

tecnica/technique

Laura Inzerillo, Cettina Santagati

Il progetto del rilievo nell'utilizzo di tecniche di modellazione dense stereo matching Using dense stereo matching techniques in survey

In recent years there has been an enormous increase in the dissemination of modelling techniques known as dense stereo matching (DSM) or image-based modelling (IBM) using free, low cost, open source software, especially applications involving the survey of cultural heritage (architecture, archaeology, and town planning). The algorithms used by these software programmes process photographic datasets and provide a 3D model of the scene in question. During complex survey and 3D restitution procedures, these techniques save time (processing) and money: these advantages can be useful to any professional. This is why it is important to identify the real potential and limits of these software programmes and how they should be used to produce the best possible results. We carried out parallel tests to verify how to use the two software packages: 123D Catch by Autodesk and VisualSfM+CMVS/PMVS2 by Changchang Wu. The photographic dataset was the most critical part of the procedure involving DSM or IBM techniques. The article illustrates three case studies in which only the third study provides an optimal dataset capable of generating a reliable model. The tests were carried out on three different architectural objects, ranging from a small to large scale; terrestrial laser scans were used for the metric comparisons of the ensuing models.

Key words: Dense stereo matching, Image based modeling, 3D modeling, Structure from Motion, Digital Heritage.

In the last few years, new, so-called Structure from Motion (SfM) algorithms have begun to be used in conjunction with the ones which digitally interpret photogrammetric techniques. These SfM algorithms use a series of photographs to recreate the internal parameters of the camera and the spatial position of the homologous points. Although these SfM techniques have been experimented primarily on tourist websites,¹ they are nonetheless very interesting because they save time and money during the acquisition and processing of cultural heritage data.² This added value can be very helpful to professionals. The complex, multifaceted stages associated with traditional and innovative survey techniques, from acquisition to 3D visualisation, require a topographic, photogrammetric³ and laser scanner survey to produce a 3D model⁴ created thanks to the integrated use of several consolidated and disseminated methodologies used in all scientific sectors involved in this field. The examples shown here concisely represent the numerous experiments performed to verify the procedures of two software

La diffusione delle tecniche di modellazione dense stereo matching (DSM) o image-based modeling (IBM) attraverso l'uso di software gratuiti, a basso costo e open source ha avuto un incremento considerevole in questi ultimi anni, specialmente nelle applicazioni che riguardano il rilievo del patrimonio culturale (architettura, archeologia, urbanistica). Gli algoritmi utilizzati all'interno dei software testati elaborano dataset fotografici per restituire il modello 3D della scena inquadrata. Queste tecniche, nelle complesse procedure di rilevamento e restituzione 3D, comportano un risparmio sia in termini di costi che di ore di elaborazione, vantaggi, questi, che potrebbero tornare utili alla figura del professionista. A tal fine, è bene individuare quali siano le reali potenzialità e i limiti di questi software e come vadano utilizzati per ottenere il miglior risultato. Nello specifico sono state condotte sperimentazioni in parallelo per la verifica delle procedure d'uso dei due pacchetti software 123D Catch della Autodesk e VisualSfM+CMVS/PMVS2 creato da Changchang Wu. La criticità maggiore riscontrata nell'utilizzo delle tecniche DSM o IBM è quella del dataset fotografico. Nell'articolo si riportano tre casi studi di cui soltanto nell'ultimo si individua un dataset ottimale in grado di generare un modello affidabile. La verifica è stata condotta su diversi oggetti architettonici dalla scala di dettaglio a quella monumentale, utilizzando scansioni laser terrestri per i confronti metrici sui modelli ottenuti.

Parole chiave: dense stereo matching, fotomodellazione, modellazione tridimensionale, Structure from Motion, patrimonio culturale.

In questi ultimi anni, agli algoritmi che traducono digitalmente le tecniche fotogrammetriche, sono stati affiancati i cosiddetti algoritmi *Structure from Motion* (SfM) che riescono a ricostruire, da una sequenza di scatti fotografici, i parametri interni della macchina fotografica e la posizione nello spazio dei punti omologhi. Se da un lato queste tecniche SfM sono state sperimentate per uno scopo prettamente turistico sul web¹, dall'altro lato costituiscono un potenziale di grande interesse per l'abbattimento dei tempi e dei costi nell'acquisizione ed elaborazione dei dati nel settore dei beni culturali². Potenziale, questo, che potrebbe essere di grande utilità per il professionista. Le complesse e articolate fasi che regolano le tradizionali e innovative tecniche di rilevamento, dall'acquisizione alla visualizzazione 3D, presuppongono, infatti, un rilievo topografico, fotogrammetrico³ e laser scanner per ottenere un modello 3D⁴ attraverso l'uso integrato di svariate metodologie ormai consolidate e ampiamente divulgate nei settori scientifici di interesse. Gli esempi riportati sono la sintesi di numerose sperimentazioni tese a verificare le procedu-

re d'uso dei due pacchetti software 123D Catch della Autodesk e Visual SfM+CMVS/PMVS2 creato da Changchang Wu⁵. Entrambi i software sono *user-friendly*, propongono elaborazioni veloci e sono utilizzabili anche da non esperti del settore del rilevamento e della modellazione. Tuttavia, al fine di ottenere un modello 3D affidabile sia sotto il profilo visivo che metrico, è indispensabile individuare un criterio per la realizzazione del *dataset* fotografico di partenza. In questo articolo si dimostrerà che se per il Catch della Autodesk, l'individuazione del *dataset* è di facile intuizione, per il Visual sono ancora in corso sperimentazioni sia in campo nazionale che internazionale: spesso si giunge al *dataset* ottimale dopo diversi tentativi e non sempre il criterio utilizzato in una tipologia architettonica ha la stessa risposta in una tipologia differente. Infine, la valutazione eseguita sul modello ottenuto è stata condotta attraverso la comparazione con la scansione laser realizzata sullo stesso oggetto. Dopo avere indagato sullo stato dell'arte, sono stati analizzati i risultati ottenuti da ciascun



1/ *Pagina precedente.* A sinistra elemento architettonico (XIX secolo), al centro Torre di Federico a Enna, a destra facciata dell'edificio INCIS a Enna.

Previous page. *Left: architectural element (nineteenth century); centre: Frederick's Tower in Enna; right: façade of the INCIS building in Enna.*

software per ogni caso studio. I casi studio presi in esame, costituiscono tre tipologie di intervento molto diverse tra loro: l'elemento architettonico/scultoreo, l'edificio intero, la facciata (fig. 1).

Stato dell'arte

In letteratura troviamo prevalentemente due categorie principali di studi sui pacchetti SfM: alcuni sono finalizzati alla presentazione delle performance dei diversi algoritmi, della loro combinazione e ottimizzazione, attraverso la ricostruzione di scene 3D complesse con applicazioni su piccola, media e grande scala fino ad arrivare alla ricostruzione del Colosseo e di piazza San Marco⁶; altri sono finalizzati alla comparazione visiva e metrica dei differenti software disponibili, al fine di comprenderne l'effettivo utilizzo.

Per quanto riguarda lo stato dell'arte sui due software da noi indagati, si riportano i lavori più significativi sia su piccola che su grande scala: digitalizzazione di piccoli oggetti⁷; ritrovamenti archeologici come frammenti o suppellettili⁸; porzioni di siti archeologici⁹; applicazioni su edifici molto semplici¹⁰; comparazioni tra diversi pacchetti SfM su oggetti a scala diversa, dalla primitiva geometrica alla piazza¹¹; verifiche su 123D Catch a scala architettonica e urbana¹²; sperimentazioni alla scala urbana su VisualSfM¹³.

Verifica delle procedure SfM

La realizzazione del *dataset* è la fase più critica nelle tecniche SfM. Da esso, infatti, dipende la qualità del risultato finale.

Se per 123D Catch è sufficiente seguire il tutorial della Autodesk, per VisualSfM, ad oggi, non si può fare riferimento a un criterio univoco secondo cui eseguire una campagna fotografica che porti a un risultato atteso privo di imprecisioni, sbavature, bassa accuratezza visiva e metrica, etc.

Nelle sperimentazioni condotte, infatti, la creazione dei *dataset* per il VisualSfM è stata molto problematica e non sempre sono stati raggiunti gli obiettivi iniziali sperati. D'altra parte, la maggior parte degli studi presenti in letteratura, compresi quelli condotti dallo stesso Furukawa¹⁴ sono incentrati sulla ricostruzione di ampie scene urbane, dove il *dataset* è costituito

principalmente da immagini disponibili sul web a cui, nelle ultime sperimentazioni, si vanno ad aggiungere foto aeree oblique e le viste *street view*¹⁵. Questo impone applicazioni su monumenti molto visitati (Colosseo, piazza San Marco, etc.) con *dataset* non strutturati e con migliaia e migliaia di immagini il più delle volte ridondanti, la cui gestione richiede accorgimenti che non sono sempre necessari nel caso di acquisizioni di beni architettonici (facciate, piccoli edifici civili, elementi architettonici, etc.).

Alla ridondanza di immagini delle sperimentazioni di Furukawa, cui corrispondono anche risorse hardware e conoscenze software non indifferenti, si contrappone la necessità per il professionista di comprendere chiaramente quale sia il criterio secondo cui realizzare il *dataset* adeguato al fine di giungere ad un risultato apprezzabile.

Nel caso della facciata dell'edificio INCIS a Enna¹⁶ degli anni Trenta, a quattro elevazioni (m 29,30 x 18,00), alla semplicità della procedura applicativa del Catch in termini di rapporto tra qualità dell'elaborato finale/tempi di elaborazione/strumentazioni necessarie/ conoscenze preliminari richieste, si è contrapposta la difficoltà operativa in Visual di pervenire ad un *dataset* utile per una ricostruzione almeno visivamente accettabile. Inoltre, è bene sottolineare che mentre il Catch restituisce un modello 3D poligonale fotografico, l'output finale fornito dal VisualSfM integrato con CMVS/PMVS2 è una nuvola di punti (*dense reconstruction*)¹⁷. Pertanto, fin tanto che le sperimentazioni hanno come finalità la visualizzazione della *dense reconstruction* ottenuta, il VisualSfM può essere considerato uno strumento molto promettente. Tuttavia, il professionista e, quindi, il settore dei Beni Culturali, ha necessità di disporre anche di una *mesh* e questo passaggio, come mostrato in questo studio, non sempre conduce a risultati utilizzabili.

Come si evince dalla figura 2 i *dataset* utilizzati per il Catch (a) e per il Visual (b) differiscono in termini di numero di immagini. Il modello poligonale ottenuto con 123D Catch è stato confrontato con la *mesh* della scansione laser all'interno di Meshlab¹⁸ (fig. 3).

La sperimentazione in Visual con il *dataset* del Catch non ha dato risultati leggibili ed è stato necessario integrarlo con altre immagini

packages: the 123D Catch produced by Autodesk, and the Visual SfM+CMVS/PMVS2 made by Changchang Wu.⁵ Both are user-friendly, provide rapid processing, and can be used by amateurs in the field of survey and modelling. However, to obtain a visually and metrically reliable 3D model it's important to establish criteria with which to create the initial photographic dataset. This article will show that it's easy to establish a dataset for Autodesk 123D Catch, while tests are still ongoing in Italy and abroad for Visual: several attempts are often needed to obtain the best dataset and the criteria used in one architectural type does not always give the same results for another type. Finally, we assessed the model by comparing it against a laser scan of the same object. After having examined the state-of-the-art, we analysed the results produced by the two software programmes for each case study. The studies in question represent three very different kinds of artefacts: an architectural/sculptural element, a building, and a façade (fig. 1).

State-of-the-art

Two main types of studies on SfM packages have been reported in literature: some present the performance, combination and optimisation of the various algorithms by recreating complex 3D scenes with small, medium and large scale applications, for example reconstructions of the Colosseum and St. Mark's Square⁶; other focus on visual and metric comparisons between the different software programmes in order to understand how they can be used in practice. As regards the state-of-the-art of the two software programmes in this study, we decided to report on the most important executed on a small and large scale: the digitalisation of small objects⁷; archaeological finds such as fragments or grave goods⁸; parts of architectural sites⁹; applications used on very simple buildings¹⁰; comparisons between objects with different scales executed using several SfM packages ranging from primitive geometry to the square¹¹; tests on an architectural and urban scale using the 123D Catch¹²; urban scale tests using VisualSfM.¹³

Tests on the SfM procedures

Creating the datasets is the most critical part of the SfM techniques because it influences the quality of the final result. For the 123D Catch,

2/ Edificio INCIS a Enna: a destra ricostruzione della scena e della posizione delle fotocamere in 123D Catch, a sinistra in VisualSfM.

The INCIS building in Enna. Right: reconstruction of the scene and position of the cameras in 123D Catch; left: in VisualSfM.

3/ Edificio INCIS a Enna: a destra visualizzazione del modello mesh in 123D Catch, a sinistra verifica metrica in Meshlab.

The INCIS building in Enna. Right: visualisation of the mesh model in 123D Catch; left: metric verification in Meshlab.

4/ Edificio INCIS a Enna: a destra visualizzazione della dense

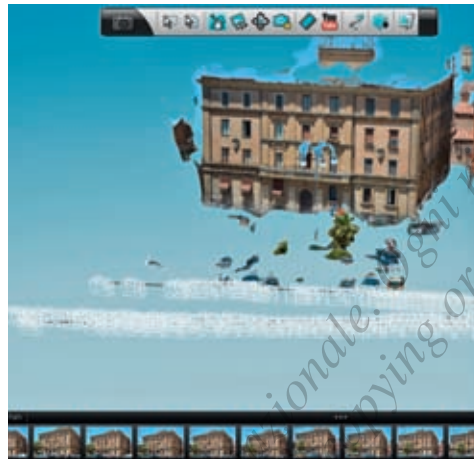
reconstruction in Visual SfM, a sinistra verifica metrica in Meshlab.

The INCIS building in Enna. Right: visualisation of the dense reconstruction in Visual SfM; left: metric verification in Meshlab.

the Autodesk tutorial is more than enough. Instead, the VisualSfM has so far not provided univocal criteria with which to perform the photographic campaign and obtain a result without imprecisions, imperfections, modest visual and metric accuracy, etc.

In fact, during the tests it was much more difficult to create VisualSfM datasets, and we were not always able to achieve the results we expected. It's important however, to remember that most of the studies reported in literature, including the ones by Furukawa himself,¹⁴ focus mainly on the reconstruction of big urban scenes in which the dataset is essentially made up of images available on the web; the most recent tests have also included oblique aerial photographs and street views.¹⁵ This requires applications involving famous monuments (the Colosseum, St. Mark's Square, etc.) with unstructured datasets and thousands and thousands of often superfluous images. Managing these images requires tricks which are not always necessary when acquiring architectural assets (façades, small civilian buildings, architectural elements, etc.). Furukawa employed important hardware resources and software expertise when executing his tests with the redundant images; on the contrary, if a professional wants to produce a significant result, he needs to clearly understand the criteria required to create the dataset.

Let's take the façade of the four storey INCIS building (29.30 x 18.00 m) built in Enna¹⁶ in the thirties. While the application procedures proposed by Catch produced a good quality final image compared to the processing times, tools and preliminary data required, it was operationally difficult to use Visual to create a dataset which could produce, at the very least, a visually acceptable reconstruction. Furthermore, we should point out that Catch provides a photographic polygonal 3D model, while the final output provided by the Visual SfM integrated with CMVS/PMVS2 is a points cloud (dense reconstruction).¹⁷ As a result, the VisualSfM can be considered a very promising tool if the tests were intended to visualise the dense reconstruction. However, professionals and the cultural heritage sector also need a mesh; this study proves that the



di dettaglio. Attraverso la ricostruzione mediante l'algoritmo di Poisson (profondità 12), si è passati dalla nuvola di punti del Visual alla relativa *mesh*, tuttavia il modello così ottenuto è risultato inutilizzabile poiché molto rumoroso, non definito e geometricamente deformato. Pertanto la verifica metrica è stata condotta in Meshlab attraverso la comparazione della *dense reconstruction* con la nuvola della scansione laser precedentemente allineate in CloudCompare¹⁹ (fig. 4). I risultati ottenuti sono sintetizzati in tabella 1.

Gli esiti raggiunti sulla facciata dell'INCIS ci hanno guidato nella realizzazione del dataset

del secondo edificio: la Torre di Federico²⁰ a Enna, del XIII secolo (lato dell'ottagono di base m 8, altezza m 27). Nonostante gli accorgimenti messi a punto nella procedura di utilizzo sia in Catch sia in Visual, l'elaborazione di questo particolare monumento ci ha condotti verso risultati non previsti. Probabilmente la geometria centrica, con base ottagonale e altezza quasi il doppio rispetto all'imposta di base, la stereotomia e l'uniformità cromatica della superficie di rivestimento hanno generato problemi di elaborazione agli algoritmi interni. Il dataset per il Catch è stato obbligato dall'esigenza di compiere un giro completo attorno

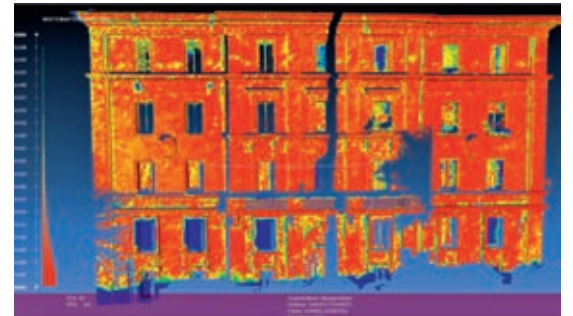


Tabella 1/ Edificio INCIS a Enna.
 Table 1/ The INCIS Building in Enna.
 Tabella 2/ Torre di Federico a Enna.
 Table 2/ Frederick's Tower in Enna.

all'edificio (fig. 5) con una sovrapposizione tra un'inquadratura e l'altra di circa il 70%²¹ con un totale di 82 scatti. La difficoltà del software è stata quella di elaborare la *mesh* in modalità *standard* e *maximum* e, questo, ha compromesso la verifica finale poiché la *mesh* in modalità *mobile* non è attendibile ai fini del confronto. L'errore medio ricavato dalla comparazione con la scansione è, infatti, di 0,0375 m e, inoltre, alcune zone presentano degli scostamenti dell'ordine di 0,14 m (tab. 2).

Se 123D Catch ha difficoltà a elaborare un numero elevato di immagini²², il VisualSfM integrato con CMVS/PMVS2, al contrario, è stato ideato per l'elaborazione di considerevoli quantità di immagini non strutturate. Le nostre sperimentazioni sono state rivolte a comprendere se e come sia possibile individuare una procedura operativa per ottimizzare le acquisizioni fotografiche e realizzare un *dataset* strutturato utilizzabile in maniera efficace in Visual. È stato pertanto eseguito un altro giro di foto attorno alla torre a quota differente rispetto alla prima (fig. 6). Anche in questo caso, l'esito raggiunto non ha dato un risultato soddisfacente con un errore medio tra i due modelli di 0,0537 m e una deviazione standard di 0,0676 m (fig. 7).

Probabilmente, l'implementazione del *dataset* con altre immagini prese a una quota più elevata/aerea e l'aggiunta di *marker* sulla superficie di rivestimento avrebbero potuto contribuire a migliorare il risultato del Visual. Tuttavia, non è stato possibile recarsi in sito per nuove elaborazioni.

Le difficoltà operative che si incontrano nello studio su un'opera architettonica, quali accessibilità, distanza, etc. (difficoltà che non ci hanno consentito, come già detto, ulteriori approfondimenti sui primi due casi indagati), sono state superate nella sperimentazione condotta su un elemento architettonico/scultoreo del XIX secolo. Si tratta di una mensola in gesso con due lati decorati da bassorilievi di intrecci fitomorfi chiusi nelle estremità da due leoni alati (larghezza 902 mm, altezza 803 mm, spessore 200 mm), disponibile presso i locali del Dipartimento di Architettura di Palermo. La sperimentazione con il Catch ha elaborato un risultato molto promettente sia sotto il profilo visivo che metrico (figg. 8-9, tab. 3).

Facciata dell'INCIS larghezza m 29,30, altezza m 18,00		
	123D Catch	VISUALSfM+CMV S/PMVS2
# di immagini	115	174
#di immagini ridondanti	0	0
Macchina fotografica	Nikon D3200	Nikon D3200
Risoluzione della camera	12 Mpix	12 Mpix
Distanza focale	18 mm	18-55 mm
Tempo totale di calcolo	45 min (sul cloud)	120 min (sul desk)
# di clusters	-	3
Dense point cloud	-	2.106.267
Mesh	1.400.662 triangoli 700.699 vertici	-
Mesh da laser scanner HDS 3000 della Leica Geosystem/accuratezza 6 mm	961.159 vertici 1920.350 triangoli	
Intervallo di calcolo (m)	0-0,05	0-0,15
Media (m)	0,010331	0,050275
Scarto quadratico medio (RMS)	0,014706	0,059693

L'accessibilità al reperto storico ha consentito di condurre una campagna di sperimentazioni molto intensa finalizzata all'individuazione di un criterio per la realizzazione di un *dataset* adeguato per il Visual. In questo articolo, tra le sperimentazioni più significative, se ne riportano tre. Il risultato ottenuto richiede un'attenzione esclusiva e pertanto si tratta l'argomento in un paragrafo dedicato.

Sperimentazione dell'elemento architettonico in VisualSfM

L'esperienza maturata sul Visual nei due casi precedenti e il confronto dei risultati ottenuti con quanto disponibile in letteratura, ha portato alla realizzazione di un primo *dataset* che fosse il più possibile compatto e corposo con 482 foto, così come si può evincere dalla figura 10). Il risultato ottenuto, tuttavia, è stato

Torre di Federico larghezza m 16,80, altezza m 27		
	123D Catch	VISUALSfM+CMV S/PMVS2
# di immagini	82	274
#di immagini ridondanti	0	0
Macchina fotografica	Nikon D3200	Nikon D3200
Risoluzione della camera	12 Mpix	12 Mpix
Lente focale	18 mm	18-55 mm
Tempo totale di calcolo	360 min (sul cloud)	180 min (sul desk)
# di clusters	-	3
Dense point cloud	-	4.206.870
Mesh	693.736 triangoli 368.357 vertici	6.810.474 triangoli 3.409.836 vertici
Mesh da laser scanner: HDS 3000 della Leica Geosystem/accuratezza 6 mm	4.112.952 triangoli 2.16.608 vertici	
Intervallo di calcolo (m)	0-0,20	0-0,20
Media (m)	0,037482	0,053757
Scarto quadratico medio (RMS)	0,057935	0,067642

VisualSfM does not always produce practical results.

Figure 2 shows how the datasets used by Catch (a) and Visual (b) differ in terms of the number of images. The polygonal model created using 123D Catch was compared with the mesh of the laser scansion in Meshlab¹⁸ (fig. 3).

Tests in Visual using the Catch dataset did not produce interpretable results and had to be integrated with other images of details. Reconstruction using Poisson's algorithm (depth 12) allowed us to pass from the Visual points cloud to the relative mesh. However the ensuing model was useless because it was extremely noisy, undefined and geometrically deformed. As a result, a metric test was conducted in

Meshlab by comparing the dense reconstruction with the cloud of the laser scansion previously aligned in CloudCompare¹⁹ (fig. 4). A summary of the results are presented in Table 1. The results regarding the façade of the INCIS building helped us create the dataset of the second building: Frederick's Tower²⁰ in Enna built in the thirteenth century (the width of the octagon 8 m, height 27 m). Despite the care taken when using Catch and Visual, processing this monument produced unexpected results. The problems arose during the processing of the internal algorithms not only due to the centric geometry with its octagonal base and height almost double that of the base layout, but also due to the stereotomy and chromatic uniformity of the surface envelope.

The Catch dataset required we executed a complete circle of the building (fig. 5) and superimposed one frame over another for about 70%²¹ of the building: a total of 82 frames. The software had difficulty processing the mesh in standard and maximum mode; this compromised the final text since the mesh in mobile mode was unreliable in a comparison. The mean error of the comparison with the scansion was, in fact, 0,0375 m; furthermore, some areas had differences of around 0,14 m (Tab. 2). If the 123D Catch had difficulty processing a large number of images,²² on the contrary the VisualSfM integrated with CMVS/PMVS2 has been designed to process a significant number of

5/ Torre di Federico a Enna: a sinistra ricostruzione della scena e della posizione delle fotocamere in 123D Catch, a destra VisualSfM.

Frederick's Tower in Enna. Left: reconstruction of the scene and the position of the cameras in 123D Catch; right: in VisualSfM.

6/ Torre di Federico a Enna: a sinistra visualizzazione del modello mesh in 123D Catch, a destra verifica metrica in Meshlab.

Frederick's Tower in Enna. Left: reconstruction of the mesh model in 123D Catch; right: metric verification in Meshlab.

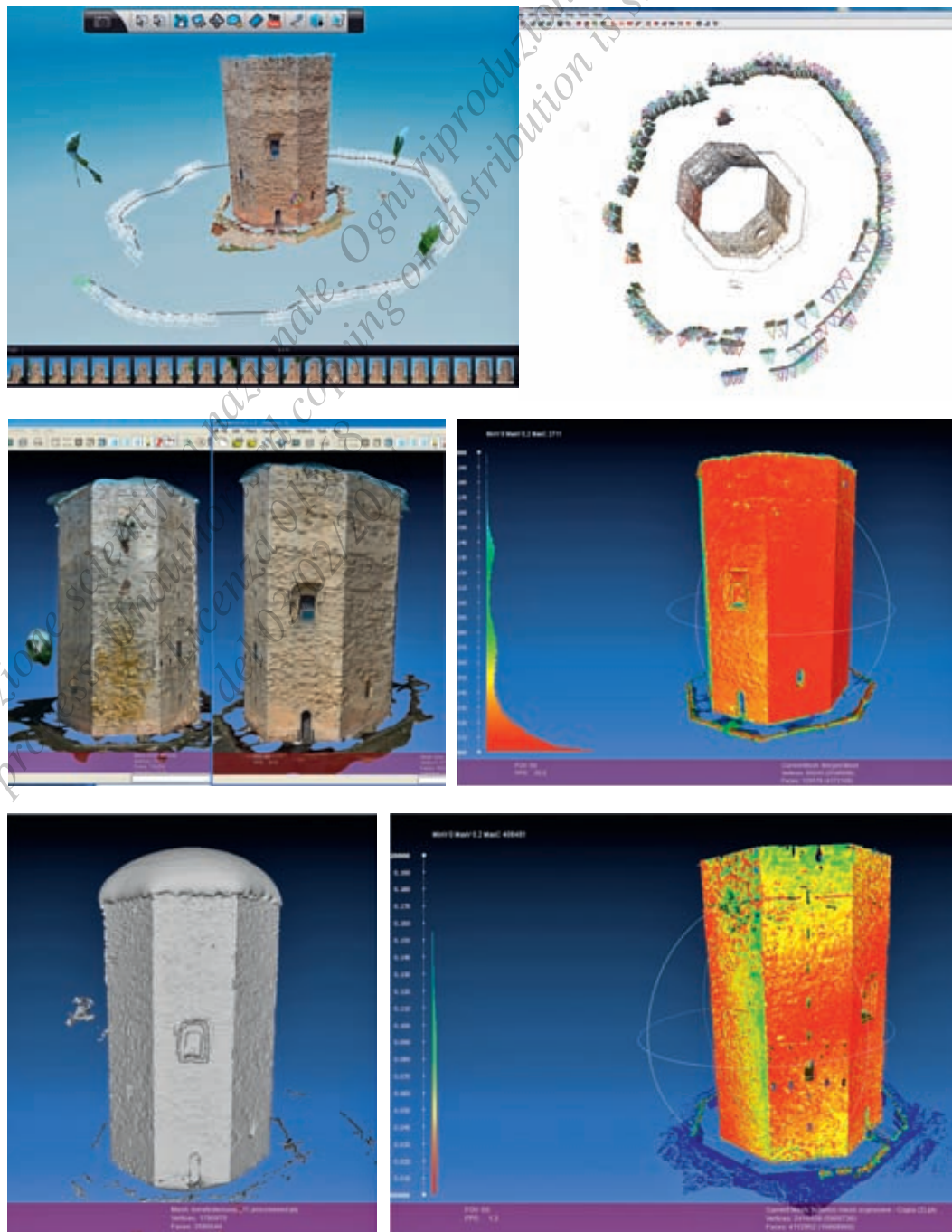
7/ Torre di Federico a Enna: a sinistra visualizzazione della dense reconstruction in Visual SfM, a destra verifica metrica in Meshlab.

Frederick's Tower in Enna. Left: visualisation of the dense reconstruction in Visual SfM; right: metric verification in Meshlab.

unstructured images. Our tests were intended to check whether and how it is possible to identify an operational procedure to optimise the photographic acquisitions and create a structured dataset which can be successfully used in Visual. So we took another set of photographs of the tower at a different level from the previous one (fig. 6). Here too, the results were not satisfactory; the mean error between the two models was 0.0537 m with a standard deviation of 0.0676 m (fig. 7). It might have been possible to improve the result of the Visual by using a dataset with other images taken at a higher level (or from the air) and by adding markers on the surface of the envelope. However, we were unable to return to the site for further tests. The operational difficulties we encountered while studying an architectural artefact (accessibility, distance, etc.) did not allow us to perform an in-depth study of the first two case studies. However we solved these difficulties during the study of a nineteenth-century architectural/sculptural element: a plaster corbel decorated on two sides by bas-reliefs of phytomorphic tracery closed at either end by two winged lions (width 902 mm, height 803 mm, depth 200 mm). The corbel is located in the Department of Architecture in Palermo. The tests with Catch provided very promising visual and metric results (figs. 8-9; Tab. 3). Accessibility of the historical artefact allowed us to carry out a very intense experimental campaign to identify the criteria needed to create the right Visual dataset. We will report on three of these artefacts in this article. The results require careful attention and are therefore illustrated in a separate paragraph.

Experimenting with the architectural element using VisualSfM

Our work with Visual in the two previous case studies, and comparison between the data we obtained and the data available in literature, allowed us to create an initial dataset as compact and dense as possible: it included 482 photographs and is shown in figure 10. However the result was very different to the one we expected; this is evident in the reconstruction of the surface using the Poisson algorithm (depth 12) visualised in Meshlab (fig. 10). The figure shows that the quality of the mesh in terms of noise, definition, etc., is not enough to carry out verification. In our



molto difforme da quello atteso come si può osservare dalla ricostruzione della superficie mediante algoritmo di Poisson (profondità 12) visualizzata in Meshlab (fig. 10). Come si evince dalla figura, la qualità della mesh, in termini di rumorosità, definizione, etc., non è sufficiente

per condurre la verifica. Nel tentativo successivo si è pensato di eseguire un rilievo fotografico su tre livelli di prese con un totale di 182 immagini secondo il criterio fotogrammetrico. Tuttavia, il terzo e più alto giro di prese (60 immagini) è stato scartato durante l'elaborazione

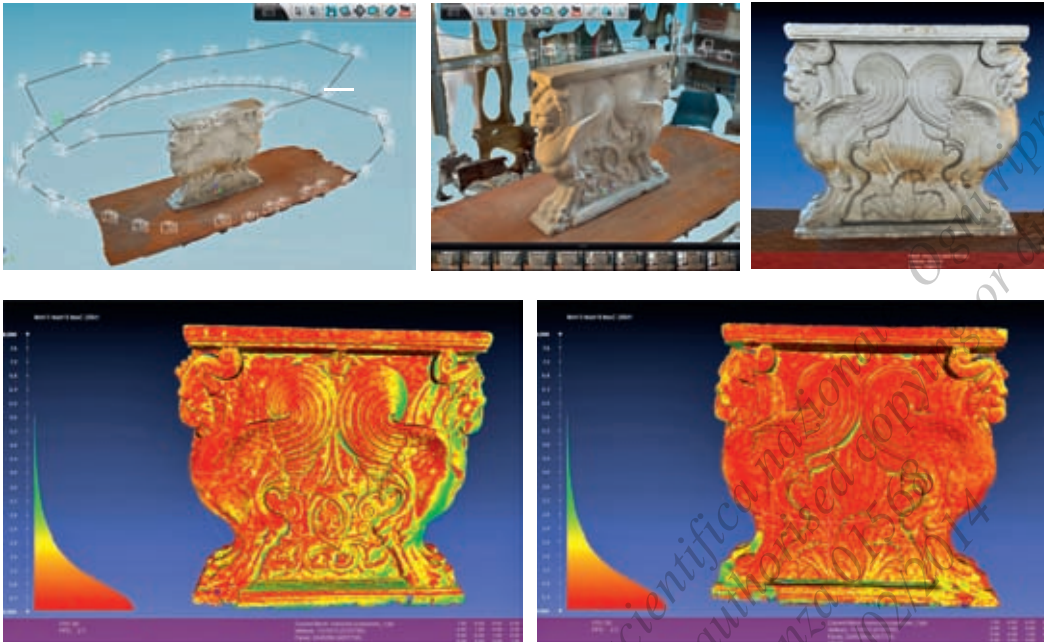
8/ Elemento architettonico: a destra e al centro ricostruzione della scena e della posizione delle fotocamere, a sinistra visualizzazione del modello mesh in 123D Catch

Architectural element. Right and centre: reconstruction of the scene and position of the cameras; left: visualisation of the mesh model in 123D Catch.

9/ Elemento architettonico: verifica metrica in Meshlab sui due lati del modello mesh ottenuto con 123D Catch

Architectural element: metric verification in Meshlab along two sides of the mesh model obtained using the 123D Catch

Tabella 3/ Elemento architettonico.
Table 3/ Architectural element.



a causa dell'inquadratura troppo angolata verso il basso, che rifletteva la cromia del pavimento in marmo. Come è chiaramente visibile in figura 11, la *mesh* (Poisson, profondità 12) ottenuta dal *dataset* con le restanti 122 immagini, riflette la conseguente perdita di informazioni. Quanto detto conferma che questo software è stato creato con algoritmi che gestiscono un numero molto elevato di immagini. Come già accennato, le sperimentazioni mostrate sono solo una piccola parte dei diversi tentativi elaborati. L'obiettivo era quello di capire quale fosse il criterio da suggerire all'utente per un corretto utilizzo del software. Tuttavia, per quanto nell'applicazione che segue dimostreremo d'essere giunti al risultato sperato, ad oggi non siamo in grado di stabilire se quanto ottenutò sia frutto del criterio seguito o semplicemente del "caso"; è necessario continuare a indagare verificando questa ipotesi su ulteriori casi studio.

Il *dataset* mostrato in figura 12 è una sintesi di tutte le sperimentazioni: risente sia del criterio fotogrammetrico (rapporto base/distanza) che di quello numerico. Al secondo *dataset* sono stati aggiunti: un giro completo di scatti alla quota superiore con un'angolazione meno accentuata rispetto alla sperimentazione precedente e un quarto giro ravvicinato per i dettagli, per un to-

tales di 233 immagini. La verifica metrica condotta ha portato ai risultati descritti in figura 13 e tabella 3 dove sono riportati i valori di spostamento su entrambi i lati dell'oggetto.

Conclusioni

Attraverso le sperimentazioni condotte e la relativa individuazione delle caratteristiche dei rispettivi *dataset* sia per il Catch che per il Visual, è possibile delineare i limiti e i vantaggi di entrambi i software. In questo modo il professionista, consapevole delle potenzialità degli strumenti di cui dispone, è nelle condizioni di scegliere quale tecnica intraprendere all'inizio del suo rilievo senza incorrere a sgra-

next attempt we decided to perform a photographic survey on three different levels (a total of 182 images) using photogrammetric criteria. However the third and highest level (60 images) was discarded during processing due to the excessively downward-pointing image which reflected the colour of the marble floor. Figure 11 clearly shows that the mesh (Poisson, depth 12), obtained from the dataset using the other 122 images reflects the ensuing loss of data, proof that this software was created with algorithms that manage a very large number of images. As mentioned earlier, these tests represent only a small number of the many attempts we made. We wanted to understand what criteria should be suggested to users for them to be able to use the software properly. However, although the demonstration illustrated below shows we obtained what we were looking for, we still cannot tell whether this depends on the fact we used the right criteria or whether it was simply a 'fluke'. To be sure, we would have to test our theory using other case studies.

The dataset shown in figure 12 summarises all the tests: it takes into account the photogrammetric criteria (base/distance ratio) and the numeric criteria. We added a second dataset: a complete round of shots at the upper level with a less accentuated angle compared to the previous test, and a close-up quarter turn for the details: a total of 233 images.

The metric verification produced the results in figure 13 and Table 3 showing the differences in value on both sides of the object.

Conclusions

The tests we carried out, and the fact we were able to identify the characteristics of the respective datasets for Catch and Visual made it possible to outline the pros and cons of both software programmes. Once a professional is aware of the potential of these tools he can make an informed decision about which techniques to use before starting his survey; this also allows him to avoid unpleasant and unexpected results. The characteristics of the two software programmes are listed below.

Dataset:

- the photographic dataset of the 123D Catch is structured and simple to create because it is based on the logic of photogrammetry. The number of images ranges from a minimum of 30 to a

Elemento architettonico larghezza mm 902, altezza mm 803, spessore mm 200		
	123D Catch	VISUAL/SIM+CM VS/PMVS2
# di immagini	67	233
# di immagini ridondanti	0	0
Macchina fotografica	Nikon D3200	Nikon D3200
Risoluzione della camera	12 Mpix	12 Mpix
Lente focale	18 mm	18-55 mm
Tempo totale di calcolo	180 min (sul cloud)	180 min (sul desk)
# di clusters		5
Dense point cloud		3,712,553
Intervallo di calcolo (mm)	0-8	0-8
Media (mm)	1.078	0.934
Scarto quadratico medio (RMS)	1.477	1.371
Mesh	1,189,176 triangoli 595,045 vertici	2,860,050 triangoli 1,439,350 vertici
Mesh da laser scanner: Minolta Vivid 9i distanza 70 cm/accuratezza 0.25-0.3 mm	2,245,298 triangoli 1,134,013 vertici	

10/ Elemento architettonico dataset 482 foto: a destra visualizzazione della ricostruzione della posizione delle camere fotografiche, a sinistra della mesh Poisson.
 Architectural element dataset 482 photographs. Right: visualisation of the reconstruction of the position of the cameras; left: of the Poisson mesh.

11/ Elemento architettonico dataset 182 foto: a destra visualizzazione della ricostruzione della posizione delle fotocamere, a sinistra della mesh Poisson.
 Architectural element dataset 182 photographs. Right:

visualisation of the reconstruction of the position of the cameras; left: of the Poisson mesh.

12/ Elemento architettonico dataset 233 foto: a destra visualizzazione della ricostruzione della posizione delle fotocamere, a sinistra della mesh Poisson.
 Architectural element dataset 233 photographs. Right: visualisation of the reconstruction of the position of the cameras; left: of the Poisson mesh.

maximum number of 70 (advised by Autodesk) even if in our trials we obtained good results with up to 200 photographs according to the scale of the object to be acquired;

- for the VisualSfM+CMVS/PMVS2, examples in literature refer mainly to unstructured photographic datasets. Our trials achieved an acceptable structured dataset only when architectural elements were involved. However further trials should be performed on similar objects and on larger scales to identify a reference criteria. The number of images ranges from a minimum of 120 to an undefined maximum. Reports in literature include tests with up to 13,800 images.²³

Processing times and hardware resources:

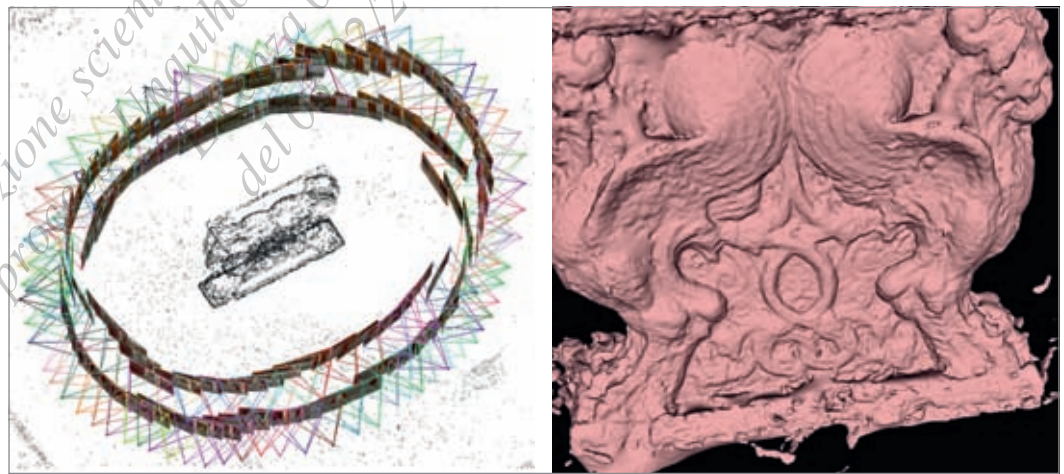
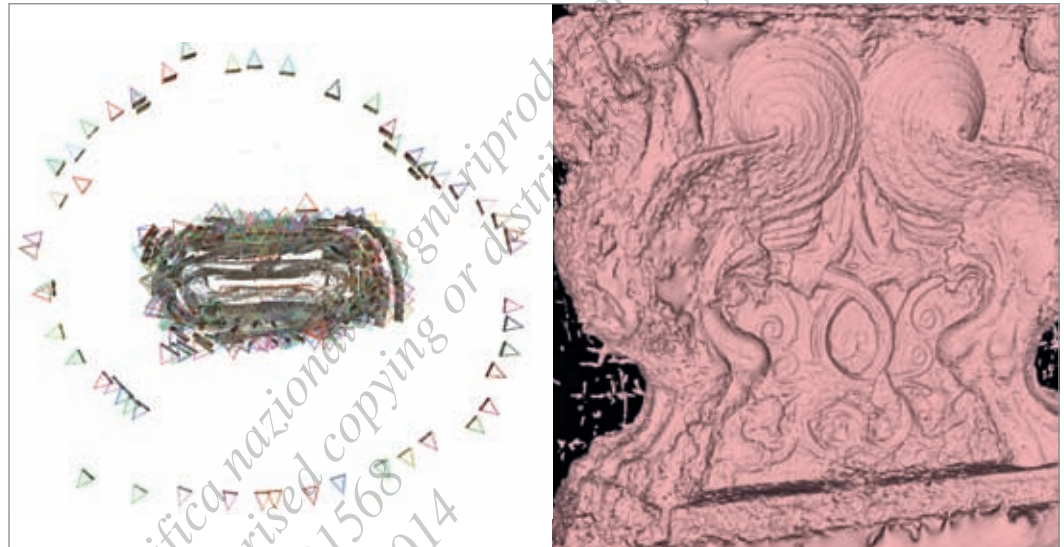
- 123D Catch processing takes place on the cloud. For a professional, but more so for a user, this is very advantageous; however, in some cases it is offset by the problem of data privacy (photographic image rights, etc.);
 - processing using VisualSfM+CMVS/PMVS2 requires long calculations and will crash if a Ram memory of at least 13 Giga and two processors are not available; these technical characteristics do not always satisfy a professional.

Output and easy use of the package:

- the end product provided by the 123D Catch is a mesh from which transversal and longitudinal sections can be extrapolated, as well as all the other drawings required by a professional; 123D Catch is also suited to users unfamiliar with computer graphics, computer vision, algorithms, etc. The system works like a 'black box' which receives the photographs as an input and provides a mesh as output. The user cannot intervene on the end product except to improve the collimation of the points;
 - the end product of the VisualSfM+CMVS/PMVS2 is an unstructured points cloud which requires specific skills to be turned into a mesh; not all professionals may possess these skills. The VisualSfM+CMVS/PMVS2 allows the user to work on the parameters and manage the reconstruction; however, despite the intuitive graphic interface, a user has to use his own skills and IT expertise to be able to understand how the system works.

Visual accuracy:

- the visual accuracy of the 123D Catch is very effective. However, the following rules have to



13/ Elemento architettonico dataset 233 foto: verifica metrica sui due lati della ricostruzione ottenuta con 123D Catch
Architectural element dataset 233 photographs. Metric verification of the two sides of the reconstruction obtained using the 123D Catch.

devoli e inattesi risultati. Si elencano di seguito le caratteristiche di entrambi i software.
Dataset:

- il *dataset* fotografico del 123D Catch è strutturato ed è semplice da realizzare perché segue la logica della fotogrammetria. Il numero di immagini va da un minimo di 30 a un massimo consigliato dalla Autodesk di 70 (anche se nelle nostre sperimentazioni si sono ottenuti buoni risultati fino a circa 200 foto) a seconda della scala dell'oggetto da acquisire;

- per il VisualSfM+CMVS/PMVS2, gli esempi presenti in letteratura si attengono a *dataset* fotografici prevalentemente non strutturati. Le nostre sperimentazioni sono pervenute a un risultato accettabile di *dataset* strutturato solo nel caso dell'elemento architettonico, ma è necessario condurre ulteriori verifiche su oggetti simili e a scala maggiore per individuare un criterio cui fare riferimento. Il numero di immagini va da un minimo di 120 a un massimo non ancora definito. In letteratura si ritrovano sperimentazioni con più di 13.800 immagini²³.

Tempi di elaborazione e risorse hardware:

- l'elaborazione di 123D Catch avviene sul *cloud* e questo per il professionista, ma più in generale per l'utente, è un vantaggio non indifferente; a questo vantaggio si contrappone, in alcuni casi, il problema della privacy dei dati (diritti sulle immagini fotografiche, etc.);

- l'elaborazione del VisualSfM+CMVS/PMVS2 richiede molte ore di calcolo e se non si dispone di una memoria Ram di almeno 13 Giga e di due processori, l'elaborazione va in *crash*; queste caratteristiche tecniche, non sempre rispondono alle aspettative del professionista.

Output e facilità di utilizzo del pacchetto:

- il prodotto finale elaborato dal 123D Catch è una *mesh* da cui è possibile estrapolare sezioni

trasversali e longitudinali e tutti gli altri elaborati, necessari per il lavoro del professionista; 123D Catch è adatto anche a utenti che non abbiano molta dimestichezza con conoscenze di *computer graphics*, *computer vision*, algoritmi e quant'altro. Il sistema lavora come una "blackbox" che riceve in input le immagini fotografiche e restituisce in output una *mesh*. L'utente quindi non può interagire con il prodotto finale se non per migliorare la collimazione dei punti;

- il prodotto finale del VisualSfM+CMVS/PMVS2 è una nuvola non strutturata di punti che, per essere usata deve essere convertita in *mesh*, passaggio, questo, che richiede competenze specifiche che non è detto che il professionista abbia. Inoltre, il VisualSfM+CMVS/PMVS2 consente di agire sui parametri che gestiscono la ricostruzione; tuttavia, nonostante l'interfaccia grafica sia intuitiva, la comprensione da parte dell'utente di come operare è lasciata alle sue abilità e conoscenze informatiche.

Accuratezza visiva:

- l'accuratezza visiva del 123D Catch è di grande effetto, tuttavia è necessario rispettare i seguenti vincoli: usare la stessa macchina fotografica e la stessa lente focale, realizzare un *dataset* fotografico strutturato che inquadra l'edificio nella sua interezza (nel rispetto dei parametri del rilievo fotogrammetrico) ed esportare la *mesh* in modalità *maximum*;

- l'accuratezza visiva del VisualSfM+CMVS/PMVS2 è di grande effetto una volta individuato il *dataset* fotografico ottimale.

Dunque, possiamo concludere che entrambi i software hanno grandi potenzialità, tuttavia se il Catch ha i limiti del sito e delle relative inquadrature, il Visual è ancora in fase di sperimentazione e richiede risorse hardware e conoscenze preliminari di un certo livello.

be observed: use the same camera and same focal lens; create a structured photographic dataset of the whole building (respecting the parameters of photogrammetric survey) and export the mesh in maximum mode;

- the visual accuracy of the VisualSfM+CMVS/PMVS2 is very effective once the optimal photographic dataset has been established. We can conclude that both software programmes have great potential, however if Catch is limited by the site and frames, Visual is still experimental and requires hardware resources and rather advanced preliminary skills.

Acknowledgements. The laser scansion of the INCIS building and Frederick's Tower in Enna were performed and aligned by Cettina Santagati and Mariateresa Galizia using the study carried out at the Laboratory of Architectural Photogrammetry and Survey 'Luigi Andreozzi' at the University of Catania. The scansion of the corbel was performed by Rosanna Sciortino from the DICAM Department of Engineering at the University of Palermo.

1. Pollefeys, Van Gool 2002; Snavely, Seitz, Szeliski 2006; Snavely, Seitz, Szeliski 2008; Vergauwen Maarten, Van Gool Luc. *Web-based 3D Reconstruction Service*. Machine Vision and Applications, 17, 2006. pp. 411-426; Pollefeys Marc, Van Gool Luc, Vergauwen Maarten, Verbiest Frank, Cornelis Kurt, Tops Jan, Koch Reinhard. *Visual modelling with a hand-held camera*. International Journal of Computer Vision, 59(3) pp. 207-232.

2. De Luca 2010; Remondino, El-Hakim 2006. Thomas Luhmann. *3D imaging: how to achieve highest accuracy*. In Proceedings SPIE 8085, Videometrics, Range Imaging, and Applications XI, 808502 (June 21, 2011).

3. Docci, Maestri 2009; Bertocci, Bini 2012; Cundari 2012.

4. Bandiera, Beraldin, Gaiani 2011; Benedetti, Gaiani, Remondino 2011; Guidi, Russo, Beraldin 2010; Migliari 2001; Andreozzi 2003. Addison, Gaiani 2000.

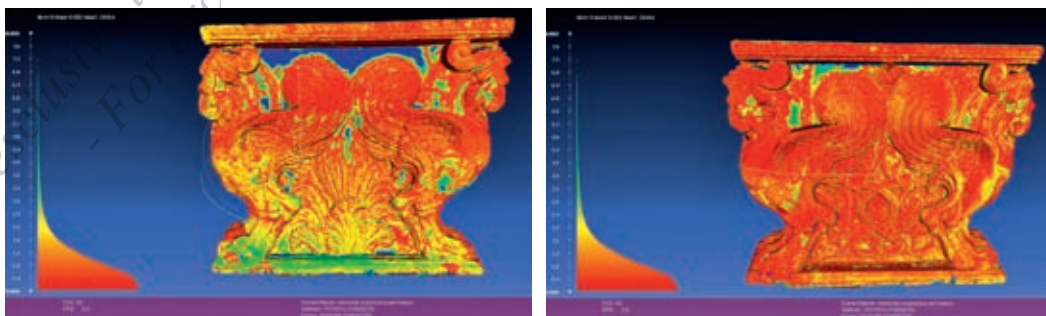
5. Wu 2011.

6. Agarwal, Furukawa, Snavely, Curless, Seitz, Szeliski 2010; Furukawa, Curless, Seitz, Szeliski 2010. Snavely, Seitz, Szeliski 2006; Snavely, Seitz, Szeliski 2008; Vu, Keriven, Labatut, Pons 2009.

7. Nguyen, Wunsche, Delmas, Lutteroth 2012.

8. Kersten, Lindstaedt 2012.

9. Lo Brutto, Meli 2012; Dellepiane, Dell'Unto, Callieri, Lindgren, Scopigno 2013.



10. Kersten, Stallmann 2012.
11. Remondino, Del Pizzo, Kersten, Troisi 2012.
12. Santagati, Inzerillo, Di Paola 2013; Santagati, Inzerillo 2013.
13. Inzerillo, Santagati, Di Paola 2013.
14. Furukawa, Curless, Seitz, Szeliski 2010; Furukawa Yasutaka, Jean Ponce. *Accurate, Dense, and Robust Multi-View Stereopsis*. IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), vol. 32, n. 8, 2010, pp. 1362-1376.
15. Shan Qi, Adams Riley, Curless Brian, Furukawa Yasutaka, Seitz Steven M. *The Visual Turing Test for Scene Reconstruction*. In 3DTV-Conference, 2013, pp. 25-32.
16. Restuccia, Galizia, Santagati 2012.
17. *The pipeline used by VisualSfM includes the following steps: 1) processing of the dataset to calculate the Sift and matchings by determining the features in the photographs; 2) processing the sparse reconstruction; 3) elimination of errors using the Multicore Bundle Adjustment algorithm; 4) calculation of the dense reconstruction using the algorithms PMVS2 and CMVS developed by Furukawa. Another process has to be added to these steps in VisualSfM: turning the points cloud into a mesh using the Poisson algorithm, as indicated in literature.* Wu 2011; Wu Changchang, Agarwal Sameer, Curless Brian, Seitz Steven M. *Multicore Bundle Adjustment*. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2011, pp.3057-3064; Wu Changchang. *SiftGPU: A GPU Implementation of Scale Invariant Feature Transform (SIFT)*, 2011, <http://www.cs.unc.edu/~ccwu/siftgpu> [November 2013].
18. *Meshlab is an open source software developed by the Visual Computing Lab of the ISTI - CNR in Pisa; it facilitates processing and managing the points clouds and meshes.* Cignoni Paolo, Callieri Marco, Corsini Massimiliano, Dellepiane Matteo, Ganovelli Fabio, Ranzuglia Guido. *MeshLab: an Open-Source Mesh Processing Tool*. In Sixth Eurographics Italian Chapter Conference, 2008, pp. 129-136.
19. *CloudCompare is an open source software facilitating processing and managing the points clouds and meshes. In this case we used CloudCompare to align the clouds, an operation not available in Meshlab.* <http://www.danielgm.net/cc/> [November 2013].
20. Galizia 2012.
21. Santagati, Inzerillo, Di Paola 2013; Santagati, Inzerillo 2013.
22. Autodesk suggests the use of a dataset with maximum 70 photos with a resolution of 3Mpix.
23. Furukawa, Curless, Seitz, Szeliski 2010.
- Ringraziamenti*
La scansione laser dell'edificio INCIS e della Torre di Federico a Enna sono state eseguite ed allineate da Cettina Santagati e Mariateresa Galizia durante l'attività di ricerca condotta presso il Laboratorio di Fotogrammetria Architettonica e Rilievo "Luigi Andreozzi" dell'Università di Catania. La scansione della mensola è stata realizzata da Rosanna Sciortino del Dipartimento DICAM di Ingegneria dell'Università di Palermo.
1. Pollefeys, Van Gool 2002; Snavely, Seitz, Szeliski 2006; Snavely, Seitz, Szeliski 2008; Vergauwen Maarten, Van Gool Luc. *Web-based 3D Reconstruction Service*. *Machine Vision and Applications*, 17, 2006, pp. 411-426; Pollefeys Marc, Van Gool Luc, Vergauwen Maarten, Verbiest Frank, Cornelis Kurt, Tops Jan, Koch Reinhard. *Visual modelling with a hand-held camera*. *International Journal of Computer Vision*, 59(3) pp. 207-232.
2. De Luca 2010; Remondino, El-Hakim 2006. Thomas Luhmann. *3D imaging: how to achieve highest accuracy*. In *Proceedings SPIE 8085, Videometrics, Range Imaging, and Applications XI*, 808502 (June 21, 2011).
3. Docci, Maestri 2009; Bertocci, Bini 2012; Cundari 2012.
4. Bandiera, Beraldin, Gaiani 2011; Benedetti, Gaiani, Remondino 2011; Guidi, Russo, Beraldin 2010; Migliari 2001; Andreozzi 2003. Addison, Gaiani 2000.
5. Wu 2011.
6. Agarwal, Furukawa, Snavely, Curless, Seitz, Szeliski 2010; Furukawa, Curless, Seitz, Szeliski 2010. Snavely, Seitz, Szeliski 2006; Snavely, Seitz, Szeliski 2008; Vu, Keriven, Labatut, Pons 2009.
7. Nguyen, Wunsche, Delmas, Lutteroth 2012.
8. Kersten, Lindstaedt 2012.
9. Lo Brutto, Meli 2012; Dellepiane, Dell'Unto, Callieri, Lindgren, Scopigno 2013.
10. Kersten, Stallmann 2012.
11. Remondino, Del Pizzo, Kersten, Troisi 2012.
12. Santagati, Inzerillo, Di Paola 2013; Santagati, Inzerillo 2013.
13. Inzerillo, Santagati, Di Paola 2013.
14. Furukawa, Curless, Seitz, Szeliski 2010; Furukawa Yasutaka, Jean Ponce. *Accurate, Dense, and Robust Multi-View Stereopsis*. IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), vol. 32, n. 8, 2010, pp. 1362-1376.
15. Shan Qi, Adams Riley, Curless Brian, Furukawa Yasutaka, Seitz Steven M. *The Visual Turing Test for Scene Reconstruction*. In 3DTV-Conference, 2013, pp. 25-32.
16. Restuccia, Galizia, Santagati 2012.
17. La pipeline definita in VisualSfM è costituita dai seguenti passaggi: 1) elaborazione del data set per il calcolo dei Sift e dei matching attraverso la determinazione delle features tra le immagini fotografiche; 2) elaborazione della sparse reconstruction; 3) eliminazione degli errori mediante l'algoritmo di *Multicore Bundle Adjustment*; 4) calcolo della dense reconstruction attraverso gli algoritmi sviluppati da Furukawa, PMVS2 e CMVS. A questi passaggi in VisualSfM si deve aggiungere un'ultima elaborazione per la conversione della nuvola di punti in mesh attraverso l'algoritmo di Poisson, così come indicato in letteratura. Wu 2011; Wu Changchang, Agarwal Sameer, Curless Brian, Seitz Steven M. *Multicore Bundle Adjustment*. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 3057-3064; Wu Changchang. *SiftGPU: A GPU Implementation of Scale Invariant Feature Transform (SIFT)*, 2011, <http://www.cs.unc.edu/~ccwu/siftgpu/> [novembre 2013].
18. Meshlab è un software *open source* sviluppato presso il Visual Computing Lab dell'ISTI - CNR di Pisa che consente di elaborare e gestire nuvole di punti e *mesh*. Cignoni Paolo, Callieri Marco, Corsini Massimiliano, Dellepiane Matteo, Ganovelli Fabio, Ranzuglia Guido. *MeshLab: an Open-Source Mesh Processing Tool*. In *Sixth Eurographics Italian Chapter Conference*, 2008, pp. 129-136.
19. CloudCompare è un software *open source* che consente di elaborare e gestire nuvole di punti e *mesh*. Nel nostro caso è stato necessario ricorrere al CloudCompare per realizzare l'allineamento tra le nuvole, operazione non consentita in Meshlab. <http://www.danielgm.net/cc/> [novembre 2013].
20. Galizia 2012.
21. Santagati, Inzerillo, Di Paola 2013; Santagati, Inzerillo 2013.
22. Le indicazioni della Autodesk suggeriscono di utilizzare un data set con massimo 70 foto con una risoluzione di 3Mpix.
23. Furukawa, Curless, Seitz, Szeliski 2010.

References

- Addison Alonzo C., Gaiani Marco. 2000. Virtualized Architectural Heritage: New Tools and Techniques for capturing Built History. *IEEE Multimedia Journal*, vol. 7, 2000, pp. 26-31.
- Agarwal Sameer, Furukawa Yasutaka, Snavely Noah, Curless Brian, Seitz Steven M., Szeliski Richard. 2011. Building Rome in a day. *Communications of the ACM*, vol. 54, no. 14, October 2011, pp. 105-112.
- Andreozzi Luigi. 2003. *Il laser scanner nel rilievo di architettura la nostra esperienza*. Enna: Il Lunario, 2003. ISBN: 978-88-8181-052-9.
- Bandiera Adriana, Beraldin Jean-Angelo, Gaiani Marco. 2011. Nascita ed utilizzo delle tecniche digitali di 3D imaging, modellazione e visualizzazione per l'architettura e i beni culturali. *IkbnoS*, 2011, pp. 81-134.
- Benedetto Benedetti, Marco Gaiani, Fabio Remondino. 2011. *Modelli digitali 3D in archeologia. Il caso di Pompei*. Pisa: Scuola Normale Superiore, 2011.
- Bertocci Stefano, Bini Marco. 2012. *Manuale di rilievo architettonico e urbano*. Novara: CittàStudi, 2012. ISBN: 978-88-2517-362-8.
- Cundari Cesare. 2012. *Il rilievo architettonico. Ragioni. Fondamenti. Applicazioni*. Roma: Aracne, 2012. ISBN: 978-88-548-4741-5.
- Dellepiane Matteo, Dell'Unto Nicolò, Callieri Marco, Lindgren Stefan, Scopigno Roberto. 2013. Archeological excavation monitoring using dense stereo matching techniques. *Journal of Cultural Heritage*, vol. 14, n. 3, pp. 201-210.
- De Luca Livio, 2011. *La fotomodellazione architettonica. Rilievo, modellazione, rappresentazione di edifici a partire da fotografie*. Palermo: D. Flaccovio, 2011. ISBN: 978-88-579-0070-4.
- Docci Mario, Maestri Diego. 2009. *Manuale di rilevamento architettonico e urbano*. Roma-Bari: Edizioni Laterza, 2009. ISBN: 978-88-4209-068-7.
- Furukawa Yasutaka, Curless Brian, Seitz Steven M., Szeliski Richard. 2010. Towards Internet-scale Multi-view Stereo. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 1434-1441.
- Galizia Mariateresa. 2012. *Il disegno delle torri medievali di Enna nel paesaggio urbano tra passato e presente*. Catania: Giuseppe Maimone Editore, 2012. ISBN: 978-88-7751-359-5.
- Guidi Gabriele, Russo Michele, Beraldin Jean-Angelo. 2010. *Acquisizione e modellazione poligonale*. Milano: McGraw Hill, 2010. ISBN: 978-88-3866-531-8.
- Inzerillo Laura, Santagati Cettina, Di Paola Francesco. 2013. Ricostruzione di una scena urbana mediante Visual SfM. *DISEGNARECON*, vol. 6, n. 12, 2013, p. XIX / 1-8.
- Kersten Thomas, Lindstaedt Maren, 2012. Image-Based Low-Cost Systems For Automatic 3d Recording And Modelling Of Archaeological Finds And Objects. In *EuroMed 2012, Progress In Cultural Heritage Preservation, LNCS 7616*, Springer-Verlag Berlin Heidelberg, 2012, pp. 1-10.
- Kersten Thomas, Stallmann Dirk, 2012. Automatic texture mapping of architectural and archaeological 3d models. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIX-B5, 2012, pp. 273-278.
- Lo Brutto Mauro, Meli Paola. 2012. Computer vision tools for 3d modelling in archaeology. *International Journal of Heritage in the Digital Era*, vol. 1, Supplement, 2012, pp. 1-6.
- Migliari Riccardo. 2001. *Frontiere del rilievo - Dalla matita alle scansioni 3D*. Vol. 5. Roma: Gangemi Editore, 2001. ISBN: 978-88-4920-215-1.
- Nguyen Hoang Minh, Wunsche Burkhard, Delmas Patrice, Lutteroth Christof. 2012. 3D Models from the Black Box: Investigating the Current State of Image-Based Modeling. In *Proceedings of the 20th International Conference on Computer Graphics, Visualization and Computer Vision (WSCG 2012)*, Pilsen, Czech Republic, June 25-28, 2012, pp 115-123.
- Pollefeys Marc, Van Gool Luc. 2002. From images to 3D models. *Communications of the ACM*, 45 (7), 2002, pp. 50-55.
- Remondino Fabio, El-Hakim Sabry. 2006. Image-based 3-D modelling: a review. *The Photogrammetric Record*, 2006, 21, pp. 269-291.
- Remondino Fabio, Del Pizzo Silvio, Kersten Thomas, Troisi Salvatore. 2012. Low-Cost and Open-Source Solutions for Automated Image Orientation—A Critical Overview. In *EuroMed 2012, Progress In Cultural Heritage Preservation, LNCS 7616*, Springer-Verlag Berlin Heidelberg 2012, pp. 40-54.
- Restuccia Francesca, Galizia Mariateresa, Santagati Cettina. 2012. Urban and Architectural Design of the “Piazza del Governo” in Enna. *Journal of Literature and Art Studies*, vol. 2 (6), 2012, pp. 606-621.
- Santagati Cettina, Inzerillo Laura. 2013. 123D Catch: efficiency, accuracy, constraints and limitations in architectural heritage field. *International Journal of Heritage in Digital Era*, vol. 2 (2), 2013, pp. 263-290.
- Santagati Cettina, Inzerillo Laura, Di Paola Francesco. 2013. Image-based modeling techniques for architectural heritage 3d digitalization: limits and potentialities. In *International Archives of Photogrammetry, Remote Sensing and Spatial and Information Sciences*. Volume XL-5/W2, 2013, pp. 550-560. ISSN 2194-9034.
- Snavely Noah, Seitz Steven M., Szeliski Richard. 2006. Photo tourism: Exploring photo collections in 3D. In *ACM Transactions on Graphics (SIGGRAPH Proceedings)*, 25 (3), 2006, pp. 835-846.
- Snavely Noah, Seitz Steven M, Szeliski Richard. 2008. Modeling the World from Internet Photo Collections. *International Journal of Computer Vision*, 80 (2), 2008, pp. 189-210.
- Vu Hoang Hiep, Keriven Renaud, Labatut Patrick, Pons Jean-Philippe. 2009. Towards high-resolution large-scale multi-view stereo. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009, pp. 1430-1437.
- Wu Changchang. 2011. *VisualSfM: A Visual Structure from Motion System, 2011*. <http://ccwu.me/vsfm/> [novembre 2013].