

Utilizzo di immagini acquisite da drone aereo per la ricostruzione tridimensionale realistica di un'area di interesse in tempo reale

Andrea Zingoni ^(a), Marco Diani ^(a), Giovanni Corsini ^(a)

^(a) Università di Pisa, Dipartimento di Ingegneria dell'Informazione, Via G. Caruso 16, 56122, Pisa, Italia,
andrea.zingoni@for.unipi.it, m.diani@iet.unipi.it, g.corsini@iet.unipi.it

Sommario

La ricostruzione tridimensionale di aree e oggetti di interesse da immagini acquisite da drone aereo riveste un ruolo fondamentale in numerosi campi, che vanno da quelli meno critici, come la pianificazione urbanistica e lo studio delle aree archeologiche, a quelli più critici, come il monitoraggio di disastri naturali e il *search & rescue*. Molte tra le tecniche proposte a tale scopo richiedono una strumentazione costosa o fanno uso di algoritmi onerosi in termini di carico computazionale o, ancora, necessitano di informazioni *a priori* non sempre reperibili in modo agevole, risultando dunque non adatte per determinate applicazioni o in alcune situazioni specifiche. Per ovviare a ciò, abbiamo implementato una tecnica che consente di creare automaticamente modelli 3D in tempo reale utilizzando soltanto una coppia di immagini della scena in esame, acquisite da una fotocamera compatta a basso costo montata sul drone, senza bisogno di alcuna informazione aggiuntiva. La tecnica proposta restituisce all'utente delle ricostruzioni tridimensionali dense e a colori reali, che danno l'impressione di essere fisicamente presenti all'interno della scena. Testata per monitorare l'avanzamento dei lavori in un cantiere edile, la tecnica è stata in grado di produrre modelli 3D realistici e facili da interpretare di aree e oggetti in meno di 1 secondo e con un'accuratezza sufficiente ad effettuare rilevazioni su larga scala.

Abstract

Tridimensional reconstruction of areas and objects of interest from images acquired from aerial drones plays a key role in numerous fields, from the less critical, like urban planning and archeological area survey, to the most critical, like natural disaster monitoring, and search & rescue. Many techniques proposed to deal with this issue require expensive instrumentation or use computationally costly algorithm or need some a priori information that are not always available. Thus, they are not suitable for certain applications or in some specific situations. In order to overcome these limitations, we designed a technique that allows automatically creating 3D models in real-time using only a couple of images of the scene one is interested in, acquired with a cheap compact camera mounted on the drone, without the needing for any additional information. The proposed technique outputs dense true-color 3D models, which give the impression to the user to be physically present within the scene. Tested for monitoring the progress of the works in a construction site, the technique has been capable to create realistic and easy-to-interpret 3D models of areas and objects in less than 1 second and with a sufficient accuracy to permit large-scale surveys.

Introduzione

Nel corso degli ultimi anni, l'utilizzo di aeromobili a pilotaggio remoto (APR) – comunemente noti come droni – è cresciuto in modo esponenziale, coinvolgendo settori di applicazione sempre più numerosi e differenti tra loro. Uno dei fattori chiave del successo degli APR risiede nella possibilità di equipaggiarli con sensori di immagini, i quali ne fanno un mezzo particolarmente efficace di

telerilevamento, sorveglianza e monitoraggio, soprattutto in presenza di aree inaccessibili o pericolose per l'uomo (Vasterling, Meyer, 2013).

Tra le numerose applicazioni di questo tipo, la ricostruzione tridimensionale può risultare di fondamentale importanza, in quanto è in grado di fornire la distribuzione spaziale delle informazioni a cui si è interessati. In particolare, ricostruzioni realistiche di ampie aree, di singole zone di interesse, di edifici, di oggetti ecc. si sono rivelate uno strumento di enorme utilità per l'analisi e la comprensione del fenomeno in esame.

In letteratura, sono state proposte numerose tecniche di generazione automatica di modelli tridimensionali usando un APR opportunamente equipaggiato. Alcune di queste mirano a raggiungere accuratezze estremamente elevate, preoccupandosi solo parzialmente del costo della strumentazione necessaria, che consiste, generalmente, in *lidar* (Bisheng, Chi, 2015) o *laserscanner* (Jutzi *et al.*, 2013), eventualmente affiancati da una o più fotocamere. Per limitare i costi, sono state proposte soluzioni che fanno uso solamente di una fotocamera. Tali soluzioni, però, per arrivare a ottenere una certa accuratezza, hanno bisogno di una lunga fase di elaborazione dei dati acquisiti, dovuta all'impiego di numerose immagini (Mayer, Bartelsen, 2008.) o all'elevato costo computazionale degli algoritmi di ricostruzione tridimensionale utilizzati (Huei-Hung *et al.*, 2011), fase che può durare diversi minuti o anche ore prima di produrre il modello 3D desiderato¹. Puntando a garantire una buona accuratezza senza sacrificare il tempo di *processing*, molte tecniche ricorrono all'uso di strumenti ausiliari (Wefelscheid, 2011), quali sistemi di navigazione inerziali (INS) e sistemi di posizionamento globale (GPS). Sfortunatamente, nella pratica, le versioni *low-cost* di tali strumenti non consentono di raggiungere l'accuratezza necessaria, imponendo l'acquisto di dispositivi piuttosto costosi. Un altro gruppo di tecniche di ricostruzione 3D si basa sullo sfruttamento di informazioni *a priori*, come per esempio i parametri intrinseci² della fotocamera (Se, 2009) o la posizione di alcuni punti di controllo (GCP) piazzati all'interno della scena da ricostruire e facilmente identificabili nelle immagini acquisite (D'Urso *et al.*, 2015). I parametri intrinseci, però, non sono sempre disponibili o calcolabili con precisione, mentre l'utilizzo di GCP richiede il posizionamento *in loco* di questi, non eseguibile in presenza di aree a rischio o inaccessibili. Tali tecniche, dunque, possono essere applicate soltanto in determinate situazioni.

A differenza degli approcci sopra elencati, abbiamo provato ad affrontare il problema della ricostruzione tridimensionale da APR contemporaneamente sia dal punto di vista del tempo computazionale che da quello del costo della strumentazione utilizzata, senza fare affidamento su informazioni note *a priori*. Abbiamo dunque elaborato una tecnica che consente di generare automaticamente modelli 3D realistici di aree o di singoli oggetti in tempo reale, usando solamente una fotocamera a basso costo montata su un APR di piccole dimensioni, senza conoscerne i parametri intrinseci e senza bisogno di posizionare GCP nell'area di interesse. La tecnica coinvolge l'intero processo di creazione del modello 3D, dalla scelta del *setup* di volo e della fotocamera, all'algoritmo di *processing* dei dati acquisiti, alla presentazione finale dell'output ottenuto. Il modello tridimensionale generato è una ricostruzione realistica della scena ripresa, che dà l'impressione all'utente di essere fisicamente presente al suo interno, la cui fruizione può avvenire tramite uno qualunque dei più comuni software di ricostruzione 3D, permettendo di variare in modo semplice, a seconda delle necessità, il punto di vista della scena.

Tecnica di ricostruzione tridimensionale

La tecnica di ricostruzione che proponiamo si basa sulla triangolazione dei punti della zona di interesse, partendo da due immagini di questa acquisite a due istanti τ_1 e τ_2 da una fotocamera montata su un APR che sorvola la scena. Non è necessario l'utilizzo di alcuna strumentazione

¹ Si noti che, utilizzando una coppia di fotocamere anziché una singola, è possibile raggiungere accuratezze migliori mantenendo contemporaneamente basso il tempo computazionale. Tale soluzione non è però ottimale per droni con una capacità di carico non elevata, che potrebbero sopportare il peso di una ma non di due fotocamere.

² Per parametri intrinseci di una fotocamera si intendono le dimensioni, orizzontale e verticale, del singolo fotorivelatore, l'*offset* tra l'asse ottico e il centro del sensore e la lunghezza focale della telecamera.

addizionale. La distanza focale (f), l'apertura del diaframma ($F\#$), il tempo di esposizione (t_{exp}) e la sensibilità del sensore (ISO) della fotocamera devono essere regolabili, condizione verificata anche su molte fotocamere compatte economiche presenti sul mercato. L'utilizzo di una fotocamera compatta dà luogo a un doppio vantaggio: da un lato il suo costo non è elevato, dall'altro, considerato il suo peso limitato, consente di utilizzare un APR di piccole dimensioni, dunque più pratico ed economico.

La procedura di ricostruzione consta di due fasi fondamentali: la definizione della modalità di acquisizione delle immagini, che consiste nel pianificare il volo dell'APR e nell'impostare opportunamente i parametri della fotocamera, per poter acquisire immagini *ad hoc* che consentano di massimizzare le *performance* globali; l'algoritmo di ricostruzione, che consente la creazione del modello 3D in tempo reale, servendosi delle sole due immagini dell'area di interesse acquisite.

Pianificazione della modalità di acquisizione delle immagini

La modalità di acquisizione delle immagini, riveste un ruolo fondamentale nell'intero processo di ricostruzione tridimensionale, influenzandone l'efficienza e fissandone i limiti. Basandoci su alcuni indici di prestazione tipici della ricostruzione 3D, andremo a vedere come sia possibile massimizzare le *performance* della tecnica proposta, mediante una corretta definizione della modalità di acquisizione. In Fig. 1 sono elencati gli indici di prestazione considerati. Per brevità, non abbiamo riportato le formule per il loro calcolo, ma abbiamo evidenziato direttamente la loro dipendenza dai parametri in gioco. Si noti che le *performance* sono tanto migliori quanto gli indici R e M diminuiscono – aumento della risoluzione – e A_{max} e H_{max} aumentano – aumento del volume ricostruibile della scena. Questo sarà dunque lo scopo di questa fase del lavoro.

Indice Di Prestazione \ Parametri in gioco	Quota di volo (h)	Velocità APR (V_{APR})	Distanza focale (f)	Baseline (b)	Apertura diaframma ($F\#$)	Tempo di esposizione (t_{exp})
Risoluzione a terra $R = (R_x, R_y)$	$R \propto h$	$R_x \propto V_{APRx}$ $R_y \propto V_{APRy}$	$R \propto 1/f$	nessuna dipendenza	$R \propto 1/F\#$	$R \propto t_{exp}$
Risoluzione modello 3D $M = (M_x, M_y, M_z)$	$M \propto h^2$	nessuna dipendenza	$M_x, M_y \propto 1/f^2$ $M_z \propto 1/f$	$M \propto 1/b$	nessuna dipendenza	nessuna dipendenza
Max area ricostruibile A_{max}	$A_{max} \propto h^2$	non direttamente dipendente	$A_{max} \propto 1/f^2$	$A_{max} \propto -b$	nessuna dipendenza	nessuna dipendenza
Max altezza ricostruibile H_{max}	$H_{max} \propto h$	non direttamente dipendente	$H_{max} \propto -f$	$H_{max} \propto -b$	nessuna dipendenza	nessuna dipendenza

Figura 1 – Indici prestazionali della ricostruzione tridimensionale e loro dipendenza dai parametri in gioco.

Prima di andare a settare correttamente i parametri, montiamo la fotocamera sull'APR in modo tale che guardi al *nadir*, semplificando così la geometria del problema e, contemporaneamente, assicurando che la distanza di campionamento a terra sia la stessa nelle due immagini acquisite. Ciò produce un miglioramento sia della R media che dell'accuratezza dell'algoritmo di ricostruzione, dal momento che questo si basa, come vedremo in seguito, sul riconoscimento di *feature* simili nelle due immagini. Dopodiché, orientiamo la fotocamera in modo tale che la dimensione maggiore del sensore sia parallela alla direzione di volo dell'APR. Grazie a questo semplice espediente, è possibile aumentare A_{max} e H_{max} senza dover modificare alcun parametro.

Il primo parametro su cui agire è $F\#$, che, se incrementato, diminuisce la sfocatura ottica, migliorando la R . $F\#$ non influisce su altri indici di prestazione, ma il suo incremento dà luogo a immagini più scure, problema non da poco in presenza di scarsa illuminazione. Per ovviare a ciò è possibile incrementare t_{exp} o ISO . La prima scelta non è raccomandabile poiché un tempo di esposizione alto genera una sfocatura dovuta al movimento, aumentando R . La seconda, invece, dà luogo a un aumento della rumorosità del sensore, di entità comunque lieve. Risultando il secondo

problema meno critico, scegliamo un t_{exp} basso e aumentiamo la *ISO*. Prima di capire come settare gli altri parametri, è opportuno notare che alcuni di questi sono vincolati da limitazioni fisiche. In particolare, h e V_{APR} sono soggette alle leggi dell'aerodinamica, mentre b non può scendere al di sotto della distanza percorsa dall'APR in $T = \frac{1}{r_{fr}} = 1 / r_{fr}$, (dove r_{fr} indica il *frame rate* della fotocamera), ovvero

$$b_{min} = \frac{V_{APR}}{r_{fr}}, \quad [1]$$

ed è opportuno che non superi 3-4 volte b_{min}^3 , per evitare che l'APR compia movimenti indesiderati tra τ_1 e τ_2 , dovuti per esempio alla presenza di vento. A questo punto, possiamo aggiustare i parametri mancanti. Innanzitutto, scegliamo la minima V_{APR} consentita all'interno di un *range* di h verosimile. In questo modo, possiamo limitare il fenomeno della sfocatura dovuta al movimento della fotocamera e dunque migliorare la risoluzione a terra. Dopodiché, considerando che un incremento di h diminuisce la risoluzione del modello 3D (ovvero fa crescere M) in modo quadratico, settiamo tale parametro al minor valore consentito dalle leggi aerodinamiche, che riesca comunque a garantire, per f pari al minimo valore permesso dalla fotocamera e per b minimo – trovato nell'Eq. [1] – una A_{max} di circa 10 volte superiore a quella richiesta dall'applicazione (A_{op}) e una H_{max} che superi di almeno 3 volte la massima altezza degli oggetti che siamo interessati a ricostruire (H_{ogg}). Infine, per andare a diminuire i valori di R e di M , andiamo ad aumentare f , fino a che A_{max} non eguagli A_{op} o H_{max} non eguagli H_{ogg} . Se una volta raggiunta la massima f consentita dalla fotocamera $A_{max} > A_{op}$ e $H_{max} > H_{ogg}$, aumentiamo b , fino a quando non si verifica che $A_{max} = A_{op}$ o che $H_{max} = H_{ogg}$, o fino a quando b non eguaglia il valore limite di 3-4 volte b_{min} .

Algoritmo di ricostruzione 3D

L'algoritmo di ricostruzione tridimensionale che abbiamo implementato permette di creare un modello 3D da due immagini acquisite come specificato nella sezione precedente. Si basa sull'algoritmo 8-punti normalizzato (Hartley, 1997), scelto in quanto migliore compromesso tra l'accuratezza e l'efficienza computazionale, quando, come nel nostro caso, non si disponga di alcuna informazione *a priori*. Lo schema a blocchi dell'algoritmo di ricostruzione è mostrato in Fig. 2.

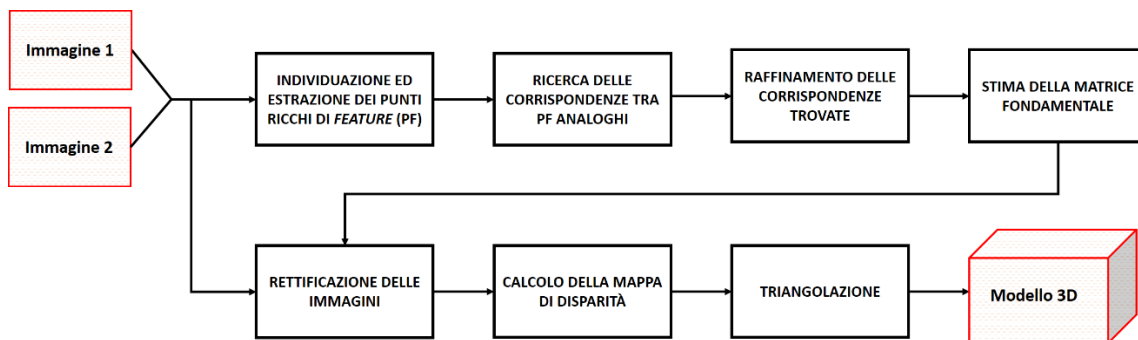


Figura 2 – Schema a blocchi della tecnica di ricostruzione tridimensionale proposta.

Il primo passo consiste nell'individuare i punti più ricchi di *feature* (che indicheremo con PF) sulle due immagini acquisite, utilizzando l'algoritmo SIFT (Lowe, 1999). Dopodiché, le *feature* di ogni PF vengono estratte e parametrizzate in forma vettoriale, sempre servendosi del SIFT, in modo da ottenere una sorta di *descrittore* delle caratteristiche del PF. Benché altri tipici algoritmi di individuazione ed estrazione di *feature* siano meno costosi dal punto di vista computazionale, come per esempio il SURF (Bay *et al.*, 2006), abbiamo preferito utilizzare un algoritmo più lento ma

³ Si noti che b varia in modo discreto, in quanto è possibile acquisire immagini solo ad intervalli di tempo multipli di $1/r_{fr}$.

anche più accurato. La scelta è dovuta al fatto che l'errore commesso durante l'individuazione/estrazione delle *feature* pesa considerevolmente sui passi successivi, come per esempio sulla ricerca delle corrispondenze tra i PF delle due immagini, che, a sua volta coinvolge direttamente il calcolo della matrice fondamentale (Prince 2012), operazione centrale nel processo di ricostruzione. Inoltre, potendo fare affidamento su un'estrazione di *feature* robusta, possiamo permetterci, durante la fase di affinamento delle corrispondenze tra PF – la quale può risultare particolarmente onerosa in termini di carico computazionale – di utilizzare una procedura più veloce anche se meno accurata, raggiungendo comunque la precisione desiderata.

Il passo successivo consiste nel trovare le coppie di PF analoghi – cioè corrispondenti allo stesso punto della scena reale – nelle due immagini, operazione che avviene considerando i descrittori dei PF della prima immagine e, per ciascuno di questi, cercando il PF della seconda il cui descrittore abbia la minima distanza vettoriale. Per gli stessi motivi visti sopra, ancora una volta abbiamo scelto di sacrificare il tempo computazionale per migliorare l'accuratezza. Infatti, una ricerca delle corrispondenze lenta ma completamente esaustiva è stata preferita a una ricerca più veloce ma approssimativa, quale per esempio la FLANN (Muja, Lowe, 2013), comunemente usata in questo genere di applicazioni.

La selezione delle coppie di PF analoghi viene successivamente raffinata. In questo caso, abbiamo scelto di privilegiare l'aspetto del tempo computazionale, rispetto all'accuratezza, confidando sulla robustezza delle *routine* usate in precedenza. Come prima cosa, vengono eliminate quelle coppie i cui descrittori abbiano una distanza superiore a una certa soglia, settata a 1.25 volte la distanza media tra tutte le coppie di PF. Dopodiché, vengono rimossi gli eventuali *outlier* mediante l'algoritmo R-RANSAC SPRT (Matas, Chum, 2005), uno dei più veloci della famiglia dei RANSAC (Sunglok *et al.*, 2009).

Sulle coppie di PF rimasti viene eseguito l'algoritmo 8-punti normalizzato, che restituisce in output il valore della matrice fondamentale (F). Si noti che gli eventuali effetti di distorsione introdotti dalle lenti della fotocamera sono stati trascurati, semplificando così notevolmente il calcolo di F. Ciò è stato possibile senza introdurre errori consistenti, dato che il fenomeno della distorsione è molto limitato nelle fotocamere compatte, come quella che abbiamo scelto di usare.

Il valore di F stimato è usato per rettificare le due immagini, ovvero per trasformarle mediante un'omografia proiettiva, in modo tale da allineare le righe della prima immagine con le righe analoghe della seconda. L'operazione di rettificazione è completamente determinata (a meno di un fattore di scala) a patto di conoscere i parametri intrinseci della fotocamera. Se questi non sono noti, la rettificazione è determinata soltanto a meno di una trasformazione proiettiva. Nel nostro caso, le dimensioni del singolo fotorivelatore e la lunghezza focale sono note, approssimativamente, dai *datasheet* della fotocamera. L'*offset* tra l'asse ottico e il centro del sensore viene invece considerato nullo, supponendo che non sia presente nessun difetto di fabbricazione. Nonostante l'introduzione di alcune approssimazioni, la loro lieve entità consente di superare l'ambiguità dovuta alla trasformazione proiettiva. Questo soprattutto grazie alla geometria di acquisizione scelta, che dà luogo a immagini praticamente già rettificate, a meno di movimenti indesiderati dell'APR.

Dalle due immagini rettificate, viene poi calcolata la mappa di disparità. Con il termine *disparità* (*d*) si intende lo scostamento orizzontale tra i pixel analoghi delle due immagini. Il primo passo per il calcolo della mappa è quello di trovare, per ogni pixel della prima immagine, il pixel analogo nella seconda. Si noti che, finora, le uniche corrispondenze trovate sono relative ai PF delle due immagini, che costituiscono un sottoinsieme sparso del numero totale dei pixel. Trovare le coppie corrispondenti per ciascun pixel permetterà di avere una ricostruzione 3D *densa*, ovvero “senza buchi” tra un punto e l'altro, e dunque più realistica. La ricerca delle corrispondenze viene eseguita mediante l'algoritmo SGBM (Hirschmuller, 2007), dimostratosi essere una delle più accurate soluzioni tra gli algoritmi *real-time* capaci di calcolare mappe di disparità dense (Scharstein, Szeliski, 2002).

L'ultimo passo dell'algoritmo di ricostruzione consiste nel produrre il modello 3D dalla mappa di disparità mediante triangolazione:

$$x = \frac{b \cdot u}{d} \cdot \Delta d, \quad y = \frac{b \cdot v}{d}, \quad z = \frac{b \cdot f}{d}, \quad [2]$$

dove u e v sono le coordinate di un determinato pixel dell'immagine acquisita all'istante τ_1 e x, y e z sono le coordinate 3D del punto ricostruito, espresse in un sistema di riferimento centrato nel fuoco della fotocamera all'istante τ_1 e i cui assi X, Y e Z sono diretti rispettivamente lungo il lato maggiore e il lato minore del sensore, e lungo l'asse ottico, seguendo la regola della mano destra.

Il modello 3D viene salvato in un file PLY, formato che associa alle coordinate di ogni punto ricostruito i suoi valori di intensità di rosso, verde e blu, ricavati dalla prima immagine acquisita, in modo da produrre una ricostruzione realistica. Il formato PLY è supportato da tutti i più comuni software di visualizzazione 3D. L'utente può dunque leggere il file contenente il modello 3D aprendolo direttamente con il programma che preferisce e, tramite questo, esplorarlo a piacimento.

Applicazione e test della tecnica proposta nel monitoraggio di un cantiere edile

La tecnica di ricostruzione proposta è stata proficuamente utilizzata nella risoluzione di alcuni *task* finalizzati al monitoraggio di un cantiere edile. In particolare, è stato richiesto di ricostruire in 3D alcune aree del cantiere, per verificare lo stato di avanzamento dei lavori. Inoltre, la presenza dei veicoli da costruzione all'interno del cantiere è servita come opportunità per testare la tecnica nell'estrazione di modelli tridimensionali di singoli oggetti, al fine di permetterne una migliore identificazione.

Per l'acquisizione delle immagini, è stato impiegato un drone ad ala fissa di piccole dimensioni (circa 1.20×2.30 metri), equipaggiato soltanto con una fotocamera compatta economica, con risoluzione di 4000×3000 pixel, che permette la regolazione dei parametri precedentemente descritti. Per decidere la modalità di acquisizione delle immagini, abbiamo applicato la procedura descritta nella sezione precedente. Le coppie di immagini acquisite sono poi state usate come input per l'algoritmo di ricostruzione 3D, che ha restituito i modelli richiesti.

In Fig. 3(a) è mostrata una delle coppie di immagini del cantiere, raffigurante un'area di interesse. Basandosi semplicemente su di esse, non è semplice capire quali zone siano "più in alto" e quali "più in basso", e quale sia l'effettiva differenza di quota. In Fig. 3(b) è presentato il modello 3D dell'area, generato con la tecnica da noi proposta e visualizzato mediante il software Meshlab. La ricostruzione appare molto realistica e permette all'osservatore di valutare facilmente lo stato dei lavori, quasi come se fosse presente sul posto.

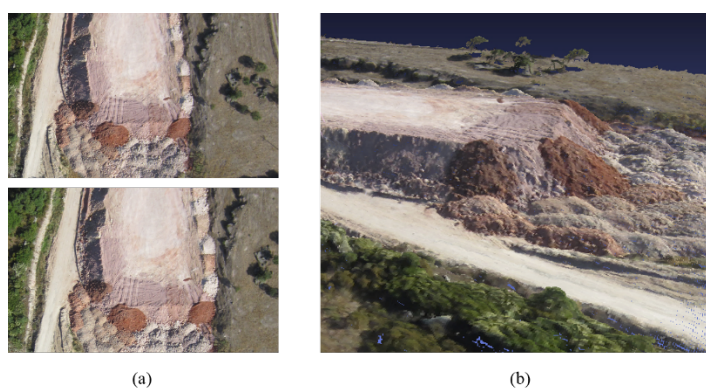


Figura 3 – Coppia di immagini di un'area del cantiere (a) e ricostruzione 3D dell'area mediante la tecnica proposta (b).

In Fig. 4(a) è mostrata un'altra coppia di immagini, stavolta di un veicolo da costruzione presente nel cantiere. Né il veicolo né la sua posizione sono identificabili con chiarezza dalle sole fotografie. Grazie alla ricostruzione 3D ottenuta dalle due immagini con la tecnica proposta (Fig. 4(b)), è invece possibile riconoscere facilmente una scavatrice, dal dettaglio del braccio meccanico, e di comprendere immediatamente che essa si trova sopra il ciglio di una cava.

I risultati ottenuti da un punto di vista qualitativo sono molto promettenti. Purtroppo, da un punto di vista quantitativo, non è stato possibile valutare in modo esaustivo l'accuratezza dei modelli 3D generati, data l'assenza di strumentazione ausiliaria a bordo dell'APR – in particolare di un GPS che permettesse di georeferenziare le immagini. Poiché però la quota dell'APR era nota, abbiamo potuto calcolare l'errore commesso nel calcolo dell'altezza di alcuni punti della scena di cui questa era conosciuta *a priori*. Mediando l'errore su tutti i punti considerati, abbiamo ottenuto un'accuratezza di 0.36m. Tale risultato ci mostra che i modelli generati con la tecnica proposta non si prestano ad essere utilizzati per applicazioni di fotogrammetria di precisione, ma permettono comunque di effettuare rilevazioni su scala più ampia.

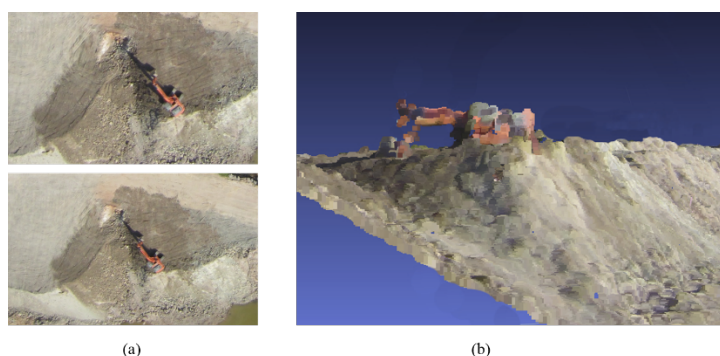


Figura 4 – Coppia di immagini di un oggetto presente all'interno del cantiere (a) e ricostruzione 3D di questo mediante la tecnica proposta (b).

Per verificare se l'algoritmo fosse in grado di generare una ricostruzione 3D in tempo reale, abbiamo calcolato il tempo computazionale per ogni modello prodotto. In scenari critici, un operatore deve poter disporre del modello 3D in pochi secondi. Dunque, abbiamo deciso di tollerare un ritardo massimo di 3s, eseguendo l'algoritmo su un normale PC, per poter considerare soddisfatto il requisito di lavorare in tempo reale. Utilizzando un PC dotato di un processore a 2GHz e con 4GB di memoria RAM, il tempo medio per produrre un modello è risultato di 0.92s, cioè ben 3 volte al di sotto del limite tollerabile, dimostrando l'idoneità dell'algoritmo a lavorare in tempo reale e dell'intera tecnica ad essere utilizzata in scenari critici.

Conclusioni

Abbiamo presentato una tecnica per creare modelli 3D realistici da immagini acquisite mediante un APR. La tecnica coinvolge l'intero processo che porta all'estrazione del modello, dalla pianificazione della modalità di acquisizione delle immagini, all'algoritmo di ricostruzione 3D, in modo da regolare *ad hoc* ogni fattore che possa influire sulle *performance*.

Utilizzando solamente coppie di immagini, acquisite da una fotocamera compatta economica montata su un APR di piccole dimensioni, anch'esso a basso costo, è stato possibile produrre ricostruzioni di aree di interesse o di singoli oggetti in tempo reale (meno di 1s), senza l'ausilio di alcuna strumentazione addizionale o di GCP piazzati sulla scena. I modelli generati, esplorabili mediante qualunque dei più comuni software di visualizzazione 3D, danno all'osservatore l'impressione di essere fisicamente presente all'interno della scena, permettendogli di recuperare informazioni non ottenibili dall'analisi delle sole immagini bidimensionali. L'accuratezza media ottenuta nella determinazione dell'altezza da terra dei punti della scena ricostruita, pari a circa 35cm, non consente l'uso dei modelli generati per la fotogrammetria di precisione, ma permette comunque di effettuare utili rilevazioni su larga scala.

La tecnica proposta si presta ad essere utilizzata in numerose applicazioni di telerilevamento, specialmente in quei contesti in cui siano richiesti, in tempi brevissimi, modelli realistici e facili da interpretare di aree o oggetti di interesse, come per esempio nel monitoraggio dei disastri naturali,

nella video-sorveglianza e nel *search & rescue*, che si vanno ad aggiungere a campi meno critici quali la topografia, la pianificazione urbanistica, lo studio delle aree archeologiche, ecc.

Riferimenti bibliografici

- Bay H., Tuytelaars T., Van Gool L. (2006), “SURF: Speeded Up Robust Feature”, *Proceedings of the 9th European Conference on Computer Vision*, Graz, Austria, 3951: 404-417.
- Bisheng Y., Chi C. (2015), “Automatic registration of UAV-borne sequent images and LiDAR data”, *ISPRS Journal of Photogrammetry and Remote Sensing*, 101: 262-274.
- D’Urso M. G., Gargiulo A., Rotondi A., Bracaglia R. (2015), “Un’applicazione al settore ambientale di fotogrammetria con piattaforma UAV”, *ASITA 2015*, Lecco, Italy, 19: 347-354.
- Hartley R. I. (1997), “In defence of eight-point algorithm”. *IEEE Transaction on Pattern Recognition and Machine Intelligence*, 19-(6): 580-593.
- Haubeck K., Prinz T. (2013) “A UAV-based low-cost stereo camera system for archaeological surveys - experiences from Doliche (Turkey)”, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40-1/W2: 195-200.
- Hirschmuller H. (2007), “Stereo processing by semiglobal matching and mutual information”, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 30-(2): 328-341.
- Huei-Hung L., Yuping L., Medioni G., (2011), “Aerial 3D reconstruction with line-constrained dynamic programming”. *Proceedings of the 2011 IEEE International Conference on Computer Vision (ICCV)*, Barcelona, Spain, pp. 1855-1862.
- Jutzi B., Weinmann M., Meidow J. (2015), “Document Improved UAV-borne 3D mapping by fusing optical and laser scanner data”, *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40-1W2: 223-228.
- Matas J., Chum, O. (2005) “Randomized RANSAC with Sequential Probability Ratio Test”, *Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV)*, Beijing, China, 2: 1727-1732.
- Mayer H., Bartelsen J. (2008). “Automated 3D Reconstruction of Urban Areas from Networks of Wide-Baseline Image Sequences”, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37-B5: 633-638.
- Muja M., Lowe D. G. (2013), “FLANN - Fast Library for Approximate Nearest Neighbors”, www.cs.ubs.ca/research/flann/.
- Prince S. J. D. (2012). *Computer Vision. Models, Learning and Inference*. Cambridge University Press, section 16.3: 361-363.
- Scharstein D., Szeliski R. (2002), “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms”, *International Journal of Computer Vision*, 47-(1-3): pp. 7-42.
- Se S., Firoozfam P., Goldstein N., Wu L., Dutkiewicz M., Pace P., Naud J. L. P. (2009), “Automated UAV-based mapping for airborne reconnaissance and video exploitation”, *Proceedings of SPIE Airborne Intelligence, Surveillance, Reconnaissance (ISR) Systems and Applications VI*, Orlando, FL, USA, vol. 7307.
- Sunglok C., Taemin K., Wonpil Y. (2009), “Performance Evaluation of RANSAC Family” *Proceedings of 2009 British Machine Vision Conference (BMVC)*, London, UK.
- Vasterling M., Meyer U. (2013), “Challenges and Opportunities for UAV-Borne Thermal Imaging”, *Remote Sensing and Digital Image Processing*, 5: 69-92.
- Wefelscheid C., Hänsch R., Hellwich O. (2011), “Three-dimensional building reconstruction using images obtained by unmanned aerial vehicles” *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38-1/C22: 1-6.