aberrazione ottica, quella cromatica, la curvatura di campo, la distorsione, ecc. sono difetti, che scaturiscono dalla composizione del cristallo ottico, dal sistema di molatura delle lenti e dal sistema di montaggio delle stesse nell'obiettivo. Alcuni difetti possono essere corretti e contenuti entro limiti accettabili, in determinati obiettivi di qualità.
- In geometria, il centro di proiezione è un punto privo di dimensioni, mentre

---

<sup>6</sup> La prospettiva è il metodo di rappresentazione, che mette in relazione gli elementi disegnati con la posizione soggettiva dell'osservatore, si traduce in un processo di proiezioni, da un punto finito che è l'osservatore, di un oggetto schematizzato in elementi geometrici, su un piano, detto piano di posa. Tale modo di rappresentare è quello che più si avvicina alla visione umana; nella fotografia, il centro di proiezione è interposto tra l'oggetto ed il quadro.

in fotografia il centro di proiezione (il centro focale del sistema ottico), è un'area circolare (disegnato dai raggi di luce, che investono l'apertura del diaframma dell'obiettivo), che pur essendo piccola, non è mai un punto adimensionale; questo comporta un problema di precisione e tolleranza delle misure, che non c'è in geometria.

- Gli enti geometrici elementari utili in geometria per schematizzare un oggetto, (ad esempio uno spigolo, che in geometria si rappresenta con una retta, individuabile da una coppia di coordinate), nella fotografia non hanno lo stesso valore, ed uno spigolo, non è più individuabile con una linea, ma si determina come confine o segno di discontinuità o disomogeneità tra due superfici a differente luminosità.

Tornando alle operazioni di rilevamento, la fotografia della facciata di un edificio, o di particolari di essa, ripresa con una prospettiva ad asse verticale, – con il piano di posa dell'elemento fotosensibile (lastra in passato, pellicola o sensore digitale) parallelo alla facciata dell'edificio –, (può essere utile) sia per una lettura delle proporzioni, che per ricavarne o verificare le misure, se è nota almeno una coppia di misure dell'oggetto reale; misure che servono per proporzionare (scalare) la fotografia, e controllarne il risultato.

Le immagini fotografiche, ottenute tramite fotocamere, impiegate per documentare e integrare le operazioni di rilevamento dell'architettura, devono avere delle distorsioni contenute entro precisi limiti, limiti che introducono vincoli che portano a scegliere tra i diversi tipi di attrezzature (camere metriche, semi metriche e camere digitali), esistenti sul mercato. Fino a pochi anni fa, le fotocamere più impiegate per le riprese d'architettura erano le camere metriche o semi metriche, ovvero fotocamere con tolleranze ammissibili, sia ottiche che meccaniche, contenute entro limiti ridotti. Per questo tipo di fotocamere, le case costruttrici forniscono un certificato di taratura, o certificato di calibrazione, con l'esatta lunghezza della distanza principale "f", cioè la distanza tra il centro di proiezione e il piano dell'immagine, che in tali tipi di fotocamere è fissa. Una camera metrica è composta di un corpo macchina montato su un supporto girevole tipo alidada, e sostenuta da un cavalletto; differisce da una comune macchina fotografica per alcune caratteristiche costruttive: dall'obiettivo, che è esente da distorsioni; l'asse ottico è perfettamente perpendicolare al piano della pellicola, e in

posizione nota rispetto al centro del fotogramma; la pellicola è piana, (in alcuni modelli di camera è presente anche un sistema di spianamento della pellicola). Oggi, con l'impiego di camere digitali, il processo di calibrazione, viene eseguito tramite *software*; tale processo di caratterizzazione della camera digitale, serve per determinare non solo la distanza principale "f", che in questo caso non è più fissa, e la posizione del punto principale, ma anche i fattori correttivi, in grado di correggere le distorsioni indotte dall'obiettivo.

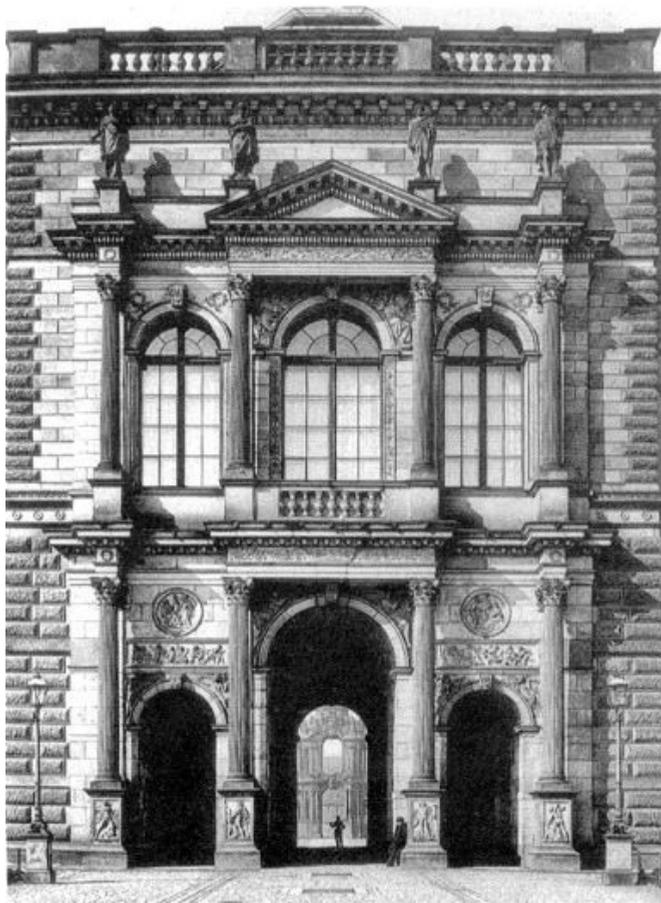


Fig. 1.7 – Ingresso alla Royal art Gallery di Dresda. La scelta del punto di vista centrale rispetto al portale, il controllo della prospettiva mediante il decentramento dell'obiettivo, e la trasparenza delle ombre nella facciata, consentono una restituzioni metriche accurate, la lettura degli allineamenti, lo stato di conservazione ecc. Fotografia dello studio fotografico Roemmler & Jonas, 1893 (M. DOCCI, 2009).



## CAPITOLO 2

## IL MODELLO DIGITALE 3D

2.1 *Riflessioni sul modello 3D in architettura*

Il modello digitale 3D di un oggetto, oggi, non è più solo la ricostruzione teorica, o pseudo-teorica, di un monumento, o di un artefatto architettonico, ma è la trascrizione, “fedele” della realtà, con tutte le caratteristiche cromatiche, formali e dimensionali, che le nuove tecnologie consentono di acquisire, offrendo la possibilità concreta di visualizzare la tridimensionalità dell’opera, di valutare quantitativamente, quello che prima era considerato in modo qualitativo e di rendere visivo e continuo, ciò che prima era discretizzato. Ancora, il modello digitale 3D, non è più un “corpo” unico, osservabile solo da alcuni punti di vista, ma un oggetto composto di varie entità – ciascuna delle quali può essere rappresentata in modi differenti – dotate di attributi diversi, ricostruiti in un *data-base* d’informazioni, consultabile, modificabile e implementabile nel tempo; *data-base* d’informazioni che consente un facile accesso ai dati e di creare e riconoscere nuove relazioni tra gli stessi dati, dando la possibilità di investigare e sviluppare differenti tipi di analisi e letture. Tali caratteristiche rendono possibili modi di esplorazione e interrogazione del modello 3D, non pensabili con i sistemi di rappresentazione tradizionale ed in grado altresì di attivare nuove forme di conoscenza. Più in generale, si può dire che i modelli digitali 3D di un manufatto architettonico, permettono oggi

diversi tipi d'indagini finora quasi impensabili, come l'analisi spaziale, le verifiche di progetti e rilievi eseguiti in epoche passate<sup>7</sup>, il confronto fra realizzazione e disegni di progetto, ecc.

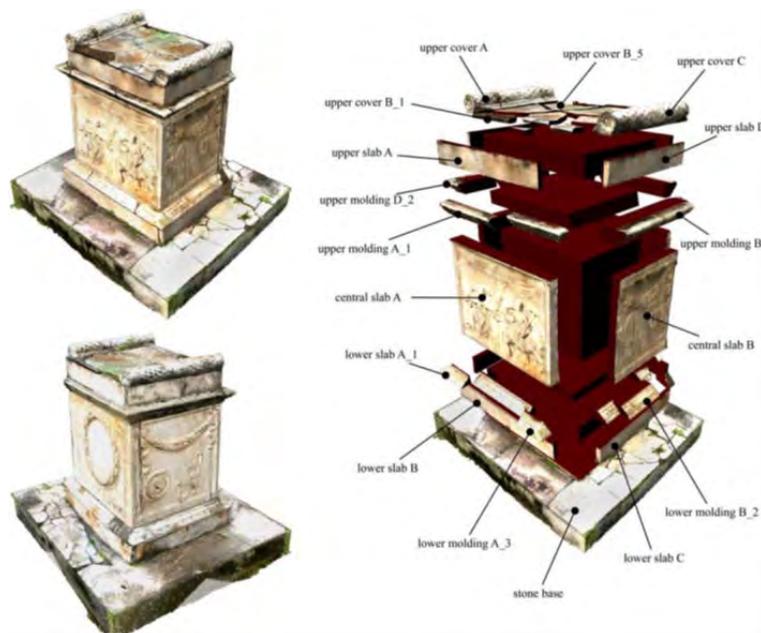


Fig.2.1 – Pompei Reperto altare con rilievo in marmo nel santuario del genio di Augusto (A. M. MANFARDINI et al., 2008).

Ripercorrere oggi con un rilievo, non solo la forma, a volte anche ricca e complessa, dei monumenti, può intendersi come una verifica del progetto originale, rispetto alla realizzazione, giunta fino a noi, o come una valutazione di quanto è stato modificato, aggiunto o sottratto nell'arco del tempo. A volte può succedere che il rilievo, è l'unica testimonianza dell'opera stessa.

<sup>7</sup>G. FORTUNATO, *The role of architectural representation for the analysis of the built. The 3d survey of San Pietro in Montorio*. Atti del convegno "X Congreso Internacional Expresión Gráfica aplicada a la Edificación Graphic Expression applied to Building", Apega (Spagna), 2,3,4 dicembre, 2010, Editorial Marfil, S. A.:Alicante (Spagna), 2010, Vol. I, pp. 451-461.

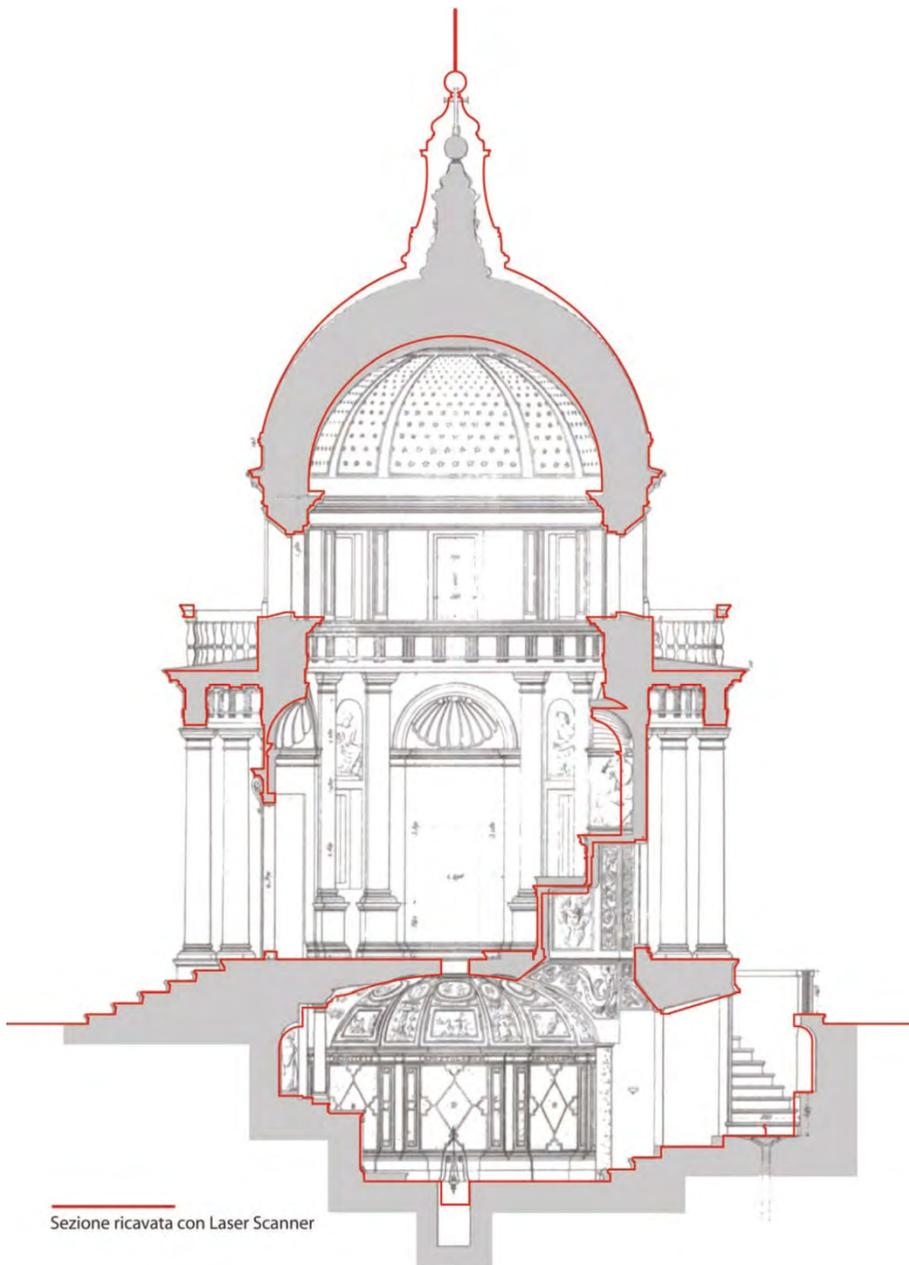


Fig. 2.2 –*Tempietto di san Pietro in Montorio Roma*. Confronto tra rilievo eseguito con *laser scanner* Leica HDS3000 ed il rilievo redatto da P. Letarouilly, pubblicato nel 1843, considerato come il principale riferimento per la conoscenza morfometrica della fabbrica. Il confronto ha riscontrato un importante scostamento tra la forma rilevata da Letarouilly e quella reale rilevata con il *laser scanner*. Tale scostamento è evidente soprattutto nel disegno del lanternino cieco, probabilmente per la difficoltà operativa da parte dell'autore di prendere misure poco accessibili (G. FORTUNATO, 2010).

Il rilievo che oggi è possibile elaborare, articola la conoscenza in un complesso d'indagini volte a individuare le qualità rilevanti (sotto l'aspetto morfologico, dimensionale, figurativo e tecnologico) di un manufatto edilizio, o di un insieme urbano, rendendo evidenti tutti i suoi "valori", da quelli geometrici a quelli dimensionali, da quelli figurativi a quelli strutturali, dai materiali impiegati, alle tecniche costruttive, dalle condizioni di degrado, alle relazioni con l'ambiente urbano. Il rilevamento del bene, diviene, cioè, l'unico modo per conoscere a fondo il monumento stesso, soprattutto in assenza di documenti di progetto, descrizioni attendibili, ecc.

La mancanza di convenzioni, standard e protocolli condivisi, riguardanti la raccolta e la presentazione di dati sul patrimonio architettonico, costituisce un problema legato alla compatibilità e al diverso tipo d'informazioni che si vogliono raccogliere. Nonostante le diverse esperienze compiute per individuare degli standard, nel campo dell'acquisizione digitale di oggetti 3D, gli esiti risultano ancora eterogenei tra loro, sia per quanto riguarda i dati acquisiti, che i modelli restituiti, rendendo quasi impossibile il loro riutilizzo in un qualsiasi conteso di studio. Le maggiori differenze si riscontrarono nell'utilizzo appropriato delle tecniche di acquisizione e degli strumenti di rappresentazione, le cui problematiche affrontate, nelle varie scuole o corsi di specializzazione<sup>8</sup>, conferenze e *workshop*, non consentono ancora di definire

---

<sup>8</sup> - Workshop di Fotogrammetria/Lidar "Metodi speditivi per il Rilievo, la Rappresentazione e la Comunicazione dell'Archeologia", organizzato da DICEA, Dipartimento di Ingegneria Civile, Edile, Architettura, Università Politecnica delle Marche di Ancona. Il lavoro si propone di mostrare come in sole due giornate di lavoro è possibile, attraverso le metodiche utilizzate, ottenere ottimi risultati in termini di dati e modelli geometrici, morfologici e comunicativi. (Riserva Naturale Statale "Gola del Furlo" Cagli (PU), 9 - 11 Novembre 2012).

- Summer School di "Rilievo e modellazione 3D", organizzato da Fondazione Bruno Kessler (FBK) e l'Università di Salerno. Il corso si pone come obiettivo l'analisi pratica e la valutazione critica del rilievo 3D "*reality-based*" finalizzato alla documentazione metrica, conservazione, restauro e valorizzazione del patrimonio culturale. (Paestum 3-9 giugno 2012).

- Scuola di "Fotogrammetria, Computer Vision e Scansione *laser*": -paradigmi e algoritmi- e -scansione laser e fotogrammetria per l'ingegneria, l'architettura e i beni culturali-, organizzato dal Centro Internazionale di Scienze Meccaniche (CISM) di Udine. Il corso articolato su due moduli didattici seriali, si proponeva di creare una sinergia fra i mondi della fotogrammetria e della computer vision alla luce delle potenzialità tecnologiche fornite dalla scansione laser, applicati ai settori dell'ingegneria, dell'architettura e i beni culturali (Udine 12 - 16 aprile 2010).

- Corso di specializzazione in "Rappresentazione e Modellazione" organizzato dalla Scuola Normale Superiore di Pisa. Il corso è stato incentrato sull'approfondimento (teorico e pratico)

linee guida, unitarie, in grado cioè, di fornire riferimenti utili da seguire nelle diverse metodologie di rilievo tridimensionale.

La penuria di tali riferimenti normativi, forse è da ricercare nella poca interazione tra le diverse discipline che si occupano di rilievo, rappresentazione e conservazione dei Beni Artistici e Culturali.

In diversi campi disciplinari, dal rilievo archeologico, al telerilevamento tramite foto aeree o satellitari, *dall'industrial design*, al rilievo topografico e infine al rilievo architettonico, anche se con finalità e scopi differenti, si utilizzano, gli stessi strumenti e metodi, di acquisizione per l'ottenimento di modelli tridimensionali. Considerando solo l'aspetto tecnico, dell'impiego e utilizzo di strumenti e tecniche, per ottenere un modello digitale 3D, da tecniche ottiche di misura 3D<sup>9</sup>, ricordiamo che nell'ultimo decennio si è assistito ad una rapida e vasta diffusione di sistemi, che arrivano anche all'automatizzazione dell'intera procedura.

L'intero *iter* procedurale, che porta dall'oggetto reale, al suo modello digitale, attraverso un rilievo tridimensionale, deve comprendere: metodi e strumenti adeguati al fine;

- una conoscenza adeguata delle caratteristiche operative degli strumenti impiegati, per ricavarne il massimo vantaggio;
- la valutazione dei limiti operativi di determinate applicazioni, per un'adeguata pianificazione del rilievo;
- capacità d'uso di *software* e algoritmi per l'elaborazione;
- l'analisi e il trattamento dei dati acquisiti. Tali aspetti influenzano direttamente sia la fase di acquisizione dei dati, che quella di restituzione.

---

di metodologie e tecnologie (laser scanner, fotogrammetria) per la rappresentazione e modellazione di manufatti e complessi monumentali; (Volterra (PI), 19-25 marzo 2009).

<sup>9</sup> Le tecniche ottiche di misura 3D, consentono di acquisire dati metrici e colorimetrici, della superficie di un oggetto, senza contatto fisico sfruttando le radiazioni luminose. Tali tecniche si possono dividere in tecniche ottiche di misura 3D attive, quei sistemi che emettono e registrano una radiazione luminosa codificata, in quanto costruiti da un emettitore di luce e da un sensore per registrare la risposta (sensori attivi), e tecniche ottiche di misura 3D passive, quei sistemi che utilizzano la radiazione luminosa (luce) presente nella scena (sensori passivi), indipendentemente dalla sorgente che genera la radiazione.

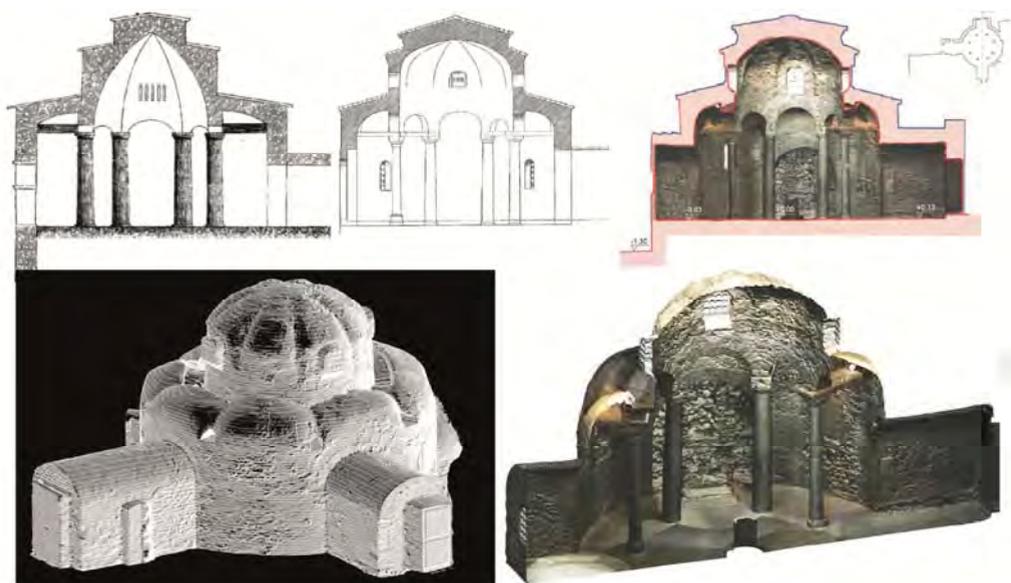


Fig. 2.3 – Battistero di Santa Severina (KR). Sopra, confronto fra rilievi precedenti con quello attuale ottenuto da scanner laser, si nota subito che essi sembrano quasi appartenere a soggetti simili, ma non alla stessa costruzione ed è chiaro che a geometrie diverse, le letture corrispondenti non possono che essere differenti; sotto una rappresentazione “per pieni” del modello con *slices* planari, e una vista sezionata della nuvola di punti. (G. FORTUNATO, 2009)

La creazione di modelli digitali 3D, oggi, può realizzarsi con diverse tecniche e metodologie, che si possono suddividere in due gruppi: metodi *reality-based*, che rilevano e restituiscono la realtà, impiegando sistemi che utilizzano sensori attivi e passivi, con l’ausilio di misure dirette e indirette dell’oggetto reale; metodi di *computergrafica*, che generano modelli digitali – con l’uso di *software*, come Autodesk 3D Studio, Cinema 4D, Google SketchUp, ecc. –, partendo da primitive grafiche (o schizzi, o da semplici elementi 3D), senza rilevare nessuna misura della realtà da restituire.

Impiegando i metodi *reality-based*, il modello digitale 3D di un monumento o di un artefatto qualsiasi, si ottiene con una successione di fasi, che inizia con la scelta di tecniche e strumenti per l’acquisizione di dati; prosegue poi con l’analisi e l’elaborazione delle informazioni acquisite e termina con la restituzione, in forma digitale, di una “copia” dell’opera in oggetto. Il modello digitale, che replica la realtà di un monumento attraverso il *monitor* di un *computer*, permette una simulazione visiva “realistica” dell’oggetto in esame; impiegando la realtà virtuale (*Virtual Reality*), si possono anche creare dei

sistemi di visualizzazione in “ambienti immersivi” o “semi immersivi” su grandi schermi, riducendo percettivamente la distanza tra l’utente e l’oggetto; in questi casi si altera il rapporto di scala, elemento caratterizzante delle rappresentazioni tradizionali. Tutte queste caratteristiche rendono il modello digitale, uno strumento appropriato per le ricostruzioni digitali, la diagnostica e la documentazione non invasiva, la pianificazione urbana, le animazioni, le repliche fisiche di oggetti, i restauri virtuali e operazioni di conservazione.

Per le applicazioni di restauro virtuale e le operazioni di conservazione, i modelli digitali, devono ottenersi con, la modellazione dello stato di fatto e la definizione delle superfici, mediante il rilievo di milioni di punti, così da ricostruire la “pelle” dell’opera; non servono, infatti ricostruzioni semplicemente congruenti dal punto di vista formale e dimensionale. Per l’ottenimento di simili superfici, dove la corrispondenza con il reale è forte, l’acquisizione dei dati in forme 3D, consente di limitare al massimo la perdita d’informazioni, che inevitabilmente si crea quando s’impiegano i metodi del rilievo tradizionale, con i quali, se ne ottiene, la costruzione di un modello semplificato, che è sostanzialmente, valido per la sola configurazione del volume dell’opera in esame.



Fig.2.4 – Sala delle Cariatidi di Palazzo Reale (Milano). Fotografia della Cariatide rovinata e viste ottenute dal modello poligonale mappato con *texture* fotorealistica per una ricostruzione virtuale dell’opera (F. BREVI et al., 2004).

Nel rilievo tradizionale, il rilevatore in una fase preliminare a quella della presa della misura, analizza il manufatto, razionalizzando le forme, per poterle misurare e rappresentare. Elabora un eidotipo, con il quale discretizza preventivamente l'opera, in funzione delle misure da prendere e del risultato che si vuole ottenere. Su questi primi disegni, che rappresentano una trascrizione sintetica dei caratteri formali dell'opera, viene predisposto il progetto di presa delle misure, ossia di quanto serve per la descrizione della morfologia e delle qualità metriche, per restituire e raffigurare l'opera. L'operazione di razionalizzare le forme, di rendere discreto ciò che è continuo, secondo un modo proprio (e quindi soggettivo) di vedere e percepire l'oggetto da analizzare (ogni uno di noi "vede" in base a propri schemi), in funzione di esperienze già affrontate, ed in funzione dei problemi che in quel momento vuole affrontare. Tutto ciò comporta una prefigurare del risultato, e i disegni che vengono eseguiti e le misure prese, risultano indirizzate ad uno specifico elaborato, e quindi non utilizzabili per altre analisi e altri rilievi. Con il rilievo tradizionale l'operatore utilizza un numero limitato di punti, e con questi ricostruisce degli elaborati riassuntivi della realtà dell'opera.

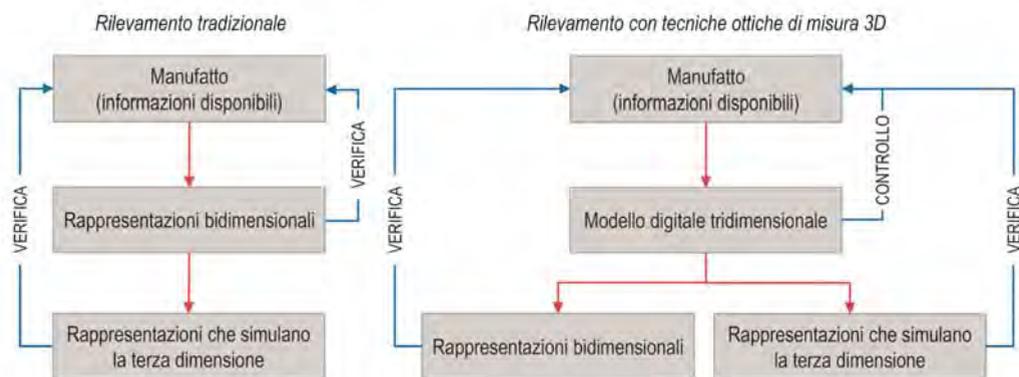


Fig. 1.5 Diagrammi esemplificativi del processo di produzione dei disegni nel rilievo tradizionale e con l'introduzione della nuvola di punti. (A. ZAPPANI, 2012)

## 2.2 Confronto tra sensori di misura attivi e passivi

Generalmente i modelli digitali 3D, sono creati con l'uso di sensori di misura attivi (*range camera* e *laser scanner 3D*) o passivi (fotocamere), integrati con dati provenienti da GPS, stazioni totali e altro. L'uso progressivo del modello, nel campo dei Beni Culturali – manufatti archeologici e architettonici, statue e dipinti – è dovuto, non solo al continuo diffondersi di nuove tecnologie per l'acquisizione digitale dei dati, ma anche all'innovarsi dello scopo e delle attese del periodo in cui si eseguono i rilievi e le indagini.

La progressiva diffusione di strumenti e tecnologie, e la conseguente messa a punto di nuove metodologie per il rilievo e la documentazione dei beni culturali, sia essi architettonici o archeologici, mette il rilevatore e lo studioso, davanti alla scelta di quale tecnica o strumentazione utilizzare, al fine di ottenere il risultato prefissato. La scelta di un metodo (strumento, tecnica ...) rispetto ad un'altro, richiede esperienza e una buona conoscenza dei diversi parametri, che caratterizzano i vari sistemi di acquisizione ed elaborazione dei dati, valutando non solo i vantaggi, ma anche i limiti e gli oneri che da essi ne derivano.

Le tecniche attive di misura 3D o semplicemente tecniche con sensori attivi, a loro volta possono essere suddivise in base al principio di funzionamento, che varia, secondo le caratteristiche e le dimensioni dell'oggetto da rilevare. I sistemi a triangolazione, sono quei sistemi adatti per acquisire piccoli o medi oggetti, con dimensioni da 1 a 500 cm, con un elevato dettaglio (statue, bassorilievi, dipinti, ...); sono caratterizzati da un'elevata accuratezza (decina di micron) e realizzano un campionamento molto denso, ed una nuvola di punti strutturata. Il principio di funzionamento è quello della triangolazione: una sorgente genera la radiazione luminosa, e un sensore digitale di solito un CCD, riprende la radiazione proiettata e riflessa dall'oggetto; essendo noti, la distanza tra il sensore ricevente e la sorgente che genera la luce, e gli angoli di proiezione e ripresa, attraverso semplici formule trigonometriche, si può determinare la distanza tra lo strumento e l'oggetto. La sorgente luminosa che proietta la *range camera*, può essere un singolo punto, una lama di luce (profilo) o un'area; i sensori che impiegano luce strutturata, proiettando un *pattern* luminoso (tipicamente delle bande verticali bianche e nere alternate),

proiettano sull'oggetto un'area. I sistemi a tempo di volo (TOF *Time-of-flight*), sono adatti per acquisire oggetti di dimensioni dell'ordine dai 1 a 200 m (manufatti architettonici, edifici, scavi archeologici, ...), con un'accuratezza dell'ordine dei millimetri; restituiscono una densa nuvola di punti non strutturata, a differenza dei sistemi a triangolazione, che restituiscono delle *range maps* già composte da triangoli (poligonalizzate). Tali sistemi sono basati sulla stima dell'intervallo di tempo impiegato da una sorgente luminosa per andare dallo strumento, all'oggetto da rilevare e tornare indietro, cioè il tempo di volo; tale intervallo è proporzionale alla distanza coperta dalla luce, che insieme alla conoscenza degli angoli di direzione del raggio luminoso, permettono di risalire alle coordinate 3D dell'area puntata dal raggio. Tali sistemi sono dotati di automatismi, che modificano la direzione del raggio *laser*, su un'area stabilita, consentendo di rilevare in automatico l'intera area grazie alla movimentazione della sorgente luminosa.

I sistemi di acquisizione di dati, che impiegano sensori attivi, emettono una sorgente luminosa (singolo raggio *laser*, fascio di raggi o *pattern* luminosi), che produce sulla superficie dell'oggetto da rilevare, un singolo punto, (o profili o ancora un'area luminosa) che viene ripreso dallo stesso strumento, che calcola la misura della distanza e la posizione dei punti battuti, rispetto al centro del sistema di acquisizione. Le caratteristiche principali di tali sistemi sono:

- velocità di restituzione in modo automatico, di una nuvola densa di punti 3D, senza l'ausilio di ulteriori attrezzature, come prismi o riflettori. Tale caratteristica è in funzione della velocità di acquisizione, cioè, il numero di punti al secondo che il sensore riesce a scansire. Per le *range camera* a triangolazione si può arrivare fino a 500000 punti/s, mentre per i *laser scanner* a tempo di volo sistemi TOF, la velocità di acquisizione varia da 1000 punti/s fino a 1000000 punti/s, ed è in funzione del metodo di valutazione della stima del tempo di volo. Questa può essere determinata in modo diretto, impiegando impulsi luminosi o *Pulsed Wave* (PW), misurando direttamente il ritardo tra l'impulso trasmesso e quello ricevuto (tempo di volo), oppure tramite l'invio di una radiazione luminosa continua a frequenza variabile o *Continuous Wave-Frequency Modulated* (CW FM), e la misura indiretta del tempo di volo, attraverso il calcolo degli sfasamenti

tra segnale trasmesso e ricevuto alle diverse frequenze. La prima categoria di strumenti (PW) è per principio adatta a misurare distanze anche molto lunghe (si possono coprire alcuni chilometri) con una intrinseca limitazione nella velocità di acquisizione e nella precisione che non può scendere sotto il millimetro; la seconda categoria (CW FM), limita il campo di lavoro, cioè, il campo di acquisizione a poche decine di metri, la precisione con tali strumenti migliora, fino a scendere sotto il millimetro, mantenendo anche una elevata velocità di acquisizione;

- l'acquisizione dei punti è indipendente dalle caratteristiche fisiche e morfologiche, della superficie da rilevare. La *range camera*, come il *laser scanner*, acquisisce milioni di punti, senza alcun tipo di interpretazione e comprensione dell'opera, senza distinguere fra gli elementi dell'architettura in esame e gli oggetti presenti sulla scena (alberi, macchine, passanti, insegne,...), senza alcun grado di discernimento tra spigoli o linee, che definiscono la geometria dell'oggetto, ciò comporta un grado di dettaglio costante sia per le parti piane che per quelle più articolate;
- la capacità di acquisire dati sia da oggetti di piccole e medie dimensioni, 1 ÷ 500 cm, quali (frammenti archeologici, statue, dipinti ed etc..) che da oggetti di grandi dimensioni, 1 ÷ 200 m, (monumenti architettonici, edifici, scavi, ed etc., in base alla tipologia di strumento impiegato).

L'utilizzo di tale metodologia, di contro presenta:

- costi elevati, relativi al costo dello strumento. Anche se si assiste ad una progressiva diffusione dei *laser scanner* e del loro utilizzo in differenti campi di applicazione, la prima significativa controindicazione è il loro costo, che va dai 40.000 a 90.000 euro per sistemi a triangolazione, e da 90.000 euro in su per i sistemi TOF più commerciali;
- *hardware* e *software* dedicati. Per l'uso di tali strumenti. È necessario che vi sia l'appoggio di un *computer*, per interfacciarsi con essi, gestire l'acquisizione e salvare i dati;
- grande quantità di tempo per la gestione e l'elaborazione della mole di dati che si generano;
- le acquisizioni richiedono un gran numero di riprese, per superare le zone di occlusione, che si vengono a creare, o per le dimensioni limitate del campo di scansione per ricoprire completamente la superficie dell'oggetto

d'acquisire (sistemi a triangolazione);

- sono poco flessibili nell'impiego, perché ingombranti e con "limitate" capacità di acquisizione, in virtù delle caratteristiche tipologiche, in quanto dedicati a un predefinito intervallo e volume di acquisizione (un sistema costruito per un impiego a corto raggio, non può essere impiegato a lungo raggio) e delle condizioni di illuminazione presente al momento dell'acquisizione; per i sistemi a triangolazione, le scansioni devono essere fatte nelle condizioni di intensità luminosa inferiore a 500 lux, se tali condizioni non si verificano, si deve utilizzare un sistema di oscuramento dell'oggetto da rilevare ;
- ha bisogno della fase di *texturizzazione*, poiché buona parte degli strumenti non acquisisce immagini adatte per le *texture*, quindi una volta ottenuto il modello 3D, si devono proiettare le immagini per ricostruire l'apparenza visiva;
- richiede una fonte di energia elettrica per l'utilizzo, il che può rendere difficile il rilievo in quei luoghi e in quelle situazioni dove ciò non ci sia.

I sistemi di acquisizione con sensori passivi, sono quei sistemi che sfruttano la luce presente nell'ambiente, sia essa naturale o artificiale, per estrarre dati. Le informazioni ricavate (contorni, forme tridimensionali, radiometria ...) sono rilevate unicamente da immagini, con metodi fotogrammetrici. I vantaggi allo stato attuale sono:

- bassi costi delle attrezzature, rispetto ai sistemi attivi; sostanzialmente per ottenere delle informazioni da immagini, per la restituzione di un modello 3D, o semplicemente la misura di una distanza sono necessari una macchina fotografica e un *software* per trasformare l'immagine in un sistema prospettico basato sulle leggi della geometria proiettiva;
- facilità d'uso e trasporto, in quanto si tratta di usare una macchina fotografica, portandosi dietro un treppiedi per le riprese con tempi di esposizione lunghi, e se le condizioni di illuminazione lo richiedono un set di lampade per ripristinare al meglio la condizione di luce naturale;
- la possibilità di ricostruire il modello di un oggetto, oggi non più disponibile o danneggiato, mediante l'utilizzo d'immagini d'archivio;
- accuratezza nella misura e un alto livello di dettaglio sia in fase di acquisizione, che di restituzione, grazie all'impiego di foto ad alta

risoluzione;

- l'ottenimento di un modello già *texturizzato*.

Le limitazioni che caratterizzano tali sistemi sono:

- richiedono buona esperienza in fase di acquisizione, cioè nella ripresa fotografica;
- molta interazione manuale per l'estrazione di dati dalle immagini, quando non si utilizza la fotogrammetria digitale con tecniche di *image matching*;
- dipendenza dalle condizioni d'illuminazione dell'ambiente in cui è posto l'oggetto;
- dipendenza dalle caratteristiche morfologiche della superficie dell'oggetto; il risultato migliora, infatti, con superfici aventi le caratteristiche geometriche definite e peggiora nel caso di geometrie irregolari o di superfici scultore e senza regolarità.

La realizzazione di un modello digitale 3D, richiede l'acquisizione di dati, che ricadono in un'ampia tipologia di casi, da piccoli oggetti con elevato dettaglio superficiale (colore, tessitura, scabrosità, fitta trama di lesioni, ...) al complesso architettonico, con la necessità di ottenere un'accuratezza/precisione, che va da qualche *micron*, al centimetro. La scelta della strumentazione più appropriata, tra metodi *range-based* (sistemi attivi) e *image-based* (sistemi passivi), dipende soprattutto dalle caratteristiche e complessità dell'opera da rilevare.

La complessità è in funzione dell'articolazione volumetrica e del livello di dettaglio; si possono avere, infatti, superfici completamente piane e semplici, fino a superfici completamente decorate, o con elementi riflettenti, o affrescati; ancora la complessità è condizionata dallo stato in cui si trova l'opera, (stato di degrado, rudere,...), dalle condizioni di accessibilità e dal grado di dettaglio richiesto per la restituzione.

Considerando le caratteristiche dei vari sistemi, l'impiego di uno solo di questi non è in grado di soddisfare tutti i requisiti, di un rilievo; spesso è necessario l'impiego di diverse metodologie e l'uso integrato di differenti strumenti, per trarre vantaggio dalle caratteristiche di ognuno.

La sequenza delle operazioni, che porta dall'oggetto reale al modello digitale, è definita come *3D acquisition pipeline*; sinteticamente, il flusso delle varie fasi operative con metodi *range-based* può essere schematizzato in:

- pianificazione della campagna di acquisizione delle prese, che sono in funzione del tipo di strumento, della complessità dell'opera, delle caratteristiche geometriche e materiche delle superfici, e dal grado di dettaglio al quale si deve arrivare, quindi dall'obiettivo della restituzione;
- acquisizione delle nuvole di punti; il prodotto di ogni singola presa dello *scanner*, è un modello costituito da una serie di punti, che descrivono la superficie dell'oggetto, in particolare per i sistemi a triangolazione viene definita *range maps*, è consiste in una matrice di punti 2,5 D, per i sistemi a tempo di volo, si preferisce il termine nuvola di punti, cioè un insieme di punti 3D; entrambe costituiscono un sistema ordinato di punti 3D, in un sistema di riferimento centrato sullo strumento. Per ogni singolo punto acquisito, vengono registrati: le tre coordinate cartesiane che ne definiscono la posizione spaziale, le tre coordinate colorimetriche nel sistema RGB e il suo valore della riflettenza<sup>10</sup>;
- allineamento o registrazione delle scansioni acquisite da differenti posizioni, per riportare le *range maps* in un unico sistema di riferimento globale. Quest'operazione può essere fatta in due modi: il primo, consiste nell'allineare le nuvole di punti, con l'ausilio di *target* (bersagli), i quali – in un numero non inferiore a tre per ogni singola posizione dello *scanner* – devono essere ripresi dallo stesso e da una stazione totale, in tal modo le singole scansioni insieme ai *target*, possono essere rototraslati rigidamente, sul sistema di riferimento globale dato dall'appoggio topografico. L'altro consiste nello sfruttare le forme stesse dell'oggetto, come riferimento per l'allineamento; in tali circostanze è necessario che vi sia una zona di sovrapposizione fra prese adiacenti pari al 30 ÷ 40 %. In fase di registrazione è l'operatore che manualmente, su una coppia di *range maps* identifica una serie di punti omologhi (almeno tre), in modo tale da bloccare una *range map* e fare rototraslare quella libera su quella bloccata. Si procede nello stesso modo per le successive, finché tutte siano registrate in un unico sistema di riferimento;
- fusione delle differenti *range maps* e creazione del modello poligonale (*mesh*), cioè di un reticolo di poligoni interconnessi (la maggior parte delle

---

<sup>10</sup> Parametro legato alla quantità di luce che viene riflessa dalla superficie su cui incide, ed è definita come il rapporto tra l'intensità della radiazione riflessa e quella incidente.

volte si usano triangoli), i cui vertici sono i punti acquisiti;

- modifica e ottimizzazione del modello (operazione di *editing*) per migliorare la qualità della *mesh*; queste operazioni si effettuano con modalità semiautomatiche, che generalmente consistono: nel chiudere lacune, controllare e invertire la direzione delle normali alla faccia delle superfici, decimazione di alcune aree dove si ritiene che siano state sovra-campionate, operazioni di lisciatura (*smoothing*);
- *texture mapping*, che aggiunge al modello geometrico, anche le proprietà della superficie (colore, finitura, grana, ...);



## CAPITOLO 3

## LA FOTOMODELLAZIONE

3.1 *Introduzione*

Le tecniche di *Image-Based Modeling and Rendering*, cioè di modellazione e resa, basate sulle immagini, sono il risultato di principi e teorie provenienti dai settori della *Computer Vision*, della Fotogrammetria e della Modellazione geometrica. Con simili tecniche si ha la possibilità di ricostruire la spazialità di una scena tridimensionale, ricavando i parametri necessari (dimensioni, posizione e forme) direttamente sull'immagine.

Nella restituzione delle forme architettoniche tramite foto, le difficoltà vengono principalmente dalla natura geometrica degli elementi da restituire; la scelta dei punti sulle foto con cui eseguire la ricostruzione geometrica è condizionata, infatti, sia dall'obiettivo della restituzione, che dal tipo di rappresentazione geometrica prescelto.

La capacità d'osservazione e comprensione della natura geometrica degli oggetti, si traduce in un processo di conoscenza, che mette in relazione due aspetti: il riconoscimento della forma e la conoscenza dei differenti tipi di rappresentazione geometrica.

In generale, si tratta di un processo di osservazione, apprendimento, scomposizione e ricostruzione delle parti di una architettura, che diviene una vera e propria fase di analisi e di interpretazione delle forme da rilevare. È

importante osservare che tale procedimento di “conoscenza” anticipata, di ciò che si deve rilevare e restituire, torna utile sia nella fase iniziale della progettazione del rilievo, che in quella successiva di restituzione. Questi due momenti sono interdipendenti: per rappresentare un elemento di architettura, bisogna conoscerne la forma, ma per conoscere la forma bisogna rilevare l’elemento stesso, decifrandone la natura geometrica.

Un limite della fotomodellazione sta nella difficoltà di riprodurre forme, che non presentano elevate discontinuità (sculture, o forme architettoniche, normalmente dette “liquide”), per la complicazione di individuare correttamente sulle foto, i punti in grado di definire la geometria delle forme stesse.

Le principali applicazioni di questa tecnica sono:

- la determinazione d’informazioni dimensionali; in tali casi, ci si limita solo ad acquisire le coordinate di alcuni punti isolati, per poi calcolarne le distanze reciproche.
- Elaborazione di piante, sezioni e prospetti, cioè di rappresentazioni bidimensionali su cui fare analisi sulla composizione architettonica, letture sullo stato di degrado, sui caratteri costruttivi, ecc.
- Ricostruzioni tridimensionali, generando un modello geometrico 3D, capace di descrivere le caratteristiche superficiali e volumetriche dell’oggetto reale.
- Restituzione dell’apparenza visiva, con l’obiettivo di riprodurre, più o meno fedelmente, le caratteristiche esteriori delle superfici reali.

In fotomodellazione l’ottenimento del modello, si raggiunge tramite un processo di elaborazione, che possiamo sintetizzare in tre fasi, strettamente relazionate tra di loro: acquisizione delle coordinate spaziali, ricostruzione tridimensionale e restituzione dell’apparenza visiva. A queste deve aggiungersi la fase di acquisizione dei dati, ovvero della ripresa fotografica e di alcune misure, utili per “scalare” il modello e introdurre dei vincoli in fase di modellazione.

La fase di acquisizione delle coordinate spaziali viene subito dopo il rilievo fotografico; considerando che si lavora con fotografie, quindi con rappresentazioni bidimensionali della realtà, per ottenere un modello 3D sono necessarie almeno due foto della stessa scena, ripresa da punti di vista differenti. Le operazioni che si compiono, sono quelle di selezionare punti sulla

prima immagine e di associarli ai punti omologhi della seconda immagine (o di altre immagini della stessa scena); i punti omologhi, possono essere dei vertici caratteristici (spigoli, punti singolari, ecc.), oppure delle discontinuità radiometriche superficiali. Raggiunto un numero sufficiente (di solito 8) di punti omologhi, possono stimarsi i parametri di orientamento delle camere impiegate, stabilendo le relazioni tra i centri ottici delle stesse camere, nelle diverse posizioni, i punti sulle immagini e le coordinate nello spazio. Con tale procedimento si arriva a raffigurare la struttura geometrica della scena in un unico sistema di riferimento spaziale. Ricostruito il modello geometrico della scena, si passa alla determinazione tridimensionale dell'oggetto, ricostruendo la natura geometrica degli elementi della scena, estraendo le informazioni per la creazione del modello, che a sua volta è in funzione del tipo di rappresentazione scelta.

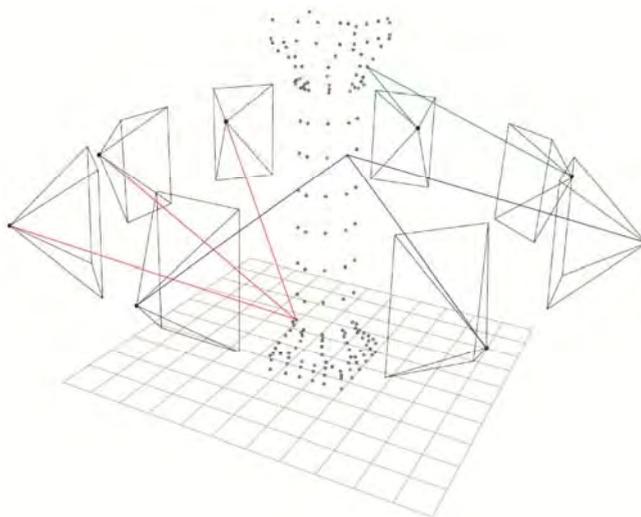


Fig. 3.1 – *Acquisizione delle coordinate spaziali*. Selezione dei punti omologhi sulle diverse immagini della stessa scena, i punti omologhi possono essere dei vertici caratteristici come spigoli, punti singolari, ecc (L. DE LUCA, 2011).

Il modello nasce in ambiente CAD, con l'aggiustamento di primitive geometriche, tramite la funzione di estrusione di facce poligonali, o con la generazione di superfici a partire dai profili.

L'ultima fase è quella di restituzione dell'apparenza visiva, che consiste nell'associare al modello 3D, ulteriori informazioni, relative ai materiali, alla

radiometria, o ad altre particolarità formali non presenti sul modello generato. L'apparenza visiva può servire anche per simulare elementi geometrici aggiuntivi, come rugosità superficiali, che per la loro modellazione richiederebbero un numero elevato di poligoni.

Questa fase prevede di proiettare le immagini (*texture*) sulle superfici del modello generato, sfruttando le relazioni proiettive – stabilite in fase di orientamento delle camere – tra i centri ottici delle camere stesse, i poligoni creati nello spazio e i piani immagine (fotografie).

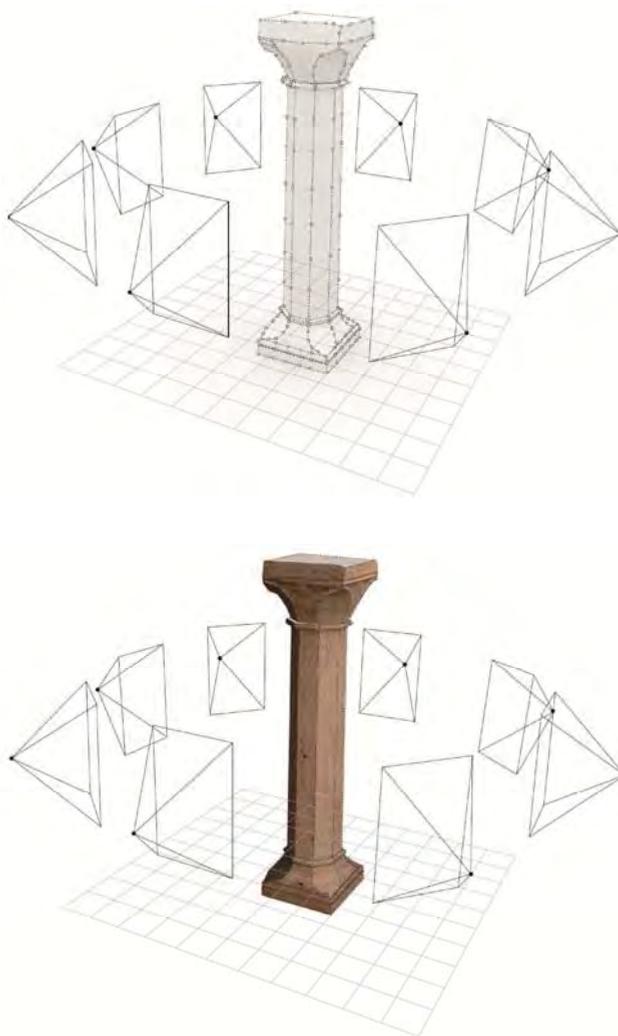


Fig. 3.2 a,b – Ricostruzione del modello geometrico e restituzione dell'apparenza visiva (L. DE LUCA, 2011).

### 3.2 La ripresa fotografica

Il processo di fotomodellazione, ha inizio con la ripresa fotografica; le immagini da utilizzare, possono essere quelle restituite direttamente da una macchina fotografica digitale, o quelle digitalizzate con uno *scanner* da foto, sviluppate da macchine fotografiche analogiche, o da immagini prese da documenti.

L'utilizzo di foto ottenute da macchine fotografiche digitali, oggi è senz'altro considerata la soluzione più diffusa e la più semplice, per l'elaborazione ed il trattamento digitale dell'immagini. L'impiego di tali macchine fotografiche è diventato uno *standard*, sia per la semplicità d'uso, che per il grado di definizione che offrono, ma soprattutto per la grande diffusione commerciale. Oggi le più comuni macchine fotografiche commerciali, hanno un tale grado di risoluzione dell'immagine (12-18 milioni di *pixel*) che consente di ottenere grandi quantità d'informazioni, necessarie per le operazioni di misura e di restituzione grafica.

Per tutte le operazioni e per tutte le fasi del processo di elaborazione – che partendo da immagini, consentono di ottenere informazioni, – la sola risoluzione dell'immagine non basta; devono essere presi in considerazione (nel condurre una campagna di rilievo fotografico) altri parametri che permettono di recuperare le informazioni necessarie, quali: la qualità delle foto, il contenuto di informazioni che devono avere e, quindi, le diverse modalità di ripresa fotografica, cioè come scattare le foto, sia in relazione alla complessità morfologica (architettonica) del soggetto da rilevare, che in relazione dello scopo del rilievo (rilievo a grande scale, o rilievo di dettaglio).

#### *Qualità della fotografia: illuminazione, nitidezza, contrasto*

Al fine di ottenere delle fotografie utili per la ricostruzione di un oggetto, la ripresa fotografica è la fase più importante di tutto il processo; occorre curarla con molta attenzione, partendo dalla scelta della macchina fotografica più adeguata, fino alla composizione delle singole fotografie ed alla regolazione di tutti quei parametri, che determinano la resa ottimale delle fotografie.



Fig. 3.3 – Macchina Fotografica compatta e bridge.

Le macchine fotografiche digitali, oggi utilizzate, si possono raggruppare in tre categorie, che vanno dal livello amatoriale a quello professionale; le tre categorie di macchine sono: le compatte, le *bridge* e le *reflex*.

Le compatte sono delle macchine fotografiche “tutto in uno”, cioè hanno il corpo macchina e l’obiettivo uniti e non è possibile cambiare obiettivo che, di solito, ha uno *zoom*, per ottimizzare la visione della scena; sono tutte *autofocus* (messa a fuoco automatica) e con un elevato grado di automatismi (non permettono regolazioni manuali avanzate, come appunto succede nelle *reflex*). Tali fotocamere sono dotate di uno schermo LCD, utile per inquadrare l’area da fotografare così come sarà poi salvata su *file*; per visualizzare la fotografia appena scattata e per vedere e regolare (quando lo permettono) alcune impostazioni della fotocamera come il tempo di posa, l’apertura del diaframma, la sensibilità delle ISO<sup>11</sup>, il bilanciamento del bianco ed altre ancora.

---

<sup>11</sup> Rappresentano la capacità di reagire più o meno sensibilmente all’azione della luce, da parte della pellicola o del sensore. Nel campo del digitale, il termine “sensibilità ISO”, si riferisce al grado di amplificazione, che a posteriori, subisce il segnale in uscita dal sensore, per essere elaborato e convertito nel valore cromatico che andrà a formare l’immagine finale. Il sensore ha una sua intrinseca sensibilità, che non può essere variata, da questo esce sempre un segnale d’intensità proporzionale all’intensità di luce incidente, la variazione del valore dell’ISO, va ad amplificare tale valore in uscita, permettendo di acquisire immagini anche in condizioni di poca illuminazione. Più alto è il valore ISO, più sensibile è “il sensore”, quindi maggiore sarà la sua capacità di catturare immagini in ambienti poco illuminati; l’aumento della sensibilità, si porta dietro anche un aumento di rumore digitale (luminance noise), sulle immagini, evidenziandosi soprattutto nelle aree uniformi o in zone scure con pochi dettagli, tali effetti sono simili a quelli delle immagini da pellicola ad alta sensibilità.



Fig. 3.4 – Macchina reflex e relativo schema ottico di funzionamento.

Le macchine *bridge* sono un compromesso tra le compatte e le *reflex*; come le compatte hanno un obiettivo fisso non inter-cambiabile, ma con ottiche migliori e permettono alcuni controlli avanzati sulle regolazioni, paragonabili alle *reflex*.

Le macchine *reflex* sono composte da un corpo macchina e da un obiettivo che può essere inter-cambiabile con altri; danno la possibilità di avere controlli avanzati su diversi tipi di parametri e consentono di regolare la messa a fuoco, l'esposizione, la sensibilità delle ISO, il tipo ed il formato del *file* in uscita (e tanti altri parametri), oltre ad avere ottiche e sensori più performanti. Tali caratteristiche rendono la macchina di tipo professionale.

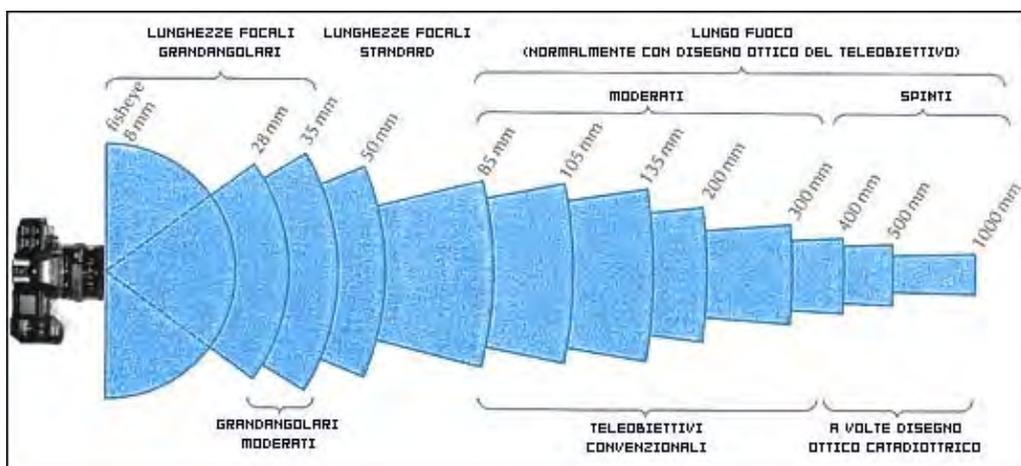


Fig. 3.5 – Schema dei differenti obiettivi delle fotocamere.

Nel caso di un rilievo fotografico per un organismo architettonico, l'uso di una macchina fotografica *reflex* digitale, rappresenta la soluzione migliore, per la possibilità di montare diversi tipi di obiettivi, in funzione delle condizioni di ripresa, delle dimensioni dell'oggetto da rilevare e per l'opportunità che offre di regolare i diversi parametri di controllo della qualità delle fotografie, (messa a fuoco, l'esposizione, il bilanciamento dei bianchi, formato di *file* in uscita...). Sinteticamente gli obiettivi vengono distinti in base alla loro lunghezza focale, in grandangolari con lunghezze focali da 8mm a 35, obiettivi *standard* con focale che va da 35mm a 85mm, e infine ci sono i teleobiettivi con lunghezze da 85mm a 600mm.

Tipico esempio di messa a fuoco, è la fotografia che presenta un primo piano perfettamente nitido e lo sfondo sfuocato. I dispositivi di messa a fuoco possono essere manuali, (la messa a fuoco si effettua ruotando la corona dell'obiettivo, su cui è inserita una ghiera con valori che variano da una distanza minima, espressa in cm, fino all'infinito), oppure automatici, attivati direttamente dalla macchina, non appena si schiaccia il pulsante di scatto, a "mezza corsa". La messa a fuoco in automatico, cioè l'*autofocus*, è regolata direttamente dalla macchina, mediante un sistema di microsensori. Il tempo di esposizione alla luce, la velocità dell'otturatore e l'apertura del diaframma, definiscono l'esposizione di una macchina fotografica, ovvero la quantità sufficiente di luce che deve raggiungere l'elemento fotosensibile (nelle fotocamere analogiche è la pellicola, nelle digitali è il sensore), affinché possa essere impressionato. Il diaframma, quindi, è quel dispositivo che è capace di modificare la quantità di luce che passa attraverso l'obiettivo; il tipo più usato è a "iride", chiamato così per la "somialtanza" con l'occhio umano. È posizionato al centro dell'obiettivo ed è costituito da lamelle mobili, che regolano la grandezza del foro dal quale passa la luce; maggiore sarà l'apertura del diaframma e maggiore sarà la quantità di luce entrante. Tale quantità, a sua volta, può essere modificata dall'otturatore, che regola il tempo di esposizione alla luce della pellicola, o del sensore. Tale tempo si misura in frazioni di secondi, (per esempio  $1/100$  = un centesimo di secondo), o in secondi. Se si sceglie una apertura del diaframma troppo grande ed una velocità di otturazione troppo rapida ( $1/8$  di secondo), si ottiene una foto sovra-esposta (troppo chiara), inversamente, con una velocità di otturazione lenta ed una

apertura piccola, si ottiene una foto sotto-esposta (troppo scura).



Fig. 3.6 – Diaframma a iride.

La sensibilità ai colori da parte della fotocamera, non è paragonabile a quella dell'occhio umano; per poter avvicinare la soglia cromatica a quella della realtà, bisogna bilanciare il punto di bianco. Con tale tecnica si tende ad individuare gli oggetti bianchi nella scena e a farli apparire completamente neutri (senza dominanti rosse, blu o gialle, che invece trasparirebbero per il tipo di luce presente nella scena). Una volta corretto il bianco, anche tutti gli altri colori appaiono naturali; con le macchine analogiche si utilizzavano dei filtri, o si sceglievano delle pellicole con diversa sensibilità. Oggi con le macchine digitali, è come se la macchina disponesse al proprio interno di una serie di filtri elettronici e li utilizzasse automaticamente. Nelle macchine digitali, anche il bilanciamento del bianco automatico; talvolta però l'automatismo non funziona e per ovviare a tale problema, si può regolare il punto di bianco in modo manuale, fotografando semplicemente una superficie bianca. In tal modo la macchina legge il tipo di luce riflessa dall'oggetto e modifica le proprie impostazioni, in modo da far apparire corretti anche tutti gli altri colori.

Quando si scatta una fotografia, prima che il *file* venga scritto e salvato sulla scheda di memoria, la fotocamera svolge diverse operazioni:

- il sensore raccoglie i dati generati dai singoli pixel e li converte da analogici in digitali; i dati sono ancora allo stato “grezzo” (RAW), non ancora elaborati;
- ai singoli pixel viene attribuito il valore di brillantezza e colore (interpolazione Bayer), in quanto i dati catturati dal sensore, sono riferiti alla sola luminanza, cioè il sensore “vede” in bianco e nero;
- le impostazioni per la ripresa (risoluzione, bilanciamento del bianco, qualità

dell'immagine, contrasto, livello di nitidezza ecc.), vengono applicate ai dati grezzi;

- il file immagine, viene elaborato da un microprocessore all'interno della fotocamera (simile a quello che si trova all'interno di un PC) ,e salvato in formato compresso (.JPEG, o altro proprietario); a questo punto, il file viene scritto sulla scheda di memoria.

Con l'utilizzo di *software* di foto *editing*, per delle elaborazioni in *post-processing* – come il bilanciamento del bianco, correggere gli errori di esposizione, o aumentare la nitidezza, lavorando su un file immagine che ha subito una “elaborazione in camera” (file salvato in formato .JPEG o altro) –, si ha una limitata possibilità d'intervento, rispetto alle elaborazioni che si possono effettuare su un *file* contenente dati “grezzi”, non compressi. Un'immagine salvata nel formato .RAW, è l'equivalente digitale del negativo della pellicola (esposta, ma non ancora sviluppata), in quanto la fotocamera non ha ancora applicato le impostazioni settate prima dello scatto; tali impostazioni vengono salvate in un *file* separato ed associato al *file* .RAW, per la visualizzazione su *PC*. Normalmente un *file* .JPEG viene codificato a 8 *bit* ( $2^8= 256$  livelli di luminosità), cioè 16,7 milioni di colori; un *file* .RAW è codificato a 12 *bit* ( $2^{12}= 4096$  livelli di luminosità), cioè 68,7 miliardi di colori. Le informazioni in più che ci sono in un *file* .RAW, permettono di effettuare in modo più efficace le operazioni di *post-processing*. L'utilizzo dei *file* .RAW, comporta anche alcuni svantaggi, come la dimensione maggiore dei *file* (una fotocamera di 12 mega *pixel* produce immagini .JPEG alla massima qualità di circa 3 MB, mentre i *file* .RAW della stessa macchina sono di oltre 9 MB), che richiede tempi maggiori per essere salvati sulla scheda di memoria, inoltre tali immagini non sono direttamente visibili e vanno elaborate singolarmente sul *PC*.

La qualità delle foto, vanno ad influenzare direttamente i risultati di calibrazione ed orientamento delle camere e di restituzione del modello geometrico, che può essere 3D oppure bidimensionale, come nel caso di un raddrizzamento per ottenere il prospetto di una facciata. Una regolazione non ottimale dell'esposizione della macchina fotografica va ad incidere sull'illuminazione della foto e considerando che alla fine del processo di fotomodellazione, le fotografie saranno utilizzate per l'estrazione e proiezione

di *texture* sul modello, è importante che tale illuminazione sia controllata ed omogenea, per evitare che vi sia discontinuità cromatica sulle diverse parti del modello. Tali difetti si accentuano soprattutto sulle parti dell'oggetto dove l'ombra è più pronunciata; per questo è opportuno eseguire il rilievo fotografico in un arco di tempo ristretto, con una proiezione delle ombre, quasi identica in tutte le foto. La condizione ideale è quella di eseguire le foto per un rilievo, con un cielo coperto, in condizioni di illuminazione omogenea; in queste condizioni, in fase di restituzione dell'apparenza visiva, si possono utilizzare delle sorgenti luminose digitali, calcolando le ombre direttamente sul modello 3D. Contrariamente, impostando una buona esposizione della fotocamera, si otterranno delle foto ben contrastate, il che facilita l'interpretazione delle forme e la leggibilità degli elementi architettonici, con delle foto ben contrastate, la ricostruzione di un elemento architettonico caratterizzato da un elevato grado di "dettaglio" e l'identificazione dei vertici caratteristici, risulta più agevole, rispetto a delle foto poco contrastate.

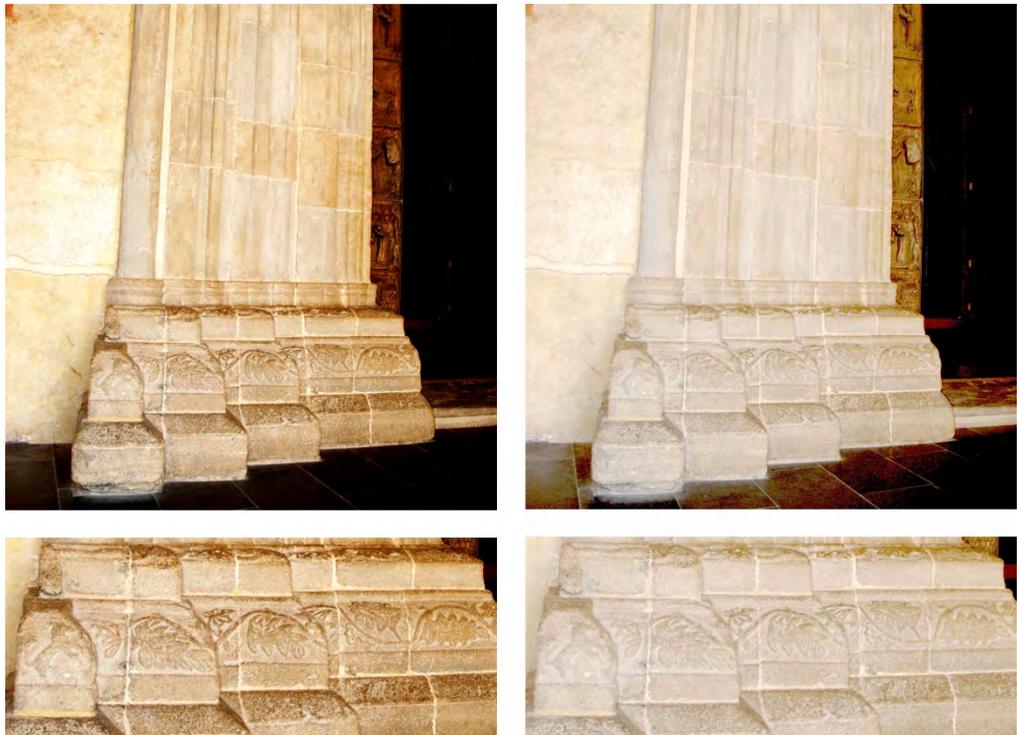


Fig. 3.7 – Due fotografie con un livello di contrasto differente.

In quest'ultimo caso, l'uso di un *software* di fotoritocco, può essere utile per aumentare il contrasto. Una fotografia, con la messa a fuoco e la velocità d'otturazione non ben regolate, ha una perdita di nitidezza ed una sfocatura della stessa foto. Processando una foto sfocata, si possono commettere maggiori errori nella calibrazione e orientamento delle camere, perché si perde precisione in fase di accoppiamento dei punti omologhi.

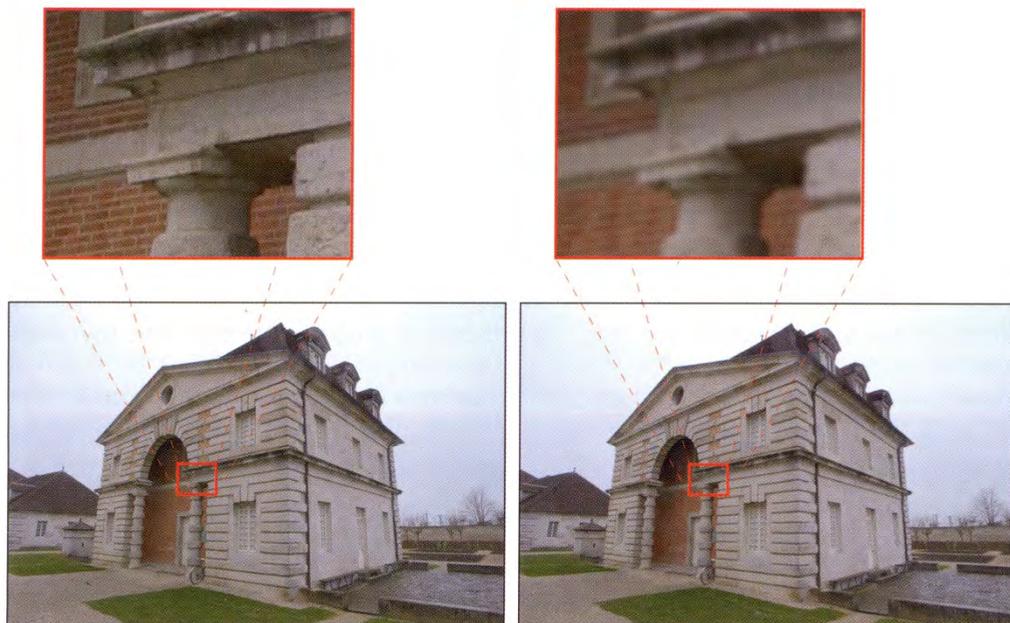


Fig. 3.8 – Due fotografie con un livello di nitidezza differente. Si noti la perdita di identificazione di alcuni vertici (L. DE LUCA, 2011).

Quando si opera in condizioni di poca luce (foto d'interni), si è costretti ad utilizzare dei tempi di otturazione molto lunghi, per far sì che il materiale fotosensibile (sensore o pellicola) possa impressionarsi; in questi casi può succedere che le foto risultino sfocate per l'instabilità della mano (si può evitare ciò utilizzando un treppiede e un sistema per lo scatto a distanza, o un temporizzatore).

### 3.3 Tecniche di ripresa fotografica

Per la restituzione delle coordinate di punti nello spazio, come detto, è necessario che gli stessi punti siano visibili in almeno due foto prese da punti di vista differenti; partendo da tali presupposti, l'acquisizione fotografica va eseguita con un insieme di foto, o meglio con gruppi di foto, scattate tutte con la stessa lunghezza focale e cercando di mantenere la stessa distanza dall'oggetto interessato dalla ripresa.

I gruppi di foto, possono essere costituiti da due o più foto, in funzione della complessità morfologica dell'oggetto, della scala e delle distanze di ripresa e dovrebbero rappresentare l'oggetto, riprendendolo interamente in ogni foto; l'oggetto dovrebbe, cioè, occupare una grande parte dello spazio fotografico, in modo da sfruttare tutta la risoluzione dell'immagine.

Tenendo conto del grado di complessità geometrica, dell'occlusione per elementi esterni, delle auto-occlusioni, delle dimensioni e del contesto nel quale l'oggetto si trova, una sola tecnica di ripresa fotografica generalmente non è sufficiente per rappresentarlo al meglio; bisogna integrare differenti tecniche di ripresa, operando con riprese ad assi convergenti, ad assi paralleli, riprese panoramiche ed eseguendo gruppi di foto con differenti obiettivi fotografici per acquisire i dettagli. L'integrazione delle differenti tecniche è anche necessaria perché si abbia una buona distribuzione delle informazioni nello spazio.

L'acquisizione ideale sarebbe quella di eseguire le foto, girando intorno all'oggetto, ad una distanza costante per avere il maggior numero di informazioni (punti caratteristici), visualizzabili su tutti i piani di profondità che rappresentano l'oggetto. Per avere a disposizione molti punti caratteristici sui diversi livelli di profondità e, quindi, sulla volumetria dell'oggetto da rilevare, le foto debbono essere eseguite con i punti di vista obliqui (inclinate di  $45^\circ$ ), rispetto ai piani principali che caratterizzano l'oggetto. Tutte le "informazioni" elencate tornano utili sia in fase di calibrazione, nella successiva fase di restituzione del modello 3D.

In campo architettonico, le condizioni di una ripresa fotografica ideale, non sono sempre applicabili, soprattutto se si vuole girare intorno all'oggetto d'indagine; le circostanze che possono impedire la non applicabilità del metodo

sono diverse e vanno dalla conformazione architettonica ed urbana in cui è inserito l'oggetto da rilevare, all'impossibilità di arretrare per eseguire foto in cui tutto l'oggetto sia inquadrato; c'è poi da considerare, la presenza di ostacoli (macchine, persone, alberi ecc.) che, come ricordato, costringe ad eseguire foto, con differenti tecniche di ripresa.

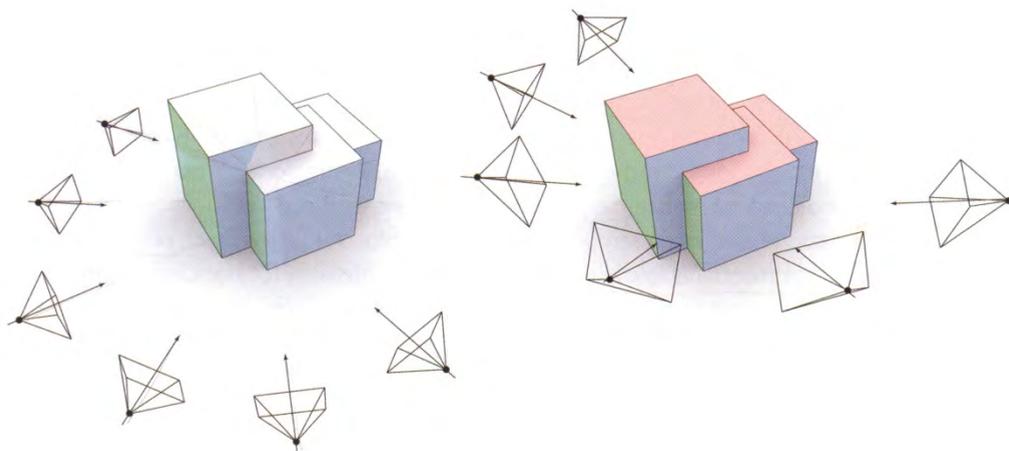


Fig. 3.9 – *Schema di ripresa ideale*. Se le condizioni lo consentono si dovrebbe riprendere l'oggetto girandovi intorno, mantenendo la stessa distanza dall'oggetto per ogni fotografia (L. DE LUCA, 2011).

Come detto, per l'individuazione delle coordinate di alcuni punti nello spazio tramite foto e per la definizione di un numero minimo di fotografie da scattare per il rilievo dei dati di un oggetto, sono necessarie almeno una coppia di fotografie, riprese con l'asse di mira inclinato rispetto al piano caratteristico su cui si trovano i punti che si vogliono individuare.

La condizione di avere l'asse di mira inclinato è necessaria per una maggiore correlazione delle immagini e per assicurare una maggiore visione spaziale. È bene ricordare che se per la determinazione delle coordinate, sono necessarie almeno una coppia di fotografie dello stesso piano – ripreso da angolature differenti con l'asse di mira inclinato – per l'estrazione di *texture*, o per la restituzione del prospetto di un edificio, sono necessarie fotografie riprese con l'asse di mira ortogonale, al piano che meglio caratterizza l'oggetto.

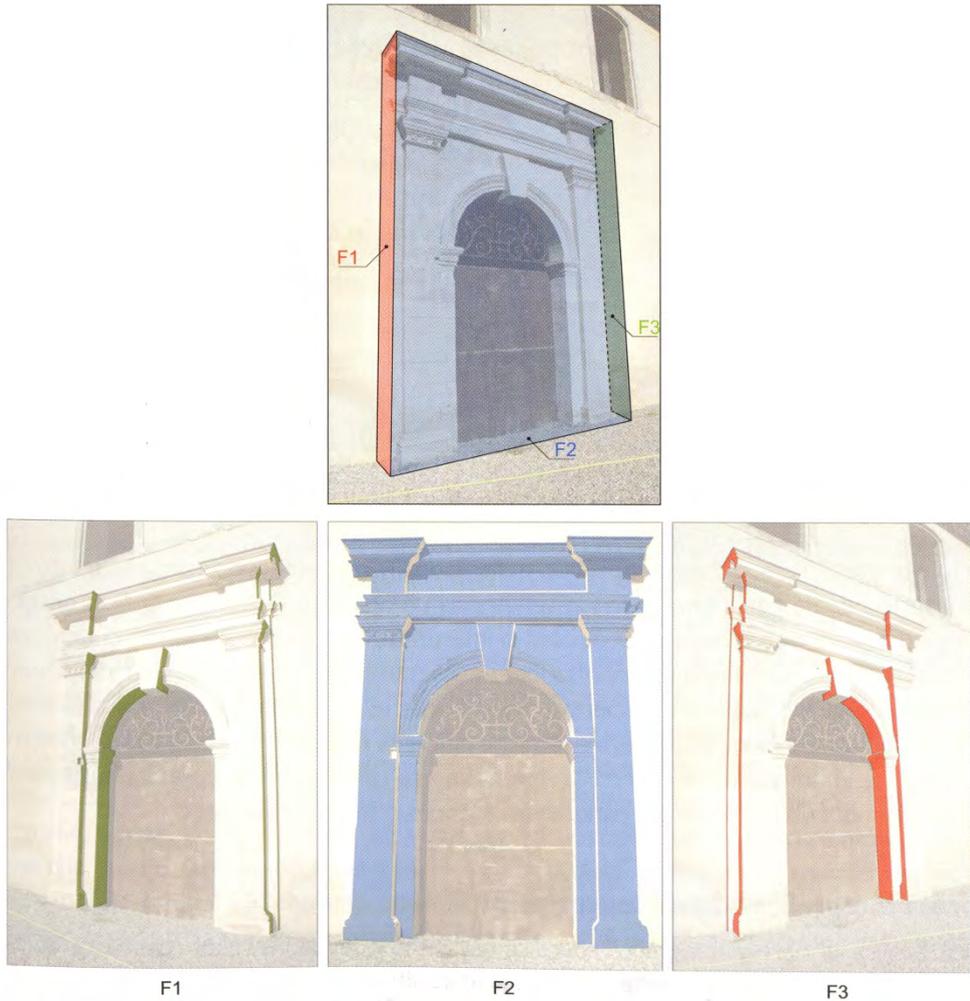


Fig. 3.10 – Complessità morfologica dell’oggetto (L. DE LUCA, 2011).

### *Ripresa ad assi convergenti*

Le caratteristiche della ripresa ad assi convergenti sono quelle, per così dire, ideali: tutte le foto vengono eseguite girando attorno all’oggetto, prendendo come riferimento l’asse verticale passante per il baricentro, mantenendo la stessa distanza dall’oggetto, inquadrando tutto l’oggetto e scattando le foto con l’asse di mira inclinato di  $30/45^\circ$ , rispetto ai principali piani caratteristi

dominanti. In alcuni casi, quando si è in presenza di un solo piano caratteristico dominante (ad esempio, rilievo o ricostruzione di un portale) – il che ci riporta alla generalizzazione del caso più semplice –, le sole due fotografie eseguite ad assi convergenti, con l’asse di mira inclinato, possono non essere sufficienti per la presenza di eventuali auto-occlusioni che possono crearsi; occlusioni che rendono impossibile l’individuazione di punti omologhi su parte delle foto stesse, con l’impossibilità di calibrazione e orientamento delle camere.

In queste circostanze, diviene necessaria una terza foto, ortogonale al piano, che servirà per raccordare le due foto scattate ad assi inclinati, tale foto risulterà anche la più indicata per l’estrazione della *texture*; inoltre, come foto raddrizzata con un piano parallelo al piano dominante, può essere utilizzata come supporto per il disegno bidimensionale degli elementi architettonici e di decoro.

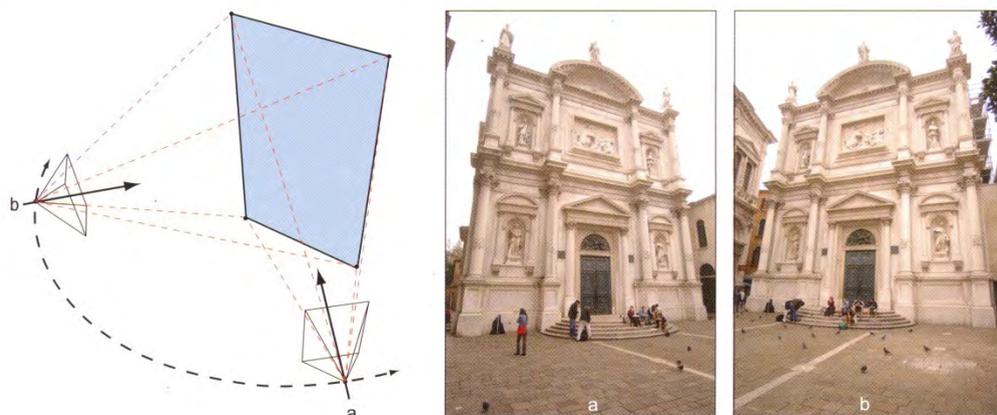


Fig. 3.11 – Ripresa fotografica ad assi convergenti (L. DE LUCA, 2011).

La scomposizione e quantificazione di un oggetto, secondo i piani caratteristici dominanti, permette di calcolare e valutare sul campo, il numero e la posizione delle riprese fotografiche necessarie per l’acquisizione, nonché di suddividere e raggruppare tra loro le foto, con la stessa distanza focale e la stessa distanza di presa rispetto all’oggetto.

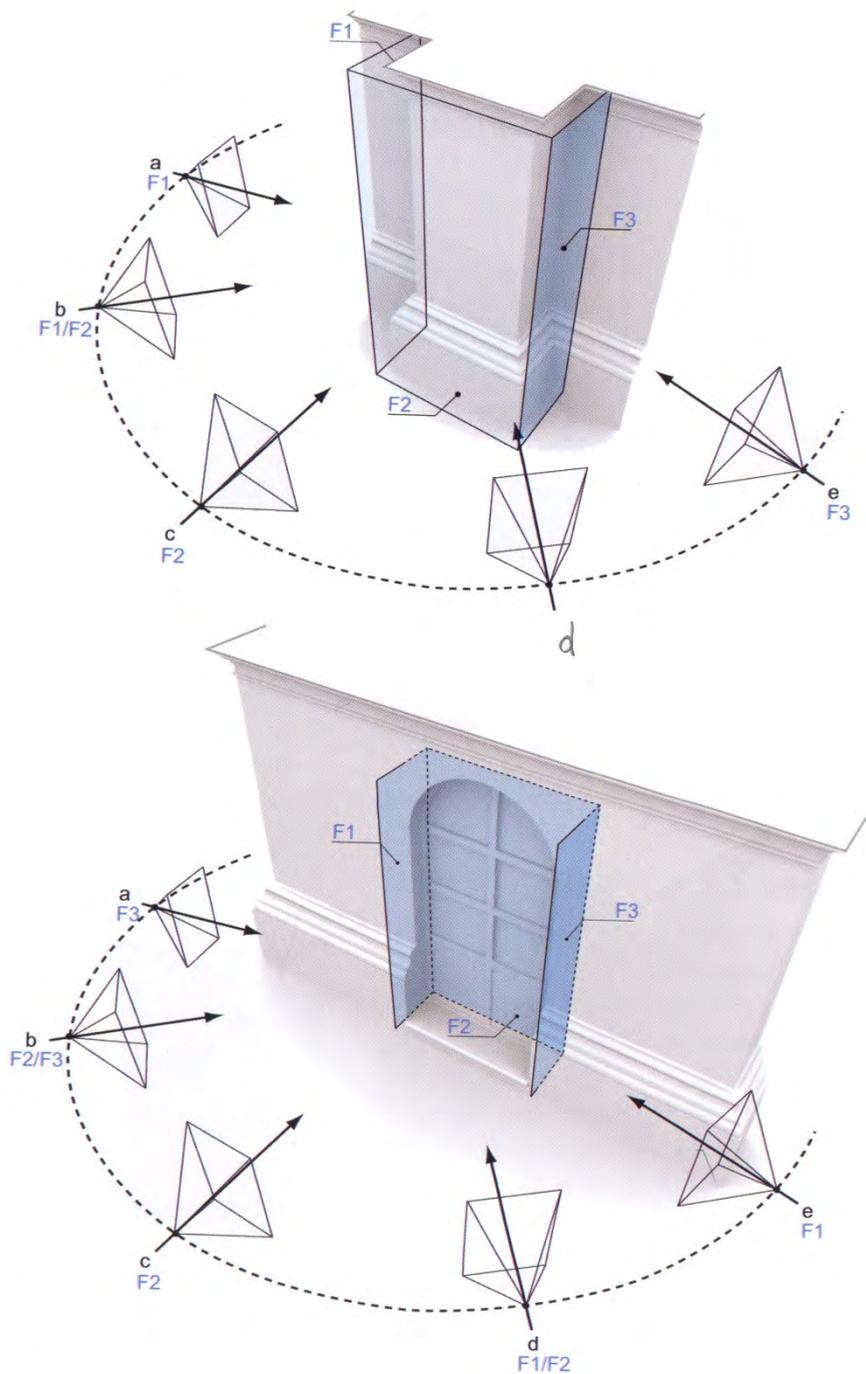


Fig. 3.12 – Individuazione dei piani caratteristici (L. DE LUCA, 2011).

### *Ripresa ad assi paralleli*

Tale tecnica di ripresa si esegue generalmente con l'asse di mira ortogonale ai piani di facciata, e spesso è parte integrante delle riprese ad assi convergenti; in particolare risulta utile quando si vuole restituire l'elaborato bidimensionale di una facciata. L'elaborazione di un "mosaico" di foto raddrizzate permette di ottenere un livello di risoluzione molto elevato, in quanto l'intera facciata risulta "coperta" da diverse foto, con una risoluzione maggiore rispetto ad una unica fotografia della facciata stessa; tale mosaico risulta utile anche per delle restituzioni dettagliate di apparati decorativi, o per elaborati sullo stato di conservazione (mappatura del degrado).

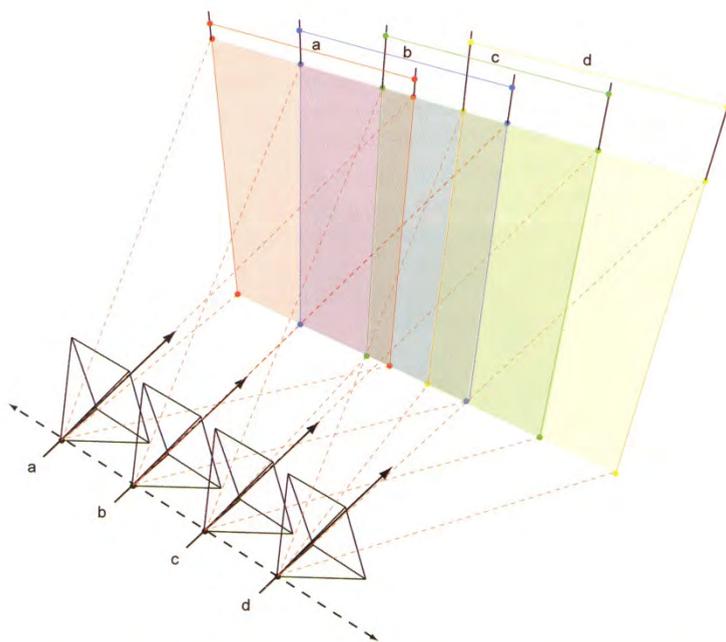


Fig. 3.13 – Ripresa fotografica ad assi paralleli (L. DE LUCA, 2011).

### Ripresa panoramica, cilindrica e sferica

La ripresa panoramica consiste nell'eseguire una rotazione di fotografie, scattate tutte da uno stesso punto, fino a completare l'angolo giro; il principio è quello di assemblare le singole foto, su uno sviluppo cilindrico o sferico. Ogni foto deve presentare una parte in sovrapposizione con la successiva, affinché possano essere assemblate in un unico punto.

Per eseguire la ripresa con tale tecnica, occorre utilizzare un treppiede con una "testa" panoramica, che permette la corretta sovrapposizione delle foto (tutte con lo stesso centro ottico).

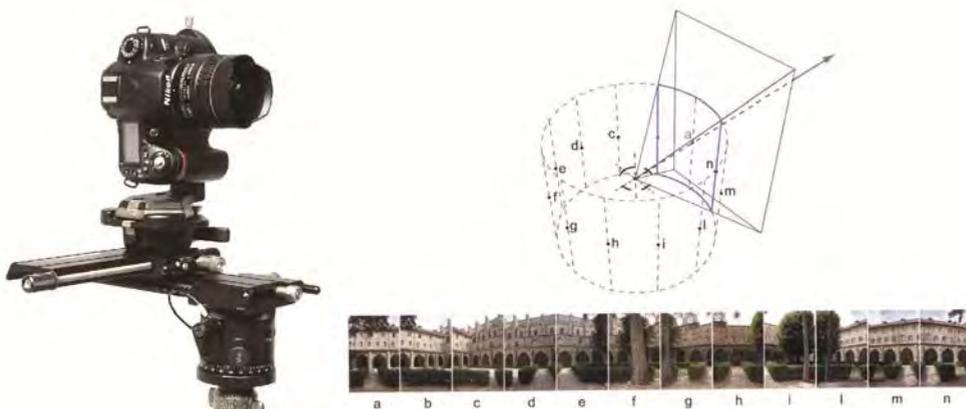


Fig. 3.14 – Testa panoramica e schema di ripresa fotografica cilindrica (immagine a dx, L. DE LUCA, 2011).

Il numero di scatti da eseguire, per catturare un giro d'orizzonte dipende da tre fattori: la distanza focale dell'obiettivo che si sta utilizzando, l'orientamento della posizione della macchina fotografica, o meglio la posizione del sensore o della pellicola (posizione di paesaggio con macchina orizzontale o ritratto con la macchina in verticale) e la percentuale di sovrapposizione tra gli scatti. La distanza focale dell'obiettivo e l'orientamento della macchina (paesaggio o ritratto), definiscono l'angolo di campo reale dell'obiettivo ( $A$ ): sostanzialmente determinano il campo visivo, con una determinata distanza focale ( $f$ ) e dimensione del sensore ( $L$ ). la figura 3.15 mostra le relazioni geometriche ed algebriche tra l'angolo di campo, le dimensioni del sensore e la distanza focale.

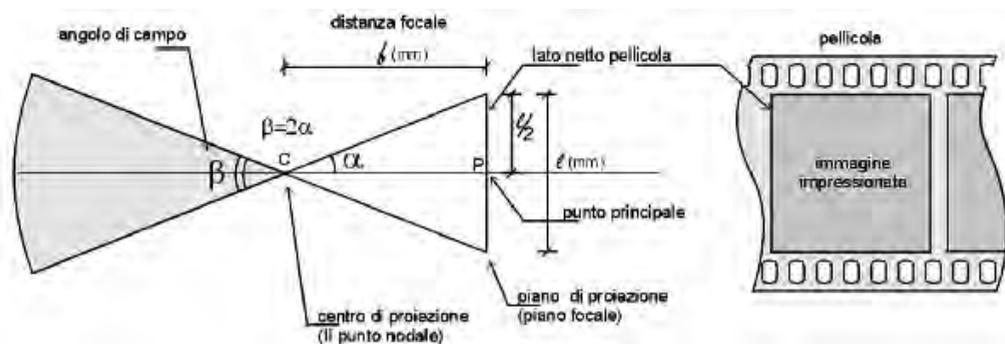


Fig. 3.15 – Relazione tra la distanza focale, il formato del sensore e l'angolo di campo  $\beta$ .

Una semplice relazione, che lega l'angolo di campo reale (A) e la percentuale di sovrapposizione (P), permette di determinare il numero di foto (N) necessario e, conseguentemente, il settore angolare tra due immagini successive:

$$N = AP / ((P/100)A) \quad \text{dove } A = 2 \times \arctan [L/2/f]$$

Nel determinare il numero di scatti da eseguire (N), è opportuno scegliere un numero che sia un divisore facile dell'angolo giro ( $5 = 75^\circ$ ;  $6 = 60^\circ$ ; ...).

Nella formula precedente il termine (AP) fa riferimento al settore angolare che si desidera riprendere nel panorama ( $180^\circ$ ;  $360^\circ$ ); mentre (P), è la percentuale di foto utile dopo la sovrapposizione desiderata (per una sovrapposizione del 20% si ha un  $P = 100 - (20 + 20) = 60$ ).

A differenza delle riprese ad assi convergenti o paralleli, indicate per il rilievo del singolo oggetto, la ripresa panoramica, è particolarmente indicata per gli spazi aperti come piazze, chiostri, cortili, etc. Anche se è possibile ricostruire un modello approssimato dello spazio con una sola immagine panoramica, per ottenere le informazioni metriche è necessario integrare tale immagine con una ripresa ad assi convergenti, spostandosi dal punto di acquisizione della ripresa panoramica.

### 3.4 Acquisizione delle coordinate 3D da foto

La conoscenza delle coordinate 3D  $(x_p, y_p, z_p)$  di un punto della scena, rispetto ad un unico sistema di riferimento, nel quale sono messi in relazione i centri ottici delle camere – nelle diverse posizioni di ripresa – ed i punti omologhi sull'immagini, passa attraverso l'accoppiamento 3D/2D degli stessi punti; per cui ad ogni foto dello spazio, ne viene assegnato uno omologo sull'immagine. Tale procedimento presuppone la conoscenza della geometria della scena fotografata, quindi la geometria interna della camera (distanza principale "c", posizione del punto principale  $(x_c, y_c)$ , le distorsioni indotte dall'obiettivo  $\Delta_x, \Delta_y$ ), le posizioni e gli orientamenti della stessa nelle varie riprese (coordinate del centro di proiezione  $(X_o, Y_o, Z_o)$  e gli angoli di rotazione intorno ai tre assi  $\Omega, \Phi, K$ ). Per restituire un modello geometrico da immagini fotografiche, la geometria dello spazio non è nota a priori; per conoscerla si deve effettuare un accoppiamento 2D/2D di punti, si devono cioè individuare punti omologhi, selezionati su una coppia di foto della stessa scena.

Nella maggior parte dei *software* di fotomodellazione, la determinazione della geometria della camera e della scena – vale a dire l'operazione di calibrazione ed orientamento della camera – viene valutata con l'individuazione di un minimo di 8 punti omologhi<sup>12</sup>, scelti su almeno due fotografie. Determinata una prima soluzione di calibrazione ed orientamento con l'accoppiamento degli 8 punti omologhi, tale soluzione può essere migliorata, con l'aggiunta di ulteriori punti omologhi sulle foto<sup>13</sup>; utilizzando le

---

<sup>12</sup> La ricostruzione della geometria della camera e della scena, equivale a calibrare la macchina fotografica, ed a determinare l'orientamento esterno; nel risolvere il problema intervengono 15 incognite, di cui 9 sono costituite dai parametri interni della camera e 6 sono di orientamento esterno. Tale incognite possono essere determinate con un sistema di 15 equazioni. Le 15 incognite che governano il sistema, sono relazionate dall'equazione di collinearità e per ogni singolo punto, si possono scrivere 2 equazioni di collinearità. Il sistema che si viene a creare, in realtà non è lineare, il che implica che per avere una soluzione, si devono avere sistemi di equazioni sovrabbondanti, cioè con un numero di equazioni maggiore del numero delle incognite. La risoluzione del sistema avviene con delle procedure iterative di minimizzazione dello scostamento tra la soluzione approssimata e quella esatta. Per avere un numero sovrabbondante di equazioni, il numero minimo dei punti necessari, per i quali scrivere l'equazione di collinearità, deve essere di almeno 8, in modo da scrivere 16 equazioni.

<sup>13</sup> L'introduzione di ulteriori punti omologhi, dopo aver eseguito la prima calibrazione e orientamento, permette di ottimizzare la soluzione trovata, andando a riproiettare sulle foto, i

conoscenze delle relazioni esistenti tra le foto e le proprietà della geometria epipolare<sup>14</sup>, il processo di calibrazione ed orientamento può ritenersi concluso quando tutte le rette epipolari si proiettano correttamente su tutti i punti omologhi considerati nelle foto.

Da un punto di vista applicativo, l'operazione di calibrazione e di orientamento della camera, resta un'operazione molto delicata, in quanto da questa dipendono la precisione e la coerenza geometrica delle restituzioni che si andranno ad eseguire. È importante avere una visione chiara del funzionamento dell'intero processo, al fine di controllare eventuali errori, ed approfondire alcuni aspetti come: la conoscenza delle caratteristiche interne della camera, l'individuazione dei punti omologhi, la loro distribuzione sull'immagine e nello spazio e i vincoli che tali punti possono ulteriormente esprimere per la calibrazione e l'orientamento della camera.

### *La geometria interna della camera: i parametri intrinseci*

Il modo in cui, il fascio di luce che attraversa l'obiettivo e si proietta sul piano immagine (pellicola o sensore), è descritto dai parametri intrinseci della macchina fotografica, che ne descrivono le proprietà proiettive. Tali parametri sono: la distanza principale (o focale dell'obiettivo), le coordinate del punto principale e le distorsioni.

La distanza principale è la distanza tra il piano di proiezione (sul quale si impressiona l'immagine) e il centro della lente frontale; questa determina il fattore d'ingrandimento dell'immagine stessa ed il campo angolare dell'obiettivo. Di solito gli obiettivi, presentano l'escursione dei valori che possono assumere (se l'obiettivo non è fisso); nelle macchine fotografiche compatte, il valore della distanza focale, può essere modificato con l'utilizzo dello *zoom* ottico; ma detto valore, riportato sull'obiettivo e sulla camera dal

---

punti prima individuati e di conseguenza, tutti i parametri calcolati saranno modificati, andando a minimizzare l'errore globale.

<sup>14</sup> La geometria epipolare individua le relazioni geometriche esistenti tra un punto della scena 3D e le relative immagini. Se la posizione relativa e la rotazione di una camera rispetto all'altra sono noti, l'analisi della geometria epipolare porta al seguente vincolo: per ogni punto sull'immagine, il corrispondente punto sull'altra immagine si trova sulla retta epipolare.

costruttore, o imposto dall'utente in fase di ripresa, non può essere impiegato, così come lo si trova<sup>15</sup>, per la descrizione della geometria della camera. Tale valore, deve essere calcolato con precisione, insieme alla posizione delle coordinate del punto principale proiettato dal centro prospettico sul piano immagine; quest'ultimo, essendo un elemento (pellicola o sensore) inserito meccanicamente nella camera, può avere piccoli errori di inserimento, e causare un disassamento tra il centro del piano immagine ed il centro di riferimento locale dell'immagini, posto sull'asse ottico.

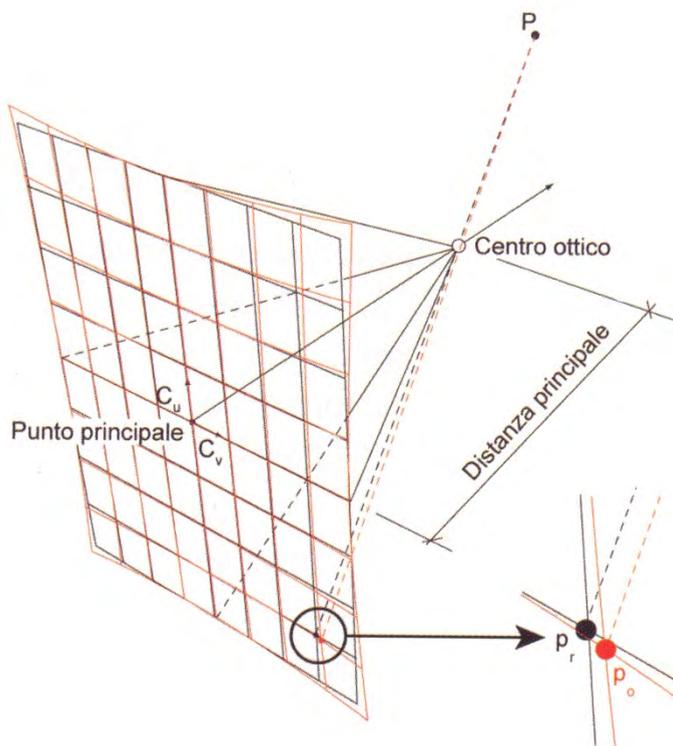


Fig. 3.16 – *Modello geometrico di una camera* (L. DE LUCA, 2011).

Le distorsioni non alterano la qualità e la nitidezza dell'immagini, ma il loro aspetto geometrico; tali alterazioni sono dovute essenzialmente al sistema ottico impiegato, e aumentano con l'aumentare della distanza dall'asse ottico. Le distorsioni più ricorrenti sono: quelle radiali, con le classiche forme a

<sup>15</sup> Viene impiegato come valore approssimato per la risoluzione del sistema di equazioni di collinearità.

“barilotto”, o a “cuscino”, che vengono generate rispettivamente da obiettivi grandangolari e da sistemi a teleobiettivi; quelle tangenziali, dovute al disallineamento del sistema di lenti che caratterizzano l’obiettivo; quelle “affini”, che vengono generate dalla non perfetta condizione di ortogonalità, tra il piano del sensore e l’asse ottico dell’obiettivo.

In fotogrammetria e, quindi, in un processo di fotomodellazione, le deformazioni indotte dalle distorsioni, possono essere controllate, e valutate, introducendo dei coefficienti, derivanti da funzioni polinomiali.

Per la determinazione della geometria della camera, nei vari *software*, in commercio si ha la possibilità di inserire alcune informazioni, in modo da ridurre il numero dei parametri da calcolare. I casi che si possono presentare sono i seguenti:

- non si ha nessuna informazione sulla macchina e sull’obiettivo utilizzati; le foto, infatti, possono essere prese da scansioni, da negativi, diapositive, o dal web. In queste situazioni, si dovrà indicare al *software* di elaborare le immagini, considerando che tutti i parametri interni possono variare da una immagine all’altra.
- Non si ha nessuna informazione sulla macchina e sull’obiettivo utilizzati, ma nelle prese delle foto si è tenuta costante la distanza focale e si è utilizzato il medesimo livello di *zoom*; in queste circostanze al *software*, si dovrà indicare che i parametri interni tra una immagine e l’altra sono costanti.
- Si ha la conoscenza di tutti i parametri interni della macchina e dell’obiettivo impiegato; in questo caso, la geometria della camera è determinata, resta solo il calcolo dell’orientamento esterno delle camere.

I parametri intrinseci delle camere, da inserire nei vari *software*, possono essere determinati con l’operazione di calibrazione delle camere, oppure andando a leggere i valori registrati nei “metadati EXIF” dei file immagine.

### *Individuazione dei punti omologhi: accoppiamento 2D/2D*

L’individuazione ed il successivo accoppiamento dei punti 2D/2D, cioè dei punti omologhi selezionati su una coppia di foto della stessa scena, è la fase del

processo di fotomodellazione in cui l'utente fornisce al *software* informazioni che legano reciprocamente le immagini tra loro, permettendo la stima del modello geometrico delle camere e l'orientamento delle stesse. I punti omologhi da selezionare, su almeno due, o più, foto della stessa scena, possono essere dei vertici caratteristici, come spigoli, punti singolari, oppure discontinuità radiometriche sulle superfici. La scelta dei punti in questa fase, non deve basarsi sulla determinazione dei punti caratteristici della forma da restituire, ma poggiarsi su aspetti puramente geometrici, distribuendo in modo equilibrato i punti nello spazio 3D della scena e sull'intera immagine.

### *Distribuzione dei punti sull'immagine e nello spazio*

Anche se nella scelta dei punti omologhi, non vi sono delle regole precise da seguire, è bene adottare una certa sistematicità nella loro scelta, cercando di individuare come primo punto, quello al centro dell'immagine; altri quattro punti dovrebbero individuarsi sulle estremità dell'immagine, in modo da ricoprire (tenere sotto controllo, espandere la ricerca dei parametri di calcolo su tutta l'area dell'immagine) tutta l'immagine. Infine, i restanti tre punti si dovrebbero selezionare su elementi appartenenti ai diversi piani di profondità nelle tre direzioni. La necessità di ricercare i punti omologhi su tutta l'immagine e sui diversi piani di profondità, deriva sostanzialmente da due motivi, entrambi legati alla precisione metrica della fase di calibrazione ed orientamento esterno

- nei parametri di orientamento interno di una camera, sono presenti le distorsioni radiali, che si evidenziano e sono direttamente legate all'allontanamento dei punti su una foto, rispetto all'asse ottico; in questa fase per effettuare un maggior controllo su tali deformazioni, è necessario introdurre nel sistema di equazioni di collinearità, i punti che si trovano all'estremità dell'immagine, caratterizzati da deformazioni maggiori.
- La distribuzione dei punti, su differenti piani di profondità nello spazio, e quindi in più direzioni, migliora la stima della distanza focale, soprattutto nei casi in cui vi è una profondità poco accentuata, come possono essere le facciate degli edifici.

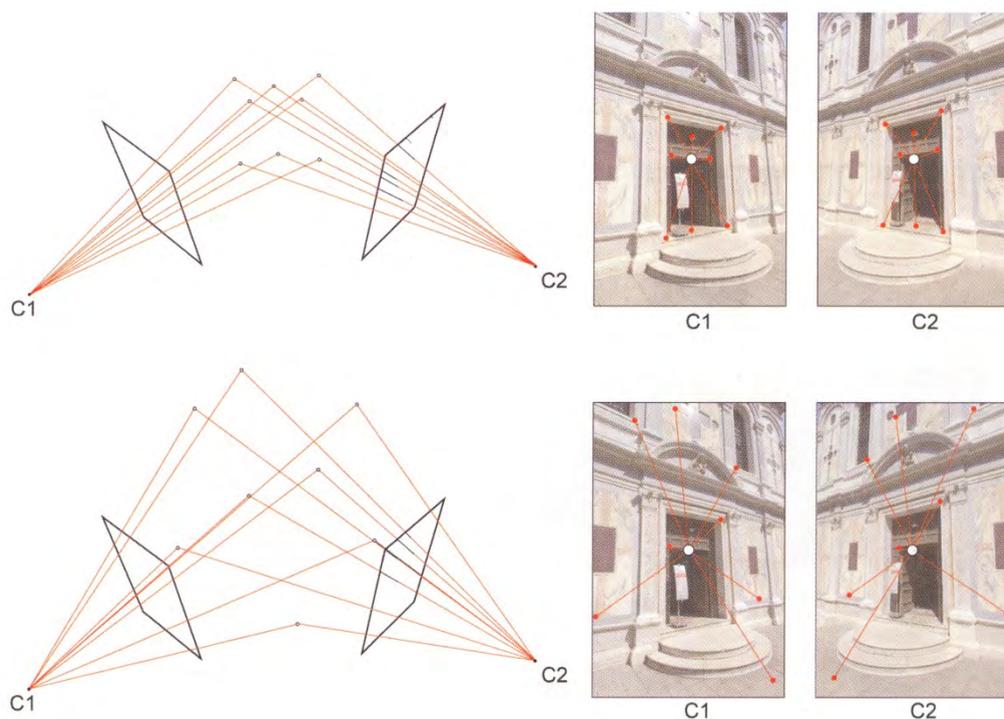


Fig. 3.17 – *Distribuzione dei punti sull'immagine* (L. DE LUCA, 2011).

In fase di accoppiamento 2D/2D, quindi di selezione di punti omologhi, altri tipi di informazioni possono fornire ulteriori indicazioni, per determinare la geometria della scena e migliorarne la qualità. Le indicazioni aggiuntive vengono espresse sotto forma di vincoli geometrici, associabili ai punti omologhi già selezionati. Alcuni di questi vincoli dipendono dalla disposizione dei punti sulla scena; si possono ad esempio individuare vincoli di complanarità – quando un certo numero di punti caratteristici (min. 3) appartengono ad un medesimo piano – ed assegnare ai punti lo stesso valore di coordinata, per una determinata direzione  $(x, y, z)$ . Altri vincoli sono quelli di ortogonalità, o perpendicolarità nello spazio, quando, ad esempio una terna di punti omologhi forma un angolo di  $90^\circ$ . Ancora, altri tipi di informazioni, che possono assegnarsi, sotto forma di vincoli, per condizionare la geometria della scena, sono le coordinate di alcuni punti omologhi, ottenute preventivamente con l'uso di strumenti topografici; in questo caso l'inserimento di vincoli metrici, comporta la restituzione dei parametri delle camere in un sistema di

riferimento solidale con quello usato per il rilievo topografico.

### 3.5 I modi per calibrare ed orientare le camere

Nella fase di acquisizione fotografica, si sono scattate le foto, impiegando differenti metodi di ripresa (assi convergenti, assi paralleli e riprese panoramiche), in funzione della complessità dell'oggetto, della scala di restituzione, di ciò che si vuole restituire e delle condizioni intrinseche dell'oggetto, (dimensioni e contesto nel quale si trova l'edificio in un centro urbano).

La scelta della tecnica più appropriata, in fase di calibrazione ed orientamento delle camere per la ricostruzione e la determinazione dei punti dell'oggetto nello spazio 3D, dipende dal numero di foto scattate e dalle differenti tecniche di ripresa adottate. Generalmente si inizia con l'individuare i punti omologhi, su una coppia o su un gruppo di foto riprese ad assi convergenti.

Tale processo richiede alcune condizioni: per determinare una prima soluzione di calibrazione ed orientamento – se non si conoscono a priori i parametri intrinseci della camera impiegata nella ripresa, per essere poi associati ai rispettivi fotogrammi – occorre che la coppia, o il gruppo di foto di partenza, sulle quali si esegue l'accoppiamento 2D/2D (individuazione dei punti omologhi), sia ripresa con la stessa camera e lo stesso obiettivo, mantenendo costante la lunghezza focale (o livello di *zoom* se si impiegano macchine compatte) e lo stesso orientamento (posizione di ritratto o paesaggio).

Queste condizioni sono necessarie per ridurre il numero di incognite, presenti nel sistema di equazioni che risolve il problema; trovata una prima soluzione di calibrazione ed orientamento delle camere, l'aggiunta di ulteriori corrispondenze (individuazione di altri punti omologhi, oltre agli otto minimi necessari), consente di migliorare la soluzione trovata.

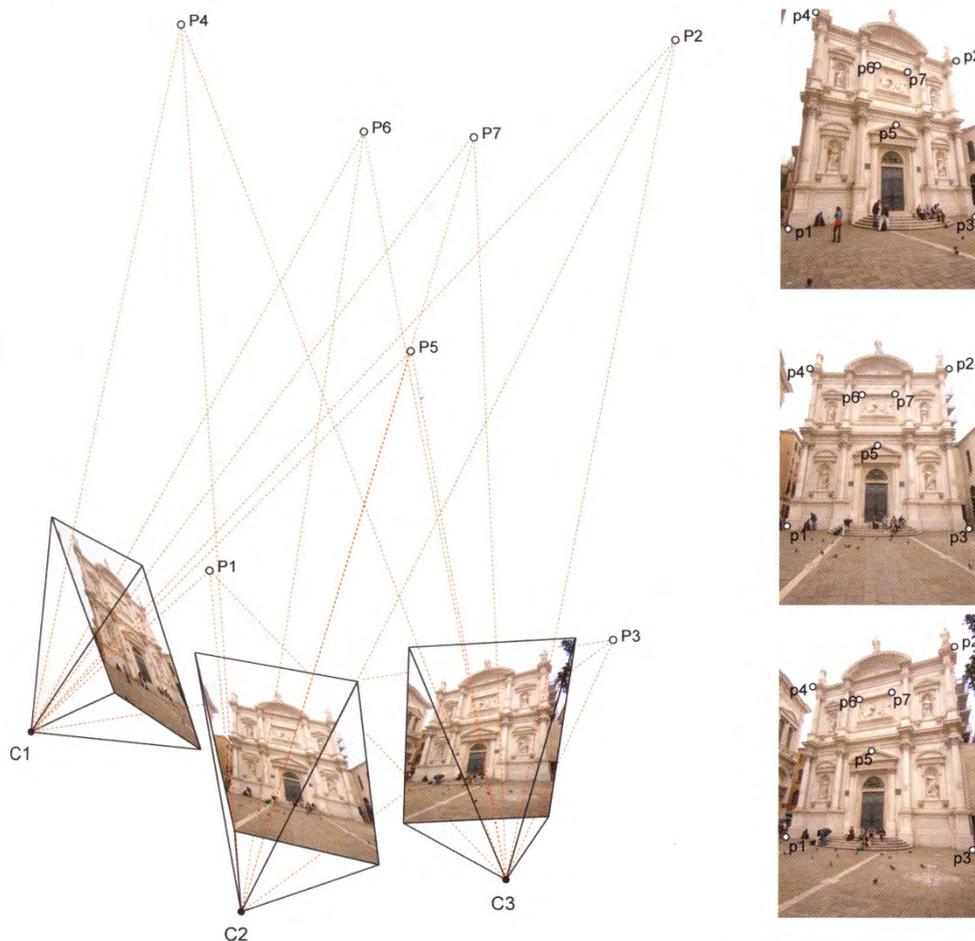


Fig. 3.18 – *Orientamento simultaneo di un blocco di tre immagini ad assi convergenti* (L. DE LUCA, 2011).

### *Calibrazione e orientamento progressivo di immagini convergenti*

La calibrazione si effettua iniziando da due foto e individuando su queste otto punti omologhi, che permettono di ottenere un primo risultato di calibrazione e orientamento, con l'identificazione della posizione dei primi punti nello spazio; alcuni di questi punti, verranno poi proiettati su una terza foto che, con l'aggiunta di ulteriori punti omologhi in comune, verrà orientata

nello stesso sistema di riferimento. Se le foto, sono suddivise in gruppi che presentano uguali distanze focali e orientamento, trovata la prima soluzione per una coppia di foto (o di un gruppo) per le restanti foto, si tratterà solo di trovare l'orientamento esterno della camera, in quanto i parametri intrinseci sono stati calcolati per la coppia di foto iniziali. La scelta della coppia di foto con cui avviare il processo di calibrazione è molto importante, poiché da questa dipenderà l'orientamento dell'intero gruppo di foto; queste dovrebbero avere i requisiti ottimali per permettere una buona calibrazione, cioè dovrebbero consentire una distribuzione dei punti omologhi su tutta l'estensione della foto e su i differenti piani di profondità.

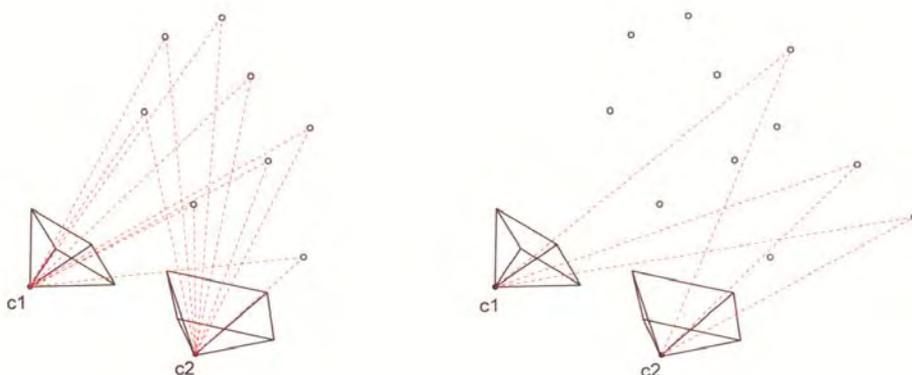


Fig. 3.19 – *Calibrazione ed orientamento progressivo di immagini ad assi convergenti*. Nella fase iniziale si individuano otto punti omologhi sulle prime due foto, successivamente si aggiungono corrispondenze per l'orientamento delle immagini successive (L. DE LUCA, 2011).

Per l'orientamento delle successive foto, come ricordato, bisogna indicare sulla terza foto, la proiezione dei punti già individuati nello spazio, sulla coppia di foto calibrata ed orientata; se alcuni punti individuati nella prima coppia, non fossero visibili sulla terza foto, allora si dovranno aggiungere ulteriori punti, tra la foto calibrata ed la terza da calibrare. Se le foto condividono lo stesso modello geometrico di camera, delle foto già calibrate, allora il numero dei punti da indicare è minore rispetto agli otto richiesti, per una prima calibrazione.

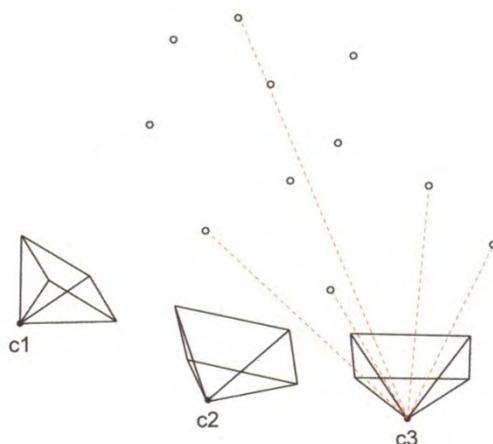


Fig. 3.20 – *Calibrazione ed orientamento progressivo di immagini ad assi convergenti, fase successiva* (L. DE LUCA, 2011).

Sinteticamente per calibrare ed orientare immagini “convergenti”, si inizia da una coppia di immagini e si determina una prima soluzione (che verrà mano mano migliorata, con l’aggiunta delle ulteriori corrispondenze derivanti dalle altre immagini). Si individuano poi gli stessi punti omologhi nella seconda e terza immagine (con eventuale aggiunta di punti, se necessario) per essere poi trasferiti su una quarta e così via.

L’impegno della ripresa ad assi convergenti (quando è possibile), coniugata alla tecnica di calibrazione sopra elencata, è particolarmente utile, nei casi in cui è possibile scattare delle foto girando attorno all’oggetto; in tal caso, nella fase finale, di calibrazione ed orientamento, è buona norma che nelle ultime immagini da calibrare, ci siano dei punti omologhi in comune con le prime immagini già calibrate del gruppo, in questo modo, si va a migliorare la posizione dei punti nello spazio, al tempo stesso, si condiziona la determinazione della geometria della scena, in quanto si ha una compensazione dei valori calcolati sulle sue estremità della maglia dei punti omologhi.

### *Calibrazione ed orientamento di immagini parallele su immagini convergenti*

Le fotografie scattate parallelamente ad un piano, possono essere calibrate

ed orientate, mettendole in relazione con un insieme di punti 3D provenienti da una calibrazione di un blocco di foto (della stessa scena) ripresa ad assi convergenti. Tale procedura si può applicare sia con una sola foto, o con una sequenza scattate in maniera parallela ad un piano dominante; la calibrazione ed orientamento di queste foto, è particolarmente adatta al raddrizzamento di immagini, o all'elaborazione di ortopiani. Le fasi in cui si articola il processo sono due, inizialmente si calibrano ed orientano, un gruppo di foto ad assi convergenti, successivamente si individuano i punti omologhi, considerati nella fase di calibrazione precedente, sulla foto ad assi paralleli (o su ogni fotogramma della sequenza), se tali punti non sono sufficienti, si devono introdurre ulteriori punti, appartenenti ad entrambi i gruppi di foto. Se le fotografie riprese ad assi paralleli, sono scattate con un obiettivo ed orientamento differenti, rispetto al gruppo di foto ad assi convergenti, bisogna stimare sia i parametri intrinseci che quelli estrinseci, il che porta ad aumentare il numero minimo di punti normalmente richiesti.

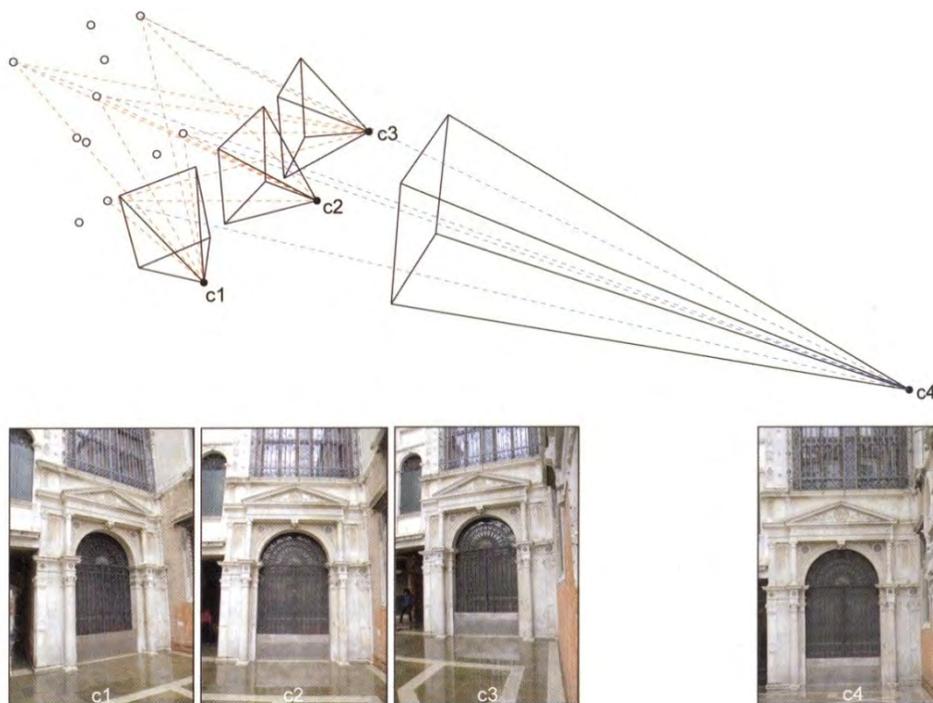


Fig. 3.21 – Calibrazione ed orientamento progressivo di un'immagine parallela su un blocco di immagini convergenti (L. DE LUCA, 2011).

La calibrazione di immagini parallele multiple, segue la stessa metodologia per quelle ad assi convergenti, una volta calibrata ed orientata una immagine, i parametri intrinseci saranno noti, quindi serviranno un numero di punti minore per orientare le altre della sequenza, alla fine per rafforzare e migliorare la fusione delle immagini è bene individuare ulteriori punti nelle zone di sovrapposizione delle stesse immagini.

Sfruttando i punti nello spazio, estratti precedentemente da un gruppo di foto, è possibile calibrare ed orientare, e quindi in un certo senso “attaccare” al modello composto da foto orientate e scalate, ulteriori foto, aventi dei livelli di *zoom* e quindi di dettaglio progressivi, in tal caso si riesce ad avere una copertura dettagliata dell’oggetto da rilevare, utilizzando al meglio la risoluzione della macchina fotografica, potendo osservare ed analizzare ulteriormente la morfologia, per poi rilevarlo e restituirlo correttamente.

## CAPITOLO 4

## LA FOTOGRAMMETRIA

4.1 *Introduzione*

Le tecniche di acquisizione dati, applicate al rilievo dell'architettura, usando i metodi di *Image-Based Modeling and Rendering*, e *Close range photogrammetry*, si occupano, sostanzialmente, di “analisi” d'immagini con il calcolatore; l'analisi è finalizzata a scoprire cosa è presente nella “scena” oggetto d'indagine, e di estrarre determinate proprietà fisiche dell'ambiente visibile, come profondità, forme tridimensionali, contorni e radiometria degli oggetti.

Tralasciando i problemi relativi alla percezione ed al riconoscimento di oggetti, il solo compito di ricostruire la struttura geometrica della “scena” non è un'operazione banale. Tale compito può essere efficacemente descritto come l'inverso della grafica al calcolatore, nella quale i dati (la descrizione geometrica della scena), quindi la forma e le posizioni delle superfici che la compongono, la descrizione radiometrica (posizione ed intensità delle fonti luminose e le proprietà delle superfici); e ancora, fornendo la descrizione completa dell'apparato di acquisizione (fotocamera), il calcolatore produce l'immagine “sintetica” della scena vista dalla fotocamera.

## 4.2 *Close range photogrammetry: fotogrammetria*

La fotogrammetria si basa sul riconoscimento di elementi comuni presenti su più immagini fotografiche (prese da posizioni differenti) di una stessa scena. Le origini di tale tecnica rinviano all'invenzione della fotografia e le basi della sua teoria possono farsi risalire allo sviluppo della prospettiva ed alle considerazioni svolte da Leon Battista Alberti, Piero della Francesca e Leonardo da Vinci.

I principi della fotogrammetria risiedono nella possibilità di ricostruire una prospettiva inversa, ovvero nella possibilità di passare dalla rappresentazione prospettica di una scena, alla stima della posizione nello spazio dei punti che la compongono. Basandosi sul concetto che un punto nello spazio è univocamente determinato con due proiezioni, questa tecnica permette di ricostruire la geometria delle strutture architettoniche attraverso la proiezione, in uno spazio tridimensionale, dei punti generati dal calcolo dell'intersezione delle linee ottiche provenienti da almeno due foto.

Dal punto di vista operativo, l'applicazione fotogrammetrica prevede di trovare una relazione univoca tra l'oggetto da rilevare (spazio 3D) e le foto dell'oggetto stesso (spazio 2D), ovvero prevede la corrispondenza di singoli punti, presenti sull'oggetto reale con gli stessi punti presenti sulle foto. Ottenuta tale corrispondenza, sarà poi possibile relazionare qualsiasi punto dello spazio immagine, con il corrispondente punto dello spazio oggetto e, quindi, ricavare le informazioni per la ricostruzione geometrica dell'oggetto stesso. In tale processo, le grandezze che entrano in gioco sono: le coordinate 3D dell'oggetto, le coordinate 2D dello spazio immagine ed i parametri di orientamento, che relazionano le coordinate 3D dell'oggetto con quelle 2D delle foto; questi esprimono le "caratteristiche" del mezzo fotografico e il suo posizionamento nel momento delle varie prese. Tra lo spazio 2D del fotogramma e l'oggetto da rilevare esiste una relazione di proiettività, quando è possibile trovare una matrice di parametri  $A$ , che funge da operatore di trasformazione, tale che individuato un generico punto  $p$ , sullo spazio 3D, esiste un corrispondente punto  $p'$  nello spazio 2D; tale relazione può essere schematizzata con la relazione

$$a=A*u$$

dove  $a$  è una matrice colonna bidimensionale, delle coordinate dei punti dello spazio immagine;  $A$  è la matrice di trasformazione, contenente tutti i parametri di orientamento ed  $u$  è la matrice colonna, con le coordinate tridimensionali dello spazio oggetto.

Il sistema così strutturato può essere esplicitato in un sistema di equazioni, da cui derivano le equazioni delle rette proiettive, o di collinearità, che rappresentano le equazioni fondamentali della fotogrammetria:

$$x_{p'} = x_h - \Delta x - c \frac{r_{11}(X - X_0) + r_{21}(Y - Y_0) + r_{31}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)}$$

$$y_{p'} = y_h - \Delta y - c \frac{r_{12}(X - X_0) + r_{22}(Y - Y_0) + r_{32}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)}$$

Da tali equazioni, si deduce che, dati la posizione del punto  $p$  nello spazio oggetto e i parametri di orientamento, si risale in maniera univoca alle coordinate del punto  $p'$  sul piano immagine.

I parametri di orientamento, sono costituiti: dalle coordinate del centro di proiezione ( $X_0, Y_0, Z_0$ ); dalle coordinate del punto principale nello spazio 2D ( $X_h, Y_h$ ); dalla distanza principale ( $c$ ), che rappresenta la lunghezza focale effettiva; dai parametri di distorsione radiale, tangenziale ed affine ( $\Delta_x, \Delta_y$ ).

In fotogrammetria quello che interessa è il processo inverso, ovvero partendo dalle coordinate sul fotogramma, devono determinarsi le dimensioni dell'oggetto reale.

Considerando le due equazioni di collinearità, partendo dal punto  $p'$  ( $x_i, y_i$ ), noti tutti i parametri della trasformazione, rimangono incognite le tre coordinate ( $X_i, Y_i, Z_i$ ), dello stesso punto  $p$  sull'oggetto, che devono essere determinate con due equazioni. Ci sono pertanto infinite possibili soluzioni del punto  $p$ , che soddisfano le equazioni di collinearità, che è proprio ciò che esprime il termine collinearità, in quanto tutti i punti nello spazio allineati (o collineari) con la retta che congiunge il punto  $p$  con il centro di proiezione, forniscono sempre la stessa proiezione sul piano immagine. Per risolvere il problema ed identificare in modo univoco le coordinate del punto  $p$ , è

necessario introdurre un secondo punto di vista, ossia un secondo fotogramma sul quale considerare lo stesso punto ( $p''$ ); in questo caso le equazioni di collinearità diventano quattro (due per ogni punto), con tre incognite ed il sistema ammette una soluzione.

Da quanto detto, emerge che da una sola immagine, rappresentante lo spazio oggetto e relazionato ad esso tramite i parametri di orientamento, non è possibile ricavare misure, se non imponendo particolari vincoli, riguardanti la forma dell'oggetto; nel caso generale, per ottenere le misure nello spazio è necessario considerare più foto dell'oggetto (almeno due), prese da differenti punti di vista.

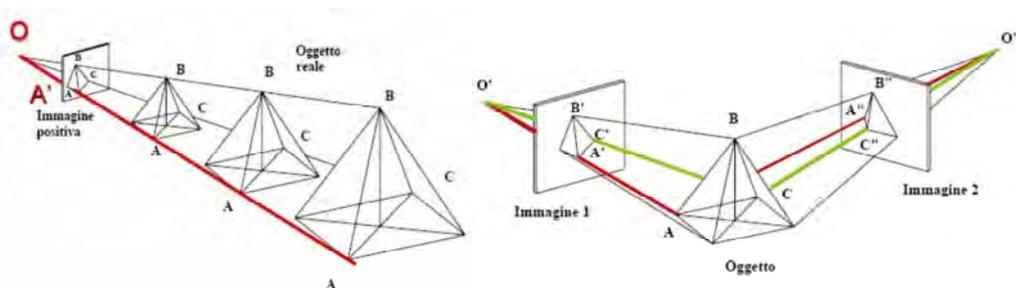


Fig. 4.1 – Da una sola immagine 2d possono essere ricavati numerosi modelli 3d, questo ci dice che una sola fotografia non basta; dall'utilizzo di due fotogrammi si può ricomporre l'oggetto originale in 3 dimensioni.

### *Parametri di orientamento interno*

Per la formulazione delle equazioni di collinearità, si considera una fotocamera ideale (priva di deformazioni), in grado di fornire una proiezione prospettica centrale. Le fotocamere normalmente impiegate nelle riprese fotogrammetriche, forniscono immagini che si discostano da quelle teoriche, per ragioni sia ottiche che meccaniche ed è per questo che è necessaria una procedura di calibrazione. Tramite tale procedura di calibrazione, si determinano i parametri di orientamento sia interno che esterno; i primi riguardano:

- le coordinate del centro principale del sistema immagine ( $X_h, Y_h$ );
- la distanza principale, o lunghezza focale effettiva ( $c$ );
- le distorsioni geometriche indotte dal sistema di lenti ( $\Delta_x, \Delta_y$ ).

I parametri di orientamento interno sono inerenti la posizione dell'elemento fotosensibile, rispetto all'asse ottico (pellicola fotografica, sensore CCD o CMOS) e la formazione dell'immagine su quest'ultimo: l'elemento fotosensibile, inserito meccanicamente nella fotocamera, può essere soggetto ad errori di montaggio (il punto principale non sta esattamente sulla normale al piano dell'immagine, passante per il centro di proiezione prospettica); anche la formazione del fascio di luce, attraverso le lenti che compongono l'obiettivo, non è mai privo di difetti (aberrazioni che influiscono sulla qualità dell'immagine, sulla geometria e quindi sulle misure).

### *Parametri di orientamento esterno*

La posizione e l'orientamento della camera, nel momento della presa, costituiscono i parametri di orientamento esterno, che sono costituiti da:

- coordinate del centro di proiezione, nel sistema esterno ( $X_0, Y_0, Z_0$ );
- assetto angolare della camera nel sistema esterno ( $\Omega, \Phi, K$ ) che definiscono la matrice di rotazione 3D ( $R$ ).

L'assetto della camera, nel momento della presa, non può essere facilmente misurato, quindi anche i parametri di orientamento esterno sono ignoti e sono valutati alla stregua dei parametri di orientamento interno. Nella procedura di calibrazione devono, dunque, determinarsi nove parametri di orientamento interno e sei di orientamento esterno, per un totale di quindici incognite.

Per eseguire la calibrazione, si deve disporre di un numero di equazioni pari, o superiore al numero d'incognite (15), quindi sono necessari almeno otto punti, per scrivere sedici equazioni di collinearità. Nella pratica correntemente si utilizzano dai 20-30 punti, uniformemente distribuiti sull'intera area del fotogramma. Tale procedura, non può essere eseguita direttamente sull'oggetto da rilevare, poiché serve una griglia di calibrazione, costituita da un insieme di punti distribuiti nello spazio, di cui sono note le coordinate, in un sistema cartesiano tridimensionale. Questi punti possono essere individuati da diverse forme geometriche bidimensionali o tridimensionali (*target*).

Il processo di calibrazione si attua con una procedura di "minimizzazione" dello scostamento, tra la soluzione approssimata e quella esatta, partendo da

una prima grossolana approssimazione. Geometricamente si può pensare ad un aggiustamento progressivo di un fascio di raggi proiettanti sull'area dell'immagine (*bundle adjustment*, o regolazione del fascio). Considerando che i parametri di orientamento interno, sono influenzati soprattutto dall'ambiente reale (condizioni climatiche, luce, ecc.), la calibrazione della camera andrebbe fatta ogni volta che si procede ad un nuovo rilievo, calibrando i diversi obiettivi.

### *Operazione di resezione*

Noti i parametri di orientamento interno, si deve ancora determinare la posizione e l'orientamento della camera (orientamento esterno), tramite la conoscenza delle coordinate di punti omologhi su almeno due foto. Per individuare e riconoscere tali punti, spesso sulla scena vengono posti dei *target* di riferimento, che sono ben riconoscibili nelle foto (7/8 punti possono essere sufficienti). Tali punti di riferimento vengono acquisiti con una stazione totale, o misurati direttamente. Note le coordinate di questi punti, si possono scrivere per ciascuno di essi, le equazioni di collinearità e ricavare gli orientamenti esterni, delle camere nelle varie prese.

Eseguita l'operazione di resezione, cioè determinati la posizione e l'orientamento delle camere, può iniziare la fase di restituzione dell'oggetto, mediante le misure (manuali, semi-automatiche o automatiche) dei punti omologhi tra le immagini.

La posizione nello spazio dei punti d'interesse, diversi dai punti di controllo, sono determinate attraverso le equazioni di collinearità, con delle procedure interattive-manuali, quando la geometria dell'oggetto da rilevare può essere discretizzata con pochi punti (strutture architettoniche semplici), oppure con procedure semi-automatiche o automatiche, quando si devono restituire superfici irregolari e complesse (spazi architettonici irregolari, statue, ...); con tali procedure si ottengono “nuvole” di migliaia di punti 3D.

### Matching

Le procedure automatiche o semi-automatiche di correlazione d'immagini, definiscono il processo di *image-matching* o anche di *automatic stereo matching*. Nei pacchetti *software* in uso, come Photomodeler della Eos Systems o ShapeCapture della ShapeQuest Inc e altri, per il processo di *image-matching* impiegano due tipi di *matching*: l'*area-based matching* e il *feature-based matching*. L'algoritmo per l'*area-based matching* considera come entità di *matching* (oggetto a confronto nelle due immagini), i livelli di grigio di due piccole aree messe a confronto; il metodo *feature-based matching* considera come entità da confrontare, elementi caratteristici (linee, spigoli, punti) delle immagini.

I passi fondamentali del processo sono: selezionare l'entità per il *matching* in una delle immagini; trovare l'entità corrispondente nell'altra immagine; determinare la posizione nello spazio dell'entità tramite la collinearità e per ultimo verificare la qualità del *matching*.

Nei *software* tali procedure sono completamente automatiche e danno notevoli vantaggi, soprattutto per il numero di fotogrammi che si possono utilizzare e per l'ottenimento di un elevato numero di "*smartpoint*", ovvero punti omologhi. Al fine di facilitare gli algoritmi di calcolo, per l'ottenimento degli *smartpoint*, si possono apportare delle modifiche alle foto (prima di processarle) per aumentare il contrasto, senza comprometterne la qualità. E' è bene ricordare, infatti, che le stesse foto ricoprono anche la funzione di *texture* e che un'eccessiva alterazione potrebbe incidere sulla verosimiglianza del modello virtuale.

### 4.3 Parametri di distorsione affine, radiale e tangenziale

Tra i molteplici difetti ottici che producono un deterioramento dell'immagine, quelli che si possono correggere per via analitica e quindi che entrano a far parte delle equazioni di collinearità, sono le distorsioni affini ( $B_1$ ,

$B_2$ ), le distorsioni tangenziali ( $K_1$ ,  $K_2$ ,  $K_3$ , il parametro  $K_3$  non è sempre utilizzato) e distorsioni tangenziali ( $P_1$ ,  $P_2$ ).

### *Distorsione provocata dalle lenti dell'obiettivo*

Il principio fisico su cui si basa una lente è quello della rifrazione, per cui se una radiazione luminosa attraversa l'interfaccia tra due mezzi caratterizzati da velocità di propagazione diverse, l'angolo di entrata e di uscita del raggio luminoso sono diversi.

Si pensi all'effetto di un cucchiaio immerso in un bicchiere d'acqua che appare spezzato; questo fenomeno accade perché la velocità di propagazione della luce nell'acqua è minore di quella nell'aria. Ogni materiale è caratterizzato da un indice di rifrazione. Immaginando di avere un sistema aria-lente-aria, se un raggio di luce attraversa una lastra di vetro obliquamente, il raggio è prima deviato in un senso e poi deviato di nuovo in senso opposto. Quindi se il raggio di luce, dopo essere penetrato nel mezzo rifrangente, emerge di nuovo, esso verrà nuovamente rifratto, secondo un angolo uguale a quello con cui era stato deviato la prima volta. Se una lente fosse costruita da una semplice lastra di vetro, ci sarebbe una semplice traslazione dei raggi luminosi, ma la focalizzazione non si otterrebbe.

### *Problema della focalizzazione dei raggi luminosi*

Per ottenere l'effetto di "concentrazione" dei raggi luminosi in uno stesso piano focale la lente deve assumere un'opportuna forma curva. La presenza di due normali diverse in entrata e in uscita, implica una deviazione complessiva del raggio luminoso che varia a seconda del punto in cui incide sulla lente. Se la forma della lente è corretta, tutti i raggi provenienti dallo stesso oggetto sono convogliati nello stesso punto, cioè focalizzati.

## Aberrazioni

La lente deve essere costruita con metodi in sostanza realizzabili. Questo provoca un allontanamento delle caratteristiche della lente da quelle previste teoricamente, il che comporta una sfocatura dell'immagine. Le cause di tali imperfezioni sono dette aberrazioni e si distinguono: aberrazioni cromatiche, aberrazioni sferiche e coma, astigmatismo

### *Aberrazioni cromatiche*

In figura 4.2 è rappresentata una focalizzazione perfetta valida per una particolare lunghezza d'onda. Poiché la velocità della luce, e quindi anche l'indice di rifrazione, varia con la lunghezza d'onda, in genere avremo una focalizzazione su piani diversi delle componenti luminose caratterizzati da lunghezze d'onda differenti. Supponiamo di usare una luce bianca, questa è composta di tutte le possibili lunghezze d'onda tra 380 e 770 nm. Per le componenti estreme dello spettro, vicine rispettivamente al blu (380 nm) e al rosso (770 nm) avremo un diverso piano di focalizzazione per i due colori. L'effetto conseguente è quello denominato *aberrazione cromatica*. Il problema dell'aberrazione cromatica è superato utilizzando una coppia di lenti detta *lente acromatica*.

Si sfrutta la capacità delle lenti divergenti di far deviare le componenti di colore in senso opposto rispetto a quanto fa la lente convergente. In questo modo le due deviazioni si compensano e le componenti cromatiche risultano nuovamente allineate.

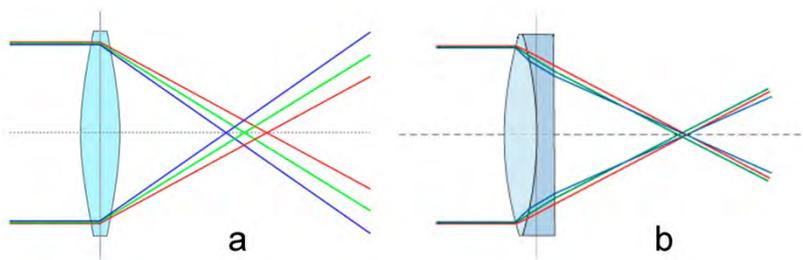


Fig. 4.2 – *Aberazione cromatica: a) lente convergente; b) lente acromatica.*

### *Aberrazioni sferiche e coma*

L'andamento ideale di una lente corrisponde a una curvatura iperbolica. Realizzare una curvatura del genere con macchine utensili è un'operazione molto complessa e assume costi assai elevati. Si preferisce realizzare una lente sferica riducendo i costi, anche se questo va a discapito della correttezza della superficie. Con una lente sferica la focalizzazione dei raggi luminosi è diversa per le zone centrali della lente rispetto a quelle periferiche. L'effetto di questi difetti consiste in un'immagine sfumata dovuta al fatto che il fuoco dei raggi cosiddetti parassiali, cioè in prossimità dell'asse ottico, è in posizione diversa rispetto a quelli marginali. Se in una lente con aberrazione si inviano raggi luminosi obliqui, evidenzia il fenomeno detto “coma”.

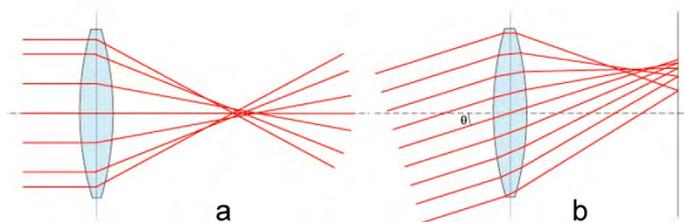


Fig. 4.3 – *Difetti geometrici di una lente: a) aberrazione sferica, i raggi luminosi più distanti dall'asse ottico tendono a focalizzarsi più vicino, formando un alone; b) “coma”, i raggi luminosi paralleli, obliqui all'asse ottico, tendono a focalizzarsi ad altezze distinte sul piano focale, formando una sequenza di aloni.*

### *Astigmatismo*

L'astigmatismo consiste nell'allungamento dell'immagine lungo una specifica direzione: un oggetto circolare e così visto attraverso un sistema astigmatico come ellittico. L'astigmatismo affligge soprattutto lenti di piccola apertura.

### *Curvatura di campo*

Poiché la lente è una calotta sferica, i diversi punti-immagine non si focalizzano su una superficie piana, ma su una superficie sferica, detta superficie di Petzval. Questo difetto prende il nome di curvatura di campo. La somma di Petzval è una grandezza che indica il raggio di curvatura di questa superficie sferica, che può essere concava o convessa, a seconda che la somma di Petzval sia positiva o negativa. I normali obiettivi anastigmatici sono caratterizzati da una somma di Petzval positiva: la calotta sferica appare concava. Al contrario in certi schemi a teleobiettivo la somma di Petzval, negativa, rende convessa la superficie sferica. Non potendo incurvare la pellicola, il progettista di lenti fa in modo che la somma di Petzval risulti il più possibile vicino allo zero.

### *Distorsione Radiale*

Le distorsioni non alterano la nitidezza dell'immagine, ma il suo aspetto. Esistono due tipi di distorsione: a cuscinetto e a barilotto. La prima è tipica degli obiettivi grandangolari, mentre quella a cuscinetto si riscontra nei teleobiettivi. La deformazione dell'immagine incide sulla stima della posizione dei punti sul piano di proiezione e influenzerebbe direttamente la qualità della misura fotogrammetrica, ma fortunatamente il suo andamento può essere rappresentato con una funzione. Per la simmetria sferica della lente la distorsione ha una simmetria radiale e per questo si indica con distorsione radiale.

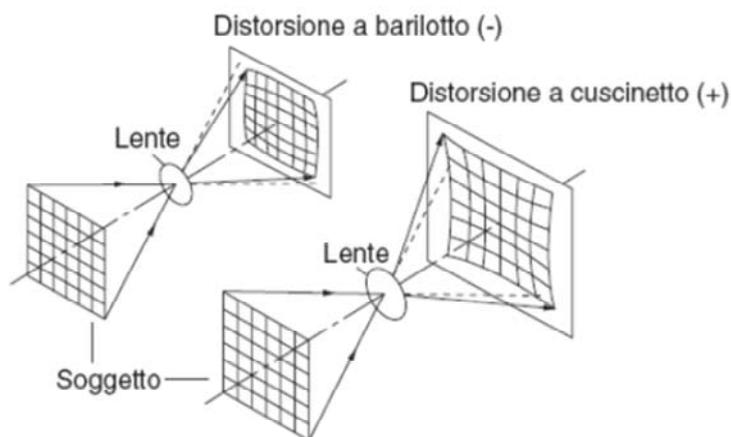


Fig. 4.4 – Rappresentazione distorsioni a barilotto e a cuscinetto.

Può essere rappresentata convenientemente in funzione della distanza dal centro ottico del sistema, che abbiamo già definito come principale  $(x'_H, y'_H)$ . Indicando con  $r$  tale distanza, avremo:

$$r = \sqrt{(x - x_H)^2 + (y - y_H)^2}$$

$$r_r = r + \Delta r_r$$

$r_r$  : Posizione radiale cui si focalizza effettivamente il punto generico che nell'immagine ideale sarebbe a una distanza  $r$

$\Delta r_r$  : Deviazione del raggio luminoso sul piano di proiezione rispetto alla posizione che avrebbe nel caso ideale

$$\Delta r_r = K_1 r^3 + K_2 r^5 + K_3 r^7$$

$K_1, K_2, K_3$  : Coefficienti dipendenti dalla distanza di messa a fuoco della macchina. In genere il contributo più significativo è quello del terzo ordine  $K_1$  mentre quelli di ordine superiore danno un aggiustamento via via meno importante.

### *Distorsione tangenziale*

È lo spostamento di un punto sull'immagine causato dal disallineamento degli elementi della lente. Anche in questo caso, dato un punto e la sua proiezione ideale  $(x, y)$  avremo una versione delle stesse coordinate alterata

dalla distorsione tangenziale:

$$x_t = x + \Delta x_t$$

$$y_t = y + \Delta y_t$$

Questo spostamento è descritto da due polinomi, uno lungo l'asse x e l'altro per lo spostamento lungo l'asse y:

$$\Delta x_t = P_1[r^2 + 2(x - x_H)^2] + 2P_2(x - x_H)(y - y_H)$$

$$\Delta y_t = P_2[r^2 + 2(y - y_H)^2] + 2P_1(x - x_H)(y - y_H)$$

$P_1, P_2$  sono coefficienti che dipendono dalla messa a fuoco della camera

### *Distorsione affine*

Per le macchine fotografiche digitali esiste un'altra causa di incertezza: il piano del sensore può non essere esattamente ortogonale rispetto all'asse ottico dell'obiettivo, il che implica che i pixel sul sensore, sebbene quadrati interessano dimensioni differenti nelle due direzioni x e y. Questa incertezza è identificata come distorsione affine e modellata dalle relazioni:

$$\Delta x_a = B_1(y - y_H)$$

$$\Delta y_a = B_2(x - x_H)$$

$B_1$  e  $B_2$  sono due coefficienti che dipendono dalla forma degli elementi del sensore della camera digitale, mentre gli altri parametri hanno lo stesso significato detto prima.

Nell'equazione di collinearità per tenere conto delle varie forme di distorsione, i termini  $\Delta x$  e  $\Delta y$  contengono i contributi relativi alle tre possibili forme di distorsione radiale tangenziale ed affine:

$$\Delta x = \Delta x_r + \Delta x_t + \Delta x_a$$

$$\Delta y = \Delta y_r + \Delta y_t + \Delta y_a$$

#### 4.4 *Caratteristiche delle immagini digitali*

Prima di addentrarci nell'argomento è opportuno fare due brevi considerazioni di carattere generale. Una riguarda la corretta interpretazione del termine immagine digitale. La parola immagine nella dizione immagine digitale può far pensare che l'immagine digitale sia di per sé un'immagine analoga a quella di tipo fotografico, con l'unica variante di essere presentata alla nostra vista sul video di un *computer* o mediante qualche altro dispositivo elettronico. In realtà quando osserviamo un'immagine sul video di un computer non osserviamo un'immagine digitale, bensì un'immagine analogica, ne più ne meno come se osservassimo una fotografia. L'unica differenza consiste nel fatto che l'immagine osservata sul video di un computer è materializzata dalla diversa illuminazione dei *pixel* del video, e quindi l'immagine che si crea sulla nostra retina è il risultato di una luce emessa, mentre quando osserviamo una fotografia l'immagine che si forma sulla nostra retina è generata dalla luce riflessa dalla fotografia stessa.

Una vera immagine digitale è in realtà una matrice di numeri che deriva dall'aver tradotto in valori numerici la quantità di energia emessa o riflessa da un oggetto, e che sarebbe in grado di produrre un'immagine percepibile dalla nostra retina o di impressionare una pellicola. Questi valori per essere nuovamente percepiti da noi come immagine nel senso comunemente attribuito al termine, devono subire un procedimento inverso di trasformazione dal numerico all'analogico.

Una seconda considerazione riguarda uno degli aspetti che a volte non vengono subito colti e cioè il fatto che il passaggio dall'analogico al digitale non riguarda solo un'evoluzione puramente tecnologica, ma un'evoluzione per quanto riguarda l'univocità e la conservazione della metricità dell'immagine. In una camera fotografica tradizionale l'immagine della realtà viene proiettata sul piano focale dove è situata la pellicola. La luce riflessa dalla scena ripresa sensibilizza la pellicola; in altre parole le cariche energetiche che costituiscono la luce (i fotoni) colpiscono i granuli di alogenuro d'argento dell'emulsione fotografica e li modificano in modo tale che quando la pellicola subisce il procedimento chimico di sviluppo essi si anneriscono tanto più quanto maggiore è il numero di fotoni che li ha colpiti. Nel momento della presa si forma quindi nell'emulsione della pellicola quella che viene detta immagine

latente, che costruisce una discretizzazione dell'immagine della realtà secondo il tessuto irregolare dei granuli di alogenuro d'argento. L'immagine latente viene ridotta in un rapporto di scala rispetto alla realtà che dipende dalla focale dell'obiettivo e dalla distanza della scena dall'obiettivo stesso e viene eventualmente deformata rispetto a quella che sarebbe una rigorosa prospettiva centrale dall'effetto della distorsione prodotta dall'obiettivo. Una ulteriore deformazione dell'immagine latente può essere generata dalla deformazione, per effetti termici o meccanici, della pellicola negativa prima di aver subito il trattamento di sviluppo.

Nel procedimento fotografico tradizionale vi è quindi la possibilità che la geometria e la radiometria dell'immagine latente vengano alterate a causa della natura del supporto su cui viene registrata e a causa dei trattamenti di sviluppo. Ma anche la conservazione nel tempo può generare in una fotografia tradizionale su pellicola delle deformazioni geometriche.

Quindi possiamo dire che nel procedimento fotografico tradizionale esiste questa mancanza di perfetta corrispondenza tra l'immagine che viene proiettata sul piano focale al momento della presa e la sua materializzazione sul supporto fotografico.

Una delle principali caratteristiche della fotografia digitale è invece quella di catturare l'immagine latente in formato numerico e di trasmetterla e conservarla nel tempo senza alterazioni. In altre parole al momento della presa si determina in modo univoco il rapporto tra la posizione di un punto nello spazio e la sua immagine e questo rapporto rimane immutato nel tempo ed è univoco per chiunque utilizzi l'immagine.

### *Il sensore*

Il sensore è l'elemento che raccoglie l'immagine sul piano focale, cioè un dispositivo elettronico che svolge la funzione di cattura e memorizzazione dell'immagine che veniva svolta dalla tradizionale pellicola. Il sensore fa parte del chip elettronico che è l'elemento fondamentale della camera digitale. Esistono oggi due tipi di chip elettronici che vengono utilizzati dalle camere digitali: il CCD, *charged-coupled device* e il C-MOS ; il CCD è quello da più

tempo utilizzato e molto più largamente usato del C-MOS.

L'invenzione del CCD, *charged-coupled device*, che in italiano significa – *dispositivo ad accoppiamento di carica*–, è avvenuta all'inizio degli anni '70 nei laboratori della Bell. Alcune volte il CCD viene anche detto dispositivo a trasferimento di carica poiché i pixel che lo compongono, oltre a svolgere la funzione di trasformare l'energia della luce incidente su di essi in una carica elettrica, svolgono anche la funzione di trasferirla in modo sequenziale dall'uno all'altro.

Il CCD non ha solo la funzione di catturare l'immagine che si forma sul sensore, ma anche quella di trasformarla in una matrice di numeri e memorizzarla su un supporto elaborabile da un computer.

Il sensore è costituito da una piastrina di silicio, quadrata o rettangolare, sulla quale vi è inciso un fitto grigliato che la suddivide in tanti elementi quadrati: i *pixel*. Un sensore così fatto, cioè costituito un grigliato, quadrato o rettangolare, di tante linee e colonne di pixel viene detto di tipo *a matrice* o, con terminologia anglosassone, di tipo *area array*.

Come è noto il formato classico delle camere fotografiche tradizionali amatoriali è di  $24 \times 36 \text{ mm}^2$ . Il sensore delle camere amatoriali digitali che si pongono in alternativa alle camere a pellicola di questo formato hanno sensori molto più piccoli. Le camere digitali che si pongono in alternativa alle camere  $6 \times 6 \text{ cm}^2$  tendono invece a raggiungere la dimensione dei sensori di  $6 \times 4 \text{ cm}^2$ . I sensori più avanzati, che si trovano anche nei dorsi digitali, hanno dunque dimensioni di  $38,8 \times 50 \text{ mm}^2$  essendo composti da un array di  $4080 \times 5440$  *pixel* (22 megapixel) di  $9 \mu\text{m}$  di lato.

Come si vede nella fig. 4.5 non vi è proporzionalità diretta tra numero di *pixel* e dimensione del sensore; infatti in genere le camere con minor numero di *pixel*, e quindi di prestazioni e prezzo inferiori a quelle con maggior numero di pixel, hanno anche pixel di dimensioni inferiori.

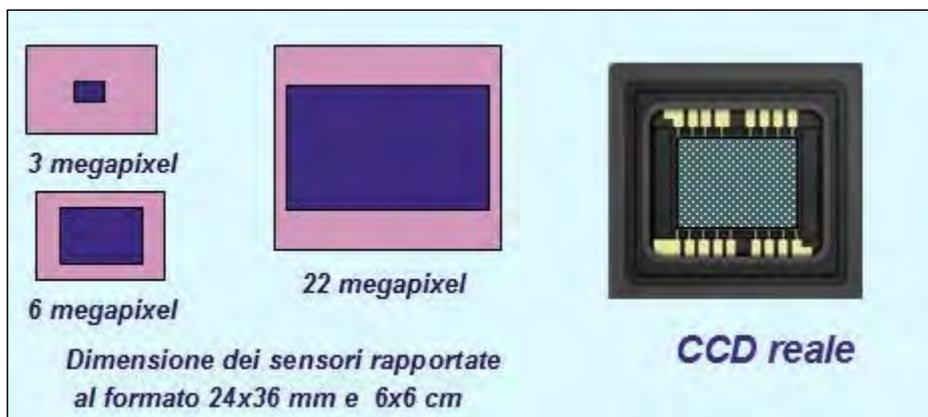


Fig. 4.5 – Schematizzazione dei sensori digitali.

La funzione del CCD è quello di catturare l'energia che cade sugli elementi fotosensibili del grigliato dei pixel del sensore per il tempo di integrazione e di trasformare questa energia in una matrice di valori numerici. Il principio di formazione di un'immagine digitale, può sintetizzarsi nel modo seguente: il sensore misura la quantità di luce incidente, che è convertita sotto forma di tensione elettrica con intensità proporzionale; quest'ultima a sua volta è trasformata da un convertitore analogico/digitale (A/D) in *digital number*, cioè in bit di informazioni.

$$g(x, y) = \begin{bmatrix} g(0,0) & g(0,1) & \dots & g(0, C - 1) \\ g(1,0) & g(1,1) & \vdots & g(1, C - 1) \\ \vdots & \vdots & \ddots & \vdots \\ g(R - 1,0) & g(R - 1,1) & \dots & g(R - 1, C - 1) \end{bmatrix}$$

Un'immagine digitale consiste, dunque, in una matrice  $g(i, j)$  di *pixel* del tipo:

- $i : 0, 1, \dots, C-1$  indice di riga
- $j : 0, 1, \dots, R-1$  indice di colonna
- $R$  : massimo numero di righe
- $C$  : massimo numero di colonne
- $g(x,y) : \{0,1,\dots,max\}$  valori radiometrici.

### Risoluzione radiometrica e geometrica

Il processo di formazione di un'immagine digitale consiste in due operazioni, la quantizzazione e il campionamento. Queste sono direttamente legate alle due caratteristiche fondamentali di questo tipo di immagini: la risoluzione radiometrica che è legata alla quantizzazione, cioè alla conversione dell'intensità del segnale luminoso (trasformata dal sensore in un valore reale di intensità elettrica) in un numero intero, che viene memorizzato in corrispondenza di ciascun *pixel* e la risoluzione geometrica che è legata all'operazione di campionamento, cioè all'area di immagine corrispondente a ciascun elemento del sensore; poiché un'immagine digitale è costituita da una matrice bidimensionale i cui elementi o *pixel* sono i contenitori dell'informazione, un'elevata risoluzione geometrica comporta la generazione di una matrice con un grande numero di righe e colonne, cioè un elevato numero di *pixel*.

Il parametro che indica il numero di *pixel* contenuti in un'unità di lunghezza e quindi la dimensione degli stessi, e la risoluzione geometrica, che solitamente è indicata in DPI, cioè *Dots Per Inch* (punti per pollice).

$$DPI = \frac{n^{\circ} \text{ pixel}}{\text{pollice}}$$

In alternativa se la risoluzione viene espressa come dimensione del *pixel* espressa in  $\mu\text{m}$ . Le espressioni seguenti danno la relazione tra dpi e dimensione del *pixel*:

$$d_{pix} = \frac{25400}{DPI}$$

Direttamente legata alla dimensione del *pixel* e la definizione del sistema di riferimento interno dell'immagine; infatti, se si vuole utilizzare un'immagine digitale per scopi fotogrammetrici, bisogna definire la relazione fra la posizione del *pixel* e il sistema di coordinate immagine.

In Figura 4.6 è mostrato il sistema di coordinate immagine con l'origine spostata di mezzo *pixel* all'esterno della matrice immagine; le coordinate immagine del centro di un *pixel*  $g(i,j)$  si ottengono moltiplicando il numero di indice per la dimensione del *pixel*.

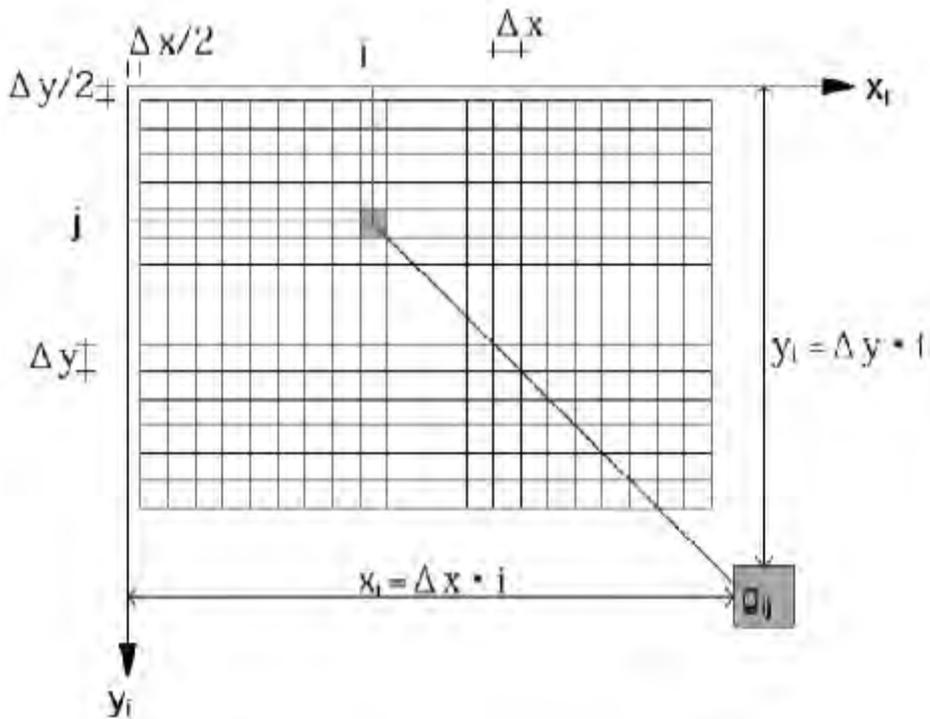


Fig. 4.6 – Corrispondenza tra le coordinate dell'immagine e i pixel immagine.

$$x_i = \Delta x \cdot i$$

$$y_i = \Delta y \cdot j$$

La tradizionale misura di coordinate è sostituita dall'individuazione del *pixel* all'interno della matrice immagine. Tale posizione si traduce in coordinate reali oggetto attraverso le relazioni analitiche della fotogrammetria. Se i *pixel* sono convenientemente piccoli, si assumono gli indici riga e colonna come fossero le coordinate immagine e anche la distanza principale  $c$  può essere espressa in unità  $\Delta x$  e  $\Delta y$ .

### Formazione di immagini digitali a colori

Gli elementi sensibili del sensore di una camera digitale, sono monocromatici; per acquisire una immagine digitale che possa dar luogo una

sua visualizzazione equivalente a una fotografia a colori, è necessario porre sopra ogni *pixel* un CFA (*Color Filter Array*), per separare la luce incidente nelle tre componenti primarie RGB (*red, blu, green* usando la terminologia anglosassone), e registrare l'energia fotonica corrispondente con i tre valori ad essa corrispondenti. La disposizione più usata dei filtri sui fotoni è quella che segue lo schema di *Bayer*, il quale prevede che il 50% dei pixel riceva la luce verde, il 25% la luce rossa e il 25% la luce blu.

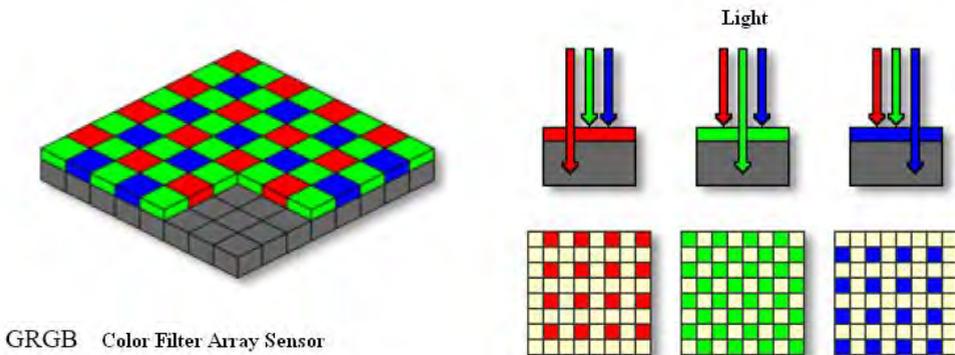


Fig. 4.7 – Filtro tipo Bayer.

Questa distribuzione in percentuali diverse dei filtri dei colori deriva dal fatto che l'occhio umano percepisce il verde come colore predominante per la definizione della nitidezza di un'immagine. Inoltre nelle camere fotografiche digitali terrestri con CCD di formato rettangolare che possono portare a inquadrature col lato lungo del CCD verticale o orizzontale, lo schema di Bayer produce lo stesso tipo di immagine indipendentemente dalla modalità di ripresa, sia essa orizzontale o verticale.

Il principio di funzionamento del filtro è semplice: esso lascia passare al sensore solo una determinata banda di lunghezza d'onda della luce, corrispondente ad un ben determinato colore; in questo modo però solamente un colore viene "misurato" per ogni *pixel* del sensore, perciò la camera deve stimare i due colori mancanti in corrispondenza di ciascun *pixel*, effettuando un'operazione che viene detta *demosaicizzazione*. A questo punto si hanno a disposizione 3 immagini, la *Red*, la *Green* e la *Blue*, che sovrapposte danno origine all'immagine RGB.

#### 4.5 *Structure and motion. Ricostruzione 3D in automatico da fotografie*

L'automazione nel campo della fotogrammetria dei vicini, o fotogrammetria terrestre (*close range photogrammetry*), negli ultimi anni ha visto lo sviluppo di interessanti soluzioni innovative. L'incidenza economica delle risorse umane, richieste per lo svolgimento delle comuni operazioni fotogrammetriche (orientamento, restituzione ed interpretazione) ed i bassi costi di *hardware* e *software*, ha spinto il mercato e la ricerca ad interessarsi maggiormente a soluzioni tecniche più efficienti, in cui dette operazioni (almeno quelle più ripetitive) possano effettuarsi in automatico.

La fotogrammetria, da sempre legata agli aspetti della misura, si è trovata a doversi confrontare con altre discipline (come ad esempio la visione artificiale o *computer vision CV*) in cui tali aspetti sono parzialmente trascurati, a favore di soluzioni automatiche in *real-time*.

Oggi, la tendenza è quella di cercare una sempre maggior integrazione dei due aspetti, in modo da avere sistemi di misura affidabili dal punto di vista metrico, ma con elevate capacità di automazione, prescindendo, a volte, dalla necessità di avere una soluzione in tempo reale.

Una delle fasi più delicate del rilievo fotogrammetrico, è l'orientamento dei fotogrammi. Se nella fotogrammetria aerea, la sequenza delle immagini ha una struttura ben definita, che deriva da tecniche e metodologie sperimentate in quasi un secolo di applicazioni, lo stesso non si può dire per le applicazioni terrestri: ciascuna applicazione presenta, infatti, i suoi aspetti singolari sia in termini di risultati attesi, che in termini di vincoli operativi per la progettazione e realizzazione delle riprese fotografiche.

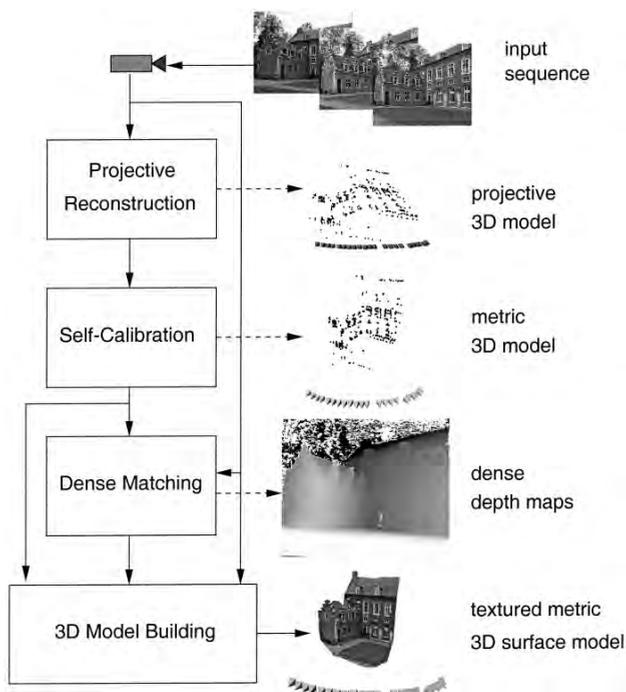


Fig. 4.8 – *Workflow Structure and Motion*, panoramica del sistema; le camere sono rappresentate da piccole piramidi, i risultati delle corrispondenze dense sono accumulati in “mappe di profondità” ed infine tutti i risultati sono integrati in una ricostruzione 3D, con texture della scena in esame (M. POLLEFEYS et al., 2000).

All’inizio degli anni ’90, con il diffondersi delle prime telecamere digitali nel settore della visione artificiale, l’attenzione si è concentrata sull’analisi di sequenze di immagini, affrontando la ricostruzione della *Structure from Motion*, ossia la ricostruzione della struttura di una scena a partire da una sequenza di fotogrammi (Fig. 4.8). Attualmente tutti i metodi di risoluzione della *Structure from Motion* si basano su metodologie analoghe, che prevedono la ricerca di punti da correlare nelle varie immagini; inizialmente si estraggono dai diversi fotogrammi una serie di elementi *features* (punti, spigoli, linee etc.) che, con un certo grado di probabilità siano riscontrabili su diversi fotogrammi, per mezzo di opportuni “operatori di interesse” (elementi ad alta riconoscibilità). Una volta individuato un numero sufficientemente elevato di elementi “interessanti”, per mezzo di opportuni criteri geometrici e radiometrici, si individuano delle corrispondenze “apparenti”, che formino un

primo tentativo di catalogazione di corrispondenze omologhe sui diversi fotogrammi.

A questo punto è possibile stimare una geometria di presa preliminare ed eliminare le corrispondenze sbagliate. Al termine delle stime la geometria individuata non verifica le condizioni reali dei fotogrammi, ma fornisce una geometria proiettivamente equivalente a quella reale. Per ottenere una ricostruzione reale dei fotogrammi e della scena, è necessaria la conoscenza dei parametri di orientamento interno. Anche se i parametri intrinseci della camera non sono noti, si possono imporre alcune condizioni, come ad esempio la forma del *pixel* rettangolare o quadrata, il rapporto d'aspetto costante in tutta la sequenza dei fotogrammi, il punto principale della camera al centro dell'immagine, etc. Imponendo tali condizioni si può eseguire la calibrazione della camera. La calibrazione facilita la ricerca dei punti corrispondenti e permette di trovare riscontri per la maggior parte dei *pixel* delle immagini.

Infine, si ricava un modello 3D ottenuto dalle “mappe di profondità”.

Per determinare una prima serie di elementi geometrici sulle immagini, possono utilizzarsi opportuni “operatori di interesse” (operatori matematici). Un operatore di interesse è un particolare operatore matematico che permette di mettere in luce elementi dell'immagine ben definiti rispetto al loro intorno e, di conseguenza, ben riconoscibili. Tali elementi sono generalmente di tipo puntuale. Il criterio fondamentale per valutare la reale efficienza di un operatore di interesse è rappresentato dalla sua capacità di individuare il medesimo particolare (ovvero di fornire punti omologhi) anche su fotogrammi differenti per scala, o per punto di vista, o per condizioni di illuminazione, etc. In altre parole un operatore di interesse è tanto più efficiente quanto maggiore è il suo grado di invarianza, rispetto a trasformazioni geometriche e radiometriche via via più generali.

Gli algoritmi *Matching*, si possono classificare in due categorie principali: algoritmi di tipo *Features-Area-Based* e algoritmi di tipo denso (*dense stereo*). I primi, danno informazioni su un limitato numero di punti ben riconoscibili tramite *Features*, ottenendo così delle mappe denominate “mappe di disparità”; gli algoritmi di tipo denso, mirano a generare mappe di disparità per ogni punto dell'immagine. Le tecniche più semplici di correlazione, cioè di *Matching* sono quelle che utilizzano la funzione *Area Based Matching* e *Features Based*

*Matching*: utilizzando *Area Based Matching*, le corrispondenze sono stabilite in funzione del grado di similitudine dei toni di grigio nell'intorno del punto stesso. Partendo da un punto noto, su una finestra rettangolare, sull'immagine di riferimento *template* (fotogramma di riferimento, che contiene i punti di cui si vuole stimare la posizione omologa su un altro fotogramma) s'individua la posizione più probabile del suo omologo su una finestra di uguali dimensioni sull'immagine *patch* (fotogramma di servizio, su cui si cercano le corrispondenze).

Con *Features Based Matching*, dopo aver estratto i punti ben riconoscibili sui vari fotogrammi, l'algoritmo di *matching* si preoccupa di individuare le corrispondenze.

Considerando l'automatismo di tali procedure – che permettono l'ottenimento di una densa nuvola di punti – è quindi il modello di un oggetto o di una scena reale, diversi enti di ricerca come Dipartimenti universitari, Centri di ricerca, Istituzioni del patrimonio, come i musei o Agenzie del patrimonio nazionale e le Imprese commerciali, che operano nel campo Beni Culturali (come la Rete europea di eccellenza EPOCH, una rete di un centinaio di istituzioni culturali europee che si interessano di tecnologie dell'informazione e della comunicazione per i Beni Culturali), da qualche tempo s'interessano alla ricostruzione 3D. Due partner di tale rete – precisamente ESAT-PSI Lab of K. U. Leuven (Belgio) e il Visual Computing Lab-CNR-ISTI (Italia) – hanno realizzato una *pipeline* di ricostruzione 3D, impiegando programmi *open source*, da utilizzare nel campo Beni Culturali.

L'idea di fondo è che una fotocamera digitale e una connessione a Internet sono sufficienti ad un utente per ricostruire una scena in 3D.

I programmi creati e impiegati sono “ARC3D”, prodotto da ESAT-PSI Lab of K. U. Leuven e “MeshLab”, del Visual Computing Lab-CNR-ISTI:

- ARC3D è un *web-service* e consente la ricostruzione 3D da immagini con tecniche di visione artificiale; tale *service* permette di calcolare per ogni foto, con vari algoritmi di correlazione automatica, una mappa di “profondità” e di “qualità” delle corrispondenze trovate.
- MeshLab è un *software* che consente di importare ed elaborare le mappe di profondità, al fine di ottenere un modello 3D.

Con l'uso combinato di ARC3D e di MeshLab, è possibile acquisire

digitalmente manufatti, monumenti e spazi urbani.

### ARC3D

I passaggi fondamentali per ricavare il modello 3D, in ARC3D sono semplici: si fotografa l'oggetto; si inviano le immagini al *server* e, dopo la segnalazione dell'avvenuta ricostruzione, si effettua il *download* dei *file* processati. L'unica accortezza pratica è verificare che non ci siano immagini sfocate, che corromperebbero la qualità del modello finale.

L'utente, una volta autenticato, può trasferire le fotografie sul *server* ESAT, dove un processo completamente automatico elabora e calcola le correlazioni 3D. Tali processi sono eseguiti su un insieme di computer connessi tra loro tramite una rete telematica (*cluster*) e quando il *server* ha terminato l'elaborazione, l'utente riceve un'*email* e può scaricare i risultati.

Lo stesso programma permette di visualizzare il modello prodotto e crea due formati: uno (.obj) in cui è registrata solo la *mesh* e la *texture* generata dall'algoritmo usato dal programma; l'altro (.v3d) può essere importato da MeshLab, cosa questa, che permette di generare la *mesh*, con maggiore versatilità.

A differenza di altri *software*, il programma non richiede la calibrazione della camera, né la referenziazione dei punti di presa, come evidenzia il nome stesso, acronimo di *Automatic Reconstruction Conduit tridimensional*, ARC3D.

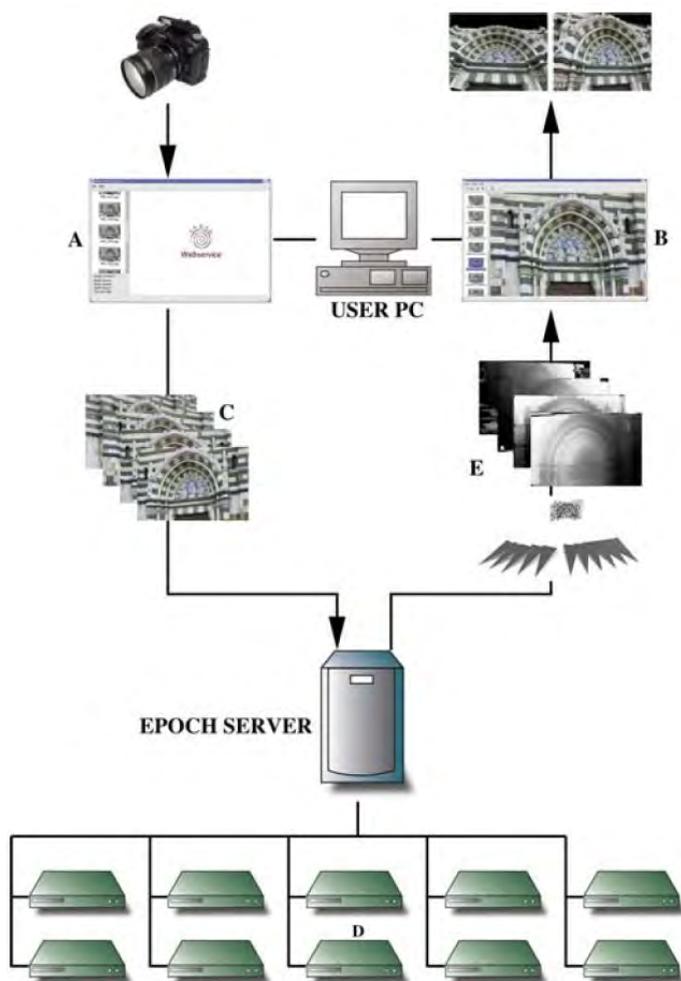


Fig. 4.9 – Schema della architettura client-server del Servizio web Arc3D . In tale schema le fotografie “C” vengono caricate dallo strumento di caricamento “A” (*Upload Images*) dall’utente al server EPOCH. In basso i cluster PC “D”. In uscita i risultati “E”, che vengono scaricati via ftp e visualizzati sul PC dell’utente con il *modelviewer* “B” (M. VERGAUWEN et al., 2006).

*Upload Images*, è il primo strumento del programma, che l’utente del *web-service* incontra per caricare le immagini nel *server*: l’utente seleziona le immagini che desidera caricare; le miniature sono visualizzate sulla sinistra dell’interfaccia e una versione più grande dell’immagine è visualizzata nella finestra principale, quando una miniatura è selezionata.

Le immagini non utili possono essere rimosse dalla lista ed altre immagini possono aggiungersi; in basso a sinistra dell’interfaccia si possono leggere le

informazioni relative all'immagine, come il numero di *pixel*, la dimensione etc.

Prima di inviare i *file* al *server*, attraverso una procedura automatica, si controllano le immagini sfocate<sup>16</sup> e si eliminano per non compromettere la qualità del modello finale.

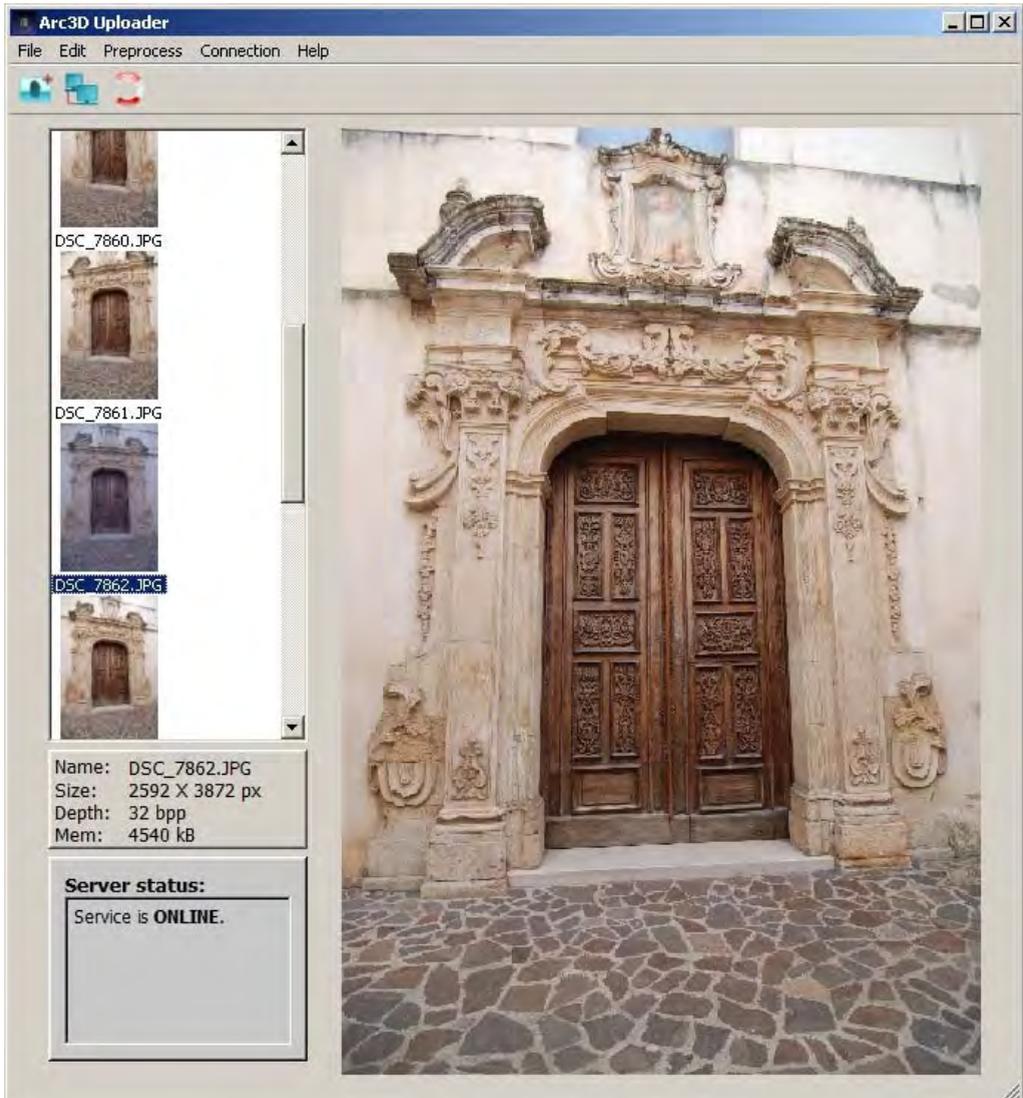


Fig. 4.10 – Interfaccia grafica del modulo Uploader di Arc3D.

---

<sup>16</sup> Check blurry images.

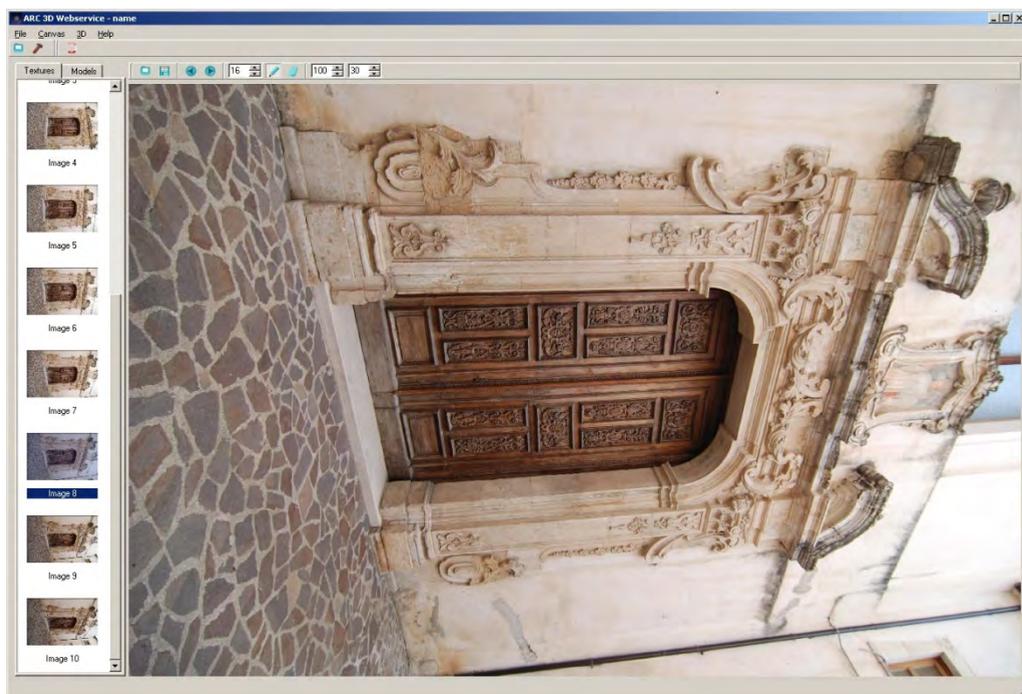


Fig. 4.11 – Interfaccia grafica del modelviewer di Arc3D.

Quando si caricano le immagini sul *server*, in automatico il programma richiede il *subsample*, cioè un ricampionamento delle immagini in percentuale, operazione che solitamente si effettua in una fase successiva, per cui inizialmente è opportuno lasciare la percentuale a 100%.

I risultati restituiti sono memorizzati in un *file zip*, contenente le immagini originali, la calibrazione della camera, le *dense maps* di profondità e le *maps* di qualità di ciascuna immagine; tutte le immagini possono essere ispezionate ed elaborate con il *modelviewer*, che ha la stessa interfaccia grafica del *Upload Images*, con miniature delle immagini sulla sinistra ed una versione più grande nella finestra principale.

In tale programma, da ogni immagine selezionata si può ottenere il modello 3D, usando le *dense maps* di profondità.

ARC3D ha lo scopo di creare una ricostruzione 3D da un'ampia varietà di immagini. Le operazioni che devono eseguirsi, sono legate alle tecniche della *Structure from Motion*, e si basano sul confronto delle corrispondenze fra tutti i *file* inviati. Ogni coppia di fotogrammi è mutuamente commisurata con un

algoritmo di *Global Image Comparison*<sup>17</sup>, con il quale, a seguito di un sottodimensionamento fra le immagini, ciascun elemento posto in corrispondenza è utilizzato per definire una geometria epipolare.

Ricordiamo che tale geometria è propria della fotogrammetria classica e correla un punto sul piano immagine della prima foto con una retta sul piano immagine della seconda foto; retta che è detta, appunto, epipolare. Questa operazione permette di ridurre la ricerca della posizione tridimensionale, vincolandola all'unidimensionalità.

Dal confronto fra ciascuna coppia di immagini si stima la geometria epipolare, la cui affidabilità è garantita da una richiesta di corrispondenza tripla<sup>18</sup>. Saranno pertanto valutati i punti che sono presenti in almeno tre fotogrammi e da qui deriva la necessità di un minimo di sei immagini per creare il modello<sup>19</sup>.

Poiché un modello è definito dalle immagini di partenza, una corretta presa dei fotogrammi non può che influenzare positivamente l'accuratezza del risultato finale. In particolare, l'elemento architettonico che deve essere trasformato in 3D è necessario che occupi l'intera area del fotogramma; i fotogrammi non devono essere troppo distanti l'uno dall'altro e, anche se non c'è bisogno di porsi in modo referenziato all'oggetto, è buona norma compiere un percorso a circonferenza<sup>20</sup>, per avere inquadrature che, in virtù della loro equidistanza, siano facilmente confrontabili<sup>21</sup>.

---

<sup>17</sup> M. VERGAUWEN, L. J. VAN GOOL, *Web-based 3D Reconstruction Service*. Mach. Vis. Appl., 17(6) (2006), pp.411-426.

<sup>18</sup> Date, ad esempio, tre immagini, vengono presi i punti comuni fra le immagini 1 e 2 e fra le immagini 2 e 3.

<sup>19</sup> L'esperienza porta anzi a dire che per effettuare una buona fotomodellazione c'è bisogno di un numero di immagini, superiore a dieci, evitando comunque di inviare *file* pesanti, che richiedono tempi lunghi di processamento. In generale, quindi, è buona norma effettuare fotografie prese con una distanza reciproca contenuta, genericamente con un angolo di trenta gradi fra le due visioni consequenziali.

<sup>20</sup> Qualora l'oggetto da riprendere richieda una presa a 360 gradi, l'esperienza suggerisce di suddividere il percorso di presa in due semicirconferenze.

<sup>21</sup> Perderebbe chiaramente di significato effettuare ingrandimenti con zoom dalla stessa posizione o da posizioni variabili, che poi sono difficilmente confrontabili; ugualmente si ha un risultato peggiore se si utilizza un cavalletto e ci si muove con uniformità, in quanto si perderebbero le discontinuità dei dati necessarie alla calibrazione della macchina.

## MeshLab

Si tratta di un programma concepito come uno strumento generale della rete del sistema di elaborazione 3D; tale programma importa ed esporta nei principali formati della modellazione 3D<sup>22</sup>, permette di elaborare e quindi lavorare sia con nuvole di punti, che con modelli generati da superfici poligonali.

I *file* ottenuti dal servizio *web* Arc3D, sono costituiti da una serie di “gamma mappe” con *texture*, corredate di una misura di qualità, che indica il numero di corrispondenze trovate tra tutte le immagini.

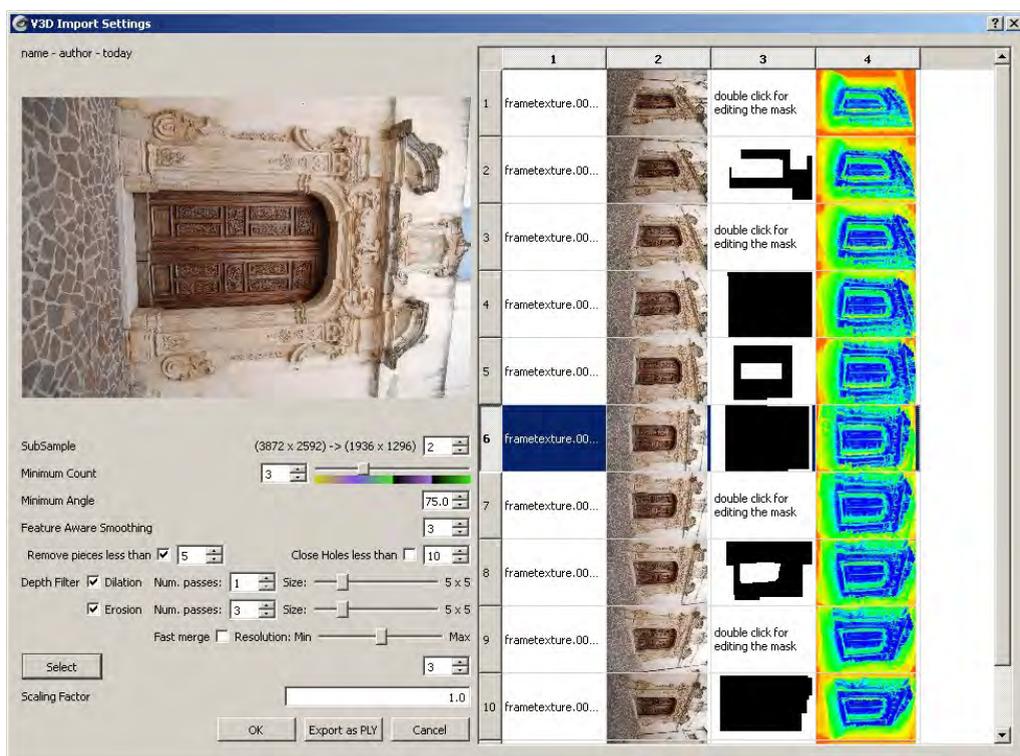


Fig. 4.12 – *Interfaccia grafica del modulo import “.v3d” di MeshLab.* Il colore di ogni mappa, raffigura la qualità di ogni “mappa gamma”, ed indica sinteticamente la quantità dei dati individuati in foto. Sono presenti determinati parametri per ottimizzare il trattamento delle mappa-gamme.

All’interno del programma c’è uno strumento specifico per l’importazione

<sup>22</sup> Importa formati: PLY, STL, OFF, OBJ, 3DS, COLLADA, PTX, X3D e VRML; mentre esporta: PLY, STL, OFF, OBJ, 3DS, COLLADA, X3D, VRML e DXF.

dei dati prodotti da Arc3D. Lo strumento è integrato all'interno MeshLab come filtro d'ingresso, quando l'utente tenta di aprire un *file* ".v3d", il *file* creato da Arc3D; l'interfaccia di figura 3 è quella che vede l'utente.

Tutte le immagini utilizzate per il processo di ricostruzione, sono visualizzate sulla destra con una misura di codici a colori della qualità in funzione della ricostruzione 3D. L'utente può selezionare alcune delle immagini e i dati 3D generati da Arc3D sono pronte per essere utilizzate all'interno di MeshLab. Per guidare il processo di conversione dei dati grezzi 3D, in "mappe gamma" di alta qualità, nel lato sinistro dell'interfaccia vi sono una serie di parametri da settare.

Inizialmente quando si importa il *file* ".v3d" in MeshLab, verrà richiesta la definizione del *subsample*, il "sottocampionamento", ovvero di un parametro di riduzione del numero di pixel in uscita (aumentandone il valore si diminuisce la definizione dei *pixel* che generano il modello). In prima istanza, sembrerebbe opportuno lasciare questo valore al suo massimo livello (pari all'unità); in realtà, poiché le camere acquisiscono anche 18 milioni di *pixel*, il modello generato al massimo della sua definizione avrà un valore molto dettagliato nel piano dell'immagine, al quale però deve essere associato puntualmente un corrispondente attributo di profondità. Ne consegue l'impossibilità di giungere alla stessa accuratezza nel valore spaziale di ogni singolo *pixel*, e poiché l'informazione sarà comunque affetta dall'errore massimo di ogni singolo *range*, è opportuno attuare il sottocampionamento, al fine di avere informazioni uniformi.

Un secondo parametro molto importante è *minimum count*, ovvero la "corrispondenza" fra gli elementi; si tratta di individuare in quante immagini il programma deve trovare il punto stimato per accettarne il valore. Bisogna fare attenzione che già la geometria e le corrispondenze sono state valutate per ciascuna immagine e ciò che il programma raffronta sono i "valori spaziali" precedentemente stimati.

Anche in questo caso si potrebbe pensare che l'assegnazione di un valore elevato porti alla creazione di un modello migliore; in realtà la selettività così impostata esclude una serie di informazioni ed il modello risulterà affetto da "buchi", o "lacune". Occorre bilanciare correttamente il parametro, in funzione dei dati a disposizione; a questo parametro si lega poi uno dei punti di forza di

MeshLab, quello della sua fruizione grafica: nell'interfaccia, ciascuna immagine caricata viene associata ad uno schema, gradualmente colorato in funzione dell'affidabilità del dato; questo valore viene ricavato in funzione della corrispondenza fra i punti (misura di codici a colori della qualità in funzione della ricostruzione 3D). Dal punto di vista pratico, una *preview* con dominante in blu scuro, evidenzia la probabile affidabilità del dato di partenza, mentre la prevalenza di toni rossi e gialli anticipa il probabile insuccesso nella creazione del modello<sup>23</sup>.

Contrariamente al rilievo tradizionale, con le tecniche di rilevamento passivo, le superfici più complesse da gestire sono quelle piatte ed uniformi (per l'assenza di punti di riferimento) e quelle traslucide e levigate, dove la variazione di riflessione della luce, crea trasformazioni cromatiche, che inficiano la ricerca delle corrispondenze al momento del confronto.

Per ovviare in parte a queste difficoltà, prima di generare il modello, si possono compiere una serie di operazioni modificando i valori di alcuni "operatori" che intervengono nella generazione del modello e che in MeshLab sono implementati. Gli "operatori" sono: *remove pieces less than*<sup>24</sup>, un filtro che elimina pezzi sparsi di modello, prodotti dal rumore; *close holes less than*, che chiude i "buchi" del modello, inferiori ad un determinato numero di *pixel*; *feature aware smoothing*, che applica una smussatura alle *mesh*. Ancora, imponendo una dilazione (*dilation*) ed erosione (*erosion*) – una sorta di *smooth* monodimensionale – si possono interpolare con maggiore uniformità i punti stimati<sup>25</sup>. In tale fase è possibile disegnare una maschera per ciascuna foto, in modo da selezionare in partenza le aree non interessate alla creazione del modello, incluse nella ripresa fotografica<sup>26</sup>, ma che possono generare corrispondenze errate.

Dopo aver regolato questi parametri, il *software* calcola e genera il modello *texturizzato*.

---

<sup>23</sup> Un pratico *escamotage* per creare il modello direttamente da ARC3D è carpire dalla *preview* di MeshLab la fotografia di maggiore affidabilità, quindi selezionarla sul visualizzatore di ARC3D.

<sup>25</sup> Quando si ha a che fare con un modello che inaspettatamente viene segnato da una serie di cuspidi è buona norma verificare l'affidabilità del dato in origine, ovvero che le foto di partenza siano state prese con cura e non si siano processate immagini sfocate.

<sup>26</sup> Gli strumenti di selezione sono simili a quelli di un qualsiasi programma di fotoritocco

La seconda fase consiste nella gestione del modello creato, modificando le *mesh* e applicando una serie di filtri che approssimano il dato sia genericamente, che puntualmente. MeshLab permette gran parte di queste operazioni, le stesse che in definitiva interessano i modelli ricavati dalla nuvola di punti prodotta dalla scansione laser<sup>27</sup>.

---

<sup>27</sup> Si può scegliere pertanto di operare con questo programma *open source* o utilizzare programmi commerciali quali *Geomagic* o *Rapidform*.



## CAPITOLO 5

## LE ESPERIENZE CONDOTTE

5.1 *Introduzione*

L'uso del *laser scanner*, è una delle più importanti innovazioni tecnologiche dell'ultimo decennio, capace di determinare cambiamenti rilevanti nel campo del rilevamento architettonico e della modellazione 3D.

Negli anni passati, il *laser scanner* è stato utilizzato intensamente per la generazione dei modelli 3D, necessari per diverse applicazioni, quali la documentazione del patrimonio culturale, la navigazione, l'esplorazione dello spazio, ecc. La tecnologia *laser* ha, cioè, superato la tradizionale fotogrammetria “da vicini” (*close range photogrammetry*) - sia riguardo alla precisione (considerando i laser scanner a triangolazione), che a livello di automazione - portando a ritenere superata la stessa fotogrammetria tradizionale ed, in futuro, completamente sostituibile da scansioni 3D (Boehler et al. 2004).

Negli ultimi anni, al contrario, la corrispondenza *multi-image* è divenuta una delle aree di ricerca più attive sia nella comunità dei “fotogrammetrici”, che in quella della *computer vision*; si sono registrati miglioramenti interessanti, per l'estrazione automatica di corrispondenze di immagine e prodotti numerosi algoritmi per la modellazione.

Oggi, una ricostruzione 3D può essere eseguita sia con il *laser scanner*, che

con la fotogrammetria, ma ci sono fattori che ancora limitano l'efficacia di entrambe le tecniche: da un lato, il *laser scanner*, è in grado di produrre dense nuvole di punti, con elevata precisione geometrica, in modo quasi completamente automatico.

Gli inconvenienti principali degli *scanner* sono l'alto costo, la portabilità, il tempo di acquisizione dei dati, il rumore da oggetti in movimento e la bassa qualità delle informazioni sul colore (*texture*).

Le tecniche IBM (*Image-based modeling*) offrono una alternativa a basso costo, applicabile in un'ampia gamma di scale (opere di architettura e scultura, spazi urbani ...), con tempi di acquisizione dei dati ridottissimi. Per questa tecnica, l'accuratezza geometrica e il livello di automazione dipendono in larga misura dall'oggetto da rilevare, dalla risoluzione delle immagini e dalla geometria della rete.

Per quanto appena detto, quale sia metodo migliore rimane ancora una questione aperta e, correntemente, si parla di integrazione tra i due metodi sia in relazione al costo delle operazioni di rilevamento, che agli scopi ed alla qualità dei "prodotti" possibili.

In questo lavoro di tesi, per approfondire la possibilità di integrazione tra i due metodi (qualità dei "prodotti" di acquisizione, incidenza degli stessi prodotti sulle restituzioni ...), vengono affrontate diverse applicazioni, impiegando tecniche *laser scanner* e tecniche IBM, al fine di valutare e confrontare i risultati ottenuti; si tratta di applicazioni condotte direttamente sia in ambienti esterni che in laboratorio. Tali applicazioni cercano di verificare, in termini disciplinari, le possibilità d'integrazione dei sistemi provenienti dalle tecniche ottiche di misura 3D passive, in modo particolare la fotogrammetria "da vicini" e dalle tecniche della *computer vision*, soprattutto in merito alla possibilità di produrre dense nuvole di punti, utilizzando anche procedure automatiche.

Per affrontare il problema in termini di precisione, affidabilità del risultato ottenuto, metodo da impiegare in funzione dello scopo e facilità d'uso si sono, pertanto, condotte una serie di sperimentazioni su casi-studio, che presentano caratteristiche differenti per dimensione, struttura geometrico-architettonica, stato di conservazione, epoca storica, ecc...: sperimentazioni di laboratorio, eseguite su oggetti selezionati, con geometria nota e, nello stesso tempo, sperimentazioni in ambienti esterni, con scene realistiche.

Le esperienze condotte in laboratorio riguardano:

- *test su un oggetto piano rettangolare 50x70 mm*; base in cartoncino, con sopra stampati elementi geometrici di vario tipo e con forte contrasto di colore, per rilevare le capacità dei vari sistemi di acquisizione.
- *Test su un capitello di polistiro*; capitello composto in polistirolo, con base circolare del diametro di 24 cm, e dimensioni di 30 cm per lato, con superficie di colore bianco, non trattata.

Le esperienze condotte in ambienti esterni riguardano:

- *Il portale della chiesa di Santa Maria delle Vergini (Cosenza)*; un portale tardo-barocco - con decorazioni floreali, a lunghi tralci, poste sull'architrave, fino alla base delle paraste - con arco ribassato e modanature sottili a rilievo, che in alcuni tratti fanno da sfondo alla decorazione esterna.
- *I mosaici della chiesa di Sant'Adriano (San Demetrio Corone, Cosenza)*; i mosaici si trovano sul pavimento della chiesa, lavorati in *opus sectile* - su base in marmo di forma tonda e rettangolare - con rappresentazioni di animali (serpenti e felini).

## 5.2 Le esperienze condotte in laboratorio

- *Test su un oggetto piano rettangolare.*
- *Metodo, strumenti e software*

Il primo *test* è stato progettato per confrontare direttamente le precisioni tra una nuvola di punti ottenuta con *laser scanner* a triangolazione, il Minolta VIVID 300 e le nuvole di punti ottenute da tecniche IBM (ARC3D/MeshLab e Agisoft PhotoScan), utilizzando un oggetto di dimensioni note e superficie piana (un foglio rettangolare di cartoncino, di 700x500 mm).

Per quel che riguarda la fase di acquisizione della nuvola di punti, fatta con il *laser scanner* a triangolazione, si è adottato il *software* a corredo dello strumento, mentre per la successiva fase di *editing* delle singole nuvole di punti

e per l'allineamento delle stesse, si è utilizzato “*Geomagic Studio 12*”. A causa del ridotto campo d'acquisizione dello scanner, si è eseguita una scansione multipla, da più punti di vista; per ricostruire l'intero modello si sono, quindi, assemblate più nuvole di punti - ognuna corrispondente ad una differente porzione dell'oggetto da acquisire - al fine di ottenere un'unica nuvola di punti, rappresentativa dell'intero soggetto. Per facilitare questa operazione si sono evidenziati, sulla superficie da rilevare, una serie di punti notevoli, utili come riferimento; tali punti notevoli (pallini di piombo incollati sul cartoncino in esame), caratterizzati sulle tre dimensioni, sono necessari per evitare che, durante la fase di allineamento, le singole *range maps* (singola superficie acquisita) slittino l'una sull'altra, rendendo l'allineamento stesso non corretto.

Per la restituzione delle nuvole di punti, ottenute da tecniche IBM, si è eseguito il rilievo fotografico con una camera digitale *reflex* Nikon D80, con un sensore di 3872x2592 *pixel*, ossia di 10 *mega-pixel*; con tale camera si sono acquisiti 10 fotogrammi.

#### *Riflessioni e risultati raggiunti*

- *Confronto tra nuvole di punti ottenute da Agisoft PhotoScan e ARC3D/meshLab*

Come si può notare dall'immagine riportata in figura 5.1, il confronto tra le due nuvole di punti è risultato inefficace: la ricostruzione eseguita da ARC3D/MeshLab, è incompleta e presenta molte parti mancanti, anche se è composta da 1.5 milioni di punti; tali parti si evidenziano laddove non sono presenti elementi contrastanti. I “buchi” sono presenti soprattutto nelle parti di colore uniforme, privi di elementi contrastanti, e sui quadrati neri.

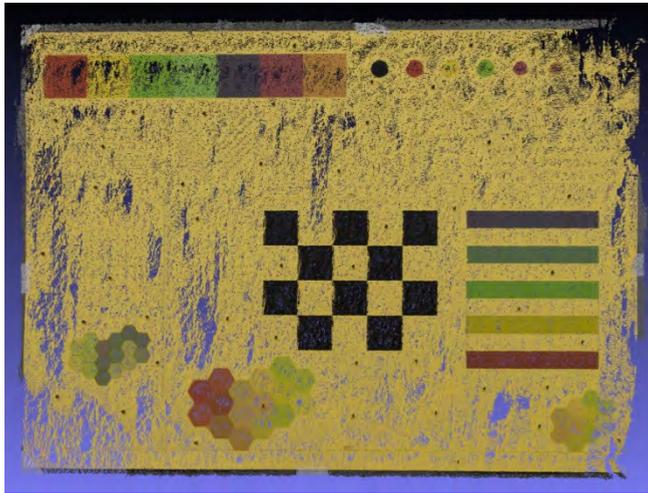


Fig.5.1 *Elemento planare rettangolare*. Nuvola di punti ottenuta con ARC3D/MeshLab, composta da 1.5 milioni di punti; presenta molti "buchi", evidenti soprattutto nelle parti di colore uniforme, privi di elementi contrastanti, e sui quadrati neri.

Questo risultato si deve all'impossibilità di trovare, in queste parti (aree bianche e aree nere), corrispondenze in fase di *matching*.

Gli stessi fotogrammi, processati con *Agisoft PhotoScan*, restituiscono, una nuvola di punti composta da 1.1 milioni di punti, e un modello completo dell'oggetto in esame, costituito da 2.2 milioni di triangoli, (fig. 5.2) e metricamente corretto, dopo il rapporto con una misura di riferimento.

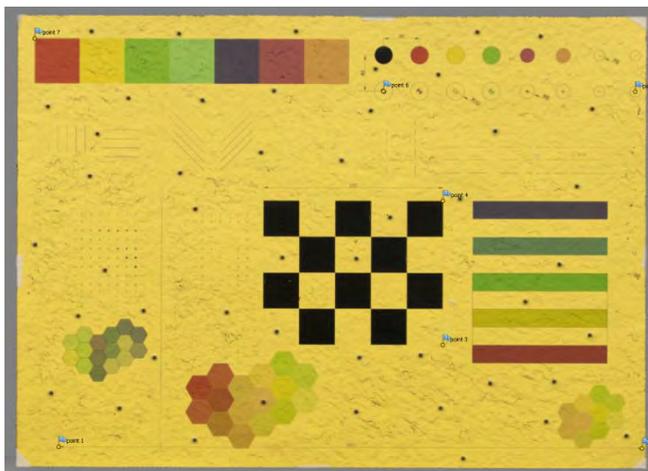


Fig.5.2 *Elemento planare rettangolare*. Modello ottenuto con *Agisoft PhotoScan*, composto da 2.2 milioni di triangoli; il modello è completo e metricamente corretto, dopo il rapporto con una misura di riferimento.

- *Confronto tra nuvole di punti ottenute da laser scanner a triangolazione, Minolta VIVID 300 e tecniche IBM ARC3D/MeshLab.*

Il confronto tra la nuvola di punti ottenuta da *laser* e la nuvola restituita da ARC3D è risultato inefficace, poiché la nuvola ottenuta da ARC3D risulta incompleta e presenta molte parti mancanti.

- *Confronto tra modello ottenuto da laser scanner a triangolazione, Minolta VIVID 300 e tecniche IBM Agisoft PhotoScan.*

Il confronto diretto tra il modello, ottenuta da *laser* ed il modello restituito da *Agisoft PhotoScan*, ha lo scopo di valutare la precisione metrica dei modelli 3D ottenuti con i due sistemi; si è eseguita una comparazione, utilizzando come modello di riferimento quello ottenuto con il *laser scanner*.

Il flusso di lavoro ha previsto le seguenti fasi:

- Riduzione in scala del modello ottenuto da *Agisoft PhotoScan*, rispetto ad una misura presa direttamente sull'oggetto, operazione eseguita direttamente in *Agisoft PhotoScan*, evidenziando sulle fotografie dei *markes*;
- allineamento dei due modelli nel medesimo sistema di riferimento. Il modello realizzato da fotogrammi, è stato allineato con quello ricavato dal laser scanner, eseguendo un allineamento tramite selezione manuale di punti corrispondenti (sette o otto) su entrambi i modelli; le operazioni di allineamento e le successive di comparazione sono state eseguite con *Geomagic Studio 12*.

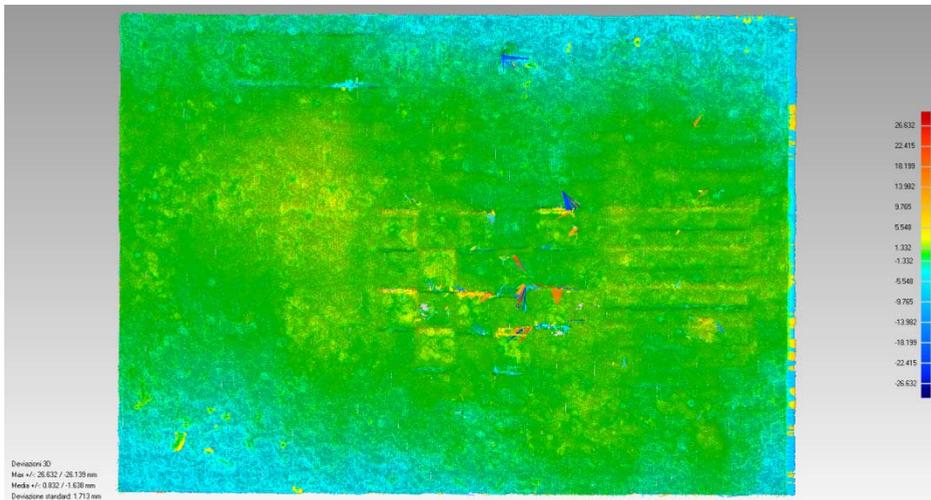


Fig.5.3 Elemento planare rettangolare. Confronto tra modello ottenuto da *laser scanner* a triangolazione, Minolta VIVID 300 e tecniche IBM Agisoft PhotoScan; gli scostamenti medi tra i due modelli si attestano in  $\pm 1,6$  mm, mentre gli scostamenti massimi sono di  $\pm 26$  mm.

Dal confronto riportato in figura 5.3 emerge che gli scostamenti medi tra i due modelli si attestano a  $\pm 1,6$  mm, mentre gli scostamenti massimi registrati sono di  $\pm 26$  mm; tali valori elevati degli scostamenti massimi, sono evidenziati dai colori blu e rosso e caratterizzano le zone ricoperte dai riquadri neri e le zone delle linee di passaggio tra nero e sfondo. Questi comportamenti in letteratura sono chiamati artefatti della misura<sup>28</sup>. Dalla figura in esame, si evidenzia come le parti periferiche del modello, ottenuto dal *laser*, si discostano da quello ottenuto da fotogrammi in quanto il supporto impiegato per acquisire l'oggetto era deformato (leggera curvatura lungo i bordi).

Successivamente, si sono confrontati in modo indiretto i due modelli, con un piano fittizio di riferimento. Tale piano è generato dalla superficie di *best-fit* del modello, ossia da una primitiva, che interpola i vertici dei triangoli che formano il modello. In questo modo è stato possibile misurare gli scostamenti dei modelli dal piano ideale ed anche fare delle sezioni, con piani perpendicolari a tale piano ideale per estrarre dei profili, per valutare

<sup>28</sup> Artefatti dovuti a forti variazioni cromatiche. In questi casi, i sistemi a triangolazione registrano un salto fittizio di quota nel punto di transizione tra i colori, anche se le due aree sono complanari. L'errore è proporzionale all'intensità luminosa utilizzata ed è dovuto ad una errata stima della posizione dello *spot* sul sensore (G. GUIDI et al., 2010).

puntualmente gli scostamenti tra piano e modello.

*Confronto tra nuvola di punti ottenuta da laser scanner e piano fittizio di riferimento.*

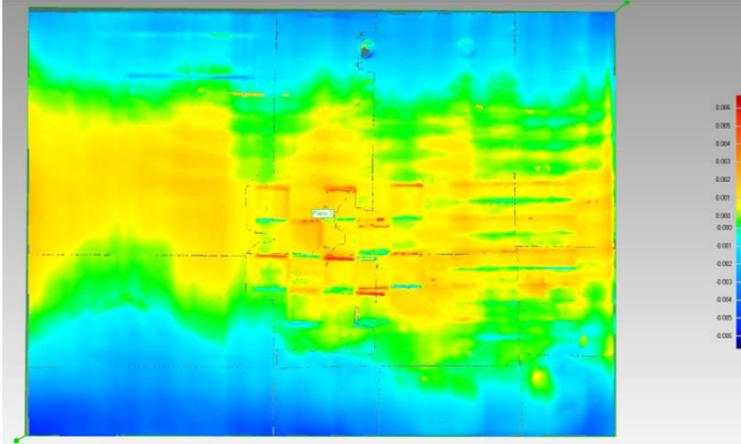


Fig. 5.4 *Elemento planare rettangolare*. Confronto tra piano fittizio e modello ottenuto da *laser* a triangolazione; gli scostamenti tra il piano fittizio e il modello sono contenuti nell'ordine dei  $\pm 6$  mm.

Come si può notare dall'immagine riportata in figura 5.4, gli scostamenti tra il piano fittizio e il modello sono contenuti nell'ordine dei  $\pm 6$  mm, questi sono evidenziati dal colore rosso, che caratterizzano le zone di passaggio tra i riquadri neri e sfondo. La parte periferica del rettangolo è da non considerare in quanto come detta in precedenza il supporto impiegato per acquisire l'oggetto era incurvato, quindi gli scostamenti in tali zone non sono da considerare.

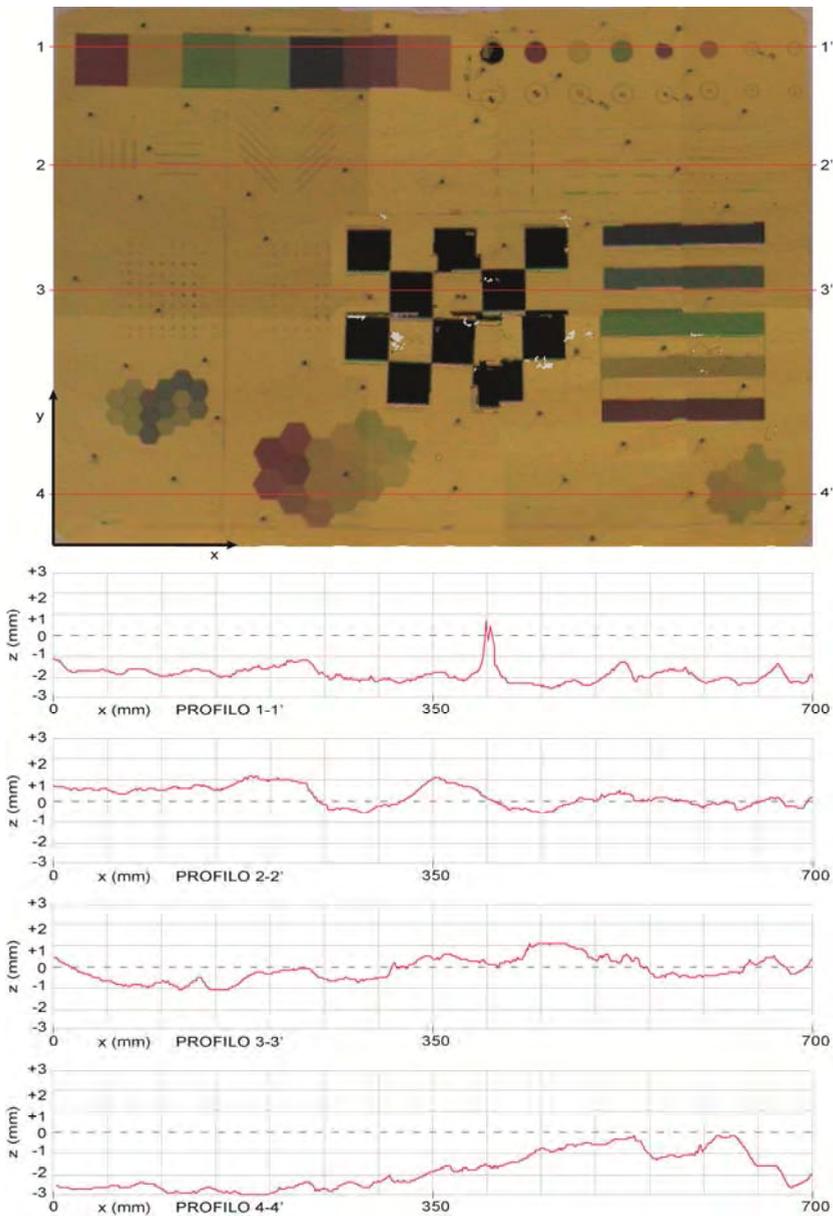


Fig. 5.5 *Elemento planare rettangolare*. Modello ottenuto da *laser scanner*. I profili ottenuti in automatico, elaborando il modello in *Geomagic Studio 12*, mostrano gli scostamenti registrati rispetto al piano; tali scostamenti sono contenuti nell'ordine dei 3 mm, i salti maggiori si evidenziano nelle zone dove vi è l'alternanza di colori. I profili 1-1' e 4-4' evidenziano l'incurvatura del supporto sul quale era poggiato il rettangolo, e di conseguenza il maggiore scostamento di questo dal piano.  
 N.B. Le altezze sono amplificate con un moltiplicatore pari a 3.

- *Confronto tra nuvola di punti ottenuta da Agisoft PhotoScan e piano fittizio di riferimento.*

Come si può notare dall'immagine riportata in figura 5.5, gli scostamenti tra il piano fittizio e il modello sono di max.  $\pm 2,3$  mm e si concentrano nelle aree bianche, su qualche riquadro nero - dove ci sono pochi elementi riconoscibili per il *software* - e nelle parti terminali, caratterizzate da rumore di bordo. Lo scostamento medio è di 0,4 mm ed ha una distribuzione quasi uniforme sull'intera area; la deviazione *standard* è di 2,6 mm.

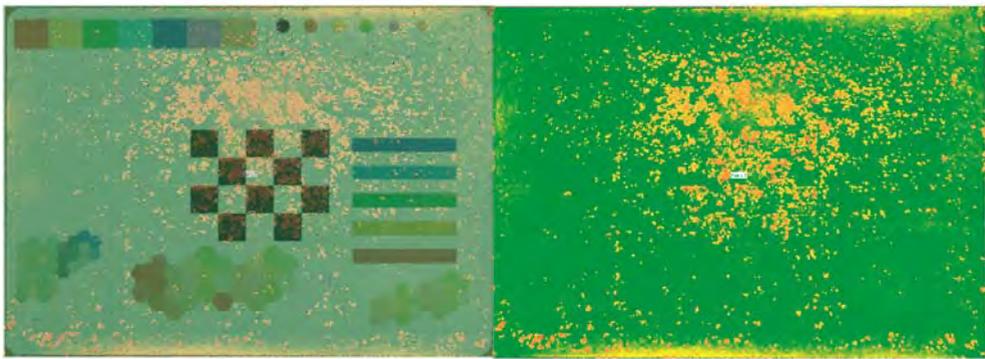


Fig. 5.5 *Elemento planare rettangolare. Confronto tra piano fittizio e nuvola di punti ottenuta da Agisoft PhotoScan.*

Scostamenti(mm)	Laser.	Agisoft.
Scost. Max	$\pm 6,2$	$\pm 2,3$
Scost.Medio	$\pm 0,3$	$\pm 0,4$
Deviazione standard	1,6	2,6

Tab. 5.1 – *Confronto tra la nuvola di punti ottenuta laser scanner a triangolazione, il Minolta VIVID 300, nuvole di punti ottenute con Agisoft PhotoScan e piano fittizio di riferimento.*

In conclusione, per gli elementi planari, la restituzione ottenuta con *Agisoft PhotoScan* è del tutto comparabile con quelle ottenute con un *laser scanner* a triangolazione.

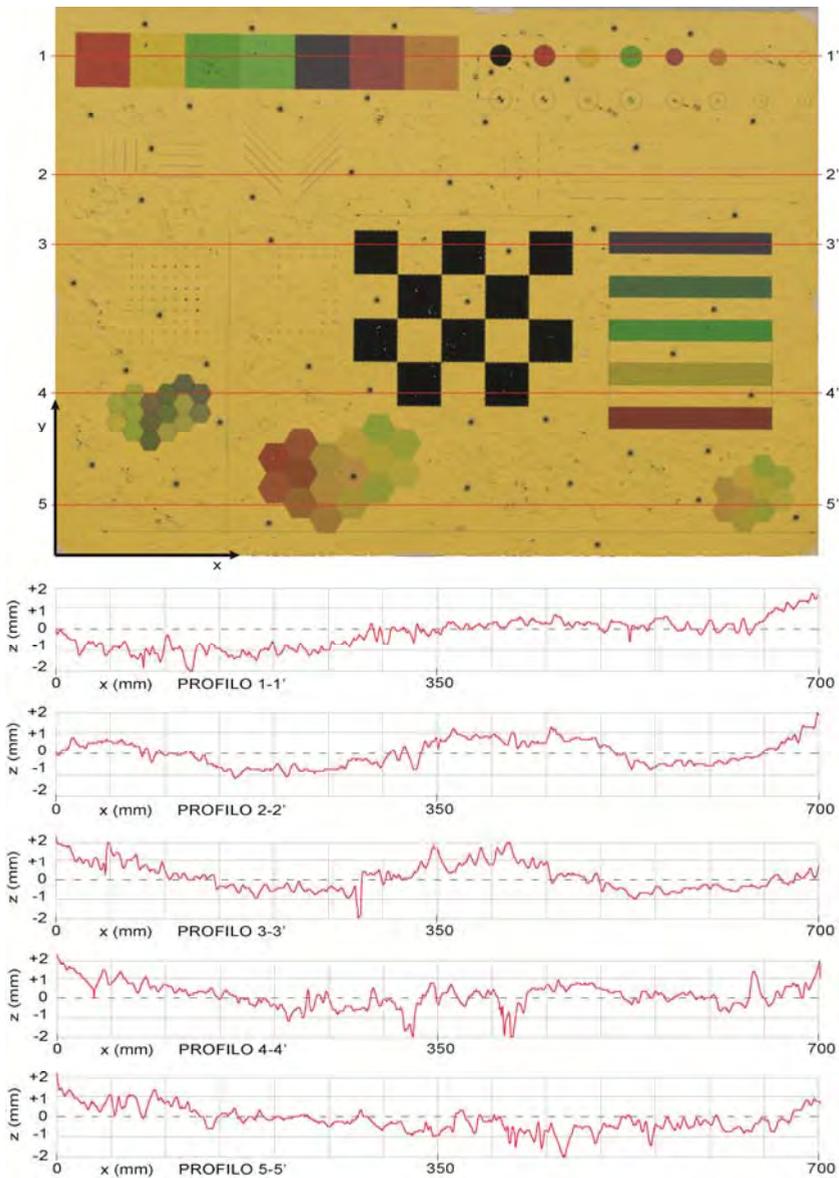


Fig. 5.6 *Elemento planare rettangolare*. Modello ottenuto con *Agisoft PhotoScan*. I profili ottenuti in automatico analizzando il modello in *Geomagic Studio 12*, mostrano gli scostamenti registrati rispetto al piano; tali scostamenti sono contenuti nell'ordine dei 2 mm; i salti maggiori si evidenziano nelle zone con colori uniformi e ciò è dovuto all'impossibilità di trovare, in queste parti, corrispondenze nelle diverse fotografie, in fase di *matching*.

N.B. Le altezze sono amplificate con un moltiplicatore pari a 3.

- *Test su un capitello di polistirolo.*
- *Metodo, strumenti e software*

Tale *test* ha avuto lo scopo di confrontare e valutare i modelli digitali 3D del capitello, ottenuti impiegando il *laser scanner* a triangolazione, (il Minolta VIVID 300) e le tecniche IBM (ARC3D e *Agisoft PhotoScan*). Una prima esperienza è stata quella di provare ad ottenere un modello del capitello così com'è, riprendendo il capitello bianco sia con il *laser scanner*, che con un set di foto: i risultati ottenuti sono stati scarsi ed i modelli incompleti. Tali risultati sono da attribuire all'effetto di "riflessione" della superficie.

Una successiva esperienza è stata condotta applicando sulla superficie del capitello una vernice a base d'acqua, di colore "ocra", per rendere la superficie meno riflettente. Per l'acquisizione di una sola faccia del capitello con il *laser scanner* a triangolazione - il Minolta VIVID 300 - sono state necessarie cinque scansioni, ognuna corrispondente a una differente parte dell'oggetto, con un'area di sovrapposizione tra le diverse scansioni superiore al 30%, tale da garantire un corretto allineamento delle singole *range maps*. A differenza del *test* eseguito sull'oggetto "piano" rettangolare, in questo caso non si sono introdotti elementi esterni al fine di ottenere un unico modello dell'intera faccia, poiché l'oggetto in esame è caratterizzato da diversi elementi a tre dimensioni, che aiutano la fase di allineamento e fusione delle differenti *range maps* acquisite.

Il modello di una faccia del capitello, elaborato con *Polygon Editing Tool Ver.2.00* della KONICA MINOLTA SENSING - *software* usato dallo *scanner* per eseguire l'operazione di acquisizione e la successiva fase di allineamento e fusione delle singole *range maps* - è stato esportato in "*Geomagic Studio 12*" e usato sia come modello di riferimento, che di confronto con gli altri modelli ottenuti con tecniche IBM.

Il rilievo fotografico è stato eseguito con una camera digitale *reflex*, Nikon D80, con un sensore di 3872x2592 *pixel*; considerate le dimensioni dell'oggetto, si è impiegato un obiettivo con una distanza focale nominale di 35 mm, scattando le foto ad una distanza di circa 1m dall'oggetto. I fotogrammi sono stati salvati nel formato grezzo NEF (Nikon Electronic Format), ovvero nel formato di ripresa digitale dei dati acquisiti direttamente dal sensore, prima che il *software* interno alla fotocamera intervenga per apportare le regolazioni

(nitidezza, contrasto, bilanciamento del colore ect...), impostate dall'utente al momento dello scatto. Impiegando il formato NEF, ogni singolo fotogramma è stato processato a posteriori, apportando i miglioramenti dovuti sulle foto scattate.

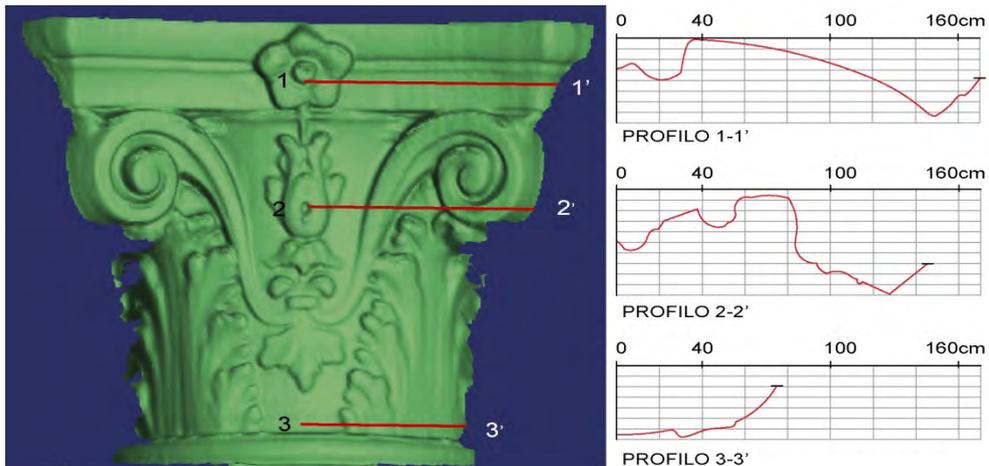


Fig. 5.7 – Modello del capitello, da laser scanner Minolta VIVID 300, ottenuto dall'elaborazione delle cinque range maps, allineate ed unite con Polygon Editing Tool Ver.2.00, software usato dallo scanner per eseguire l'operazione di acquisizione e la successiva fase di editing. A destra sono riportati 3 profili, del modello poligonale, estratti in automatico con "Geomagic Studio 12", usato per la fase di allineamento dei diversi modelli e per il confronto.

I modelli ottenuti con le con tecniche *Structure and motion*, sono stati realizzati con *Agisoft PhotoScan* (software commerciale) e con il sistema ARC3D / MeshLab (servizio web e *software open source*), impiegando 13 fotogrammi in entrambi i sistemi.

Il modello restituito con *Agisoft PhotoScan* è composto da 5 milioni di triangoli, che ricostruiscono per intero, la faccia del capitello.

Impiegando il servizio web ARC3D ed elaborando i dati restituiti con MeshLab, si è ottenuto un modello della stessa faccia considerata, composta da 3 milioni di triangoli; in questo caso compaiono alcune lacune, cioè alcune parti del modello non ricostruite.

Nelle figure 5.7, 5.8 e 5.9 sono riportati i modelli poligonali 3D, ottenuti con le diverse tecniche. A lato, di ogni modello, vengono riportati tre profili, ottenuti in modo automatico; gli stessi saranno confrontati tra loro per leggere gli scostamenti che si registrano, impiegando le diverse tecniche di

acquisizione e creazione dei modelli 3D.

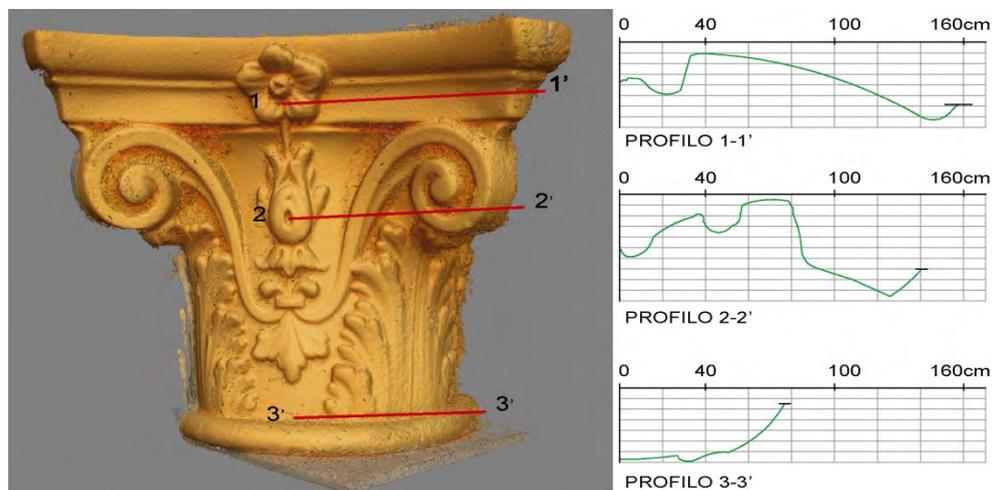


Fig. 5.8 – Modello poligonale del capitello, restituito da Agisoft PhotoScan, composto da 5 milioni di triangoli. A destra sono riportati 3 profili, del modello poligonale, estratti in automatico con “Geomagic Studio 12”, usato per la fase di allineamento dei diversi modelli e per il confronto.

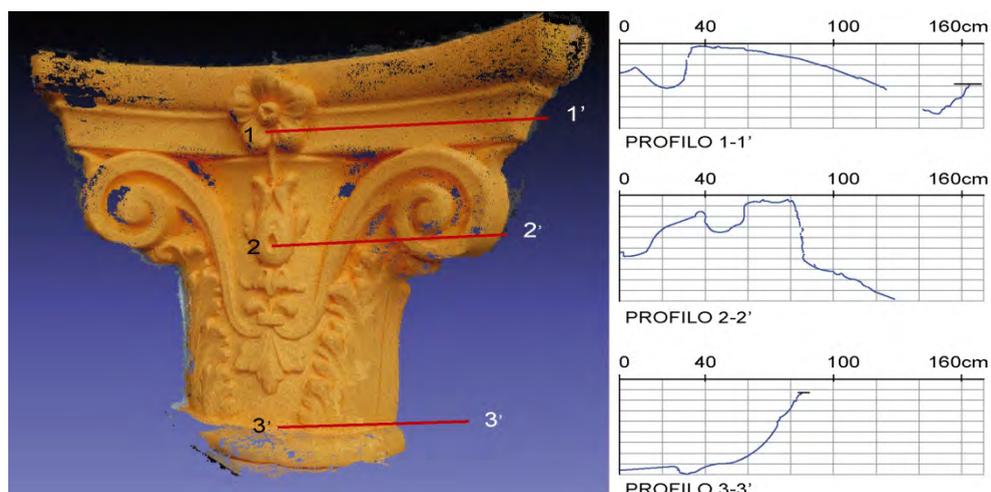


Fig. 5.9 – Modello poligonale del capitello, restituito da ARC3D/MeshLab, composto da 3 milioni di triangoli. A destra sono riportati 3 profili, del modello poligonale, estratti in automatico con “Geomagic Studio 12”, usato per la fase di allineamento dei diversi modelli e per il confronto.

Al fine di valutare la precisione metrica dei modelli 3D ottenuti con i diversi sistemi, si è eseguita una comparazione, utilizzando come modello di riferimento quello ottenuto con il *laser scanner*.

Il flusso di lavoro ha previsto le seguenti fasi:

- riduzione scalare i modelli attraverso le tecniche *Structure and motion*, rispetto ad una misura presa direttamente sull'oggetto;
- allineamento dei tre modelli nel medesimo sistema di riferimento. I due modelli realizzati da fotogrammi, sono stati allineati con quello ricavato dal *laser scanner*, eseguendo un allineamento preliminare tramite selezione manuale di sette o otto punti corrispondenti; in seguito si è stata seguita un'implementazione ICP, procedendo con una registrazione automatica dell'allineamento, utile per minimizzare l'errore residuo. Le operazioni di allineamento in un unico sistema di riferimento si sono eseguite con "Geomagic Studio 12", e sempre nello stesso *software* si sono eseguiti i confronti che sono:
- confronto tra modelli ottenuti da *Agisoft PhotoScan* e *ARC3D/MeshLab*;
- confronto tra il modello ottenuto da *laser scanner* Minolta VIVID 300 e i modelli ottenuti con tecniche *Structure and motion*.

### *Riflessioni e risultati raggiunti*

- *Confronto tra modelli ottenuti da Agisoft PhotoScan e ARC3D/MeshLab*

Dal confronto riportato in figura 5.10, emerge che gli scostamenti medi tra i due modelli ottenuti con tecniche *Structure and motion*, si attestano a  $\pm 0,9$  mm, mentre gli scostamenti massimi registrati sono di  $\pm 15$  mm; tali valori elevati degli scostamenti massimi, si evidenziano nelle parti periferiche del modello, parti in cui uno dei modelli è incompleto (mancante), perché non restituito. Sempre nella stessa figura le parti in grigio scuro, sono quelle in cui il modello ottenuto da *ARC3D/MeshLab* è incompleto, come si può notare anche nella figura 5.9.

Analizzando i profili dei due modelli, messi a confronto, emerge che gli scostamenti maggiori, registrati puntualmente, sono dell'ordine di 1.8 mm.

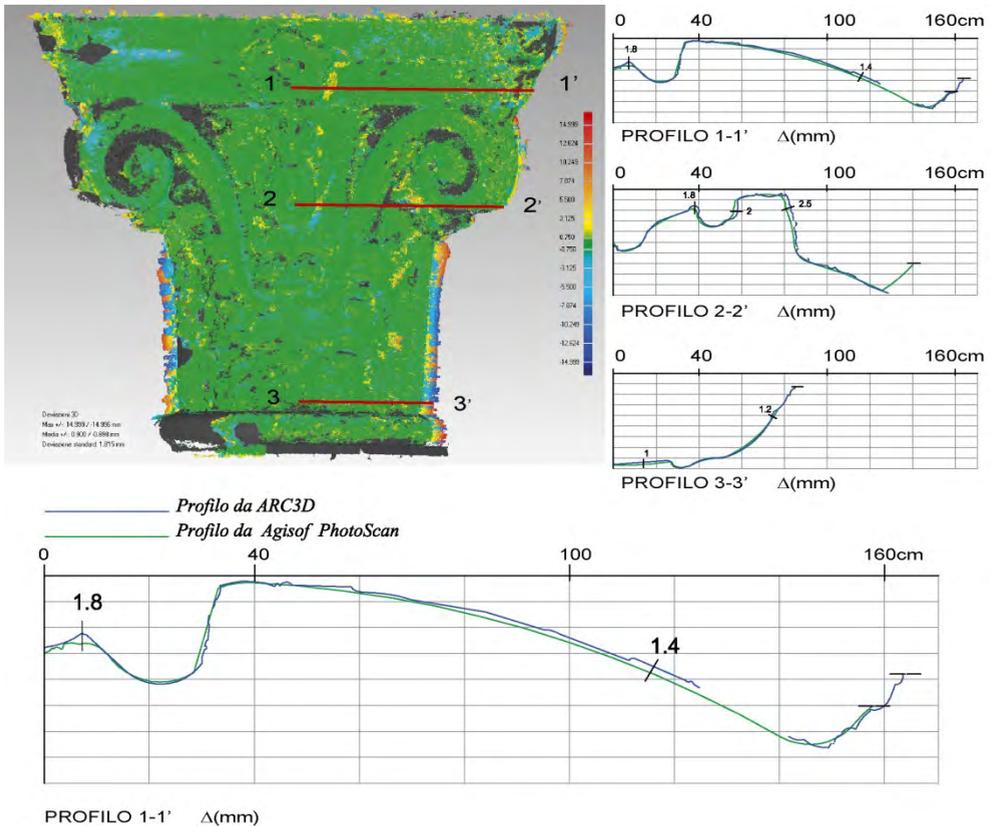


Fig. 5.10 – Confronto tra modelli ottenuti da Agisoft PhotoScan e ARC3D/MeshLab. Gli scostamenti medi tra i due modelli sono di  $\pm 0,9$  mm, gli scostamenti maggiori si sono registrati nelle parti periferiche dei modelli; in tali parti si ha la mancata ricostruzione del modello. Le aree rappresentate in grigio scuro, sono parti in cui il modello ottenuto da ARC3D/MeshLab è incompleto. Dal confronto tra i profili emerge che gli scostamenti maggiori rilevati in punti singoli sono di 1.8 mm.

- Confronto tra il modello ottenuto da laser scanner Minolta VIVID 300 e i modelli ottenuti da ARC3D/MeshLab e Agisoft PhotoScan

Gli scostamenti massimi tra il modello restituito da *laser scanner* e i modelli restituiti da tecniche *Structure and motion*, sono rispettivamente di  $\pm 18$  mm con il modello restituito da ARC3D/MeshLab, e di  $\pm 15$  mm il modello restituito dal *software* commerciale Agisoft PhotoScan; come per il confronto tra i due sistemi fotografici, anche in questo, i picchi di scostamento si evidenziano nelle zone periferiche, dove i modelli non coincidono. Lo

scostamento medio è nei due casi di  $\pm 1,7$  mm, nel confronto tra modello restituito da *laser scanner* e modello elaborato con *ARC3D/MeshLab*, mentre diventa  $\pm 1,9$  mm, nel confronto tra modello restituito con *laser scanner* e modello elaborato con *Agisoft PhotoScan*. Nelle figure 5.11 e 5.12, vengono riportati i confronti eseguiti in “*Geomagic Studio 12*”, riportando a colori gli scostamenti tra i modelli; a lato sono riportati, per confronto, i profili analizzati con l’indicazione degli scostamenti maggiori.

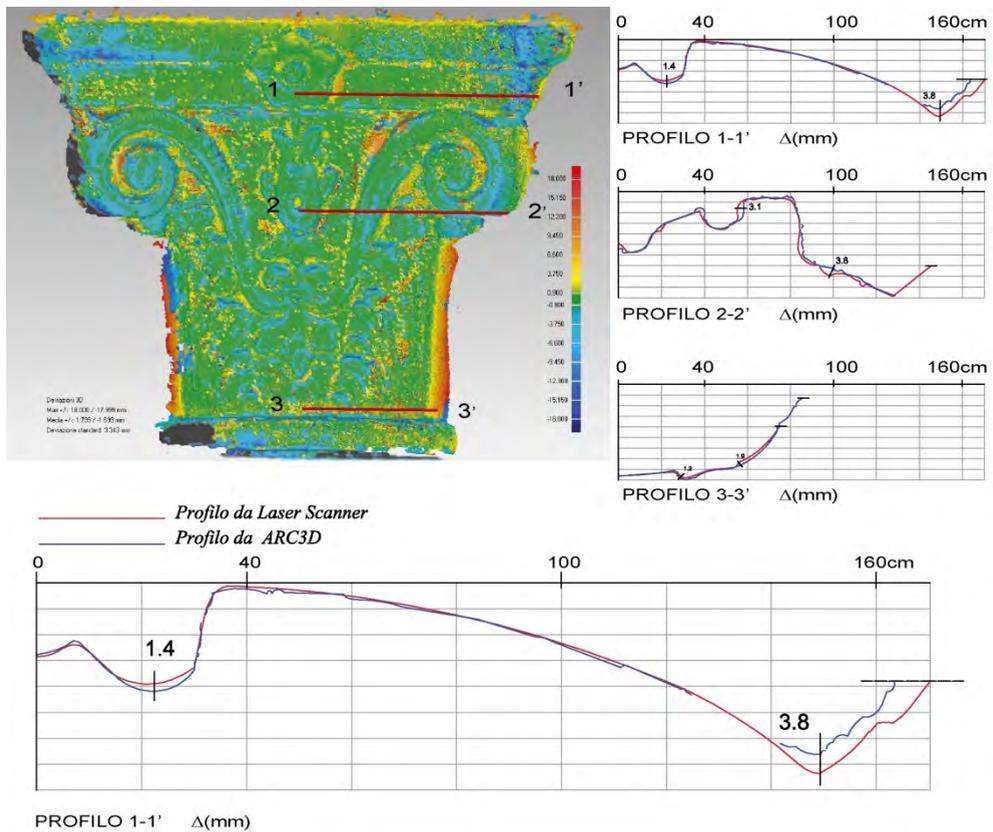


Fig. 5.11 – Confronto tra modelli restituiti da *laser scanner Minolta VIVID 300* e *ARC3D/MeshLab*. Gli scostamenti medi tra i due modelli sono di  $\pm 1,7$  mm, gli scostamenti massimi sono di  $\pm 17$  mm; tali differenze si sono registrate nelle parti periferiche dei modelli. In queste parti si ha la mancata ricostruzione del modello. Dal confronto tra i profili emerge che gli scostamenti maggiori rilevati, in punti singolari, sono di 3.8 mm.

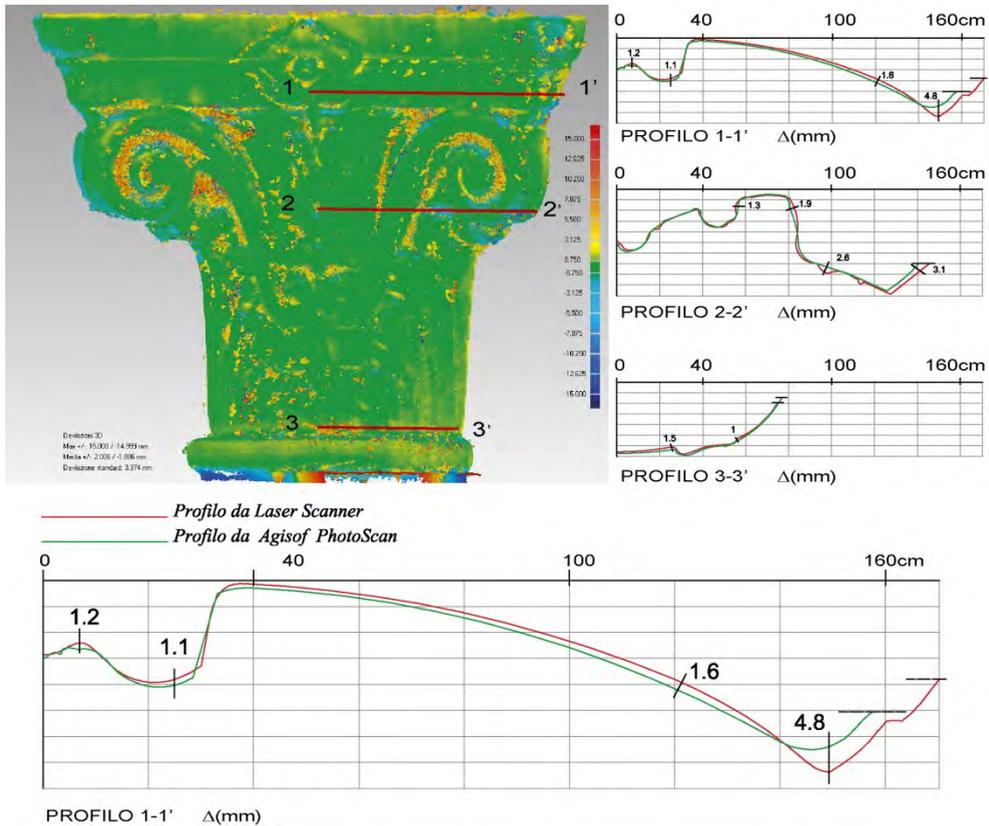


Fig. 5.12 – Confronto tra modelli restituiti da laser scanner Minolta VIVID 300 e Agisoft PhotoScan. Gli scostamenti medi tra i due modelli sono di  $\pm 1,9$  mm, gli scostamenti massimi sono di  $\pm 15$  mm, mentre gli scostamenti massimi, in punti singoli, rilevabili dal confronto dei profili sono di 4,8 mm.

Nella tabella 5.2, sono riepilogati i valori degli scostamenti emersi dal confronto tra il modello elaborato dalla nuvola di punti, ottenuta dal laser Minolta VIVID 300, presa come riferimento ed i modelli ottenuti in automatico da tecniche *Structure and motion*.

Scostamenti(mm)	ARC3D/MeshLab 3 milioni di tr.	Agisoft. 5 milioni di tr.
Scost. Max	$\pm 18$	$\pm 15$
Scost.Medio	$\pm 1,7$	$\pm 1,9$
Deviazione standard	3,34	3,37

Tab. 5.2 – Confronto tra il modello restituito da laser scanner Minolta VIVID 300, e modelli restituiti con ARC3D/MeshLab e Agisoft PhotoScan.

In conclusione, anche per questo tipo di oggetti, la restituzione ottenuta con tecniche *Structure and motion* è del tutto comparabile con quelle ottenute con un *laser scanner* a triangolazione; in più tali modelli derivanti da fotografie sono completi di *texture*, che permettono di restituire l'apparenza visiva dell'oggetto reale.

### 5.3 Le esperienze condotte all'esterno

#### *Il portale della chiesa di Santa Maria delle Vergini (Cosenza)*

##### *Descrizione del manufatto*

La nascita del monastero e, conseguentemente, della chiesa di S. Maria delle Vergini, nella città di Cosenza, si può collocare nel secondo decennio del Cinquecento, molto probabilmente negli anni compresi tra il 1512 e il 1520, sotto la giurisdizione dell'arcivescovo Giovanni Ruffo Teodoli (1511-1527).

Per restringere il campo delle possibilità in cui collocare la nascita del complesso, in genere, si fa riferimento a due differenti tipi di documenti: considerando la cosiddetta "Cronaca di Bosco", C. Minicucci ritiene che la fondazione del monastero delle Vergini, sotto il titolo di S. Maria delle Vergini, sia dell'anno 1515; mentre D. Andreotti, nella *Storia dei Cosentini*, fa risalire la data di fondazione del monastero al 1517<sup>29</sup>.

Nel 1639, il monastero fu danneggiato notevolmente dal terremoto che colpì la città di Cosenza. In seguito, durante il cosiddetto "decennio francese", venne decretata la soppressione degli Ordini religiosi, che, per il nostro monastero, avvenne il 22 maggio 1809; dopo tale evento, nei suoi locali, venne istituito l'omonimo Orfanotrofio femminile, per le fanciulle provenienti da altri

---

<sup>29</sup> In B. CANONACO, *Cosentia. Il progetto dell'antica città calabra attraverso i documenti d'archivio e le vedute storiche*. Soveria Mannelli (CZ) Rubbettino, 2007, p. 65.

preesistenti istituti della città.

Dal 1887, alle Vergini, ha sede la prima casa cosentina della Congregazione delle Figlie di S. Anna, mentre, in altri locali, quelli propri dell'antico monastero, cioè quelli contigui al chiostro, si insedia la Scuola "Fratelli Bandiera".

Per l'edificazione dell'intero complesso architettonico fu necessario provvedere all'innalzamento di una platea per compensare il dislivello tra l'attuale Via del Liceo, alle spalle del monastero e via G. Argentò.

Architettonicamente, il monastero si caratterizza, con ambienti voltati ed archi; attraversando il portale, a bugne, si accede ad un cortile quadrangolare, che precede la chiesa e fa da disimpegno per il chiostro e per i locali di tutto complesso.

L'accesso avviene attraverso un grande portale, caratterizzante il prospetto sia in termini dimensionali, che compositivi; portale che può leggersi come formato da due parti ben definite: la prima si caratterizza con una larga fascia di blocchi rettangolari lisci, che reggono l'architrave; superiormente compare una lunetta. La seconda, esterna, ha una mostra formata da una fascia tripartita che, partendo dal basamento, segue la ghiera dell'arco, decorata da motivi a dentelli ed a "catenella". La stessa mostra è composta da una cornice a gola interna, con foglie di acanto, da una fila di bugne sovrapposte, quadrangolari e lavorate a punta di diamante, da un pilastro decorato con ovoli. La data di costruzione del portale coincide con quella del monastero.

Il portale d'ingresso della Chiesa risale al periodo compreso tra il sec. XVII e il XVIII; d'impronta tardo-barocca, con decorazioni floreali a lunghi tralci, poste sull'architrave e fino alla base delle paraste, lavorate a scasso e sormontate da due capitelli compositi.

L'arco è ribassato e modanato ed è costituito da sottili superfici a rilievo sovrapposte, le quali, in alcuni tratti, fanno da sfondo alla decorazione esterna. Lo stile fa pensare alle forme utilizzate dagli scalpellini operanti nell'*hinterland* cosentino (Rogliano e Rende), per cui l'opera potrebbe essere attribuibile a questi. Sopra l'architrave compare un'edicola con l'immagine, probabile, di S. Bernardo da Chiaravalle. L'ipotesi si giustifica con il fatto che, nelle edicole esterne delle chiese dei monasteri, solitamente si effigia il fondatore dell'Ordine.

La porta in legno, a due battenti, è databile tra la fine del sec. XVI e i primi anni del XVII; è decorata con 14 riquadri rettangolari, dodici dei quali con ornamenti legati alle devozioni cistercensi e benedettine. Tra le chiare connotazioni iconografiche, si possono individuare S. Bernardo da Chiaravalle, a destra, con i simboli della Passione e S. Egidio, a sinistra. La chiesa ha un impianto ad aula unica, con un presbiterio quadrato, coperto da una pseudo-cupola impostata senza tamburo. Il primo elemento che si scorge, entrando nella chiesa, è il coro che si affaccia sull'aula. La volta che lo sovrasta è decorata con stucchi, che alternano superfici lisce e bianche a rilievi in oro. La decorazione, risalente alla fine del sec. XVII, o agli inizi del successivo, parte dal centro, dove è posta la Colomba, simbolo dello Spirito Santo e si allarga in centri concentrici; ci sono poi altre ripartizioni con una serie di ricchi intrecci floreali. Il prospetto ligneo del coro, appartenente ai primi anni del Settecento, è un'opera di intaglio che presenta un fitto decoro fitomorfo, insieme a teste d'angelo e festoni; il parapetto del coro è chiuso in alto da una griglia realizzata da archetti intrecciati. Mentre il pannello centrale reca il Nome di Gesù, in uno stemma raggiato, quelli laterali sono dipinti con un motivo floreale a medaglione.

Cuore del complesso, composto da chiesa, monastero e l'annesso palazzo Sersale, è il chiostro, di forma romboidale, forse scaturita dalla forma del lotto. I costruttori cosentini<sup>30</sup> hanno dimensionato lo spazio in modo da suddividere un lato in quattro moduli uguali. Il chiostro si compone di uno spazio porticato coperto da volte a crociera sui quattro lati e si sviluppa su due livelli coperti, tranne il versante adiacente la chiesa. I prospetti si caratterizzano dal susseguirsi delle quattro campate prima menzionate, con colonne in asse tra loro, pressoché identiche per tutti i quattro lati (alcune differenze tra i lati, derivano dall'assenza di ortogonalità).

---

<sup>30</sup> Si pensa che i lavori del Complesso delle Vergini siano stati guidati dal "capomastro" Domenico La Cava con l'ausilio di Pietro Celeste. I due furono abili costruttori della "Valle del Crati", grandi conoscitori delle più aggiornate tipologie e degli stili della loro epoca e della loro tradizione.

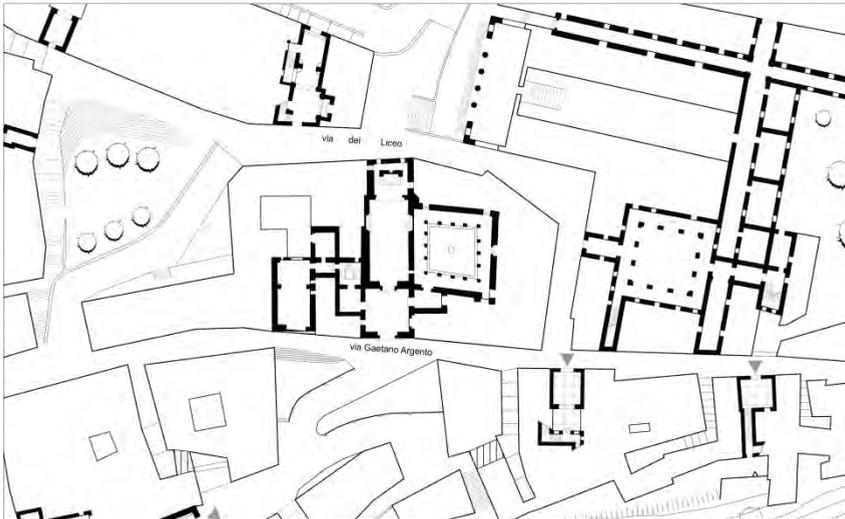


Fig. 5.13 – Pianta del complesso monastico di S. Maria delle Vergini (Cosenza).

#### *Metodo, strumenti e software*

Per la campagna di acquisizione del portale della chiesa di Santa Maria delle Vergini, si è usato un *laser scanner* TOF HDS 3000 della Leica Geosystems; mediante 3 *scanworld* (posizionamenti dello scanner) si è acquisita l'intera facciata della chiesa ed il suo immediato intorno, differenziando, l'acquisizione del solo portale, con una griglia di campionamento dal passo teorico di 0,5x0,5 cm, mentre l'acquisizione dell'intorno ha una griglia di 1x1 cm.

Per la registrazione delle nuvole di punti sono stati usati due differenti approcci e poi si sono confrontati i risultati: nel primo, l'allineamento è guidato da *target* piani ad elevata riflettanza, distribuiti efficacemente nella scena e ripresi dai tre posizionamenti del *laser*. Per la modesta dimensione dell'area interessata dal rilevamento e grazie alla particolare disposizione dei *target*, non si è fatto uso della stazione totale.

Il *software* utilizzato è il Cyclone 5.2 della Leica; l'allineamento delle tre scansioni, elaborato con il modulo *registrer* del *software*, è stato eseguito in modalità automatica, impostando lo *scanworld*, ripreso dalla posizione centrale rispetto al portone, come *scanworld* di riferimento e rototraslando gli altri su questo. Questa operazione è stata possibile, poiché in tutte le scansioni i *target* piani non sono stati mai spostati. L'errore di allineamento massimo, sui *target*, è risultato pari a 3 mm.

Nel secondo approccio, si sono filtrate e registrate le riprese, eseguendo un allineamento preliminare manuale, mediante punti comuni a due scansioni adiacenti ed, in seguito, un allineamento globale automatico di tutte le scansioni, per minimizzare l'errore residuo; per questa seconda procedura si è adoperato il *software* “*Geomagic Studio 12*”.

In entrambi i casi, abbiamo verificato la qualità metrica dei due allineamenti<sup>31</sup>, confrontando la posizione di una serie di punti significativi presi, sia sulla nuvola di punti, che *in situ*. Lo scostamento massimo registrato risulta essere di circa 5 mm.

La nuvola di punti, unita in un unico sistema di riferimento, esportata in “*Geomagic Studio 12*”, è stata usata sia come modello di riferimento, che di confronto con le altre nuvole di punti ottenute da fotografie; sempre in *Geomagic Studio 12* si è realizzato un modello poligonale del portale, avente circa 1.6 milioni di punti e 3.2 milioni di triangoli.



Fig. 5.14 – Portale della chiesa di *S. Maria delle Vergini* (Cosenza). Nuvola di punti da *Laser scanner* (1.6 milioni di punti.) e modello poligonale costituito da 2 milioni di triangoli.

---

<sup>31</sup> I due approcci hanno dato esiti sostanzialmente simili e non abbiamo registrato differenze significative dei risultati ottenuti.

Il rilievo fotografico ha utilizzato una camera digitale *reflex* Nikon D80, con un sensore di 3872x2592 *pixel*, ossia di 10 *mega-pixel*, alla risoluzione predefinita. Per la restituzione del modello si sono impiegati due sistemi *web-service* (ARC3D e Autodesk/123D Catch) e il *software* commerciale Agisoft *PhotoScan 0.83* (64 bit).

Si è acquisita, una sequenza di 11 immagini convergenti, utilizzando un obiettivo grandangolare con una distanza focale nominale di 12 mm; tutte le immagini sono state riprese da una distanza di circa 5 m, scattando le foto su un percorso a semicirconferenza. I fotogrammi sono stati tutti salvati nel formato grezzo RAW/NEF, per essere in seguito processati.

Il sistema ARC3D, permette di inviare le immagini da processare su *server* dedicati e restituisce delle *rang maps* di profondità per ciascuna immagine. Per la creazione del modello si è utilizzato MeshLab, attraverso il quale si sono importate le *rang maps* create da ARC3D, unite in un unico modello.



Fig. 5.15 – Portale della chiesa di S. Maria delle Vergini (Cosenza). Nuvola di punti, ottenuta da fotogrammi processati con il sistema web ARC3D ed elaborando i dati restituiti con MeshLab (1.5 milioni di punti); il modello poligonale elaborato dalla nuvola di punti, è stato realizzato in Geomagic Studio 12 ed è composto da 2.5 milioni di triangoli.

Per eseguire il confronto con i vari modelli ottenuti con i diversi *software*, si è scelto di esportare il modello creato nel formato “.PLY” che, importato in “*Geomagic Studio 12*”, può essere scalato, allineato con il modello di riferimento e sottoposto a confronto. La nuvola di punti della parte inerente il portale, esportata in *Geomagic Studio 12*, è costituita da 1.5 milioni di punti ed ha restituito un modello poligonale composto da 3 milioni di triangoli. In figura 5.15 è riportata la nuvola di punti ed il relativo modello ottenuto, del solo portale.

Il servizio *web* Autodesk/123D Catch, funziona in modo simile ad ARC3D, cioè ci si autentica, si inoltrano le foto da processare e si ottiene un modello poligonale della scena in oggetto; il modello ottenuto è stato esportato nel formato “.OBJ”, per essere a sua volta importato in *Geomagic Studio 12* e, successivamente, per essere allineato con il modello di riferimento e sottoposto a confronto. Il modello restituito è composto da 63420 triangoli.

In figura 5.16 viene riportato il modello ottenuto, impiegando sempre le stesse fotografie usate anche con gli altri *software*, con il servizio *web* Autodesk/123D Catch; il modello presenta parti con lacune ed ha una campionatura di triangoli inferiore.



Fig. 5.16 – Portale della chiesa di S. Maria delle Vergini (Cosenza). Modello poligonale del solo portale, ottenuto con il servizio *web* Autodesk/123D Catch, composto da 63240 triangoli.

In *Agisoft PhotoScan*, dopo aver allineato le immagini ed estratta la geometria (corrispondenze su ogni fotogramma), si ottiene il modello, che esportato nel formato “.PLY”, è stato importato in *Geomagic Studio 12*, per essere scalato, allineato con il modello di riferimento e confrontato; la nuvola di punti della parte messa a confronto, è costituita da 1 milione di punti; da questa si è generato un modello poligonale, composto da 2 milioni di triangoli.



Fig. 5.17 – Portale della chiesa di S. Maria delle Vergini (Cosenza). Nuvola di punti, ottenuta da fotogrammi processati con *Agisoft PhotoScan* (1 milioni di punti); il modello poligonale, è stato realizzato in *Geomagic Studio 12* ed è composto da 2 milioni di triangoli.

Al fine di valutare la precisione metrica dei modelli 3D, si è eseguita una comparazione, utilizzando il modello del *laser scanner* 3D, come riferimento.

Il flusso di lavoro è stato:

- riduzione scalare dei modelli, con una misura di riferimento;
- allineamento dei modelli, in un sistema globale di riferimento;
- confronto del modello di riferimento, con i modelli ottenuti con gli altri sistemi.

Avendo scelto come riferimento, il modello ottenuto dalla scansione *laser*, da questo si sono prese le misure per scalare tutti i modelli ottenuti da

immagini, confrontando direttamente la distanza tra due punti e moltiplicando per il relativo fattore di scala.<sup>32</sup>

Tutte le operazioni di messa in scala e di allineamento delle varie nuvole di punti, nel medesimo sistema di riferimento, si sono eseguite con *Geomagic Studio 12*. Ciascuno dei tre modelli, è stato allineato con quello di riferimento, eseguendo un allineamento preliminare tramite selezione manuale di sette o otto punti corrispondenti<sup>33</sup>. Il passaggio manuale è stato seguito da un'implementazione ICP, per raggiungere la raffinatezza di registrazione automatica, utile per minimizzare l'errore residuo.

La terza fase, ossia il confronto, è stato eseguito sempre in *Geomagic Studio 12*; i confronti effettuati sono stati:

- confronto tra nuvole di punti ottenute da *Agisoft PhotoScan* e *ARC3D/MeshLab*;
- confronto tra nuvole di punti ottenute un *laser scanner* TOF Leica HDS 3000 e le nuvole di punti ottenute con tecniche *Structure and motion*<sup>34</sup>.

#### *Riflessioni e risultati raggiunti*

- *Confronto tra nuvole di punti ottenute da Agisoft PhotoScan e ARC3D/MeshLab*

Come si può notare dall'immagine riportata in figura 5.18, lo scostamento massimo tra le due nuvole di punti confrontate, risulta abbastanza contenuto, mentre lo scostamento medio risulta pari a  $\pm 0,22$  cm, con una deviazione *standard* di 0,65 cm.

---

<sup>32</sup> La messa in scala del modello ottenuto tramite immagini, è l'unica operazione che l'utente può controllare; considerata l'importanza dell'operazione, è opportuno che la misura di riferimento sia "grande" e che si abbiano altre due misure di controllo, prese nelle altre due direzioni.

<sup>33</sup> In queste circostanze, l'importanza di avere nuvole di punti "colorate" è notevole. In caso di acquisizione con *laser scanner* senza una fotocamera digitale integrata, la fase di allineamento e successiva registrazione manuale diventa molto complicata, se non quasi impossibile.

<sup>34</sup> Il confronto tra i dati processati da servizio web Autodesk/123D Catch e la nuvola di punti, ottenuta da *laser scanner*, non è stato possibile farlo, perché Autodesk/123D Catch, non restituisce una nuvola di punti, ma una *mesh*.

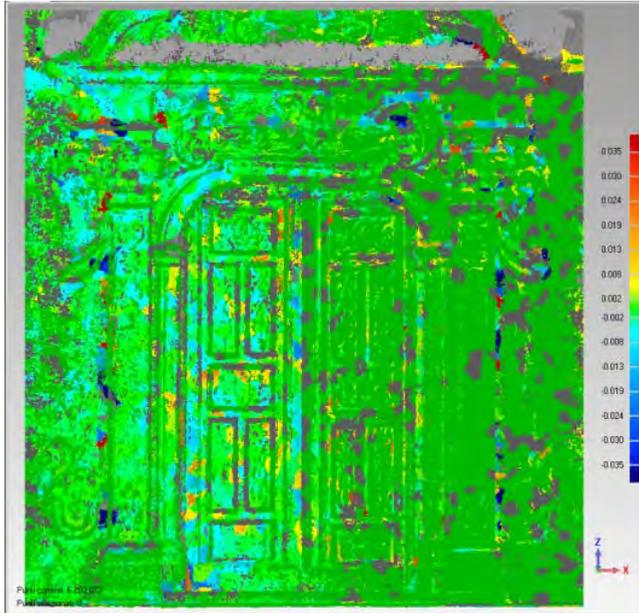


Fig. 5.18 – Portale della chiesa di S. Maria delle Vergini (Cosenza). Confronto tra nuvole di punti ottenute da Agisoft PhotoScan e ARC3D/MeshLab.

Il confronto effettuato è ottenuto impiegando gli algoritmi automatici per analizzare le deviazioni degli scostamenti - estesi alle intere nuvole di punti - tramite una gamma di colori, che associa a ciascun colore un valore numerico degli scostamenti.

Dai modelli poligonali, elaborati dalle rispettive nuvole di punti, si sono estratti in automatico dei profili - uno verticale e uno orizzontale - e confrontati tra loro; su cui sono stati evidenziati i valori massimi di scostamento. I valori degli scostamenti massimi sono comparabili con quelli che emergono dal confronto tra le nuvole di punti ottenute con gli stessi sistemi.

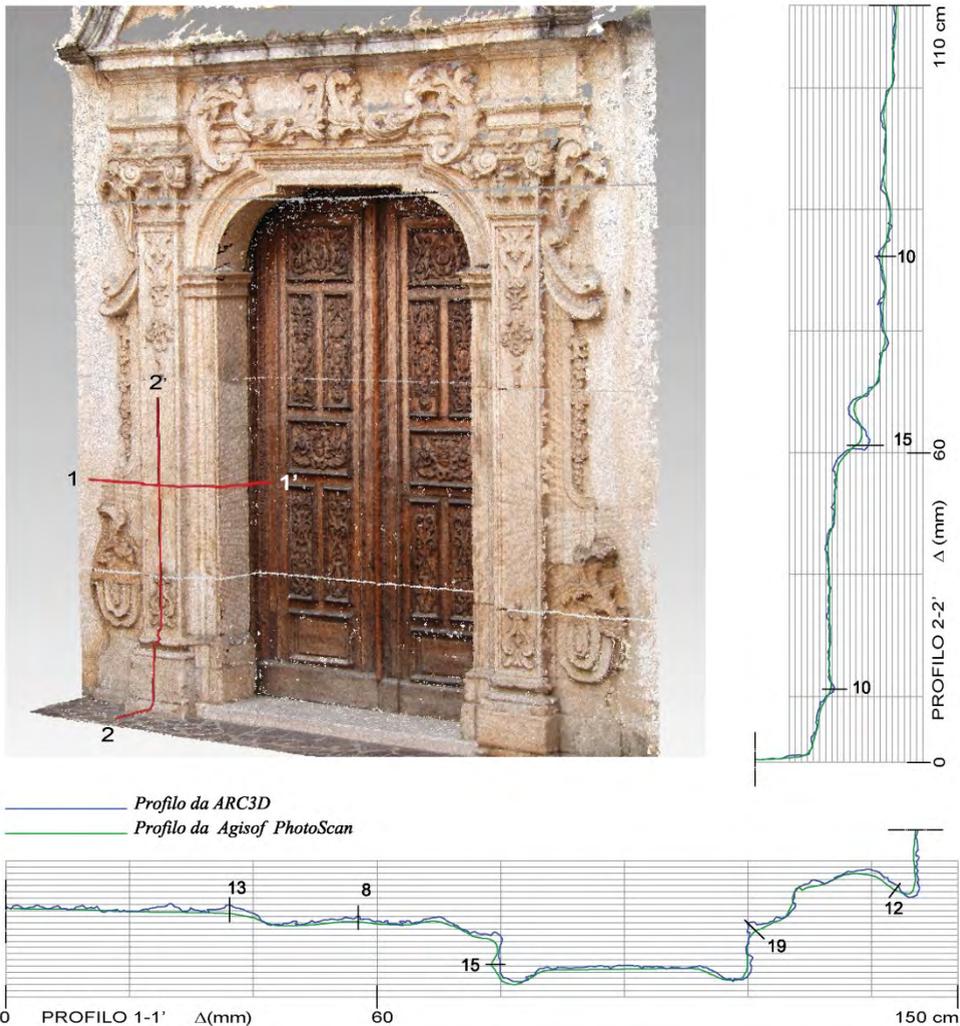


Fig. 5.19 – Portale della chiesa di S. Maria delle Vergini (Cosenza). Modello poligonale del portale, ottenuto da ARC3D/MeshLab. I due profili riportati in figura, estratti in automatico dai rispettivi modelli poligonali, ottenuti con ARC3D/MeshLab e Agisoft PhotoScan, vengono confrontati tra loro per evidenziare gli scostamenti massimi registrati in alcuni punti; questi risultano comparabili con quelli che emergono dal confronto tra le nuvole di punti, ottenute con gli stessi software.

- Confronto tra nuvole di punti ottenute con laser scanner TOF Leica HDS 3000 e le nuvole di punti ottenute con tecniche IBM

I valori elevati degli scostamenti massimi, riscontrati tra la nuvola di punti ottenuta da laser scanner e le nuvole di punti ottenute da tecniche Structure

*and motion*, molto probabilmente sono dovuti ad artefatti della misura, sono dovuti, cioè, a particolari scostamenti della superficie (dislivelli improvvisi, bruschi cambi di orientamento della superficie, ...), che incidono sul modo in cui la posizione dello *spot* viene rilevata dal *laser scanner* TOF, dando luogo a misurazioni non corrette.

Lo scostamento medio risulta essere nei due casi di  $\pm 0,15$  cm, nel confronto nuvola di punti ottenuta da *laser scanner* e nuvola di punti elaborata da ARC3D, mentre diventa il doppio ( $\pm 0,34$  cm), nel confronto nuvola di punti ottenuta da *laser scanner* e nuvola di punti elaborata da *Agisoft PhotoScan*.

La deviazione *standard* (precisione) è un indice della dispersione delle misure, intorno al loro valore medio: a valori più bassi della deviazione *standard*, corrisponde una maggiore affidabilità della misura.

Il valore registrato da *Geomagic Studio 12* risulta essere di 0,43 cm per ARC3D, contro un valore di 0,68 cm di *Agisoft PhotoScan*; se si tiene come riferimento il dato proveniente da *laser scanner*. Come sinora fatto, riguardo alla misura risulta maggiormente affidabile l'elaborazione della nuvola di punti eseguita con ARC3D, anche se le differenze tra gli spostamenti relativi tra le due nuvole sono contenuti.

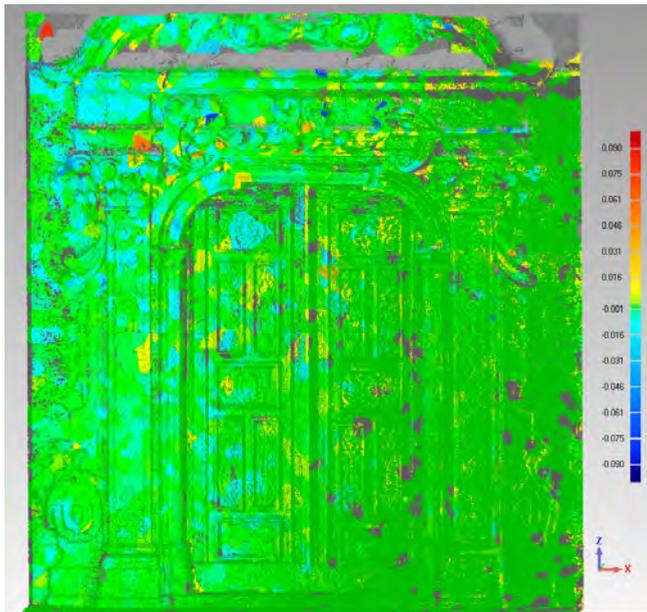


Fig. 5.20 – Portale della chiesa di S. Maria delle Vergini (Cosenza). Confronto tra la nuvola di punti ottenuta da *laser scanner* Leica HDS3000 e quella ottenuta da ARC3D/MeshLab.

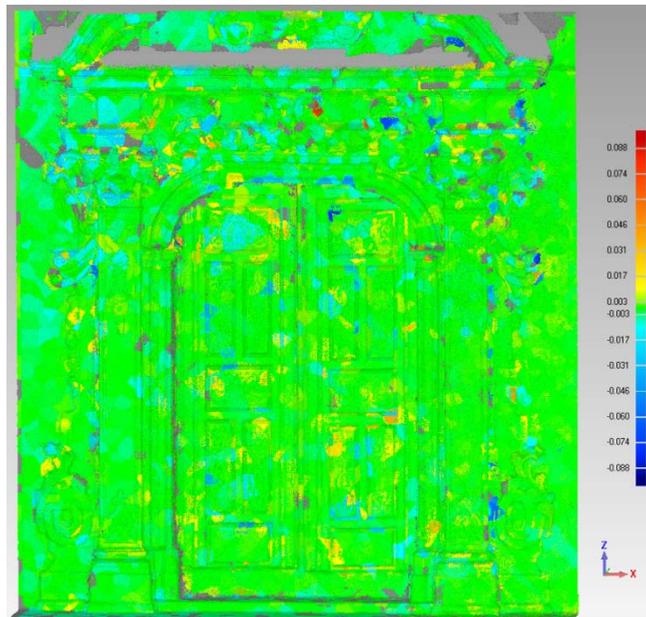


Fig. 5.21 – Portale della chiesa di S. Maria delle Vergini (Cosenza). Confronto tra la nuvola ottenuta da laser scanner Leica HDS3000 e quella ottenuta da Agisoft PhotoScan.

Nelle figure 5.20 e 5.21, sono riportati il confronto esteso sulle intere nuvole di punti, rappresentato tramite una gamma di colori per analizzare le deviazioni degli scostamenti, tra le nuvole considerate.

Nella tabella 5.2, sono riepilogati i valori degli scostamenti emersi dal confronto tra la nuvola di punti ottenuta dal *laser*, presa come riferimento e le due nuvole di punti ottenute in automatico da tecniche *Structure and motion*.

Scostamenti (cm)	ARC3D 1.5 milioni di pt.	Agisoft. 1 milione di pt.
Scost. Max	9.1	8.76
Scost.Medio	0.15	0.34
Deviazione standard	0.43	0.68

Tab. 5.2 – Confronto tra la nuvola di punti (1,6 milioni di punti) ottenuta con laser scanner TOF Leica HDS 3000 e le nuvole di punti ottenute con ARC3D/MeshLab e Agisoft PhotoScan.

Di seguito sono riportati i profili estratti in automatico ed il confronto tra loro; sui profili sono evidenziati i valori massimi di scostamento.

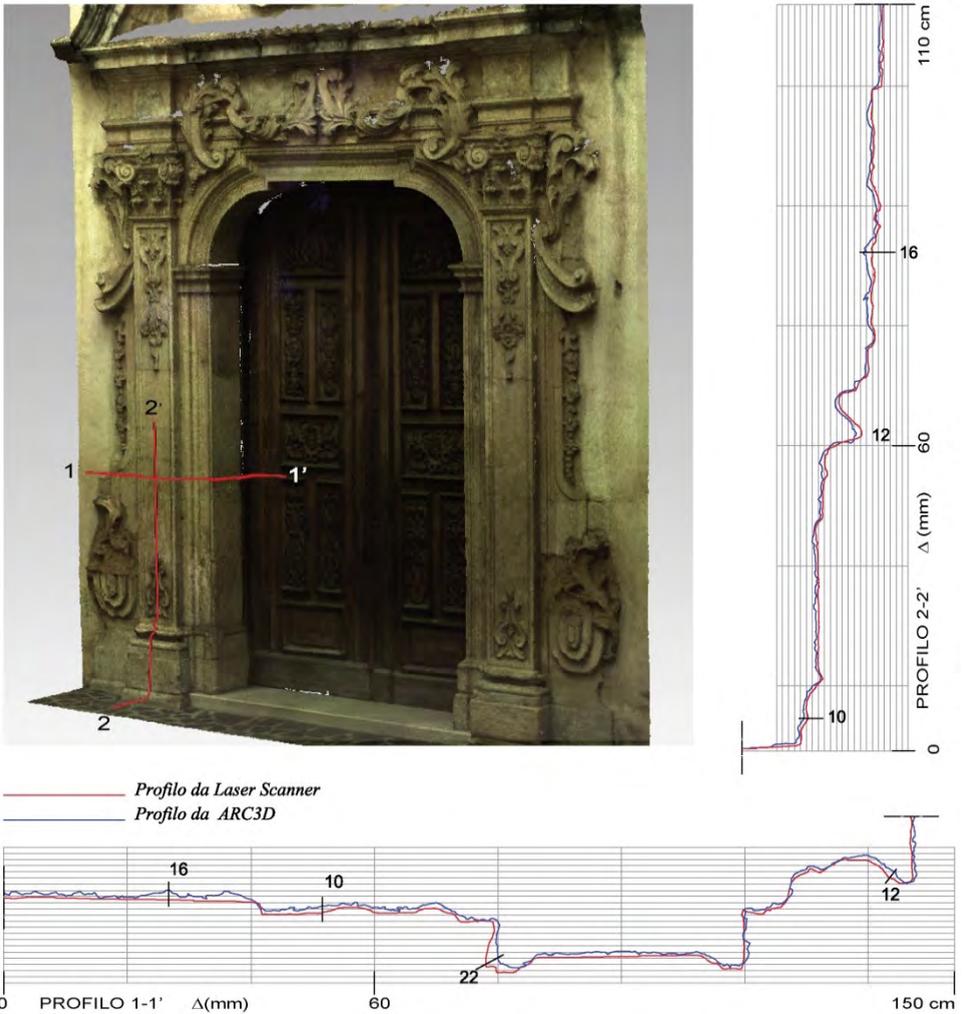


Fig. 5.22 – Portale della chiesa di *S. Maria delle Vergini* (Cosenza). Modello poligonale del portale, ottenuto dalla nuvola di punti acquisita con laser scanner. I due profili riportati in figura, estratti in automatico dai rispettivi modelli poligonali, ottenuti da laser scanner e da ARC3D/MeshLab, sono confrontati tra loro; sono anche evidenziati gli scostamenti massimi registrati in alcuni punti. Tali scostamenti sono comparabili con quelli medi che emergono dal confronto tra le nuvole di punti, ottenute con gli stessi sistemi.

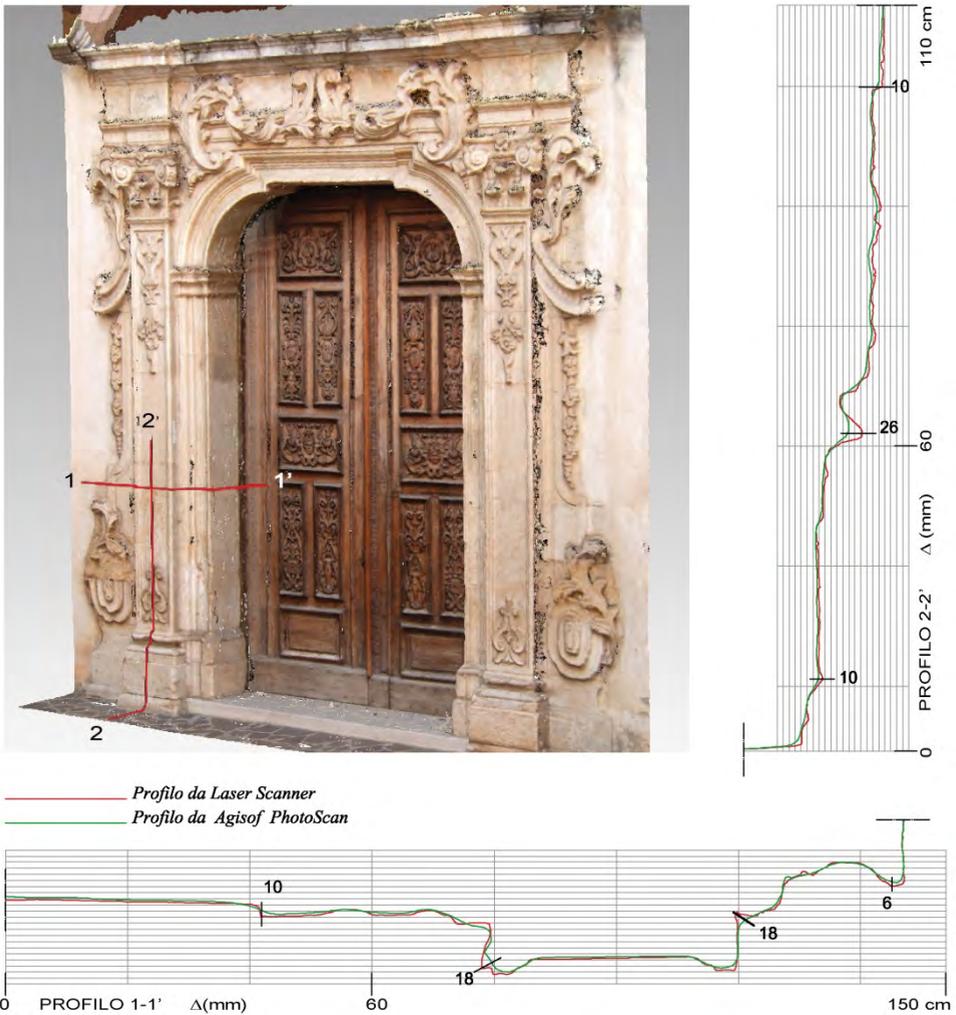


Fig. 5.23 – Portale della chiesa di S. Maria delle Vergini (Cosenza). Modello poligonale del portale ottenuto con *Agisoft PhotoScan*. I due profili riportati in figura, estratti in automatico dai rispettivi modelli poligonali, ottenuti da laser e da *Agisoft PhotoScan*, sono confrontati tra loro e sono anche evidenziati gli scostamenti massimi registrati in alcuni punti. Tali scostamenti sono comparabili con gli scostamenti medi che emergono dal confronto tra le nuvole di punti, ottenute con gli stessi sistemi.

*I mosaici della chiesa di Sant'Adriano (San Demetrio Corone, Cosenza)*

*Descrizione del manufatto*

La chiesa di Sant'Adriano, dedicata ai santi Adriano e Natalia, vissuti a cavallo tra il terzo ed il quarto secolo d.C., in Asia minore, è uno dei monumenti di valore dell'arte normanno-bizantina, in Calabria. L'edificio monastico sorse intorno al 995, per opera di Nicola Malena, divenuto poi San Nilo, sui resti di un antico sacello dei monaci orientali, nei pressi di Rossano.

San Nilo e i suoi discepoli, costruirono una chiesetta bizantina, dedicata ai santi martiri, ma della primitiva chiesa non restano tracce. Alcuni ritengono che sia stata inglobata nella nuova costruzione del tempio, avvenuta nei secoli XI e XII. L'arrivo dei saraceni in Calabria, non risparmiò il piccolo monastero, che fu distrutto e la successiva ricostruzione avvenne per opera di un altro monaco basiliano, San Vitale da Castronovo, nel 980.

Nel 1115 il monastero divenne sede di un Archimandrita e dotata di ricchi feudi; la stessa chiesa fu abbellita con mosaici e affreschi in stile bizantino.

Nel 1794, con un decreto di Ferdinando IV di Borbone, la chiesa fu tolta ai monaci ed assegnata con tutti i suoi beni al Collegio italo-greco, trasferitosi a San Demetrio Corona. Con il passaggio di proprietà, anche la chiesa subì profondi rifacimenti, che ne alterarono le caratteristiche, come l'addossamento del fabbricato del Collegio italo-greco sulla facciata principale e l'ampliamento del medesimo istituto sulla facciata sud-est.

La costruzione del collegio a ridosso della facciata principale, provocò la chiusura dell'ingresso principale e l'ostruzione di tre finestre, poste sulla medesima facciata; in tale occasione fu anche realizzato l'attuale portone d'ingresso, sul lato sud-est della chiesa.

Negli anni Ottanta viene eliminato il fabbricato del collegio, addossato alla chiesa e la facciata principale ritorna allo stato originario, con l'apertura delle tre finestre ostruite ed i segni sulla muratura dell'antico portone d'ingresso. L'interno della chiesa presenta una combinazione di elementi contrastanti, sia dal punto di vista cronologico, che da quello stilistico (dal romanico, al bizantino, al barocco).

La chiesa è composta da tre navate (25x13.6 m), divisa da due file di colonne e sei pilastri reggenti degli archi di conformazione e dimensioni diverse tra loro.

Di stile bizantino, probabilmente del XI e XII secolo, sono gli affreschi, di cui è ricca la chiesa; affreschi che furono riportati alla luce nel 1939.

L'intradosso di ogni arco, che separa le navate laterali da quella centrale, presenta figure di Santi, vescovi e martiri eremiti di identità sconosciuta, per la mancanza di iscrizioni.

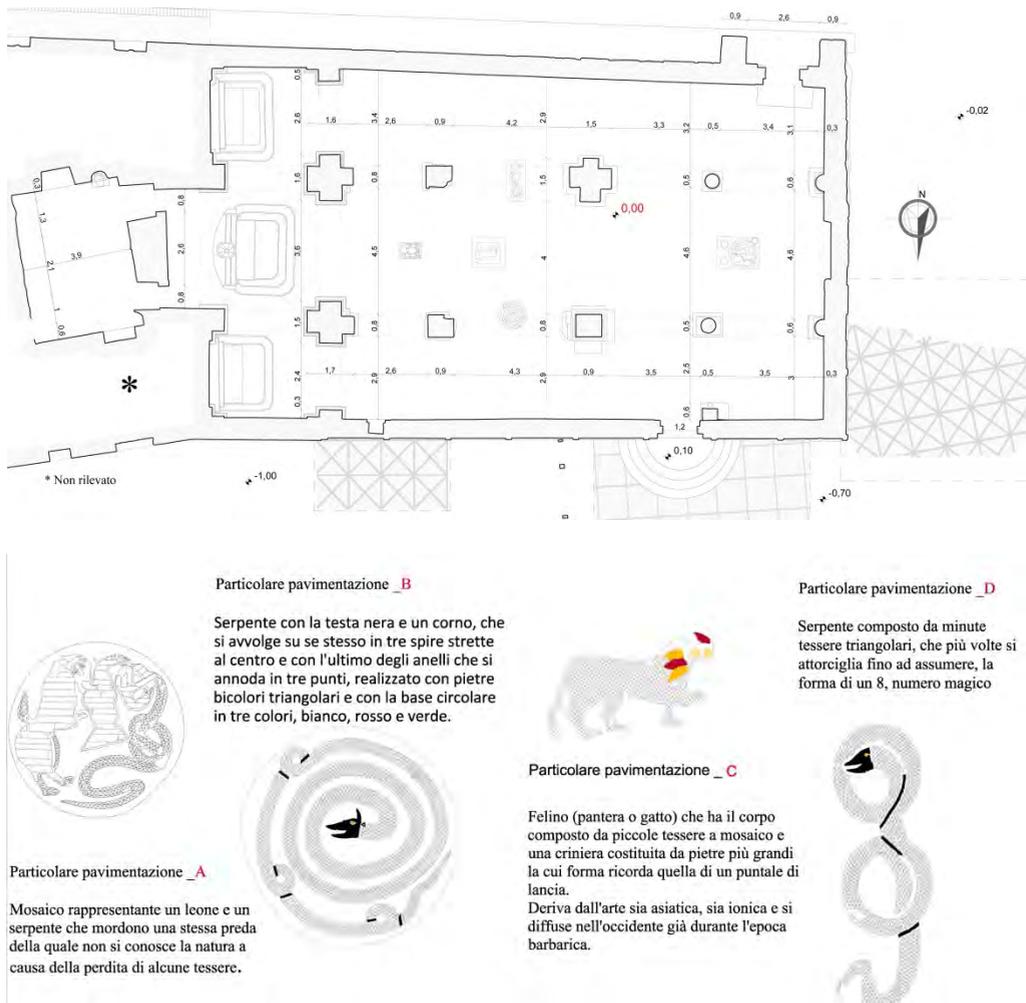


Fig. 5.24 – Chiesa di Sant’Adriano (San Demetrio Corone, Cosenza); pianta e particolari della pavimentazione.

Il pavimento è in *opus sectile*, realizzato tra la fine del XII secolo e l’inizio del XIII, originariamente doveva estendersi su tutta l’area della chiesa; oggi è

presente fino all'iconostasi bizantina. Tra grandi riquadri marmorei e formelle a disegno geometrico, compaiono quattro lastre con animali, i cui corpi sono realizzati con un motivo 'a scacchiera', alternando tessere quadrate o triangolari colorate e tessere bianche, su un fondo chiaro. Le immagini rappresentano un leone ed un serpente nell'atto di contendersi una preda ed un serpente con il corpo a spirale, inseriti su due lastre circolari; ci sono poi una pantera ed un altro serpente dal corpo annodato, con due code, inseriti su lastre rettangolari.

#### *Metodo, strumenti e software*

Per rilevare il complesso architettonico (chiesa, ambienti contermini e immediato intorno) abbiamo usato prevalentemente il *laser scanner* TOF ad impulsi (Leica HDS 3000) integrato, con metodi e strumenti tradizionali (stazione totale Leica TCR 407p, metodo diretto, fotocamera)<sup>35</sup>. L'interno e l'esterno della chiesa sono stati campionati con un passo teorico di 1x1 cm, arrivando fino a 0,5x0,5 cm, nella scansione dei particolari; mentre il terreno circostante è stato ripreso con una "maglia" teorica di 5x5 cm<sup>36</sup>. L'allineamento dei singoli *scanworld* è stato guidato da *target* piani ad elevata riflettanza, distribuiti efficacemente nella scena e ripresi dai 17 posizionamenti del laser.

Una mirata pianificazione della fase di rilevamento, ha permesso la registrazione dei 17 *scanworld*, senza l'uso della stazione totale, grazie alla particolare disposizione dei *target* ed alle diverse zone di sovrapposizione dei differenti *scanworld*.

Il *software* utilizzato è il Cyclone 5.2 della Leica e l'allineamento dei 17 *scanworld*, è stato eseguito in modalità automatica, impostando uno *scanworld* come riferimento e rototraslando gli altri su questo; questa operazione è stata possibile, poiché in tutte le scansioni, i *target* piani non sono stati mai spostati.

Dalla nuvola di punti allineata, tramite l'applicativo Leica Cloudworks per Autocad, abbiamo realizzato gli elaborati tradizionali (piante, prospetti, sezioni e configurazione planimetrica generale). Tali elaborati, oltre a restituire gli

---

<sup>35</sup> Il rilevamento strumentale, integrato parzialmente con il metodo diretto e le riprese fotografiche, ha interessato: gli altari, alcuni particolari decorativi, gli affreschi e i mosaici.

<sup>36</sup> Per un totale di 17 *scanworld* ed una nuvola di circa 72 milioni di punti.

aspetti geometrici dimensionali, sono stati impiegati, come mezzo di confronto, con rilievi eseguiti in passato.

La successiva fase di *meshing* e di *editing* (condotta con *Geomagic Studio 12*) ha dato luogo ad un modello poligonale, con risoluzione geometrica di 1 cm; da questo abbiamo estratto, in maniera automatica, i profili ed altri elementi, necessari per ottenere un modello semplificato della chiesa, che supportato da fonti bibliografiche, ha permesso di simulare un'ipotesi ricostruttiva della chiesa originaria. La generazione del modello semplificato è stata eseguita tramite *Rhinoceros 4.0*.

Il rilievo fotografico è stato eseguito con una camera digitale *reflex* Nikon D80, con un sensore di 3872x2592 *pixel*, ossia di 10 *mega-pixel*, alla risoluzione predefinita. Per la restituzione del modello si sono impiegati due sistemi, il *web-sevice* *ARC3D* e il *software* commerciale *Agisoft PhotoScan 0.83 (64 bit)*.

In questa esperienza i modelli ottenuti da tecniche *Structure and motion*, sono stati quelli relativi a due particolari della pavimentazione in *opus sectile* e, precisamente, il mosaico rappresentante un leone e un serpente che mordono la stessa preda; preda di cui non si conosce la natura a causa della perdita di alcune tessere. Il secondo particolare è il mosaico di un serpente con il corpo a spirale.

Per la restituzione del mosaico relativo al leone ed al serpente, è stata scattata una sequenza di 19 immagini convergenti, utilizzando un obiettivo con distanza focale nominale di 18 mm; tutte le immagini sono state riprese da una distanza di circa 2 m, scattando le foto su un percorso a semicirconferenza. I fotogrammi sono stati tutti salvati nel formato grezzo RAW/NEF, per essere in seguito processati.

### *Riflessioni e risultati raggiunti*

La ripresa con *laser scanner*, è stata mirata ad acquisire la morfologia della chiesa con misurazioni precise e coordinate tra di loro, in modo di ottenere delle restituzioni (modelli, piante, prospetti, sezioni, ecc.), utili per verificare (o mettere in discussione) le ipotesi che si formulano sulla geometria degli antichi monumenti, ma anche sulla loro evoluzione storico-architettonica.

Nel caso specifico, l'acquisizione eseguita evidenzia i limiti degli strumenti

impiegati (*laser*) per la generazione di un modello geometrico, con un elevato grado di restituzione dell'apparenza visiva; nella chiesa sono, infatti, presenti elementi con una elevata diversità di materiali e colori (pavimento a mosaico, affreschi ...), che richiedono, necessariamente, l'integrazione con tecniche ottiche, per la restituzione dell'apparenza visiva.

Nella figura 5.25 è riportata la nuvola di punti acquisita con il *laser scanner* TOF ad impulsi (Leica HDS 3000) e il corrispondente modello elaborato con *Geomagic Studio 12*, del mosaico interessato; l'immagine mostra appunto i limiti dello strumento per tali applicazioni, rispetto ai modelli elaborati da fotogrammi.

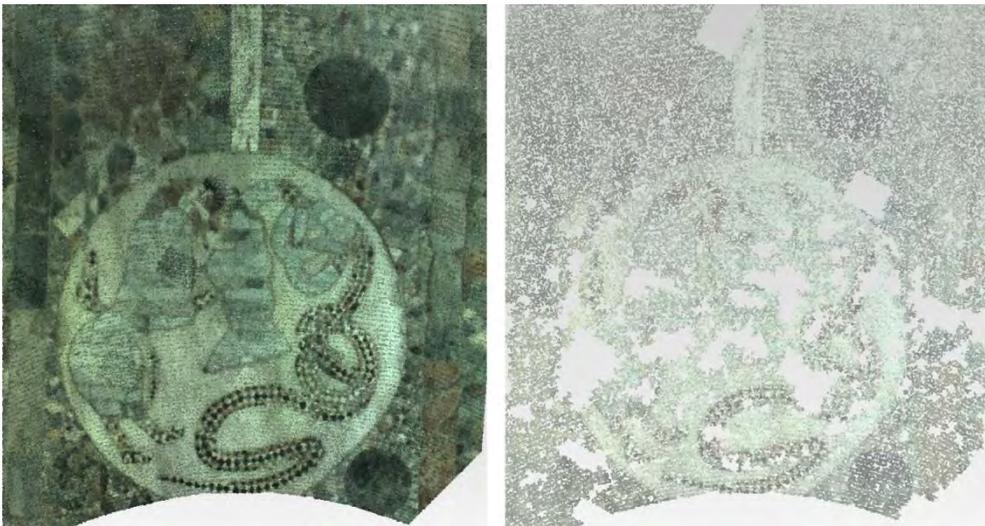


Fig. 5.25 – Chiesa di Sant'Adriano (San Demetrio Corone, Cosenza); nuvola di punti e relativo modello poligonale ottenuto, con acquisizione da *laser scanner* TOF (Leica HDS 3000), del mosaico con leone e serpente. L'immagine mostra i limiti di tale strumentazione, per la generazione di un modello geometrico con un elevato grado di restituzione dell'apparenza visiva.

Le *rang maps* di profondità, restituite dal sistema *web-sevice* ARC3D per ciascuna immagine, sono state elaborate con MeshLab, che ha fornito un modello, dell'area interessata dal mosaico, composto da 3 milioni di punti ed un modello poligonale costituito da 4.8 milioni di triangoli. Le stesse immagini elaborate con il *software* commerciale *Agisoft PhotoScan*, hanno restituito una nuvola di circa 1 milione di punti ed un corrispondente modello poligonale di 1.7 milioni di triangoli.

Nella figura 5.26 sono riportati i modelli poligonali del mosaico ottenuti con entrambi i sistemi.



Fig. 5.26 – Chiesa di Sant’Adriano (San Demetrio Corone, Cosenza); mesh (modello poligonale) del mosaico con leone e serpente ottenuta da ARC3D/MeshLab (sinistra) e da Agisoft PhotoScan (destra).

- *Confronto tra nuvole di punti ottenute da Agisoft PhotoScan e ARC3D*

Come si può notare dall’immagine riportata in figura 5.27, non risulta alcuna differenza tra le due nuvole di punti ottenute con i *software* sopra indicati.

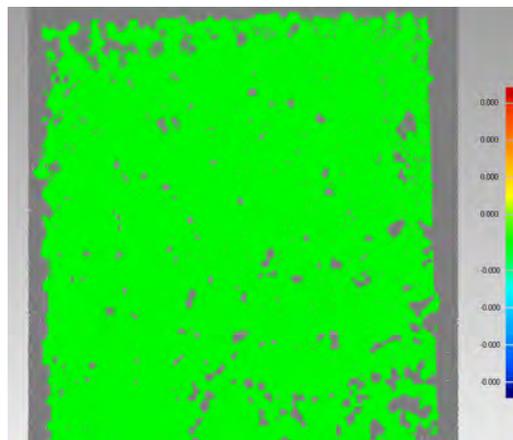


Fig. 5.27 – Chiesa di Sant’Adriano (San Demetrio Corone, Cosenza), mosaico con leone e serpente. Confronto tra nuvole di punti ottenute da Agisoft PhotoScan e ARC3D.

- *Confronto tra nuvole di punti ottenute con laser scanner TOF Leica HDS 3000 e le nuvole di punti ottenute con tecniche IBM.*

Tale confronto non ha dato risultati apprezzabili, in quanto le dimensioni degli elementi che costituiscono il tondo marmoreo intarsiato (in particolare le tessere triangolari del serpente e i giunti tra le varie tessere) risultano essere paragonabili, o inferiori, alla risoluzione di scansione massima del *laser scanner* impiegato; inoltre, risultano essere paragonabili alla precisione sul singolo punto (posizione 6 mm, distanza 4 mm): da questo caso, in particolare, risulta evidente la necessità di integrare la nuvola di punti ottenuta con *laser scanner TOF* con i sistemi *image based* (IBM).

Incidentalmente, ricordiamo che porzioni di pavimento, sostanzialmente piane, possono essere ricostruite geometricamente con pochi punti e l'utilizzo di un'immagine ad alta risoluzione, in grado di restituire l'apparenza visiva.

In altre parole, la gestione della *texture mapping*, in maniera indipendente rispetto alla risoluzione geometrica, permette di avere:

- superfici geometricamente semplici, definite con pochi punti, ma con una risoluzione elevata dell'immagine (quando le superfici sono sostanzialmente piane);
- superfici geometricamente complesse, definite con molti punti e con una risoluzione elevata dell'immagine (ad esempio bassorilievi ...);
- superfici geometricamente complesse, definite con pochi punti, ma con una risoluzione elevata dell'immagine (ad muro in mattoni a vista);
- superfici geometricamente complesse, definite con pochi punti e con una *texture standard* per rappresentare il materiale (ad esempio una pavimentazione a ciottoli).

Per la restituzione del mosaico rappresentante il serpente con il corpo a spirale sono stati acquisiti 20 fotogrammi impiegando un obiettivo con distanza focale nominale di 18 mm; le foto sono prese ad una distanza di circa 2 m e sono tutte convergenti verso il centro del mosaico. I fotogrammi sono stati tutti salvati nel formato grezzo RAW/NEF.

I fotogrammi elaborati con tecniche *Structure and motion*, hanno restituito un modello di 2.3 milioni di punti e di 3.5 milioni di triangoli, impiegando ARC3D/MesLab; impiegando il *software* commerciale *Agisoft PhotoScan*, si è

invece ottenuta una nuvola di 400.000 di punti, con un modello poligonale di 800.000 triangoli.

Nella figura 5.28 sono riportati i modelli del mosaico ottenuti con entrambi i sistemi.



Fig. 5.28 – Chiesa di Sant'Adriano (San Demetrio Corone, Cosenza); mesh (modello poligonale) del mosaico rappresentante il serpente con il corpo a spirale, ottenuta da ARC3D/MeshLab (sinistra) e da Agisoft PhotoScan (destra).



## Conclusioni

Il rilievo d'architettura, lo ricordiamo, non prevede mai la semplice riproduzione di un'opera, ma cerca di formulare restituzioni orientate secondo uno scopo e la loro adeguatezza risiede proprio nella rispondenza allo scopo che le determina. Non conta la completezza (cromatica, di particolari ...), conta invece la coerenza degli elementi che si scelgono per restituire e per rispondere alle motivazioni dell'azione di rilevamento. Inoltre, di volta in volta, cambiano gli operatori e con loro cambiano le aspettative con cui si eseguono i rilievi. Ancora, possono cambiare sia il contesto di riferimento che gli assetti dell'architettura ed è per simili motivi che, di solito, si svolgono nuovi rilievi e nuove restituzioni. Anche gli strumenti ed il linguaggio delle restituzioni si modificano nel tempo, proponendo nuove possibilità per analizzare ed interpretare un'opera, attraverso nuovi dispositivi grafici, schematizzazioni e modelli, mai sperimentati prima.

Le tecniche *range-based* e *image-based* consentono l'uso di maggiori precisioni nel descrivere le geometrie dell'oggetto reale, favorendo restituzioni "prossime" alle conformazioni reali, riproducendo, cioè, forme, colori e materiali delle conformazioni in esame ed offrendo, inoltre, l'opportunità di analizzare al meglio i fenomeni di degrado, l'analisi dei dissesti, l'interpretazione del quadro fessurativo, la lettura degli elevati, ecc.

La tecnologia *laser scanner*, come detto in precedenza, è una delle più importanti innovazioni tecnologiche dell'ultimo decennio nel campo del rilevamento; un'innovazione che ha determinato cambiamenti sostanziali sia rispetto al rilievo tradizionale – architettonico, archeologico ed urbano - che nel campo della modellazione 3D. Ha superato la tradizionale fotogrammetria per precisione e livello di automazione, portando a pensare di potere sostituirla completamente con scansioni 3D.

Negli ultimi anni, invece, la corrispondenza *multi-image* è divenuta, una delle aree di ricerca più attive, sia nella comunità dei "fotogrammetrici", che in

quella della *computer vision*: sono migliorate le procedure di estrazione automatica di corrispondenze da immagine ed i relativi algoritmi per la modellazione 3D. Ed è per tali motivi che l'uso dei metodi provenienti dalla fotogrammetria e dalla *computer vision*, ha iniziato a sollevare discussioni; in particolare, riguardo alla possibilità di produrre nuvole dense di punti - e quindi un attendibile modello 3D dell'opera in esame - analoghe a quelle che si ottengono con i sistemi *laser scanner*.

In questo lavoro di tesi, per affrontare il problema dell'integrazione dei differenti sistemi di acquisizione - in termini di precisione, di affidabilità del risultato ottenuto, di metodologia da impiegare e di facilità d'uso - abbiamo condotto direttamente una serie di sperimentazioni su casi-studio, che presentano caratteristiche differenti per dimensione, struttura geometrico-architettonica, stato di conservazione, epoca storica e così via, in ambienti esterni, con scene realistiche e, nello stesso tempo, diverse esperienze di laboratorio su oggetti artificiali con geometria nota.

I risultati ottenuti dalle varie esperienze, si possono sintetizzare in:

1/ esperienza su un oggetto caratterizzato da una conformazione assimilabile a 2,5 D; un oggetto in tre dimensioni, ma caratterizzato da profondità poco pronunciate e ricco di decorazioni.

Il confronto tra nuvole di punti ottenute da tecniche *image-based*, ha registrato scostamenti massimi contenuti e riportati in pochi punti isolati; gli scostamenti medi sono pari a  $\pm 0,22$  cm e, sostanzialmente, i due sistemi restituiscono lo stesso risultato.

2/ Confronto tra nuvole di punti ottenute con laser scanner TOF e le nuvole di punti ottenute con tecniche *image-based*.

In tale prova si è imposto come riferimento, il modello ottenuto da *laser* e si sono confrontati gli altri due su questo. I risultati presentano scostamenti accentuati, registrati su entrambi i modelli, in pochi punti isolati, probabilmente dovuti ad artefatti della misura; gli scostamenti medi sono contenuti nell'ordine di  $\pm 1,5$  cm, per la nuvola di punti restituita dal server ARC3D e di  $\pm 3,4$  cm, per la nuvola di punti restituita da *Agisoft PhotoScan*.

Valutando anche la deviazione *standard* (la precisione), la nuvola di punti restituita dal server ARC3D presenta scostamenti di 0,43 cm, contro 0,68 cm per quella restituita da *Agisoft PhotoScan*; anche se i due valori sono molto

prossimi, la restituzione eseguita con ARC3D presenta una maggiore precisione.

3/ Esperienza su un oggetto caratterizzato da una conformazione assimilabile a una superficie pseudo-piana, ricca di elementi colorati e decorati.

Il confronto tra nuvole di punti ottenute da tecniche *image-based*, non ha riportato grandi differenze; tale risultato ci permette di affermare, che entrambe le tecniche offrono, in tali casi, le medesime restituzioni.

4/ Il confronto tra nuvole di punti ottenute con *laser scanner* TOF e le nuvole di punti ottenute con tecniche *image-based*, non ha dato risultati apprezzabili, in quanto, molte delle dimensioni degli elementi che costituiscono la superficie in esame sono paragonabili o inferiori alla risoluzione di scansione massima del *laser scanner* impiegato, quindi da questo caso particolare risulta evidente la necessità di integrare la nuvola di punti ottenuta con *laser scanner* TOF, con i sistemi *image based*.

*Esperienze condotte in laboratorio, su un elemento planare con forte contrasto cromatico.*

5/ Il confronto tra nuvole di punti ottenute con tecniche *image-based*, non si è potuto eseguirlo, in quanto la ricostruzione da ARC3D, risulta incompleta e presenta molte parti mancanti; tali parti sono presenti dove non compaiono elementi contrastanti. Ci sono “buchi”, soprattutto, nelle parti bianche e sui quadrati neri; tale risultato si deve all’impossibilità di trovare, in queste parti, corrispondenze in fase di *matching*.

Gli stessi fotogrammi, processati con *Agisoft PhotoScan*, restituiscono, invece, una nuvola di punti completa dell’oggetto.

6/ Confronto indiretto, tra nuvola di punti ottenuta da *laser scanner*, nuvola di punti ottenuta con *Agisoft PhotoScan* e piano fittizio di riferimento.

Data l’impossibilità di eseguire un corretto allineamento delle due nuvole di punti, si sono confrontate, in modo indiretto, le due nuvole di punti con un piano fittizio di riferimento. I risultati numerici sono comparabili, ma ad un’analisi approfondita della distribuzione degli scostamenti, notiamo che le tecniche *image-based*, presentano problemi nelle parti dove non ci sono elementi caratteristici, necessari per individuare punti omologhi (corrispondenti

su due o più immagini) e per effettuare la fase di *matching* per la restituzione della nuvola di punti.

Dall'insieme delle prove eseguite, emergono alcune considerazioni importanti:

- la prima quasi ovvia, ma non per questo banale, è relativa alla necessità di conoscere a fondo le differenti tecniche, al fine di utilizzarle al meglio delle loro possibilità (potenzialità e limiti di ogni tecnica, “qualità” che sono in grado di restituire ...); al fine, cioè, di impiegare con rigore e competenza le potenzialità che le nuove tecnologie possono offrire; nuove tecnologie che spesso derivano da settori differenti da quelli dell'architettura e dei BB.CC. in genere e, dunque, presentano possibilità applicative indirette (di solito da verificare) per simili settori.

- La seconda, è relativa alla necessità di integrazione dei metodi di acquisizione oggi in uso; integrazione spesso indispensabile sia in termini di risultati (precisione, affidabilità, differenze di risultati ...), che di “prodotti” utili per il rilievo d'architettura. I *test* che abbiamo condotto, confermano tale necessità.

Relativamente ai “prodotti”, come in parte accennato in precedenza, ricordiamo che nell'azione di rilevamento, il tentativo che si compie non è solo quello di produrre immagini sempre più somiglianti con la realtà, né quello di conseguire misure sempre più accurate. “*Il rilevamento è ricerca ed interpretazione, applicazione razionale e creativa insieme*”<sup>37</sup>; serve per presentare l'architettura seguendone le forme, ma agisce anche seguendo le intenzioni di chi rileva e le indicazioni che, via via, il procedere delle elaborazioni propone.

Il rilievo d'architettura, cioè, ha valore per le restituzioni che permette di congegnare e per le attenuazioni/enfatizzazioni (di particolari, di assetti ...) che consente di compiere sulle componenti dell'opera e tra opera e contesto di riferimento; è per loro tramite – per il tramite delle elaborazioni che dà a vedere - che divengono possibili nuove interpretazioni e nuove conoscenze.

In altri termini, non bastano solo i progressi della tecnica per ottenere un

---

<sup>37</sup> De Sanctis A., *Restituzioni e conoscenza architettonica*, in “Strategie didattiche e di ricerca” (a cura di De Sanctis A, Fortunato G., Parise F.), Roma Edizioni Kappa 2011, p. 11.

rilievo di qualità; quello che, maggiormente, occorrono sono raffigurazioni evidenti dell'opera in esame, in grado di suscitare i contenuti, di cui la stessa opera è depositaria.

Da qui nasce l'esigenza dell'integrazione delle tecniche, ovvero dalla necessità di scegliere ed elaborare le restituzioni migliori (il più coerente, esatto, facile da ottenersi e manipolarsi ...), per operare ai fini dell'interpretazione e della conoscenza architettonica.

- L'ultima considerazione riguarda gli sviluppi della fotogrammetria e delle procedure automatiche per l'elaborazione di nuvole dense di punti; vale a dire, è relativa ai progressi della tecnica che, forse a breve, sarà in grado di sostituire in gran parte delle applicazioni l'uso dello *scanner* 3D.

Un problema che, evidentemente, si discosta dalle tematiche di questo lavoro, ma che può avere riflessi importanti sia per quanto riguarda l'approfondimento disciplinare (realizzazione di nuvole dense di punti, grado di affidabilità ...), che le applicazioni sul campo (facilità d'uso degli strumenti, economia delle operazioni, diffusione di procedure ...).

Le stesse prove che abbiamo effettuato, in laboratorio ed *in situ*, ci portano a considerare questa eventualità come possibile e, al tempo stesso, come foriera di nuove opportunità di risultati (ad esempio, quelli relativi alla creazione di modelli con *texture* ad alta ed altissima definizione).



## NOTA BIBLIOGRAFICA

*Teoria, riferimenti*

R. BARTHES, *La camera chiara. Nota sulla fotografia*, Torino Piccola Biblioteca Einaudi, 2003.

U. GALIMBERTI, *Psiche e tecne. L'uomo nell'età della tecnica*, Milano Feltrinelli, 2007.

R. L. GREGORY, *Occhio e cervello. La psicologia del vedere*, Milano, Raffaello Cortina, 1998 (ed. or. *Eye and Brain*, 1998).

U. NEISSER, *Conoscenza e realtà. Un esame critico del cognitivismo*, Bologna Il Mulino, 2000 (I ed. it. 1981; ed. or. *Cognition and Reality. Principles and Implications of Cognitive Psychology*, San Francisco, W. H. Freeman and Company, 1976).

*Analisi, rilievo, rappresentazione*

AA.VV., *Atti del Convegno: Il Rilievo dei Beni Architettonici per la Conservazione*, Roma MG Sistemi Editoriali, 2005.

M. DOCCI, D. MAESTRI, *Manuale di rilevamento architettonico e urbano*, Roma-Bari Editori Laterza, 2009.

M. DOCCI, D. MAESTRI, *Storia del rilevamento architettonico e urbano*, Roma-Bari Editori Laterza, 1993.

G. FORTUNATO, *The role of architectural representation for the analysis of the built. The 3D survey of San Pietro in Montorio*. Atti del convegno "X Congreso Internacional Expresión Gráfica aplicada a la Edificación Graphic Expression applied to Building", Apega (Spagna), 2,3,4 dicembre, 2010, Editorial Marfil, S. A.:Alicante (Spagna), 2010, Vol. I, pp. 451-461.

G. FORTUNATO, *Indagine e immagine. Integrazioni tra tecniche innovative e metodi di rappresentazione grafica: il rilievo del Battistero di Santa Severina (KR)*, Atti del convegno "convegno U.I.D.", Lerici (SP), 2008, Graphic Sector Genova, 2009, pp. 84-93.

R. MIGLIARI (a cura di), *Geometria descrittiva. Volume II - Tecniche e applicazioni*, Novara CittàStudiEdizioni De Agostini, 2009.

R. MIGLIARI, *Metodi e tecniche della rappresentazione digitale*, in M. UNALI, *New Lineamenta*, Roma Edizioni Kappa, 2009, pp.19-31.

A. SDEGNO, *Architettura e rappresentazione digitale*, Venezia Libreria Editrice Cafoscarina, 2002.

G. TESTA, A. DE SANCTIS, *Rappresentazione e architettura. Linguaggi per il rilievo ed il progetto*, Roma Gangemi Editore, 2003.

A. ZAPPANI, *Le tecniche ottiche di misura 3D nel rilevamento architettonico*, UNIVERSITA' DELLA CALABRIA 2012 (tesi di dottorato).

### *Tecniche range-based e image-based*

B. BENEDETTI, M. GAIANI, F. REMONDINO (a cura di), *Modelli digitali 3D in archeologia: il caso di Pompei*, Pisa Edizioni della Normale, 2010.

J.-A. BERARDIN, *Basic Theory on Surface Measurement Uncertainty of 3D Imaging Systems*, in Proceedings of "IS&T, Three-Dimensional Imaging Metrology Conference", San Jose (USA), 18-22 gennaio 2009.

F. BERNARDINI, H. RUSHMEIER, *The 3D model acquisition pipeline*, in "Computer Graphics Forum", vol. XXI, n. 2/2002, pp. 149-172.

F. BLAIS, *Review of 20 Years of Range Sensor Development*, in "Journal of Electronic Imaging", vol. XIII, n. 1/2004, pp. 231-243.

F. CROSILLA, S. DEQUAL (a cura di), *Laser scanning terrestre*, Udine CISM, 2006 (Collana di Geodesia e Cartografia, 8).

P. CIGNONI, M. CORSINI, M. DELLEPIANE, G. RANZUGLIA, M. VERGAUVEN, L. VAN GOOL, *MeshLab and Arc3D: Photo-Reconstruction and Processing 3D meshes*. EPOCH Conference on Open Digital Cultural Heritage Systems (2008), pp. 1-6.

F. CROSILLA, R. GALETTO (a cura di), *La tecnica del laser scanning: Teoria ed Applicazioni*, Udine CISM, 2003 (Collana di Geodesia e Cartografia, 7).

C. CUNDARI, *Fotogrammetria architettonica*, Napoli Edizione Kappa, 1983.

G. GODIN, M. RIOUX, J.-A. BERARDIN, M. LEVOY, L. COURNOYER, F. BLAIS, *An assessment of laser range measurement on marble surfaces*, in Proceedings of "V Conference on Optical 3D Measurement Techniques", Vienna (Austria), 1-4 ottobre 2001, pp. 49-56.

L. DE LUCA, *La fotomodellazione architettonica. Rilievo, modellazione, rappresentazione di edifici a partire da fotografie*, Palermo Dario Flaccovio Editore, 2011.

E. CHIAVONI, M. FILIPPI (a cura di), *Metodologie integrate per il rilievo, il disegno, la modellazione dell'architettura e della città*, (ricerca Prin 2007, coord. naz. Mario Docci), Roma Gangemi Editore, 2011.

M. DOCCI (a cura di), *Metodologie innovative integrate per il rilevamento dell'architettura e dell'ambiente*, (ricerca Cofin 2002, coord. naz. Mario Docci), Roma Gangemi Editore, 2005.

G. FORLANI, R. RONCELLA, F. REMONDINO, *Structure and Motion Reconstruction of*

*Short Mobile Mapping Image Sequences*. Optical 3-D Measurement Techniques, Vienna, Austria, 1 (2005), pp. 265-274.

J. FRYER, H. MITCHELL, J. CHANDLER (editors), *Applications of 3D Measurement from Images*, Dunbeath (UK), Whittles Publishing, 2007.

M. GAIANI (a cura di), *La rappresentazione riconfigurata. Un viaggio lungo il processo di produzione del progetto di disegno industriale*, Milano Edizioni POLI.DESIGN, 2006.

G. GUIDI, M. RUSSO, J.-A. BERARDIN, *Acquisizione 3D e modellazione poligonale*, Milano McGraw-Hill, 2010.

G. GUIDI, J.-A. BERARDIN, *Acquisizione 3D e modellazione poligonale. Dall'oggetto fisico al suo calco digitale*, Milano Edizioni POLI.DESIGN, 2004.

R. KOCH, M. P. L. VAN GOOL, *Realistic surface reconstruction of 3D scenes from uncalibrated image sequences*. Journal of visualization and computer animation, 11 (2000), pp.115-127.

R. MIGLIARI (a cura di), *Frontiere del rilievo. Dalla matita alle scansioni 3D*, Roma Gangemi Editore, 2001.

A. M. MANFERDINI, F. REMONDINO, S. BALDISSINI, M. GAIANI, B. BENEDETTI, *3D modeling and semantic classification of archaeological finds for management and visualization in 3D archaeological databases*. Proc. 14th Int. Conference on Virtual Systems and MultiMedia (VSMM), Limassol, Cyprus 2008, pp. 221-228.

M. POLLEFEYS, R. KOCH, M. VERGAUWEN, L. VAN GOOL, *Automated reconstruction of 3D scenes from sequences of images*. ISPRS The Journal of Photogrammetry & Remote Sensing, 55 (2000), pp. 251–267.

M. VERGAUWEN, LUC J. VAN GOOL, *Web-based 3D Reconstruction Service*. Mach. Vis. Appl., 17(6) (2006), pp 411-426.

G. VOSSELMAN, H.-G. MAAS (editors), *Airborne and Terrestrial Laser Scanning*, Dunbeath (UK), Whittles Publishing, 2010.