

L'uso di dispositivi mass market per la documentazione e la modellazione dei beni culturali

Irene Aicardi, Andrea Lingua, Marco Piras

Dipartimento di Ingegneria dell'ambiente del Territorio e delle Infrastrutture
(irene.aicardi, andrea.lingua, marco.piras)

Riassunto

Il patrimonio culturale e artistico è da sempre al centro di attività finalizzate alla sua salvaguardia e valorizzazione. Nel nostro Paese, noto ed apprezzato per le sue ricchezze da tutelare, all'alto rischio naturale dovuto ai terremoti, alle frane e alle alluvioni si sommano le attività dell'uomo che contribuiscono a rendere il patrimonio ancora più fragile. Pertanto, il territorio per poter essere valorizzato e protetto, deve essere innanzitutto conosciuto, sia da un punto di vista storico e culturale, ma anche metrico, vale a dire essere in grado di descriverlo e rappresentarlo in maniera geometricamente corretta, per migliorare la conservazione. In questo scenario, le nuove tecnologie possono offrire strumenti davvero innovativi per la valorizzazione dei beni culturali: un esempio potrebbe essere l'utilizzo di dispositivi come *smartphone* e videocamere *mass-market* per realizzare la documentazione e il censimento dei beni esistenti.

Questi sensori sono ad oggi facilmente reperibili e largamente utilizzati dalla maggior parte degli utenti, specialmente durante i loro viaggi e le loro attività, e negli ultimi anni sono stati ulteriormente implementati fornendo la possibilità di ottenere, oltre a dati di tipo fotografico a risoluzioni sempre migliori, anche informazioni di posizione e assetto derivanti da ricevitori GPS e sensori inerziali integrati nei dispositivi stessi.

L'obiettivo di questo lavoro è quello di testare e validare l'utilizzo di alcuni dispositivi *mass-market* per la documentazione dei beni culturali.

Abstract

The cultural and artistic heritage has always been at the center of activities aimed at enhancing its preservation. In particular, in Italy, a country full of heritage to be protected, the high risk due to earthquakes, landslides and floods must be added to the human activities that contribute to make the heritage even more fragile.

Therefore, in order to value and protect the land, it must first be known, and not only in a historical and cultural heritage, but also from a metric point of view, so as to be able to describe and represent it at the best manner.

In this scenario, the new technology can offer truly innovative instruments for the promotion of cultural heritage and the use of mass market devices could be a new opportunity to realize this documentation. These sensors are nowadays easily available and widely used by the majority of users, especially during their travels and their activities, and in recent years they have been expanded by providing the possibility to obtain, in addition to photos at resolutions always higher, even position and attitude information derived from GPS receivers and data from inertial sensors embedded in the devices themselves. In addition, through an Internet connection, this data can be shared in real time and published on the Web.

Therefore, the authors' objective is to test and validate the application of certain mass market devices, those that are currently more present on the market, for the documentation of cultural heritage.

Introduzione

“Most of the world’s significant sites have been photographed under many different conditions, both from ground and from the air” (Seith et.al, 2008).

La possibilità di acquisire dati fotografici, e non solo, sta diventando sempre più alla portata di tutti poiché, la maggior parte di noi, o nel tempo libero o durante le attività lavorative ha a disposizione uno *smartphone* che consente di acquisire immagini con risoluzioni sempre più elevate.

A questo, si aggiunge lo sviluppo notevole a cui abbiamo assistito negli ultimi anni delle videocamere compatte per riprese sportive, le cosiddette “*action cam*”, sempre più maneggevoli e ad elevata risoluzione, in grado di consentire una documentazione dei beni culturali, e non solo, presenti nel nostro territorio (Brown et al., 2005) (Debevec et al., 1996).

Ad oggi, i moderni *smartphone* includono diversi sensori che spesso vengono adottati per applicazioni di tipo geomatico, come per esempio la fotocamera digitale, i ricevitori GNSS e le piattaforme inerziali basate su giroscopi, accelerometri e magnetometri e sistemi RFID per quel che riguarda gli *smartphone* (Piras et al., 2014).

Le potenziali prestazioni di questi sensori potrebbero estendere il ruolo degli utenti nelle attività di documentazione del patrimonio artistico e culturale, in quanto esistono già, ad esempio, delle applicazioni di *crowd-mapping* in cui le persone stesse interagiscono nello scambio di informazioni e di dati, soprattutto di dati spaziali. In effetti, un gran numero di immagini di parti del mondo e dei monumenti più famosi è già disponibile sul Web in diverse condizioni di illuminazione e inquadratura grazie alle potenzialità offerte da Internet e la possibilità che gli utenti hanno di condividerle.

Proprio in quest’ottica, al giorno d’oggi, esistono già diverse possibilità offerte agli utenti di creare il proprio modello tridimensionale da immagini acquisite da loro stessi. E’ infatti possibile utilizzare software *user-friendly*, usufruibili anche da persone non esperte di fotogrammetria, oppure caricare le proprie immagini sul Web (ad esempio, attraverso Autodesk 123D Catch) e ricevere il modello 3D già elaborato via e-mail (Vergauwen et al., 2006).

Il coinvolgimento degli utenti e la loro curiosità in queste applicazioni stanno diventando pertanto molto importanti, quindi lo scopo in questo lavoro è proprio quello di testare la qualità dei dati acquisibili e dei risultati che si possono ottenere utilizzando i sensori *mass-market* disponibili e più utilizzati ad oggi presenti sul mercato (Chandler et .al, 2005).

La calibrazione dei sensori

Nei test svolti sono stati utilizzati i sensori più popolari ad oggi disponibili sul mercato, in particolare *smartphone* e videocamere che, pur rientrando nella categoria dei sensori *mass-market*, consentono di acquisire video e immagini ad alte risoluzioni. I sensori sono stati singolarmente analizzati per studiarne le caratteristiche tecniche e, per quanto riguarda la parte relativa alle immagini, si riportano alcune caratteristiche fondamentali nella Tabella 1:

				
Modello	Samsung Galaxy S Advance	iPhone4	Go Pro Hero2	Contour Plus2
Categoria	Smartphone	Smartphone	Video camera	Video camera
Costo [€]	~ 170	~ 400	~ 200	~ 400
Lungh. focale [mm]	4,310	3,850	4,300	3,000
Dim. pixel [µm]	1,75	2,45	1,40	4,50
Dim. immagine [pixel]	1920x2560	1936x2592	1920x1080	1280x960

Tabella 1 - I sensori utilizzati e alcune loro caratteristiche

I dispositivi analizzati contengono foto/video-camere integrate, non metriche, che occorre calibrare per analizzare i parametri di distorsione e valutare gli errori ad essi imputabili.

Soprattutto per applicazioni di tipo fotogrammetrico e per la modellazione 3D, è fondamentale una corretta calibrazione delle lenti e dei sensori per consentire risultati più accurati e corretti delle deformazioni. Come ben noto (Kraus, 1997), il sistema ottico delle camere è costituito da una serie di lenti di forma curva diversa e nella realtà le lenti fotogrammetriche presentano alcune differenze significative rispetto al modello ideale/teorico dovute a:

- il disallineamento nel montaggio delle lenti;
- l'asse fotogrammetrico di riferimento non coincide con l'asse ottico;
- gli angoli di rifrazione e di incidenza non corrispondono;
- la distanza principale è leggermente diversa dalla distanza principale ottica;
- il piano immagine non è perfettamente perpendicolare all'asse ottico.

Queste incongruenze danno origine a distorsioni di tipo diverso che potremmo riassumere come:

- distorsioni radiali:

$$\Delta\xi_1 = \Delta\rho \frac{\xi - \xi_0}{\rho} = (\xi - \xi_0) \cdot (k_1\rho^2 + k_2\rho^4 + k_3\rho^6 + \dots) \quad [1]$$

$$\Delta\eta_1 = \Delta\rho \frac{\eta - \eta_0}{\rho} = (\eta - \eta_0) \cdot (k_1\rho^2 + k_2\rho^4 + k_3\rho^6 + \dots) \quad [2]$$

- distorsioni tangenziali:

$$\Delta\xi_2 = [b_1(\rho^2 + 2\xi^2) + 2b_2\xi\eta](1 + b_3\rho^2) \quad [3]$$

$$\Delta\eta_2 = [b_2(\rho^2 + 2\eta^2) + 2b_1\xi\eta](1 + b_3\rho^2) \quad [4]$$

- distorsioni anche dette trasformazioni affini:

$$\Delta\xi_3 = c_1(\xi - \xi_0) + c_2(\eta - \eta_0) \quad [5]$$

$$\Delta\eta_3 = 0 \quad [6]$$

Per caratterizzare le conseguenze di tutti questi effetti occorre quindi calibrare la fotocamera e analizzare, in particolare, la distorsione radiale che più influenza la distorsione delle immagini delle fotocamere degli smartphone.

Per questa applicazione si è utilizzata una procedura di autocalibrazione (Clarke, 1998) attraverso il poligono situato all'esterno del Laboratorio di Geomatica del Politecnico di Torino, nel quale sono presenti una serie di marker quadrati di lato 8 cm (Figura 1).



Figura 1 - Poligono utilizzato per l'autocalibrazione dei dispositivi

Il reticolo è formato da 26 target disposti in modo da coprire uniformemente le facciate. Questi target sono stati tutti rilevati topograficamente, mediante uno schema ridondante per un miglior controllo dei risultati (le coordinate dei target) e una stima affidabile della precisione ottenuta mediante soluzione ai minimi quadrati.

La calibrazione è stata effettuata per tutti i sensori attraverso il software *Leica Photogrammetric Suite* (LPS) di ERDAS che consente di stimare i parametri di orientamenti interno e le distorsioni. In particolare, le distorsioni radiali degli *smartphone* sono riportate in Figura 2:

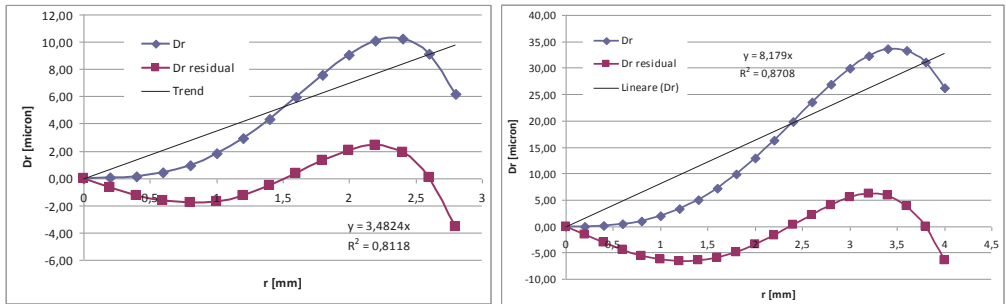


Figura 2 - Distorsioni radiali degli smartphone: Samsung (sinistra) e iPhone (destra)

Mentre, per le videocamere, sono stati acquisiti dei video (Akbarzadeh et al., 2006) dai quali, attraverso una procedura implementata in Matlab, si sono poi estratti singoli *frames* per realizzare la calibrazione, di cui i risultati in termini radiali sono riportati in Figura 3:

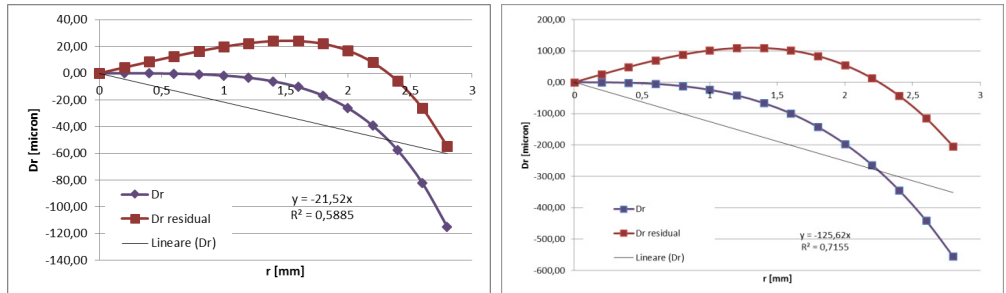


Figura 3 - Distorsioni radiali delle videocamere: GoPro (sinistra) e Contour (destra)

Come è possibile notare dai grafici, l'iPhone ha una distorsione radiale circa quattro volte più grande rispetto quella del Samsung e, correggendo i parametri con l'andamento lineare, la sua distorsione è comunque doppia rispetto quella dell'altro smartphone.

Invece i risultati delle videocamere sono molto diversi rispetto a quelli ottenuti con gli smartphone. È possibile vedere come le distorsioni radiali non siano assolutamente trascurabili per scopi fotogrammetrici e, soprattutto per la Contour, possono raggiungere valori di circa 0,5 mm se non si applica la correzione del trend lineare. Inoltre, va specificato che è stato molto difficile realizzare un'autocalibrazione delle videocamere, perché le loro grandi distorsioni hanno reso il sistema molto instabile e qualsiasi piccolo cambiamento nei punti di controllo rendeva il risultato finale molto diverso. Per tal motivo sono state riportate esclusivamente le distorsioni radiali e si stanno valutando altri software e procedure per effettuare una nuova calibrazione più rigorosa.

Per tale motivo abbiamo calibrato i dispositivi prima di effettuare i test e, visti i risultati, si può affermare che, prima di usarli, è meglio fare sempre una nuova calibrazione poiché, probabilmente, questi parametri non sono stabili nel tempo.

Possiamo quindi riassumere i risultati ottenuti dalla calibrazione in Tabella 2 per avere un'indicazione dei parametri di orientamento interno e le distorsioni radiali:

	Samsung	iPhone4	Go Pro Hero2	Contour Plus2
c [mm]	4,3099	3,8700	4,5069	3,7615
c_{cor} [mm]	4,3249	3,9017	4,4099	3,2890
ξ_0 [mm]	0,0868	-0,0162	-0,9064	-0,6031
η_0 [mm]	0,0704	-0,0359	2,0577	1,1397
k_1	2,01E-03	2,01E-03	-1,20E-03	-2,38E-02
k_2	-2,21E-04	-1,00E-04	-5,15E-04	-1,92E-04
Δr_{max} [μ m]	-3,6	-6,6	-54,7	-204,1

Tabella 2 - Risultati principali delle calibrazioni

Dopo la procedura di calibrazione, possiamo quindi eseguire alcuni test per valutare i risultati ottenibili utilizzando questi dispositivi.

La modellazione 3D

La procedura di modellazione 3D è stata testata per realizzare il modello di un elemento architettonico complesso: un capitello a disposizione presso il laboratorio di Geomatica (Dick et al., 2004). Questo elemento è stato appositamente scelto per la sua struttura geometrica piuttosto complicata con lo scopo di verificare la bontà dei modelli ottenibili da tali dispositivi nella ricostruzione della geometria.

Per fare questo, abbiamo acquisito una serie di foto con gli *smartphone* (Grzeszczuk et al., 2002) e due video utilizzando le videocamere. In questi ultimi casi le immagini sono poi state estratte dai video attraverso una procedura di Matlab ed esportati in formato "tif" (72 dpi).

Per la generazione del modello 3D abbiamo usato il software Photoscan di Agisoft. La scelta di questa soluzione è stata fatta in modo tale da simulare un comune utente che vuole realizzare un modello 3D con le sue foto. Analizzando le soluzioni oggi presenti sul mercato, riteniamo che Photoscan sia il modo migliore per ottenere buoni risultati in un modo molto semplice.

Esistono sul mercato altri software, come ad esempio MicMac, che sono in grado di generare modelli più sofisticati, ma non sono di facile utilizzo soprattutto per gli utenti non esperti in fotogrammetria.

La procedura seguita nel lavoro è stata la seguente:

1. allineamento delle foto: al fine di generare una nuvola di punti 3D a partire dalle immagini attraverso un *bundle block adjustment*;
2. selezione dell'oggetto: le immagini acquisite contengono dati anche esterni all'oggetto di interesse. Quindi, se vogliamo solo il capitello, PhotoScan offre la possibilità di mascherare le altre parti delle immagini, ma richiede molto tempo, quindi in alternativa è possibile ridurre il box da analizzare intorno all'oggetto. È importante dire che a volte le maschere si sono rivelate fondamentali al fine di ottenere un modello del capitello. Infatti, come meglio specificato in seguito, nel caso specifico delle videocamere non è stato possibile eseguire correttamente l'allineamento delle immagini senza l'inserimento preventivo delle maschere;
3. costruzione del modello: ovvero triangolazione della nuvola di punti in modo da interpolare e realizzare la mesh;
4. applicazione della texture: infine il software offre anche la possibilità di applicare il colore associato alle immagini per avere una nuvola di punti 3D e una mesh colorata.

A titolo esemplificativo riportiamo in Figura 4 la mesh e l'applicazione della texture del modello ottenuto con le immagini acquisite con l'iPhone.

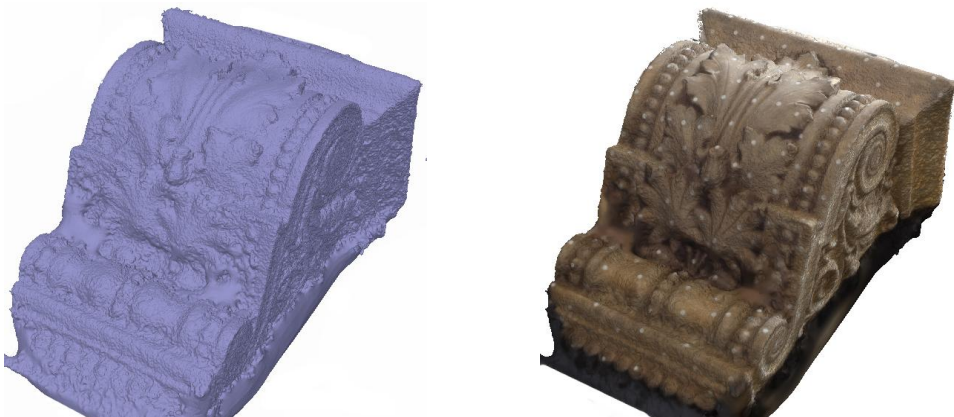


Figura 4 - Mesh e applicazione della texture del modello realizzato con le immagini dell'iPhone

I quattro test effettuati hanno condotto tutti alla generazione di modelli 3D con precisioni diverse e con procedure più o meno automatiche, come schematizzato in Tabella 3:

	Samsung	iPhone4	Go Pro Hero2	Contour Plus2
Foto	26	33	150	119
Punti	6683	9190	14957	8775
Facce	6429167	10901706	2981021	1821502
Vertici	3214932	5451137	1490854	910954
Dim. [MB]	268	436	124	76

Tabella 3 - Dati sull'elaborazione delle immagini

Come precedentemente accennato, occorre precisare che nel caso degli *smartphone* i modelli sono stati generati in modo automatico dal software, mentre nel caso delle videocamere (Pollefeys et al., 2004), probabilmente a causa delle forti distorsioni, è stato necessario, in entrambi i casi, un pretrattamento di ciascuna immagine al fine di mascherare le aree esterne non facenti parte dell'oggetto studio.

L'elaborazione dei dati è stata realizzata con un PC portatile con Windows8 Pro x64, Intel Core i7, 8 GB di RAM. Il processo di elaborazione dei dati ha richiesto mediamente circa 5-6 ore per i modelli degli *smartphone* e 6-7 ore per le videocamere (poiché circa un'ora è stata destinata alla mascheratura delle immagini prima dell'allineamento).

La validazione dei risultati

Al fine di poter effettuare delle analisi sulla bontà dei risultati ottenuti, è stato generato un modello 3D di riferimento attraverso tecnica laser scanner. Nello specifico, sono state acquisite 4 scansioni con il laser Faro Cam 3D che hanno consentito di ottenere un modello 3D completo del capitello. La nuvola di punti è stata elaborata attraverso il software Scene v.5.2 e la generazione della mesh, assunta poi come dato di riferimento per le successive comparazioni, è stata realizzata in 3DReshaper.

Pertanto, ogni modello fotogrammetrico è stato confrontato con il modello di riferimento laser al fine di verificarne la completezza e la bontà, soprattutto nella definizione dei particolari.

Tale comparazione (Figura 5 e Figura 6) è avvenuta nel software 3DReshaper. Un primo tentativo era stato effettuato anche in CloudCompare che, tuttavia, non riusciva a supportare adeguatamente le mole di dati.

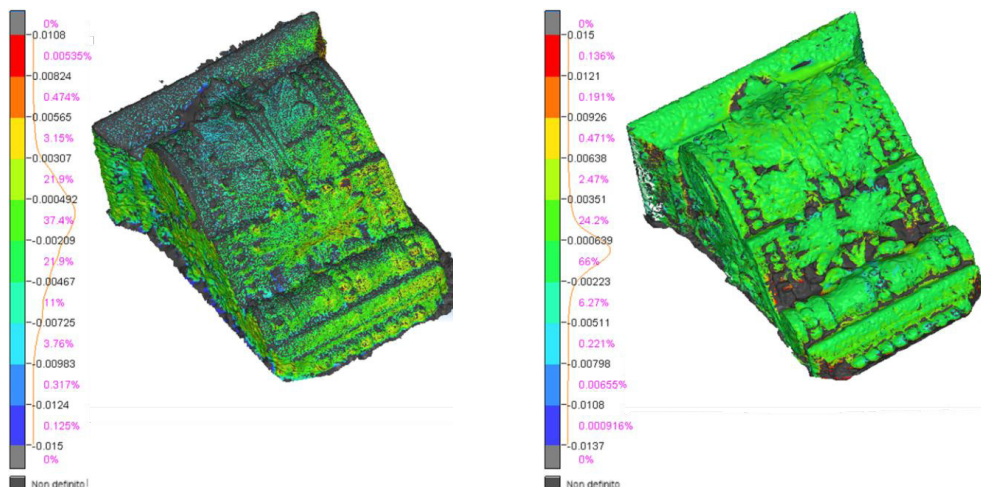


Figura 5 - Comparazione: laser - Samsung (sinistra), laser - iPhone (destra)

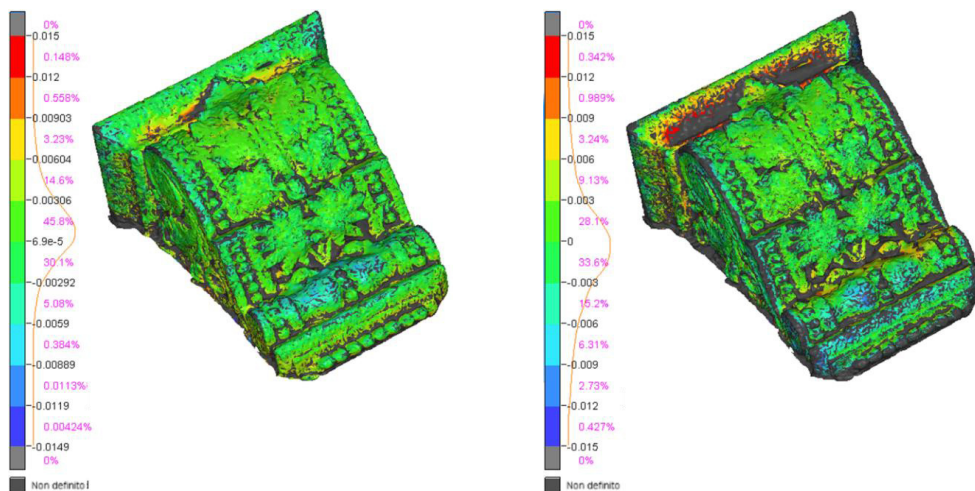


Figura 6 - Comparazione: laser - GoPro (sinistra), laser - Contour (destra)

Com'è possibile notare dalle immagini, il software restituisce anche un istogramma sulla comparazione dei due dati che consente un tipo di analisi più intuitiva e rapida.

Per le comparazioni, il limite massimo di differenza tra le due *mesh* è stato impostato a 1,5 cm; in questo modo, le parti oltre tale limite non sono state considerate perché sono assimilate come errori grossolani. Analizzando gli istogrammi possiamo notare che la maggior parte dei modelli presentano piccole differenze rispetto al modello di riferimento e, in particolare, possiamo riassumere i risultati ottenuti in Tabella 4:

	Samsung [%]	iPhone4 [%]	Go Pro Hero2 [%]	Contour Plus2 [%]
D < 3 [mm]	~ 81,20	~ 90,20	~ 75,90	~ 61,70
3 < D < 9 [mm]	~ 18,35	~ 9,43	~ 23,28	~ 33,82
9 < D < 15 [mm]	~ 0,40	~ 0,33	~ 0,72	~ 4,43
D > 15 [mm]	~ 0,05	~ 0,04	~ 0,10	~ 0,05

Tabella 4 - Percentuale di differenze tra il modello laser e gli altri

Il miglior modello ottenuto è stato creato utilizzando lo smartphone iPhone che è quello più coerente rispetto al modello laser, mentre le altre *mesh* presentano errori tra i 3-9 mm rispetto al modello di riferimento laser. Ma è anche importante notare come il numero di punti utilizzati per effettuare il confronto sia diverso per ciascun dispositivo. Pertanto, le percentuali riportate in Tabella 4 si riferiscono ad un diverso numero di punti, ma è comunque significativo rispetto all'intero modello.

Conclusioni

I risultati ottenuti dimostrano che i sensori testati possono essere impiegati per acquisire immagini utili nella realizzazione di modelli 3D e, in particolare con gli *smartphone*, è possibile ottenere buone precisioni. Le immagini acquisite con dispositivi *mass-market*, sia *smartphone* che videocamere, hanno risoluzioni inferiori a quelle normalmente utilizzate per scopi fotogrammetrici, pertanto per ottenere un buon risultato, è necessario aumentare il numero di scatti e le posizioni delle acquisizioni.

In particolare per le videocamere, è necessario intensificare la fase di elaborazione delle immagini, perché, come abbiamo testato, il software non è in grado di generare in modo automatico i modelli tridimensionali. Questo è probabilmente dovuto alle grandi distorsioni di questi dispositivi rispetto quelle degli *smartphone*. In particolare la Contour ha un GPS integrato, in grado di offrire anche informazioni di posizionamento, per cui lo sviluppo di tali dispositivi è stato volto non tanto al

migliorare le risoluzioni delle focali, quanto a produrre sistemi integrati e completi. Inoltre, per le videocamere la caratteristica delle lenti è che le distorsioni aumentano allontanandosi dal centro dell'oggetto, pertanto tale aspetto deve essere preso in considerazione quando si costruisce un modello 3D utilizzando questi dispositivi.

E' infine molto importante, per stimare al meglio le distorsioni e il loro monitoraggio nel tempo, tenere in conto che, essendo i sensori non-metriche, probabilmente i parametri di calibrazione non sono stabili nel tempo.

Ringraziamenti

Il progetto è stato realizzato grazie ai fondi del programma di ricerca nazionale PRIN “*Tecniche geomatiche innovative ed emergenti di rilievo, telerilevamento (da aereo, satellite, UAV) e WEBGIS per la mappatura del rischio in tempo reale e la prevenzione del danno ambientale*”.

Bibliografia

- Akbarzadeh A., Frahm J.-M., Mordohai P., Clipp B., Engels C., Gallup D., Merrell P., Phelps M., Sinha S., Talton B., Wang L., Yang Q., Stewenius H., Yang R., Welch G., Towles H., Nistér D., Pollefeys M. (2006), “Towards urban 3D reconstruction from video”. In *Proceedings of the international symposium on 3D data processing, visualization, and transmission*.
- Brown M., Lowe D. G. (2005), “Unsupervised 3D object recognition and reconstruction in unordered datasets”. In *Proceedings of the international conference on 3D digital imaging and modelling*.
- Chandler J., Fryer J.G., Jack A. (2005), “Metric capabilities of low cost digital cameras for close range surface measurement”. *The Photogrammetric Record*.
- Clarke T.A., Fryer J.G. (1998), “The development of camera calibration methods and models”. *The Photogrammetric Record*.
- Debevec P. E., Taylor C. J., Malik J. (1996), “Modeling and rendering architecture from photographs: a hybrid geometry- and image based approach”. In *SIGGRAPH conference proceedings* (pp. 11–20).
- Dick A. R., Torr P. H. S., Cipolla R. (2004), “Modelling and interpretation of architecture from several images”. *International Journal of Computer Vision*, 60(2), 111–134.
- Grzeszczuk R. (2002), “Course 44: image-based modelling”. In *SIGGRAPH*.
- Kraus K., (1997), “Photogrammetry II, Advanced methods and applications”. Ummeler/ Bonn, ISBN 3-427-78694-3.
- Piras M., Dabove P., Lingua A.M., Aicardi I. (2014), “Indoor Navigation Using Smartphone Technology: A Future Challenge Or An Actual Possibility?”. In 2014 Plans Conference proceedings.
- Pollefeys M., Van Gool L., Vergauwen M., Verbiest F., Cornelis K., Tops J., Koch R. (2004), “Visual modeling with a hand-held camera”. *International Journal of Computer Vision*, 59(3), 207–232.
- Seith S., Snavely N., Szeliski R. (2008), “Modeling the World from Internet Photo Collections”. In: *International Journal of Computer Vision*, Vol. 80, pp.189-210.
- Vergauwen M., Van Gool L. (2006), “Web-based 3D reconstruction service”. *Machine Vision and Applications*, 17(2), 321–329.